

PREDICTING FIFA 2022 QUALIFIERS

TEAM
25

Alex Kim, Steve Jang, Jamie Tay, Janette Lin

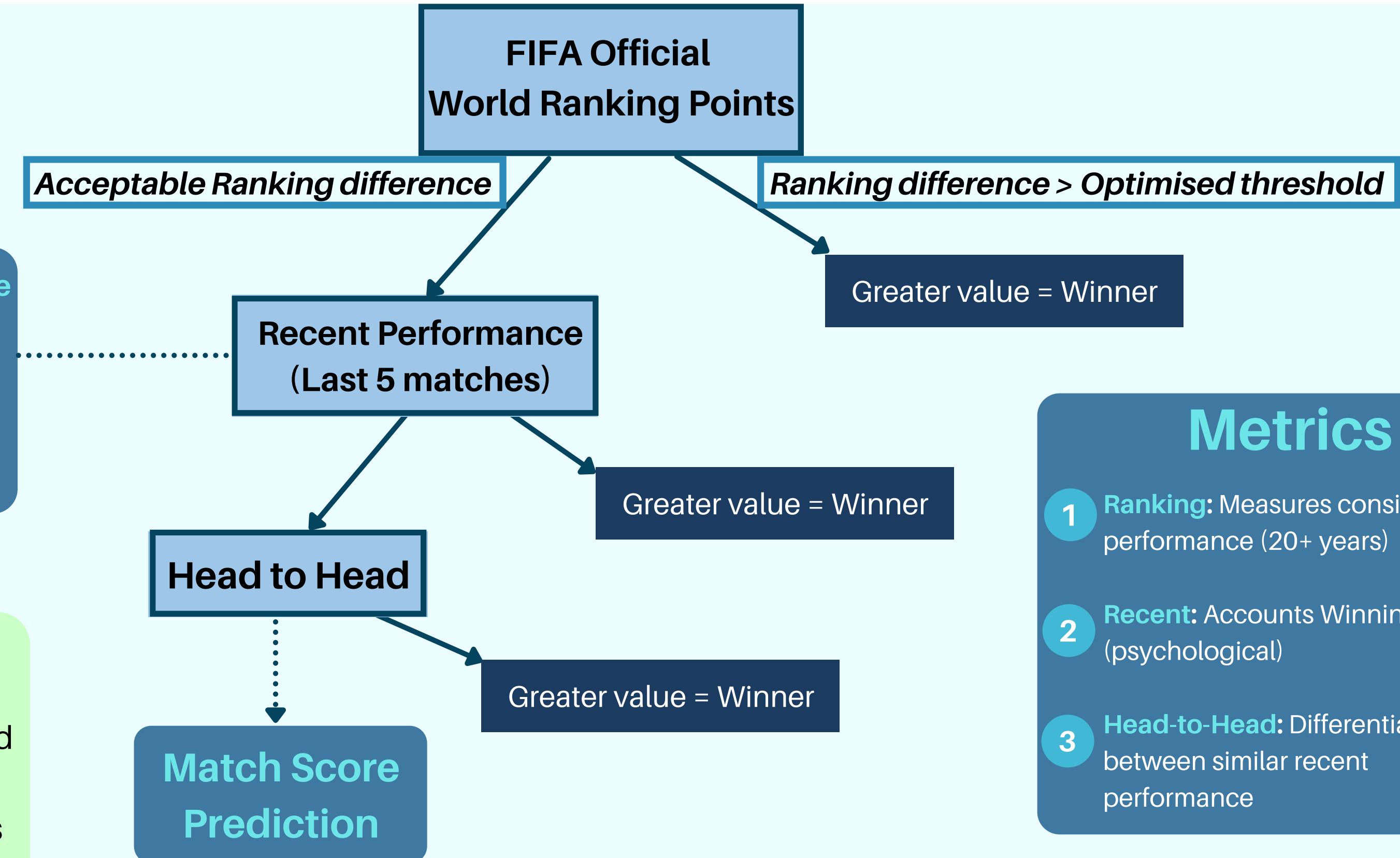
QUALIFICATION PROCESS (AS OF CURRENT)



ALGORITHM BREAKDOWN

FIFA Performance Formulae
 $P = M \times I \times T \times C$
P: Points
M: Win/Loss/Draw etc
I: Match Importance
T: Opponent Team Strength
C: Federation Strength

78.13%
25/32 Teams Predicted
when tested against
Russia 2018 qualifiers
with ~2017 data



- ## Metrics
- 1 **Ranking:** Measures consistent performance (20+ years)
 - 2 **Recent:** Accounts Winning Streaks (psychological)
 - 3 **Head-to-Head:** Differentiate between similar recent performance

QUALIFIER PREDICTIONS

North & Central America (CONCACAF)



United States	
Mexico	
Canada	
Costa Rica	

South America (CONMEBOL)



Argentina	
Brazil	
Uruguay	
Colombia	
Peru	

Africa (AFC)



South Africa	
Senegal	
Algeria	
Tunisia	
Ivory Coast	

Asia (AFC)



Iran	
South Korea	
Australia	
Japan	
Qatar	

Europe (UEFA)



