The Pennsylvania State University

League of Legends: What Increases Damage to Other Players?

Using Regression Tools to Research In-Game Damage Output

Ishan Muzumdar, Deric Liang, Atharv Gupte

Statistics 462

Professor Cremona

9 December 2018

Abstract

The video game *League of Legends* has gained traction over the last decade as a pioneer of professional Esports. With the rapid market growth of the Esports industry, many players want to figure out how to maximize their damage in the game. This was the main goal of this study. The data used in this paper contained 1,000,000 player statistics, and a sample of 1000 player observations with 12 predictors of interest were used to perform various regression analyses on the response sqrt('totdmgtochamp'). Three models were constructed: an SLR model with the 'goldearned' predictor; an MLR model chosen using best subset selection methods; and an IRLS Robust Regression model. Both the MLR and IRLS models included the predictors 'kills,' 'totminionkills,' 'champlvl,' and 'deaths.' The main findings of interest were the models showed extremely high adjusted R-squared, and that all four predictors demonstrated a significant, positive relationship to sqrt('totdmgtochamp').

Introduction

The dataset contains data concerning match statistics of individual players of the popular eSports game "League of Legends."

This dataset, obtained at https://www.kaggle.com/paololol/league-of-legends-ranked-matches, was developed by Paolo Campanelli, drawing data from 184,070 ranked matches over several years. A ranked match means that the game was played competitively to move up the player hierarchy, as opposed to a normal match which is just for fun. In the description, Campanelli acknowledges that they found this data on an SQL database.

Since each match contains 10 players, the data contains an extremely large number of individuals. Campanelli split the match statistics for individual players into two datasets, and in our analysis, we are looking at the first part. This contains statistics for 1,000,000 individuals. Although this dataset has 56 columns, we will only mention the variables of interest regarding our research questions, which we will discuss later. These variables were chosen based of previous experience playing the game.

- totdmgtochamp: The aggregate of all the damage done to other players by an individual in a single game. This will be the response variable in analysis.
- kills: The number of other players killed onto other players by an individual in a game.
- longesttimespentliving: The longest period of time an individual spends in that game alive. Based off of previous game knowledge and inspection of the dataset, this is measured in seconds.
- totminionskilled: The number of minions an individual kills in the game. Minions are small, non-player characters which are killed throughout the game to get gold. See 'goldearned' description to learn more about gold.
- goldearned: The amount of gold a player earns in the game. Gold allows a player to buy items to make themselves stronger.
- champlvl: A measure of experience earned, ranging from 1-18. All players start at level 1, and gain experience through behaviors related to success in the game, such as killing minions or champions.
- dmgtoobj: The measure of damage done by a player to elements of the other team's base. Most of these elements serve as defense mechanisms to keep players from destroying the main objective, the Nexus. Destroying the Nexus wins the destroying team the game. Objectives also include the two large monsters on the map, which give power boosts to an entire team.
- totdmgtaken: The aggregate of all the damage taken by a player that game from any source.
- totheal: The aggregate of all the healing done by a player. Healing gives health back to a player.
- deaths: The number of times a player dies in a game.

- largestcrit: Measures the damage of the largest critical strike a player deals in a game. Each player has a basic attack, and a critical strike is when the damage from this basic attack doubles. The chances of a critical strike occurring can be purchased with gold.
- visionscore: A player's Vision Score in a game. Vision Score can be increased by placing wards or destroying wards. In the game, most parts of the map are not visible, allowing other players to sneak up on you; wards allow players and their teammates see these parts of the map.
- firstblood: A Bernoulli variable (0: No or 1: Yes) of whether or not a player gets First Blood in the game. First Blood is given to whoever gets the first kill.

In playing the game, all players are interested in maximizing their damage to other players, as it allows them and their team to win fights and build advantages over the other team; as such, it will be used as a metric to measure success in the game. In this light, the main objective of this analysis is to explore how other individual player statistics related to success in the game (based off previous game experience) can help predict how much damage a player does to other players, and the degree to which these variables are related. Considering the objective, these are some research questions of interest:

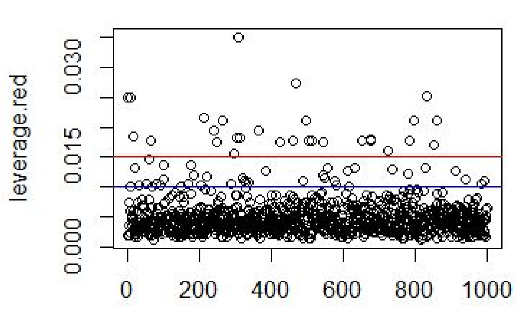
- Will proactivity in vision control (which leads to higher vision score) show a higher ability to damage other players?
- Does killing objectives serve as a good time investment, if the player wants to maximize the damage they do to other players?
- Many players will earn a great deal of gold but never spend it. Therefore, how related will their gold earned be to damage done to other champions?
- What are the most useful metrics in predicting the amount of damage an individual will do to other players?

