Hyperiondev

TASK

Exploratory Data Analysis on the Forbes List Top 10 Highest Paid Athletes from 1990 - 2020 Dataset

Visit our website

Introduction

Summary of the data set

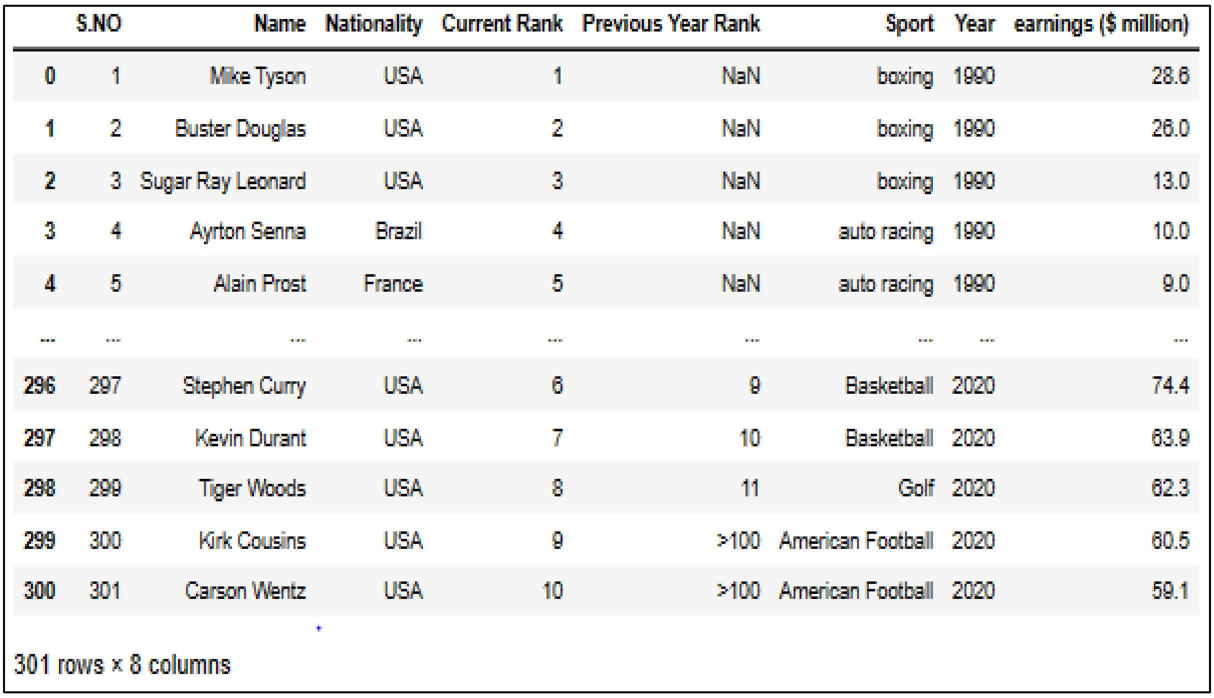
This dataset contains the top 10 highest paid athletes according to Forbes between the years 1990 – 2020.

IMPORTANT NOTE:

In 2002, Forbes changed the reporting period from the full financial calendar year to June-to-June, thus there are not records for 2001.

The data contains the following 8 columns of data:

- **S.NO** – [Integer] [Continuous] Unique number to identify each entry in the dataset.
- **Name** – [String] [Categorical] Name of the athlete.
- **Nationality** – [String] [Categorial] Country the athlete represents legally.
- **Current Rank** – [Integer] [Categorical] Athletes rank out of 10 for that year.
- **Previous Year Rank** – [Integer] [Continuous] Previous ranking of the athlete, if applicable, in relation to the entire Forbes list rankings.
- **Sport** – [String] [Categorical] The sport the athlete participates in as a profession.
- **Year** – [Integer] [Categorical] The year of the ranking.
- **Earnings** – [Float] [Continuous] The total earnings of the athlete in $ millions.



	S.NO	Name	Nationality	Current Rank	Previous Year Rank	Sport	Year	earnings ($ million)
0	1	Mike Tyson	USA	1	NaN	boxing	1990	28.6
1	2	Buster Douglas	USA	2	NaN	boxing	1990	26.0
2	3	Sugar Ray Leonard	USA	3	NaN	boxing	1990	13.0
3	4	Ayrton Senna	Brazil	4	NaN	auto racing	1990	10.0
4	5	Alain Prost	France	5	NaN	auto racing	1990	9.0
...	...	...	...	...	...	...	...	...
296	297	Stephen Curry	USA	6	9	Basketball	2020	74.4
297	298	Kevin Durant	USA	7	10	Basketball	2020	63.9
298	299	Tiger Woods	USA	8	11	Golf	2020	62.3
299	300	Kirk Cousins	USA	9	>100	American Football	2020	60.5
300	301	Carson Wentz	USA	10	>100	American Football	2020	59.1

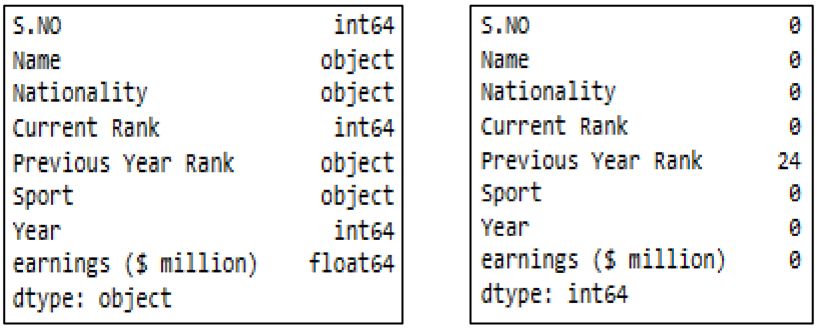
301 rows × 8 columns

Figure 1: Snapshot of the raw dataset showing the first and last 5 entries.

MISSING DATA

Summary of how missing data was identified and dealt with

Initially the data types of the columns where checked to ensure they were of the correct expected formats. Column ‘Previous Year Rank’ had the datatype ‘object’, however this was expected to the data type ‘int64’. Thus, to explore this further the entire dataset was checked for NaN values, in which there were 24 missing values in the ‘Previous year Rank’ column.



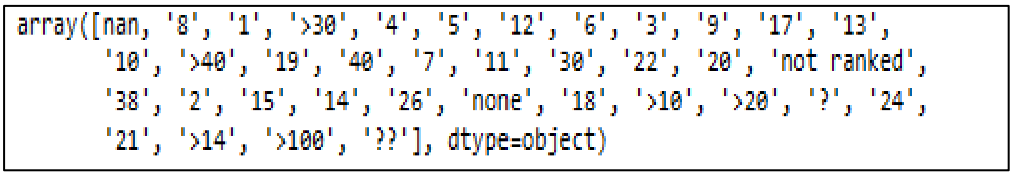
```
S.NO                       int64
Name                      object
Nationality               object
Current Rank               int64
Previous Year Rank        object
Sport                     object
Year                       int64
earnings ($ million)     float64
dtype: object
```

```
S.NO                      0
Name                      0
Nationality               0
Current Rank              0
Previous Year Rank       24
Sport                     0
Year                      0
earnings ($ million)      0
dtype: int64
```

Figure 2: Datatypes for each column in the data set (left) and count of NaN values in the dataset for each column (right)

It appears that the data has not been pre-processed to a degree where all missing values have been input as NaN. As missing values can take many forms such as a question mark (?), a zero (0), a minus one (-1) or a blank cell, each columns unique values where explored.

All other columns excluding the ‘Previous Year Rank’ had the appropriate unique values, whilst the ‘Previous Year Rank’ column had the following entries throughout the column:



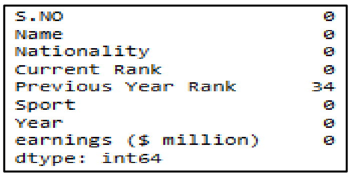
```
array([nan, '8', '1', '>30', '4', '5', '12', '6', '3', '9', '17', '13',
       '10', '>40', '19', '40', '7', '11', '30', '22', '20', 'not ranked',
       '38', '2', '15', '14', '26', 'none', '18', '>10', '>20', '?', '24',
       '21', '>14', '>100', '??'], dtype=object)
```

Figure 3: Array of all the unique entries in the 'Previous Year Rank' column

From Figure 3 missing values are appering as the following values:

- NaN
- ?
- ??
- not ranked
- none

By exploring these values, it was apporiate to conclude that any entries with the above entries could be deemed as an missing value and thus all ‘Previous Year Rank’ entries with the above missing value types where set to NaN. The missing value cout was checked once more for a total of 34 missing values in the entire dataset accompanied with an visualisation of the missing values:



```
S.NO                        0
Name                        0
Nationality                 0
Current Rank                0
Previous Year Rank         34
Sport                       0
Year                        0
earnings ($ million)        0
dtype: int64
```

