



League of Legends Analysis

BLAKE CASEY, SOPHIE PANAGROSSI,
BENJAMIN TISINGER, DEVENE BELLAMY,
CRISTOBAL GAVALDON

IST-687

Intro/ Fields

- League of Legends (LoL): a team-based strategy game consisting of two teams, each containing five players/people
- There are over 100 characters (champions) to choose from, and each team will fight one another, and pursue objectives to gain advantages to assist in achieving victory.
- Every champion begins at level 1 at the start of each match
- The objective of this investigation is to use statistical analysis to help us determine whether certain champions are inherently better picks than others.
- Champions can take on 1 of 5 roles/classes. These roles are as follows:
 - Top, jungle, middle, bottom, utility

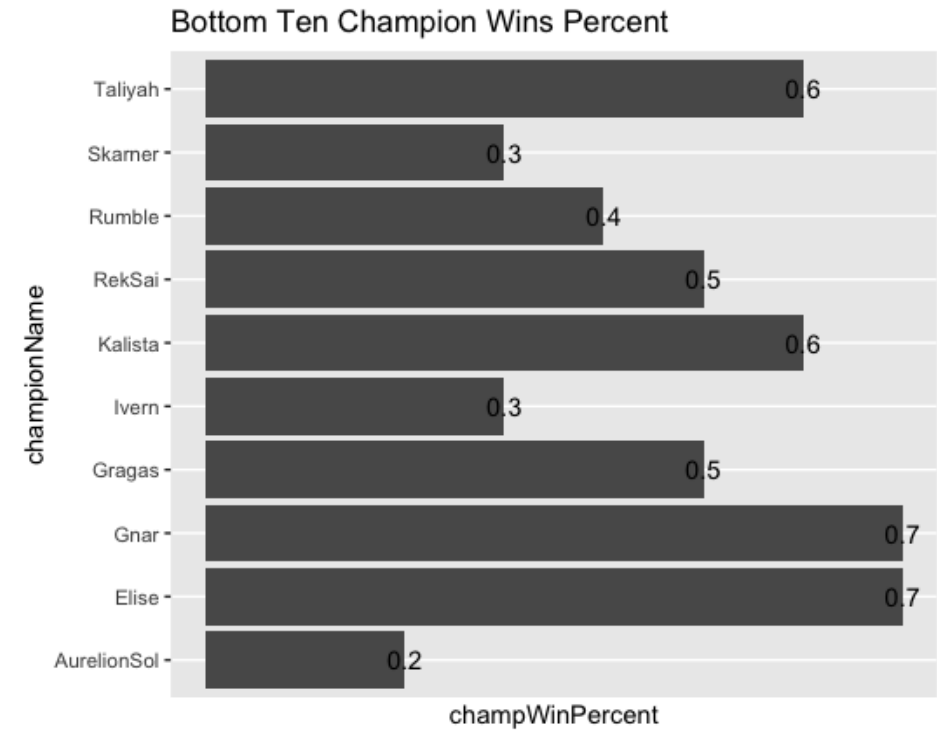
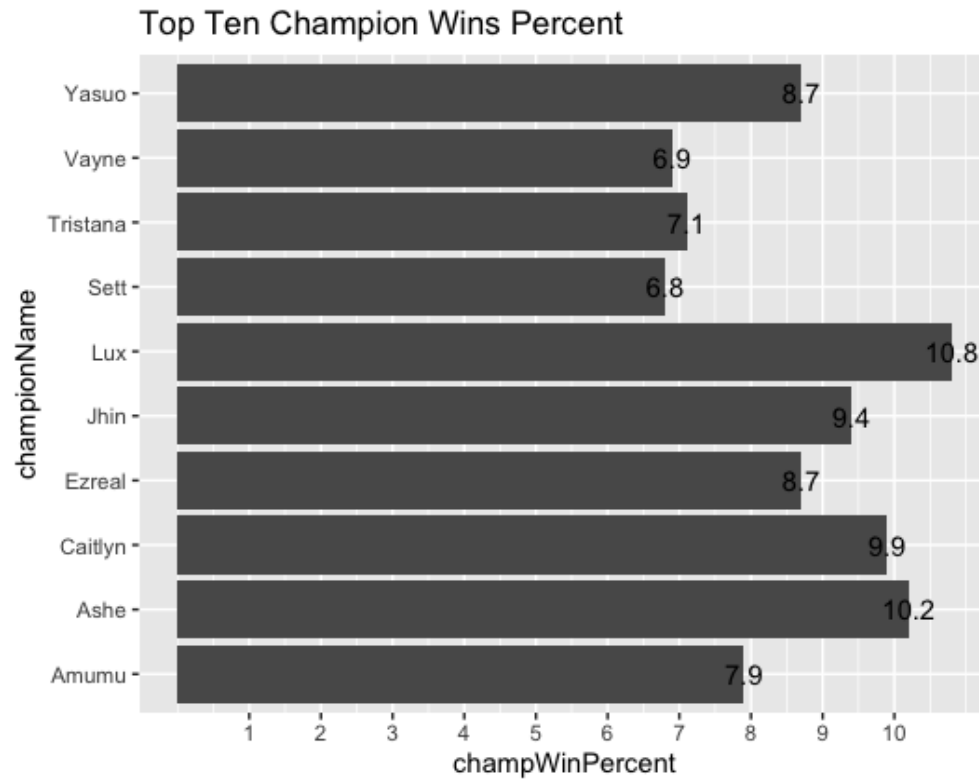
Data source