Exploratory Data Analysis

For ease of analysis, a random sample of 1,000 observations was taken from the dataset. All analyses were performed on this chosen subset.

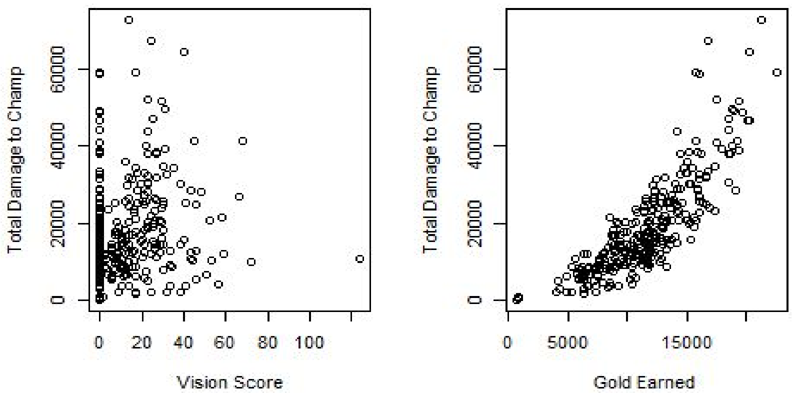
The response variable, the total damage to champion ('totdmgtochamp'), has a mean of 17,828, and a median of 15,538. Since the median is considerably less than the mean, it is likely that the response is skewed to the right (i.e. a few players have caused significantly more damage than the rest of the players).

Before a reliable regression model for the 'totdmgtochamp' can be constructed, it is essential to understand potential strong correlations between predictors, and see possible relationships in their distributions. Upon analysis of the correlation matrix between all combinations of 2 predictors, most pairs displayed little to no possibility of multicollinearity. However, the 'goldearned' and 'champlvl' predictors displayed a significantly stronger linear relationship than any other pair of predictors. It turned out, in fact, the correlation between the two was .8995. Additionally, it turned out that another instance with high correlation between two predictors also occurred with the goldearned, only this time with 'kills' (with a correlation of .7409). The fact that 'goldearned' is highly correlated with two other predictors could cause potential errors in the analysis on how related their gold earned will be to damage done to other champions.



Several points in the dataset exhibited high leverage, which means that they had an average predictor value considerably different than the bulk of the data. The existence of this could have severe impacts on future analysis on causational relationships between the predictors and total damage done.

On the previous graph, the blue line represents the cutoff threshold, above which leverage is considered significant enough for further analysis. As shown in the plot above, several points lie above this threshold..



Consider the 2 graphs to the left, which are scatter plots between two predictors ('visionscore' and 'goldearned,' and 'longesttime spentliving'), and the response.

The first research question, on whether proactivity in vision control shows higher ability to damage can be partially answered by the leftmost scatterplot. It is likely from this plot that an accurate relationship between the vision score and response would not be able to be constructed, without at least some transformation. Note how individuals with a vision score close to 0 have a range of total damage performed, from 0 to over 60,000. Most likely, there is little to no relationship between the two variables, and a majority of the data points skew to lower vision scores.

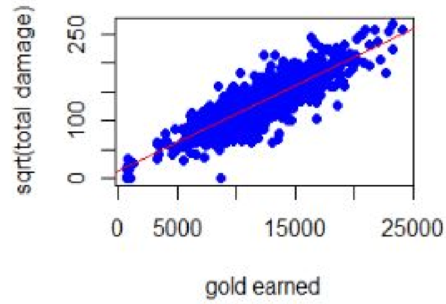
As for the relationship between gold earned and total damage done, there is evidence that the relationship is positive, but the shape of the band of points suggests that such a relationship may not be linear. Note how the points roughly follow an upwards-facing parabola, with increased spread in the response as the gold earned increase. It is likely here that both the assumptions of linearity and equal variance would be violated when constructing linear models. Also, note that the gold earned is highly correlated with two other predictors, so potential errors in further analysis of this relationship are possible without modifications to the predictor space.

