LeTicia Cancel

DATA622 Homework #4

December 22, 2022

Introduction & Data

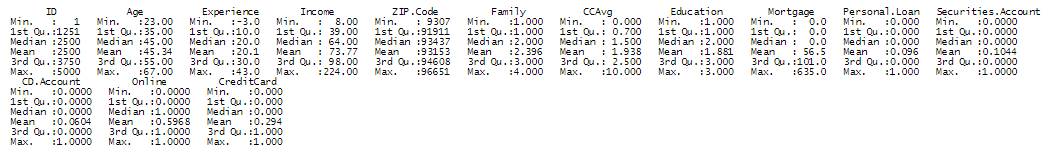
Loan Predictions

Data

Data column descriptions

|  |  |
| --- | --- |
| Age | Customer's age in completed years |
| Experience | #years of professional experience |
| Income | Annual income of the customer ($000) |
| Zip | Home Address ZIP code. |
| Family | Family size of the customer |
| CCAvg | Avg. spending on credit cards per month ($000) |
| Education | Education Level.  1: Undergrad;  2: Graduate;  3: Advanced/Professional |
| Mortgage | Value of house mortgage if any. ($000) |
| Personal Loan | Did this customer accept the personal loan offered in the last campaign? |
| Securities Account | Does the customer have a securities account with the bank? |
| CD Account | Does the customer have a certificate of deposit (CD) account with the bank? |
| Online | Does the customer use internet banking facilities? |
| CeditCard | Does the customer use a credit card issued by UniversalBank? |

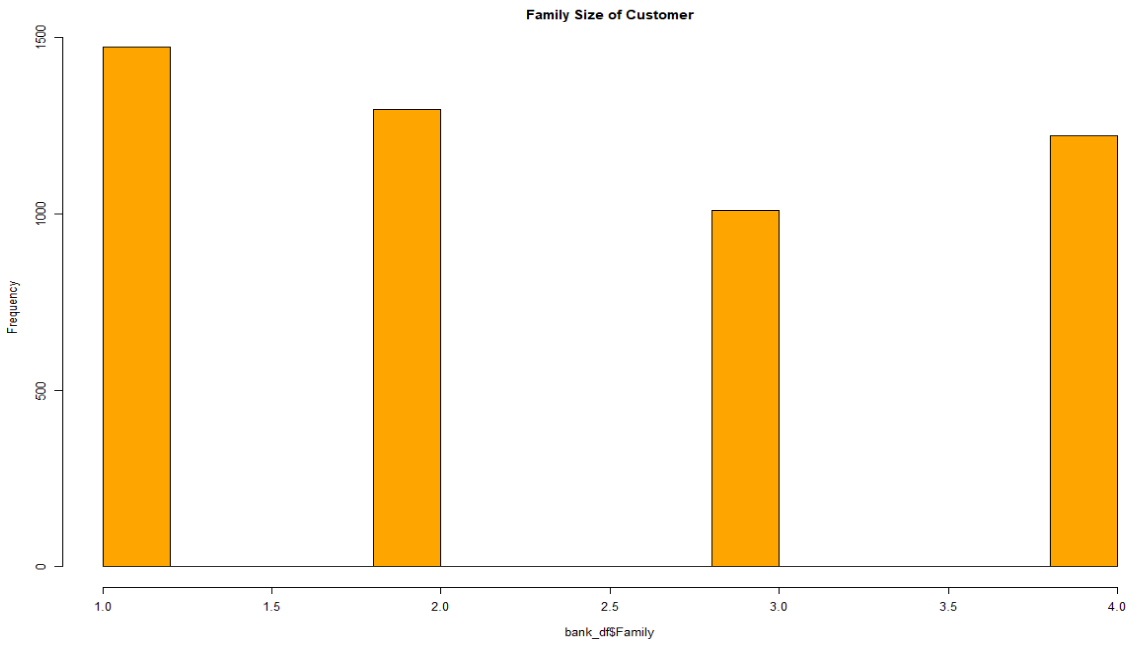
Summary of the dataframe. There are no N/A values

