Bank Churner

Title	Objective	Variables	Numeric variables - T-test	Total Relationship Count	Month Inactive	Contacts count

BANK CHURNER

[Reduce Churn Rate]

Presented by Seul Lee
Nov.08.2021

Bank Churner

Title	Objective	Variables	Numeric variables - T-test	Total Relationship Count	Month Inactive	Contacts count

Problem Statement

Churn is a serious issue because it costs a lot for customer to stop doing business.
For example, the Havard Business School report claims that on average, a **5% increase** in customer retention rates results in 25% - 95% increase of profits.

The bank decided to **find a group of customers that is more likely to leave** than others and **offer better marketing service to reduce churn rate.**

This data consists of **10,127** customers mentioning their gender, salary, marital satatus, credit card limit, etc.

Current churn rate is **16.07%.**

Churn Rate

Attrition Flag	
Attrited Customer	16.07% 1,627
Existing Customer	83.93% 8,500

Bank Churner

Variables in dataset

Divded the variables into numeric and categoric variables.

Numeric Variables
Customer Age
Months on book
Total Relationship Count
Months Inactive 12 mon
Contacts Count 12 mon
Credit Limit
Total Revolving Bal
Avg Open To Buy
Total Amt Chng Q4 Q1
Total Trans Amt
Total Trans Ct
Total Ct Chng Q4 Q1
Avg Uilization Ratio

Categoric Variables
Gender
Education Level
Marital Status
Income Category
Card Category
Dependent Count

Bank Churner

T-test

Implemeted T-test with numeric variables to establish significance.

*Attrited customer and existing customer classes are releveled to 1 and 0 respectively.

Significant numeric variables are :
Total_Relationship_Count / Months_Inactive_12_mon
Contacts_Count_12_mon / Credit_Limit
Total_Revolving_Bal / Avg_Utilization_Ratio
Total_Amt_Chng_Q4_Q1 / Total_Ct_Chng_Q4_Q1
Total_Trans_Amt / Total_Trans_Ct

t-Test: Two-Sample Assuming Equal Variances

	Contacts_Count_12_mon_0	Contacts_Count_12_mon_1
Mean	2.356352941	2.972341733
Variance	1.169503111	1.18927145
Observations	8500	1627
Pooled Variance	1.17267776	
Hypothesized Mean Difference	0	
df	10125	
t Stat	-21.0206577	
P(T<=t) one-tail	2.34874E-96	
t Critical one-tail	1.645004137	
P(T<=t) two-tail	4.69749E-96	
t Critical two-tail	1.96019831	

t-Test: Two-Sample Assuming Unequal Variances

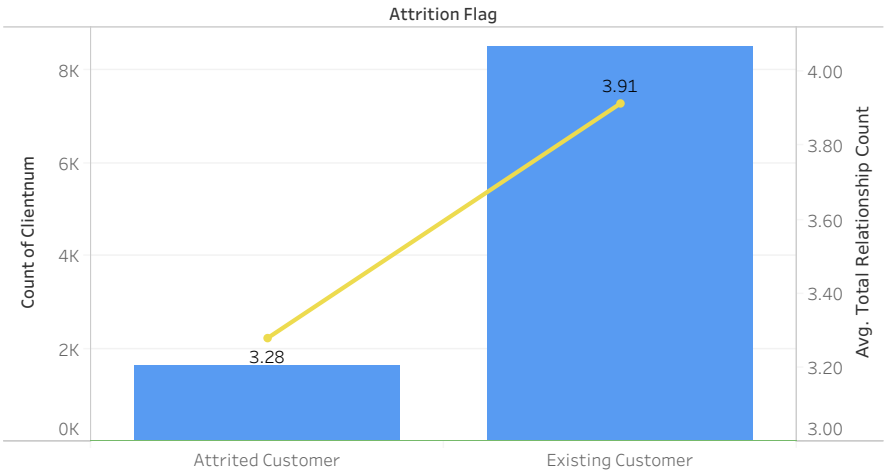
	Total_Trans_Ct_0	Total_Trans_Ct_1
Mean	68.67258824	44.93362016
Variance	525.2810699	212.2391334
Observations	8500	1627
Hypothesized Mean Difference	0	
df	3386	
t Stat	54.14187251	
P(T<=t) one-tail	0	
t Critical one-tail	1.645303771	
P(T<=t) two-tail	0	
t Critical two-tail	1.960664842	

