Analyzing and visualizing a social-media phenomena

FIFA WORLD CUP Qatar2022

Sanjeewa Kulathunga

Idea and Plan

- The idea is to scrape tweets related to the 2022 FIFA World Cup and analyse the data retrieved
  - Scraping tweets related to the social media phenomenon using different python libraries
  - Building network graphs with nodes for countries and edges for tweets that mention both countries
  - Applying sentiment analysis algorithm on tweets using NLP tools
  - Analyzing the graph with sentiment and visualizing the results

Tools Used

- Snscrape
- Plotly
- Natural Language Toolkit
- Pandas
- Sklearn Sentiment Analyzer
- NetworkX and Gephi tool

Scrapping

- Why Twitter?
  - It is a social media platform with rich source of **text** data.
  - It is **legal** to scrape publicly available data.

- Why Scraping by Snscrape library?
  - Limitations of Twitter API with free API keys:
    - It is not possible to get more than 3200 tweets.
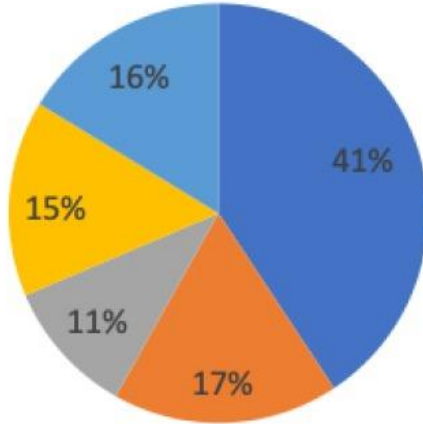   - It can get only tweets within the last 7 days.

Data Processing

- Webscraping
- Data Cleaning
- Data Processing
- General Analysis
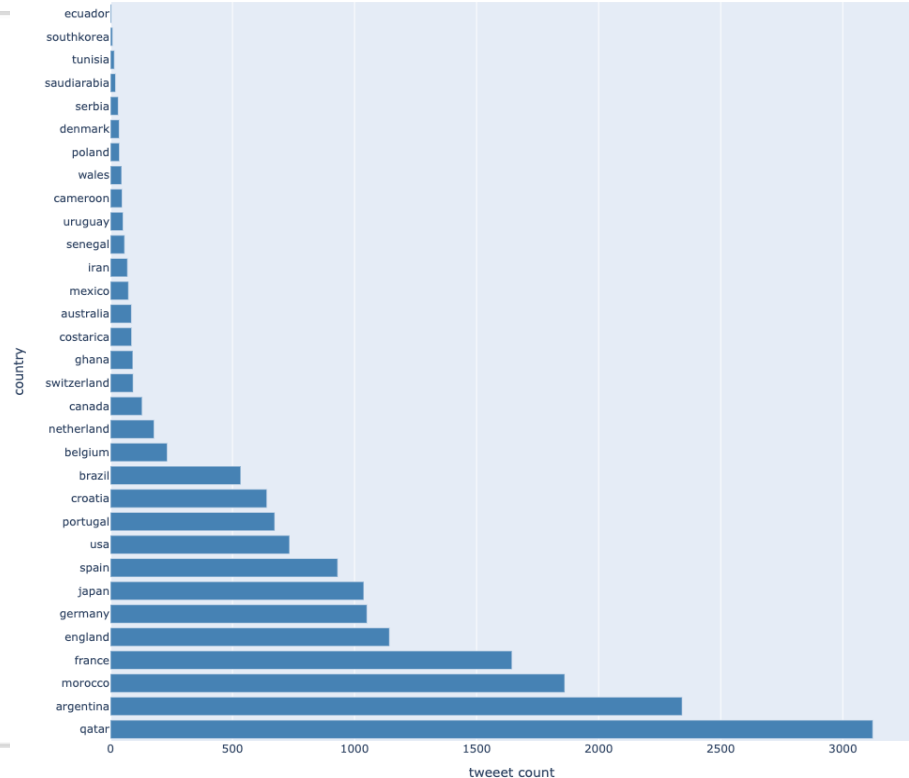- Network Analysis
- Sentiment Analysis

