

EDA- CREDIT CARD CHURN DATASET

Nitish Baradwaj, Nuo Chen, Yueru Zhang, Jianing Liu

ISE 535 Fall 2022 Group-3

University of Southern California

Analysis Plan

- Data cleaning & Pre-processing & Possible Attributions Removal
- EDA (Univariate and Bivariate analysis)
- Use statistical analysis (Chi-Sq test) on the factors that may correlate with the attrition flag of customers.
- Compare two classification models to find the best fit one
 - Logistic Regression
 - Decision Tree
 - Could also use unsupervised learning
- Find the factors related to the reasons customers leaving the credit card service
- Conclusion

Credit Card Customer Churn Analysis

Overview

Business Objectives

- Credit cards are a good source of income for banks, so the bank wants to analyze the data and identify who will leave or stay on their credit card services and reason for leaving – so that bank could improve upon those areas

Data Source

- A dataset of bank credit card customer churn depend on different customers attribution

Credit Card Customer Churn Analysis

Data Overview

Data Source

- A dataset of bank credit card customer churn depend on different customers attribution
- Data consisting of 21 attributes and 10,127 observations where each observation represents single customer information.

Credit Card Customer Churn Analysis

Data Dictionary: Category

Variable	Description
Attrition_Flag	Customer activity variable
Gender	Gender of the customer
Education_level	Customer Educational Qualification: Doctorate, Post-graduate, Graduate, College, High School, Unknown, Uneducated
Marital_Status	Marital Status of account holders
Income_Category	Annual Income Category in customer
Card_Category	Type of Card

Credit Card Customer Churn Analysis

Data Dictionary: Measures

Variable	Description
CLIENTNUM	Client number
Customer_Age	Customer Age in Years
Months_on_book	Period of relationship with the bank
Total_Relationship_Count	Total count of products held by the customer
Months_Inactive_12_mon	Numbers of months inactive in the last 12 months
Contacts_Count_12_mon	Numbers of Contacts between the customer and bank in the last 12 months
Credit_Limit	Credit Limit on the Credit Card
Total_Revolving_Bal	The balance that carries over from one month to the next is the revolving balance
Avg_Open_To_Buy	Open to Buy refers to the amount left on the credit card to use
Total_Amt_Chng_Q4_Q1	Ratio of the total transaction amount in 4th quarter and the total transaction amount in 1st quarter
Total_Trans_Amt	Total Transaction Amount
Total_Trans_Ct	Total Transaction Count
Total_Ct_Chng_Q4_Q1	Ratio of the total transaction count in 4th quarter and the total transaction count in 1st quarter
Avg_Utilization_Ratio	Represents how much of the available credit the customer spent
Dependent_count	Number of dependents

Initial Data Review

Numeric Variables

Attribute	Missing Values	Unique Values	Mean	Min	Max	SD
CLIENTNUM	0	10127	7.391776e+08	708082083.0	8.283431e+08	3.690378e+07
Customer_Age	0	45	4.632596e+01	26.0	7.300000e+01	8.016814e+00
Dependent_count	0	6	2.346203e+00	0.0	5.000000e+00	1.298908e+00
Months_on_book	0	44	3.592841e+01	13.0	5.600000e+01	7.986416e+00
Total_Relationship_Count	0	6	3.812580e+00	1.0	6.000000e+00	1.554408e+00
Months_Inactive_12_mon	0	7	2.341167e+00	0.0	6.000000e+00	1.010622e+00
Contacts_Count_12_mon	0	7	2.455317e+00	0.0	6.000000e+00	1.106225e+00
Credit_Limit	0	6205	8.631954e+03	1438.3	3.451600e+04	9.088777e+03
Total_Revolving_Bal	0	1974	1.162814e+03	0.0	2.517000e+03	8.149873e+02
Avg_Open_To_Buy	0	6813	7.469140e+03	3.0	3.451600e+04	9.090685e+03

Attribute	Missing Values	Unique Values	Mean	Min	Max	SD
Total_Amt_Chng_Q4_Q1	0	1158	7.599407e-01	0.0	3.397000e+00	2.192068e-01
Total_Trans_Amt	0	5033	4.404086e+03	510.0	1.848400e+04	3.397129e+03
Total_Trans_Ct	0	126	6.485869e+01	10.0	1.390000e+02	2.347257e+01
Total_Ct_Chng_Q4_Q1	0	830	7.122224e-01	0.0	3.714000e+00	2.380861e-01
Avg_Utilization_Ratio	0	964	2.748936e-01	0.0	9.990000e-01	2.756915e-01

Credit limit and Average open to buy are left skewed distributed with majority of customers having a value of <10,000

Initial Data Review

Character Variables

Attribute <chr>	Missing Values <int>	Unique Values <int>
Attrition_Flag	0	2
Gender	0	2
Education_Level	0	7
Marital_Status	0	4
Income_Category	0	6
Card_Category	0	4

Initial Data Review

Factor Variables

Attribute `<chr>`	Missing Values `<int>`	Unique Values `<int>`	Mode `<chr>`
Attrition_Flag	0	2	character
Gender	0	2	character
Education_Level	0	7	character
Marital_Status	0	4	character
Income_Category	0	6	character
Card_Category	0	4	character

```
credit_card_cust$Attrition_Flag = as_factor(credit_card_cust$Attrition_Flag)
credit_card_cust$Gender = as_factor(credit_card_cust$Gender)
credit_card_cust$Education_Level = as_factor(credit_card_cust$Education_Level)
credit_card_cust$Marital_Status = as_factor(credit_card_cust$Marital_Status)
credit_card_cust$Income_Category = as_factor(credit_card_cust$Income_Category)
credit_card_cust$Card_Category = as_factor(credit_card_cust$Card_Category)

summarize_factor(credit_card_cust)
```

Initial Data Review

Numeric Variables

Attribute `<chr>`	Missing Values `<int>`	Unique Values `<int>`	Mean `<dbl>`	Min `<dbl>`	Max `<dbl>`	SD `<dbl>`
Customer_Age	0	45	46.3259603	26.0	73.000	8.0168140
Dependent_count	0	6	2.3462032	0.0	5.000	1.2989083
Months_on_book	0	44	35.9284092	13.0	56.000	7.9864163
Total_Relationship_Count	0	6	3.8125802	1.0	6.000	1.5544079
Months_Inactive_12_mon	0	7	2.3411672	0.0	6.000	1.0106224
Contacts_Count_12_mon	0	7	2.4553175	0.0	6.000	1.1062251
Credit_Limit	0	6205	8631.9536980	1438.3	34516.000	9088.7766502
Total_Revolving_Bal	0	1974	1162.8140614	0.0	2517.000	814.9873352
Avg_Open_To_Buy	0	6813	7469.1396366	3.0	34516.000	9090.6853237
Total_Amt_Chng_Q4_Q1	0	1158	0.7599407	0.0	3.397	0.2192068

Attribute `<chr>`	Missing Values `<int>`	Unique Values `<int>`	Mean `<dbl>`	Min `<dbl>`	Max `<dbl>`	SD `<dbl>`
Total_Trans_Amt	0	5033	4404.0863039	510.0	18484.000	3397.1292536
Total_Trans_Ct	0	126	64.8586946	10.0	139.000	23.4725704
Total_Ct_Chng_Q4_Q1	0	830	0.7122224	0.0	3.714	0.2380861
Avg_Utilization_Ratio	0	964	0.2748936	0.0	0.999	0.2756915

Delete LIENTNUM

Logical Groupings of Variables

Category (6)

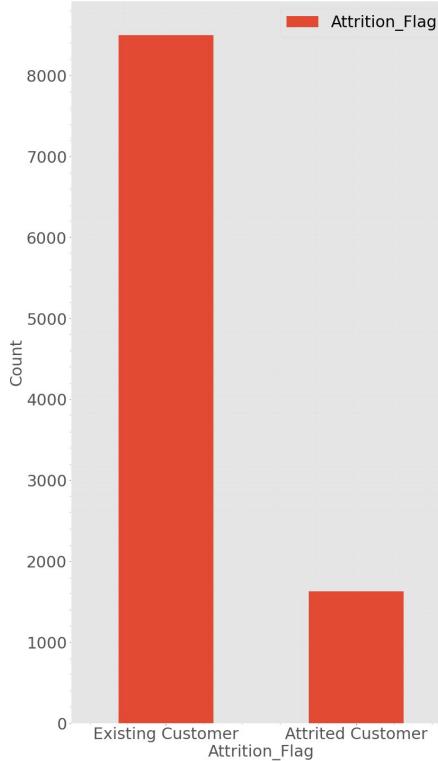
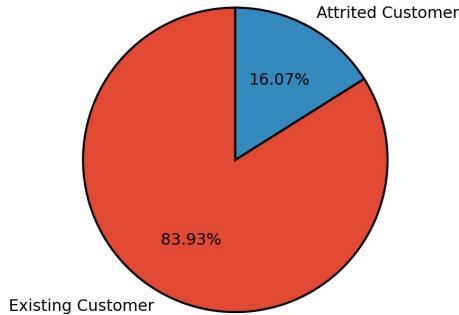
- **Outcome**
 - Attrition_Flag
- **Customer information**
 - Gender, Education_Level, Marital_Status, Income_Category
- **Card information**
 - Card_Category

Measures (14)

- **Customer information**
 - Customer_Age, Dependent_count
- **The relationship between bank and customer**
 - Months_on_book, Total_Relationship_Count, Months_Inactive_12_mon, Contacts_Count_12_mon
- **The credit card information**
 - Credit_Limit, Total_Revolving_Bal, Avg_Open_To_Buy, Total_Amt_Chng_Q4_Q1, Total_Ct_Chng_Q4_Q1, Total_Trans_Amt, Total_Trans_Ct, Avg_Utilization_Ratio

Univariate Analysis

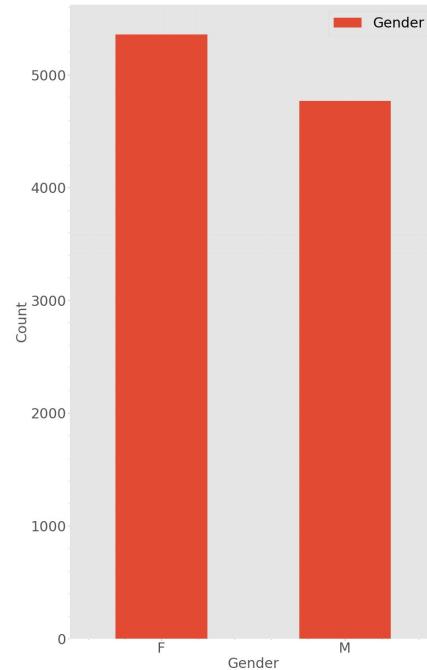
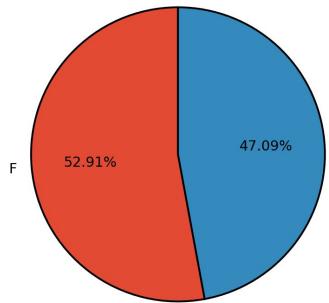
Target Outcome Variable



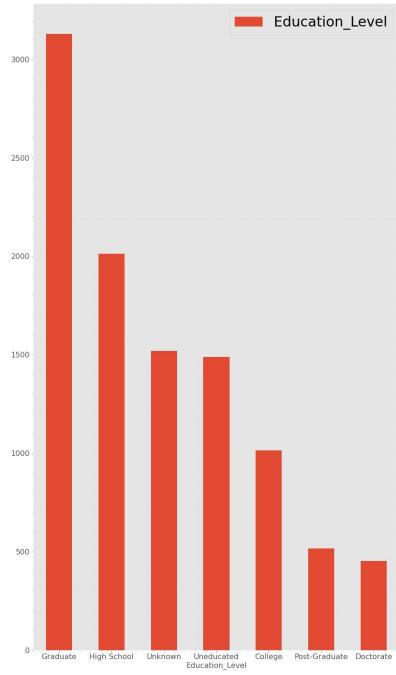
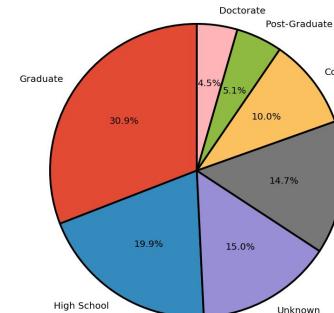
Attrited Customer: About 84% of the customers are existing (active) customers. High Imbalance in data.