Bank Churner

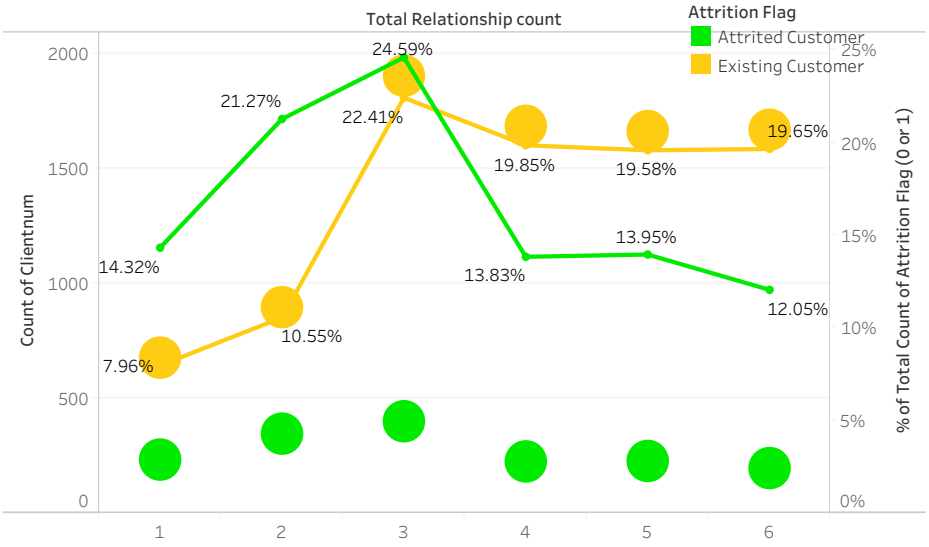
Objective	Variables	Numeric variables - T-test	Total Relationship Count	Month Inactive	Contacts count	Avg. Total Amount & Count Change (Q4 o..

Total Relationship Count

Attrition customer class has **lower** average total relationship count.



Customers with **3 or less** total relationship count has **a higher proprotion of attrition.**

