# Tweet Analysis: Russia & Ukraine War

Group 1: Kata Mezo, Yaxin Liu, Qifei (Freda) Yuan,

Kemal Olgun, Mary Diermier

Columbia University

QMSS GR5067 Natural Language Processing in Social Sciences

Spring 2022 Final Project

# Table of Contents

## Abstract

This study examined sentiment analysis on Twitter data about Ukraine and Russia Conflict. Analyzing people's reactions toward the conflict can help us discover the sentiments behind their opinions in the text and see the pattern of the distribution of sentiment on similar issues. The contributions of this paper are: (1) We discussed how the sentiment of the tweets changed over a month. (2) We examined the pattern of the sentiment distribution between countries. (3) We explored how the length of the tweets change for the different sentiment.

## 1. Introduction

Online Social Networks (OSNs) have often been used to gauge public sentiment in response to wars or crises; they are uniquely able to capture rapidly changing mass opinions, as well as very authentic responses. With growing attention and popularity on OSN analysis, there may be pressure to change old school newsroom structures limiting information sharing, maybe leading to a more free flowing and honest discourse from the top down (Sacco, V. & Bossio, D., 2015). In this project, we chose to analyze the public's changing response to the conflict in Ukraine over time using Twitter as a tool to gauge large scale opinions and sentiment.

On February 24th, 2022, Russian President Valdimir Putin initiated the invasion of Ukraine, calling the attack a 'special military operation'. Many cities were targeted and more than 1.5 million civilians have been displaced. Many people took to social media to join the conversation, motivations including the spreading of propaganda, desire to rally for a social cause, and to join or create information discourse on this topic (Haq, Ehsan-Ul, et al, 2022). Our goal was to map how topics, sentiment, and objectivity changed as the crisis developed. We chose to compare data from three different days as the crisis continued. We examined data from shortly after the initial invasion, a week after that date, and then a month following. Over the course of this span, the conflict evolved, world leaders voiced opinions, and communities rallied over social media. Throughout this paper, we will employ natural language processing methods to analyze user sentiment, the changing patterns of discourse over time, and explanations as to why the patterns may exist.

## 2. Dataset description

### 2.1 Data Description

As it is discussed in the previous chapter these tweets started to be collected on February 27 so three days after the Ukraine invasion officially started. We wanted to look at the day closest to the first day of the invasion, therefore we chose February 27 as our day 1. March 4th, so a week after the Ukraine-Russia war started, was chosen to be day 2 for our analysis. Finally we looked at March 24th as our day three, a month after the war started.Altogether we examined one month long period to see how the sentiments and the main topics of these tweets changed.

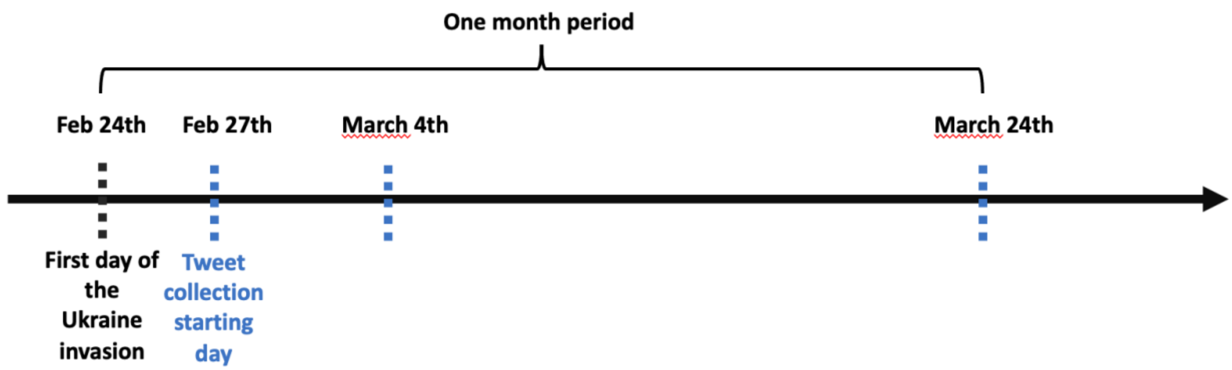



*Figure 1. Dataset timeline*

The dataset utilized in this exploration was compiled by Haq, Ehsan-Ul, et al. whose mission was to collect data in a time of crisis to allow for large scale analysis into public sentiment and trends in general populous response (Haq, Ehsan-Ul, et al, 2022). As is common with OSN data collection, the team compiling the Russo-Ukrainian conflict data used a Twitter Streaming API to continuously collect data directly from twitter in real time. A list of keywords were established to detect tweets associated with the conflict. The keywords evolved as the crisis further developed, but at the time of the release of the report, the keywords included: russia, ukraine, putin, zelensky, kiev, and kyiv (Haq, Ehsan-Ul, et al, 2022).

| | Unnamed: 0 | location | username | followers | totaltweets | retweetcount | text | hashtags | language | clean |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | NaN | CuttingEdgeDino | 205.0 | 7624.0 | 0.0 | @SenFeinstein you and your #oligarch cronies (... | [{'text': 'oligarch', 'indices': [27, 36]}, {'... | en | senfeinstein oligarch croni r guilti war crime... |
| 1 | 3 | NaN | ShaunaO2011 | 266.0 | 18115.0 | 1.0 | Their bloods are on the Russian hands.\nMilita... | [{'text': 'Mariupol', 'indices': [77, 86]}, {'... | en | blood russian hand militari hospit mariupol uk... |
| 2 | 4 | Manchester, England | thewelshtudor | 12.0 | 809.0 | 265.0 | 9:32 pm in #Ukraine. \n\nMy colleague from Sum... | [{'text': 'Ukraine', 'indices': [30, 38]}] | en | pm ukrain colleagu sumi name daryna left heart... |
| 3 | 6 | NaN | ArabNewsjp | 25079.0 | 38653.0 | 0.0 | #ICYMI: Loud applause resounded in the #UN Gen... | [{'text': 'ICYMI', 'indices': [0, 6]}, {'text'... | en | icymi loud applaus resound un gener assembl ha... |
| 4 | 8 | Brooklyn, NY | paultaylor__ | 367.0 | 326.0 | 13.0 | I'm donating 3 x signed A3 prints for #PhotoPr... | [{'text': 'PhotoPrintDay', 'indices': [57, 71]}] | en | donat x sign print photoprintday friday th mar... |