Importantly, this previous analysis was not rigorous in nature, and no explicit regression model was checked and its significance verified by a hypothesis test. Nevertheless, a starting point to answering the exploratory questions on Page 2 can certainly aim to stratify future analysis and modeling of the data to draw more reliable conclusions.

Methods

Before performing any model analysis, the regression assumptions of the model needed to be checked. After evaluating the residuals vs. fitted values and the Normal Q-Q Plot for the full model, it was clear the linearity, equivariance and normality assumptions were violated. To address these issues, a square root transformation of the response variable 'totdmgtochamp' was performed. The VIF of each predictor was then calculated to measure potential collinearity effects on the model's coefficients. Most notably, the predictor 'goldearned' indicated a VIF of 11.09, and accordingly, was removed from the full model. The resulting model demonstrated a maximum VIF of 3.28, under the standard collinearity threshold of 4. To answer the research questions about 'goldearned,' it was necessary to fit a simple linear model measuring its relationship with square root of ‘totdmgtochamp.’ Below is the R summary output of this model, as well as a regression plot charting the relationship between these two variables. The model successfully satisfied all necessary regression assumptions and thus, all interpretations and results were deemed legitimate and valid.

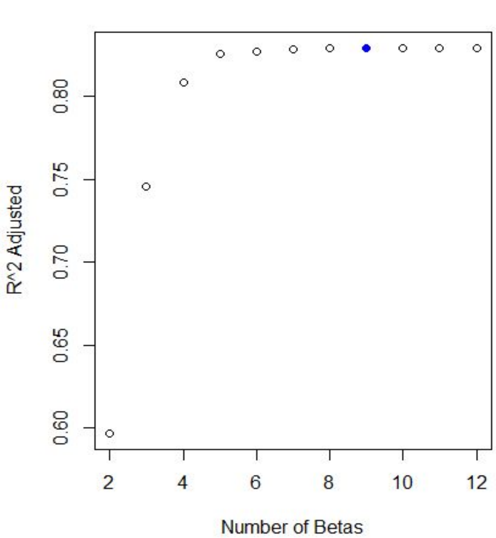
```
Call:
lm(formula = sqrt(totdmgtochamp) ~ goldearned, data = df.new)

Residuals:
     Min       1Q   Median       3Q      Max 
-100.272  -13.249    0.786   15.154   77.113 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.535e+01  2.158e+00   7.109 2.22e-12 ***
goldearned  9.752e-03  1.804e-04  54.045  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 21.53 on 998 degrees of freedom
Multiple R-squared:  0.7453,    Adjusted R-squared:  0.7451 
F-statistic:  2921 on 1 and 998 DF,  p-value: < 2.2e-16
```

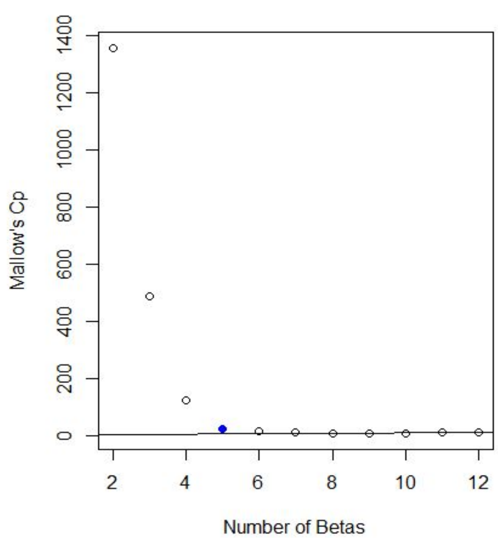


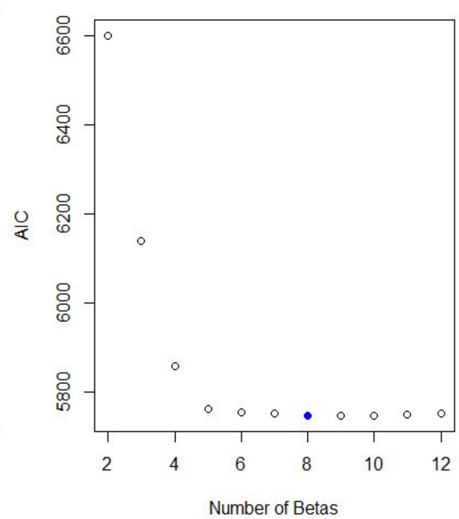
From the above regression summary, the fitted equation for this model is:

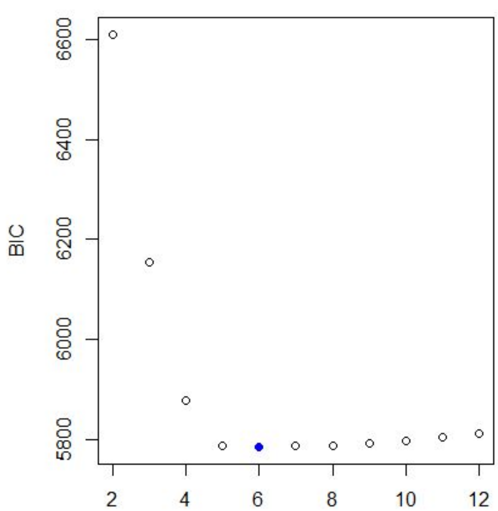
$$\sqrt{Y_{i-hat}} = 15.35 + .00975\, x_{goldearned,i}$$

Note that the p-value for the ‘goldearned’ coefficient is $2(10)^{-16}$, implying that it is indeed significant and nonzero.



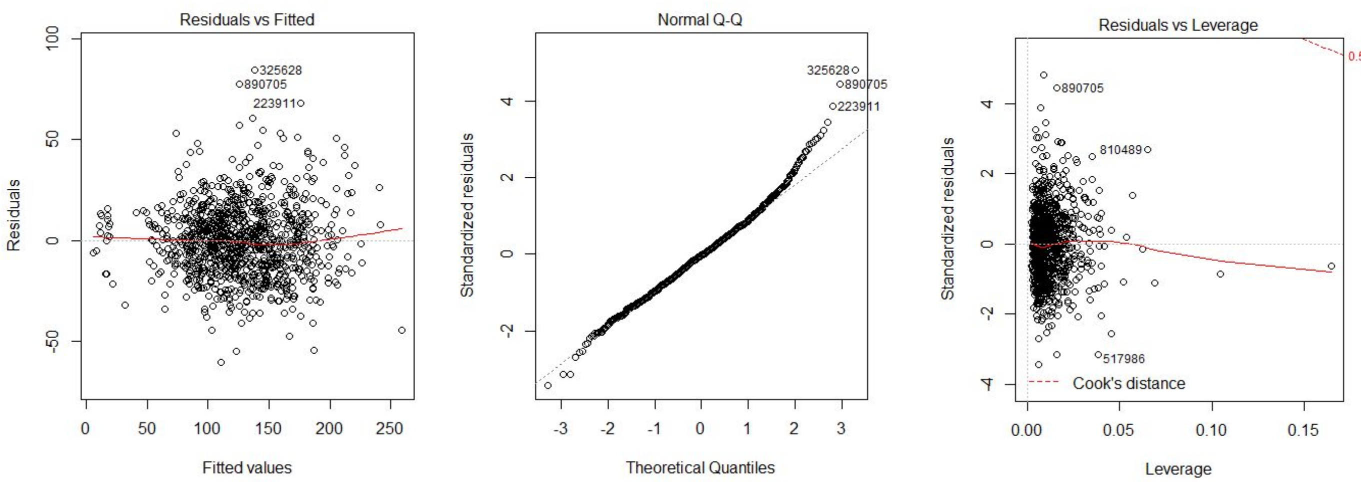






With the revised full model, a best subset matrix is necessary in order to compare adjusted R-squared, AIC, BIC, and Mallow’s Cp values across a different number of predictors. These four measure of fit all serve as indicators of how good a model is, with a high adjusted R-squared, low AIC and BIC, and a Mallow’s Cp close to the number of predictors (p-1) desired. Adjusted R-squared gives the model with 8 predictors, AIC gives the model with 7 predictors, BIC gives the model with 5 predictors, and Mallow’s Cp gives the model with 4 predictors. The adjusted R-squared of these four different models were found to be similar, so the model with the least predictors (4) was chosen to improve the interpretability of the model. The predictors ‘champlvl,’ ‘deaths,’ ‘kills,’ and ‘totminionskilled’ were were selected. Below is a data table with the adjusted R-squared values for the models selected by each measure of fit.

