

League of Legends Snowballing Analysis

By Nicole Quentin, and Charles Sandahl

Introduction



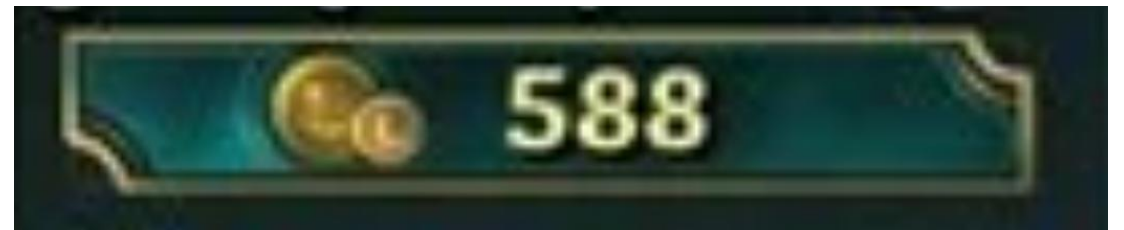


Background Information

- What is League of Legends?
- What is snowballing? Why does it matter?
 - How does snowballing occur?
 - Getting kills against the enemy team
 - Getting a lead in creep score / experience
 - Objectives

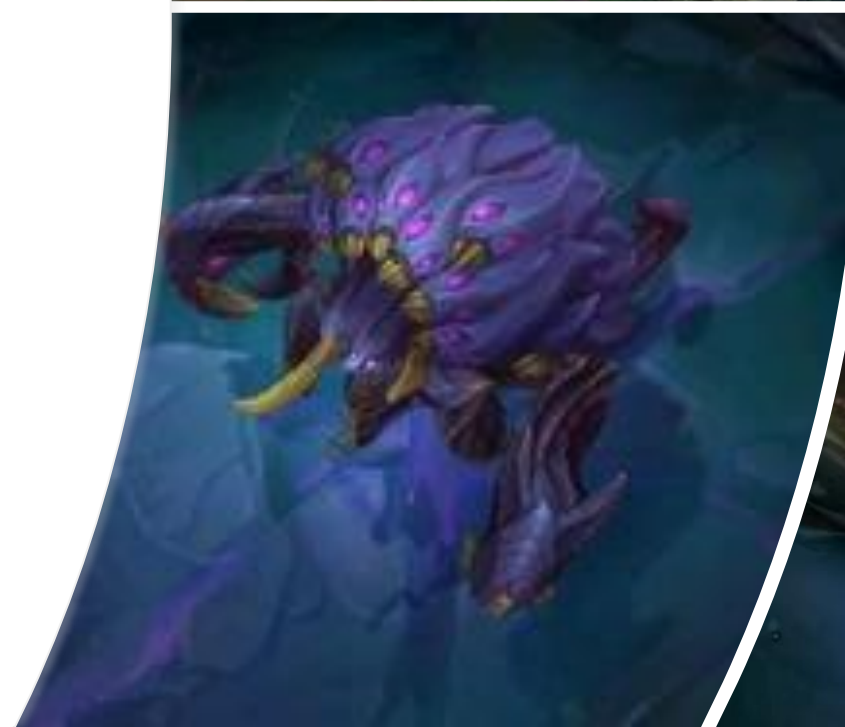
Background Information (cont.)

- Creep score increases by killing minions
 - They grant gold and experience!



Background Information (cont.)

- Creep score increases by killing minions
 - They grant gold and experience!
- What are objectives?
 - Dragon
 - Towers
 - Rift Herald
 - First Blood