Portugal	
Spain	
Italy	
France	

Belgium	
Denmark	
Netherlands	
Croatia	
England	

Germany	
Switzerland	
Sweden	
Ukraine	



FIFA 2022 INFOGRAPHIC

TEAM
25

Alex Kim, Steve Jang, Jamie Tay, Janette Lin

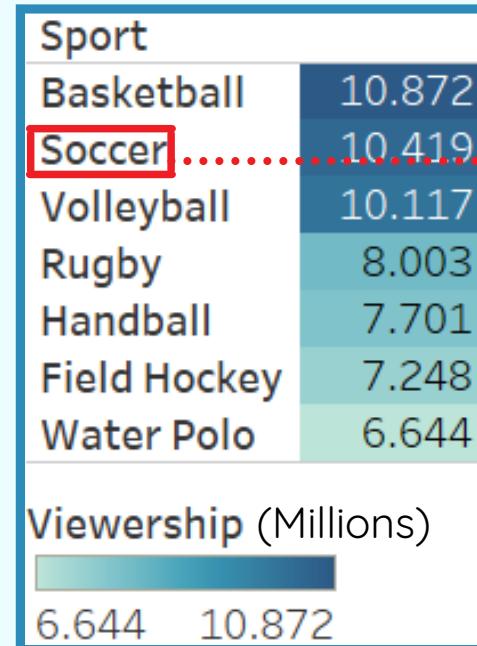
