# **StarCraft II Replay Analysis – Final Report**



**Prepared by:**

Michael King

Akshay Dirisala

Aditya Guli

Advait Satish Deodikar

# **Executive Overview:**

The data used for our multiple linear regression modeling is a gameplay analysis from Starcraft 2.

**About StarCraft 2:**

***StarCraft*** is a military science fiction real-time strategy video game developed and published by Blizzard Entertainment and released for Microsoft Windows on March 31, 1998. Set in a fictitious timeline during the Earth's 25th century, the game revolves around three species fighting for dominance in a distant part of the Milky Way galaxy known as the Koprulu Sector: the Terrans, humans exiled from Earth skilled at adapting to any situation; the Zerg, a race of insectoid aliens in pursuit of genetic perfection, obsessed with assimilating other races; and the Protoss, a humanoid species with advanced technology and psionic abilities, attempting to preserve their civilization and strict philosophical way of living from the Zerg.

### **Questions worth exploring:**

* How do the replay attributes differ by level of player expertise?

A player belonging to league index 8 has significantly higher no. of contributing factors like worker units made, complex units made, mnimap right clicks which eventually were factored into the Actions per Minute compared to a player from league index 1 or 3. We predicted the interval of APM that players from each of league indices should have and then went back into the data to compare if the actual captured APM fell in this interval to determine what area of the game the player needs to focus on to move from one skill level to the next. This concept was better visually demonstrated in the Residual plot.

Through the process of modeling, we decided to approach this analysis by sub-setting the data by the league index since predicting APM for different levels of skilled players with equally varying predictor variables turned out to be non – linear. So, we decided to predict the APM for each of these buckets (or subset). After fitting a model for each of the sub groups of players, we found the variables listed below are significant for league index 8 players:

* SelectbyHotKeys
* GapBetweenPACs
* ComplexUnitsMade
* ActionsInPAC
* NumberOfPACs
* WorkersMade

After running statistical tests, we found that this model explained maximum variation in APM. Residual plots endorsed this understanding.

# **Data Collection:**

We downloaded the dataset from Kaggle competitions database. This data has been posted by Simon Fraser University – Summit.

Note: You must give attribution to the work. You may not use this work for commercial purposes. You may not alter, transform, or build upon this work. Any further uses require the permission of the rights holder.

Main Data variables:

**GameID**: Unique ID for each game.

**LeagueIndex**: 1-8 for Bronze, Silver, Gold, Diamond, Master, GrandMaster, Professional leagues.

**Age**: Age of each player.

**HoursPerWeek**: Hours spent playing per week.

**TotalHours**: Total hours spent playing.

**APM**: Action per minute.

**SelectByHotkeys**: Number of unit selections made using hotkeys per timestamp.

**AssignToHotkeys**: Number of units assigned to hotkeys per timestamp.

**UniqueHotkeys**: Number of unique hotkeys used per timestamp.

**MinimapAttacks**: Number of attack actions on minimal per timestamp.

**MinimapRightClicks**: Number of right-clicks on minimal per timestamp.

**NumberOfPACs**: Number of PACs per timestamp.

**GapBetweenPACs**: Mean duration between PACs (milliseconds).

**ActionLatency**: Mean latency from the onset of PACs to their first action (milliseconds).

**ActionsInPAC**: Mean number of actions within each PAC.

**TotalMapExplored**: Number of 24x24 game coordinate grids viewed by player per timestamp.

**WorkersMade**: Number of SCVs, drones, probes trained per timestamp.

**UniqueUnitsMade**: Unique units made per timestamp.

**ComplexUnitsMade**: Number of ghosts, investors, and high templars trained per timestamp.

**ComplexAbilityUsed**: Abilities requiring specific targeting instructions used per timestamp.

**MaxTimeStamp**: Time stamp of game's last recorded event.

This data has a mix of continuous and categorical variables. There are 21 columns with 3000 plus rows. Working on this dataset provides us with a unique opportunity and experience of analytics application in the gaming industry especially with the growing e-Sports community. Building different linear models and inter model comparisons, variable selections for these MLR models and then choosing the most accurate of predictive models based on statistics like Multiple R squared, Adjusted R squared and the corresponding significance and effects of the chosen predictor variables.

