The U.S. 2020 election and Twitter analysis

Alex Zhou

UMBC Data Science Department

DATA606 Capstone in Data Science

Professor Barber

12/15/2022

# Abstract

The 2020 US Presidential election was a highly anticipated and closely contested race between the incumbent President Donald Trump and the challenger, former Vice President Joe Biden. As with most modern elections, social media played a significant role in the campaign, with both candidates using Twitter extensively to reach out to voters and share their messages.

One way that data scientists and researchers can analyze the impact of Twitter on the election is by looking at the number of votes each candidate received in relation to their Twitter activity. For example, if a candidate had a large number of Twitter followers and generated a lot of engagement on the platform, one might expect that they would receive a larger number of votes on election day. Another approach to analyzing Twitter data from the 2020 election is through sentiment analysis, which involves using natural language processing techniques to determine the overall sentiment of a piece of text. By applying sentiment analysis to tweets from the candidates and their supporters, researchers can gain insight into how positively or negatively voters felt about each candidate and their campaign. Overall, the use of Twitter data in the 2020 US Presidential election highlights the increasing importance of data science and digital technologies in the democratic process. By leveraging the vast amounts of data generated on social media platforms, researchers can gain a better understanding of voter sentiment and the impact of online campaigning on election outcomes.

Keywords: US 2020 election, Twitter data, Biden, Trump, votes, democratization, data science, sentiment analysis

# Introduction

The United States Presidential election is considered one of the most important elections in the world. This is because the President of the United States is not only the leader of the country, but also a global leader and the head of the world's largest economy. The decisions made by the President have a significant impact not only on the American people, but on people all over the world. One of the main reasons the US Presidential election is considered so important is the role of the United States in the global political and economic systems. The US is a superpower, with a strong military and significant influence on international affairs. As a result, the policies and decisions of the President can have far-reaching consequences for the global community.

Another reason the US Presidential election is important is the impact it has on domestic policy. The President has the power to shape the direction of the country through executive orders, appointments to key positions, and the creation of laws and policies. The President also has the power to veto legislation passed by Congress, making the office a powerful force in shaping the future of the country.

Furthermore, the US Presidential election is seen as a key indicator of the state of democracy in the country. The peaceful transfer of power between elected officials is a cornerstone of democratic societies, and the US Presidential election is a high-profile example of this process in action. The election also serves as a forum for the American people to express their views and preferences, and to hold their elected officials accountable for their actions.

In short, the US Presidential election is a crucial event that has a significant impact on the United States and the world. The outcome of the election can shape the future of the country and the global community, making it a highly important event for citizens and leaders alike.

Twitter is a social media platform that allows users to send and read short messages known as "tweets." These tweets can be up to 280 characters in length and can include text, photos, videos, and links. Twitter is used by millions of people around the world, including individuals, organizations, and businesses.

For many people, Twitter is a way to stay connected with friends and family, share their thoughts and experiences, and stay up-to-date with the latest news and trends. It allows people to have conversations, express their opinions, and engage with others who have similar interests.

Twitter is also a valuable tool for journalists and news organizations, as it allows them to quickly disseminate information and gather information from sources in real-time. This can be especially useful during breaking news events, when traditional news outlets may not have all the information.

Additionally, Twitter is a powerful platform for activism and social change. It allows individuals and organizations to raise awareness about important issues and mobilize support for causes they care about. It has also been used to organize and promote political movements and protests.

Overall, Twitter has many different meanings and uses for different people. For some, it is a way to stay connected and share information, while for others it is a tool for activism and social change. Regardless of how it is used, Twitter has become a key part of the way people communicate and interact in the digital age.

Data analysis and sentiment analysis are important tools for understanding and interpreting information. Data analysis involves using statistical and computational techniques to identify patterns and trends in data. This can help researchers and organizations make more informed decisions by providing a clearer picture of the data they are working with.

Sentiment analysis, also known as opinion mining, involves using natural language processing techniques to determine the overall sentiment of a piece of text. This can be useful for understanding how people feel about a particular topic or issue, and can be applied to a wide range of fields, including marketing, customer service, and politics.

