2020

Data Analysis Report

SUMBITTED BY:

MUHAMMAD JEHANZEB MIRZA

Contents

1. Revenue Analysis

The first analysis which we have done is with respect to the revenue which has been logged for different timestamps. We will share some visuals which have been created by Power BI and by using Python's Matplotlib and then we will comment on the most general trends which we could find by those visuals.

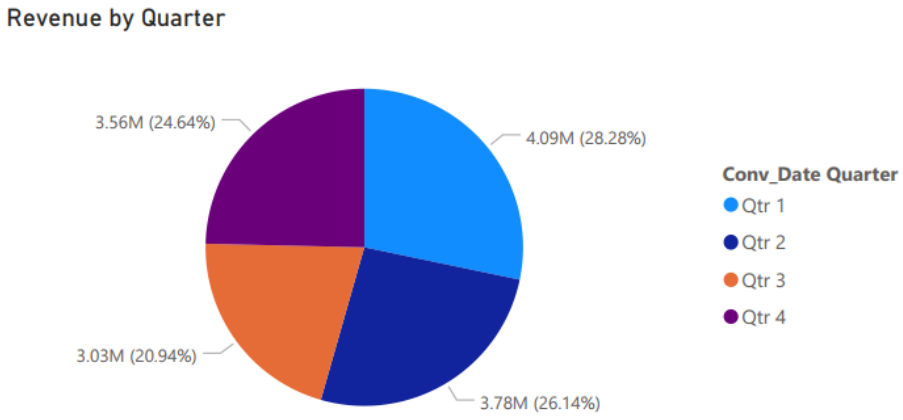

Fig 1: Revenue by Quarter

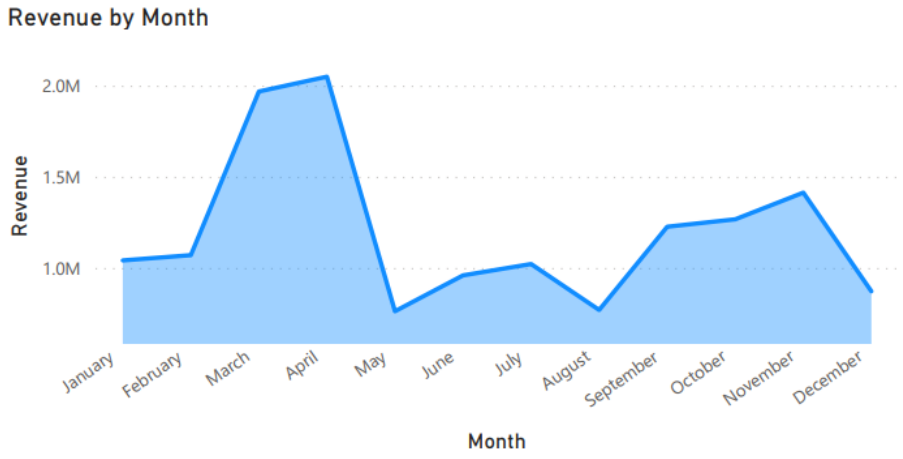

Fig 2: Revenue by Month

Fig 1 and Fig 2 give us a visual representation on how the revenue is changing with each passing month and tells us an important insight about how the revenue is fluctuating with each quarter for the company. The overall trend for the revenue generation is indeed on the downside for the company.