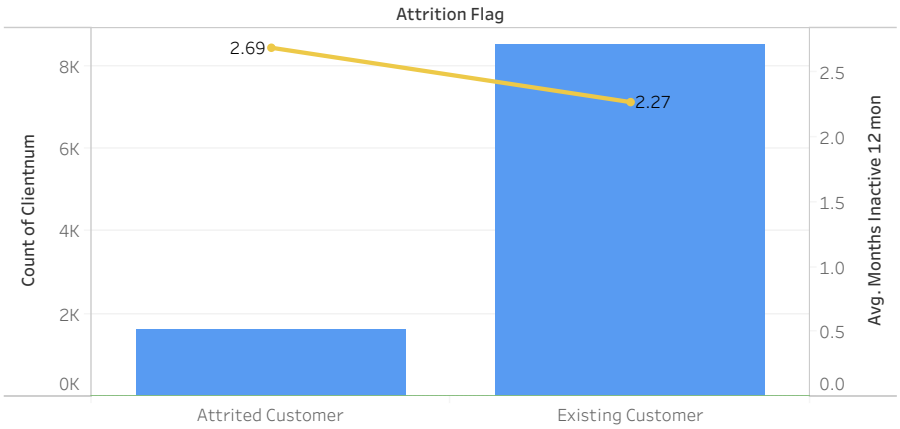


Bank Churner

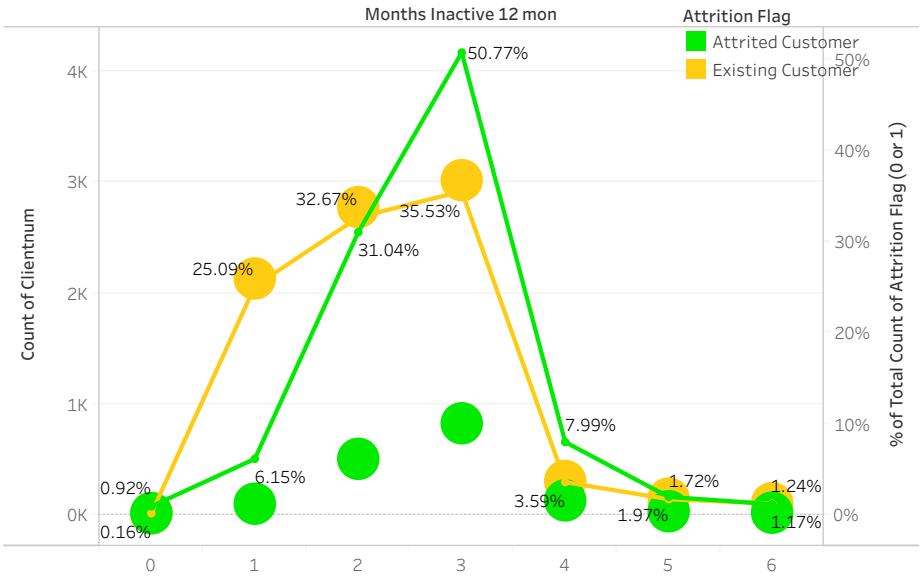
Variables	Numeric variables - T-test	Total Relationship Count	Month Inactive	Contacts count	Avg. Total Amount & Count Change (Q4 ove..	Avg. Utilization Ratio

Month Inactive Count

Attrition customer class has **higher** average inactive months count in the past 12 months.



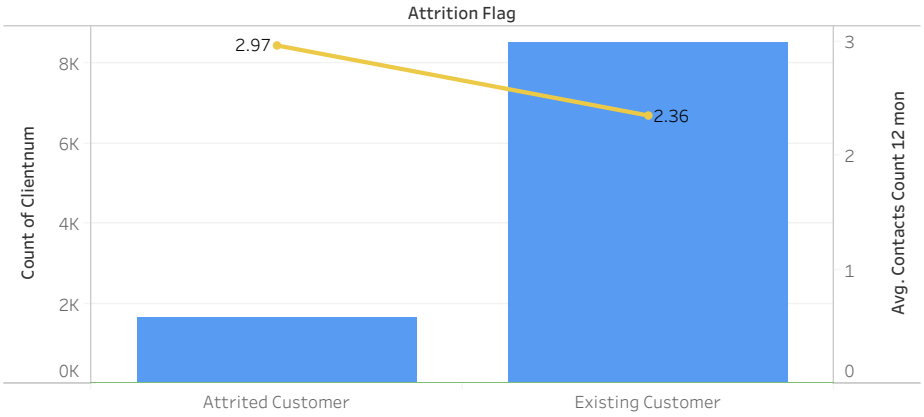
Customers with **3 or more** inactive in the past 12 months have a higher proportion of attrition.