*Figure 2. Example of dataset after cleaning 'text'*

Above is an example of what is included in the dataset after it had been cleaned by members of our team. As can be seen, the data included information on the Twitter user like tweet and follower counts, the content of the tweet, what hashtags were included, and the language of the tweet, which was helpful in our data selection. As mentioned, we are looking at data from 3 days spaced throughout the ongoing crisis to gauge changing sentiment. On average, 200k tweets were scraped per day for analysis. As a whole, more than 900,000 users have contributed to the discourse around this crisis on twitter, amassing around 500,000 unique tweets. The most popular hashtags include clear identifiers of the topic, like #Russia or #UkraineRussiaWar, and popular mentions include @NATO, @POTUS, and @KyivIndependent (Haq, Ehsan-Ul, et al, 2022).

## 2.2 Preliminary Analysis about the Dataset

The bar chart below shows how the number of tweets changed over this month period. As we can see the number of tweets fluctuated a lot depending on the hot topics of each day during the Ukrainian invasion. The yellow columns represent the days that we analyzed in our project. Each day, the number of tweets were between 200,000 and 350,000 tweets.  It is important to note that on day 2, we had 700,000 tweets but after further analyzes our team realized most of the tweets were coming from the same ID and had the exact same messages. Therefore decided to remove these suspicious tweets, that is how we ended up with close to 350,000  tweets for Day 2.

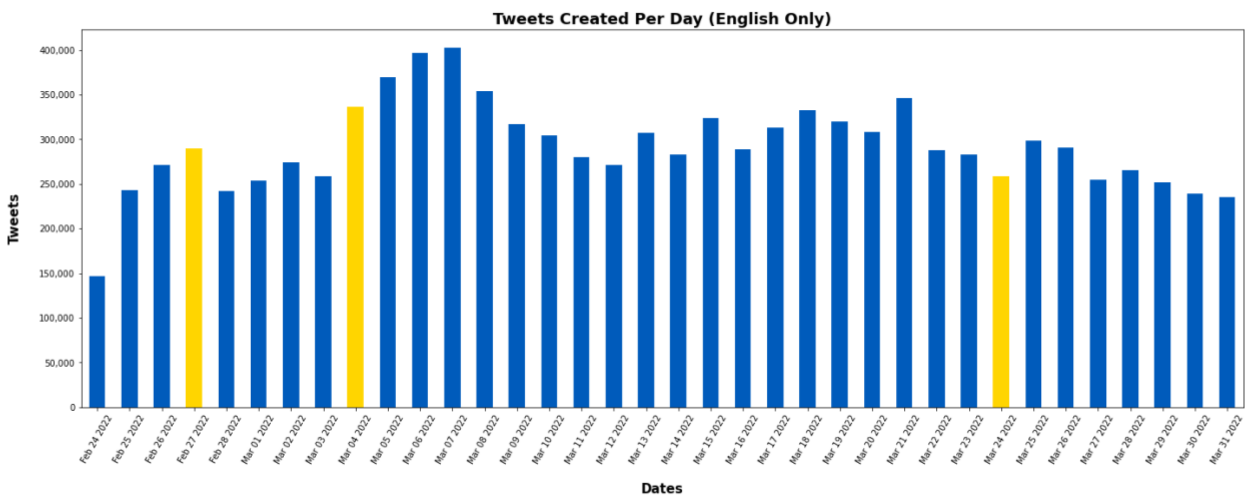
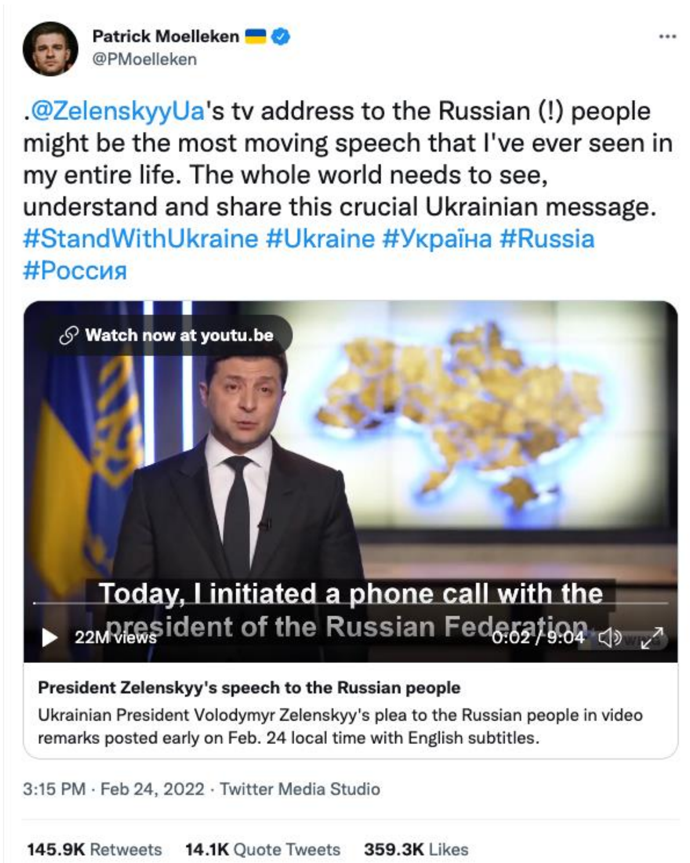



*Figure 3. The number of English tweets created per day*

Another interesting observation that we had in the preliminary analyses is the most retweeted posts throughout this time period. The tweet that you can see on the left was retweeted the most on every single day. This is the post where Zelensky ask the Russian people to stand against the Russian autocratic behavior and stand with the Ukrainian people during this crisis.