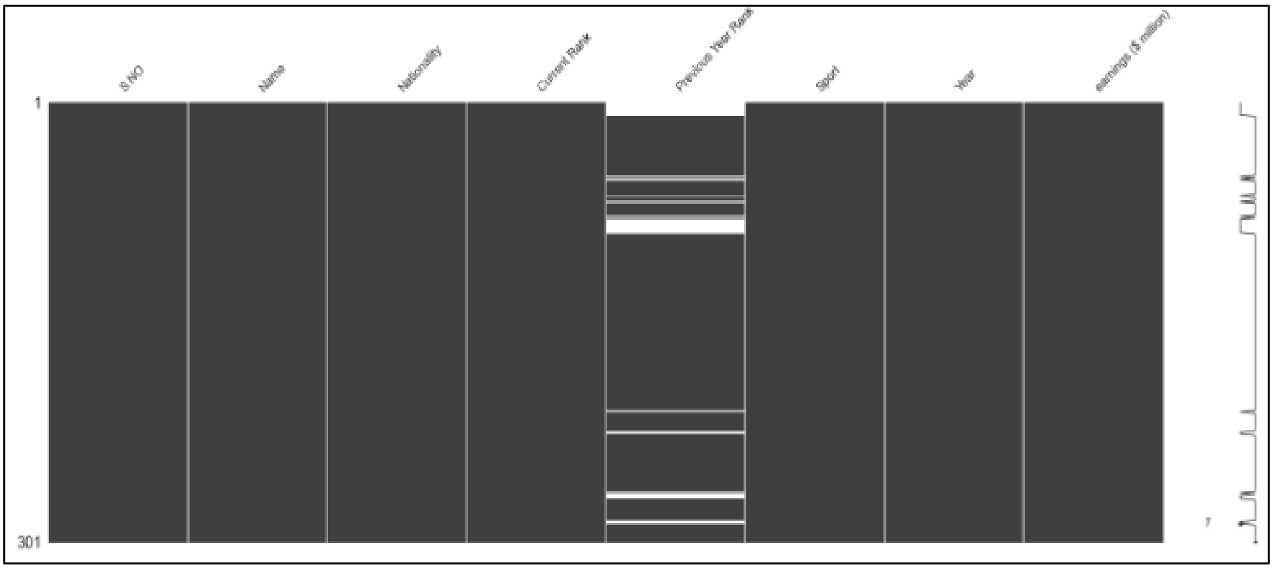
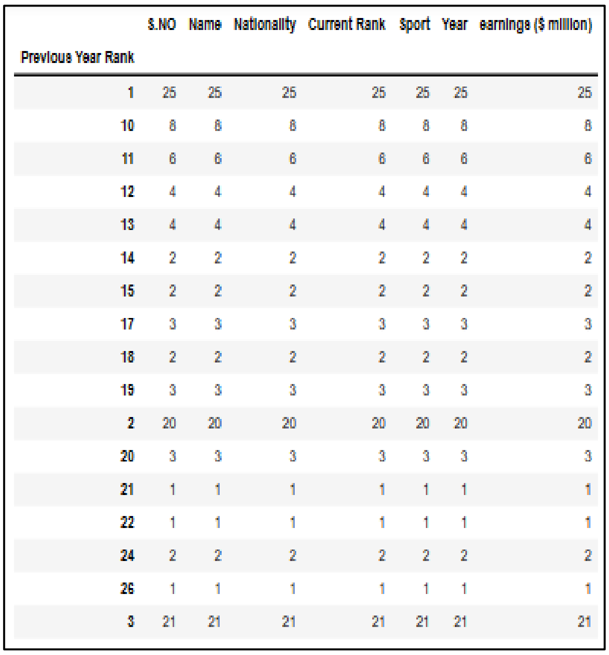
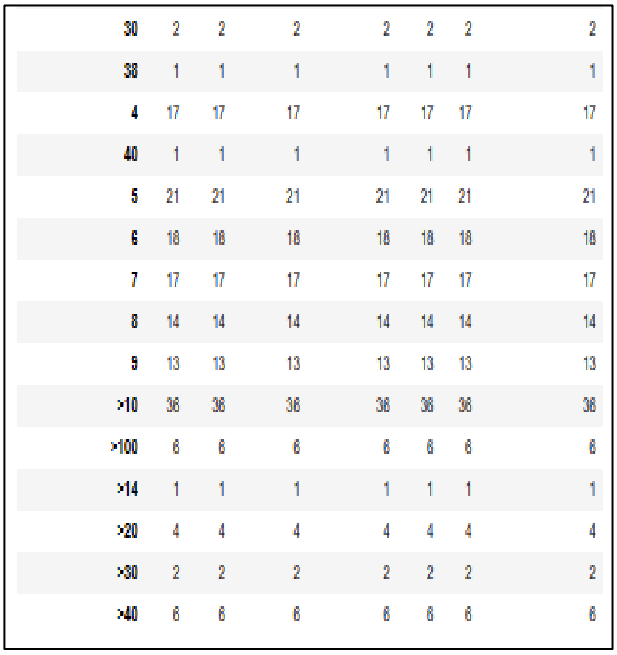


Figure 4: Missing value count of NaN values in the entire dataset (top) and a missingno matrix to demonstrate where sparse data in located in the dataset (bottom)

To determine whether the data is complete enough to peform data analysis, the total percentage of missing values was calculated to be 1.4%. This was deemed to be within an acceptable range and thus data analsis could be carried out on the dataset.

From the ‘Previous Year Column’ the percentage of missing values was 11.3% which was higher then desired. Thus, to explore these missing values further all unique values in the ‘Previous Year Rank’ column was explored further.

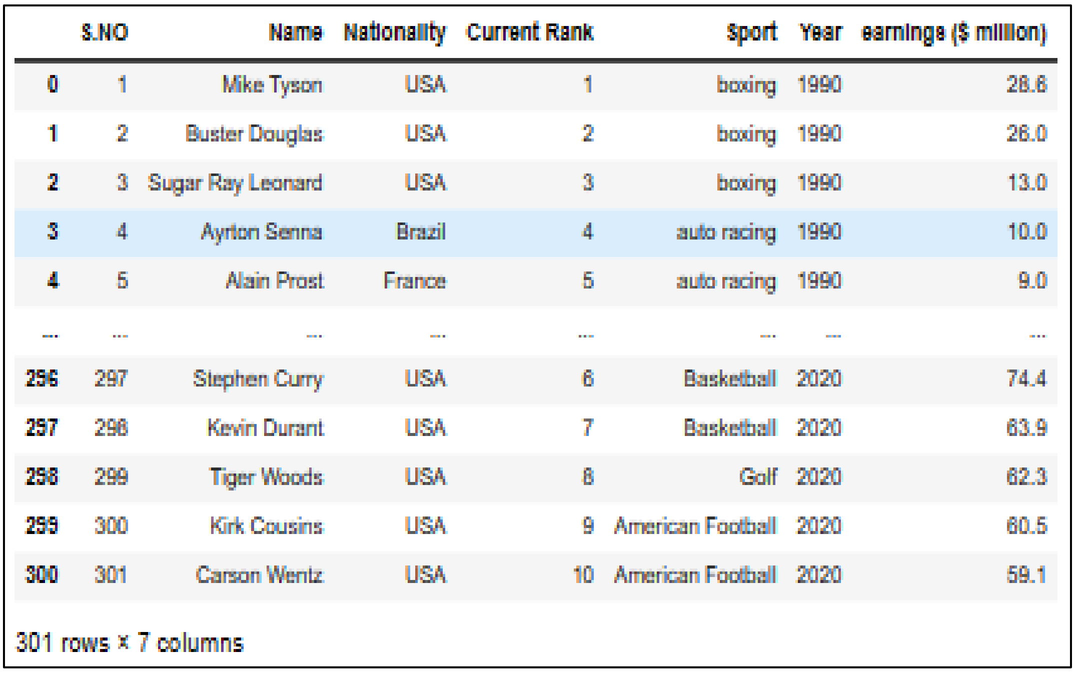




	S.NO	Name	Nationality	Current Rank	Sport	Year	earnings ($ million)
Previous Year Rank							
1	25	25	25	25	25	25	25
10	8	8	8	8	8	8	8
11	6	6	6	6	6	6	6
12	4	4	4	4	4	4	4
13	4	4	4	4	4	4	4
14	2	2	2	2	2	2	2
15	2	2	2	2	2	2	2
17	3	3	3	3	3	3	3
18	2	2	2	2	2	2	2
19	3	3	3	3	3	3	3
2	20	20	20	20	20	20	20
20	3	3	3	3	3	3	3
21	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1
24	2	2	2	2	2	2	2
26	1	1	1	1	1	1	1
3	21	21	21	21	21	21	21
30	2	2	2	2	2	2	2
38	1	1	1	1	1	1	1
4	17	17	17	17	17	17	17
40	1	1	1	1	1	1	1
5	21	21	21	21	21	21	21
6	18	18	18	18	18	18	18
7	17	17	17	17	17	17	17
8	14	14	14	14	14	14	14
9	13	13	13	13	13	13	13
>10	36	36	36	36	36	36	36
>100	6	6	6	6	6	6	6
>14	1	1	1	1	1	1	1
>20	4	4	4	4	4	4	4
>30	2	2	2	2	2	2	2
>40	6	6	6	6	6	6	6

Figure 5: All unique values in the 'Previous Year Rank' column with data on how many entries appeared in each rank value.