Most Mentioned Countries

Tweets by Schedule

Pie chart legend:
- Knockout stage — 41%
- Round of 16 — 17%
- Quarter-finals — 11%
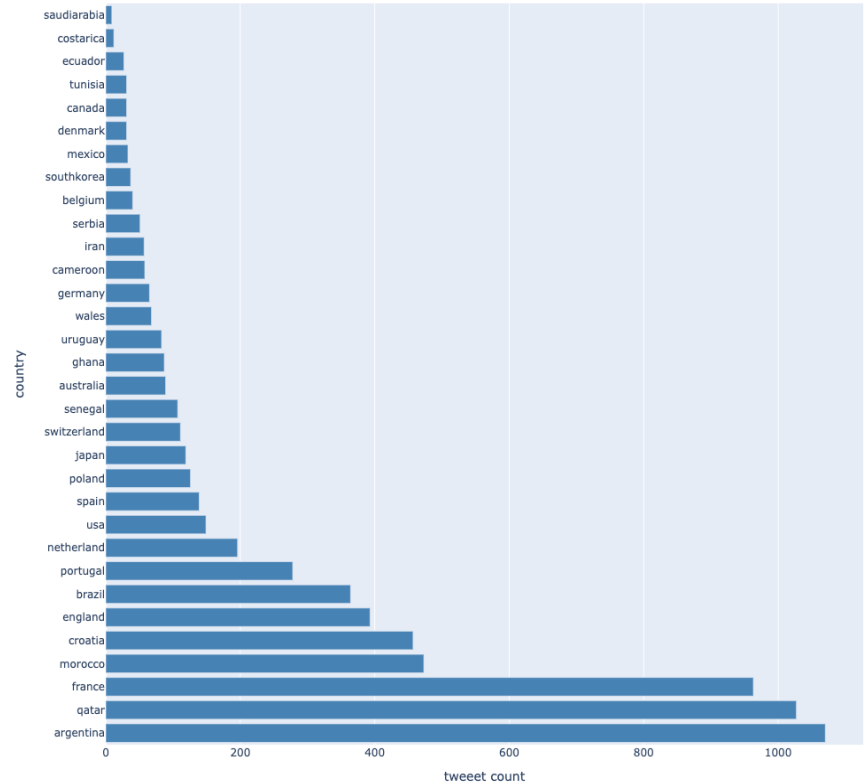- Semi-finals — 15%
- Final — 16%

Top Mentioned Countries

Most Mentioned Countries

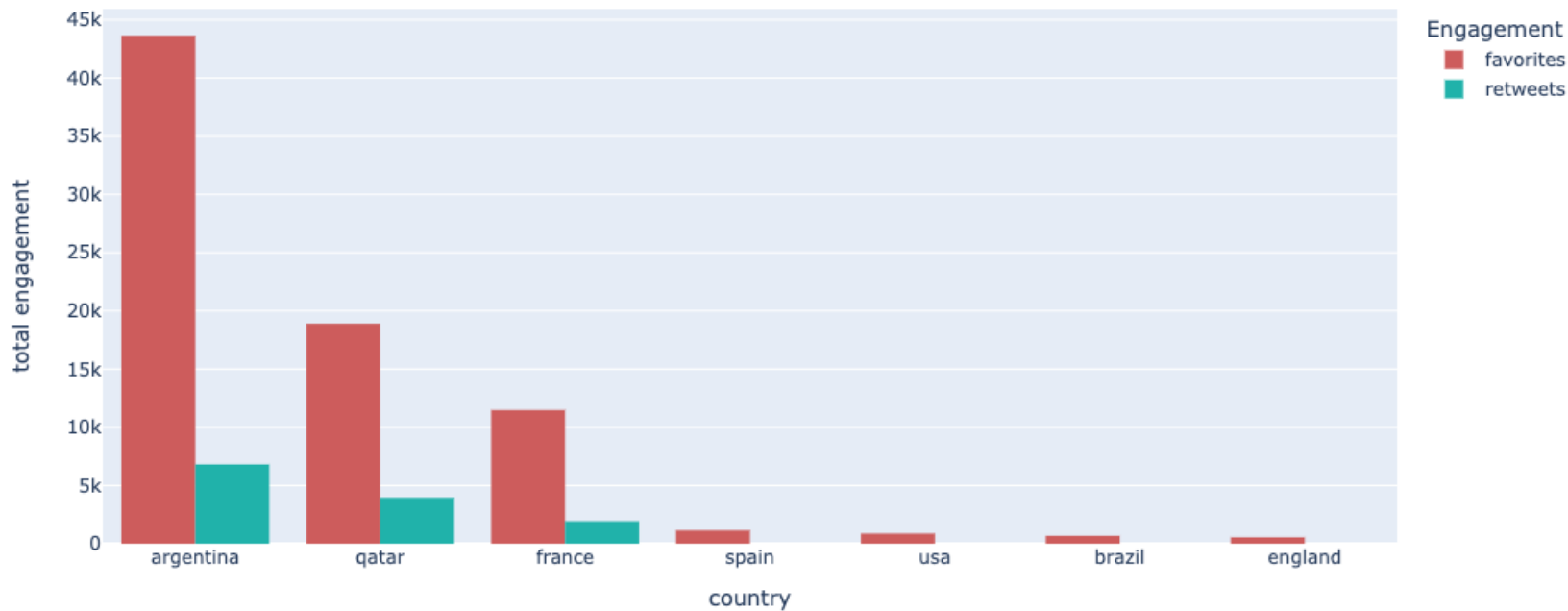- Analysed tweets that mentioned the keyword predict or prediction