Univariate Analysis

Low Cardinality Category



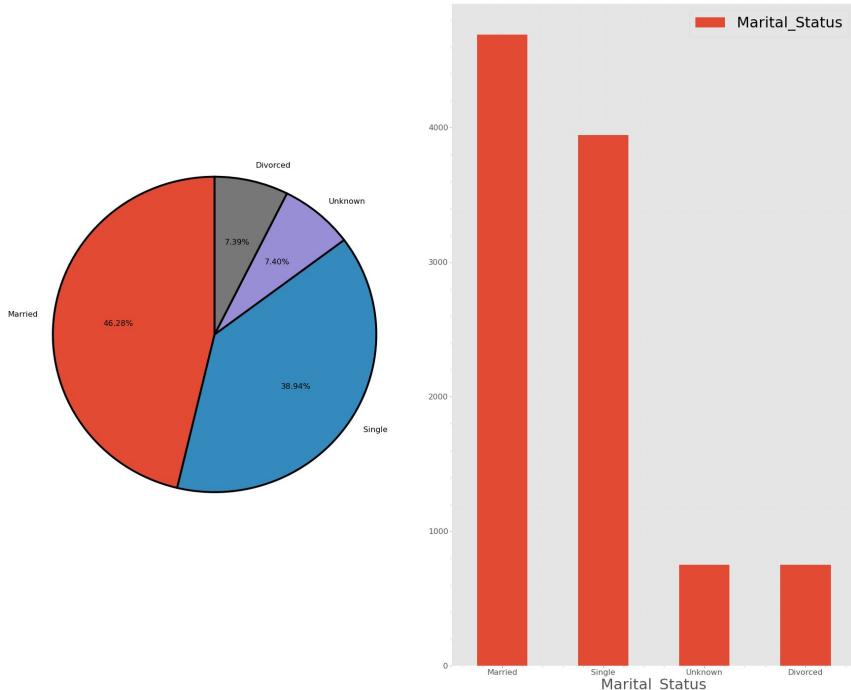
Gender: 53% of the customers are female



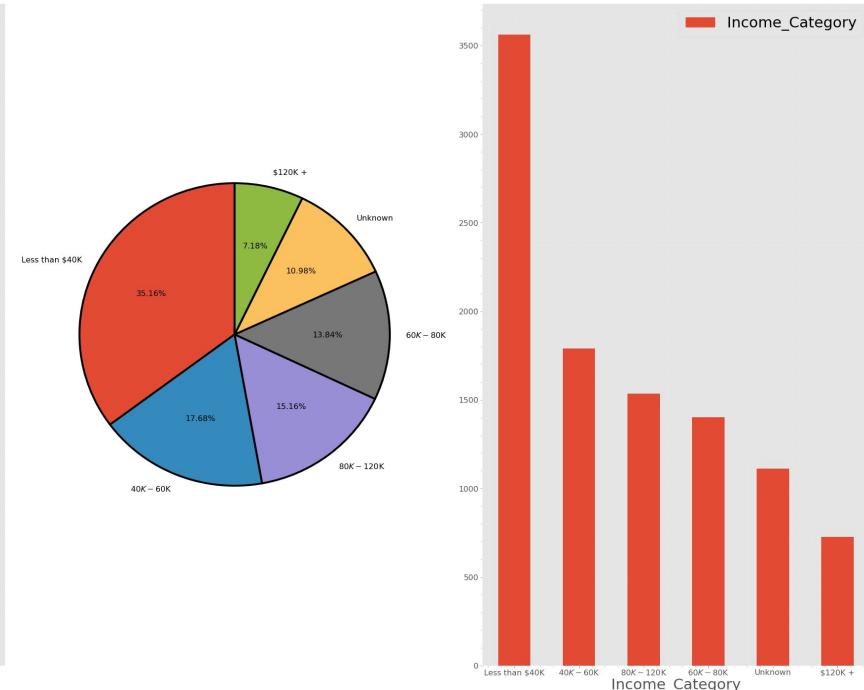
Education level: 31% customers are graduate, around 20% customers have high school diploma; only few customers have doctorate and post doctorate degree

Univariate Analysis

Low Cardinality Category



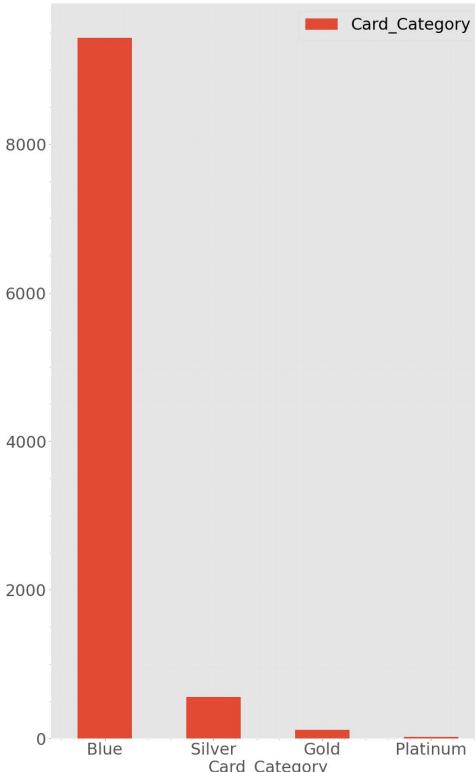
Marital status: Married and Single account for 84% of the total



Income: 35% customers earn less than \$40K; very few customers (7.18%) have extremely high income that over \$120K

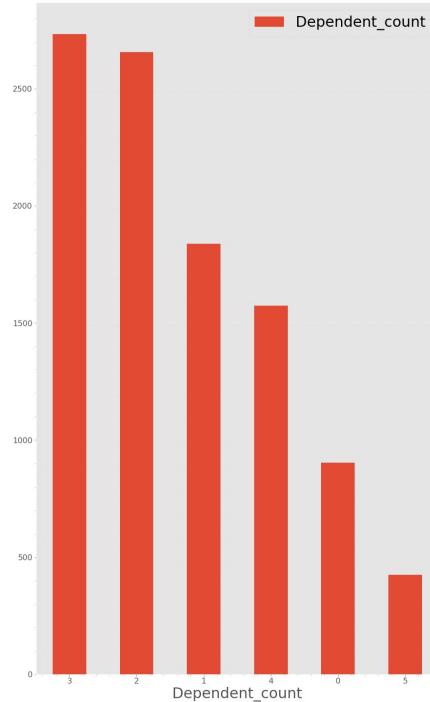
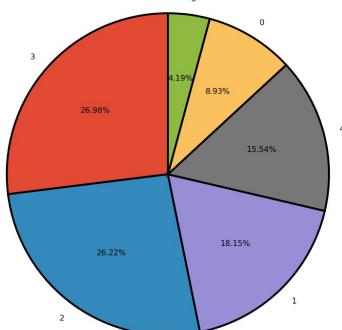
Univariate Analysis

Low Cardinality Category

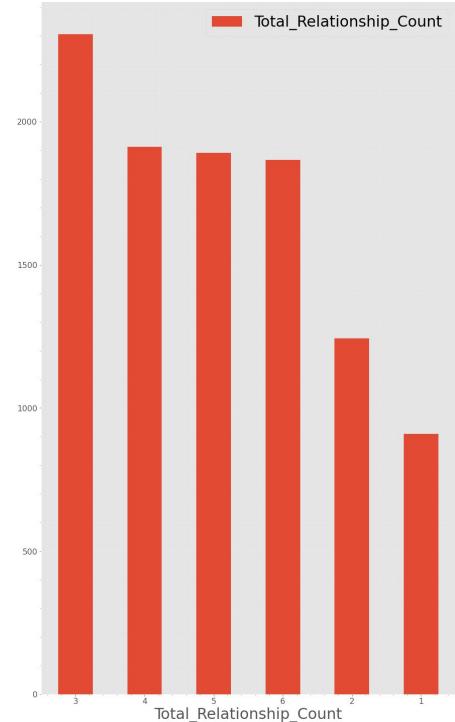
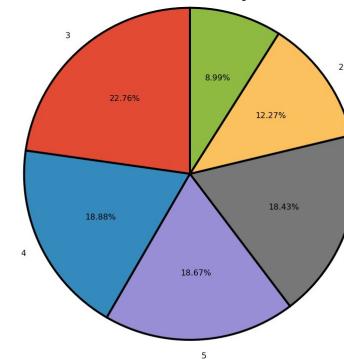


Card Category: 90% customers own Blue or Silver Card. **We need to perform bivariate analysis to see if this is related to the income of customers.**

Univariate Analysis - Measures

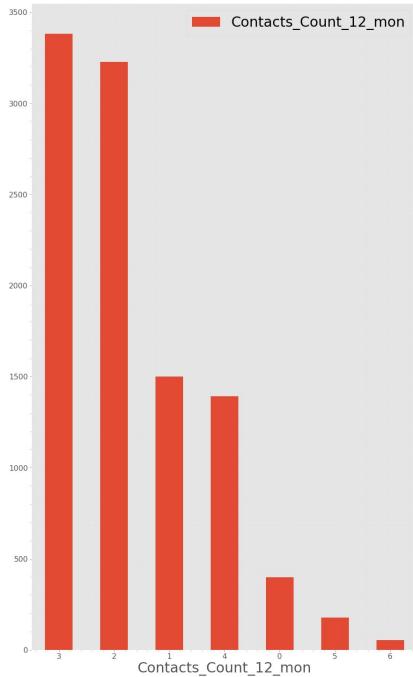


Dependent Count: Mostly 2 or 3

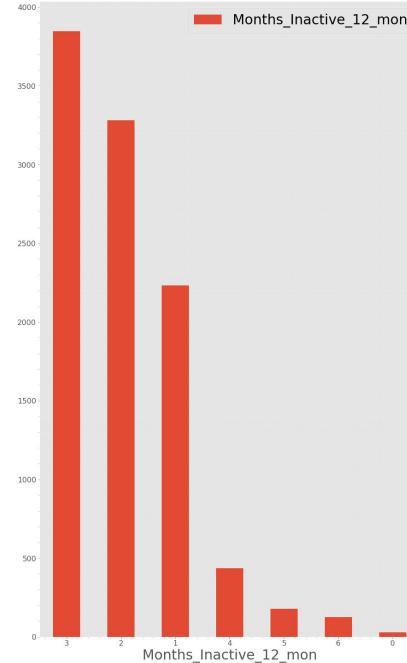


Total Relationship Count: 78.74% customers have 3+ relations with the bank

Univariate Analysis - Measures

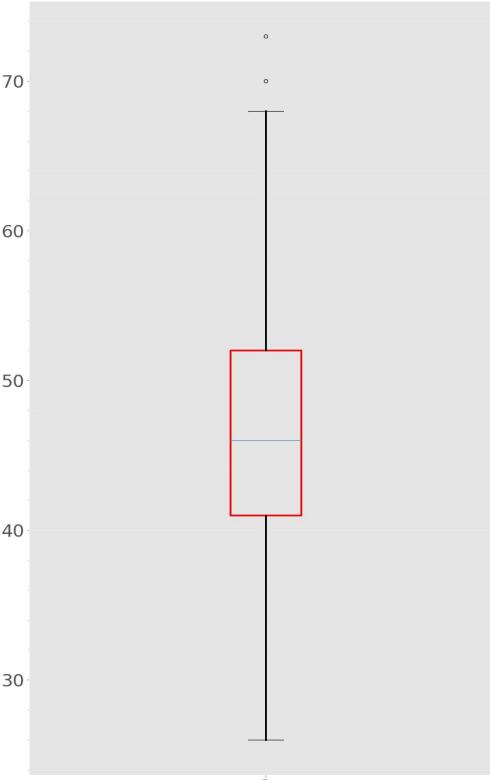
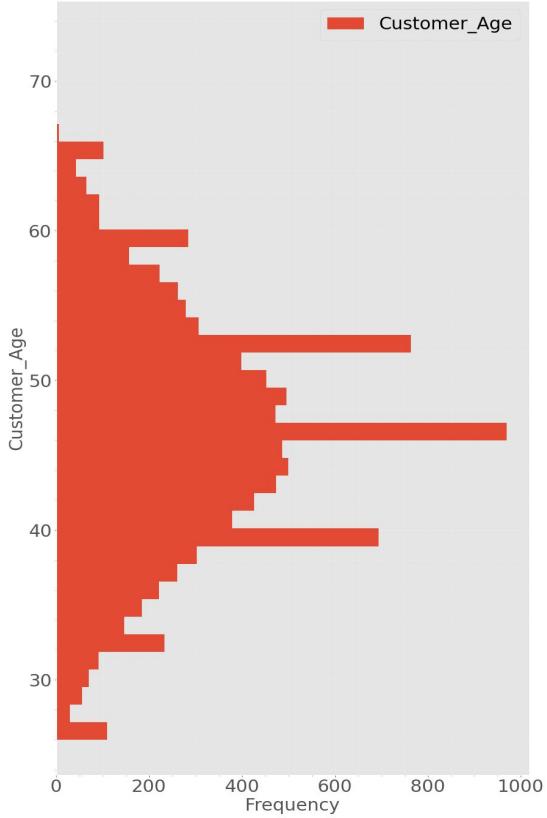


Contacts Count: Most customers have regular contacts with the bank in the last 12 months



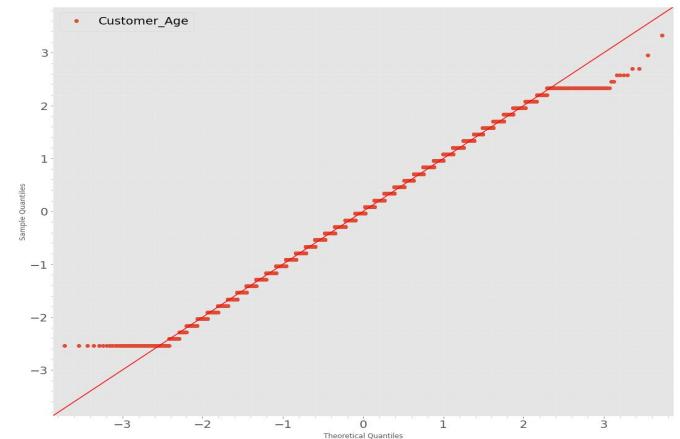
Month Inactive: Most customers inactive for 3 months in the last 12 months, **we need to check if this variable is related to customer status**

