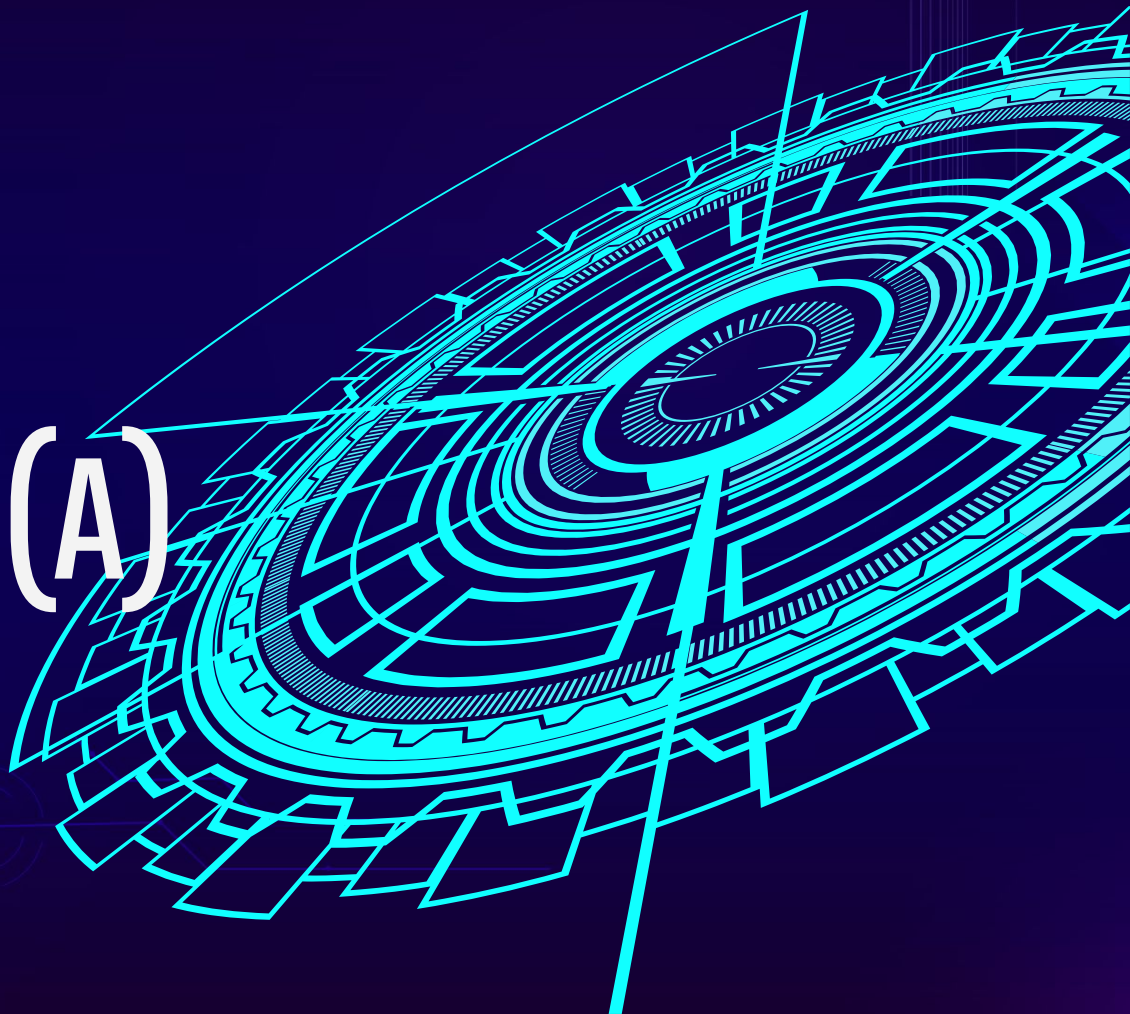
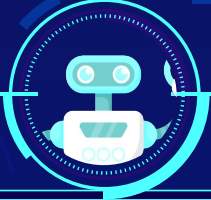


How to Win League of Legends 101(A)