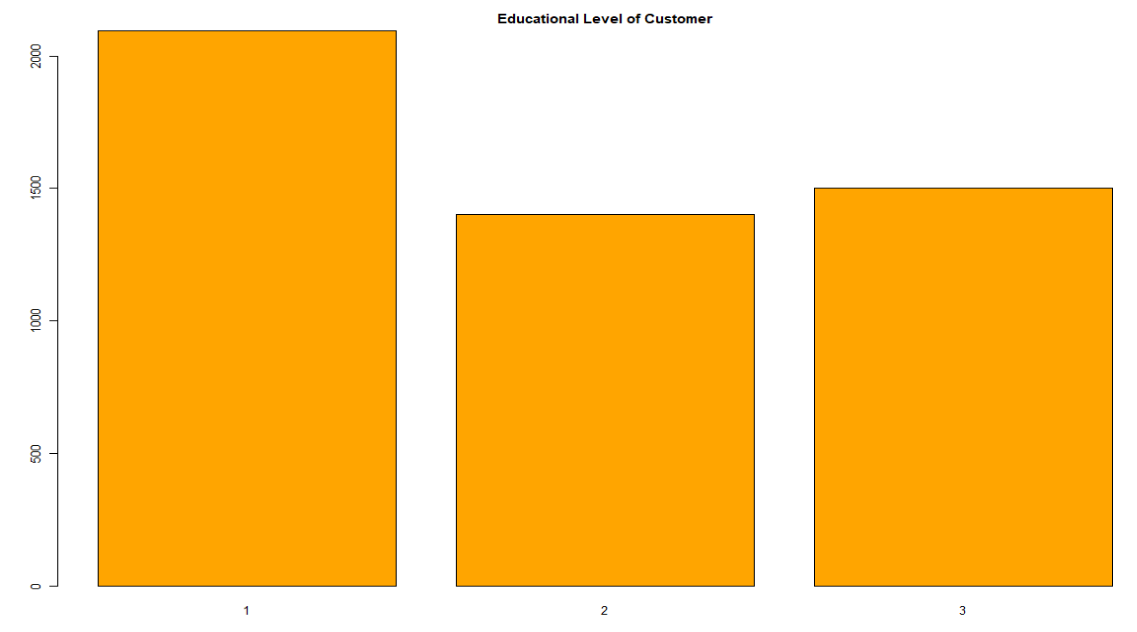


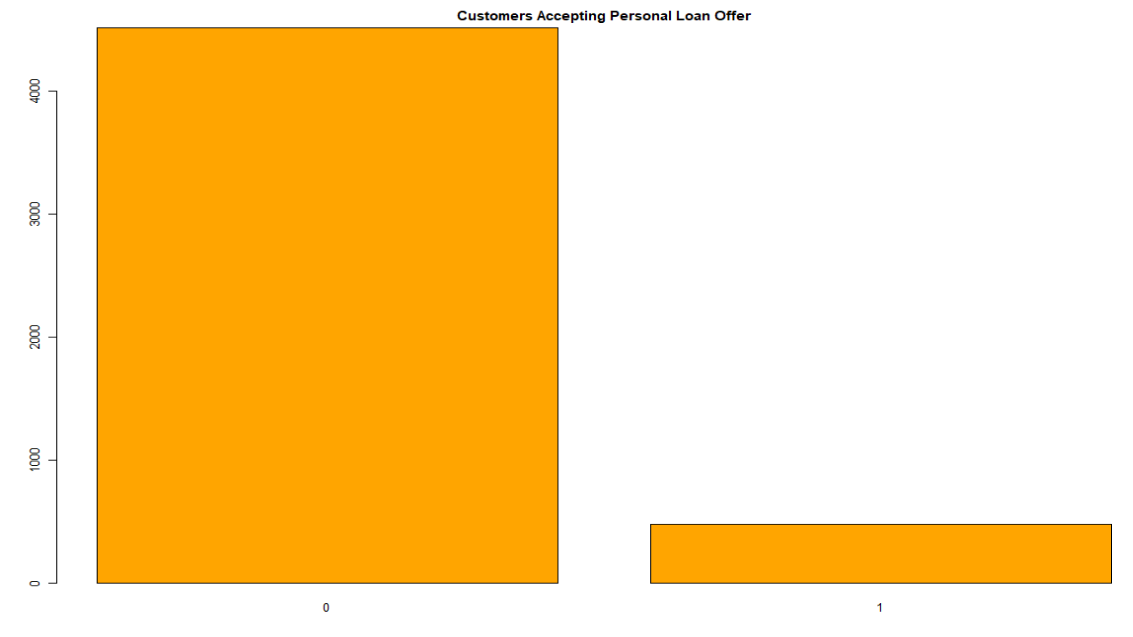
Convert categorical and bit columns to Factor:

