

Capstone Project -1

Telecom Churn Analysis

By
V Bhavya Reddy

Introduction

Customer churn occurs when a customer switches from one service provider to its competitor.

Since the number of telecom service providers has increased substantially, companies are competing to generate revenue.

The three main methods of generating revenue are acquiring new customers, convincing existing customers to buy more plans, and increasing customer retention.

Surveys have shown that losing an existing customer is more costly than acquiring a new one. As a result, churn prediction is critical.

Problem Statement

The Churn dataset is provided by a French multinational telecommunications corporation, Orange. The Orange Telecom's Churn Dataset consists of 3333 rows and 20 columns.

Every row corresponds to a unique account and in each column, there is a unique attribute indicating several services, account information, churn, etc.

The main objective is to analyse the attributes, their dependence on customer churn, and ways to increase customer retention.

Attributes Summary

To analyze Telecom Churn dataset file, I believe the first step is to understand all the attributes defined in the file.

State: 51 US states as a two-letter abbreviation in which the account holder(customer) resides.

Account Length: The number of days the account has been active.

Area Code: The three-digit area code of the corresponding customer's phone number.

Attributes Summary

International plan: whether the account has an international calling plan - yes/no.

Voice mail plan: whether the account has a voice mail plan- yes/no.

Number vmail messages: Average number of voice mail messages per month.

Total day minutes: The average number of minutes spent on calls throughout the day.

Attributes Summary

Total day calls: The number of calls placed during the day.

Total day charge: The amount charged during the daytime calls.

Total eve minute: The average number of minutes spent on calls throughout the evening.

Total eve calls: The number of calls placed during the day.

Total eve charge: The amount charged during the evening time calls.

Attributes Summary

Total night calls: The number of calls placed during the night.

Total night charge: The amount charged during the night time calls.

Total night minute: The average number of minutes spent on calls throughout the night.

Total intl minutes: The average number of minutes spent on international calls.

Total intl call: The number of international calls placed.

Attributes Summary

Total intl charge: The amount charge for international calls.

Customer service calls: The number calls to customer service.

Churn: Whether the customer left the service - Yes/No.

Identifying the key factors for churn

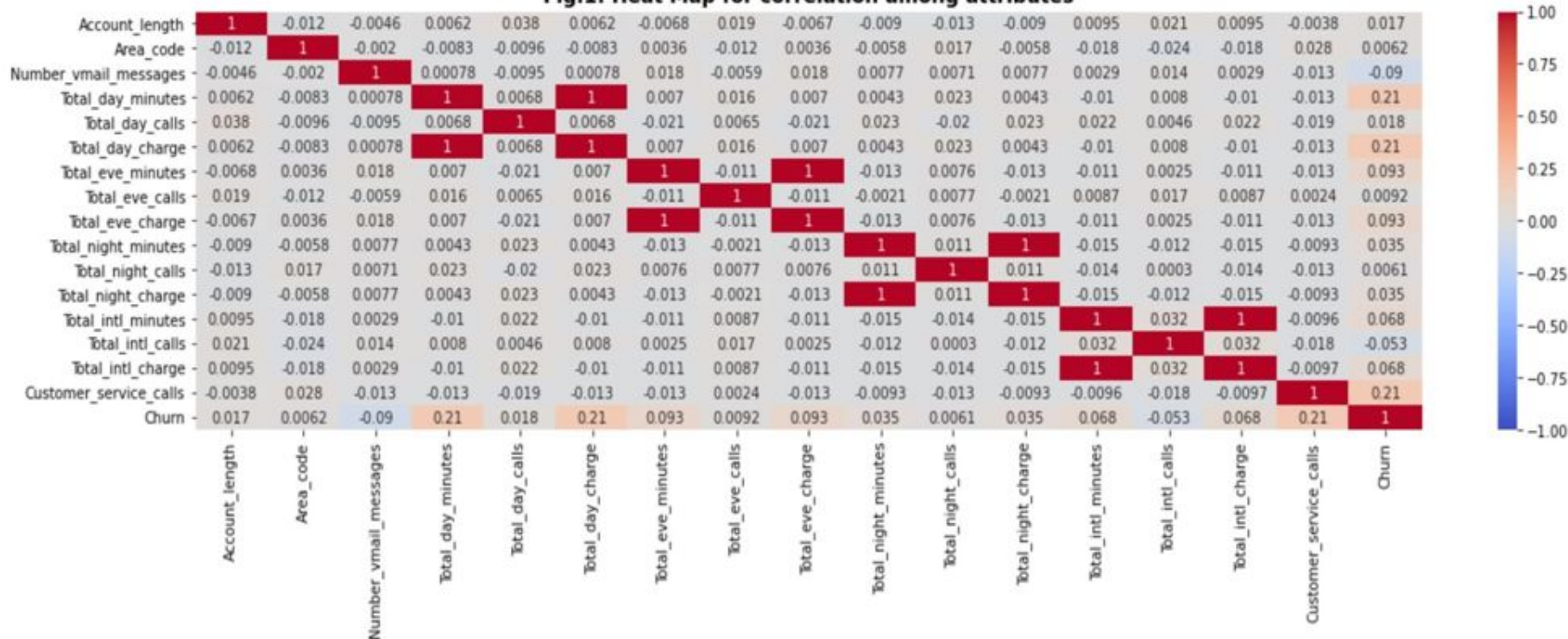
I believe the best way to identify the key attributes is by correlating the data and plotting the heatmap.