By the Non-Constant
Variants

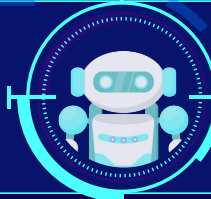


OUR TEAM



Brandon Louie

2nd Year Statistics Major



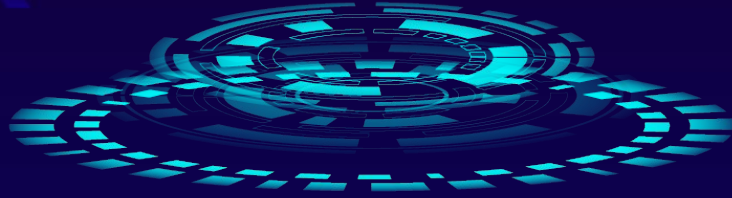
Emma Chi

2nd Year Statistics Major



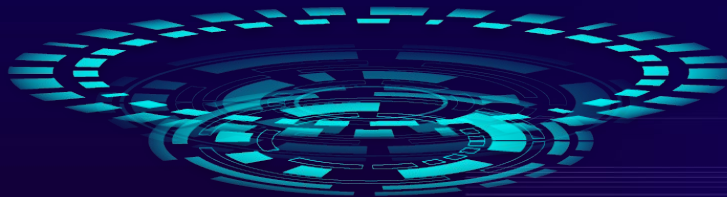
Tiffany Minami

2nd Year Statistics Major



INTRODUCTION

In this popular multiplayer online battle arena game, two teams face off, utilizing a variety of roles, lanes, minions, and bonuses to win the title of Champion. In this presentation, we will use real game data to analyze the performance of the blue team and demonstrate why our model explaining the difference between the gold acquired by the blue and red team is accurate.



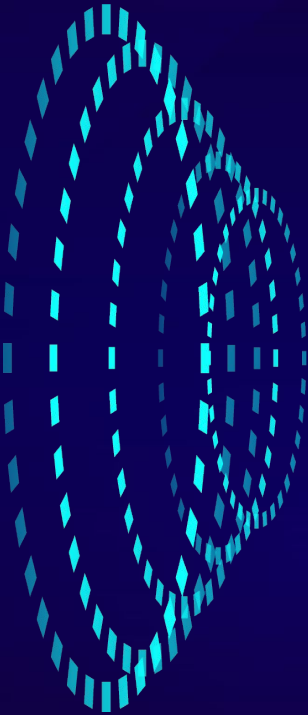


THE DATA

The data has 24912 League of Legends matches and 55 features.

There are 53 predictors.

The response is goldDiff, which is the Blue team gold difference.



METHODOLOGY AND RESULTS

descriptive analysis,
assessment of multicollinearity,
initial model and evaluation,
and final model and evaluation

METHODS

Multicollinearity

We used color maps to check for multicollinearity.

BIC

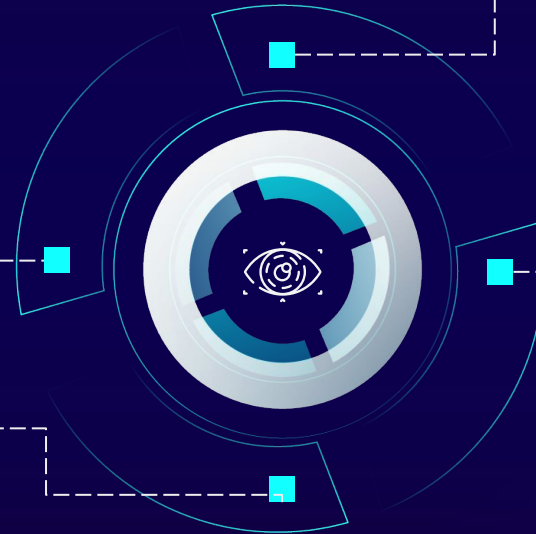
We determined the models with the smallest Bayes Information Criterion and selected the best model based on these values.