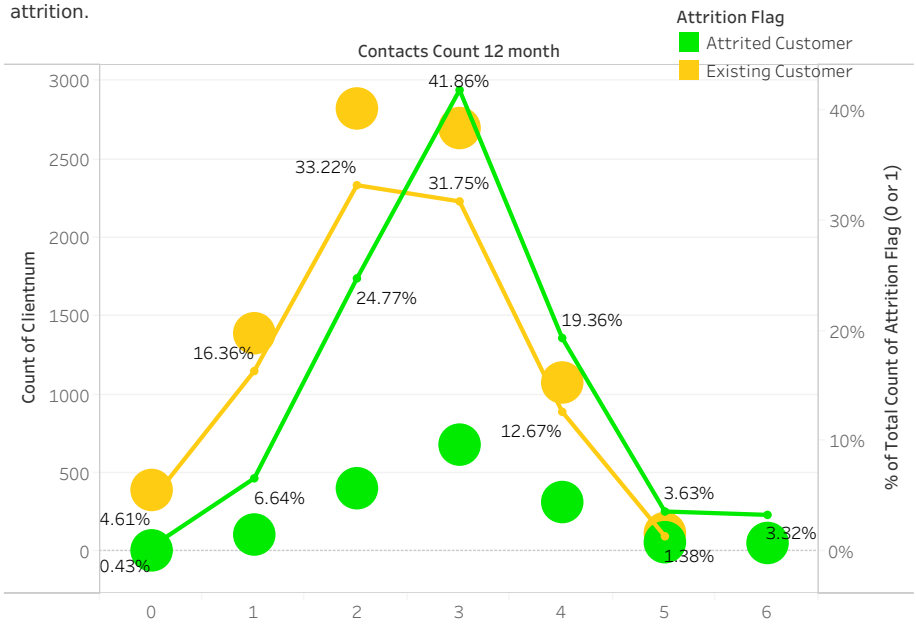
Bank Churner

Contacts Count

Attrition customer class has **higher** average contacts count in the past 12 months.



Customers with **3 or more** contacts count in the past 12 months have a higher proportion of attrition.

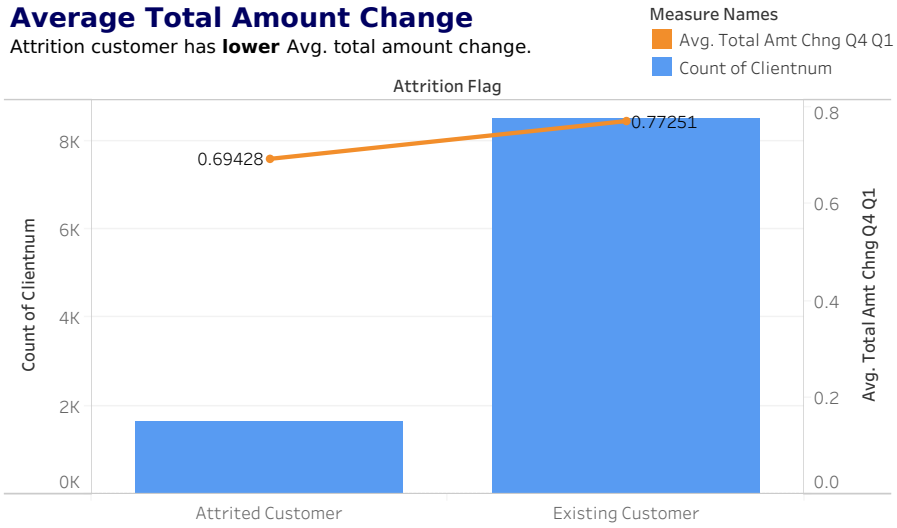


Bank Churner

Total Relationship Co..	Month Inactive	Contacts count	Avg. Total Amount & Count Change (Q4 ove..	Avg. Utilization Ratio	Transaction Count & Amount Scatter Plot	Transaction Count Box Plot

