SkillCraft Multivariant Regression

Arrington Walters

10/7/2020

# Introduction

Videogames are one of my favorite past times. As a player who participates in ranked play, I’ve always kept my eye on the forefront of global competitions. One of the most notable games to establish a global competitive scene backed by paid professionals was the real-time strategy game (RTS) Starcraft II (SC2). In 2013 the top 10 starcraft players by earnings made nearly four-million dollars from their combined winnings (“Winnings: 2019 - Liquipedia - the Starcraft Ii Encyclopedia” n.d.). Watching top-caliber players reflexes and control is astonishing to even seasoned videogame enthusiasts. At the 2019 StarCraft II World Championship Finale, many others and I packed into the arena to see what these professions could do firsthand.



‘(“Congrats to the Starcraft Ii Wcs Global Finals Champion! - Blizzcon” n.d.)’

The eye watering speeds they perform at is universally referenced in gaming terminology as actions per minute (APMS). Professionals take actions at such fastspeeds (high APMS) it becomes challenging to follow their overall strategy. So past pondering their sheer speed, I found it difficult to distinctly define what made these players professionals.

To learn more about what defines talent in SC2 this analysis we will explore in game metrics in attempts to explain rank in competitive mode. The dataset used was provided by [‘(“UCI Machine Learning Repository: SkillCraft1 Master Table Dataset Data Set” n.d.)’](https://archive.ics.uci.edu/ml/datasets/SkillCraft1+Master+Table+Dataset).

## Goal

To model the response LeagueIndex a sample of player data from a 2013 ranked season of Starcraft will be explored. The predictors provided summarize in game performance metrics for a season by player (GameID). The modeling process will consider all the predictor variables and then trim down until only significant predictors remain. Variables will be vetted for multicolinearity and finally the model will be explored for to see if the BLUE assumptions hold.

## Limitations of the Model

The multivariate regression model used for the midterm 2 portions of this study explores the linear estimation of mean response of LeagueIndex estimated by predictors in the design matrix .

The assumptions of this model’s explanatory power depends on the residual error being gaussian. Considering LeagueIndex is ordinal variable it is doubtful if not impossible for the residuals to be statistically normal.

A more suitable form of model for this regression would be based off a Polytonomous Logistic Regression for Ordinal Response (Proportional Odds Model) (“Ordinal Logistic Regression | R Data Analysis Examples” n.d.). These methods will be revisited for the final portion of this analysis.

# The Data Exploring

This dataset is a sample of averaged in-game metrics of Starcraft II players who participate in 2013 ranked play. The variables are as follows:

## [1] "GameID" "LeagueIndex" "Age"   
## [4] "HoursPerWeek" "TotalHours" "APM"   
## [7] "SelectByHotkeys" "AssignToHotkeys" "UniqueHotkeys"   
## [10] "MinimapAttacks" "MinimapRightClicks" "NumberOfPACs"   
## [13] "GapBetweenPACs" "ActionLatency" "ActionsInPAC"   
## [16] "TotalMapExplored" "WorkersMade" "UniqueUnitsMade"   
## [19] "ComplexUnitsMade" "ComplexAbilitiesUsed"

The appendix covers each in depth but the following are highlighted as a preface.

**LeagueIndex** The levels of LeagueIndex range (1:8) corresponding to player ranks Bronze, Silver, Gold, Platinum, Diamond, Master, Grand Master, and Professional league. Game ranking systems frequently are based on ELO/MMR that varies over a much larger range typically ~1200-3000. These ranges are then masked with medal ranks as listed above and then further subdivided into divisions within each medal (“Leagues: 2019 - Liquipedia - the Starcraft Ii Encyclopedia” n.d.) . Using either MMR or having the subdivisions of each player would provide some much needed continuity but unfortunately neither of these metrics are available. These limitation will be revisited more specifically along the exploration, modeling, and the predictions of the values.. The following are the icon’s earned for players who achieve related rank by the end of a given season. The legends for the following plots are styled to match.