From Figure 5, there are 32 different values used to describe an athlete's 'Previous Year Rank'. There are 6 values that are pre-fix with the greater than operator (>) and therefore are too ambiguous to pin-point an athlete's actual previous year rank. This would thus increasing the missing values from 34 to 89 as the 6 ambigous values contain a total of 55 athlete entries and therefore increasing the percentage of missing values in the 'Previous Year Rank' column to 29.5%. This value was deemed to be far to high, and setting these values based on a mean for example would potential introduce too much bias, and as there are no significant questions to be explored using this particular column this column was removed from the dataset.



	S.NO	Name	Nationality	Current Rank	Sport	Year	earnings ($ million)
0	1	Mike Tyson	USA	1	boxing	1990	28.6
1	2	Buster Douglas	USA	2	boxing	1990	26.0
2	3	Sugar Ray Leonard	USA	3	boxing	1990	13.0
3	4	Ayrton Senna	Brazil	4	auto racing	1990	10.0
4	5	Alain Prost	France	5	auto racing	1990	9.0
...	...	...	...	...	...	...	...
296	297	Stephen Curry	USA	6	Basketball	2020	74.4
297	298	Kevin Durant	USA	7	Basketball	2020	63.9
298	299	Tiger Woods	USA	8	Golf	2020	62.3
299	300	Kirk Cousins	USA	9	American Football	2020	60.5
300	301	Carson Wentz	USA	10	American Football	2020	59.1

301 rows × 7 columns

Figure 6: Preview of the first and last 5 rows of the semi-processed dataset

DATA CLEANING

Summary of the methods and visualisations done during the data cleaning

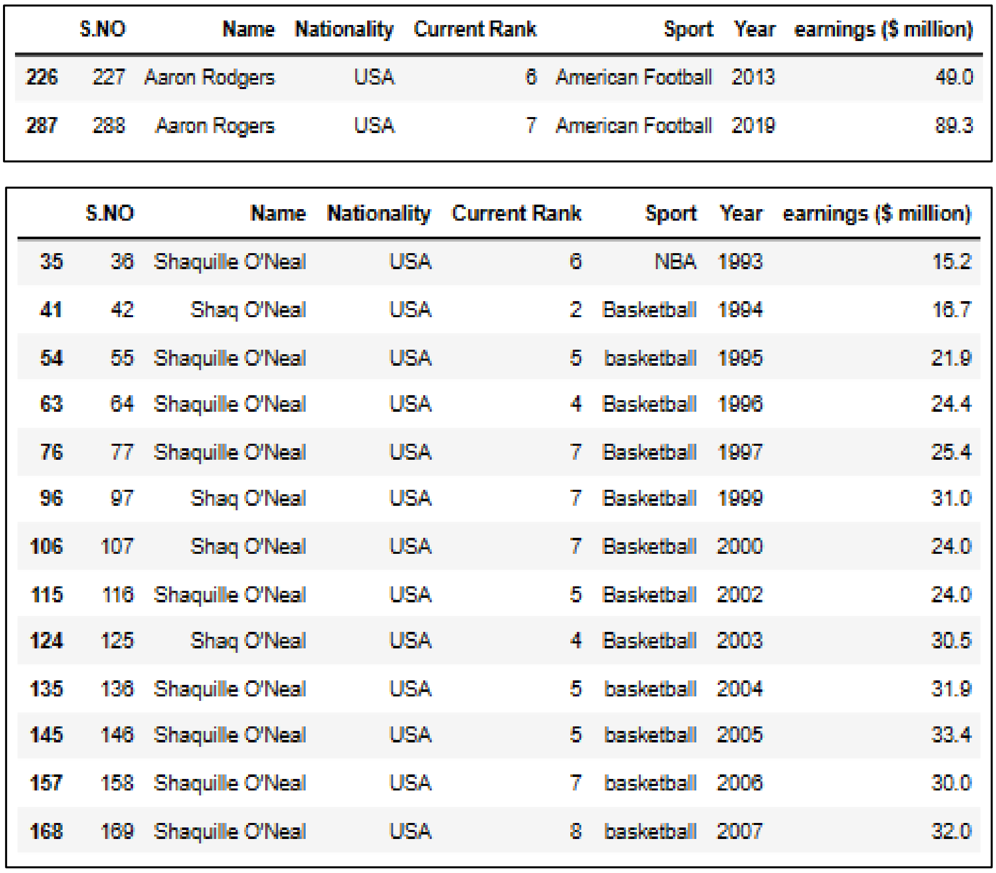
The check the dataset for inconsistent data entries the 'Name', 'Nationality' and 'Sport' columns which all have the 'object' datatype where checked for duplicates.

'Name' column

The 'Name' columns unique values are sorted and counted, in which it contained 82 different athletes. By examining the sorted names of athletes two entries appeared to have been entered inconsistently which were:

1. 'Aaron Rodgers' and 'Aaron Rogers'
2. 'Shaq O'Neal' and 'Shaquille O'Neal'

All entries for each of the two athletes were extracted from the dataset to check how many of the inconsistent data entries have occurred:



	S.NO	Name	Nationality	Current Rank	Sport	Year	earnings ($ million)
226	227	Aaron Rodgers	USA	6	American Football	2013	49.0
287	288	Aaron Rogers	USA	7	American Football	2019	89.3

	S.NO	Name	Nationality	Current Rank	Sport	Year	earnings ($ million)
35	36	Shaquille O'Neal	USA	6	NBA	1993	15.2
41	42	Shaq O'Neal	USA	2	Basketball	1994	16.7
54	55	Shaquille O'Neal	USA	5	basketball	1995	21.9
63	64	Shaquille O'Neal	USA	4	Basketball	1996	24.4
76	77	Shaquille O'Neal	USA	7	Basketball	1997	25.4
96	97	Shaq O'Neal	USA	7	Basketball	1999	31.0
106	107	Shaq O'Neal	USA	7	Basketball	2000	24.0
115	116	Shaquille O'Neal	USA	5	Basketball	2002	24.0
124	125	Shaq O'Neal	USA	4	Basketball	2003	30.5
135	136	Shaquille O'Neal	USA	5	basketball	2004	31.9
145	146	Shaquille O'Neal	USA	5	basketball	2005	33.4
157	158	Shaquille O'Neal	USA	7	basketball	2006	30.0
168	169	Shaquille O'Neal	USA	8	basketball	2007	32.0

Figure 7: All rows contain the inconsistent name entries of the first athlete (top) and all rows containing the second inconsistent data of the name of the second athlete (bottom)

From further research ‘Aaron Rogers’ and ‘Aaron Rodgers’ is the same person and there must have been an spelling error when the athlete was entered into the dataset, thus all entries where set to ‘Aaron Rodgers’.

All entries for ‘Shaq O’Neal’ where set to the athlete full government name of ‘Shaquille O’Neal’, thus reducing the number of athletes in the dataset to 80.

‘Nationality’ column

Due to the title given to the column there has been some inconsistent data entry into this column, although most have entered the Country the athlete represents some have entered the athletes nationality instead. As a result there are 22 unique entries in this particular column.

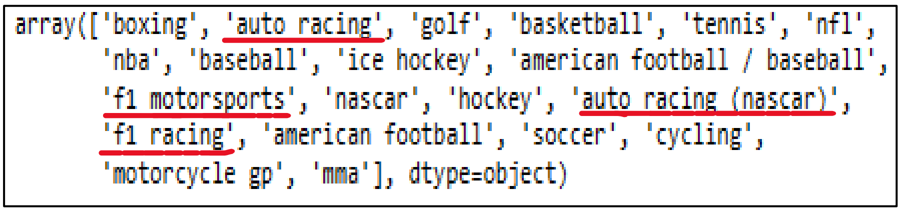
The following Nationality values will be adjusted:

1. Filipino -to- Philippines
2. Dominican -to- USA
   - As Alex Rodriguez an Baseball player represents the USA when competing.

Therefore, reducing the number of unique entries in the Nationality column from 22 to 20.

‘Sport’ column

All values in the ‘Sport’ column where standarised by covering them all to lower case which reduced the number of unique entries from 29 to 20.

