BF4 Capstone Project – Statistical Data Analysis

**Dataset:** "collected data9.6.2019/merged\_data.csv"

**Jupyter Notebook Files:** “Data Story Telling.ipynb”, “Capstone Project 1 Statistical Data Analysis.ipynb”

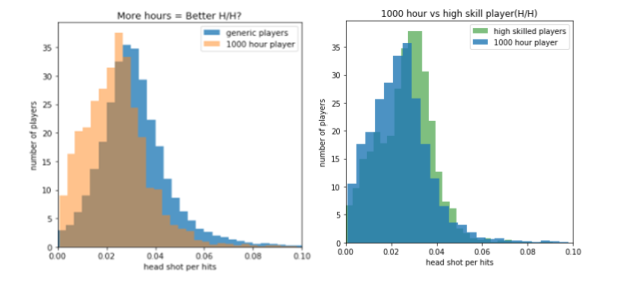
**Methods:** mean, median, percentiles, standard deviation, correlations, t-test, chi-square

Within the BF4 Capstone Project, we used a series of methods to analyze player data based on skill level, hours played, win ratio, gun usage etc.

# Generic Performance

In the beginning of the analysis, we mostly focused on overall player performance. By looking at the skill level and player frequency, we found that majority of our players had a skill level below 237 and most players had less than 1000 hours of play time. This meant that our data is heavily skewed left when it comes to hours. Top players are in general very dedicated to their game.

So in order to better understand this data, we split the players into two groups based on the 1000 hour play time, and used histograms to see if 1000 hour players did better than generic player who has less than 1000 hours of play time. The results showed that contributing to more hours within the game did lead to better wins, higher skills and better Kill/Death ratios. However one thing that stood out was headshots. In typical first person shooters, headshots are considered the most effective way of eliminating the enemy as it correlates with high damage. Yet in this scenario, players who have long playtime seems to be actively avoiding making headshots and going for body-shots instead.



Even highly skilled players seem to only make a little more effort in getting more headshots. This may help us infer that headshots are not an effective way to make more gains within the BF4 game. A few other conclusions drawn from comparing groups and also introducing a correlation matrix:

1. Deaths per minute are nearly the same regardless of skill level;
2. The win-rate for every group is nearly the same since the server does a relatively good job at auto-match making
3. High level players are more likely to target and destroy vehicles which is more likely to contribute to higher win-rates

# Gun Usage

Within the gun usage data, we focused on the mean accuracy and standard deviation for all weapon categories. We found that Shotguns have the highest average accuracy, but a high std, which means, this weapon takes a lot of effort to master; LMG is the worst weapon in terms of accuracy; The best and most stable weapon seems to be SNIPER RIFLE, with low standard deviation and a relatively high accuracy.

Since the mean and std of carbines were only slightly better than assault rifles, we compared the two groups using a student t-test. The null hypothesis for our t-test was H0: There is no statistically significant difference between the carbine and assault rifle accuracy.

The t-test resulted in the results below which meant the null hypothesis was rejected and that carbines were indeed better than assault rifles when it came to accuracy.

Ttest\_indResult(statistic=-21.71692396390152, pvalue=3.8026573725064386e-104)

Another data group we looked at was weapon usage amongst the top 25% and bottom 75% of players. We used score as a way to look at usage frequency since score highly correlates with weapon usage time.

The null hypothesis was that there isn't a significant difference between top 25% of players and bottom 75% in terms of gun preference. We used chi-square to compare the two groups and the results showed that the null hypothesis couldn’t be rejected and there was no evidence that the two groups had significant differences.

Power\_divergenceResult(statistic=4.55, pvalue=0.804410169231813)