Univariate Analysis - Measures

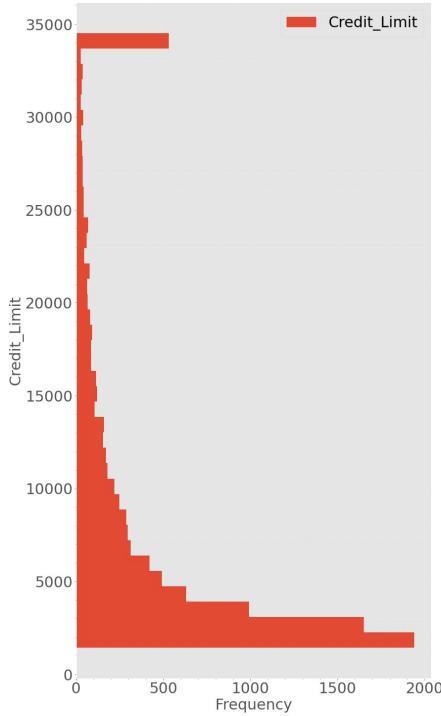


Customer Age:

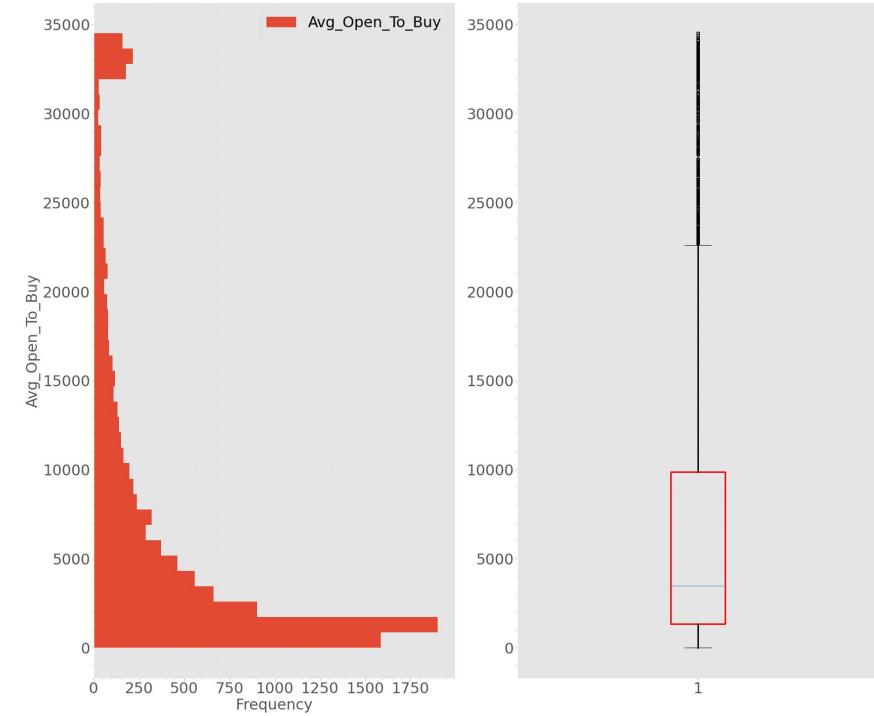
- Data is normally distributed.
- The average age of customers is around 46, most customers are younger than 50.
- Two outliers at the upper bound side.



Univariate Analysis - Measures

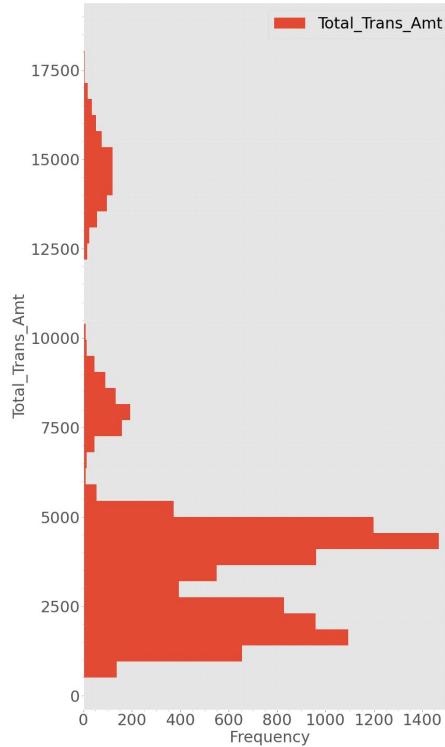


Credit Limit: Data is right skewed. Most customers have less than 5,000 credit limit. Plus, there are many outliers beyond upper bound, **we need to do further analysis to see the relationship between credit limit and income.**

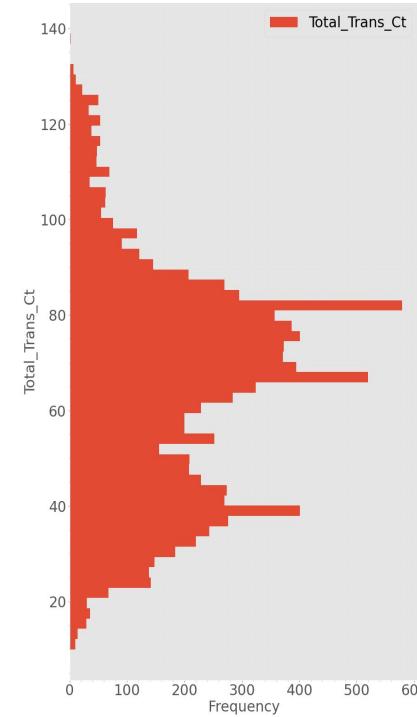
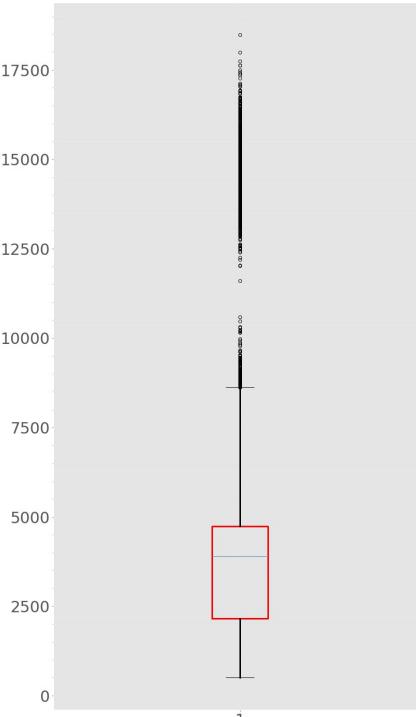


Average Open to Buy: Same trends as Credit_Limit

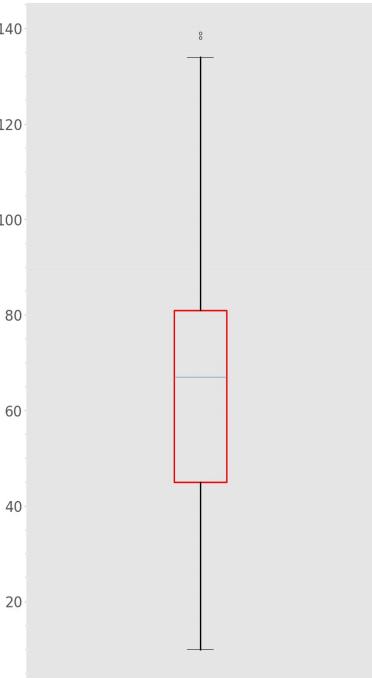
Univariate Analysis - Measures



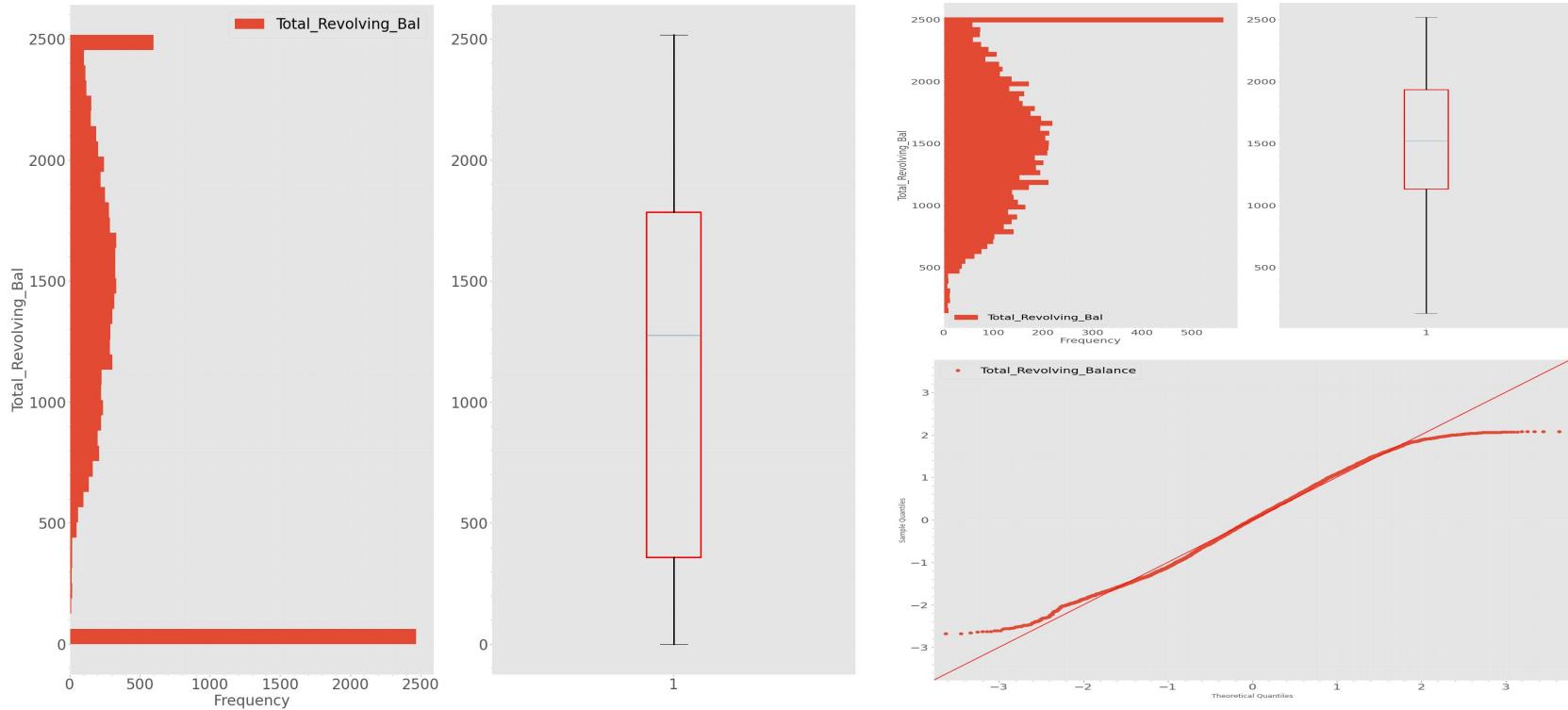
Total Transaction Amount: Most total transaction amount per customer in the last 12 months are under \$5k. Lots of outliers beyond upper bound



Total Transaction Count: From the Frequency distribution histogram, the total transaction count has two parts

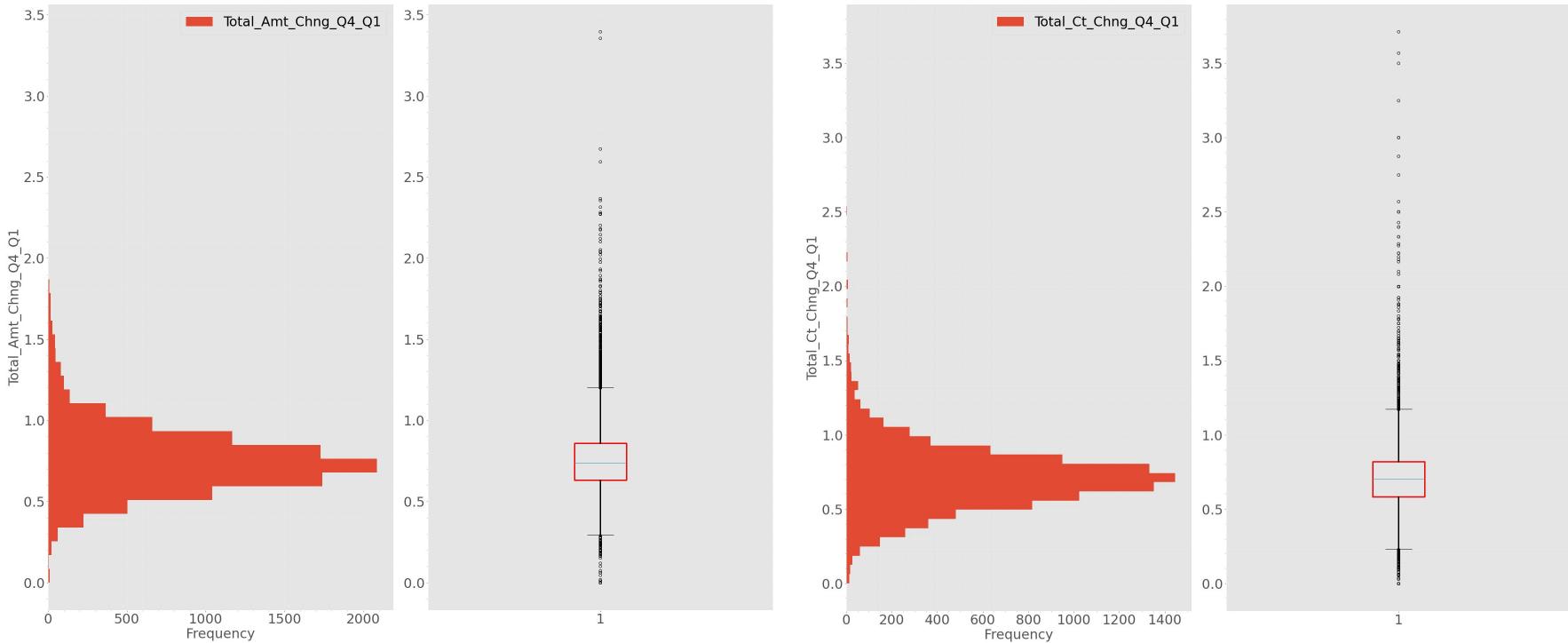


Univariate Analysis - Measures



Total Revolving Balance: Some customers never use credit cards. After removing People who do not carry over balance into next month, the data looks **normally distributed**. Although it has tails at the end as evidenced by the QQ plot.