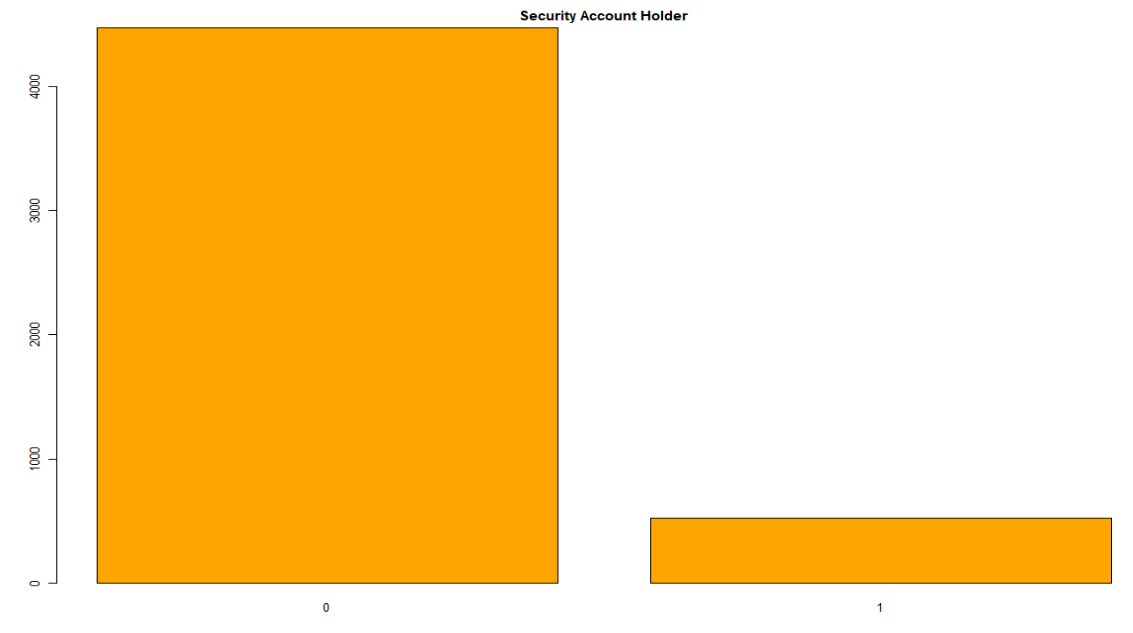
* Education
* Personal Loan
* Securities Account
* CD Account
* Online
* CreditCard

