

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | |
|  | | Customer Churn Analysis | | | | |  | |
|  |  | | | | | | |  |
|  | | | |  |  | | | |
|  | | | | Bhathiya Bandara239147H |  | | | |
|  | | | | 21-3-2023Data Science— |  | | | |
|  | | |  | | |  | | |



|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | | |  |  | | |  |
|  | INTRODUCTION | | | | | | |  |
|  |  | | |  |  | | |  |
|  |  | |  | | |  | |  |
|  |  |  | Customer churn analysis is a critical analysis of many businesses who aimed at retaining customers and improving profitability. Churn, which refers to the rate at which customers stop doing business with a company, is a common problem faced by organizations across all industries. In today's competitive business landscape, understanding why customers churn and taking proactive steps to reduce churn can make a significant impact on a company's profitability.  In this analysis, we will examine the factors that contribute to customer churn in telecommunication industry, explore methods for predicting churn, and identify strategies for retaining customers and increasing customer loyalty. | | |  |  |  |
|  | | | | |

# Data Preparation

As data transformations, I created 2 new attributes from the existing attributes.

1. **Customer Age Group**: I have categorized each subscriber based on their age decade (Eg: if age is 22 years, then 20s group. If age is 34, then 30s group.)
2. **Monthly Charge Group**: It is easy to identify the user groups if we categorize their behavior based on the amount they spent monthly.

*Figure 1* shows the spearman correlation between all the metric variables in the dataset. We can see that “Revenue” and charges related columns have high positive correlation. This is obvious because revenue is proportional to charges. Therefore in some analysis, we will reduce some attributes since they have a high correlation between them.

there is significant correlation between the “Age” and the “Avg Monthly GB Download” attributes.

