CS9.422 Behavioural Research: Statistical Methods

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY

HYDERABAD

Final Project Report

Topic Name: In the Octagon: Statistical Insights into UFC Fighters' Age and Performance metrics

Team Name: исследовать

Submitted By: Nilesh Keshwani (2023201077)

Other Team Members: Shubham Jaiswal (202320170)

Amey Kunte (2020111024)

1. Introduction

The Ultimate Fighting Championship (UFC) commands a global audience with its electrifying mix of martial arts styles showcased within the octagon. This project zooms in on the statistical analysis of UFC fighters' performance metrics and demographics (which includes stance, reach, age and other vital characteristics for a fighter). We aim to test and realize the general trends that are obvious in such field of sports by performing statistical methods taught to us during the coursework and will try to distil key insights into factors that influence fighter performance.

2. Dataset Description

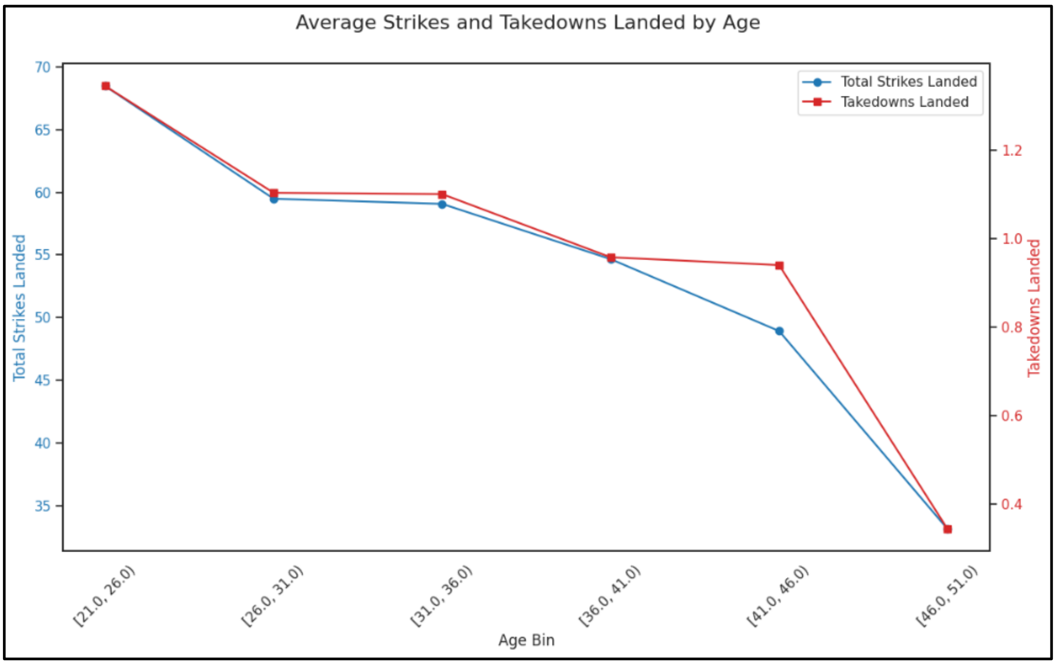
We have used this dataset from kaggle (UFC/MMA Biggest Dataset With Differentials) that uses ufcstats using the masterdataframe.csv dataset with 530 features and 13322 entries (6661 fights since each pair of entry is for similar fight it's just that it is represented from the other way around), with matches from year 1994 to 2022 (almost 3 decades). It covers the matches from across all divisions (both male and female) with the peak number of matches in the year 2019 and the least in the year 1998. The dataset is about fight across the year with each entry having metrics that corresponds to fighter's demographics, their attributes, fight outcome and various performance metrics. The dataset also has additional differential stats calculated for each fight (differential stats are calculated by dividing the two individual's technique stats).

3. EDA & Data Cleaning

With the challenge of selecting the suitable features from available nearly 1000 feature for each fight (since each pair represent the same fight, we can neglect the common fight attributes for each fight), we tried to come up with some data cleaning and EDA based on our understanding of the domain (influenced by the hypothesis we want to test).