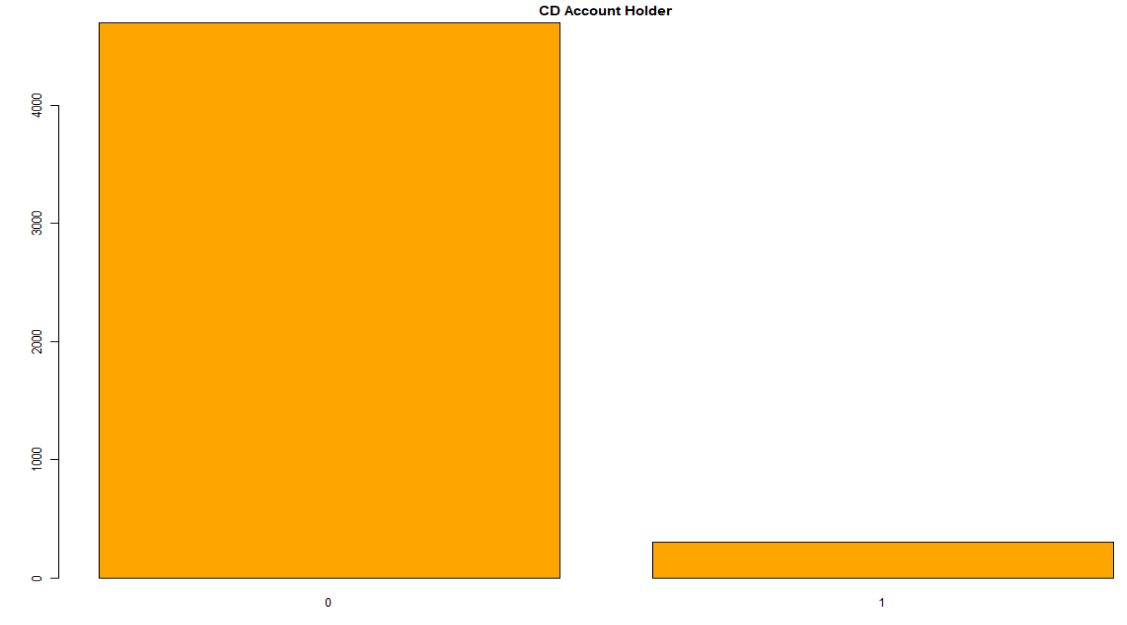
Exploring the columns

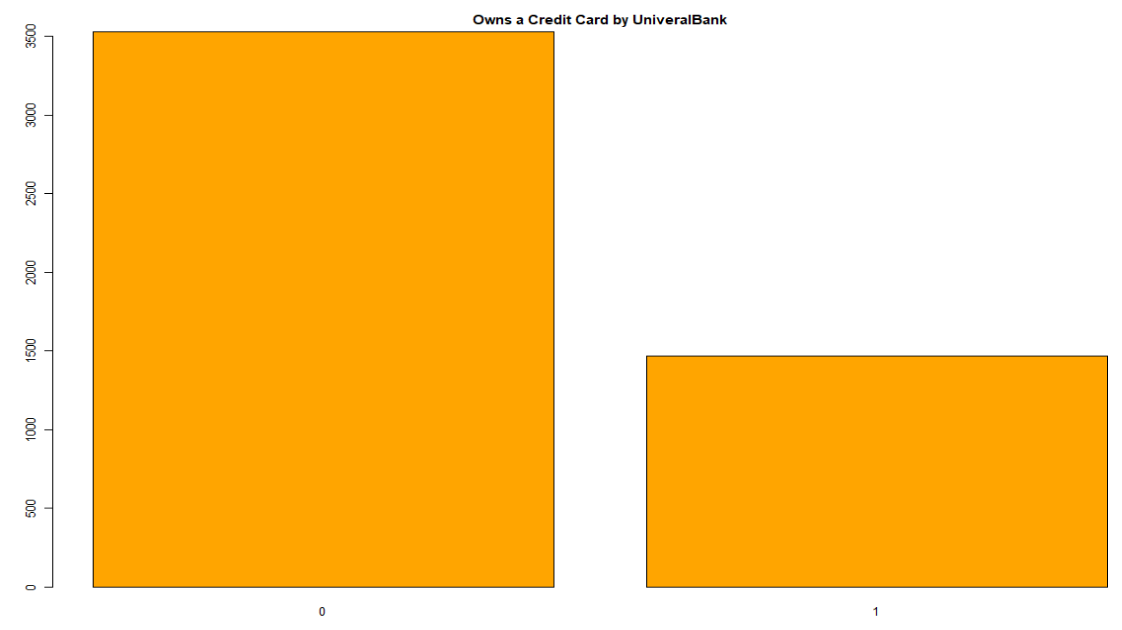