Overall, data analysis and sentiment analysis are important because they allow researchers and organizations to gain deeper insights from data and text. They can help identify trends and patterns, and provide a more complete understanding of complex issues. As a result, these tools are increasingly being used in a variety of fields to inform decision-making and drive change.

Based on the above points, this thesis will conduct EDA and sentiment analysis on the Twitter data of the US 2020 election to observe the importance and authenticity of the election, the meaning of Twitter to the public, and the sentiment analysis of the tweets to verify that Biden's victory was due to more support.

# Background, Methods and Materials

# Background

This research was motivated primarily by papers, "The 2017 UK general election: a Twitter analysis", "Location-based sentiment analysis and visualization of Twitter election data". The research in this paper is mainly based on EDA and sentiment analysis, using a relatively basic python library and two sentiment analysis tools. Exploratory Data Analysis (EDA) is a method of exploring and discovering patterns in data in order to better understand a dataset and provide a foundation for further analysis and model building. EDA is an important part of the data analysis process because it helps to identify interesting trends and patterns in the dataset and to uncover potential issues with data quality. Overall, EDA is very important for data analysis because it provides a foundation for further research and modeling. Sentiment analysis is a method of analyzing the emotional content of text, often with the goal of determining the overall emotional tone of the text. This can be used to identify the sentiment of a piece of text as positive, negative, or neutral, and can be applied to a wide range of applications, such as social media monitoring and customer feedback analysis. Sentiment analysis is often performed using natural language processing and machine learning techniques, and can be useful for gaining insights into the opinions and emotions of large groups of people. Data source kaggle. Regarding the description of the data, the data comes from collected tweets using the Twitter API statuses\_lookup and snsscrape for keywords, the time frame will eventually cover 15.10.2020 to the end of November 6.

# Data and Results

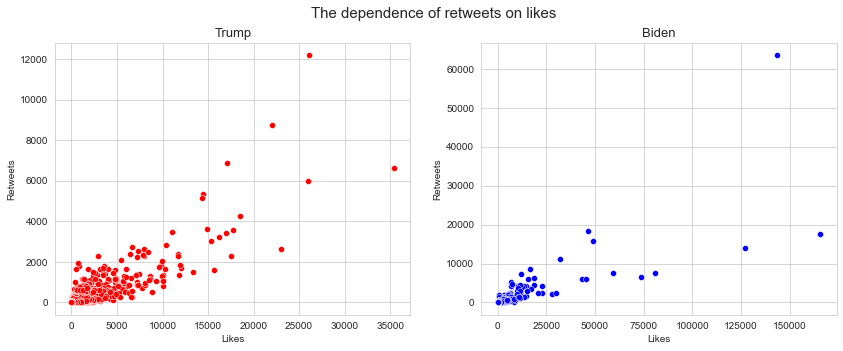
The following paragraph of the article will explain the results obtained from EDA and sentiment analysis, and the data will be described in the context of the actual situation. The main focus is on the number of likes, tweets, countries of interest, the number of users analyzed, and the sentiment analysis.

## Number of likes and comments on tweets

On Twitter, a "like" is a way for users to show their appreciation for a particular tweet. When a user likes a tweet, they are adding their endorsement to that tweet and letting the author know that they found the content valuable, interesting, or funny. Liking a tweet is similar to "thumbing up" or "upvoting" a post on other social media platforms, and serves as a way for users to engage with the content on Twitter.

