ECONOMIC FORECASTING DATABASE USING GDELT PROJECT DATA

YUSUF KAYA KUZU

SEPTEMBER 2019

ECONOMIC FORECASTING DATABASE USING
GDELT PROJECT DATA

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF NATURAL AND
APPLIED SCIENCES
OF
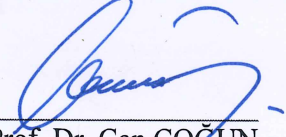ÇANKAYA UNIVERSITY

BY

YUSUF KAYA KUZU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COMPUTER ENGINEERING
DEPARTMENT

SEPTEMBER 2019

Title of the Thesis: **Economic Forecasting Database Using GDELT Project Data**

Submitted by **Yusuf Kaya KUZU**

Approval of the Graduate School of Natural and Applied Sciences, Çankaya University.

Prof. Dr. Can ÇOĞUN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Sıtkı Kemal İDER
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Hayri SEVER
Supervisor

Examination Date: 13.09.2019

Prof. Dr. Hayri SEVER (Çankaya Univ.)

Assist. Prof. Dr. Tülin ERÇELEBİ AYYILDIZ (Başkent Univ.)

Assist. Prof. Dr. Abdül Kadir GÖRÜR (Çankaya Univ.)

STATEMENT OF NON-PLAGIARISM

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Yusuf Kaya KUZU

Signature:

Date:13.09.2019

ABSTRACT

Economic Forecasting Database using GDELT Project Data

KUZU, Yusuf Kaya

MSc, Department of Computer Engineering

Supervisor: Prof. Dr. Hayri SEVER

September 2019

In this thesis, GDELT Global Knowledge Graph (GKG) data is used to analyse the sentiment values of our economic dictionary and for individual countries in a specific timeframe. The tone values of the individual online articles were used as basis for sentiment values for individual economic terms, which are then used on the original file resources to find their total sentiment values, negative sentiment values, positive sentiment values, and polarity of the files. Polarity of the files are calculated using the amount of matching terms divided by total terms in files. First 5 sample results of all these fields are calculated for individual files and location tags in the files initially by using the terms individually, and then by clustering the terms in groups.

Keywords: Sentiment analysis, Text analysis, Economic sentiment, Word clustering

ÖZ

GDELT Projesi Verisini kullanan Ekonomik Tahmin Veri Tabanı

KUZU, Yusuf Kaya

Yüksek Lisans, Bilgisayar Mühendisliği Anabilim Dalı

Tez Yöneticisi: Prof. Dr. Hayri SEVER

Eylül 2019

Bu tezde GDELT Küresel Bilgi Grafiği (GKG) verisi ekonomik sözlüğümüz ve tekil ülkeler için belirli bir zaman diliminde duygusallık değerlerini ölçmek için kullanılmıştır. Tekil çevrimiçi makalelerin ton değerleri tekil ekonomik terimlerin duygusallık değerlerinin temeli olarak kullanılmıştır, bu değerler daha sonra orijinal dosya kaynaklarında dosyaların toplam duygusallık değerleri, olumsuz duygusallık değerleri, olumlu duygusallık değerleri ve kutupluluk değerlerinin ölçülmesi için kullanılmıştır. Kutupluluk değerleri eşleşen değerlerin sayısının toplam kelime sayısına bölümüyle hesaplanmıştır. Bütün bu alanların ilk beş örnek değerleri tekil dosyalar ve dosyalardaki konum etiketleri ilk olarak terimleri ayrı olarak, daha sonra terimleri gruplar halinde kümeleyerek hesaplanmıştır.

Anahtar Kelimeler: Duygusallık tahlili, Metin çözümlemesi, Ekonomik duygusallık, Kelime kümelemesi

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Prof. Dr. Hayri SEVER for his direction and recommendations throughout the development of this thesis.

I also would like to express my gratitude for my family and my other lecturers for supporting me throughout this work.

TABLE OF CONTENTS

LIST OF TABLES

APPENDICES:

CHAPTER 1

INTRODUCTION

Our research is focused on developing an economic relational database using the sentiment values found from GDELT Global Knowledge Graph tables. We aim to achieve this goal by identifying sources with relevant economic information using the preconstructed tags of GDELT database, and process this information using our preconstructed word library. We use the source URLs from GDELT to extract their source texts and see matching words in these texts. If they do, we give the sentiment value of the GDELT GKG to these words, and then we find the arithmetic mean of these values to calculate the individual sentiment values of these words. Then we apply these words' sentiment values to articles to see if they match closely with their original values.

In this project, GDELT database was chosen due to its ease of use and extensive internal documentation. GDELT database was also chosen because its entries have its own calculated sentiment values and tags, which we utilize in our project. This also allows us to cross-reference our own sentiment calculations, both those that utilize GDELT's values and our own calculated values.

a. Background:

Previous sentiment analysis using GDELT have only focused on specific industries or specific countries in a specific timeframe. In our approach, we look at multiple locations in a specific timeframe to compare them with each other. This would allow us to see the macroeconomic situations of different countries and territories. This difference is important as different economic sentiment outlooks of different countries

and territories can give us information regarding the competitiveness and well-being of a country or territory in comparison to others.

Most sentiment analyses focus on a file-by-file basis, whereas we aim to look in a term-by-term basis. We do this as while there might be negatively charged news articles regarding the situation in a country or a territory, the economic situation might be unaffected or vice versa. Thus, we cannot consider overall sentiment value of the articles regarding a country or territory to be relevant to the economic situation. Since there can also be news articles that have economically relevant terms without being focused on the macroeconomic situation of a country or territory, we must collect articles that are relevant to the economic situation of the country or territory. These articles will also contain more economically relevant terms, which would allow us to encounter more matching terms than the average article.

Terms are also used because our files are too sparse to be used individually for sentiment analysis, whereas terms allow us to do sentiment analysis more accurately, as we do not have to concern with terms that are sentimentally relevant but economically irrelevant. Terms related to natural disasters such as "earthquake" or "land collapse" or terms related to rule of law such as "corruption" and "scandal" do not have immediate economic concerns as "tax" or "foreign exchange". This will cause differences with GDELT's original values but focusing on accuracy with GDELT would give us less accurate economic results.

GDELT Project's GKG database's TONE values contain both positive and negative values, alongside the total values. This allows us to calculate the sentiment values for both positive sentiment values of terms, files and locations, and the negative sentiment values of terms, files and locations. This also allows us to compare our calculations to GDELT's original values to see differences in calculations, as well as

the biases of the articles towards negativity or positivity. These TONE values also allow us to compare the positive charge or negative charge of a term, location or file.

b. Objectives:

In our calculations, we aim to calculate sentiment values for locations in the LOCATION tags in the GDELT GKG database. To accomplish this, we use an economic term dictionary for which we obtain sentiment values from the articles stored in the GDELT GKG database. We use the TONE values of the GDELT GKG database to serve as our basis for the terms, as we take these values as their sentiment values for each file, and then calculate their average as the sentiment value of the term.

We aim to calculate our results in two manners, first we use individual terms for recalculating the sentiment values of the files we obtained from GDELT's entries. This would allow us to cross-reference our results with GDELT's more general sentiment analysis. After accomplishing this, we use the files to calculate the average sentiment values of the LOCATION tags in GDELT entries, and we repeat this process after clustering our terms according to their cooccurrence values.

Our initial calculations aim to find sentiment values for files that approximate to GDELT's original TONE values to showcase if we can reverse engineer GDELT's sentiment calculations using our limited scope dictionary. Once this is fulfilled, the average sentiment values for LOCATION tags is used and cross-referenced by the original averages obtained from GDELT GKG database.

Our term clustering uses a clustering method based on k-means algorithm, and we cluster our terms to find terms that are close in terms of sentiment values, even if they might not be semantically similar. This difference is important as semantically different terms might be used in similar articles describing the same events, for example "debt" and "unemployment" can exist in similar files and can both reference

a recession, but high debt and high unemployment can exist or not exist alongside each other (a government that has high foreign debt can enter in debt to solve unemployment through government spending, or a government can avoid foreign debt but not spending for economy, thus causing high unemployment). Instead of finding semantically close terms, our clustering methodology intends to find terms that give similar positive or negative results, and thus have the same "value" within the texts.

CHAPTER 2

DEFINITIONS AND LITERATURE RESEARCH:

a. Definition of Relational Database:

Relational databases are databases that store data in a format that has data points in relation to each other. Each row in a relational database are stored with a unique ID, which serves as their key value. Because the views of the tables are separate from the tables themselves, the data can be edited without changing the view, and vice versa. Relational databases can also have stored procedures that can be accessed at any time [1].

While relational databases can have issues with multiple users viewing the same data, this can be solved by either locking, which would stop other users from viewing the data while data is updated and concurrency when users look at the same database at the same time [1].

b. Definition of Economic Sentiment:

Economic sentiment is generally calculated with consumer confidence [2] and market sentiment [3], but Eurostat uses a different index, Economic Sentiment Indicator. Eurostat's Economic Sentiment Indicator is a composite index composed of several different indices with different weights, which are:

1. Industrial confidence indicator,

2. Services confidence indicator,

3. Consumer confidence indicator,

4. Construction confidence indicator

5. Retail trade confidence indicator.

This index is calculated for the entire European Community [4].

Our approach aims to form a similar sentiment value by looking at the sentiment values of the GDELT GKG entries, by analysing the source texts of the entries, using our own economic vocabulary. We aim to find results that are like these indices.

c. Literature research:

One of our early literature research includes "Data exchange: semantics and query answering" [5], which discusses the data exchange problem, specifically in reference to universal solutions for query answering. This paper was significant in conceptualizing our research model, regarding extracting information from GDELT. In this paper, Fagin et. al describes a difference between data exchange and data integration, and describes the data exchange as limited exchange of data for a given query from a database J to an incomplete database I, whereas data integration is described as a view over a database J for a given query that does not have a specific limitation or requires direct reference to the database J [5]. According to this definition, our methodology is closer to data integration, as we do not directly reference to main database, which is GDELT, after our initial calculations for possible connectivity issues.

Our second research paper was "Data integration: a theoretical perspective" [6], which was influential in utilizing the data we obtained from GDELT. In Lenzerini's paper, the differences between "local as view" and "global as view" methods are described. In this paper, "local as view" refers to views where each source is used for a specific view, where as in "global as view", we have each element in multiple databases are referred by the same view. While the latter view is described to more efficient in multiple databases with similar information, the former is described to be

more efficient in forming ontologies [6]. Since we intend to find the cooccurrence values of terms to see their closeness to each other, our methodology is closer to "local as view" perspective.

Our third paper was "Searching for Interacting Features "[7], which was influential in conceptualizing our dataset. Zhao and Liu describe the issue of removing features in various databases to cut down processing time without causing a loss. Zhao and Liu describe the issues of assuming all features to be independent from each other, as well as possible issues of removing features that might be considered irrelevant, despite having high correlation with a relevant feature. In this paper, Zhao and Liu order the relevant features in terms using their inconsistency rate and feature relevance to calculate "consistency contribution", iteratively eliminating the irrelevant features [7]. The ideas described in this paper is used for trimming our cooccurrence matrix from terms without cooccurring values.

i. Similar works:

"Using Four Different Online Media Sources to Forecast Crude Oil Price" [8] uses GDELT as one of the four methods to predict crude oil prices in major crude oil indices, one of which is GDELT. In this paper by Elshendy et. al GDELT's predictive power over WTI Crude Oil index is compared with Wikipedia, Twitter and Google Trends. In Elshandy et. al's paper, Twitter is shown to be the fastest correlating source for sentiment analysis, whereas Google Trends is shown to be the most accurate in a three-day delay whereas GDELT is effective at a two-day lag. Elshandy et. al describe that these differences are present because of the different update time for these platforms, as GDELT updates at a two day delay, Wikipedia and Google Trends update on average three days, and Twitter updates almost simultaneously, but Twitter contains much more negative bias as a reflection of the perception of oil traders rather than an

actual reflection of the oil market [8].

"Growing pains for global monitoring of societal events" [9] uses GDELT data as an example of the issues in using sentiment analysis in researching real life events. This paper by Wang et. al showcases the issues of using automated codes for sentiment analysis using the example of Venezuelan protests. In this paper, it is shown that GDELT has multiple copy entries of the same event, causing an unnatural spike of articles describing the same event. In Wang et. al's paper this is described to because different articles reported different aspects of the same protests in January 2014, in contrast to three other databases, International Crisis Early Warning System (ICEWS), Gold Standard Report (GSR) and Social, Political and Economic Event Database (SPEED). In this paper, it is recommended to test different possible automated databases to find the best solution [9]. In our application, we use this insight in order to avoid duplicate entries in GDELT.

"Multi-level Analysis of Peace and Conflict Data in GDELT" [10] uses GDELT data to see correlation by two real life events, Sri Lankan Civil War and Fijian coup d'état. In this paper Keertipati et. al searches for possible relevant events before the capture of the town of Kilinochchi , the de facto capital of Liberation Tigers of Tamil Eelam (LTTE) by the Sri Lankan military and the relationship between New Zealand and Fiji before the Fijian coup d'etat. In the former, the aggregate events involving Sri Lanka are searched to see if there is a correlation with the spike of violence in Sri Lankan Civil War, as recapture of Kilinochchi from LTTE is considered as a turning point in Sri Lankan Civil War. Keertipati et. al's work shows a great correlation with the spike of violence and the prevalence of articles related to Sri Lanka, though there isn't a significant correlation with the end of the Civil War. In the latter case, former Prime Minister of Fiji Laisenia Qarase asked New Zealand government for protection

by New Zealand, but as that meeting ended without a resolution, head of the Fijian military, Frank Bainimarama went ahead with the government takeover, causing New Zealand to denounce the military coup and cut ties with Fijian government. In Keertipati's work we see the correlation between the spike of the articles about New Zealand-Fijian relations, and a decline after New Zealand denounces the action of the Fijian military and cut ties with the Fijian government [10]. This paper, while in retrospect, show great predictive power for GDELT in relation to crisis events, which is relevant to our research as we can utilize GDELT to predict possible recessions.

ii.   Other inspirations:

"Word Clustering and Disambiguation Based on Co-occurrence Data" [11] uses a cooccurrence matrix to cluster semantically similar terms. This paper by Li and Abe, cooccurrence matrix is created with two sets, one with a set of nouns N, and other with a set of verbs V. In this form, each noun is considered in relation to its use alongside a verb. In this application, each noun is clustered alongside each verb cooccur with most. This model is chosen as simpler cooccurrence matrices can have the issue of clustering terms with similar cooccurrence patterns but different absolute frequencies, that is, terms that might cooccur with other terms might not appear in texts as commonly as each other [11]. In our implementation of the cooccurrence matrix, we simply convert our cooccurrence matrix into a distance matrix, using the term that represents the given row as the base value.

d. GDELT GKG table structure:

GDELT GKG tables are constructed from the following values:

1. DATE, which contains the DATE information of the given entry. This corresponds to the date of the article in the entry itself, not the date it was added to the GDELT database.

2. NUMARTS, which contains the number of mentions the given url was given, specifically, in other news articles.

3. COUNTS, which contains a name value for a count of occurrences, such as number of deaths or arrests in a given article. The occurrences are described alongside the count.

4. THEMES, which contains the theme tags of the given source. These, just like occurrences, exist in GDELT's master list.

5. LOCATIONS, which contains the locations described in the source article. These are generally focused on the country or territories, but in other cases it also includes US states and international cities.

6. PERSONS, which contains the names of important people in the article, such as state leaders.

7. ORGANIZATIONS, which contains the companies and organizations described in the article.

8. TONE, which contains the TONE values of the article. These are the sentiment values calculated automatically by GDELT. This value is separated as:

   a. Tone: which contains the sentiment value of the article

   b. Positive tone: which contains the positive sentiment value of the article.

   c. Negative tone: which contains the negative sentiment value of the article.

d. Polarity: Which contains the emotional charge of the article.

e. Activity Reference Density: which contains the "activeness" (GDELT, 2013) of the article. This refers to whether the article is written in a personal manner.

f. Self/Group Reference Density: which contains the "group reference" (GDELT, 2013) of the article. This refers to whether the article uses specific terminology of a specific study, such as biology, engineering etc.

9. CAMEOEVENTIDS, which contains the CAMEO Event ID codes of the article. These are Ids unique to GDELT that describe the events in the article.

10. SOURCES, which contains the names of the articles. Specifically, it contains the top-level headers.

11. SOURCEURLS, which contains the URL links of the articles. Many articles have duplicates of themselves in different URL links [12].

For our sentiment studies, we will not use Activity Reference Density and Self/Group Reference Density values. GDELT's GKG archives the file format of version 1, whose codebook can be found here(http://data.gdeltproject.org/documentation/GDELT-Global_Knowledge_Graph_Codebook.pdf). Public tables in Google BigQuery instead use the version 2 framework, whose codebook can be found here (http://data.gdeltproject.org/documentation/GDELT-Global_Knowledge_Graph_Codebook-V2.1.pdf).

Notes:

1. GDELT's polarity values can be negative or positive. The documentation does not clarify how the negative and positive values are calculated, as such we take our polarity values as only positive.

CHAPTER 3

DATA MODELLING:

a. Approach:

We approach our problem by utilizing GDELT's predetermined tone values as our base, and we look for the relevant entries in the GDELT database. Our approach requires the use of both a reference array of the Themes we will look for in the GDELT database as well as a folder for documenting the text files we looked at from the source URLs. Our aim is to convert these data into an excel file and if/when possible, into a searchable SQL database.

Our relational database is intended to be composed of the following parts: sentiment values for our vocabulary, sentiment values for our files (calculated and obtained), sentiment values for our word clusters, and finally sentiment values for economies(countries). We also aim to store the relevant entries in the GDELT database including the text files of the source articles and their URLs.

b. Design iterations:

Our design initially focused on the development of a SQL database, but as SQL was too restrictive in its file formats, we focused on making necessary calculations in Java first, and then import these results to a SQL database when applicable. Therefore, one of our classes is GKGprosql, as it was originally conceived with SQL data format in the mind.

As this focus shifted, we looked on how to document our relevant documents. As we moved away from the SQL format, we first documented our initial results in txt files, especially in the case of the sentiment calculations for words and files. However,

we finally decided on csv format, as the vast majority of classes and results were stored in arrays or ArrayLists.

For our project, our goal was initially to simply document the GDELT articles in a searchable format for SQL, based on the Theme tags they had, and other relevant tags we would later add through our research on the articles. Instead, our focus shifted to study these articles according to their sentiment values, so we can get a coherent sentiment value for the general macroeconomic situation of a given country. This idea came to be as our initial research was focused on international trade and investment projects, but economic sentiment research had more resources related to it.

For this, we could not rely on GDELT's own sentiment values for these articles exclusively, as we had to calculate our own sentiment values for our own dictionary. Since we could not give manual sentiment values using expert opinion, we instead decided to look at the sentiment values of the articles itself and obtain the sentiment values for our dictionary through that method.

For creating clusters, we initially focused on adjusting our sentiment values by checking matching values within the dictionary we created. If there were similar words, or words used in a similar context, this could negatively impact our research by giving us false positives. To avoid this, we initially looked the commonly occurring words in each text, and later, once we did that, we decided to combine these words into clusters that would be enough for our calculations. This was decided as we could not simply remove the similar sounding words or words with a similar context, instead we had to act as if they were one collective group.

As our clusters were not enough, we employed a different method. Using our cooccurrence matrix, we extrapolated a distance matrix out of our cooccurrence matrix, using the cooccurrence values for each document and simplifying them. Once we

obtain the distance matrix, we use it to create clusters using k-means algorithm, and then use these clusters for processing our values.

Our displays for country sentiment values have changed from our initial design. In our initial design, our country data was used as our primary class for our operations. As our operations have shifted from simply documenting and analysing articles to obtaining sentiment values for our terms, we initially used the ISO2 country codes for our country sentiment calculations. Since these could not be used for obtaining their full names, as self-governing territories such as Gibraltar and non-governing areas such as "Oceans" exist as location codes exist within the GDELT data. Instead, we used the country names that existed in the tags of the GDELT data, which also considers such regions as analogous to sovereign countries.

c. Design problems:

Our initial design problem was the lack of access to the BigQuery database. Google BigQuery requires paid accounts for bulk access, as such we needed a reliable alternative to this. We used the archived files of GDELT database to accomplish this, as GDELT makes daily archives of the new entries for both GKG and Event database. GKG files can be found in http://data.gdeltproject.org/gkg/index.html.

Second major issue was reorganizing csv files into human-readable format. For this, we had to separate the comma separated values into different variables, and then separate those values that had multiple values in each column. This was necessary for us to separate relevant information from irrelevant ones.

Third issue was removing duplicate entries. This was caused with three issues, first was the issue of the duplicate entries with similar articles. In order to eliminate these easily, we had to look up to the headers of the entries, and for this, we needed an Internet connection to check them. As some of the websites couldn't be checked in

Java platform (some because they were defunct, others were due to incompatibility issues), we had to eliminate those entries as well. Secondly was the issue of the duplicate articles that passed the first check. These had to eliminated by looking up the articles themselves, and for this we had to compare previous two articles with the given article, using the first 14 lines (initially this was 5, but that proved insufficient). Thirdly was the issue of duplicate links in the given entries, as many entries have multiple URL sources. In order to remove these, we had to check them using the previous two methods.

Fourth issue was utilizing these articles without connection, as there were several cases of connection breakdown. For this issue, we had to store these articles as text files, and document them according their numbers in the classes we had. After doing this, we could calculate our dictionary's words' sentiment values using these texts and then reapply them to our texts to see if they match correctly.

Fifth issue was creating clusters of words according to their relevance to their each other. For this reason, we calculated their co-occurrence in the texts we studied. We gave each word in the dictionary and a float array that was large enough to contain spaces for all words in the dictionary, and we inserted the times they co-occurred per file. Finally, we found the total values of co-occurrence for each word by summing all the co-occurrence values according the number of files they cooccurred in. After this, we created clusters for each word that had co-occurrence with other words and used their average sentiment values as the values of the clusters.

Our initial form for establishing these clusters did not avoid terms that were put in other clusters, but we established a checklist for each used term, and then we ignored the terms that were used in a previous cluster. However, we did not add the primary terms for each connectivity array, as that could cause issues for our

calculations.

Sixth issue was separating the file sentiment values for each country. For this we used our calculated values for each file and stored them according to their country tags. After this, we added their sentiment values for each country tag they had and then found the average sentiment value for each country tag, using the average sentiment values of the files that have that country tag.

Seventh issue was reconsidering our clusters. Our initial clusters used their cooccurrence values to create ArrayLists composed of each term for each row of the cooccurrence matrix we used. Instead, we changed this idea to trim our cooccurrence matrix by not using the terms without connections to create a symmetric cooccurrence matrix, and then convert this symmetric matrix into a distance matrix that contains the distance values between terms. We find this by subtracting the cooccurrence values of each term of a row by subtracting it from the cooccurrence value of the term that row corresponds to. After this, we use k-means clustering to form less than 24 clusters and apply those clusters to our calculations.

CHAPTER 4

SOFTWARE DESIGN

Our code contains several classes. Full details can be found in APPENDIX A:

a. GKGcodes: which contains the relevant Theme codes from GDELT's master list. These are described below in GDELT Themes.

b. GKGrow: This class contains the elements of the given row in the GKG tables.

c. GKGsemipro: This class contains the elements of the given row in the GKG tables, but in a semi-processed form. In this class we separate relevant variables with multiple entries such as Tone to arrays.

d. GKGrowlist: Which is composed of an ArrayList of GKGsemipro and GKGrow classes.

e. GKGprosql: This class contains the GKG values that were processed from the GKG csv file. As we cannot access BigQuery directly, GDELT's csv file archives had to be used.

f. Sentcon: which is the class used for sentiment calculations. It contains:

g. Econterncon: which is the class used for storing the connectivity between terms.

h. Econcluster: This class contains the sentiment values and words for our word clusters. We form our word clusters by calculating their cooccurrence in the articles we obtained from the GDELT GKG database. We also calculate the sentiment values for our clusters in this class.

i. Econdictionary: This class is where we store our main operations, including our term dictionary. We store the sentiment values for our terms in the dictionary, as well as the calculated sentiment values for the files we utilized.

j. HTTPCLienttest: Which serves as our main class. This class only stores our main operations.

Notes:

1. Our distance matrix can have false positives as the cooccurrence values of terms in relation to themselves are not unique. That is, if a term b exists in 3 files, it is possible that there is a term c that also exists in 3 files. This does not mean these two terms are semantically similar even if they exist in the same files in our dataset. To solve this, we add 1 more cooccurrence for each term that is checked, so they have a unique value in their own row.

2. Because our normalization efforts have caused larger error margins compared to GDELT's Tone values, we decided against normalizing our finals results. Even though GDELT's Tone values are between 0 and 100, vast majority of the sentiment values are actually between 0-10, but we did not need to normalize our values to that range as our values are already in that range for small datasets.

CHAPTER 5

FORMAT

a. Economic dictionary sources:

Our economic dictionary was cultivated from various different sources. Our first source is the economics glossary of The Economist, which is found here: https://www.economist.com/economics-a-to-z/. This glossary largely contains financial terms that related to the news articles written in The Economist, but it also contains some macroeconomic terms such as "stagflation" and "depression".

Second source is The Conference Board of Canada, which is a non-profit think tank [13], founded by the former American National Industrial Conference Board, now known as The Conference Board. This association was initially formed out of tensions between the labour force and the industrialists of U.S. in 1916 [15]. Their link is as follows:

 https://www.conferenceboard.ca/topics/economics/economic-terms.aspx?AspxAutoDetectCookieSupport=1#A. This lexicon is primarily focused on terminology related to relationship between governments and private companies, as well as terminology related to employment.

Our third source is the glossary from Economics Classroom by Jason Walker, their glossary can be found here: https://econclassroom.com/glossary/. This glossary largely contains terms that are relevant to college students studying in economic departments.

Our fourth source is Economics Help by Tejan Pettinger, an UK based college teacher. Their glossary is found here:

https://www.economicshelp.org/economics-a-z/. This glossary is largely focused on terms and organizations related to UK economy, such as "Chartists movements", but it also contains terminology related to different schools of economics, such as "Chicago School of Economics" and "Classical Economics Model". This lexicon also contains major historical economic events, such as "Chartist movements" and "Asian debt crisis of 1990s".

Our fifth source is Guardian's Economic Jargon glossary. Much like The Economist's glossary, their glossary is focused on financial news. Their glossary can be found here: https://www.theguardian.com/business/glossary-business-terms-a-z-jargon. Much like Pettinger's glossary, this list contains terminology related to UK economics, but it also contains financial slang, such as "poison pill" and "grey knight".

Our sixth source is an old glossary by Eshargh Motahar from Union College. Their glossary is an old glossary focused on international trade dating back to 2000: http://minerva.union.edu/motahare/Eco354/glossary.pdf. While some of the terms such as those relating to EU are outdated, others are not. We use the currently relevant ones for our dictionary, but these also contain values that are relevant to articles relating to older events, such as "gold standard" or "Deutsche mark".

Our seventh and currently final source is the glossary by St. Louis Federal Reserve: https://www.stlouisfed.org/education/glossary. Their glossary is focused on U.S. economy and regulation legislations. It also contains terms relevant to how US government manages its economy at federal level, such as "Federal Reserve System".
Notes:

1. Our economic dictionary does not contain terms for energy sources and prices, even though Bloomberg.com's data format has been one of our inspirations and GDELT's economic themes contains tags related to energy

market. We have done this because while energy market is a good predictor for national economies, literature for energy market and commodity markets is different from general economic terminology.

2. Our economic dictionary does not contain specific terminology for other non-energy heavy industries and agriculture. Similar to energy markets, these industries need specific terminology, especially for more valuable items such as coffee and specific metal alloys, which could potentially clutter our dataset.

3. Though infrastructure was a major topic we expected to discuss in our research, we could not find sufficient resources for literature regarding infrastructure investment. We could find sufficient terminology for international trade, but to properly utilize it we would also need to organize the articles we obtained from GDELT and organize the information according to each country, which would increase our complexity and time cost.

b. GDELT Themes:

As we previously discussed, we have utilized theme tags that are relevant to our economic research. For our analysis, we have used tags that that contain the "ECON" prefix in GDELT's Master Theme list. The themes we used are:

1. "ECON_BANKRUPTCY", which covers discussions regarding bankruptcies.

2. "ECON_BITCOIN", which covers discussions regarding cryptocurrencies.

3. "ECON_BOYCOTT", which covers discussions regarding boycotts.

4. "ECON_BUBBLE", which covers discussions regarding economic bubbles, which a periodical expansions and deflations in the economy.

5. "ECON_CENTRALBANK", which covers discussions regarding central

banks.

6. "ECON_COST_OF_LIVING", which covers discussions regarding cost of living issues.

7. "ECON_COUNTERFEITMONEY", which covers discussions regarding counterfeit money.

8. "ECON_CURRENCY_EXCHANGE_RATE", which covers discussions regarding currency exchange rates.

9. "ECON_CURRENCY_RESERVES", which covers discussions regarding national currency reserves.

10. "ECON_CUTOUTLOOK", which covers discussions regarding cuts in economic outlook. [15] Economic outlook is the predictions regarding a national economy.

11. "ECON_DEBT", which covers discussions regarding debt. This can be national debt (public or private) or private debt.

12. "ECON_DEFLATION", which covers discussions regarding deflation.

13. "ECON_DEREGULATION", which covers discussions regarding deregulation.

14. "ECON_DEVELOPMENTORGS", which covers discussions regarding organizations for economic development, such as the World Bank.

15. "ECON_DIESELPRICE", which covers discussions regarding diesel prices.

16. "ECON_EARNINGSREPORT", which covers discussions regarding company earnings reports. These are important in predicting the prices of specific goods, such as RAM prices and processors.

17. "ECON_ELECTRICALDEMAND", which covers discussions regarding electrical demand.

18. "ECON_ELECTRICALGENERATION", which covers discussions regarding electric generation from power plants.

19. "ECON_ELECTRICALGRID", which covers discussions regarding electric grids.

20. "ECON_ELECTRICALLOADSHEDDING", which covers discussions regarding electrical load shedding. This refers to intentional blackouts by power producers due to work overload [16].

21. "ECON_ELECTRICALPRICE", which covers discussions regarding electric prices.

22. "ECON_EMERGINGECON", which covers discussions regarding emerging economies.

23. "ECON_ENTREPRENEURSHIP", which covers discussions regarding entrepreneurships.

24. "ECON_FOREIGNBANKS", which covers discussions regarding foreign banks. This generally refers to non-US banks.

25. "ECON_FOREIGNINVEST", which covers discussions regarding foreign investment.

26. "ECON_FREETRADE", which covers discussions regarding free trade.

27. "ECON_GASOLINEPRICE", which covers discussions regarding gasoline price.

28. "ECON_GOLDPRICE", which covers discussions regarding gold price.

29. "ECON_HEATINGOIL", which covers discussions regarding heating oil. Heating oil is generally used in countries without natural gas resources, such as U.S. and U.K. [17]

30. "ECON_HEATINGOILPRICE", which covers discussions regarding heating

oil price.

31. "ECON_HOUSING_PRICES", which covers discussions regarding housing prices. This includes both purchased houses and rented houses.

32. "ECON_IDENTITYTHEFT", which covers discussions identity theft.

33. "ECON_INFLATION", which covers discussions regarding inflation.

34. "ECON_INFORMAL_ECONOMY", which covers discussions regarding informal economy. This includes non-regulated markets such as the psychedelic drugs and border smuggling.

35. "ECON_INTEREST_RATES", which covers discussions regarding interest rates.

36. "ECON_IPO", which covers discussions regarding Initial Public Offerings (IPOs) [12]. This refers to the process of private companies offering their shares to the public in a new stock issuance [18].

37. "ECON_MIDDLECLASS", which covers discussions regarding the middle class.

38. "ECON_MONEYLAUNDERING", which covers discussions regarding money laundering.

39. "ECON_MONOPOLY", which covers discussions regarding monopolies and uncompetitive practices.

40. "ECON_MOU", which covers discussions regarding memorandums of understanding [19]. This is generally related to international trade discussions.

41. "ECON_NATGASPRICE", which covers discussions regarding national gas price. This generally covers US.

42. "ECON_NATIONALIZE", which covers discussions regarding nationalization of industries.

43. "ECON_NEWPOWERPLANT", which covers discussions regarding new power plants, such as new coal plants and new nuclear plants.

44. "ECON_OILPRICE", which covers discussions regarding oil price.

45. "ECON_PRICECONTROL", which covers discussions regarding price control. This generally refers to government price controls, rather than private price controls.

46. "ECON_PRICEGOUGE", which covers discussions regarding price gouging.

47. "ECON_PROPANE", which covers discussions regarding propane, such as propane production, and propane price.

48. "ECON_PROPANEPRICE", which covers discussions regarding propane price.

49. "ECON_QUITRATE", which covers discussions regarding quit rate. Quit rate, also pronounced quits rate, refers to the rate of people leaving employment [20], which is either alongside or in place of unemployment rate.

50. "ECON_REMITTANCE", which overs discussions regarding remittances.

51. "ECON_STOCKMARKET", which covers discussions regarding the stock market. This also includes discussions regarding government bonds [15].

52. "ECON_SUBSIDIES", which covers discussions regarding government subsidies to key businesses.

53. "ECON_SUSPICIOUSACTIVITYREPORT", which covers discussions regarding corruption and business malpractice.

54. "ECON_TAXATION", which covers discussions regarding taxation practices.

55. "ECON_TRADE_DISPUTE", which covers discussions regarding trade disputes, such as the current US vs. China tariff war.

56. "ECON_UNDEREMPLOYMENT", which covers discussions regarding unemployment.

57. "ECON_UNIONS", which covers discussions regarding unions.

58. "ECON_WORKINGCLASS", which covers discussions regarding the labour force.

59. "ECON_WORLDCURRENCIES", which covers discussions regarding non-US currencies.

As we did not use GDELT's Event database in addition to Global Knowledge Graph, we did not utilize the ID tags for events. We have done this because the Event database was not enough for our search for economics related articles, though we could feasibly include them in a future iteration of the work.

c. Output format:

Our outputs are listed here. For full details, check Appendix B below.

Our initial output prints the sentiment values we calculated for each term in .txt file format. We use this to see which terms were calculated during our runs of the software.

Our second output prints the values we used in Tables 1.0-1.3 and Tables 2.0-2.3. This prints the initial run of the sentiment calculations for individual files using term-by-term calculations.

Our third output prints the values we used in Tables 3.0-3.3 and Tables 4.0-4.3. This prints the initial run of the sentiment calculations using term-by-term calculations.

Our fourth output prints the connectivity values between the terms in a txt format, where each term is printed with the connected term, the file they cooccur in and the respective connectivity value. Our file format is in .txt.

Our fifth output prints the simplified connectivity matrix in a csv format. The cooccurrence values are listed in separate columns to the terms.

Our sixth output is similar to our fourth output, but we do not list the files they cooccur in, and our file format is in .csv.

Our seventh output prints the clustered terms according to their cluster number in a .csv format, alongside their sentiment values.

Our eighth output prints the sentiment values we calculated using clustered terms for our calculations. This output is the same as Tables 5.0-5.3 and Tables 7.0-7.3.

Our ninth output prints the sentiment values we calculated using clustered terms for our calculations. This output is the same as Tables 6.0-6.3 and Tables 8.0-8.3.

CHAPTER 6

RESULTS

As our database uses economic terms exclusively, our results do not match with GDELT's own values as closely as we hoped. Instead, we have a large gap that might be covered by either enlarging our dictionary or increasing the amount of the files we have in our dataset. Increasing our file size has not worked sufficiently so far, so we can rely on increasing our vocabulary, either by our economic vocabulary or our initial sentiment vocabulary.

Second issue we have is the fact our studies largely focused one a small dataset to due to lack of our computing resources. Due to several issues with the software we developed, we couldn't spend time on focusing a large dataset which left our results somewhat inconclusive.

We aim to perfect our results by increasing the capacity of our clusters and our datasets. We cannot increase our dictionary's base without halting the process speed significantly, nor can we rely on improving our internet connection due to lack of material resources.

a. Results with single terms in comparison GDELT's original tone values:

As discussed above, our values vary considerably from GDELT's Tone values as our dictionary is much smaller and specific than GDELT's dictionary. This is evidenced by our different polarity scores, which means the terms that match differ in each dictionary. Nonetheless, the results for 22 August 2019 rundown for 5 of the 19 documents:

Table 1.0: Calculated sentiment values vs. Original tone values

File Number	Original File Result	Calculated Result
0	Original file's tone is: -8.51485148514852	Calculated result is: 2.0820258
1	Original file's tone is: -5.22088353413655	Calculated result is: 2.0050862
2	Original file's tone is: -2.7027027027027	Calculated result is: 2.0306797
3	Original file's tone is: 2.02898550724638	Calculated result is: 2.0788028
4	Original file's tone is: 0.420168067226891	Calculated result is: 2.1741009

Our calculated values are different from the original GDELT values, as our negative values are much smaller than original GDELT values. This will be improved when we use a larger dataset as seen in Table 3.0, but it will not be a satisfying change. In this Table 1.0, our results are all positive, even though most of the original results of our tables are negative. Our initial results also do not differ significantly with each other, but this will change in the larger dataset results in Table 3.0.

Table 1.1: Calculated negative sentiment values vs. Original negative tone values

File Number	Original File Result (Negativity)	Calculated Result (Negativity)
0	Original file's negativity is: 9.5049504950495	Calculated negativity is: -0.5668151
1	Original file's negativity is: 8.43373493975904	Calculated negativity is: -0.725307
2	Original file's negativity is: 4.05405405405405	Calculated negativity is: -0.38039628
3	Original file's negativity is: 0.869565217391304	Calculated negativity is: -1.2790592
4	Original file's negativity is: 1.5406162464986	Calculated negativity is: -0.27572516

Our results in Table 1.1 differ significantly from the original values, as the original negative values in the given files are much larger than the results we will discuss in the later tables. Though Table 2.1 will show marginal improvement over these results that used the smaller database, our improvement will be significant, just as how our improvement from Table 1.0 to Table 2.1 won't be significant.

Table 1.2: Calculated positive sentiment values vs. Original positive tone values

File Number	Original File Result (Positivity)	Calculated Result (Positivity)
0	Original file's positivity is: 0.99009900990099	Calculated positivity is: 2.648841
1	Original file's positivity is: 3.21285140562249	Calculated positivity is: 2.7303932
2	Original file's positivity is: 1.35135135135135	Calculated positivity is: 2.411076
3	Original file's positivity is: 2.89855072463768	Calculated positivity is: 3.357862
4	Original file's positivity is: 1.96078431372549	Calculated positivity is: 2.449826

In Table 1.2 biggest difference occurs in File 0, while File 1 is the only result where the original has a higher value than the calculated value. The differences here are not as significant as Table 1.1, and our discrepancies will be improved in Table 2.1, but our initial positive results being closer to original is consistent with our later calculations.

Table 1.3: Calculated polarity values vs. Original polarity values

Full rundown is available in Appendix C as Tables A1.0-A.1.3

File Number	Original File Result (Polarization)	Calculated Result (Polarization, different calculation than original)
0	Original file's polarity is: -8.51485148514852	Calculated polarity is: 3.215656
1	Original file's polarity is: -5.22088353413655	Calculated polarity is: 3.4557002
2	Original file's polarity is: -2.7027027027027	Calculated polarity is: 2.7914724
3	Original file's polarity is: 2.02898550724638	Calculated polarity is: 4.636921
4	Original file's polarity is: 0.420168067226891	Calculated polarity is: 2.7255511

As our sentiment results are obtained from the positive and negative values Results in comparison GDELT's original tone values, thus as our positive and negative sentiment values differ, our sentiment values also vary differently. Mainly, our negative sentiment values are generally much lower than the negative sentiment values for GDELT entries, while our positive sentiment values are generally similar. For 29 August 2019, our first 5 file results for 178 entries is as follows:

Table 2.0: Calculated sentiment values for larger dataset vs. Original tone values

File Number	Original File Result	Calculated Result
0	Original file's tone is: -1.93298969072165	Calculated result is: -1.7215965
1	Original file's tone is: -2.90758047767394	Calculated result is: -1.6267359
2	Original file's tone is: -6.95970695970696	Calculated result is: -1.772471
3	Original file's tone is: 0.0496770988574271	Calculated result is: -1.6913092
4	Original file's tone is: -2.11946050096339	Calculated result is: -1.4819407

As seen in Table 2.0, our negative sentiment values different significantly from the GDELT's results. However instead of having smaller negative sentiment values than GDELT's original values, as seen in the first initial calculation, our negative results end up much higher than GDELT's original values. It can be inferred that our sentiment calculations become biased towards negativity in larger datasets, which could be due to bias towards negativity in the current articles. To have a more varied view, we could take articles from different time periods, but this would conflict with our desire to look at a specific timeframe.

Table 2.1: Calculated negative sentiment values for larger dataset vs. Original tone values

File Number	Original File Result (Negativity)	Calculated Result (Negativity)
0	Original file's negativity is: 4.7680412371134	Calculated negativity is: -3.9450452
1	Original file's negativity is: 5.81516095534787	Calculated negativity is: -4.0306535
2	Original file's negativity is: 8.60805860805861	Calculated negativity is: -3.9734616
3	Original file's negativity is: 2.0367610531545	Calculated negativity is: -4.066751
4	Original file's negativity is: 4.17469492614001	Calculated negativity is: -4.049916

In Table 2.1, we see a massive difference in negative values that go from twice the negative values to 8 times the negative values. This discrepancy is too large for our calculations to be comfortable. Our calculated negative values for each file is also too similar to each other, which could be because we haven't accounted for the frequencies of the terms we calculated. In other words, even if a term a appears twice, we counted as appearing once. However, as our positive sentiment values do not show as much as a discrepancy, the reason for this discrepancy is likely different.

Table 2.2: Calculated positive sentiment values for larger dataset vs. Original tone

values

File Number	Original File Result (Positivity)	Calculated Result (Positivity)
0	Original file's positivity is: 2.83505154639175	Calculated positivity is: 2.2234488
1	Original file's positivity is: 2.90758047767394	Calculated positivity is: 2.4039176
2	Original file's positivity is: 1.64835164835165	Calculated positivity is: 2.2009907
3	Original file's positivity is: 2.08643815201192	Calculated positivity is: 2.3754418
4	Original file's positivity is: 2.05523442517662	Calculated positivity is: 2.567975

Our positive sentiment values in Table 2.2 do not show as significant difference as our negative values, but this is consistent with the smaller dataset. Our results are closer in our larger dataset, but it is still too different from the original GDELT values to be comfortable. These values could be improved if we account for the frequency of the terms, but more likely we would need to increase amount of the terms in our dictionary.

Table 2.3: Calculated polarity values for larger dataset vs. Original polarity values

Full rundown is available in Appendix C as Tables A2.0-A2.3.

File Number	Original File Result (Polarization)	Calculated Result (Polarization - different calculation than original)
0	Original file's polarity is: 7.60309278350515	Calculated polarity is: 6.168494
1	Original file's polarity is: 8.72274143302181	Calculated polarity is: 6.4345713
2	Original file's polarity is: 10.2564102564103	Calculated polarity is: 6.1744523
3	Original file's polarity is: 4.12319920516642	Calculated polarity is: 6.442193
4	Original file's polarity is: 6.22992935131663	Calculated polarity is: 6.617891

As we don't take negative or positive terms into account in our polarity scores, our polarity scores end up higher than original GDELT results. Our polarity values differ greatly as they calculated as aggregate values for all the documents the location exists in the tags of the GDELT entries. While our values are closer than values of individual files, values are not close enough to be comfortable. Our results for first 5 results for 87 locations from 180 files is as follows:

Table 3.0: Calculated sentiment values for locations vs. Original GDELT averages

Location	Original File Result (Average)	Calculated Result
Location is: South Korea	Original files' average tone is: - 2.5908592	Calculated result is: 0.11933136
Location is: Japan	Original files' average tone is: - 2.1076438	Calculated result is: 0.11933088
Location is: United States	Original files' average tone is: - 1.4894167	Calculated result is: 0.118980885
Location is: China	Original files' average tone is: - 3.106933	Calculated result is: 0.11933136
Location is: North Korea	Original files' average tone is: - 3.840702	Calculated result is: 0.11933112

In Table 3.0 we see there is a significant difference in our calculated results and the GDELT average values. The difference is larger than the negative results for individual files, as the difference in values change from around 10 times to 30 times, but positive. In this case, it seems our negative results have lower values due to larger dataset, which can be seen in comparison with Table 3.1 and Table 4.1 below. These results confirm that our calculated values bias towards negativity as our dataset increases, but our positive values do not show a similar bias, instead they generally become closer to the original values.