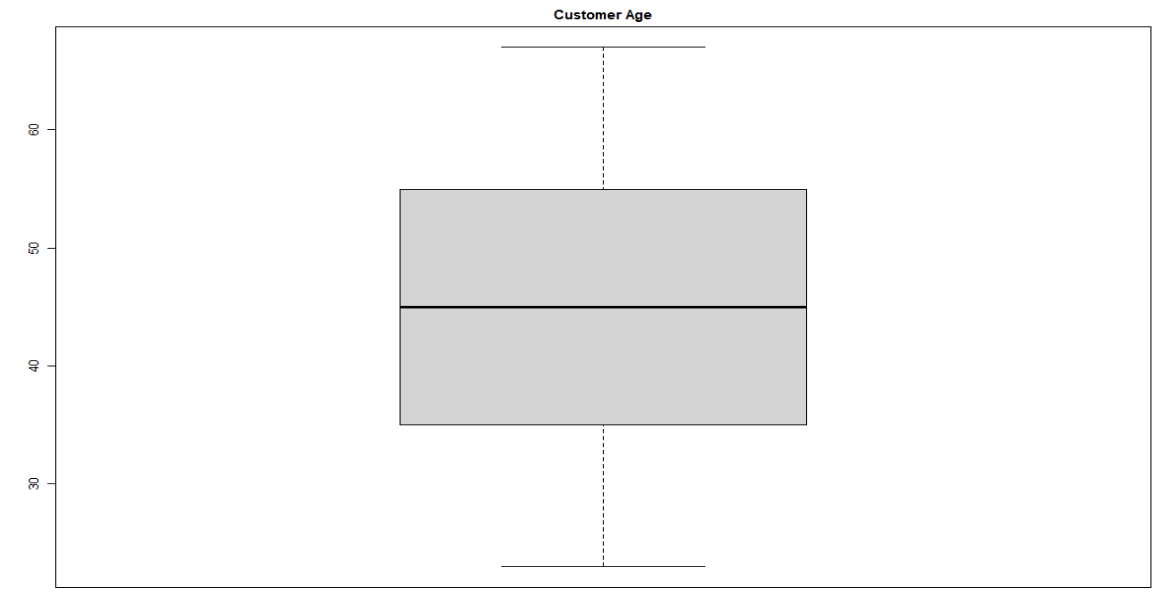


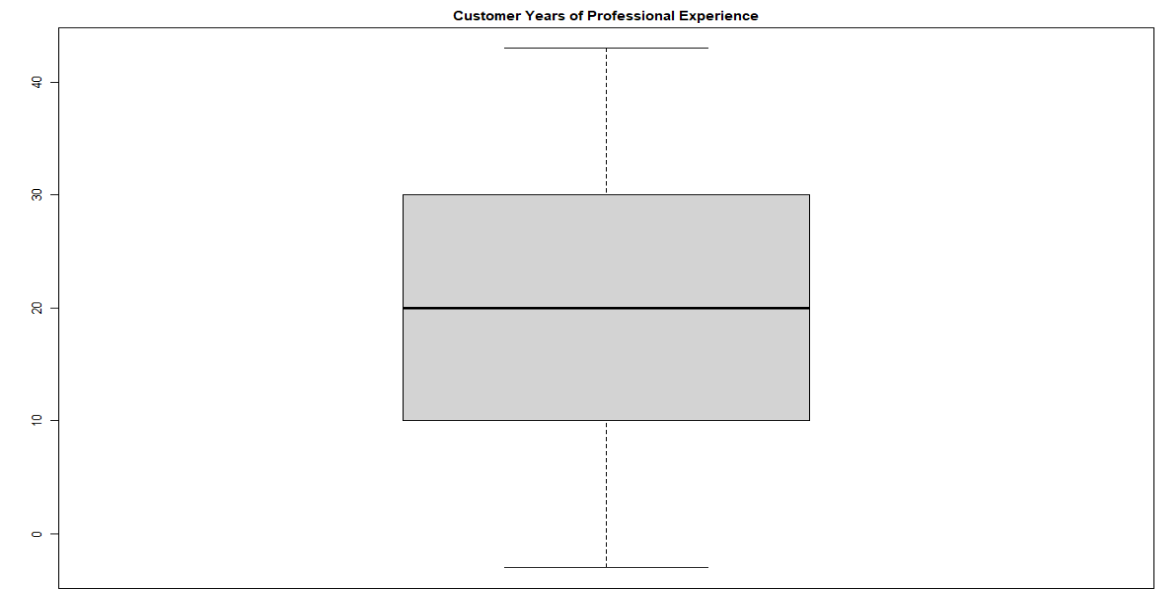


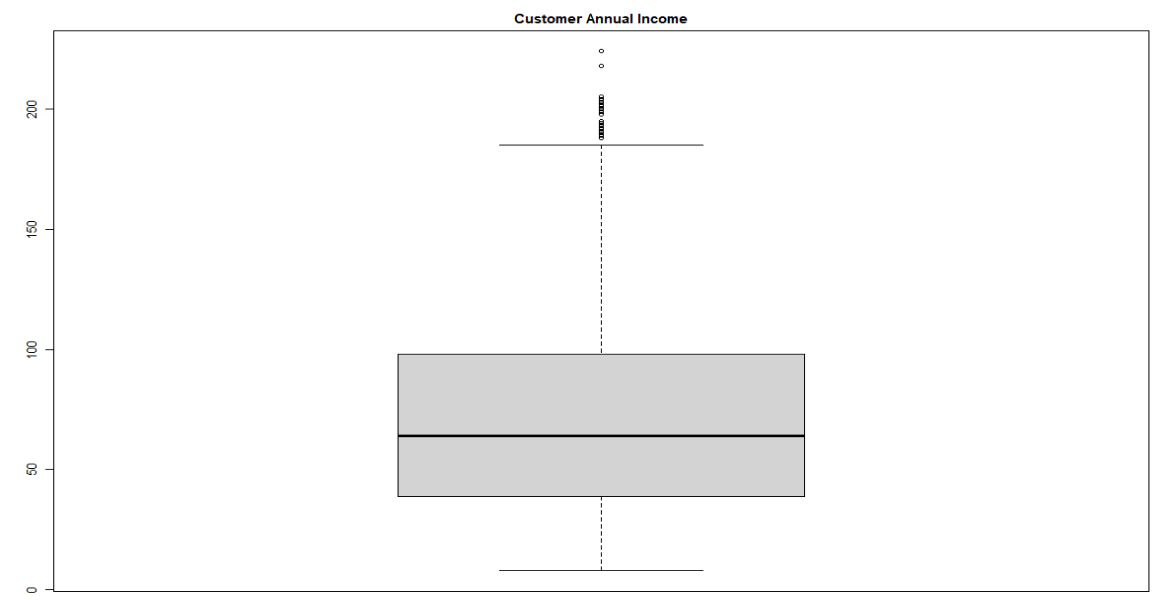