From our limited knowledge about the dataset and the domain itself we decided to clean the data for the purpose of testing several hypothesis, as we believed certain factor such as the differentials in the statistics of fighter for each fight could be difficult to handle and currently out of scope for our project work and we scaled down to 34 features suiting our need of the project. We also trimmed down the number of data points to 8238 (considering the most recent fights and removing those fights with a NaN value for any of the 37 selected features, left with 4119 matches).

We performed the following EDA after performing data cleaning to get an estimate of what analysis or hypothesis we could work around with:



We can see the trend of reducing total strikes and total takedowns down the age bin (of 6 years) this can include two possibilities that either the fights down the age bins are reducing leading to this misleading analysis or that the fighter suffers injuries or loses their physical attributes (endurance, agility, etc.) that cannot be measured accurately while keeping the record of fights.

4. Hypothesis & Results

The following hypothesis were formulated and tested on our cleaned dataset and the observed results were reported. The inspiration for the statistical method was also tried to explain for each hypothesis.

<u>**Hypothesis 1**: There exists an identifiable relation between reach and match win across divisions</u>

Null Hypothesis: There is no significant relationship between reach and match win across divisions.

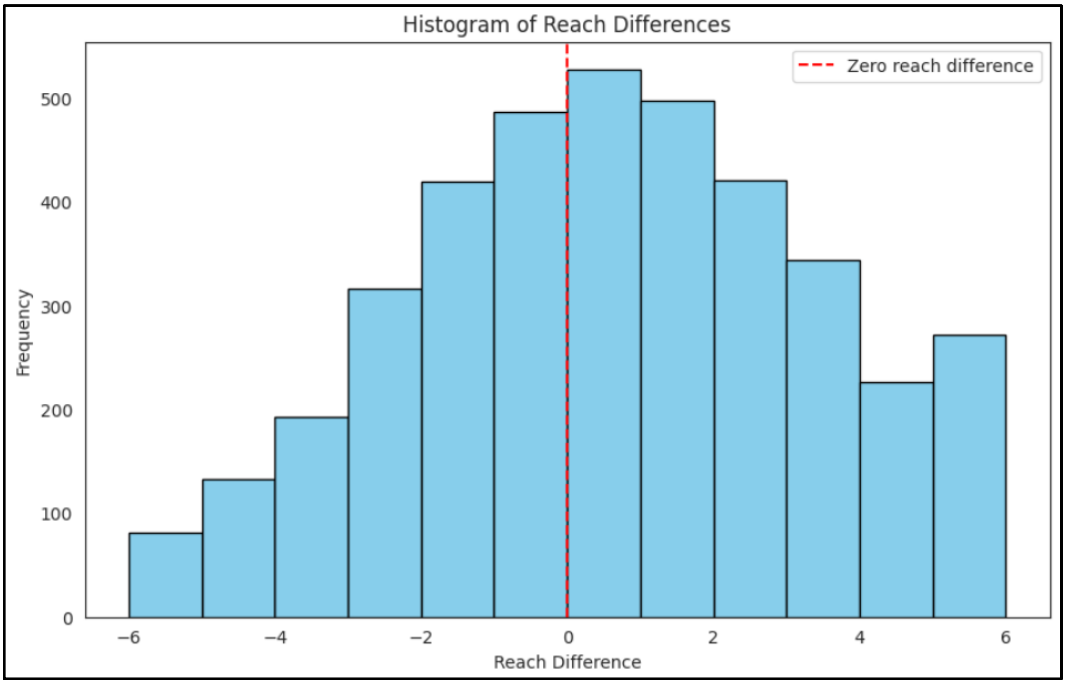
Alternate Hypothesis: There is a significant relationship between reach and match win across divisions.

Statistical Test: Binomial Test

Inspiration: We wanted to see whether the positive reach difference is a common characteristic among the winners, as opposed to winners with a reach disadvantage. The sign test is appropriate in this context because it doesn't assume any specific distribution and is suitable for situation we want to compare medians rather than means.