Table 3.1: Calculated negative sentiment values for locations vs. Original GDELT

negative sentiment averages

Location	Original File Result (Negativity - Average)	Calculated Result (Negativity)
Location is: South Korea	Original files' average negativity is: -5.170285	Calculated negativity is: -2.0286405
Location is: Japan	Original files' average negativity is: -4.5477605	Calculated negativity is: -2.0286405
Location is: United States	Original files' average negativity is: -3.749911	Calculated negativity is: -2.0226052
Location is: China	Original files' average negativity is: -5.201587	Calculated negativity is: -2.0286405
Location is: North Korea	Original files' average negativity is: -6.592665	Calculated negativity is: -2.0286405

Though our calculated values in Table 3.1 are closer to original values than in Table 1.1 and Table 2.1, they are still 2 or 3 times lesser than the original GDELT average values. Our average GDELT results for countries have generally larger negative sentiment values than our previous calculations, and though they cover different datasets from different days, it seems our aggregate results have a bias towards negativity.

Table 3.2: Calculated positive sentiment values for locations vs. Original GDELT

positive sentiment averages

Location	Original File Result (Positivity - Average)	Calculated Result (Positivity)
Location is: South Korea	Original files' average positivity is: 2.579426	Calculated positivity is: 2.1479719
Location is: Japan	Original files' average positivity is: 2.4401162	Calculated positivity is: 2.1479714
Location is: United States	Original files' average positivity is: 2.260491	Calculated positivity is: 2.141586
Location is: China	Original files' average positivity is: 2.0946546	Calculated positivity is: 2.1479719
Location is: North Korea	Original files' average positivity is: 2.7519634	Calculated positivity is: 2.1479716

Our calculated positive values are almost identical to the average values obtained from GDELT, as seen in Table 3.2. These results are not consistent with Table 4.2 below, as the smaller dataset shows a variance between the GDELT averages and our calculated values. Nonetheless, the fact that our positive values being closer to the original than our negative sentiment values is consistent with the previous results.

Table 3.3: Calculated polarity values for locations vs. Original GDELT polarity

value averages

Full rundown is available in Appendix C as Tables A3.0-3.3

Location	Original File Result (Polarization - Average)	Calculated Result (Polarization - different calculation than original)
Location is: South Korea	Original files' average polarity is: 7.7497125	Calculated polarity is: 45.57143
Location is: Japan	Original files' average polarity is: 6.9878764	Calculated polarity is: 45.07143
Location is: United States	Original files' average polarity is: 6.010395	Calculated polarity is: 39.17919
Location is: China	Original files' average polarity is: 7.296241	Calculated polarity is: 49.48077
Location is: North Korea	Original files' average polarity is: 9.344628	Calculated polarity is: 39.0

Our polarity values are radically different, because instead of average values for all the files, we use the articles the location tags appear in and collect them as one single file. Then we divide the amount of the matching terms with the total amount of words in the combined document. While all the polarity values listed here and in Table 4.3 are positive, as GDELT's polarity values for files can be negative, our average polarity values can also be negative, though our calculated results cannot.

Below are the results for a smaller dataset to compare with the larger dataset discussed above. Our first 5 results for 47 locations in 72 files is as follows (these results are for 24 August 2019 like the ones above):

Table 4.0: Calculated sentiment values for locations using larger dataset vs. Original

GDELT average values

Location	Original File Result (Average)	Calculated Result
Location is: Germany	Original files' average tone is: -2.3911307	Calculated result is: -3.370785
Location is: United States	Original files' average tone is: -1.8895149	Calculated result is: -3.2793386
Location is: China	Original files' average tone is: -3.5094118	Calculated result is: -3.3458247
Location is: India	Original files' average tone is: -3.4675782	Calculated result is: -3.3639889
Location is: United Kingdom	Original files' average tone is: -1.8051269	Calculated result is: -3.3707845

Our calculated results for smaller dataset gives us much closer results, but the results that do not match differ greatly, with the results for US and UK having 2 times the negative results than the GDELT Average values. Differences seem to be caused primarily due to US and UK differing in the negative results, as seen in Table 4.1. Differences seem to be due to our calculations ignoring the frequencies of the matching terms in the documents.

Table 4.1: Calculated negative sentiment values for locations using larger dataset vs.

Original GDELT sentiment values.

Location	Original File Result (Negativity - Average)	Calculated Result (Negativity)
Location is: Germany	Original files' average negativity is: -5.208263	Calculated negativity is: -6.629212
Location is: United States	Original files' average negativity is: -3.9072924	Calculated negativity is: -6.449366
Location is: China	Original files' average negativity is: -5.543452	Calculated negativity is: -6.5801244
Location is: India	Original files' average negativity is: -5.2850695	Calculated negativity is: -6.6158466
Location is: United Kingdom	Original files' average negativity is: -4.1516275	Calculated negativity is: -6.629212

Our results with the smaller dataset match closer to GDELT values than the results to Table 3.1, but as our positive results in Table 4.2 differ greatly with the original results than the positive results in Table 3.2, we cannot say that the smaller dataset has given us more accurate results. It is much more likely that the results do not differ significantly in these first 5 locations because of the lesser amount of matching terms, but since polarity values do not differ with each other in Table 4.3, it is more likely this is because the lesser matching documents the other locations have.

Table 4.2: Calculated positive sentiment values for locations using larger dataset vs.

Original GDELT sentiment values

Location	Original File Result (Positivity - Average)	Calculated Result (Positivity)
Location is: Germany	Original files' average positivity is: 2.8171325	Calculated positivity is: 3.258427
Location is: United States	Original files' average positivity is: 2.0177817	Calculated positivity is: 3.1700275
Location is: China	Original files' average positivity is: 2.0340402	Calculated positivity is: 3.2342997
Location is: India	Original files' average positivity is: 1.8174914	Calculated positivity is: 3.2518578
Location is: United Kingdom	Original files' average positivity is: 2.3465002	Calculated positivity is: 3.2584274

Table 4.2 shows bigger variance in GDELT values than Table 3.2, showing that a smaller dataset will cause less accurate results for locations' average sentiment values, as it has done for individual files in Tables 1.2 vs. 2.2. But since our negative values are similar to the original values, this does not affect our final values as greatly as the differences in the negative values for previous calculations.

Table 4.3: Calculated polarity values for locations using larger dataset vs. Original

polarity value averages.

Full rundown is available in Appendix C as Tables A4.0-A4.3.

Location	Original File Result (Polarization - Average)	Calculated Result (Polarization - different calculation than original)
Location is: Germany	Original files' average polarity is: 8.025395	Calculated polarity is: 56.714287
Location is: United States	Original files' average polarity is: 5.925073	Calculated polarity is: 37.55634
Location is: China	Original files' average polarity is: 7.577492	Calculated polarity is: 47.75
Location is: India	Original files' average polarity is: 7.1025605	Calculated polarity is: 44.25
Location is: United Kingdom	Original files' average polarity is: 6.4981265	Calculated polarity is: 34.136364

While our polarity values are calculated differently than the original values, our calculated results do not show a great variance between each other, just like the original average values. Nonetheless, the biggest difference occurs with US and UK values, as US's polarity value in Table 4.3 is the second smallest value in the calculated results, when it should be the smallest value like the original GDELT values.

While not immediately apparent for the first 5 results, our smaller sample size of files has caused more different results than original averages from GDELT's Tone values. This confirms the necessity of a larger sample size for calculations.

b. Results for cluster-based calculations:

Our results for calculations using clusters as our terms collectively instead of terms individually is described here. Here's the first 5 results in 59 files for 5 clusters (though our k is 24, we could only create 4 clusters with our calculation with two clusters having only one member):

Table 5.0: Calculated sentiment values using clusters vs. Original tone values

File Number	Original File Result	Calculated Result
0	Original file's tone is: -3.29512893982808	Calculated result is: -9.734554
1	Original file's tone is: 3.68349249658936	Calculated result is: -9.2402525
2	Original file's tone is: 1.42302716688228	Calculated result is: -9.779709
3	Original file's tone is: 1.15942028985507	Calculated result is: -9.325769
4	Original file's tone is: -1.93187595322827	Calculated result is: -9.828964

Our calculated results in Table 5.0 are significantly inaccurate compared to other calculations we have in the other tables. This is because of the limited number of clusters we could form using our algorithm. However, other results for clusters will further reinforce our calculations using the term clusters is insufficient for sparse texts like the ones we use in our calculations.

Table 5.1: Calculated negative sentiment values using clusters vs. Original negative

tone values.

File Number	Original File Result (Negativity)	Calculated Result (Negativity)
0	Original file's negativity is: 5.58739255014327	Calculated negativity is: -10.857763
1	Original file's negativity is: 1.63710777626194	Calculated negativity is: -10.306429
2	Original file's negativity is: 0.517464424320828	Calculated negativity is: -10.908128
3	Original file's negativity is: 2.17391304347826	Calculated negativity is: -10.401813
4	Original file's negativity is: 3.76207422470768	Calculated negativity is: -10.963066

Our results for Table 5.1 are significantly different from the original results. Because our calculated results are very close to each other and the GDELT negative Tone values are significantly different from each other, these results show that creating clusters with aggregate sentiment values is insufficient for our intentions.

Table 5.2: Calculated positive values using clusters vs. Original positive tone values.

File Number	Original File Result (Positivity)	Calculated Result (Positivity)
0	Original file's positivity is: 2.29226361031519	Calculated positivity is: 1.1232088
1	Original file's positivity is: 5.3206002728513	Calculated positivity is: 1.066176
2	Original file's positivity is: 1.9404915912031	Calculated positivity is: 1.1284188
3	Original file's positivity is: 3.33333333333333	Calculated positivity is: 1.0760429
4	Original file's positivity is: 1.83019827147941	Calculated positivity is: 1.134102

Our Table 5.2 results are similar to Table 5.0 and Table 5.1, as our calculated results do not differ significantly from each other, but the original GDELT values vary different from each other. While Table 5.2's original values do not differ from each other considerably compared to values in Table 5.1 and Table 5.0, this does not mean that our calculated are fitting for the files.

Table 5.3: Calculated polarity values using clusters vs. Original polarity values

Full rundown is available in Appendix C as Tables A5.0-A5.3.

File Number	Original File Result (Polarization)	Calculated Result (Polarization, different calculation than original)
0	Original file's polarity is: 7.87965616045845	Calculated polarity is: 0.1458886
1	Original file's polarity is: 6.95770804911323	Calculated polarity is: 0.011923688
2	Original file's polarity is: 2.45795601552393	Calculated polarity is: 0.08906882
3	Original file's polarity is: 5.50724637681159	Calculated polarity is: 0.110843375
4	Original file's polarity is: 5.59227249618709	Calculated polarity is: 0.021182701

Our calculated polarity values for Table 5.3 are significantly less than the calculated polarity values for Table 1.3 and Table 3.3, because we consider the number of matching clusters divided by the total words by the respective GDELT entries' texts. Because the maximum amount of our clusters, 24, is significantly lower than our total number of terms, 2477, our polarity values start off significantly lower than GDELT's values. Though the calculated polarity averages in this table are positive, because GDELT polarity values can be negative, our polarity averages can be negative, but our calculated polarity values using our clusters cannot.

Our results for clusters are less accurate than our results for individual terms, as our sentiment values are not discrete. As our initial values were also inaccurate to original GDELT values, these clusters strongly confirm that we cannot use a separate dictionary to find matching results. Our results for first 5 locations in 51 location results using 16 clusters (with 7 clusters having only one member) used on 58 files are:

Table 6.0: Calculated sentiment values for locations using clusters vs. Original GDELT average tone values

Location	Original File Result (Average)	Calculated Result
Location is: China	Original files' average tone is: -1.3909457	Calculated result is: -3.5340407
Location is: Canada	Original files' average tone is: -0.4365438	Calculated result is: -3.6735737
Location is: United Kingdom	Original files' average tone is: -0.7342701	Calculated result is: -3.696995
Location is: United States	Original files' average tone is: -1.1494569	Calculated result is: -3.6296322
Location is: Brazil	Original files' average tone is: -6.534622	Calculated result is: -3.6776626

Our calculated results using clusters differ greatly from the original average values, which is consistent with the inaccuracies with individual files. Most notably, both Table 5.0 and Table 6.0 shows our calculated results not differing from each other significantly. This is likely due to the lack of multiple clusters for our terms.

Table 6.1: Calculated negative sentiment values for locations using clusters vs.

Original GDELT average negative tone values.

Location	Original File Result (Negativity - Average)	Calculated Result (Negativity)
Location is: China	Original files' average negativity is: -3.2600229	Calculated negativity is: -6.1042337
Location is: Canada	Original files' average negativity is: -2.5134428	Calculated negativity is: -6.3452435
Location is: United Kingdom	Original files' average negativity is: -3.6720748	Calculated negativity is: -6.3856997
Location is: United States	Original files' average negativity is: -3.8065732	Calculated negativity is: -6.2693424
Location is: Brazil	Original files' average negativity is: -7.4426174	Calculated negativity is: -6.3523088

Despite the original values, our calculated results in Table 6.1 do not differ significantly from each other. This is likely because our clusters give us a bias towards the overall average of the files, rather than giving us a more simplified result. This could be improved by increasing the number of clusters, but as the amount of matching terms is significantly smaller than our total dictionary, creating clusters of terms will only benefit in significantly larger datasets.

Table 6.2: Calculated positive sentiment values for locations using clusters vs.

Original GDELT average positive tone values

Location	Original File Result (Positivity - Average)	Calculated Result (Positivity)
Location is: China	Original files' average positivity is: 1.8690774	Calculated positivity is: 2.570193
Location is: Canada	Original files' average positivity is: 2.0768988	Calculated positivity is: 2.6716697
Location is: United Kingdom	Original files' average positivity is: 2.9378045	Calculated positivity is: 2.6887047
Location is: United States	Original files' average positivity is: 2.657113	Calculated positivity is: 2.6397102
Location is: Brazil	Original files' average positivity is: 0.9079947	Calculated positivity is: 2.6746461

Table 6.2's calculated results do not differ significantly from each other much like Table 6.0 and Table 6.1, which causes in the case of Brazil, which has a much smaller original GDELT average than other locations, but our results conform to the other locations' overall average value. This proves clusters are inefficient in the case of terms with significant difference from each other, and since GDELT values are already normalized, re-normalizing them will not solve our miscalculation.

Table 6.3: Calculated polarity values for locations using clusters vs. Original

GDELT average polarity values

Full results are available in Appendix C as Tables A6.0-A6.3.

Location	Original File Result (Polarization - Average)	Calculated Result (Polarization - different calculation than original)
Location is: China	Original files' average polarity is: 5.1291	Calculated polarity is: 0.0753389
Location is: Canada	Original files' average polarity is: 4.590341	Calculated polarity is: 0.072262146
Location is: United Kingdom	Original files' average polarity is: 6.609879	Calculated polarity is: 0.048169825
Location is: United States	Original files' average polarity is: 6.4636803	Calculated polarity is: 0.064564176
Location is: Brazil	Original files' average polarity is: 8.350612	Calculated polarity is: 0.0652855

Our calculated polarity results match with the previous results as they do not significantly differ from each other, but in our calculations polarity values differ significantly, since we do not account each time a cluster has a matching term, but instead the amount of the matching cluster divided with the total words the files with the given country tags. This causes our polarity values to radically smaller other polarity values in Tables 1.3, 2.3, 3.3 and 4.3.

For country averages our positive scores are similar but our negative scores differ significantly, meaning our total values also differ significantly. For comparison, our first 5 results within 180 files for 11 clusters is as thus:

Table 7.0: Calculated sentiment values using clusters on a larger dataset vs. Original tone values.

File Number	Original File Result	Calculated Result
0	Original file's tone is: 1.73410404624277	Calculated result is: -5.6077147
1	Original file's tone is: -1.94489465153971	Calculated result is: -5.857382
2	Original file's tone is: 3.97553516819572	Calculated result is: -5.9918346
3	Original file's tone is: -1.65413533834587	Calculated result is: -5.5697117
4	Original file's tone is: -2.0979020979021	Calculated result is: -5.548524

Our calculated results in Table 7.0 are similar to each other, much like Table 5.0, but due to original GDELT values being noticeably different from each other, our cluster-based results do not give us a close result to the original, especially since not all values are negative in the original GDELT values. This suggests that in the cases where original dataset do not have a significant variance in terms of absolute distance, but have different signs, our cluster-based results have a tendency to yield more inaccurate results.

Table 7.1: Calculated negative sentiment values using clusters on a larger dataset vs.

Original negative tone values.

File Number	Original File Result (Negativity)	Calculated Result (Negativity)
0	Original file's negativity is: 1.15606936416185	Calculated negativity is: -7.9281516
1	Original file's negativity is: 5.02431118314425	Calculated negativity is: -8.281131
2	Original file's negativity is: 0	Calculated negativity is: -8.471218
3	Original file's negativity is: 4.66165413533835	Calculated negativity is: -7.874423
4	Original file's negativity is: 2.0979020979021	Calculated negativity is: -7.8444686

Our negative sentiment calculations show greater change between each other than previous tables, which suggests that should our clustering process be improved upon, we could find more accurate results. Nonetheless, the values are still significantly off compared to the original values, especially in the case of File 2, which has 0 negative values. However, we could improve our results by ignoring these 0 positive or negative tone values, much like how we ignore the results with 0 total tone values.

Table 7.2: Calculated positive sentiment values using clusters on a larger dataset vs.

Original positive tone values

File Number	Original File Result (Positivity)	Calculated Result (Positivity)
0	Original file's positivity is: 2.89017341040462	Calculated positivity is: 2.3204367
1	Original file's positivity is: 3.07941653160454	Calculated positivity is: 2.4237487
2	Original file's positivity is: 3.97553516819572	Calculated positivity is: 2.4793837
3	Original file's positivity is: 3.00751879699248	Calculated positivity is: 2.3047116
4	Original file's positivity is: 0	Calculated positivity is: 2.2959447

Much like Table 7.1, we have a 0-positive value, which we can ignore in future iterations of our model. Table 7.2 is also like Table 5.2 as though our positive values do not differ much from the original results, because they do not differ with each other enough to be comparable to original results, we cannot consider our calculated results to be enough.

Table 7.3: Calculated polarity values using clusters on a larger dataset vs. Original

polarity values

Full results are available in Appendix C as Tables A7.0-A7.3.

File Number	Original File Result (Polarization)	Calculated Result (Polarization, different calculation than original)
0	Original file's polarity is: 4.04624277456647	Calculated polarity is: 0.17910448
1	Original file's polarity is: 8.10372771474878	Calculated polarity is: 0.1109215
2	Original file's polarity is: 3.97553516819572	Calculated polarity is: 0.037037037
3	Original file's polarity is: 7.66917293233083	Calculated polarity is: 0.05357143
4	Original file's polarity is: 2.0979020979021	Calculated polarity is: 0.2991453

Our polarity values are much lower than the results with individual sentiment values for terms, much like the previous tables Table 5.3 and Table 6.3, as we do account for matching clusters, rather than matching terms. Nonetheless, as our dataset is larger and we have more clusters to work on, our polarity values generally increase compared to Table 5.3. This increase is not satisfying, but it is important. Much like previous Tables 5.3 and Table 6.3, our original average polarity values are all positive, but they can be negative as GDELT polarity values can be negative.

Much like our initial cluster results, our positive results are more accurate than our negative results, which impacts our final results. Our first 5 locations in 99 locations with 12 clusters (including 4 clusters that have single members) for 183 files is as thus:

Table 8.0: Calculated sentiment values for locations using clusters on a larger dataset vs. Original GDELT average sentiment values

Location	Original File Result (Average)	Calculated Result
Location is: China	Original files' average tone is: -2.0371635	Calculated result is: -5.7690387
Location is: Canada	Original files' average tone is: -0.57163346	Calculated result is: -5.798462
Location is: United Kingdom	Original files' average tone is: -0.8766754	Calculated result is: -5.8459535
Location is: United States	Original files' average tone is: -0.56520724	Calculated result is: -5.801475
Location is: Brazil	Original files' average tone is: -3.3280118	Calculated result is: -5.7032747

Our results in Table 8.0 show that despite the increase in our clusters and our dataset our calculated results lean towards a specific value range. In the case of Table 8.0 our calculated result is not like original values or their average value, unlike Table 6.0. This would suggest increasing our dataset or cluster count will not yield more accurate results as our sentiment values will be aggregate averages of our clusters, but the sparsity of the matching terms requires specific sentiment values.

Table 8.1: Calculated negative sentiment values for locations using clusters on larger dataset vs. Original GDELT average negative tone values

Location	Original File Result (Negativity - Average)	Calculated Result (Negativity)
Location is: China	Original files' average negativity is: -3.8480163	Calculated negativity is: -8.156233
Location is: Canada	Original files' average negativity is: -2.7043316	Calculated negativity is: -8.197832
Location is: United Kingdom	Original files' average negativity is: -3.2977118	Calculated negativity is: -8.264974
Location is: United States	Original files' average negativity is: -3.4962513	Calculated negativity is: -8.20209
Location is: Brazil	Original files' average negativity is: -5.0211368	Calculated negativity is: -8.063258

Table 8.1 shows similar results to Table 8.0, but our negative results are significantly different from the original values. Much like the Table 8.0 results our calculated results do not differ significantly from each other due to our clusters taking average values of the terms. Our larger dataset does not give us a comparative advantage when compared to Table 6.1, as our calculated negative sentiment values are more than double of most of the values presented in the first 5 results. Instead, our results in Table 8.1 are slightly more inaccurate compared to Table 6.1.

Table 8.2: Calculated positive sentiment values for locations using clusters on a

larger dataset vs. Original GDELT average positive tone values.

Location	Original File Result (Positivity - Average)	Calculated Result (Positivity)
Location is: China	Original files' average positivity is: 1.8108537	Calculated positivity is: 2.3871942
Location is: Canada	Original files' average positivity is: 2.1326973	Calculated positivity is: 2.39937
Location is: United Kingdom	Original files' average positivity is: 2.421037	Calculated positivity is: 2.4190202
Location is: United States	Original files' average positivity is: 2.931045	Calculated positivity is: 2.4006152
Location is: Brazil	Original files' average positivity is: 1.6931261	Calculated positivity is: 2.3599832

In Table 8.2, our calculated results are closer to each other than the results in Table 6.2, which suggests increasing our dataset and cluster count will yield results closer to average, which will cause more significant difference than original GDELT values. These values are consistent with other cluster-based results, which further showcases the unreliability of clusters for sparse documents.

Table 8.3: Calculated polarity results for locations using clusters on a larger dataset

vs. Original GDELT average polarity values

Full results are available in APPENDIX C as Tables A8.0-A8.3

Location	Original File Result (Polarization - Average)	Calculated Result (Polarization - different calculation than original)
Location is: China	Original files' average polarity is: 5.65887	Calculated polarity is: 0.071994275
Location is: Canada	Original files' average polarity is: 4.83703	Calculated polarity is: 0.058414306
Location is: United Kingdom	Original files' average polarity is: 5.718749	Calculated polarity is: 0.051593244
Location is: United States	Original files' average polarity is: 6.427304	Calculated polarity is: 0.055679962
Location is: Brazil	Original files' average polarity is: 6.714264	Calculated polarity is: 0.06560743

Our polarity results are lower than our results for Table 6.3, as the total word count of the files that contain the given location tag is much larger, due to larger dataset. Much like in the case of Table 6.3, the only way to find results closer to the original values is by checking for each term in each document, rather than checking for each cluster once for the total sum of the documents.

Much like our file results, our positive scores remain similar, but our negative scores remain different. Nonetheless, increasing the file count increases the amount and closeness of results significantly.

CHAPTER 7

CONCLUSIONS:

Our results primarily differ in negative calculations. This might be due to our negative sentiment values being stored as negative, instead of being stored as positive values like GDELT's values. Our polarity values differ because as mentioned before the reason for negative values are not presented in GDELT's GKG documentation. Having clusters for terms has cut down the processing time but constructing the clusters has taken time. Clusters do not increase our accuracy, instead they can cause greater discrepancies.

Increasing word size and dataset count has increased our accuracy compared to our source material in GDELT data, but it also increased the complexity of our calculations. Not to mention, even with increased values our negative calculations and polarity calculations have turned significantly different. It is possible our accuracy could be increased by having an additional dictionary for common non-economic terms and using machine learning techniques to utilize them as a secondary sentiment check, but this would be irrelevant to our economic analysis directly.

CHAPTER 8

FUTURE WORK:

Our main objective for future implementations is converting our outputs to SQL database format as it was originally planned. To accomplish this, we would need to create separate entries for each location tags in the GDELT database, if not separate tables. This would archive our previous rundowns of our application, as well as archiving other relevant data for these locations taken from other sources such as World Bank and IMF.

Second main objective would be increasing the term count of economic dictionary by including terms relevant to energy markets, as presented in the GDELT Themes, as well as infrastructure investment terminology, as we originally intended. Our economic dictionary can also have specific terminology relevant to other major economies, such as European Union (EU) and People's Republic of China (PRC).

Our third objective would be to try different term clustering methods to see which one would be better suited in either calculating sentiment values for files and locations or finding terms with similar sentiment value. Such an approach can either use common clustering methods like fuzzy clustering or alternate clustering methods such as those proposed by Li and Abe [11].

Our fourth objective would be to increase our viable THEME tags from simply economics related subjects to subjects that are partially related to economics, such as rule of law issues and international conflicts. We can also add CAMEO Event IDs as possible checks for relevant articles, especially if we also add GDELT EVENT database as another source in addition to our current use of the Global Knowledge

Graph.

Our fifth objective would be to compare our results with other public sources, such as news aggregation websites, social media platforms such as Twitter and other public event databases. We could not find an additional source to compare to in our limited timeframe, but in a future implementation we can compare the efficiency of GDELT with other sources, as several papers in our literature research has done.

REFERENCES:

1. "What Is a Relational Database", Oracle Türkiye,
    https://www.oracle.com/tr/database/what-is-a-relational-database/ [Accessed:
    Aug. 26, 2019]

2. "Consumer confidence", Wikipedia,
    https://en.wikipedia.org/wiki/Consumer_confidence [Accessed: Aug. 26,
    2019]

3. "Market sentiment", Wikipedia,
    https://en.wikipedia.org/wiki/Market_sentiment [Accessed: Aug. 26, 2019]

4. "Economic sentiment indicator", Eurostat,
    https://ec.europa.eu/eurostat/web/products-datasets/-/teibs010 [Accessed: Aug.
    26, 2019]

5. Fagin, R., Kolaitis, P., Miller, R., Popa, Ln. (2005). "Data Exchange:
    Semantics and Query Answering", *Theoretical Computer Science*, vol. 336,
    no. 1, pp. 89-124. Internet:
    https://users.soe.ucsc.edu/~kolaitis/bio11/papers11/icdt03-tcs.pdf [Accessed:
    Aug. 26, 2019]

6. Lenzerini, M. (2002). "Data Integration: A Theoretical Perspective",
    Symposium on Principles of Database Systems, pp. 233-246. Internet:
    https://www.cs.ubc.ca/~rap/teaching/534a/readings/Lenzerini-pods02.pdf
    [Accessed: Aug. 26, 2019]

7. Zhao, Z. and Liu, H. (2007). "Searching for Interacting Features", *IJCAI
    International Joint Conference on Artificial Intelligence*, pp. 1156-116.
    Internet: http://www.aaai.org/Papers/IJCAI/2007/IJCAI07-187.pdf [Accessed:
    Aug. 26, 2019]

8. Elshendy, M., Colladon, A. F., Battistoni, E., Gloor, P. A. (2018). "Using Four
    Different Online Media Sources to Forecast Crude Oil Price", *Journal of
    Information Science*, vol. 44., pp. 408-421. Internet:
    http://www.ickn.org/documents/OipPricePredictionv7.pdf [Accessed: Aug. 26,
    2019]

9. Wang, W., Kennedy, R., Lazer, D., Ramakrishnan, N. (2016). "Growing pains for global monitoring of societal events", *Science,* vol. 353., pp. 1502-1503. Internet: http://people.cs.vt.edu/~naren/papers/1502.full.pdf [Accessed: Aug. 26, 2019]

10. Keertipati, S., Savarimuthu, B. T. R., Purvis, M., Purvis, M. (2014). "Multi-level Analysis of Peace and Conflict Data in GDELT", *MLSDA'14 Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, pp. 34. Internet: https://www.researchgate.net/publication/269097432_Multi-level_Analysis_of_Peace_and_Conflict_Data_in_GDELT [Accessed: Aug. 26, 2019]

11. Li, H. and Abe, N. (1998), "Word Clustering and Disambiguation Based on Co-occurrence Data", *ACL '98/COLING '98 Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 2*, pp.749-755. Internet: https://www.aclweb.org/anthology/P98-2124 [Accessed: Sep. 7, 2019]

12. "THE GDELT GLOBAL KNOWLEDGE GRAPH (GKG) DATA FORMAT CODEBOOK V 1.0ALPHAEXP (Alpha Release - Experimental)", GDELT Project (2013). http://data.gdeltproject.org/documentation/GDELT-Global_Knowledge_Graph_Codebook.pdf [Accessed: Aug. 26, 2019]

13. "Conference Board of Canada", Wikipedia. https://en.wikipedia.org/wiki/Conference_Board_of_Canada [Accessed: Aug. 26, 2019]

14. "The Conference Board", Wikipedia. https://en.wikipedia.org/wiki/The_Conference_Board [Accessed: Aug. 26, 2019]

15. "GDELT- Global Knowledge Graph Category List", (2014) http://data.gdeltproject.org/documentation/GDELT-Global_Knowledge_Graph_CategoryList.xlsx [Accessed: Aug. 27, 2019]

16. Rouse, M., "What is load shedding?", seachdatacenter.techtarget.com, https://searchdatacenter.techtarget.com/definition/load-shedding [Accessed: Aug. 26, 2016]

17. "Heating oil", Wikipedia, https://en.wikipedia.org/wiki/Heating_oil [Accessed:

Aug. 27, 2019]

18. "Initial Public Offering (IPO)", Investopedia,
    https://www.investopedia.com/terms/i/ipo.asp [Accessed: Aug. 27, 2019]

19. "Memorandum of Understanding (MOU)", Investopedia,
    https://www.investopedia.com/terms/m/mou.asp [Accessed: Aug. 27, 2019]

20. "quit rate", Lexico, https://www.lexico.com/en/definition/quit_rate [Accessed:
    Aug. 27, 2019]

APPENDIX A: SOFTWARE DESIGN WITH JAVA CLASSES:

a. GKGcodes: which contains the relevant Theme codes from GDELT's master list. These are described below in GDELT Themes.

b. GKGrow: This class contains the elements of the given row in the GKG tables. It contains:

1. String DATE, which contains the DATE information of the given entry. This corresponds to the date of the article in the entry itself, not the date it was added to the GDELT database. We currently do not use this value, as we obtain the archives by their date.

2. String NUMARTS, which contains the number of mentions the given url was given in other websites. We currently do not use this value.

3. String COUNTS, which contains a name value for a count of occurrences, such as number of deaths or arrests in each article. We currently do not use this value.

4. String THEMES, which contains the theme tags of the given source. As described above, we currently use this value to segment related articles from irrelevant ones.

5. String LOCATIONS, which contains the locations described in the source article. We use this for finding the macroeconomic sentiment values for countries.

6. String PERSONS, which contains the names of important people in the article, such as state leaders. We currently do not use this value.

7. String ORGANIZATIONS, which contains the companies and organizations described in the article. We currently do not use this value.

8. String TONE, which contains the TONE values of the article. These are the sentiment values calculated automatically by GDELT. As described above, we use this value for

9. String CAMEOEVENTIDS, which contains the CAMEO Event ID codes of the article. These are Ids unique to GDELT that describe the events in the article. Though we have the relevant Event IDs, we currently do not use this value as Themes are sufficient enough.

10.   String SOURCES, which contains the names of the articles. Instead of this, we look the header values by going to the source urls.

11.   String SOURCEURLS, which contains the urls of the articles.

c. GKGsemipro: This class contains the elements of the given row in the GKG tables, but in a semi-processed form. In this class the following values are split into arrays:

1. String THEMES,

2. String LOCATIONS,

3. String PERSONS,

4. String ORGANIZATIONS,

5. String TONE,

6. String CAMEOEVENTIDS,

7. String SOURCES,

The rest of the variables are the same as GKGrow.

d. GKGrowlist: Which is composed of an ArrayList of GKGsemipro and GKGrow classes.

e. GKGprosql: This class contains the GKG values that were processed from the GKG csv file. As we cannot access BigQuery directly, GDELT's csv file

archives had to be used. GKGprosql contains following terms:

1. String[] Themes, which contains the THEME values of GKG data. We use these detect which files are relevant to our needs.

2. String[] Country, which contains the COUNTRY codes of GKG data. We aim to use these for calculating the macroeconomic sentiment values of a given country in a given time period that exists in GKG archives.

3. String[] Sourceurl, which contains the source urls of the given entry. As some entries have multiple urls of the similar articles, we utilize them all.

4. String[] Tone, which is the sentiment values used by GKG and thus by us. They are composed of overall sentiment value, positive sentiment value, negative sentiment value, and polarity.

5. GKGsemipro resource, which is the original resource class of this class.

f. Sentcon: which is the class used for sentiment calculations. It contains:

1. float sentcon, which contains of the final sentiment value of a file or word. This is the sentpos added by the sentneg value.

2. float sentneg, which contains the negative sentiment value of a file or word. Unlike GDELT's negative sentiment value, our sentiment value is stored as a negative float value.

3. Float sentpos, which contains the positive sentiment value of a file or word.

4. Float sentpol, which contains the polarity of a term. GDELT calculates the polarity of a file by finding the sentiment-related terms and divide them to the total amount of words in the file. In our calculation, we do something similar, we divide the amount of times terms in our dictionary appear to the total amount of words in the file.

g. Econterncon: which is the class used for connectivity between words.

1. String term, which is the term itself.

2. float[] connectivity, which is the connection values for words. The array is the same size as the size of the word dictionary, so connectivity values are inserted according to their index number in our dictionary.

3. Int[] conn, which is the simplified connectivity value for the terms. This simplified value only contains the cooccurrence values without the frequencies they appear in the documents, as such it is only initialized after the connectivity values for each file is found.

h. Econcluster: This class contains the sentiment values and words for our word clusters. We form our word clusters by calculating their cooccurance in the articles we obtained from the GDELT GKG database. This contains the following variables:

1. List<String> clusterterms, which is the ArrayList containing our terms in the cluster.

2. Sentcon clustersenti, which is the collective sentiment values of the cluster using the Tone values of GDELT GKG entries.

3. float clusterpos, which is collective positive sentiment value of the cluster using the Tone values of GDELT GKG entries. We use this for its calculations.

4. float clusterneg, which is collective negative sentiment value of the cluster using the Tone values of GDELT GKG entries. Much like cluster pos, we use this for its calculations.

5. Sentcon clustcalci, which is the collective sentiment values of the cluster using our independent sentiment analysis.

6. Float clustercalpos, which is the collective positive sentiment value of the cluster calculated according to our independent sentiment analysis.

7. Float clustercalneg, which is the collective negative sentiment value of the cluster calculated according to our independent sentiment analysis.

This class contains calccluster, which calculates the arithmetic average of the cluster using the terms that were found to exist in the same files and gets their arithmetic average as the cluster's sentiment value.

i. Countsent: This class contains the sentiment values of a given file alongside its country tags as a string array. This class is used to calculate the individual sentiment values of given countries.

1. String [] country, which contains the country tags in the GDELT entries.

2. Sentcon con, which is the calculated sentiment values of the GDELT entry using our sentiment values from our clusters.

3. Sentcon ton, which is the original Tone values obtained from GDELT entries.

j. Econdictionary: This class is where we store our main operations. This class composed of the following variables;

1. String[] econdic, which stores our 2477 terms. This count currently includes duplicates and related terms.

Sentcon[][] ecsent, this stores the sentiment values calculated through our own independent sentiment calculation

2. Sentcon[][] filesen, this stores the sentiment values obtained from the GDELT file.

3. Econtercon[][] ecsenti, this contains the connection values for each term with each other. This 2d array has the size of the length of economic

dictionary x the amount of files processed.

4. Sentcon[]ecfinal, this contains the final calculated sentiment values for each word

5. Sentcon[]filefinal, this contains the final obtained sentiment values for each word.

6. Sentcon[]wordsent, this contains the final calculated sentiment values for each word, without using the word clusters.

7. Sentcon[]wordfile, this contains the final calculated sentiment values for each word, without using the word clusters.

8. List <String> uo, this contains the words that were matched in the text of the article from GDELT's entries.

9. List<Econcluster> list=new ArrayList(), this contains the word clusters.

10. int size, this term contains the number of files that were processed during the calculation. This variable used as not only all entries aren't relevant to us, some of the relevant entries have url sources that have either link issues, or 0 sentiment values (GDELT adds 0 sentiment values for personal pages and advertisements).

This class contains the following methods:

1. calcterncon, which adds the connectivity value of the matching word according to given file number. Connectivity value is taken as an object of Econtercon, which contains the term as a string and a float array of connectivity values.

2. calcse, which adds the sentiment value of the matching word according to given the file number. These sentiment values are stored as Sentcon classes.

3. sentcalc, which adds the matching terms with the given string from the

dictionary to ArrayList uo for the other calculations.

4. printdic, which prints the sentiment values for the dictionary. These values are printed as a txt file and we utilize both the sentiment values we obtained from the GDELT tone values, as well as our simple sentiment analysis.

5. printesent, which prints the output for the cooccurring terms by printing the terms, the terms they are connected to, and their connectivity values.

6. putsen(String string, Sentcon sentcount, int i), which puts the sentiment value for words in the dictionary that match with the given line for the given file number.

7. wordcalc, which calculates the individual sentiment values of the given words by taking the average sentiment values of the files they were present in.

8. setfloat(int size), which sets the Econdictionary's variables according to the amount of files.

9. worldcalcfile, which puts the obtained sentiment value for the given numbered file. Sentiment values are stored as Sentcon classes.

10. worldcalcse, which puts the calculated sentiment value for the given numbered file. Sentiment values are stored as Sentcon classes.

11. printcsvword, which prints the output of our calculations for each file using the sentiment values we associated to our dictionary as a csv file. Details are in Output format section.

12. Printcsvcount, which prints the output of our calculations for each country or territory that exists in the tags we processed with printcsvword as csv file. Details are in the Output format section.

13. ecsentfinal, which calculates the average connectivity values between the terms in the dictionary. These averages are calculated by taking the sum of the connectivity values for each term and divide them by the number of files they appear in.

14. printectern, which prints the output for ecsentfinal in a csv format. The table's headers are: Main term, which covers the term the index corresponds to, Matching term, which corresponds to the term that matches to our main term, and Connectivity, which is the connectivity value between these terms, which is calculated as the average value of the connectivity values between these two terms per document.

15. Newset, which resets the sizes of ecsenti, wordsent, wordfile as the size value if the calculated count of files is larger or smaller than the initial number of files.

16. Cooccurnorm, which simplifies our cooccurrence matrix to a matrix that does not include the frequencies of the words in our dictionary. We do this, as our original co-occurrence matrix counts the amount of times a term appears in all of the documents, not the number of documents it appears. We use this calculation for clustering.

    a. This method gives us both a further simplified matrix, which contains only the terms that appeared in the documents we processed, as well as a distance matrix, that uses this co-occurrence matrix as its source. Our distance matrix uses the co-occurrence value of the given term for each row and subtracts it from other terms' co-occurrence values in that row to give us the distance matrix.

17. Kcluster, which uses the distance matrix to create k-means clusters of terms. These clusters are created as a result of applying the distance matrix's results as the distances between terms, and for this clustering process, we use k random terms as our centroid values. We then cluster according to the shortest distance between the random terms and iterate until each cluster has equal total distance between their terms (we do not use the distances between the clusters themselves, as while such a process is possible, would take much more time).

18. Printcluster, which prints the terms of the clusters we created with kcluster and their average sentiment values. Details are in the Output format below.

k. HTTPCLienttest: Which serves as our main class. This class contains the following methods:

1. getdates, which gets the recent GKG archive files according to the amount of the files that will be obtained. GKG files are stored for each day's entries, as such we look at the links that correspond to nearest days.

2. unzip, which unzips the zipped files obtained from GDELT. Files are stored as xls files.

3. filereaderxls, which reads the xls files obtained from the unzipped folders and puts the given rows of the tables into GKGrow classes and then adds them to GKGrowlist.

4. filesemipro, which reorganizes the GKGrow classes into GKGsemipro classes and adds them to GKGrowlist.

5. fileprinthemes, which selects the relevant entries and puts them into an ArrayList GKGprosql classes. These themes are selected from the GKGcodes class.

6. cleanurls, which eliminates the matching or similar urls by checking their headers from the websites they link to. It also eliminates url sources with 0 tone values.

7. cleantext, which eliminates the matching or similar urls by checking their first 14 lines. This is done by checking the first 14 lines of the file the previous text and the text before the previous text, as GDELT can contain elements between duplicate events.

8. matchingtext, which is the method cleantext uses to match the files. If there are 7 or more lines that match within the first 14 lines between the article before it and the article two articles behind it, then the text is considered a duplicate text.

9. getsent, which uses the remaining files to calculate the sentiment values using either its own sentiment check, or using GDELT's own data, and stores the text files from the links as txt files.

10. sentcount, which calculates the sentiment values of the files using a basic 20 word dictionary that contains common expressions. Negative sentiment value is increased by each time a negative term is matched, and positive sentiment value is increased by each time a positive term is matched with the given line. Final values are then assigned for each word in the file.

11. checkdic, which checks the connections between words using the stored text files. We do this, by constructing a 2d array of "Econtercon" classes, where each "Econtercon" has a float array of the length of the econdic called "connectivity". For each term in the file, the index value of the file gets as much as the times the term appears in the file, and then this array is added to the 2d array for each term that appears in the file, with the first

index being the term's index, and second index being the index of the file.

12. checkword, which checks the sentiment values of the files according to the calculated values for each word in the dictionary. We use this to compare our results to the tone values of GDELT.

13. clustecono, which checks the connectivity values, and creates clusters for each word that has a connection with another. For this reason, this method includes an ArrayList called uol which serves as a checklist for the terms that were covered in previous clusters.

14. clustersent, which calculates the final values for files using the clusters' collective values, by checking matching values. We do this by collecting all of the matching clusters and finding their average positive and negative values, and then adds these up.

15. Clustcont, which calculates the sentiment values for each country using the country tags from GDELT GKG database. We accomplish this by utilizing the values in clustersent, and reorganize them according to the files' country tags, and find the average sentiment values for each country tag available. We initially used the 2-letter ISO2 tags for each country, but as that was insufficient, we used GDELT's own country tags, which includes non-country territories such as "Oceans" and semi-governing regions such as Gibraltar and Hong Kong.

APPENDIX B: OUTPUT FORMAT DETAILS

Our initial output, printed by printdic, for our vocabulary is composed of two parts. This text file presents the outputs of the sentiment values we calculated, and we obtained from GDELT. Sentiment values we obtained are called "sent", and the sentiment values we obtained are called "filesent". These values are printed for each value in their averages.

Our second output, printed by printcsvword, prints a csv file with the following header:

1. "File Number", which is the number of the calculated file.

2. "Original File Result", which is the Tone value of GDELT entry.

3. "Calculated Result", which is the total sentiment value calculated using GDELT data.

4. "Original File Result (Negativity)", which is the negative sentiment Tone from GDELT entry.

5. "Calculated Result (Negativity)", which is the negative sentiment value calculated using the GDELT data.

6. "Original File Result (Positivity)", which is the positive sentiment Tone from GDELT entry.

7. "Calculated Result (Positivity)", which is the positive sentiment value calculated using the GDELT data.

8. "Original File Result (Polarization)", which is the polarity value obtained from GDELT entry.

9. "Calculated Result (Polarization, different calculation than original)", which is the polarity value calculated using the GDELT data and our dictionary. The description refers to the fact that our dictionary is different from the GDELT

dictionary.