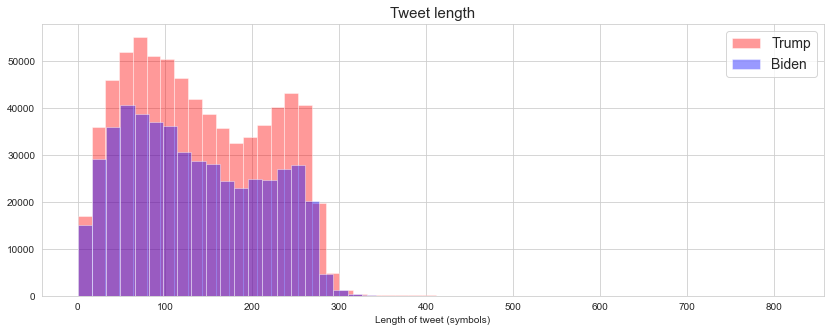
Liking a tweet has several effects. First, it increases the visibility of the tweet to other users, as it signals to the Twitter algorithm that the content is valuable and should be shown to more people. This can help the tweet to gain more traction and potentially go viral. Second, it allows the user who liked the tweet to save it to their own personal collection of liked tweets, which they can access at any time. This can be useful for keeping track of tweets that the user found particularly interesting or useful. Finally, liking a tweet can also serve as a way for users to show their support for the author and their content.

It's important to note that liking a tweet is different from retweeting it. When a user likes a tweet, they are indicating that they enjoyed the content of the tweet, but they are not sharing it with their own followers. Retweeting a tweet, on the other hand, shares the tweet with the user's own followers, which can help it to reach a wider audience. Both actions are valuable ways for users to engage with content on Twitter, and can help to promote interesting and valuable content on the platform.

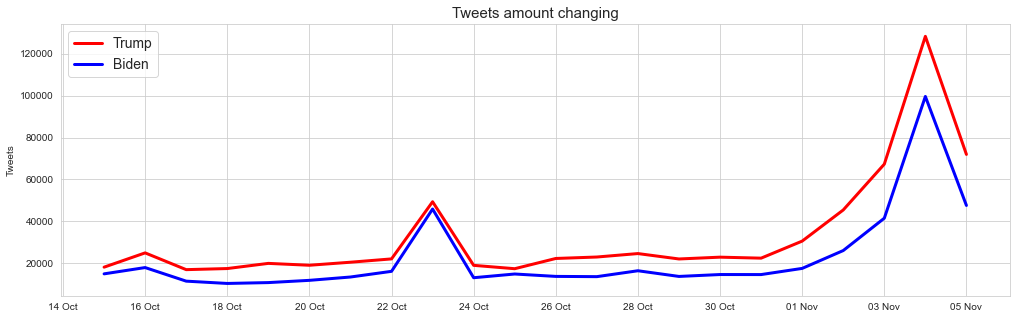


*Number of Like and retweets*

The number and length of tweets by a public figure can have a number of meanings. For example, a public figure who tweets frequently and in long form may be seen as being very engaged with their audience and willing to share their thoughts and opinions. This can help to foster a strong connection between the public figure and their followers, and can also help to generate buzz and attention for the public figure's work.

On the other hand, a public figure who tweets infrequently or in short form may be seen as being less engaged with their audience. This can lead to a weaker connection with followers and less attention for their work. Additionally, the content of a public figure's tweets can also be important. Tweets that are interesting, informative, or funny can help to engage followers and generate attention, while tweets that are dull, repetitive, or controversial can have the opposite effect.

Overall, the number and length of tweets by a public figure can be a useful indicator of their level of engagement with their audience, and can also provide insight into the content and tone of their tweets. However, it is important to consider a wide range of factors when evaluating the significance of a public figure's tweets, and not to rely solely on their quantity and length.