LOCATION: FIFA HEADQUARTERS

CURRENTLY: SPONSORSHIP SEMINAR IN PROGRESS





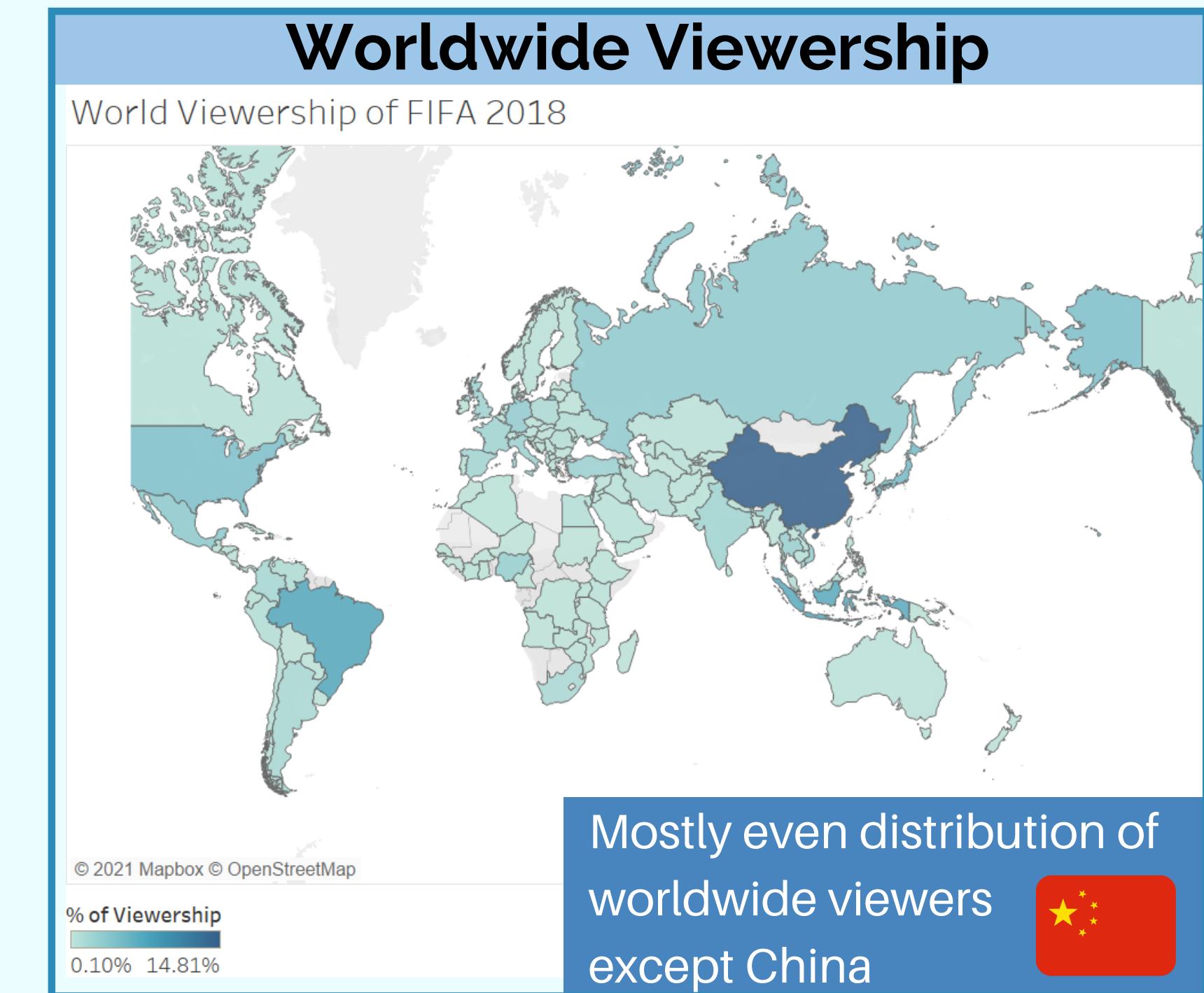
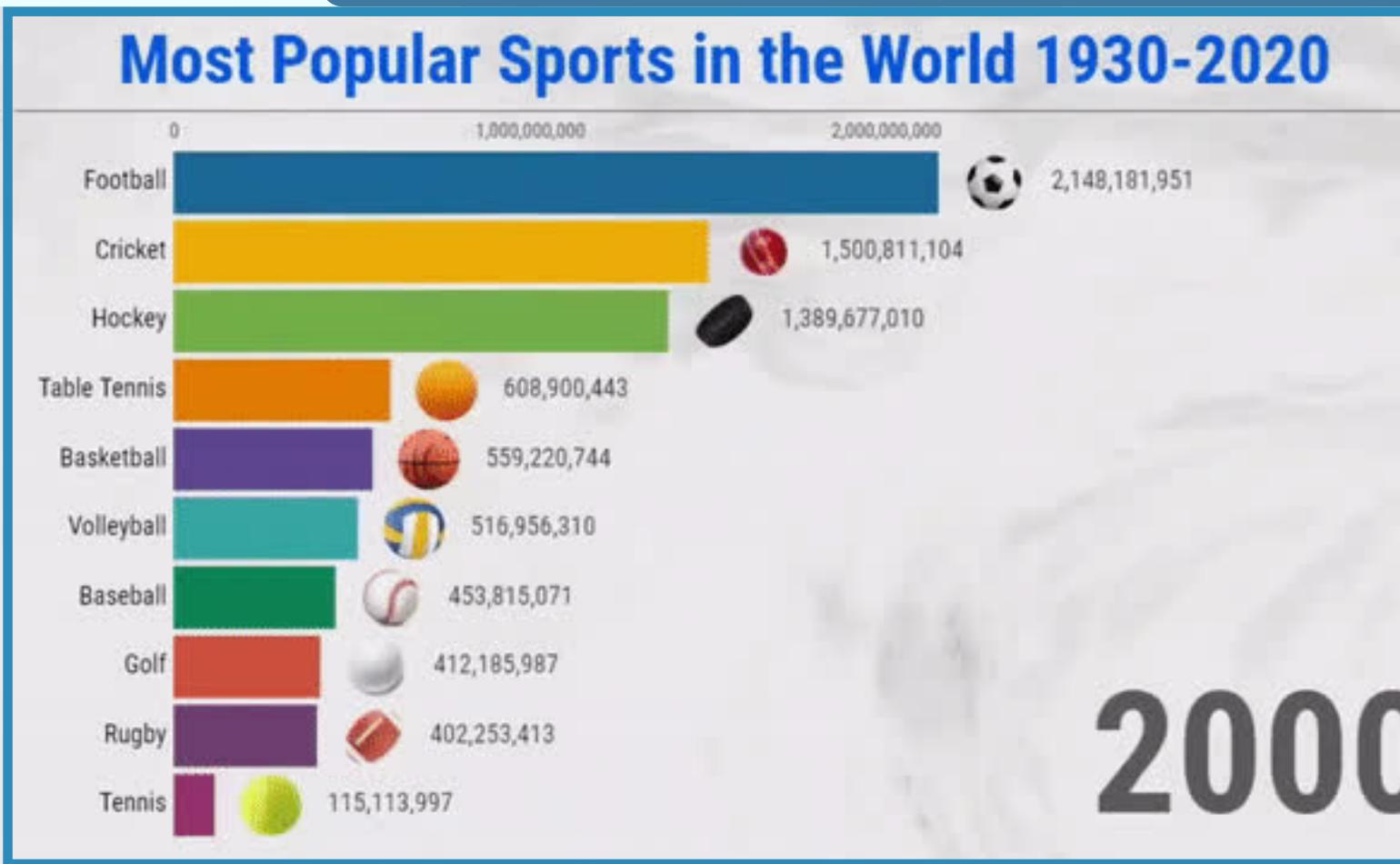
INTERNATIONAL APPEAL



2nd most watched team sport in Tokyo
2021 Olympics (worldwide)

