Churn Prediction

Denis Orina

11/15/2021

Contents

[Introduction 1](#_Toc91571737)

[Objectives 1](#_Toc91571738)

[Data Source 1](#_Toc91571739)

[Data Analysis 1](#_Toc91571740)

[Preliminary Analysis 1](#_Toc91571741)

[Overview 1](#_Toc91571742)

[Exploratory Data Analysis 6](#_Toc91571743)

[Univariate Distribution 6](#_Toc91571744)

[Bivariate Analysis 12](#_Toc91571745)

[Data Pre processing 41](#_Toc91571746)

[Drop Unncecessary Columns 41](#_Toc91571747)

[Dealing with Missing Values 41](#_Toc91571748)

[Dealing with Skewness and abnormality 41](#_Toc91571749)

[Train test Split 41](#_Toc91571750)

[Upsampling Minority Class 41](#_Toc91571751)

[Modeling 41](#_Toc91571752)

[Logistic regression 41](#_Toc91571753)

[Decision Trees 43](#_Toc91571754)

[Random Forest 43](#_Toc91571755)

# Introduction

## Objectives

* Predict Churn probabilities of customers and minimize.

## Data Source

1. Customer Listing
2. Transactions table
3. Loan Book

# Data Analysis

## Preliminary Analysis

## Overview

## Warning: The `x` argument of `as\_tibble.matrix()` must have unique column names if `.name\_repair` is omitted as of tibble 2.0.0.  
## Using compatibility `.name\_repair`.

| Variable | Value |
| --- | --- |
| rows | 32,506 |
| columns | 39 |
| discrete\_columns | 3 |
| continuous\_columns | 36 |
| all\_missing\_columns | 0 |
| total\_missing\_values | 56,095 |
| complete\_rows | 28,191 |
| total\_observations | 1,267,734 |

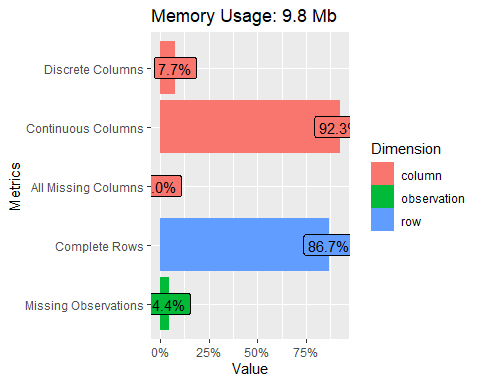


Table 1: Dataset Variables

| **Column** | **Explanation** | **Datatype** |
| --- | --- | --- |
| ClientID |  |  |
| NumberOfProducts |  |  |
| Tenure |  |  |
| No\_of\_transactions |  |  |
| No\_of\_transactions\_Deposits |  |  |
| TotalDepositsMade |  |  |
| Max\_DepositsMade |  |  |
| Average\_DepositsMade |  |  |
| No\_of\_transactions\_Withdrawals |  |  |
| WithdrawalMade |  |  |
| Max\_WithdrawalMade |  |  |
| LastTrxDate |  |  |
| Days\_Since\_LastTransaction |  |  |
| transactions\_made\_12 |  |  |
| transactions\_made\_6 |  |  |
| transactions\_made\_3 |  |  |
| deposits\_12 |  |  |
| deposits\_6 |  |  |
| deposits\_3 |  |  |
| withdrawals\_12 |  |  |
| withdrawals\_6 |  |  |
| withdrawals\_3 |  |  |
| last\_deposit\_amount |  |  |
| last\_withdrawal\_amount |  |  |
| Max\_LoanTaken |  |  |
| Min\_LoanTaken |  |  |
| ave\_LoanTaken |  |  |
| no\_of\_LoanTaken |  |  |
| LastLoanDate |  |  |
| DaysSince\_Last\_Loan |  |  |
| LastLoanTaken |  |  |
| FirstLoanTaken |  |  |
| Has\_ActiveLoan |  |  |
| Max\_ArrearsDays |  |  |
| First\_Loan\_Last\_Loan\_Ratio |  |  |
| First\_Loan\_Last\_Loan\_diff |  |  |
| Last\_Loan\_First\_Loan\_Ratio |  |  |
| Trx\_Status |  |  |
| Churn |  |  |

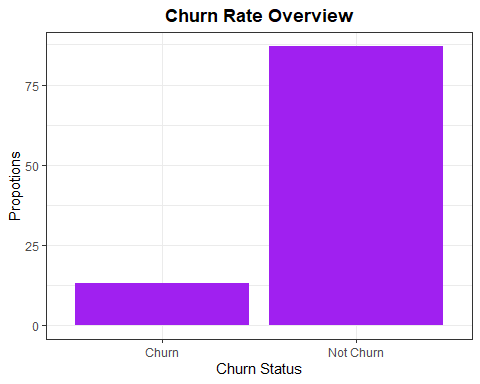
## Exploratory Data Analysis

### Univariate Distribution

#### Target Variable

Table 2: Churn Rate

| **Churn** | **Count** | **Propotions** |
| --- | --- | --- |
| Churn | 4,315 | 13 |
| Not Churn | 28,191 | 87 |
| Total | 32,506 | 100 |



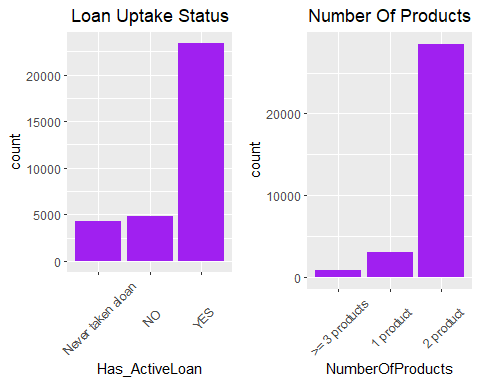
**Observations and inference**

* The churn rate for this institution is very low as compared to the overall client count at 13% and 87% respectively.

#### Explanatory variables.

##### Categorical variables

| variable | value | Count | Proportion |
| --- | --- | --- | --- |
| Has\_ActiveLoan | Never taken aloan | 4,315 | 13 % |
| Has\_ActiveLoan | NO | 4,788 | 15 % |
| Has\_ActiveLoan | YES | 23,403 | 72 % |
| Total | - | 32,506 | - |
| NumberOfProducts | >= 3 products | 838 | 3 % |
| NumberOfProducts | 1 product | 3,128 | 10 % |
| NumberOfProducts | 2 product | 28,540 | 88 % |
| Total | - | 32,506 | - |



**Observations and inferences**

* Majority of clients have 2 products as compared to the remaining < 50% with more than one product.
* Majority of clients have existing running loans with the institution.