Results: P-value from the Binomial Test: 0.02674965133682936

Visualization: The following is the histogram of the reach differences plotted against the frequency of wins for that particular bin



Inference: With a p-value of 0.024 below the standard significance level of 0.05, we can reject the null hypothesis. It indicates a statistically significant contrast between the winners possessing a positive reach advantage and those with a negative reach advantage (or none). Put differently, a notable proportion of winners enjoyed a reach advantage over their opponents, suggesting that reach may play a pivotal role in determining outcomes in UFC bouts.

<u>**Hypothesis 2**: Method of Victory as an Indicator of Division Characteristics</u>

Null Hypothesis: The method of victory (e.g., submission, decision) is independent of the division.

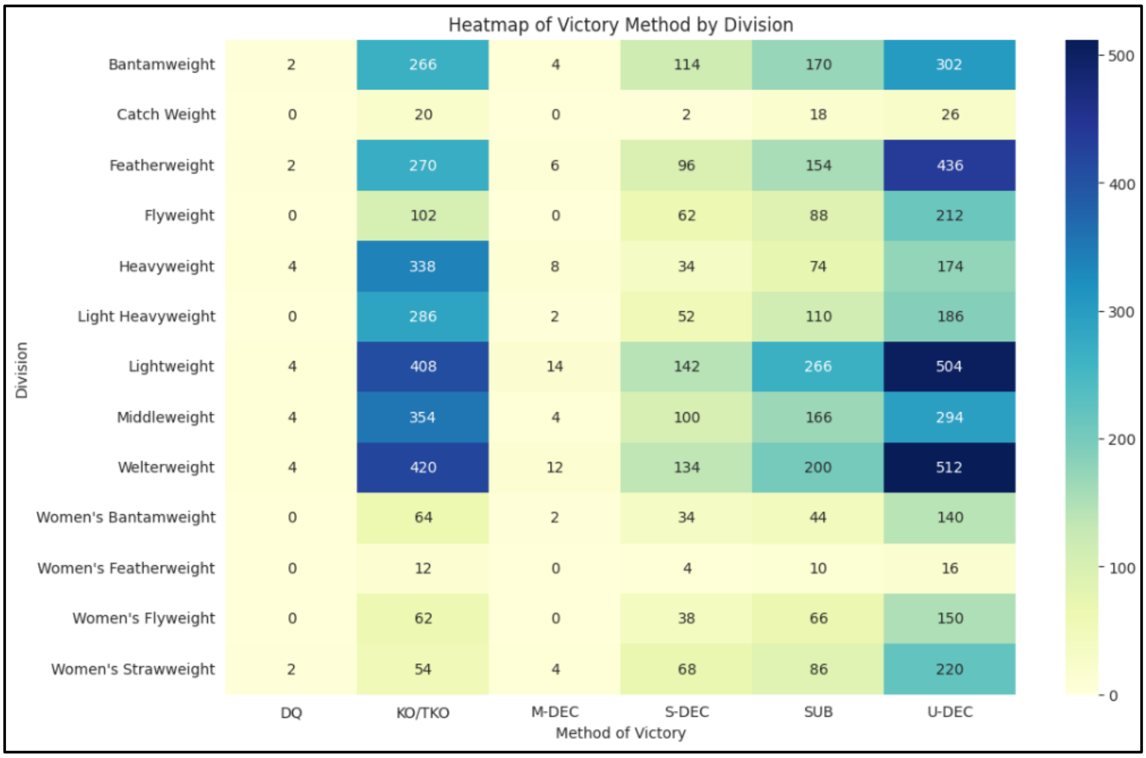
Alternate Hypothesis: Certain victory methods are more common in specific divisions, suggesting that division characteristics influence the fighting style or strategy.

Statistical Test: Chi-Square Test

Inspiration: We wanted to see what trends we could observe between two categorical variable namely division (13 division in all the UFC fights we have considered, divisions are based on the weight of the fighter) and the method of fight decision (6 method of decision) which seems to be an obvious curiosity to the audience of this sport. We tried to statistically prove this using the chi-square test and then plotting a contingency table for the same.