Top Mentioned Countries with Keyword Predict

Engagement Analysis of Tweets

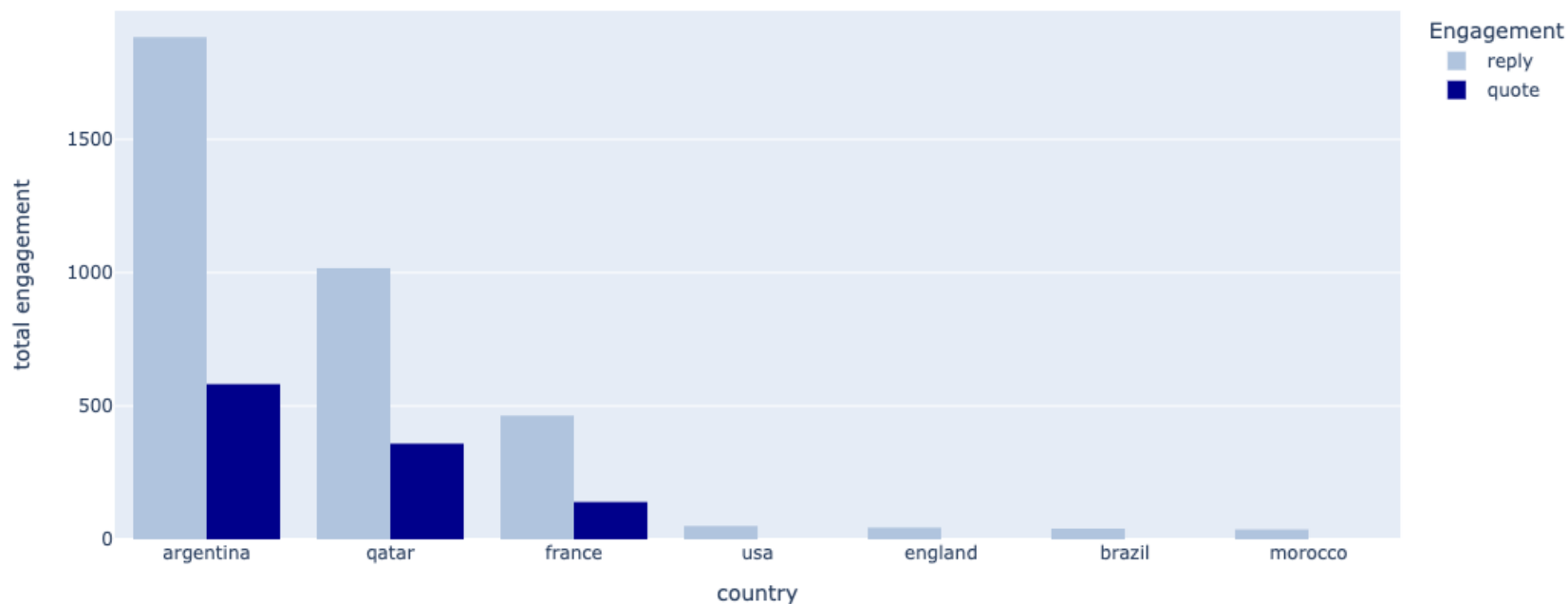Engagement Analysis: Retweets and Likes
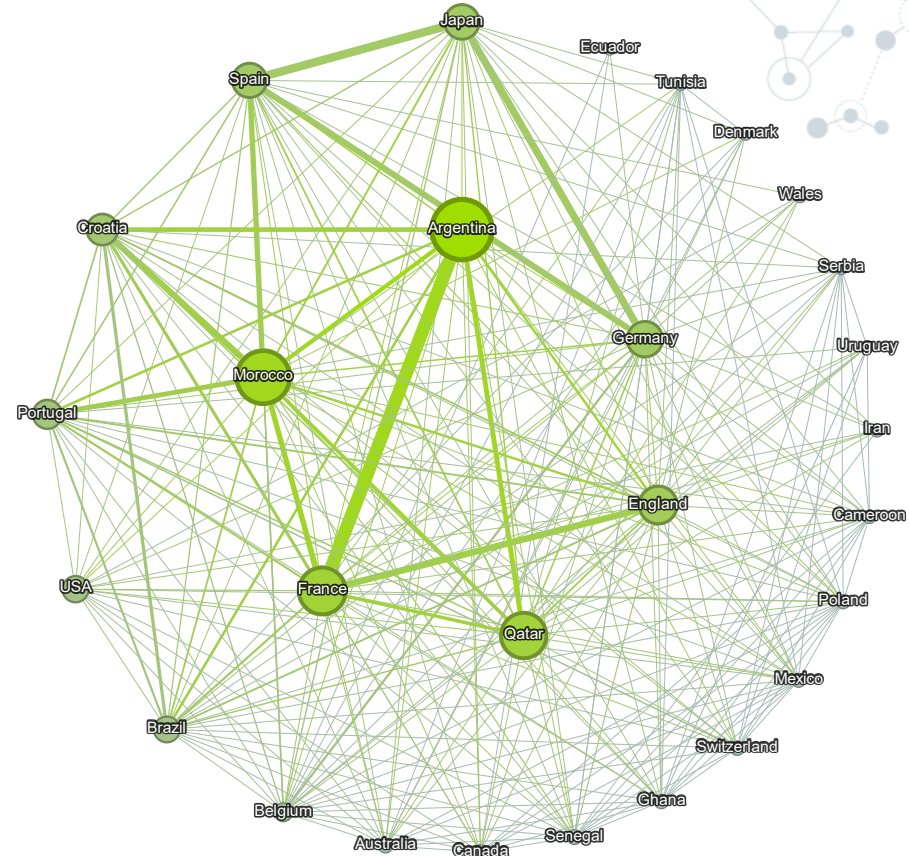
Engagement Analysis Cont'd