##### Numerical Variables

| value | mean | variance | sd | max | min | 0% | 100% | 25% | 50% | 75% |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Tenure | 19.57 | 10.70 | 3.27 | 65.00 | 0.00 | 0.00 | 65.00 | 20.00 | 20.00 | 21.00 |
| No\_of\_transactions | 22.16 | 1,419.17 | 37.67 | 3,114.00 | 1.00 | 1.00 | 3,114.00 | 7.00 | 8.00 | 20.00 |
| No\_of\_transactions\_Deposits | 6.22 | 173.35 | 13.17 | 939.00 | 0.00 | 0.00 | 939.00 | 1.00 | 2.00 | 6.00 |
| TotalDepositsMade | 40,773.15 | 30,191,976,859.33 | 173,758.39 | 15,869,967.98 | 0.00 | 0.00 | 15,869,967.98 | 1,120.00 | 6,272.50 | 27,618.75 |
| Max\_DepositsMade | 7,850.02 | 1,390,436,850.47 | 37,288.56 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 1,120.00 | 4,145.00 | 10,080.00 |
| Average\_DepositsMade | 1,507.12 | 6,958,201.71 | 2,637.84 | 275,108.80 | 0.00 | 0.00 | 275,108.80 | 318.57 | 890.29 | 2,006.21 |
| No\_of\_transactions\_Withdrawals | 15.58 | 666.20 | 25.81 | 2,161.00 | 0.00 | 0.00 | 2,161.00 | 6.00 | 6.00 | 13.00 |
| WithdrawalMade | 45,625.00 | 30,689,245,553.01 | 175,183.46 | 16,171,410.72 | 0.00 | 0.00 | 16,171,410.72 | 2,240.00 | 10,960.00 | 38,080.00 |
| Max\_WithdrawalMade | 7,732.10 | 1,344,203,060.35 | 36,663.37 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 1,120.00 | 4,070.00 | 10,080.00 |
| Days\_Since\_LastTransaction | 415.69 | 55,963.45 | 236.57 | 751.00 | 2.00 | 2.00 | 751.00 | 185.00 | 567.00 | 615.00 |
| transactions\_made\_12 | 8.57 | 485.18 | 22.03 | 665.00 | 0.00 | 0.00 | 665.00 | 0.00 | 0.00 | 3.00 |
| transactions\_made\_6 | 3.94 | 142.77 | 11.95 | 579.00 | 0.00 | 0.00 | 579.00 | 0.00 | 0.00 | 0.00 |
| transactions\_made\_3 | 1.98 | 45.82 | 6.77 | 408.00 | 0.00 | 0.00 | 408.00 | 0.00 | 0.00 | 0.00 |
| deposits\_12 | 607.79 | 906,971,233.26 | 30,115.96 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 0.00 | 0.00 | 0.00 |
| deposits\_6 | 544.88 | 856,756,427.81 | 29,270.40 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 0.00 | 0.00 | 0.00 |
| deposits\_3 | 538.85 | 856,685,933.61 | 29,269.20 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 0.00 | 0.00 | 0.00 |
| withdrawals\_12 | 2,844.73 | 50,513,498.36 | 7,107.28 | 700,000.00 | 0.00 | 0.00 | 700,000.00 | 0.00 | 0.00 | 3,472.00 |
| withdrawals\_6 | 2,003.19 | 27,379,037.20 | 5,232.50 | 220,000.00 | 0.00 | 0.00 | 220,000.00 | 0.00 | 0.00 | 0.00 |
| withdrawals\_3 | 1,435.58 | 20,616,978.23 | 4,540.59 | 220,000.00 | 0.00 | 0.00 | 220,000.00 | 0.00 | 0.00 | 0.00 |
| last\_deposit\_amount | 5,269.24 | 155,371,068.92 | 12,464.79 | 1,434,950.00 | 0.00 | 0.00 | 1,434,950.00 | 200.00 | 1,120.00 | 5,998.75 |
| last\_withdrawal\_amount | 12,516.80 | 1,348,495,850.51 | 36,721.87 | 2,480,237.00 | 0.00 | 0.00 | 2,480,237.00 | 2,240.00 | 6,272.00 | 15,680.00 |
| Max\_LoanTaken | 7,849.72 | 387,205,799.23 | 19,677.55 | 1,200,000.00 | 590.00 | 590.00 | 1,200,000.00 | 1,120.00 | 5,488.00 | 9,900.00 |
| Min\_LoanTaken | 7,701.02 | 546,603,477.05 | 23,379.55 | 1,250,000.00 | 1,100.00 | 1,100.00 | 1,250,000.00 | 1,120.00 | 4,816.00 | 11,000.00 |
| ave\_LoanTaken | 7,781.98 | 363,352,280.82 | 19,061.80 | 1,200,000.00 | 992.25 | 992.25 | 1,200,000.00 | 1,120.00 | 5,264.00 | 10,825.77 |
| no\_of\_LoanTaken | 2.78 | 15.30 | 3.91 | 40.00 | 1.00 | 1.00 | 40.00 | 1.00 | 1.00 | 2.00 |
| LastLoanDate | 44,011.18 | 44,466.22 | 210.87 | 44,510.00 | 43,753.00 | 43,753.00 | 44,510.00 | 43,882.00 | 43,901.00 | 44,035.00 |
| DaysSince\_Last\_Loan | 499.82 | 44,466.22 | 210.87 | 758.00 | 1.00 | 1.00 | 758.00 | 476.00 | 610.00 | 629.00 |
| LastLoanTaken | 8,150.94 | 429,223,788.71 | 20,717.72 | 1,200,000.00 | 1,100.00 | 1,100.00 | 1,200,000.00 | 1,120.00 | 5,488.00 | 11,088.00 |
| FirstLoanTaken | 7,760.17 | 574,032,255.44 | 23,958.97 | 1,300,000.00 | 590.00 | 590.00 | 1,300,000.00 | 1,120.00 | 5,040.00 | 10,416.00 |
| Max\_ArrearsDays | 389.85 | 68,070.84 | 260.90 | 615.00 | -28.00 | -28.00 | 615.00 | 16.00 | 573.00 | 594.00 |
| First\_Loan\_Last\_Loan\_Ratio | 0.98 | 1.35 | 1.16 | 159.44 | 0.03 | 0.03 | 159.44 | 1.00 | 1.00 | 1.00 |
| First\_Loan\_Last\_Loan\_diff | -390.76 | 203,180,954.87 | 14,254.16 | 1,272,500.00 | -600,000.00 | -600,000.00 | 1,272,500.00 | 0.00 | 0.00 | 0.00 |
| Last\_Loan\_First\_Loan\_Ratio | 1.14 | 0.87 | 0.93 | 31.00 | 0.01 | 0.01 | 31.00 | 1.00 | 1.00 | 1.00 |
| Trx\_Status | 1.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

