Customer Churn Analysis

To get started, I first took a look at what variables were correlated with each other, if at all, as a starting point. I created a simple correlation matrix for numeric variables in the copy of the Excel Sheet provided, named "Correlation Matrix.xlsx", shown below. I grouped the correlation scores based on their strength. Most variables were not correlated to each other all that much, as you can see. A few fields such as Age, Support Calls, Payment Delay and Last Interaction showed some correlation with churn, Support Calls having the largest score.

L	M	N	O	P	Q	R	S	T
Column1	**Age**	**Tenure**	**Usage Frequency**	**Support Calls**	**Payment Delay**	**Total Spend**	**Last Interaction**	**Churn**
Age	1	-0.00637	-0.010767653	0.159632161	0.060684844	-0.08885352	0.03073528	0.220574
Tenure	-0.006370215	1	-0.027605604	-0.022471565	-0.013822495	0.015123777	-0.008322481	-0.05113
Usage Frequency	-0.010767653	-0.02761	1	-0.026300114	-0.016129259	0.021091078	-0.000152136	-0.05035
Support Calls	0.159632161	-0.02247	-0.026300114	1	0.163478455	-0.218292441	0.080122811	0.571989
Payment Delay	0.060684844	-0.01382	-0.016129259	0.163478455	1	-0.123674885	0.046097575	0.314476
Total Spend	-0.08885352	0.015124	0.021091078	-0.218292441	-0.123674885	1	-0.057481189	-0.42744
Last Interaction	0.03073528	-0.00832	-0.000152136	0.080122811	0.046097575	-0.057481189	1	0.153179
Churn	0.22057442	-0.05113	-0.050351659	0.57198938	0.314476205	-0.427435012	0.153178742	1

Correlation Score	**Strength**
0-0.2	Almost no correlation
0.2-0.5	Weak
0.5+	Strong
Less than 0	Negative Relationship

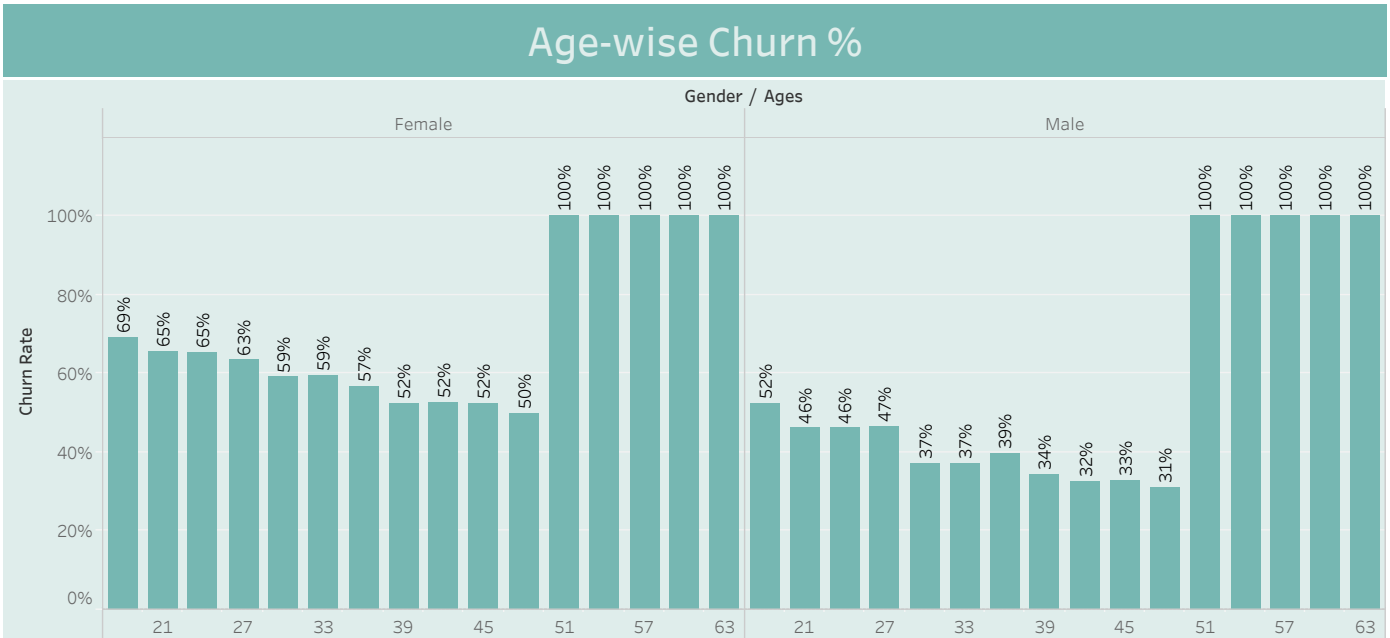
With that in mind, I started doing basic EDA and tried to find reasons for customers churning out. In the cases where certain variables did not affect churn as much, I had left them out. I used Tableau Desktop as my visualization and data manipulation tool. The Tableau workbook is attached to the accompanying email.

