Does hosting the World Cup create a better global image for a country? A Large-scale Comparative analysis

Akaran Sivakumar, AU720104

202208344@post.au.dk

Cognition and Communication, Cognitive Science, Aarhus University

Jens Chr. Skous Vej 2, 8000 Aarhus, Denmark

Supervisor: Kristian Tylén

04/01/2023

Preparing for a mega-event like hosting a world cup requires immense preparation and infrastructural change. With Qatar being awarded the 2022 world cup in 2010, the time was ticking, and they went straight to work. This however backfired as the pressure they presented on these workers and their poor human rights adherence led to multiple deaths sparking outrage around the world. This has led the public to voice their opinions on social media like Twitter. It has however been difficult to capture the sentiment of Twitter users discussing and commenting on the issue. This study targets this problem by investigating the differences in sentiment and discourse in tweets concerning the World Cup in Qatar. To do so, a Python scraper was used to collect tweets using the hashtag "Qatar2022" (n = *679,440*) and then analysed using sentiment analysis. Through hypothesis testing, this paper finds a statistically significant change in the sentiment in the tweets between the period before the event and during the event. The change was elaborated on and discussed through further evaluation of the results. This work contributes to a better understanding of the sentiment on tweets regarding the world cup in Qatar and provides a statistical and theoretical framework for the linguistic reaction to the event in the domain of Twitter.

Keywords: Qatar; World Cup; Twitter; sentiment analysis; discourse

Table of Contents

1. Introduction

1.1 The World Cup

The FIFA World Cup is one of if not the biggest sporting events in the whole world. Once every four years hundreds of millions of people tune in for a total of 64 games all leading up to the World Cup final which is the most-watched sporting event in the world. There are several potential benefits for a nation that hosts the World Cup. One of the main benefits is the opportunity to showcase the country to the world. Hosting the World Cup can provide a platform for a nation to promote its culture, tourism, and economic opportunities to a global audience. This can lead to increased international visibility and attractiveness as a destination for tourism and investment. Hosting the World Cup can also provide a boost to the local economy, as the event can generate significant economic activity through the construction of new infrastructure, the hiring of local workers, and the influx of tourists. Additionally, hosting the World Cup can bring a sense of pride and national unity, as the event can bring people together and serve as a source of national pride and identity. A great example of this is the 2010 World Cup hosted by South Africa. They proclaimed the hosting of the world cup was a sign of a unified nation and a global communication platform for all their progress in all sectors of society. (Al-Emadi et al., 2017)

Qatar was announced as the host nation for the 2022 World Cup on December 2, 2010, by The Fédération Internationale de Football Association (FIFA), marking the first opportunity for the Middle East to host a so-called mega event. The country has a rich cultural history and is home to a diverse population that includes Qatari nationals, as well as expatriates from around the world. Qatari culture is heavily influenced by Islam, with Islam playing a central role in the country's laws, customs, and daily life. Since being awarded the opportunity to host the 2022 world cup, the nation has undergone major infrastructural developments, accompanied by global criticism about the event award process and several sociocultural issues in Qatari society. Hosting a mega event such as the world cup often requires great infrastructural development to accompany the great masses of crowds that attend these events. To help expedite the process, Qatar chose to hire many migrant workers. This has been met with a large amount of global backlash as these migrant workers have supposedly been subjected to terrible conditions.

1.2 Prior research

A study published in 2017 by (Al-Emadi et al., 2017) investigated the impact of the preparations for the world cup on the Qatari population. Overall, 2,163 interviews with 1,058 Qatari nationals and 1,105 white-collar expatriates. The socio-cultural impacts of hosting the events were the most impactful for both Qatari and expatriate samples. This included the pride of hosting the event, the opportunity to embrace other cultures, developing a sports culture/community, the potential disruption of everyday lives, and peace during the event. There were however differences between the two groups as the expatriates viewed the possible immediate and future economic growth as a central pillar for their support and attitude toward the event, and the Qataris viewed the more long-term benefits as most important such as building a sports culture and creating opportunity for Qatar to achieve greater tourism following the event. The reasoning behind this difference can be explained by the social exchange theory. The social exchange theory is a social psychological and sociological perspective that explains social change and stability as a process of negotiated exchanges between individuals and groups. According to this theory, people weigh the potential costs and benefits of social interactions and relationships and decide whether to engage in them based on their perceived outcomes. (Markovsky & Cook, 1989) As the white-collar expatriates are in Qatar primarily for financial reasons, they weigh the monetary and immediate possibilities and benefits as superior, whereas the Qatari residents who have lived their entire life in Qatar and are expected to continue living their life in Qatar, can see past the immediate benefits and focus on the posterior and potentially more fruitful benefits of hosting the event.