### Bivariate Analysis

##### Descriptives

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

## Warning in min(x, na.rm = T): no non-missing arguments to min; returning Inf

## Warning in max(x, na.rm = T): no non-missing arguments to max; returning -Inf

| Variable | ChurnStatus | mean | variance | sd | max | min | 0% | 100% | 25% | 50% | 75% |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ave\_LoanTaken | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| ave\_LoanTaken | Not\_Churn | 7,781.98 | 363,352,280.82 | 19,061.80 | 1,200,000.00 | 992.25 | 992.25 | 1,200,000.00 | 1,120.0 | 5,264.00 | 10,825.77 |
| Average\_DepositsMade | Churn | 2.77 | 82.17 | 9.06 | 241.00 | 0.00 | 0.00 | 241.00 | 0.0 | 1.00 | 3.00 |
| Average\_DepositsMade | Not\_Churn | 17.55 | 726.61 | 26.96 | 2,161.00 | 1.00 | 1.00 | 2,161.00 | 6.0 | 6.00 | 16.00 |
| Days\_Since\_LastTransaction | Churn | 438.64 | 35,931.92 | 189.56 | 751.00 | 2.00 | 2.00 | 751.00 | 272.0 | 538.00 | 600.00 |
| Days\_Since\_LastTransaction | Not\_Churn | 412.18 | 58,937.98 | 242.77 | 694.00 | 2.00 | 2.00 | 694.00 | 157.0 | 576.00 | 619.00 |
| DaysSince\_Last\_Loan | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| DaysSince\_Last\_Loan | Not\_Churn | 499.82 | 44,466.22 | 210.87 | 758.00 | 1.00 | 1.00 | 758.00 | 476.0 | 610.00 | 629.00 |
| deposits\_12 | Churn | 1,263.59 | 3,746,228,103.52 | 61,206.44 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 0.0 | 0.00 | 50.00 |
| deposits\_12 | Not\_Churn | 507.42 | 472,427,531.02 | 21,735.40 | 1,300,000.00 | 0.00 | 0.00 | 1,300,000.00 | 0.0 | 0.00 | 0.00 |
| deposits\_3 | Churn | 1,150.35 | 3,745,104,674.52 | 61,197.26 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 0.0 | 0.00 | 0.00 |
| deposits\_3 | Not\_Churn | 445.25 | 414,626,967.69 | 20,362.39 | 1,300,000.00 | 0.00 | 0.00 | 1,300,000.00 | 0.0 | 0.00 | 0.00 |
| deposits\_6 | Churn | 1,186.84 | 3,745,315,100.96 | 61,198.98 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 0.0 | 0.00 | 0.00 |
| deposits\_6 | Not\_Churn | 446.62 | 414,669,311.59 | 20,363.43 | 1,300,000.00 | 0.00 | 0.00 | 1,300,000.00 | 0.0 | 0.00 | 0.00 |
| First\_Loan\_Last\_Loan\_diff | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| First\_Loan\_Last\_Loan\_diff | Not\_Churn | -390.76 | 203,180,954.87 | 14,254.16 | 1,272,500.00 | -600,000.00 | -600,000.00 | 1,272,500.00 | 0.0 | 0.00 | 0.00 |
| First\_Loan\_Last\_Loan\_Ratio | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| First\_Loan\_Last\_Loan\_Ratio | Not\_Churn | 0.98 | 1.35 | 1.16 | 159.44 | 0.03 | 0.03 | 159.44 | 1.0 | 1.00 | 1.00 |
| FirstLoanTaken | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| FirstLoanTaken | Not\_Churn | 7,760.17 | 574,032,255.44 | 23,958.97 | 1,300,000.00 | 590.00 | 590.00 | 1,300,000.00 | 1,120.0 | 5,040.00 | 10,416.00 |
| last\_deposit\_amount | Churn | 635.37 | 480,653,690.81 | 21,923.82 | 1,434,950.00 | 0.00 | 0.00 | 1,434,950.00 | 5.0 | 50.00 | 100.00 |
| last\_deposit\_amount | Not\_Churn | 5,978.51 | 101,807,723.22 | 10,089.98 | 530,000.00 | 0.00 | 0.00 | 530,000.00 | 1,000.0 | 1,568.00 | 7,761.00 |
| Last\_Loan\_First\_Loan\_Ratio | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| Last\_Loan\_First\_Loan\_Ratio | Not\_Churn | 1.14 | 0.87 | 0.93 | 31.00 | 0.01 | 0.01 | 31.00 | 1.0 | 1.00 | 1.00 |
| last\_withdrawal\_amount | Churn | 625.29 | 482,871,137.23 | 21,974.33 | 1,434,950.00 | 0.00 | 0.00 | 1,434,950.00 | 0.0 | 0.00 | 6.00 |
| last\_withdrawal\_amount | Not\_Churn | 14,336.95 | 1,456,054,672.03 | 38,158.28 | 2,480,237.00 | 0.00 | 0.00 | 2,480,237.00 | 2,240.0 | 8,960.00 | 18,815.00 |
| LastLoanDate | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| LastLoanDate | Not\_Churn | 44,011.18 | 44,466.22 | 210.87 | 44,510.00 | 43,753.00 | 43,753.00 | 44,510.00 | 43,882.0 | 43,901.00 | 44,035.00 |
| LastLoanTaken | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| LastLoanTaken | Not\_Churn | 8,150.94 | 429,223,788.71 | 20,717.72 | 1,200,000.00 | 1,100.00 | 1,100.00 | 1,200,000.00 | 1,120.0 | 5,488.00 | 11,088.00 |