- <https://www.kaggle.com/datasets/prestonrobertson7/league-of-legends-data-9292022>
- Original dataset contains 69 columns and 10001 observations
- Filtered dataset used for this analysis contains 15 columns and 10000 observations

Methods

- Linear regression
 - `lm()`
 - Ex. `lm(champEndedInSur~ champTotalDamTaken, champstats)`
 - $\text{champEndedInSur} = b_0 + b_1 * \text{champTotalDamTaken}$
- Correlation test
 - `Cor.test()`
 - Ex. `cor.test(champstats$champTotalDam, champstats$win, method = c("pearson"))`
- Tidyverse functions
 - `%>%`
 - Ex. `champstats <- champstats %>%
group_by(championName) %>%
summarise(win = sum(win))`
- Ggplot2
 - `Ggplot()`
 - `ggplot(champstats, aes(x=win, y=champTotalDam))+
geom_point()+
geom_smooth()`

Question 1



Question 2

```
-----
lm(formula = champTotalDam ~ win, data = champstats)

Residuals:
    Min       1Q   Median       3Q      Max
-1064501 -112660    8978   149546  1004796

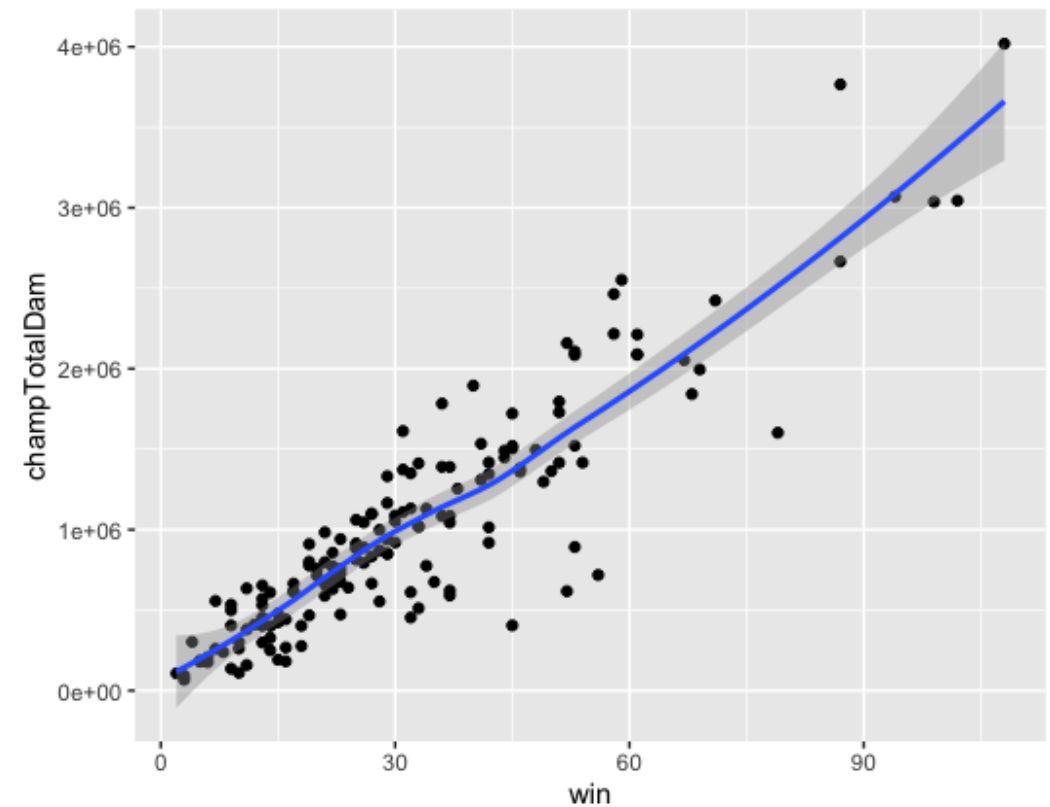
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   15719     44966   0.35   0.727
win          31560     1176   26.84  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 306300 on 154 degrees of freedom
Multiple R-squared:  0.8238,    Adjusted R-squared:  0.8227
F-statistic: 720.2 on 1 and 154 DF,  p-value: < 2.2e-16
```

