

### **Mini Project 3 – Jens Sabel – Assignment Proposal**

#### **Project statement:**

In these times of great uncertainty, I would like to see how the discussions have trended regarding some of the aspects contributing to the current sense of crisis around the world. For this project, I thought of looking mainly at aspects surrounding the Russian-Ukrainian war.

With higher electricity prices, higher cost of living, higher inflation and overall, more social uncertainty, this would be interesting to see how the discussions about this have trended this year (2022). The cost of living has become more expensive in Europe and the United States as inflation rises, so this project will mainly focus on these two areas.

#### **Problem importance:**

This project could be important because it would provide statistics about what people are most concerned about regarding the fallout of the Russian-Ukrainian war. Regarding the war itself, or about some aspect of the war that is reflected in people's life outside the immediate warring countries. As an example, the current energy crisis in Europe could be one aspect that one might want to look closer at it.

#### **Proposed approach:**

Firstly, I will set up the necessary tools and equipment for me to mine the data from Twitter. Here I will use Tweepy which is a great tool in Python to access the data using Twitters own API.

Secondly, I will target some of the words that are a bi-product of the Russian-Ukrainian war. These words are: “War, Conflict, Food Shortage, Electricity, Gas, Power, Inflation, Nuclear Weapons “. In order to get more accurate data from Europe, I will also mine for words that are the French and German translations of the words. After this, I will try to combine the translated words with its English counterpart.

Thirdly, I will extract the timestamps from when these tweets were posted on Twitter to be able to compare at which dates/times tweets were written about the different subjects. After this I will try to look at the news from days where some words are trending more than others. By doing this, I would try to tie trends with specific events that took place on the different days.

Lastly, I will visualize the data to easier see what words trended at which time. This should create a clearer picture of what issues were discussed when.

#### **The expected outcome:**

I expect that the discussions revolve mostly about the following subjects: war, electricity, gas, inflation, food shortage and nuclear weapons. I also expect that these different subjects will correlate to some major news from the Russian-Ukrainian war. Exact expected dates of some trending words are harder to try to specify, since a lot of events do happen close to each other and there can be some delay from when the discussion starts on Twitter after the news break.