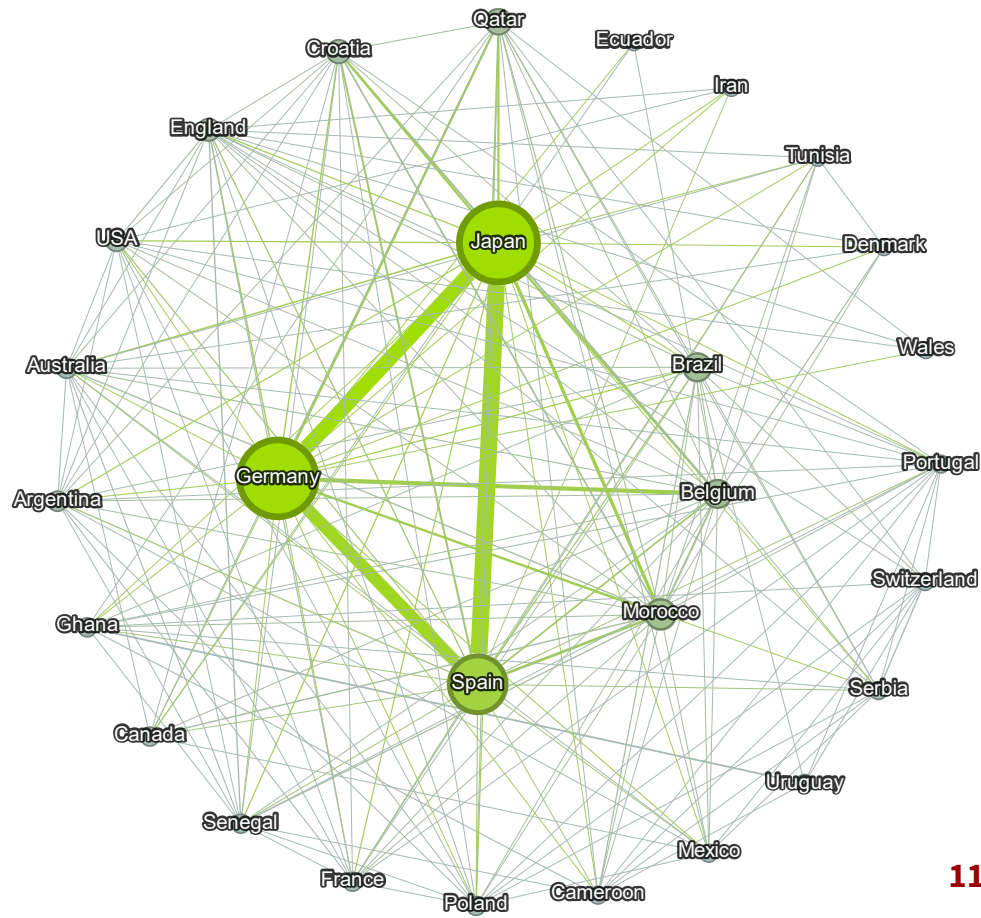Engagement Analysis: Replies and Quotes

Network Analysis

- Node : country

- Edge: a tweet that mentions both countries

- Node size: tweets count that mentions the country

Network Analysis

Knockout stage

- Germany was knocked out due to Japan's shock win over Spain.