Total Number of Customers
65.53K

Average Tenure in Days
31

Churn Rate
57%

Insights:
1. The e-commerce company in question struggles to retain customers due to its high overall churn rate and low tenure across the board.
2. Over half the customers churn out and, on average, do not stay with the company for more than a month.



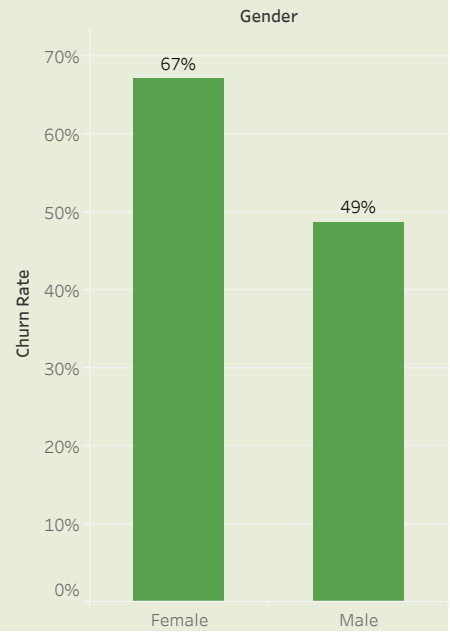
Insights

1. Each age group receives a different value proposition from the e-commerce platform, causing them to churn out at different rates.
2. Common reasons may be the absence of an intuitive user experience and lack of customer education for the platform.