Univariate Analysis - Measures



Total Transaction Amount & Count in Q4-Q1(Ratio): Most customers did less transactions in Q4 compared to Q1, also spent less total amount of money

Univariate Outlier Analysis

Contacts_Count_12_mon `n()` <i><dbl> <int></i>	Months_Inactive_12_mon `n()` <i><dbl> <int></i>	Customer_Age `n()` <i><dbl> <int></i>	Credit_Limit `n()` <i><dbl> <int></i>
6 54	6 124	73 1	34516 507
5 175	5 178	70 1	34496 1
4 1389	4 434	68 2	34458 1
3 3379	3 3841	67 4	34427 1
2 3226	2 3281	66 2	34198 1
1 1498	1 2233	64 43	34173 1
0 399	0 29	63 65	34162 1
		62 93	34140 1
		61 93	34058 1
		.. with 35 more rows	34010 1
			. with 6,192 more rows

Avg_Open_To_Buy `n()` <i><dbl> <int></i>	Total_Trans_Amt `n()` <i><dbl> <int></i>	Total_Trans_Ct `n()` <i><dbl> <int></i>
34516 98	18484 1	139 1
34362 1	17995 1	138 1
34300 1	17744 1	134 1
34297 1	17634 1	132 1
34286 1	17628 1	131 6
34238 1	17498 1	130 5
34227 1	17437 1	129 6
34140 1	17390 1	128 10
34119 1	17350 1	127 12
34117 1	17258 1	126 10
. with 6,798 more rows	. with 5,022 more rows	. with 116 more rows

- None of these outliers appears out of the range of being possible, so we will leave them in the dataset.

Univariate Outlier Analysis

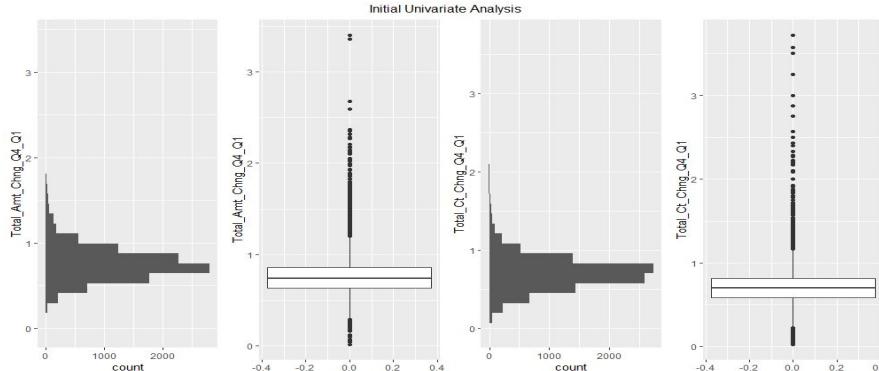
Maximum values

Total_Amt_Chng_Q4_Q1 `n()`	Total_ct_Chng_Q4_Q1 `n()`
<dbl> <int>	<dbl> <int>
3.40 1	3.71 1
3.36 1	3.57 1
2.68 1	3.5 1
2.59 1	3.25 1
2.37 1	3 2
2.36 1	2.88 1
2.32 1	2.75 1
2.28 1	2.57 1
2.28 1	2.5 3
2.27 1	2.43 1
with 1,145 more rows	
with 819 more rows	

Minimum values

Attrition_Flag	Total_Ct_Chng_Q4_Q1	Total_Amt_Chng_Q4_Q1	Credit_Limit	Avg_Open_To_Buy	Card_Category	Total_Trans_Amt	Total_Trans_Ct	Avg_Utilization_Ratio
Attrited Customer	0.000	0.153	13662.0	13662.0	Blue	725	22	0.000
Attrited Customer	0.000	0.000	9904.0	7391.0	Blue	1152	28	0.254
Attrited Customer	0.000	0.000	8258.0	6487.0	Blue	1447	23	0.214
Attrited Customer	0.000	0.000	20974.0	20974.0	Blue	1246	27	0.000
Attrited Customer	0.000	0.010	5590.0	5590.0	Blue	1507	32	0.000
Attrited Customer	0.000	0.000	34516.0	34302.0	Silver	1201	22	0.006
Attrited Customer	0.000	0.000	3233.0	716.0	Blue	1339	32	0.779
Existing Customer	0.028	0.459	3242.0	2443.0	Blue	1424	37	0.246
Attrited Customer	0.029	0.046	1760.0	1760.0	Blue	1554	35	0.000
Attrited Customer	0.038	1.214	4075.0	2089.0	Blue	2449	27	0.487

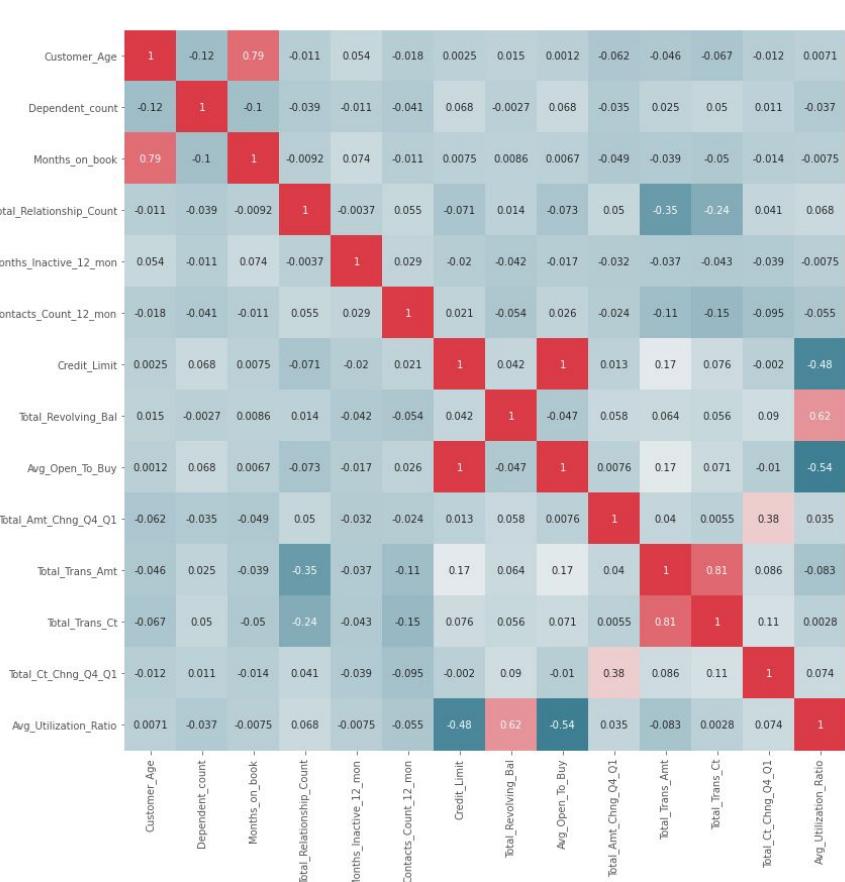
Remove 7 outliers, which in the rows that
 Total_Ct_Chng_Q4_Q1 is 0.00, and
 Attrition_Flag is Attrited_Flag



Univariate Analysis Summary

- Total revolving balance contains lots of 0 and the distribution of both credit limit and average open to buy are right-skewed. Indicates that some customers open an account but seldom use credit card services.
- Total transaction amount has 3 parts, total transaction count has 2 parts, we need to do further research.
- After we removed the outliers in Total Transaction Amount & Count in Q4-Q1(Ratio), there are 10,120 rows left.
- We need to perform bivariate analysis to find inter-dependencies between variables.

Bivariate Analysis

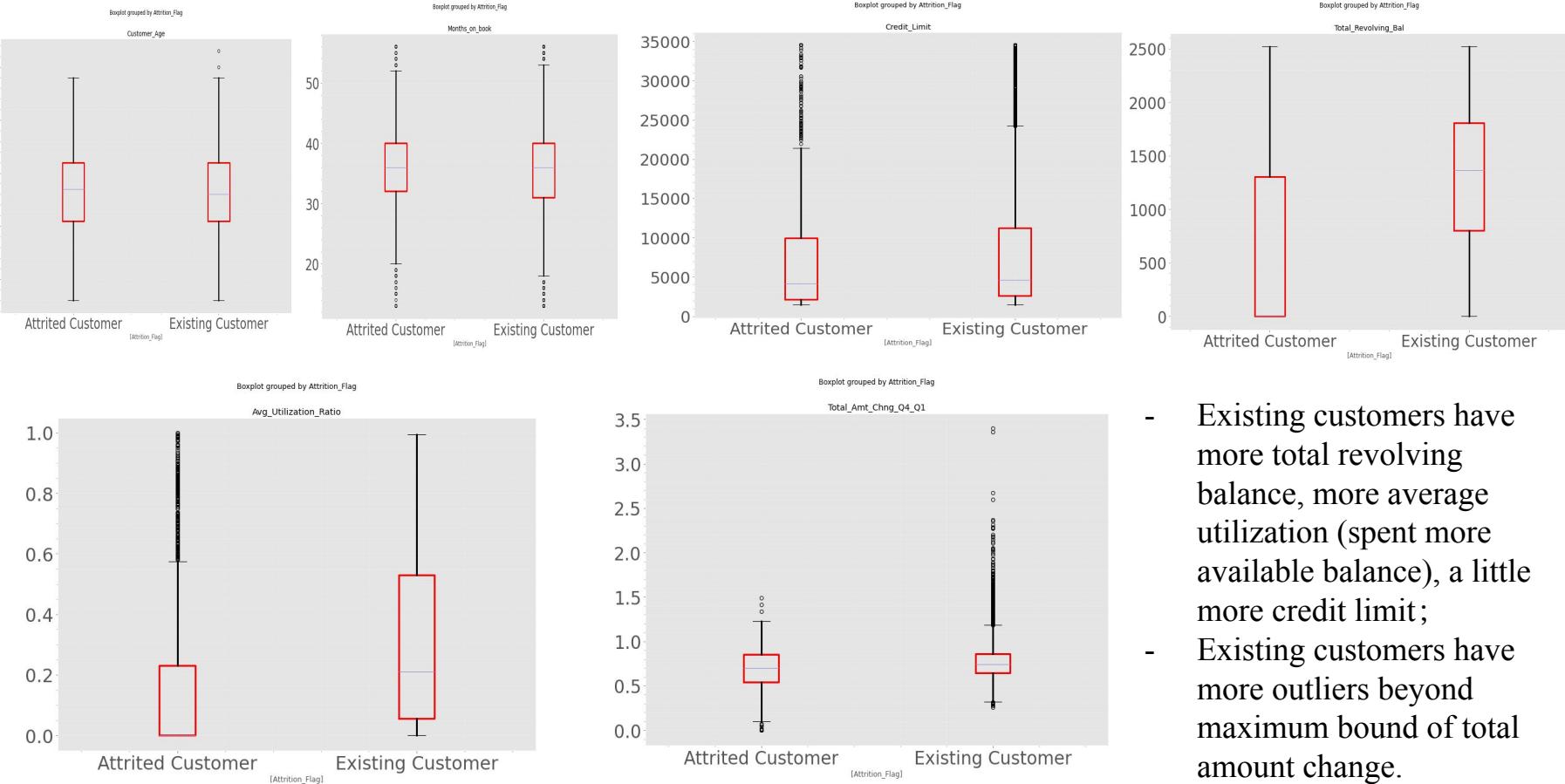


- Avg_Open_To_Buy and Credit_Limit have 100% collinearity
- Months_on_book and Customer_Age, Total_Trans_Ct and Total_Trans_Amt have quite strong correlation
- Total_Revolving_Bal and Avg_Utilization_Ratio also have positive correlation

$$\text{Avg_Utilization_Ratio} = \frac{\text{Total_Revolving_Bal}}{\text{Credit_Limit}}$$