ORACLE'S ELIXIR

ORACLESELIXIR.COM

ADVANCED ESPORTS ANALYTICS

Dataset

- Oracle's Elixir
 - 2025 season Analytics
- Filtered using Python Script



Research Question

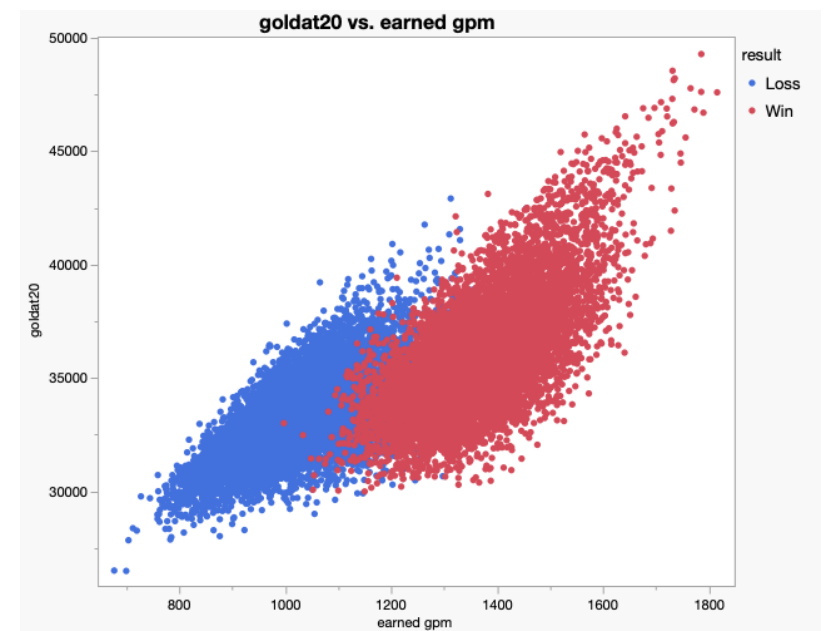
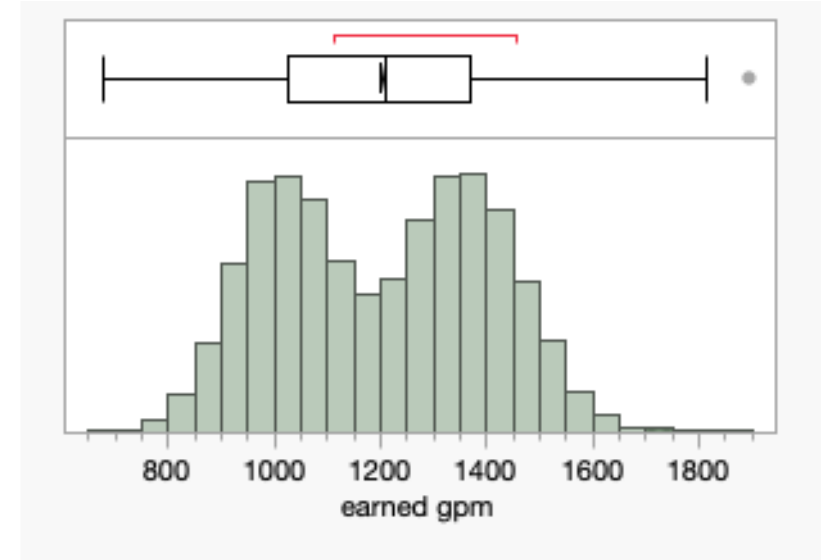
Can we accurately predict how much gold a team has 20 minutes into a League of Legends game, by using variables measured at 10 minutes into the game?

Exploratory Data Analysis

A thick, hand-drawn style orange line that underlines the title "Exploratory Data Analysis". It starts under the 'E' and ends under the 's', following the baseline of the text.