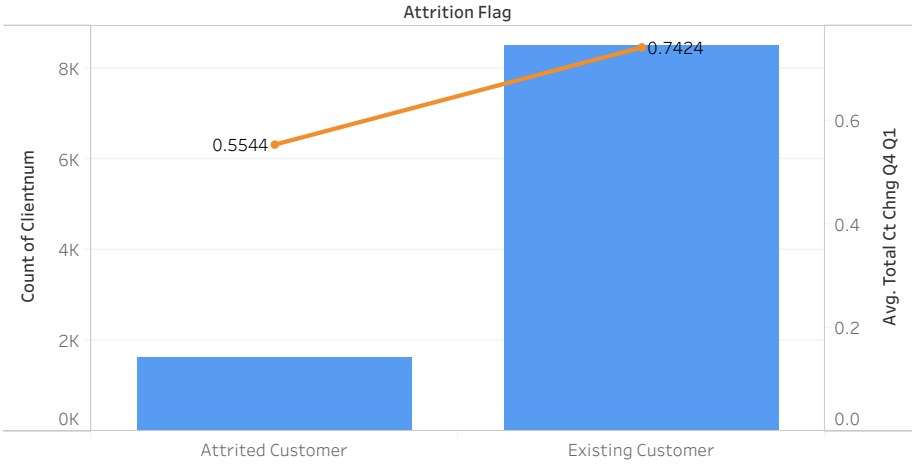
Average Total Amount Change

Attrition customer has **lower** Avg. total amount change.



Average Total Count Change

Attrition customer has **lower** Avg. total count change.

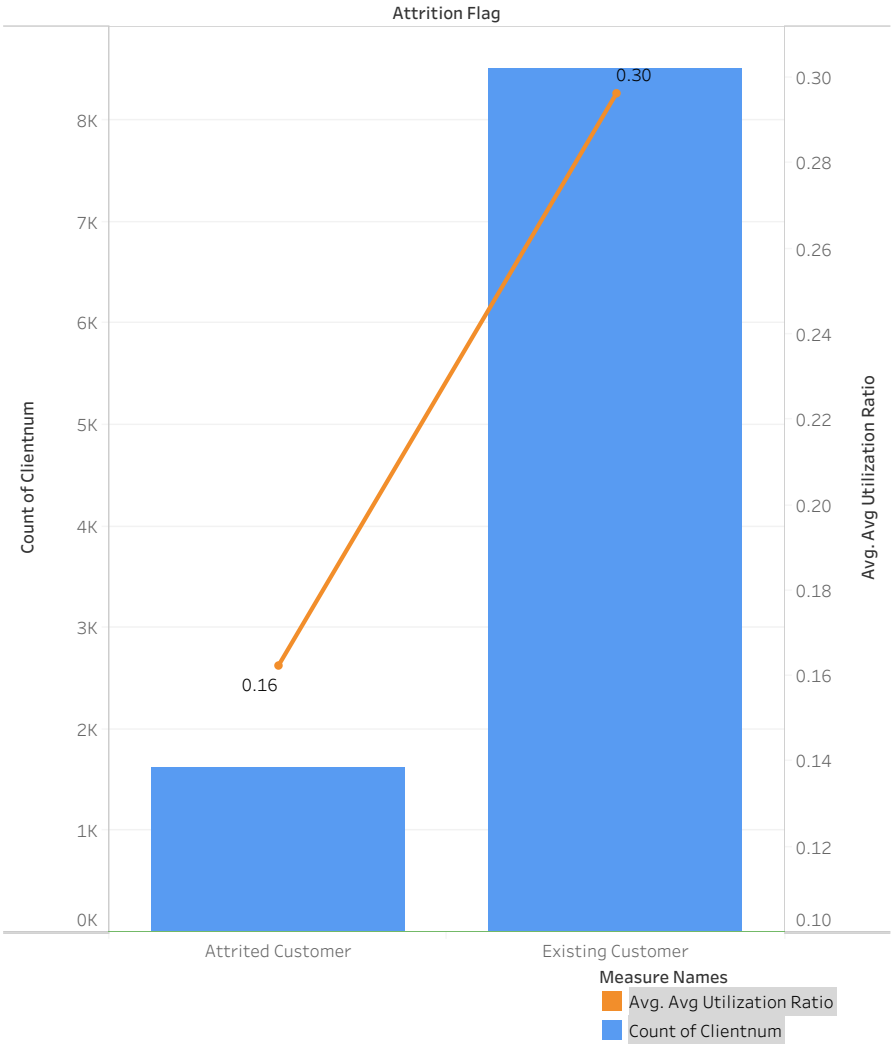


Bank Churner

Month Inactive	Contacts count	Avg. Total Amount & Count Change (Q4 ove..	Avg. Utilization Ratio	Transaction Count & Amount Scatter Plot	Transaction Count Box Plot	Categoric Variables - IV

Avg. Utilization Ratio

Attrition customer class has **lower** average utilization ration.