Round of 16

NED	3 – 1	USA	
ENG	3 – 0	SEN	
ARG	2 – 1	AUS	
FRA	3 – 1	POL	
JPN	1 – 1	CRO	
Croatia wins 3 – 1 on penalties			
MAR	0 – 0	ESP	
Morocco wins 3 – 0 on			
BRA	4 – 1	KOR	
POR	6 – 1	SUI	
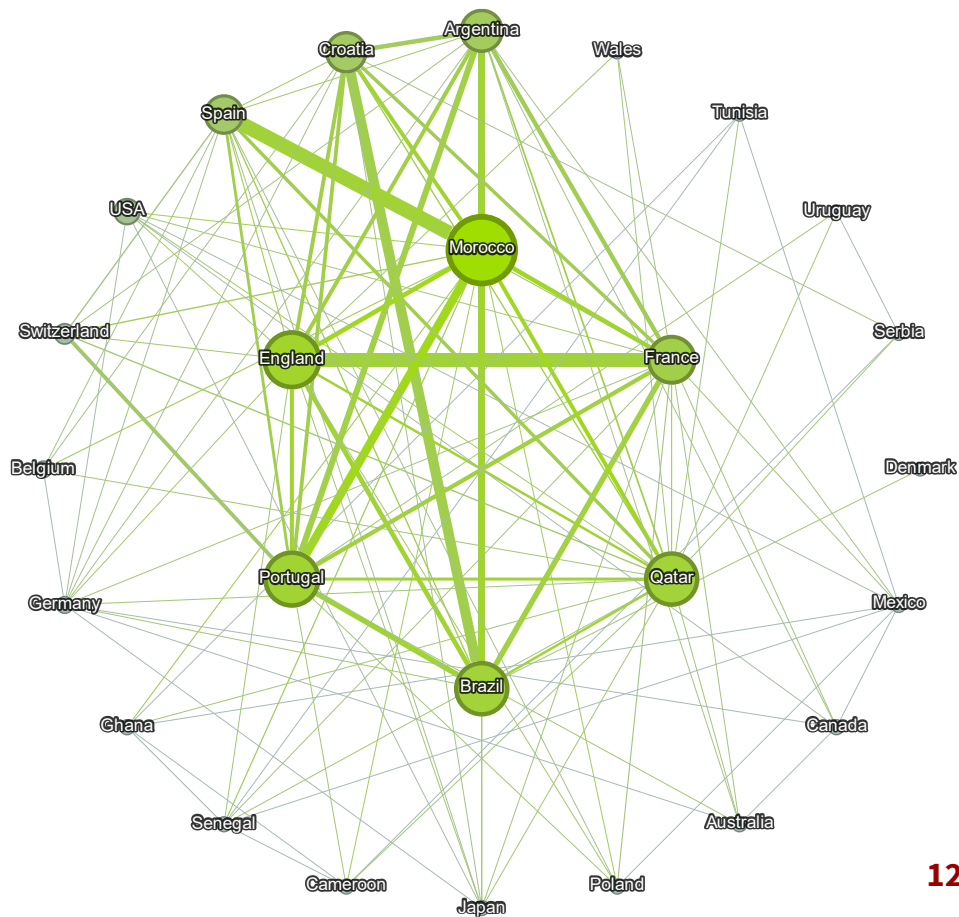
Network Analysis

