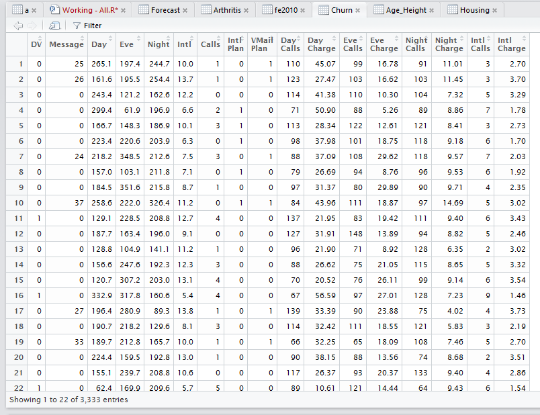
**Final project – Churn Prediction**

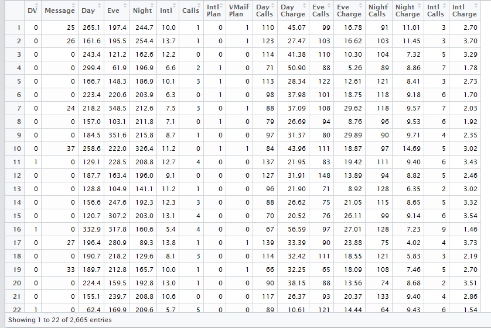
1. Use R to do the pre-processing and training. Your final submission should have all the R code and results coming from Training Model. Show all the accuracy measures and ROC curve for different probability cut-off. Find the optimal Threshold.

**Answer**

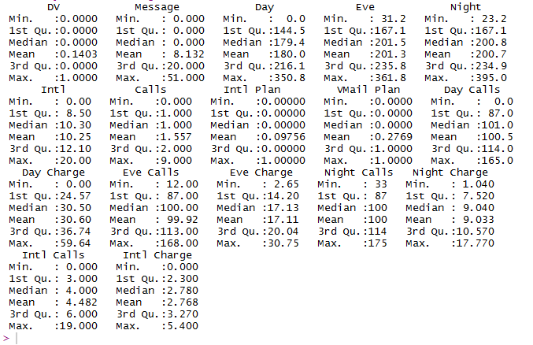
> View(Churn) – 3333 entries



> View(Churn\_Training) – 2665 entries

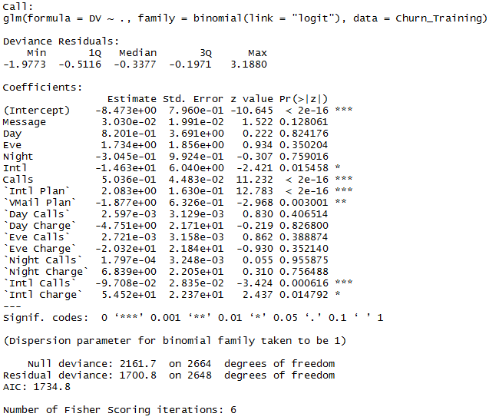


> summary(Churn\_Training)



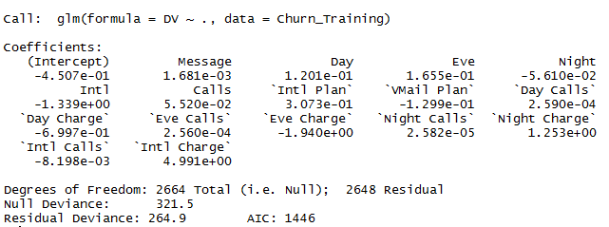
> Model=glm(DV~., data = Churn\_Training, family = binomial(link = "logit"))

> summary(Model)

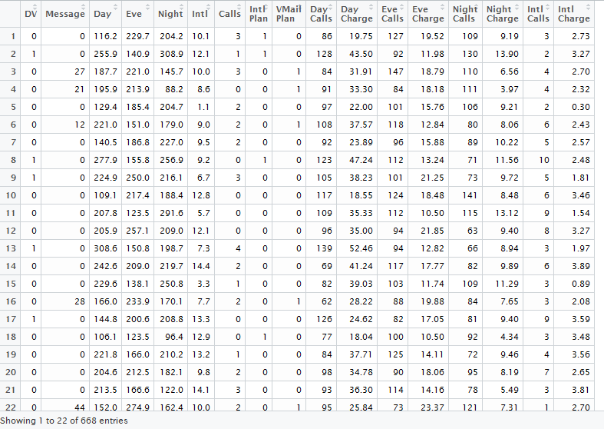


> reduced=glm(DV~.,data = Churn\_Training)