As mentioned earlier in the executive summary, the goal was to identify most significant and influential predictor variables in the model. We hope that this study will help professional coaches to identify areas of improvement depending on their skill level and the residuals we may find.

We first built a full model to gain insight into which variables had high significance w.r.t our null hypothesis’ H0. Then we looked at the correlation between the predictor variables and multicollinearity and eliminated the ones that affected our statistics. Then we tried running the partial F test on multiple other reduced models to see if they fared well compared to the full model but that didn’t lead us to a proper set of indicators. Then we decided to use the stepwise selection method which incorporates the advantages of both forward selection and backward elimination methods and the method selected all predictor variables which maximized the multiple R squared and adjusted values, the F statistic from the summary.

Once we interpreted the model summary and drew residuals plots, we noted that they were scattered with no specific pattern. So, we concluded that there was a linear relation between the predictor variable and response variables.

# **Regression Modelling:**

1. Preprocessing:

Early in this stage, we discovered that there were a lot of missing values in the column “Age” for league index 8 players. We hoped to use this in our initial stages of modeling so we fixed this by approximating the age by taking the nearest group. We used the triangular distribution to approximate and impute the trimean values into the missing “Age” fields. We had removed the variables total hours, total hours per week as these variables are had missing fields so to avoid corruption of our model.

1. Identify categorical variables and sub setting the data:

We had identified that league index is the categorical variables having the players categorized into eight levels. We then subset the players based on their league index.

1. Full model: A few important statistics from the summary of the full model are listed below which looks good but we had to eliminate a few variables which didn’t have significance based on the corresponding P – values.

Residual standard error: 5.496 on 38 degrees of freedom

Multiple R-squared: 0.9933, Adjusted R-squared: 0.9904

F-statistic: 350.3 on 16 and 38 DF, p-value: < 2.2e-16

1. Lack of FIT F test.

We used the lack of fit test on the models initially only to note abnormally large F-statistic values in the order of 10^24 due to which we reject the null hypothesis that our full model has a linear relationship between the response and predictor variables. This means the linearity assumption is violated and we fixed this by sub setting the data based on the league indices.

1. Reduced-model: One of the 8 reduced models’ summary is pasted below for insight from an analytics standpoint.

summary(StepwiseAic8)

Call:

Lm (formula = APM ~ SelectByHotkeys + GapBetweenPACs + ComplexUnitsMade +

ActionsInPAC + NumberOfPACs + WorkersMade, data = starcraft8)

Residuals:

Min 1Q Median 3Q Max

-14.0046 -2.6707 0.6941 3.8054 7.5906

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -126.1831 15.4910 -8.146 1.33e-10 \*\*\*

SelectByHotkeys 5390.3904 78.9430 68.282 < 2e-16 \*\*\*

GapBetweenPACs -0.2981 0.1712 -1.742 0.0880 .

ComplexUnitsMade 18842.9660 10435.0548 1.806 0.0772 .

ActionsInPAC 23.5838 1.2049 19.572 < 2e-16 \*\*\*

NumberOfPACs 27933.2635 1377.9995 20.271 < 2e-16 \*\*\*

WorkersMade 5027.0802 1907.7001 2.635 0.0113 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.149 on 48 degrees of freedom

Multiple R-squared: 0.9925, Adjusted R-squared: 0.9916

F-statistic: 1063 on 6 and 48 DF, p-value: < 2.2e-16

1. Residual analysis and outlier detection – Assumptions checks.

For league index 8 our model has few outliers based on the residual plot but they don’t cross the cooks line and they are not high leverage cases. Most of the residuals are on the normal Q-Q line which confirms the normality assumption in the residuals. In the variance, we can see that most of the residuals fall under the same range except for the instances labeled 3360,3356,3343.

1. Variable selection and refinement in variable selection.

We did variable selection on our first model comparing stepwise, backward and forward selection using both AIC and BIC. After using the lack of fitness test, we decided to subset our data by league index and then use stepwise AIC to select variables.

1. Find best possible models

Predicting APM for different levels of skilled players with equally varying predictor variables turned out to be non – linear and logically didn’t make much sense to us. So, the sensible thing to do was to build eight different models one for each group. We used the stepwise AIC method determine the predictor variables for each of these models. These are attached in the appendix section of this document.

We copied all the code in appendices which follows all above steps approximately same order as above.