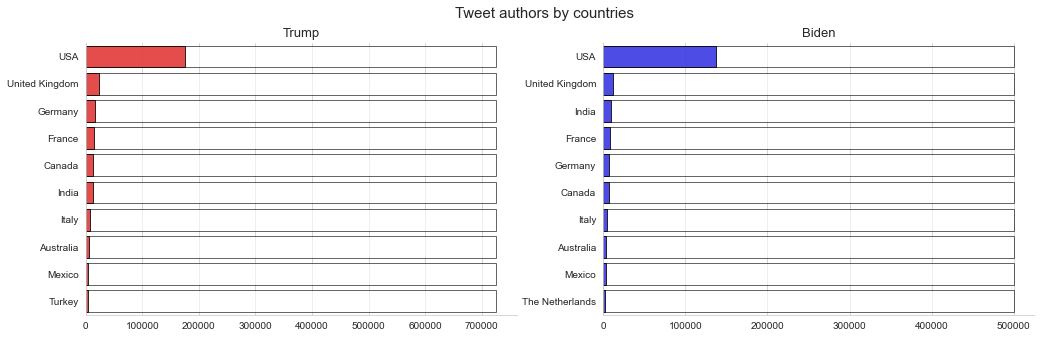
### How it affects other Country

The United States presidential election can have significant impacts on other countries around the world, particularly those that have strong political, economic, and military ties with the US. For example, the election results can affect the policies and decisions of countries like Germany and India, which are both major players on the global stage.

In the case of Germany, the US presidential election can have a major impact on its economy. Germany is one of the largest export economies in the world, and the US is one of its major trading partners. If the US were to implement protectionist trade policies, it could hurt German exports and put pressure on its economy. Additionally, if the US were to withdraw from international agreements and organizations, such as NATO or the World Trade Organization, it could also have negative implications for Germany's security and trade.

India, on the other hand, has a more complex relationship with the US. The two countries have a strong economic partnership, with the US being one of India's largest trading partners. However, there are also significant political and strategic differences between the two countries, particularly when it comes to issues like immigration and foreign policy. For example, if the US were to adopt a more isolationist stance, it could affect India's ability to access certain markets and resources. Additionally, if the US were to change its stance on issues like climate change or terrorism, it could also impact India's foreign policy decisions.

Overall, the US presidential election can have both positive and negative effects on other countries, depending on the policies and decisions of the incoming administration. It is important for countries like Germany and India to monitor the election closely and be prepared to adapt to any changes that may arise.



**Users of Twitter: number of likes and replies to comments**

Twitter users can influence the outcome of the United States presidential election in a number of ways. For example, they can use their accounts to spread information and opinions about the candidates and their policies, which can influence the views of other users and shape the public discourse around the election.

The likes and replies that users give to tweets can also have an impact on the election. When a user likes or replies to a tweet, it can help to promote the tweet to a wider audience and increase its visibility. This can be especially important during the election season, when there is a high level of interest and engagement around the candidates and their platforms.

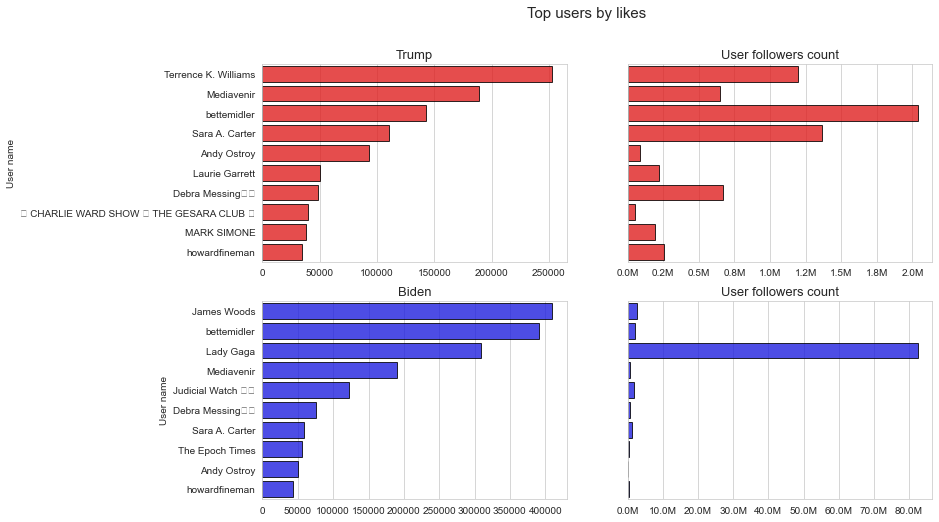
# 

Additionally, the likes and replies that users give to tweets can also signal to the candidates and their campaigns which issues and messages are resonating with the public. This can help the candidates to tailor their messaging and focus on the issues that are most important to their supporters.

