Tawatchai (Andy) Holmes tho78 81662793

Project Proposal

Summary

In this Project Proposal, I will be analyzing the response to the Ukraine and Russia War in the US compared to the top countries around the world. The dataset that I will be using to accomplish this will be a dataset that will be pulled from GDELT 2.0 Events Database using gdeltPyR or GDELT Summary using the URL search. I will be using the cosine similarity and vector algorithms to compare each of the response events (Positive, Negative and Neutral). The Intended result that I am hoping for is more countries to support Ukraine more than Russia.

Motivation

I chose this Research Question, because the Ukraine and Russia War is still currently ongoing and I am interested in how different top countries from all around the world are responding to the Ukraine and Russia War compared to the US (Positive or Negative or Neutral). This would be an interesting research topic to see which country is more positive towards Russia or Ukraine or which country is more negative towards Russia or Ukraine.

Background

GDELT 2.0 Events Database is an open dataset that attempts to make human society itself “computable”, leveraging the power of Google Cloud to fundamentally reimagine how we study the human world in real time at a planetary scale. GDELT 2.0 Events Database gathers information of major events from news reports around the world and converts it into datasets. GDELT 2.0 datasets update every 15 minutes, have real time translation of 65 languages, tracking events, discussion progression and much more. Cosine similarity is the measurement between two vectors of an inner product space, it is most often used to measure document similarity in text analysis. This will be used to compare the Positive, Negative and Neutral events between countries.

Research Question

The question that I chose is “The response to the Ukraine and Russia War in the US compared to the top countries around the world”. To make the research question relevant to the data set, I will be comparing the positive event response, negative event response and neutral event response for each of the countries that I chose. Using a vector algorithm, I can split the data into 3 sections Positive, Negative and Neutral. I can then compare the response for each of the events, using cosine similarity to see which country supports more or supports less.

Design and Methods

To accomplish my Research Question, First I will try to collect the dataset that is necessary to answer the question from GDELT 2.0 Event Database or GDELT Summary by early next week. I then will sort through the data/articles and pick out countries that I have selected to be compared (mid next week), then I will run the output through the neutral, negative and positive events codes that I have created, to collect the response for each of the countries (end of next week). If everything went smoothly, I will then use a cosine similarity to compare the response to each country, but run a vector convert function first to be used in the cosine similarity function.

Technical difficulties that I can see coming across while trying to sort through the dataset is that I will be dealing with a huge amount of data as at the moment Ukraine and Russia War is quite a hot topic for different News source to cover and the Limitation that I can see that I might come across is getting a positive/negative and neutral response from each of the countries about Russia and Ukraine to be compared to one another, as each country would only support the war or don't support the war.

References

GDELT Summary
https://api.gdeltproject.org/api/v2/summary/summary

GDELT 2.0 Event Database
https://blog.gdeltproject.org/gdelt-2-0-our-global-world-in-realtime/

Cosine Similarity
https://www.sciencedirect.com/topics/computer-science/cosine-similarity#:~:text=Cosine%20similarity%20measures%20the%20similarity,document%20similarity%20in%20text%20analysis.

gdeltPyR
https://linwoodc3.github.io/gdeltPyR/

This website to determine which Top Country supports Russia or Ukraine
https://dmerharyana.org/countries-supporting-russia-ukraine/

Tawatchai (Andy) Holmes tho78 81662793

Sample GDELT Project that I will be following to gather the dataset that I need from GDELT 2.0 Event Database
https://colab.research.google.com/drive/1sTsl_-f2ipgzqM6htsVdKjf4MZ3Ds2CW#scrollTo=yR53wBVVi02z