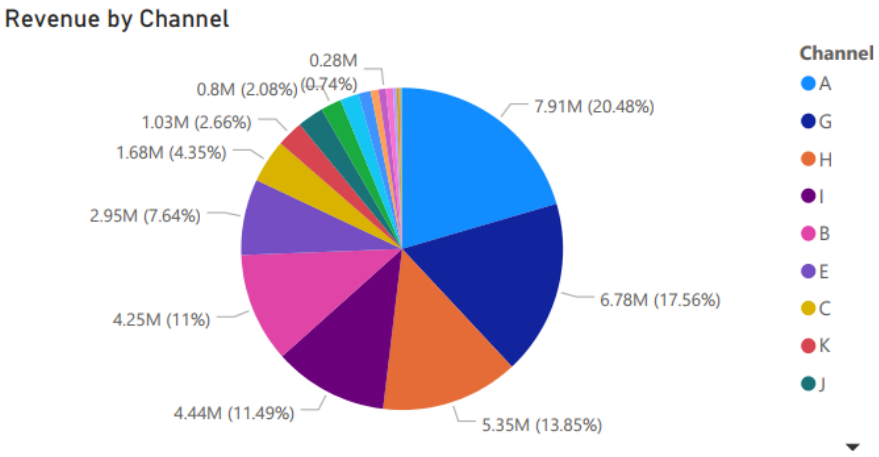

Fig 3: Revenue by Channels

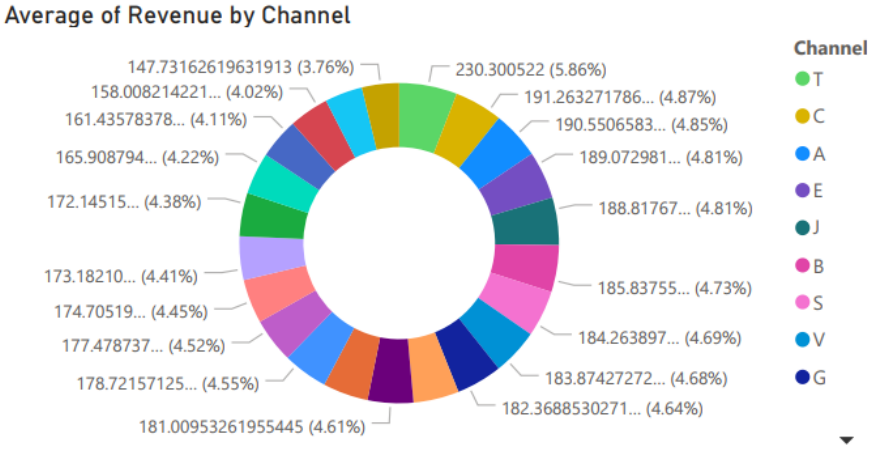

Fig 4: Average Revenue by Channels

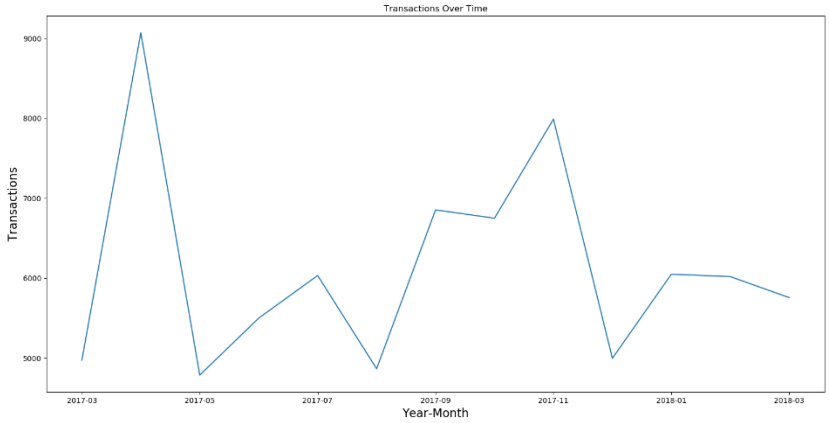

Fig 5: Transaction Over Time

The Fig 3 and Fig 4 give us an insight on the revenue generation w.r.t the different channels which have been employed for marketing. It can give the company a potential look into what kind of marketing strategies to use in the future. The Fig: 5 tells us about the fluctuations of the number of transactions which have been performed over the period for which the data has been provided. An interesting thing to see is how the trend for the number of transactions is relating to the drop-in revenue as we move further into the timeline.

2. Customer Analysis

For analysis of different customers, we have divided the customers in to three main categories:

1. Total Customers
2. Unique Customers
3. Returning Customers

How we are extracting these three customers can be seen in detail in the code. However, we can share our finding through some visuals here for discussing these three types of customers.

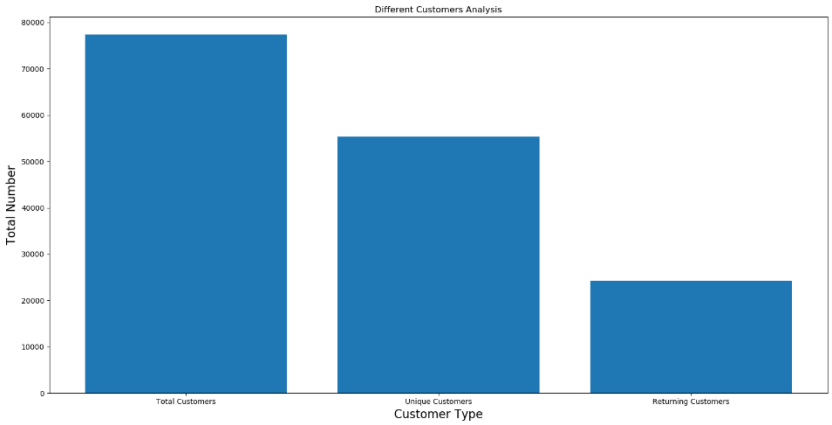

Fig 6: Customer Segmentation

The Fig 6 gives us the number of these three types of customers. If we run the code, we can see that the percentage of returning customers is around 30%. This gives us a very important insight that around 1/3rd of the customers are coming back again to buy again from the company which is a good sign.