Results: Chi-squared Test: chi2 = 430.0132828416177, p-value = 2.4039841881870444e-57

Visualization: The following is the heatmap for the contingency table plotted for the two categorical variables:



Inference: The chi-squared test result reveals a significantly high chi-square statistic and an almost negligible p-value, suggesting that the distribution of winning methods (e.g., decisions, submissions) across divisions is not random. Therefore, we reject the null hypothesis (H0) and accept the alternative hypothesis (H1) indicates that certain victory methods are notably more prevalent in specific divisions. This observation implies that divisional characteristics potentially exert a significant influence on fighting styles or strategies adopted by UFC fighters.

Hypothesis 3: Head strike defence is more significant than other defence techniques in influencing the outcome of UFC fights

Null Hypothesis: The effectiveness of head strike defence is not more significant in determining fight outcomes than the effectiveness of other defence techniques (e.g., body strikes defence, leg strikes defence, etc).

Alternate Hypothesis: The effectiveness of head strike defence is more significant in determining fight outcomes than the effectiveness of other defence techniques.

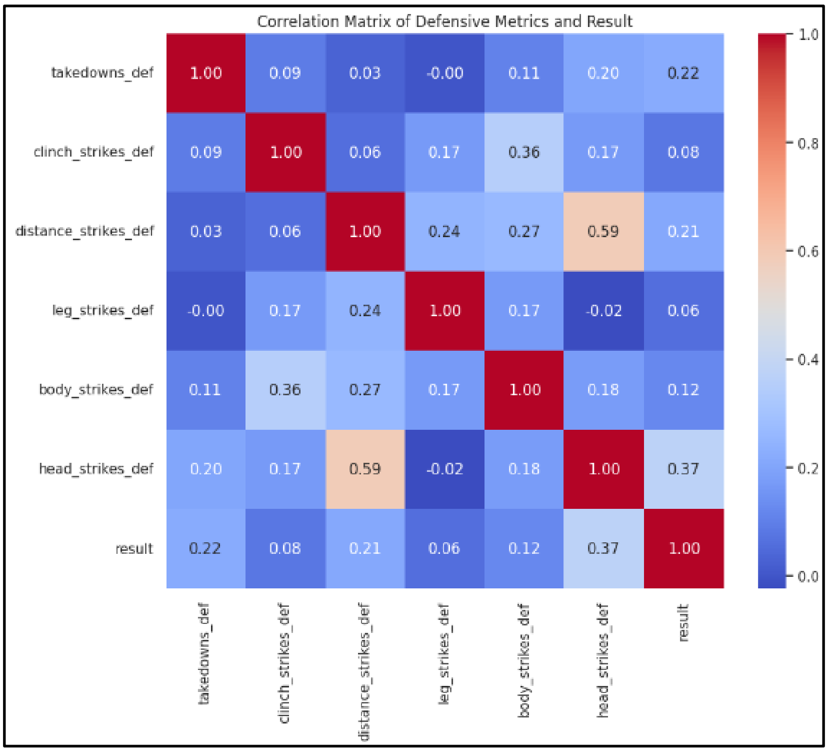
Statistical Test: Logistic Regression

Inspiration: We intend to find that a fighter that employs a defensive approach in their fights especially defending against the head strike attempts, the likelihood of their winning the match increases as compared to other defensive approach such as defending body, leg, clinch strike, etc. Logistic Regression proves to be a better approach since we can see the control the head strike defence has over the winning rate in such an approach.