| Max\_ArrearsDays | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| Max\_ArrearsDays | Not\_Churn | 389.85 | 68,070.84 | 260.90 | 615.00 | -28.00 | -28.00 | 615.00 | 16.0 | 573.00 | 594.00 |
| Max\_DepositsMade | Churn | 210.54 | 21,685,108.90 | 4,656.73 | 275,108.80 | 0.00 | 0.00 | 275,108.80 | 0.0 | 2.00 | 44.44 |
| Max\_DepositsMade | Not\_Churn | 1,705.58 | 4,408,030.14 | 2,099.53 | 95,998.40 | 45.23 | 45.23 | 95,998.40 | 320.0 | 1,142.86 | 2,249.71 |
| Max\_WithdrawalMade | Churn | 1,874.72 | 3,907,725,918.45 | 62,511.81 | 4,000,000.00 | 0.00 | 0.00 | 4,000,000.00 | 0.0 | 5.00 | 140.00 |
| Max\_WithdrawalMade | Not\_Churn | 8,628.64 | 945,891,749.16 | 30,755.35 | 2,000,000.00 | 1,100.00 | 1,100.00 | 2,000,000.00 | 1,120.0 | 5,500.00 | 11,200.00 |
| no\_of\_LoanTaken | Churn | 0.00 | 0.00 | 0.00 | -Inf | Inf | 0.00 | 0.00 | 0.0 | 0.00 | 0.00 |
| no\_of\_LoanTaken | Not\_Churn | 2.78 | 15.30 | 3.91 | 40.00 | 1.00 | 1.00 | 40.00 | 1.0 | 1.00 | 2.00 |
| No\_of\_transactions | Churn | 5.52 | 177.53 | 13.32 | 298.00 | 1.00 | 1.00 | 298.00 | 1.0 | 3.00 | 5.00 |
| No\_of\_transactions | Not\_Churn | 24.70 | 1,560.41 | 39.50 | 3,114.00 | 1.00 | 1.00 | 3,114.00 | 7.0 | 10.00 | 23.00 |
| No\_of\_transactions\_Deposits | Churn | 2.48 | 26.37 | 5.14 | 121.00 | 0.00 | 0.00 | 121.00 | 1.0 | 1.00 | 2.00 |
| No\_of\_transactions\_Deposits | Not\_Churn | 6.79 | 193.38 | 13.91 | 939.00 | 0.00 | 0.00 | 939.00 | 1.0 | 2.00 | 7.00 |
| No\_of\_transactions\_Withdrawals | Churn | 5,562.32 | 18,572,693,134.62 | 136,281.67 | 8,268,800.00 | 1.00 | 1.00 | 8,268,800.00 | 50.0 | 100.00 | 360.00 |
| No\_of\_transactions\_Withdrawals | Not\_Churn | 46,162.63 | 31,752,359,452.70 | 178,191.92 | 15,869,967.98 | 0.00 | 0.00 | 15,869,967.98 | 2,296.5 | 9,556.00 | 30,383.00 |
| Tenure | Churn | 16.77 | 29.27 | 5.41 | 56.00 | 0.00 | 0.00 | 56.00 | 14.0 | 20.00 | 20.00 |
| Tenure | Not\_Churn | 20.00 | 6.46 | 2.54 | 65.00 | 0.00 | 0.00 | 65.00 | 20.0 | 20.00 | 21.00 |
| TotalDepositsMade | Churn | 1,954.86 | 3,900,852,835.37 | 62,456.81 | 4,000,000.00 | 1.00 | 1.00 | 4,000,000.00 | 50.0 | 100.00 | 200.00 |
| TotalDepositsMade | Not\_Churn | 8,752.35 | 1,000,175,910.50 | 31,625.56 | 2,000,000.00 | 0.00 | 0.00 | 2,000,000.00 | 1,200.0 | 5,500.00 | 11,200.00 |
| transactions\_made\_12 | Churn | 2.37 | 101.84 | 10.09 | 298.00 | 0.00 | 0.00 | 298.00 | 0.0 | 0.00 | 1.00 |
| transactions\_made\_12 | Not\_Churn | 9.52 | 537.09 | 23.18 | 665.00 | 0.00 | 0.00 | 665.00 | 0.0 | 0.00 | 4.00 |
| transactions\_made\_3 | Churn | 0.52 | 17.65 | 4.20 | 126.00 | 0.00 | 0.00 | 126.00 | 0.0 | 0.00 | 0.00 |
| transactions\_made\_3 | Not\_Churn | 2.21 | 49.75 | 7.05 | 408.00 | 0.00 | 0.00 | 408.00 | 0.0 | 0.00 | 0.00 |
| transactions\_made\_6 | Churn | 1.00 | 49.13 | 7.01 | 221.00 | 0.00 | 0.00 | 221.00 | 0.0 | 0.00 | 0.00 |
| transactions\_made\_6 | Not\_Churn | 4.39 | 155.58 | 12.47 | 579.00 | 0.00 | 0.00 | 579.00 | 0.0 | 0.00 | 1.00 |
| Trx\_Status | Churn | 1.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.0 | 1.00 | 1.00 |
| Trx\_Status | Not\_Churn | 1.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.0 | 1.00 | 1.00 |
| WithdrawalMade | Churn | 4,452.20 | 6,685,237,766.87 | 81,763.30 | 4,126,632.00 | 0.00 | 0.00 | 4,126,632.00 | 0.0 | 6.00 | 260.00 |
| WithdrawalMade | Not\_Churn | 51,927.03 | 34,064,540,710.91 | 184,565.82 | 16,171,410.72 | 1,120.00 | 1,120.00 | 16,171,410.72 | 2,884.0 | 14,526.00 | 44,800.00 |
| withdrawals\_12 | Churn | 162.81 | 113,558,124.91 | 10,656.37 | 700,000.00 | 0.00 | 0.00 | 700,000.00 | 0.0 | 0.00 | 0.00 |
| withdrawals\_12 | Not\_Churn | 3,255.23 | 39,597,892.87 | 6,292.69 | 220,000.00 | 0.00 | 0.00 | 220,000.00 | 0.0 | 0.00 | 5,183.40 |
| withdrawals\_3 | Churn | 0.46 | 927.00 | 30.45 | 2,000.00 | 0.00 | 0.00 | 2,000.00 | 0.0 | 0.00 | 0.00 |
| withdrawals\_3 | Not\_Churn | 1,655.24 | 23,409,138.46 | 4,838.30 | 220,000.00 | 0.00 | 0.00 | 220,000.00 | 0.0 | 0.00 | 0.00 |
| withdrawals\_6 | Churn | 0.54 | 947.81 | 30.79 | 2,000.00 | 0.00 | 0.00 | 2,000.00 | 0.0 | 0.00 | 0.00 |
| withdrawals\_6 | Not\_Churn | 2,309.73 | 30,861,891.81 | 5,555.35 | 220,000.00 | 0.00 | 0.00 | 220,000.00 | 0.0 | 0.00 | 1,120.00 |