**Actions Per Minute (APMs)** - APMs apply to variety of games but are common metric for analyzing proficiency of players at RTS games, its theorized skills like this provide a great advantage to players (**???**). Action quickness alone does not capture the strategy or macro/micro skills so these additional predictors may add some unique color in hopes of further explaining what makes players skilled.

**Perception Action Cycles (PACs)** - are the circular flow of information between an organism and its environment where a sensor-guided sequence of behaviors are iteratively followed towards a goal (“Perception-Action Cycle - Models, Architectures, and Hardware | Vassilis Cutsuridis | Springer” n.d.). In this data-set PACs are aggregate of screen movements where PAC is a screen fixation of containing at least one action (“UCI Machine Learning Repository: SkillCraft1 Master Table Dataset Data Set” n.d.).

# Cleanliness

The missing values are related exclusive to players with LeagueIndex equivalent to Professional Players (8). Where the 55 players with LeagueIndex==8 the age data is NA and the HoursPerWeek are 0. LeagueIndexes 1-7 are obtainable natural game play, to be a professional you would have to be part of a team. The study is aiming to understand how players go from being average to good, less so elite to best.The 55 values associated with professionals will be dropped to resolve both issues.

Another issue with **LeagueIndex** is that **LeagueIndex** 1-6 may contain any number of players, while **LeagueIndex** 7-Grandmaster may only contain some set range of players targeted at 1000 total per region (“Leagues: 2019 - Liquipedia - the Starcraft Ii Encyclopedia” n.d.). Dropping **LeagueIndex** 7 would be a step towards normality.Considering this multivariate non-polylogistical model is already hampered by its selected application, **LeagueIndex**==7 will be kept to preserve a potential insight into the true population of Starcraft II players.

In addition to the missing values we have a clear error with the **TotalHours** of one player. has 1,000,000 **TotalHours** that equates to 114 years of game time.

If we assumed one extra zeros was added the end of the player’s **TotalHours** it equates to 14 years of played time on a game that is only 10 years old as of a 2020. Removing two zeros equates to 1.4 years of played time, and 3 zeros in 51.1 days of played time both that seems just as realistic. There is not a clear path to extrapolate this player’s true **TotalHours** so their data will be dropped from the analysis. This was originally detected during modeling, but brought earlier into that analysis because how obviously unintentional this value is.

Finally, preforming basic a inspection on **HoursPerWeek** a max value of 168 was discovered. Considering there are 168 hours in a week its not plausible for a individual player to do this. There could be multiple players using this account making this possible. Another prospect is that this player is actually an AI like google’s DeepMind (“AlphaStar: Mastering the Real-Time Strategy Game Starcraft Ii | Deepmind” n.d.). Either way this observation will be kept because what is realistic cutoff for the hours per week is not apparent and after removing the observation the next max value is 140 which seems almost as unrealistic.

Its worth noting that dropping any amount of high hour outliers still far from combats all the issues encountered through the use of **HoursPerWeek** and **TotalHours**. Multiple players could be using any of the accounts even either variable is not relatively high. Potentially exacerbating the left-extrema is that nothing prevents one player from smurfing. Smurfing is when a player makes a additional accounts (“What Is a Smurf Account? Everything You Need to Know | Lol-Smurfs” n.d.). A common reason for doing this is to dominate the competition until their rating adapts to their skill level.

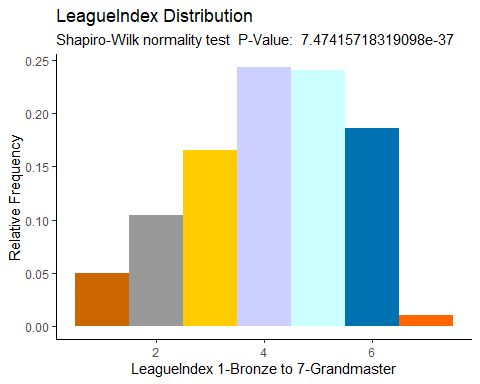
## Converting Units

Some of the time averaged metrics are per SC2 timestamp while other are per milisecond. To make these metrics more interpretable each metric will be converted into seconds. There are roughly 88.5 timestamps per second so each metric in timestamps will be multiplied that as a coefficient (“UCI Machine Learning Repository: SkillCraft1 Master Table Dataset Data Set” n.d.). Some of the time averaged metrics are per milisecond. These will be transformed into seconds so the time units are completely uniform. Both of these transformations are linear and will not affect our model’s assumptions.