Prior research has shown that hosting a mega-sport event exerts a positive influence on the destination image among football fans. (Florek et al., 2008) A survey was conducted in Camp Sweden, A Swedish football fan group on Facebook, surrounding the likelihood of them attending the world cup in Qatar compared with the 2018 world cup in Russia and the 2016 European Championship in France. Results showed that Qatar had a negative image among Swedish football supporters. (Andersson et al., 2021) The reasoning behind this was explained by Qatar not being associated with aspects that were important for supporters traveling to earlier mega football events, such as the cost to travel to the destination, convenience to travel to the country, attractions, and activities, and climate). The paper also

emphasizes the extreme heat and restrictions on alcohol as factors that potentially could be beneficial or detrimental for Qatar to improve its destination image. (Henderson, 2014)

1.3 Twitter, the micro-blogging platform

Twitter is a social media platform that is widely used by people all around the world to share their thoughts, opinions, and ideas. It has become a popular platform for people to voice their opinions and engage in discussions about a wide range of topics, passing 237 million daily active users worldwide in 2022. (*Twitter Global MDAU 2022*, n.d.) One of the reasons why Twitter is such a valuable platform for understanding public sentiment is the real-time nature of the platform. This means that tweets are posted and shared in real-time, providing a snapshot of what people are thinking and feeling about a particular issue at a specific moment in time. Additionally, the open access provided by the company offers insight into massive amounts of authentic text from the public. With the widespread public interest, the almost daily news about Qatar and their handling of the event, and the rapid increase in the number of tweets, people tend to utter their opinions on various media, making platforms like Twitter a great resource to investigate the attitude of the public.

1.4 Natural language processing and sentiment analysis

In the present day, it is not physically possible to perceive an emotion or a thought in its rawest form, so language can be used as a means of translating complex thoughts and emotions. Various methods have been used to investigate how the use of language can reflect cognitive processes such as attentional focus, social cognition, and thinking styles (Tausczik & Pennebaker, 2010)

With more and more of our daily communication being digital, natural language processing has presented a new way of studying the human mind. Natural language processing (NLP) is a field of artificial intelligence (AI) that focuses on the interaction between computers and humans with natural language. It involves the development of algorithms and systems that can understand and interpret human language. With NLP, language is thought of as being divided into different “levels”. Examples of the levels are the lexical level: where the meaning of the individual words is interpreted, the syntactic level: where the focus is on the grammatical structure of sentences, and the semantic level: where the semantics of a text is analysed. (Liddy, 2001)

One approach to investigating the semantic level of language is opinion mining, also known as sentiment analysis. With the immense amount of data available online, it is now possible the understand public opinion and views through language analytics. Sentiment analysis provides the opportunity to detect large-scale emotional valence. There are various ways to work with sentiment analysis, but the two main approaches are either using a pre-made sentiment lexicon or a machine learning approach. The premade lexicon has already assigned a semantic value, whereas the machine learning approach classifies sentiment by using word embedding. (Kaur et al., 2021) This paper uses opinion mining methods as a window for looking into the public's emotional valence and reaction towards Qatar hosting the World Cup. This paper uses the lexicon-based approach for computational reasons

1.5 Hypothesis

On these grounds, the paper will aim to investigate the following hypothesis:

- The general sentiment surrounding the event will change as people arrive in Qatar and get to experience the nation/environment for themselves. The reasoning behind this hypothesis is as the news is often depicted to focus on bad and negative news as this generates more clicks. We, therefore, expect a more positive sentiment after people have travelled to the nation and experienced it first-hand.