## 3. Methodology

### 3.1 Preprocessing and cleaning the dataset

Before we started the sentimental analysis we cleaned the body of the tweets to be in an appropriate form. Most of the tweets in our data set were in English however we found some tweets in other languages as well. Therefore, the first step of the cleaning part was to filter our data set to only English ones. We did that by filtering the language column in pandas. After that we made sure we only had English tweets before we started the process of cleaning corpuses.

The text cleaning process is usually very similar, starting with removing non-English characters, such as symbols. The next step is to remove every stop word. This is important because words like a, an or the can influence the result of word clouds  or sentimental analysis. Most of the times these words are only there for grammatical reasons therefore removing them is necessary. Removing stop words can also help to fasten up the running time of each model.

The next step is to stem the words in the text of the tweet. Stemming is the process of removing part of the word so only the root of the word means. In other ways, it is the process to make the word into its dictionary form. This step is crucial for word cloud and sentimental analyses. Finally, we ended the pre-processing part with tokenizing each corpus. Meaning we created a list of words out of the stem corpuses.

**3.2 Sentiment Analysis**

Using nltk package's Sentiment Intensity Analyzer function, we calculated the compound score for each of the tweets. We classify all tweets into three categories based on their score. We regarded a positive compound score as positive sentiment tweets; negative compound score as negative sentiment tweets; a score of 0 as neutral sentiments. The method for finding the trends of sentiment score for three days is using Kernel Density Estimation. Kernel Density Estimation is a technique that helps us to create a smooth estimation of the distribution of a given set of data. In this study, we tried to see what the distributions of the sentiment score look like for Ukraine and the United States. Using Kernel Density Estimation to create the curve of the distribution can better help us to analyze the distribution than when just using a histogram. And at last, we use TextBlob to find the subjectivity of all tweets in the dataset. Subjective texts mean the texts contain more personal opinions, and objective texts mean the texts contain more factual information. The TextBlob subjectivity score lies between 0 and 1, and the higher the score is, the more subjective the text will be. Therefore, we classify all tweets into two categories; Subjective Tweets with a score greater than 0.5, and Objective Tweets with a subjectivity score less than 0.5.

## 4. Results

### 4.1 Word Clouds


*Figure 4. Word Cloud for February 27th*

The first word cloud is made for february 27th and it showed that initial tweets were more of a shock response and people were trying to understand what is going on. And the tweets had no definitive objectivity or subjectivity; they were stating what was happening. You can see words such as ukraine, war, russia and putin.


*Figure 5. Word Cloud for March 4th*

The second word cloud was made for March 4th and it had the most objective tweets. They were calling for humanitarian assistance and people were thinking that there was a humanitarian crisis going on. As you can see there are words in the cloud such as innocent, defend, stand with ukraine, assistance, needs, provide.


*Figure 6. Word Cloud for March 24th*

The last word cloud was made for March 24th and the tweets were mostly in the middle ground. The tweets were mostly siding with the ukraine and you can see words such as nato, mariupol, stand with ukraine.

### 4.2 Hashtags

After visualizing the most common words, we delved into hashtags as the next step. Hashtag is entrenched in the culture of Twitter. By its original design, hashtag is used to categorize content and allow users to search for all messages within the same topic.

By looking at the 20 most frequent hashtags each day, we found that people used hashtags for two purposes: (1) to express personal emotions or (2) reporting trending news topics. Therefore, we classified these hashtags into two categories and examined the composition each day. Besides, we noticed a lot of city names as hashtags, which tracks war updates and reflect the current regions in conflict. On each day, one new city name emerges, indicating the most recent attack locations.

Below is the summary of top 20 hashtags on each day, including hashtag terms, frequency, and category.

On day 1, most of the hashtags are general terms such as “ukraine”, “russia”, and “ukraineunderattack”. These general topics indicate the emergence of the war. The capital city of Ukraine “Kyiv”/ “Kiev” is also very frequent, reflecting public concerns of invasion in the major city. By contrast, there was only one sentimental hashtag “standwithukraine.”

On day 2, hashtags started to get more sentimental. There were 5 strong emotional hashtags such as “stopputin” and “stoprussia.” The usage of the strong verb “stop” indicates a call for action. Noticeably, a new city name “Zaporizhzhia” reflects the invasion progress of Russia.

As the war progressed to day 3, hashtags carried more sided messages. Most of the hashtags support Ukraine, with 6 out of 20 containing personal emotions. New hashtags was created such as “weareallukrainians”, “freeukraine.” At the same time, new city name “mariupol” becomes popular because Mariupol was surrounded and bombarded by Russian artillery one week before the Day 3 (March 24).