The likes and replies that Twitter users give to tweets can help to shape the public discourse around the election and influence the candidates and their campaigns. By engaging with tweets and expressing their opinions, Twitter users can play a key role in the electoral process. In addition to likes and replies, Twitter users can also influence the election by retweeting tweets and sharing them with their own followers. This can help to further promote a particular message or perspective, and can also help to amplify the voices of individuals and organizations that may not have as large of a following on their own.

Furthermore, Twitter users can also participate in hashtags and Twitter chats related to the election. This can help to bring together users with similar interests and views, and can provide a platform for them to discuss and debate the issues that matter most to them. This can also help to raise awareness about certain candidates or issues, and can potentially sway the opinions of other users who are still undecided.

In summary, Twitter users have the potential to play a significant role in the United States presidential election by expressing their opinions, engaging with content, and participating in online discussions. By using their voices and their social media platforms, they can help to shape the public discourse and potentially influence the outcome of the election.



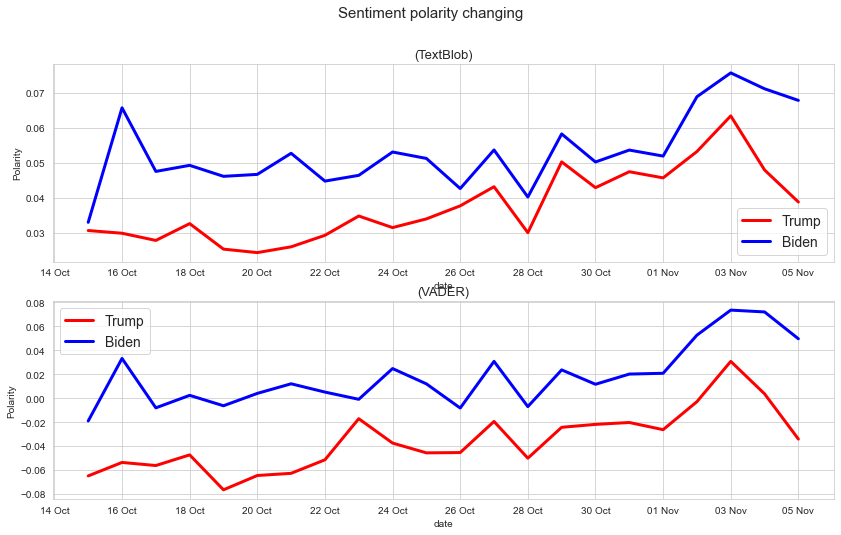
**Results and Analysis of Sentiment Analysis**

# Sentiment analysis is a method of analyzing and interpreting the emotions, attitudes, and opinions expressed in text data. It is often used in natural language processing and other fields to help understand and interpret large volumes of text data.

# The purpose of sentiment analysis is to provide a quantitative and objective measure of the overall sentiment expressed in a piece of text. This can help to identify the general tone of a text, as well as the specific emotions and opinions expressed within it.

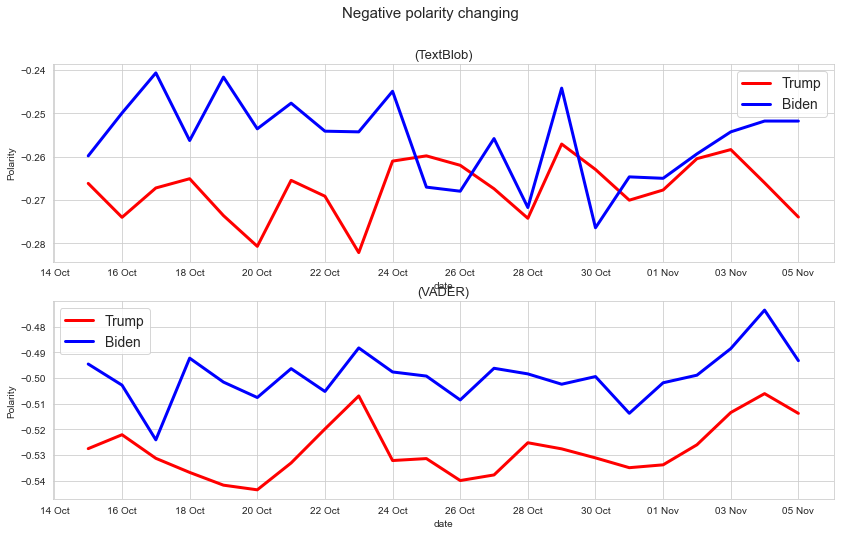
# To perform sentiment analysis, algorithms and models are used to analyze the words, phrases, and sentences in a piece of text and assign a sentiment score. This score can range from very positive to very negative, with a neutral score in the middle. The specific method and approach used to calculate the sentiment score will depend on the specific application and the goals of the analysis.