Similarly, in Fig 7 we can see the number of customers over time. The plot contains the total and unique customers over time. If we relate it to our revenue curve in Fig 2 then we can say that as the customers over time is also decreasing as the revenue is decreasing. The difference between the graphs in Fig 7 give us the returning customer percentage over that period. It can also be argued that increase in revenue is also associated with the increase in number of returning customers which is a very important insight from these graphs.

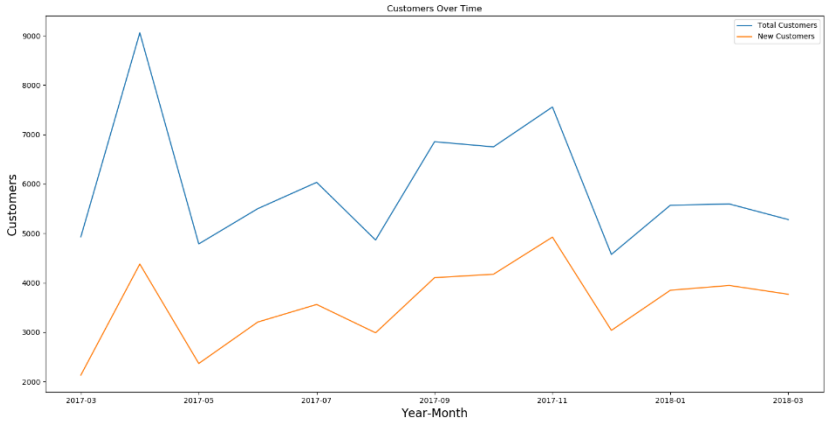

Fig 7: New and Total Customers over time

3. Channel Analysis

This is a very important analysis in which we have seen how different marketing channels a direct effect on the revenue and which kind of channels are most beneficial for the company. We have already related the revenue with the channels in Fig 3 and Fig 4. We can further probe into the importance of the channels used by the company by discussing the following graphic.

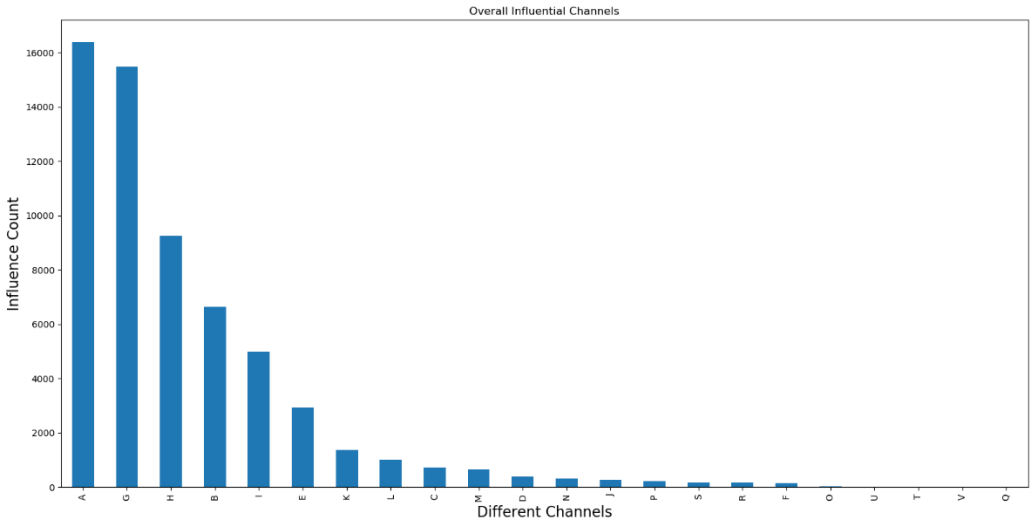

Fig 8: Influential Channels

We can see from Fig 8 that the most important channel which can be used by the company is channel A. This gives us some sort of customer segmentation as well. Customers are responding more towards channel A than any other channel. Further analysis about finding the most important channel for individual user can be found in the python implementation of the project. However, a

sample of the analysis is shown in Fig 9. It tells us the most influential channels for a group of 10 users. For finding influential channels for all the customers we would recommend having a look at the python implementation of the project.