**20 Most Frequent Hashtags by Day**

| Day 1 | | | | Day 2 | | | | Day 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ranking | Hashtag | Frequency | Category | Ranking | Hashtag | Frequency | Category | Ranking | Hashtag | Frequency | Category |
| 1 | ukraine | 95576 | fact | 1 | ukraine | 112935 | fact | 1 | ukraine | 83030 | fact |
| 2 | russia | 43777 | fact | 2 | putin | 86094 | fact | 2 | russia | 33690 | fact |
| 3 | putin | 20824 | fact | 3 | safeairliftukraine | 64038 | emotional | 3 | mariupol | 20188 | fact (city name) |
| 4 | ukrainerussiawar | 14068 | fact | 4 | stopputin | 63776 | emotional | 4 | nato | 16636 | fact |
| 5 | russian | 13549 | fact | 5 | russia | 50529 | fact | 5 | putin | 15913 | fact |
| 6 | kyiv | 13518 | fact (city name) | 6 | ukrainerussianwar | 16511 | fact | 6 | standwithukraine | 15253 | emotional |
| 7 | anonymous | 11379 | fact | 7 | standwithukraine | 13715 | emotional | 7 | russian | 12671 | fact |
| 8 | ukraineunderattack | 8571 | emotional | 8 | russian | 12011 | fact | 8 | anonymous | 9690 | fact |
| 9 | ukrainewar | 8275 | fact | 9 | zaporizhzhia | 11043 | fact (city name) | 9 | kyiv | 8595 | fact (city name) |
| 10 | russiaukrainewar | 7064 | fact | 10 | russianukrainianwar | 10569 | fact | 10 | slavaukraini | 7980 | emotional |
| 11 | nato | 5271 | fact | 11 | nato | 8949 | fact | 11 | zelenskyy | 7668 | fact |
| 12 | standwithukraine | 5228 | fact | 12 | kyiv | 8444 | fact (city name) | 12 | ukrainewar | 6788 | fact |
| 13 | russianarmy | 4982 | fact | 13 | stopputinnow | 8069 | emotional | 13 | ukrainian | 5132 | fact |
| 14 | ukrainerussia | 4478 | fact | 14 | ukrainerussiawar | 6042 | fact | 14 | stoprussia | 4650 | emotional |
| 15 | ukrainian | 3580 | fact | 15 | russiaukraine | 5268 | fact | 15 | war | 4537 | fact |
| 16 | breaking | 3515 | fact | 16 | war | 3818 | fact | 16 | ukrainerussiawar | 4228 | fact |
| 17 | zelensky | 3460 | fact | 17 | ukrainian | 3606 | fact | 17 | freeukraine | 3583 | emotional |
| 18 | kharkiv | 2615 | fact | 18 | stoprussia | 3514 | emotional | 18 | stopputinswar | 3051 | emotional |
| 19 | russiaukraineconflict | 2600 | fact | 19 | zelensky | 3416 | fact | 19 | weareallukrainians | 3034 | emotional |
| 20 | kiev | 2567 | fact (city name) | 20 | ukriane | 3371 | fact | 20 | natosummit | 2890 | fact |

*Figure 7. Top 20 Frequent Hashtags per Day*

## 4.3 Sentiment Analysis

### (1) Positive vs. Negative

After seeing the changes in hashtag contents overtime, we confirmed that the tweets' sentiment evolved as the war progressed. Therefore, we continued to develop more rigorous statistics to capture even the minute changes in tweets sentiment.

To further confirm the validity of our classification approach, we extracted examples of tweets in each category, as the screenshots below. With a compound score of 0.9747, the first tweet is from a ukrainian, who said many “thank you” and expressed grateful and positive sentiments.


*Figure 8. Example of Positive Tweet*

The second tweet is classified as negative, with a -0.9913 compound score. The extreme negativity of his sentiment is a representation of his impolite and disrespectful attitude.

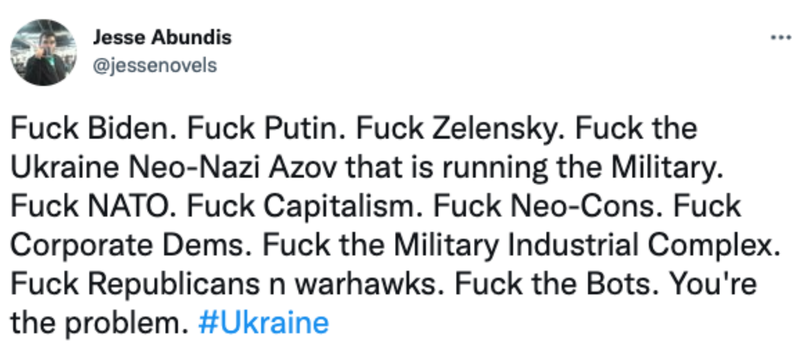



*Figure 9. Example of Negative Tweet*

The last tweet is an example of neutral tweets. Similar to lot of other neutral tweets, this short news report is conveyed in a neutral and objective manner.


*Figure 10. Example of Neutral Tweet*

As a result of our validity test, our sentimental classification method aligns with the actual sentiment expressed in these example tweets.

As the next step, we examined the composition of sentiment on each of three days and visualized the result using pie charts. The pie charts represent a common trend over three days, with the negative emotions in the largest percentage.