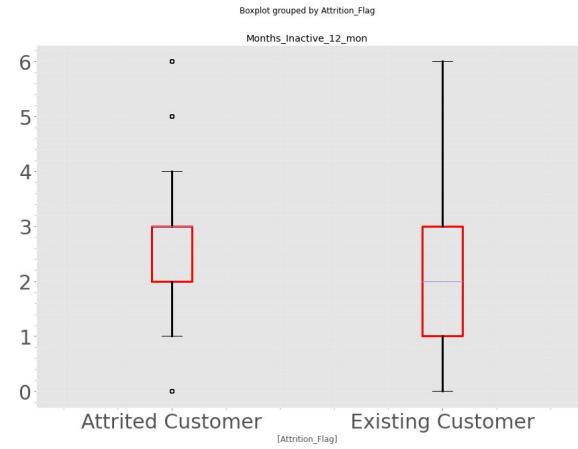
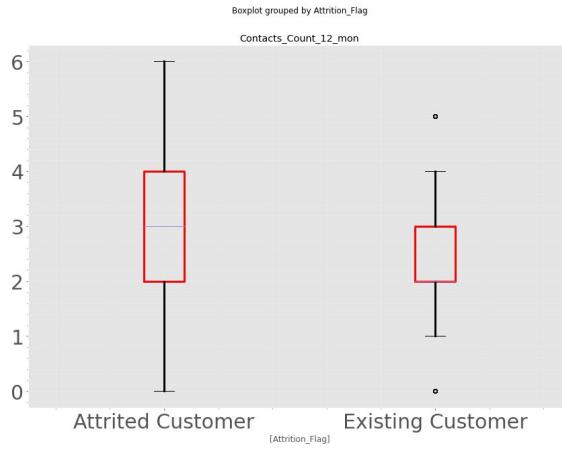
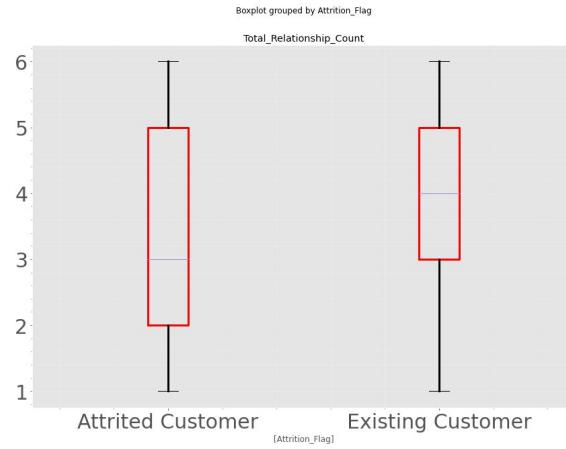
$$\text{Total_Revolving_Bal} = \text{Credit_Limit} - \text{Avg_Open_To_Buy}$$

Bivariate Analysis - Attrition Flag



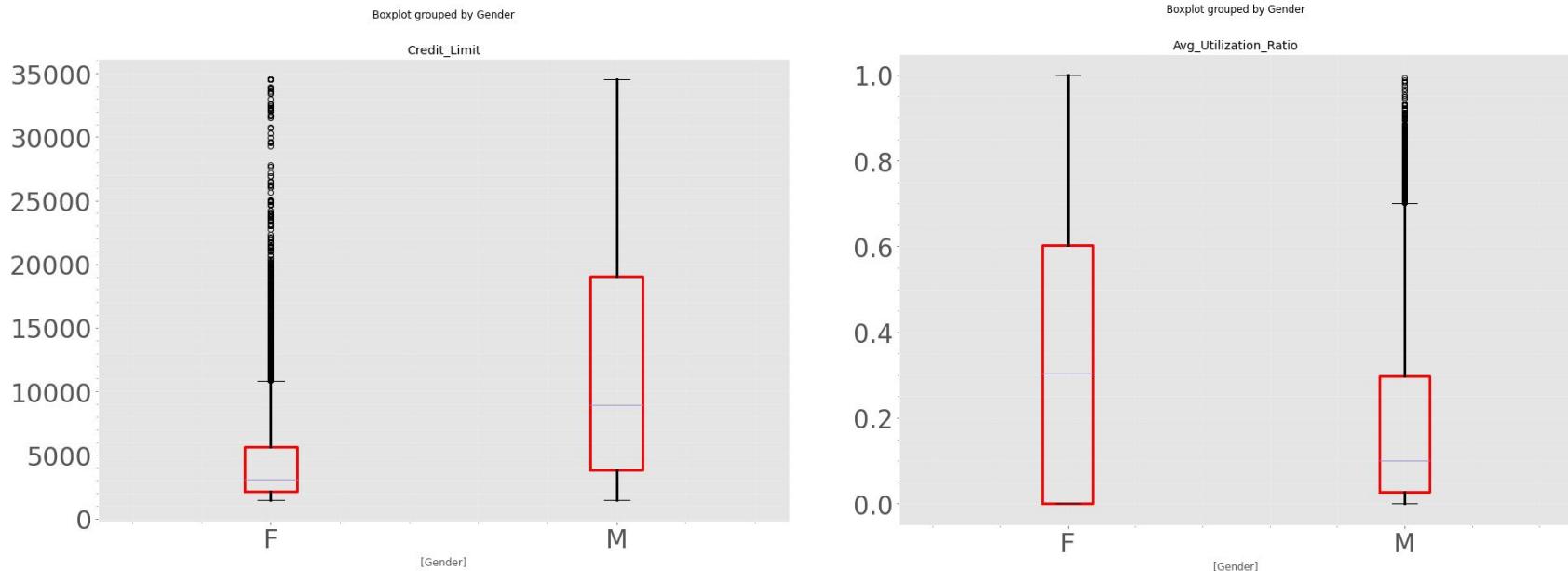
- Existing customers have more total revolving balance, more average utilization (spent more available balance), a little more credit limit;
- Existing customers have more outliers beyond maximum bound of total amount change.

Bivariate Analysis - Attrition Flag



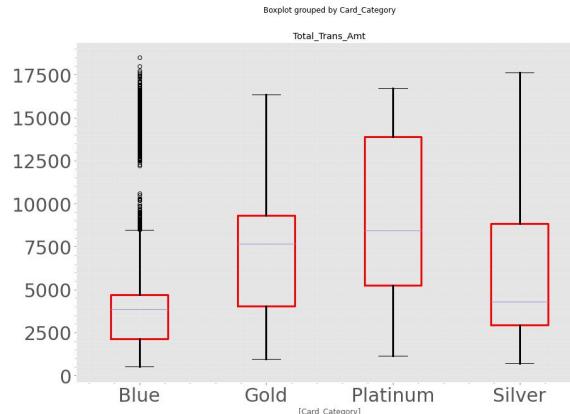
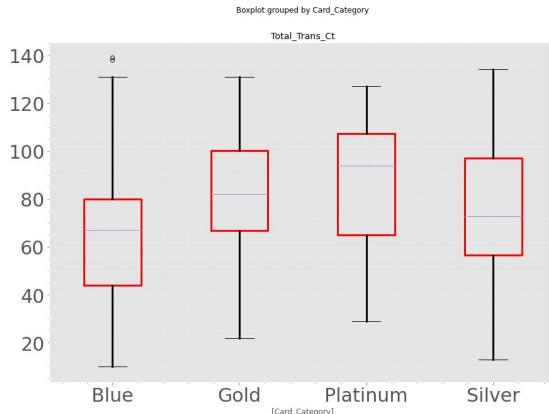
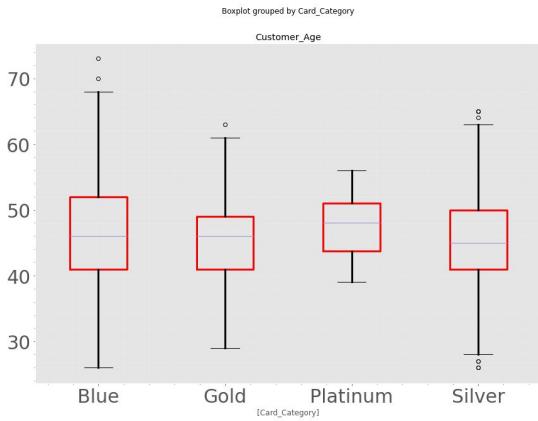
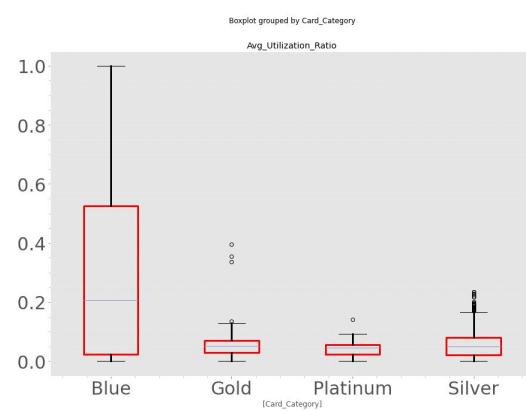
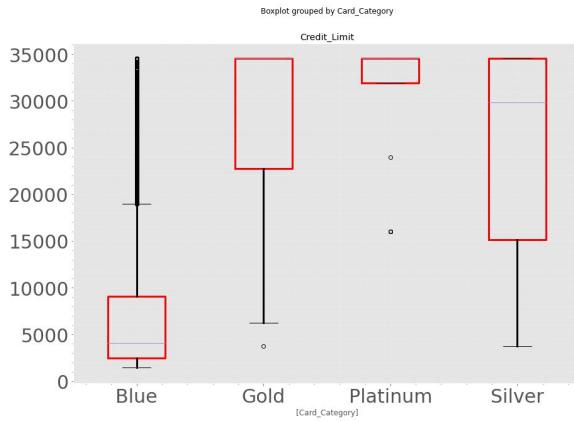
- Existing customers have more relations with the bank, and relatively less contacts with bank and inactivity in the last 12 months.
- ***Existing customers have more total transaction amount and counts.***

Bivariate Analysis - Gender & Marriage Status



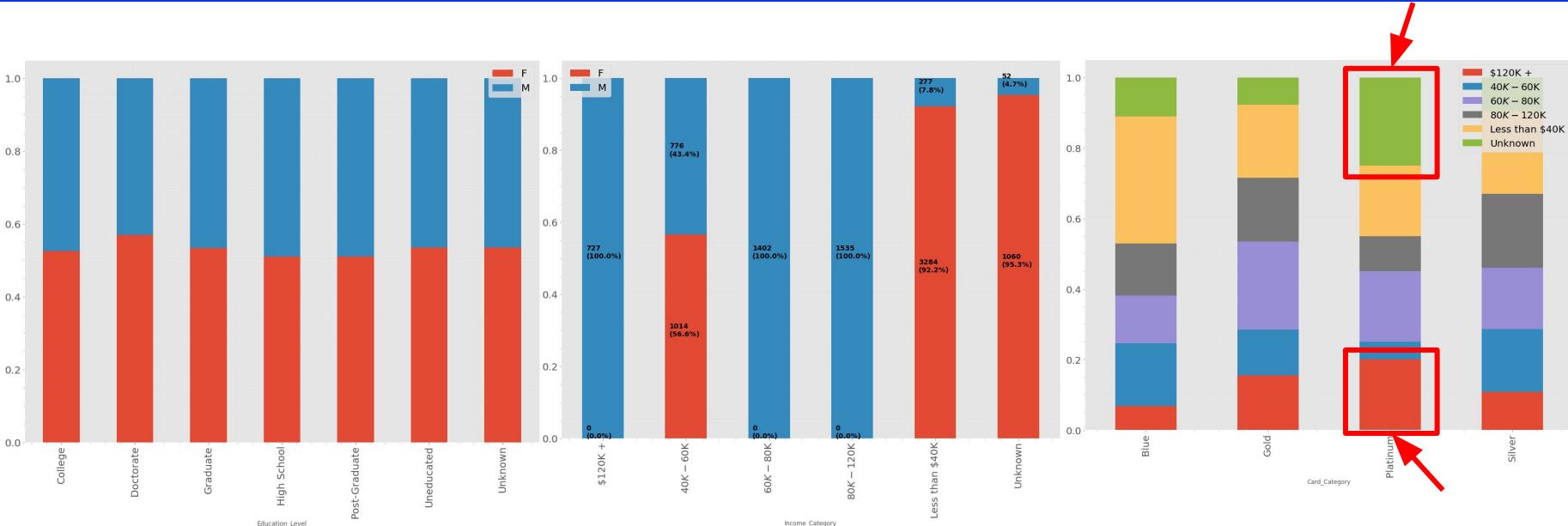
Male customers have more credit limit than female customers. But female customers have more average utilization ratio.