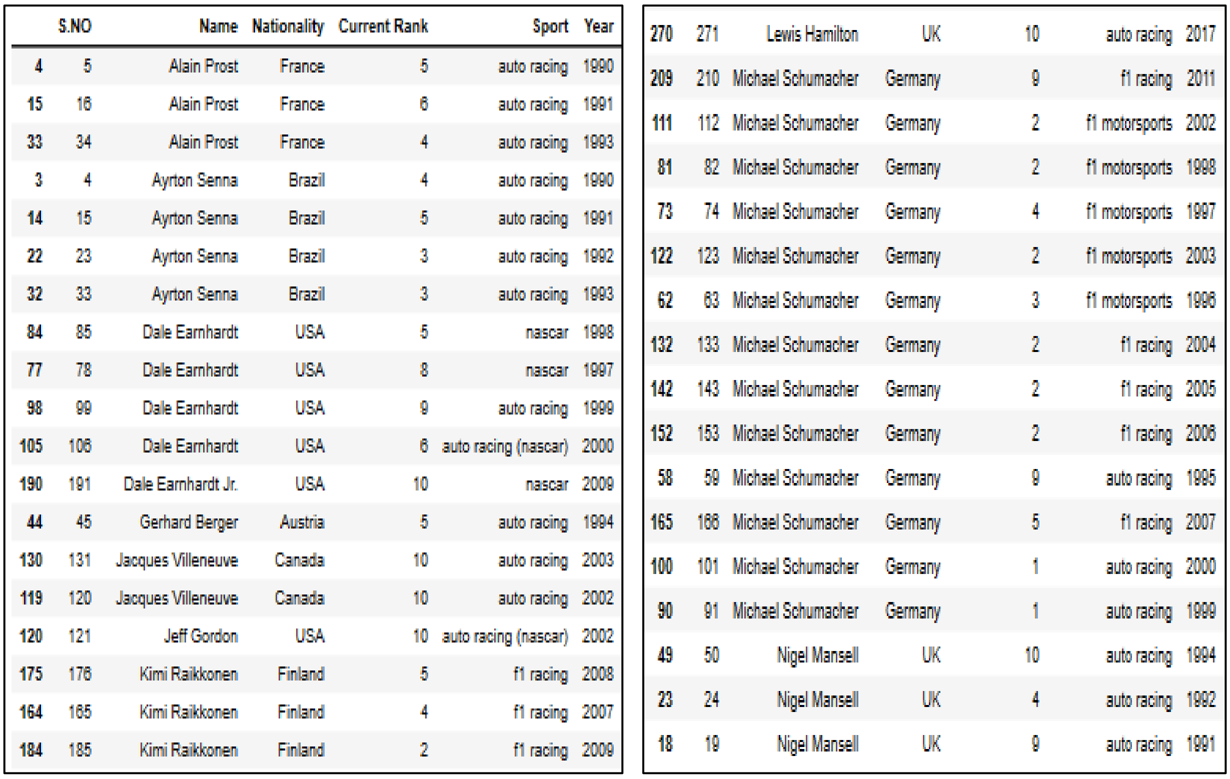


```
array(['boxing', 'auto racing', 'golf', 'basketball', 'tennis', 'nfl',
       'nba', 'baseball', 'ice hockey', 'american football / baseball',
       'f1 motorsports', 'nascar', 'hockey', 'auto racing (nascar)',
       'f1 racing', 'american football', 'soccer', 'cycling',
       'motorcycle gp', 'mma'], dtype=object)
```

Figure 8: All unique entries from the 'Sport' column converted to lowercase.

From Figure 8, we can see that different names have been assigned for athletes in motorsports, thus each athlete where their ‘Sport’ was any of the following where selecting into a temporay dataframe ordererd by the athlete’s name (Figure 9) so further reasearch could be done to add clarity to which sport they where athletes of:

- ‘auto racing’
- ‘nascar’
- ‘auto racing (nascar)’
- ‘f1 motorsports’
- ‘f1 racing’

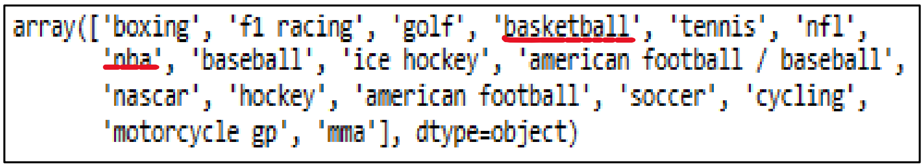


	S.NO	Name	Nationality	Current Rank	Sport	Year
4	5	Alain Prost	France	5	auto racing	1990
15	16	Alain Prost	France	6	auto racing	1991
33	34	Alain Prost	France	4	auto racing	1993
3	4	Ayrton Senna	Brazil	4	auto racing	1990
14	15	Ayrton Senna	Brazil	5	auto racing	1991
22	23	Ayrton Senna	Brazil	3	auto racing	1992
32	33	Ayrton Senna	Brazil	3	auto racing	1993
84	85	Dale Earnhardt	USA	5	nascar	1998
77	78	Dale Earnhardt	USA	8	nascar	1997
98	99	Dale Earnhardt	USA	9	auto racing	1999
105	106	Dale Earnhardt	USA	6	auto racing (nascar)	2000
190	191	Dale Earnhardt Jr.	USA	10	nascar	2009
44	45	Gerhard Berger	Austria	5	auto racing	1994
130	131	Jacques Villeneuve	Canada	10	auto racing	2003
119	120	Jacques Villeneuve	Canada	10	auto racing	2002
120	121	Jeff Gordon	USA	10	auto racing (nascar)	2002
175	176	Kimi Raikkonen	Finland	5	f1 racing	2008
164	165	Kimi Raikkonen	Finland	4	f1 racing	2007
184	185	Kimi Raikkonen	Finland	2	f1 racing	2009
270	271	Lewis Hamilton	UK	10	auto racing	2017
209	210	Michael Schumacher	Germany	9	f1 racing	2011
111	112	Michael Schumacher	Germany	2	f1 motorsports	2002
81	82	Michael Schumacher	Germany	2	f1 motorsports	1998
73	74	Michael Schumacher	Germany	4	f1 motorsports	1997
122	123	Michael Schumacher	Germany	2	f1 motorsports	2003
62	63	Michael Schumacher	Germany	3	f1 motorsports	1996
132	133	Michael Schumacher	Germany	2	f1 racing	2004
142	143	Michael Schumacher	Germany	2	f1 racing	2005
152	153	Michael Schumacher	Germany	2	f1 racing	2006
58	59	Michael Schumacher	Germany	9	auto racing	1995
165	166	Michael Schumacher	Germany	5	f1 racing	2007
100	101	Michael Schumacher	Germany	1	auto racing	2000
90	91	Michael Schumacher	Germany	1	auto racing	1999
49	50	Nigel Mansell	UK	10	auto racing	1994
23	24	Nigel Mansell	UK	4	auto racing	1992
18	19	Nigel Mansell	UK	9	auto racing	1991

Figure 9: Dataset of all athletes in motorsports organised in alphabetical order by their name.

From external research the following decision where made to each athlete’s ‘Sport’ classifcation, reducing the number of unique entries in the ‘Sport’ column from 20 to 17:

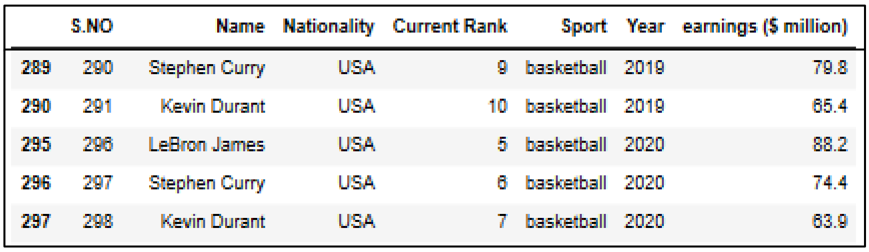
- Alain Prost – f1 racing
- Ayrton Senna – f1 racing
- Dale Earnhardt – nascar
- Dale Earnhardt Jr. – nascar (no changes)
- Gerhand Berger – f1 racing
- Jacques Villeneuve – f1 racing
- Jeff Gordon – nascar
- Kimi Raikkonen – f1 racing (no changes)
- Lewis Hamilton – f1 racing
- Michael Schumacher – f1 racing
- Nigel Mansell – f1 racing



```
array(['boxing', 'f1 racing', 'golf', 'basketball', 'tennis', 'nfl',
       'nba', 'baseball', 'ice hockey', 'american football / baseball',
       'nascar', 'hockey', 'american football', 'soccer', 'cycling',
       'motorcycle gp', 'mma'], dtype=object)
```

Figure 10: All unique values in the 'Sport' column with values associated with basketball highlighted.

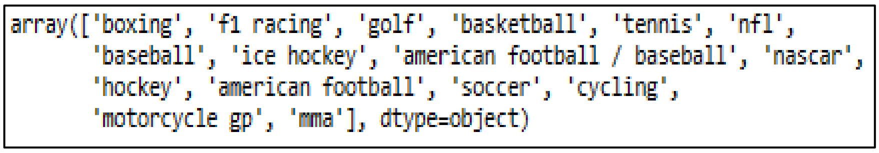
Another ‘Sport’ classifcation that needed to be fixed were athletes where their ‘Sport’ was either ‘nba’ or ‘basketball’ (Figure 10). There were 82 athletes with either of these classifcations and as the most recent classification for this sport in 2020 was ‘basketball’ (Figure 11), all athletes with the value ‘nba’ where changed to ‘basketball’. This reduced the number of unqiue entries in the ‘Sport’ column from 17 to 16.



