

Info 450 Data science project

By: Michael Otalora



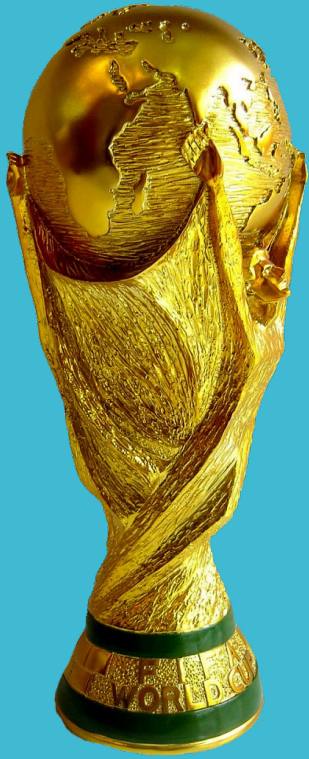
FIFA WORLD CUP
Qatar 2022

The Problem Statement

- I've been hired to research the history of the previous FIFA world cups that have passed. More specifically what I am working on in this project is to predict this year's world cup revenue by using the total attendance of the tournament history to predict if it'll bring more revenue for the next host.



Benefits



- The benefits of this project is to see how much increase of profit can be funded for the upcoming world cups to see if its worth hosting a world cup.

Methodology

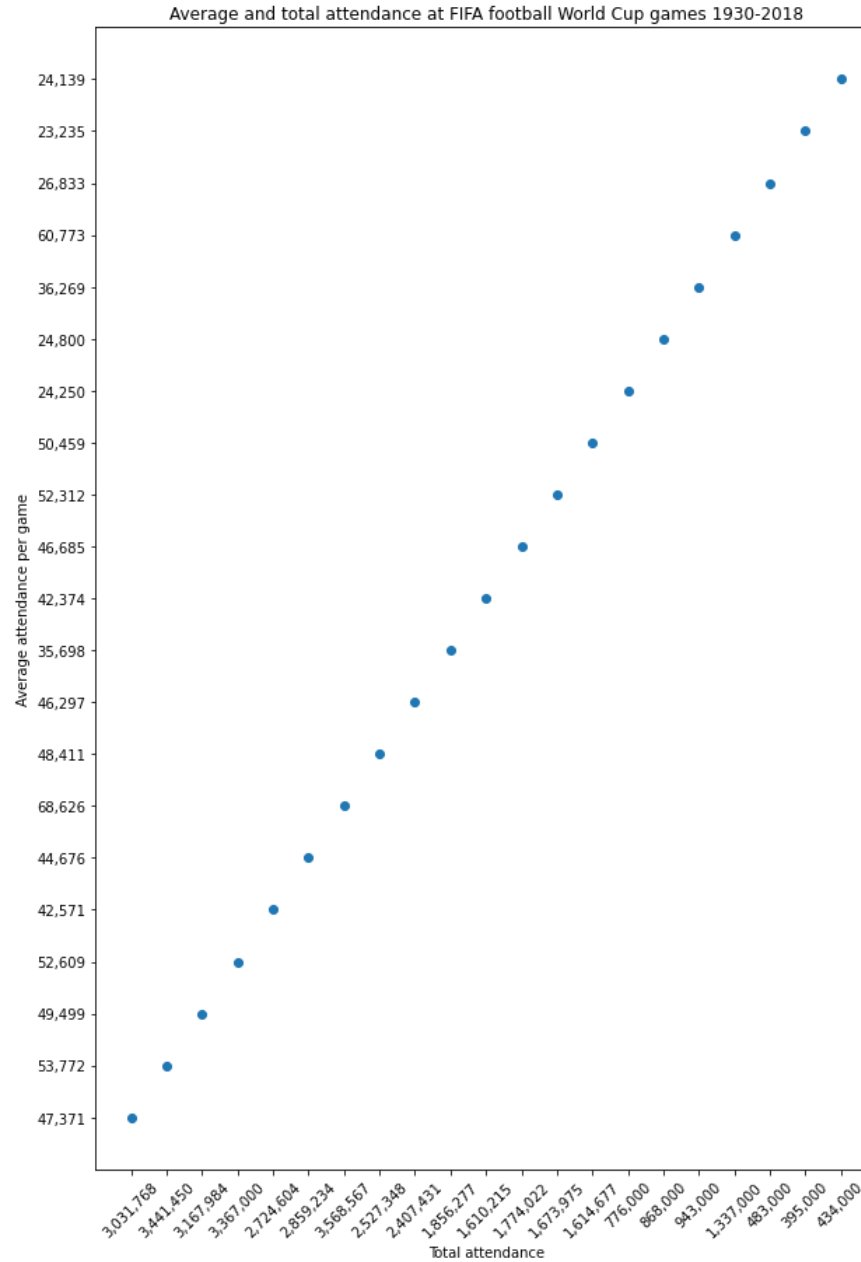
- Started off by finding the data
- Data visualizations
- Group by
- Two machine learning models