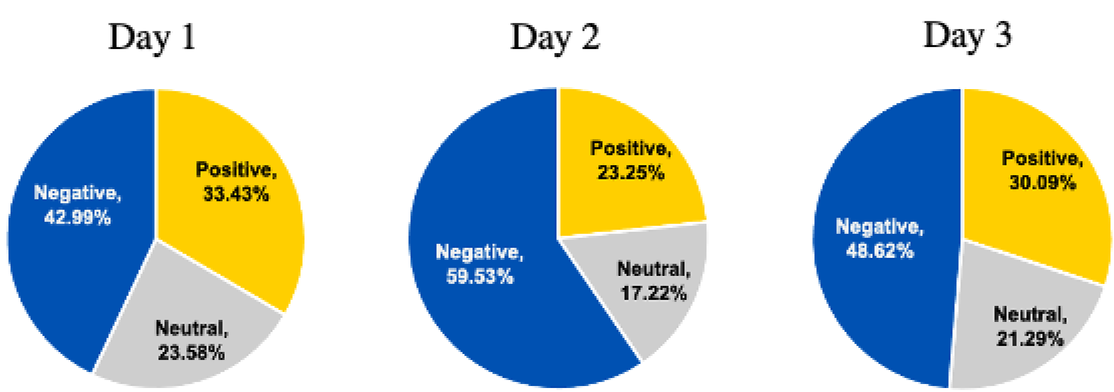



*Figure 11. Pie Chart of Sentiment Compound Score per day*

**Statistics of Sentiment Compound Score (Positive vs. Negative)**

| | day1 | day2 | day3 |
|---|---|---|---|
| count | 227513 | 336614 | 258060 |
| mean | -0.076951 | -0.263531 | -0.147206 |
| std | 0.470019 | 0.51784 | 0.512451 |
| min | -0.9967 | -0.9954 | -0.9894 |
| 25% | -0.4767 | -0.8074 | -0.5994 |
| 50% | 0 | -0.34 | 0 |
| 75% | 0.296 | 0 | 0.2732 |
| max | 0.9826 | 0.9774 | 0.984 |

*Figure 12. Statistics of Sentiment Compound Score*

Among three days, day 2 has the most negative tone, with the lowest mean of -0.2635 and median of -0.34 compound score. Day 1 and Day 3 are more similar in score statistics and category percentages.

Lastly, we also compared the length between positive and negative tweets. Length is measured by the number of words in a tweet before removing stop words and stemming. Figure 13 and 14 are density distributions of tweet length of positive and negative tweets respectively. Positive tweet has a range between 2 to 93 words, while negative tweet has a slightly larger range between 2 to 97 words.

The length distribution of positive tweets has two peaks at 19 and 24 words. By contrast, the distribution of negative tweets have a longer right tail, indicating the existence of long and negative tweet. Besides, there is only one noticable peak at 19 words in negative tweet, indicating a greater concentration in the shorter end.

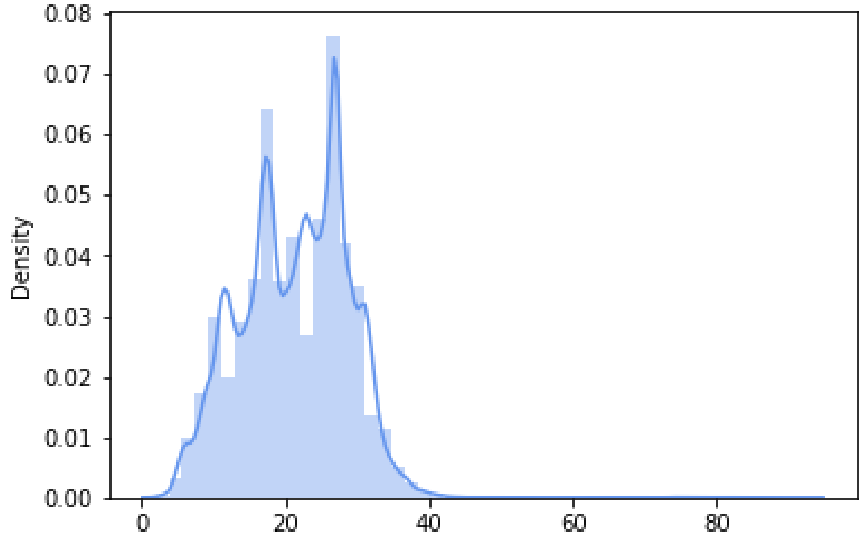



*Figure 13. The distribution of positive tweets*

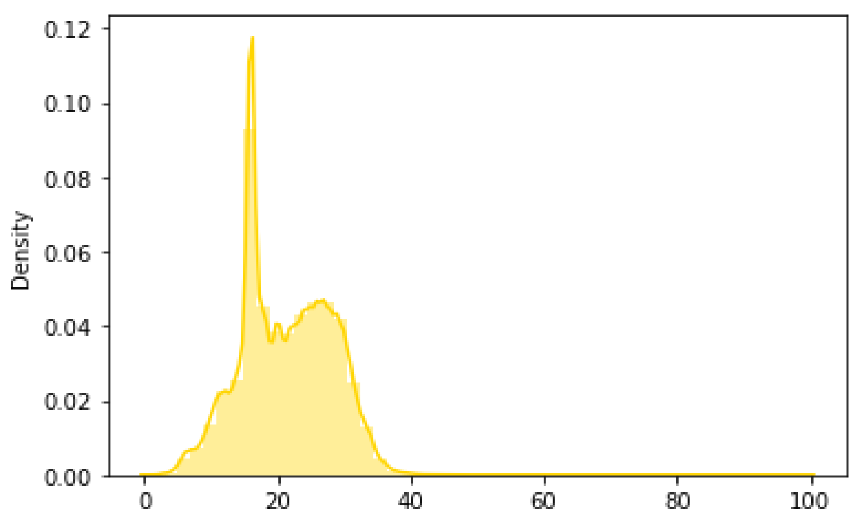



*Figure 14. The distribution of negative tweets*

Moreover, to further understand the trends of sentiment score, the estimated distribution of the score for Ukraine and the U.S. were plotted. The distributions of the sentiment score for Ukraine were shown below. The blue line is for Ukraine, and the red line is for other countries.