Many significant observations were made about the attributes:

1. There is a strong correlation between Total_day/night/eve/intl_minutes and Total_day/night/eve/intl_charge and it makes sense because we are charged by the length of our call.
2. Account_length, Area_code, Total_day/eve/night_calls are quite insignificant to the Churn.

Identifying the key factors for churn

Fig.1: Heat Map for correlation among attributes



Identifying the key factors for churn

3. Many columns are uncorrelated, which can be understandable through intuition, however, the important correlation should be noticed between the column "Churn" and various columns "Total_day_minutes", "Total_day_charge", "Total_evening_minutes", "Total_evening_charge", "Total_intl_minutes", and "Customer_service_calls", since correlation is significant.
4. Further analysis of these columns will be done using various visualization tools.

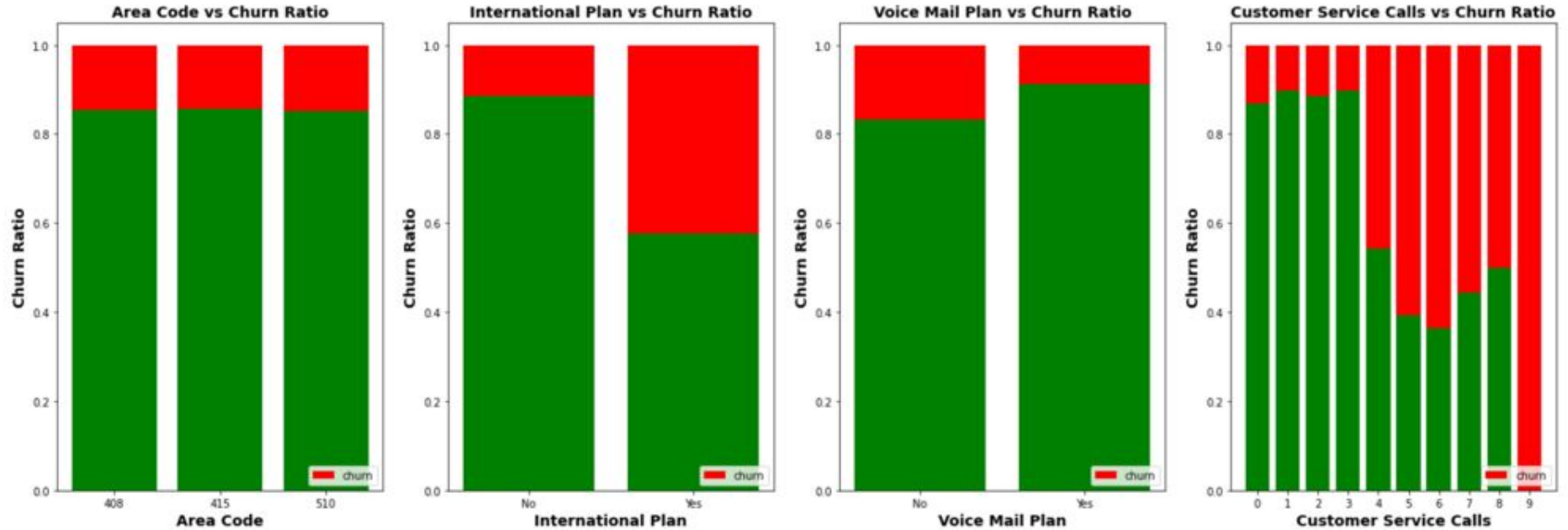
Identifying the key factors for churn

Since correlation only works on numerical data, I chose stacked bar graph for the non- numerical attributes.