2. Methodology

A total of 679,440 tweets from Twitter search queries including the hashtag "Qatar2022" were scraped in two intervals before and during the Qatar WC (World Cup):

INTERVAL DATE	PERIOD	NUMBER OF TWEETS	TWEETS AFTER CLEANING
20TH SEPTEMBER TO 20TH NOVEMBER 2022	Before Qatar WC	62,982 tweets	49,676 tweets
20TH NOVEMBER TO 20TH DECEMBER 2022	During Qatar WC	616,458 tweets	464,415 tweets

Table 1 Table showing the number of tweets scraped pre- and post-processing

These tweets were all scraped from the micro-blogging platform Twitter using the Python package "snscrape". The search query was limited to only obtaining tweets written in English and included all mentions of Qatar2022 in both intervals to make the sample as homogenous

as possible. The hashtag "Qatar2022" was chosen as this was the most trending hashtag about the World Cup and therefore presented the best opportunity for capturing the largest sample of tweets. Tweets were only scraped in English as the dictionaries used in the later analysis are in English.

2.1 Cleaning and pre-processing

All 679,440 tweets were obtained in the file format comma-separated values (CSV). Before analysis of the tweets was possible, the tweets had to be pre-processed. This was a 4-step process. 1. Lemmatizing the tweets. Lemmatization is the process of reducing inflected words to their base form, also called lemma. For example, the lemma of the word "running" is "run". To lemmatize, one must identify the part of speech and inflection of each word, and then use this information to determine the lemma. The process of lemmatization is often used in natural language processing tasks and data pre-processing for topic modelling or word count. This made each tweet into a list of strings. 2. Filtering out all the special characters and links. 3. Filtering out Stop words from the premade list in the Python package "nltk".(Bird et al., 2009) 4. Removing all the tweets about football games and football players. This was done as these tweets do not contain information about the conditions in Qatar, but rather are focused on the conditions with players and teams. The now cleaned data frames consisted only of lemmatized tweets concerning "Qatar2022" as an event and nation.

Original tweet	#Qatar2022 The selection seemed corrupt. The infrastructure seems built on modern-day slavery. The organizer seems to fail to welcome all.
Pre-processed tweet	['selection', 'seem', 'corrupt', 'infrastructure', 'seem', 'build', 'modern-day', 'slavery', 'organizer', 'seem', 'fail', 'welcome']

Table 2 Example of pre-processes of a tweet

3. Analysis

3.1 Sentiment analysis

To test the hypothesis, a word-level sentiment score was calculated for each tweet with the use of the lexicon-based sentiment, instead of a machine learning approach. This was chosen as it was not seen as beneficial to train the data before mining the data. The dictionary for this operation was provided from the Python package "Textblob" (Loria, 2018). The package is an API (Application Programming Interface) that contains dictionaries and other relevant resources for NLP projects. The dictionary *Textblob* was applied to the two Twitter data sets, which returned a compound sentiment score for each tweet, ranging from -1 to 1. This resulted in a more grained sentiment score than that of a binary outcome of either positive or negative. The sentiment score was then calculated and added to the data frame for the respective periods, resulting in a data frame consisting of the following columns: Date (*date and time for the tweet)*, Tweet (*lemmatized version of tweet)*, sentiment (*the sentiment score of the tweet ranging from -1 to 1)*, and period (*Used to distinguish between tweets before and during the event)*.

As the aim of this analysis was to compare the general sentiment between the two periods, the data was divided into two periods (*before and during*), and a mean sentiment score (MSS) was assigned to each period.

Based on the hypotheses prior stated, we expect a significant difference between the period before the event and during the event. Through assumption testing on the data, the data was concluded to be normally distributed with unequal sample size. Using sentiment as the value, a Levene's test with unequal sample sizes was conducted. This led to Welch's two-sample t-test being conducted to further examine the difference. All statistical tests were conducted using RStudio (R Core Team, 2020)

3.2 Exploratory analysis

As a follow-up after the hypothesis-testing section, an exploratory approach was taken where the most frequent words were visualized to visually see the vocabulary used in this area of the Twitter space and how common phrases and words perhaps differed between the two periods.