## Summary Statistics and Plots

### Gaussianity of the Response

When using Shapiro-Wilk W test on response **LeagueIndex** the null hypothesis that the sample comes from normally distribution can be rejected. Besides the obvious issues with performing a W test with an ordinal response, the response has a negative skew with a mean of 4.12. Additionally we cannot assume the levels between LeagueIndexes are uniforming spaced. As previously mentioned **LeagueIndex** like Grandmaster are capped forcibly prevents **LeagueIndex** from having a gaussian distribution.



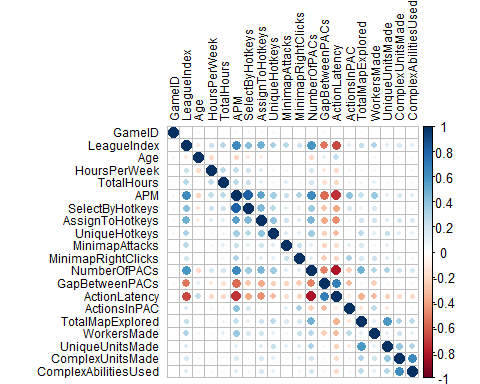
### Correlation Plot

Visually we can see that LeagueIndex has a relatively strong correlation with **APM, SelectByHotkeys, AssignToHotkeys, NumberofPACs, GapBetweenPACs, and Action Latency**. Some of these predictors may be the best choices for model, although its worth noting at this point many of the predictor values also have fairly strong correlations within themselves which may cause multiplecolinearity in a model. This is not too surprising because many of these metrics capture rate of actions in slightly difference forms. For example **APM** and **NumberOfPacs** likely have a strict mathematical relationship where approximately.

The slight differences between these metrics them could have some deep explanatory power but that level of explore ration is beyond the scope of this analysis.

The following columns will be dropped because plausible confoundance with **APMs, ActionLatency, GapBetweenPACs**,[[1]](#footnote-33) **NumberofPACS, SelectbyHotkey, and ActionsInPAC.**

Focusing exclusively on **APM** fits into an Occam’s razor approach by minimizing .



### Visual Relations

Visually determining trends between the predictors and the response with a ordinal response is is best done with alternatives to scatter plots. Violinplots will be used to gauge the linearity in relation to the response and distribution with variable at the varying levels (“A Complete Guide to Violin Plots | Tutorial by Chartio” n.d.).

**MinimapAttacks, HoursPerweek, TotalHours, MinimapRightClicks, ComplexUnitsMade, ComplexAbilitiesUsed** all have very long right tails. In search of gaussian predictors the listed variables were considered transformations, but this would have affect the affected the simplicity of the explanation.[[2]](#footnote-36)

#### No Relationship

**Age** the mean age of 21.6531775 does not vary much across **LeaguIndex** such that there is no stark linear relationship. Although the variance at the highest level seems to be much narrower then that at the lower levels.