	S.NO	Name	Nationality	Current Rank	Sport	Year	earnings ($ million)
289	290	Stephen Curry	USA	9	basketball	2019	79.8
290	291	Kevin Durant	USA	10	basketball	2019	65.4
295	296	LeBron James	USA	5	basketball	2020	88.2
296	297	Stephen Curry	USA	6	basketball	2020	74.4
297	298	Kevin Durant	USA	7	basketball	2020	63.9

Figure 11: Last 5 entries of athletes with the 'Sport' classification of 'nba' or 'basketball'.

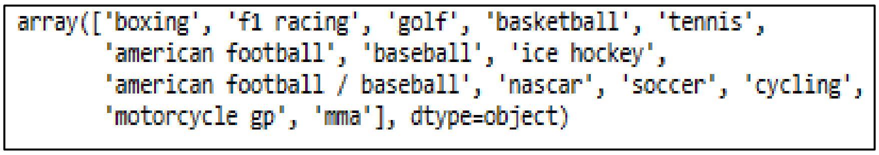
The same process was done for athletes where their ‘Sport’ was classified as ‘nfl’ or ‘american football’, with all athletes sport being set to ‘american football’ as it was the latest classification used in the dataset. This reduced the number of unique ‘Sport’ entries from 16 to 15 (Figure 12).



```
array(['boxing', 'f1 racing', 'golf', 'basketball', 'tennis', 'nfl',
       'baseball', 'ice hockey', 'american football / baseball', 'nascar',
       'hockey', 'american football', 'soccer', 'cycling',
       'motorcycle gp', 'mma'], dtype=object)
```

Figure 12: All unique values in the 'Sport' column with the fixed values for 'american football'.

Finally, clarity for athletes where their ‘Sport’ classification was ‘ice hockey’ or ‘hockey’ was need, and through further research all players in the dataset are considered ‘ice hockey’ players and thus those with the value of ‘hockey’ where changes to ‘ice hockey’. This reduced the total number of unique values in the ‘Sport’ column from 15 to 14 (Figure 13).



```
array(['boxing', 'f1 racing', 'golf', 'basketball', 'tennis',
       'american football', 'baseball', 'ice hockey',
       'american football / baseball', 'nascar', 'soccer', 'cycling',
       'motorcycle gp', 'mma'], dtype=object)
```

Figure 13: All unique classifications for the 'Sport' column that has been processed and cleaned.

DATA STORIES AND VISUALISATION

Stories and assumptions based on the visualisations of the data.

These are the following questions I was interested in investigating based on viewing the initial dataset:

1. Which athlete appears on the Forbes highest paid athletes list the most over the 30 year period?
2. Which nation has the most highest paid athletes appearing on the Forbes list?
3. Which sport has the most highest paid athletes appearing on the Forbes list?
4. What sport has the highest earners on average?
5. Which sport has the highest earner ranked number 1 over the 30 year period?
6. How do ‘earnings $ (millions)’ change over time?
7. Is there any correlation between an athlete’s current rank, year and earnings?

1. Which athlete appears on the Forbes highest paid athletes list the most over the 30 year period?

As shown in the wordcloud visualisation in Figure 14, which shows the top 5 athletes who have appeared on the Forbes list the most over the 30 year period. It shows that Michael Jordan and Tiger Woods appeared on the Forbes list the most with 19 appearances for both athletes. Closely followed by Kobe Bryant with 14 appearances, and with Michael Schumacher and Shaquille O’Neal both having 13 appearances on the list.

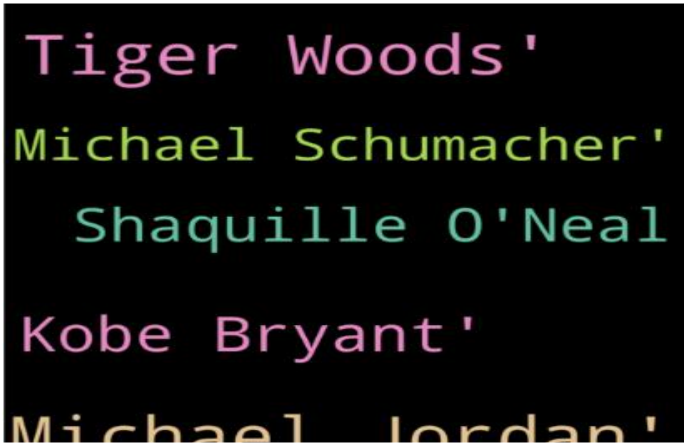


Figure 14: Wordcloud visualisation of the top 5 most frequent athletes who appear on the Forbes lists highest paid athletes.

2. **Which nation has the most highest paid athletes appearing on the Forbes list?**

The USA accounts for 207 of the 301 athletes that have appeared on the list, thus the USA accounts for 68.8% of the highest paid athletes that have appeared within the top 10 over a 30 year period.

The other 19 nations numbers dwindle in comparison with Germany and the UK each having 13 athletes appearing on the list and the other nations having <10 athletes on the list.

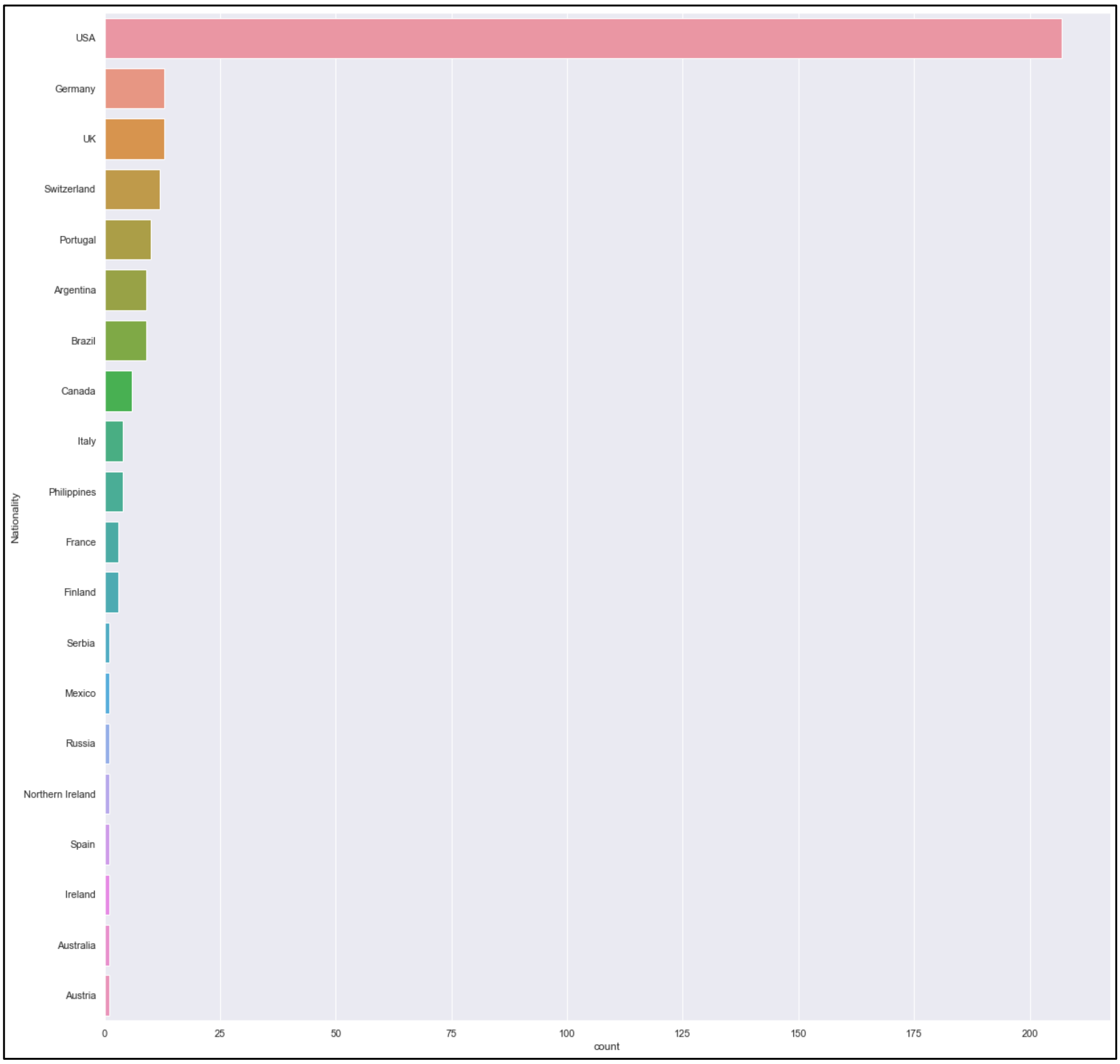


Figure 15: Histogram of the number of athletes that have appeared on the Forbes list from each of the 20 nations.