Our third output, printed by printcsvcount, prints the average sentiment values for the countries and territories that exist in the locations tags of the documents that were calculated in printcsvword. We accomplish this by using the calculated sentiment values for files and then applying them to the locations we find in the location tags. Our headers are thus;

1. "Location", which is the location tagged in the GDELT entry.

2. "Original File Result (Average)", which is the original average sentiment value of the documents the given location appears in.

3. "Calculated Result", which is the calculated average sentiment value of the documents the given location appears in.

4. "Original File Result (Negativity, Average)", which is the original average negative sentiment value of the documents the given location appears in.

5. "Calculated Result (Negativity)", which is the calculated average negative sentiment value of the documents the given location appears in.

6. "Original File Result (Positivity, Average)", which is the original average positive sentiment value of the documents the given location appears in.

7. "Calculated Result (Positivity)", which is the calculated average positive sentiment value of the documents the given location appears in.

8. "Original File Result (Polarization, Average)", which is the original average polarization value of the documents the location appears in.

9. "Calculated Result (Polarization, different calculation than original)" which is the calculated average polarization value of the documents the location appears in. As our dictionary is different from the GDELT's dictionary, our calculations are different.

Our fourth output for vocabulary, printed by printesent, is composed of two parts, one covering our term and its file, and another which has our connected term, the file it is looked for, and the connectivity value. First the main term is printed, and then the connected term is printed alongside the file it cooccurs with the initial term, and the connectivity value is printed as "ecsenti". Our output is in the txt format.

Our fifth output is a csv file, printed by cooccurnorm, that contains each element in our distance matrix, which is calculated by our trimmed and simplified cooccurrence matrix. This csv file contains the value "i", which is the row of the element, "j", which is the column of the element, the respective terms for I and j, and their distance value. Rather than printing it as a table, we printed these values as a list, because our table was too large to fit the Excel format.

Our sixth output for vocabulary, printed by printectern, is a csv file which has the headers, Main term, Matching term, Matching value. Then our rows are our term, the term it is connected to and their average connectivity value, calculated from all the files.

Our seventh output, printed by printcluster, is a csv file which has the following headers:

1. "Cluster number", which is the index number of the cluster.

2. "Main term", which is the first term of the cluster.

3. "Matching terms", which is the matching terms to the first term.

4. "Cluster Sentiment", which is the average sentiment value of the cluster.

5. "Cluster Negativity", which is the average negative sentiment value of the cluster.

6. "Cluster Positivity", which is the average positive sentiment value of the cluster.

7. "Cluster Polarity", which is the average polarity value of the cluster;

Our eighth output, printed by clustsent, prints a csv file similar to printcsvword, but instead of using the individual sentiment values of the terms, we use the sentiment values of the clusters we gathered, and use the average sentiment values of the clusters that contain the matching terms.

Our ninth output, printed by clustcount, prints a csv file that is list of the matching locations and their average sentiment values, using the values obtained from clustsent and finding the average sentiment values for the countries and territories that exist in the location tags of GDELT.

a. Tables A1.0-A1.3:

Table A1.0: Calculated sentiment values vs. Original tone values

File Number	Original File Result	Calculated Result
0	Original file's tone is: -8.51485148514852	Calculated result is: 2.0820258
1	Original file's tone is: -5.22088353413655	Calculated result is: 2.0050862
2	Original file's tone is: -2.7027027027027	Calculated result is: 2.0306797
3	Original file's tone is: 2.02898550724638	Calculated result is: 2.0788028
4	Original file's tone is: 0.420168067226891	Calculated result is: 2.1741009
5	Original file's tone is: -3.1304347826087	Calculated result is: 1.5508312
6	Original file's tone is: 1.29032258064516	Calculated result is: 1.9892522
7	Original file's tone is: -8.58974358974359	Calculated result is: 2.054471
8	Original file's tone is: 0.965250965250966	Calculated result is: 1.9531429
9	Original file's tone is: -2.60416666666667	Calculated result is: 2.0086756
10	Original file's tone is: 0.519480519480519	Calculated result is: 2.2505715
11	Original file's tone is: -3.98089171974522	Calculated result is: 2.1138678
12	Original file's tone is: -3.93700787401575	Calculated result is: 2.8414235
13	Original file's tone is: -3.24574961360124	Calculated result is: 2.3438342
14	Original file's tone is: 0.91533180778032	Calculated result is: 2.0856097
15	Original file's tone is: -2.54872563718141	Calculated result is: 2.16573
16	Original file's tone is: -0.952380952380952	Calculated result is: 3.085727
17	Original file's tone is: -0.647948164146869	Calculated result is: 2.064246
18	Original file's tone is: -5.87449933244326	Calculated result is: 1.8526478
19	Original file's tone is: 2.05479452054794	Calculated result is: 2.1622648

Table A1.1: Calculated negative sentiment values vs. Original negative tone values

File Number	Original File Result (Negativity)	Calculated Result (Negativity)
0	Original file's negativity is: 9.5049504950495	Calculated negativity is: -0.5668151
1	Original file's negativity is: 8.43373493975904	Calculated negativity is: -0.725307
2	Original file's negativity is: 4.05405405405405	Calculated negativity is: -0.38039628
3	Original file's negativity is: 0.869565217391304	Calculated negativity is: -1.2790592
4	Original file's negativity is: 1.5406162464986	Calculated negativity is: -0.27572516
5	Original file's negativity is: 5.21739130434783	Calculated negativity is: -1.4686843
6	Original file's negativity is: 3.87096774193548	Calculated negativity is: -0.33653414
7	Original file's negativity is: 9.74358974358974	Calculated negativity is: -0.6037601
8	Original file's negativity is: 3.08880308880309	Calculated negativity is: -0.701272
9	Original file's negativity is: 3.99305555555556	Calculated negativity is: -0.43336254
10	Original file's negativity is: 1.81818181818182	Calculated negativity is: -0.5926388
11	Original file's negativity is: 5.57324840764331	Calculated negativity is: -0.31931144
12	Original file's negativity is: 4.7244094488189	Calculated negativity is: -1.5288612
13	Original file's negativity is: 4.94590417310665	Calculated negativity is: -0.4520607
14	Original file's negativity is: 2.97482837528604	Calculated negativity is: -0.53389555
15	Original file's negativity is: 4.94752623688156	Calculated negativity is: -0.7371818
16	Original file's negativity is: 0.952380952380952	Calculated negativity is: -0.46571407
17	Original file's negativity is: 3.67170626349892	Calculated negativity is: -0.3416636
18	Original file's negativity is: 7.61014686248331	Calculated negativity is: -0.5585799
19	Original file's negativity is: 0.228310502283105	Calculated negativity is: -0.5122396

Table A1.2: Calculated positive sentiment values vs. Original positive tone values

File Number	Original File Result (Positivity)	Calculated Result (Positivity)
0	Original file's positivity is: 0.99009900990099	Calculated positivity is: 2.648841
1	Original file's positivity is: 3.21285140562249	Calculated positivity is: 2.7303932
2	Original file's positivity is: 1.35135135135135	Calculated positivity is: 2.411076
3	Original file's positivity is: 2.89855072463768	Calculated positivity is: 3.357862
4	Original file's positivity is: 1.96078431372549	Calculated positivity is: 2.449826
5	Original file's positivity is: 2.08695652173913	Calculated positivity is: 3.0195155
6	Original file's positivity is: 5.16129032258065	Calculated positivity is: 2.3257864
7	Original file's positivity is: 1.15384615384615	Calculated positivity is: 2.658231
8	Original file's positivity is: 4.05405405405405	Calculated positivity is: 2.654415
9	Original file's positivity is: 1.38888888888889	Calculated positivity is: 2.442038
10	Original file's positivity is: 2.33766233766234	Calculated positivity is: 2.8432102
11	Original file's positivity is: 1.59235668789809	Calculated positivity is: 2.4331791
12	Original file's positivity is: 0.78740157480315	Calculated positivity is: 4.3702846
13	Original file's positivity is: 1.70015455950541	Calculated positivity is: 2.7958949
14	Original file's positivity is: 3.89016018306636	Calculated positivity is: 2.6195052
15	Original file's positivity is: 2.39880059970015	Calculated positivity is: 2.9029117
16	Original file's positivity is: 0	Calculated positivity is: 3.551441
17	Original file's positivity is: 3.02375809935205	Calculated positivity is: 2.4059095
18	Original file's positivity is: 1.73564753004005	Calculated positivity is: 2.4112277
19	Original file's positivity is: 2.28310502283105	Calculated positivity is: 2.6745043

Table A1.3: Calculated polarity values vs. Original polarity values

File Number	Original File Result (Polarization)	Calculated Result (Polarization, different calculation than original)
0	Original file's polarity is: -8.51485148514852	Calculated polarity is: 3.215656
1	Original file's polarity is: -5.22088353413655	Calculated polarity is: 3.4557002
2	Original file's polarity is: -2.7027027027027	Calculated polarity is: 2.7914724
3	Original file's polarity is: 2.02898550724638	Calculated polarity is: 4.636921
4	Original file's polarity is: 0.420168067226891	Calculated polarity is: 2.7255511
5	Original file's polarity is: -3.1304347826087	Calculated polarity is: 4.4881997
6	Original file's polarity is: 1.29032258064516	Calculated polarity is: 2.6623206
7	Original file's polarity is: -8.58974358974359	Calculated polarity is: 3.261991
8	Original file's polarity is: 0.965250965250966	Calculated polarity is: 3.355687
9	Original file's polarity is: -2.60416666666667	Calculated polarity is: 2.8754005
10	Original file's polarity is: 0.519480519480519	Calculated polarity is: 3.435849
11	Original file's polarity is: -3.98089171974522	Calculated polarity is: 2.7524905
12	Original file's polarity is: -3.93700787401575	Calculated polarity is: 5.8991456
13	Original file's polarity is: -3.24574961360124	Calculated polarity is: 3.2479556
14	Original file's polarity is: 0.91533180778032	Calculated polarity is: 3.1534007
15	Original file's polarity is: -2.54872563718141	Calculated polarity is: 3.6400933
16	Original file's polarity is: -0.952380952380952	Calculated polarity is: 4.017155
17	Original file's polarity is: -0.647948164146869	Calculated polarity is: 2.7475731
18	Original file's polarity is: -5.87449933244326	Calculated polarity is: 2.9698076
19	Original file's polarity is: 2.05479452054794	Calculated polarity is: 3.1867437

b. Tables A2.0-A2.3:

Table A2.0: Calculated sentiment values for larger dataset vs. Original tone values

File Number	Original File Result	Calculated Result
0	Original file's tone is: -1.93298969072165	Calculated result is: -1.7215965
1	Original file's tone is: -2.90758047767394	Calculated result is: -1.6267359
2	Original file's tone is: -6.95970695970696	Calculated result is: -1.772471
3	Original file's tone is: 0.0496770988574271	Calculated result is: -1.6913092
4	Original file's tone is: -2.11946050096339	Calculated result is: -1.4819407
5	Original file's tone is: -1.16279069767442	Calculated result is: -1.8108361
6	Original file's tone is: 1.45089285714286	Calculated result is: -1.5362127
7	Original file's tone is: -1.86721991701245	Calculated result is: -1.7954762
8	Original file's tone is: -2.4024024024024	Calculated result is: -1.8023677
9	Original file's tone is: -2.08333333333333	Calculated result is: -1.7323928
10	Original file's tone is: 3.41614906832298	Calculated result is: -1.737787
11	Original file's tone is: 2.14723926380368	Calculated result is: -1.8613929
12	Original file's tone is: -4.33070866141732	Calculated result is: -1.7700226
13	Original file's tone is: -3.51288056206089	Calculated result is: -1.8417344
14	Original file's tone is: 3.15091210613599	Calculated result is: -1.8078046
15	Original file's tone is: 1.07334525939177	Calculated result is: -1.7430315
16	Original file's tone is: 0.217627856365615	Calculated result is: -1.7960346
17	Original file's tone is: 3.4647550776583	Calculated result is: -1.8123057
18	Original file's tone is: 2.46305418719212	Calculated result is: -1.8254852
19	Original file's tone is: 2.44215938303342	Calculated result is: -1.7829044
20	Original file's tone is: 3.88007054673721	Calculated result is: -1.6491928
21	Original file's tone is: 1.00671140939597	Calculated result is: -1.7563753
22	Original file's tone is: -1.02543068088597	Calculated result is: -1.6200261
23	Original file's tone is: -0.552486187845304	Calculated result is: -1.9529488
24	Original file's tone is: 3.73831775700935	Calculated result is: -1.8306975
25	Original file's tone is: -1.96531791907514	Calculated result is: -1.3782654
26	Original file's tone is:	Calculated result is: -1.5784509

	0.59676044330758	
27	Original file's tone is: 0.428082191780822	Calculated result is: -1.6221955
28	Original file's tone is: -1.92982456140351	Calculated result is: -1.4550216
29	Original file's tone is: -5.75221238938053	Calculated result is: -1.8280646
30	Original file's tone is: -1.28755364806867	Calculated result is: -1.6918199
31	Original file's tone is: -7.09515859766277	Calculated result is: -1.6208956
32	Original file's tone is: -9.01287553648069	Calculated result is: -1.8382702
33	Original file's tone is: 2.93577981651376	Calculated result is: -1.8439801
34	Original file's tone is: -5.24412296564195	Calculated result is: -1.8805971
35	Original file's tone is: -0.336197235711618	Calculated result is: -1.7185099
36	Original file's tone is: 1.87353629976581	Calculated result is: -1.7949431
37	Original file's tone is: -3.40425531914894	Calculated result is: -1.8819817
38	Original file's tone is: -0.787401574803149	Calculated result is: -1.8477256
39	Original file's tone is: -5.13347022587269	Calculated result is: -1.8544606
40	Original file's tone is: -2.45098039215686	Calculated result is: -1.7768979
41	Original file's tone is: -1.64765525982256	Calculated result is: -1.7850003
42	Original file's tone is: 0.485436893203884	Calculated result is: -1.6981664
43	Original file's tone is: -1.6497461928934	Calculated result is: -1.7061646
44	Original file's tone is: 2.89855072463768	Calculated result is: -1.9124285
45	Original file's tone is: -2.34806629834254	Calculated result is: -1.7408123
46	Original file's tone is: -2.36794171220401	Calculated result is: -1.4789052
47	Original file's tone is: 0.602409638554217	Calculated result is: -1.8223215
48	Original file's tone is: -0.278164116828929	Calculated result is: -1.7615077
49	Original file's tone is: -3.16622691292876	Calculated result is: -1.7621901
50	Original file's tone is: 1.63934426229508	Calculated result is: -1.1077597
51	Original file's tone is: -1.93236714975845	Calculated result is: -1.49225
52	Original file's tone is: 1.26582278481013	Calculated result is: -1.8358157
53	Original file's tone is: 0.710479573712256	Calculated result is: -1.7239552
54	Original file's tone is: 3.30188679245283	Calculated result is: -1.2658672
55	Original file's tone is: -1.11642743221691	Calculated result is: -1.6948936

56	Original file's tone is: 3.23529411764706	Calculated result is: -1.6352253
57	Original file's tone is: -0.530973451327434	Calculated result is: -1.7411563
58	Original file's tone is: 2.86975717439294	Calculated result is: -1.7141519
59	Original file's tone is: -2.12765957446809	Calculated result is: -1.8680677
60	Original file's tone is: -0.13262599469496	Calculated result is: -1.8444707
61	Original file's tone is: -1.94552529182879	Calculated result is: -1.7558961
62	Original file's tone is: -3.16301703163017	Calculated result is: -1.8164055
63	Original file's tone is: 2.06766917293233	Calculated result is: -1.7476285
64	Original file's tone is: -0.111111111111111	Calculated result is: -1.7667935
65	Original file's tone is: -9.48905109489051	Calculated result is: -1.6583676
66	Original file's tone is: -3.47003154574133	Calculated result is: -1.6186152
67	Original file's tone is: 1.83387270765912	Calculated result is: -1.8066106
68	Original file's tone is: -3.3167495854063	Calculated result is: -1.6933765
69	Original file's tone is: -2.54129606099111	Calculated result is: -1.6256154
70	Original file's tone is: 0.273972602739726	Calculated result is: -1.8245022
71	Original file's tone is: 0.221075902726603	Calculated result is: -1.8642603
72	Original file's tone is: 1.97183098591549	Calculated result is: -2.3151639
73	Original file's tone is: 0.28735632183908	Calculated result is: -1.6952288
74	Original file's tone is: -2.5	Calculated result is: -1.8625603
75	Original file's tone is: -0.92899800928998	Calculated result is: -1.7356117
76	Original file's tone is: -1.03896103896104	Calculated result is: -1.8049426
77	Original file's tone is: -3.83561643835616	Calculated result is: -1.9194732
78	Original file's tone is: -1.68243953732913	Calculated result is: -1.8511567
79	Original file's tone is: -0.613496932515337	Calculated result is: -1.6862469
80	Original file's tone is: -3.14814814814815	Calculated result is: -1.2798634
81	Original file's tone is: -0.815217391304348	Calculated result is: -1.853636
82	Original file's tone is: -1.08932461873638	Calculated result is: -1.7914667
83	Original file's tone is: -4.4987146529563	Calculated result is: -1.6057315
84	Original file's tone is: 1.18243243243243	Calculated result is: -1.5394251
85	Original file's tone is: 0.533617929562433	Calculated result is: -1.5437155

86	Original file's tone is: 0.945179584120983	Calculated result is: -1.7723927
87	Original file's tone is: -2.21606648199446	Calculated result is: -1.8423502
88	Original file's tone is: 1.69491525423729	Calculated result is: -1.0488114
89	Original file's tone is: -1.51228733459357	Calculated result is: -1.35203
90	Original file's tone is: -7.72727272727273	Calculated result is: -1.8470678
91	Original file's tone is: 1.78571428571429	Calculated result is: -1.7982612
92	Original file's tone is: 2.02492211838006	Calculated result is: -1.7371502
93	Original file's tone is: 0.493827160493827	Calculated result is: -1.8671858
94	Original file's tone is: -3.10457516339869	Calculated result is: -1.3812466
95	Original file's tone is: -1.77147918511958	Calculated result is: -1.3781362
96	Original file's tone is: 1.57728706624606	Calculated result is: -1.6573398
97	Original file's tone is: 0.254452926208651	Calculated result is: -1.804338
98	Original file's tone is: 3.53302611367127	Calculated result is: -1.7999
99	Original file's tone is: 2.67295597484277	Calculated result is: -1.7445841
100	Original file's tone is: -2.51798561151079	Calculated result is: -1.4828835
101	Original file's tone is: 0.533333333333333	Calculated result is: -1.7592354
102	Original file's tone is: -0.172637030643073	Calculated result is: -1.2989924
103	Original file's tone is: 4.63446475195822	Calculated result is: -1.8439603
104	Original file's tone is: -1.84100418410042	Calculated result is: -1.7561526
105	Original file's tone is: -0.545851528384279	Calculated result is: -1.7887766
106	Original file's tone is: -0.753295668549906	Calculated result is: -1.6675742
107	Original file's tone is: 2.58467023172906	Calculated result is: -1.7795455
108	Original file's tone is: -1.29449838187702	Calculated result is: -1.7523916
109	Original file's tone is: -3.98126463700234	Calculated result is: -1.1532164
110	Original file's tone is: -4.51977401129944	Calculated result is: -1.4358282
111	Original file's tone is: 2.20264317180617	Calculated result is: -1.8497951
112	Original file's tone is: -6.18556701030928	Calculated result is: -1.4378593
113	Original file's tone is: 1.74563591022444	Calculated result is: -1.8976407
114	Original file's tone is: 1.47954743255004	Calculated result is: -1.6456156
115	Original file's tone is:	Calculated result is: -1.7168207

	3.99274047186933	
116	Original file's tone is: 2.0108275328693	Calculated result is: -1.8057971
117	Original file's tone is: 1.90114068441065	Calculated result is: -1.5580502
118	Original file's tone is: 0.0354861603974452	Calculated result is: -1.5795722
119	Original file's tone is: -7.04387990762125	Calculated result is: -1.6952264
120	Original file's tone is: 2.8735632183908	Calculated result is: -1.7845914
121	Original file's tone is: -4.10654827968923	Calculated result is: -1.5302222
122	Original file's tone is: -1.41935483870968	Calculated result is: -1.4356003
123	Original file's tone is: -0.3003003003003	Calculated result is: -1.703109
124	Original file's tone is: -2.02429149797571	Calculated result is: -1.7279959
125	Original file's tone is: 0.12531328320802	Calculated result is: -1.6396921
126	Original file's tone is: 1.74563591022444	Calculated result is: -1.6459048
127	Original file's tone is: -1.61290322580645	Calculated result is: -1.6728053
128	Original file's tone is: 2.47619047619048	Calculated result is: -1.6648948
129	Original file's tone is: -1.13122171945701	Calculated result is: -1.4431839
130	Original file's tone is: -0.308641975308642	Calculated result is: -1.8153923
131	Original file's tone is: -6.29370629370629	Calculated result is: -1.9967263
132	Original file's tone is: -5.1440329218107	Calculated result is: -1.6264687
133	Original file's tone is: -10.2803738317757	Calculated result is: -1.7243752
134	Original file's tone is: -2.79955207166853	Calculated result is: -1.6850257
135	Original file's tone is: -1.08262108262108	Calculated result is: -1.7236288
136	Original file's tone is: -5.48302872062663	Calculated result is: -1.7163634
137	Original file's tone is: 1.59045725646123	Calculated result is: -1.8466839
138	Original file's tone is: -2.56410256410256	Calculated result is: -1.366863
139	Original file's tone is: 0.735294117647059	Calculated result is: -1.8769879
140	Original file's tone is: 3.2967032967033	Calculated result is: -1.2523978
141	Original file's tone is: -2.6431718061674	Calculated result is: -1.7703424
142	Original file's tone is: 0.874125874125874	Calculated result is: -1.8589338
143	Original file's tone is: -1.69628432956381	Calculated result is: -1.7153647
144	Original file's tone is: -1.69491525423729	Calculated result is: -1.0622616

145	Original file's tone is: -0.853889943074004	Calculated result is: -1.7088826
146	Original file's tone is: -1.61812297734628	Calculated result is: -1.8336823
147	Original file's tone is: -1.5748031496063	Calculated result is: -1.6410911
148	Original file's tone is: -2.6431718061674	Calculated result is: -1.8245239
149	Original file's tone is: -0.409836065573771	Calculated result is: -1.7875245
150	Original file's tone is: 3.41151385927505	Calculated result is: -1.8017747
151	Original file's tone is: -5.52995391705069	Calculated result is: -1.8450522
152	Original file's tone is: -2.07100591715976	Calculated result is: -1.8191538
153	Original file's tone is: -1.1070110701107	Calculated result is: -1.7170918
154	Original file's tone is: -4.38247011952191	Calculated result is: -1.8251872
155	Original file's tone is: 3.84615384615385	Calculated result is: -1.8881128
156	Original file's tone is: -0.432900432900433	Calculated result is: -1.7120636
157	Original file's tone is: -5.60344827586207	Calculated result is: -1.831363
158	Original file's tone is: 3.486529318542	Calculated result is: -1.7660971
159	Original file's tone is: 2.74914089347079	Calculated result is: -1.8242989
160	Original file's tone is: 6.61322645290581	Calculated result is: -1.8682395
161	Original file's tone is: -0.490196078431373	Calculated result is: -1.532877
162	Original file's tone is: -2.56529850746269	Calculated result is: -1.6159534
163	Original file's tone is: 1.54617634768074	Calculated result is: -1.4705615
164	Original file's tone is: 0.823045267489712	Calculated result is: -1.5902824
165	Original file's tone is: -4.2563845768653	Calculated result is: -1.7696455
166	Original file's tone is: -4.09090909090909	Calculated result is: -1.585042
167	Original file's tone is: -1.19225037257824	Calculated result is: -1.8022969
168	Original file's tone is: 1.61662817551963	Calculated result is: -1.7127564
169	Original file's tone is: -0.177935943060498	Calculated result is: -1.7395093
170	Original file's tone is: 3.28282828282828	Calculated result is: -1.8501878
171	Original file's tone is: -1.57303370786517	Calculated result is: -1.767448
172	Original file's tone is: -3.00353356890459	Calculated result is: -1.7189476
173	Original file's tone is: 2.64255910987483	Calculated result is: -1.7393868
174	Original file's tone is:	Calculated result is: -1.6487806

	2.1356783919598	
175	Original file's tone is: -1.10062893081761	Calculated result is: -1.7562268
176	Original file's tone is: -2.94117647058824	Calculated result is: -1.5578802
177	Original file's tone is: -1.91780821917808	Calculated result is: -1.6839867

Table A2.1: Calculated negative sentiment values for larger dataset vs. Original tone

values

File Number	Original File Result (Negativity)	Calculated Result (Negativity)
0	Original file's negativity is: 4.7680412371134	Calculated negativity is: -3.9450452
1	Original file's negativity is: 5.81516095534787	Calculated negativity is: -4.0306535
2	Original file's negativity is: 8.60805860805861	Calculated negativity is: -3.9734616
3	Original file's negativity is: 2.0367610531545	Calculated negativity is: -4.066751
4	Original file's negativity is: 4.17469492614001	Calculated negativity is: -4.049916
5	Original file's negativity is: 2.03488372093023	Calculated negativity is: -3.863818
6	Original file's negativity is: 0.111607142857143	Calculated negativity is: -4.1414485
7	Original file's negativity is: 4.149377593361	Calculated negativity is: -3.8411942
8	Original file's negativity is: 3.3033033033033	Calculated negativity is: -3.8749514
9	Original file's negativity is: 4.76190476190476	Calculated negativity is: -3.846437
10	Original file's negativity is: 0.62111801242236	Calculated negativity is: -3.9907389
11	Original file's negativity is: 2.14723926380368	Calculated negativity is: -3.8602362
12	Original file's negativity is: 6.49606299212598	Calculated negativity is: -3.7976077
13	Original file's negativity is: 6.08899297423888	Calculated negativity is: -3.8771992
14	Original file's negativity is: 0.66334991708126	Calculated negativity is: -3.9966414
15	Original file's negativity is: 1.25223613595707	Calculated negativity is: -4.6030273
16	Original file's negativity is: 3.04678998911861	Calculated negativity is: -3.9415865
17	Original file's negativity is: 1.4336917562724	Calculated negativity is: -3.8770635
18	Original file's negativity is: 0.492610837438424	Calculated negativity is: -3.8257594
19	Original file's negativity is: 2.95629820051414	Calculated negativity is: -3.8477733
20	Original file's negativity is: 1.23456790123457	Calculated negativity is: -3.9102411
21	Original file's negativity is: 3.69127516778524	Calculated negativity is: -4.1297884
22	Original file's negativity is: 3.99917965545529	Calculated negativity is: -4.1780367
23	Original file's negativity is: 3.59116022099448	Calculated negativity is: -4.1040707
24	Original file's negativity is: 1.51869158878505	Calculated negativity is: -3.844997
25	Original file's negativity is: 3.8150289017341	Calculated negativity is: -4.4555836
26	Original file's negativity is:	Calculated negativity is: -

	2.21653878942881	4.2165923
27	Original file's negativity is: 0.513698630136986	Calculated negativity is: -4.1955504
28	Original file's negativity is: 4.21052631578947	Calculated negativity is: -3.9938767
29	Original file's negativity is: 7.30088495575221	Calculated negativity is: -3.791578
30	Original file's negativity is: 3.00429184549356	Calculated negativity is: -3.8411639
31	Original file's negativity is: 10.016694490818	Calculated negativity is: -4.1949906
32	Original file's negativity is: 12.0171673819742	Calculated negativity is: -3.8307455
33	Original file's negativity is: 0.36697247706422	Calculated negativity is: -3.8244634
34	Original file's negativity is: 5.78661844484629	Calculated negativity is: -3.9551144
35	Original file's negativity is: 4.48262980948823	Calculated negativity is: -3.9517965
36	Original file's negativity is: 1.63934426229508	Calculated negativity is: -3.8415728
37	Original file's negativity is: 4.8936170212766	Calculated negativity is: -3.7442834
38	Original file's negativity is: 4.92125984251969	Calculated negativity is: -3.8567512
39	Original file's negativity is: 6.16016427104723	Calculated negativity is: -3.796451
40	Original file's negativity is: 4.41176470588235	Calculated negativity is: -3.817896
41	Original file's negativity is: 2.53485424588086	Calculated negativity is: -3.8445551
42	Original file's negativity is: 2.42718446601942	Calculated negativity is: -4.1547556
43	Original file's negativity is: 3.68020304568528	Calculated negativity is: -3.882855
44	Original file's negativity is: 0	Calculated negativity is: -3.7259119
45	Original file's negativity is: 4.55801104972376	Calculated negativity is: -4.0200424
46	Original file's negativity is: 3.82513661202186	Calculated negativity is: -4.3894553
47	Original file's negativity is: 2.40963855421687	Calculated negativity is: -3.816038
48	Original file's negativity is: 4.3115438108484	Calculated negativity is: -3.9305537
49	Original file's negativity is: 3.95778364116095	Calculated negativity is: -3.9032586
50	Original file's negativity is: 0.655737704918033	Calculated negativity is: -4.950142
51	Original file's negativity is: 4.10628019323672	Calculated negativity is: -4.029361
52	Original file's negativity is: 2.0253164556962	Calculated negativity is: -3.84274
53	Original file's negativity is: 3.55239786856128	Calculated negativity is: -4.0438275
54	Original file's negativity is: 2.35849056603774	Calculated negativity is: -4.3493156
55	Original file's negativity is: 3.34928229665072	Calculated negativity is: -3.9257598

56	Original file's negativity is: 2.05882352941176	Calculated negativity is: -4.0672455
57	Original file's negativity is: 4.7787610619469	Calculated negativity is: -3.9250975
58	Original file's negativity is: 0.441501103752759	Calculated negativity is: -3.904578
59	Original file's negativity is: 3.19148936170213	Calculated negativity is: -3.9079287
60	Original file's negativity is: 2.91777188328912	Calculated negativity is: -3.8484468
61	Original file's negativity is: 3.50194552529183	Calculated negativity is: -3.8781495
62	Original file's negativity is: 3.64963503649635	Calculated negativity is: -3.8425024
63	Original file's negativity is: 0	Calculated negativity is: -3.8353527
64	Original file's negativity is: 3.22222222222222	Calculated negativity is: -4.067042
65	Original file's negativity is: 9.48905109489051	Calculated negativity is: -4.086741
66	Original file's negativity is: 5.99369085173502	Calculated negativity is: -4.0719676
67	Original file's negativity is: 2.48112189859763	Calculated negativity is: -3.794835
68	Original file's negativity is: 5.30679933665008	Calculated negativity is: -4.0093937
69	Original file's negativity is: 4.44726810673443	Calculated negativity is: -4.1967087
70	Original file's negativity is: 1.0958904109589	Calculated negativity is: -3.8425374
71	Original file's negativity is: 0.957995578481945	Calculated negativity is: -3.8284175
72	Original file's negativity is: 2.53521126760563	Calculated negativity is: -5.19875
73	Original file's negativity is: 2.1551724137931	Calculated negativity is: -4.1486406
74	Original file's negativity is: 3.46153846153846	Calculated negativity is: -3.9285617
75	Original file's negativity is: 3.91506303915063	Calculated negativity is: -4.062572
76	Original file's negativity is: 3.50649350649351	Calculated negativity is: -3.8329742
77	Original file's negativity is: 4.10958904109589	Calculated negativity is: -4.5073385
78	Original file's negativity is: 3.89064143007361	Calculated negativity is: -3.800733
79	Original file's negativity is: 3.06748466257669	Calculated negativity is: -3.9027276
80	Original file's negativity is: 5.74074074074074	Calculated negativity is: -4.0812974
81	Original file's negativity is: 2.98913043478261	Calculated negativity is: -3.8401172
82	Original file's negativity is: 2.61437908496732	Calculated negativity is: -4.0263815
83	Original file's negativity is: 8.3547557840617	Calculated negativity is: -4.0003467
84	Original file's negativity is: 0.675675675675676	Calculated negativity is: -4.3411994
85	Original file's negativity is:	Calculated negativity is: -

	1.06723585912487	4.1862235
86	Original file's negativity is: 2.93005671077505	Calculated negativity is: -3.8937545
87	Original file's negativity is: 3.8781163434903	Calculated negativity is: -3.8997183
88	Original file's negativity is: 0.651890482398957	Calculated negativity is: -4.3135624
89	Original file's negativity is: 3.21361058601134	Calculated negativity is: -4.051097
90	Original file's negativity is: 9.09090909090909	Calculated negativity is: -3.823367
91	Original file's negativity is: 0.446428571428571	Calculated negativity is: -3.9696565
92	Original file's negativity is: 1.09034267912773	Calculated negativity is: -3.9267368
93	Original file's negativity is: 1.48148148148148	Calculated negativity is: -4.1916385
94	Original file's negativity is: 5.06535947712418	Calculated negativity is: -4.318384
95	Original file's negativity is: 4.162976085031	Calculated negativity is: -4.3356867
96	Original file's negativity is: 1.57728706624606	Calculated negativity is: -4.109063
97	Original file's negativity is: 3.05343511450382	Calculated negativity is: -3.8115993
98	Original file's negativity is: 1.53609831029186	Calculated negativity is: -3.9662833
99	Original file's negativity is: 1.10062893081761	Calculated negativity is: -3.8562753
100	Original file's negativity is: 4.31654676258993	Calculated negativity is: -4.1593018
101	Original file's negativity is: 1.33333333333333	Calculated negativity is: -3.9454443
102	Original file's negativity is: 3.6685369011653	Calculated negativity is: -4.0644503
103	Original file's negativity is: 1.43603133159269	Calculated negativity is: -3.7787638
104	Original file's negativity is: 3.76569037656904	Calculated negativity is: -4.0371814
105	Original file's negativity is: 1.52838427947598	Calculated negativity is: -3.8886473
106	Original file's negativity is: 3.07595731324545	Calculated negativity is: -4.114437
107	Original file's negativity is: 1.15864527629234	Calculated negativity is: -3.8126984
108	Original file's negativity is: 3.17152103559871	Calculated negativity is: -3.9001207
109	Original file's negativity is: 6.08899297423888	Calculated negativity is: -4.461862
110	Original file's negativity is: 7.34463276836158	Calculated negativity is: -4.0417533
111	Original file's negativity is: 1.3215859030837	Calculated negativity is: -3.8726113
112	Original file's negativity is: 6.18556701030928	Calculated negativity is: -4.160269
113	Original file's negativity is: 1.49625935162095	Calculated negativity is: -3.7864156
114	Original file's negativity is: 2.69799825935596	Calculated negativity is: -4.0266533

115	Original file's negativity is: 1.4519056261343	Calculated negativity is: - 3.9784703
116	Original file's negativity is: 1.85614849187935	Calculated negativity is: - 3.8362796
117	Original file's negativity is: 0.760456273764259	Calculated negativity is: - 3.928366
118	Original file's negativity is: 3.1582682753726	Calculated negativity is: - 3.9470544
119	Original file's negativity is: 9.46882217090069	Calculated negativity is: - 3.9214907
120	Original file's negativity is: 2.75862068965517	Calculated negativity is: - 3.891588
121	Original file's negativity is: 7.32519422863485	Calculated negativity is: - 3.906094
122	Original file's negativity is: 3.61290322580645	Calculated negativity is: - 4.3945913
123	Original file's negativity is: 1.5015015015015	Calculated negativity is: - 3.8735533
124	Original file's negativity is: 3.64372469635628	Calculated negativity is: - 4.065532
125	Original file's negativity is: 3.38345864661654	Calculated negativity is: - 3.925884
126	Original file's negativity is: 0.997506234413965	Calculated negativity is: - 3.9922676
127	Original file's negativity is: 3.2258064516129	Calculated negativity is: - 3.9507172
128	Original file's negativity is: 1.33333333333333	Calculated negativity is: - 4.223694
129	Original file's negativity is: 2.26244343891403	Calculated negativity is: - 3.9738095
130	Original file's negativity is: 3.08641975308642	Calculated negativity is: - 3.8159332
131	Original file's negativity is: 7.16783216783217	Calculated negativity is: - 4.265844
132	Original file's negativity is: 6.68724279835391	Calculated negativity is: - 3.9366531
133	Original file's negativity is: 11.9626168224299	Calculated negativity is: - 3.8990457
134	Original file's negativity is: 3.58342665173572	Calculated negativity is: - 3.999698
135	Original file's negativity is: 3.58974358974359	Calculated negativity is: - 4.124714
136	Original file's negativity is: 6.39686684073107	Calculated negativity is: - 3.9003396
137	Original file's negativity is: 1.59045725646123	Calculated negativity is: - 3.801336
138	Original file's negativity is: 4.3956043956044	Calculated negativity is: - 4.2869444
139	Original file's negativity is: 0	Calculated negativity is: - 4.270624
140	Original file's negativity is: 1.0989010989011	Calculated negativity is: - 3.9602473
141	Original file's negativity is: 3.9647577092511	Calculated negativity is: - 3.9619396
142	Original file's negativity is: 4.37062937062937	Calculated negativity is: - 3.8121572
143	Original file's negativity is: 5.81583198707593	Calculated negativity is: - 3.896603
144	Original file's negativity is:	Calculated negativity is: -

	4.14312617702448	4.2117386
145	Original file's negativity is: 3.98481973434535	Calculated negativity is: - 3.9823716
146	Original file's negativity is: 6.47249190938511	Calculated negativity is: - 3.860935
147	Original file's negativity is: 3.54330708661417	Calculated negativity is: - 4.0109186
148	Original file's negativity is: 3.08370044052863	Calculated negativity is: - 3.8513288
149	Original file's negativity is: 1.22950819672131	Calculated negativity is: - 3.9070678
150	Original file's negativity is: 1.27931769722815	Calculated negativity is: - 3.8241384
151	Original file's negativity is: 7.48847926267281	Calculated negativity is: - 4.1026154
152	Original file's negativity is: 3.25443786982249	Calculated negativity is: - 3.8057988
153	Original file's negativity is: 2.9520295202952	Calculated negativity is: - 3.8912582
154	Original file's negativity is: 7.17131474103586	Calculated negativity is: - 3.9422772
155	Original file's negativity is: 0.678733031674208	Calculated negativity is: - 3.7754965
156	Original file's negativity is: 2.5974025974026	Calculated negativity is: - 3.8918014
157	Original file's negativity is: 7.75862068965517	Calculated negativity is: - 3.7410042
158	Original file's negativity is: 0	Calculated negativity is: - 4.1330466
159	Original file's negativity is: 1.03092783505155	Calculated negativity is: - 3.9990613
160	Original file's negativity is: 0.200400801603206	Calculated negativity is: - 3.8299465
161	Original file's negativity is: 2.20588235294118	Calculated negativity is: - 4.144668
162	Original file's negativity is: 5.92350746268657	Calculated negativity is: - 4.0051575
163	Original file's negativity is: 1.29544504805683	Calculated negativity is: - 4.093984
164	Original file's negativity is: 0.411522633744856	Calculated negativity is: - 3.8636212
165	Original file's negativity is: 5.35803705558337	Calculated negativity is: - 3.9655797
166	Original file's negativity is: 7.12121212121212	Calculated negativity is: - 4.180957
167	Original file's negativity is: 5.36512667660209	Calculated negativity is: - 3.8247821
168	Original file's negativity is: 2.3094688221709	Calculated negativity is: - 3.9484541
169	Original file's negativity is: 1.95729537366548	Calculated negativity is: - 4.010098
170	Original file's negativity is: 3.03030303030303	Calculated negativity is: - 3.8991873
171	Original file's negativity is: 3.82022471910112	Calculated negativity is: - 4.020044
172	Original file's negativity is: 5.8303886925795	Calculated negativity is: - 3.8238947
173	Original file's negativity is: 1.25173852573018	Calculated negativity is: - 3.9256258

174	Original file's negativity is: 0.753768844221105	Calculated negativity is: -4.1181893
175	Original file's negativity is: 3.61635220125786	Calculated negativity is: -3.9220219
176	Original file's negativity is: 5.3475935828877	Calculated negativity is: -4.333268
177	Original file's negativity is: 3.6986301369863	Calculated negativity is: -3.9070914

Table A2.2: Calculated positive sentiment values for larger dataset vs. Original tone

values

File Number	Original File Result (Positivity)	Calculated Result (Positivity)
0	Original file's positivity is: 2.83505154639175	Calculated positivity is: 2.2234488
1	Original file's positivity is: 2.90758047767394	Calculated positivity is: 2.4039176
2	Original file's positivity is: 1.64835164835165	Calculated positivity is: 2.2009907
3	Original file's positivity is: 2.08643815201192	Calculated positivity is: 2.3754418
4	Original file's positivity is: 2.05523442517662	Calculated positivity is: 2.567975
5	Original file's positivity is: 0.872093023255814	Calculated positivity is: 2.0529819
6	Original file's positivity is: 1.5625	Calculated positivity is: 2.6052358
7	Original file's positivity is: 2.28215767634855	Calculated positivity is: 2.045718
8	Original file's positivity is: 0.900900900900901	Calculated positivity is: 2.0725837
9	Original file's positivity is: 2.67857142857143	Calculated positivity is: 2.1140442
10	Original file's positivity is: 4.03726708074534	Calculated positivity is: 2.2529519
11	Original file's positivity is: 4.29447852760736	Calculated positivity is: 1.9988433
12	Original file's positivity is: 2.16535433070866	Calculated positivity is: 2.027585
13	Original file's positivity is: 2.57611241217799	Calculated positivity is: 2.0354648
14	Original file's positivity is: 3.81426202321725	Calculated positivity is: 2.1888368
15	Original file's positivity is: 2.32558139534884	Calculated positivity is: 2.8599958
16	Original file's positivity is: 3.26441784548422	Calculated positivity is: 2.145552
17	Original file's positivity is: 4.8984468339307	Calculated positivity is: 2.0647578
18	Original file's positivity is: 2.95566502463054	Calculated positivity is: 2.0002742
19	Original file's positivity is: 5.39845758354756	Calculated positivity is: 2.064869
20	Original file's positivity is: 5.11463844797178	Calculated positivity is: 2.2610483
21	Original file's positivity is: 4.69798657718121	Calculated positivity is: 2.373413
22	Original file's positivity is: 2.97374897456932	Calculated positivity is: 2.5580106
23	Original file's positivity is: 3.03867403314917	Calculated positivity is: 2.1511219
24	Original file's positivity is: 5.25700934579439	Calculated positivity is: 2.0142994
25	Original file's positivity is: 1.84971098265896	Calculated positivity is: 3.0773182
26	Original file's positivity is:	Calculated positivity is:

	2.81329923273657	2.6381414
27	Original file's positivity is: 0.941780821917808	Calculated positivity is: 2.573355
28	Original file's positivity is: 2.28070175438596	Calculated positivity is: 2.538855
29	Original file's positivity is: 1.54867256637168	Calculated positivity is: 1.9635135
30	Original file's positivity is: 1.71673819742489	Calculated positivity is: 2.149344
31	Original file's positivity is: 2.92153589315526	Calculated positivity is: 2.574095
32	Original file's positivity is: 3.00429184549356	Calculated positivity is: 1.9924753
33	Original file's positivity is: 3.30275229357798	Calculated positivity is: 1.9804833
34	Original file's positivity is: 0.54249547920434	Calculated positivity is: 2.0745173
35	Original file's positivity is: 4.14643257377662	Calculated positivity is: 2.2332866
36	Original file's positivity is: 3.51288056206089	Calculated positivity is: 2.0466297
37	Original file's positivity is: 1.48936170212766	Calculated positivity is: 1.8623017
38	Original file's positivity is: 4.13385826771654	Calculated positivity is: 2.0090256
39	Original file's positivity is: 1.02669404517454	Calculated positivity is: 1.9419905
40	Original file's positivity is: 1.96078431372549	Calculated positivity is: 2.040998
41	Original file's positivity is: 0.887198986058302	Calculated positivity is: 2.0595548
42	Original file's positivity is: 2.9126213592233	Calculated positivity is: 2.4565892
43	Original file's positivity is: 2.03045685279188	Calculated positivity is: 2.1766903
44	Original file's positivity is: 2.89855072463768	Calculated positivity is: 1.8134834
45	Original file's positivity is: 2.20994475138122	Calculated positivity is: 2.27923
46	Original file's positivity is: 1.45719489981785	Calculated positivity is: 2.91055
47	Original file's positivity is: 3.01204819277108	Calculated positivity is: 1.9937164
48	Original file's positivity is: 4.03337969401947	Calculated positivity is: 2.169046
49	Original file's positivity is: 0.79155672823219	Calculated positivity is: 2.1410685
50	Original file's positivity is: 2.29508196721311	Calculated positivity is: 3.8423822
51	Original file's positivity is: 2.17391304347826	Calculated positivity is: 2.5371108
52	Original file's positivity is: 3.29113924050633	Calculated positivity is: 2.0069244
53	Original file's positivity is: 4.26287744227353	Calculated positivity is: 2.3198724
54	Original file's positivity is: 5.66037735849057	Calculated positivity is: 3.0834484
55	Original file's positivity is: 2.23285486443381	Calculated positivity is: 2.2308662