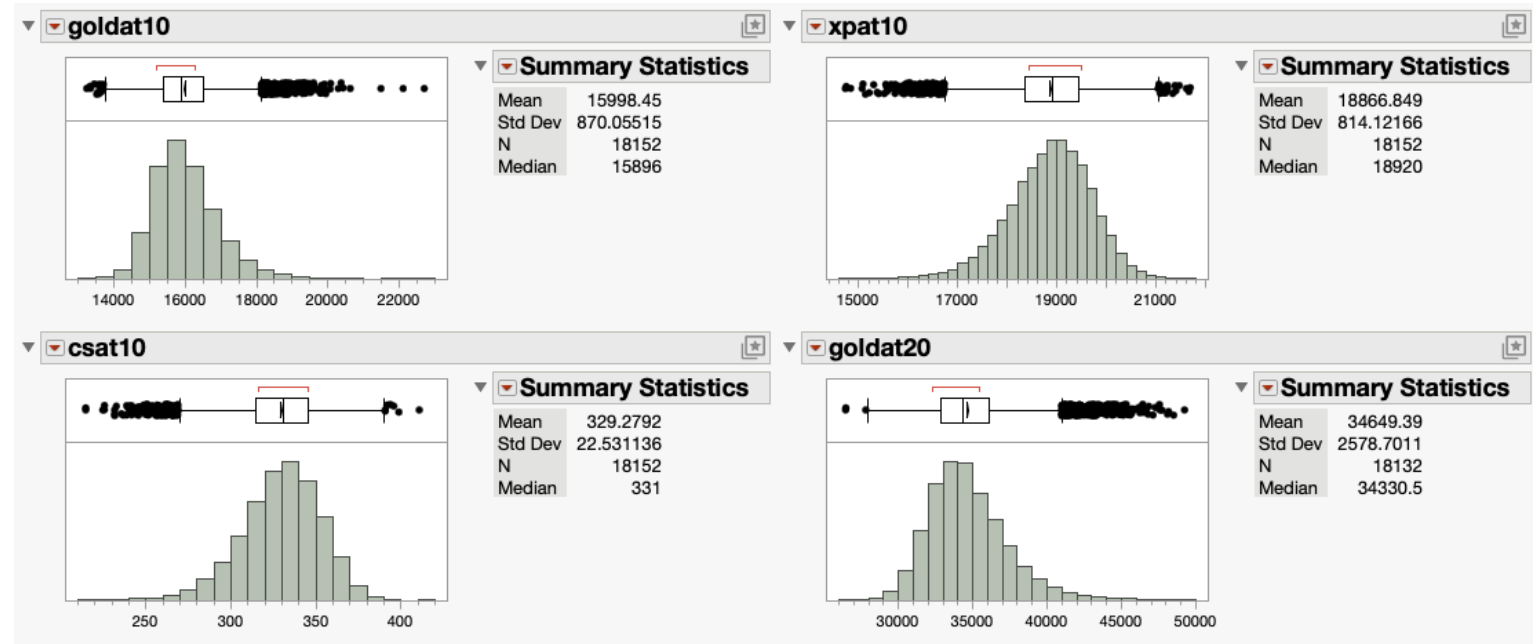
Exploratory Data Analysis

While creating histograms, we noticed one variable with a bimodal distribution:



Exploratory Data Analysis (cont.)

- Created histograms for continuous variables:



Methods

Checking For Multicollinearity

- XP at 10 & CS at 10 have a moderate positive correlation
- We kept XP at 10, since it has a higher R^2 value
 - $0.3666^2 > 0.1557^2$

Correlations

	goldat10	xpat10	csat10
goldat10	1.0000	0.3666	0.1557
xpat10	0.3666	1.0000	0.7602
csat10	0.1557	0.7602	1.0000

Stepwise Regression

- Stepwise Regression with 70/20/10 and 60/20/20 cross-validation had these two models in common:
- 5-Fold Cross Validation removed xpat10

70/20/10 Stepwise Predictors	R^2	BIC	C_p
goldat10, xpat10, firsttower, firstdragon, firstherald	0.5374	~225564	5.0029
goldat10, firsttower, firstdragon, firstherald	0.5374	~225562	3.0691

60/20/20 Stepwise Predictors	R^2	BIC	C_p
goldat10, xpat10, firsttower, firstdragon, firstherald	0.5440	~193240	7
goldat10, firsttower, firstdragon, firstherald	0.5440	~193238	5.0005

Predictors
goldat10, xpat10, firsttower, firstdragon, firstherald
goldat10, firsttower, firstdragon, firstherald

Model 1

- For Model 1, we tried a first order model.
 - $R_a^2 \approx 0.5427$, $s = 1743.795$, $2s = 3487.59$