Pearson's product-moment correlation

```
data: champstats$champTotalDam and champstats$win
t = 26.836, df = 154, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8753656 0.9318807
sample estimates:
      cor
0.9076536
```

Correlation Between Total Damage and Win Rate



Question 3

```
Call:
lm(formula = champEndedInSur ~ champTotalDamTaken, data = champstats)
```

Residuals:

Min	1Q	Median	3Q	Max
-21.385	-3.969	-0.449	3.184	36.895

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.221e+00	1.062e+00	3.033	0.00284 **
champTotalDamTaken	1.083e-05	6.266e-07	17.277	< 2e-16 ***

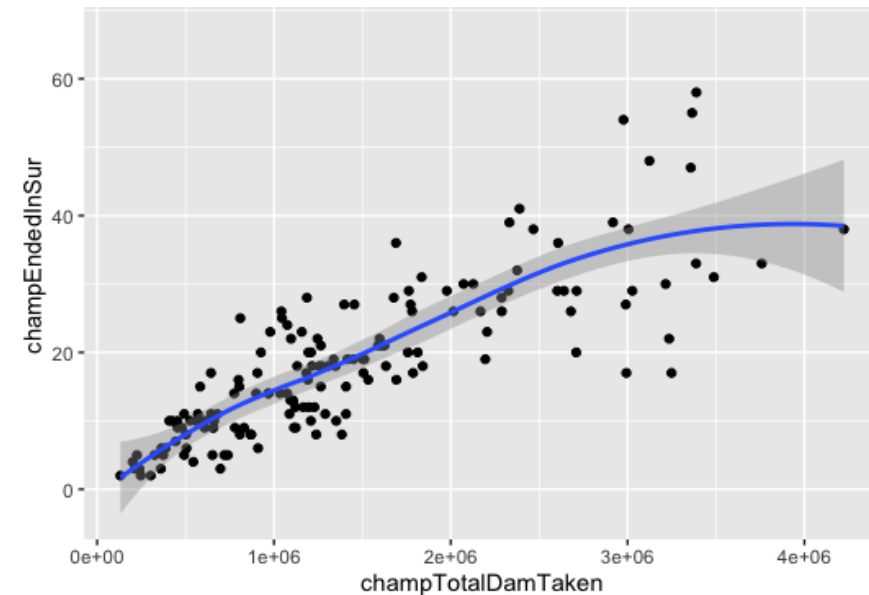
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.025 on 154 degrees of freedom
Multiple R-squared: 0.6597, Adjusted R-squared: 0.6574
F-statistic: 298.5 on 1 and 154 DF, p-value: < 2.2e-16