56	Original file's positivity is: 5.29411764705882	Calculated positivity is: 2.4320202
57	Original file's positivity is: 4.24778761061947	Calculated positivity is: 2.1839411
58	Original file's positivity is: 3.3112582781457	Calculated positivity is: 2.190426
59	Original file's positivity is: 1.06382978723404	Calculated positivity is: 2.039861
60	Original file's positivity is: 2.78514588859416	Calculated positivity is: 2.003976
61	Original file's positivity is: 1.55642023346304	Calculated positivity is: 2.1222534
62	Original file's positivity is: 0.48661800486618	Calculated positivity is: 2.0260968
63	Original file's positivity is: 2.06766917293233	Calculated positivity is: 2.0877242
64	Original file's positivity is: 3.11111111111111	Calculated positivity is: 2.3002484
65	Original file's positivity is: 0	Calculated positivity is: 2.4283733
66	Original file's positivity is: 2.52365930599369	Calculated positivity is: 2.4533525
67	Original file's positivity is: 4.31499460625674	Calculated positivity is: 1.9882245
68	Original file's positivity is: 1.99004975124378	Calculated positivity is: 2.3160172
69	Original file's positivity is: 1.90597204574333	Calculated positivity is: 2.5710933
70	Original file's positivity is: 1.36986301369863	Calculated positivity is: 2.0180352
71	Original file's positivity is: 1.17907148120855	Calculated positivity is: 1.9641572
72	Original file's positivity is: 4.50704225352113	Calculated positivity is: 2.8835862
73	Original file's positivity is: 2.44252873563218	Calculated positivity is: 2.4534118
74	Original file's positivity is: 0.961538461538462	Calculated positivity is: 2.0660014
75	Original file's positivity is: 2.98606502986065	Calculated positivity is: 2.3269603
76	Original file's positivity is: 2.46753246753247	Calculated positivity is: 2.0280316
77	Original file's positivity is: 0.273972602739726	Calculated positivity is: 2.5878654
78	Original file's positivity is: 2.20820189274448	Calculated positivity is: 1.9495764
79	Original file's positivity is: 2.45398773006135	Calculated positivity is: 2.2164807
80	Original file's positivity is: 2.59259259259259	Calculated positivity is: 2.801434
81	Original file's positivity is: 2.17391304347826	Calculated positivity is: 1.9864812
82	Original file's positivity is: 1.52505446623094	Calculated positivity is: 2.2349148
83	Original file's positivity is: 3.8560411311054	Calculated positivity is: 2.3946152
84	Original file's positivity is: 1.85810810810811	Calculated positivity is: 2.8017743
85	Original file's positivity is:	Calculated positivity is: 2.642508

		1.6008537886873	
86	Original file's positivity is: 3.87523629489603		Calculated positivity is: 2.1213617
87	Original file's positivity is: 1.66204986149584		Calculated positivity is: 2.057368
88	Original file's positivity is: 2.34680573663625		Calculated positivity is: 3.264751
89	Original file's positivity is: 1.70132325141777		Calculated positivity is: 2.6990669
90	Original file's positivity is: 1.36363636363636		Calculated positivity is: 1.9762993
91	Original file's positivity is: 2.23214285714286		Calculated positivity is: 2.1713953
92	Original file's positivity is: 3.11526479750779		Calculated positivity is: 2.1895866
93	Original file's positivity is: 1.97530864197531		Calculated positivity is: 2.3244526
94	Original file's positivity is: 1.96078431372549		Calculated positivity is: 2.9371376
95	Original file's positivity is: 2.39149689991143		Calculated positivity is: 2.9575505
96	Original file's positivity is: 3.15457413249211		Calculated positivity is: 2.4517233
97	Original file's positivity is: 3.30788804071247		Calculated positivity is: 2.0072613
98	Original file's positivity is: 5.06912442396313		Calculated positivity is: 2.1663833
99	Original file's positivity is: 3.77358490566038		Calculated positivity is: 2.1116912
100	Original file's positivity is: 1.79856115107914		Calculated positivity is: 2.6764183
101	Original file's positivity is: 1.86666666666667		Calculated positivity is: 2.186209
102	Original file's positivity is: 3.49589987052223		Calculated positivity is: 2.7654579
103	Original file's positivity is: 6.07049608355091		Calculated positivity is: 1.9348035
104	Original file's positivity is: 1.92468619246862		Calculated positivity is: 2.2810287
105	Original file's positivity is: 0.982532751091703		Calculated positivity is: 2.0998707
106	Original file's positivity is: 2.32266164469554		Calculated positivity is: 2.446863
107	Original file's positivity is: 3.74331550802139		Calculated positivity is: 2.0331528
108	Original file's positivity is: 1.87702265372168		Calculated positivity is: 2.1477292
109	Original file's positivity is: 2.10772833723653		Calculated positivity is: 3.3086457
110	Original file's positivity is: 2.82485875706215		Calculated positivity is: 2.605925
111	Original file's positivity is: 3.52422907488987		Calculated positivity is: 2.0228162
112	Original file's positivity is: 0		Calculated positivity is: 2.7224095
113	Original file's positivity is: 3.24189526184539		Calculated positivity is: 1.8887749
114	Original file's positivity is: 4.17754569190601		Calculated positivity is: 2.3810377

115	Original file's positivity is: 5.44464609800363	Calculated positivity is: 2.2616496
116	Original file's positivity is: 3.86697602474865	Calculated positivity is: 2.0304825
117	Original file's positivity is: 2.6615969581749	Calculated positivity is: 2.3703158
118	Original file's positivity is: 3.19375443577005	Calculated positivity is: 2.3674822
119	Original file's positivity is: 2.42494226327945	Calculated positivity is: 2.2262642
120	Original file's positivity is: 5.63218390804598	Calculated positivity is: 2.1069965
121	Original file's positivity is: 3.21864594894562	Calculated positivity is: 2.375872
122	Original file's positivity is: 2.19354838709677	Calculated positivity is: 2.958991
123	Original file's positivity is: 1.2012012012012	Calculated positivity is: 2.1704443
124	Original file's positivity is: 1.61943319838057	Calculated positivity is: 2.3375363
125	Original file's positivity is: 3.50877192982456	Calculated positivity is: 2.286192
126	Original file's positivity is: 2.7431421446384	Calculated positivity is: 2.3463628
127	Original file's positivity is: 1.61290322580645	Calculated positivity is: 2.277912
128	Original file's positivity is: 3.80952380952381	Calculated positivity is: 2.558799
129	Original file's positivity is: 1.13122171945701	Calculated positivity is: 2.5306256
130	Original file's positivity is: 2.77777777777778	Calculated positivity is: 2.000541
131	Original file's positivity is: 0.874125874125874	Calculated positivity is: 2.2691176
132	Original file's positivity is: 1.54320987654321	Calculated positivity is: 2.3101845
133	Original file's positivity is: 1.68224299065421	Calculated positivity is: 2.1746705
134	Original file's positivity is: 0.783874580067189	Calculated positivity is: 2.3146722
135	Original file's positivity is: 2.50712250712251	Calculated positivity is: 2.4010851
136	Original file's positivity is: 0.913838120104439	Calculated positivity is: 2.1839762
137	Original file's positivity is: 3.18091451292246	Calculated positivity is: 1.9546522
138	Original file's positivity is: 1.83150183150183	Calculated positivity is: 2.9200814
139	Original file's positivity is: 0.735294117647059	Calculated positivity is: 2.3936362
140	Original file's positivity is: 4.3956043956044	Calculated positivity is: 2.7078495
141	Original file's positivity is: 1.3215859030837	Calculated positivity is: 2.1915972
142	Original file's positivity is: 5.24475524475525	Calculated positivity is: 1.9532233
143	Original file's positivity is: 4.11954765751212	Calculated positivity is: 2.1812384
144	Original file's positivity is:	Calculated positivity is: 3.149477

	2.44821092278719	
145	Original file's positivity is: 3.13092979127135	Calculated positivity is: 2.273489
146	Original file's positivity is: 4.85436893203883	Calculated positivity is: 2.0272527
147	Original file's positivity is: 1.96850393700787	Calculated positivity is: 2.3698275
148	Original file's positivity is: 0.440528634361234	Calculated positivity is: 2.026805
149	Original file's positivity is: 0.819672131147541	Calculated positivity is: 2.1195433
150	Original file's positivity is: 4.6908315565032	Calculated positivity is: 2.0223637
151	Original file's positivity is: 1.95852534562212	Calculated positivity is: 2.257563
152	Original file's positivity is: 1.18343195266272	Calculated positivity is: 1.986645
153	Original file's positivity is: 1.8450184501845	Calculated positivity is: 2.1741664
154	Original file's positivity is: 2.78884462151394	Calculated positivity is: 2.11709
155	Original file's positivity is: 4.52488687782805	Calculated positivity is: 1.8873837
156	Original file's positivity is: 2.16450216450216	Calculated positivity is: 2.1797378
157	Original file's positivity is: 2.1551724137931	Calculated positivity is: 1.9096413
158	Original file's positivity is: 3.486529318542	Calculated positivity is: 2.3669496
159	Original file's positivity is: 3.78006872852234	Calculated positivity is: 2.1747625
160	Original file's positivity is: 6.81362725450902	Calculated positivity is: 1.961707
161	Original file's positivity is: 1.7156862745098	Calculated positivity is: 2.6117911
162	Original file's positivity is: 3.35820895522388	Calculated positivity is: 2.389204
163	Original file's positivity is: 2.84162139573757	Calculated positivity is: 2.6234226
164	Original file's positivity is: 1.23456790123457	Calculated positivity is: 2.2733388
165	Original file's positivity is: 1.10165247871808	Calculated positivity is: 2.1959343
166	Original file's positivity is: 3.03030303030303	Calculated positivity is: 2.5959148
167	Original file's positivity is: 4.17287630402385	Calculated positivity is: 2.0224853
168	Original file's positivity is: 3.92609699769053	Calculated positivity is: 2.2356977
169	Original file's positivity is: 1.77935943060498	Calculated positivity is: 2.2705886
170	Original file's positivity is: 6.31313131313131	Calculated positivity is: 2.0489995
171	Original file's positivity is: 2.24719101123596	Calculated positivity is: 2.252596
172	Original file's positivity is: 2.82685512367491	Calculated positivity is: 2.104947
173	Original file's positivity is: 3.89429763560501	Calculated positivity is: 2.186239

174	Original file's positivity is: 2.8894472361809	Calculated positivity is: 2.4694088
175	Original file's positivity is: 2.51572327044025	Calculated positivity is: 2.165795
176	Original file's positivity is: 2.40641711229947	Calculated positivity is: 2.775388
177	Original file's positivity is: 1.78082191780822	Calculated positivity is: 2.2231047

Table A2.3: Calculated polarity values for larger dataset vs. Original polarity values

File Number	Original File Result (Polarization)	Calculated Result (Polarization - different calculation than original)
0	Original file's polarity is: 7.60309278350515	Calculated polarity is: 6.168494
1	Original file's polarity is: 8.72274143302181	Calculated polarity is: 6.4345713
2	Original file's polarity is: 10.2564102564103	Calculated polarity is: 6.1744523
3	Original file's polarity is: 4.12319920516642	Calculated polarity is: 6.442193
4	Original file's polarity is: 6.22992935131663	Calculated polarity is: 6.617891
5	Original file's polarity is: 2.90697674418605	Calculated polarity is: 5.9167995
6	Original file's polarity is: 1.67410714285714	Calculated polarity is: 6.746684
7	Original file's polarity is: 6.43153526970954	Calculated polarity is: 5.8869123
8	Original file's polarity is: 4.2042042042042	Calculated polarity is: 5.947535
9	Original file's polarity is: 7.44047619047619	Calculated polarity is: 5.960481
10	Original file's polarity is: 4.6583850931677	Calculated polarity is: 6.2436905
11	Original file's polarity is: 6.44171779141104	Calculated polarity is: 5.8590794
12	Original file's polarity is: 8.66141732283465	Calculated polarity is: 5.8251925
13	Original file's polarity is: 8.66510538641686	Calculated polarity is: 5.912664
14	Original file's polarity is: 4.47761194029851	Calculated polarity is: 6.185478
15	Original file's polarity is: 3.5778175313059	Calculated polarity is: 7.463023
16	Original file's polarity is: 6.31120783460283	Calculated polarity is: 6.087138
17	Original file's polarity is: 6.33213859020311	Calculated polarity is: 5.941821
18	Original file's polarity is: 3.44827586206897	Calculated polarity is: 5.8260336
19	Original file's polarity is: 8.3547557840617	Calculated polarity is: 5.9126425
20	Original file's polarity is: 6.34920634920635	Calculated polarity is: 6.1712894
21	Original file's polarity is: 8.38926174496644	Calculated polarity is: 6.5032015
22	Original file's polarity is: 6.97292863002461	Calculated polarity is: 6.7360473
23	Original file's polarity is: 6.62983425414365	Calculated polarity is: 6.2551928
24	Original file's polarity is: 6.77570093457944	Calculated polarity is: 5.8592963
25	Original file's polarity is: 5.66473988439306	Calculated polarity is: 7.532902
26	Original file's polarity is: 5.02983802216539	Calculated polarity is: 6.8547335

27	Original file's polarity is: 1.45547945205479	Calculated polarity is: 6.7689056
28	Original file's polarity is: 6.49122807017544	Calculated polarity is: 6.532732
29	Original file's polarity is: 8.84955752212389	Calculated polarity is: 5.7550917
30	Original file's polarity is: 4.72103004291846	Calculated polarity is: 5.990508
31	Original file's polarity is: 12.9382303839733	Calculated polarity is: 6.769086
32	Original file's polarity is: 15.0214592274678	Calculated polarity is: 5.8232207
33	Original file's polarity is: 3.6697247706422	Calculated polarity is: 5.804947
34	Original file's polarity is: 6.32911392405063	Calculated polarity is: 6.0296316
35	Original file's polarity is: 8.62906238326485	Calculated polarity is: 6.1850834
36	Original file's polarity is: 5.15222482435597	Calculated polarity is: 5.8882027
37	Original file's polarity is: 6.38297872340426	Calculated polarity is: 5.606585
38	Original file's polarity is: 9.05511811023622	Calculated polarity is: 5.865777
39	Original file's polarity is: 7.18685831622177	Calculated polarity is: 5.7384415
40	Original file's polarity is: 6.37254901960784	Calculated polarity is: 5.858894
41	Original file's polarity is: 3.42205323193916	Calculated polarity is: 5.90411
42	Original file's polarity is: 5.33980582524272	Calculated polarity is: 6.611345
43	Original file's polarity is: 5.71065989847716	Calculated polarity is: 6.0595455
44	Original file's polarity is: 2.89855072463768	Calculated polarity is: 5.5393953
45	Original file's polarity is: 6.76795580110497	Calculated polarity is: 6.2992725
46	Original file's polarity is: 5.28233151183971	Calculated polarity is: 7.3000054
47	Original file's polarity is: 5.42168674698795	Calculated polarity is: 5.8097544
48	Original file's polarity is: 8.34492350486787	Calculated polarity is: 6.0996
49	Original file's polarity is: 4.74934036939314	Calculated polarity is: 6.044327
50	Original file's polarity is: 2.95081967213115	Calculated polarity is: 8.792524
51	Original file's polarity is: 6.28019323671498	Calculated polarity is: 6.5664716
52	Original file's polarity is: 5.31645569620253	Calculated polarity is: 5.8496647
53	Original file's polarity is: 7.81527531083481	Calculated polarity is: 6.3637
54	Original file's polarity is: 8.0188679245283	Calculated polarity is: 7.432764
55	Original file's polarity is: 5.58213716108453	Calculated polarity is: 6.1566257
56	Original file's polarity is:	Calculated polarity is: 6.4992657

	7.35294117647059	
57	Original file's polarity is: 9.02654867256637	Calculated polarity is: 6.1090384
58	Original file's polarity is: 3.75275938189845	Calculated polarity is: 6.095004
59	Original file's polarity is: 4.25531914893617	Calculated polarity is: 5.9477897
60	Original file's polarity is: 5.70291777188329	Calculated polarity is: 5.8524227
61	Original file's polarity is: 5.05836575875486	Calculated polarity is: 6.000403
62	Original file's polarity is: 4.13625304136253	Calculated polarity is: 5.868599
63	Original file's polarity is: 2.06766917293233	Calculated polarity is: 5.9230766
64	Original file's polarity is: 6.33333333333333	Calculated polarity is: 6.3672905
65	Original file's polarity is: 9.48905109489051	Calculated polarity is: 6.5151143
66	Original file's polarity is: 8.51735015772871	Calculated polarity is: 6.52532
67	Original file's polarity is: 6.79611650485437	Calculated polarity is: 5.7830596
68	Original file's polarity is: 7.29684908789386	Calculated polarity is: 6.325411
69	Original file's polarity is: 6.35324015247776	Calculated polarity is: 6.7678022
70	Original file's polarity is: 2.46575342465753	Calculated polarity is: 5.860573
71	Original file's polarity is: 2.13706705969049	Calculated polarity is: 5.792575
72	Original file's polarity is: 7.04225352112676	Calculated polarity is: 8.082336
73	Original file's polarity is: 4.59770114942529	Calculated polarity is: 6.6020527
74	Original file's polarity is: 4.42307692307692	Calculated polarity is: 5.994563
75	Original file's polarity is: 6.90112806901128	Calculated polarity is: 6.389532
76	Original file's polarity is: 5.97402597402597	Calculated polarity is: 5.861006
77	Original file's polarity is: 4.38356164383562	Calculated polarity is: 7.095204
78	Original file's polarity is: 6.09884332281809	Calculated polarity is: 5.7503095
79	Original file's polarity is: 5.52147239263804	Calculated polarity is: 6.1192083
80	Original file's polarity is: 8.33333333333333	Calculated polarity is: 6.8827314
81	Original file's polarity is: 5.16304347826087	Calculated polarity is: 5.826598
82	Original file's polarity is: 4.13943355119826	Calculated polarity is: 6.2612963
83	Original file's polarity is: 12.2107969151671	Calculated polarity is: 6.394962
84	Original file's polarity is: 2.53378378378378	Calculated polarity is: 7.142974
85	Original file's polarity is: 2.66808964781217	Calculated polarity is: 6.8287315

86	Original file's polarity is: 6.80529300567108	Calculated polarity is: 6.015116
87	Original file's polarity is: 5.54016620498615	Calculated polarity is: 5.9570866
88	Original file's polarity is: 2.9986962190352	Calculated polarity is: 7.5783134
89	Original file's polarity is: 4.91493383742911	Calculated polarity is: 6.750164
90	Original file's polarity is: 10.4545454545455	Calculated polarity is: 5.7996664
91	Original file's polarity is: 2.67857142857143	Calculated polarity is: 6.141052
92	Original file's polarity is: 4.20560747663551	Calculated polarity is: 6.1163235
93	Original file's polarity is: 3.45679012345679	Calculated polarity is: 6.5160913
94	Original file's polarity is: 7.02614379084967	Calculated polarity is: 7.255522
95	Original file's polarity is: 6.55447298494243	Calculated polarity is: 7.293237
96	Original file's polarity is: 4.73186119873817	Calculated polarity is: 6.5607862
97	Original file's polarity is: 6.36132315521628	Calculated polarity is: 5.8188605
98	Original file's polarity is: 6.60522273425499	Calculated polarity is: 6.1326666
99	Original file's polarity is: 4.87421383647799	Calculated polarity is: 5.9679666
100	Original file's polarity is: 6.11510791366906	Calculated polarity is: 6.83572
101	Original file's polarity is: 3.2	Calculated polarity is: 6.1316533
102	Original file's polarity is: 7.16443677168753	Calculated polarity is: 6.8299084
103	Original file's polarity is: 7.5065274151436	Calculated polarity is: 5.7135673
104	Original file's polarity is: 5.69037656903766	Calculated polarity is: 6.31821
105	Original file's polarity is: 2.51091703056769	Calculated polarity is: 5.9885178
106	Original file's polarity is: 5.39861895794099	Calculated polarity is: 6.5613003
107	Original file's polarity is: 4.90196078431373	Calculated polarity is: 5.845851
108	Original file's polarity is: 5.04854368932039	Calculated polarity is: 6.0478497
109	Original file's polarity is: 8.19672131147541	Calculated polarity is: 7.770508
110	Original file's polarity is: 10.1694915254237	Calculated polarity is: 6.6476784
111	Original file's polarity is: 4.84581497797357	Calculated polarity is: 5.8954277
112	Original file's polarity is: 6.18556701030928	Calculated polarity is: 6.882678
113	Original file's polarity is: 4.73815461346633	Calculated polarity is: 5.6751904
114	Original file's polarity is: 6.87554395126197	Calculated polarity is: 6.407691
115	Original file's polarity is: 6.89655172413793	Calculated polarity is: 6.24012

116	Original file's polarity is: 5.723124516628	Calculated polarity is: 5.866762
117	Original file's polarity is: 3.42205323193916	Calculated polarity is: 6.2986817
118	Original file's polarity is: 6.35202271114265	Calculated polarity is: 6.3145366
119	Original file's polarity is: 11.8937644341801	Calculated polarity is: 6.1477547
120	Original file's polarity is: 8.39080459770115	Calculated polarity is: 5.9985847
121	Original file's polarity is: 10.5438401775805	Calculated polarity is: 6.281966
122	Original file's polarity is: 5.80645161290323	Calculated polarity is: 7.3535824
123	Original file's polarity is: 2.7027027027027	Calculated polarity is: 6.043998
124	Original file's polarity is: 5.26315789473684	Calculated polarity is: 6.4030685
125	Original file's polarity is: 6.8922305764411	Calculated polarity is: 6.212076
126	Original file's polarity is: 3.74064837905237	Calculated polarity is: 6.3386307
127	Original file's polarity is: 4.83870967741935	Calculated polarity is: 6.228629
128	Original file's polarity is: 5.14285714285714	Calculated polarity is: 6.7824926
129	Original file's polarity is: 3.39366515837104	Calculated polarity is: 6.504435
130	Original file's polarity is: 5.8641975308642	Calculated polarity is: 5.816474
131	Original file's polarity is: 8.04195804195804	Calculated polarity is: 6.5349617
132	Original file's polarity is: 8.23045267489712	Calculated polarity is: 6.2468376
133	Original file's polarity is: 13.6448598130841	Calculated polarity is: 6.073716
134	Original file's polarity is: 4.36730123180291	Calculated polarity is: 6.31437
135	Original file's polarity is: 6.0968660968661	Calculated polarity is: 6.525799
136	Original file's polarity is: 7.31070496083551	Calculated polarity is: 6.084316
137	Original file's polarity is: 4.7713717693837	Calculated polarity is: 5.755988
138	Original file's polarity is: 6.22710622710623	Calculated polarity is: 7.2070255
139	Original file's polarity is: 0.735294117647059	Calculated polarity is: 6.6642604
140	Original file's polarity is: 5.49450549450549	Calculated polarity is: 6.6680965
141	Original file's polarity is: 5.2863436123348	Calculated polarity is: 6.153537
142	Original file's polarity is: 9.61538461538462	Calculated polarity is: 5.7653804
143	Original file's polarity is: 9.93537964458804	Calculated polarity is: 6.0778418
144	Original file's polarity is: 6.59133709981168	Calculated polarity is: 7.3612156
145	Original file's polarity is:	Calculated polarity is: 6.2558603

	7.1157495256167	
146	Original file's polarity is: 11.3268608414239	Calculated polarity is: 5.8881874
147	Original file's polarity is: 5.51181102362205	Calculated polarity is: 6.380746
148	Original file's polarity is: 3.52422907488987	Calculated polarity is: 5.878134
149	Original file's polarity is: 2.04918032786885	Calculated polarity is: 6.0266113
150	Original file's polarity is: 5.97014925373134	Calculated polarity is: 5.8465023
151	Original file's polarity is: 9.44700460829493	Calculated polarity is: 6.3601785
152	Original file's polarity is: 4.43786982248521	Calculated polarity is: 5.7924438
153	Original file's polarity is: 4.7970479704797	Calculated polarity is: 6.065425
154	Original file's polarity is: 9.9601593625498	Calculated polarity is: 6.059367
155	Original file's polarity is: 5.20361990950226	Calculated polarity is: 5.66288
156	Original file's polarity is: 4.76190476190476	Calculated polarity is: 6.071539
157	Original file's polarity is: 9.91379310344828	Calculated polarity is: 5.6506453
158	Original file's polarity is: 3.486529318542	Calculated polarity is: 6.499996
159	Original file's polarity is: 4.81099656357388	Calculated polarity is: 6.173824
160	Original file's polarity is: 7.01402805611222	Calculated polarity is: 5.7916536
161	Original file's polarity is: 3.92156862745098	Calculated polarity is: 6.756459
162	Original file's polarity is: 9.28171641791045	Calculated polarity is: 6.3943615
163	Original file's polarity is: 4.1370664437944	Calculated polarity is: 6.7174067
164	Original file's polarity is: 1.64609053497942	Calculated polarity is: 6.13696
165	Original file's polarity is: 6.45968953430145	Calculated polarity is: 6.1615143
166	Original file's polarity is: 10.1515151515152	Calculated polarity is: 6.7768717
167	Original file's polarity is: 9.53800298062593	Calculated polarity is: 5.847267
168	Original file's polarity is: 6.23556581986143	Calculated polarity is: 6.1841516
169	Original file's polarity is: 3.73665480427046	Calculated polarity is: 6.2806864
170	Original file's polarity is: 9.34343434343434	Calculated polarity is: 5.948187
171	Original file's polarity is: 6.06741573033708	Calculated polarity is: 6.2726398
172	Original file's polarity is: 8.65724381625442	Calculated polarity is: 5.9288416
173	Original file's polarity is: 5.14603616133519	Calculated polarity is: 6.111865
174	Original file's polarity is: 3.64321608040201	Calculated polarity is: 6.587598

175	Original file's polarity is: 6.13207547169811	Calculated polarity is: 6.087817
176	Original file's polarity is: 7.75401069518717	Calculated polarity is: 7.108656
177	Original file's polarity is: 5.47945205479452	Calculated polarity is: 6.130196

c. Tables A3.0-A3.3:

Table A3.0: Calculated sentiment values for locations vs. Original GDELT averages

Location	Original File Result (Average)	Calculated Result
Location is: South Korea	Original files' average tone is: -2.5908592	Calculated result is: 0.11933136
Location is: Japan	Original files' average tone is: -2.1076438	Calculated result is: 0.11933088
Location is: United States	Original files' average tone is: -1.4894167	Calculated result is: 0.118980885
Location is: China	Original files' average tone is: -3.106933	Calculated result is: 0.11933136
Location is: North Korea	Original files' average tone is: -3.840702	Calculated result is: 0.11933112
Location is: Russia	Original files' average tone is: -1.4011174	Calculated result is: 0.119330645
Location is: Germany	Original files' average tone is: -2.7648642	Calculated result is: 0.11933088
Location is: India	Original files' average tone is: -3.1561415	Calculated result is: 0.11933136
Location is: United Kingdom	Original files' average tone is: -1.4649134	Calculated result is: 0.11926055
Location is: Canada	Original files' average tone is: -0.5019273	Calculated result is: 0.1193316
Location is: Mexico	Original files' average tone is: -1.1577206	Calculated result is: 0.11933136
Location is: Brazil	Original files' average tone is: -3.3911896	Calculated result is: 0.11933136
Location is: Colombia	Original files' average tone is: -0.11297866	Calculated result is: 0.11933136
Location is: Israel	Original files' average tone is: -1.7165874	Calculated result is: 0.119330645
Location is: Afghanistan	Original files' average tone is: -3.9722493	Calculated result is: 0.11933088
Location is: Djibouti	Original files' average tone is: -1.931876	Calculated result is: 0.119330645
Location is: West Bank	Original files' average tone is: -1.931876	Calculated result is: 0.119330645
Location is: Nigeria	Original files' average tone is: 0.53077996	Calculated result is: 0.11933112
Location is: South Africa	Original files' average tone is: 0.8006354	Calculated result is: 0.11933112
Location is: Australia	Original files' average tone is: 0.4756126	Calculated result is: 0.11933136
Location is: Hong Kong	Original files' average tone is: -4.072198	Calculated result is: 0.11933088
Location is: France	Original files' average tone is: -3.4478614	Calculated result is: 0.11918044
Location is: Congo	Original files' average tone is: 0.990099	Calculated result is: 0.1193316
Location is: Netherlands	Original files' average tone is: 0.7448665	Calculated result is: 0.11933136
Location is: Malaysia	Original files' average tone is: -1.1439828	Calculated result is: 0.11933112
Location is: Chile	Original files' average tone is: -2.2275198	Calculated result is: 0.11933136

Location is: Indonesia	Original files' average tone is: 0.7368421	Calculated result is: 0.11932993
Location is: Singapore	Original files' average tone is: -0.24311836	Calculated result is: 0.11933088
Location is: Zimbabwe	Original files' average tone is: -4.195804	Calculated result is: 0.11933136
Location is: Ukraine	Original files' average tone is: -0.5245304	Calculated result is: 0.1193316
Location is: Switzerland	Original files' average tone is: -0.3889721	Calculated result is: 0.119330406
Location is: Jamaica	Original files' average tone is: 1.8459015	Calculated result is: 0.11933136
Location is: Greenland	Original files' average tone is: 0.56484646	Calculated result is: 0.11933136
Location is: Denmark	Original files' average tone is: -1.2262305	Calculated result is: 0.11933136
Location is: Peru	Original files' average tone is: -5.7042065	Calculated result is: 0.119330645
Location is: Bolivia	Original files' average tone is: -6.545961	Calculated result is: 0.119330645
Location is: Iran	Original files' average tone is: -3.579773	Calculated result is: 0.11933088
Location is: Italy	Original files' average tone is: -3.3906515	Calculated result is: 0.11933088
Location is: New Zealand	Original files' average tone is: -0.9850718	Calculated result is: 0.11933088
Location is: Poland	Original files' average tone is: 6.0465117	Calculated result is: 0.11933136
Location is: Turkey	Original files' average tone is: -4.001675	Calculated result is: 0.11933112
Location is: Honduras	Original files' average tone is: -2.0697167	Calculated result is: 0.11933136
Location is: El Salvador	Original files' average tone is: -3.3900917	Calculated result is: 0.11933112
Location is: Guatemala	Original files' average tone is: -2.0697167	Calculated result is: 0.11933136
Location is: Belgium	Original files' average tone is: -3.2739012	Calculated result is: 0.1193316
Location is: Kenya	Original files' average tone is: -7.6648507	Calculated result is: 0.11933112
Location is: Austria	Original files' average tone is: 2.819212	Calculated result is: 0.11933088
Location is: Puerto Rico	Original files' average tone is: 0.64794815	Calculated result is: 0.11933112
Location is: Mauritania	Original files' average tone is: -0.47619048	Calculated result is: 0.11933088
Location is: Luxembourg	Original files' average tone is: -0.47619048	Calculated result is: 0.11933088
Location is: Spain	Original files' average tone is: -1.7948879	Calculated result is: 0.1044147
Location is: Iceland	Original files' average tone is: -0.973236	Calculated result is: 0.11933112
Location is: Bahrain	Original files' average tone is: -3.030303	Calculated result is: 0.11933136
Location is: Norway	Original files' average tone is: -4.929163	Calculated result is: 0.11933112
Location is: Argentina	Original files' average tone is: -1.2855	Calculated result is: 0.11933112
Location is:	Original files' average tone is: -	Calculated result is: 0.1193316

Venezuela	4.9342103	
Location is: Ghana	Original files' average tone is: -1.1211467	Calculated result is: 0.11933112
Location is: Philippines	Original files' average tone is: 1.7006803	Calculated result is: 0.11933088
Location is: Ireland	Original files' average tone is: -2.4630437	Calculated result is: 0.11933088
Location is: Syria	Original files' average tone is: -7.4934278	Calculated result is: 0.11822605
Location is: Croatia	Original files' average tone is: 0.53333336	Calculated result is: 0.1193316
Location is: Sierra Leone	Original files' average tone is: 0.95419854	Calculated result is: 0.11933136
Location is: United Arab Emirates	Original files' average tone is: 0.9541985	Calculated result is: 0.11933136
Location is: American Samoa	Original files' average tone is: 4.316547	Calculated result is: 0.11933112
Location is: Qatar	Original files' average tone is: 0.37950665	Calculated result is: 0.11933136
Location is: Saudi Arabia	Original files' average tone is: 0.37950665	Calculated result is: 0.11933136
Location is: Kuwait	Original files' average tone is: 0.37950665	Calculated result is: 0.11933136
Location is: Portugal	Original files' average tone is: -0.7092199	Calculated result is: 0.11933112
Location is: Kazakhstan	Original files' average tone is: 0.22675736	Calculated result is: 0.11933136
Location is: Kyrgyzstan	Original files' average tone is: 0.22675736	Calculated result is: 0.11933136
Location is: Macau	Original files' average tone is: 1.1235955	Calculated result is: 0.11933112
Location is: Papua New Guinea	Original files' average tone is: 0.22953328	Calculated result is: 0.11933112
Location is: Taiwan	Original files' average tone is: -8.764045	Calculated result is: 0.1193316
Location is: Hungary	Original files' average tone is: -4.6117463	Calculated result is: 0.11933112
Location is: Bulgaria	Original files' average tone is: -4.6117463	Calculated result is: 0.11933112
Location is: Brunei	Original files' average tone is: -4.6117463	Calculated result is: 0.11933112
Location is: Romania	Original files' average tone is: -4.6117463	Calculated result is: 0.11933112
Location is: Egypt	Original files' average tone is: 2.0164986	Calculated result is: 0.11933136
Location is: Liberia	Original files' average tone is: -3.721374	Calculated result is: 0.11933112
Location is: Paraguay	Original files' average tone is: 1.6722409	Calculated result is: 0.119330645
Location is: Uruguay	Original files' average tone is: 1.6722409	Calculated result is: 0.119330645
Location is: Bangladesh	Original files' average tone is: 1.6722409	Calculated result is: 0.119330645
Location is: Greece	Original files' average tone is: 9.159348	Calculated result is: 0.11933112
Location is:	Original files' average tone is: 3.988604	Calculated result is: 0.11305082

Sweden		
Location is: Pakistan	Original files' average tone is: 3.053435	Calculated result is: 0.113780856

Table A3.1: Calculated negative sentiment values for locations vs. Original GDELT

negative sentiment averages

Location	Original File Result (Negativity - Average)	Calculated Result (Negativity)
Location is: South Korea	Original files' average negativity is: -5.170285	Calculated negativity is: -2.0286405
Location is: Japan	Original files' average negativity is: -4.5477605	Calculated negativity is: -2.0286405
Location is: United States	Original files' average negativity is: -3.749911	Calculated negativity is: -2.0226052
Location is: China	Original files' average negativity is: -5.201587	Calculated negativity is: -2.0286405
Location is: North Korea	Original files' average negativity is: -6.592665	Calculated negativity is: -2.0286405
Location is: Russia	Original files' average negativity is: -4.278405	Calculated negativity is: -2.0286405
Location is: Germany	Original files' average negativity is: -4.683919	Calculated negativity is: -2.0286405
Location is: India	Original files' average negativity is: -4.908395	Calculated negativity is: -2.0286405
Location is: United Kingdom	Original files' average negativity is: -4.160953	Calculated negativity is: -2.0274417
Location is: Canada	Original files' average negativity is: -3.0799656	Calculated negativity is: -2.0286403
Location is: Mexico	Original files' average negativity is: -3.8297997	Calculated negativity is: -2.0286405
Location is: Brazil	Original files' average negativity is: -5.336324	Calculated negativity is: -2.0286405
Location is: Colombia	Original files' average negativity is: -2.6298583	Calculated negativity is: -2.0286405
Location is: Israel	Original files' average negativity is: -3.9469204	Calculated negativity is: -2.0286405
Location is: Afghanistan	Original files' average negativity is: -6.5994606	Calculated negativity is: -2.0286405
Location is: Djibouti	Original files' average negativity is: -3.7620742	Calculated negativity is: -2.0286405
Location is: West Bank	Original files' average negativity is: -3.7620742	Calculated negativity is: -2.0286405
Location is: Nigeria	Original files' average negativity is: -2.9405801	Calculated negativity is: -2.0286403
Location is: South Africa	Original files' average negativity is: -0.88959134	Calculated negativity is: -2.0286403
Location is: Australia	Original files' average negativity is: -2.225084	Calculated negativity is: -2.0286405
Location is: Hong Kong	Original files' average negativity is: -6.2089157	Calculated negativity is: -2.0286405
Location is: France	Original files' average negativity is: -5.428293	Calculated negativity is: -2.0260792
Location is: Congo	Original files' average negativity is: -0.4950495	Calculated negativity is: -2.02864
Location is: Netherlands	Original files' average negativity is: -2.151424	Calculated negativity is: -2.0286403
Location is: Malaysia	Original files' average negativity is: -3.6972005	Calculated negativity is: -2.0286403
Location is:	Original files' average negativity is: -	Calculated negativity is: -

Chile	4.0016575	2.0286405
Location is: Indonesia	Original files' average negativity is: -1.3684211	Calculated negativity is: -2.0286405
Location is: Singapore	Original files' average negativity is: -2.307103	Calculated negativity is: -2.0286405
Location is: Zimbabwe	Original files' average negativity is: -4.895105	Calculated negativity is: -2.02864
Location is: Ukraine	Original files' average negativity is: -2.6029525	Calculated negativity is: -2.02864
Location is: Switzerland	Original files' average negativity is: -2.2998793	Calculated negativity is: -2.0286405
Location is: Jamaica	Original files' average negativity is: -2.2839665	Calculated negativity is: -2.0286403
Location is: Greenland	Original files' average negativity is: -2.0736701	Calculated negativity is: -2.0286403
Location is: Denmark	Original files' average negativity is: -3.9152725	Calculated negativity is: -2.0286403
Location is: Peru	Original files' average negativity is: -7.262943	Calculated negativity is: -2.0286405
Location is: Bolivia	Original files' average negativity is: -7.8922935	Calculated negativity is: -2.0286403
Location is: Iran	Original files' average negativity is: -5.997152	Calculated negativity is: -2.0286405
Location is: Italy	Original files' average negativity is: -5.893038	Calculated negativity is: -2.0286405
Location is: New Zealand	Original files' average negativity is: -2.6638317	Calculated negativity is: -2.0286405
Location is: Poland	Original files' average negativity is: -0.81395346	Calculated negativity is: -2.0286405
Location is: Turkey	Original files' average negativity is: -6.095144	Calculated negativity is: -2.0286403
Location is: Honduras	Original files' average negativity is: -5.2287583	Calculated negativity is: -2.0286405
Location is: El Salvador	Original files' average negativity is: -5.5604243	Calculated negativity is: -2.0286403
Location is: Guatemala	Original files' average negativity is: -5.2287583	Calculated negativity is: -2.0286405
Location is: Belgium	Original files' average negativity is: -5.3838196	Calculated negativity is: -2.0286403
Location is: Kenya	Original files' average negativity is: -8.455624	Calculated negativity is: -2.0286403
Location is: Austria	Original files' average negativity is: -2.1316383	Calculated negativity is: -2.0286405
Location is: Puerto Rico	Original files' average negativity is: -2.1598272	Calculated negativity is: -2.0286405
Location is: Mauritania	Original files' average negativity is: -0.95238096	Calculated negativity is: -2.0286405
Location is: Luxembourg	Original files' average negativity is: -0.95238096	Calculated negativity is: -2.0286405
Location is: Spain	Original files' average negativity is: -4.071936	Calculated negativity is: -1.77506
Location is: Iceland	Original files' average negativity is: -3.406326	Calculated negativity is: -2.0286403
Location is: Bahrain	Original files' average negativity is: -3.030303	Calculated negativity is: -2.0286405
Location is: Norway	Original files' average negativity is: -6.2491155	Calculated negativity is: -2.0286403
Location is: Argentina	Original files' average negativity is: -4.096164	Calculated negativity is: -2.0286405

Location is: Venezuela	Original files' average negativity is: -6.9078946	Calculated negativity is: -2.0286403
Location is: Ghana	Original files' average negativity is: -2.6718814	Calculated negativity is: -2.0286403
Location is: Philippines	Original files' average negativity is: -0.6802721	Calculated negativity is: -2.0286403
Location is: Ireland	Original files' average negativity is: -4.2798853	Calculated negativity is: -2.0286403
Location is: Syria	Original files' average negativity is: -9.046331	Calculated negativity is: -2.0098565
Location is: Croatia	Original files' average negativity is: -1.3333334	Calculated negativity is: -2.0286403
Location is: Sierra Leone	Original files' average negativity is: -2.2900763	Calculated negativity is: -2.0286403
Location is: United Arab Emirates	Original files' average negativity is: -2.2900763	Calculated negativity is: -2.0286403
Location is: American Samoa	Original files' average negativity is: -0.7194245	Calculated negativity is: -2.0286403
Location is: Qatar	Original files' average negativity is: -0.7590133	Calculated negativity is: -2.0286398
Location is: Saudi Arabia	Original files' average negativity is: -0.7590133	Calculated negativity is: -2.0286398
Location is: Kuwait	Original files' average negativity is: -0.7590133	Calculated negativity is: -2.0286398
Location is: Portugal	Original files' average negativity is: -2.8368795	Calculated negativity is: -2.0286403
Location is: Kazakhstan	Original files' average negativity is: -2.9478457	Calculated negativity is: -2.0286403
Location is: Kyrgyzstan	Original files' average negativity is: -2.9478457	Calculated negativity is: -2.0286403
Location is: Macau	Original files' average negativity is: -1.9662921	Calculated negativity is: -2.0286403
Location is: Papua New Guinea	Original files' average negativity is: -1.3771996	Calculated negativity is: -2.0286405
Location is: Taiwan	Original files' average negativity is: -9.588015	Calculated negativity is: -2.0286403
Location is: Hungary	Original files' average negativity is: -7.6862435	Calculated negativity is: -2.0286405
Location is: Bulgaria	Original files' average negativity is: -7.6862435	Calculated negativity is: -2.0286405
Location is: Brunei	Original files' average negativity is: -7.6862435	Calculated negativity is: -2.0286405
Location is: Romania	Original files' average negativity is: -7.6862435	Calculated negativity is: -2.0286405
Location is: Egypt	Original files' average negativity is: -3.116407	Calculated negativity is: -2.0286403
Location is: Liberia	Original files' average negativity is: -5.057252	Calculated negativity is: -2.0286403
Location is: Paraguay	Original files' average negativity is: -1.3377926	Calculated negativity is: -2.0286405
Location is: Uruguay	Original files' average negativity is: -1.3377926	Calculated negativity is: -2.0286405
Location is: Bangladesh	Original files' average negativity is: -1.3377926	Calculated negativity is: -2.0286405
Location is: Greece	Original files' average negativity is: -0.12547052	Calculated negativity is: -2.0286405

Location is: Sweden	Original files' average negativity is: -0.56980056	Calculated negativity is: -1.9218696
Location is: Pakistan	Original files' average negativity is: -1.0178117	Calculated negativity is: -1.934285