The following observations were made:

1. It appears that the churn ratio for each area code does not vary from the overall churn ratio.
2. With an international plan, the churn ratio is 3 times larger.
3. Customers without voice mail plans seem to experience higher churn than average.
4. For customer service calls exceeding four, the churn rate is significantly higher.

Fig.2: Stacked bar graph for various attributes against Churn Ratio



Identifying the key factors for churn

Heatmap identifies that area code has no impact on churn. The stacked bar graph also demonstrates that area code has no impact on churn.

1. Fig.1(heatmap) already shows that area code has no impact on churn. Fig.2(Stacked bar graph) also demonstrates that area code is not affected by churn.
2. It is possible that accounts with international plans chose to cancel because the international plans might be costlier than the other operators.
3. Churn for voice mail plan could be an unintentional coincidence, but further exploration will yield a better understanding.
4. Those accounts with more than 4 customer service calls may have a problem with their services and chose to churn.

State-wise Analysis

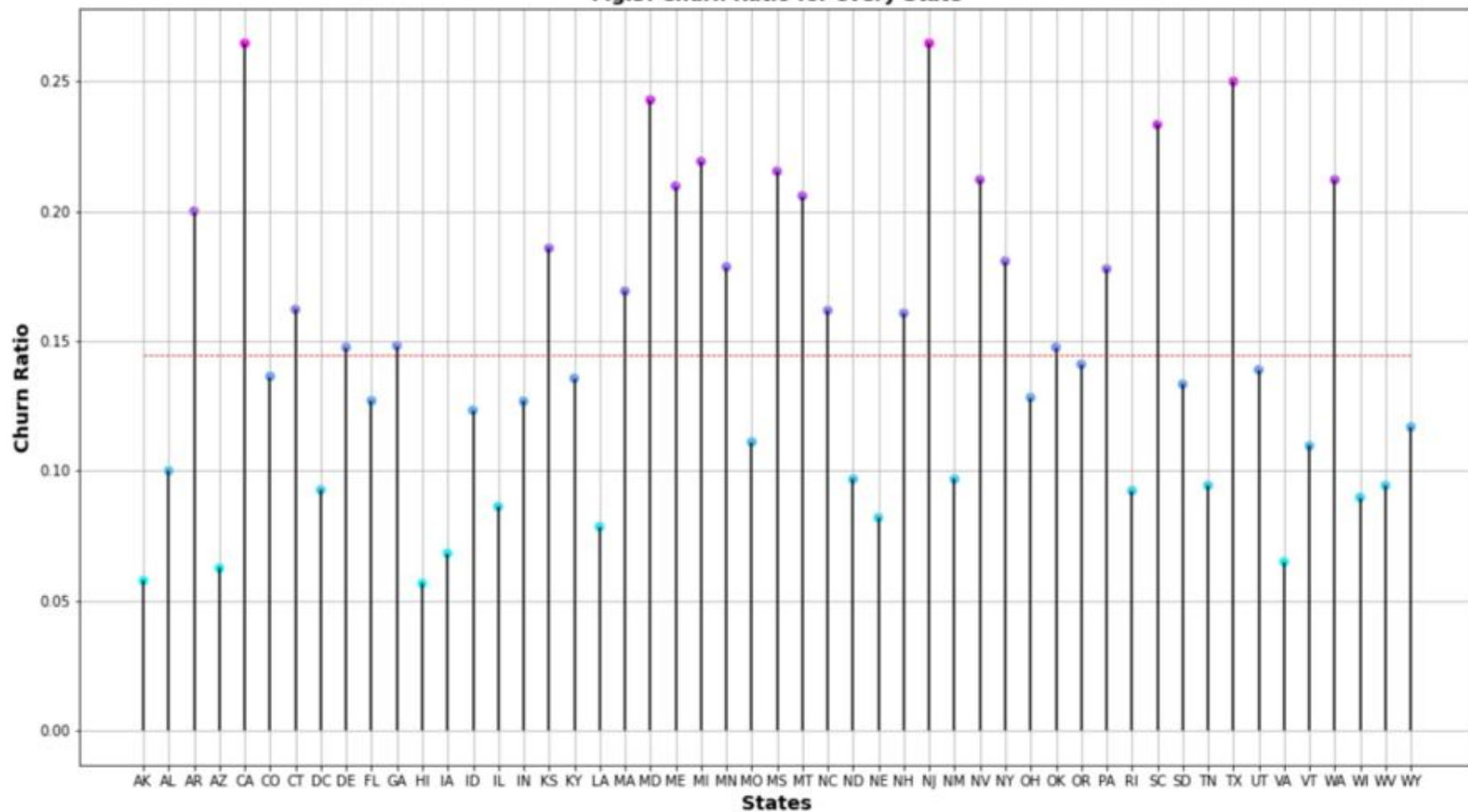
State vs Churn

Here, a line plot with mean churn ratio is laid on top of the state vs churn ratio scatter plot.

The red dotted line represents the average churn ratio. Of the 51 states, 23 have churn rates that are higher than average.

The 23 states must be offered a variety of incentives and plans by the company to retain their customers.

Fig.3: Churn Ratio for every State

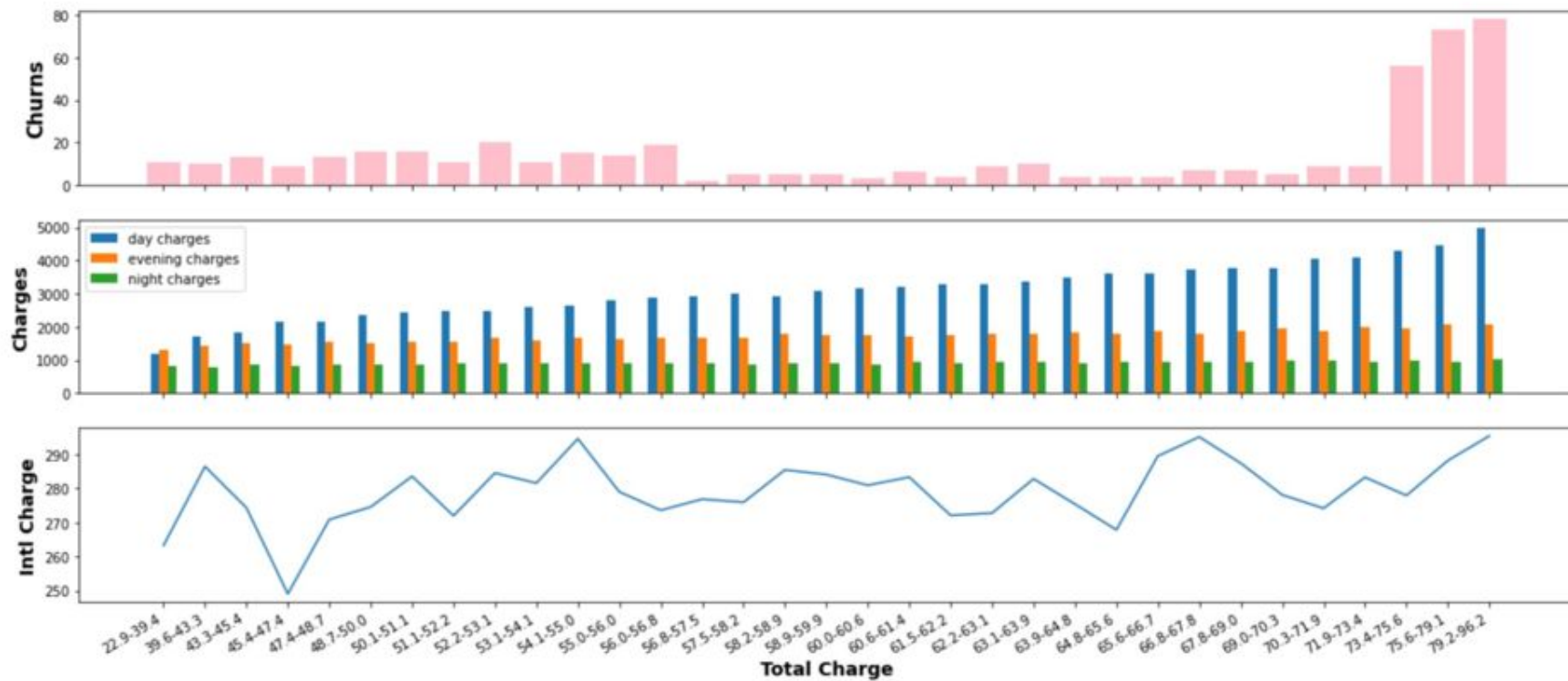


Effect of total charge on the churn

There is a high churn when the total charge exceeds 73 dollars, and it appears that the day charge is responsible for most of the increase.