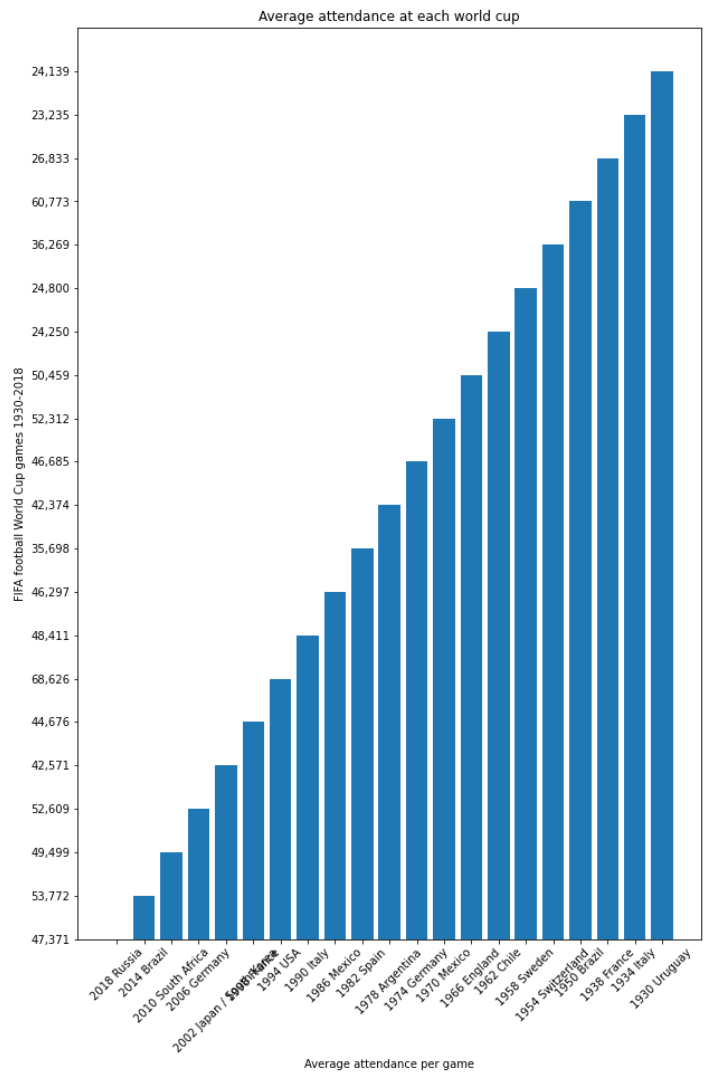
Data

	Total attendance	Average attendance per game
FIFA football World Cup games 1930-2018		
2018 Russia	3,031,768	47,371
2014 Brazil	3,441,450	53,772
2010 South Africa	3,167,984	49,499
2006 Germany	3,367,000	52,609
2002 Japan / South Korea	2,724,604	42,571
1998 France	2,859,234	44,676
1994 USA	3,568,567	68,626
1990 Italy	2,527,348	48,411
1986 Mexico	2,407,431	46,297
1982 Spain	1,856,277	35,698