Parameter Estimates

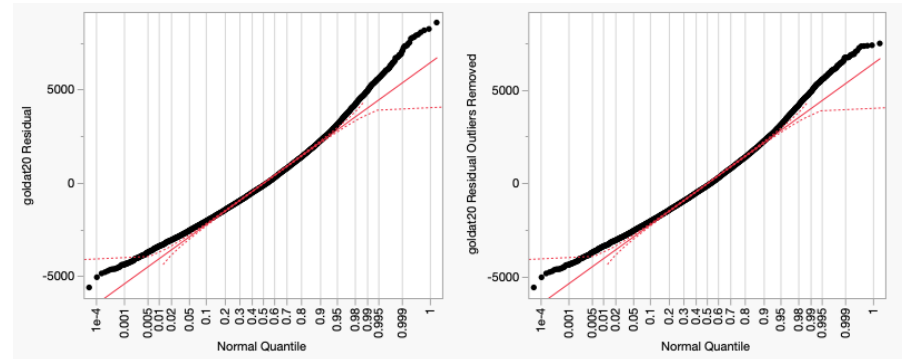
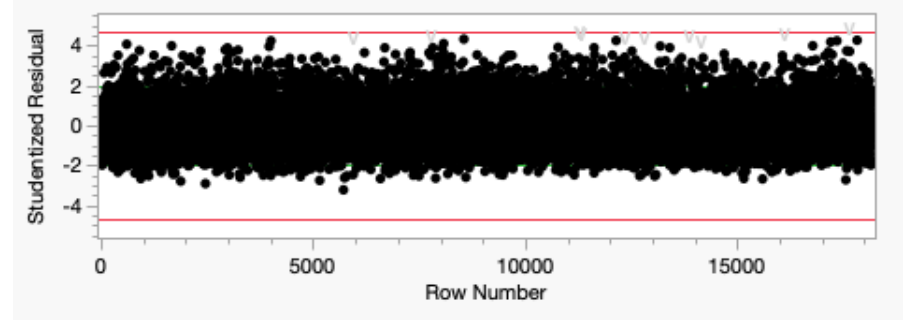
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	8265.0626	251.9749	32.80	<.0001*
goldat10	1.5727706	0.016103	97.67	<.0001*
firsttower	1220.0515	28.19724	43.27	<.0001*
firstdragon	218.54918	25.94483	8.42	<.0001*
firstherald	1010.8811	27.5548	36.69	<.0001*

Summary of Fit

RSquare	0.542814
RSquare Adj	0.542713
Root Mean Square Error	1743.795
Mean of Response	34649.39
Observations (or Sum Wgts)	18132

Model 1 (cont.)

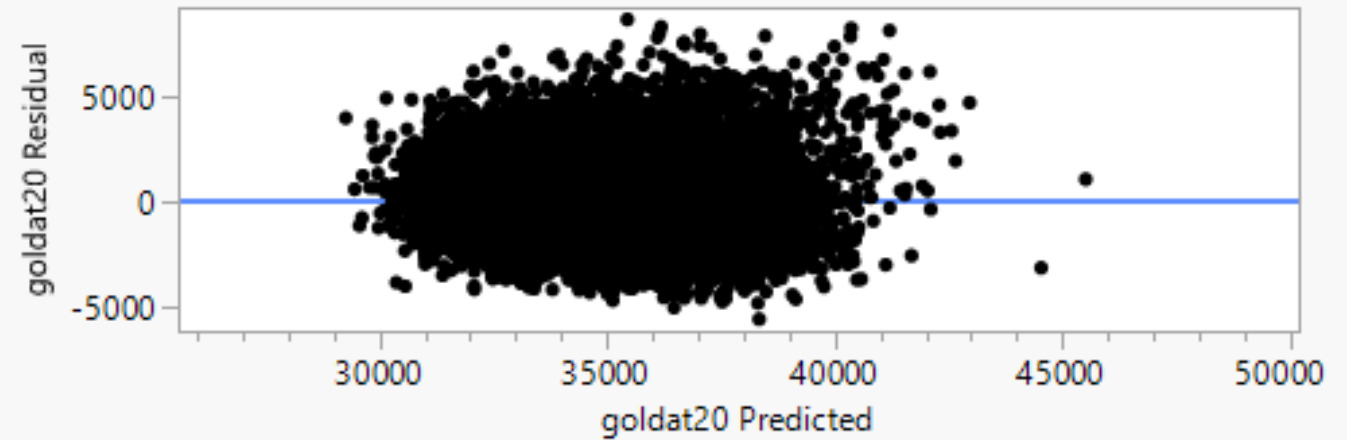
- Removing potential outliers didn't have a significant effect



Model 1 (cont.)