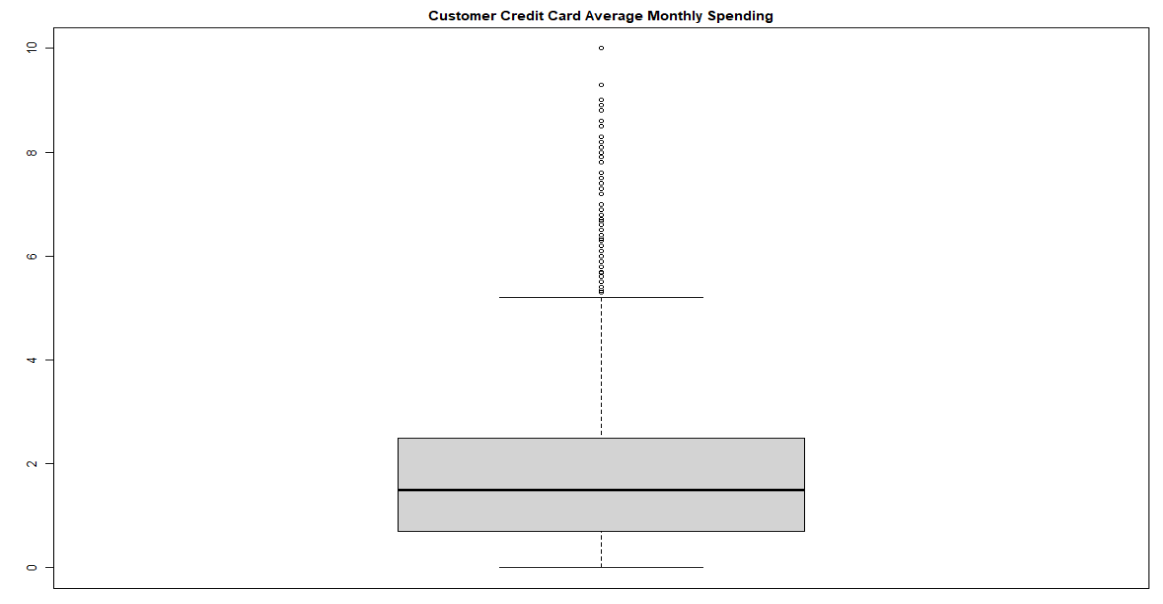


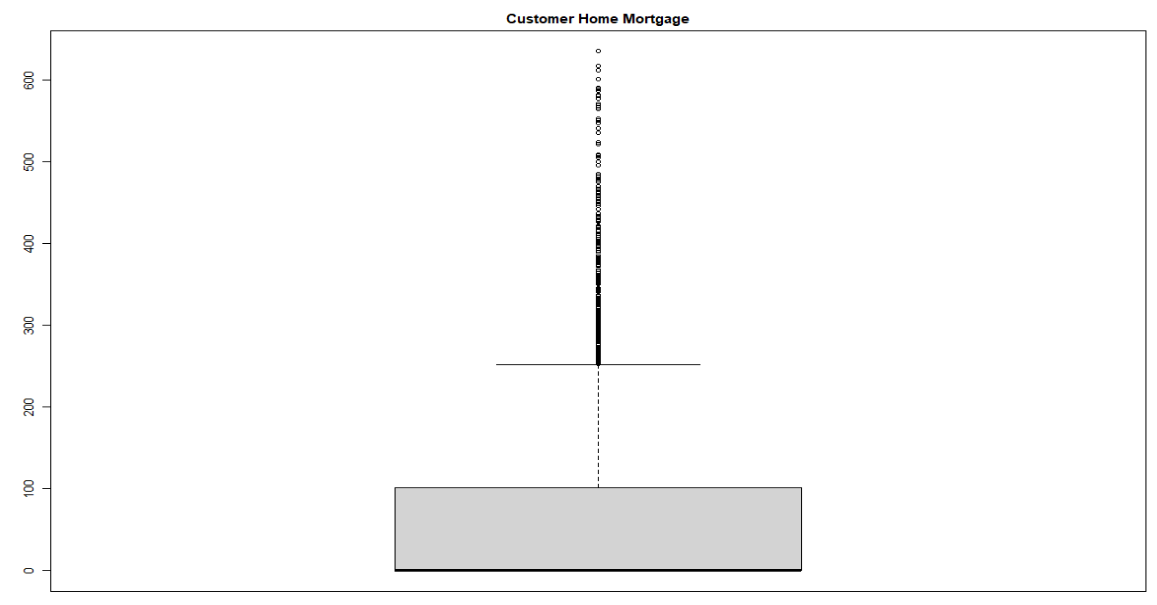
Boxplot of continuous variables