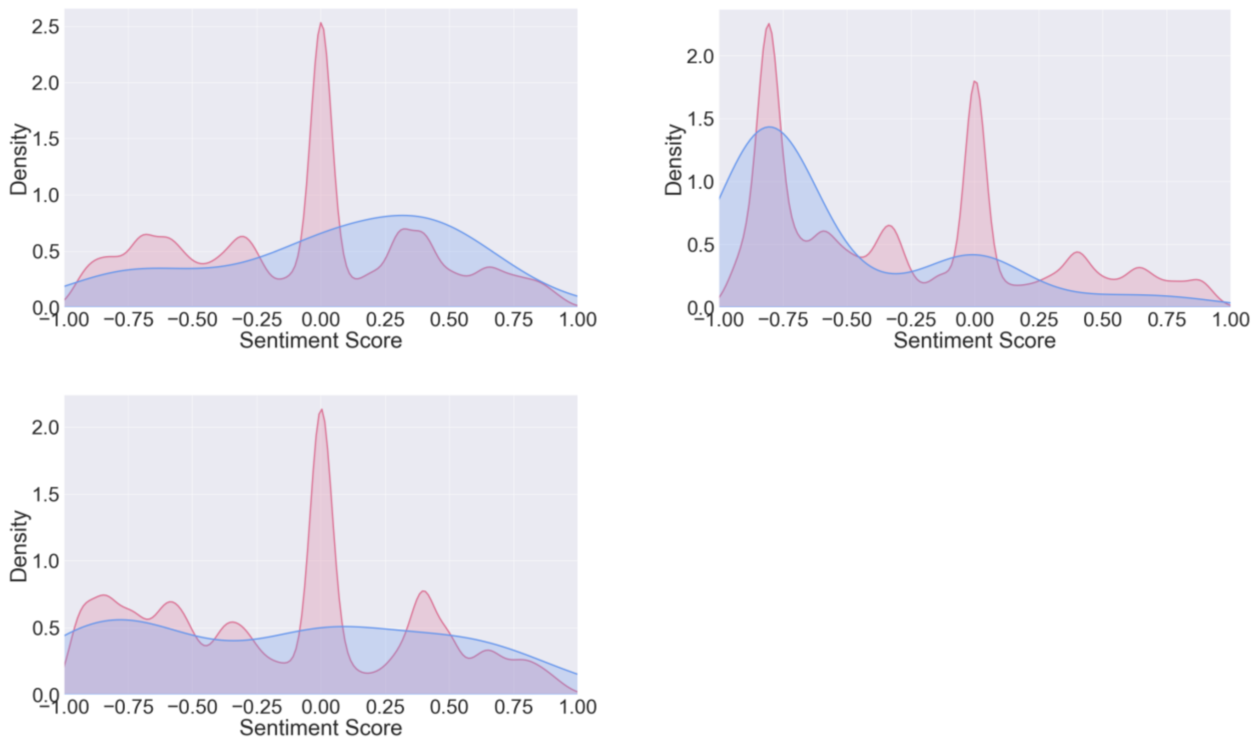



*Figure 15. The distribution of sentiment score for Ukraine and Other countries for 3 days*

The above graph show the distribution of the sentiment score for Ukraine and other countries except for Ukraine. The first graph is the distribution on February 27th. It is not hard to tell that the distribution for Ukraine is skewed to the left, and the peak of the distribution is at sentiment score 0.3 which is positive sentiment. Hence, more proportion of positive tweets are sent on Feb 27th for Ukraine. The distribution for other countries looks symmetric, and the peak is at sentiment score 0 which is neutral sentiment. Thus, the density for positive tweets and for negative tweets for other countries are about the same due to its symmetric distribution.

The second graph is the distribution on March 4th. It is obvious that the distribution for Ukraine is right-skewed, and its peak is around the sentiment score -0.8. Meanwhile, for each sentiment score over 0, they all only have a little density. Thus, most of the tweets have negative and mostly extremely negative sentiment for Ukraine. The distribution for other countries is a multimodal distribution, which has two peaks at sentiment score 0 and score -0.8. Thus, Ukraine and other countries all have lots of negative tweets on Marth 4th.

The distribution on March 24th is the third graph here. The distribution for Ukraine is more like a uniform distribution. So, for almost every sentiment score, the number of tweets sent from Ukraine is about the same except for sentiment score over 0.75. And again, the distribution for other countries looks symmetric, and the peak is at sentiment score 0 which is neutral sentiment. Hence, the proportion for positive tweets and for negative tweets for other countries are about the same due to its symmetric distribution .Overall, there are some special patterns found for sentiment score in Ukraine for all days, and they all looks different from the other countries' distribution.