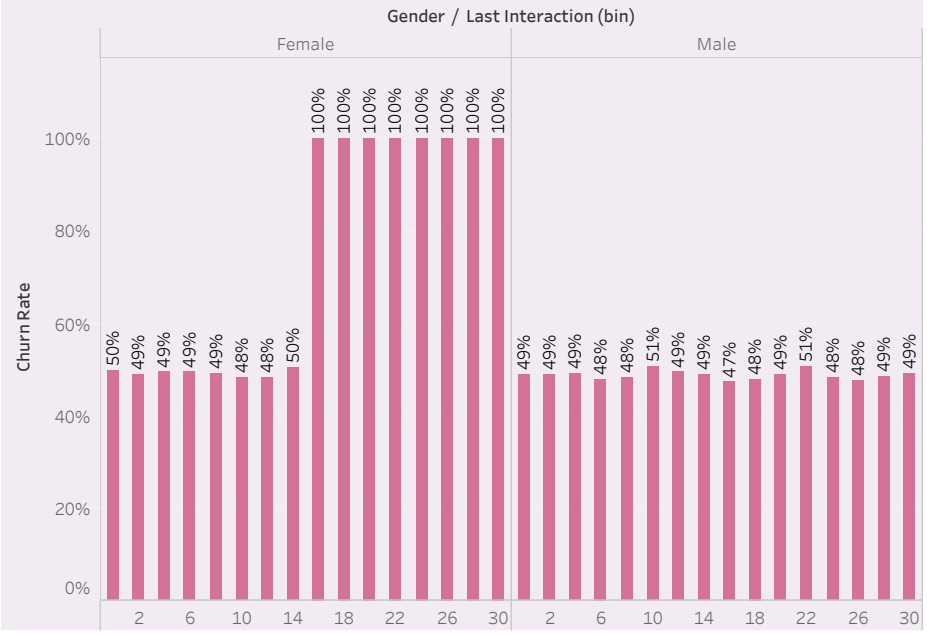
Solutions

1. Create advertisements and content aimed at customer education and value proposition to show what the user is missing out on.
2. Improve customer experience to be more responsive for older age groups, as that segment has a 100% churn rate.

Gender-Wise Churn %



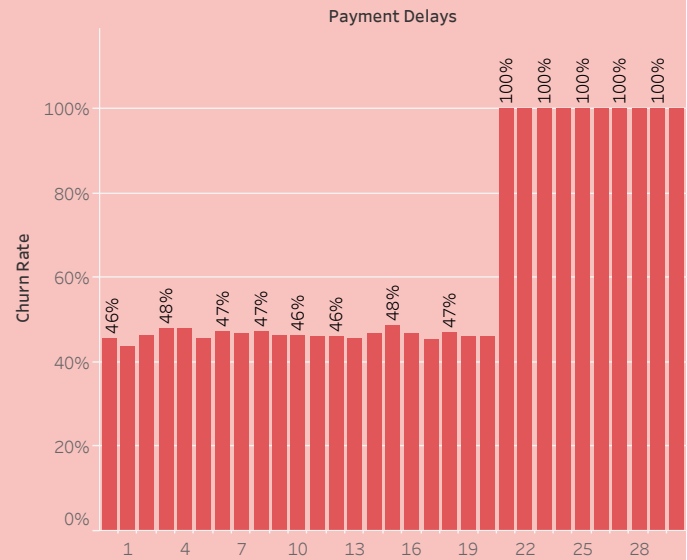
Last Interaction Churn %



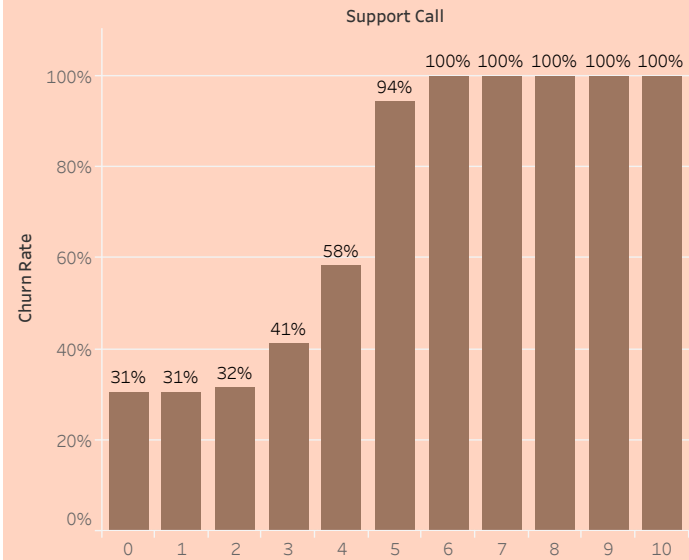
Insights:
Women are more likely to churn out from the platform than men, especially when they reach 15 days since their last usage, regardless of the age groups.

Solutions:
1. CRM pushes can be sent and churn-specific incentives can be created to bring these customers back to the platform.
2. For high value customers an attempt could be made to find out qualitative reasons for their churn through surveys and phone calls.

