# DEMOGRAPHIC ANALYSIS OF FIFA WORLD CUP USING TWITTER

## Objective & Motivation

- The 22nd FIFA World Cup is taking place in Doha, Qatar
- People across the world have different opinions about their favorite team over social media
- We utilize this data to check the sentiments of the people and identify a general voice of the demographic over a certain topic, in this case the FIFA World Cup
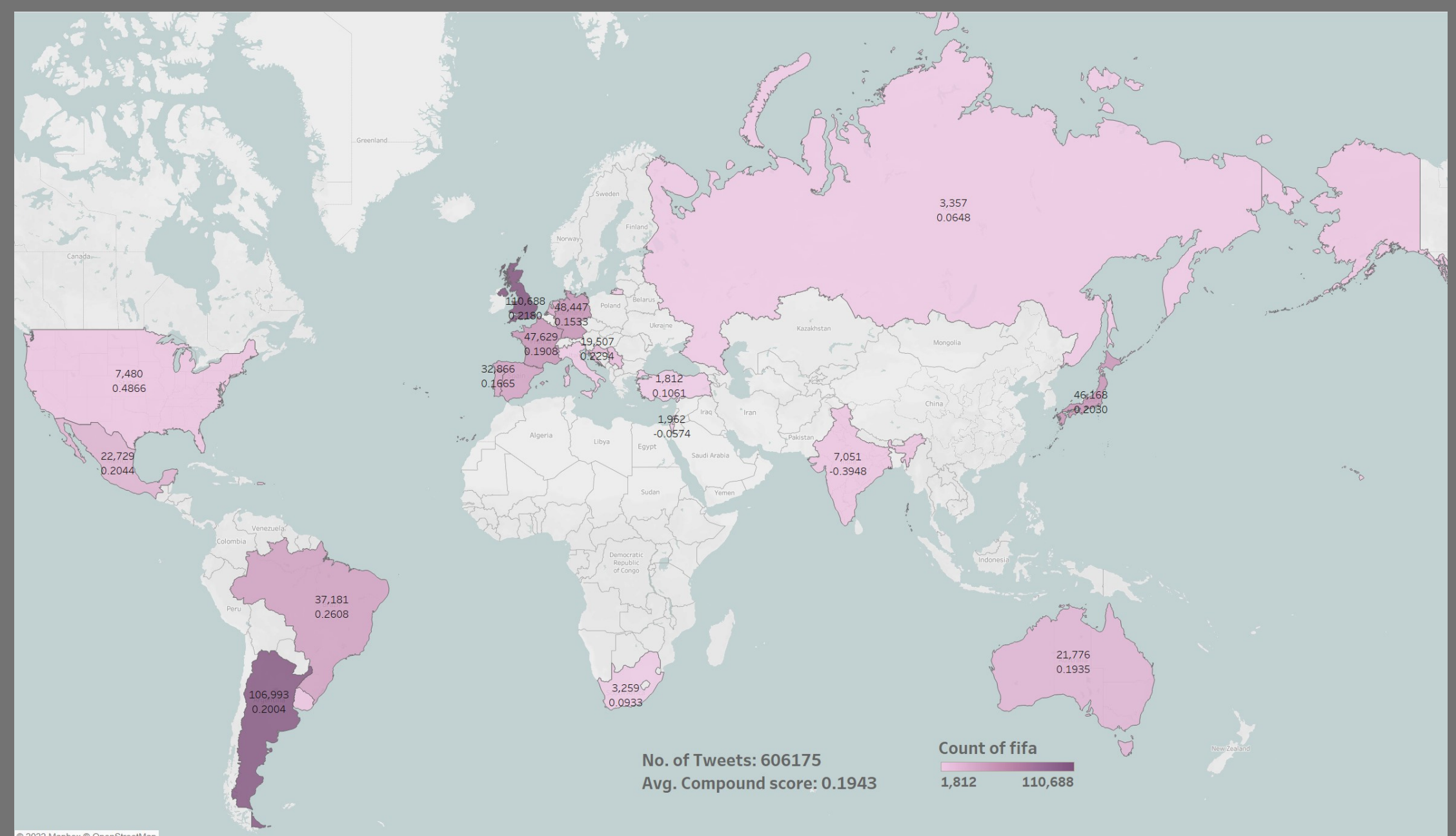
## Challenges Faced