Quarter-finals

NED 2 – 2 ARG
Argentina wins 4 – 3 on penalties

ENG 1 – 2 FRA
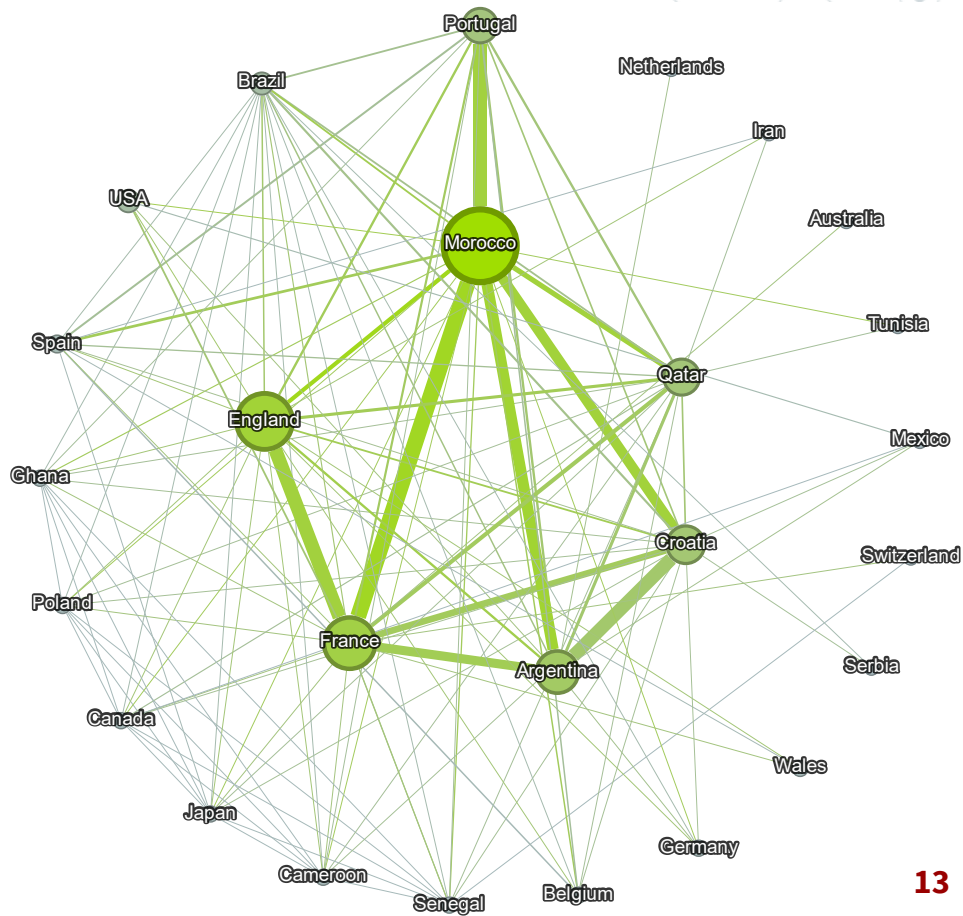
CRO 1 – 1 BRA
Croatia wins 4 – 2 on penalties