Payment Delay-wise Churn %



Support Call Churn % Distribution

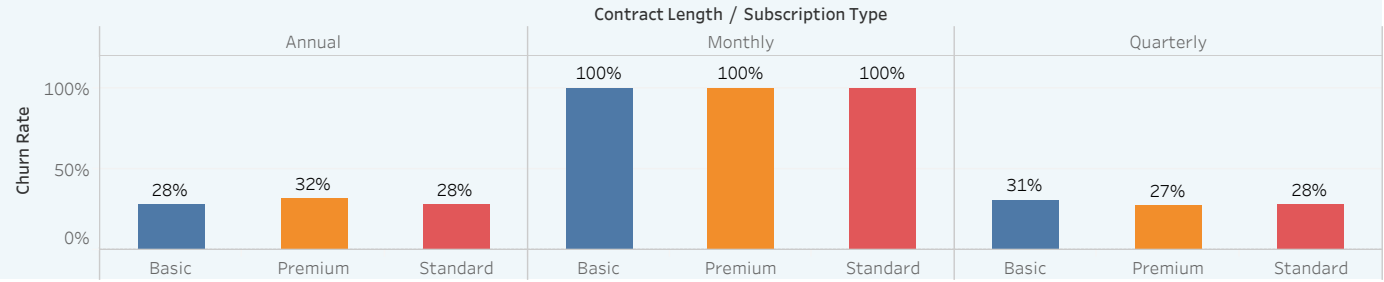


Insights

Delayed payments and lack of sufficient responses from support calls cause customers to unwillingly churn out. Since the platform is not always on the customer's mind, they require constant reminders from to interact with the customer.

Solutions
1. Aiming CRM or push notifications to remind customers delaying their payments close to 21 days.
2. Acquire feedback and trouble shoot problems of those customers approaching their 4th support call.

Subscription Type & Contract Length Churn Rate



Average Monthly Contract User Support Calls
5

Average Platform Support Calls
3.6

Average Retained User Support Calls
1.6

Insights:
1. Every customer who has subscribed to the monthly contract has always churned out.
2. The inefficacy of the monthly contract can also be seen in its churned customers placing a higher average number of support calls than both the platform and retained customers average.

Possible Solution:
1. Revise the constructs and pricing of the monthly contract.
2. Target customers churning out from the monthly contracted subscriptions with an initial reduced price offer and a personal call for improved customer support.
3. For higher value customers try to dig deeper as to why they churned and find qualitative reasons.

Recommended Strategy

We have looked at multiple reasons for churn and their possible solutions. Based on the available dataset and the problem statement, a simple RFM methodology can be created to address customer churn by placing the customers based on their

1. Recency - Last Interaction
2. Frequency - Usage Frequency
3. Monetary Value Score - Total Spend

I created another copy of the dataset, named "RFM Analysis.xlxs", to bucket each customer on a simple segment based on an RFM method.