Fig.1(heatmap) already illustrates that the total daily charge (or total daily minutes) on Churn is higher than any other charge. This can be verified in Fig.4. Having more total day minutes will increase the network's traffic, which may result call drops, and hence the churn. Another possibility is that the prices of other networks are lower for more call duration.

Fig.4: A detailed plot for various charges vs churn



International plan and Voice mail plan for Churned and Active accounts

1. Fig.5 clarifies Fig.2's ambiguity regarding higher churn for no Voice Mail Plan customers. Voice Mail plans account for 16.6% of churned accounts, while Voice Mail plans account for 28.6% of active accounts.
2. There are 4 times fewer active international plan accounts than churned international plan accounts.

Fig.5: Pie chart for IP and VP for Churned and Active accounts

Fig.5.1: Nested pie chart for International plan and Voice Mail Plan

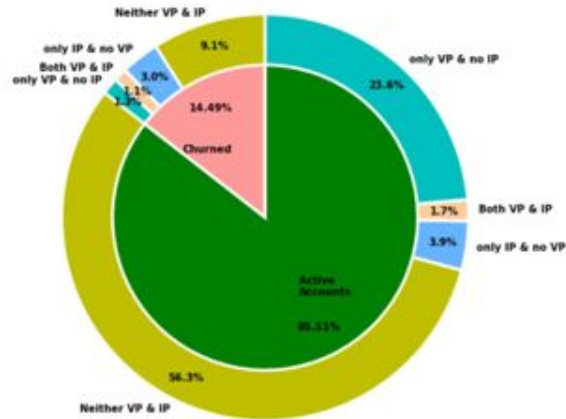


Fig.5.2: International plan and Voice Mail Plan for Churned Accounts

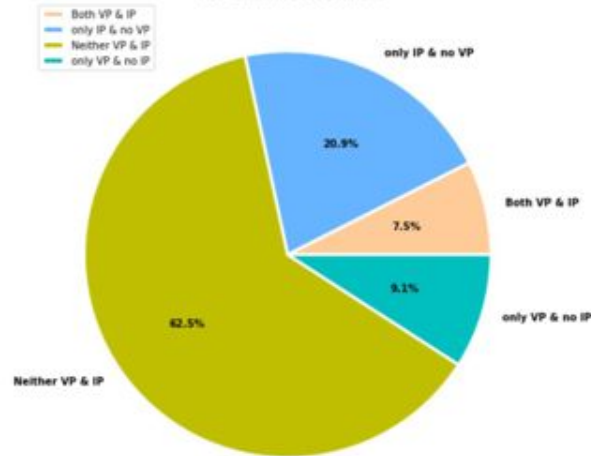
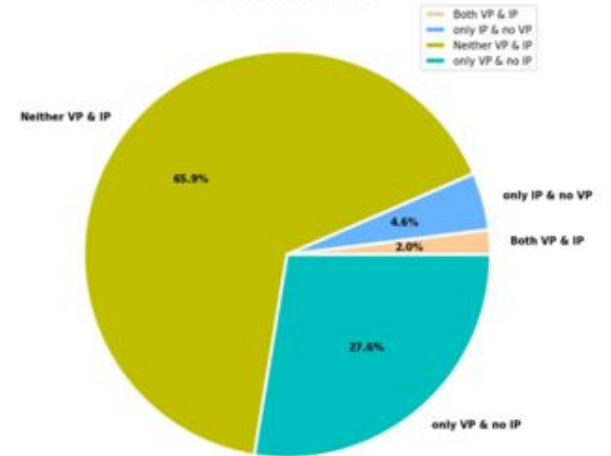