Table A3.2: Calculated positive sentiment values for locations vs. Original GDELT

positive sentiment averages

Location	Original File Result (Positivity - Average)	Calculated Result (Positivity)
Location is: South Korea	Original files' average positivity is: 2.579426	Calculated positivity is: 2.1479719
Location is: Japan	Original files' average positivity is: 2.4401162	Calculated positivity is: 2.1479714
Location is: United States	Original files' average positivity is: 2.260491	Calculated positivity is: 2.141586
Location is: China	Original files' average positivity is: 2.0946546	Calculated positivity is: 2.1479719
Location is: North Korea	Original files' average positivity is: 2.7519634	Calculated positivity is: 2.1479716
Location is: Russia	Original files' average positivity is: 2.8772883	Calculated positivity is: 2.1479712
Location is: Germany	Original files' average positivity is: 1.9190552	Calculated positivity is: 2.1479714
Location is: India	Original files' average positivity is: 1.7522541	Calculated positivity is: 2.1479719
Location is: United Kingdom	Original files' average positivity is: 2.6960397	Calculated positivity is: 2.1467023
Location is: Canada	Original files' average positivity is: 2.578039	Calculated positivity is: 2.1479719
Location is: Mexico	Original files' average positivity is: 2.6720796	Calculated positivity is: 2.1479719
Location is: Brazil	Original files' average positivity is: 1.9451336	Calculated positivity is: 2.1479719
Location is: Colombia	Original files' average positivity is: 2.5168793	Calculated positivity is: 2.1479719
Location is: Israel	Original files' average positivity is: 2.2303333	Calculated positivity is: 2.1479712
Location is: Afghanistan	Original files' average positivity is: 2.6272118	Calculated positivity is: 2.1479714
Location is: Djibouti	Original files' average positivity is: 1.8301983	Calculated positivity is: 2.1479712
Location is: West Bank	Original files' average positivity is: 1.8301983	Calculated positivity is: 2.1479712
Location is: Nigeria	Original files' average positivity is: 3.4713602	Calculated positivity is: 2.1479714
Location is: South Africa	Original files' average positivity is: 1.6902268	Calculated positivity is: 2.1479714
Location is: Australia	Original files' average positivity is: 2.7006958	Calculated positivity is: 2.1479719
Location is: Hong Kong	Original files' average positivity is: 2.1367176	Calculated positivity is: 2.1479714
Location is: France	Original files' average positivity is: 1.9804314	Calculated positivity is: 2.1452596
Location is: Congo	Original files' average positivity is: 1.4851485	Calculated positivity is: 2.1479716
Location is: Netherlands	Original files' average positivity is: 2.8962905	Calculated positivity is: 2.1479716
Location is: Malaysia	Original files' average positivity is: 2.5532181	Calculated positivity is: 2.1479714
Location is:	Original files' average positivity is:	Calculated positivity is:

Chile	1.774138	2.1479719
Location is: Indonesia	Original files' average positivity is: 2.1052632	Calculated positivity is: 2.1479704
Location is: Singapore	Original files' average positivity is: 2.0639846	Calculated positivity is: 2.1479714
Location is: Zimbabwe	Original files' average positivity is: 0.6993007	Calculated positivity is: 2.1479714
Location is: Ukraine	Original files' average positivity is: 2.0784223	Calculated positivity is: 2.1479716
Location is: Switzerland	Original files' average positivity is: 1.9109074	Calculated positivity is: 2.147971
Location is: Jamaica	Original files' average positivity is: 4.129868	Calculated positivity is: 2.1479716
Location is: Greenland	Original files' average positivity is: 2.638517	Calculated positivity is: 2.1479716
Location is: Denmark	Original files' average positivity is: 2.6890419	Calculated positivity is: 2.1479716
Location is: Peru	Original files' average positivity is: 1.5587351	Calculated positivity is: 2.1479712
Location is: Bolivia	Original files' average positivity is: 1.3463324	Calculated positivity is: 2.147971
Location is: Iran	Original files' average positivity is: 2.4173784	Calculated positivity is: 2.1479714
Location is: Italy	Original files' average positivity is: 2.5023859	Calculated positivity is: 2.1479714
Location is: New Zealand	Original files' average positivity is: 1.6787599	Calculated positivity is: 2.1479714
Location is: Poland	Original files' average positivity is: 6.860465	Calculated positivity is: 2.1479719
Location is: Turkey	Original files' average positivity is: 2.093469	Calculated positivity is: 2.1479714
Location is: Honduras	Original files' average positivity is: 3.1590414	Calculated positivity is: 2.1479719
Location is: El Salvador	Original files' average positivity is: 2.1703322	Calculated positivity is: 2.1479714
Location is: Guatemala	Original files' average positivity is: 3.1590414	Calculated positivity is: 2.1479719
Location is: Belgium	Original files' average positivity is: 2.1099186	Calculated positivity is: 2.1479719
Location is: Kenya	Original files' average positivity is: 0.7907729	Calculated positivity is: 2.1479714
Location is: Austria	Original files' average positivity is: 4.9508505	Calculated positivity is: 2.1479714
Location is: Puerto Rico	Original files' average positivity is: 2.8077753	Calculated positivity is: 2.1479716
Location is: Mauritania	Original files' average positivity is: 0.47619048	Calculated positivity is: 2.1479714
Location is: Luxembourg	Original files' average positivity is: 0.47619048	Calculated positivity is: 2.1479714
Location is: Spain	Original files' average positivity is: 2.277048	Calculated positivity is: 1.8794748
Location is: Iceland	Original files' average positivity is: 2.43309	Calculated positivity is: 2.1479714
Location is: Bahrain	Original files' average positivity is: 0.0	Calculated positivity is: 2.1479719
Location is: Norway	Original files' average positivity is: 1.3199527	Calculated positivity is: 2.1479714
Location is: Argentina	Original files' average positivity is: 2.8106644	Calculated positivity is: 2.1479716

Location is: Venezuela	Original files' average positivity is: 1.9736842	Calculated positivity is: 2.1479719
Location is: Ghana	Original files' average positivity is: 1.5507349	Calculated positivity is: 2.1479714
Location is: Philippines	Original files' average positivity is: 2.3809524	Calculated positivity is: 2.1479712
Location is: Ireland	Original files' average positivity is: 1.8168418	Calculated positivity is: 2.1479712
Location is: Syria	Original files' average positivity is: 1.5529038	Calculated positivity is: 2.1280825
Location is: Croatia	Original files' average positivity is: 1.8666667	Calculated positivity is: 2.1479719
Location is: Sierra Leone	Original files' average positivity is: 3.2442749	Calculated positivity is: 2.1479716
Location is: United Arab Emirates	Original files' average positivity is: 3.2442749	Calculated positivity is: 2.1479716
Location is: American Samoa	Original files' average positivity is: 5.035971	Calculated positivity is: 2.1479714
Location is: Qatar	Original files' average positivity is: 1.1385199	Calculated positivity is: 2.1479712
Location is: Saudi Arabia	Original files' average positivity is: 1.1385199	Calculated positivity is: 2.1479712
Location is: Kuwait	Original files' average positivity is: 1.1385199	Calculated positivity is: 2.1479712
Location is: Portugal	Original files' average positivity is: 2.1276596	Calculated positivity is: 2.1479714
Location is: Kazakhstan	Original files' average positivity is: 3.1746032	Calculated positivity is: 2.1479716
Location is: Kyrgyzstan	Original files' average positivity is: 3.1746032	Calculated positivity is: 2.1479716
Location is: Macau	Original files' average positivity is: 3.0898876	Calculated positivity is: 2.1479714
Location is: Papua New Guinea	Original files' average positivity is: 1.606733	Calculated positivity is: 2.1479716
Location is: Taiwan	Original files' average positivity is: 0.82397	Calculated positivity is: 2.1479719
Location is: Hungary	Original files' average positivity is: 3.0744975	Calculated positivity is: 2.1479716
Location is: Bulgaria	Original files' average positivity is: 3.0744975	Calculated positivity is: 2.1479716
Location is: Brunei	Original files' average positivity is: 3.0744975	Calculated positivity is: 2.1479716
Location is: Romania	Original files' average positivity is: 3.0744975	Calculated positivity is: 2.1479716
Location is: Egypt	Original files' average positivity is: 5.1329055	Calculated positivity is: 2.1479716
Location is: Liberia	Original files' average positivity is: 1.3358779	Calculated positivity is: 2.1479714
Location is: Paraguay	Original files' average positivity is: 3.0100334	Calculated positivity is: 2.1479712
Location is: Uruguay	Original files' average positivity is: 3.0100334	Calculated positivity is: 2.1479712
Location is: Bangladesh	Original files' average positivity is: 3.0100334	Calculated positivity is: 2.1479712
Location is: Greece	Original files' average positivity is: 9.284818	Calculated positivity is: 2.1479716

| Location is: Sweden | Original files' average positivity is: 4.5584044 | Calculated positivity is: 2.0349205 |
| Location is: Pakistan | Original files' average positivity is: 4.0712466 | Calculated positivity is: 2.048066 |

Table A3.3: Calculated polarity values for larger dataset vs. Original polarity values

Location	Original File Result (Polarization - Average)	Calculated Result (Polarization - different calculation than original)
Location is: South Korea	Original files' average polarity is: 7.7497125	Calculated polarity is: 45.57143
Location is: Japan	Original files' average polarity is: 6.9878764	Calculated polarity is: 45.07143
Location is: United States	Original files' average polarity is: 6.010395	Calculated polarity is: 39.17919
Location is: China	Original files' average polarity is: 7.296241	Calculated polarity is: 49.48077
Location is: North Korea	Original files' average polarity is: 9.344628	Calculated polarity is: 39.0
Location is: Russia	Original files' average polarity is: 7.1556935	Calculated polarity is: 49.0
Location is: Germany	Original files' average polarity is: 6.602975	Calculated polarity is: 45.7
Location is: India	Original files' average polarity is: 6.6606483	Calculated polarity is: 44.468086
Location is: United Kingdom	Original files' average polarity is: 6.856991	Calculated polarity is: 35.48936
Location is: Canada	Original files' average polarity is: 5.6580048	Calculated polarity is: 44.065216
Location is: Mexico	Original files' average polarity is: 6.501879	Calculated polarity is: 39.46154
Location is: Brazil	Original files' average polarity is: 7.2814584	Calculated polarity is: 37.61905
Location is: Colombia	Original files' average polarity is: 5.1467376	Calculated polarity is: 47.333332
Location is: Israel	Original files' average polarity is: 6.1772532	Calculated polarity is: 33.0
Location is: Afghanistan	Original files' average polarity is: 9.226672	Calculated polarity is: 42.333332
Location is: Djibouti	Original files' average polarity is: 5.5922723	Calculated polarity is: 27.0
Location is: West Bank	Original files' average polarity is: 5.5922723	Calculated polarity is: 27.0
Location is: Nigeria	Original files' average polarity is: 6.4119396	Calculated polarity is: 24.666666
Location is: South Africa	Original files' average polarity is: 2.5798182	Calculated polarity is: 19.5
Location is: Australia	Original files' average polarity is: 4.9257803	Calculated polarity is: 45.875
Location is: Hong Kong	Original files' average polarity is: 8.345633	Calculated polarity is: 38.916668
Location is: France	Original files' average polarity is: 7.408726	Calculated polarity is: 44.909092
Location is: Congo	Original files' average polarity is: 1.980198	Calculated polarity is: 16.0
Location is: Netherlands	Original files' average polarity is: 5.047714	Calculated polarity is: 38.0
Location is: Malaysia	Original files' average polarity is: 6.250418	Calculated polarity is: 39.0
Location is: Chile	Original files' average polarity is: 5.7757955	Calculated polarity is: 45.0

Location is: Indonesia	Original files' average polarity is: 3.4736843	Calculated polarity is: 73.0
Location is: Singapore	Original files' average polarity is: 4.3710876	Calculated polarity is: 43.333332
Location is: Zimbabwe	Original files' average polarity is: 5.5944057	Calculated polarity is: 29.0
Location is: Ukraine	Original files' average polarity is: 4.6813755	Calculated polarity is: 17.5
Location is: Switzerland	Original files' average polarity is: 4.210787	Calculated polarity is: 48.42857
Location is: Jamaica	Original files' average polarity is: 6.4138346	Calculated polarity is: 26.75
Location is: Greenland	Original files' average polarity is: 4.7121873	Calculated polarity is: 26.6
Location is: Denmark	Original files' average polarity is: 6.6043143	Calculated polarity is: 31.666666
Location is: Peru	Original files' average polarity is: 8.821677	Calculated polarity is: 58.384617
Location is: Bolivia	Original files' average polarity is: 9.238626	Calculated polarity is: 66.0
Location is: Iran	Original files' average polarity is: 8.414531	Calculated polarity is: 46.666668
Location is: Italy	Original files' average polarity is: 8.395424	Calculated polarity is: 47.77778
Location is: New Zealand	Original files' average polarity is: 4.342592	Calculated polarity is: 57.285713
Location is: Poland	Original files' average polarity is: 7.6744184	Calculated polarity is: 53.0
Location is: Turkey	Original files' average polarity is: 8.188614	Calculated polarity is: 30.2
Location is: Honduras	Original files' average polarity is: 8.387799	Calculated polarity is: 44.0
Location is: El Salvador	Original files' average polarity is: 7.7307563	Calculated polarity is: 32.666668
Location is: Guatemala	Original files' average polarity is: 8.387799	Calculated polarity is: 44.0
Location is: Belgium	Original files' average polarity is: 7.4937387	Calculated polarity is: 48.666668
Location is: Kenya	Original files' average polarity is: 9.246397	Calculated polarity is: 27.5
Location is: Austria	Original files' average polarity is: 7.082489	Calculated polarity is: 25.75
Location is: Puerto Rico	Original files' average polarity is: 4.9676027	Calculated polarity is: 57.0
Location is: Mauritania	Original files' average polarity is: 1.4285715	Calculated polarity is: 29.0
Location is: Luxembourg	Original files' average polarity is: 1.4285715	Calculated polarity is: 29.0
Location is: Spain	Original files' average polarity is: 6.3489847	Calculated polarity is: 19.625
Location is: Iceland	Original files' average polarity is: 5.839416	Calculated polarity is: 29.0
Location is: Bahrain	Original files' average polarity is: 3.030303	Calculated polarity is: 56.0
Location is: Norway	Original files' average polarity is: 7.5690684	Calculated polarity is: 27.0
Location is: Argentina	Original files' average polarity is: 6.9068294	Calculated polarity is: 41.666668
Location is:	Original files' average polarity is:	Calculated polarity is: 54.0

Venezuela	8.881579	
Location is: Ghana	Original files' average polarity is: 4.222616	Calculated polarity is: 20.666666
Location is: Philippines	Original files' average polarity is: 3.0612245	Calculated polarity is: 33.0
Location is: Ireland	Original files' average polarity is: 6.0967274	Calculated polarity is: 27.5
Location is: Syria	Original files' average polarity is: 10.599236	Calculated polarity is: 20.0
Location is: Croatia	Original files' average polarity is: 3.2	Calculated polarity is: 31.0
Location is: Sierra Leone	Original files' average polarity is: 5.5343513	Calculated polarity is: 40.0
Location is: United Arab Emirates	Original files' average polarity is: 5.5343513	Calculated polarity is: 40.0
Location is: American Samoa	Original files' average polarity is: 5.755396	Calculated polarity is: 29.0
Location is: Qatar	Original files' average polarity is: 1.8975332	Calculated polarity is: 17.0
Location is: Saudi Arabia	Original files' average polarity is: 1.8975332	Calculated polarity is: 17.0
Location is: Kuwait	Original files' average polarity is: 1.8975332	Calculated polarity is: 17.0
Location is: Portugal	Original files' average polarity is: 4.964539	Calculated polarity is: 36.0
Location is: Kazakhstan	Original files' average polarity is: 6.122449	Calculated polarity is: 41.0
Location is: Kyrgyzstan	Original files' average polarity is: 6.122449	Calculated polarity is: 41.0
Location is: Macau	Original files' average polarity is: 5.05618	Calculated polarity is: 28.0
Location is: Papua New Guinea	Original files' average polarity is: 2.9839327	Calculated polarity is: 58.0
Location is: Taiwan	Original files' average polarity is: 10.411985	Calculated polarity is: 37.0
Location is: Hungary	Original files' average polarity is: 10.760741	Calculated polarity is: 44.0
Location is: Bulgaria	Original files' average polarity is: 10.760741	Calculated polarity is: 44.0
Location is: Brunei	Original files' average polarity is: 10.760741	Calculated polarity is: 44.0
Location is: Romania	Original files' average polarity is: 10.760741	Calculated polarity is: 44.0
Location is: Egypt	Original files' average polarity is: 8.249312	Calculated polarity is: 29.0
Location is: Liberia	Original files' average polarity is: 6.39313	Calculated polarity is: 27.0
Location is: Paraguay	Original files' average polarity is: 4.347826	Calculated polarity is: 26.0
Location is: Uruguay	Original files' average polarity is: 4.347826	Calculated polarity is: 26.0
Location is: Bangladesh	Original files' average polarity is: 4.347826	Calculated polarity is: 26.0
Location is: Greece	Original files' average polarity is: 9.410289	Calculated polarity is: 28.0
Location is:	Original files' average polarity is:	Calculated polarity is: 38.0

Sweden	5.1282053	
Location is: Pakistan	Original files' average polarity is: 5.0890584	Calculated polarity is: 43.0

d. Tables A4.0-4.3:

Table A4.0: Calculated sentiment values for locations using larger dataset vs.

Original GDELT average values

Location	Original File Result (Average)	Calculated Result
Location is: Germany	Original files' average tone is: -2.3911307	Calculated result is: -3.370785
Location is: United States	Original files' average tone is: -1.8895149	Calculated result is: -3.2793386
Location is: China	Original files' average tone is: -3.5094118	Calculated result is: -3.3458247
Location is: India	Original files' average tone is: -3.4675782	Calculated result is: -3.3639889
Location is: United Kingdom	Original files' average tone is: -1.8051269	Calculated result is: -3.3707845
Location is: Canada	Original files' average tone is: -0.6047509	Calculated result is: -3.3707838
Location is: Mexico	Original files' average tone is: -1.0764394	Calculated result is: -3.3707845
Location is: Brazil	Original files' average tone is: -3.1127605	Calculated result is: -3.370785
Location is: Colombia	Original files' average tone is: -1.3961606	Calculated result is: -3.3707845
Location is: Israel	Original files' average tone is: -1.2765633	Calculated result is: -3.3707843
Location is: Afghanistan	Original files' average tone is: -1.931876	Calculated result is: -3.370784
Location is: Djibouti	Original files' average tone is: -1.931876	Calculated result is: -3.370784
Location is: West Bank	Original files' average tone is: -1.931876	Calculated result is: -3.370784
Location is: Nigeria	Original files' average tone is: -0.6820391	Calculated result is: -3.370785
Location is: South Africa	Original files' average tone is: 0.8006354	Calculated result is: -3.3707848
Location is: France	Original files' average tone is: -3.7385807	Calculated result is: -3.370785
Location is: Congo	Original files' average tone is: 0.990099	Calculated result is: -3.370785
Location is: Netherlands	Original files' average tone is: 2.198953	Calculated result is: -3.3707833
Location is: Malaysia	Original files' average tone is: 0.5208333	Calculated result is: -3.3707852
Location is: Chile	Original files' average tone is: 0.5208333	Calculated result is: -3.3707852
Location is: South Korea	Original files' average tone is: 0.5208333	Calculated result is: -3.3707852
Location is: Australia	Original files' average tone is: 3.6663735	Calculated result is: -3.3707836
Location is: Indonesia	Original files' average tone is: 0.7368421	Calculated result is: -3.3707862
Location is: Singapore	Original files' average tone is: -1.0656092	Calculated result is: -3.3707848
Location is: Zimbabwe	Original files' average tone is: -4.195804	Calculated result is: -3.3707838

Location is: Ukraine	Original files' average tone is: -2.1978023	Calculated result is: -3.370785
Location is: Switzerland	Original files' average tone is: -2.877698	Calculated result is: -3.3707845
Location is: Jamaica	Original files' average tone is: 3.2355497	Calculated result is: -3.3707838
Location is: Greenland	Original files' average tone is: 0.9493671	Calculated result is: -3.370784
Location is: Denmark	Original files' average tone is: 0.9493671	Calculated result is: -3.370784
Location is: Russia	Original files' average tone is: -0.731189	Calculated result is: -3.3707852
Location is: Peru	Original files' average tone is: -6.5459604	Calculated result is: -3.3707862
Location is: Bolivia	Original files' average tone is: -6.545961	Calculated result is: -3.3707855
Location is: Japan	Original files' average tone is: -1.2354937	Calculated result is: -3.3707843
Location is: Hong Kong	Original files' average tone is: -5.0700054	Calculated result is: -3.370785
Location is: Iran	Original files' average tone is: -4.9603176	Calculated result is: -3.3707857
Location is: Italy	Original files' average tone is: -4.438955	Calculated result is: -3.3707852
Location is: New Zealand	Original files' average tone is: -1.6958852	Calculated result is: -3.370784
Location is: Poland	Original files' average tone is: 6.0465117	Calculated result is: -3.3707848
Location is: Turkey	Original files' average tone is: -0.7751938	Calculated result is: -3.3707845
Location is: Honduras	Original files' average tone is: -2.0697167	Calculated result is: -3.3707838
Location is: El Salvador	Original files' average tone is: -2.0697167	Calculated result is: -3.3707838
Location is: Guatemala	Original files' average tone is: -2.0697167	Calculated result is: -3.3707838
Location is: Belgium	Original files' average tone is: -4.178273	Calculated result is: -3.370785
Location is: Kenya	Original files' average tone is: -6.5656567	Calculated result is: -3.3707848
Location is: Austria	Original files' average tone is: -6.5656567	Calculated result is: -3.3707848

Table A4.1: Calculated negative sentiment values for locations using larger dataset

vs. Original GDELT sentiment values

Location	Original File Result (Negativity - Average)	Calculated Result (Negativity)
Location is: Germany	Original files' average negativity is: -5.208263	Calculated negativity is: -6.629212
Location is: United States	Original files' average negativity is: -3.9072924	Calculated negativity is: -6.449366
Location is: China	Original files' average negativity is: -5.543452	Calculated negativity is: -6.5801244
Location is: India	Original files' average negativity is: -5.2850695	Calculated negativity is: -6.6158466
Location is: United Kingdom	Original files' average negativity is: -4.1516275	Calculated negativity is: -6.629212
Location is: Canada	Original files' average negativity is: -3.0672312	Calculated negativity is: -6.629212
Location is: Mexico	Original files' average negativity is: -2.4530036	Calculated negativity is: -6.629212
Location is: Brazil	Original files' average negativity is: -3.9105785	Calculated negativity is: -6.629212
Location is: Colombia	Original files' average negativity is: -1.9197208	Calculated negativity is: -6.629212
Location is: Israel	Original files' average negativity is: -3.4532435	Calculated negativity is: -6.629212
Location is: Afghanistan	Original files' average negativity is: -3.7620742	Calculated negativity is: -6.629212
Location is: Djibouti	Original files' average negativity is: -3.7620742	Calculated negativity is: -6.629212
Location is: West Bank	Original files' average negativity is: -3.7620742	Calculated negativity is: -6.629212
Location is: Nigeria	Original files' average negativity is: -3.3942556	Calculated negativity is: -6.6292124
Location is: South Africa	Original files' average negativity is: -0.88959134	Calculated negativity is: -6.629212
Location is: France	Original files' average negativity is: -5.7037525	Calculated negativity is: -6.629212
Location is: Congo	Original files' average negativity is: -0.4950495	Calculated negativity is: -6.629212
Location is: Netherlands	Original files' average negativity is: -1.4659686	Calculated negativity is: -6.629212
Location is: Malaysia	Original files' average negativity is: -2.2569444	Calculated negativity is: -6.629212
Location is: Chile	Original files' average negativity is: -2.2569444	Calculated negativity is: -6.629212
Location is: South Korea	Original files' average negativity is: -2.2569444	Calculated negativity is: -6.629212
Location is: Australia	Original files' average negativity is: -1.0215389	Calculated negativity is: -6.629212
Location is: Indonesia	Original files' average negativity is: -1.3684211	Calculated negativity is: -6.629212
Location is: Singapore	Original files' average negativity is: -3.62498	Calculated negativity is: -6.629212
Location is: Zimbabwe	Original files' average negativity is: -4.895105	Calculated negativity is: -6.629212
Location is:	Original files' average negativity is: -	Calculated negativity is: -

Ukraine	3.8461537	6.629212
Location is: Switzerland	Original files' average negativity is: -5.035971	Calculated negativity is: -6.629212
Location is: Jamaica	Original files' average negativity is: -1.0386789	Calculated negativity is: -6.629212
Location is: Greenland	Original files' average negativity is: -1.7405063	Calculated negativity is: -6.629212
Location is: Denmark	Original files' average negativity is: -1.7405063	Calculated negativity is: -6.629212
Location is: Russia	Original files' average negativity is: -4.3588533	Calculated negativity is: -6.629212
Location is: Peru	Original files' average negativity is: -7.8922935	Calculated negativity is: -6.6292124
Location is: Bolivia	Original files' average negativity is: -7.8922935	Calculated negativity is: -6.629212
Location is: Japan	Original files' average negativity is: -3.5391152	Calculated negativity is: -6.629212
Location is: Hong Kong	Original files' average negativity is: -7.1094604	Calculated negativity is: -6.6292124
Location is: Iran	Original files' average negativity is: -7.4404764	Calculated negativity is: -6.629212
Location is: Italy	Original files' average negativity is: -6.751282	Calculated negativity is: -6.629212
Location is: New Zealand	Original files' average negativity is: -3.5875041	Calculated negativity is: -6.629212
Location is: Poland	Original files' average negativity is: -0.81395346	Calculated negativity is: -6.629212
Location is: Turkey	Original files' average negativity is: -0.7751938	Calculated negativity is: -6.629212
Location is: Honduras	Original files' average negativity is: -5.2287583	Calculated negativity is: -6.629212
Location is: El Salvador	Original files' average negativity is: -5.2287583	Calculated negativity is: -6.629212
Location is: Guatemala	Original files' average negativity is: -5.2287583	Calculated negativity is: -6.629212
Location is: Belgium	Original files' average negativity is: -6.4066854	Calculated negativity is: -6.629212
Location is: Kenya	Original files' average negativity is: -7.323232	Calculated negativity is: -6.629212
Location is: Austria	Original files' average negativity is: -7.323232	Calculated negativity is: -6.629212

Table A4.2: Calculated positive sentiment values for locations using larger dataset

vs. Original GDELT sentiment values

Location	Original File Result (Positivity - Average)	Calculated Result (Positivity)
Location is: Germany	Original files' average positivity is: 2.8171325	Calculated positivity is: 3.258427
Location is: United States	Original files' average positivity is: 2.0177817	Calculated positivity is: 3.1700275
Location is: China	Original files' average positivity is: 2.0340402	Calculated positivity is: 3.2342997
Location is: India	Original files' average positivity is: 1.8174914	Calculated positivity is: 3.2518578
Location is: United Kingdom	Original files' average positivity is: 2.3465002	Calculated positivity is: 3.2584274
Location is: Canada	Original files' average positivity is: 2.4624798	Calculated positivity is: 3.258428
Location is: Mexico	Original files' average positivity is: 1.3765641	Calculated positivity is: 3.2584274
Location is: Brazil	Original files' average positivity is: 0.79781765	Calculated positivity is: 3.258427
Location is: Colombia	Original files' average positivity is: 0.5235602	Calculated positivity is: 3.2584274
Location is: Israel	Original files' average positivity is: 2.1766803	Calculated positivity is: 3.2584276
Location is: Afghanistan	Original files' average positivity is: 1.8301983	Calculated positivity is: 3.2584279
Location is: Djibouti	Original files' average positivity is: 1.8301983	Calculated positivity is: 3.2584279
Location is: West Bank	Original files' average positivity is: 1.8301983	Calculated positivity is: 3.2584279
Location is: Nigeria	Original files' average positivity is: 2.7122164	Calculated positivity is: 3.2584274
Location is: South Africa	Original files' average positivity is: 1.6902268	Calculated positivity is: 3.2584271
Location is: France	Original files' average positivity is: 1.9651715	Calculated positivity is: 3.258427
Location is: Congo	Original files' average positivity is: 1.4851485	Calculated positivity is: 3.258427
Location is: Netherlands	Original files' average positivity is: 3.6649215	Calculated positivity is: 3.2584286
Location is: Malaysia	Original files' average positivity is: 2.7777777	Calculated positivity is: 3.2584267
Location is: Chile	Original files' average positivity is: 2.7777777	Calculated positivity is: 3.2584267
Location is: South Korea	Original files' average positivity is: 2.7777777	Calculated positivity is: 3.2584267
Location is: Australia	Original files' average positivity is: 4.687912	Calculated positivity is: 3.2584283
Location is: Indonesia	Original files' average positivity is: 2.1052632	Calculated positivity is: 3.2584257
Location is: Singapore	Original files' average positivity is: 2.559371	Calculated positivity is: 3.2584271
Location is: Zimbabwe	Original files' average positivity is: 0.6993007	Calculated positivity is: 3.258428
Location is:	Original files' average positivity is:	Calculated positivity is:

Ukraine	1.6483517	3.258427
Location is: Switzerland	Original files' average positivity is: 2.1582735	Calculated positivity is: 3.2584274
Location is: Jamaica	Original files' average positivity is: 4.2742286	Calculated positivity is: 3.258428
Location is: Greenland	Original files' average positivity is: 2.6898735	Calculated positivity is: 3.2584279
Location is: Denmark	Original files' average positivity is: 2.6898735	Calculated positivity is: 3.2584279
Location is: Russia	Original files' average positivity is: 3.627664	Calculated positivity is: 3.2584267
Location is: Peru	Original files' average positivity is: 1.3463324	Calculated positivity is: 3.2584262
Location is: Bolivia	Original files' average positivity is: 1.3463324	Calculated positivity is: 3.2584264
Location is: Japan	Original files' average positivity is: 2.3036215	Calculated positivity is: 3.2584276
Location is: Hong Kong	Original files' average positivity is: 2.039455	Calculated positivity is: 3.2584274
Location is: Iran	Original files' average positivity is: 2.4801588	Calculated positivity is: 3.2584262
Location is: Italy	Original files' average positivity is: 2.3123276	Calculated positivity is: 3.2584267
Location is: New Zealand	Original files' average positivity is: 1.8916187	Calculated positivity is: 3.2584279
Location is: Poland	Original files' average positivity is: 6.860465	Calculated positivity is: 3.2584271
Location is: Turkey	Original files' average positivity is: 0.0	Calculated positivity is: 3.2584274
Location is: Honduras	Original files' average positivity is: 3.1590414	Calculated positivity is: 3.258428
Location is: El Salvador	Original files' average positivity is: 3.1590414	Calculated positivity is: 3.258428
Location is: Guatemala	Original files' average positivity is: 3.1590414	Calculated positivity is: 3.258428
Location is: Belgium	Original files' average positivity is: 2.2284122	Calculated positivity is: 3.258427
Location is: Kenya	Original files' average positivity is: 0.75757575	Calculated positivity is: 3.2584271
Location is: Austria	Original files' average positivity is: 0.75757575	Calculated positivity is: 3.2584271

Table A4.3: Calculated polarity values for locations using larger dataset vs. Original

polarity value averages

Location	Original File Result (Polarization - Average)	Calculated Result (Polarization - different calculation than original)
Location is: Germany	Original files' average polarity is: 8.025395	Calculated polarity is: 56.714287
Location is: United States	Original files' average polarity is: 5.925073	Calculated polarity is: 37.55634
Location is: China	Original files' average polarity is: 7.577492	Calculated polarity is: 47.75
Location is: India	Original files' average polarity is: 7.1025605	Calculated polarity is: 44.25
Location is: United Kingdom	Original files' average polarity is: 6.4981265	Calculated polarity is: 34.136364
Location is: Canada	Original files' average polarity is: 5.5297112	Calculated polarity is: 25.59091
Location is: Mexico	Original files' average polarity is: 3.8295674	Calculated polarity is: 33.5
Location is: Brazil	Original files' average polarity is: 4.7083955	Calculated polarity is: 55.333332
Location is: Colombia	Original files' average polarity is: 2.443281	Calculated polarity is: 50.0
Location is: Israel	Original files' average polarity is: 5.6299233	Calculated polarity is: 32.0
Location is: Afghanistan	Original files' average polarity is: 5.5922723	Calculated polarity is: 27.0
Location is: Djibouti	Original files' average polarity is: 5.5922723	Calculated polarity is: 27.0
Location is: West Bank	Original files' average polarity is: 5.5922723	Calculated polarity is: 27.0
Location is: Nigeria	Original files' average polarity is: 6.106472	Calculated polarity is: 20.75
Location is: South Africa	Original files' average polarity is: 2.5798182	Calculated polarity is: 19.5
Location is: France	Original files' average polarity is: 7.668924	Calculated polarity is: 53.0
Location is: Congo	Original files' average polarity is: 1.980198	Calculated polarity is: 16.0
Location is: Netherlands	Original files' average polarity is: 5.13089	Calculated polarity is: 36.0
Location is: Malaysia	Original files' average polarity is: 5.0347223	Calculated polarity is: 59.0
Location is: Chile	Original files' average polarity is: 5.0347223	Calculated polarity is: 59.0
Location is: South Korea	Original files' average polarity is: 5.0347223	Calculated polarity is: 59.0
Location is: Australia	Original files' average polarity is: 5.7094507	Calculated polarity is: 41.285713
Location is: Indonesia	Original files' average polarity is: 3.4736843	Calculated polarity is: 77.0
Location is: Singapore	Original files' average polarity is: 6.184351	Calculated polarity is: 54.0
Location is: Zimbabwe	Original files' average polarity is: 5.5944057	Calculated polarity is: 29.0

Location is: Ukraine	Original files' average polarity is: 5.4945054	Calculated polarity is: 15.0
Location is: Switzerland	Original files' average polarity is: 7.1942444	Calculated polarity is: 16.0
Location is: Jamaica	Original files' average polarity is: 5.312907	Calculated polarity is: 28.0
Location is: Greenland	Original files' average polarity is: 4.43038	Calculated polarity is: 27.0
Location is: Denmark	Original files' average polarity is: 4.43038	Calculated polarity is: 27.0
Location is: Russia	Original files' average polarity is: 7.986517	Calculated polarity is: 56.0
Location is: Peru	Original files' average polarity is: 9.238626	Calculated polarity is: 66.0
Location is: Bolivia	Original files' average polarity is: 9.238626	Calculated polarity is: 66.0
Location is: Japan	Original files' average polarity is: 5.8427367	Calculated polarity is: 48.666668
Location is: Hong Kong	Original files' average polarity is: 9.148915	Calculated polarity is: 34.6
Location is: Iran	Original files' average polarity is: 9.920635	Calculated polarity is: 72.0
Location is: Italy	Original files' average polarity is: 9.06361	Calculated polarity is: 61.333332
Location is: New Zealand	Original files' average polarity is: 5.479122	Calculated polarity is: 45.333332
Location is: Poland	Original files' average polarity is: 7.6744184	Calculated polarity is: 53.0
Location is: Turkey	Original files' average polarity is: 0.7751938	Calculated polarity is: 17.0
Location is: Honduras	Original files' average polarity is: 8.387799	Calculated polarity is: 44.0
Location is: El Salvador	Original files' average polarity is: 8.387799	Calculated polarity is: 44.0
Location is: Guatemala	Original files' average polarity is: 8.387799	Calculated polarity is: 44.0
Location is: Belgium	Original files' average polarity is: 8.6350975	Calculated polarity is: 56.0
Location is: Kenya	Original files' average polarity is: 8.080808	Calculated polarity is: 18.0
Location is: Austria	Original files' average polarity is: 8.080808	Calculated polarity is: 18.0

e. Tables A5.0-A5.3:

Table A5.0: Calculated sentiment values using clusters vs. Original tone values

File Number	Original File Result	Calculated Result
0	Original file's tone is: -3.29512893982808	Calculated result is: -9.734554
1	Original file's tone is: 3.68349249658936	Calculated result is: -9.2402525
2	Original file's tone is: 1.42302716688228	Calculated result is: -9.779709
3	Original file's tone is: 1.15942028985507	Calculated result is: -9.325769
4	Original file's tone is: -1.93187595322827	Calculated result is: -9.828964
5	Original file's tone is: 1.41176470588235	Calculated result is: -9.848506
6	Original file's tone is: 3.96825396825397	Calculated result is: -9.393738
7	Original file's tone is: 3.6096256684492	Calculated result is: -9.848506
8	Original file's tone is: -0.67842605156038	Calculated result is: -9.626197
9	Original file's tone is: -2.07715133531157	Calculated result is: -9.848506
10	Original file's tone is: -4.04040404040404	Calculated result is: -9.848506
11	Original file's tone is: 0.99009900990099	Calculated result is: -9.812429
12	Original file's tone is: -3.16490838423098	Calculated result is: -9.759781
13	Original file's tone is: -1.91176470588235	Calculated result is: -9.848506
14	Original file's tone is: 2.19895287958115	Calculated result is: -9.835479
15	Original file's tone is: 0.520833333333333	Calculated result is: -9.609956
16	Original file's tone is: 6.07734806629834	Calculated result is: -9.525196
17	Original file's tone is: -5.4270462633452	Calculated result is: -9.658428
18	Original file's tone is: 1.81818181818182	Calculated result is: -9.500324
19	Original file's tone is: -1.77215189873418	Calculated result is: -9.640662
20	Original file's tone is: -4.1958041958042	Calculated result is: -9.565773
21	Original file's tone is: 0.193050193050192	Calculated result is: -9.724231
22	Original file's tone is: -0.78740157480315	Calculated result is: -9.490377
23	Original file's tone is: -3.88349514563107	Calculated result is: -9.379526
24	Original file's tone is: 0.888888888888889	Calculated result is: -9.829745
25	Original file's tone is: -5.26315789473684	Calculated result is: -9.191935
26	Original file's tone is:	Calculated result is: -9.757317

	1.97132616487455	
27	Original file's tone is: 2	Calculated result is: -9.510421
28	Original file's tone is: -3.80499405469679	Calculated result is: -9.686166
29	Original file's tone is: -2.1978021978022	Calculated result is: -9.8094225
30	Original file's tone is: -2.87769784172662	Calculated result is: -9.675724
31	Original file's tone is: -2.13178294573643	Calculated result is: -9.7165785
32	Original file's tone is: 0.641025641025641	Calculated result is: -9.616384
33	Original file's tone is: -0.23696682464455	Calculated result is: -9.848506
34	Original file's tone is: 6.21468926553672	Calculated result is: -9.444054
35	Original file's tone is: -3.83036935704514	Calculated result is: -9.517527
36	Original file's tone is: 0.949367088607595	Calculated result is: -9.750801
37	Original file's tone is: -6.54596100278552	Calculated result is: -9.727161
38	Original file's tone is: -3.52112676056338	Calculated result is: -9.834712
39	Original file's tone is: -7.50636132315522	Calculated result is: -9.848506
40	Original file's tone is: -1.67014613778706	Calculated result is: -9.65104
41	Original file's tone is: -0.294550810014728	Calculated result is: -9.509368
42	Original file's tone is: -4.18848167539267	Calculated result is: -9.658588
43	Original file's tone is: 0.857142857142857	Calculated result is: -9.80587
44	Original file's tone is: 1.27551020408163	Calculated result is: -9.848506
45	Original file's tone is: -2.9723991507431	Calculated result is: -9.848506
46	Original file's tone is: -7.92682926829268	Calculated result is: -9.848506
47	Original file's tone is: 0.256410256410256	Calculated result is: -9.80587
48	Original file's tone is: -5.01792114695341	Calculated result is: -9.660913
49	Original file's tone is: -4.78723404255319	Calculated result is: -9.400845
50	Original file's tone is: 7.88608981380066	Calculated result is: -9.666529
51	Original file's tone is: -0.775193798449612	Calculated result is: -9.629649
52	Original file's tone is: -2.06971677559913	Calculated result is: -9.6140175
53	Original file's tone is: -4.17827298050139	Calculated result is: -9.848506
54	Original file's tone is: -0.359066427289048	Calculated result is: -9.622658
55	Original file's tone is: -3.52112676056338	Calculated result is: -9.472211

56	Original file's tone is: -3.52112676056338	Calculated result is: -9.472211
57	Original file's tone is: -4	Calculated result is: -9.675781
58	Original file's tone is: -6.56565656565657	Calculated result is: -9.643328

Table A5.1: Calculated negative sentiment values using clusters vs. Original

negative tone values

File Number	Original File Result (Negativity)	Calculated Result (Negativity)
0	Original file's negativity is: 5.58739255014327	Calculated negativity is: -10.857763
1	Original file's negativity is: 1.63710777626194	Calculated negativity is: -10.306429
2	Original file's negativity is: 0.517464424320828	Calculated negativity is: -10.908128
3	Original file's negativity is: 2.17391304347826	Calculated negativity is: -10.401813
4	Original file's negativity is: 3.76207422470768	Calculated negativity is: -10.963066
5	Original file's negativity is: 0.470588235294118	Calculated negativity is: -10.984862
6	Original file's negativity is: 0.198412698412698	Calculated negativity is: -10.477623
7	Original file's negativity is: 0.935828877005348	Calculated negativity is: -10.984862
8	Original file's negativity is: 1.22116689280868	Calculated negativity is: -10.736903
9	Original file's negativity is: 3.26409495548961	Calculated negativity is: -10.984862
10	Original file's negativity is: 6.06060606060606	Calculated negativity is: -10.984862
11	Original file's negativity is: 0.495049504950495	Calculated negativity is: -10.944624
12	Original file's negativity is: 4.10882842865075	Calculated negativity is: -10.8859
13	Original file's negativity is: 5	Calculated negativity is: -10.984862
14	Original file's negativity is: 1.46596858638743	Calculated negativity is: -10.970332
15	Original file's negativity is: 2.25694444444444	Calculated negativity is: -10.718788
16	Original file's negativity is: 0.828729281767956	Calculated negativity is: -10.6242485
17	Original file's negativity is: 6.76156583629893	Calculated negativity is: -10.772853
18	Original file's negativity is: 0	Calculated negativity is: -10.596507
19	Original file's negativity is: 4.55696202531646	Calculated negativity is: -10.753037
20	Original file's negativity is: 4.89510489510489	Calculated negativity is: -10.669507
21	Original file's negativity is: 6.37065637065637	Calculated negativity is: -10.846249
22	Original file's negativity is: 1.83727034120735	Calculated negativity is: -10.585412
23	Original file's negativity is: 4.85436893203883	Calculated negativity is: -10.461771
24	Original file's negativity is: 0.444444444444444	Calculated negativity is: -10.963938
25	Original file's negativity is: 6.57894736842105	Calculated negativity is: -10.252535
26	Original file's negativity is: 0.716845878136201	Calculated negativity is: -10.883151

27	Original file's negativity is: 1.90909090909091	Calculated negativity is: -10.607769
28	Original file's negativity is: 5.82639714625446	Calculated negativity is: -10.803792
29	Original file's negativity is: 3.84615384615385	Calculated negativity is: -10.94127
30	Original file's negativity is: 5.03597122302158	Calculated negativity is: -10.792145
31	Original file's negativity is: 4.45736434108527	Calculated negativity is: -10.837714
32	Original file's negativity is: 2.24358974358974	Calculated negativity is: -10.725958
33	Original file's negativity is: 1.18483412322275	Calculated negativity is: -10.984862
34	Original file's negativity is: 0.282485875706215	Calculated negativity is: -10.533745
35	Original file's negativity is: 5.19835841313269	Calculated negativity is: -10.615695
36	Original file's negativity is: 1.74050632911392	Calculated negativity is: -10.875885
37	Original file's negativity is: 7.89229340761374	Calculated negativity is: -10.849517
38	Original file's negativity is: 6.8075117370892	Calculated negativity is: -10.969477
39	Original file's negativity is: 9.03307888040712	Calculated negativity is: -10.984862
40	Original file's negativity is: 3.34029227557411	Calculated negativity is: -10.764613
41	Original file's negativity is: 2.35640648011782	Calculated negativity is: -10.606594
42	Original file's negativity is: 5.23560209424084	Calculated negativity is: -10.773032
43	Original file's negativity is: 0.571428571428571	Calculated negativity is: -10.937308
44	Original file's negativity is: 1.02040816326531	Calculated negativity is: -10.984862
45	Original file's negativity is: 5.09554140127389	Calculated negativity is: -10.984862
46	Original file's negativity is: 9.14634146341463	Calculated negativity is: -10.984862
47	Original file's negativity is: 1.79487179487179	Calculated negativity is: -10.937308
48	Original file's negativity is: 7.34767025089606	Calculated negativity is: -10.775625
49	Original file's negativity is: 6.91489361702128	Calculated negativity is: -10.485549
50	Original file's negativity is: 0.657174151150055	Calculated negativity is: -10.781889
51	Original file's negativity is: 0.775193798449612	Calculated negativity is: -10.740753
52	Original file's negativity is: 5.22875816993464	Calculated negativity is: -10.723318
53	Original file's negativity is: 6.4066852367688	Calculated negativity is: -10.984862
54	Original file's negativity is: 2.69299820466786	Calculated negativity is: -10.732956
55	Original file's negativity is: 4.5271629778672	Calculated negativity is: -10.56515
56	Original file's negativity is:	Calculated negativity is: -10.56515