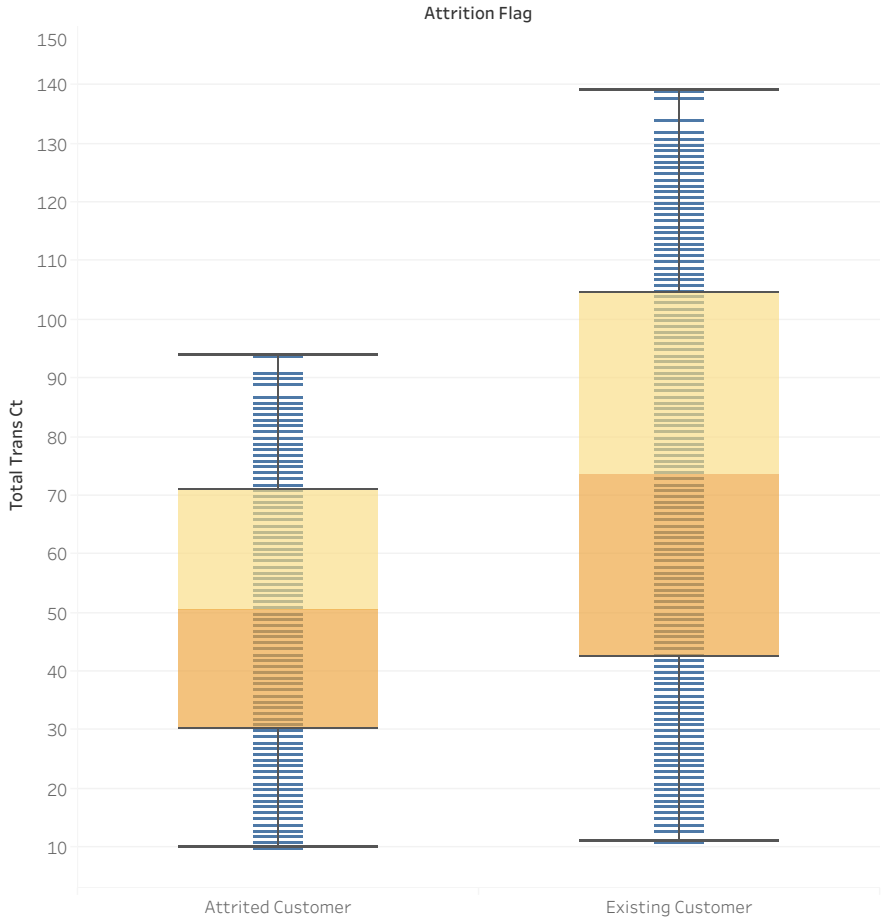


Bank Churner

Bank Churner

Totlal Transaction Count - Box Plot

Attrited Customer
Min : 10 Max : 94 Median : 50.5
Existing Customer
Min : 11 Max : 139 Median : 73.5



Bank Churner

Categoric Variables - Information Value

Tested Infomation Value to establish significance of categoric variables.
All the categoric variables are useless for prediction.

Edcuation Level

Row Labels	Attrited C	Existing C	Total	Churn rate	Attrited %	Existing %	Population%	IV
College	154	859	1013	15%	9%	10%	10%	0.000419522
Doctorate	95	356	451	21%	6%	4%	4%	0.005484962
Graduate	487	2641	3128	16%	30%	31%	31%	0.00042478
High School	306	1707	2013	15%	19%	20%	20%	0.00083596
Post-Graduate	92	424	516	18%	6%	5%	5%	0.000835485
Uneducated	237	1250	1487	16%	15%	15%	15%	1.3238E-05
Unknown	**256**	**1263**	**1519**	17%	16%	15%	15%	0.000501407
Grand Total	**1627**	**8500**	**10127**	16%	100%	100%	100%	**0.008515354**