Adjusted R-squared	**Mallow’s C_p Crtierion**	**AIC**	**BIC**
.829	.826	.829	.827



Above are diagnostic plots for the best selected overall model. As shown in the diagnostic plots, the residuals in the Q-Q plot of the reduced model do not follow a straight line, and the Shapiro-Wilk test produces a statistically significant p-value to reject normality in the residuals, which is a requirement to ensure a reliable regression model given any sample of data. In the case of the reduced model, the residuals show a heavy right tail, indicating the existence of upper outliers in the dataset, and the fact that they may have a less significant relationship on the square root of the response than the regression model would suggest. Importantly, however, these points have been analysed not to be errors in the data, but rather a subset of the population of players that the model may not fully take into account. Using prior game experience, none of the observations appeared erroneous or abnormal, and it was deemed unnecessary to remove them.

A method that aims to be a compromise between removing outlying data points and treating all data points equally in the LSE regression is known as Robust Regression. There are several different variations of Robust Regression, but all aim to perform an iterative reweighted least squares algorithm to reweight the points to account for potential outlying observations.

For the iterative reweighted least squares algorithm (IRLS), the weights are calculated as a function of the slopes, but the slopes also need to be calculated in terms of the weights. Different versions of the algorithm start of with different weights, from which:

$$\beta_1 = [X'W_0X]^{-1}X'W_0Y \text{ , and } \sum_{i=1}^{j} w_j(y_j - x^T b)x_i^T = 0 \text{ , for iteration } j > 1 .$$

In general, the coefficient-matrix for the iterative reweighted least squares algorithm at time j becomes:

$$\beta_j = [X'W_{j-1}X]^{-1}X'W_{j-1}Y$$

The vector sequence $\{\beta_j\}$ and weight matrix sequence $\{W_j\}$ will be converging sequences for a particular element set of numbers in $\Re$. The specific form of IRLS employed in the Robust Regression method used for the League of Legends data set set the initial weights equal to an exponential function of the data. The IRLS algorithm was run to converge at values for the β which aimed to minimise the effect of extreme, but non-erroneous values in the data.

An ordinary least squares model and a robust regression model were obtained. For the ordinary least squares regression, the fitted model for the square root of total damage to champions is:

$$\sqrt{Y_{i-hat}} = -1.400 + 3.901x_{kills,i} + 0.154x_{minions,i} + 5.325x_{champlvl,i} + 1.917x_{deaths,i}$$

Based on the above IRLS, the fitted model for the square root of total damage to champions is:

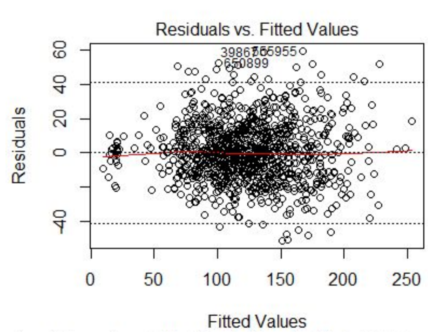
$$\sqrt{Y_{i-hat}} = 4.383 + 4.091x_{kills,i} + .1332x_{minions,i} + 4.872x_{champlvl,i} + 1.8363x_{deaths,i}$$