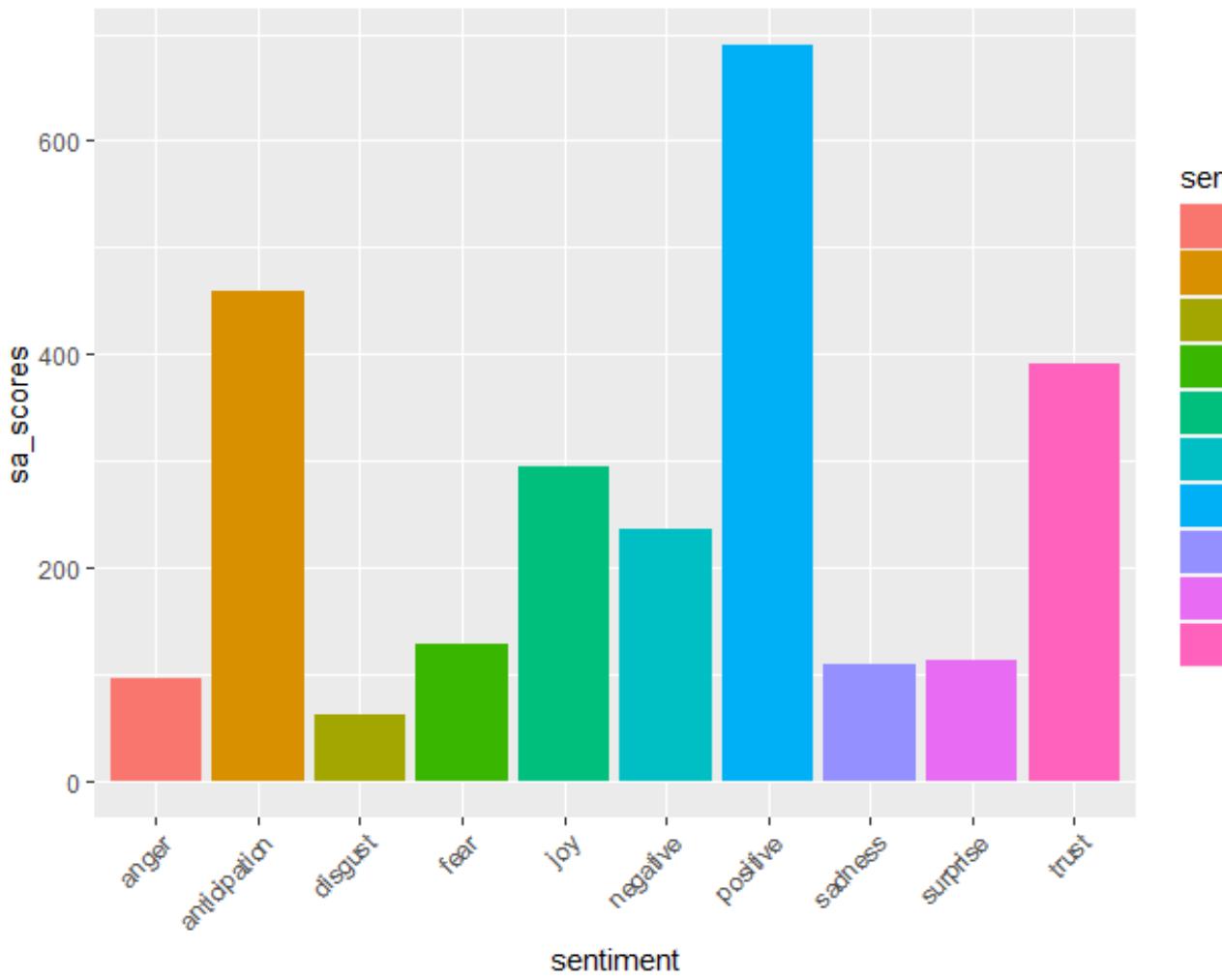


Undisputed Popularity over 30 years



HIGH ANTICIPATION

Sentiment Analysis of 3000 Recent Tweets on 'FIFA World Cup'



76% of Tweets
(Sept '21) on
'FIFA World Cup'
Positive

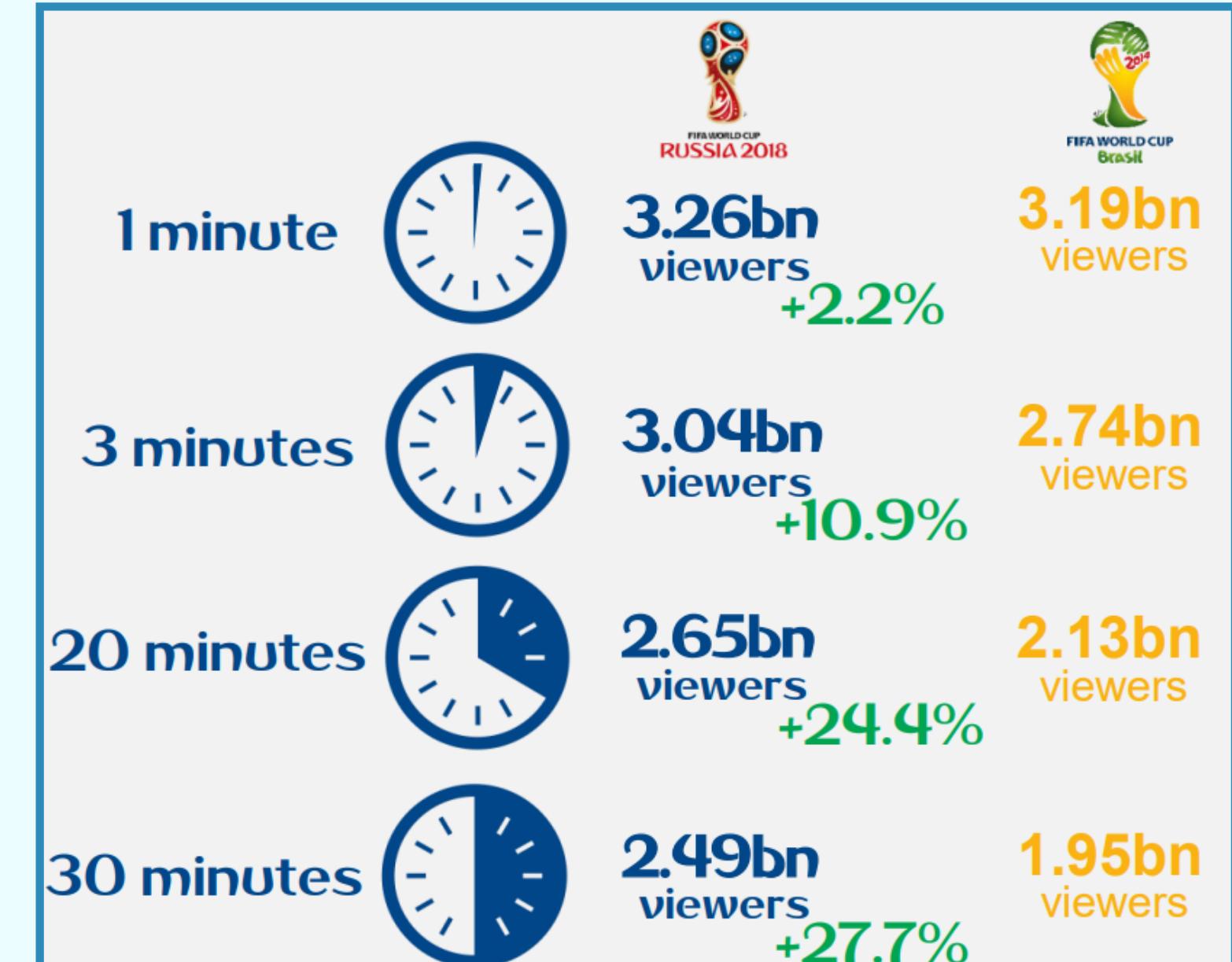
Natural Language Processing (NLP)

