



TASK

Exploratory Data Analysis on the Forbes Richest Athlete Dataset

[Visit our website](#)

Introduction

I was given the Forbes Richest Athletes dataset to analyse. This dataset has details about the top earning athletes in the world. Most of the data in the csv file is numeric in nature. Using various data analysis tools I have explored this dataset and I am presenting the findings in this report.

This report includes findings on the different nationalities of the athletes and how well they earned over the years.

DATA CLEANING

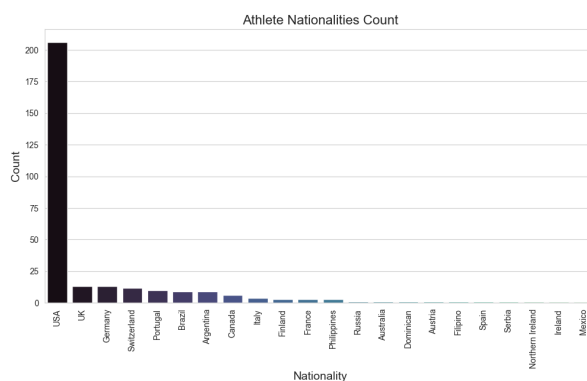
The previous year rankings column had a lot of missing data so I dropped the column altogether. The Sport column had words that meant the same thing but are spelt differently. I have corrected the different spellings.

MISSING DATA

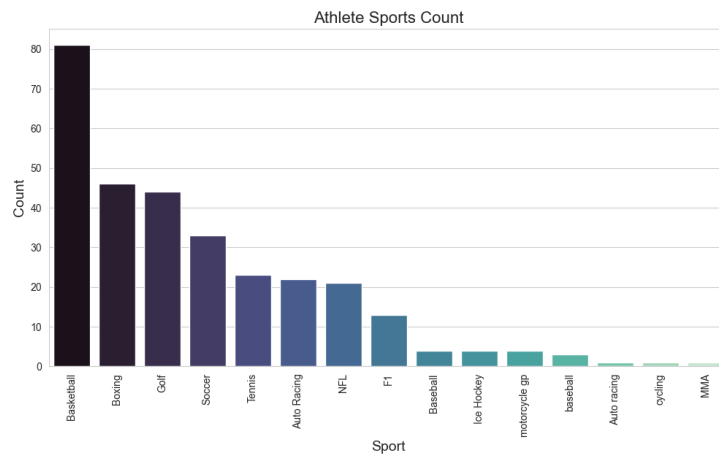
It was removed using pandas

DATA STORIES AND VISUALISATIONS

To get an idea of the dataset I explored the different nationalities of the athletes that were included in the dataset. According to the data, most of the athletes in the dataset originate from the USA, UK and Germany.

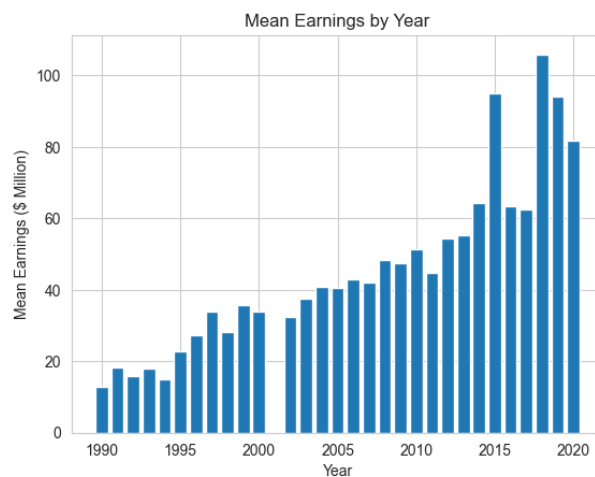


I also explored which sports these athletes participate in. The dataset shows that most of the athletes included in this dataset participate in basketball, soccer and golf.

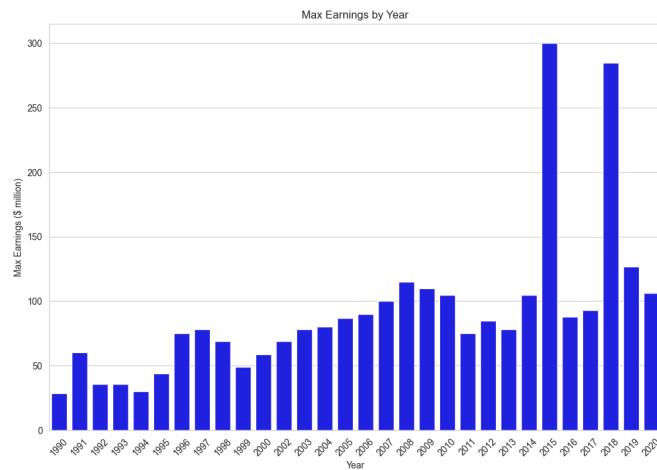


Once this was established I went on to explore the earnings data. I explored the average earnings in the dataset. The information in the dataset states that the average earnings is \$45.52mil.