##### Categorical Vs Target

##### Numerical Vs Target

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

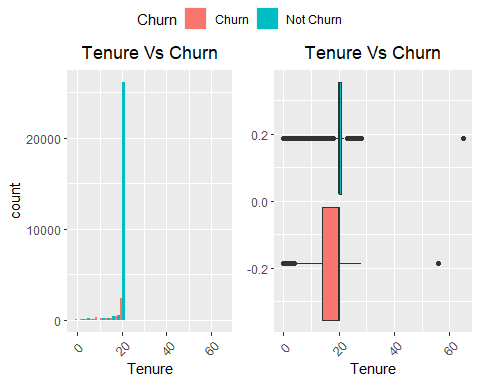
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

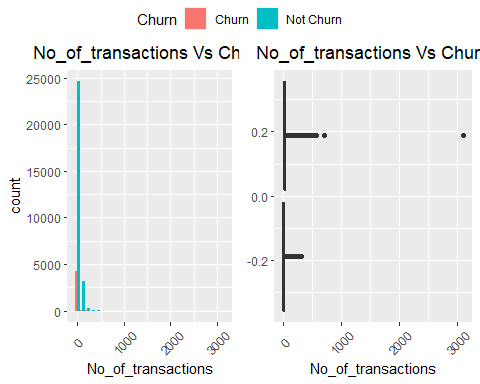
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).  
  
## Warning: Removed 4315 rows containing non-finite values (stat\_bin).

## Warning: Removed 4315 rows containing non-finite values (stat\_boxplot).

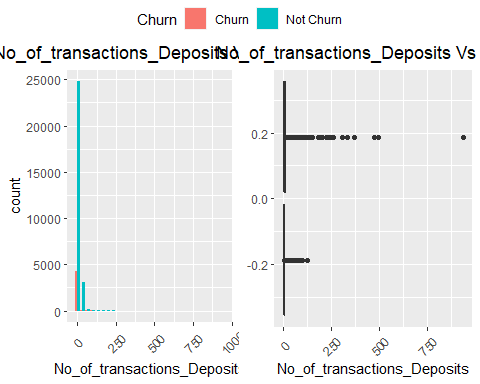
## [[1]]



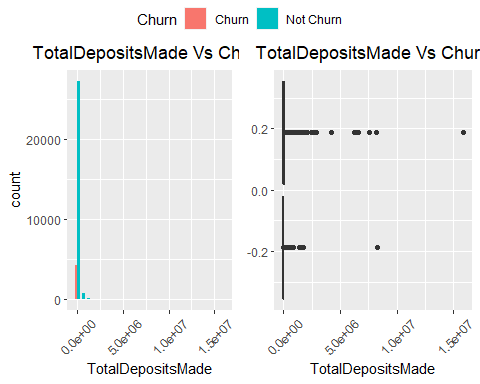
##   
## [[2]]



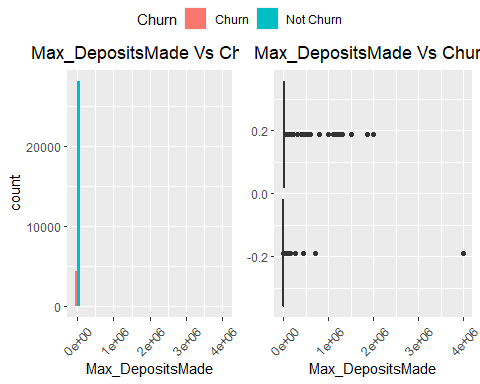
##   
## [[3]]



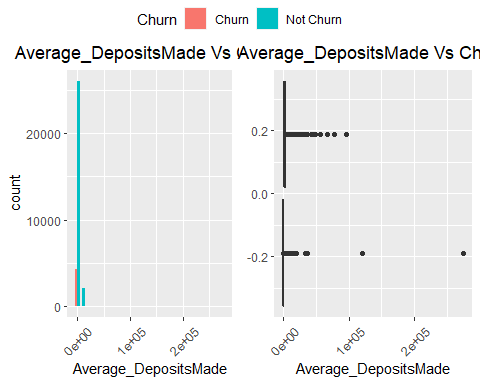
##   
## [[4]]



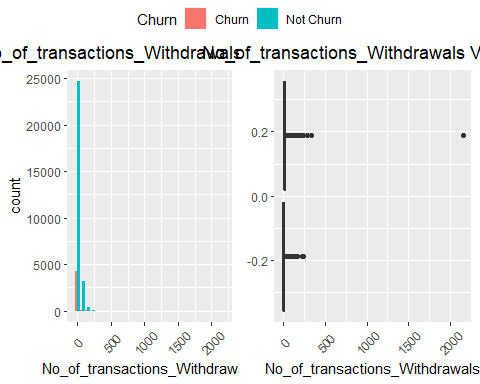
##   
## [[5]]



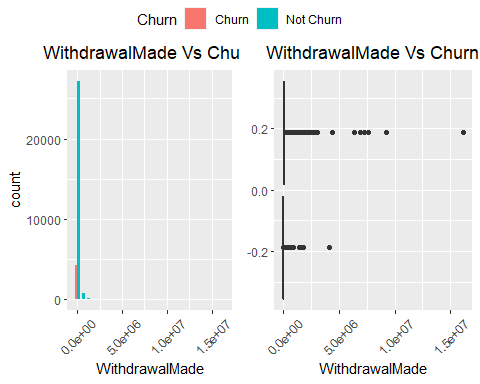
##   
## [[6]]



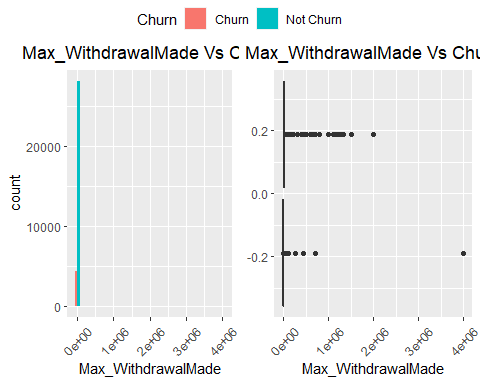
##   
## [[7]]