65%

15,000 out of 23,000 people across
all 6 regions prefer
Biennial FIFA
IRIS Survey (July 2021)

Russia 2018 exceeded
forecasted \$6.4B
income by:

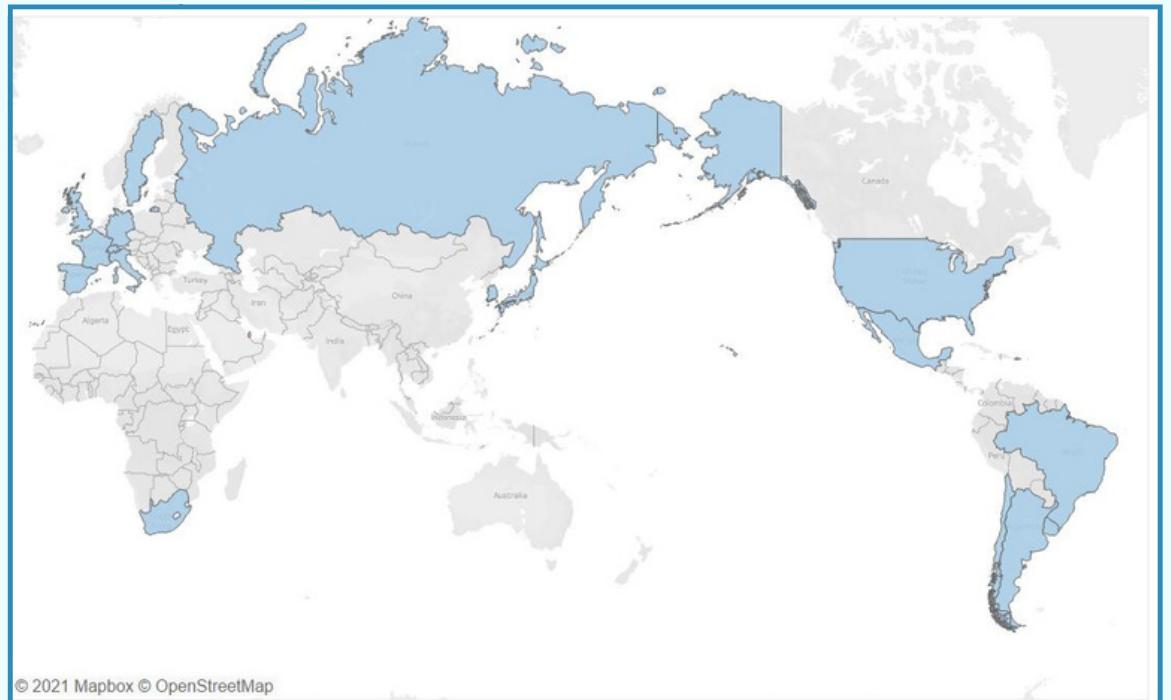
234%



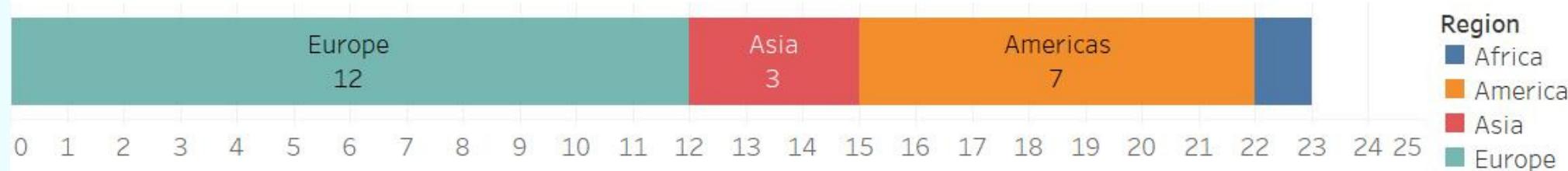
2018 World Cup has upped
2014 in audience engagement