Bivariate Analysis - Card Category



- Blue card has less credit (but many higher end outliers), less transactions count and transaction amount although its average utilization ratio is high (due to small amount of credit, customers can reach the limit easily)

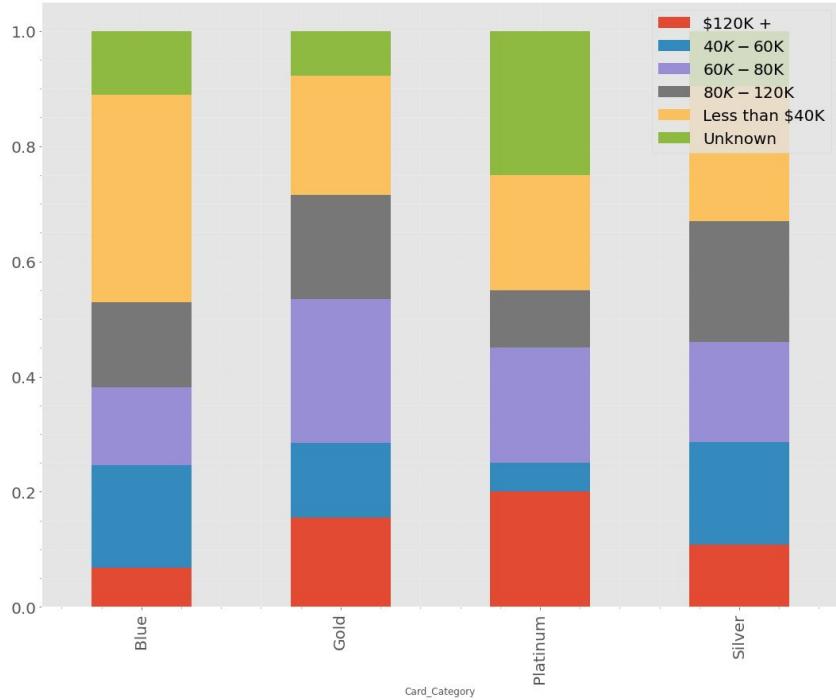
Bivariate Analysis



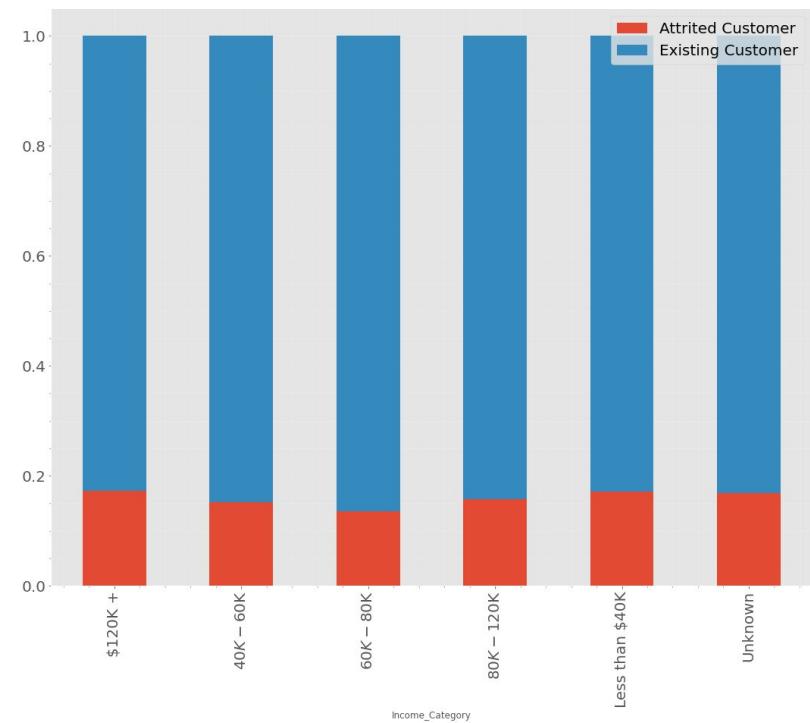
```
> credit_card_cust %>% filter(Income_category == 'Unknown') %>% group_by(Gender) %>% summarize(n())
# A tibble: 2 × 2
  Gender `n()`
  <fct> <int>
1 M      52
2 F     1059
```

There are 1,059 unknown values in Income Category of female, so we cannot compare the actual education level and income by gender.

Bivariate Analysis

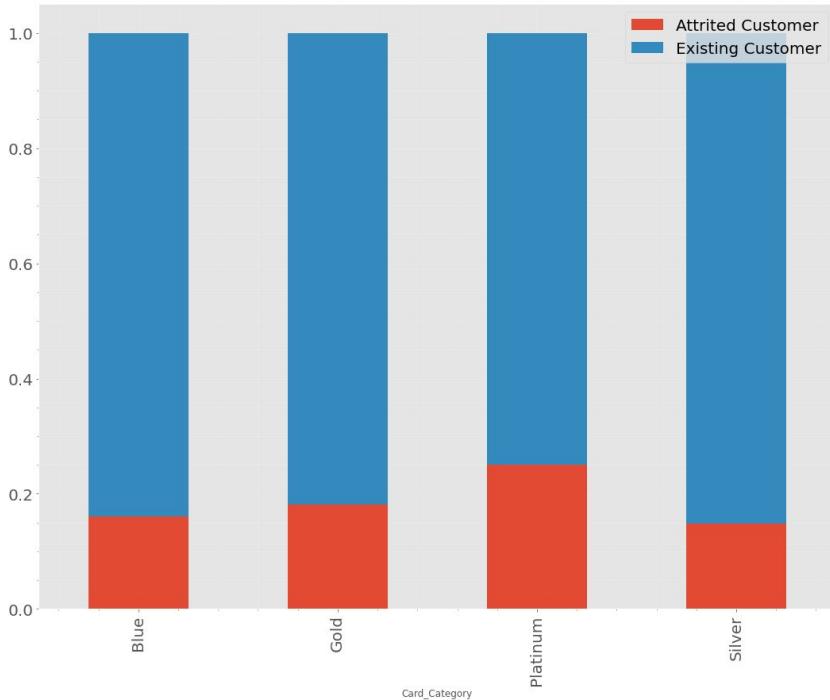


Blue cards are owned more by customers whose income is less than 40k, Platinum card owners contains more customers with income more than 120k. Customers with higher income have more advanced card.

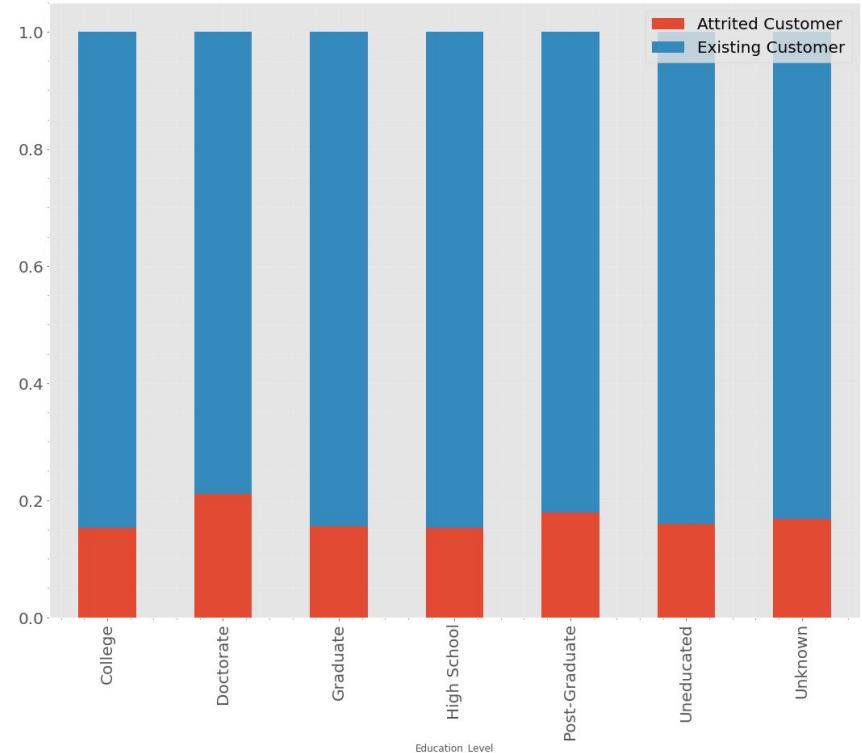


Higher attrition rates among people earning < \$40K and >\$120K. They are looking for better deals elsewhere.

Bivariate Analysis

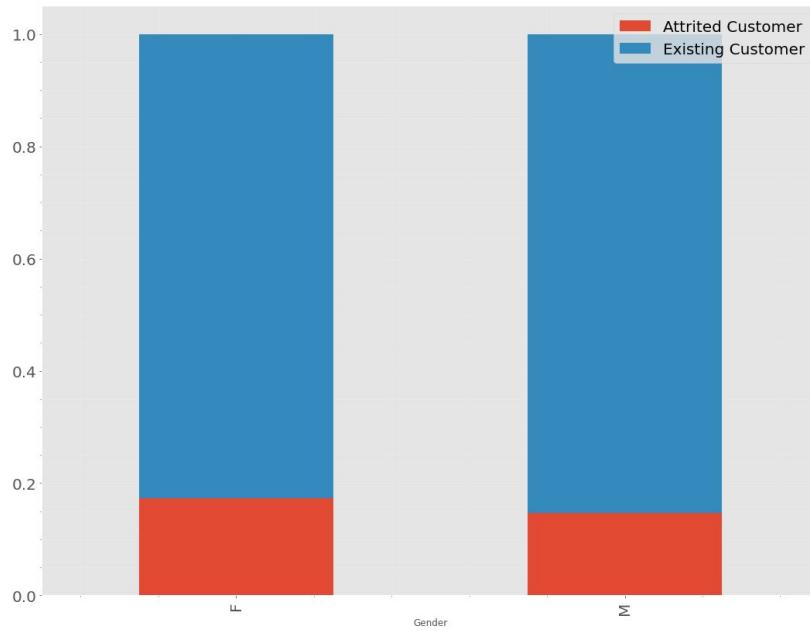


Customers with higher credit limits (Platinum card holders) seem more likely to attrite.

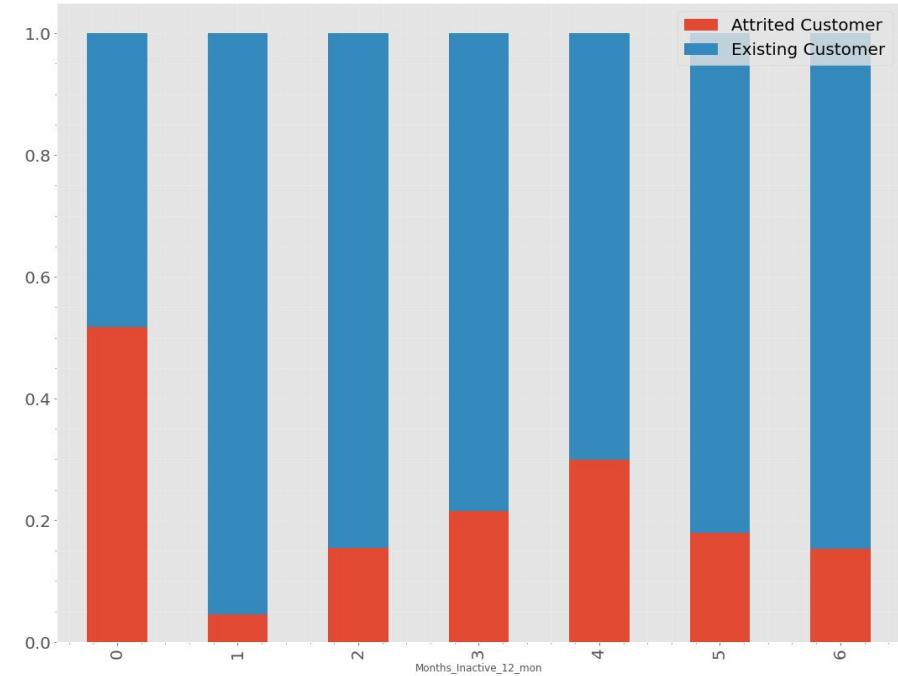


Customers with higher degrees (doctorate card post-graduate) seem more likely to attrite.

Bivariate Analysis

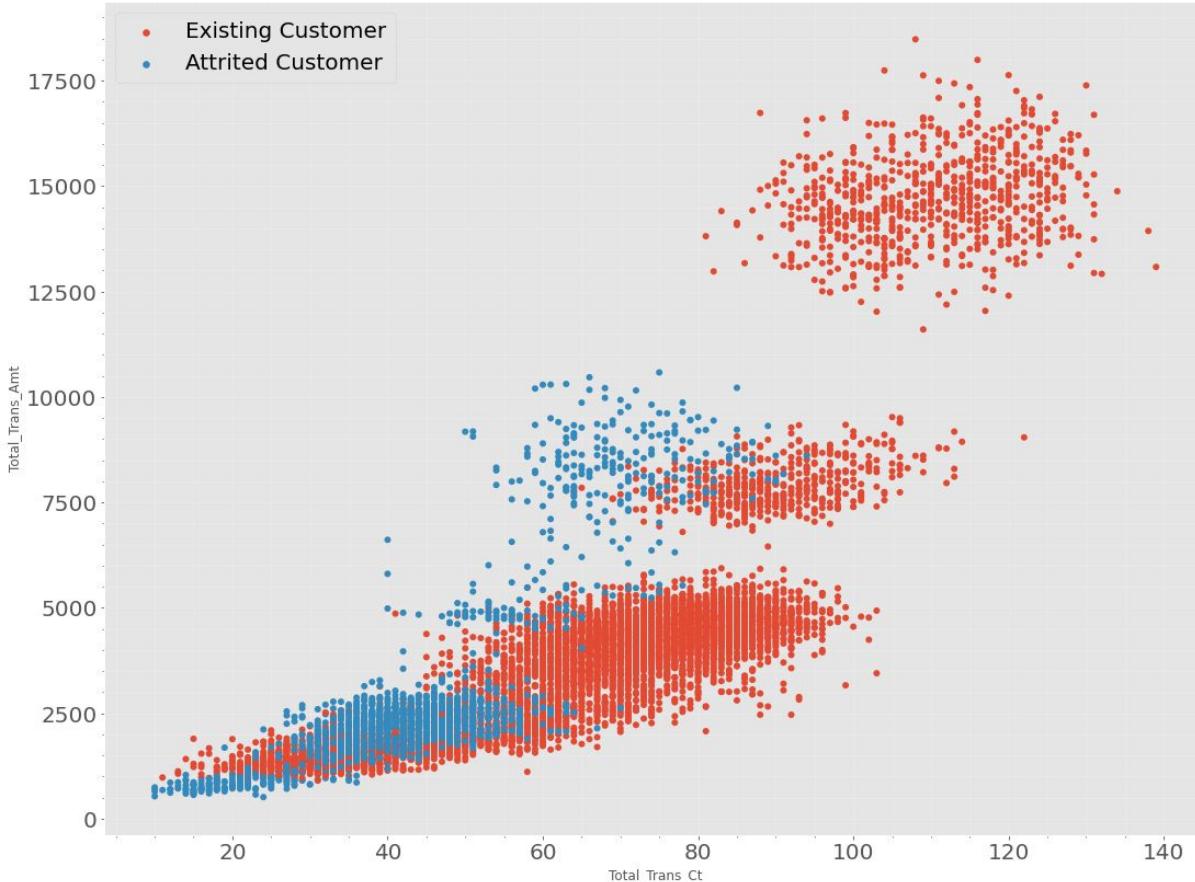


Higher rate of attrition among females than males.