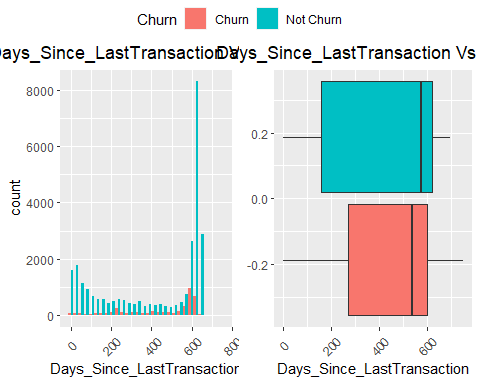
##   
## [[8]]



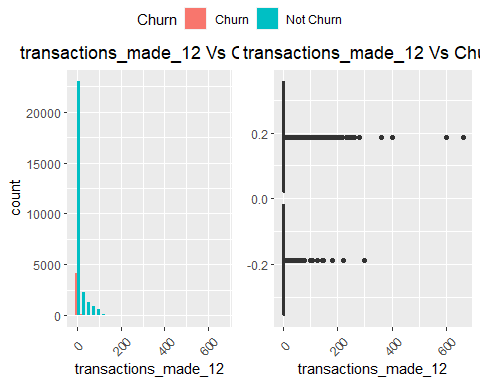
##   
## [[9]]



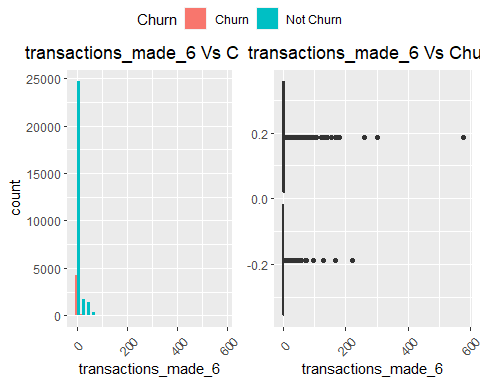
##   
## [[10]]



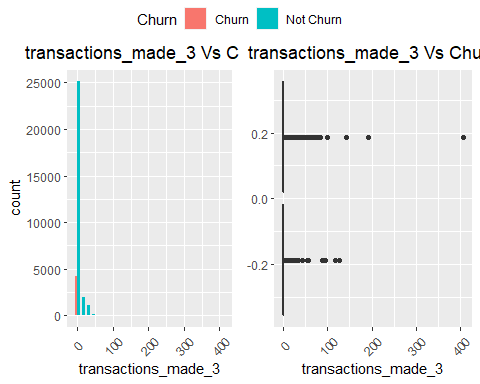
##   
## [[11]]



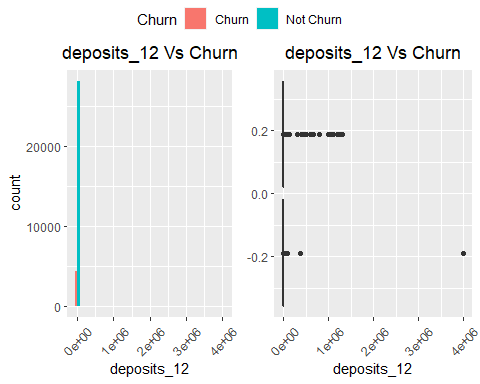
##   
## [[12]]



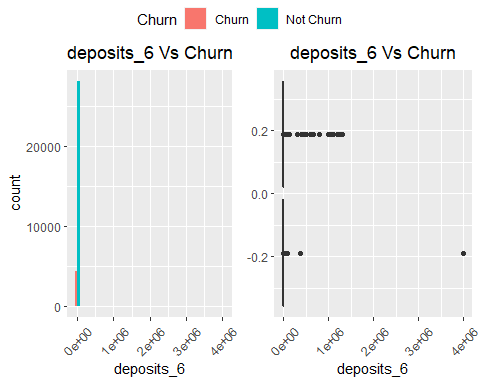
##   
## [[13]]



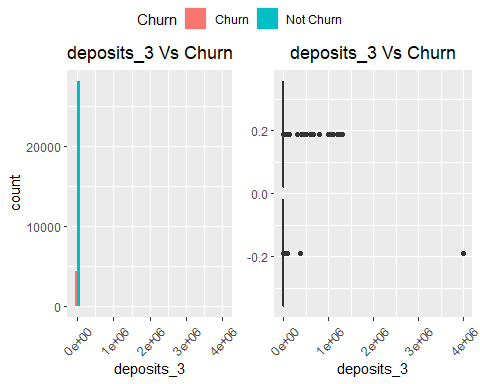
##   
## [[14]]



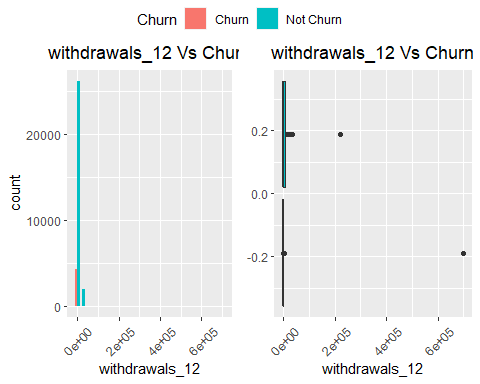
##   
## [[15]]



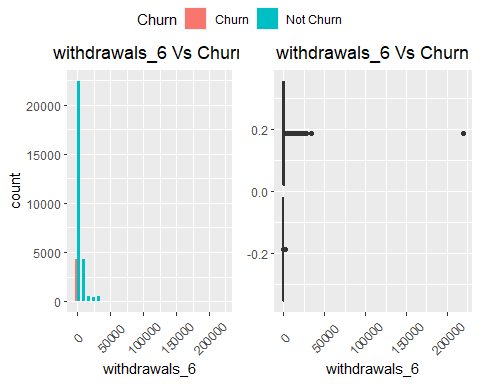
##   
## [[16]]



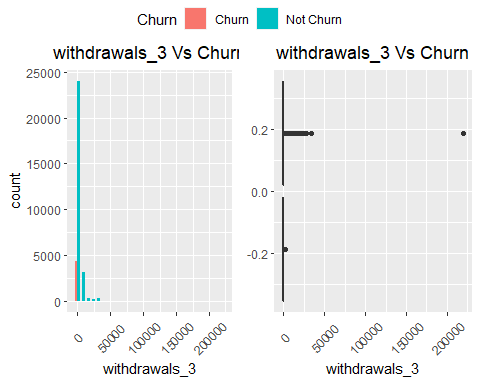
##   
## [[17]]