Card Category

Row Labels	Attrited C	Existing C	Total	Churn rate	Attrited %	Existing %	Population%	IV
Blue	1519	7917	9436	16%	93%	93%	93%	5.22995E-06
Gold	21	95	116	18%	1%	1%	1%	0.000249179
Platinum	5	15	20	25%	0%	0%	0%	0.00072581
Silver	82	473	555	15%	5%	6%	5%	0.000519758
Grand Total	1627	8500	10127	16%	100%	100%	100%	**0.001499977**

Card Category

Row Labels	Attrited C	Existing C	Total	Churn rate	Attrited %	Existing %	Population%	IV
Blue	1519	7917	9436	16%	93%	93%	93%	5.22995E-06
Gold	21	95	116	18%	1%	1%	1%	0.000249179
Platinum	5	15	20	25%	0%	0%	0%	0.00072581
Silver	82	473	555	15%	5%	6%	5%	0.000519758
Grand Total	1627	8500	10127	16%	100%	100%	100%	**0.001499977**

Income Category

Row Labels	Attrited C	Existing C	Total	Churn rate	Attrited %	Existing %	Population%	IV
$120K +	126	601	727	17%	8%	7%	7%	0.000613194
$40K - $60K	271	1519	1790	15%	17%	18%	18%	0.000854291
$60K - $80K	189	1213	1402	13%	12%	14%	14%	0.005461549
$80K - $120K	242	1293	1535	16%	15%	15%	15%	7.58425E-05
Less than $40K	612	2949	3561	17%	38%	35%	35%	0.002361414
Unknown	187	925	1112	17%	11%	11%	11%	0.000333976
Grand Total	1627	8500	10127	16%	100%	100%	100%	**0.009700267**

Bank Churner

Logistic Regression

Combine significant variables and come up with equation building a predictive model.

Regression Statistics	
Multiple R	0.603291195
R Square	0.363960266
Adjusted R Square	0.363043771
Standard Error	0.293065674
Observations	10127

ANOVA

	df	SS	MS	F	Significance F
Regression	14	497.026612	35.50190086	445.150004	0
Residual	10113	868.5801817	0.08588749		
Total	10127	1365.606794			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.973848813	0.028422205	34.26366122	1.597E-243	0.918135647	1.02956198	0.918135647	1.02956198
Customer_Age	-0.000162143	0.00059423	-0.272861979	0.784964885	-0.001326951	0.001002666	-0.001326951	0.001002666
Dependent_count	0.012014899	0.002273745	5.284189284	1.2889E-07	0.007557907	0.01647189	0.007557907	0.01647189
Months_on_book	-0.000403519	0.000594569	-0.678674116	0.4973599	-0.001568992	0.000761955	-0.001568992	0.000761955
Total_Relationship_Count	-0.043210702	0.002014136	-21.45371321	6.953E-100	-0.047158809	-0.039262595	-0.047158809	-0.039262595
Months_Inactive_12_mon	0.042271509	0.002897927	14.58680727	1.0383E-47	0.036590996	0.047952023	0.036590996	0.047952023
Contacts_Count_12_mon	0.041579147	0.00268123	15.50749273	1.27598E-53	0.036323405	0.046834889	0.036323405	0.046834889
Credit_Limit	0	0	65535	#NUM!	0	0	0	0
Total_Revolving_Bal	-9.91982E-05	5.13943E-06	-19.30141019	#NUM!	-0.000109273	-8.91239E-05	-0.000109273	-8.91239E-05
Avg_Open_To_Buy	-1.48132E-06	4.25696E-07	-3.479773417	0.000503953	-2.31577E-06	-6.46876E-07	-2.31577E-06	-6.46876E-07
Total_Amt_Chng_Q4_Q1	-0.062076728	0.014506088	-4.27935688	1.8916E-05	-0.090511542	-0.033641915	-0.090511542	-0.033641915
Total_Trans_Amt	3.47641E-05	1.54869E-06	22.44748544	6.0376E-109	3.17284E-05	3.77999E-05	3.17284E-05	3.77999E-05
Total_Trans_Ct	-0.009658154	0.00021578	-44.75931213	0	-0.010081126	-0.009235183	-0.010081126	-0.009235183
Total_Ct_Chng_Q4_Q1	-0.295666247	0.013446065	-21.98905325	1.0163E-104	-0.322023206	-0.269309289	-0.322023206	-0.269309289
Avg_Utilization_Ratio	0.006545377	0.017989649	0.363841289	0.715984135	-0.028717908	0.041808663	-0.028717908	0.041808663