UNIQUE HOST

Third **Asian** country to host,
First time in **Middle East**



No. of times region's countries have hosted



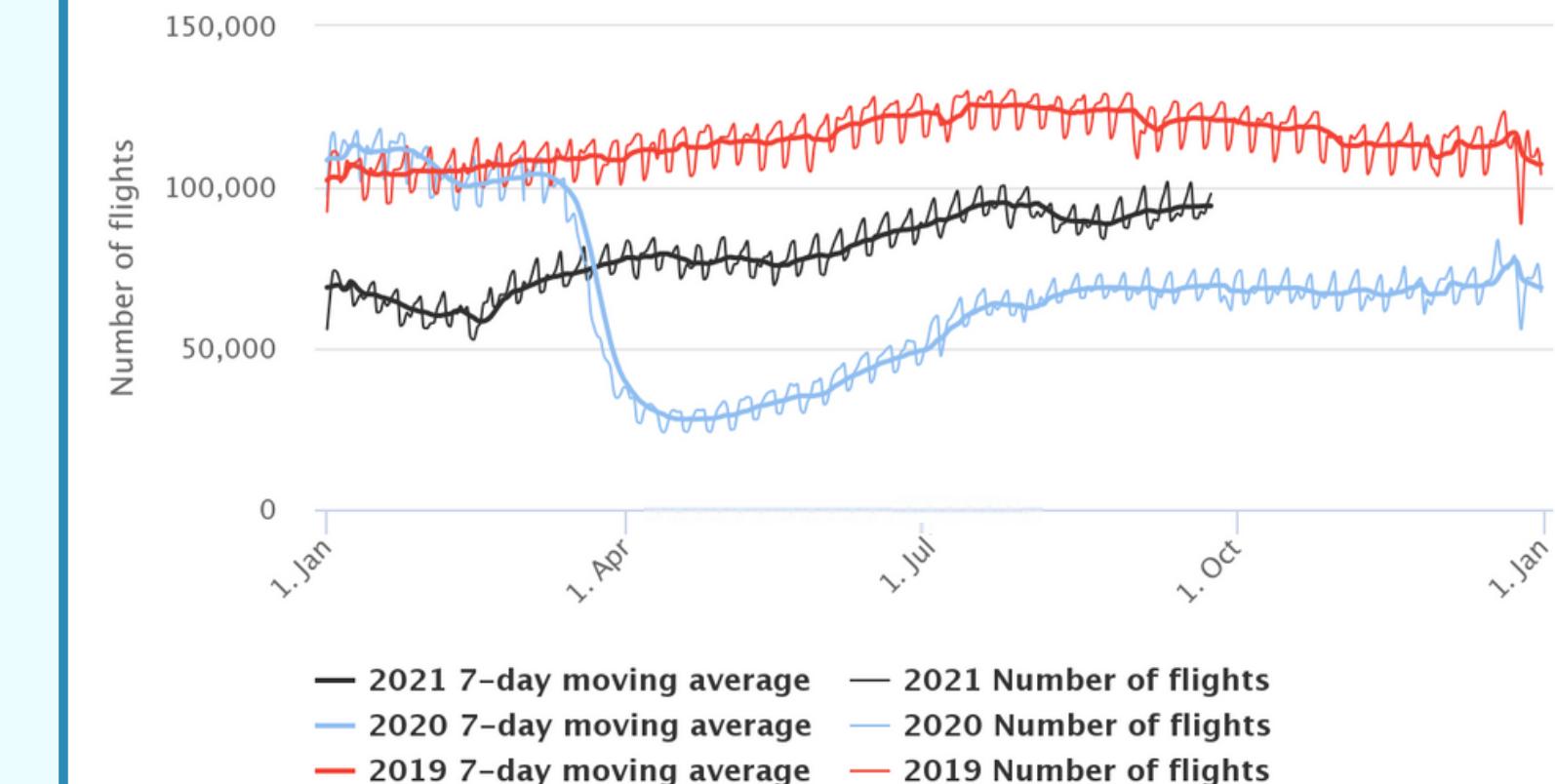
Encouraging Diplomacy of
Middle East & Europe



FIFA WORLD CUP
Qatar2022



Number of commercial flights tracked by Flightradar24, per day (UTC time), 2019 vs 2020 vs 2021