Two word clouds were created, one for the tweets before the world cup, and one for the tweets during the world cup. These were created to explore the potential discourse change in the Twitter environment because of the event and how these potential changes could relate to the sentiment scores. The Python package "Wordcloud" and "matplotlib" were utilized (Hunter, 2007; Oesper et al., 2011)

3. Results

4.1 Sentiment Classification

The results of the F-test for equality of variances indicated that the variances were not equal ($p < 0.01$, df = 514,089). The level of significance was set at 0.05. Welch's t-test was then performed to test if the difference in sentiment scores was statistically significant. A Welch's t-test was conducted to compare the means of sentiment scores before the event (M = 0.083, SD = 0.26, n = 49,676) and sentiment scores during the event (M = 0.063, SD = 0.27, n = 464,415). The results indicated that the difference between the means was statistically significant (t (61558) = 16.5, $p < 0.01$). The level of significance was set at 0.05.

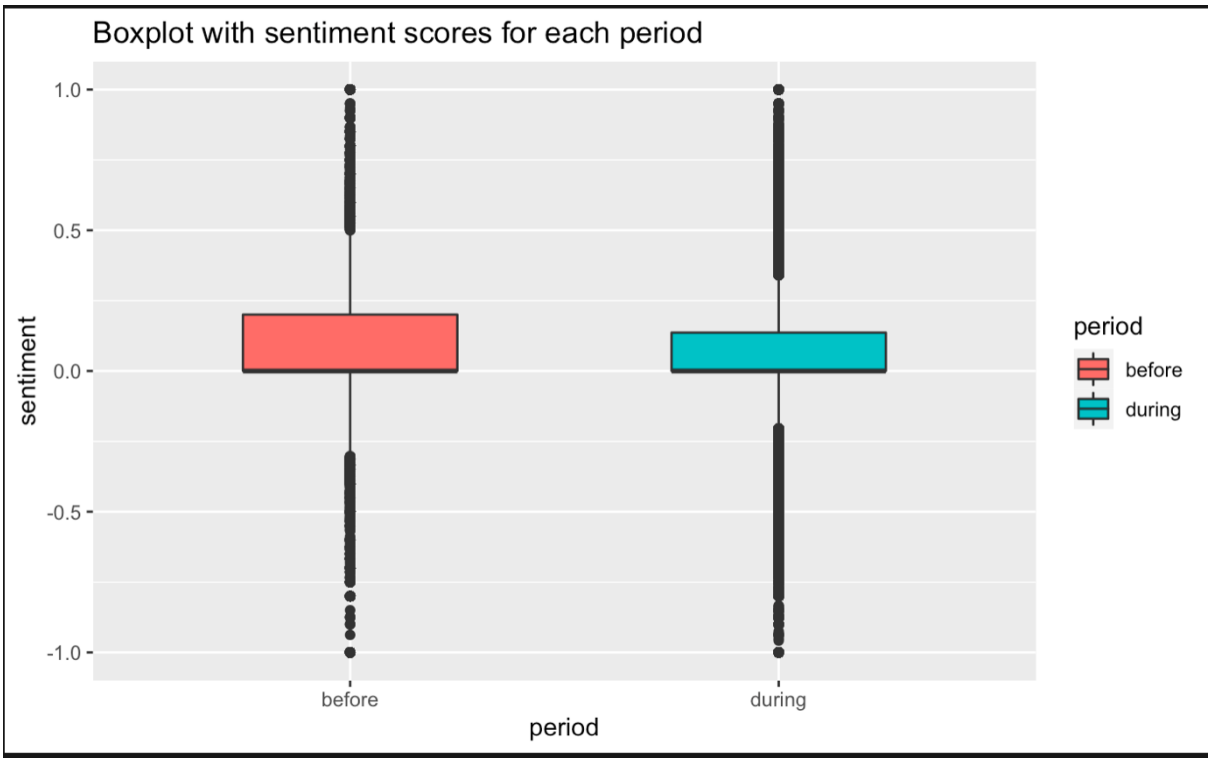

Figure 4. 1 Boxplot with sentiment scores for each period

4.2 Exploratory inspection

To test if there was any discoursal change between the two periods, two word clouds were created. The following words were the most frequent before the tournament:

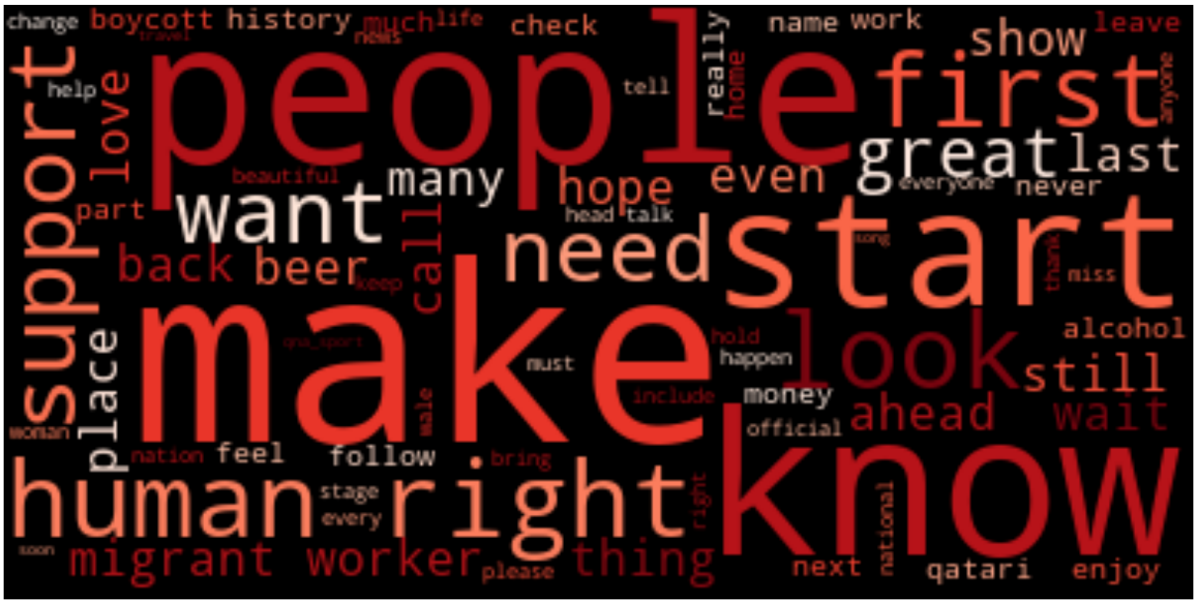

Figure 4.2 Word cloud showing the most frequent words in the period before the event

The visualization displayed in figure 4. 2 gives an understanding of the most frequent words tweeted before the event.