Results: Following are the results that we observed when we tried to apply Logistic Regression on the different defensive metrics for each fighter.

```
Optimization terminated successfully.
         Current function value: 0.601011
         Iterations 5
                           Logit Regression Results
==============================================================================
Dep. Variable:                 result   No. Observations:                 8238
Model:                          Logit   Df Residuals:                     8232
Method:                           MLE   Df Model:                            5
Date:                Wed, 08 May 2024   Pseudo R-squ.:                  0.1329
Time:                        16:35:27   Log-Likelihood:                -4951.1
converged:                       True   LL-Null:                       -5710.1
Covariance Type:            nonrobust   LLR p-value:                     0.000
=====================================================================================
                        coef    std err          z      P>|z|      [0.025      0.975]
-------------------------------------------------------------------------------------
const                -0.0116      0.024     -0.478      0.632      -0.059       0.036
takedowns_def         0.3535      0.025     13.952      0.000       0.304       0.403
clinch_strikes_def   -0.0445      0.027     -1.672      0.095      -0.097       0.008
leg_strikes_def       0.1533      0.026      6.007      0.000       0.103       0.203
body_strikes_def      0.1008      0.027      3.699      0.000       0.047       0.154
head_strikes_def      0.8448      0.030     28.411      0.000       0.786       0.903
=====================================================================================
```

Visualization: The following is the heatmap for the correlation matrix for different defensive techniques based on which we selected the most suitable features



Inference: The results of the logistic regression analysis support the alternate hypothesis. Head strike defence was found to be the most statistically significant factor (coefficient: 0.8448, p-value: 0.000) compared to other defence techniques in influencing the outcome of UFC fights. This suggests that effective head strike defence has a stronger association with winning a fight than effective defence against body strikes or leg strikes.

Hypothesis 4: Stance Advantage in Fight Outcomes

Null Hypothesis: Fighter stance (Orthodox, Southpaw or Switch) does not give an advantage in fights

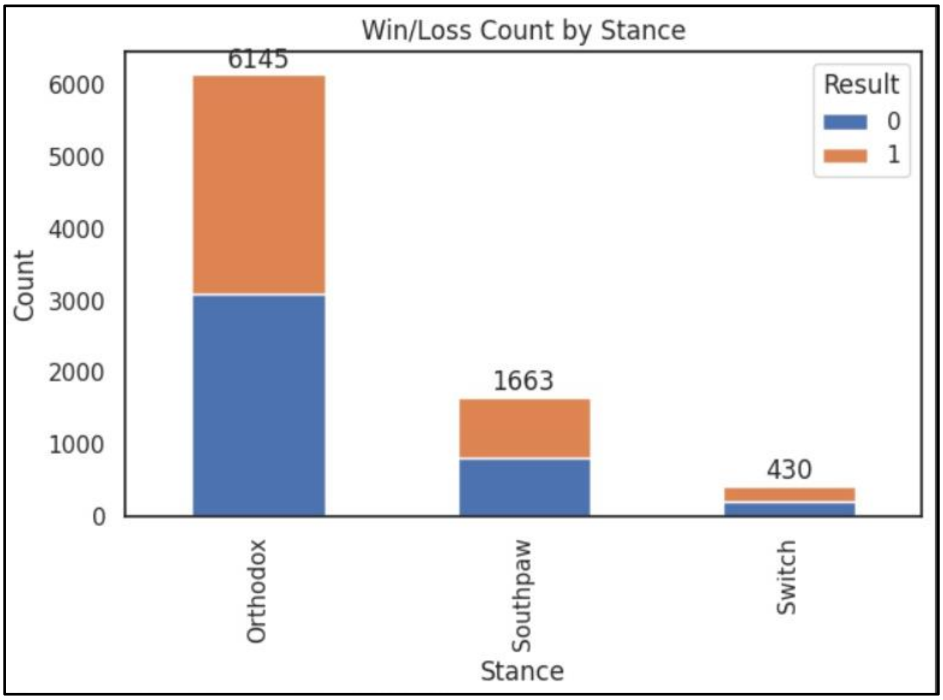
Alternate Hypothesis: Southpaw fighters have an advantage in fights, reflected in higher win rates across divisions.

Statistical Test: Chi-Square Test

Inspiration: In combat sports like UFC, wrestling, etc. the initial fighting stance can hold a major effect on the win outcome of the fighter. On analysis of our clean dataset we observed that there are three types of fighting stance and since we wanted to observe whether southpaw can be ruled as the superior fighting stance or not we tried to apply a chi-square test between two nominal categorical variable to test the hypothesis.