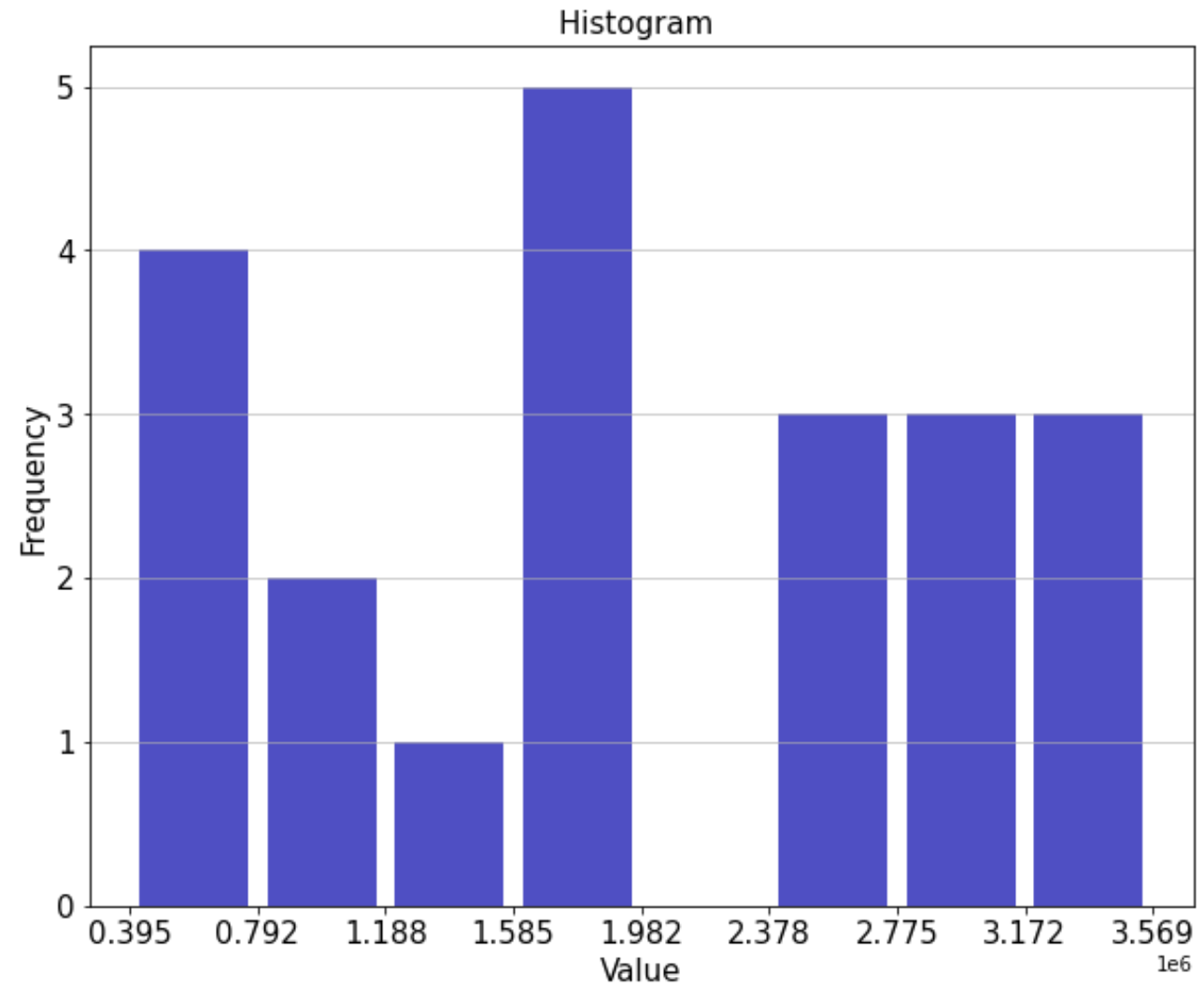
Scatterplot



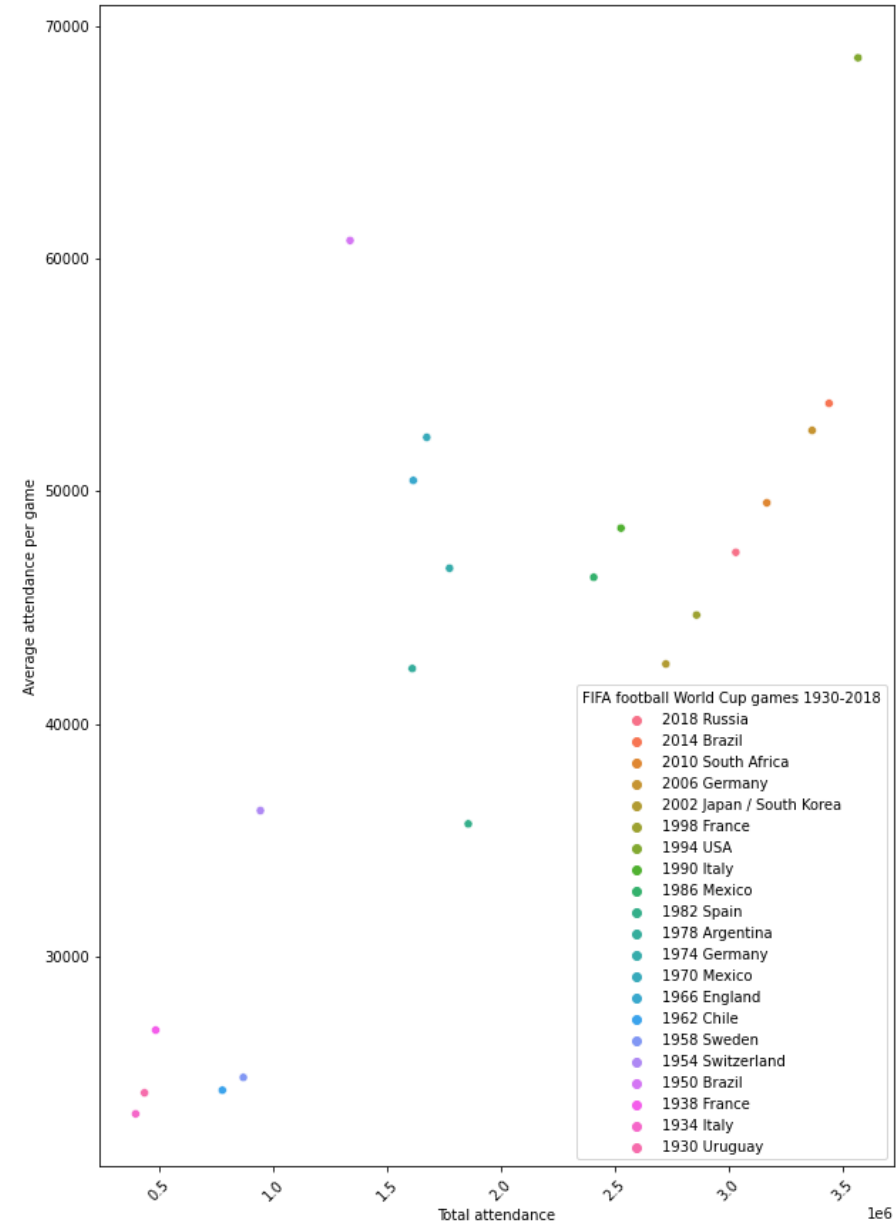
Bar grapgh



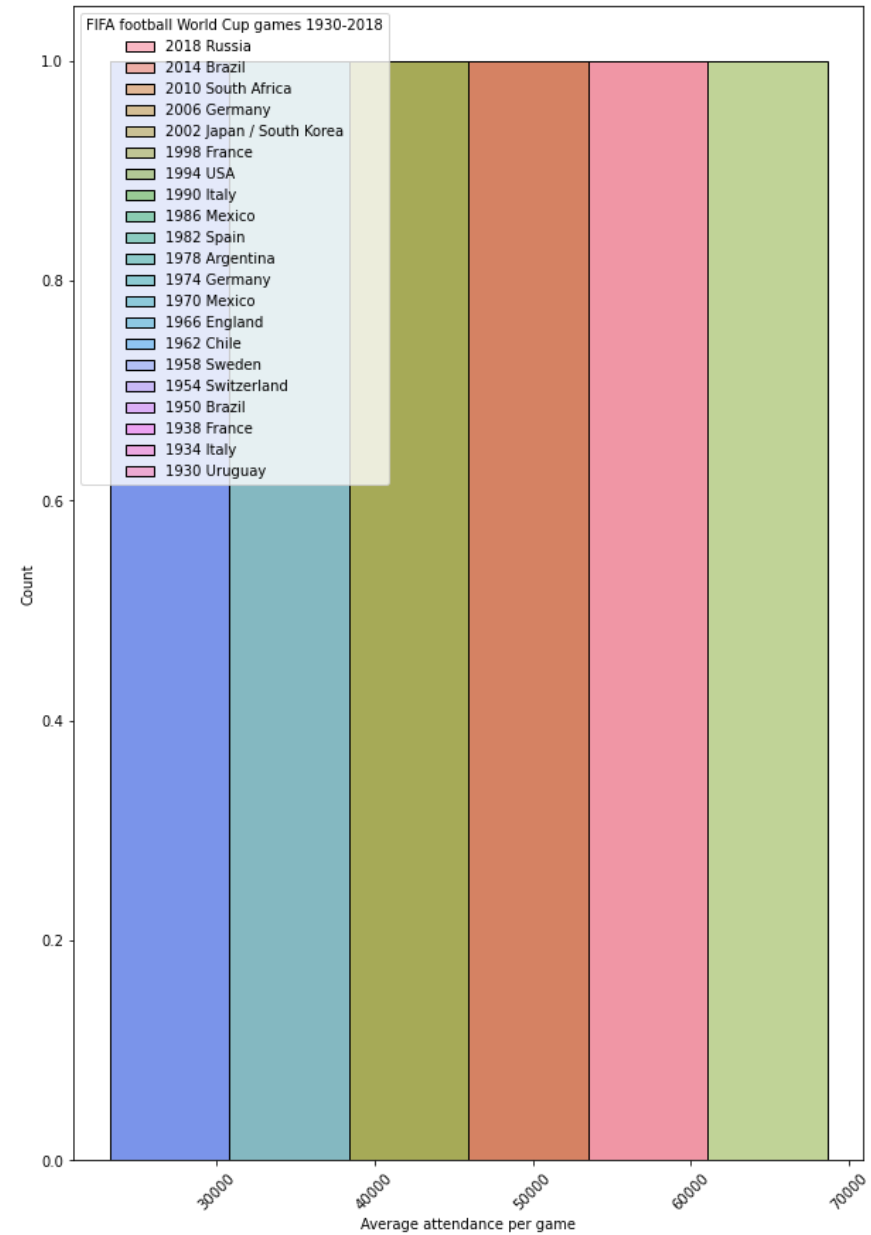
Histogram



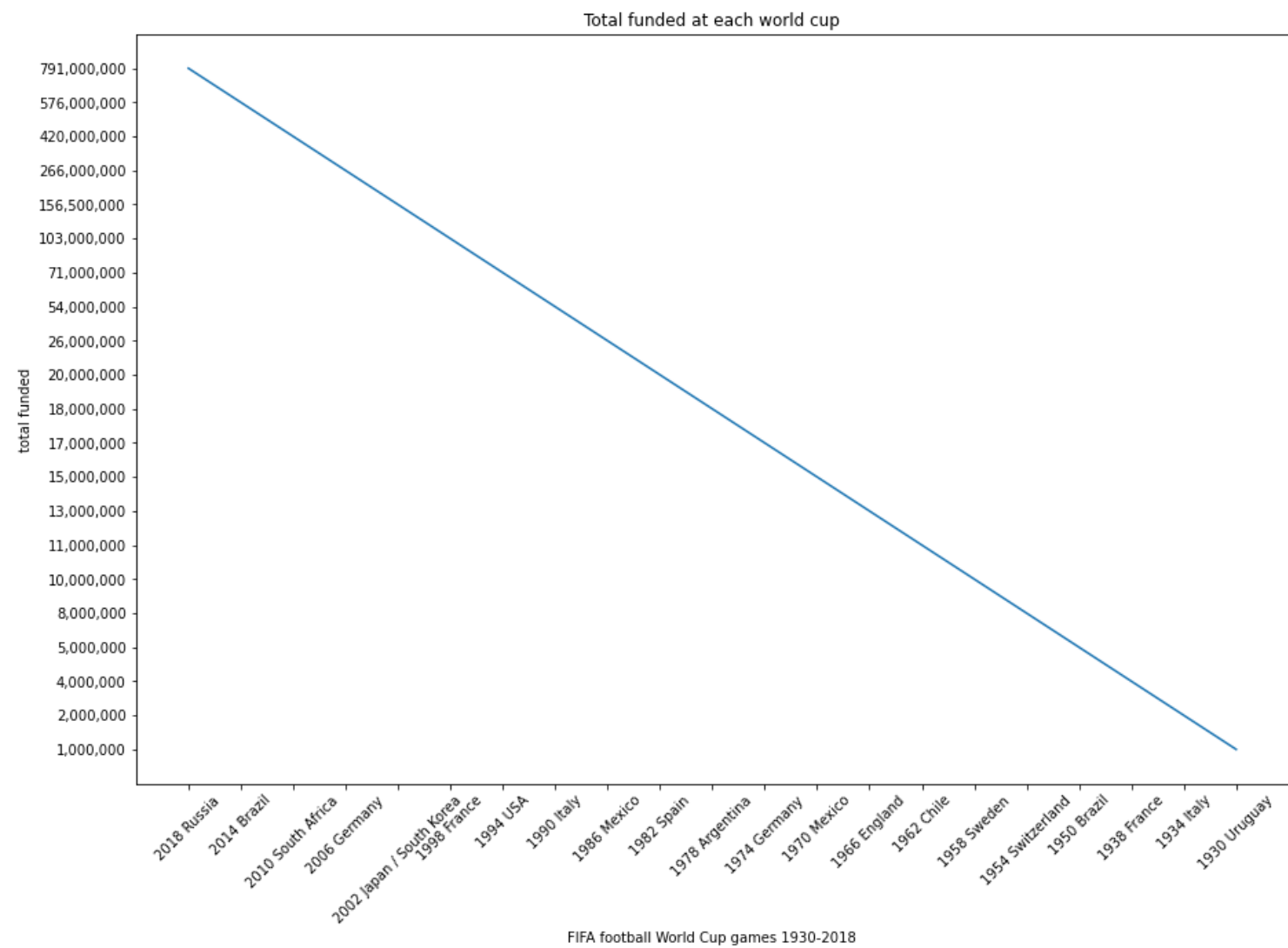
Dot plot



bar chart



Line chart



Linear regression models

Used data from total funded to predict winning team prize money

```
print(regressor.score(x_test, y_test))
```

0.8768336080965257

Used data from the winning team prize money to predict total funded

```
print(regressor.score(x_test, y_test))
```

0.7946871256552213

Recommendation

- Based on data and project it will be good to host the world cup to have much more revenue then previous tournaments.
- There is an increase in revenue every 4 years

Future Work

- This project should be taken a look at again in the future
- Can be improved and not

Enjoy Qatar!