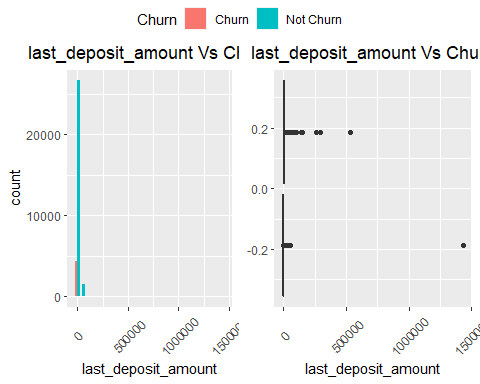
##   
## [[18]]



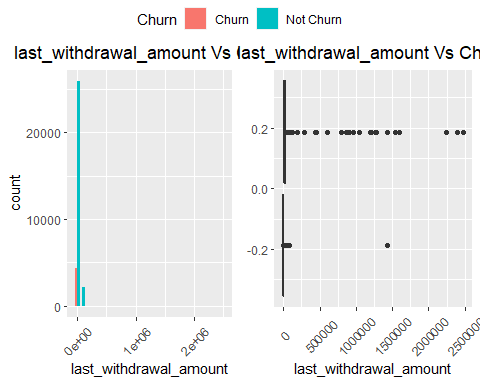
##   
## [[19]]



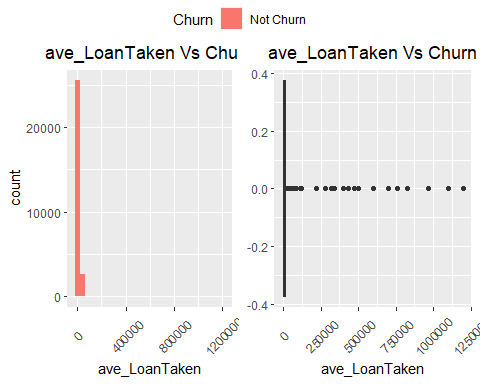
##   
## [[20]]



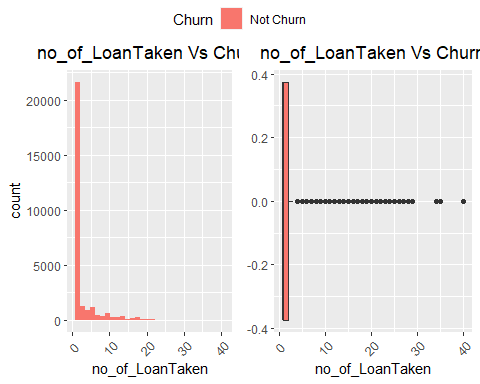
##   
## [[21]]



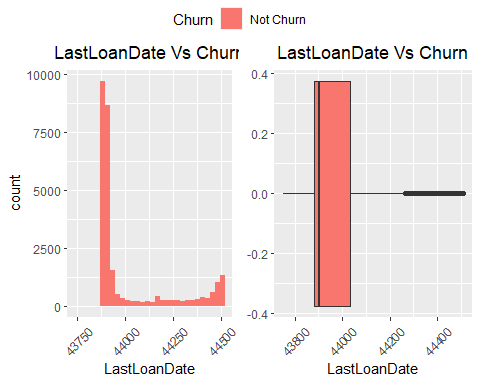
##   
## [[22]]



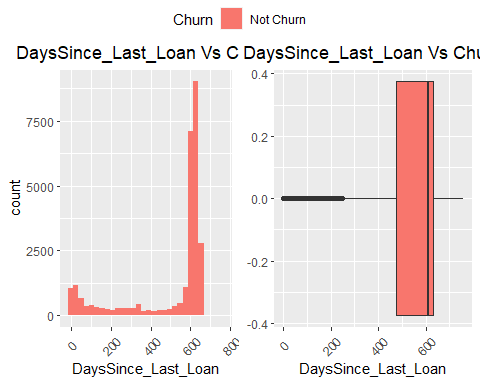
##   
## [[23]]



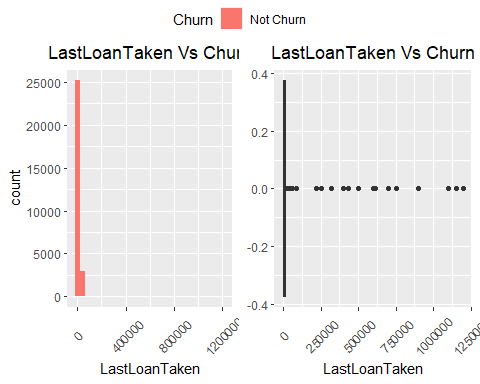
##   
## [[24]]



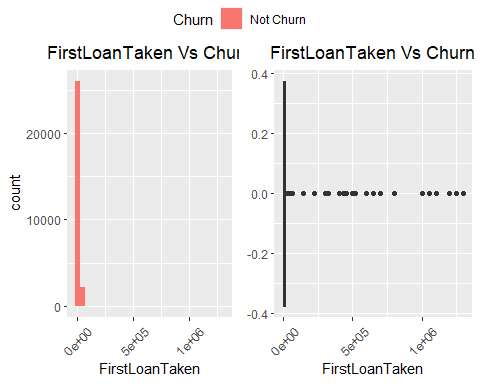
##   
## [[25]]



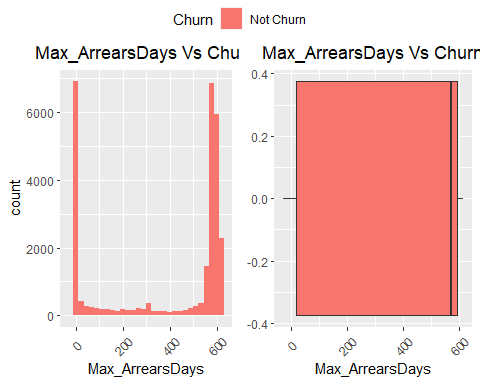
##   
## [[26]]



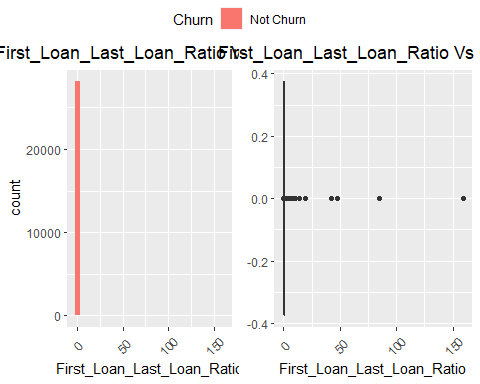
##   
## [[27]]