#### Bimodal

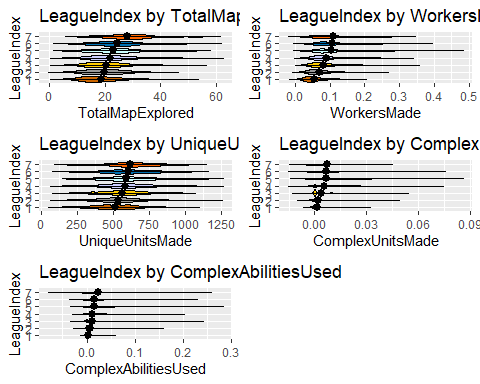
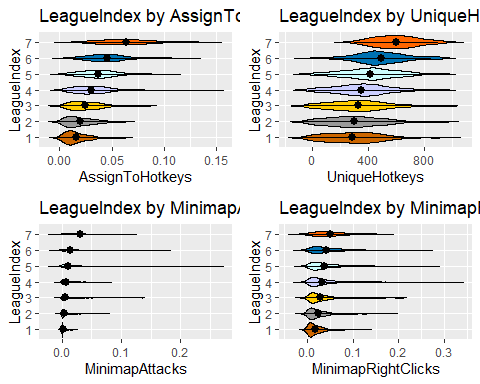
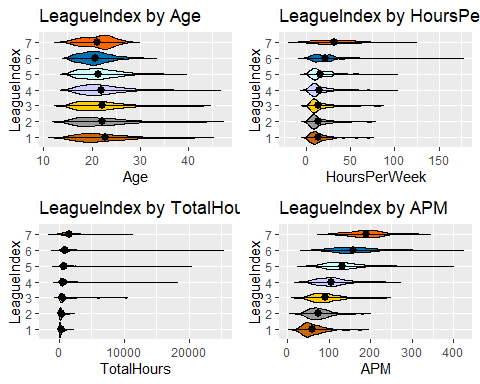
**HourPerWeek** has visibly little or no relation to LeagueIndex 1-4 where LeagueIndex 4-7 seems to have a visible linear trend resulting in 0.03 vs 0.25. If HoursPerWeek survives the model trimming, it’s bimodality may cause issues with the model’s assumption .**Workersmade,ComplexUnitsMades, and ComplexAbilityUsed** both have similar differences between LeagueIndex 1-4 and 4-7 with the portions that have no relation and a linear relation swapped in comparison to HoursPerWeek.

#### Linear

**TotalHours, MinimapRightClicks, TotalMapExplored, and UniiqueUnitsMade** have a positive linear trend with the response. **APM** has a strong linear relationship with the response.

#### Root

**AssignToHotkeys, UniqueHotkeys, and MinimapAttacks** have a notable square root relationship with the response. This also may cause issues with the Gaussianity of the model’s residuals.



# Model Specifications

## Manual Regression Iterations

Before beginning diagnostic tests, a model with all predictors will be made. Subsequently predictors will be dropped one by one until only predictors with significance of at least remain starting with and end with . The results are as follows:

The all inclusive model summary:

Summary of Starting Manual Stepwise Backward Model Selection

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| (Intercept) | 1.4609 | 0.1366 | 10.7 | 0.000 |
| Age | -0.0023 | 0.0046 | -0.5 | 0.617 |
| HoursPerWeek | 0.0051 | 0.0016 | 3.1 | 0.002 |
| TotalHours | 0.0002 | 0.0000 | 7.2 | 0.000 |
| APM | 0.0123 | 0.0005 | 23.4 | 0.000 |
| AssignToHotkeys | 13.4224 | 1.2323 | 10.9 | 0.000 |
| UniqueHotkeys | 0.0005 | 0.0001 | 4.7 | 0.000 |
| MinimapAttacks | 11.1976 | 1.3853 | 8.1 | 0.000 |
| MinimapRightClicks | -0.7756 | 0.6249 | -1.2 | 0.215 |
| TotalMapExplored | 0.0059 | 0.0031 | 1.9 | 0.060 |
| WorkersMade | 2.7113 | 0.4408 | 6.2 | 0.000 |
| UniqueUnitsMade | 0.0000 | 0.0001 | 0.2 | 0.810 |
| ComplexUnitsMade | 1.9539 | 2.5076 | 0.8 | 0.436 |
| ComplexAbilitiesUsed | 1.4389 | 1.0069 | 1.4 | 0.153 |

**Age,UniqueUnitsMade,ComplexUnitsMade, MinimapRightClicks, and TotalMapExplored** were removed in that order across the 5 iterations of the model all remaining predictors were significant to the predetermined . The final iteration provided:

Summary of Final Manual Stepwise Backward Model Selection

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| (Intercept) | 1.5073 | 0.0584 | 25.8 | 0.000 |
| HoursPerWeek | 0.0053 | 0.0016 | 3.2 | 0.001 |
| TotalHours | 0.0002 | 0.0000 | 7.3 | 0.000 |
| APM | 0.0123 | 0.0005 | 24.4 | 0.000 |
| AssignToHotkeys | 13.5401 | 1.2313 | 11.0 | 0.000 |
| UniqueHotkeys | 0.0005 | 0.0001 | 5.2 | 0.000 |
| MinimapAttacks | 11.1888 | 1.3551 | 8.3 | 0.000 |
| WorkersMade | 2.7573 | 0.4331 | 6.4 | 0.000 |
| ComplexAbilitiesUsed | 2.3409 | 0.7983 | 2.9 | 0.003 |

### Assessing Fit and Overall Significance

Before proceeding with further analysis of the model, a statistical test will be performed to see if the predictors provide statistically significant better model than where .Thus the null model is there is no systematic structure to the response LeagueIndex given the predictors such that . Our alternative is there is some relation such that , where X all other variables with the exception of index GameID. Without much surprise using 13 predictor variables results of a very small p-value of ~0. Over the iteration this does not change in a notable fashion across the 9 other models as the last model also results in a p-value0. Thus all iterations of the model we can reject the null hypothesis suggesting that we should further investigate the explanatory power of our alternative hypothesis.

## Testing for Signifance Between Models

If both models had normal residuals a F-test could be used on and to determine if the model’s have significantly different residuals. and both have Shapiro-Wilk’s test results that reject the null at shown in a later section that examines the normality of each models residuals. Without gaussian residuals conducting a ANOVA will be a only practice exercise for the final where hypothesis are provided by:

Performing an ANOVA test we find that there is not significant difference in the models at such that we cannot reject the .(ANOVA see table below).The implications not rejecting is that regardless of trimming the predictor space by predictors, is expected to produce comparable residuals with a 5% chance this is a result of the sampling. Additionally their is barely different. Where 0.462 and 0.4614. Further analysis will be conducted to see if the predict and explanatory power of the models vary past the magnitude of their residuals.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Res.Df | RSS | Df | Sum of Sq | F | Pr(>F) |
| 3322 | 3748.29 | NA | NA | NA | NA |
| 3327 | 3757.68 | -5 | -9.39 | 1.66 | 0.14 |

The confidence Intervals Table[[3]](#footnote-48) show only a few subtleties between models coefficients confidence intervals. For comparison two summary statistics were added as follows:

The OMEGA model includes many 0 values. Amongst delta the most interesting difference notable difference

Confidence Intervals Statistics at 0.05

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Row.names | LL\_s | UL\_s | LL\_f | UL\_f | mean\_s | mean\_f | delta | delta\_width |
| (Intercept) | 1.193 | 1.729 | 1.393 | 1.622 | 1.461 | 1.507 | 0.03 | -1.34 |
| Age | -0.011 | 0.007 | NA | NA | -0.002 | NA | NA | NA |
| APM | 0.011 | 0.013 | 0.011 | 0.013 | 0.012 | 0.012 | 0.00 | -0.04 |
| AssignToHotkeys | 11.006 | 15.838 | 11.126 | 15.954 | 13.422 | 13.540 | 0.01 | 0.00 |
| ComplexAbilitiesUsed | -0.535 | 3.413 | 0.776 | 3.906 | 1.439 | 2.341 | 0.38 | -0.26 |
| ComplexUnitsMade | -2.963 | 6.870 | NA | NA | 1.954 | NA | NA | NA |
| HoursPerWeek | 0.002 | 0.008 | 0.002 | 0.008 | 0.005 | 0.005 | 0.03 | -0.01 |
| MinimapAttacks | 8.481 | 13.914 | 8.532 | 13.846 | 11.198 | 11.189 | 0.00 | -0.02 |
| MinimapRightClicks | -2.001 | 0.450 | NA | NA | -0.776 | NA | NA | NA |
| TotalHours | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.00 | 0.00 |
| TotalMapExplored | 0.000 | 0.012 | NA | NA | 0.006 | NA | NA | NA |
| UniqueHotkeys | 0.000 | 0.001 | 0.000 | 0.001 | 0.000 | 0.001 | 0.08 | -0.03 |
| UniqueUnitsMade | 0.000 | 0.000 | NA | NA | 0.000 | NA | NA | NA |
| WorkersMade | 1.847 | 3.575 | 1.908 | 3.607 | 2.711 | 2.757 | 0.02 | -0.02 |