Results: Chi-square statistic: 2.272692183101421, p-value: 0.32098974880444432

Visualization: The following is the bar graph for the different stances vs Win Loss Frequency (total frequency written in the black label).



Inference: Since the p-value (0.321) is greater than the typical significance level of 0.05, we do not have enough evidence to reject the null hypothesis. Therefore, based on this analysis, we cannot conclude that there is a statistically significant association between stance and win/loss outcomes. In other words, there is insufficient evidence to support the hypothesis that certain stances offer a strategic advantage leading to higher win rates.

Hypothesis 5: Older (Positive Age difference) Winning fighters prefer more head strikes than body strikes

Null Hypothesis: There is no difference in the preference for head strikes over body strikes between older winning fighters and younger winning fighters. This means that age does not influence a fighter's strategy regarding the preference for head versus body strikes.

Alternate Hypothesis: Older winning fighters have a different preference for head strikes over body strikes compared to younger winning fighters. Specifically, older fighters prefer head strikes more than body strikes relative to their younger counterparts.

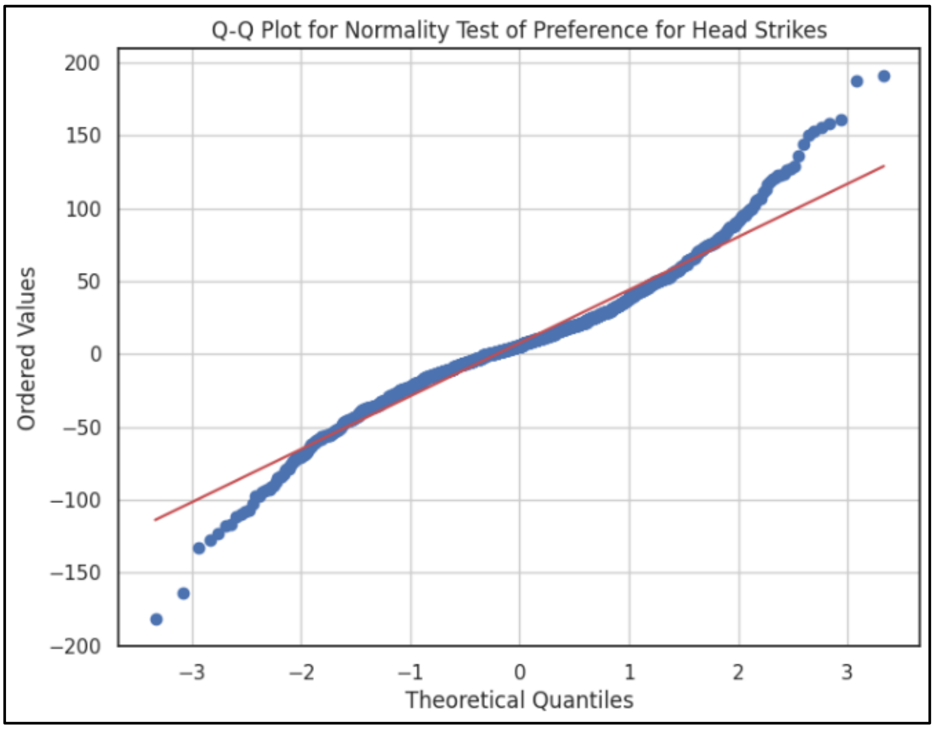
Statistical Test: t-Test if the data in normal and homogeneous else Mann-Whitney U test.

Inspiration: As demonstrated in the EDA above that there was a decline in the total strikes and takedown landed across the age bins, but we also stated the reason why that could be misleading, to prove something relevant to it we conducted this test.

Results: Following were the results reported when we performed a t-test on our cleaned data:

Normality p-value: 1.1516245979271934e-20
Data is not normally distributed. Using non-parametric test.
Mann-Whitney U test p-value: 1.9799682811963305e-15

Visualization: The following is the Q-Q plot which shows the deviation in the normality test.