3. **Which sport has the most highest paid athletes appearing on the Forbes list?**

82 out of the 301 entries on the list have been Basketball players appearing the list almost double that of Boxing with 46 entries (Figure 16). As explored earlier, Michael Jordan, Kobe Bryant and Shaquille O'Neal are large contributors to Basketball appearing on the list around 27% of the time as all three Basketball players total appearances on the list

sum to 46 which accounts for 56% of the total number of Baseball players on the list over the 30 year period.

The sports with the least amount of athletes appearing within the top 10 highest paid athletes are cycling and mma. This could be to external factors such as Basketball players having extra sponsorship deals with big brands such as Nike in which they have their own apparel lines such as ‘Jordan’ brand. Kobe Bryant also had his own line of trainers with Nike for example. Whereas cyclists and mma athletes might not have as much endorsement deals thus their yearly earnings are lower than that of a basketball player for example.

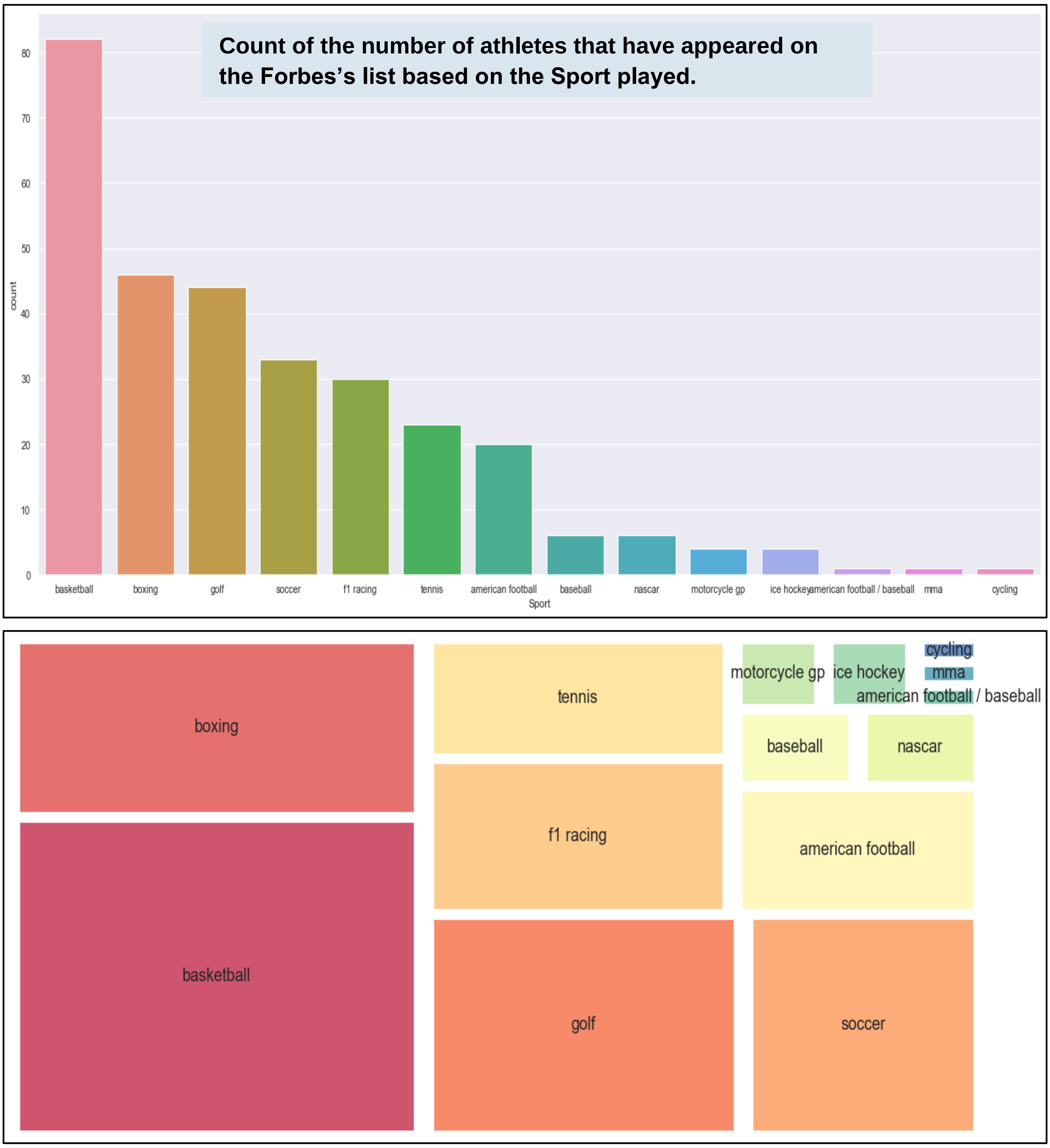


Figure 16: Bar graph ordered in descending order of the most frequent Sport's to occur on the Forbes’s list (top) and an accompanying treemap displaying the proportion of athletes from each Sport on the list (bottom).

4. **What sport has the highest earners on average?**

MMA fighters would appear to have the highest average salary according to the bar graph, however by exploring the corresponding boxplot we can see that this value is attributed to one single entry. The same observation could be said for Cycling and American Football/Baseball in which only one entry in the Forbes's list is contributing to the Sport's appearance on the list.

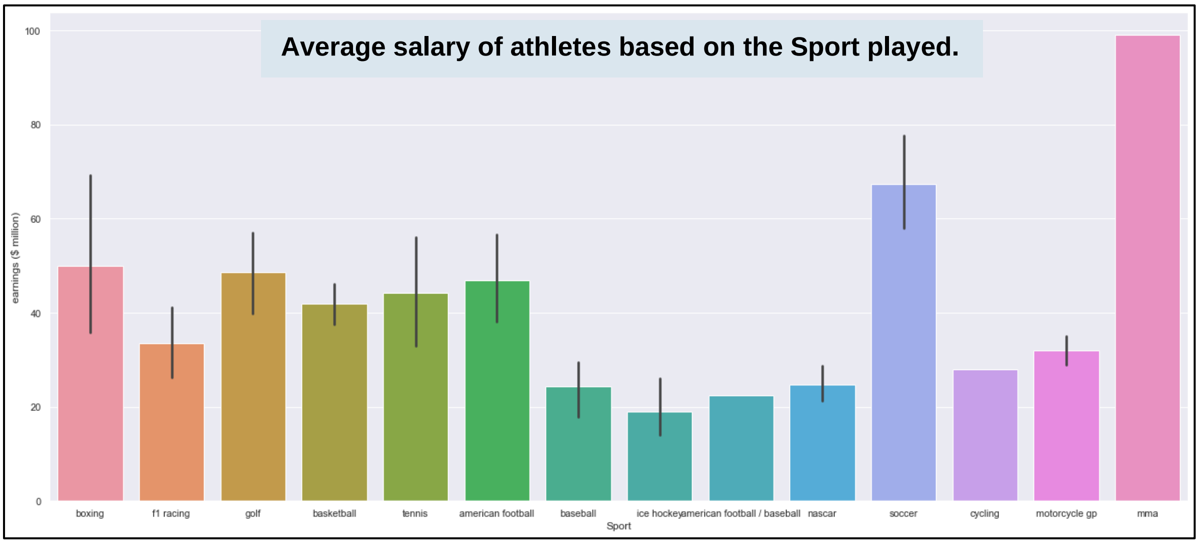
However, all values could be considered accurate representations for each sport as the value for American Football/Baseball for example is an accurate depiction for this Sport category as this athlete did participate in both sports in a single season.

Furthermore, as these values are actual salaries of athletes and only covers the highest earners across multiple Sports it could still be concluded that MMA fighters have the highest average salaries among all other Sports.

However, the manor is which MMA fighters and boxers salaries are determined are extremely different in comparison to the other 12 sports. As by exploring the box plot, we can see that Boxing is the only Sport with outliers with the top earner earning more than double that of the top earner in Soccer for instance. This is because boxers earnings are mainly based on who their opponents are and the title and thus engagement they are receiving for views to watch and pay to view the fight, which is why their salaries are astronomically high in comparison to other athletes as they don't have a set salary year round or for multiple years.

By consider the median, Soccer players would earn the highest average salary which is true for the maximum salary, median salary and minimum salary. Although, Soccer players do appear to have one of the largest IQR however as the salaries have not be adjusted to reflect inflation this could explain some of the large IQR's. As the Soccer players appearing on the list span from 2004 – 2020.

As these values are real salaries we can conclude that if we consider all sport's MMA fighters have the highest average salary. However, if we consider the way salaries are processed it could be argued that Soccer players have the highest average salaries.