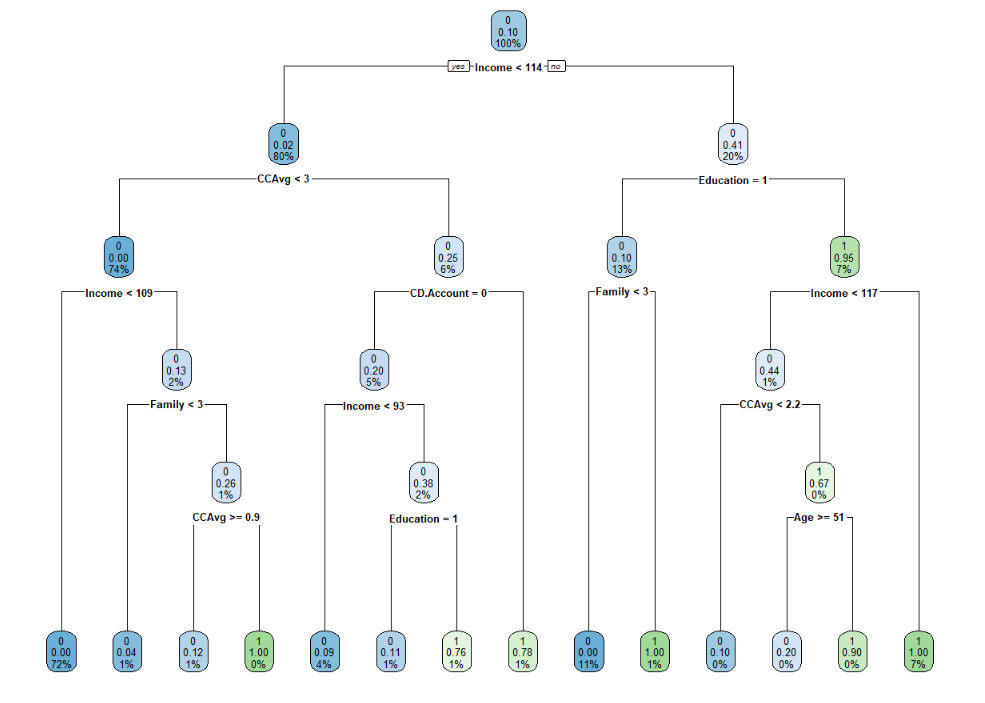








Modeling



Confusion Matrix of Decision Tree has a 98.5% accuracy.