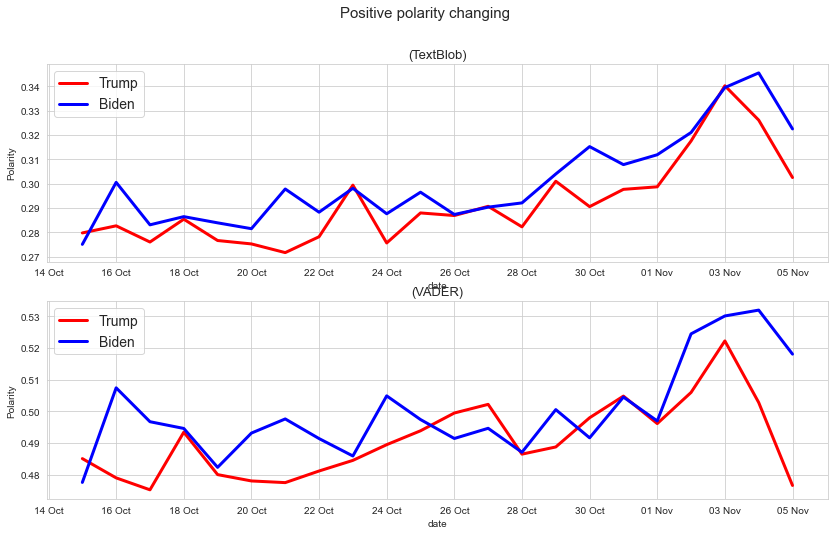
# One of the key advantages of sentiment analysis is that it can help to provide a more accurate and objective measure of the sentiment expressed in a piece of text. This can be particularly useful when analyzing large volumes of data, or when trying to identify subtle differences in the sentiment of different texts.Sentiment analysis is a valuable tool for interpreting and understanding the emotions and attitudes expressed in text data. By providing a quantitative measure of sentiment, it can help to improve the accuracy and objectivity of analyses and decision-making.



During the United States presidential election, people's emotional fluctuations can come from a variety of sources. One of the main sources of emotional fluctuation is the campaign itself, as people may become excited or energized by the speeches, debates, and other events that take place during the election season.

Additionally, people's emotional fluctuations can also be influenced by the media coverage of the election. The news and other media outlets play a crucial role in informing the public about the candidates and their policies, and the way that they present this information can have a significant impact on people's emotions. For example, if a news outlet presents a particularly negative or sensationalized story about a candidate, it can cause people to become more emotional and reactive.

Another source of emotional fluctuation during the election is the social media and online platforms that people use to discuss and engage with the election. These platforms can be highly influential in shaping people's opinions and emotions, as they provide a space for people to share their thoughts, opinions, and experiences with others. The discussions and interactions that take place on these platforms can be highly emotional, and can contribute to people's emotional fluctuations.

Overall, the emotional fluctuations that people experience during the United States presidential election can come from a variety of sources, including the campaign itself, the media, and online platforms. These sources can all play a role in shaping people's emotions and influencing the way that they engage with the election.

# Discussion

The United States presidential election and Twitter have a complex and interconnected relationship. On the one hand, Twitter can be a powerful tool for candidates and campaigns to reach and engage with voters. Candidates and their teams can use Twitter to share their messages, policies, and opinions with a large and diverse audience. This can be especially useful during the election season, when people are paying close attention to the candidates and their platforms.