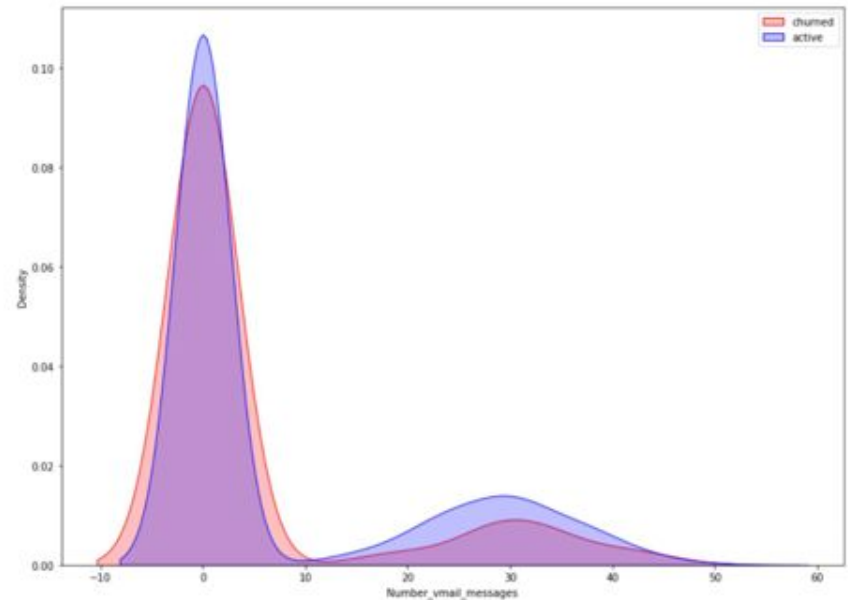
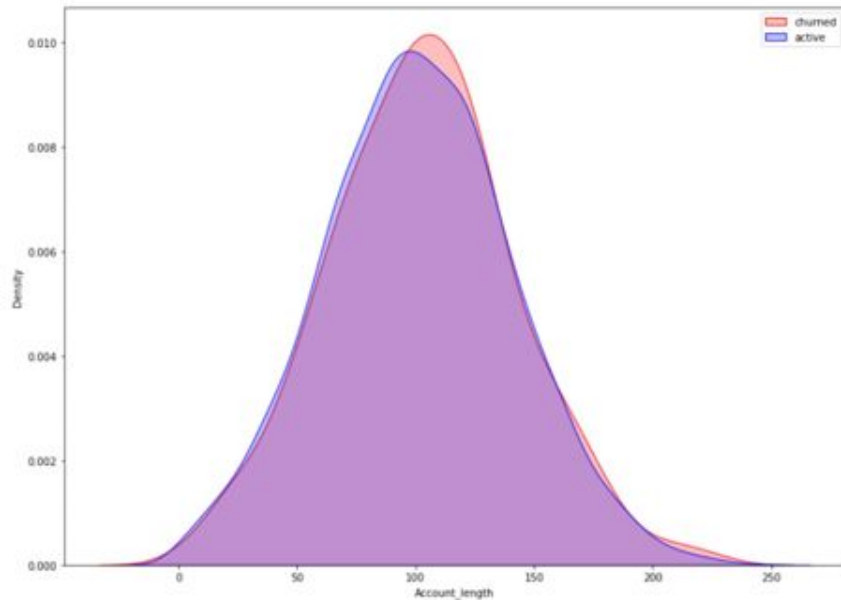
Fig.5.3: International plan and Voice Mail Plan for Active Accounts



Analyzing account length effect on churn

Account length, number of voice mail messages are indifferent w.r.t. churn throughout.

Fig.6: Account length, number of voice mail messages are indifferent w.r.t. churn



Way for Customer Retention

It is important to note that the loss percent is much higher than just 15.93%, and therefore customer retention is vital.

1. Surveys have shown that the cost efficient way to generate revenue is by increasing the customer's retention period.
2. A few ways to increase customer retention period- better price offers, lucrative incentives, better service experiences, etc.
3. The earlier analysis could also help the company to retain its customers by reducing the cost of international calls, improving service in a few areas, prioritizing customers who called customer service more than four times, eliminating call drops, etc.

Conclusion

1. One is charged based on the amount of time one spends on the call.
2. The attributes which are not contributing to the churn are -
 1. Area code
 2. Account length
 3. Voice Mail Plan
 4. Number of voice messages.
3. Either the international call services are poor, or the prices are higher than those of other operators.

Conclusion

4. It would make sense to offer lucrative incentives to customers who make more than four service calls, such as free limited data or OTT services.
5. Of the 51 states, 23 states have churn rates higher than the average. This requires more attention to understanding local customer needs.
6. The churn rate was highest when the total charge (day charges + night charges + evening charges + international charges) exceeded 73 dollars.
7. In the daytime, more calls are made than at night or in the evening, which means more traffic, which results in call drops, which ultimately leads to poor service and churn.