International Flights gradually increasing to 2019 levels



Qatar is opening up for tourism, if adhering to COVID measures

A HISTORICAL EVENT FOR SOCCER



Generational Talents,
MESSI & RONALDO'S
Final Appearance



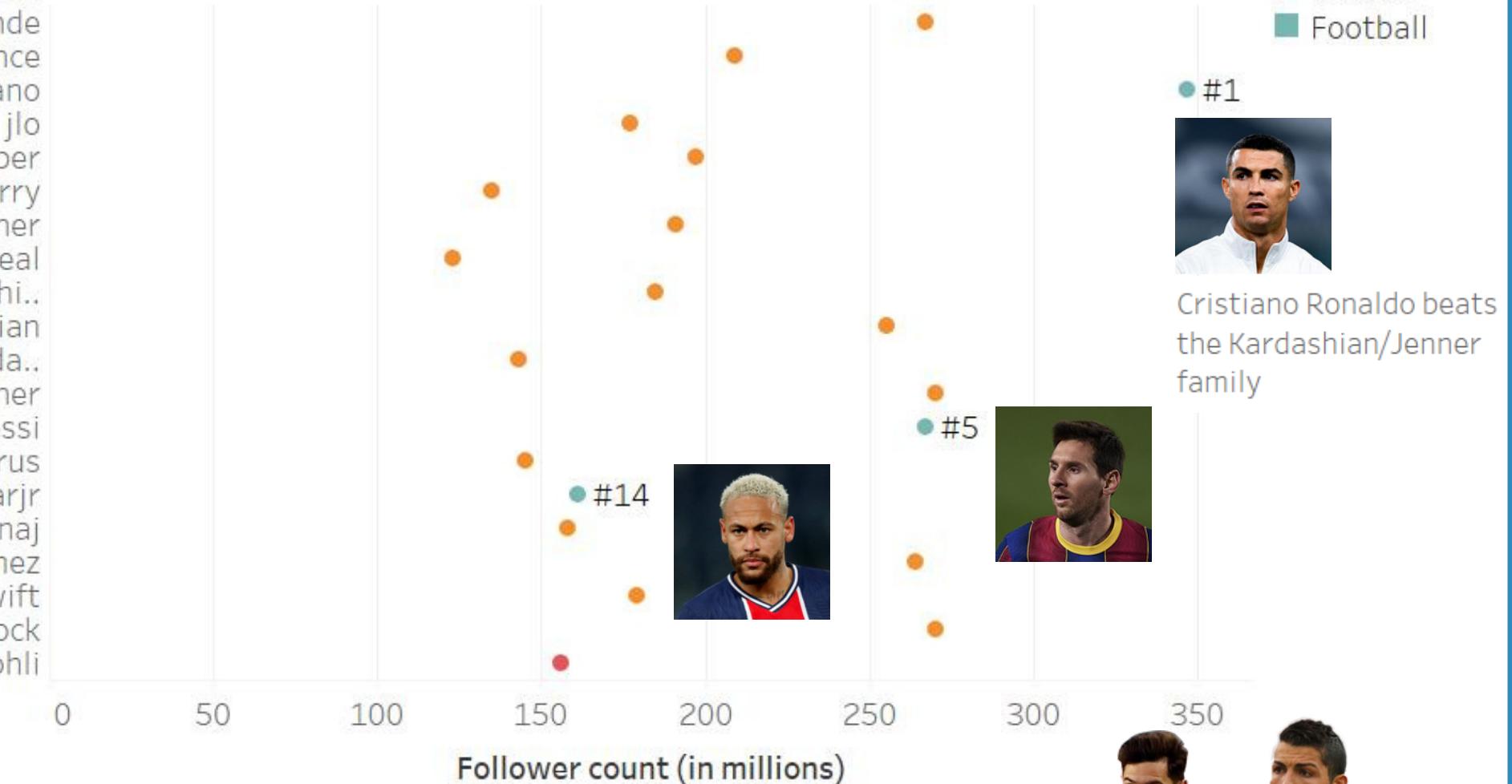
Soccer Players have the highest **follower count** of all sports



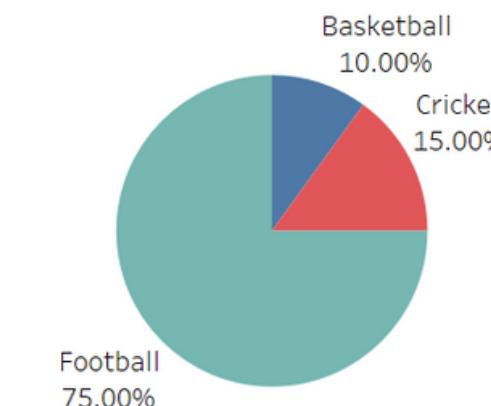
\$400M
*AFP / FIFA feasibility study
July 2021*

Top 20 Individuals on Instagram

Instagram user
arianagrande
beyonce
cristiano
jlo
justinbieber
katyperry
kendalljenner
kevinhart4real
khloekardash..
kimkardashian
kourtneykarda..
kylijenner
leomessi
mileycyrus
neymarjr
nickiminaj
selenagomez
taylorswift
therock
virat.kohli



