2015 Discover Cup University Modeling Contest

Template

****

Table of Contents

[Key Personnel (Required) 3](#_Toc429489441)

[1. Abstract (Required) 4](#_Toc429489442)

[2. Model Methodology (Required) 4](#_Toc429489443)

[3. Model Development Data (Required) 4](#_Toc429489444)

[4. Tools, Codes and Appendices (Required) 4](#_Toc429489445)

[5. Model Assumptions (optional) 4](#_Toc429489446)

[6. Model Dependent & Independent Variable definitions (optional) 4](#_Toc429489447)

[7. Model Limitations and Weaknesses (optional) 4](#_Toc429489448)

[8. Other Information (optional) 4](#_Toc429489449)

[9. References (optional) 4](#_Toc429489450)

# Key Personnel (Required)

|  |  |
| --- | --- |
| Personnel | Developer Name/ School Name/Expected Graduation date/Expected Degree/Major/Email/Cell Phone |
| Developer YU Hao/Fudan University/Jun-2015/Bachelor/Psychology/teddytiome@gmail.com/1881-730-1907  ZHENG Lina/ Fudan University/Jun-2015/Bachelor/Psychology/nana-coin@163.com/1391-600-3255 | |

# 1. Abstract (Required)

*We took the Seasonal Autogressive Integrated Moving Average Models with disticint parameters for Questions 1-3. To predict future call volumes and handling times, we assumed that there exist both one-week periods in the datasets after our previous ANOVA tests and anlayses from time-series plots. For every-day volumes, we extracted specific call streams in each period and predicted at daily level.*

# 2. Model Methodology (Required)

*We assumed the seasonal change period as one week with stationarity of time series. The average value and the variation of the series should be constant with respect to time, with first difference transformation.*

*The methodology was introduced by Box and Jenkins in 1970, including:*

*Auto Regressive (AR) model*

*Moving Average (MA) model*

*With combination of the both models, ARIMA (p, d, q) (P, D, Q) is a regular from, where p is the order of AR process, q is the order of MA process, d is the order of differencing. And (P, D, Q) are seasonal elements.*

*In our model, parameters values are assigned as that p is 7, d is 1, q is 2 and P is 1, D is 1, Q is 0, with which the model gets a lowest AIC (Akaike’s Information Criterion) values.*

# 3. Model Development Data (Required)

*For daily call volumes in Q1, We used the numbers of row each day in original data to calculate every-day volumes from 1999-01-01 to 1999-11-31. For interval call volumes in Q2, we chose numbers of row by daily time period from selected data in Q1. And for Q3, we flitered our data with call outcome as “AGENT” and sum up all the service times of each call.*

*As a whole, we utilized several variables – Call\_date, IVR\_entry, outcome and servc\_time – to selected and reconstructed our particular databases at each level for different questions with other variables excluded.*

# 4. Tools, Codes and Appendices (Required)

*We used R as our software for applying statistics measures and modeling. To illustrate the model, we contain appendices (PDF file) including tool screenshots, codes.*

# 5. Model Assumptions (optional)

*The data do has a one-week period, and when there was no extreme days (like festivals or holidays) and went stable on November, the model may fit well.*

# 6. Model Dependent & Independent Variable definitions (optional)

-Dependent Variable

*Daily and intervel call volumes as well as daily handling time.*

-Independent Variables

*Weekdays and date as time series variable.*

# 7. Model Limitations and Weaknesses (optional)

*We have missing values in our data, since we distracted this part, maybe there are little unconvincing estimates for the past. And all are done by mathematically modeling, we havenot involved into specific details in data.*

*For time series analysis, more data will improve our results.*

# 8. Other Information (optional)

*Introduciotn documents for R packages:astsa, forecast, chron.*

# 9. References (optional)

[*http://www.stat.pitt.edu/stoffer/tsa2/Rissues.htm*](http://www.stat.pitt.edu/stoffer/tsa2/Rissues.htm)

[*https://stat.ethz.ch/R-manual/R-patched/library/stats/html/ts.html*](https://stat.ethz.ch/R-manual/R-patched/library/stats/html/ts.html)

[*https://stat.ethz.ch/R-manual/R-patched/library/stats/html/predict.arima.html*](https://stat.ethz.ch/R-manual/R-patched/library/stats/html/predict.arima.html)

[*https://stat.ethz.ch/R-manual/R-devel/library/stats/html/arima.html*](https://stat.ethz.ch/R-manual/R-devel/library/stats/html/arima.html)

[*http://www.talkstats.com/showthread.php/27541-Time-Series-ARIMA(p-d-q)-how-to-determine-what-p-d-and-q-equal*](http://www.talkstats.com/showthread.php/27541-Time-Series-ARIMA(p-d-q)-how-to-determine-what-p-d-and-q-equal)

[*http://www.inside-r.org/r-doc/stats/arima*](http://www.inside-r.org/r-doc/stats/arima)