- Model 1 met all regression assumptions
 - Unsure about unequal variance

Residual by Predicted Plot



Model 2

- Added a second order term to Model 1:
 - $E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_2 + \beta_4 x_3 + \beta_5 x_4$
 - $x_1 = \text{goldat10}$
 - $x_2 = \begin{cases} 1 & \text{if team took first tower} \\ 0 & \text{if not} \end{cases}$,
 - $x_3 = \begin{cases} 1 & \text{if team took first dragon} \\ 0 & \text{if not} \end{cases}$,
 - $x_4 = \begin{cases} 1 & \text{if team took first herald} \\ 0 & \text{if not} \end{cases}$
- $R_a^2 \approx 0.547, s = 1735.568, 2s = 3471.136$

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	45979.632	2876.196	15.99	<.0001*
goldat10	-3.06387	0.352618	-8.69	<.0001*
goldat10*goldat10	0.000142	1.079e-5	13.16	<.0001*
firsttower	1237.8132	28.09663	44.06	<.0001*
firstdragon	211.40195	25.82813	8.18	<.0001*
firstherald	1019.7993	27.43317	37.17	<.0001*

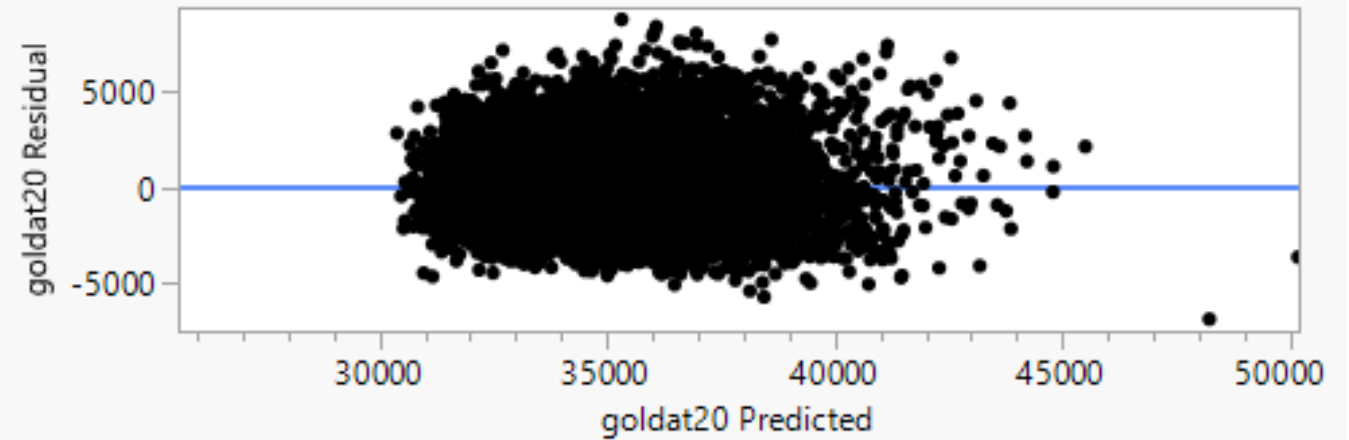
Summary of Fit

RSquare	0.547143
RSquare Adj	0.547018
Root Mean Square Error	1735.568
Mean of Response	34649.39
Observations (or Sum Wgts)	18132

Model 2

- Model 2 meets the Lack of Fit and Normality assumptions
 - We're unsure about Unequal Variances

Residual by Predicted Plot



Model 3

- Added interactions between goldat10 and indicator variables to Model 2
 - $E(y) = \beta_0 + \beta_1x_1 + \beta_2x_1^2 + \beta_3x_2 + \beta_4x_1x_2 + \beta_5x_3 + \beta_6x_1x_3 + \beta_7x_4 + \beta_8x_1x_4$
 - $R_a^2 \approx 0.5486, s = 1732.585, 2s = 3465.17$