User_ID	Channel	IHC_Conv
5094298f068196c5349d43847de5afc9125cf989	H	1.0
59853624ad85789cc5c865405e741c612da334da	I	1.0
8b14ce0002a2a7adbe61c6386dd5b8b8a8435222	A	1.0
a1557f0cc6b413c971c2f41087bb6111660f5e1f	A	1.0
c9a125ddacc4a6e6f6cb2c7e20f16270c464bf10	A	0.5753628000000001
8878878228b4a06fa4badf7ce1d255828538ab00	G	1.0
ad5edc88eb2e8b4c2dd2d33ccf13064ba2402b79	H	0.6999893
90ca7c786279258de0ae3933f72a0e4eda2426a2	H	1.0
e3f89ca72b7283dd3a9291e0d24b9630030146f4	A	1.0
31d668ddcd0aca8a8771b6e1e74ea0dafeca6b3c	H	1.0

Fig 9: Influential Channels for Individual Users

4. Cohort Analysis

One of the most important analysis in this project has been the cohort analysis of the data. Even though the details can be found in the python implementation of the project. Here we can try to visualize the Monthly Cohort analysis by adding some visualizations.

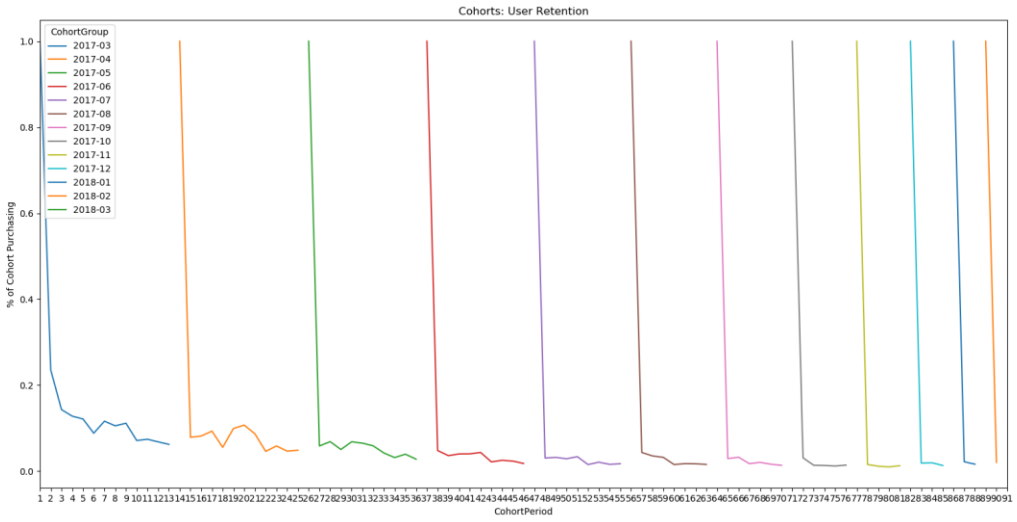

Fig 10: User Retention Cohort

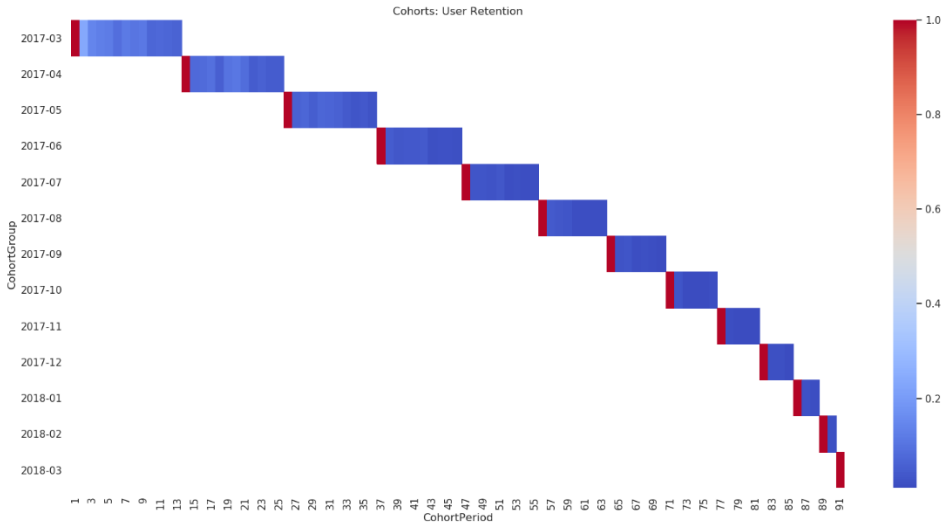

Fig 11: Heat Map for User Retention Cohort

The main goal behind this analysis was to find that how the company is retaining users/customers over time. It is one of the most important strategic goals of the company to retain its customers over a period. This analysis divides the users into cohort groups depending on the first purchase date of the user and then tries to track that how successfully the company has been able to hold these customers. We can see from Fig 10 and Fig 11 that the company has been lacking the ability to retain the customers over a huge period. As the time goes on, the company has the tendency to lose its returning customers. This is something the company can look in to going forward.