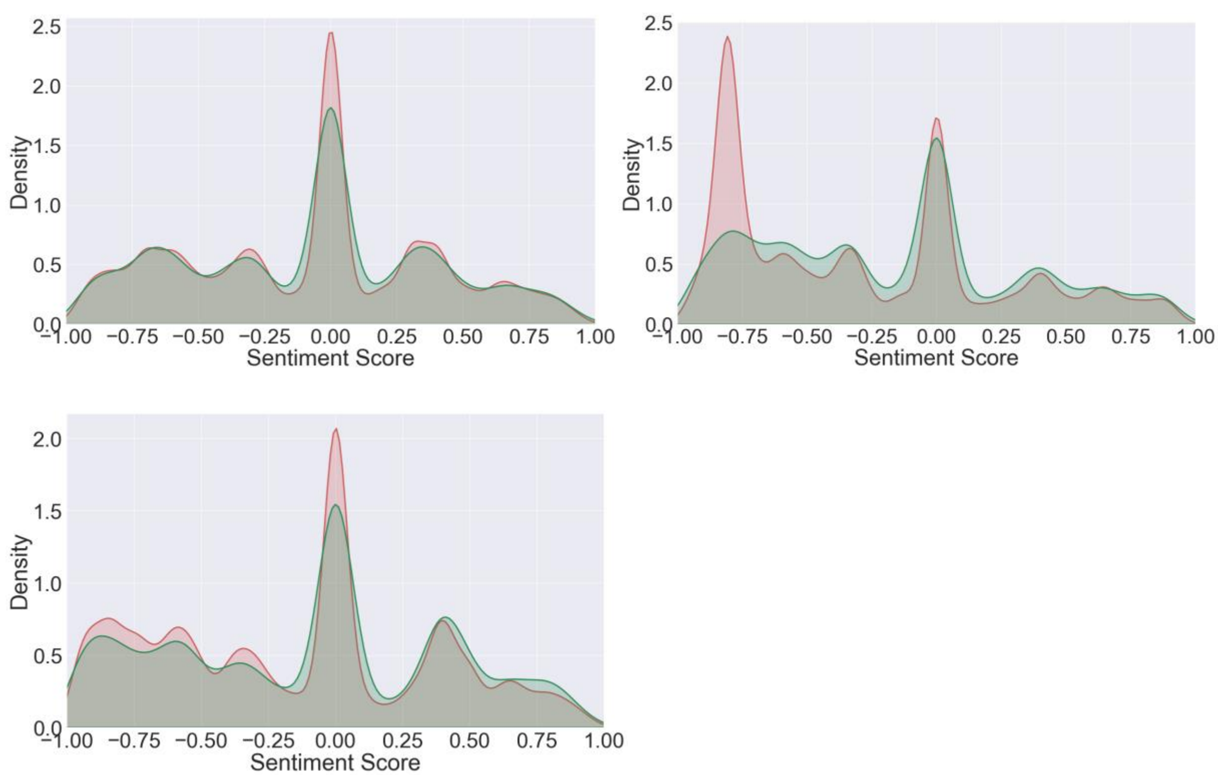



*Figure 16. The distribution of sentiment score for the U.S. and Other countries for 3 days*

The above three graphs are the distribution of sentiment score between the U.S. and other countries. Clearly, the distribution of the U.S. looks similar to the other countries except for the second plot, the distribution on Marth $4^{th}$. Noticed that for other countries, the distribution is a multimodal distribution, which has two peaks at sentiment score 0 and score -0.8. Thus, other countries have lots of negative tweets on Marth $4^{th}$. And for the U.S., the peak of its distribution is at sentiment score 0, and the density for all negative sentiment score is higher than the density for all positive sentiment score. Thus, tweets sent from the U.S. still have more negative sentiment than positive sentiment. Overall, the distributions for the U.S. are all symmetric, but they all share similar sentiment distribution with all other countries.

**(2) Subjective vs. Objective**

At last, we calculate the subjectivity score using TextBlob. The basic statistics of subjective score and the visualization of the result are shown below.

| | **Day 1** | **Day 2** | **Day 3** |
|---|---|---|---|
| count | 227513 | 336614 | 258060 |
| mean | 0.203491 | 0.167642 | 0.206486 |
| std | 0.266051 | 0.249922 | 0.258618 |
| min | 0 | 0 | 0 |
| 25% | 0 | 0 | 0 |
| 50% | 0.1 | 0 | 0.1 |
| 75% | 0.4 | 0.3 | 0.383333 |
| max | 1 | 1 | 1 |

*Figure 17. Statistics of TextBlob Subjectivity Score*

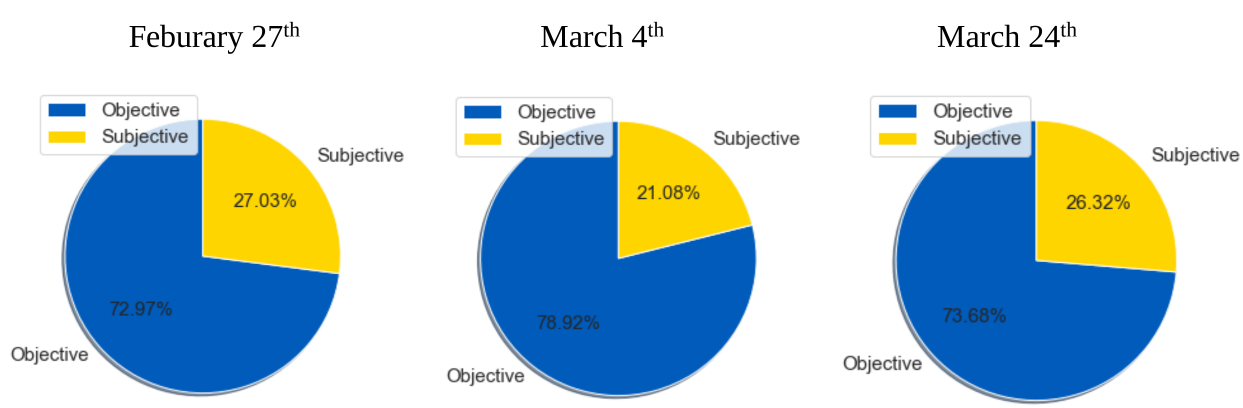



*Figure 18. Pie Chart of Subjectivity Score per day*

The mean value for three days are all smaller than 0.5, and March 4th has the lowest mean score 0.167642. Thus, on average, tweets are more objective for all three days. The above pie charts help us to see the proportion of subjective and objective tweets among three days. Based on the pie charts above, over 70% of the tweets are objective among three days. March 4th has the highest proportion of objective tweets while Feburary 27th has the highest proportion of subjective tweets.