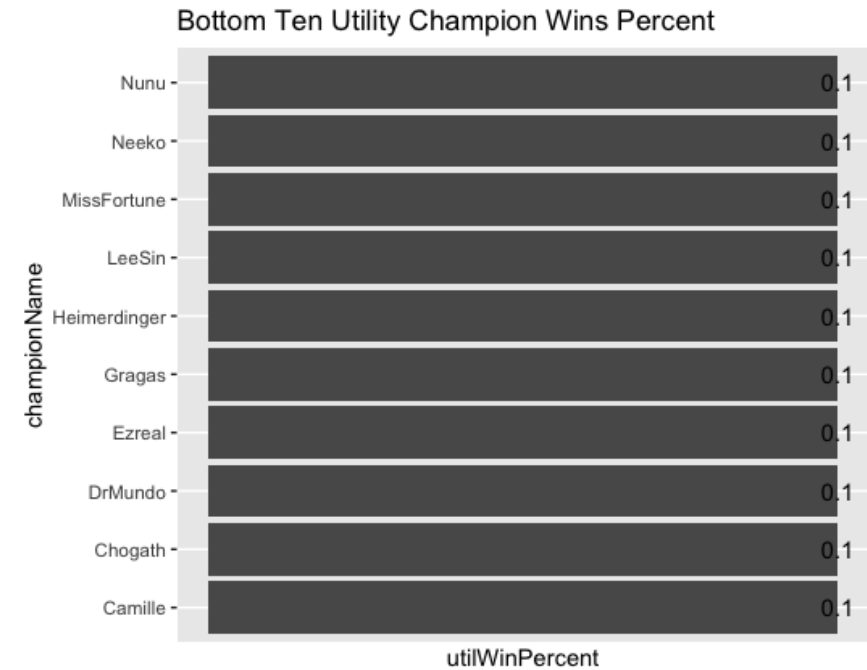
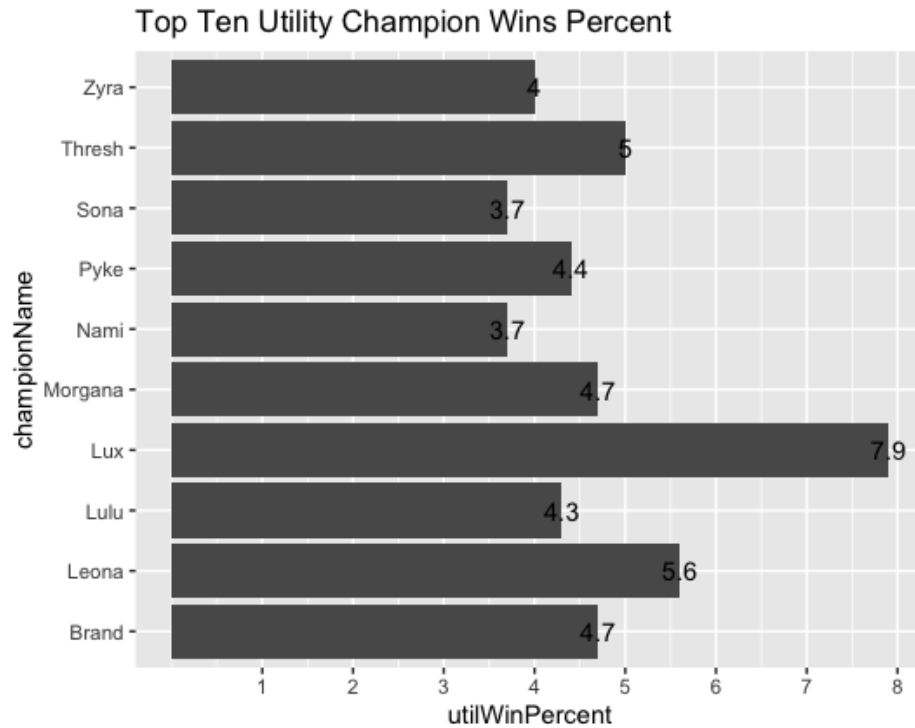
Pearson's product-moment correlation

```
data: champstats$champTotalDamTaken and champstats$champEndedInSur
t = 17.277, df = 154, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.7508824 0.8596188
sample estimates:
cor
0.8121895
```

