**Model Code Summary**

**John Lance Abbott**

**Data Loading**

>library(readr)

>data <- read.csv(“data path”)  
#read.excel is also acceptable if the file type is excel

**Data Cleaning**

>duplicated(data)  
#checks for duplicated values in data

>unique(data)  
#outputs count of unique values in data

>which(is.na(data))  
#specifies any missing values

>newdata <- data[-c(1,2,3),]  
#removes specified rows. 1,2,3 are examples, specify row number using the OBS value. Use whatever name for newdata as new data frame.

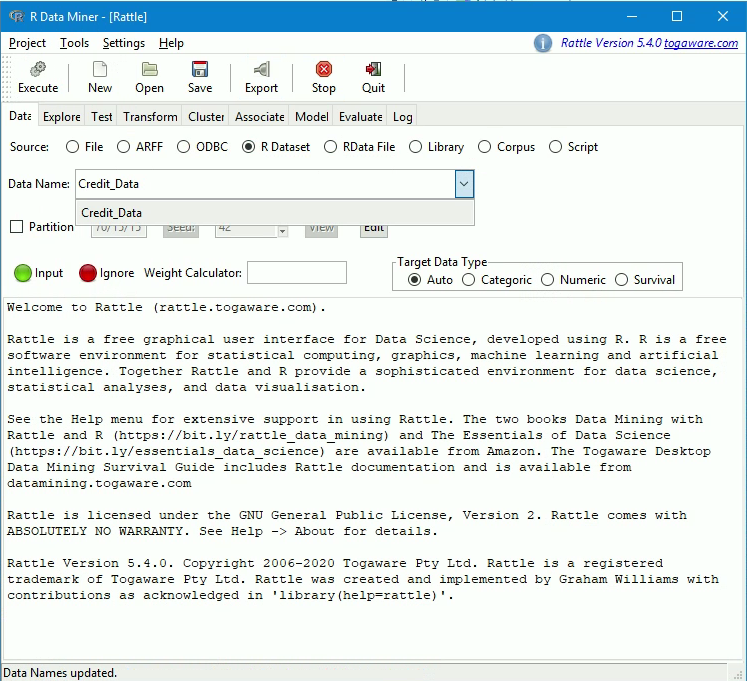
**Correlation Testing**

>cor(newdata$variableA, newdata$variableB)  
#outputs correlation coefficient between two specified variables. This may determine what variables are useful for the model.

**Building the Model**

>library(rattle)

>rattle()  
#this will open rattle. From here, instructions will be interface based

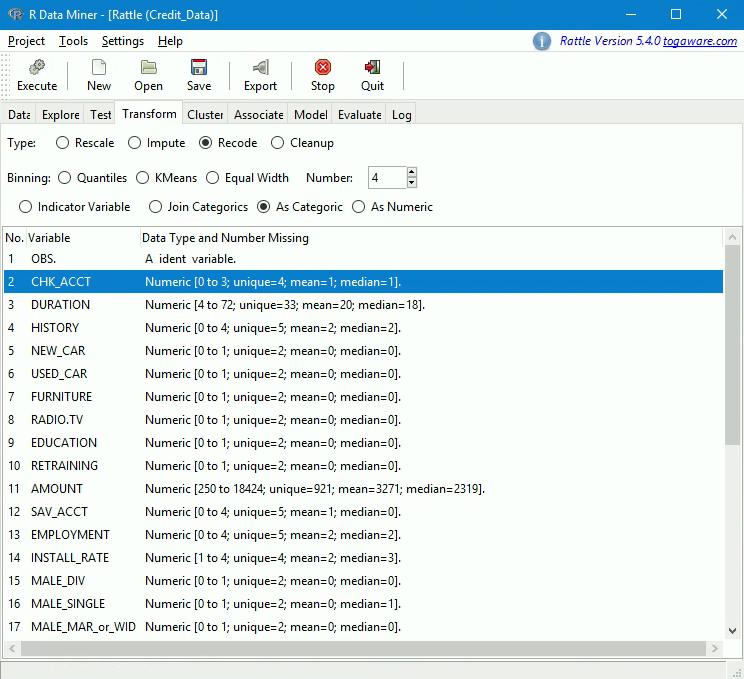


-Using R Dataset, load in your newdata (displayed here as Credit\_Data)

-Check “partition”

- Execute

-Go to Transform tab



-Recode, As Categoric, Select categorical variables and hit Execute. Do this for all relevant categorical variables to be included in the model

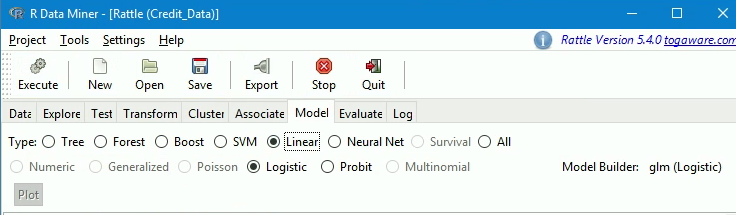
-Return to Data tab and set all newly transformed variables to “Input” as well as any other variables to be used in the model. Ensure OBS is set to “Ident” and TFC\_DEFAULT is set to “Target”

-Input variables: AMOUNT, INSTALL\_RATE, TFC\_CHK\_ACCT, TFC\_HISTORY,  
TFC\_SAV\_ACCT, TFC\_EMPLOYMENT, TFC\_CO.APPLICANT, TFC\_GUARANTOR, DURATION

-Set all others to “Ignore”

-Execute

-Enter Model tab



-Set to Linear: Logistic

-Execute

-The model will run its outputs

-Save

**Validation/New Data**

>library(readr)

>valdata <- read.csv(“file path”)

-Follow the same loading and cleaning steps as the training data. Prepare the validation data as follows

>library(rattle)

>rattle()

-The rattle window will open. Follow the same loading and transform steps as before, but do not go to the model tab.

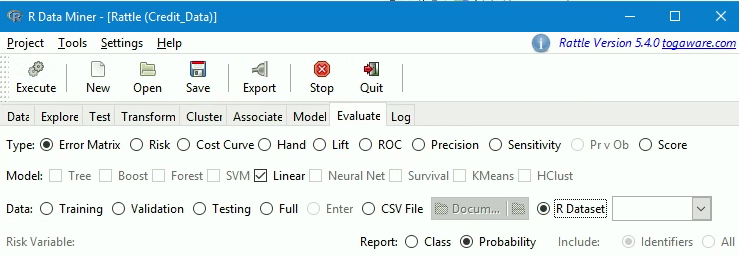
-Instead, Export and save the transformed datafile. Then load that datafile into R.

>valdata\_transform <- read.csv(“file path”)

>rattle()

-Opening rattle once more. Load the model using “Open”.

-Enter Evaluate tab



-Select Error Matrix, Linear, R Dataset: valdata\_transform

-Execute

-Rattle will present an error matrix on the validation data

-Select Pr v Ob to generate a csv making a prediction for every entry in the validation set