### Exploring the chosen model

Evaluate the goodness of fit of the model in terms of R2 and the standard errors, and the major sources of uncertainty. This includes parameter uncertainty, as well as structural uncertainty in the model.

### Model Analysis Reform (Trim and then confound)

Describe your model, how you arrived at it, its goodness of fit, its significance versus other choices of models, and its uncertainty. Describe the predictive power, and the uncertainty. Include relevant tables and figures.

## Predict

Evaluate the predictive power of the model – particularly, how effective does the model appear to be at making predictions of future observations or the mean response. How might these predictions be unreliable? What are the limits of the prediction power, and where do we fall into extrapolation?

## Prelude to final

*7 Describe your proposed research question for the final. How will you revise your original research question? What issues have you encountered so far? What assumptions do you think you need to (re-)evaluate?*

For the final, the logistical regression will remodel the same problem with a different set of techniques and assumptions that fit the ordinal response.

The goal is for the analysis pull in additional regression techniques while still integrating the previous exploratory exercises.

# Appendix

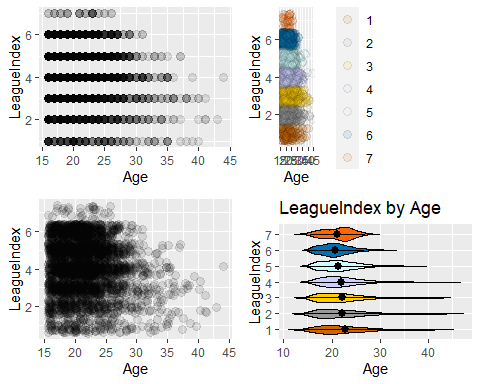
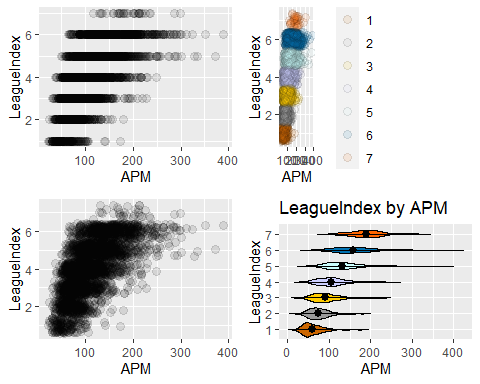
### About Columns

Attribute Information:

1. GameID: Unique ID number for each game (integer)
2. LeagueIndex: Bronze, Silver, Gold, Platinum, Diamond, Master, GrandMaster, and Professional leagues coded 1-8 (Ordinal)
3. Age: Age of each player (integer)
4. HoursPerWeek: Reported hours spent playing per week (integer)
5. TotalHours: Reported total hours spent playing (integer)
6. APM: Action per minute (continuous)
7. SelectByHotkeys: Number of unit or building selections made using hotkeys per timestamp (continuous)
8. AssignToHotkeys: Number of units or buildings assigned to hotkeys per timestamp (continuous)
9. UniqueHotkeys: Number of unique hotkeys used per timestamp (continuous)
10. MinimapAttacks: Number of attack actions on minimap per timestamp (continuous)
11. MinimapRightClicks: number of right-clicks on minimap per timestamp (continuous)
12. NumberOfPACs: Number of PACs per timestamp (continuous)
13. GapBetweenPACs: Mean duration in milliseconds between PACs (continuous)
14. ActionLatency: Mean latency from the onset of a PACs to their first action in milliseconds (continuous)
15. ActionsInPAC: Mean number of actions within each PAC (continuous)
16. TotalMapExplored: The number of 24x24 game coordinate grids viewed by the player per timestamp (continuous)
17. WorkersMade: Number of SCVs, drones, and probes trained per timestamp (continuous)
18. UniqueUnitsMade: Unique unites made per timestamp (continuous)
19. ComplexUnitsMade: Number of ghosts, infestors, and high templars trained per timestamp (continuous)
20. ComplexAbilitiesUsed: Abilities requiring specific targeting instructions used per timestamp (continuous)