Attrited customers can also be active in the last 12 months.

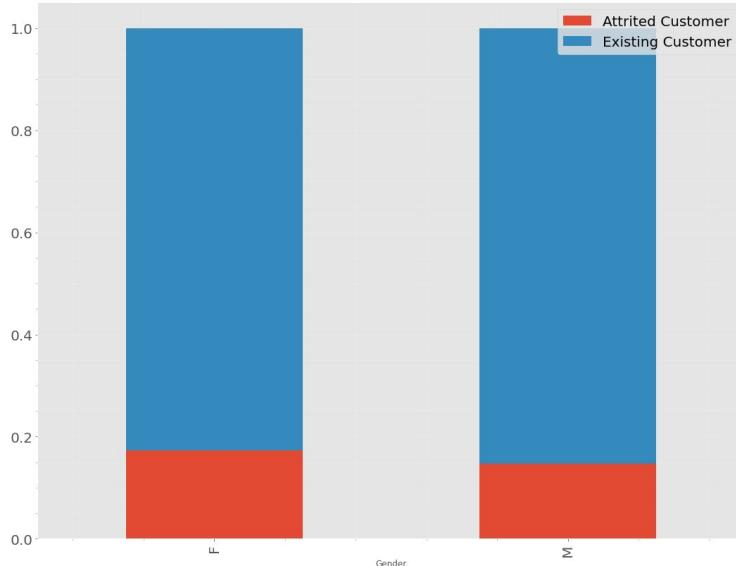
Bivariate Analysis



- Existing customers have higher transaction count and transaction amount.
- Customers with higher transaction count and higher transaction amount do not attrition (Not even one blue dot in the top right corner).

Chi-Squared Test For Category Variables (Gender)

- **Null Hypothesis:** Female and male customers have same attrition rates.
- **Alternate Hypothesis:** Female and male customers have different attrition rates.



Contingency Table

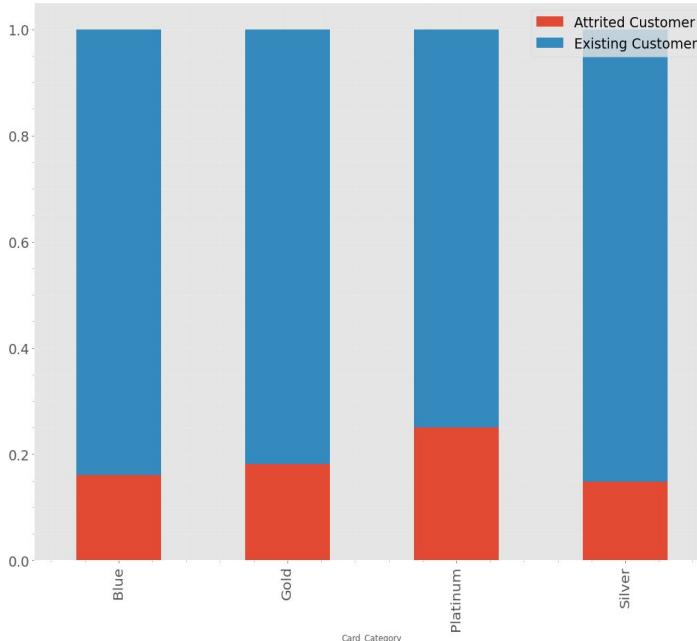
Gender	Attrited_Customer	Existing_Customer	rate
Male	930	4428	0.174
Female	697	4072	0.146

χ^2 Statistic = 13.856
p-Value = 0.0002
Dof = 1

P-value < 0.05, reject Null hypothesis.
Female and male customers have different attrition rate.

Chi-Squared Test For Category Variables (Credit Card Category)

- **Null Hypothesis:** Different card categories have same attrition rates.
- **Alternate Hypothesis:** Different card categories have different attrition rates.



Contingency Table

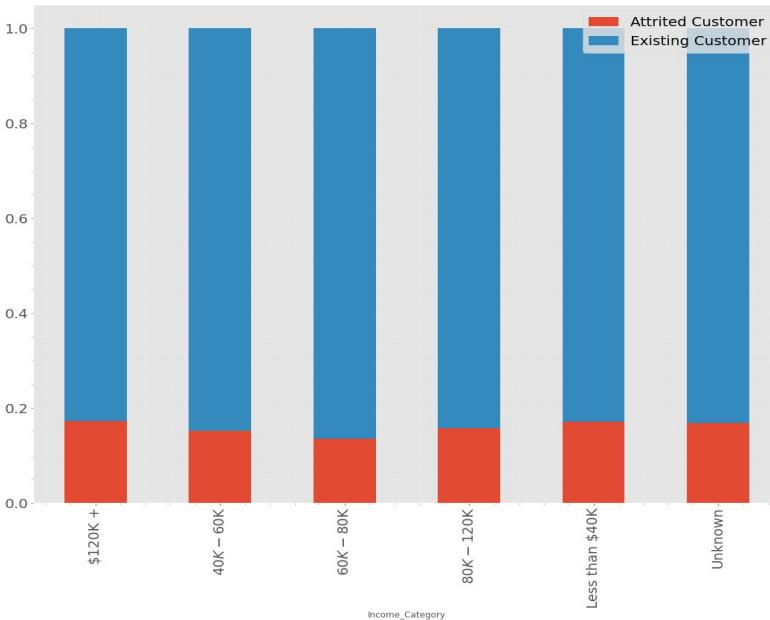
Card category	Attrited_Customer	Existing_Customer	rate
Blue	1519	7917	0.161
Gold	21	95	0.181
Platinum	5	15	0.250
Silver	82	473	0.148

χ^2 Statistic = 2.23
p-Value = 0.5252
Dof = 3

P-value > 0.05, can't reject null hypothesis.
Different card categories have same attrition rates.

Chi-Squared Test For Category Variables (Income Category)

- **Null Hypothesis:** Different income categories have same attrition rates.
- **Alternate Hypothesis:** Different income categories have different attrition rates.



Contingency Table

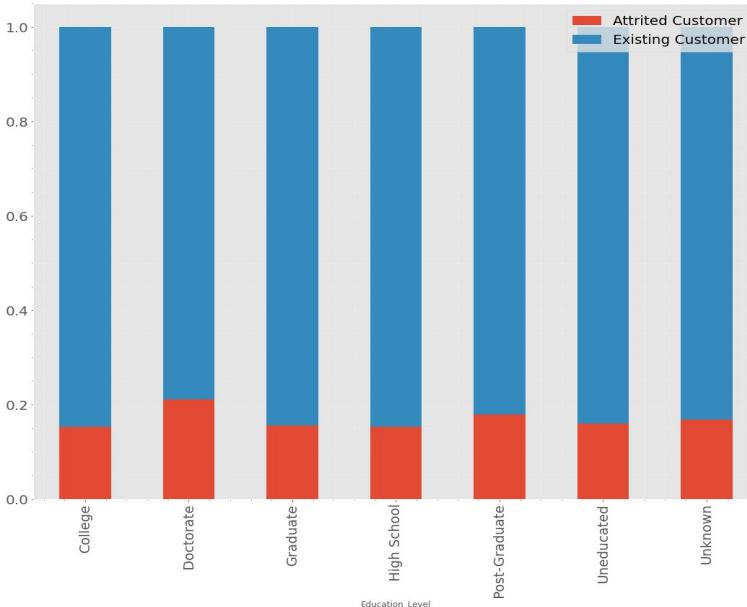
Income	Attrited_Customer	Existing_Customer	rate
>\$120k	176	601	0.227
\$40-\$60k	271	1519	0.151
\$60-\$80k	189	1213	0.135
\$80-\$120k	242	1293	0.158
<\$40k	612	2949	0.172
Unknown	187	925	0.168

$$\begin{aligned}\chi^2 \text{ Statistic} &= 12.83 \\ p\text{-Value} &= 0.025 \\ \text{Dof} &= 5\end{aligned}$$

P-value < 0.05, reject Null hypothesis.
Different income categories have different attrition rates.

Chi-Squared Test For Category Variables (Education Level)

- **Null Hypothesis:** Different education levels have same attrition rates.
- **Alternate Hypothesis:** Different education level have different attrition rates.



Contingency Table

Education level	Attrited_Customer	Existing_Customer	rate
College	154	859	0.152
Doctorate	95	356	0.211
Graduate	487	2641	0.156
High School	306	1707	0.152
Post-Graduate	92	424	0.178
Uneducated	237	1250	0.159
Unknown	256	1263	0.169

$$\begin{aligned}\chi^2 \text{ Statistic} &= 12.51 \\ p\text{-Value} &= 0.0514 \\ \text{Dof} &= 6\end{aligned}$$

P-value > 0.05, can't reject null hypothesis.
Different education levels have same attrition rates.

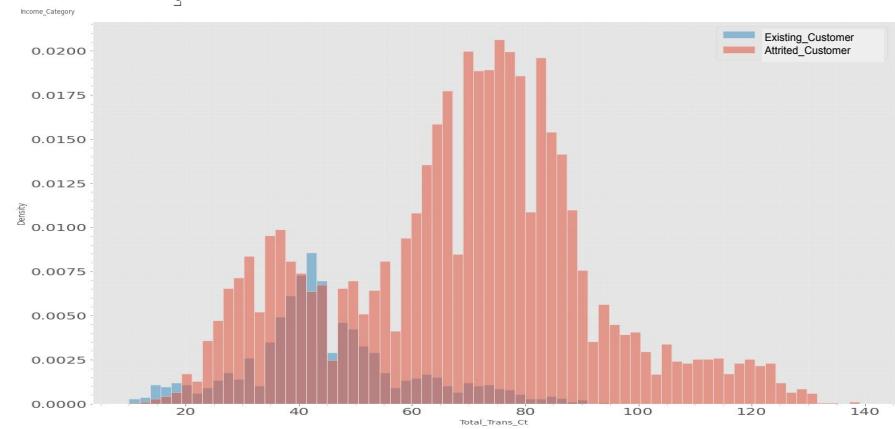
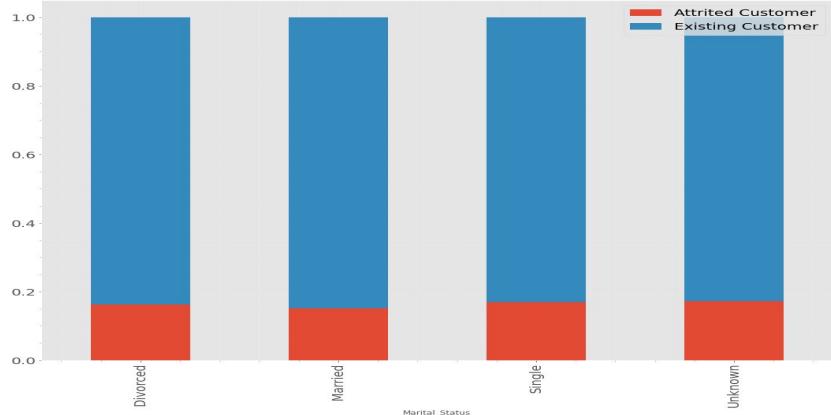
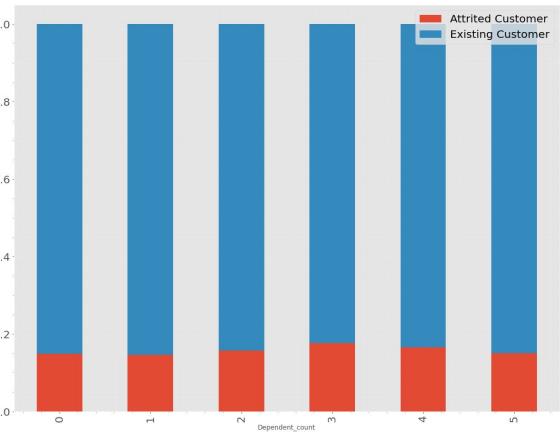
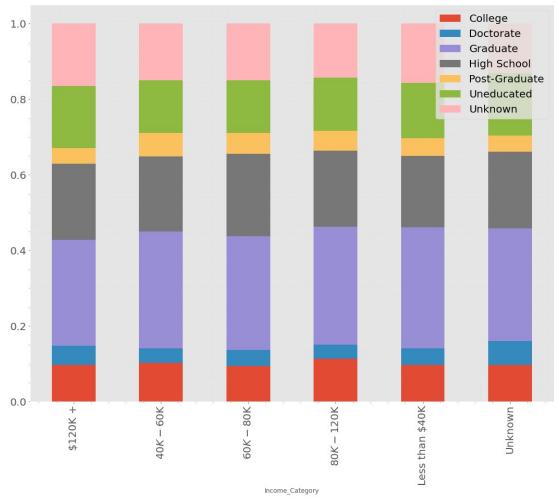
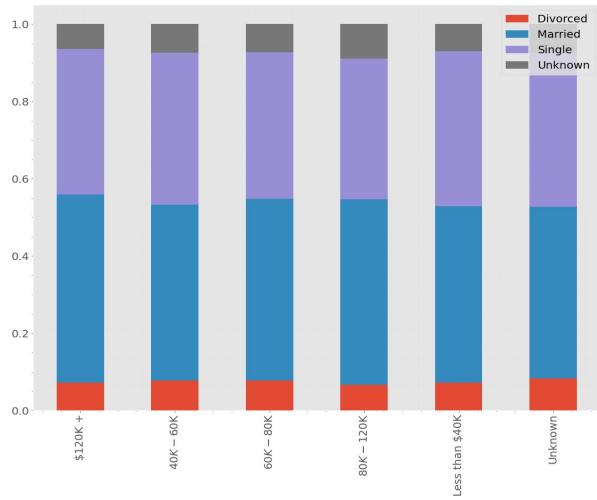
Bivariate Analysis Summary

- Months_on_book and Customer_Age, Total_Trans_Ct and Total_Trans_Ct have strong correlation, need to further explore their relationship by model.
- Existing customers have higher transaction count and transaction amount (more total revolving balance and average utilization). They are more likely to connect with the bank.
- Male customers have higher credit limit than female customers. But female customers have higher average utilization ratio.
- Attrition rates are significantly different for customers of different genders and income categories.