- Language barriers while collecting data
- Limitation on gathering tweets due to API timeout
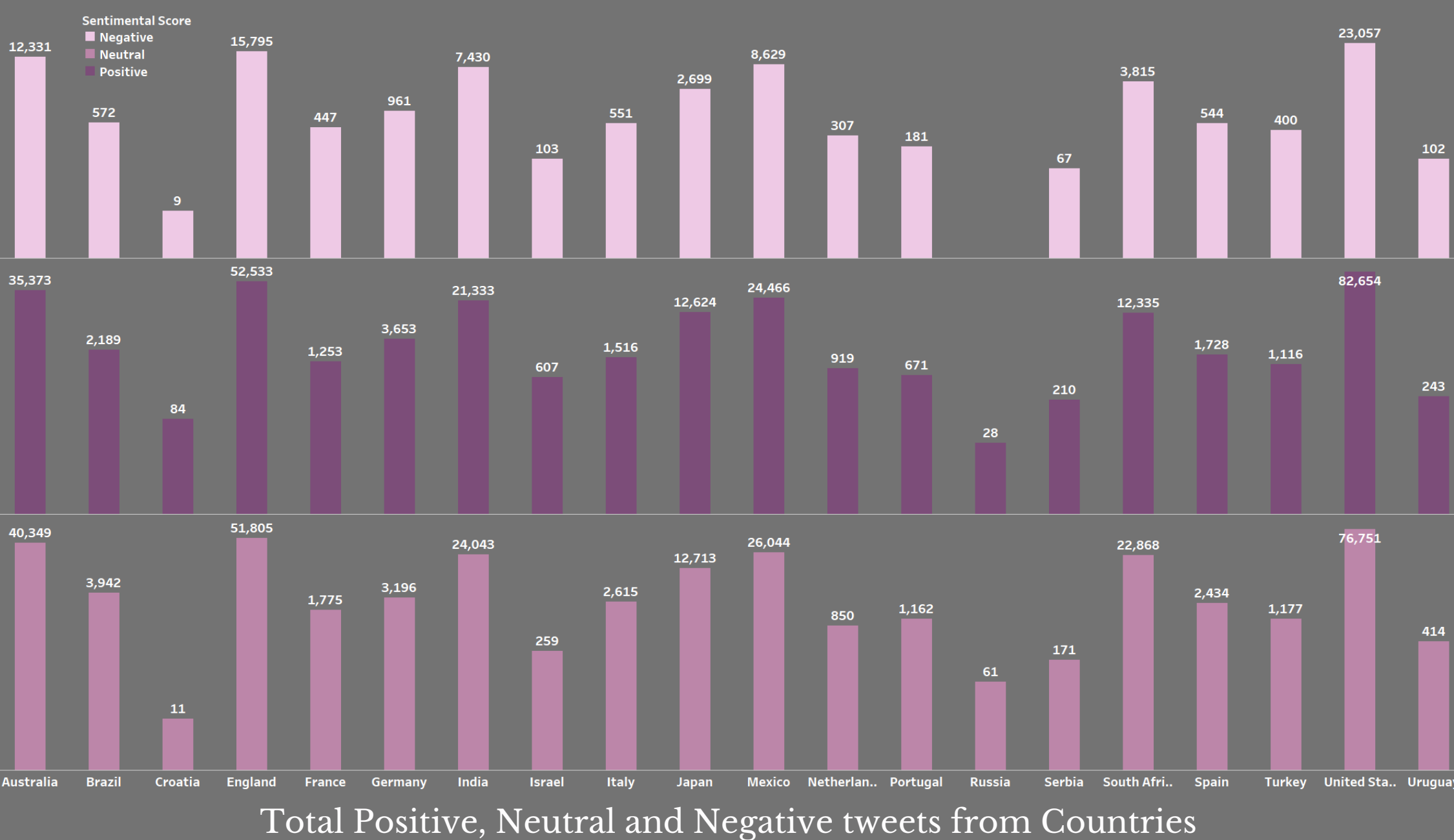- Cleaning the ASCII characters and spams in tweets