Summary of Fit

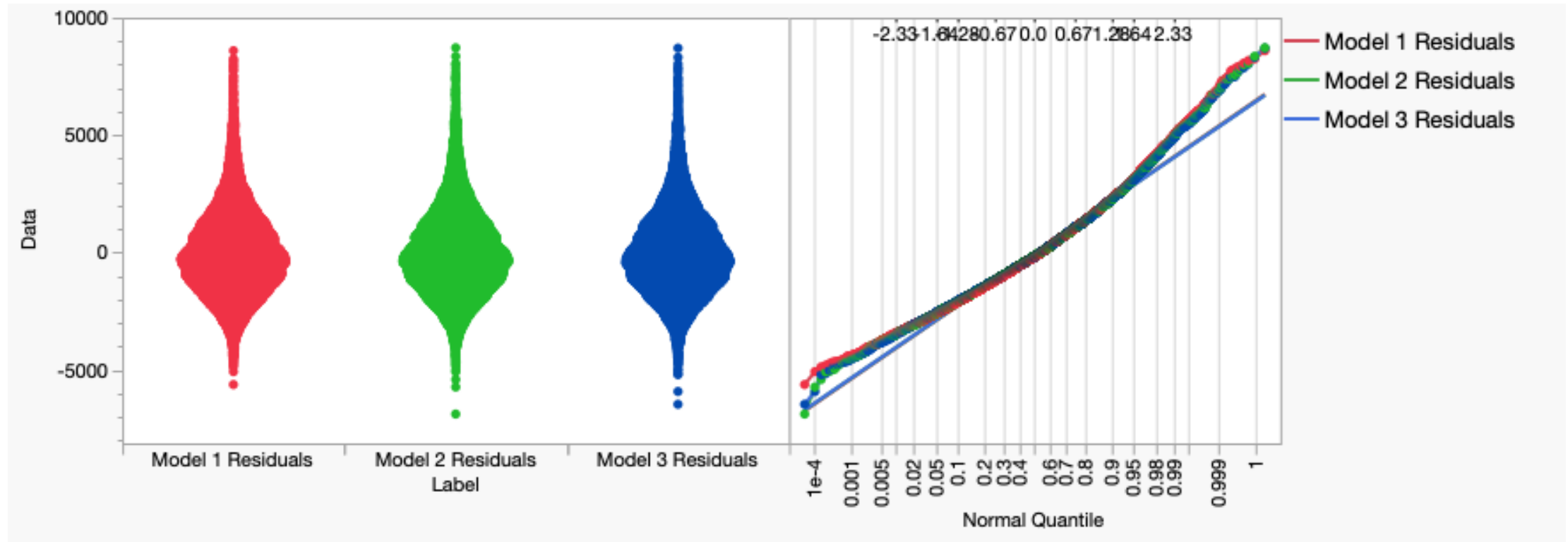
RSquare	0.548772
RSquare Adj	0.548573
Root Mean Square Error	1732.585
Mean of Response	34649.39
Observations (or Sum Wgts)	18132