Correlation b/w Damage Received and Probability of Game Ending in Surrender

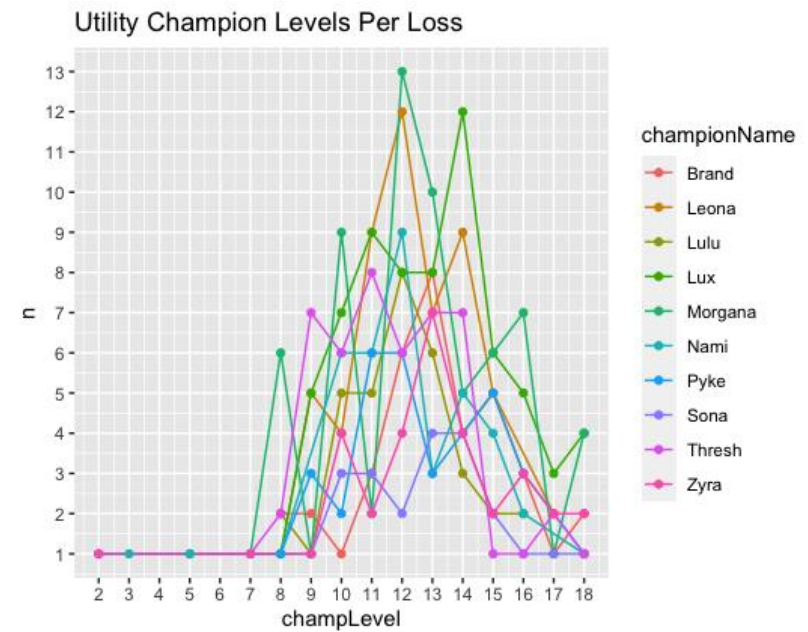
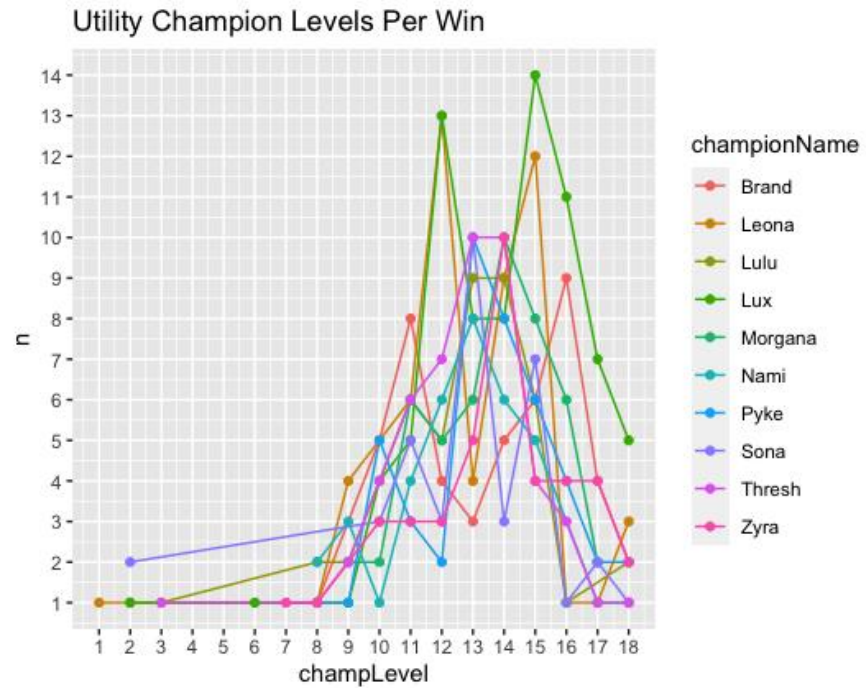


Question 4



Question 5

What is the win/loss ratio of utility champions per level?



Conclusion

- Lux is statistically shown to be a better pick than other characters.
- Odds of winning against Lux is best before level 15.
 - Most losses for Lux are at level 14
- After level 15, Lux wins more than all other utility champions.
- Aurelion Sol has the lowest number of wins for all champions
- Camille has the lowest number of wins for **utility** champions
- The limitations of this project had a focus on utility champions, however, if we were to expand the scope and include all champions and all roles, it may reveal how teams could strategically compose their roster to emphasize certain points of the game (such as aiming to win early in a short game or play a longer game and win late).
- More forward-looking ideas would include normalizing the dataset to only include games that the specific champion participated in.

Q&A

Thank you!