## Our Solution

- We shortlisted a few countries where FIFA world cup is popular and extracted tweets from those countries using Twitter API
- We collected tweets originating from the selected countries based on the #FIFAWorldCup followed by the name of the country playing in the World Cup
- Once we extracted the tweets, we processed the data and then used NLP techniques like Stemming, Lemmatization, Removal of Stop Words, Sentiment Analysis, and Word Summarization
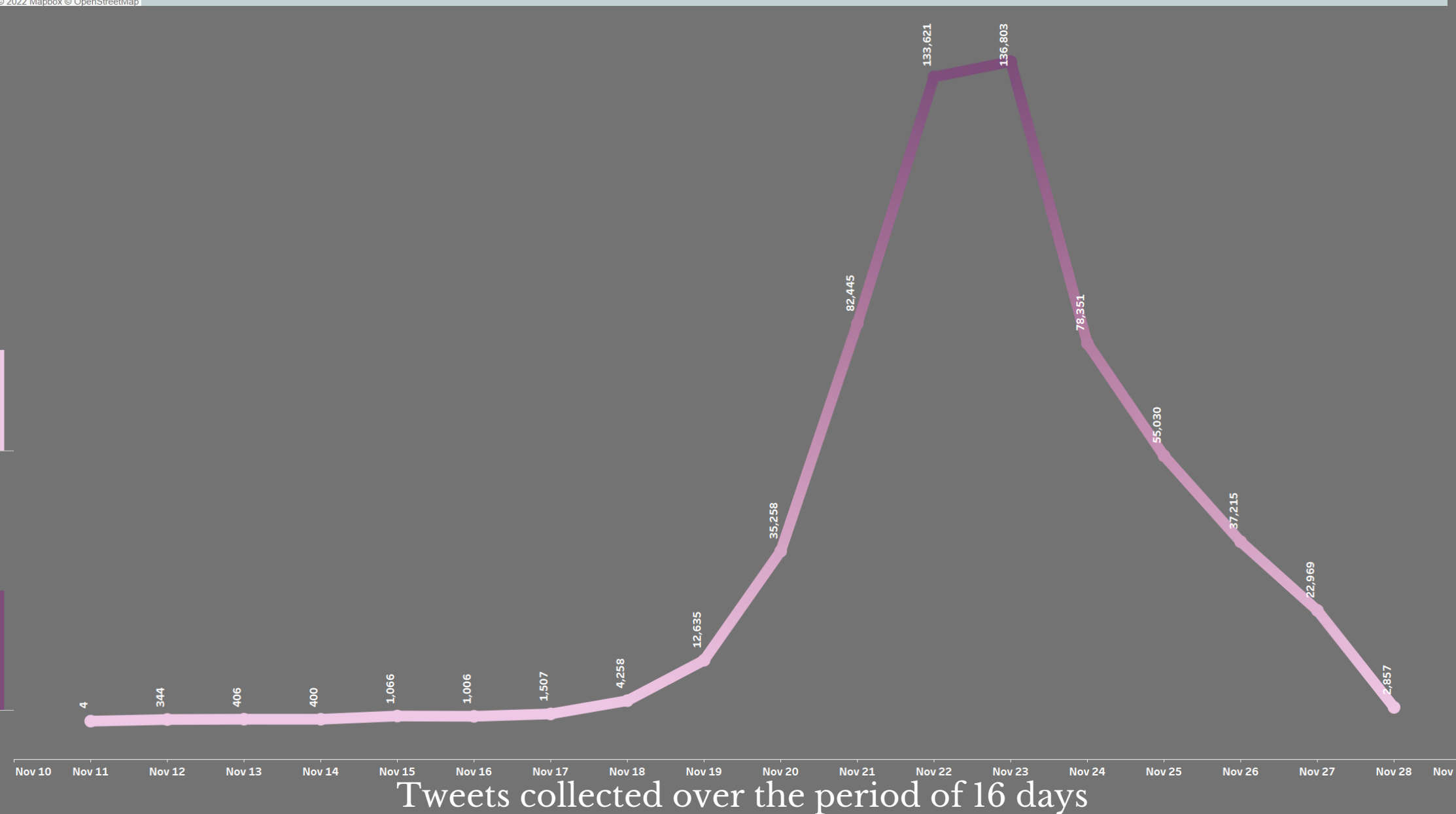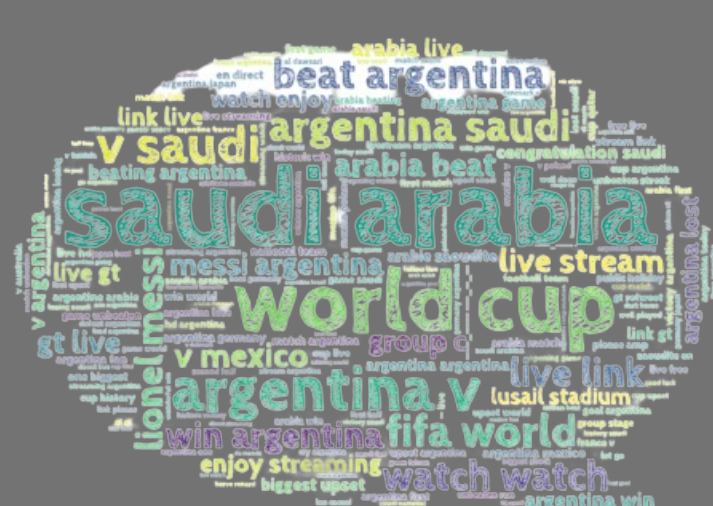
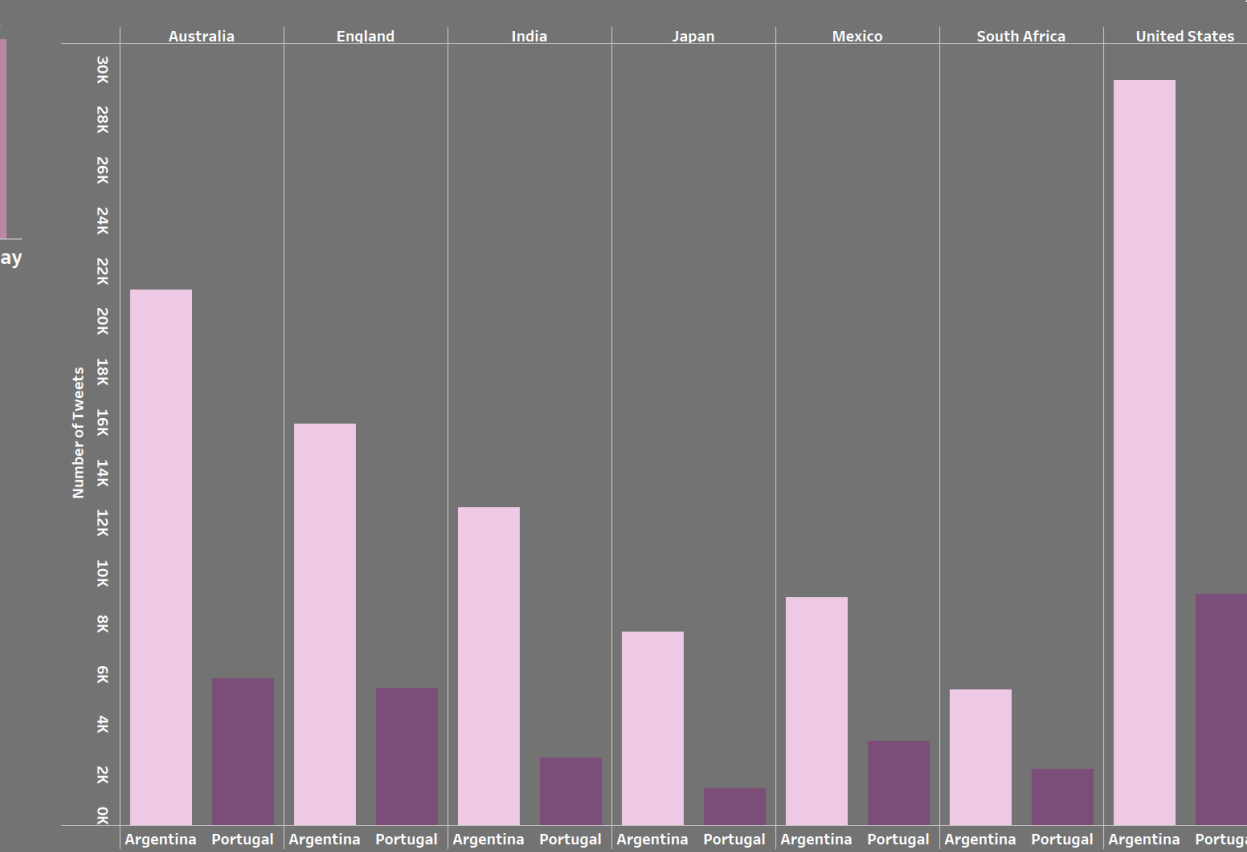## Demographic Analysis



No. of Tweets: 606175
Avg. Compound score: 0.1943

Count of fifa
1,812    110,688

## Conclusion



Sentimental Score: Negative, Neutral, Positive

Total Positive, Neutral and Negative tweets from Countries

| Tweet Location | % of Total Tweets | Count of tweets |
|---|---|---|
| United States | 30.10% | 182,462 |
| England | 19.82% | 120,133 |
| Australia | 14.53% | 88,053 |
| Mexico | 9.76% | 59,139 |
| India | 8.71% | 52,806 |

Top 5 countries based on highest tweet count



Tweets collected over the period of 16 days

Tweet comparison Argentina vs Portugal

Authors: Nihar Deshmukh, Aditya Ahale, Shubhada Mohite, Siva Ganesh Polepalli
Instructor: Mei Zhang