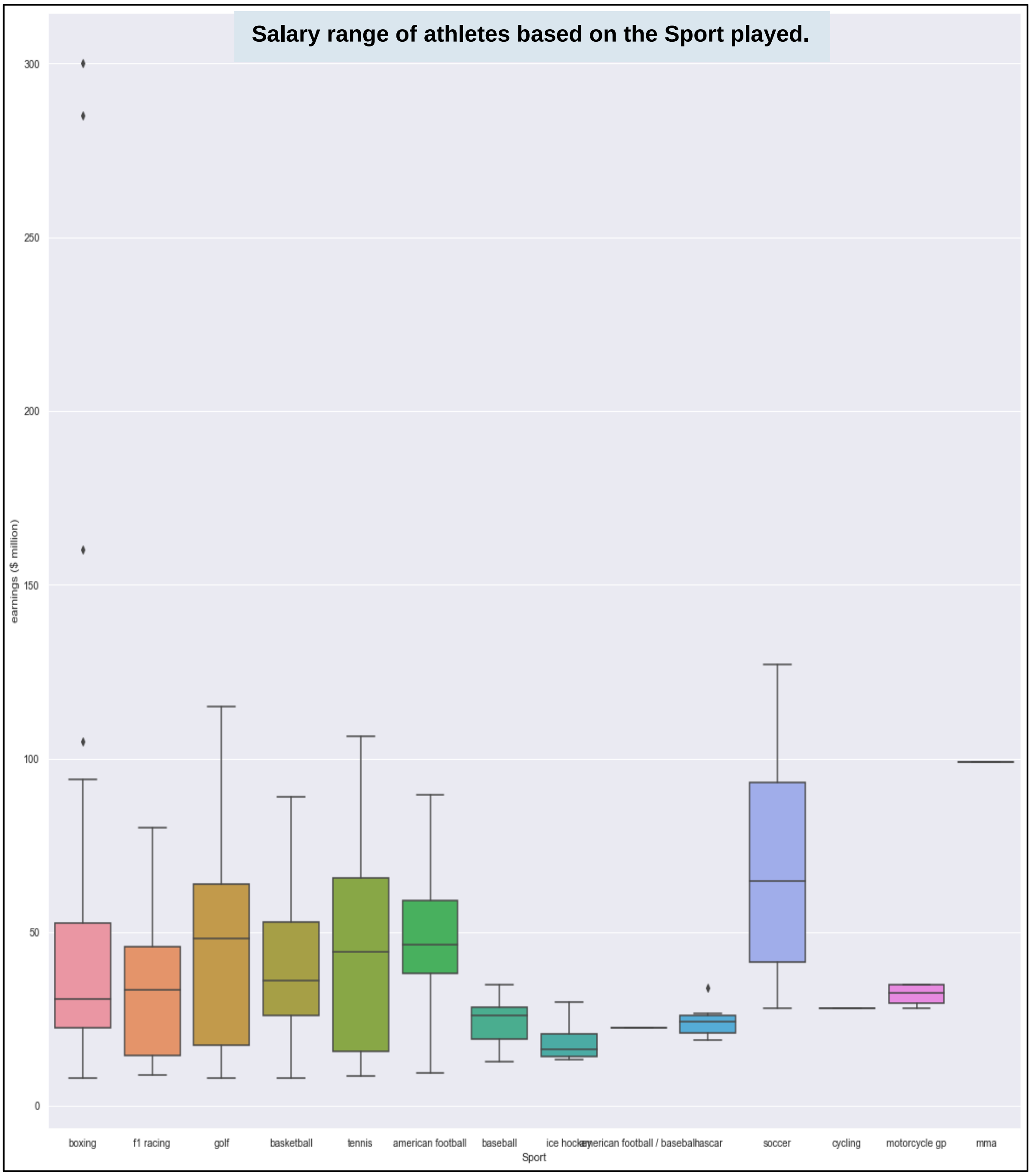


Figure 17: Barplot of average salary earned by athletes based on their Sport (top) and an accompanying boxplot show the range in salaries based on an athletes Sport (bottom).

5. **Which sport has the highest earner ranked number 1 over the 30 year period?**

From the scatterplot shown in Figure 18, we can see that only 6 of the 14 Sports are outputting the number one earnings. This appears to change for different periods in which a certain sport holds this value consecutively and then a new sport will hold the number one earner for some consecutive years. The only sport to obtain the number one sport after a period of time is Boxing, again this can be as a result of their uncapped pay. In general if Sport loses the top earning it appears to be due to Boxing.

The Sport with the most athlete(s) that have held the number one spot for the highest earning athlete is Golf due to Tiger Woods with 11 different years at the top spot. Followed by Boxing, Basketball, Soccer, F1 racing and Tennis.

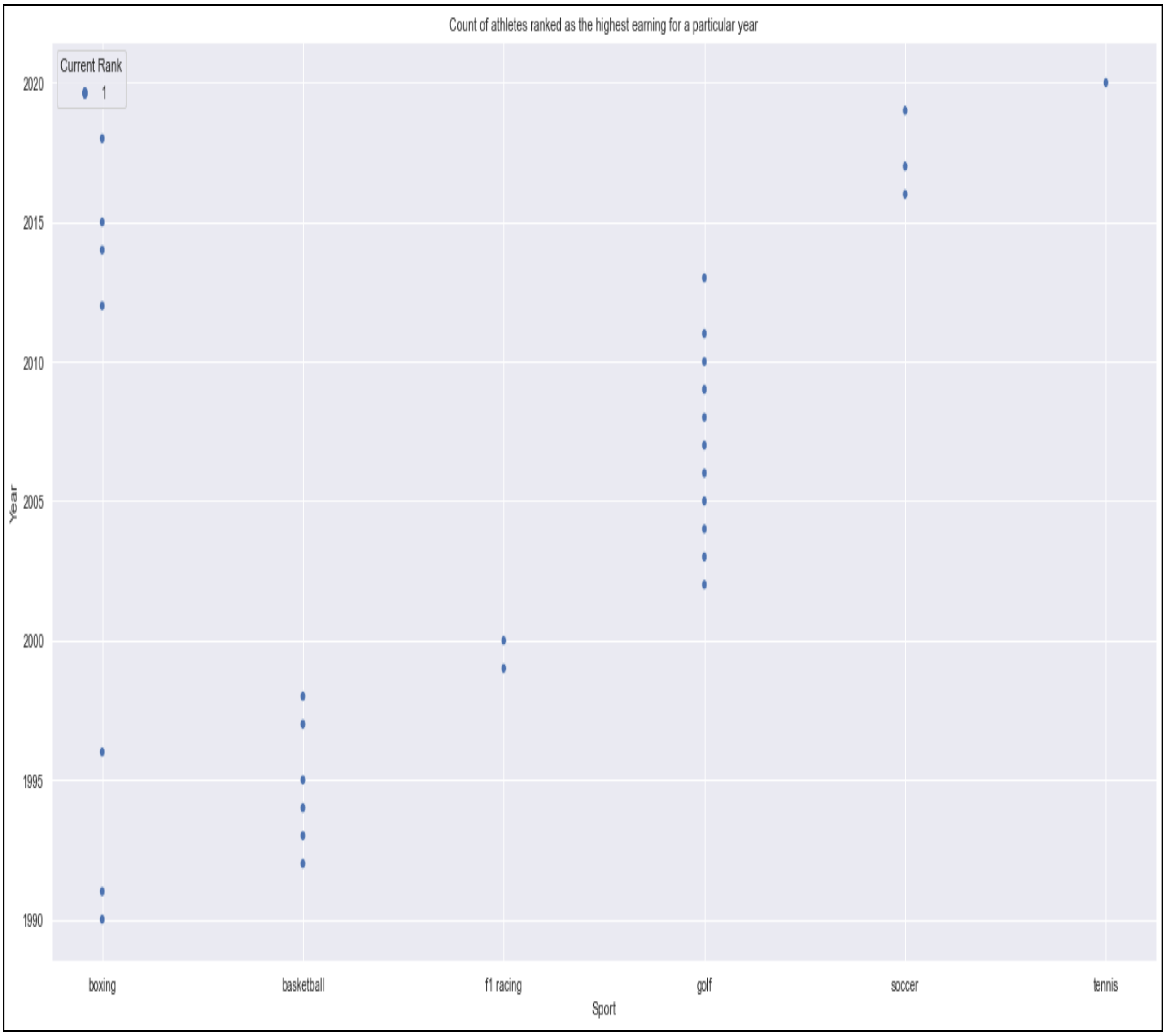


Figure 18: Scatterplot showing the number of times a sport has had an athlete hold the top spot for the highest earning athlete for each year.

6. How do ‘earnings $ (millions)’ change over time?

As expected as the years have moved on the average earnings of athletes have increased in general, as between some consecutive years there are small fluctuations in the total average earnings of the top 10 athletes in certain years. This could be as a result of an particular athlete earning significantly more than the average in comparison to other athletes. This is supported by the error bars which when very long show a large variance in the earnings of all athletes for that particular year.

The most obvious example being 2015 in which Floyd Mayweather made $300 million in earnings, thus resulting in the large error bar in Figure 19.