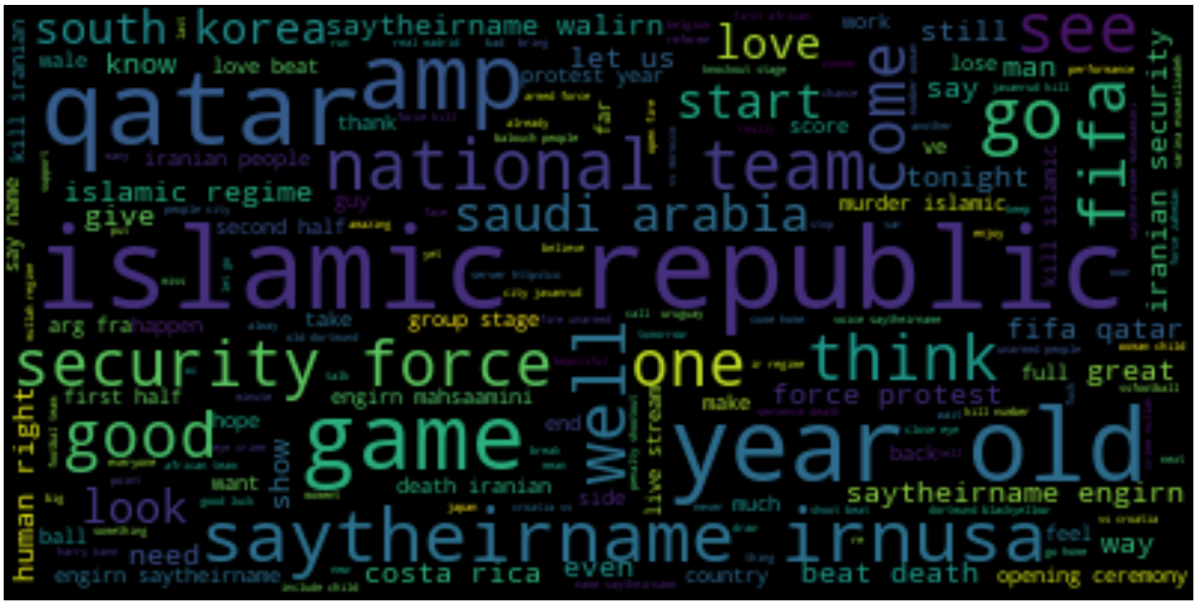

Figure 4.3 Word cloud showing the most frequent words in the period during the event

The visualization displayed in figure 4.3 gives an understanding of the most frequent words tweeted before the event. An important observation regarding the word clouds is the appearance of seemingly normal words such as “need, know, one, and think”. This is a fault

caused by the stop words implementation as these words are not included in the stop words corpus. These words are "ignored" as there are other words, we are more focused on.

5. Discussion

5.1 Hypothesis discussion

In the following section, the prior mentioned hypothesis will be discussed based on the results and findings. As the results showed there was a statistically significant difference in the sentiment regarding tweets before the event and during the event.

Figure 5.1 Tweet showing incentive for Companies to positively talk about the event. (@worldcupguide, 2022)

We see a mean sentiment score of 0.083 before the event which can be explained by the increased attention in Qatar and their violations of human rights paired with tweets from companies or other affiliates who wish to endorse the event for their gain. This is shown in figure 5.1. We see a company called Timeout Doha providing lots of positive sentiment through this tweet. If we then look at figure 5.2, we will then see the personal view of a Twitter user providing negative sentiment to the mean score. These two opposing objectives could be one of the main reasons for the seemingly neutral mean sentiment score for the period before the world cup. Leading up to the start of the event many problems were discussed, but 3 were the most discussed: Human rights violations in infrastructure building, Human rights for the LGBTQ+ community, and the implications of hosting the world cup in a Muslim country

Figure5. 2A twitter user sharing their opinion on the World cup. (@TBTFL_Chief, personal communication, November 20, 2022)

Figure 5.3 Fan raising concern over ridiculously priced alcohol in the Muslim state. (Hughes, 2022)

like Qatar. Figure 5.3 depicts the people's opinion on the last matter as we see a football fan unhappy about the ludicrous prices for alcoholic beverages. The alcoholic beverages offered in these fan zones are only 0,1 % alcohol which could be seen as a bad thing, as earlier studies showed that the availability of alcohol generally also raised the perception of the event(Andersson et al., 2021).

During the event, we see a mean sentiment score of 0.063, which is a statistically significant difference. We first see tenfold the total tweets (*n = 616,458 tweets),* which can be explained by the increased media attention and public attention on the event. The hypothesis was however that we would see an increase in mean sentiment score as we hypothesized that the media painted a negative picture of Qatar to generate more clicks. So, even though we see a statistically significant difference, it is not the difference we expected.

5.2 Exploratory research

To explore the reasoning behind the lowered mean sentiment score during the world, we examined the two generated word clouds to find that even though we only extracted tweets with the hashtag #qatar2022, that does not mean that all our tweets are related to the event. In the word cloud during the event, we see loads of mentions of Islamic republic, say their names, Islamic regime. This can be explained by the conflict in Iran which ignited again during the event. As the event provided the greatest channel to reach the largest amount of people, many people used the hashtag #qatar2022 to ensure their post got more traction and therefore a greater chance of going viral. As Twitter is a fast-paced micro-blogging platform people want to get as many likes, retweets, and shares for their tweets as possible to go viral and become trending. We, therefore, see people "exploiting" trending topics to gain traction.