# **Summary and Conclusion:**

We have chosen model 1 as final model that predicts APM with the help of below mentioned predictor variables:

1) SelectByHotkeys

2) GapBetweenPACs

3) NumberOfPACs

4) ActionsInPAC

5) WorkersMade

6) MinimapRightClicks

7) UniqueUnitsMade

8) ComplexAbilityUsed

We chose this model over others because it explains variability in APM with

Help of identified predictor variables. R squared is 97.7% indicating better explicability. P value is approaching 0 and F statistics is very high which further confirms goodness of fit.

Using this model, we then can predict the players APM based on the variables. We can then find the residuals which can tell which players in a specific league have unusual APM values which could be due to factors like lack of knowledge about the game or game sense.

# **Appendices:**

> summary(StepwiseAic8)

Call:

lm(formula = APM ~ SelectByHotkeys + GapBetweenPACs + ComplexUnitsMade +

ActionsInPAC + NumberOfPACs + WorkersMade, data = starcraft8)

Residuals:

Min 1Q Median 3Q Max

-14.0046 -2.6707 0.6941 3.8054 7.5906

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -126.1831 15.4910 -8.146 1.33e-10 \*\*\*

SelectByHotkeys 5390.3904 78.9430 68.282 < 2e-16 \*\*\*

GapBetweenPACs -0.2981 0.1712 -1.742 0.0880 .

ComplexUnitsMade 18842.9660 10435.0548 1.806 0.0772 .

ActionsInPAC 23.5838 1.2049 19.572 < 2e-16 \*\*\*

NumberOfPACs 27933.2635 1377.9995 20.271 < 2e-16 \*\*\*

WorkersMade 5027.0802 1907.7001 2.635 0.0113 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.149 on 48 degrees of freedom

Multiple R-squared: 0.9925, Adjusted R-squared: 0.9916

F-statistic: 1063 on 6 and 48 DF, p-value: < 2.2e-16

> summary(StepwiseAic7)

Call:

lm(formula = APM ~ SelectByHotkeys + ActionLatency + ActionsInPAC +

NumberOfPACs + ComplexAbilityUsed + WorkersMade + Age + MinimapRightClicks +

ComplexUnitsMade, data = starcraft7)

Residuals:

Min 1Q Median 3Q Max

-12.7684 -3.0186 -0.0087 3.8188 8.4330

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -66.9867 33.6378 -1.991 0.057469 .

SelectByHotkeys 5071.2841 212.9053 23.819 < 2e-16 \*\*\*

ActionLatency -0.8935 0.3264 -2.737 0.011238 \*

ActionsInPAC 22.4535 1.6451 13.649 4.34e-13 \*\*\*

NumberOfPACs 18776.6633 2734.7756 6.866 3.39e-07 \*\*\*

ComplexAbilityUsed -8063.1936 2135.1720 -3.776 0.000878 \*\*\*

WorkersMade 3413.7218 1763.7931 1.935 0.064324 .

Age 1.1158 0.3694 3.020 0.005748 \*\*

MinimapRightClicks 7130.2200 3174.5458 2.246 0.033781 \*

ComplexUnitsMade 21391.7898 10077.3959 2.123 0.043855 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.739 on 25 degrees of freedom

Multiple R-squared: 0.9854, Adjusted R-squared: 0.9802

F-statistic: 188.1 on 9 and 25 DF, p-value: < 2.2e-16

> summary(StepwiseAic6)

Call:

lm(formula = APM ~ SelectByHotkeys + ActionLatency + ActionsInPAC +

NumberOfPACs + MinimapRightClicks + WorkersMade + ComplexAbilityUsed +

GapBetweenPACs + UniqueUnitsMade + Age, data = starcraft6)

Residuals:

Min 1Q Median 3Q Max

-71.482 -3.561 0.607 4.233 35.435

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.608e+01 7.939e+00 -3.285 0.00108 \*\*