Top 20 Athletes on Instagram







APPENDIX

TEAM
25

Alex Kim, Steve Jang, Jamie Tay, Janette Lin

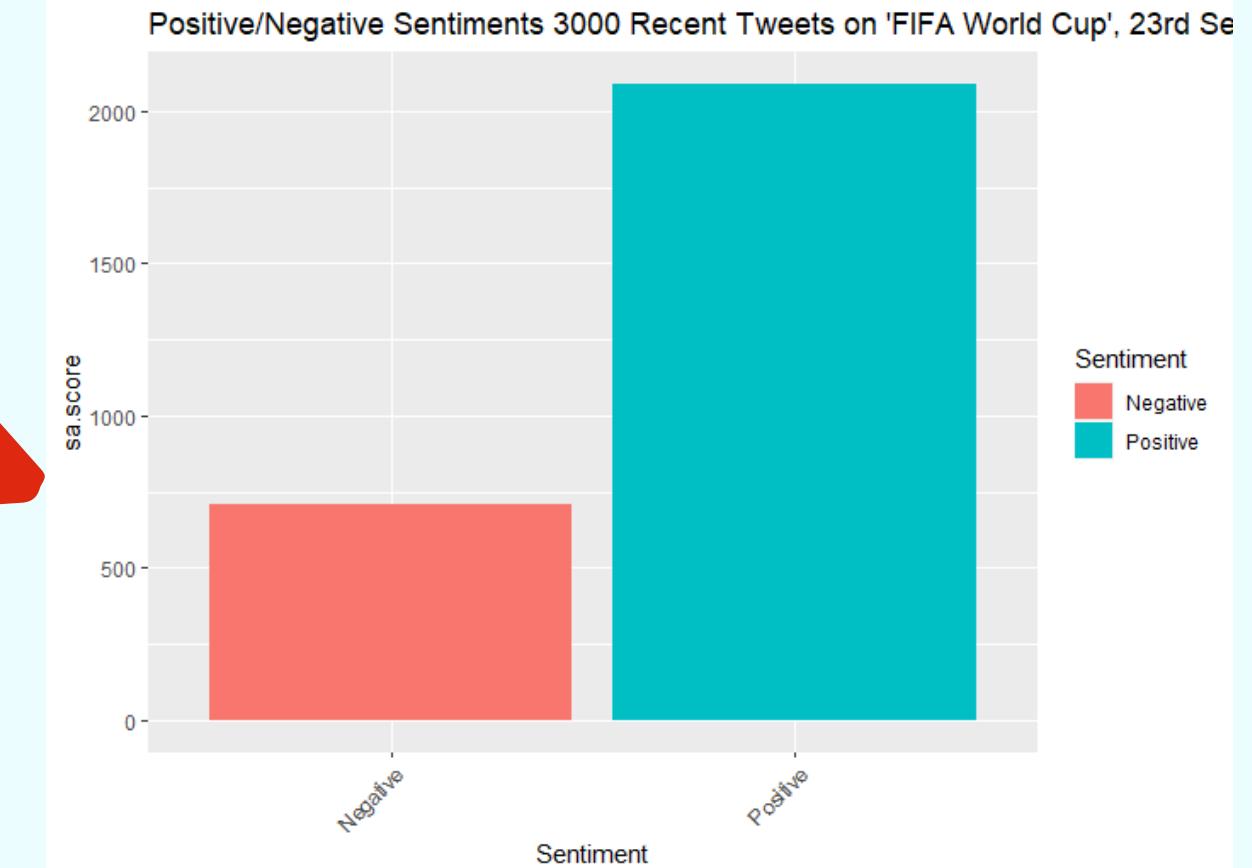
```

# Extracting actual twitter text
# Search for tweets about lockdown, exclude retweets
fifa_tweets <- search_tweets("fifa 2022", n = 18000, include_rts = F, lang = 'en')

# Sentiment analysis(sa): Looks at sentiment of lockdown tweets
sa_value <- get_nrc_sentiment(fifa_tweets$text)
# View sentiment scores
sa_value[1:5, 1:7]
# calculate sum of scores
sa_scores <- colsums(sa_value[,])
# Convert to data frame
sa_scores_df <- data.frame(sa_scores)
# Create sentiment column in the dataframe
sa_scores_final <- cbind(sentiment=row.names(sa_scores_df),
                           sa_scores_df, row.names = NULL)
# Plot the sentiment scores
ggplot(data = sa_scores_final, aes(x = sentiment, y = sa_scores,
                                    fill = sentiment)) +
  geom_bar(stat = "identity") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  ggtitle("Sentiment Analysis of 3000 Recent Tweets on 'FIFA World Cup', 23rd Sept 2021")
| 

# Create new plot with ONLY negative and positive column
# Find data with only negative emotions
neg.df <- sa_scores_final[sa_scores_final$sentiment %in%
  c("anger", "disgust",
    "fear", "sadness", "negative"), ]

```



NLP Analysis of Tweets, R/Twitter API

Head-to-head match history and stats

```

# Example
h2h_dict['Brazil']['Germany']

{'GA_away_losses': 3,
 'GA_away_wins': 2,
 'GA_home_losses': 7,
 'GA_home_wins': 0,
 'GF_away_losses': 2,
 'GF_away_wins': 6,
 'GF_home_losses': 1,
 'GF_home_wins': 0,
 'away_games': 4,
 'away_losses': 1,
 'away_wins': 3,
 'draws': 1,
 'games_played': 6,
 'home_games': 1,
 'home_losses': 1,
 'home_wins': 0,
 'losses': 2,
 'wins': 3}

```



North & Central America (CONCACAF)		South America (CONMEBOL)		Africa (AFC)	
United States		Argentina		South Africa	
Mexico		Brazil		Senegal	
Canada		Uruguay		Algeria	
Costa Rica		Colombia		Tunisia	
		Peru		Ivory Coast	
Asia (AFC)		Europe (UEFA)			
Iran		Portugal		Belgium	
South Korea		Spain		Denmark	
Australia		Italy		Netherlands	
Japan		France		Croatia	
Qatar				Sweden	
				Ukraine	
				England	