Exploring the relationship between League of Legends players' non-professional play and professional play.

DVG ESports: msiordia, bsohn2, knguyen50, bguan2

Prediction Task

League of Legends is one of the most popular games in ESports, a rapidly growing industry. Professional teams are always on the lookout for players that can win and bring home the prize money. We want to explore if a professional player's performance in non-professional play is related to their performance in professional play. We will be examining if there is a significant linear relationship between a player's average kills, deaths, assists, rank, and win rate in their non-professional play and their win rate in professional play. We believe that at least one of these statistics will be significantly related to win rate. That being said, it will not be surprising if there is not a significant relationship between these statistics and professional win rate. Professional play is more team-oriented than non-professional play and many factors outside of individual non-professional play such as team chemistry and coaching play an important role.

Data

Our non-professional data is from the developer's (Riot Games) API.
We scraped Oracle's Elixir for player names which we then used to scrape the player 's most recent in-game ID using trackingthepros.com. With player IDs, we then scraped the developer's (Riot Games) API for the players' match history (kills, deaths, assists, win/loss) and rank. The data was clean and did not have any missing data. We did have to edit our data throughout the project as we added and removed other statistics (e.g., matches in other game modes). We only scraped 40 players because that is the roster size for a single professional season. Riot's API has a rate limit that prevented us from making too many calls, so we limited each person to 25 matches.

Findings

Claim 1: Our statistics have a high R-squared; it fits the model well. Our lines of best fit predict a player's professional win rate accurately given a non-professional statistic.
Claim 2: Our statistics have high p-values; the statistics do not hold a significant relation with win rate.
Support: Our high R-squared values mean that our non-professional statistics fit our models well. The low p-values mean that the non-professional statistics are not significantly related to professional win rate. Overall, none of the non-professional statistics had a significant linear relationship with professional win rate.

	R-squared	p-values
Average Kills	0.767	0.463
Average Assists	0.769	0.214
Average Deaths	0.851	0.716
Rank	0.795	0.433
Win Rate	0.759	0.179

Claim 3: Our model's accuracy of predictions is generally low.
Support: We split our data into a training and testing set (20/80) and found that the MSE for every predictor is high, due to our data being too scattered.

	Training MSE	Testing MSE
All	360.48	292.90
Average Kills	349.84	476.71
Average Deaths	408.26	275.03
Average Assists	417.23	259.89
Rank	402.51	279.62
Win Rate	377.75	348.91

Note: Our R-squared and our MSE values are high; these are conflicting measures of prediction accuracy. We couldn't figure out how to reconcile the two, so we defaulted to R-squared and concluded that our model accurately predicted professional win rate.