CustomerID	Age	Tenure	Usage Frequency	Support Calls	Payment Delay	Total Spend	Last Interaction	Churn	Recency Score	Monetary Value Score	Frequency Score	RFM Score	Segmented Score	Customer Segment
1	64	3	25	2	11	415	29	1	1	2	7	10	1	Immediate Attention
2	27	52	8	7	3	434	19	1	4	2	2	8	0	Immediate Attention
3	59	26	21	0	10	822	17	1	5	7	6	18	7	Loyal Customer
4	32	6	22	3	12	413	20	1	4	2	6	12	2	At Risk/Need Attention
5	23	26	21	7	24	988	20	1	4	9	6	19	7	Loyal Customer
6	42	7	9	1	11	196	18	1	4	0	2	6	0	Immediate Attention
7	36	2	15	3	29	501	4	1	9	2	4	15	4	At Risk/Need Attention
8	22	43	16	8	21	966	30	1	1	9	4	14	4	At Risk/Need Attention
9	32	3	5	1	6	282	20	1	4	1	1	6	0	Immediate Attention
10	44	43	1	8	28	548	28	1	1	3	0	4	0	Immediate Attention
11	45	25	4	8	13	681	26	1	2	5	0	7	0	Immediate Attention
12	55	54	2	5	0	159	29	1	1	0	0	1	0	Immediate Attention
13	55	34	6	2	7	392	5	1	9	1	1	11	2	At Risk/Need Attention
14	28	33	14	7	29	856	6	1	9	7	4	20	8	Top Customer
15	60	53	23	1	10	917	12	1	6	8	7	21	8	Top Customer
16	18	46	30	0	5	269	10	1	7	1	9	17	6	Loyal Customer
17	58	32	6	4	5	684	28	1	1	5	1	7	0	Immediate Attention
18	51	2	6	1	23	959	23	1	3	9	1	13	3	At Risk/Need Attention
19	49	18	18	8	20	517	1	1	10	2	5	17	6	Loyal Customer
20	63	35	13	5	9	947	10	1	7	9	3	19	7	Loyal Customer
21	53	51	4	9	1	325	18	1	4	1	0	5	0	Immediate Attention
22	54	54	4	4	15	855	27	1	2	7	0	9	1	Immediate Attention
23	61	49	27	0	2	937	25	1	2	9	8	19	7	Loyal Customer
24	43	49	22	8	29	261	11	1	7	1	6	14	4	At Risk/Need Attention
25	43	59	8	9	27	535	22	1	3	3	2	8	0	Immediate Attention
26	34	19	17	1	5	201	1	1	10	0	5	15	4	At Risk/Need Attention
27	55	12	13	1	7	167	11	1	7	0	3	10	1	Immediate Attention
28	48	2	30	1	19	408	8	1	8	2	9	19	7	Loyal Customer
29	63	48	10	3	9	448	18	1	4	2	2	8	0	Immediate Attention
30	54	38	7	7	14	715	8	1	8	5	1	14	4	At Risk/Need Attention
31	59	24	24	2	3	516	25	1	2	2	7	11	2	At Risk/Need Attention
32	56	15	18	3	11	471	12	1	6	2	5	13	3	At Risk/Need Attention
33	52	56	14	10	29	719	4	1	9	5	4	18	7	Loyal Customer
34	44	2	27	7	12	936	3	1	10	9	8	27	9	Top Customer
35	31	46	13	3	20	211	7	1	8	0	3	11	2	At Risk/Need Attention
36	54	33	19	1	19	802	6	1	9	7	5	21	8	Top Customer
37	44	50	19	5	28	618	24	1	3	4	5	12	2	At Risk/Need Attention
38	29	56	12	9	25	144	19	1	4	0	3	7	0	Immediate Attention
39	58	29	5	5	26	262	2	1	10	1	1	12	2	At Risk/Need Attention
40	32	9	12	4	30	936	23	1	3	9	3	15	4	At Risk/Need Attention
41	35	57	11	4	21	532	12	1	6	3	3	12	2	At Risk/Need Attention
42	54	3	17	4	27	401	21	1	3	1	5	9	1	Immediate Attention

The recency, frequency and monetary value scores were given based on the percentile rank and then converted to 10. The "Segmented Score" in column N is the combined RFM score out of 10. Based on that, each customer was profiled according to the chart below. (Any method found online can be used). Given the simple nature of the data, I opted for the following.

RFM Score	Customer Segment
10	Top Customer
9	Top Customer
8	Top Customer
7	Loyal Customer
6	Loyal Customer
5	Loyal Customer
4	At Risk/Need Attention
3	At Risk/Need Attention
2	At Risk/Need Attention
1	Immediate Attention
0	Immediate Attention

From this, I generated a small summary, as shown below. Top customers tend to pay more and stay longer than other segments. However, the At Risk/Need Attention group has historically spent more, possibly coming to the platform only during promotions or incentivised campaigns, which show tell-tale signs of being an unprofitable segment.

Row Labels	Number of Customers	Total Spends	Average Spends	Average of Usage	Average of Last Interaction
At Risk/Need Attention	22712	13,641,842	601	14	16
Immediate Attention	14853	6,368,749	429	9	21
Loyal Customer	16890	12,171,090	721	19	12
Top Customer	11075	9,230,845	833	23	7
Grand Total	**65530**	**41,412,526**	**632**	**16**	**15**

With the overall distribution of the data now under the RFM method, we can start devising more detailed plans and campaigns, along with revising our CRM to retain and increase the profitability of each segment.