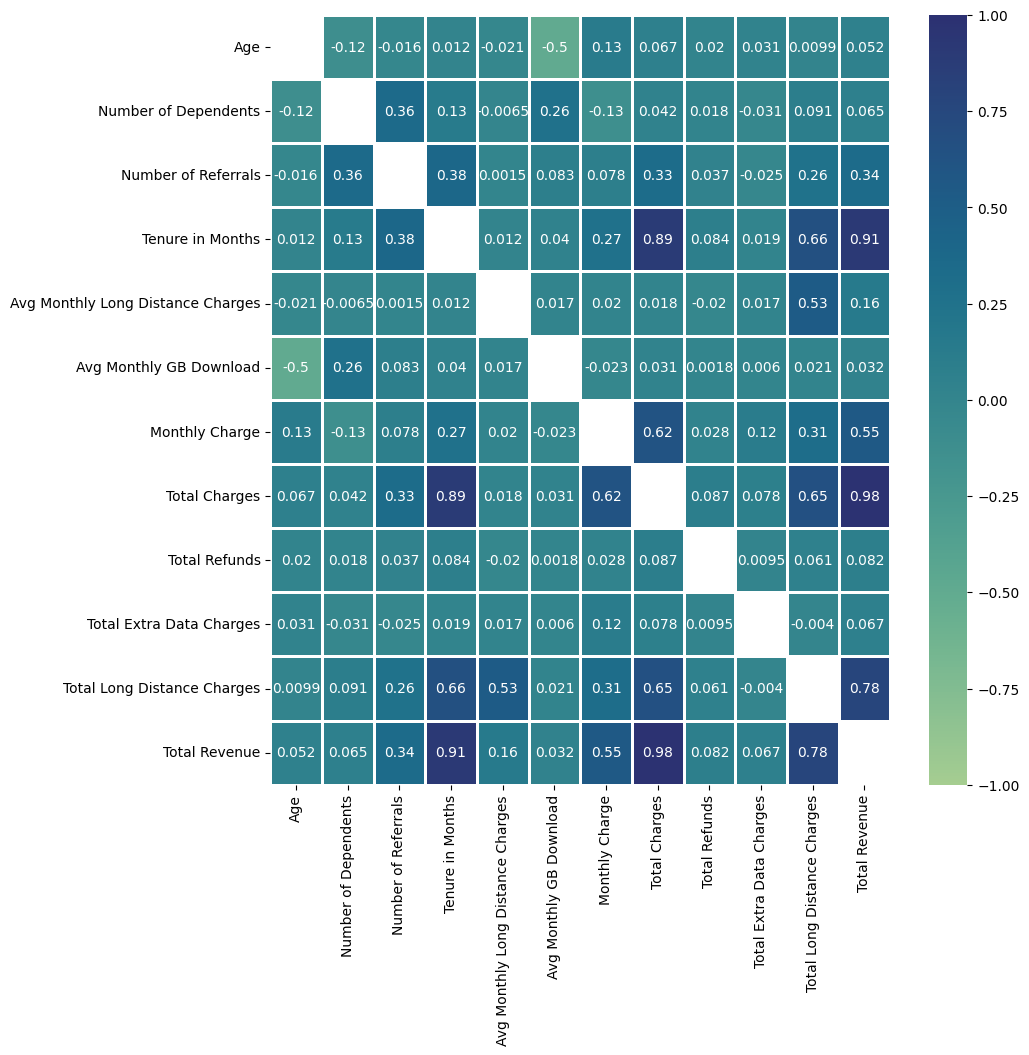


Figure : spearman correlation

# Data Analysis

As mentioned in the previous chapter, we can see a strong correlation between “Age” and “Avg Monthly GB Download”. Before analyzing the churned users, lets check this behavior.  *Figure 2* shows the variation between “Age” and “Avg Monthly GB Download”. When the age is below 30, users are tend to use more data compared to the other users. In comparison users whose age under 30 use **55.62 GB** while users with age above 30 use only **18.4 GB**. As a percentage this is **66.92%** drop in data usage. *Figure 3* shows the age wise data usage boxplot for **churned** users and **stayed** users. Both category users show similar trend. Therefore, this behavior is not cause to churn but rather than this is due to the user’s usage behavior. Even though this is not related to churn, telecom company can further analyze the cause for this behavior so that they can providing solutions or introduce new marketing strategies among users whose age is above 30 to increase the data usage.

Chart, histogram

Description automatically generatedChart, box and whisker chart

Description automatically generated

*Figure 2: Bar plot Figure 3: Box plot*

Among customers in the dataset, **67.0 %** customers **stayed**, **6.4%** **joined** and **26.5% churned.**

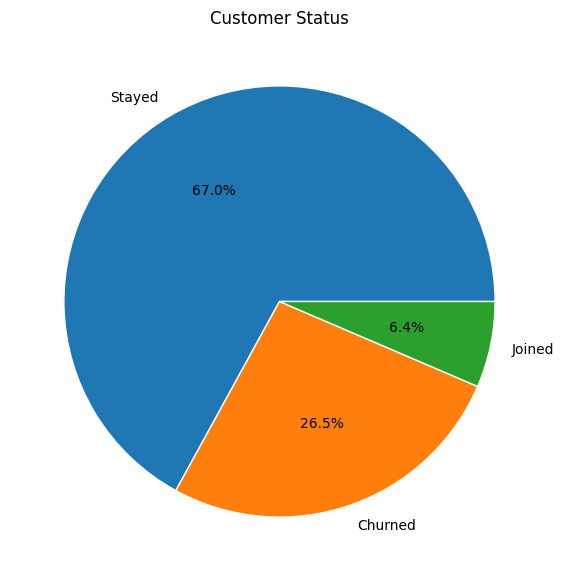


Figure 4: Customer Status

For the telecommunication company, revenue generated by the monthly charges of its subscribers. From the *Figure 5* we can see that, churned subscribers average monthly charge is higher than the currently staying subscribers.

Chart, box and whisker chart

Description automatically generated

* Average Monthly Charge for a Churned user = $73.3
* Average Monthly Charge for a Staying user = $61.7

Not only that but also, 50% of the churned users monthly charge is between $55 - $94, while staying customers 50% monthly charge is between $25 - $89.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Customer Status** | **count** | **mean** | **std** | **min** | **25%** | **50%** | **75%** | **max** |
| **Churned** | 1869 | 73.34 | 26.37 | -10 | 55.3 | 79.5 | 94.2 | 118.35 |
| **Joined** | 454 | 42.77 | 24.61 | -8 | 20.16 | 43.97 | 64.35 | 107.95 |
| **Stayed** | 4720 | 61.73 | 32.19 | -10 | 25.1 | 65.6 | 89.55 | 118.75 |

Figure 5:Customer Status vs Monthly Charge

Chart, line chart, histogram

Description automatically generated

Figure 6: Distribution

We can identify main two groups based on the **Monthly Charge** for staying customers based on the *Figure 6* distribution plot (Blue line)

* First group average monthly charge is around $20 and their variation is small.
* Second group average monthly charge is around $80 and their variation is higher compared to first group.

As a summery, we can say that, company have 2 customer groups which one group is generating low revenue while other group is generating more revenue.

#### Churned users

Operator has categorized churn users based their feedback to 5 categories. 45% of the churns due to “Competitor”, 17.2% due to Dissatisfaction and 16.8% due to poor attitude. Apart from that, churned users tenure in months is small compared to staying users. Majority of the users churned during first few months after joining, while staying customer tenure is uniformly distributed.

Chart, histogram

Description automatically generated Chart, pie chart

Description automatically generated

*Figure 7: Tenure in months Figure 8: Churn Category*

Geological area based analyze: Following figures shows the geographical based churned users and staying users distribution.

Green dots represent the staying customers and Blue dots represents the churned users. Background color of the cities represents the city population. This help us to identify which cities have higher population and how many users are churned from that area.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Figure 9: Competitor | | Map  Description automatically generated  Figure 10: Dissatisfaction | | Map  Description automatically generated  Figure 11: Attitude | |
|  | Figure 1: Price | | Map  Description automatically generated  Figure 1: Other | |  |

By analyzing the maps, we can see that most of the churned users lived in high population areas. Most probably in city areas. Among five churn categories, Attitude category can see a clear difference in user distribution.

# Conclusion