t-tests

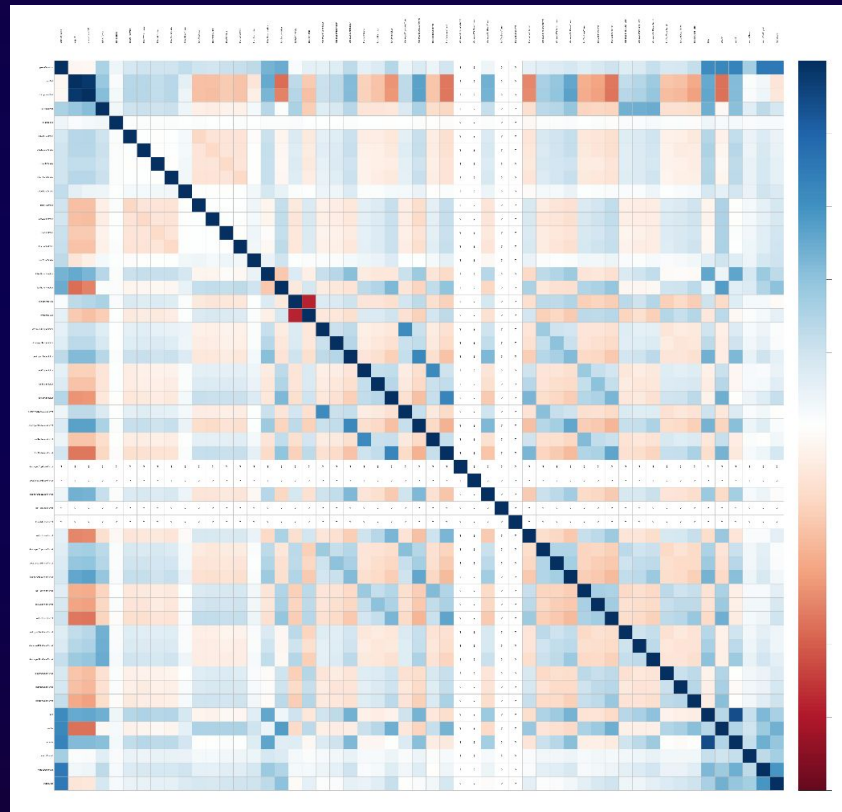
We conducted t-tests for each regression coefficient to determine which predictors have significant association with the blue team's gold difference when all other predictors are in the model.

Regression Diagnostics

Looking at the residual plots, we conducted residual analysis and accounted for pesky points to assess the fit of our final model.



COLOR MAP



After testing for multicollinearity and removing variables, we fit our initial model with 37 predictors.

These predictors returned collinearity values greater than $|0.6|$, so we will remove them:

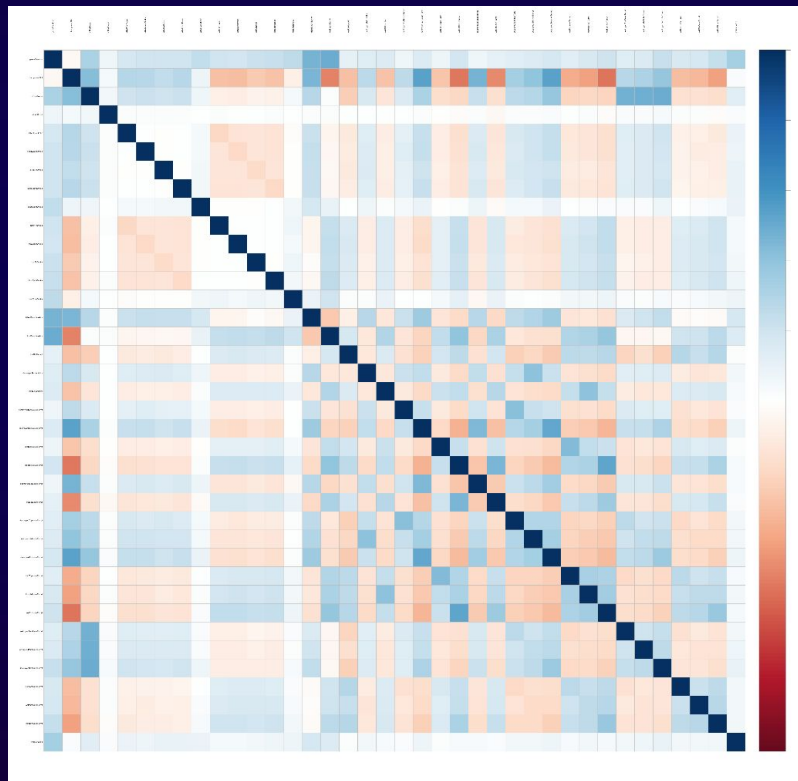
- gameDuration
- expDiff
- killedRiftHerald
- destroyedTopInhibitor
- destroyedBotInhibitor
- lostTopInhibitor
- lostBotInhibitor
- kills
- deaths
- assists
- wardsDestroyed
- wardsLost

These predictors returned NA values for their collinearity, so we will also remove them:

- destroyedTopBaseTurret
- destroyedMidBaseTurret
- lostTopBaseTurret
- lostMidBaseTurret

FIT THE INITIAL MODEL

UPDATED COLOR MAP WHEN DESIGNATED PREDICTORS ARE REMOVED



F-Test

We conducted a full model F-test to determine whether the predictors in the model have a significant explanatory power. The p-value is calculated using a reference distribution F with 37 and 24874 degrees of freedom. The p-value $< 2.2e-16$ determined by this test was less than $\alpha = 0.001$, so we rejected the null hypothesis and determined that at least one of the predictors has a significant association with the blue team's gold difference, however at this point we are not sure which predictor.