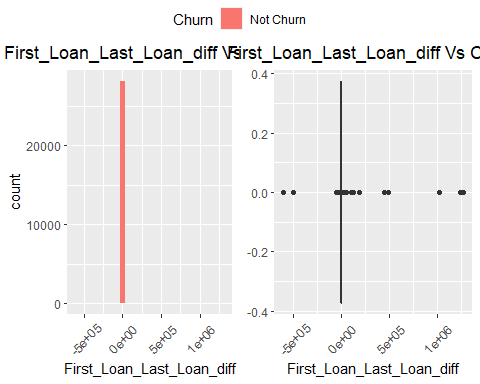
##   
## [[28]]



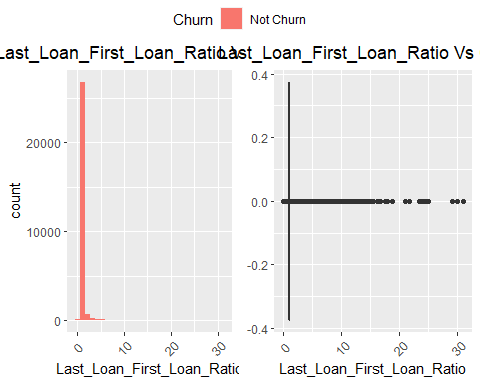
##   
## [[29]]



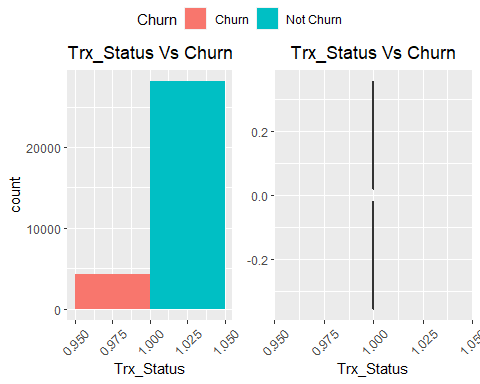
##   
## [[30]]



##   
## [[31]]

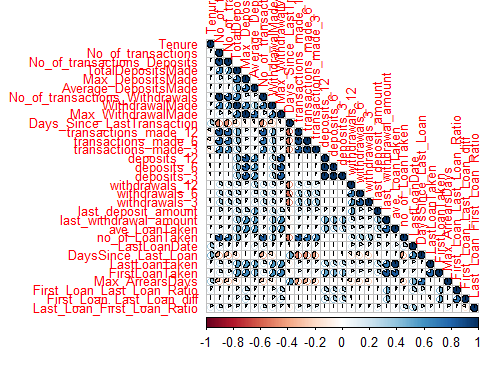


##   
## [[32]]



##### Variable Correlation

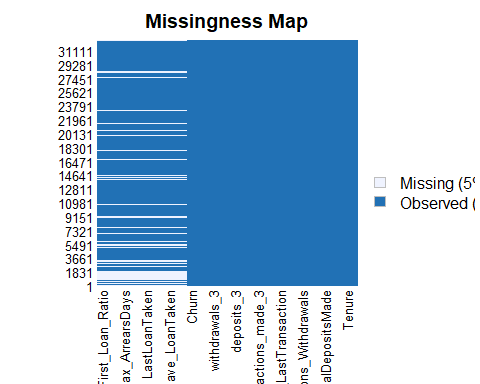
## corrplot 0.84 loaded



## Data Pre processing

### Drop Unncecessary Columns

### Dealing with Missing Values



There are so many ways to deal with missing values.For instance in this dataset most variables are missing because of no transaction history.In this case they will be imputed with value zero.

From the above chart,its evident that Loan Related variables are more likely to have a missing value than any other transaction related variable.Such cases would only imply that the said clients do not have any loan transaction history with the financial institution.

### Dealing with Skewness and abnormality

As observed from above,the variables were heavily skewed,to correct this ,min max scaling was used.

### Train test Split

### Upsampling Minority Class

## Modeling

### Logistic regression

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: 'tidy.numeric' is deprecated.  
## See help("Deprecated")

## Warning: `data\_frame()` was deprecated in tibble 1.1.0.  
## Please use `tibble()` instead.

## # A tibble: 34 x 2  
## names coeff  
## <chr> <dbl>  
## 1 (Intercept) -26.0   
## 2 NumberOfProducts 2.14   
## 3 Tenure -0.340  
## 4 No\_of\_transactions 118.   
## 5 No\_of\_transactions\_Deposits -48.3   
## 6 TotalDepositsMade 435.   
## 7 Max\_DepositsMade -61.8   
## 8 Average\_DepositsMade 0.987  
## 9 No\_of\_transactions\_Withdrawals -98.8   
## 10 WithdrawalMade -547.   
## # ... with 24 more rows

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction Churn Not Churn  
## Churn 863 1  
## Not Churn 0 5637  
##   
## Accuracy : 0.9998   
## 95% CI : (0.9991, 1)  
## No Information Rate : 0.8673   
## P-Value [Acc > NIR] : <2e-16   
##   
## Kappa : 0.9993   
##   
## Mcnemar's Test P-Value : 1   
##   
## Sensitivity : 1.0000   
## Specificity : 0.9998   
## Pos Pred Value : 0.9988   
## Neg Pred Value : 1.0000   
## Prevalence : 0.1327   
## Detection Rate : 0.1327   
## Detection Prevalence : 0.1329   
## Balanced Accuracy : 0.9999   
##   
## 'Positive' Class : Churn   
##

### Decision Trees

#### Model definition

##   
## Classification tree:  
## tree(formula = Churn ~ ., data = trainsmote)  
## Variables actually used in tree construction:  
## [1] "Max\_LoanTaken"  
## Number of terminal nodes: 2   
## Residual mean deviance: 0 = 0 / 24160   
## Misclassification error rate: 0 = 0 / 24164

#### Model evaluation

### Random Forest

## [1] 1

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction Churn Not Churn  
## Churn 863 0  
## Not Churn 0 5638  
##   
## Accuracy : 1   
## 95% CI : (0.9994, 1)  
## No Information Rate : 0.8673   
## P-Value [Acc > NIR] : < 2.2e-16   
##   
## Kappa : 1   
##   
## Mcnemar's Test P-Value : NA   
##   
## Sensitivity : 1.0000   
## Specificity : 1.0000   
## Pos Pred Value : 1.0000   
## Neg Pred Value : 1.0000   
## Prevalence : 0.1327   
## Detection Rate : 0.1327   
## Detection Prevalence : 0.1327   
## Balanced Accuracy : 1.0000   
##   
## 'Positive' Class : Churn   
##