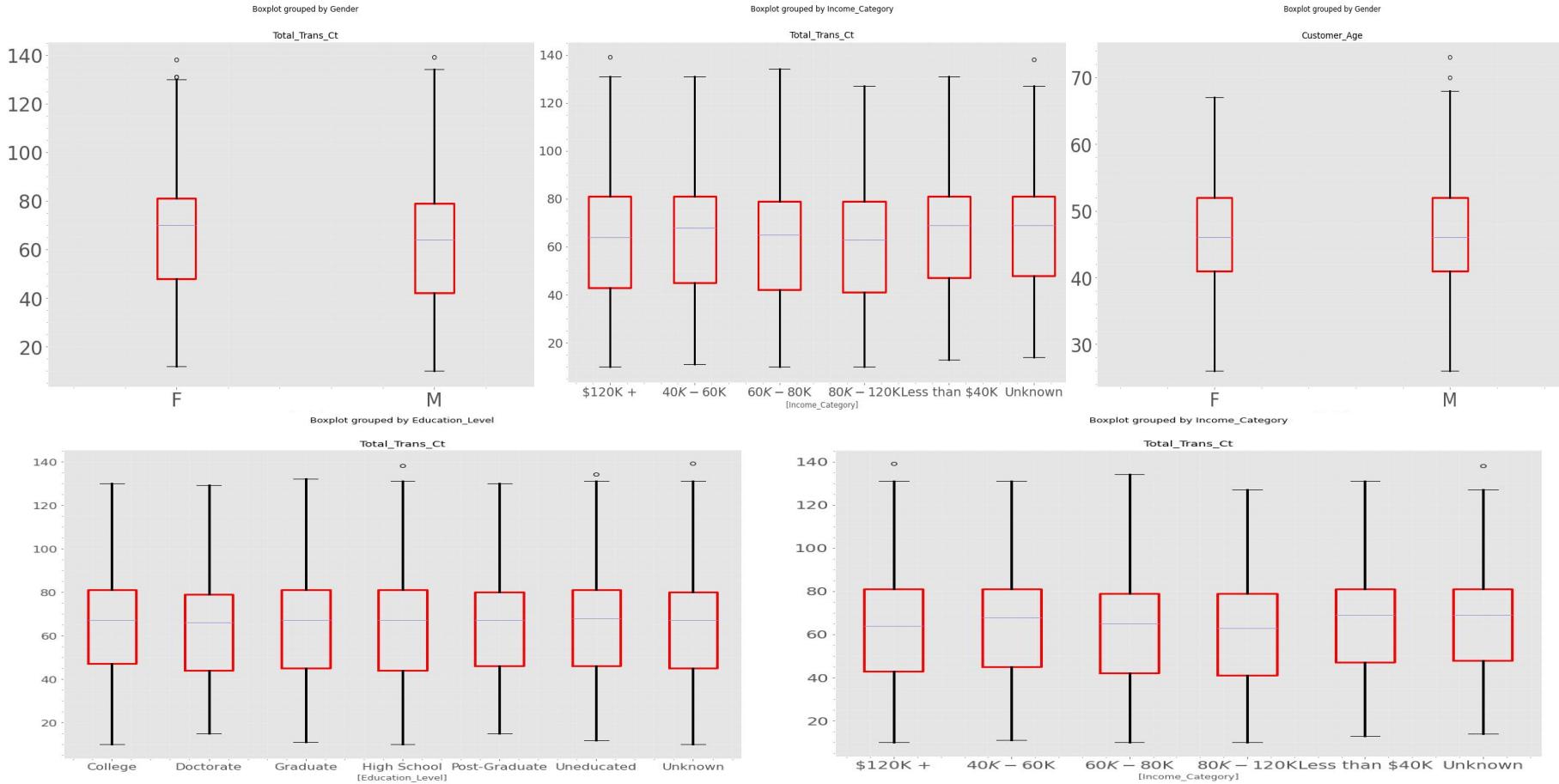
Analysis Plan - Noteworthy Features

- The total number of transactions in the last 12 months is the most important feature to keep in mind while building the model.
- Many unknown data in Income_Category of female, we need to do further exploration.
- Platinum Card customers have the highest attrition rate of 25% [Refer slide 33].
- Customers who have a doctorate degree also have a high attrition rate.
- Our efforts must focus on these factors two categories [Platinum Card and Doctorate degree customers] to reduce attrition.

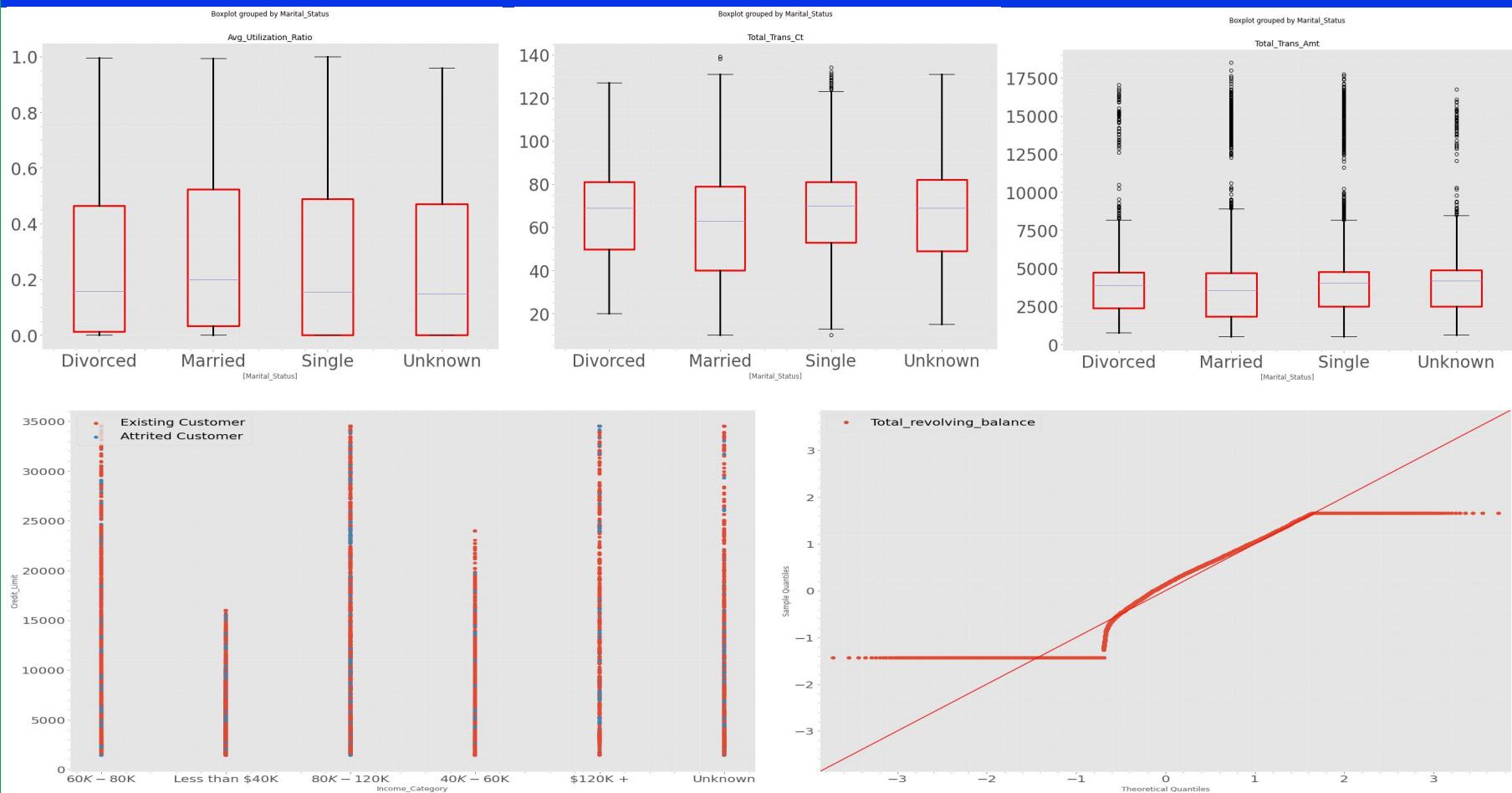
Appendix



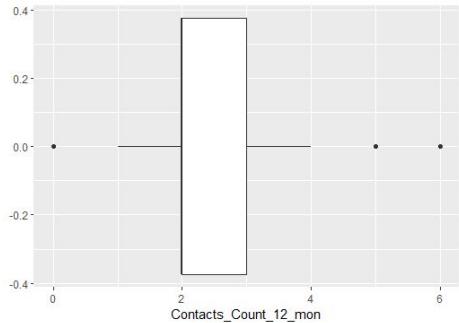
Appendix



Appendix

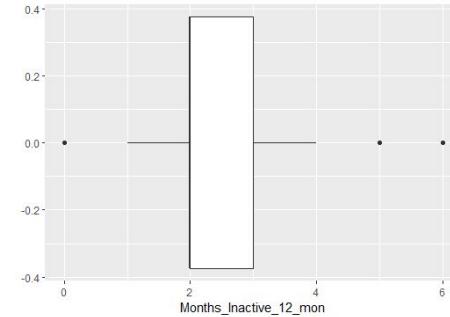


Appendix



Keep in the dataset

Contacts_Count_12_mon	n
3	3380
2	3227
1	1499
4	1392
0	399
5	176
6	54



Keep in the dataset

Months_Inactive_12_mon	n
3	3846
2	3282
1	2233
4	435
5	178
6	124
0	29

Attrition_Flag	Customer_Age	Gender	Education_Level	Income_Category	Card_Category	Total_Trans_Amt	Total_Trans_Ct	Customer_Age`n()
	<dbl>	<fctr>	<fctr>	<fctr>	<fctr>	<dbl>	<dbl>	<dbl>
Existing Customer	73	M	High School	\$40K - \$60K	Blue	1765	34	73 1
Existing Customer	70	M	High School	Less than \$40K	Blue	1227	15	70 1
Existing Customer	68	M	Graduate	Unknown	Blue	1910	32	68 2
Attrited Customer	68	M	High School	Less than \$40K	Blue	760	21	67 4
Existing Customer	67	F	Graduate	Less than \$40K	Blue	1661	32	66 2
Existing Customer	67	M	Graduate	\$40K - \$60K	Blue	2133	55	65 101
Existing Customer	67	M	Uneducated	\$40K - \$60K	Blue	1512	38	64 43
Existing Customer	67	F	Unknown	Unknown	Blue	1365	34	63 65

.. with 35 more rows

Existing customers, keep in the dataset

Appendix

A tibble: 50 × 9

Attrition_Flag <fctr>	Total_Amt_Chng_Q4_Q1 <dbl>	Customer_Age <dbl>	Gender <fctr>	Education_Level <fctr>	Income_Category <fctr>			
Existing Customer	3.397	56	M	College	\$80K - \$120K			
Existing Customer	3.355	37	M	Uneducated	\$60K - \$80K			
Existing Customer	2.675	61	M	Post-Graduate	Unknown			
Existing Customer	2.594	51	M	Graduate	\$80K - \$120K			
Existing Customer	2.368	44	F	Uneducated	Less than \$40K			
Existing Customer	2.357	59	M	Doctorate	\$40K - \$60K			
Existing Customer	2.316	56	M	Doctorate	\$60K - \$80K			
Existing Customer	2.282	46	M	Graduate	\$60K - \$80K			
Existing Customer	2.275	44	F	Graduate	Unknown			
Existing Customer	2.271	63	M	Graduate	\$60K - \$80K			

Attrition_Flag <fctr>	Total_Ct_Chng_Q4_Q1 <dbl>	Customer_Age <dbl>	Gender <fctr>	Education_Level <fctr>	Income_Category <fctr>			
Existing Customer	3.714	49	F	Graduate	Less than \$40K			
Existing Customer	3.571	61	M	Post-Graduate	Unknown			
Existing Customer	3.500	54	M	Graduate	\$60K - \$80K			
Existing Customer	3.250	56	M	College	\$80K - \$120K			
Existing Customer	3.000	54	F	Uneducated	Less than \$40K			
Existing Customer	3.000	57	M	Graduate	\$80K - \$120K			
Existing Customer	2.875	41	F	Graduate	Less than \$40K			
Existing Customer	2.750	36	F	Graduate	\$40K - \$60K			
Existing Customer	2.571	53	M	Unknown	\$80K - \$120K			
Existing Customer	2.500	40	M	Uneducated	\$60K - \$80K			

Appendix