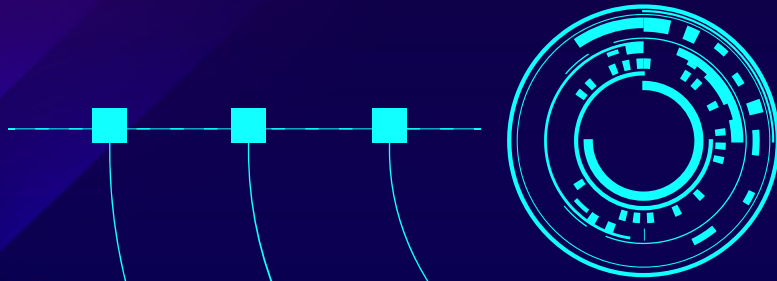
Residual standard error: 2307 on 24874 degrees of freedom
Multiple R-squared: 0.9409, Adjusted R-squared: 0.9408
F-statistic: 1.07e+04 on 37 and 24874 DF, p-value: $< 2.2e-16$

t-test

We conducted a t-test for each individual β_j 's which determines the significance in association when adding X_j to a model where the other predictors are present. Using $\alpha = 0.001$, the variable wardsPlaced has a p-value of 0.00775, which is greater than α . Thus, wardsPlaced does not have a significant association with the blue team's gold difference when all the other predictors are in the model.

destroyedTopInnerTurret	867.2754	38.5203	22.515	< 2e-16	***
destroyedMidInnerTurret	940.5953	38.0737	24.705	< 2e-16	***
destroyedBotInnerTurret	724.7731	38.3308	18.908	< 2e-16	***
lostTopInnerTurret	-886.0552	39.5611	-22.397	< 2e-16	***
lostMidInnerTurret	-947.8275	38.3729	-24.700	< 2e-16	***
lostBotInnerTurret	-732.3331	38.8391	-18.856	< 2e-16	***
destroyedTopOuterTurret	437.1985	34.2296	12.773	< 2e-16	***
destroyedMidOuterTurret	552.7431	34.4494	16.045	< 2e-16	***
destroyedBotOuterTurret	492.1667	36.0730	13.644	< 2e-16	***
lostTopOuterTurret	-667.6252	32.6561	-20.444	< 2e-16	***
lostMidOuterTurret	-707.2468	32.6272	-21.677	< 2e-16	***
lostBotOuterTurret	-754.8115	34.6461	-21.786	< 2e-16	***
wardsPlaced	-0.5560	0.2088	-2.663	0.00775	**

VARIABLE SELECTION



Best-Subset Selection

a method that assesses all submodels to find the best model

Forward Selection

a method that evaluates models by adding one variable at a time, creating nested models

Backward Selection

a method that evaluates models by removing one variable at a time, creating nested models

BIC FACTORS OF SUBMODELS

Bayes Information Criterion (BIC) is a commonly used model selection criterion. The best model corresponds with the smallest BIC value.

For our project, the smallest BIC value is -70066.82, which corresponds to a model with 36 predictors. Therefore, we will use the 36 predictors in the submodel evaluated by the best-subset selection.

BIC values of submodels

[1]	-51208.82	-54025.11	-56381.55	-58389.51	-59875.55	-61391.13	-62725.69	-63792.11	-64711.18	-65440.82
[11]	-65980.34	-66437.78	-66858.64	-67244.02	-67624.05	-68008.51	-68312.83	-68565.19	-68790.82	-68969.60
[21]	-69128.48	-69287.38	-69423.11	-69527.83	-69611.18	-69686.29	-69786.02	-69882.73	-69954.64	-70024.42
[31]	-70037.82	-70042.47	-70047.23	-70054.25	-70061.00	-70066.82				