### Why Violin Plots.

I decided to use Violin plots because I found with less tweaking they provided almost all the information I was looking for compared to scatter plots. Head to head a limitation of violin plots is that make it seems as though the LeagueIndex level size contains the same . The histogram earlier in this analysis shows clearly that at each level of the **LeagueIndex** is not equal so choosing a tool on the basis of reiterating that point seems redundant. The benefit of Violin plots is that they provide a smoothed density plot at each **LeagueIndex** with a single point that represents the mean. This same thing could be done with scatter plots but I found it took much more staring and plot to plot variation.

The following is some head to head varieties plotting the data. 

# References

“A Complete Guide to Violin Plots | Tutorial by Chartio.” n.d. Accessed November 2, 2020. <https://chartio.com/learn/charts/violin-plot-complete-guide/>.

“AlphaStar: Mastering the Real-Time Strategy Game Starcraft Ii | Deepmind.” n.d. Accessed November 2, 2020. <https://deepmind.com/blog/article/alphastar-mastering-real-time-strategy-game-starcraft-ii>.

“Congrats to the Starcraft Ii Wcs Global Finals Champion! - Blizzcon.” n.d. Accessed October 26, 2020. <https://blizzcon.com/en-us/news/23198508/congrats-to-the-starcraft-ii-wcs-global-finals-champion>.

“Leagues: 2019 - Liquipedia - the Starcraft Ii Encyclopedia.” n.d. Accessed October 26, 2020. <https://liquipedia.net/starcraft2/Battle.net_Leagues>.

“Ordinal Logistic Regression | R Data Analysis Examples.” n.d. Accessed October 27, 2020. <https://stats.idre.ucla.edu/r/dae/ordinal-logistic-regression/>.

“Perception-Action Cycle - Models, Architectures, and Hardware | Vassilis Cutsuridis | Springer.” n.d. Accessed October 29, 2020. <https://www.springer.com/gp/book/9781441914514>.

“UCI Machine Learning Repository: SkillCraft1 Master Table Dataset Data Set.” n.d. Accessed October 26, 2020. <https://archive.ics.uci.edu/ml/datasets/SkillCraft1+Master+Table+Dataset>.

“What Is a Smurf Account? Everything You Need to Know | Lol-Smurfs.” n.d. Accessed November 2, 2020. <https://www.lol-smurfs.com/blog/what-is-a-smurf-account>.

“Winnings: 2019 - Liquipedia - the Starcraft Ii Encyclopedia.” n.d. Accessed October 26, 2020. <https://liquipedia.net/starcraft2/Winnings/2019>.

1. An additional issues with this predictor is that it does not seem to line up with the time units in the description. Before and after the unit transformation **GapBetweenPACs** results in a mean is 11.3094444 hours.) [↑](#footnote-ref-33)
2. A log transformation would be preferable, but enough observations by GameID that contain at least 0 in the related predictor entry a would have to drop observation containing . If a transformation is persued for the second-half of this analysis it will likely be a square root transformation [↑](#footnote-ref-36)
3. Delta values are normalized by dividing by the mean of the model by related predictors confidence interval [↑](#footnote-ref-48)