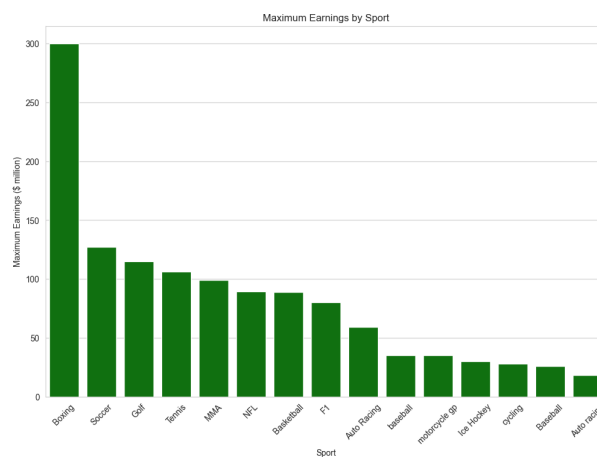
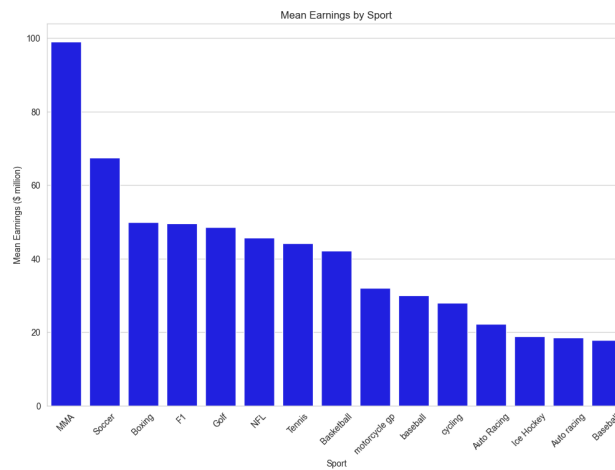
I then went on to explore the highest average earnings on a yearly basis. The data suggests that the highest average earnings was in 2018 with an average earning of \$105 mil.



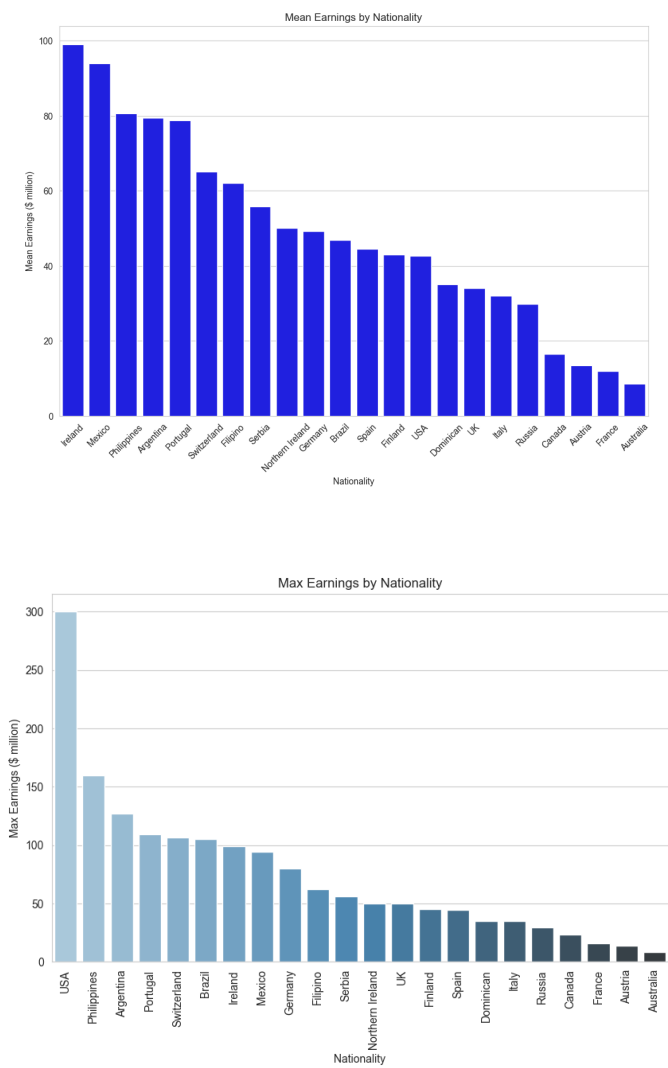
However, the highest earnings received in this dataset was in 2015 where an athlete received an earnings of \$300mil.



I then went on to explore the earnings by the sport an athlete participates in. The top 3 sports with the highest average earnings are: MMA, soccer and boxing. However, the athlete with the highest earnings in this dataset is a boxing athlete.



Lastly, I went on to explore the earnings for every nationality of the athlete in the dataset. On average the highest paid athletes originate from Ireland, Mexico or the Philippines but the highest earning athlete originates from the US.



THIS REPORT WAS WRITTEN BY :Pule Molefe