Minimum BIC value

[1] -70066.82

REGRESSION DIAGNOSTICS

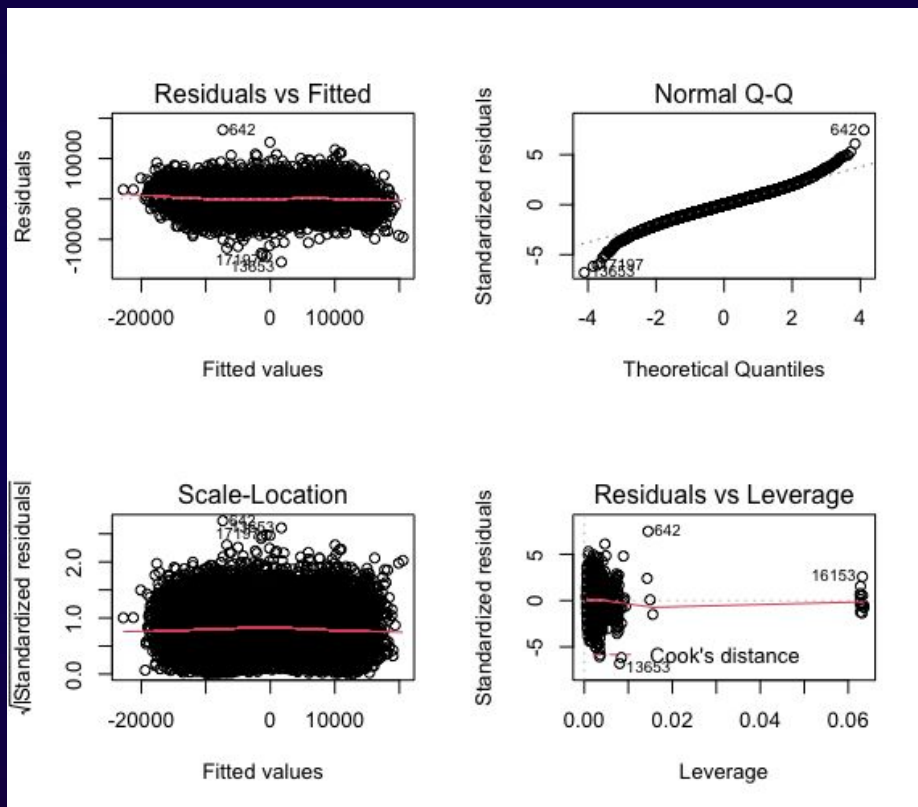
RESIDUAL ANALYSIS

We must check if the residuals have constant variance, are around 0, and follow a normal distribution as they may make the model fall apart.

PESKY POINTS

We must check for high leverage points and outliers as they may influence the model.

RESIDUAL PLOTS



CURRENT PREDICTORS

champLevelDiff
isFirstTower
isFirstBlood
killedFireDrake
killedWaterDrake
killedAirDrake
killedEarthDrake
killedElderDrake
lostFireDrake
lostWaterDrake
lostAirDrake
lostEarthDrake
lostElderDrake
killedBaronNashor
lostBaronNashor
lostRiftHerald
destroyedMidInhibitor
lostMidInhibitor

destroyedTopNexusTurret
destroyedBotNexusTurret
lostTopNexusTurret
lostBotNexusTurret
destroyedBotBaseTurret
lostBotBaseTurret
destroyedTopInnerTurret
destroyedMidInnerTurret
destroyedBotInnerTurret
lostTopInnerTurret
lostMidInnerTurret
lostBotInnerTurret
destroyedTopOuterTurret
destroyedMidOuterTurret
destroyedBotOuterTurret
lostTopOuterTurret
lostMidOuterTurret
lostBotOuterTurret

CONDENSED MODEL

```

champLevelDiff
isFirstTower
isFirstBlood
killedDrake
lostDrake
killedBaronNashor
lostBaronNashor
lostRiftHerald
destroyedMidInhibitor
lostMidInhibitor
destroyedNexusTurret
lostNexusTurret
destroyedBotBaseTurret
lostBotBaseTurret
destroyedInnerTurret
lostInnerTurret
destroyedOuterTurret
lostOuterTurret