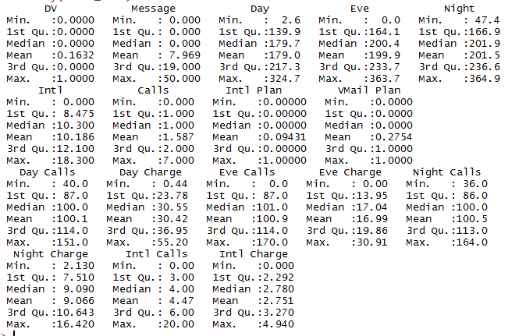
> reduced



> View(Churn\_Test) – 668 entries

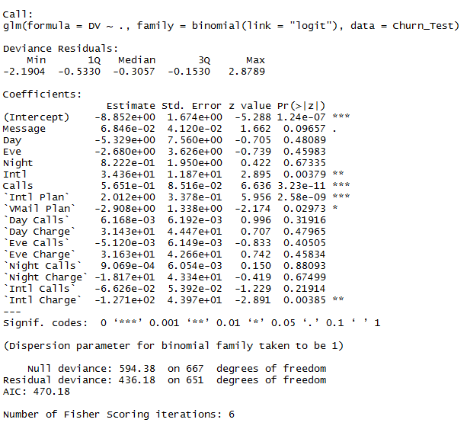


> summary(Churn\_Test)



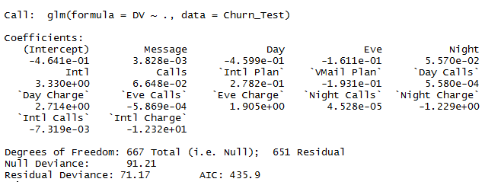
> Model=glm(DV~., data = Churn\_Test, family = binomial(link = "logit"))

> summary(Model)

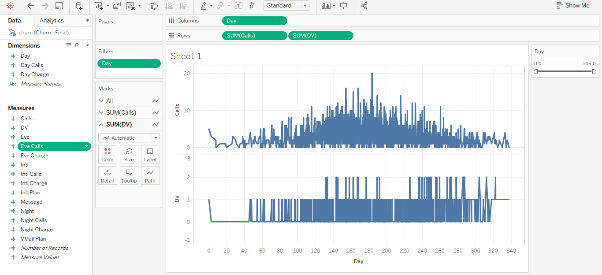


> reduced=glm(DV~.,data = Churn\_Test)

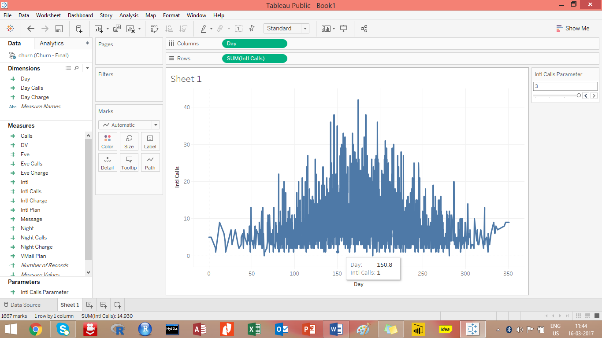
> reduced



1. Use R for giving all the results and statistics of the Trained Model performance on the Test Dataset. Your final report should contain all these results.
2. Use Tableau for Visualization Reporting.
   1. Perform the training and threshold optimization in tableau. Prepare required worksheet to perform the same. You should be able to change the probability threshold using a slider. When changing the threshold, you must show the accuracy measures accordingly.



* 1. Once fixed on a threshold, you should use R script to run the testing on the test data from Tableau.
  2. One you get the test result. Show the accuracy measure as well.



* 1. How the odds ratio on the viz, i.e. show the effect of change of input variables to probability/chances of churn. You should provide option such that, we can choose any input variable and change the value to see the effect. Change on numerical variable would be offset change (some unit change), and categorical variable would be change from one category to other (Example: If binary then change from 0 to 1).