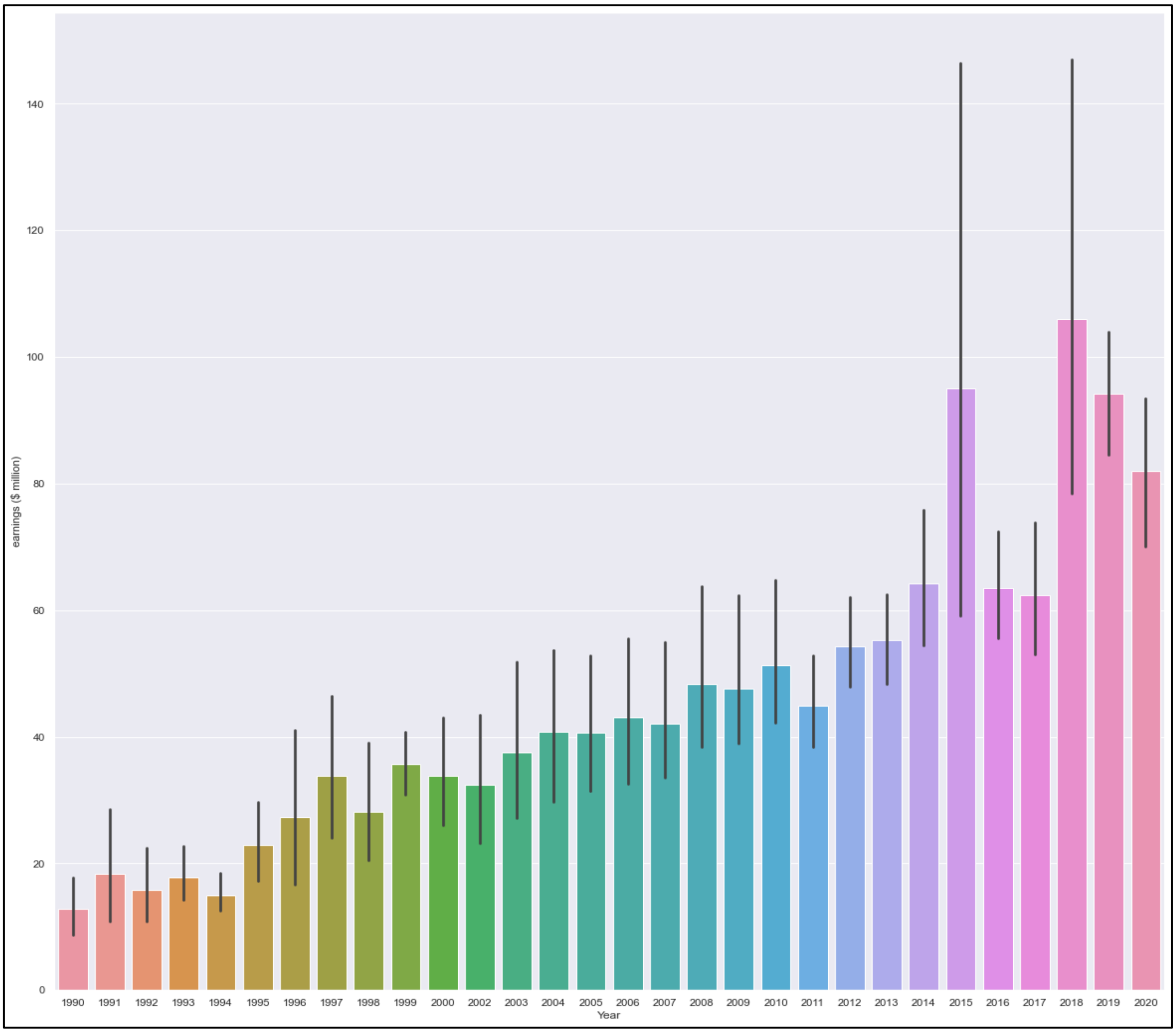


Figure 19: Histogram showing the total average salary of the top 10 highest paid athletes for each year.

7. Is there any correlation between an athlete's current rank, the year and earnings?

From Figure 20, there appears to be a negative correlation between an athlete's 'Current Rank' and their 'earnings $ (millions)' which makes sense because as the 'Current Rank' increases from 1 to 10 the earnings of the athletes decreases as the top earning athlete holds rank 1.

There a no significant correlation between an athletes 'Current Rank' and 'Year' as these values are categorical values that are standard to occur and increment in a set manor, therefore I wouldn't expected for there to be an correlation between the two variables. However, there does appear to be a strong positive correlation between the 'Year' and an athlete's 'earnings $ (millions)' which would again be expected as the values have not been adjusted for inflation and therefore it would be expected for the salaries of the athletes to increase over the 30 year period. I would also expect for this to be true even with the salaries adjusted to account for inflation.

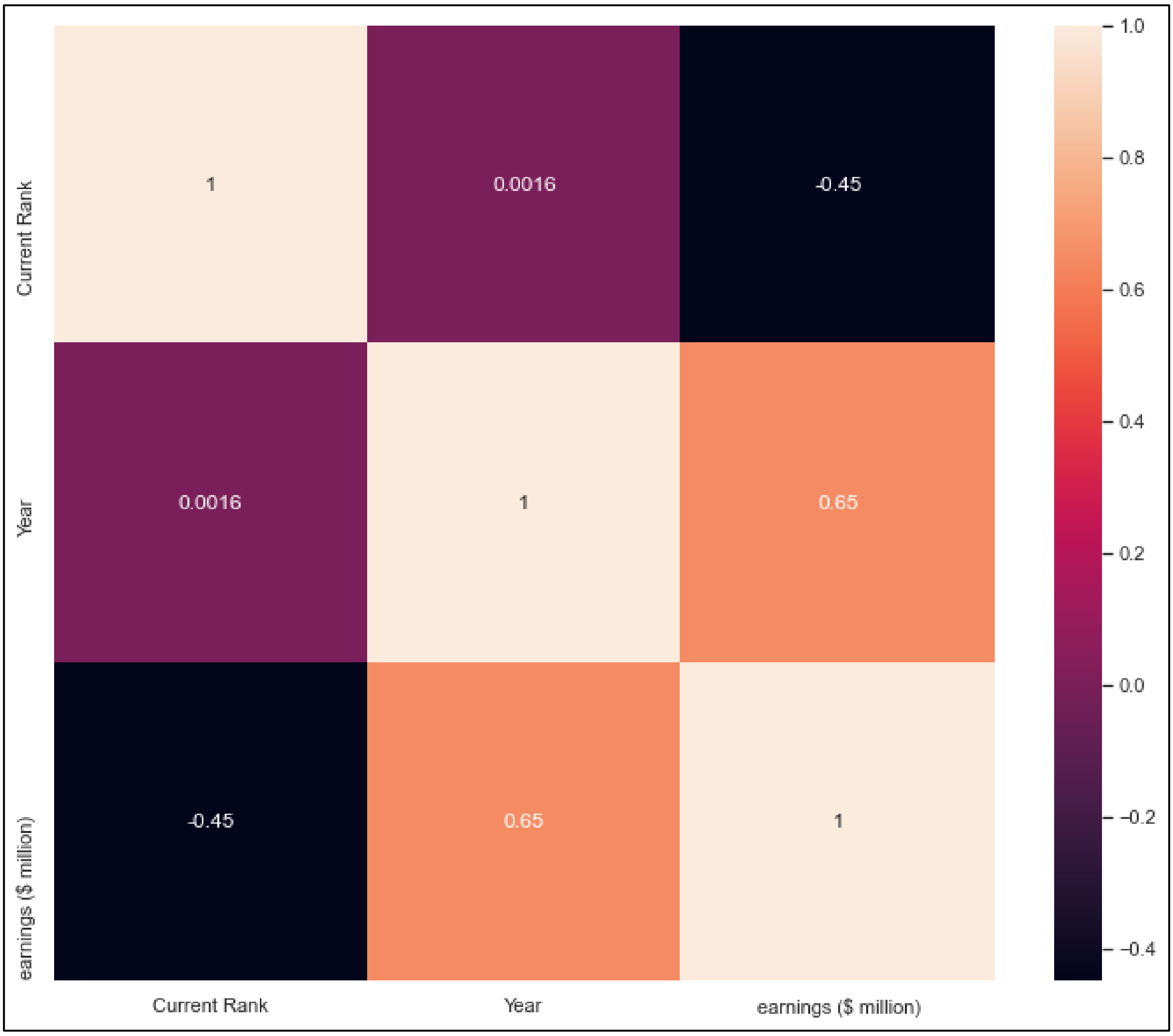


Figure 20: Correlation matrix of the numerical variables in the dataset (earnings, year, current rank)

Summary

- Most frequent athlete to occur on the list is Michael Jordan.
- Most frequent sport to occur on the list is Basketball.
- The USA accounts for 68.8% of the athletes appearing on the Forbes Top 10 Highest Paid Athletes over the 30 year period from 1990 – 2020.
- The sport with the highest paid athletes is Boxing if salary caps and processes are ignored or else Soccer players would be considered the highest paid athletes on average.
- Golf has had the most athlete(s) to hold the ‘Current Rank’ of 1 a total of 11 times over the 30 year period due to Tiger Woods success at the sport.
- There is a strong positive correlation between the Year and an athlete’s earnings and a moderate negative correlation between an athletes Current Rank and earnings.

THIS REPORT WAS WRITTEN BY :

Sophia Dorothy Powell-Morris