

# Info 201: Project Proposal

Section AD, Group AD5

16th November, 2017

## **About the Dataset**

We will be working with a League of Legends dataset (set of csv files). League of Legends is one of the biggest Esports in the world currently, with over 100 million active monthly users. We found this dataset on Kaggle, uploaded by Chuck Ephron. The dataset contains data from all competitive games played between 2015 and 2017. The dataset can be found using this link.

## **Target Audience**

Our target audience is anyone who is interested in Esports. As the dataset is about competitive League of Legends matches, our *main* target audience would be people who play League of Legends, or those who are interested in playing it. We'd like to address several questions through this project. Some of them are:

1. Do the top pro-players actually perform well when it matters the most?
2. What are the factors affecting the game's win rate the most? Is it the number of wards placed by the team? Or is it other factors such as which team is on the blue side?
3. We would also like to address how different champions selected is affecting the outcome of games.
4. Lastly, how important are objectives in winning the game? Do teams who take baron first always win the game?

## **Technical Description**

We plan to use Shiny app, as it provides an amazing tool kit to make visually appealing web-pages.

The dataset we are using is a bunch of csv files, with different kinds of data such as champions banned, objectives taken and other aspects of the game in each csv file. Thus,

we will read these static csv files in R and analyze the data.

Currently the data presented in the csv files is not properly formatted. For example, there are columns which contain lists of game times where a team has taken an objective. However, we must reshape this data in order to make plots, and thus must use data-wrangling techniques to do so- We plan to use dplyr and other data-wrangling techniques to reshape this data. Also, as there are multiple csv files which have common game ids, we'll join these csv files into a single data frame. The major libraries we plan to use are:

1. ggplot2/plotly for data visualizations.
2. We might use matrixExtra to perform different functions on the rows and columns.
3. Lastly, we will use shiny app for making the web page.

The major challenge we will have is to wrangle and present the data in such a way that people who do not play League of Legends can also comprehend our analysis. We plan to explain everything in detail, and also plot several data visualizations.

The link to the github repository can be found [here](#).