MAR 1 – 0 POR

Network Analysis

Semi-finals

ARG	3 - 0	CRO
FRA	2 - 0	MAR

Network Analysis

FINAL

ARG 3 – 3 FRA

Argentina wins 4 – 2 on penalties
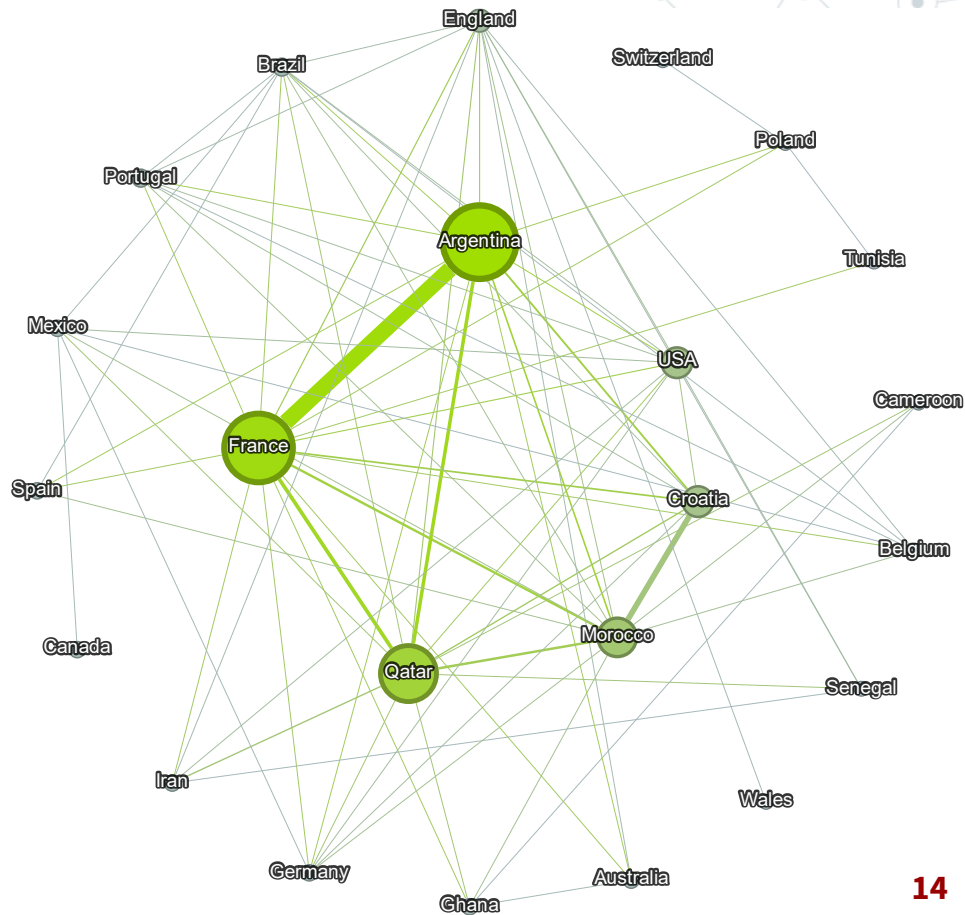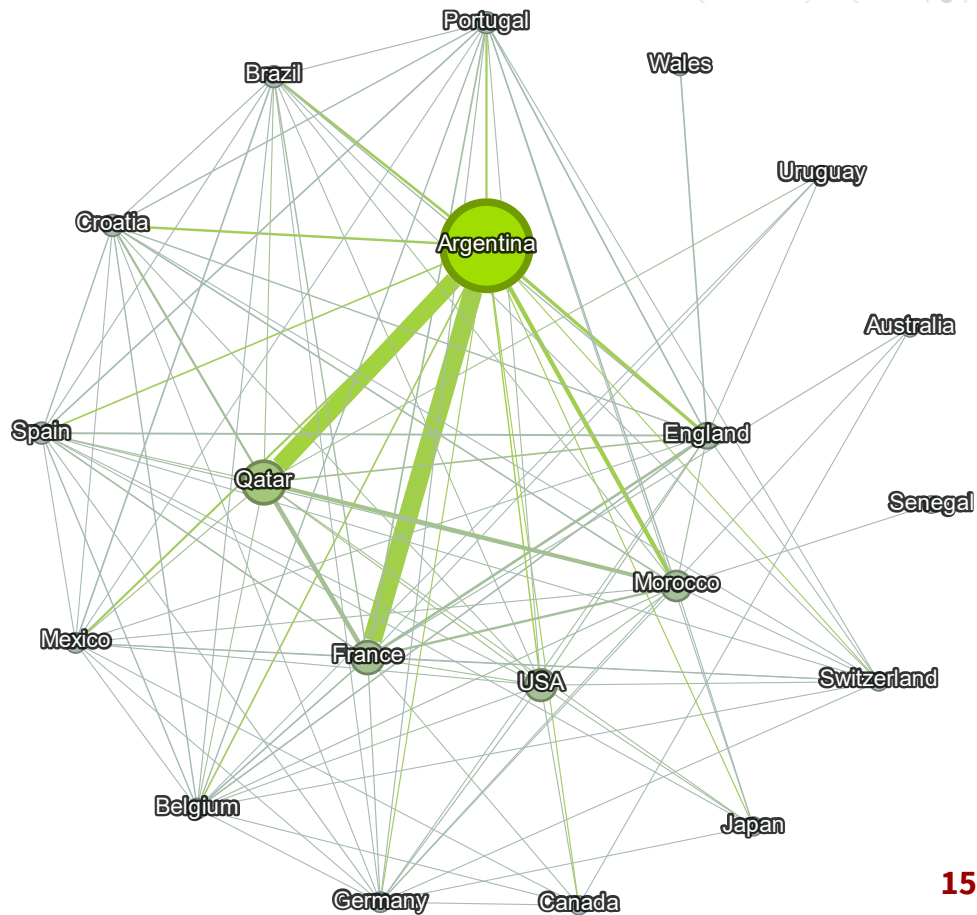
Sentiment Analysis by Country

