

Analysing UFC Data

Kevin Kontchou

DScourseS18

05-1-2018

What is the UFC?

- The Ultimate Fighting Championship (UFC) was created in 1993 to to identify the most effective martial art in a contest.
- Was the between competitors of different fighting disciplines like boxing, Brazilian jiu-jitsu, Sambo, wrestling, Muay Thai, karate, judo, ect.
- UFC has grown to the world's leading mixed martial arts (MMA) promotion company, putting on the most events in the industry by a large margin.

What data did I use?

- An MMA and data science enthusiast, Jian Qiao, scraped the data from sherdog.com
- Two data sets: Fights and Fighters
- The two data sets used for this 390 events and 4058 matches has been held around the globe

Problems with the data

- Needed some cleaning and wrangling
- Does not include events after 2/4/17. 44 fights and over 100 debuting fighters
- UFC has absorbed other MMA promotions like StrikeForce and Pride MMA
- Overwhelming majority of fighters do not spend entire career UFC.

What questions am I looking to answer?

- Who is the greatest fighter of all-time?
- What is the distribution of the different methods of victory?
- How does age affect success in the UFC?

What methods did I use?

- Merged the two data sets by number of fights and wins to get win percentage for each fighter
- Used Empirical Bayes Estimation to adjust for fighters with very few fights and win percentages of 0 or 1

What methods did I use?

- Merged two data sets by method of victory to get show the distribution of different methods.
- Did the same for each individual weight class

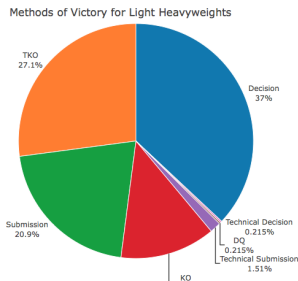
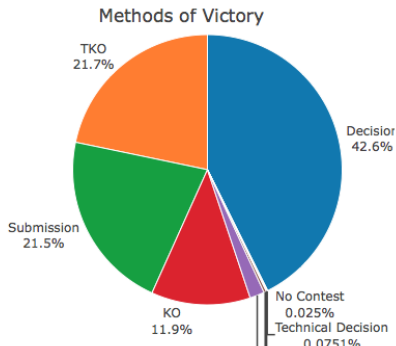
What methods did I use?

- Merged two data sets by the birth date of each fighter to get an age for each fighter.
- Subtracted the date of the event by the birth date of each fighter to get a "FightingAge" variable
- Used FightingAge to get distribution of number of wins of UFC fights by age of UFC fighters and create a graph that measures success with age.

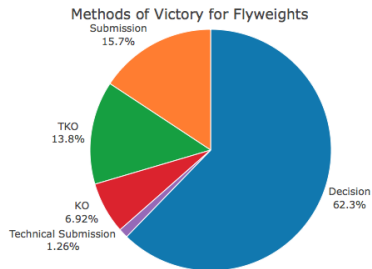
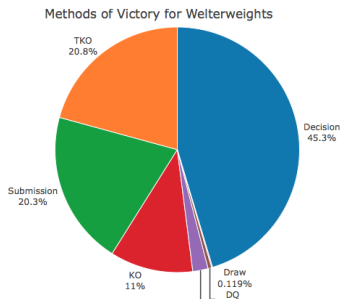
What were some of the findings?

Rank	Name	Class	Bouts	Wins	Win Percentage
1	Jon Jones	Light Heavyweight	17	16	0.83239
2	Georges St. Pierre	Welterweight	21	19	0.81999
3	Demetrious Johnson	Flyweight	14	13	0.80669

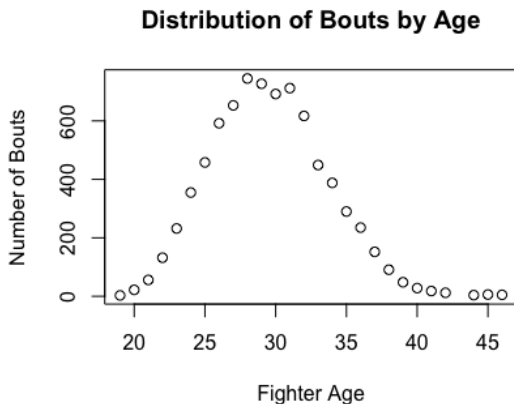
What were some of the findings?



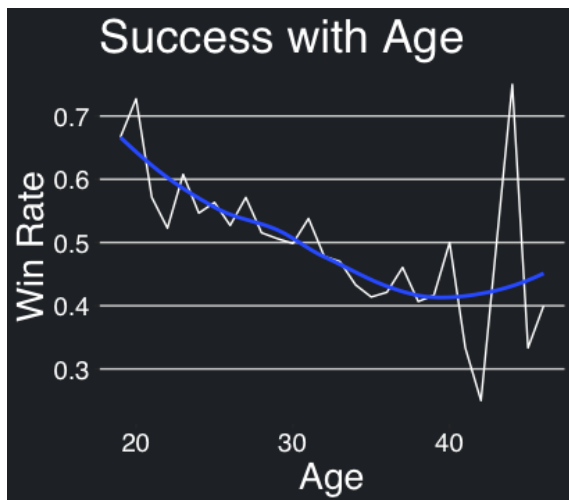
What were some of the findings?



What were some of the findings?



What were some of the findings?



Conclusion

- There is room for argument in the top 3, but Jon Jones has the edge as the GOAT.
- Most fights in the UFC are taken by fighters in their late 20s and early 30s, which indicates that the UFC is a destination that fighters are building their careers towards.
- Win rate declines with age, which says something about youth vs. experience.