There have been many advancements in the statistical analysis of traditional sports over the past 30 years. Sabermetrics has been popularized by the work of Bill James and entered the consciousness of popular culture through Fantasy Sports and movies such as Moneyball. Advanced analytics attempts to empirically analyze the performance of a team or individuals in a given sport. It allows for insights into patterns, anomalies, and efficiency of a participant or a group of participants in a sport that is not inferable through visual observation. My project will be an attempt at evaluating the current state of statistical analysis as well as creating a new set of metrics for performance for the video game League of Legends.

League of Legends is the most widely played game in the world with over 70 million users and approximately 30 million active users a month. It is a game which has made the most progress in creating and fostering a competitive environment that culminates in the world finals which draws millions of viewers across the globe.

The study will begin by mining data through the official Riot Games API. Correlation analysis and a statistical summary of available match data will be explored. Using a large database of game datasets, the matches will be separated by divisions which already exist within the competitive ranking system. Initial evaluation will consist of identifying various numerical win condition thresholds for ranked matches. Comprehending which stats should qualify for initial observation requires some domain expertise about the nature of the game.

The primary focus of the study will then shift to improving on the KDA(Kills/Deaths/Assists) ratio which is most commonly used as an indicator of player performance. This traditional KDA is the sum of kills and assists divided by the number of deaths. While this metric is commonly accepted as an accurate indicator of performance, it intrinsically has biases and omissions about the relative performance and the context.

The tools that are required to perform these tasks will primarily be Python for scraping, RStudio for analysis, and various regression and correlation techniques. An attempt will be made to create a predictive model that determines the winning team based on the newly developed performance metric. The successful creation of a new metric has a lot of commercial value as the game is popular and other third party services are already in regular use amongst the League community.