**UFC – The Data Analysis**

This all started, when I watched Conor McGregor’s post-fight interview after his fight with Eddie Alvarez. To my amazement, he said ‘I want my share, I want equity’. The media was well aware that he had some big announcements to make. I didn’t expect him to want equity in the UFC!

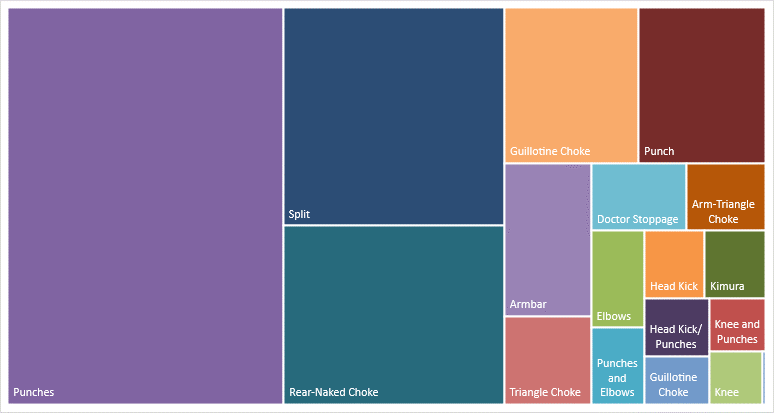
After all, he is the face of UFC at the moment, why should he not use that as leverage? Also he was in the top 100 paid athletes of 2016 and he made history that day in New York by winning his second UFC title.

How much money has he actually made for UFC, though?

To my astonishment, I found two datasets. One contained all fights from 1993 to January 2016 and the other contained pay per view values of the main events in UFC from 2001 to August 2016. Note, I can not vouce for the qualilty or correctness of the data set, but I enjoyed working with it!

Before looking at the pay per view data. I decided to check out the fight’s data – after all, this is where all the valuable data is. After a load of wrangling with the data using Spark and then with the use of excel I came up with a few charts!

Below is a link to my GitHub where you can find the datasets that I used and the code.



*Figure shows the most dominant technique fighters adopted for the finishing move – the fatal blow.*

The data was collected from 3,570 fights. For presentation purposes I removed those techniques if their frequency was below 20.

From history we know that ground techniques are devastating if executed correctly, no surprise that 8 of the top 17 techniques used are either BJJ or wrestling techniques. They represent 47.3% of the data shown the above treemap.

699 fights finished with punches being the move that ended the match, is this evidence to how lethal a punch can be or how worn out the fighters must have been? Having been through a number of rounds. Note non-championship matches last only 3 round (minimum 5min round) and championship matches being 5 rounds (minimum 5min round).

Now let’s get to the interesting part – the pay per view data.

*Figure shows the total revenue from ppv that each fighter has brought in for the UFC*

When dealing with this dataset the ppv value given was for the main event for each UFC tournament taken place since 2001. I split the ppv value 50/50 between the contenders.

Shockingly McGregor doesn’t even reach the top 5!

The top earner for UFC (income from ppv) is Anderson Silva, who holds the longest tittle streak in the UFC history, he held the Middleweight Championship title for well over 2,000 days. He defended his title 10 times!

This being put this aside, McGregor earnt $4 million from endorsements this year, which includes the sporting giant Reebok. His charisma is far superior to any other fighter in the UFC, and has attracted a large and loyal fan base, which has in turn boosted his ego even further! Additional evidence to support his demand for equity.

*Figure shows wins as a percentage for those fighters that have taken part in 20 or more fights*

Here I wanted to look how successful a fighter has been in their career, given all the fights they have taken part in.

The top performing fighter is Georges St. Pierre, the coveted welterweight fighter of all time. He also holds the record for the most wins in title bouts.

Other interesting findings:

* 62% of fighters aged over 32 will lose
* 64% of fighters without a wrestling background are likely to lose
* 62% chance of winning if the opponent is American

GitHub Link: <https://github.com/hemesh-patel/UFC>