	Coefficients
Intercept	0.901868012
Total_Relationship_Count	-0.044310603
Months_Inactive_12_mon	0.044669653
Contacts_Count_12_mon	0.041803019
Total_Amt_Chng_Q4_Q1	-0.069834001
Total_Trans_Amt	2.85766E-05
Total_Trans_Ct	-0.009119931
Total_Ct_Chng_Q4_Q1	-0.305434882
Avg_Utilization_Ratio	-0.157258245

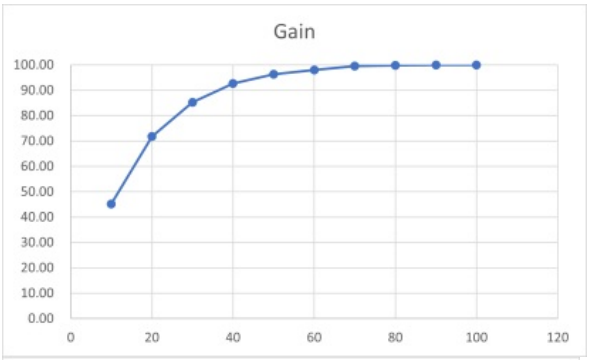
Bank Churner

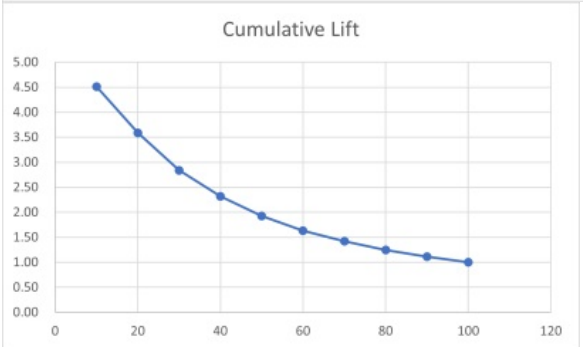
Gain and Lift Charts

Validate the predictive model with claasification evaluation.

These gain and lift charts proves that the predictive model performances better than random selection.

% of data sets	Decile	Number of Cases	Number of Responses	Cumulative Responses	% of events	Gain	Cumulative Lift
10	0.9~1	1012.7	734	734	45.11	45.11	4.51
20	0.8 ~ 0.899	1012.7	434	1168	26.67	71.79	3.59
30	0.7~0.799	1012.7	218	1386	13.40	85.19	2.84
40	0.6~0.699	1012.7	122	1508	7.50	92.69	2.32
50	0.5~0.599	1012.7	59	1567	3.63	96.31	1.93
60	0.4~0.499	1012.7	28	1595	1.72	98.03	1.63
70	0.3~0.399	1012.7	25	1620	1.54	99.57	1.42
80	0.2~0.299	1012.7	4	1624	0.25	99.82	1.25
90	0.1~0.199	1012.7	3	1627	0.18	100.00	1.11
100	0~0.099	1012.7	0	1627	0.00	100.00	1.00
	Total	**10127**	**1627**				





Bank Churner

Summary

The bank needs to focus more on the customers with below conditions and offer better marketing service.

Total Relationship Count : 3 or less
Inactive Months in the last 12 months : 3 or more
Contancts Count in the last 12 months : 3 or more
Total Transcation Amount : Under 71 transactions
Total Transaction Count : Lower counts
Total Amount change (O4 over O1) : Lower average
Total Count Change (O4 over O1) : Lower average
Average Utilization Ratio : Lower average