Parameter Estimates

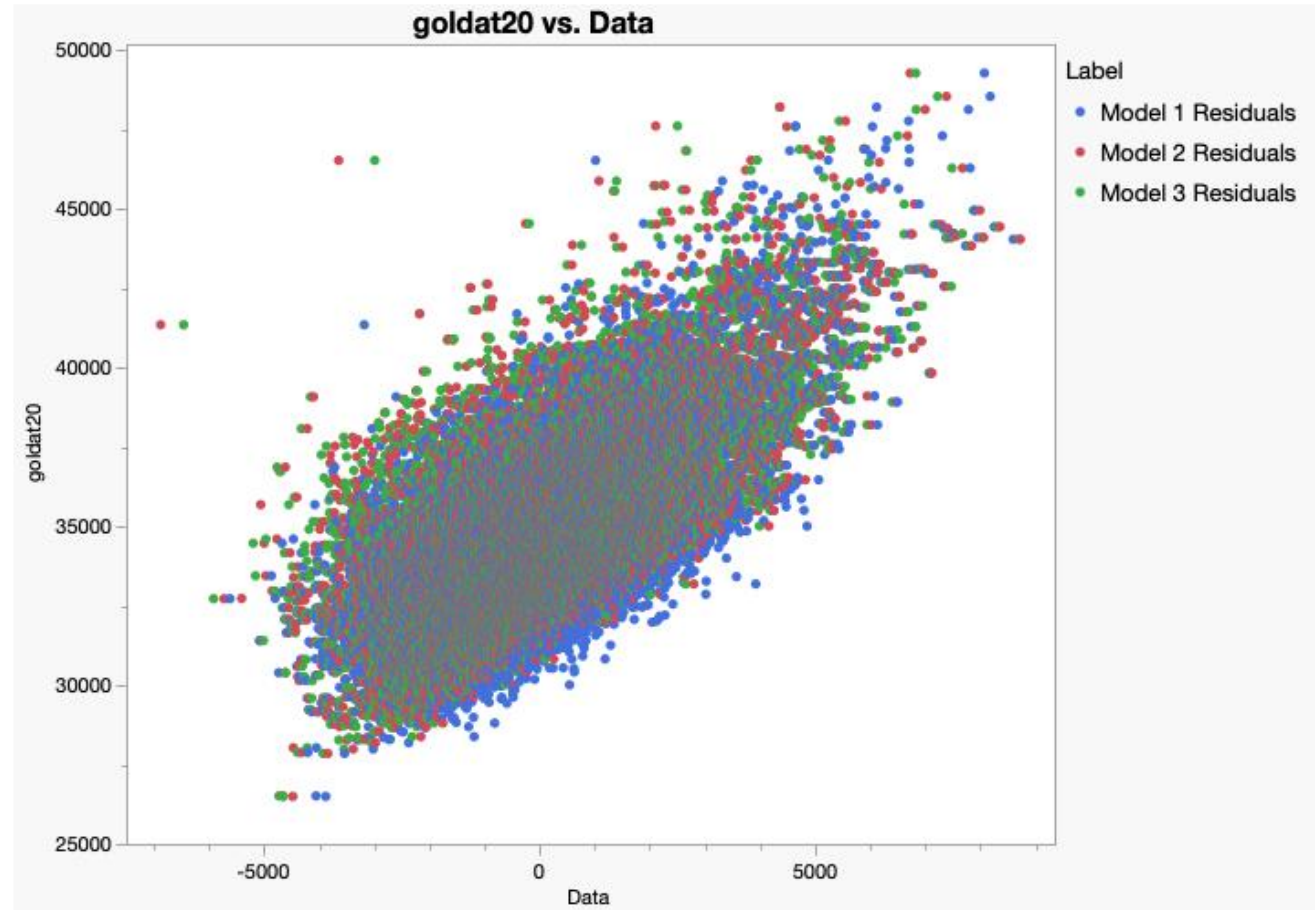
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	36182.604	3166.978	11.42	<.0001*
goldat10	-1.648743	0.399515	-4.13	<.0001*
goldat10*goldat10	0.0000918	1.264e-5	7.26	<.0001*
firsttower	-887.3494	591.1261	-1.50	0.1333
goldat10*firsttower	0.1330144	0.036982	3.60	0.0003*
firstdragon	-630.4189	480.5197	-1.31	0.1896
goldat10*firstdragon	0.0526157	0.030002	1.75	0.0795
firstherald	-2290.059	549.2735	-4.17	<.0001*
goldat10*firstherald	0.2070633	0.034341	6.03	<.0001*

Results

Q-Q and Violin Plots



Unequal Variance Assumption



Variance

- Our models explain about half of the variability in goldat20
- League of Legends games can be unpredictable, so we believe our R_a^2 values are acceptable

Model	R_a^2	MSE
Model 1	~0.5427	1743.795
Model 2	~0.547	1735.568
Model 3	~0.5486	1732.585

Model	Lack of Fit	Normality	Unequal Variances
Model 1	✓	✓	?
Model 2	✓	✓	?
Model 3	-	-	-

Conclusion

- Final Model: $\hat{y} = 8265.063 + 1.573x_1 + 1220.052x_2 + 218.549x_3 + 1010.881x_4$
 - $R_a^2 \approx 0.5427$

GitHub Repository

[lolSnowballEffect](#) Repository

References

Oracle's Elixir. Frequently Asked Questions. n.d. Document. 14 November 2025. <https://oracleselixir.com/faq>.

Oracle's Elixir. Google Drive. 14 January 2025. CSV. 5 November 2025. <https://drive.google.com/file/d/1v6LRphp2kYciU4SXp0PCjEMuev1bDejc/view?usp=sharing>.