SelectByHotkeys 5.412e+03 5.152e+01 105.031 < 2e-16 \*\*\*

ActionLatency -4.416e-01 6.341e-02 -6.964 8.57e-12 \*\*\*

ActionsInPAC 1.499e+01 3.795e-01 39.496 < 2e-16 \*\*\*

NumberOfPACs 2.146e+04 7.903e+02 27.152 < 2e-16 \*\*\*

MinimapRightClicks 4.963e+03 8.161e+02 6.082 2.10e-09 \*\*\*

WorkersMade 3.481e+03 6.256e+02 5.565 3.93e-08 \*\*\*

ComplexAbilityUsed -4.436e+03 1.087e+03 -4.081 5.07e-05 \*\*\*

GapBetweenPACs -1.402e-01 4.533e-02 -3.092 0.00208 \*\*

UniqueUnitsMade -3.619e+04 1.358e+04 -2.665 0.00791 \*\*

Age -2.452e-01 1.069e-01 -2.294 0.02212 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.772 on 610 degrees of freedom

Multiple R-squared: 0.9745, Adjusted R-squared: 0.9741

F-statistic: 2334 on 10 and 610 DF, p-value: < 2.2e-16

> summary(StepwiseAic5)

Call:

lm(formula = APM ~ SelectByHotkeys + ActionLatency + ActionsInPAC +

NumberOfPACs + WorkersMade + MinimapRightClicks + UniqueUnitsMade +

GapBetweenPACs + AssignToHotkeys + Age + MaxTimeStamp + ComplexAbilityUsed,

data = starcraft5)

Residuals:

Min 1Q Median 3Q Max

-77.900 -3.900 0.253 4.317 27.938

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -3.290e+00 6.399e+00 -0.514 0.60734

SelectByHotkeys 5.560e+03 6.080e+01 91.448 < 2e-16 \*\*\*

ActionLatency -4.820e-01 4.730e-02 -10.189 < 2e-16 \*\*\*

ActionsInPAC 1.273e+01 2.847e-01 44.731 < 2e-16 \*\*\*

NumberOfPACs 2.034e+04 7.824e+02 25.998 < 2e-16 \*\*\*

WorkersMade 3.336e+03 5.358e+02 6.227 7.72e-10 \*\*\*

MinimapRightClicks 4.033e+03 7.846e+02 5.140 3.46e-07 \*\*\*

UniqueUnitsMade -6.087e+04 1.486e+04 -4.098 4.60e-05 \*\*\*

GapBetweenPACs -1.035e-01 3.135e-02 -3.301 0.00101 \*\*

AssignToHotkeys -4.782e+03 1.600e+03 -2.988 0.00289 \*\*

Age -1.986e-01 7.552e-02 -2.629 0.00872 \*\*

MaxTimeStamp -2.251e-05 1.258e-05 -1.789 0.07398 .

ComplexAbilityUsed -1.549e+03 1.023e+03 -1.514 0.13054

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.653 on 793 degrees of freedom

Multiple R-squared: 0.9668, Adjusted R-squared: 0.9663

F-statistic: 1925 on 12 and 793 DF, p-value: < 2.2e-16

> summary(StepwiseAic4)

Call:

lm(formula = APM ~ SelectByHotkeys + ActionLatency + ActionsInPAC +

NumberOfPACs + WorkersMade + ComplexAbilityUsed + UniqueUnitsMade +

MinimapRightClicks + Age + GapBetweenPACs + MinimapAttacks,

data = starcraft4)

Residuals:

Min 1Q Median 3Q Max

-42.714 -2.819 0.185 2.995 33.051

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -4.577e+01 4.844e+00 -9.449 < 2e-16 \*\*\*

SelectByHotkeys 5.391e+03 6.611e+01 81.544 < 2e-16 \*\*\*

ActionLatency -1.597e-01 3.087e-02 -5.173 2.91e-07 \*\*\*

ActionsInPAC 1.376e+01 2.220e-01 61.967 < 2e-16 \*\*\*

NumberOfPACs 2.374e+04 6.356e+02 37.344 < 2e-16 \*\*\*

WorkersMade 2.544e+03 4.871e+02 5.223 2.25e-07 \*\*\*

ComplexAbilityUsed -3.541e+03 8.744e+02 -4.050 5.62e-05 \*\*\*

UniqueUnitsMade -3.556e+04 9.061e+03 -3.925 9.42e-05 \*\*\*

MinimapRightClicks 2.270e+03 6.521e+02 3.481 0.000528 \*\*\*

Age -1.742e-01 5.278e-02 -3.300 0.001011 \*\*

GapBetweenPACs -4.783e-02 1.990e-02 -2.404 0.016458 \*

MinimapAttacks 3.209e+03 2.023e+03 1.586 0.113121

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.936 on 799 degrees of freedom