FIFA 2022 Sentiment Analysis
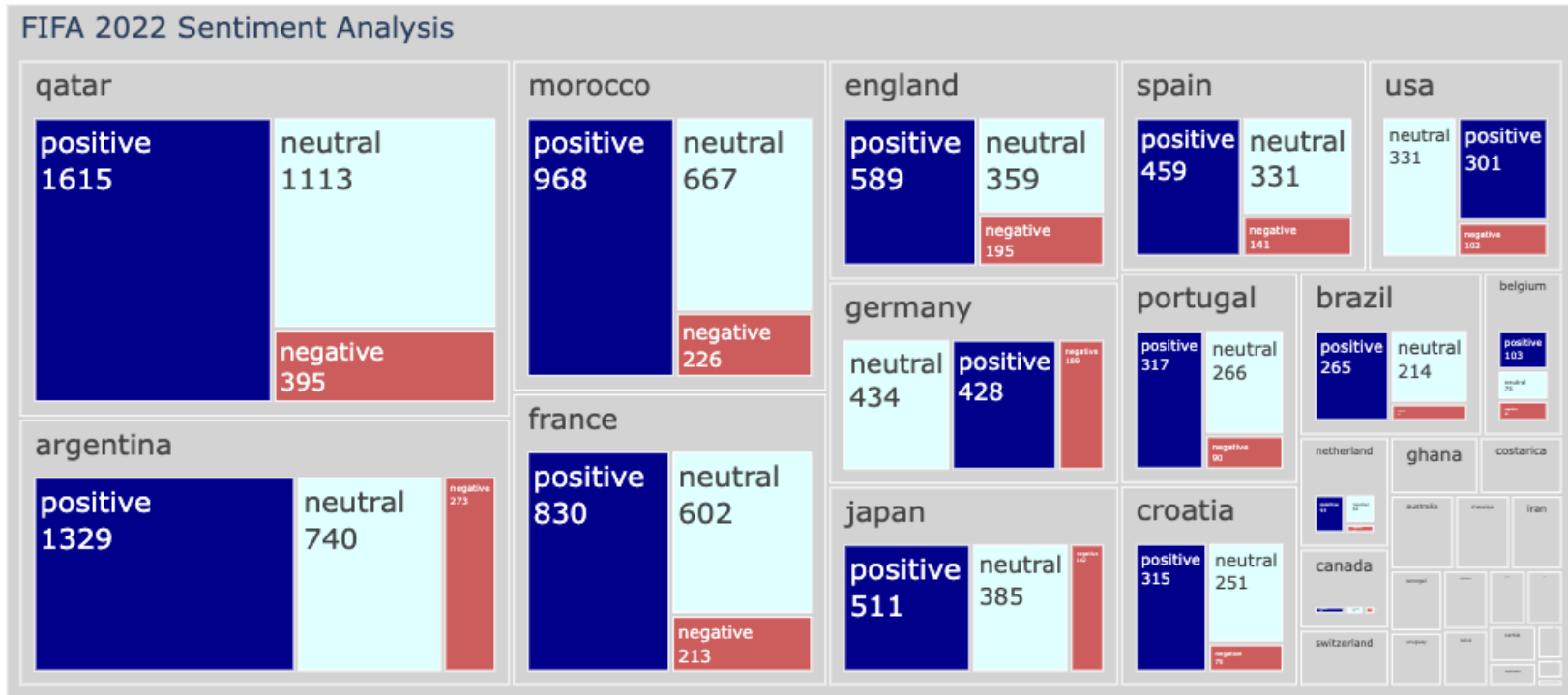
Sentiment Analysis Cont'd

Analysis by Schedule

Percentage Sentiment Analysis on All Tweets
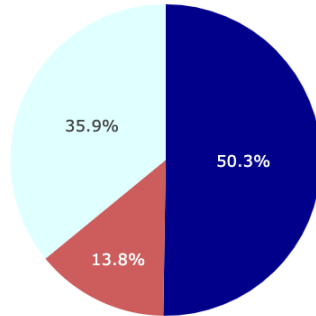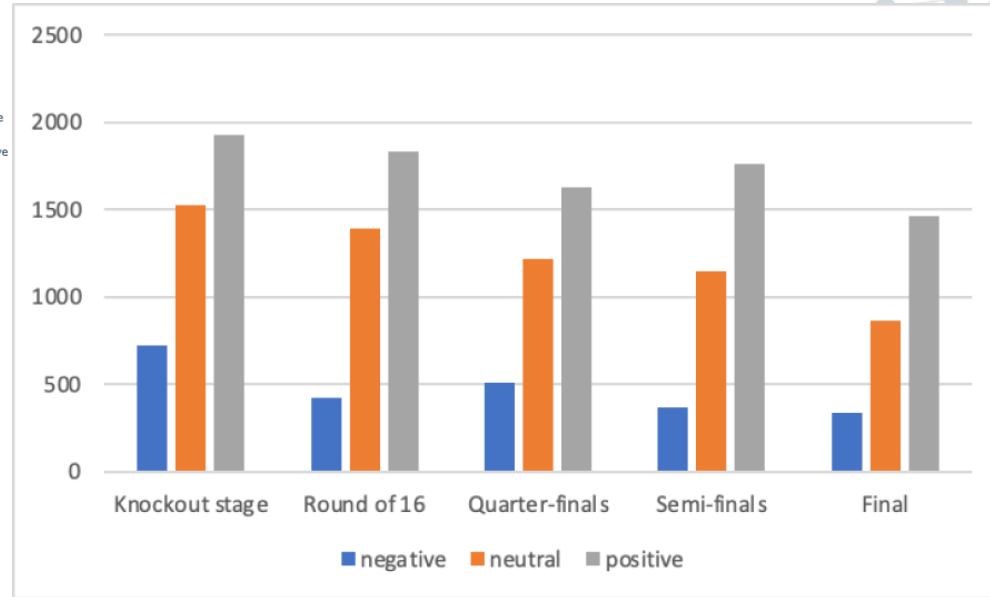
Thanks for listening!

Sanjeewa Kulathunga