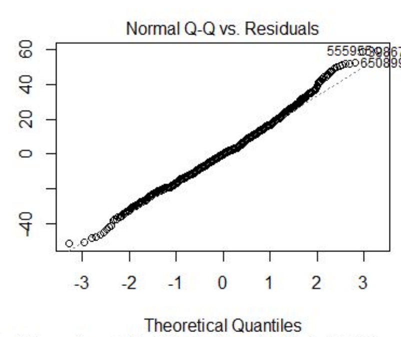
The coefficient of determination of this model is .8481, and all of the 4 predictors are significant, since the p-values of the respective βs are extremely small. However, before this model can be employed, the model assumptions need to be verified. As with the pre-IRLS model, the residuals v. fitted plot shows a horizontal band of points, with an equal height across all fitted values. Therefore, equivariance and linearity can both be assumed. However, while the heavy right tail has become less heavy in the new Q-Q plot for the IRLS model, the Shapiro-Wilk test still produces a (albeit lower) statistically significant p-value against normality of .0001981, so unfortunately, normality cannot be assumed from this model either.

```
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      -1.399832   3.043422   -0.46    0.646
kills             3.900532   0.152433   25.59   <2e-16 ***
totminionskilled  0.154091   0.007941   19.40   <2e-16 ***
champlvl          5.324702   0.265985   20.02   <2e-16 ***
deaths            1.917479   0.191280   10.02   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.8 on 995 degrees of freedom
Multiple R-squared:  0.8265,    Adjusted R-squared:  0.8258
F-statistic:  1185 on 4 and 995 DF,  p-value: < 2.2e-16
```

```
Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      4.382885   2.056161   2.132   0.0333 *
kills            4.091471   0.153542  26.647   <2e-16 ***
totminionskilled 0.133227   0.008825  15.097   <2e-16 ***
champlvl         4.872378   0.205690  23.688   <2e-16 ***
deaths           1.836320   0.174540  10.521   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Robust residual standard error: 16.57
Multiple R-squared:  0.8481,    Adjusted R-squared:  0.8475
Convergence in 11 IRWLS iterations
```

data: lm.robust.reduced.2$residuals
W = 0.99337, p-value = 0.0001981

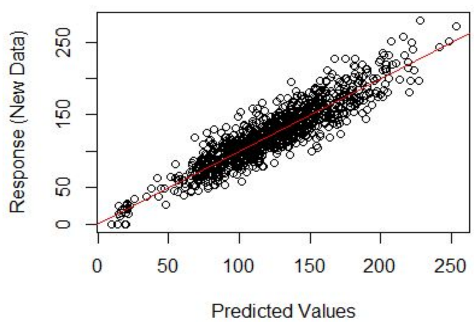




Note that the p-value of the Shapiro test did increase from .0001486 to .0001981. However, this is still a very slight increase, and well below the threshold of 5% to accept normality of the data with respect to the employed fitted model. It is evident that the robust regression algorithm employed did decrease the degree of nonnormality in the residuals, but nonnormality nonetheless persists.

The R-squared of this robust model was very high, at .8481. This shows that nearly 85% of the variation in the response sqrt(total damage) can be explained by the robust model. While a high coefficient of determination may signify that a model is very reliable to predict new responses, it is important to note that such high values may signify overfitting to this particular sample of the data. In that case, the model may not have a strong ability to generalise against different samples of the data.

To check that this robust model could be used as a generalisation of the relationship analysed between different subsets of the data, a new subset of 1000 data points was sampled from the initial data frame, where none of these 1000 data points coincided with the 1000 chosen from the initial fitting.



The goal of any model is to minimise the mean square prediction error (MSPE) across different data sets. One way to determine this is to rescale this value to provide a cross-validation R^2, which is defined to be the RSS between these sampled points and the previously fitted regression line, divided by the total sum squares (TSS) of these sampled points. The cross-validation R^2 using these new 1000 data points was nearly the same as before, at .8412, which indicates that the previously fitted robust model can reliably generalise predicted response values across new sampled data. As shown on the plot to the left, there is a strong linear mapping between the new predicted values (from the 1000 new sampled data) and the actual response of these sampled values.

Results

In the end, two models were selected: the simple model with 'goldearned' as the sole predictor, and the robust regression model. The following is the fitted simple model with gold earned:

$$\sqrt{Y_{i-hat}} = 15.35 + .00975\, x_{goldearned,i}$$

Thus, for every extra coin of gold earned, on average, the square root of total damage to champion increases by .00975. This relationship is statistically significant. Although this appears to be a trivial rise in square root of total damage, such a small value is in fact the result of scaling differences between the response and 'goldearned' predictor. Typically, gold is earned in increments far greater than one coin, and consequently, a collection of earned coins can have a far greater impact in inflicting damage than one might intuit. Hence, a more reasonable interpretation of this model suggests that for every 1,000 gold coins earned, the mean square root of total damage dealt climbs by 9.75 points.

Even with the IRLS model generated by the Robust Regression, the Shapiro test violates normality. Importantly, however, of the four assumptions required to satisfy reliable, unbiased estimates for a regression model slope parameters, the normality requirement is the least severe. It is especially nonsevere in the case where the total dataset size is large (over 300), where the Shapiro test may overplay the estimate of a larger number of outliers that nonetheless forms a lower percentage of the overall data. Besides the slightly heavy right tail, the residuals on the Q-Q plot nearly perfectly follow a linear line, showing that within 2 standard deviations of the mean of the predictor space (or about 95% of the data), accurate inferences can be drawn from the fitted model above.