Multiple R-squared: 0.9696, Adjusted R-squared: 0.9692

F-statistic: 2317 on 11 and 799 DF, p-value: < 2.2e-16

> summary(StepwiseAic3)

Call:

lm(formula = APM ~ SelectByHotkeys + ActionLatency + ActionsInPAC +

NumberOfPACs + WorkersMade + ComplexAbilityUsed + AssignToHotkeys +

UniqueUnitsMade + TotalMapExplored, data = starcraft3)

Residuals:

Min 1Q Median 3Q Max

-30.025 -2.814 0.355 3.328 29.338

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -4.474e+01 5.484e+00 -8.160 2.35e-15 \*\*\*

SelectByHotkeys 5.597e+03 1.279e+02 43.756 < 2e-16 \*\*\*

ActionLatency -1.721e-01 3.123e-02 -5.510 5.55e-08 \*\*\*

ActionsInPAC 1.251e+01 2.465e-01 50.739 < 2e-16 \*\*\*

NumberOfPACs 2.348e+04 8.369e+02 28.061 < 2e-16 \*\*\*

WorkersMade 2.654e+03 7.143e+02 3.715 0.000224 \*\*\*

ComplexAbilityUsed -2.359e+03 1.090e+03 -2.164 0.030870 \*

AssignToHotkeys -2.991e+03 1.769e+03 -1.691 0.091342 .

UniqueUnitsMade -2.972e+04 1.339e+04 -2.220 0.026861 \*

TotalMapExplored 6.983e+03 3.997e+03 1.747 0.081193 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.354 on 543 degrees of freedom

Multiple R-squared: 0.9577, Adjusted R-squared: 0.957

F-statistic: 1368 on 9 and 543 DF, p-value: < 2.2e-16

> summary(StepwiseAic2)

Call:

lm(formula = APM ~ ActionLatency + SelectByHotkeys + ActionsInPAC +

NumberOfPACs + MinimapRightClicks + WorkersMade + AssignToHotkeys +

GapBetweenPACs + ComplexAbilityUsed + UniqueUnitsMade + TotalMapExplored +

Age + MinimapAttacks, data = starcraft2)

Residuals:

Min 1Q Median 3Q Max

-66.794 -3.520 0.514 3.970 22.173

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -1.716e+01 7.924e+00 -2.166 0.031018 \*

ActionLatency -1.214e-01 4.301e-02 -2.824 0.005034 \*\*

SelectByHotkeys 5.876e+03 2.913e+02 20.173 < 2e-16 \*\*\*

ActionsInPAC 8.489e+00 3.757e-01 22.594 < 2e-16 \*\*\*

NumberOfPACs 2.109e+04 1.262e+03 16.709 < 2e-16 \*\*\*

MinimapRightClicks 5.015e+03 1.502e+03 3.340 0.000934 \*\*\*

WorkersMade 3.993e+03 1.184e+03 3.373 0.000831 \*\*\*

AssignToHotkeys -9.688e+03 3.166e+03 -3.060 0.002392 \*\*

GapBetweenPACs -6.498e-02 2.710e-02 -2.397 0.017071 \*

ComplexAbilityUsed -3.356e+03 2.118e+03 -1.584 0.114069

UniqueUnitsMade -5.305e+04 1.951e+04 -2.719 0.006890 \*\*

TotalMapExplored 1.233e+04 5.787e+03 2.130 0.033865 \*

Age -1.593e-01 8.345e-02 -1.908 0.057207 .

MinimapAttacks 7.113e+03 4.665e+03 1.524 0.128335

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.528 on 333 degrees of freedom

Multiple R-squared: 0.9042, Adjusted R-squared: 0.9004

F-statistic: 241.7 on 13 and 333 DF, p-value: < 2.2e-16

> summary(StepwiseAic1)

Call:

lm(formula = APM ~ ActionLatency + ActionsInPAC + SelectByHotkeys +

NumberOfPACs + TotalMapExplored + ComplexAbilityUsed + Age +

UniqueHotkeys + AssignToHotkeys + MaxTimeStamp, data = starcraft1)

Residuals:

Min 1Q Median 3Q Max

-19.0900 -2.6205 0.4475 2.5688 13.4519

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.814e+01 7.120e+00 -3.952 0.000117 \*\*\*

ActionLatency -9.730e-02 3.077e-02 -3.162 0.001880 \*\*

ActionsInPAC 1.037e+01 3.217e-01 32.244 < 2e-16 \*\*\*

SelectByHotkeys 5.653e+03 3.553e+02 15.911 < 2e-16 \*\*\*

NumberOfPACs 1.982e+04 1.285e+03 15.421 < 2e-16 \*\*\*

TotalMapExplored -1.023e+04 6.002e+03 -1.705 0.090170 .