Inference: With an exceptionally low p-value obtained from the Mann-Whitney U test, we confidently reject the null hypothesis. This indicates a statistically significant inclination among older winning fighters towards head strikes rather than body strikes, compared to their younger counterparts. Such a discovery corroborates the hypothesis, implying that older fighters potentially adopt a fighting style prioritizing head-targeting techniques over body strikes, likely influenced by factors such as accumulated experience or strategic adjustments throughout their careers.

Hypothesis 6: Aggression Leads to More Wins

Null Hypothesis: There is no significant relationship between the aggressive attributes and the winning rate (the p-value for these attributes will be greater than 0.05 and for those less than 0.05 the contributing factor in the likelihood of winning would be marginal)

Alternate Hypothesis: The attributes that shows the aggression of the fighter would have a significant effect on the likelihood of winning (there p-value will be less than 0.05 and their contributing factor would also be significant).

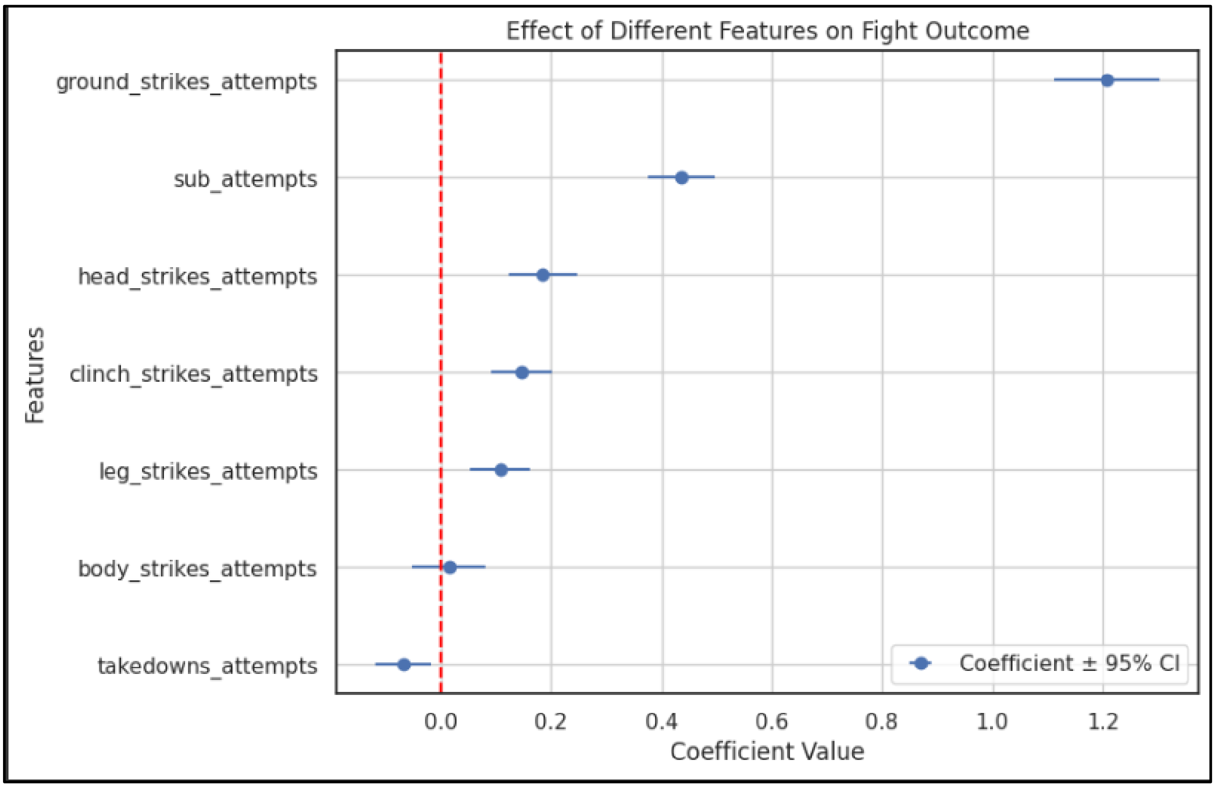
Statistical Test: Logistic Regression.