Even though we found a statistically significant difference in the mean sentiment score between the two periods, it was not the difference we expected. So, we could confirm our hypothesis of there being a change in sentiment between the two periods, but we could not accurately predict the change.

5.3 Methodological implications

In the study of language, there are many approaches and assumptions to utilize. The primary and consistent approach throughout this study has been the “bag-of-words” - approach, which assumes exchangeability. As this assumption disregards the order of words, context, connotation, and irony, it can be classified as a simplification of language. A potential solution to this problem would be to train a machine learning (ML) model to be sentence-based in the classification of tweets in the context of the Qatar world cup - as some words could have negative/positive connotations. This would then result in a change of assumptions from “bag-of-words” to another computationally heavy approach.

As the primary data was collected on Twitter, it brings some limitations. One of them is the spelling errors and extensive use of slang users are prone to have due to Twitter being a casual media platform. This could cause issues in the analysis of the data as these words would just be disregarded in the lexicon-based approach, even though they could prove significant in the outcome. A solution could be the ML model explained in the previous section.

The scraped tweets were all in English and from the same hashtag (#Qatar2022) to maintain some consistency in methods, but this resulted in a lot of unnecessary tweets being scraped as well (unnecessary in the sense of the hypothesis). As all tweets were in English, most of the scraped tweets originated from either UK or USA, or they were the targeted audience. So, to research a broader and more global sentiment analysis, one should scrape in different languages as this includes many more people. Other hashtags could also have been scraped in unison to gather more data. Another point of interest could be the tweets that are geotagged in Qatar. When you geotag a tweet, you tag a location together with the tweet, so viewers know where you are. As the hypothesis was to research the change in sentiment from people who had visited the country this method could have proven instrumental, but unfortunately most tweets posted to Twitter do not geotag a location, so the samples would have been tiny compared to the current samples

5.4 Further work

Further research within this field could contribute by using more advanced models to investigate the language used on Twitter. This has worked with a sentiment dictionary with a

binary view of sentiment, either assigning a negative or positive sentiment score. Language, however, consists of multiple dimensions. Therefore, it would be worth considering using a dictionary with a polynary sentiment scale, such as the *NRC* dictionary provided by the *Syuzhet* package. This dictionary would provide the opportunity to assign a score to a word in the dimension of ten different emotions, such as anger, fear, and surprise. Examining multiple dimensions at once could contribute to a more well-rounded and accurate interpretation of the emotional reactions on Twitter connected to the World Cup. As previously mentioned, using a ML model instead of the "bag-of-words"- approach could also contribute to another understanding of the data. This could be done by using a method, where the contexts are not disregarded.

Due to the timing of this assignment and the event, it was unfortunately impossible for me to research the period after the world, and thereby have 3 periods to compare with each other. We could then run an ANOVA test instead. This could be accompanied by a survey presented to football fans, who traveled down to Qatar to watch the event so that we have more first-hand experience. This would prove more effective as scraping tweets from a hashtag does not guarantee that the tweeters have experienced the things they tweet about.

6. Conclusion

Since 1930 Countries around the world have met every four years to determine the world's greatest footballing nation. Such a big tournament garners loads of attention both from the public as well as the media. However, hosting such a big tournament also requires years of preparations and sometimes also infrastructural changes. With Qatar as the host nation for the 2022 world cup, the nation had to make drastic changes quickly unfortunately leading to the deaths of migrant workers. This was only one of the issues in Qatar as human rights for the LGBTQ+ community were in danger as well as the alcohol consumption the west typically connect with football was still up for debate. The popular social media Twitter was used to research public sentiment regarding the event. Twitter offers open access, which offers insight into massive amounts of authentic text from the public. This paper aimed to investigate the emotional valence of these tweets surrounding the World Cup in Qatar, by examining if there was a difference in sentiment before and during the event. This was done by scraping Twitter for tweets commenting on Qatar2022 using the hashtag "Qatar2022", with a total of 679,440 tweets collected in two periods: 20th September 2022 - 20th November

2022 and 20th November 2022 - 20th December 2022. After pre-processing and cleaning the data, sentiment analysis was applied to the collected tweets. A Welch's t-test showed a statistically significant difference in mean sentiment score between the two periods, as the mean sentiment score was significantly lower during the tournament compared to before. Upon visual inspection of the word clouds generated for each period, we could see that the data during the event was "tainted" as people used the trending hashtag "Qatar2022" to promote other news such as the conflict in Iran igniting during the same period. As a future direction, it would be worth considering some of the methodological limitations faced in this paper, one of them being the loss of context when using the "bag-of-words"-approach.