The fitted multiple regression robust model was:

$$\sqrt{Y_{i-hat}} = 4.383 + 4.091x_{kills,i} + .1332x_{minions,i} + 4.872x_{champlvl,i} + 1.8363x_{deaths,i} .$$

From this model, one can see that kills, total damage to minions, champ level, and deaths all have a positive relationship with respect to total damage caused, and all of these relationships are significant. One would expect the square root of total damage to increases by 4.091 if the kills increased by one, and the square root of the damage to increase by .1332 for each minion killed. Additionally, the square root of the damage is expected to increase by 4.872 for every champ level gained, and is expected to increase by 1.8363 for every death. These four values are therefore the four most useful metrics to reliably predict total damage caused. Killing objectives are therefore a good time investment, as is indicated by the model above. However, a proactivity to vision control will likely not have any significant impact on the damage caused, as the respective predictor, vision score, was not included in the final reduced MLR model and showed a statistically insignificant relationship.

Conclusions/Discussions

Considering there were 12 potential predictors of a player's damage to champions in a game, we ended up with a relatively smaller amount of predictors. As a concept, using total damage to champions as the response variable was meant to serve as the closest measure to a player's success in the game in doing analysis. Often in the League of Legends community, many high-level players stress the importance of killing minions, damaging objectives, and increasing vision score as a means to succeed in the game, rather than mindlessly chasing after kills. However, our models suggest that player kills and minion kills are significant in predicting damage done to champions, while damage to objectives and vision score are not. Hence, perhaps the best advice emphasized in the community is that the minions killed leads to more successful game.

It was unexpected to find that deaths had a positive relationship with total damage to champions, as a higher number of deaths usually means poor performance in the game, while a higher damage dealt to champions usually means high performance in the game. A possible explanation for this phenomenon is some players will constantly search for kills, do a great deal of damage, and end up getting the fight turned on them and dying in the process.

Regarding the gold earned in the game as a predictor for damage dealt to champions, the model demonstrated a significant positive relationship, contradicting the original thought that players sitting on gold and not spending it would cause the gold earned to not be related to damage dealt to champions. However, more gold allows users to purchase items which increase their inflicted damage. This translates into more damage dealt to champions, and thus more success in the game.

These models' results may indicate two conclusions. The first conclusion is that success can be measured in other ways than how much damage a player deals to champions; for example, there are role players in the game called support, and their success is based purely on vision score and attempting to get their teammates ahead in total map vision, instead of doing damage. This would likely explain how vision score is not highly relate to damage done to champions. However, assuming total damage to champions is a good metric by which to measure game success, these results could simply indicate that vision score and damaging objectives are not as important as many players make them out to be. This could interest many players, as some players believe it is a waste of time and gold to concern themselves with vision control, when they could be spending gold to directly boost their damage. Due to game experience and the complexity of the issue at hand, the first conclusion is most likely closer to the truth.

To further improve the model's predictive power, future researchers may wish to implement a layered neural network. Such a model would more effectively leverage the vast amount of data observations and predictors in the original dataset, without compromising the model's ability to generalize to new data (overfitting). However, complex machine learning methodologies would obfuscate the parsimony and interpretability of the model's components. Hence, this form of analysis should be used in conjunction with other statistical tools to formulate a more holistic evaluation of the data.

Team Member Contributions

Ishan was responsible for choosing methods and performing regression analysis in R, as well as creating the presentation. He also contributed to the 'Abstract,' 'Methods,' 'Results,' and 'Conclusion' sections. Deric was responsible for finding the data set, contributing context to the group, and contributing to 'Abstract,' 'Introduction,' 'Methods,' 'Results,' and 'Conclusion.' He also contributed to the presentation. Atharv was responsible for writing the 'Exploratory Analysis,' explaining the robust regression process in 'Methods,' and 'Results,' and formatting the content in this paper in a neat and legible manner, as well as contributing to the diagnostics and robust regression sections in the presentation.

References

https://www.kaggle.com/paololol/league-of-legends-ranked-matches
https://stats.idre.ucla.edu/r/dae/robust-regression/