Inspiration: As fight result/outcome is a categorical feature we can view this problem as a Logistic Regression problem. To measure the aggression of the fighter we use the attributes that defines the number of attempts the fighter took on the opponent (be it head strikes, body strikes, leg strikes, etc.)

Results: Following were the results reported when we performed a Logistic Regression test on our cleaned data:

```
                           Logit Regression Results
==============================================================================
Dep. Variable:                 result   No. Observations:                 8238
Model:                          Logit   Df Residuals:                     8230
Method:                           MLE   Df Model:                            7
Date:                Wed, 08 May 2024   Pseudo R-squ.:                  0.1401
Time:                        17:07:11   Log-Likelihood:                -4910.2
converged:                       True   LL-Null:                       -5710.1
Covariance Type:            nonrobust   LLR p-value:                     0.000
===========================================================================================
                              coef    std err          z      P>|z|      [0.025      0.975]
-------------------------------------------------------------------------------------------
const                       0.1618      0.027      6.098      0.000       0.110       0.214
ground_strikes_attempts     1.2069      0.049     24.711      0.000       1.111       1.303
body_strikes_attempts       0.0149      0.034      0.440      0.660      -0.052       0.081
head_strikes_attempts       0.1848      0.031      5.921      0.000       0.124       0.246
leg_strikes_attempts        0.1076      0.028      3.904      0.000       0.054       0.162
sub_attempts                0.4367      0.031     14.276      0.000       0.377       0.497
clinch_strikes_attempts     0.1461      0.029      5.125      0.000       0.090       0.202
takedowns_attempts         -0.0668      0.026     -2.590      0.010      -0.117      -0.016
===========================================================================================
```

Visualization: Following is a coefficient value plot for different features that we plotted for easy presentation and perception:



Inference: The logistic regression analysis partially supports the alternate hypothesis. Ground strikes and clinch strikes attempts are statistically significant predictors of winning fights, while other factors such as head strikes, leg strikes, body strikes, and takedowns show no significant associations with fight outcomes. Aggression, particularly through ground strikes and clinch strikes, influences UFC fight outcomes. However, the significance of other strike types and takedowns remains inconclusive based on this analysis.

5. Contributions & Discussion

- Based on the valuable feedbacks provided in the final presentation by both the professor we revisited some of our hypothesis improved there results (by performing alternative test) and where needed we tried to reformulate our hypothesis.
- The Data Cleaning part and EDA was done by all three of the group member with the help from online kaggle notebooks, again we took them as a resource and not as a reference to revolve our hypothesis around their work.
- The work shown in this report was majorly done by me (Nilesh Keshwani) and Shubham Jaiswal which includes formulating the hypothesis on the EDA we performed on the cleaned data, stating the underlying hypothesis, deciding the statistical test based on inspiration (the domain knowledge we have in the field to decide the hypothesis), implementing and visualizing them and finally drawing inference from them.

- The other part of our project includes prime age range analysis of UFC fighters. Due to limited knowledge about the domain (as mentioned in the report repeatedly), me and Shubham helped in determining that the win rate (adjusted using Laplace smoothing over different fight thresholds) could be a contributing factor in our in the determination of prime age range, but other work on the same was done by Amey Kunte, which includes deciding other performance metrics that would help in such an interesting study (strikes landed, strikes defended, takedowns landed, etc.) and then showing them as trend lines or trend curves.

6. References

- **Reference Notebook for EDA on the original UFC dataset:** **https://www.kaggle.com/code/miroslavkirnak/ufc-eda-and-prediction**

- **Prime Age Range Analysis Inspiration :** **https://youtu.be/iOf20O49zig?si=bksumsZzY-zQ5a-V**

- **UFC Original Dataset :** **https://www.kaggle.com/datasets/danmcinerney/mma-differentials-and-elo/data?select=masterdataframe.csv**

- **StatsModel API documentation for referring to different hypothesis testing methods available in the module:** **https://www.statsmodels.org/stable/api.html**