	4.5271629778672	
57	Original file's negativity is: 5.05263157894737	Calculated negativity is: -10.792209
58	Original file's negativity is: 7.32323232323232	Calculated negativity is: -10.75601

Table A5.2: Calculated positive values using clusters vs. Original positive tone

values

File Number	Original File Result (Positivity)	Calculated Result (Positivity)
0	Original file's positivity is: 2.29226361031519	Calculated positivity is: 1.1232088
1	Original file's positivity is: 5.3206002728513	Calculated positivity is: 1.066176
2	Original file's positivity is: 1.9404915912031	Calculated positivity is: 1.1284188
3	Original file's positivity is: 3.33333333333333	Calculated positivity is: 1.0760429
4	Original file's positivity is: 1.83019827147941	Calculated positivity is: 1.134102
5	Original file's positivity is: 1.88235294117647	Calculated positivity is: 1.1363564
6	Original file's positivity is: 4.16666666666667	Calculated positivity is: 1.083885
7	Original file's positivity is: 4.54545454545455	Calculated positivity is: 1.1363564
8	Original file's positivity is: 0.542740841248304	Calculated positivity is: 1.110706
9	Original file's positivity is: 1.18694362017804	Calculated positivity is: 1.1363564
10	Original file's positivity is: 2.02020202020202	Calculated positivity is: 1.1363564
11	Original file's positivity is: 1.48514851485149	Calculated positivity is: 1.1321944
12	Original file's positivity is: 0.943920044419767	Calculated positivity is: 1.1261183
13	Original file's positivity is: 3.08823529411765	Calculated positivity is: 1.1363564
14	Original file's positivity is: 3.66492146596859	Calculated positivity is: 1.1348535
15	Original file's positivity is: 2.77777777777778	Calculated positivity is: 1.1088325
16	Original file's positivity is: 6.9060773480663	Calculated positivity is: 1.0990524
17	Original file's positivity is: 1.33451957295374	Calculated positivity is: 1.1144251
18	Original file's positivity is: 1.81818181818182	Calculated positivity is: 1.0961832
19	Original file's positivity is: 2.78481012658228	Calculated positivity is: 1.1123754
20	Original file's positivity is: 0.699300699300699	Calculated positivity is: 1.1037343
21	Original file's positivity is: 6.56370656370656	Calculated positivity is: 1.1220175
22	Original file's positivity is: 1.0498687664042	Calculated positivity is: 1.0950351
23	Original file's positivity is: 0.970873786407767	Calculated positivity is: 1.0822449
24	Original file's positivity is: 1.33333333333333	Calculated positivity is: 1.1341922
25	Original file's positivity is: 1.31578947368421	Calculated positivity is: 1.0606003
26	Original file's positivity is:	Calculated positivity is: 1.1258347

	2.68817204301075	
27	Original file's positivity is: 3.90909090909091	Calculated positivity is: 1.0973483
28	Original file's positivity is: 2.02140309155767	Calculated positivity is: 1.117626
29	Original file's positivity is: 1.64835164835165	Calculated positivity is: 1.1318476
30	Original file's positivity is: 2.15827338129496	Calculated positivity is: 1.1164205
31	Original file's positivity is: 2.32558139534884	Calculated positivity is: 1.1211355
32	Original file's positivity is: 2.88461538461538	Calculated positivity is: 1.1095742
33	Original file's positivity is: 0.947867298578199	Calculated positivity is: 1.1363564
34	Original file's positivity is: 6.49717514124294	Calculated positivity is: 1.0896908
35	Original file's positivity is: 1.36798905608755	Calculated positivity is: 1.0981684
36	Original file's positivity is: 2.68987341772152	Calculated positivity is: 1.1250836
37	Original file's positivity is: 1.34633240482823	Calculated positivity is: 1.1223557
38	Original file's positivity is: 3.28638497652582	Calculated positivity is: 1.134765
39	Original file's positivity is: 1.52671755725191	Calculated positivity is: 1.1363564
40	Original file's positivity is: 1.67014613778706	Calculated positivity is: 1.113573
41	Original file's positivity is: 2.06185567010309	Calculated positivity is: 1.0972264
42	Original file's positivity is: 1.04712041884817	Calculated positivity is: 1.1144437
43	Original file's positivity is: 1.42857142857143	Calculated positivity is: 1.1314378
44	Original file's positivity is: 2.29591836734694	Calculated positivity is: 1.1363564
45	Original file's positivity is: 2.12314225053079	Calculated positivity is: 1.1363564
46	Original file's positivity is: 1.21951219512195	Calculated positivity is: 1.1363564
47	Original file's positivity is: 2.05128205128205	Calculated positivity is: 1.1314378
48	Original file's positivity is: 2.32974910394265	Calculated positivity is: 1.114712
49	Original file's positivity is: 2.12765957446809	Calculated positivity is: 1.0847045
50	Original file's positivity is: 8.54326396495071	Calculated positivity is: 1.1153598
51	Original file's positivity is: 0	Calculated positivity is: 1.1111044
52	Original file's positivity is: 3.15904139433551	Calculated positivity is: 1.1093006
53	Original file's positivity is: 2.22841225626741	Calculated positivity is: 1.1363564
54	Original file's positivity is: 2.33393177737881	Calculated positivity is: 1.110298
55	Original file's positivity is: 1.00603621730382	Calculated positivity is: 1.092939

56	Original file's positivity is: 1.00603621730382	Calculated positivity is: 1.092939
57	Original file's positivity is: 1.05263157894737	Calculated positivity is: 1.1164275
58	Original file's positivity is: 0.757575757575758	Calculated positivity is: 1.1126826

Table A5.3: Calculated polarity values using clusters vs. Original polarity values

File Number	Original File Result (Polarization)	Calculated Result (Polarization, different calculation than original)
0	Original file's polarity is: 7.87965616045845	Calculated polarity is: 0.1458886
1	Original file's polarity is: 6.95770804911323	Calculated polarity is: 0.011923688
2	Original file's polarity is: 2.45795601552393	Calculated polarity is: 0.08906882
3	Original file's polarity is: 5.50724637681159	Calculated polarity is: 0.110843375
4	Original file's polarity is: 5.59227249618709	Calculated polarity is: 0.021182701
5	Original file's polarity is: 2.35294117647059	Calculated polarity is: 0.024742268
6	Original file's polarity is: 4.36507936507936	Calculated polarity is: 0.05722071
7	Original file's polarity is: 5.48128342245989	Calculated polarity is: 0.011976048
8	Original file's polarity is: 1.76390773405699	Calculated polarity is: 0.06410257
9	Original file's polarity is: 4.45103857566766	Calculated polarity is: 0.04347826
10	Original file's polarity is: 8.08080808080808	Calculated polarity is: 0.04379562
11	Original file's polarity is: 1.98019801980198	Calculated polarity is: 0.08280255
12	Original file's polarity is: 5.05274847307052	Calculated polarity is: 0.017461067
13	Original file's polarity is: 8.08823529411765	Calculated polarity is: 0.024742268
14	Original file's polarity is: 5.13089005235602	Calculated polarity is: 0.036108326
15	Original file's polarity is: 5.03472222222222	Calculated polarity is: 0.16470589
16	Original file's polarity is: 7.73480662983425	Calculated polarity is: 0.17751479
17	Original file's polarity is: 8.09608540925267	Calculated polarity is: 0.091603056
18	Original file's polarity is: 1.81818181818182	Calculated polarity is: 0.077922076
19	Original file's polarity is: 7.34177215189873	Calculated polarity is: 0.20512821
20	Original file's polarity is: 5.59440559440559	Calculated polarity is: 0.1557377
21	Original file's polarity is: 12.9343629343629	Calculated polarity is: 0.11547912
22	Original file's polarity is: 2.88713910761155	Calculated polarity is: 0.05172414
23	Original file's polarity is: 5.8252427184466	Calculated polarity is: 0.020989506
24	Original file's polarity is: 1.77777777777778	Calculated polarity is: 0.10460251
25	Original file's polarity is: 7.89473684210526	Calculated polarity is: 0.030487806
26	Original file's polarity is: 3.40501792114695	Calculated polarity is: 0.10465116

27	Original file's polarity is: 5.81818181818182	Calculated polarity is: 0.053926207
28	Original file's polarity is: 7.84780023781213	Calculated polarity is: 0.043551087
29	Original file's polarity is: 5.49450549450549	Calculated polarity is: 0.053333335
30	Original file's polarity is: 7.19424460431655	Calculated polarity is: 0.058103975
31	Original file's polarity is: 6.78294573643411	Calculated polarity is: 0.09447415
32	Original file's polarity is: 5.12820512820513	Calculated polarity is: 0.033210333
33	Original file's polarity is: 2.13270142180095	Calculated polarity is: 0.057142857
34	Original file's polarity is: 6.77966101694915	Calculated polarity is: 0.13857678
35	Original file's polarity is: 6.56634746922025	Calculated polarity is: 0.06972477
36	Original file's polarity is: 4.43037974683544	Calculated polarity is: 0.048241206
37	Original file's polarity is: 9.23862581244197	Calculated polarity is: 0.050485436
38	Original file's polarity is: 10.093896713615	Calculated polarity is: 0.0648855
39	Original file's polarity is: 10.559796437659	Calculated polarity is: 0.013613861
40	Original file's polarity is: 5.01043841336117	Calculated polarity is: 0.080168776
41	Original file's polarity is: 4.41826215022091	Calculated polarity is: 0.06413994
42	Original file's polarity is: 6.282722513089	Calculated polarity is: 0.0882353
43	Original file's polarity is: 2	Calculated polarity is: 0.015580736
44	Original file's polarity is: 3.31632653061224	Calculated polarity is: 0.023054754
45	Original file's polarity is: 7.21868365180467	Calculated polarity is: 0.015010722
46	Original file's polarity is: 10.3658536585366	Calculated polarity is: 0.225
47	Original file's polarity is: 3.84615384615385	Calculated polarity is: 0.02606635
48	Original file's polarity is: 9.67741935483871	Calculated polarity is: 0.07782101
49	Original file's polarity is: 9.04255319148936	Calculated polarity is: 0.048672568
50	Original file's polarity is: 9.20043811610077	Calculated polarity is: 0.07649254
51	Original file's polarity is: 0.775193798449612	Calculated polarity is: 0.011135858
52	Original file's polarity is: 8.38779956427015	Calculated polarity is: 0.061269145
53	Original file's polarity is: 8.63509749303621	Calculated polarity is: 0.08108108
54	Original file's polarity is: 5.02692998204668	Calculated polarity is: 0.13962264
55	Original file's polarity is: 5.53319919517103	Calculated polarity is: 0.056234717
56	Original file's polarity is: 5.53319919517103	Calculated polarity is: 0.056234717

| 57 | Original file's polarity is: 6.10526315789474 | Calculated polarity is: 0.1331058 |
| 58 | Original file's polarity is: 8.08080808080808 | Calculated polarity is: 0.083333336 |

f. Tables A6.0-A6.3:

Table A6.0: Calculated sentiment values for locations using clusters vs. Original

GDELT average tone values

Location	Original File Result (Average)	Calculated Result
Location is: China	Original files' average tone is: -1.3909457	Calculated result is: -3.5340407
Location is: Canada	Original files' average tone is: -0.4365438	Calculated result is: -3.6735737
Location is: United Kingdom	Original files' average tone is: -0.7342701	Calculated result is: -3.696995
Location is: United States	Original files' average tone is: -1.1494569	Calculated result is: -3.6296322
Location is: Brazil	Original files' average tone is: -6.534622	Calculated result is: -3.6776626
Location is: Pakistan	Original files' average tone is: 2.7128341	Calculated result is: -3.831054
Location is: Jersey	Original files' average tone is: -3.7920845	Calculated result is: -3.6868412
Location is: Ireland	Original files' average tone is: 0.14568973	Calculated result is: -3.6953187
Location is: Egypt	Original files' average tone is: -2.7211466	Calculated result is: -3.6918263
Location is: Germany	Original files' average tone is: -3.2287197	Calculated result is: -3.669894
Location is: Japan	Original files' average tone is: -1.6279894	Calculated result is: -3.6407628
Location is: Italy	Original files' average tone is: -4.256711	Calculated result is: -3.6675048
Location is: France	Original files' average tone is: -2.8513541	Calculated result is: -3.6568756
Location is: Australia	Original files' average tone is: 0.42652646	Calculated result is: -3.6155734
Location is: New Zealand	Original files' average tone is: -0.007379815	Calculated result is: -3.624111
Location is: South Korea	Original files' average tone is: -0.4518088	Calculated result is: -3.6138105
Location is: North Korea	Original files' average tone is: -1.6435986	Calculated result is: -3.5527492
Location is: Russia	Original files' average tone is: -0.27747166	Calculated result is: -3.6299043
Location is: Taiwan	Original files' average tone is: -2.7149322	Calculated result is: -3.7227817
Location is: Switzerland	Original files' average tone is: -2.7149322	Calculated result is: -3.7227817
Location is: Belgium	Original files' average tone is: -1.0752689	Calculated result is: -3.695092
Location is: Netherlands	Original files' average tone is: -0.93544227	Calculated result is: -3.741347
Location is: Croatia	Original files' average tone is: -1.0752689	Calculated result is: -3.695092
Location is: Philippines	Original files' average tone is: -0.15082957	Calculated result is: -3.50647
Location is: India	Original files' average tone is: -0.6280015	Calculated result is: -3.615723

Location is: Sweden	Original files' average tone is: 2.5910711	Calculated result is: -3.6847572
Location is: Honduras	Original files' average tone is: -2.527646	Calculated result is: -3.6211138
Location is: Guatemala	Original files' average tone is: -2.527646	Calculated result is: -3.6211138
Location is: El Salvador	Original files' average tone is: -2.527646	Calculated result is: -3.6211138
Location is: Portugal	Original files' average tone is: 3.6779325	Calculated result is: -3.6637623
Location is: Greenland	Original files' average tone is: -1.7268445	Calculated result is: -3.5671215
Location is: Denmark	Original files' average tone is: -1.7268445	Calculated result is: -3.5671215
Location is: Syria	Original files' average tone is: -1.7268445	Calculated result is: -3.5671215
Location is: Mexico	Original files' average tone is: -0.13229895	Calculated result is: -3.6407142
Location is: Madagascar	Original files' average tone is: -0.04405286	Calculated result is: -3.724013
Location is: Bolivia	Original files' average tone is: -8.655333	Calculated result is: -3.6290138
Location is: Paraguay	Original files' average tone is: -8.655333	Calculated result is: -3.6290138
Location is: Peru	Original files' average tone is: -4.4442163	Calculated result is: -3.636307
Location is: Norway	Original files' average tone is: -2.6415095	Calculated result is: -3.9376667
Location is: Iran	Original files' average tone is: -4.343048	Calculated result is: -3.670866
Location is: Trinidad And Tobago	Original files' average tone is: 0.8152174	Calculated result is: -3.5595593
Location is: Malaysia	Original files' average tone is: -0.23310024	Calculated result is: -3.6435997
Location is: Thailand	Original files' average tone is: 0.73998094	Calculated result is: -3.6748714
Location is: Singapore	Original files' average tone is: 0.73998094	Calculated result is: -3.6748714
Location is: Republic Of	Original files' average tone is: -0.23310024	Calculated result is: -3.6435997
Location is: Brunei	Original files' average tone is: -0.23310024	Calculated result is: -3.6435997
Location is: Chile	Original files' average tone is: 0.24614911	Calculated result is: -3.5892537
Location is: Ukraine	Original files' average tone is: 2.3046093	Calculated result is: -3.675734
Location is: Kenya	Original files' average tone is: -0.81135905	Calculated result is: -3.6855125
Location is: Malawi	Original files' average tone is: -0.81135905	Calculated result is: -3.6855125
Location is: Ghana	Original files' average tone is: -0.81135905	Calculated result is: -3.6855125

Table A6.1: Calculated negative sentiment values for locations using clusters vs.

Original GDELT average negative tone values

Location	Original File Result (Negativity - Average)	Calculated Result (Negativity)
Location is: China	Original files' average negativity is: -3.2600229	Calculated negativity is: -6.1042337
Location is: Canada	Original files' average negativity is: -2.5134428	Calculated negativity is: -6.3452435
Location is: United Kingdom	Original files' average negativity is: -3.6720748	Calculated negativity is: -6.3856997
Location is: United States	Original files' average negativity is: -3.8065732	Calculated negativity is: -6.2693424
Location is: Brazil	Original files' average negativity is: -7.4426174	Calculated negativity is: -6.3523088
Location is: Pakistan	Original files' average negativity is: -1.0752689	Calculated negativity is: -6.617254
Location is: Jersey	Original files' average negativity is: -4.823012	Calculated negativity is: -6.3681617
Location is: Ireland	Original files' average negativity is: -3.247582	Calculated negativity is: -6.3828073
Location is: Egypt	Original files' average negativity is: -4.4731584	Calculated negativity is: -6.376771
Location is: Germany	Original files' average negativity is: -5.0939956	Calculated negativity is: -6.3388877
Location is: Japan	Original files' average negativity is: -4.21349	Calculated negativity is: -6.288569
Location is: Italy	Original files' average negativity is: -5.1725693	Calculated negativity is: -6.3347588
Location is: France	Original files' average negativity is: -4.731679	Calculated negativity is: -6.3164
Location is: Australia	Original files' average negativity is: -2.4847982	Calculated negativity is: -6.2450604
Location is: New Zealand	Original files' average negativity is: -2.4534054	Calculated negativity is: -6.2598095
Location is: South Korea	Original files' average negativity is: -3.6784513	Calculated negativity is: -6.242014
Location is: North Korea	Original files' average negativity is: -5.363322	Calculated negativity is: -6.136547
Location is: Russia	Original files' average negativity is: -3.410898	Calculated negativity is: -6.2698164
Location is: Taiwan	Original files' average negativity is: -3.9215686	Calculated negativity is: -6.4302406
Location is: Switzerland	Original files' average negativity is: -3.9215686	Calculated negativity is: -6.4302406
Location is: Belgium	Original files' average negativity is: -4.435484	Calculated negativity is: -6.3824134
Location is: Netherlands	Original files' average negativity is: -4.027888	Calculated negativity is: -6.4623075
Location is: Croatia	Original files' average negativity is: -4.435484	Calculated negativity is: -6.3824134
Location is: Philippines	Original files' average negativity is: -3.6199095	Calculated negativity is: -6.0566134
Location is: India	Original files' average negativity is: -2.929152	Calculated negativity is: -6.245319
Location is:	Original files' average negativity is: -	Calculated negativity is: -6.364555

Sweden	2.0827076	
Location is: Honduras	Original files' average negativity is: -5.21327	Calculated negativity is: -6.2546344
Location is: Guatemala	Original files' average negativity is: -5.21327	Calculated negativity is: -6.2546344
Location is: El Salvador	Original files' average negativity is: -5.21327	Calculated negativity is: -6.2546344
Location is: Portugal	Original files' average negativity is: -1.689861	Calculated negativity is: -6.3282995
Location is: Greenland	Original files' average negativity is: -4.2386184	Calculated negativity is: -6.161377
Location is: Denmark	Original files' average negativity is: -4.2386184	Calculated negativity is: -6.161377
Location is: Syria	Original files' average negativity is: -4.2386184	Calculated negativity is: -6.161377
Location is: Mexico	Original files' average negativity is: -2.3295915	Calculated negativity is: -6.2884836
Location is: Madagascar	Original files' average negativity is: -3.0396476	Calculated negativity is: -6.432367
Location is: Bolivia	Original files' average negativity is: -8.964452	Calculated negativity is: -6.26828
Location is: Paraguay	Original files' average negativity is: -8.964452	Calculated negativity is: -6.26828
Location is: Peru	Original files' average negativity is: -5.2980766	Calculated negativity is: -6.280875
Location is: Norway	Original files' average negativity is: -3.7735848	Calculated negativity is: -6.801407
Location is: Iran	Original files' average negativity is: -5.9080324	Calculated negativity is: -6.340565
Location is: Trinidad And Tobago	Original files' average negativity is: -0.54347825	Calculated negativity is: -6.14831
Location is: Malaysia	Original files' average negativity is: -1.6317016	Calculated negativity is: -6.29347
Location is: Thailand	Original files' average negativity is: -1.993581	Calculated negativity is: -6.3474803
Location is: Singapore	Original files' average negativity is: -1.993581	Calculated negativity is: -6.3474803
Location is: Republic Of	Original files' average negativity is: -1.6317016	Calculated negativity is: -6.29347
Location is: Brunei	Original files' average negativity is: -1.6317016	Calculated negativity is: -6.29347
Location is: Chile	Original files' average negativity is: -1.0065422	Calculated negativity is: -6.199599
Location is: Ukraine	Original files' average negativity is: -1.002004	Calculated negativity is: -6.348975
Location is: Kenya	Original files' average negativity is: -4.665314	Calculated negativity is: -6.3658686
Location is: Malawi	Original files' average negativity is: -4.665314	Calculated negativity is: -6.3658686
Location is: Ghana	Original files' average negativity is: -4.665314	Calculated negativity is: -6.3658686

Table A6.2: Calculated positive sentiment values for locations using clusters vs.

Original GDELT average positive tone values

Location	Original File Result (Positivity - Average)	Calculated Result (Positivity)
Location is: China	Original files' average positivity is: 1.8690774	Calculated positivity is: 2.570193
Location is: Canada	Original files' average positivity is: 2.0768988	Calculated positivity is: 2.6716697
Location is: United Kingdom	Original files' average positivity is: 2.9378045	Calculated positivity is: 2.6887047
Location is: United States	Original files' average positivity is: 2.657113	Calculated positivity is: 2.6397102
Location is: Brazil	Original files' average positivity is: 0.9079947	Calculated positivity is: 2.6746461
Location is: Pakistan	Original files' average positivity is: 3.7881029	Calculated positivity is: 2.7861998
Location is: Jersey	Original files' average positivity is: 1.0309278	Calculated positivity is: 2.6813204
Location is: Ireland	Original files' average positivity is: 3.3932717	Calculated positivity is: 2.6874886
Location is: Egypt	Original files' average positivity is: 1.7520117	Calculated positivity is: 2.6849446
Location is: Germany	Original files' average positivity is: 1.8652757	Calculated positivity is: 2.6689937
Location is: Japan	Original files' average positivity is: 2.5855005	Calculated positivity is: 2.6478062
Location is: Italy	Original files' average positivity is: 0.9158584	Calculated positivity is: 2.667254
Location is: France	Original files' average positivity is: 1.8803257	Calculated positivity is: 2.6595244
Location is: Australia	Original files' average positivity is: 2.9113247	Calculated positivity is: 2.629487
Location is: New Zealand	Original files' average positivity is: 2.4460256	Calculated positivity is: 2.6356986
Location is: South Korea	Original files' average positivity is: 3.2266426	Calculated positivity is: 2.6282034
Location is: North Korea	Original files' average positivity is: 3.7197232	Calculated positivity is: 2.583798
Location is: Russia	Original files' average positivity is: 3.1334262	Calculated positivity is: 2.6399121
Location is: Taiwan	Original files' average positivity is: 1.2066365	Calculated positivity is: 2.707459
Location is: Switzerland	Original files' average positivity is: 1.2066365	Calculated positivity is: 2.707459
Location is: Belgium	Original files' average positivity is: 3.360215	Calculated positivity is: 2.6873214
Location is: Netherlands	Original files' average positivity is: 3.0924454	Calculated positivity is: 2.7209604
Location is: Croatia	Original files' average positivity is: 3.360215	Calculated positivity is: 2.6873214
Location is: Philippines	Original files' average positivity is: 3.46908	Calculated positivity is: 2.5501435
Location is: India	Original files' average positivity is: 2.30115	Calculated positivity is: 2.629596
Location is:	Original files' average positivity is:	Calculated positivity is: 2.6797976

Sweden	4.6737785	
Location is: Honduras	Original files' average positivity is: 2.6856241	Calculated positivity is: 2.6335206
Location is: Guatemala	Original files' average positivity is: 2.6856241	Calculated positivity is: 2.6335206
Location is: El Salvador	Original files' average positivity is: 2.6856241	Calculated positivity is: 2.6335206
Location is: Portugal	Original files' average positivity is: 5.3677936	Calculated positivity is: 2.6645372
Location is: Greenland	Original files' average positivity is: 2.5117738	Calculated positivity is: 2.5942554
Location is: Denmark	Original files' average positivity is: 2.5117738	Calculated positivity is: 2.5942554
Location is: Syria	Original files' average positivity is: 2.5117738	Calculated positivity is: 2.5942554
Location is: Mexico	Original files' average positivity is: 2.1972928	Calculated positivity is: 2.6477695
Location is: Madagascar	Original files' average positivity is: 2.9955947	Calculated positivity is: 2.7083538
Location is: Bolivia	Original files' average positivity is: 0.30911902	Calculated positivity is: 2.6392663
Location is: Paraguay	Original files' average positivity is: 0.30911902	Calculated positivity is: 2.6392663
Location is: Peru	Original files' average positivity is: 0.8538602	Calculated positivity is: 2.6445682
Location is: Norway	Original files' average positivity is: 1.1320754	Calculated positivity is: 2.8637402
Location is: Iran	Original files' average positivity is: 1.5649848	Calculated positivity is: 2.6696992
Location is: Trinidad And Tobago	Original files' average positivity is: 1.3586956	Calculated positivity is: 2.5887508
Location is: Malaysia	Original files' average positivity is: 1.3986014	Calculated positivity is: 2.6498702
Location is: Thailand	Original files' average positivity is: 2.733562	Calculated positivity is: 2.6726089
Location is: Singapore	Original files' average positivity is: 2.733562	Calculated positivity is: 2.6726089
Location is: Republic Of	Original files' average positivity is: 1.3986014	Calculated positivity is: 2.6498702
Location is: Brunei	Original files' average positivity is: 1.3986014	Calculated positivity is: 2.6498702
Location is: Chile	Original files' average positivity is: 1.2526913	Calculated positivity is: 2.6103451
Location is: Ukraine	Original files' average positivity is: 3.3066132	Calculated positivity is: 2.6732411
Location is: Kenya	Original files' average positivity is: 3.8539553	Calculated positivity is: 2.680356
Location is: Malawi	Original files' average positivity is: 3.8539553	Calculated positivity is: 2.680356
Location is: Ghana	Original files' average positivity is: 3.8539553	Calculated positivity is: 2.680356

Table A6.3: Calculated polarity values for locations using clusters vs. Original

GDELT average polarity values

Location	Original File Result (Polarization - Average)	Calculated Result (Polarization - different calculation than original)
Location is: China	Original files' average polarity is: 5.1291	Calculated polarity is: 0.0753389
Location is: Canada	Original files' average polarity is: 4.590341	Calculated polarity is: 0.072262146
Location is: United Kingdom	Original files' average polarity is: 6.609879	Calculated polarity is: 0.048169825
Location is: United States	Original files' average polarity is: 6.4636803	Calculated polarity is: 0.064564176
Location is: Brazil	Original files' average polarity is: 8.350612	Calculated polarity is: 0.0652855
Location is: Pakistan	Original files' average polarity is: 4.863372	Calculated polarity is: 0.023047324
Location is: Jersey	Original files' average polarity is: 5.8539395	Calculated polarity is: 0.057537626
Location is: Ireland	Original files' average polarity is: 6.640853	Calculated polarity is: 0.07093917
Location is: Egypt	Original files' average polarity is: 6.22517	Calculated polarity is: 0.09800049
Location is: Germany	Original files' average polarity is: 6.959271	Calculated polarity is: 0.061293483
Location is: Japan	Original files' average polarity is: 6.7989907	Calculated polarity is: 0.069202974
Location is: Italy	Original files' average polarity is: 6.0884275	Calculated polarity is: 0.062385738
Location is: France	Original files' average polarity is: 6.6120057	Calculated polarity is: 0.06559183
Location is: Australia	Original files' average polarity is: 5.3961234	Calculated polarity is: 0.07488532
Location is: New Zealand	Original files' average polarity is: 4.899431	Calculated polarity is: 0.11306525
Location is: South Korea	Original files' average polarity is: 6.9050937	Calculated polarity is: 0.07288149
Location is: North Korea	Original files' average polarity is: 9.083045	Calculated polarity is: 0.033023115
Location is: Russia	Original files' average polarity is: 6.544324	Calculated polarity is: 0.038717855
Location is: Taiwan	Original files' average polarity is: 5.1282053	Calculated polarity is: 0.061585836
Location is: Switzerland	Original files' average polarity is: 5.1282053	Calculated polarity is: 0.061585836
Location is: Belgium	Original files' average polarity is: 7.795699	Calculated polarity is: 0.06812339
Location is: Netherlands	Original files' average polarity is: 7.120333	Calculated polarity is: 0.07230573
Location is: Croatia	Original files' average polarity is: 7.795699	Calculated polarity is: 0.06812339
Location is: Philippines	Original files' average polarity is: 7.0889893	Calculated polarity is: 0.050678734
Location is: India	Original files' average polarity is: 5.2303023	Calculated polarity is: 0.06347145
Location is:	Original files' average polarity is:	Calculated polarity is: 0.093592286

Sweden	6.756486	
Location is: Honduras	Original files' average polarity is: 7.8988943	Calculated polarity is: 0.05931418
Location is: Guatemala	Original files' average polarity is: 7.8988943	Calculated polarity is: 0.05931418
Location is: El Salvador	Original files' average polarity is: 7.8988943	Calculated polarity is: 0.05931418
Location is: Portugal	Original files' average polarity is: 7.0576544	Calculated polarity is: 0.027004216
Location is: Greenland	Original files' average polarity is: 6.7503924	Calculated polarity is: 0.031504065
Location is: Denmark	Original files' average polarity is: 6.7503924	Calculated polarity is: 0.031504065
Location is: Syria	Original files' average polarity is: 6.7503924	Calculated polarity is: 0.031504065
Location is: Mexico	Original files' average polarity is: 4.526884	Calculated polarity is: 0.090055175
Location is: Madagascar	Original files' average polarity is: 6.035242	Calculated polarity is: 0.026876738
Location is: Bolivia	Original files' average polarity is: 9.27357	Calculated polarity is: 0.04559088
Location is: Paraguay	Original files' average polarity is: 9.27357	Calculated polarity is: 0.04559088
Location is: Peru	Original files' average polarity is: 6.1519365	Calculated polarity is: 0.080190584
Location is: Norway	Original files' average polarity is: 4.90566	Calculated polarity is: 0.10336239
Location is: Iran	Original files' average polarity is: 7.4730177	Calculated polarity is: 0.07356422
Location is: Trinidad And Tobago	Original files' average polarity is: 1.9021739	Calculated polarity is: 0.11247637
Location is: Malaysia	Original files' average polarity is: 3.030303	Calculated polarity is: 0.11479029
Location is: Thailand	Original files' average polarity is: 4.7271433	Calculated polarity is: 0.11273986
Location is: Singapore	Original files' average polarity is: 4.7271433	Calculated polarity is: 0.11273986
Location is: Republic Of	Original files' average polarity is: 3.030303	Calculated polarity is: 0.11479029
Location is: Brunei	Original files' average polarity is: 3.030303	Calculated polarity is: 0.11479029
Location is: Chile	Original files' average polarity is: 2.2592335	Calculated polarity is: 0.12500446
Location is: Ukraine	Original files' average polarity is: 4.308617	Calculated polarity is: 0.063467495
Location is: Kenya	Original files' average polarity is: 8.51927	Calculated polarity is: 0.04261954
Location is: Malawi	Original files' average polarity is: 8.51927	Calculated polarity is: 0.04261954
Location is: Ghana	Original files' average polarity is: 8.51927	Calculated polarity is: 0.04261954

g. Tables A7.0-A7.3:

Table A7.0: Calculated sentiment values using clusters on a larger dataset vs.

Original tone values

File Number	Original File Result	Calculated Result
0	Original file's tone is: 1.73410404624277	Calculated result is: -5.6077147
1	Original file's tone is: -1.94489465153971	Calculated result is: -5.857382
2	Original file's tone is: 3.97553516819572	Calculated result is: -5.9918346
3	Original file's tone is: -1.65413533834587	Calculated result is: -5.5697117
4	Original file's tone is: -2.0979020979021	Calculated result is: -5.548524
5	Original file's tone is: 3.72340425531915	Calculated result is: -5.8491974
6	Original file's tone is: -4.26666666666667	Calculated result is: -5.57008
7	Original file's tone is: -1.15183246073298	Calculated result is: -5.675329
8	Original file's tone is: 1.96078431372549	Calculated result is: -5.6592827
9	Original file's tone is: 1.63934426229508	Calculated result is: -5.9170246
10	Original file's tone is: -2.08604954367666	Calculated result is: -5.5715666
11	Original file's tone is: -4.46650124069479	Calculated result is: -5.687276
12	Original file's tone is: 8.02919708029197	Calculated result is: -5.7964544
13	Original file's tone is: -2.71493212669683	Calculated result is: -5.5504045
14	Original file's tone is: -1.0752688172043	Calculated result is: -5.90483
15	Original file's tone is: -0.150829562594269	Calculated result is: -5.1803894
16	Original file's tone is: 3.46907993966817	Calculated result is: -5.962897
17	Original file's tone is: -1.69971671388102	Calculated result is: -5.959142
18	Original file's tone is: -2.52764612954186	Calculated result is: -5.7159796
19	Original file's tone is: -0.749625187406297	Calculated result is: -5.693126
20	Original file's tone is: -3.27332242225859	Calculated result is: -5.6816325
21	Original file's tone is: 3.6779324055666	Calculated result is: -5.49611
22	Original file's tone is: -3.9426523297491	Calculated result is: -5.516957
23	Original file's tone is: -1.72684458398744	Calculated result is: -5.509794
24	Original file's tone is: -0.3003003003003	Calculated result is: -5.437614

25	Original file's tone is: -2.9126213592233	Calculated result is: -5.7428045
26	Original file's tone is: -3.0448717948718	Calculated result is: -5.6214705
27	Original file's tone is: -0.0440528634361237	Calculated result is: -5.7282515
28	Original file's tone is: -4.98177399756987	Calculated result is: -5.625373
29	Original file's tone is: -9.09090909090909	Calculated result is: -5.6085386
30	Original file's tone is: -2.64150943396226	Calculated result is: -5.7466307
31	Original file's tone is: -1.0752688172043	Calculated result is: -5.742298
32	Original file's tone is: -1.44628099173554	Calculated result is: -5.8306055
33	Original file's tone is: 0.641025641025641	Calculated result is: -5.853413
34	Original file's tone is: -5.2568697729988	Calculated result is: -5.705863
35	Original file's tone is: -4.04692082111437	Calculated result is: -5.585783
36	Original file's tone is: -1.985559566787	Calculated result is: -5.5430183
37	Original file's tone is: -2.18978102189781	Calculated result is: -5.8811417
38	Original file's tone is: 1.94174757281553	Calculated result is: -5.321988
39	Original file's tone is: 0.815217391304348	Calculated result is: -5.5361776
40	Original file's tone is: 1.88679245283019	Calculated result is: -5.7870827
41	Original file's tone is: -1.51717983043284	Calculated result is: -5.6894093
42	Original file's tone is: -0.233100233100233	Calculated result is: -5.4559317
43	Original file's tone is: 1.71306209850107	Calculated result is: -5.693721
44	Original file's tone is: -6.16197183098592	Calculated result is: -5.7728643
45	Original file's tone is: 0.196463654223968	Calculated result is: -5.7138414
46	Original file's tone is: 1.12044817927171	Calculated result is: -5.151408
47	Original file's tone is: 1.51187904967603	Calculated result is: -5.5458574
48	Original file's tone is: 2.30460921843687	Calculated result is: -5.880371
49	Original file's tone is: 0.485773768216516	Calculated result is: -5.50107
50	Original file's tone is: -5.12820512820513	Calculated result is: -5.7249727
51	Original file's tone is: 9.67741935483871	Calculated result is: -5.6681356
52	Original file's tone is: 2.6578073089701	Calculated result is: -5.7932887
53	Original file's tone is: 1.26984126984127	Calculated result is: -5.945775
54	Original file's tone is: -	Calculated result is: -5.7707033

	4.63917525773196	
55	Original file's tone is: 3.21257689678742	Calculated result is: -5.637211
56	Original file's tone is: 1.22850122850123	Calculated result is: -5.8563795
57	Original file's tone is: -2.47191011235955	Calculated result is: -5.6799536
58	Original file's tone is: -0.249221183800624	Calculated result is: -5.7447343
59	Original file's tone is: 2.27272727272727	Calculated result is: -5.9391956
60	Original file's tone is: 1.4895057549086	Calculated result is: -5.600623
61	Original file's tone is: -0.985221674876847	Calculated result is: -5.565797
62	Original file's tone is: -3.80577427821522	Calculated result is: -5.7701817
63	Original file's tone is: 3.79746835443038	Calculated result is: -5.875473
64	Original file's tone is: 3.6319612590799	Calculated result is: -5.4321194
65	Original file's tone is: -0.316455696202531	Calculated result is: -5.463617
66	Original file's tone is: 1.48957298907646	Calculated result is: -5.323694
67	Original file's tone is: -4.91803278688525	Calculated result is: -5.493974
68	Original file's tone is: -3.84615384615385	Calculated result is: -5.886318
69	Original file's tone is: 5.59006211180124	Calculated result is: -5.2074375
70	Original file's tone is: -5.57377049180328	Calculated result is: -5.9634905
71	Original file's tone is: -6.62650602409638	Calculated result is: -5.6463413
72	Original file's tone is: 0.606060606060606	Calculated result is: -5.9120255
73	Original file's tone is: 2.33100233100233	Calculated result is: -5.417015
74	Original file's tone is: -1.72413793103448	Calculated result is: -5.782058
75	Original file's tone is: 2.16216216216216	Calculated result is: -5.814427
76	Original file's tone is: -0.219780219780219	Calculated result is: -5.6097937
77	Original file's tone is: -6.34920634920635	Calculated result is: -5.71442
78	Original file's tone is: -2.96296296296296	Calculated result is: -5.4793415
79	Original file's tone is: -0.924214417744917	Calculated result is: -5.688584
80	Original file's tone is: 1.28331688055281	Calculated result is: -5.930422
81	Original file's tone is: -1.57384987893462	Calculated result is: -5.7565165
82	Original file's tone is: 1.88405797101449	Calculated result is: -5.686994
83	Original file's tone is: 3.20855614973262	Calculated result is: -5.7660947

84	Original file's tone is: 0.819672131147541	Calculated result is: -5.6556635
85	Original file's tone is: 6.97674418604651	Calculated result is: -5.4771123
86	Original file's tone is: 3.86996904024768	Calculated result is: -5.5375767
87	Original file's tone is: 2.06896551724138	Calculated result is: -5.7434273
88	Original file's tone is: -0.440528634361234	Calculated result is: -5.5678654
89	Original file's tone is: -2.01612903225806	Calculated result is: -5.7536674
90	Original file's tone is: -0.980392156862745	Calculated result is: -5.6646767
91	Original file's tone is: -0.276243093922652	Calculated result is: -5.4978466
92	Original file's tone is: -0.813008130081301	Calculated result is: -5.4303617
93	Original file's tone is: -0.465657741559953	Calculated result is: -5.449726
94	Original file's tone is: -1.42857142857143	Calculated result is: -5.403536
95	Original file's tone is: -1	Calculated result is: -5.9918365
96	Original file's tone is: 3.16622691292876	Calculated result is: -5.5304832
97	Original file's tone is: -1.55038759689922	Calculated result is: -5.7707987
98	Original file's tone is: -3.78548895899054	Calculated result is: -5.7939653
99	Original file's tone is: -2.85714285714286	Calculated result is: -5.858906
100	Original file's tone is: 0.213675213675214	Calculated result is: -5.3971357
101	Original file's tone is: -4.53752181500873	Calculated result is: -5.4305143
102	Original file's tone is: 1.18918918918919	Calculated result is: -5.6766806
103	Original file's tone is: -11.9565217391304	Calculated result is: -5.6516266
104	Original file's tone is: -2.11946050096339	Calculated result is: -5.799161
105	Original file's tone is: 0.575539568345324	Calculated result is: -5.1941233
106	Original file's tone is: -2.77544154751892	Calculated result is: -5.8024673
107	Original file's tone is: -3.43137254901961	Calculated result is: -5.6904106
108	Original file's tone is: 1.81488203266788	Calculated result is: -5.5927377
109	Original file's tone is: -1.55945419103314	Calculated result is: -5.6441555
110	Original file's tone is: 0.34965034965035	Calculated result is: -5.682334
111	Original file's tone is: 2.46305418719212	Calculated result is: -5.991834
112	Original file's tone is: -4.84261501210654	Calculated result is: -5.991836
113	Original file's tone is: -2.02020202020202	Calculated result is: -5.5341215

114	Original file's tone is: -0.268096514745308	Calculated result is: -5.564045
115	Original file's tone is: -4.52107279693486	Calculated result is: -6.053048
116	Original file's tone is: 1.68269230769231	Calculated result is: -5.222988
117	Original file's tone is: -2.82413350449294	Calculated result is: -5.8719244
118	Original file's tone is: 0.816326530612245	Calculated result is: -5.4965124
119	Original file's tone is: -6.06060606060606	Calculated result is: -5.6328025
120	Original file's tone is: 0.294334069168506	Calculated result is: -5.3991737
121	Original file's tone is: -2.03045685279188	Calculated result is: -5.597522
122	Original file's tone is: -4.5045045045045	Calculated result is: -5.6436462
123	Original file's tone is: -2.12201591511936	Calculated result is: -5.4101973
124	Original file's tone is: 0.833333333333333	Calculated result is: -5.832687
125	Original file's tone is: -2.57731958762886	Calculated result is: -5.991834
126	Original file's tone is: -0.614754098360656	Calculated result is: -5.7351832
127	Original file's tone is: -2.68317853457172	Calculated result is: -5.6298056
128	Original file's tone is: -4.08163265306122	Calculated result is: -5.9918346
129	Original file's tone is: 2.53807106598985	Calculated result is: -5.7260766
130	Original file's tone is: -2.10909090909091	Calculated result is: -5.710155
131	Original file's tone is: -4.34782608695652	Calculated result is: -5.84026
132	Original file's tone is: -1.79392824287029	Calculated result is: -5.3402042
133	Original file's tone is: 3.35413416536661	Calculated result is: -5.827594
134	Original file's tone is: 5.61224489795918	Calculated result is: -5.599635
135	Original file's tone is: -1.12596762843068	Calculated result is: -5.564517
136	Original file's tone is: -1.87207488299532	Calculated result is: -5.4371223
137	Original file's tone is: 0.587084148727984	Calculated result is: -5.6860094
138	Original file's tone is: 2.24525043177893	Calculated result is: -5.436451
139	Original file's tone is: -5.25164113785558	Calculated result is: -5.698877
140	Original file's tone is: -1.40449438202247	Calculated result is: -5.6697416
141	Original file's tone is: -6.56620021528525	Calculated result is: -5.619113
142	Original file's tone is: -2.76338514680484	Calculated result is: -5.6617813
143	Original file's tone is: -	Calculated result is: -5.710951