ComplexAbilityUsed 1.035e+04 4.693e+03 2.205 0.028888 \*

Age -1.438e-01 7.721e-02 -1.863 0.064351 .

UniqueHotkeys 3.099e+04 1.381e+04 2.244 0.026215 \*

AssignToHotkeys -5.623e+03 3.914e+03 -1.437 0.152799

MaxTimeStamp 2.326e-05 1.632e-05 1.425 0.156172

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.273 on 156 degrees of freedom

Multiple R-squared: 0.9524, Adjusted R-squared: 0.9493

F-statistic: 312.1 on 10 and 156 DF, p-value: < 2.2e-16

**Tests for assumptions:**

> dwtest(StepwiseAic1)

Durbin-Watson test

data: StepwiseAic1

DW = 2.0767, p-value = 0.6924

alternative hypothesis: true autocorrelation is greater than 0

> ad.test(StepwiseAic1$residuals)

Anderson-Darling normality test

data: StepwiseAic1$residuals

A = 1.3544, p-value = 0.001597

>

> dwtest(StepwiseAic2)

Durbin-Watson test

data: StepwiseAic2

DW = 2.0284, p-value = 0.608

alternative hypothesis: true autocorrelation is greater than 0

> ad.test(StepwiseAic2$residuals)

Anderson-Darling normality test

data: StepwiseAic2$residuals

A = 4.8538, p-value = 4.925e-12

>

> dwtest(StepwiseAic3)

Durbin-Watson test

data: StepwiseAic3

DW = 1.8908, p-value = 0.09973

alternative hypothesis: true autocorrelation is greater than 0

> ad.test(StepwiseAic3$residuals)

Anderson-Darling normality test

data: StepwiseAic3$residuals

A = 9.153, p-value < 2.2e-16

>

> dwtest(StepwiseAic4)

Durbin-Watson test

data: StepwiseAic4

DW = 1.9471, p-value = 0.2241

alternative hypothesis: true autocorrelation is greater than 0

> ad.test(StepwiseAic4$residuals)

Anderson-Darling normality test

data: StepwiseAic4$residuals

A = 10.755, p-value < 2.2e-16

>

> dwtest(StepwiseAic5)

Durbin-Watson test

data: StepwiseAic5

DW = 2.0504, p-value = 0.7619

alternative hypothesis: true autocorrelation is greater than 0

> ad.test(StepwiseAic5$residuals)

Anderson-Darling normality test

data: StepwiseAic5$residuals

A = 7.6359, p-value < 2.2e-16

>

> dwtest(StepwiseAic6)

Durbin-Watson test

data: StepwiseAic6

DW = 2.0965, p-value = 0.8875

alternative hypothesis: true autocorrelation is greater than 0

> ad.test(StepwiseAic6$residuals)

Anderson-Darling normality test

data: StepwiseAic6$residuals

A = 7.2781, p-value < 2.2e-16

>

> dwtest(StepwiseAic7)

Durbin-Watson test

data: StepwiseAic7

DW = 1.1126, p-value = 0.002886

alternative hypothesis: true autocorrelation is greater than 0

> ad.test(StepwiseAic7$residuals)

Anderson-Darling normality test

data: StepwiseAic7$residuals

A = 0.40177, p-value = 0.3414

>

> dwtest(StepwiseAic8)

Durbin-Watson test

data: StepwiseAic8

DW = 2.1474, p-value = 0.6551

alternative hypothesis: true autocorrelation is greater than 0

> ad.test(StepwiseAic8$residuals)

Anderson-Darling normality test

data: StepwiseAic8$residuals

A = 0.77994, p-value = 0.0403

**Supporting Graphs and visualizations:**