7. References

Al-Emadi, A., Kaplanidou, K., Diop, A., Sagas, M., Le, K. T., & Al-Ali Mustafa, S. (2017). 2022 Qatar World Cup: Impact Perceptions among Qatar Residents. *Journal of Travel Research*, *56*(5), 678–694. https://doi.org/10.1177/0047287516652502

Andersson, S., Bengtsson, L., & Svensson, Å. (2021). Mega-sport football events' influence on destination images: A study of the of 2016 UEFA European Football Championship in France, the 2018 FIFA World Cup in Russia, and the 2022 FIFA World Cup in Qatar. *Journal of Destination Marketing & Management*, *19*, 100536. https://doi.org/10.1016/j.jdmm.2020.100536

Bird, Klein, & Loper. (2009). *Natural language processing with Python: Analysing text with the natural language toolkit.* (3.8). Reilly Media, Inc.

Florek, M., Breitbarth, T., & Conejo, F. (2008). Mega Event = Mega Impact? Travelling Fans' Experience and Perceptions of the 2006 FIFA World Cup Host Nation. *Journal of Sport & Tourism*, *13*(3), 199–219. https://doi.org/10.1080/14775080802310231

Henderson, J. C. (2014). Hosting the 2022 FIFA World Cup: Opportunities and challenges for Qatar. *Journal of Sport & Tourism*, *19*(3–4), 281–298. https://doi.org/10.1080/14775085.2015.1133316

Hughes, G. (2022, November 20). *"Oh, and one last thing… UNLESS YOU'RE LUDICROUSLY RICH, THERE WILL BE NO BOOZE FOR YOU!! Ha! Good luck with that FIFA, you stupid miserable contemptible – (and easily bribed) – morons." Graham Hughes 5 December 2010*. https://twitter.com/EveryCountry/status/1594104865978044417

Hunter. (2007). *Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3), 90–95.* (3.6) [Python].

Kaur, H., Ahsaan, S. U., Alankar, B., & Chang, V. (2021). A Proposed Sentiment Analysis Deep Learning Algorithm for Analyzing COVID-19 Tweets. *Information Systems Frontiers*, *23*(6), 1417–1429. https://doi.org/10.1007/s10796-021-10135-7

Liddy, E. (2001). Natural Language Processing. *School of Information Studies - Faculty Scholarship*. https://surface.syr.edu/istpub/63

Loria, S. (2018). *TextBlob* (0.15.2) [Python].

Markovsky, B., & Cook, K. S. (1989). Social Exchange Theory. *Social Forces*, *68*(2), 647. https://doi.org/10.2307/2579267

Oesper, Merico, Isserlin, & Bader. (2011). *WordCloud: A Cytoscape plugin to create a visual semantic summary of networks.* (1.8.1) [Python].

R Core Team. (2020). *R: A language and environment for statistical computing.* [R]. R Foundation for Statistical Computing. https://www.R-project.org/.

Tausczik, Y., & Pennebaker, J. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*, 24–54. https://doi.org/10.1177/0261927X09351676

@TBTFL_Chief. (2022, November 20). *The selection seemed corrupt. The infrastructure seems built on modern-day slavery. The organizer seems to fail to welcome all.* https://bit.ly/3IlmY1e

Twitter global mDAU 2022. (n.d.). Statista. Retrieved January 4, 2023, from https://www.statista.com/statistics/970920/monetizable-daily-active-twitter-users-worldwide/

@worldcupguide. (2022, November 20). *This small, Happy Hours list from Timeout Doha may be of interest to a few out there. Especially the one that is " unlimited drinking for QR185 between 6pm and 8pm almost all week.* https://bit.ly/3CjUFwx