	3.63321799307958	
144	Original file's tone is: 1.17647058823529	Calculated result is: -5.512495
145	Original file's tone is: -3.45911949685535	Calculated result is: -5.670857
146	Original file's tone is: -0.462962962962963	Calculated result is: -5.709406
147	Original file's tone is: -7.62331838565022	Calculated result is: -5.738144
148	Original file's tone is: 0.904977375565611	Calculated result is: -5.5216427
149	Original file's tone is: 0.353606789250354	Calculated result is: -5.7906837
150	Original file's tone is: -4.24242424242424	Calculated result is: -5.9918356
151	Original file's tone is: -0.887311446317657	Calculated result is: -5.644725
152	Original file's tone is: -1.42700329308452	Calculated result is: -5.742593
153	Original file's tone is: -3.69068541300527	Calculated result is: -5.7085876
154	Original file's tone is: -4.85933503836317	Calculated result is: -5.682018
155	Original file's tone is: -2.92887029288703	Calculated result is: -5.9918356
156	Original file's tone is: -3.95408163265306	Calculated result is: -5.9398794
157	Original file's tone is: -0.0779423226812159	Calculated result is: -5.9292374
158	Original file's tone is: 4.36681222707424	Calculated result is: -5.781569
159	Original file's tone is: -0.322061191626409	Calculated result is: -5.6981134
160	Original file's tone is: 0.282485875706215	Calculated result is: -5.83304
161	Original file's tone is: 0.927643784786642	Calculated result is: -5.695365
162	Original file's tone is: -2.33545647558386	Calculated result is: -5.9918365
163	Original file's tone is: -1.63934426229508	Calculated result is: -5.4035363
164	Original file's tone is: 3.40030911901082	Calculated result is: -5.6279306
165	Original file's tone is: -1.94174757281553	Calculated result is: -5.8537226
166	Original file's tone is: -0.71563088512241	Calculated result is: -5.5046997
167	Original file's tone is: -0.2909796314258	Calculated result is: -5.474162
168	Original file's tone is: 2.3696682464455	Calculated result is: -5.6336985
169	Original file's tone is: -1.21951219512195	Calculated result is: -5.5369987
170	Original file's tone is: -1.16959064327485	Calculated result is: -5.7007074
171	Original file's tone is: -3.25443786982249	Calculated result is: -5.9324474
172	Original file's tone is: -1.63265306122449	Calculated result is: -5.5398026

173	Original file's tone is: 1.28676470588235	Calculated result is: -5.336373
174	Original file's tone is: -3.72960372960373	Calculated result is: -5.7085876
175	Original file's tone is: 2.80898876404494	Calculated result is: -5.677414
176	Original file's tone is: -2.1978021978022	Calculated result is: -5.367125
177	Original file's tone is: 1.87617260787992	Calculated result is: -5.427455
178	Original file's tone is: -0.189933523266856	Calculated result is: -5.815537
179	Original file's tone is: -3.63864491844417	Calculated result is: -5.6194835

Table A7.1: Calculated negative sentiment values using clusters on a larger dataset

vs. Original negative tone values

File Number	Original File Result (Negativity)	Calculated Result (Negativity)
0	Original file's negativity is: 1.15606936416185	Calculated negativity is: -7.9281516
1	Original file's negativity is: 5.02431118314425	Calculated negativity is: -8.281131
2	Original file's negativity is: 0	Calculated negativity is: -8.471218
3	Original file's negativity is: 4.66165413533835	Calculated negativity is: -7.874423
4	Original file's negativity is: 2.0979020979021	Calculated negativity is: -7.8444686
5	Original file's negativity is: 2.12765957446809	Calculated negativity is: -8.269558
6	Original file's negativity is: 6.66666666666667	Calculated negativity is: -7.8749437
7	Original file's negativity is: 2.82722513089005	Calculated negativity is: -8.0237465
8	Original file's negativity is: 1.44478844169247	Calculated negativity is: -8.001059
9	Original file's negativity is: 1.78837555886736	Calculated negativity is: -8.365452
10	Original file's negativity is: 4.30247718383312	Calculated negativity is: -7.8770456
11	Original file's negativity is: 4.71464019851117	Calculated negativity is: -8.040634
12	Original file's negativity is: 0.291970802919708	Calculated negativity is: -8.194991
13	Original file's negativity is: 3.92156862745098	Calculated negativity is: -7.847127
14	Original file's negativity is: 4.43548387096774	Calculated negativity is: -8.348211
15	Original file's negativity is: 3.61990950226244	Calculated negativity is: -7.324002
16	Original file's negativity is: 1.80995475113122	Calculated negativity is: -8.430306
17	Original file's negativity is: 3.82436260623229	Calculated negativity is: -8.424997
18	Original file's negativity is: 5.21327014218009	Calculated negativity is: -8.081216
19	Original file's negativity is: 3.59820089955022	Calculated negativity is: -8.048905
20	Original file's negativity is: 4.74631751227496	Calculated negativity is: -8.032656
21	Original file's negativity is: 1.68986083499006	Calculated negativity is: -7.7703657
22	Original file's negativity is: 4.83870967741935	Calculated negativity is: -7.799838
23	Original file's negativity is: 4.23861852433281	Calculated negativity is: -7.7897124
24	Original file's negativity is: 2.4024024024024	Calculated negativity is: -7.6876645
25	Original file's negativity is: 4.07766990291262	Calculated negativity is: -8.119142
26	Original file's negativity is: 4.6474358974359	Calculated negativity is: -7.947599

27	Original file's negativity is: 3.03964757709251	Calculated negativity is: -8.098566
28	Original file's negativity is: 7.53341433778858	Calculated negativity is: -7.953117
29	Original file's negativity is: 9.09090909090909	Calculated negativity is: -7.929316
30	Original file's negativity is: 3.77358490566038	Calculated negativity is: -8.124551
31	Original file's negativity is: 4.3010752688172	Calculated negativity is: -8.118425
32	Original file's negativity is: 5.37190082644628	Calculated negativity is: -8.243274
33	Original file's negativity is: 2.24358974358974	Calculated negativity is: -8.275518
34	Original file's negativity is: 7.16845878136201	Calculated negativity is: -8.066917
35	Original file's negativity is: 5.63049853372434	Calculated negativity is: -7.8971477
36	Original file's negativity is: 3.79061371841155	Calculated negativity is: -7.836684
37	Original file's negativity is: 4.66030320044919	Calculated negativity is: -8.314721
38	Original file's negativity is: 0.485436893203883	Calculated negativity is: -7.524193
39	Original file's negativity is: 0.543478260869565	Calculated negativity is: -7.8270125
40	Original file's negativity is: 1.71526586620926	Calculated negativity is: -8.181741
41	Original file's negativity is: 3.6590807675145	Calculated negativity is: -8.0436535
42	Original file's negativity is: 1.63170163170163	Calculated negativity is: -7.713563
43	Original file's negativity is: 2.35546038543897	Calculated negativity is: -8.04975
44	Original file's negativity is: 7.3943661971831	Calculated negativity is: -8.161639
45	Original file's negativity is: 1.96463654223969	Calculated negativity is: -8.078193
46	Original file's negativity is: 0	Calculated negativity is: -7.2830286
47	Original file's negativity is: 0.647948164146868	Calculated negativity is: -7.8406987
48	Original file's negativity is: 1.00200400801603	Calculated negativity is: -8.313632
49	Original file's negativity is: 0.693962526023595	Calculated negativity is: -7.77738
50	Original file's negativity is: 5.41310541310541	Calculated negativity is: -8.093931
51	Original file's negativity is: 0.379506641366224	Calculated negativity is: -8.013575
52	Original file's negativity is: 1.32890365448505	Calculated negativity is: -8.1905155
53	Original file's negativity is: 5.3968253968254	Calculated negativity is: -8.406099
54	Original file's negativity is: 6.18556701030928	Calculated negativity is: -8.158585
55	Original file's negativity is: 0.75187969924812	Calculated negativity is: -7.969853
56	Original file's negativity is: 0.245700245700246	Calculated negativity is: -8.279713

57	Original file's negativity is: 4.26966292134831	Calculated negativity is: -8.030284
58	Original file's negativity is: 2.36760124610592	Calculated negativity is: -8.121873
59	Original file's negativity is: 0	Calculated negativity is: -8.396797
60	Original file's negativity is: 2.2342586323629	Calculated negativity is: -7.9181275
61	Original file's negativity is: 7.14285714285714	Calculated negativity is: -7.868889
62	Original file's negativity is: 6.2992125984252	Calculated negativity is: -8.157846
63	Original file's negativity is: 0.281293952180028	Calculated negativity is: -8.306706
64	Original file's negativity is: 2.17917675544794	Calculated negativity is: -7.679896
65	Original file's negativity is: 2.21518987341772	Calculated negativity is: -7.7244263
66	Original file's negativity is: 0.893743793445879	Calculated negativity is: -7.5266066
67	Original file's negativity is: 5.46448087431694	Calculated negativity is: -7.767346
68	Original file's negativity is: 7.17948717948718	Calculated negativity is: -8.322039
69	Original file's negativity is: 0	Calculated negativity is: -7.3622417
70	Original file's negativity is: 6.22950819672131	Calculated negativity is: -8.431146
71	Original file's negativity is: 7.2289156626506	Calculated negativity is: -7.982762
72	Original file's negativity is: 2.28163992869875	Calculated negativity is: -8.358385
73	Original file's negativity is: 1.51515151515152	Calculated negativity is: -7.658542
74	Original file's negativity is: 1.72413793103448	Calculated negativity is: -8.174637
75	Original file's negativity is: 3.10810810810811	Calculated negativity is: -8.2204
76	Original file's negativity is: 3.95604395604396	Calculated negativity is: -7.9310913
77	Original file's negativity is: 7.67195767195767	Calculated negativity is: -8.079011
78	Original file's negativity is: 4.44444444444444	Calculated negativity is: -7.7466583
79	Original file's negativity is: 4.4362292051756	Calculated negativity is: -8.042484
80	Original file's negativity is: 2.56663376110563	Calculated negativity is: -8.384393
81	Original file's negativity is: 3.38983050847458	Calculated negativity is: -8.138527
82	Original file's negativity is: 2.39130434782609	Calculated negativity is: -8.040236
83	Original file's negativity is: 2.31729055258467	Calculated negativity is: -8.152069
84	Original file's negativity is: 2.45901639344262	Calculated negativity is: -7.995942
85	Original file's negativity is: 0.290697674418605	Calculated negativity is: -7.743507
86	Original file's negativity is: 0.154798761609907	Calculated negativity is: -7.8289914

87	Original file's negativity is: 0.689655172413793	Calculated negativity is: -8.120022
88	Original file's negativity is: 1.76211453744493	Calculated negativity is: -7.871816
89	Original file's negativity is: 3.89784946236559	Calculated negativity is: -8.134499
90	Original file's negativity is: 3.18627450980392	Calculated negativity is: -8.008684
91	Original file's negativity is: 3.10773480662983	Calculated negativity is: -7.7728214
92	Original file's negativity is: 0.813008130081301	Calculated negativity is: -7.677412
93	Original file's negativity is: 1.9790454016298	Calculated negativity is: -7.7047877
94	Original file's negativity is: 4.76190476190476	Calculated negativity is: -7.6394854
95	Original file's negativity is: 3.85714285714286	Calculated negativity is: -8.47122
96	Original file's negativity is: 0.527704485488127	Calculated negativity is: -7.8189626
97	Original file's negativity is: 3.10077519379845	Calculated negativity is: -8.158719
98	Original file's negativity is: 5.20504731861199	Calculated negativity is: -8.191472
99	Original file's negativity is: 4.41558441558442	Calculated negativity is: -8.283284
100	Original file's negativity is: 4.05982905982906	Calculated negativity is: -7.630436
101	Original file's negativity is: 6.98080279232112	Calculated negativity is: -7.677627
102	Original file's negativity is: 0.216216216216216	Calculated negativity is: -8.025656
103	Original file's negativity is: 11.9565217391304	Calculated negativity is: -7.9902334
104	Original file's negativity is: 4.23892100192678	Calculated negativity is: -8.198817
105	Original file's negativity is: 1.58273381294964	Calculated negativity is: -7.343419
106	Original file's negativity is: 4.54163162321278	Calculated negativity is: -8.203492
107	Original file's negativity is: 5.49019607843137	Calculated negativity is: -8.045067
108	Original file's negativity is: 1.81488203266788	Calculated negativity is: -7.9069777
109	Original file's negativity is: 3.31384015594542	Calculated negativity is: -7.979676
110	Original file's negativity is: 0.699300699300699	Calculated negativity is: -8.033648
111	Original file's negativity is: 0	Calculated negativity is: -8.471218
112	Original file's negativity is: 7.26392251815981	Calculated negativity is: -8.47122
113	Original file's negativity is: 3.7037037037037	Calculated negativity is: -7.824105
114	Original file's negativity is: 2.41286863270777	Calculated negativity is: -7.866412
115	Original file's negativity is: 6.97318007662835	Calculated negativity is: -8.557761
116	Original file's negativity is: 0	Calculated negativity is: -7.3842273

117	Original file's negativity is: 4.36456996148909	Calculated negativity is: -8.301694
118	Original file's negativity is: 0	Calculated negativity is: -7.770935
119	Original file's negativity is: 7.57575757575758	Calculated negativity is: -7.9636207
120	Original file's negativity is: 1.61883738042678	Calculated negativity is: -7.633322
121	Original file's negativity is: 4.06091370558376	Calculated negativity is: -7.913741
122	Original file's negativity is: 5.85585585585586	Calculated negativity is: -7.9789524
123	Original file's negativity is: 3.71352785145889	Calculated negativity is: -7.648902
124	Original file's negativity is: 3.75	Calculated negativity is: -8.246216
125	Original file's negativity is: 7.21649484536082	Calculated negativity is: -8.471218
126	Original file's negativity is: 1.84426229508197	Calculated negativity is: -8.108366
127	Original file's negativity is: 4.3343653250774	Calculated negativity is: -7.9593835
128	Original file's negativity is: 4.48979591836735	Calculated negativity is: -8.471218
129	Original file's negativity is: 0.50761421319797	Calculated negativity is: -8.095491
130	Original file's negativity is: 3.85454545454545	Calculated negativity is: -8.072982
131	Original file's negativity is: 6.64961636828645	Calculated negativity is: -8.256923
132	Original file's negativity is: 4.46182152713891	Calculated negativity is: -7.5499473
133	Original file's negativity is: 1.71606864274571	Calculated negativity is: -8.239017
134	Original file's negativity is: 1.27551020408163	Calculated negativity is: -7.9167295
135	Original file's negativity is: 5.70021111893033	Calculated negativity is: -7.8670783
136	Original file's negativity is: 2.65210608424337	Calculated negativity is: -7.6869683
137	Original file's negativity is: 2.73972602739726	Calculated negativity is: -8.038844
138	Original file's negativity is: 1.89982728842832	Calculated negativity is: -7.6860204
139	Original file's negativity is: 6.34573304157549	Calculated negativity is: -8.057036
140	Original file's negativity is: 4.11985018726592	Calculated negativity is: -8.015848
141	Original file's negativity is: 7.10441334768568	Calculated negativity is: -7.944267
142	Original file's negativity is: 5.35405872193437	Calculated negativity is: -8.004591
143	Original file's negativity is: 4.84429065743945	Calculated negativity is: -8.074106
144	Original file's negativity is: 0.588235294117647	Calculated negativity is: -7.7935305
145	Original file's negativity is: 5.13626834381551	Calculated negativity is: -8.017426
146	Original file's negativity is: 3.24074074074074	Calculated negativity is: -8.071922

147	Original file's negativity is: 9.41704035874439	Calculated negativity is: -8.112551
148	Original file's negativity is: 3.16742081447964	Calculated negativity is: -7.806463
149	Original file's negativity is: 2.19236209335219	Calculated negativity is: -8.186832
150	Original file's negativity is: 4.84848484848485	Calculated negativity is: -8.471219
151	Original file's negativity is: 5.9449866903283	Calculated negativity is: -7.9804764
152	Original file's negativity is: 3.07354555433589	Calculated negativity is: -8.118846
153	Original file's negativity is: 5.53602811950791	Calculated negativity is: -8.070768
154	Original file's negativity is: 5.62659846547315	Calculated negativity is: -8.033201
155	Original file's negativity is: 5.02092050209205	Calculated negativity is: -8.471219
156	Original file's negativity is: 5.42091836734694	Calculated negativity is: -8.397763
157	Original file's negativity is: 1.09119251753702	Calculated negativity is: -8.382718
158	Original file's negativity is: 0.436681222707424	Calculated negativity is: -8.173945
159	Original file's negativity is: 2.57648953301127	Calculated negativity is: -8.055958
160	Original file's negativity is: 1.41242937853107	Calculated negativity is: -8.246716
161	Original file's negativity is: 3.8961038961039	Calculated negativity is: -8.0520735
162	Original file's negativity is: 3.39702760084926	Calculated negativity is: -8.471221
163	Original file's negativity is: 2.45901639344262	Calculated negativity is: -7.639486
164	Original file's negativity is: 0.154559505409583	Calculated negativity is: -7.956733
165	Original file's negativity is: 2.9126213592233	Calculated negativity is: -8.275956
166	Original file's negativity is: 2.4105461393597	Calculated negativity is: -7.7825127
167	Original file's negativity is: 1.09925638538636	Calculated negativity is: -7.73934
168	Original file's negativity is: 1.18483412322275	Calculated negativity is: -7.964887
169	Original file's negativity is: 1.21951219512195	Calculated negativity is: -7.828173
170	Original file's negativity is: 3.50877192982456	Calculated negativity is: -8.059624
171	Original file's negativity is: 4.43786982248521	Calculated negativity is: -8.387257
172	Original file's negativity is: 2.44897959183673	Calculated negativity is: -7.832137
173	Original file's negativity is: 1.37867647058824	Calculated negativity is: -7.5445304
174	Original file's negativity is: 4.89510489510489	Calculated negativity is: -8.070768
175	Original file's negativity is: 0	Calculated negativity is: -8.026691
176	Original file's negativity is: 3.84615384615385	Calculated negativity is: -7.588007

177	Original file's negativity is: 1.31332082551595	Calculated negativity is: -7.673301
178	Original file's negativity is: 2.46913580246914	Calculated negativity is: -8.22197
179	Original file's negativity is: 5.14429109159347	Calculated negativity is: -7.944791

Table A7.2: Calculated positive sentiment values using clusters on a larger dataset

vs. Original positive tone values

File Number	Original File Result (Positivity)	Calculated Result (Positivity)
0	Original file's positivity is: 2.89017341040462	Calculated positivity is: 2.3204367
1	Original file's positivity is: 3.07941653160454	Calculated positivity is: 2.4237487
2	Original file's positivity is: 3.97553516819572	Calculated positivity is: 2.4793837
3	Original file's positivity is: 3.00751879699248	Calculated positivity is: 2.3047116
4	Original file's positivity is: 0	Calculated positivity is: 2.2959447
5	Original file's positivity is: 5.85106382978723	Calculated positivity is: 2.4203606
6	Original file's positivity is: 2.4	Calculated positivity is: 2.304864
7	Original file's positivity is: 1.67539267015707	Calculated positivity is: 2.3484175
8	Original file's positivity is: 3.40557275541796	Calculated positivity is: 2.3417761
9	Original file's positivity is: 3.42771982116244	Calculated positivity is: 2.448427
10	Original file's positivity is: 2.21642764015645	Calculated positivity is: 2.305479
11	Original file's positivity is: 0.248138957816377	Calculated positivity is: 2.3533585
12	Original file's positivity is: 8.32116788321168	Calculated positivity is: 2.3985364
13	Original file's positivity is: 1.20663650075415	Calculated positivity is: 2.2967222
14	Original file's positivity is: 3.36021505376344	Calculated positivity is: 2.4433813
15	Original file's positivity is: 3.46907993966818	Calculated positivity is: 2.1436124
16	Original file's positivity is: 5.2790346907994	Calculated positivity is: 2.4674096
17	Original file's positivity is: 2.12464589235127	Calculated positivity is: 2.4658551
18	Original file's positivity is: 2.68562401263823	Calculated positivity is: 2.3652365
19	Original file's positivity is: 2.84857571214393	Calculated positivity is: 2.3557792
20	Original file's positivity is: 1.47299509001637	Calculated positivity is: 2.3510232
21	Original file's positivity is: 5.36779324055666	Calculated positivity is: 2.274256
22	Original file's positivity is: 0.896057347670251	Calculated positivity is: 2.2828813
23	Original file's positivity is: 2.51177394034537	Calculated positivity is: 2.2799184
24	Original file's positivity is: 2.1021021021021	Calculated positivity is: 2.2500505
25	Original file's positivity is: 1.16504854368932	Calculated positivity is: 2.3763368
26	Original file's positivity is: 1.6025641025641	Calculated positivity is: 2.3261287

27	Original file's positivity is: 2.99559471365639	Calculated positivity is: 2.3703146
28	Original file's positivity is: 2.55164034021871	Calculated positivity is: 2.3277438
29	Original file's positivity is: 0	Calculated positivity is: 2.3207772
30	Original file's positivity is: 1.13207547169811	Calculated positivity is: 2.3779202
31	Original file's positivity is: 3.2258064516129	Calculated positivity is: 2.376127
32	Original file's positivity is: 3.92561983471074	Calculated positivity is: 2.412668
33	Original file's positivity is: 2.88461538461538	Calculated positivity is: 2.4221053
34	Original file's positivity is: 1.9115890083632	Calculated positivity is: 2.3610544
35	Original file's positivity is: 1.58357771260997	Calculated positivity is: 2.3113647
36	Original file's positivity is: 1.80505415162455	Calculated positivity is: 2.293666
37	Original file's positivity is: 2.47052217855138	Calculated positivity is: 2.4335797
38	Original file's positivity is: 2.42718446601942	Calculated positivity is: 2.2022047
39	Original file's positivity is: 1.35869565217391	Calculated positivity is: 2.2908351
40	Original file's positivity is: 3.60205831903945	Calculated positivity is: 2.394658
41	Original file's positivity is: 2.14190093708166	Calculated positivity is: 2.354244
42	Original file's positivity is: 1.3986013986014	Calculated positivity is: 2.2576313
43	Original file's positivity is: 4.06852248394004	Calculated positivity is: 2.3560295
44	Original file's positivity is: 1.23239436619718	Calculated positivity is: 2.3887749
45	Original file's positivity is: 2.16110019646365	Calculated positivity is: 2.3643515
46	Original file's positivity is: 1.12044817927171	Calculated positivity is: 2.1316202
47	Original file's positivity is: 2.15982721382289	Calculated positivity is: 2.2948413
48	Original file's positivity is: 3.30661322645291	Calculated positivity is: 2.433261
49	Original file's positivity is: 1.17973629424011	Calculated positivity is: 2.2763102
50	Original file's positivity is: 0.284900284900285	Calculated positivity is: 2.3689582
51	Original file's positivity is: 10.0569259962049	Calculated positivity is: 2.345439
52	Original file's positivity is: 3.98671096345515	Calculated positivity is: 2.3972268
53	Original file's positivity is: 6.66666666666667	Calculated positivity is: 2.460324
54	Original file's positivity is: 1.54639175257732	Calculated positivity is: 2.3878813
55	Original file's positivity is: 3.96445659603554	Calculated positivity is: 2.3326423
56	Original file's positivity is: 1.47420147420147	Calculated positivity is: 2.423333

57	Original file's positivity is: 1.79775280898876	Calculated positivity is: 2.35033
58	Original file's positivity is: 2.1183800623053	Calculated positivity is: 2.3771386
59	Original file's positivity is: 2.27272727272727	Calculated positivity is: 2.4576015
60	Original file's positivity is: 3.7237643872715	Calculated positivity is: 2.3175046
61	Original file's positivity is: 6.1576354679803	Calculated positivity is: 2.3030922
62	Original file's positivity is: 2.49343832020997	Calculated positivity is: 2.3876648
63	Original file's positivity is: 4.07876230661041	Calculated positivity is: 2.4312336
64	Original file's positivity is: 5.81113801452785	Calculated positivity is: 2.2477767
65	Original file's positivity is: 1.89873417721519	Calculated positivity is: 2.2608094
66	Original file's positivity is: 2.38331678252234	Calculated positivity is: 2.2029123
67	Original file's positivity is: 0.546448087431694	Calculated positivity is: 2.2733717
68	Original file's positivity is: 3.33333333333333	Calculated positivity is: 2.4357207
69	Original file's positivity is: 5.59006211180124	Calculated positivity is: 2.1548045
70	Original file's positivity is: 0.655737704918033	Calculated positivity is: 2.467655
71	Original file's positivity is: 0.602409638554217	Calculated positivity is: 2.3364203
72	Original file's positivity is: 2.88770053475936	Calculated positivity is: 2.4463596
73	Original file's positivity is: 3.84615384615385	Calculated positivity is: 2.2415268
74	Original file's positivity is: 0	Calculated positivity is: 2.3925793
75	Original file's positivity is: 5.27027027027027	Calculated positivity is: 2.405973
76	Original file's positivity is: 3.73626373626374	Calculated positivity is: 2.3212976
77	Original file's positivity is: 1.32275132275132	Calculated positivity is: 2.3645911
78	Original file's positivity is: 1.48148148148148	Calculated positivity is: 2.2673168
79	Original file's positivity is: 3.51201478743068	Calculated positivity is: 2.3539004
80	Original file's positivity is: 3.84995064165844	Calculated positivity is: 2.453971
81	Original file's positivity is: 1.81598062953995	Calculated positivity is: 2.3820105
82	Original file's positivity is: 4.27536231884058	Calculated positivity is: 2.3532424
83	Original file's positivity is: 5.52584670231729	Calculated positivity is: 2.3859744
84	Original file's positivity is: 3.27868852459016	Calculated positivity is: 2.3402784
85	Original file's positivity is: 7.26744186046512	Calculated positivity is: 2.2663946
86	Original file's positivity is: 4.02476780185759	Calculated positivity is: 2.2914147

87	Original file's positivity is: 2.75862068965517	Calculated positivity is: 2.3765945
88	Original file's positivity is: 1.3215859030837	Calculated positivity is: 2.3039508
89	Original file's positivity is: 1.88172043010753	Calculated positivity is: 2.3808312
90	Original file's positivity is: 2.20588235294118	Calculated positivity is: 2.3440075
91	Original file's positivity is: 2.83149171270718	Calculated positivity is: 2.2749746
92	Original file's positivity is: 0	Calculated positivity is: 2.24705
93	Original file's positivity is: 1.51338766006985	Calculated positivity is: 2.2550619
94	Original file's positivity is: 3.33333333333333	Calculated positivity is: 2.2359493
95	Original file's positivity is: 2.85714285714286	Calculated positivity is: 2.4793837
96	Original file's positivity is: 3.69393139841689	Calculated positivity is: 2.288479
97	Original file's positivity is: 1.55038759689922	Calculated positivity is: 2.3879201
98	Original file's positivity is: 1.41955835962145	Calculated positivity is: 2.397507
99	Original file's positivity is: 1.55844155844156	Calculated positivity is: 2.4243782
100	Original file's positivity is: 4.27350427350427	Calculated positivity is: 2.2333002
101	Original file's positivity is: 2.44328097731239	Calculated positivity is: 2.2471128
102	Original file's positivity is: 1.40540540540541	Calculated positivity is: 2.348975
103	Original file's positivity is: 0	Calculated positivity is: 2.338607
104	Original file's positivity is: 2.11946050096339	Calculated positivity is: 2.3996565
105	Original file's positivity is: 2.15827338129496	Calculated positivity is: 2.1492956
106	Original file's positivity is: 1.76619007569386	Calculated positivity is: 2.4010246
107	Original file's positivity is: 2.05882352941176	Calculated positivity is: 2.3546562
108	Original file's positivity is: 3.62976406533575	Calculated positivity is: 2.31424
109	Original file's positivity is: 1.75438596491228	Calculated positivity is: 2.3355203
110	Original file's positivity is: 1.04895104895105	Calculated positivity is: 2.3513138
111	Original file's positivity is: 2.46305418719212	Calculated positivity is: 2.479384
112	Original file's positivity is: 2.42130750605327	Calculated positivity is: 2.479384
113	Original file's positivity is: 1.68350168350168	Calculated positivity is: 2.2899835
114	Original file's positivity is: 2.14477211796247	Calculated positivity is: 2.3023672
115	Original file's positivity is: 2.45210727969349	Calculated positivity is: 2.5047133
116	Original file's positivity is: 1.68269230769231	Calculated positivity is: 2.1612394

117	Original file's positivity is: 1.54043645699615	Calculated positivity is: 2.4297693
118	Original file's positivity is: 0.816326530612245	Calculated positivity is: 2.2744224
119	Original file's positivity is: 1.51515151515152	Calculated positivity is: 2.3308182
120	Original file's positivity is: 1.91317144959529	Calculated positivity is: 2.2341478
121	Original file's positivity is: 2.03045685279188	Calculated positivity is: 2.3162193
122	Original file's positivity is: 1.35135135135135	Calculated positivity is: 2.335306
123	Original file's positivity is: 1.59151193633952	Calculated positivity is: 2.238705
124	Original file's positivity is: 4.58333333333333	Calculated positivity is: 2.413529
125	Original file's positivity is: 4.63917525773196	Calculated positivity is: 2.479384
126	Original file's positivity is: 1.22950819672131	Calculated positivity is: 2.3731828
127	Original file's positivity is: 1.65118679050568	Calculated positivity is: 2.329578
128	Original file's positivity is: 0.408163265306122	Calculated positivity is: 2.4793837
129	Original file's positivity is: 3.04568527918782	Calculated positivity is: 2.3694148
130	Original file's positivity is: 1.74545454545455	Calculated positivity is: 2.3628268
131	Original file's positivity is: 2.30179028132992	Calculated positivity is: 2.4166627
132	Original file's positivity is: 2.66789328426863	Calculated positivity is: 2.209743
133	Original file's positivity is: 5.07020280811232	Calculated positivity is: 2.4114225
134	Original file's positivity is: 6.88775510204082	Calculated positivity is: 2.317094
135	Original file's positivity is: 4.57424349049965	Calculated positivity is: 2.3025615
136	Original file's positivity is: 0.78003120124805	Calculated positivity is: 2.2498462
137	Original file's positivity is: 3.32681017612524	Calculated positivity is: 2.3528345
138	Original file's positivity is: 4.14507772020725	Calculated positivity is: 2.2495692
139	Original file's positivity is: 1.09409190371991	Calculated positivity is: 2.3581595
140	Original file's positivity is: 2.71535580524345	Calculated positivity is: 2.3461065
141	Original file's positivity is: 0.538213132400431	Calculated positivity is: 2.3251536
142	Original file's positivity is: 2.59067357512953	Calculated positivity is: 2.3428097
143	Original file's positivity is: 1.21107266435986	Calculated positivity is: 2.3631554
144	Original file's positivity is: 1.76470588235294	Calculated positivity is: 2.2810354
145	Original file's positivity is: 1.67714884696017	Calculated positivity is: 2.3465686
146	Original file's positivity is:	Calculated positivity is: 2.3625162

	2.77777777777778	
147	Original file's positivity is: 1.79372197309417	Calculated positivity is: 2.374407
148	Original file's positivity is: 4.07239819004525	Calculated positivity is: 2.2848198
149	Original file's positivity is: 2.54596888260255	Calculated positivity is: 2.3961484
150	Original file's positivity is: 0.606060606060606	Calculated positivity is: 2.4793837
151	Original file's positivity is: 5.05767524401065	Calculated positivity is: 2.3357518
152	Original file's positivity is: 1.64654226125137	Calculated positivity is: 2.3762531
153	Original file's positivity is: 1.84534270650264	Calculated positivity is: 2.3621805
154	Original file's positivity is: 0.767263427109974	Calculated positivity is: 2.3511834
155	Original file's positivity is: 2.09205020920502	Calculated positivity is: 2.4793837
156	Original file's positivity is: 1.46683673469388	Calculated positivity is: 2.457884
157	Original file's positivity is: 1.01325019485581	Calculated positivity is: 2.453481
158	Original file's positivity is: 4.80349344978166	Calculated positivity is: 2.3923764
159	Original file's positivity is: 2.25442834138486	Calculated positivity is: 2.357844
160	Original file's positivity is: 1.69491525423729	Calculated positivity is: 2.413675
161	Original file's positivity is: 4.82374768089054	Calculated positivity is: 2.3567083
162	Original file's positivity is: 1.06157112526539	Calculated positivity is: 2.4793844
163	Original file's positivity is: 0.819672131147541	Calculated positivity is: 2.2359495
164	Original file's positivity is: 3.5548686244204	Calculated positivity is: 2.3288023
165	Original file's positivity is: 0.970873786407767	Calculated positivity is: 2.4222336
166	Original file's positivity is: 1.69491525423729	Calculated positivity is: 2.277813
167	Original file's positivity is: 0.808276753960556	Calculated positivity is: 2.2651775
168	Original file's positivity is: 3.55450236966825	Calculated positivity is: 2.3311887
169	Original file's positivity is: 0	Calculated positivity is: 2.2911747
170	Original file's positivity is: 2.33918128654971	Calculated positivity is: 2.3589163
171	Original file's positivity is: 1.18343195266272	Calculated positivity is: 2.4548094
172	Original file's positivity is: 0.816326530612245	Calculated positivity is: 2.2923348
173	Original file's positivity is: 2.66544117647059	Calculated positivity is: 2.2081575
174	Original file's positivity is: 1.16550116550117	Calculated positivity is: 2.3621805
175	Original file's positivity is: 2.80898876404494	Calculated positivity is: 2.3492777

176	Original file's positivity is: 1.64835164835165	Calculated positivity is: 2.220882
177	Original file's positivity is: 3.18949343339587	Calculated positivity is: 2.2458465
178	Original file's positivity is: 2.27920227920228	Calculated positivity is: 2.4064326
179	Original file's positivity is: 1.50564617314931	Calculated positivity is: 2.3253074

Table A7.3: Calculated polarity values using clusters on a larger dataset vs. Original

polarity values

File Number	Original File Result (Polarization)	Calculated Result (Polarization, different calculation than original)
0	Original file's polarity is: 4.04624277456647	Calculated polarity is: 0.17910448
1	Original file's polarity is: 8.10372771474878	Calculated polarity is: 0.1109215
2	Original file's polarity is: 3.97553516819572	Calculated polarity is: 0.037037037
3	Original file's polarity is: 7.66917293233083	Calculated polarity is: 0.05357143
4	Original file's polarity is: 2.0979020979021	Calculated polarity is: 0.2991453
5	Original file's polarity is: 7.97872340425532	Calculated polarity is: 0.1171875
6	Original file's polarity is: 9.06666666666667	Calculated polarity is: 0.19402985
7	Original file's polarity is: 4.50261780104712	Calculated polarity is: 0.11642743
8	Original file's polarity is: 4.85036119711042	Calculated polarity is: 0.015276631
9	Original file's polarity is: 5.21609538002981	Calculated polarity is: 0.020547945
10	Original file's polarity is: 6.51890482398957	Calculated polarity is: 0.034136545
11	Original file's polarity is: 4.96277915632754	Calculated polarity is: 0.06413994
12	Original file's polarity is: 8.61313868613139	Calculated polarity is: 0.031825796
13	Original file's polarity is: 5.12820512820513	Calculated polarity is: 0.07467532
14	Original file's polarity is: 7.79569892473118	Calculated polarity is: 0.061833687
15	Original file's polarity is: 7.08898944193062	Calculated polarity is: 0.06387665
16	Original file's polarity is: 7.08898944193062	Calculated polarity is: 0.19266056
17	Original file's polarity is: 5.94900849858357	Calculated polarity is: 0.033163264
18	Original file's polarity is: 7.89889415481833	Calculated polarity is: 0.05870021
19	Original file's polarity is: 6.44677661169415	Calculated polarity is: 0.20353982
20	Original file's polarity is: 6.21931260229133	Calculated polarity is: 0.115384616
21	Original file's polarity is: 7.05765407554672	Calculated polarity is: 0.01968504
22	Original file's polarity is: 5.73476702508961	Calculated polarity is: 0.02749141
23	Original file's polarity is: 6.75039246467818	Calculated polarity is: 0.031313818
24	Original file's polarity is: 4.5045045045045	Calculated polarity is: 0.04718417
25	Original file's polarity is: 5.24271844660194	Calculated polarity is: 0.09406953

26	Original file's polarity is: 6.25	Calculated polarity is: 0.07581227
27	Original file's polarity is: 6.0352422907489	Calculated polarity is: 0.029152542
28	Original file's polarity is: 10.0850546780073	Calculated polarity is: 0.06838906
29	Original file's polarity is: 9.09090909090909	Calculated polarity is: 0.10071942
30	Original file's polarity is: 4.90566037735849	Calculated polarity is: 0.0546875
31	Original file's polarity is: 7.52688172043011	Calculated polarity is: 0.017766498
32	Original file's polarity is: 9.29752066115702	Calculated polarity is: 0.05514706
33	Original file's polarity is: 5.12820512820513	Calculated polarity is: 0.030035336
34	Original file's polarity is: 9.08004778972521	Calculated polarity is: 0.09979424
35	Original file's polarity is: 7.21407624633431	Calculated polarity is: 0.083
36	Original file's polarity is: 5.5956678700361	Calculated polarity is: 0.035842296
37	Original file's polarity is: 7.13082537900056	Calculated polarity is: 0.03543743
38	Original file's polarity is: 2.9126213592233	Calculated polarity is: 0.10169491
39	Original file's polarity is: 1.90217391304348	Calculated polarity is: 0.12636165
40	Original file's polarity is: 5.31732418524871	Calculated polarity is: 0.084745765
41	Original file's polarity is: 5.80098170459616	Calculated polarity is: 0.11311054
42	Original file's polarity is: 3.03030303030303	Calculated polarity is: 0.2659176
43	Original file's polarity is: 6.42398286937901	Calculated polarity is: 0.14965035
44	Original file's polarity is: 8.62676056338028	Calculated polarity is: 0.116071425
45	Original file's polarity is: 4.12573673870334	Calculated polarity is: 0.030075189
46	Original file's polarity is: 1.12044817927171	Calculated polarity is: 0.08139535
47	Original file's polarity is: 2.80777537796976	Calculated polarity is: 0.09934854
48	Original file's polarity is: 4.30861723446894	Calculated polarity is: 0.07602339
49	Original file's polarity is: 1.87369882026371	Calculated polarity is: 0.1502783
50	Original file's polarity is: 5.6980056980057	Calculated polarity is: 0.047898337
51	Original file's polarity is: 10.4364326375712	Calculated polarity is: 0.045513652
52	Original file's polarity is: 5.3156146179402	Calculated polarity is: 0.12686567
53	Original file's polarity is: 12.0634920634921	Calculated polarity is: 0.026143791
54	Original file's polarity is: 7.7319587628866	Calculated polarity is: 0.033143938
55	Original file's polarity is: 4.71633629528366	Calculated polarity is: 0.034412954

56	Original file's polarity is: 1.71990171990172	Calculated polarity is: 0.054054055
57	Original file's polarity is: 6.06741573033708	Calculated polarity is: 0.08571429
58	Original file's polarity is: 4.48598130841121	Calculated polarity is: 0.09843937
59	Original file's polarity is: 2.27272727272727	Calculated polarity is: 0.21538462
60	Original file's polarity is: 5.95802301963439	Calculated polarity is: 0.17548387
61	Original file's polarity is: 13.3004926108374	Calculated polarity is: 0.114241004
62	Original file's polarity is: 8.79265091863517	Calculated polarity is: 0.062322944
63	Original file's polarity is: 4.36005625879044	Calculated polarity is: 0.10982659
64	Original file's polarity is: 7.99031476997579	Calculated polarity is: 0.08615384
65	Original file's polarity is: 4.11392405063291	Calculated polarity is: 0.03432137
66	Original file's polarity is: 3.27706057596822	Calculated polarity is: 0.11255411
67	Original file's polarity is: 6.01092896174863	Calculated polarity is: 0.036277603
68	Original file's polarity is: 10.5128205128205	Calculated polarity is: 0.022222223
69	Original file's polarity is: 5.59006211180124	Calculated polarity is: 0.041666668
70	Original file's polarity is: 6.88524590163934	Calculated polarity is: 0.021487603
71	Original file's polarity is: 7.83132530120482	Calculated polarity is: 0.06635071
72	Original file's polarity is: 5.16934046345811	Calculated polarity is: 0.028422274
73	Original file's polarity is: 5.36130536130536	Calculated polarity is: 0.039583333
74	Original file's polarity is: 1.72413793103448	Calculated polarity is: 0.6847826
75	Original file's polarity is: 8.37837837837838	Calculated polarity is: 0.024647888
76	Original file's polarity is: 7.69230769230769	Calculated polarity is: 0.06323185
77	Original file's polarity is: 8.99470899470899	Calculated polarity is: 0.22772278
78	Original file's polarity is: 5.92592592592593	Calculated polarity is: 0.019512195
79	Original file's polarity is: 7.94824399260628	Calculated polarity is: 0.043184884
80	Original file's polarity is: 6.41658440276407	Calculated polarity is: 0.04054054
81	Original file's polarity is: 5.20581113801453	Calculated polarity is: 0.024509804
82	Original file's polarity is: 6.66666666666667	Calculated polarity is: 0.051454138
83	Original file's polarity is: 7.84313725490196	Calculated polarity is: 0.15402299
84	Original file's polarity is: 5.73770491803279	Calculated polarity is: 0.030701755
85	Original file's polarity is:	Calculated polarity is: 0.21451104

	7.55813953488372	
86	Original file's polarity is: 4.17956656346749	Calculated polarity is: 0.05065666
87	Original file's polarity is: 3.44827586206897	Calculated polarity is: 0.055031445
88	Original file's polarity is: 3.08370044052863	Calculated polarity is: 0.11646587
89	Original file's polarity is: 5.77956989247312	Calculated polarity is: 0.15544042
90	Original file's polarity is: 5.3921568627451	Calculated polarity is: 0.031939164
91	Original file's polarity is: 5.93922651933702	Calculated polarity is: 0.05222437
92	Original file's polarity is: 0.813008130081301	Calculated polarity is: 0.25174826
93	Original file's polarity is: 3.49243306169965	Calculated polarity is: 0.05851064
94	Original file's polarity is: 8.09523809523809	Calculated polarity is: 0.02877698
95	Original file's polarity is: 6.71428571428571	Calculated polarity is: 0.04290429
96	Original file's polarity is: 4.22163588390501	Calculated polarity is: 0.011167512
97	Original file's polarity is: 4.65116279069767	Calculated polarity is: 0.04411765
98	Original file's polarity is: 6.62460567823344	Calculated polarity is: 0.13114753
99	Original file's polarity is: 5.97402597402597	Calculated polarity is: 0.04950495
100	Original file's polarity is: 8.33333333333333	Calculated polarity is: 0.035264485
101	Original file's polarity is: 9.42408376963351	Calculated polarity is: 0.09090909
102	Original file's polarity is: 1.62162162162162	Calculated polarity is: 0.15841584
103	Original file's polarity is: 11.9565217391304	Calculated polarity is: 0.022315202
104	Original file's polarity is: 6.35838150289017	Calculated polarity is: 0.07174888
105	Original file's polarity is: 3.7410071942446	Calculated polarity is: 0.125
106	Original file's polarity is: 6.30782169890664	Calculated polarity is: 0.03373016
107	Original file's polarity is: 7.54901960784314	Calculated polarity is: 0.036036037
108	Original file's polarity is: 5.44464609800363	Calculated polarity is: 0.1780303
109	Original file's polarity is: 5.0682261208577	Calculated polarity is: 0.17114094
110	Original file's polarity is: 1.74825174825175	Calculated polarity is: 0.14634146
111	Original file's polarity is: 2.46305418719212	Calculated polarity is: 0.009727626
112	Original file's polarity is: 9.68523002421307	Calculated polarity is: 0.06926407
113	Original file's polarity is: 5.38720538720539	Calculated polarity is: 0.036057692
114	Original file's polarity is: 4.55764075067024	Calculated polarity is: 0.07109005