Additionally, Twitter can also be a useful tool for journalists and other members of the media to cover the election. Journalists can use Twitter to stay up-to-date on the latest developments and to share their own reporting and analysis with their audience. This can help to provide a more comprehensive and nuanced picture of the election and the candidates.

On the other hand, Twitter can also be a source of controversy and misinformation during the election. The platform is known for its fast-paced and often chaotic nature, which can make it difficult to verify the accuracy and reliability of the information that is shared on it. This can lead to the spread of false or misleading information, which can have negative consequences for the election and the candidates.

The relationship between the United States presidential election and Twitter is complex and multifaceted. Twitter can be a useful tool for candidates and the media, but it can also be a source of controversy and misinformation.

The results of sentiment analysis cannot accurately predict the outcome of the United States presidential election. While sentiment analysis can provide valuable insights into the emotions, attitudes, and opinions expressed in text data, it is not a perfect or comprehensive measure of public opinion.

There are several reasons why the results of sentiment analysis cannot accurately predict the outcome of the election. First, sentiment analysis is based on a limited and specific dataset, which may not be representative of the overall population. This means that the results of sentiment analysis may not accurately reflect the views and opinions of all voters.

Second, sentiment analysis is based on algorithms and models, which can be subject to bias and other limitations. These algorithms and models may not be able to accurately capture the full range of emotions, attitudes, and opinions that are expressed in text data.

Third, sentiment analysis does not take into account other factors that can influence the outcome of an election, such as voter turnout, the electoral college system, and the specific issues and policies that are being discussed.

Overall, while sentiment analysis can provide valuable insights into the emotions and attitudes expressed in text data, it cannot accurately predict the outcome of the United States presidential election.

**Conclusion**

The above is the analysis of the U.S. election and Twitter data, although on the surface Biden won the sentiment analysis and the election, but we can not deny that there are too many factors affecting the U.S. election, these data is not controlled and regulated we do not know. We hope that these Twitter data are true to represent people's emotions.