```

```

(Intercept)      2394.97    580.55   4.125 3.71e-05 ***
champLevelDiff   4307.47     20.67 208.419 < 2e-16 ***
isFirstTower     1272.70     60.28  21.113 < 2e-16 ***
isFirstBlood    -2570.57    581.31  -4.422 9.82e-06 ***
killedDrake       177.14     20.37   8.698 < 2e-16 ***
lostDrake        -199.47     20.45  -9.754 < 2e-16 ***
killedBaronNashor  139.19     35.05   3.972 7.16e-05 ***
lostBaronNashor  -458.98     35.05 -13.096 < 2e-16 ***
lostRiftHerald   -386.87     22.53 -17.169 < 2e-16 ***
destroyedMidInhibitor  593.54     43.44  13.663 < 2e-16 ***
lostMidInhibitor -542.26     43.39 -12.498 < 2e-16 ***
destroyedNexusTurret  664.42     30.97  21.453 < 2e-16 ***
lostNexusTurret  -600.44     31.24 -19.218 < 2e-16 ***
destroyedBotBaseTurret  456.88     24.36  18.756 < 2e-16 ***
lostBotBaseTurret -442.63     24.48 -18.078 < 2e-16 ***
destroyedInnerTurret  836.75     21.01  39.836 < 2e-16 ***
lostInnerTurret   -849.41     21.08 -40.304 < 2e-16 ***
destroyedOuterTurret  488.04     21.09  23.137 < 2e-16 ***
lostOuterTurret   -700.68     17.96 -39.024 < 2e-16 ***

```

```

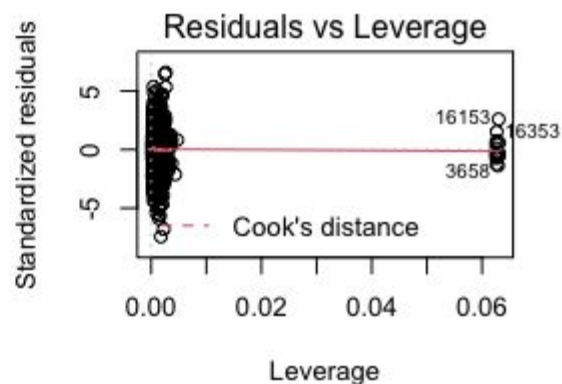
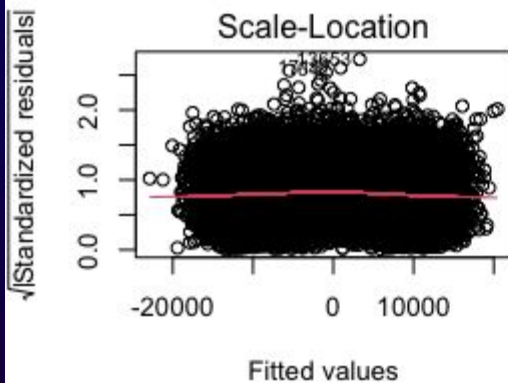
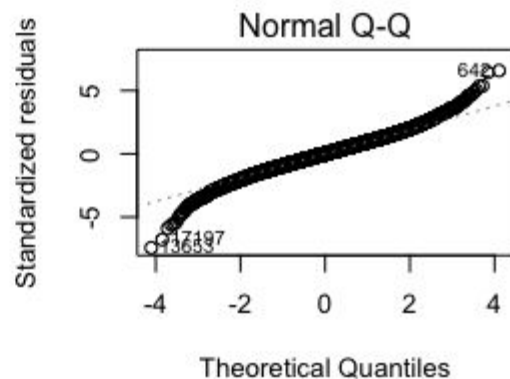
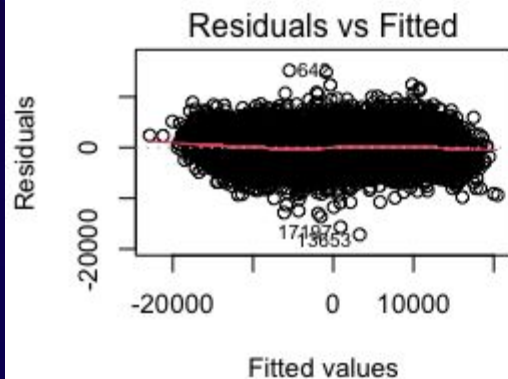
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 2314 on 24893 degrees of freedom
Multiple R-squared:  0.9405,    Adjusted R-squared:  0.9404
F-statistic: 2.186e+04 on 18 and 24893 DF,  p-value: < 2.2e-16

```



STRENGTHS

✓ We were able to reduce our model from 53 predictors to 18 predictors.

Balanced

Model is Interpretable



Passed the Regression Diagnostics so the conclusions for our model are reliable and valid.

Includes All Significant Predictors

RECOMMENDATIONS



INTERPRETABILITY

In our project, we combined similar variables to reduce the number of predictors, as this will allow us to explain the model better.



PREDICTABILITY

We want a model that is able to predict the response. Without prediction power, the model is basically useless. We want to make conclusions with our model.



THANK YOU