115	Original file's polarity is: 9.42528735632184	Calculated polarity is: 0.039179105
116	Original file's polarity is: 1.68269230769231	Calculated polarity is: 0.117117114
117	Original file's polarity is: 5.90500641848524	Calculated polarity is: 0.16791044
118	Original file's polarity is: 0.816326530612245	Calculated polarity is: 0.192
119	Original file's polarity is: 9.09090909090909	Calculated polarity is: 0.058641974
120	Original file's polarity is: 3.53200883002207	Calculated polarity is: 0.15286624
121	Original file's polarity is: 6.09137055837564	Calculated polarity is: 0.04680468
122	Original file's polarity is: 7.20720720720721	Calculated polarity is: 0.055658627
123	Original file's polarity is: 5.30503978779841	Calculated polarity is: 0.05506608
124	Original file's polarity is: 8.33333333333333	Calculated polarity is: 0.14084508
125	Original file's polarity is: 11.8556701030928	Calculated polarity is: 0.0029411765
126	Original file's polarity is: 3.07377049180328	Calculated polarity is: 0.038229376
127	Original file's polarity is: 5.98555211558307	Calculated polarity is: 0.012884043
128	Original file's polarity is: 4.89795918367347	Calculated polarity is: 0.06451613
129	Original file's polarity is: 3.55329949238579	Calculated polarity is: 0.043392505
130	Original file's polarity is: 5.6	Calculated polarity is: 0.07070707
131	Original file's polarity is: 8.95140664961637	Calculated polarity is: 0.026624069
132	Original file's polarity is: 7.12971481140754	Calculated polarity is: 0.058768656
133	Original file's polarity is: 6.78627145085803	Calculated polarity is: 0.08655617
134	Original file's polarity is: 8.16326530612245	Calculated polarity is: 0.010619469
135	Original file's polarity is: 10.27445460943	Calculated polarity is: 0.14888337
136	Original file's polarity is: 3.43213728549142	Calculated polarity is: 0.08431373
137	Original file's polarity is: 6.0665362035225	Calculated polarity is: 0.11278196
138	Original file's polarity is: 6.04490500863558	Calculated polarity is: 0.06407323
139	Original file's polarity is: 7.43982494529541	Calculated polarity is: 0.16477273
140	Original file's polarity is: 6.83520599250936	Calculated polarity is: 0.14215687
141	Original file's polarity is: 7.64262648008611	Calculated polarity is: 0.02038925
142	Original file's polarity is: 7.9447322970639	Calculated polarity is: 0.08830022
143	Original file's polarity is: 6.05536332179931	Calculated polarity is: 0.08442777
144	Original file's polarity is: 2.35294117647059	Calculated polarity is: 0.09146342

145	Original file's polarity is: 6.81341719077568	Calculated polarity is: 0.12939522
146	Original file's polarity is: 6.01851851851852	Calculated polarity is: 0.067857146
147	Original file's polarity is: 11.2107623318386	Calculated polarity is: 0.10796915
148	Original file's polarity is: 7.23981900452489	Calculated polarity is: 0.06344411
149	Original file's polarity is: 4.73833097595474	Calculated polarity is: 0.031083481
150	Original file's polarity is: 5.45454545454545	Calculated polarity is: 0.021598272
151	Original file's polarity is: 11.002661934339	Calculated polarity is: 0.0704698
152	Original file's polarity is: 4.72008781558727	Calculated polarity is: 0.25
153	Original file's polarity is: 7.38137082601054	Calculated polarity is: 0.1451049
154	Original file's polarity is: 6.39386189258312	Calculated polarity is: 0.296063
155	Original file's polarity is: 7.11297071129707	Calculated polarity is: 0.025706941
156	Original file's polarity is: 6.88775510204082	Calculated polarity is: 0.03732504
157	Original file's polarity is: 2.10444271239283	Calculated polarity is: 0.07862408
158	Original file's polarity is: 5.24017467248908	Calculated polarity is: 0.023622047
159	Original file's polarity is: 4.83091787439614	Calculated polarity is: 0.060763888
160	Original file's polarity is: 3.10734463276836	Calculated polarity is: 0.027707808
161	Original file's polarity is: 8.71985157699443	Calculated polarity is: 0.04849601
162	Original file's polarity is: 4.45859872611465	Calculated polarity is: 0.029962547
163	Original file's polarity is: 3.27868852459016	Calculated polarity is: 0.021276595
164	Original file's polarity is: 3.70942812982998	Calculated polarity is: 0.09293681
165	Original file's polarity is: 3.88349514563107	Calculated polarity is: 0.058165547
166	Original file's polarity is: 4.10546139359699	Calculated polarity is: 0.09585121
167	Original file's polarity is: 1.90753313934691	Calculated polarity is: 0.06286081
168	Original file's polarity is: 4.739336492891	Calculated polarity is: 0.041666668
169	Original file's polarity is: 1.21951219512195	Calculated polarity is: 0.10714286
170	Original file's polarity is: 5.84795321637427	Calculated polarity is: 0.07072368
171	Original file's polarity is: 5.62130177514793	Calculated polarity is: 0.122977346
172	Original file's polarity is: 3.26530612244898	Calculated polarity is: 0.036190476
173	Original file's polarity is: 4.04411764705882	Calculated polarity is: 0.117338
174	Original file's polarity is:	Calculated polarity is: 0.1451049

	6.06060606060606	
175	Original file's polarity is: 2.80898876404494	Calculated polarity is: 0.032352943
176	Original file's polarity is: 5.49450549450549	Calculated polarity is: 0.099378884
177	Original file's polarity is: 4.50281425891182	Calculated polarity is: 0.14662756
178	Original file's polarity is: 4.74833808167141	Calculated polarity is: 0.05940594
179	Original file's polarity is: 6.64993726474278	Calculated polarity is: 0.062898815

h. Tables A8.0-A8.3:

Table A8.0: Calculated sentiment values for locations using clusters on a larger

dataset vs. Original GDELT average sentiment values

Location	Original File Result (Average)	Calculated Result
Location is: China	Original files' average tone is: -2.0371635	Calculated result is: -5.7690387
Location is: Canada	Original files' average tone is: -0.57163346	Calculated result is: -5.798462
Location is: United Kingdom	Original files' average tone is: -0.8766754	Calculated result is: -5.8459535
Location is: United States	Original files' average tone is: -0.56520724	Calculated result is: -5.801475
Location is: Brazil	Original files' average tone is: -3.3280118	Calculated result is: -5.7032747
Location is: Pakistan	Original files' average tone is: 2.797428	Calculated result is: -5.77931
Location is: Jersey	Original files' average tone is: -3.7920845	Calculated result is: -5.9172125
Location is: Ireland	Original files' average tone is: 0.14568973	Calculated result is: -5.8128166
Location is: Germany	Original files' average tone is: -2.4737775	Calculated result is: -5.767595
Location is: Japan	Original files' average tone is: -1.3979262	Calculated result is: -5.7632194
Location is: France	Original files' average tone is: -2.018263	Calculated result is: -5.760124
Location is: Iran	Original files' average tone is: -3.102877	Calculated result is: -5.8220787
Location is: Italy	Original files' average tone is: -2.0734982	Calculated result is: -5.775991
Location is: Egypt	Original files' average tone is: -1.586472	Calculated result is: -5.840026
Location is: Australia	Original files' average tone is: -0.8530742	Calculated result is: -5.804755
Location is: New Zealand	Original files' average tone is: -1.0188055	Calculated result is: -5.7760267
Location is: South Korea	Original files' average tone is: -0.4518088	Calculated result is: -5.7209845
Location is: North Korea	Original files' average tone is: -1.6435986	Calculated result is: -5.642949
Location is: Russia	Original files' average tone is: -0.7793609	Calculated result is: -5.723633
Location is: Taiwan	Original files' average tone is: -2.6276898	Calculated result is: -5.765448
Location is: Switzerland	Original files' average tone is: -4.6405663	Calculated result is: -5.8360887
Location is: Belgium	Original files' average tone is: -0.5775709	Calculated result is: -5.826577
Location is: Netherlands	Original files' average tone is: -0.93544227	Calculated result is: -5.725334
Location is: Croatia	Original files' average tone is: -1.0752689	Calculated result is: -5.847474
Location is: Philippines	Original files' average tone is: -0.15082957	Calculated result is: -5.754113

Location is: India	Original files' average tone is: -1.1626252	Calculated result is: -5.8352094
Location is: Sweden	Original files' average tone is: 2.3554	Calculated result is: -5.821808
Location is: Honduras	Original files' average tone is: -1.3884336	Calculated result is: -5.752496
Location is: Guatemala	Original files' average tone is: -1.3884336	Calculated result is: -5.752496
Location is: El Salvador	Original files' average tone is: -2.527646	Calculated result is: -5.9853344
Location is: Greece	Original files' average tone is: -0.7496252	Calculated result is: -5.8654156
Location is: Portugal	Original files' average tone is: 3.6779325	Calculated result is: -5.843486
Location is: Greenland	Original files' average tone is: -2.6391895	Calculated result is: -5.746065
Location is: Denmark	Original files' average tone is: -0.65062577	Calculated result is: -5.7503767
Location is: Syria	Original files' average tone is: -1.9179677	Calculated result is: -5.7293506
Location is: Mexico	Original files' average tone is: 1.6731753	Calculated result is: -5.8063745
Location is: Madagascar	Original files' average tone is: -0.04405286	Calculated result is: -5.7192307
Location is: Bolivia	Original files' average tone is: -3.7392924	Calculated result is: -5.849245
Location is: Paraguay	Original files' average tone is: -8.655333	Calculated result is: -5.829954
Location is: Peru	Original files' average tone is: -3.1096537	Calculated result is: -5.817692
Location is: Norway	Original files' average tone is: -1.132987	Calculated result is: -5.8497486
Location is: Nigeria	Original files' average tone is: 3.2965508	Calculated result is: -5.8392124
Location is: Trinidad And Tobago	Original files' average tone is: 0.8152174	Calculated result is: -5.701764
Location is: Hong Kong	Original files' average tone is: -3.004282	Calculated result is: -5.807829
Location is: Malaysia	Original files' average tone is: 0.07425678	Calculated result is: -5.738511
Location is: Thailand	Original files' average tone is: 0.41024688	Calculated result is: -5.7058992
Location is: Singapore	Original files' average tone is: -0.18349831	Calculated result is: -5.8344183
Location is: Republic Of	Original files' average tone is: 0.6109917	Calculated result is: -5.8125134
Location is: Brunei	Original files' average tone is: -0.23310024	Calculated result is: -5.7401905
Location is: Chile	Original files' average tone is: -0.3622344	Calculated result is: -5.8619986
Location is: Kenya	Original files' average tone is: -2.9863026	Calculated result is: -5.8102612
Location is: Malawi	Original files' average tone is: -0.81135905	Calculated result is: -5.884449
Location is: Ghana	Original files' average tone is: -0.81135905	Calculated result is: -5.884449
Location is: Colombia	Original files' average tone is: -0.24922119	Calculated result is: -5.519656

Location is: Bahrain	Original files' average tone is: 0.8802817	Calculated result is: -5.7589846
Location is: Afghanistan	Original files' average tone is: 0.91901654	Calculated result is: -5.758267
Location is: Rwanda	Original files' average tone is: 1.7404857	Calculated result is: -5.905871
Location is: Angola	Original files' average tone is: 5.590062	Calculated result is: -6.0262938
Location is: Democratic Republic Of The Congo	Original files' average tone is: 5.590062	Calculated result is: -6.0262938
Location is: Uganda	Original files' average tone is: 5.590062	Calculated result is: -6.0262938
Location is: Spain	Original files' average tone is: -1.092583	Calculated result is: -5.8202324
Location is: Slovenia	Original files' average tone is: -5.5737705	Calculated result is: -5.7170105
Location is: Sri Lanka	Original files' average tone is: -1.7241379	Calculated result is: -5.8741565
Location is: Indonesia	Original files' average tone is: -1.6269712	Calculated result is: -5.7679653
Location is: Ethiopia	Original files' average tone is: -2.1836064	Calculated result is: -5.7696238
Location is: Cameroon	Original files' average tone is: 1.884058	Calculated result is: -5.792429
Location is: Finland	Original files' average tone is: 1.884058	Calculated result is: -5.792429
Location is: Venezuela	Original files' average tone is: 1.884058	Calculated result is: -5.792429
Location is: Lebanon	Original files' average tone is: 0.8196721	Calculated result is: -5.967621
Location is: Turkey	Original files' average tone is: 1.76347	Calculated result is: -5.7988443
Location is: Argentina	Original files' average tone is: -2.489025	Calculated result is: -5.8365226
Location is: Poland	Original files' average tone is: -0.44052863	Calculated result is: -5.8829317
Location is: South Africa	Original files' average tone is: -2.9838448	Calculated result is: -5.742217
Location is: Azerbaijan	Original files' average tone is: -0.8130081	Calculated result is: -5.974862
Location is: Austria	Original files' average tone is: -3.7854888	Calculated result is: -5.8207264
Location is: Cuba	Original files' average tone is: -2.7754416	Calculated result is: -5.8490562
Location is: Qatar	Original files' average tone is: -0.5834694	Calculated result is: -5.7731237
Location is: Iraq	Original files' average tone is: -4.521073	Calculated result is: -5.8545833
Location is: Sudan	Original files' average tone is: -4.521073	Calculated result is: -5.8545833
Location is: Saudi Arabia	Original files' average tone is: -4.521073	Calculated result is: -5.8545833
Location is: Libya	Original files' average tone is: -3.3150818	Calculated result is: -5.820015
Location is: Serbia	Original files' average tone is: 0.81632656	Calculated result is: -5.992132
Location is:	Original files' average tone is:	Calculated result is: -5.992132

Montenegro	0.81632656	
Location is: Oceans	Original files' average tone is: 0.29433408	Calculated result is: -5.5929146
Location is: Ukraine	Original files' average tone is: 1.0221021	Calculated result is: -5.7012978
Location is: Mauritius	Original files' average tone is: -2.0304568	Calculated result is: -5.905216
Location is: Senegal	Original files' average tone is: -4.3376455	Calculated result is: -5.7811112
Location is: Israel	Original files' average tone is: -0.02692753	Calculated result is: -5.777399
Location is: Somalia	Original files' average tone is: -1.7939283	Calculated result is: -5.7748423
Location is: Tunisia	Original files' average tone is: -5.2516413	Calculated result is: -5.9740534
Location is: Niger	Original files' average tone is: 2.2006896	Calculated result is: -5.832275
Location is: Chad	Original files' average tone is: -6.5662003	Calculated result is: -5.776774
Location is: Namibia	Original files' average tone is: -6.5662003	Calculated result is: -5.776774
Location is: Mali	Original files' average tone is: -6.5662003	Calculated result is: -5.776774
Location is: Zimbabwe	Original files' average tone is: -6.5662003	Calculated result is: -5.776774
Location is: Botswana	Original files' average tone is: -6.5662003	Calculated result is: -5.776774
Location is: Central African Republic	Original files' average tone is: -6.5662003	Calculated result is: -5.776774
Location is: United Arab Emirates	Original files' average tone is: -3.254438	Calculated result is: -5.6925793
Location is: Aruba	Original files' average tone is: 1.2867647	Calculated result is: -5.8100376

Table A8.1: Calculated negative sentiment values for locations using clusters on

larger dataset vs. Original GDELT average negative tone values

Location	Original File Result (Negativity - Average)	Calculated Result (Negativity)
Location is: China	Original files' average negativity is: -3.8480163	Calculated negativity is: -8.156233
Location is: Canada	Original files' average negativity is: -2.7043316	Calculated negativity is: -8.197832
Location is: United Kingdom	Original files' average negativity is: -3.2977118	Calculated negativity is: -8.264974
Location is: United States	Original files' average negativity is: -3.4962513	Calculated negativity is: -8.20209
Location is: Brazil	Original files' average negativity is: -5.0211368	Calculated negativity is: -8.063258
Location is: Pakistan	Original files' average negativity is: -1.6376235	Calculated negativity is: -8.170753
Location is: Jersey	Original files' average negativity is: -4.823012	Calculated negativity is: -8.36572
Location is: Ireland	Original files' average negativity is: -3.247582	Calculated negativity is: -8.218123
Location is: Germany	Original files' average negativity is: -4.36491	Calculated negativity is: -8.154194
Location is: Japan	Original files' average negativity is: -3.5080044	Calculated negativity is: -8.148006
Location is: France	Original files' average negativity is: -4.1840563	Calculated negativity is: -8.14363
Location is: Iran	Original files' average negativity is: -5.343085	Calculated negativity is: -8.23122
Location is: Italy	Original files' average negativity is: -4.414668	Calculated negativity is: -8.166062
Location is: Egypt	Original files' average negativity is: -4.0909104	Calculated negativity is: -8.256594
Location is: Australia	Original files' average negativity is: -3.2028916	Calculated negativity is: -8.206729
Location is: New Zealand	Original files' average negativity is: -2.9832842	Calculated negativity is: -8.166113
Location is: South Korea	Original files' average negativity is: -3.6784513	Calculated negativity is: -8.088296
Location is: North Korea	Original files' average negativity is: -5.363322	Calculated negativity is: -7.977969
Location is: Russia	Original files' average negativity is: -2.7769914	Calculated negativity is: -8.09204
Location is: Taiwan	Original files' average negativity is: -4.5401344	Calculated negativity is: -8.151155
Location is: Switzerland	Original files' average negativity is: -5.512991	Calculated negativity is: -8.251026
Location is: Belgium	Original files' average negativity is: -3.3088508	Calculated negativity is: -8.237579
Location is: Netherlands	Original files' average negativity is: -4.027888	Calculated negativity is: -8.094444
Location is: Croatia	Original files' average negativity is: -4.435484	Calculated negativity is: -8.267122
Location is: Philippines	Original files' average negativity is: -3.6199095	Calculated negativity is: -8.135129
Location is:	Original files' average negativity is:	Calculated negativity is: -8.249784

India	-3.6216843	
Location is: Sweden	Original files' average negativity is: -2.1855733	Calculated negativity is: -8.23084
Location is: Honduras	Original files' average negativity is: -3.7904358	Calculated negativity is: -8.132846
Location is: Guatemala	Original files' average negativity is: -3.7904358	Calculated negativity is: -8.132846
Location is: El Salvador	Original files' average negativity is: -5.21327	Calculated negativity is: -8.462028
Location is: Greece	Original files' average negativity is: -3.5982008	Calculated negativity is: -8.292489
Location is: Portugal	Original files' average negativity is: -1.689861	Calculated negativity is: -8.261483
Location is: Greenland	Original files' average negativity is: -4.19432	Calculated negativity is: -8.123753
Location is: Denmark	Original files' average negativity is: -3.4948227	Calculated negativity is: -8.1298485
Location is: Syria	Original files' average negativity is: -4.0465817	Calculated negativity is: -8.100122
Location is: Mexico	Original files' average negativity is: -1.5592895	Calculated negativity is: -8.209019
Location is: Madagascar	Original files' average negativity is: -3.0396476	Calculated negativity is: -8.085814
Location is: Bolivia	Original files' average negativity is: -4.8133645	Calculated negativity is: -8.269628
Location is: Paraguay	Original files' average negativity is: -8.964452	Calculated negativity is: -8.2423525
Location is: Peru	Original files' average negativity is: -4.1194224	Calculated negativity is: -8.2250185
Location is: Norway	Original files' average negativity is: -3.3128245	Calculated negativity is: -8.270339
Location is: Nigeria	Original files' average negativity is: -1.4964145	Calculated negativity is: -8.255444
Location is: Trinidad And Tobago	Original files' average negativity is: -0.54347825	Calculated negativity is: -8.061121
Location is: Hong Kong	Original files' average negativity is: -4.715481	Calculated negativity is: -8.211079
Location is: Malaysia	Original files' average negativity is: -1.52217	Calculated negativity is: -8.113074
Location is: Thailand	Original files' average negativity is: -2.1182544	Calculated negativity is: -8.066971
Location is: Singapore	Original files' average negativity is: -2.6826918	Calculated negativity is: -8.248669
Location is: Republic Of	Original files' average negativity is: -1.7352954	Calculated negativity is: -8.217698
Location is: Brunei	Original files' average negativity is: -1.6317016	Calculated negativity is: -8.115448
Location is: Chile	Original files' average negativity is: -1.7272574	Calculated negativity is: -8.287661
Location is: Kenya	Original files' average negativity is: -5.217534	Calculated negativity is: -8.214513
Location is: Malawi	Original files' average negativity is: -4.665314	Calculated negativity is: -8.319397
Location is: Ghana	Original files' average negativity is: -4.665314	Calculated negativity is: -8.319397
Location is: Colombia	Original files' average negativity is: -2.3676012	Calculated negativity is: -7.8036637
Location is:	Original files' average negativity is:	Calculated negativity is: -8.142017

Bahrain	-1.4084507	
Location is: Afghanistan	Original files' average negativity is: -3.3204992	Calculated negativity is: -8.141004
Location is: Rwanda	Original files' average negativity is: -1.9272727	Calculated negativity is: -8.349685
Location is: Angola	Original files' average negativity is: 0.0	Calculated negativity is: -8.5199375
Location is: Democratic Republic Of The Congo	Original files' average negativity is: 0.0	Calculated negativity is: -8.5199375
Location is: Uganda	Original files' average negativity is: 0.0	Calculated negativity is: -8.5199375
Location is: Spain	Original files' average negativity is: -3.8479786	Calculated negativity is: -8.228609
Location is: Slovenia	Original files' average negativity is: -6.2295084	Calculated negativity is: -8.082674
Location is: Sri Lanka	Original files' average negativity is: -1.7241379	Calculated negativity is: -8.304848
Location is: Indonesia	Original files' average negativity is: -2.7718506	Calculated negativity is: -8.1547165
Location is: Ethiopia	Original files' average negativity is: -4.4560294	Calculated negativity is: -8.157061
Location is: Cameroon	Original files' average negativity is: -2.3913043	Calculated negativity is: -8.1893015
Location is: Finland	Original files' average negativity is: -2.3913043	Calculated negativity is: -8.1893015
Location is: Venezuela	Original files' average negativity is: -2.3913043	Calculated negativity is: -8.1893015
Location is: Lebanon	Original files' average negativity is: -2.4590163	Calculated negativity is: -8.436985
Location is: Turkey	Original files' average negativity is: -1.2786773	Calculated negativity is: -8.198373
Location is: Argentina	Original files' average negativity is: -4.371459	Calculated negativity is: -8.251642
Location is: Poland	Original files' average negativity is: -1.7621145	Calculated negativity is: -8.317257
Location is: South Africa	Original files' average negativity is: -4.688898	Calculated negativity is: -8.118312
Location is: Azerbaijan	Original files' average negativity is: -0.8130081	Calculated negativity is: -8.447224
Location is: Austria	Original files' average negativity is: -5.205047	Calculated negativity is: -8.229305
Location is: Cuba	Original files' average negativity is: -4.5416317	Calculated negativity is: -8.269361
Location is: Qatar	Original files' average negativity is: -4.3446245	Calculated negativity is: -8.162006
Location is: Iraq	Original files' average negativity is: -6.9731803	Calculated negativity is: -8.277173
Location is: Sudan	Original files' average negativity is: -6.9731803	Calculated negativity is: -8.277173
Location is: Saudi Arabia	Original files' average negativity is: -6.9731803	Calculated negativity is: -8.277173
Location is: Libya	Original files' average negativity is: -5.4138627	Calculated negativity is: -8.228302
Location is: Serbia	Original files' average negativity is: 0.0	Calculated negativity is: -8.47164
Location is: Montenegro	Original files' average negativity is: 0.0	Calculated negativity is: -8.47164

Location is: Oceans	Original files' average negativity is: -1.6188374	Calculated negativity is: -7.907235
Location is: Ukraine	Original files' average negativity is: -0.95244396	Calculated negativity is: -8.060461
Location is: Mauritius	Original files' average negativity is: -4.0609136	Calculated negativity is: -8.34876
Location is: Senegal	Original files' average negativity is: -5.4794793	Calculated negativity is: -8.1733
Location is: Israel	Original files' average negativity is: -3.2910404	Calculated negativity is: -8.168054
Location is: Somalia	Original files' average negativity is: -4.4618216	Calculated negativity is: -8.16444
Location is: Tunisia	Original files' average negativity is: -6.345733	Calculated negativity is: -8.44608
Location is: Niger	Original files' average negativity is: -2.1195254	Calculated negativity is: -8.245635
Location is: Chad	Original files' average negativity is: -7.1044135	Calculated negativity is: -8.167167
Location is: Namibia	Original files' average negativity is: -7.1044135	Calculated negativity is: -8.167167
Location is: Mali	Original files' average negativity is: -7.1044135	Calculated negativity is: -8.167167
Location is: Zimbabwe	Original files' average negativity is: -7.1044135	Calculated negativity is: -8.167167
Location is: Botswana	Original files' average negativity is: -7.1044135	Calculated negativity is: -8.167167
Location is: Central African Republic	Original files' average negativity is: -7.1044135	Calculated negativity is: -8.167167
Location is: United Arab Emirates	Original files' average negativity is: -4.43787	Calculated negativity is: -8.048134
Location is: Aruba	Original files' average negativity is: -1.3786764	Calculated negativity is: -8.214197

Table A8.2: Calculated positive sentiment values for locations using clusters on a

larger dataset vs. Original GDELT average positive tone values

Location	Original File Result (Positivity - Average)	Calculated Result (Positivity)
Location is: China	Original files' average positivity is: 1.8108537	Calculated positivity is: 2.3871942
Location is: Canada	Original files' average positivity is: 2.1326973	Calculated positivity is: 2.39937
Location is: United Kingdom	Original files' average positivity is: 2.421037	Calculated positivity is: 2.4190202
Location is: United States	Original files' average positivity is: 2.931045	Calculated positivity is: 2.4006152
Location is: Brazil	Original files' average positivity is: 1.6931261	Calculated positivity is: 2.3599832
Location is: Pakistan	Original files' average positivity is: 4.4350514	Calculated positivity is: 2.391442
Location is: Jersey	Original files' average positivity is: 1.0309278	Calculated positivity is: 2.448507
Location is: Ireland	Original files' average positivity is: 3.3932717	Calculated positivity is: 2.4053068
Location is: Germany	Original files' average positivity is: 1.8911324	Calculated positivity is: 2.386599
Location is: Japan	Original files' average positivity is: 2.1100776	Calculated positivity is: 2.384787
Location is: France	Original files' average positivity is: 2.1657927	Calculated positivity is: 2.383506
Location is: Iran	Original files' average positivity is: 2.240208	Calculated positivity is: 2.4091418
Location is: Italy	Original files' average positivity is: 2.3411694	Calculated positivity is: 2.3900714
Location is: Egypt	Original files' average positivity is: 2.5044384	Calculated positivity is: 2.416568
Location is: Australia	Original files' average positivity is: 2.3498173	Calculated positivity is: 2.4019737
Location is: New Zealand	Original files' average positivity is: 1.9644789	Calculated positivity is: 2.390086
Location is: South Korea	Original files' average positivity is: 3.2266426	Calculated positivity is: 2.3673115
Location is: North Korea	Original files' average positivity is: 3.7197232	Calculated positivity is: 2.3350198
Location is: Russia	Original files' average positivity is: 1.9976305	Calculated positivity is: 2.368407
Location is: Taiwan	Original files' average positivity is: 1.9124441	Calculated positivity is: 2.3857074
Location is: Switzerland	Original files' average positivity is: 0.87242484	Calculated positivity is: 2.4149375
Location is: Belgium	Original files' average positivity is: 2.7312796	Calculated positivity is: 2.4110024
Location is: Netherlands	Original files' average positivity is: 3.0924454	Calculated positivity is: 2.36911
Location is: Croatia	Original files' average positivity is: 3.360215	Calculated positivity is: 2.419648
Location is: Philippines	Original files' average positivity is: 3.46908	Calculated positivity is: 2.3810158
Location is:	Original files' average positivity is:	Calculated positivity is: 2.414575

India	2.4590592	
Location is: Sweden	Original files' average positivity is: 4.540973	Calculated positivity is: 2.4090319
Location is: Honduras	Original files' average positivity is: 2.402002	Calculated positivity is: 2.3803504
Location is: Guatemala	Original files' average positivity is: 2.402002	Calculated positivity is: 2.3803504
Location is: El Salvador	Original files' average positivity is: 2.6856241	Calculated positivity is: 2.4766932
Location is: Greece	Original files' average positivity is: 2.8485758	Calculated positivity is: 2.4270737
Location is: Portugal	Original files' average positivity is: 5.3677936	Calculated positivity is: 2.4179976
Location is: Greenland	Original files' average positivity is: 1.5551306	Calculated positivity is: 2.3776877
Location is: Denmark	Original files' average positivity is: 2.844197	Calculated positivity is: 2.379472
Location is: Syria	Original files' average positivity is: 2.1286142	Calculated positivity is: 2.370772
Location is: Mexico	Original files' average positivity is: 3.232465	Calculated positivity is: 2.402644
Location is: Madagascar	Original files' average positivity is: 2.9955947	Calculated positivity is: 2.3665838
Location is: Bolivia	Original files' average positivity is: 1.0740722	Calculated positivity is: 2.4203827
Location is: Paraguay	Original files' average positivity is: 0.30911902	Calculated positivity is: 2.4123986
Location is: Peru	Original files' average positivity is: 1.0097688	Calculated positivity is: 2.407327
Location is: Norway	Original files' average positivity is: 2.179838	Calculated positivity is: 2.4205904
Location is: Nigeria	Original files' average positivity is: 4.792965	Calculated positivity is: 2.4162312
Location is: Trinidad And Tobago	Original files' average positivity is: 1.3586956	Calculated positivity is: 2.3593566
Location is: Hong Kong	Original files' average positivity is: 1.7111994	Calculated positivity is: 2.4032497
Location is: Malaysia	Original files' average positivity is: 1.5964267	Calculated positivity is: 2.3745635
Location is: Thailand	Original files' average positivity is: 2.5285013	Calculated positivity is: 2.3610713
Location is: Singapore	Original files' average positivity is: 2.4991934	Calculated positivity is: 2.4142501
Location is: Republic Of	Original files' average positivity is: 2.346287	Calculated positivity is: 2.4051845
Location is: Brunei	Original files' average positivity is: 1.3986014	Calculated positivity is: 2.3752575
Location is: Chile	Original files' average positivity is: 1.3650229	Calculated positivity is: 2.4256623
Location is: Kenya	Original files' average positivity is: 2.2312307	Calculated positivity is: 2.4042516
Location is: Malawi	Original files' average positivity is: 3.8539553	Calculated positivity is: 2.4349482
Location is: Ghana	Original files' average positivity is: 3.8539553	Calculated positivity is: 2.4349482
Location is: Colombia	Original files' average positivity is: 2.11838	Calculated positivity is: 2.2840075
Location is:	Original files' average positivity is:	Calculated positivity is: 2.383033

Bahrain	2.2887323	
Location is: Afghanistan	Original files' average positivity is: 4.239516	Calculated positivity is: 2.3827367
Location is: Rwanda	Original files' average positivity is: 3.6677585	Calculated positivity is: 2.4438138
Location is: Angola	Original files' average positivity is: 5.590062	Calculated positivity is: 2.4936435
Location is: Democratic Republic Of The Congo	Original files' average positivity is: 5.590062	Calculated positivity is: 2.4936435
Location is: Uganda	Original files' average positivity is: 5.590062	Calculated positivity is: 2.4936435
Location is: Spain	Original files' average positivity is: 2.755396	Calculated positivity is: 2.4083767
Location is: Slovenia	Original files' average positivity is: 0.6557377	Calculated positivity is: 2.3656638
Location is: Sri Lanka	Original files' average positivity is: 0.0	Calculated positivity is: 2.4306912
Location is: Indonesia	Original files' average positivity is: 1.1448791	Calculated positivity is: 2.3867512
Location is: Ethiopia	Original files' average positivity is: 2.2724226	Calculated positivity is: 2.3874369
Location is: Cameroon	Original files' average positivity is: 4.2753625	Calculated positivity is: 2.3968728
Location is: Finland	Original files' average positivity is: 4.2753625	Calculated positivity is: 2.3968728
Location is: Venezuela	Original files' average positivity is: 4.2753625	Calculated positivity is: 2.3968728
Location is: Lebanon	Original files' average positivity is: 3.2786884	Calculated positivity is: 2.4693644
Location is: Turkey	Original files' average positivity is: 3.0421474	Calculated positivity is: 2.3995287
Location is: Argentina	Original files' average positivity is: 1.8824334	Calculated positivity is: 2.4151196
Location is: Poland	Original files' average positivity is: 1.3215859	Calculated positivity is: 2.434325
Location is: South Africa	Original files' average positivity is: 1.7050532	Calculated positivity is: 2.3760948
Location is: Azerbaijan	Original files' average positivity is: 0.0	Calculated positivity is: 2.4723613
Location is: Austria	Original files' average positivity is: 1.4195584	Calculated positivity is: 2.408579
Location is: Cuba	Original files' average positivity is: 1.76619	Calculated positivity is: 2.4203043
Location is: Qatar	Original files' average positivity is: 3.7611551	Calculated positivity is: 2.3888829
Location is: Iraq	Original files' average positivity is: 2.4521072	Calculated positivity is: 2.4225898
Location is: Sudan	Original files' average positivity is: 2.4521072	Calculated positivity is: 2.4225898
Location is: Saudi Arabia	Original files' average positivity is: 2.4521072	Calculated positivity is: 2.4225898
Location is: Libya	Original files' average positivity is: 2.0987809	Calculated positivity is: 2.408287
Location is: Serbia	Original files' average positivity is: 0.81632656	Calculated positivity is: 2.4795072
Location is: Montenegro	Original files' average positivity is: 0.81632656	Calculated positivity is: 2.4795072

Location is: Oceans	Original files' average positivity is: 1.9131714	Calculated positivity is: 2.3143203
Location is: Ukraine	Original files' average positivity is: 1.9745461	Calculated positivity is: 2.3591635
Location is: Mauritius	Original files' average positivity is: 2.0304568	Calculated positivity is: 2.4435437
Location is: Senegal	Original files' average positivity is: 1.1418338	Calculated positivity is: 2.3921885
Location is: Israel	Original files' average positivity is: 3.264113	Calculated positivity is: 2.3906543
Location is: Somalia	Original files' average positivity is: 2.6678932	Calculated positivity is: 2.389598
Location is: Tunisia	Original files' average positivity is: 1.0940919	Calculated positivity is: 2.4720266
Location is: Niger	Original files' average positivity is: 4.3202147	Calculated positivity is: 2.41336
Location is: Chad	Original files' average positivity is: 0.53821313	Calculated positivity is: 2.3903928
Location is: Namibia	Original files' average positivity is: 0.53821313	Calculated positivity is: 2.3903928
Location is: Mali	Original files' average positivity is: 0.53821313	Calculated positivity is: 2.3903928
Location is: Zimbabwe	Original files' average positivity is: 0.53821313	Calculated positivity is: 2.3903928
Location is: Botswana	Original files' average positivity is: 0.53821313	Calculated positivity is: 2.3903928
Location is: Central African Republic	Original files' average positivity is: 0.53821313	Calculated positivity is: 2.3903928
Location is: United Arab Emirates	Original files' average positivity is: 1.183432	Calculated positivity is: 2.3555548
Location is: Aruba	Original files' average positivity is: 2.6654413	Calculated positivity is: 2.4041593

Table A8.3: Calculated polarity results for locations using clusters on a larger

dataset vs. Original GDELT average polarity values

Location	Original File Result (Polarization - Average)	Calculated Result (Polarization - different calculation than original)
Location is: China	Original files' average polarity is: 5.65887	Calculated polarity is: 0.071994275
Location is: Canada	Original files' average polarity is: 4.83703	Calculated polarity is: 0.058414306
Location is: United Kingdom	Original files' average polarity is: 5.718749	Calculated polarity is: 0.051593244
Location is: United States	Original files' average polarity is: 6.427304	Calculated polarity is: 0.055679962
Location is: Brazil	Original files' average polarity is: 6.714264	Calculated polarity is: 0.06560743
Location is: Pakistan	Original files' average polarity is: 6.0726757	Calculated polarity is: 0.053413842
Location is: Jersey	Original files' average polarity is: 5.8539395	Calculated polarity is: 0.053386647
Location is: Ireland	Original files' average polarity is: 6.640853	Calculated polarity is: 0.058484096
Location is: Germany	Original files' average polarity is: 6.256042	Calculated polarity is: 0.07402672
Location is: Japan	Original files' average polarity is: 5.618082	Calculated polarity is: 0.05988
Location is: France	Original files' average polarity is: 6.3498497	Calculated polarity is: 0.054738652
Location is: Iran	Original files' average polarity is: 7.5832925	Calculated polarity is: 0.050744206
Location is: Italy	Original files' average polarity is: 6.755837	Calculated polarity is: 0.05866594
Location is: Egypt	Original files' average polarity is: 6.5953493	Calculated polarity is: 0.07409099
Location is: Australia	Original files' average polarity is: 5.552709	Calculated polarity is: 0.06023559
Location is: New Zealand	Original files' average polarity is: 4.947763	Calculated polarity is: 0.07507379
Location is: South Korea	Original files' average polarity is: 6.9050937	Calculated polarity is: 0.07005579
Location is: North Korea	Original files' average polarity is: 9.083045	Calculated polarity is: 0.033474065
Location is: Russia	Original files' average polarity is: 4.7746224	Calculated polarity is: 0.071558766
Location is: Taiwan	Original files' average polarity is: 6.452578	Calculated polarity is: 0.052200228
Location is: Switzerland	Original files' average polarity is: 6.385416	Calculated polarity is: 0.039688446
Location is: Belgium	Original files' average polarity is: 6.04013	Calculated polarity is: 0.070879884
Location is: Netherlands	Original files' average polarity is: 7.120333	Calculated polarity is: 0.06744336
Location is: Croatia	Original files' average polarity is: 7.795699	Calculated polarity is: 0.06169666
Location is: Philippines	Original files' average polarity is: 7.0889893	Calculated polarity is: 0.043438915
Location is:	Original files' average polarity is:	Calculated polarity is: 0.05737722

India	6.080743	
Location is: Sweden	Original files' average polarity is: 6.726546	Calculated polarity is: 0.07948921
Location is: Honduras	Original files' average polarity is: 6.192438	Calculated polarity is: 0.07250502
Location is: Guatemala	Original files' average polarity is: 6.192438	Calculated polarity is: 0.07250502
Location is: El Salvador	Original files' average polarity is: 7.8988943	Calculated polarity is: 0.053753477
Location is: Greece	Original files' average polarity is: 6.4467764	Calculated polarity is: 0.035755478
Location is: Portugal	Original files' average polarity is: 7.0576544	Calculated polarity is: 0.02359882
Location is: Greenland	Original files' average polarity is: 5.7494507	Calculated polarity is: 0.03394187
Location is: Denmark	Original files' average polarity is: 6.33902	Calculated polarity is: 0.04318145
Location is: Syria	Original files' average polarity is: 6.175196	Calculated polarity is: 0.040862553
Location is: Mexico	Original files' average polarity is: 4.791755	Calculated polarity is: 0.062365305
Location is: Madagascar	Original files' average polarity is: 6.035242	Calculated polarity is: 0.029174427
Location is: Bolivia	Original files' average polarity is: 5.8874364	Calculated polarity is: 0.04351895
Location is: Paraguay	Original files' average polarity is: 9.27357	Calculated polarity is: 0.037322275
Location is: Peru	Original files' average polarity is: 5.129191	Calculated polarity is: 0.06758378
Location is: Norway	Original files' average polarity is: 5.4926624	Calculated polarity is: 0.08152712
Location is: Nigeria	Original files' average polarity is: 6.2893796	Calculated polarity is: 0.043153763
Location is: Trinidad And Tobago	Original files' average polarity is: 1.9021739	Calculated polarity is: 0.097353496
Location is: Hong Kong	Original files' average polarity is: 6.4266806	Calculated polarity is: 0.08919119
Location is: Malaysia	Original files' average polarity is: 3.1185966	Calculated polarity is: 0.06745098
Location is: Thailand	Original files' average polarity is: 4.6467557	Calculated polarity is: 0.10151053
Location is: Singapore	Original files' average polarity is: 5.1818857	Calculated polarity is: 0.08727291
Location is: Republic Of	Original files' average polarity is: 4.081583	Calculated polarity is: 0.059935875
Location is: Brunei	Original files' average polarity is: 3.030303	Calculated polarity is: 0.10299003
Location is: Chile	Original files' average polarity is: 3.0922801	Calculated polarity is: 0.0806876
Location is: Kenya	Original files' average polarity is: 7.4487643	Calculated polarity is: 0.05288921
Location is: Malawi	Original files' average polarity is: 8.51927	Calculated polarity is: 0.038541667
Location is: Ghana	Original files' average polarity is: 8.51927	Calculated polarity is: 0.038541667
Location is: Colombia	Original files' average polarity is: 4.4859815	Calculated polarity is: 0.09125657
Location is:	Original files' average polarity is:	Calculated polarity is: 0.056270096

Bahrain	3.6971831	
Location is: Afghanistan	Original files' average polarity is: 7.5600147	Calculated polarity is: 0.06608347
Location is: Rwanda	Original files' average polarity is: 5.595031	Calculated polarity is: 0.037272148
Location is: Angola	Original files' average polarity is: 5.590062	Calculated polarity is: 0.022408964
Location is: Democratic Republic Of The Congo	Original files' average polarity is: 5.590062	Calculated polarity is: 0.022408964
Location is: Uganda	Original files' average polarity is: 5.590062	Calculated polarity is: 0.022408964
Location is: Spain	Original files' average polarity is: 6.603374	Calculated polarity is: 0.037617866
Location is: Slovenia	Original files' average polarity is: 6.885246	Calculated polarity is: 0.03711559
Location is: Sri Lanka	Original files' average polarity is: 1.7241379	Calculated polarity is: 0.074291304
Location is: Indonesia	Original files' average polarity is: 3.9167295	Calculated polarity is: 0.04038433
Location is: Ethiopia	Original files' average polarity is: 6.728452	Calculated polarity is: 0.044487346
Location is: Cameroon	Original files' average polarity is: 6.6666665	Calculated polarity is: 0.047819234
Location is: Finland	Original files' average polarity is: 6.6666665	Calculated polarity is: 0.047819234
Location is: Venezuela	Original files' average polarity is: 6.6666665	Calculated polarity is: 0.047819234
Location is: Lebanon	Original files' average polarity is: 5.7377048	Calculated polarity is: 0.022033898
Location is: Turkey	Original files' average polarity is: 4.3208246	Calculated polarity is: 0.07146092
Location is: Argentina	Original files' average polarity is: 6.253892	Calculated polarity is: 0.072136775
Location is: Poland	Original files' average polarity is: 3.0837004	Calculated polarity is: 0.062439024
Location is: South Africa	Original files' average polarity is: 6.393951	Calculated polarity is: 0.042278036
Location is: Azerbaijan	Original files' average polarity is: 0.8130081	Calculated polarity is: 0.058912385
Location is: Austria	Original files' average polarity is: 6.6246057	Calculated polarity is: 0.0554371
Location is: Cuba	Original files' average polarity is: 6.3078218	Calculated polarity is: 0.050955415
Location is: Qatar	Original files' average polarity is: 8.10578	Calculated polarity is: 0.034935296
Location is: Iraq	Original files' average polarity is: 9.425287	Calculated polarity is: 0.027355623
Location is: Sudan	Original files' average polarity is: 9.425287	Calculated polarity is: 0.02735562
Location is: Saudi Arabia	Original files' average polarity is: 9.425287	Calculated polarity is: 0.027355623
Location is: Libya	Original files' average polarity is: 7.512644	Calculated polarity is: 0.039745476
Location is: Serbia	Original files' average polarity is: 0.81632656	Calculated polarity is: 0.03958692
Location is: Montenegro	Original files' average polarity is: 0.81632656	Calculated polarity is: 0.03958692

Location is: Oceans	Original files' average polarity is: 3.532009	Calculated polarity is: 0.11158493
Location is: Ukraine	Original files' average polarity is: 2.92699	Calculated polarity is: 0.06932268
Location is: Mauritius	Original files' average polarity is: 6.0913706	Calculated polarity is: 0.048543688
Location is: Senegal	Original files' average polarity is: 6.621313	Calculated polarity is: 0.03998688
Location is: Israel	Original files' average polarity is: 6.555153	Calculated polarity is: 0.055084493
Location is: Somalia	Original files' average polarity is: 7.129715	Calculated polarity is: 0.053833604
Location is: Tunisia	Original files' average polarity is: 7.439825	Calculated polarity is: 0.048920862
Location is: Niger	Original files' average polarity is: 6.4397407	Calculated polarity is: 0.04145206
Location is: Chad	Original files' average polarity is: 7.6426263	Calculated polarity is: 0.027838428
Location is: Namibia	Original files' average polarity is: 7.6426263	Calculated polarity is: 0.027838428
Location is: Mali	Original files' average polarity is: 7.6426263	Calculated polarity is: 0.027838428
Location is: Zimbabwe	Original files' average polarity is: 7.6426263	Calculated polarity is: 0.027838428
Location is: Botswana	Original files' average polarity is: 7.6426263	Calculated polarity is: 0.027838428
Location is: Central African Republic	Original files' average polarity is: 7.6426263	Calculated polarity is: 0.027838428
Location is: United Arab Emirates	Original files' average polarity is: 5.6213017	Calculated polarity is: 0.045643155
Location is: Aruba	Original files' average polarity is: 4.0441175	Calculated polarity is: 0.055028465