# 

# 

# 

# References

* Lorien C. Abroms and R. Craig Lefebvre. 2009. Obama's wired campaign: Lessons for public health communication. *J. Health Commun.* 14, 5 (2009), 415–423.
* Sitaram Asur and Bernardo A. Huberman. 2010. Predicting the future with social media. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT’10)*, Vol. 1. IEEE, 492–499.
* Steven Bird, Ewan Klein, and Edward Loper. 2009. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O'Reilly Media, Inc.
* Javier Borondo, A. J. Morales, Juan Carlos Losada, and Rosa M. Benito. 2012. Characterizing and modeling an electoral campaign in the context of Twitter: 2011 Spanish presidential election as a case study. *Chaos Interdisc. J. Nonlin. Sci.* 22, 2 (2012), 023138.
* Lori Brainard and Mariglynn Edlins. 2015. Top 10 US municipal police departments and their social media usage. *Amer. Rev. Pub. Admin.* 45, 6 (2015), 728–745.
* Guido Caldarelli, Alessandro Chessa, Fabio Pammolli, Gabriele Pompa, Michelangelo Puliga, Massimo Riccaboni, and Gianni Riotta. 2014. A multi-level geographical study of Italian political elections from Twitter data. *PloS One* 9, 5 (2014), e95809.
* Nadya A. Calderon, Brian Fisher, Jeff Hemsley, Billy Ceskavich, Greg Jansen, Richard Marciano, and Victoria L. Lemieux. 2015. Mixed-initiative social media analytics at the World Bank: Observations of citizen sentiment in Twitter data to explore “trust” of political actors and state institutions and its relationship to social protest. In *Proceedings of the IEEE International Conference on Big Data (Big Data’15)*. IEEE, 1678–1687.
* Zhiyuan Cheng, James Caverlee, and Kyumin Lee. 2010. You are where you tweet: A content-based approach to geo-locating Twitter users. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. ACM, 759–768.
* Zi Chu, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. 2012. Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?*IEEE Trans. Depend. Sec. Comput.* 9, 6 (2012), 811–824.
* Elanor Colleoni, Alessandro Rozza, and Adam Arvidsson. 2014. Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *J. Commun.* 64, 2 (2014), 317–332.
* Brett Edkin. 2016. Study: Trump Benefited From ‘Overwhelmingly Negative’ Tone Of Election News Coverage. <https://www.forbes.com/sites/brettedkins/2016/12/13/trump-benefited-from-overwhelmingly-negative-tone-of-election-news-coverage-study-finds/#43ab79f93202>.
* Emilio Ferrara and Zeyao Yang. 2015. Quantifying the effect of sentiment on information diffusion in social media. *PeerJ Comput. Sci.* 1 (2015), e26.
* Daniel Gayo Avello, Panagiotis T. Metaxas, and Eni Mustafaraj. 2011. Limits of electoral predictions using Twitter. In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence.
* Matthew Glassman, Jacob R. Straus, and Colleen J. Shogan. 2010. Social networking and constituent communication: Member use of Twitter during a two-week period in the 111th Congress. 66.
* Jennifer Golbeck, Justin M. Grimes, and Anthony Rogers. 2010. Twitter use by the US Congress. *J. Amer. Soc. Inf. Sci. Technol.* 61, 8 (2010), 1612–1621.
* Stefanie Haustein, Timothy D. Bowman, Kim Holmberg, Andrew Tsou, Cassidy R. Sugimoto, and Vincent Larivière. 2016. Tweets as impact indicators: Examining the implications of automated “bot” accounts on Twitter. *J. Assoc. Inf. Sci. Technol.* 67, 1 (2016), 232–238.
* Sara Hofmann, Michael Räckers, Daniel Beverungen, and Jörg Becker. 2013. Old blunders in new media? How local governments communicate with citizens in online social networks. In *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS’13)*. IEEE, 2023–2032.
* Vishal Kharde, Prof Sonawane et al. 2016. Sentiment analysis of Twitter data: A survey of techniques. *Arxiv Preprint Arxiv:1601.06971* (2016).
* Soonhee Kim and Jooho Lee. 2012. E-participation, transparency, and trust in local government. *Publ. Admin. Rev.* 72, 6 (2012), 819–828.
* Rob Kitchin. 2014. The real-time city? Big data and smart urbanism. *GeoJournal* 79, 1 (2014), 1–14.
* Nan Li and Desheng Dash Wu. 2010. Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Dec. Supp. Syst.* 48, 2 (2010), 354–368.
* Ana Carolina E. S. Lima, Leandro Nunes de Castro, and Juan M. Corchado. 2015. A polarity analysis framework for Twitter messages. *Appl. Math. Comput.* 270 (2015), 756–767.
* David Lorenzi, Jaideep Vaidya, Basit Shafiq, Soon Chun, Neelima Vegesna, Zamil Alzamil, Nabil Adam, Seth Wainer, and Vijayalakshmi Atluri. 2014. Utilizing social media to improve local government responsiveness. In *Proceedings of the 15th International Conference on Digital Government Research*. ACM, 236–244.
* Laura McKenna and Antoinette Pole. 2008. What do bloggers do: An average day on an average political blog. *Pub. Choice* 134, 1 (2008), 97–108.
* Panagiotis T. Metaxas, Eni Mustafaraj, and Dani Gayo-Avello. 2011. How (not) to predict elections. In *Proceedings of the IEEE 3rd International Conference on Privacy, Security, Risk and Trust (PASSAT’11) and the IEEE 3rd International Conference on Social Computing (SocialCom’11)*. IEEE, 165–171.
* Anshul Mittal and Arpit Goel. 2012. Stock prediction using Twitter sentiment analysis. Retrieved from <http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf>).
* A. J. Morales, J. C. Losada, and R. M. Benito. 2012. Users structure and behavior on an online social network during a political protest. *Physica A: Stat. Mech. Applic.* 391, 21 (2012), 5244–5253.
* Fred Morstatter, Jürgen Pfeffer, Huan Liu, and Kathleen M. Carley. 2013. Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose. In *Proceedings of the International Conference on Web and Social Media (ICWSM’13)*.