At last, we also compare the length between subjective and objective tweets. The figures below show the distribution of the length for subjective and objective tweets among all three days. The peak of the distribution of subjective tweets is at 28 to 30 words while the peak of the distribution of objective tweets is at 18 to 20 words. Therefore, tweets tend to be longer when they are more subjective.

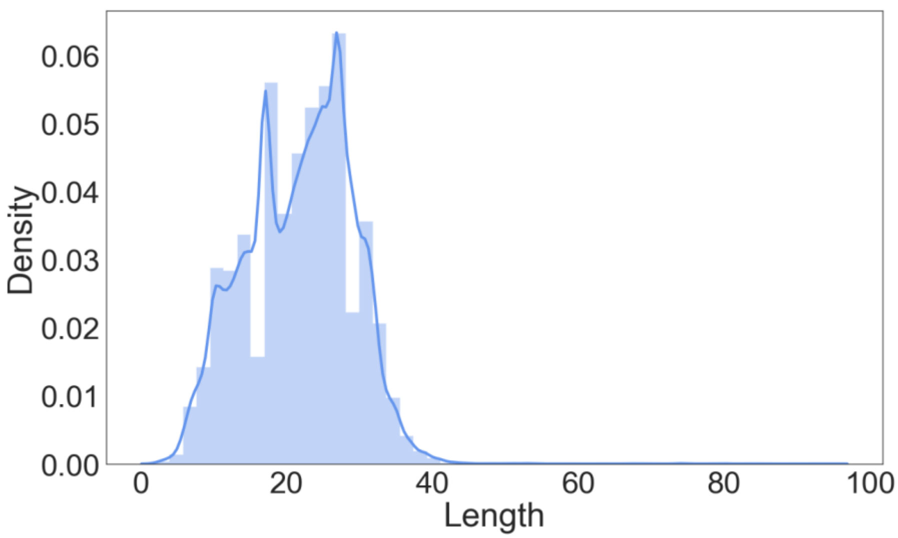



*Figure 19. The distribution of Subjective tweets*

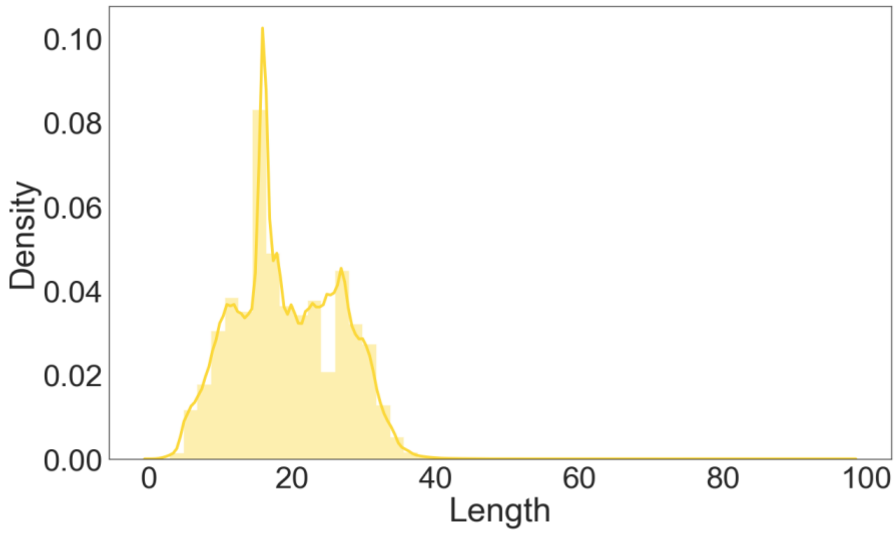



*Figure 20. The distribution of Objective tweets*

## 5. Conclusion

In our research we explored how the tweets related to the Ukraine invasion. We wanted to see how the topics and the sentiment of the tweet changed over a month. For these analyses we used a Twitter dataset that was collected by college students with a Twitter API. We examined the first day when they started collecting these tweets, and the day after a week and the day after a month the Ukrainian war started. In this part we review our findings and give our assumptions about what exactly happened during this month.

During day 1, we can see how the sentiment of the analyses represented an initial shock but also a gratitude. Lot of tweets were about being grateful that the word stands with Ukraine. This is the day when we had the highest number of subjective tweets and the highest positive number of tweets as well. These results resonate with the shock and the gratefulness. However, in all three days the objective tweets were prominent, day 2 had the highest amount of objective tweets and the highest amount of negative tweets. It shows that people started to get involved into the war and wanted to know more details about what exactly is happening in Ukraine. Tweets that call for humanitarian actions started appearing which as well also reflects the involvement of people. We found that on day 3, people found a middle ground. By this day people took sides in this war and majority of them realize that the good side is the Ukrainian side. On the other hand, we can see that people lost interest in the details of the war. It looks like tweets stood back one step further from the war and rather than talking about the details of this crisis, they were rather focusing on sympathizing with the Ukrainian people.

The tweets that we analyzed showed how the trend changed during this month. At the beginning, people were tweeting subjective and hopeful tweets which changed to more factual reports and negative tweets by day 2. After a month of war, we saw how people lost their interest in the details of the war and they settled in a middle ground between day 1 and day 2.

# References

Haq, Ehsan-Ul, et al. (2022). Twitter Dataset for 2022 Russo-Ukrainian Crisis.

Sacco, V. & Bossio, D. (2015). Using social media in the news reportage of War & Conflict: Opportunities and Challenges. the Journal of Media Innovations. 2. 59-76.