

Capstone Project - 1

EDA - Telecom Churn Analysis

Team members:

Aakash Patil

Akshay Arsude

Aparna Patil

Content

- Introduction
- Churn in telecommunication industry
- Problem definition
- EDA on given data
- Churn percentage of states (EDA)
- Relation between Churn and Others
- Conclusion
- References

Introduction

- Usage of mobile devices has increased by a tremendous amount. In a survey conducted by the Socio Economic council of United Nations, the world population was estimated to be about 7.7 billion users in 2019, with an almost equivalent number of cell phones.
- Several developing countries exhibit a penetration level of 97%–100%
- If one analyses the data from the telecommunications industry, the relationship between the number of non-churners and churners is high, which demonstrates an immense imbalance. Hence the churners corresponding data is a minority and the non-churners corresponding data is a plurality.[1]

Introduction

- Customer churn shifting from one service provider to the next competitor in the market, is a key challenge in highly competitive markets and is very much observed in telecommunication sector.
- Customer churns are those targeted customers who have decided to leave a service provider, product, or even a company and shifted to the other competitor in the market.
- If the churn rate of company is 50% then the company will be going to shut in two years. As company with churn rate 25% then it will shut in four years so Churn is the major factor to be taken in consideration.

Churn in telecommunication industry

- Currently organizations have their major focus on reducing the churn by focusing on customers independently.
- Churn can be defined as the propensity of a customer to cease business transactions with an organization.
- Churn occurs mainly due to customer dissatisfaction. Identifying customer dissatisfaction requires several parameters. A customer usually does not churn due to a single dissatisfaction scenario.
- There usually exist several dissatisfaction cases before a customer completely ceases to do transactions with an organization. [3]

Problem definition

As we know creating new customer is not easy in telecom industry so we have to focus on existing customers.

For Telecom churn analysis problem can be defined as:

- Customers Churn
 - In this we have to look into the problems why customers are leaving.
 - For different states churn rate is different.
 - Dependent variable is Churn in the data set.
 - Finally we have to look for Customer satisfaction.

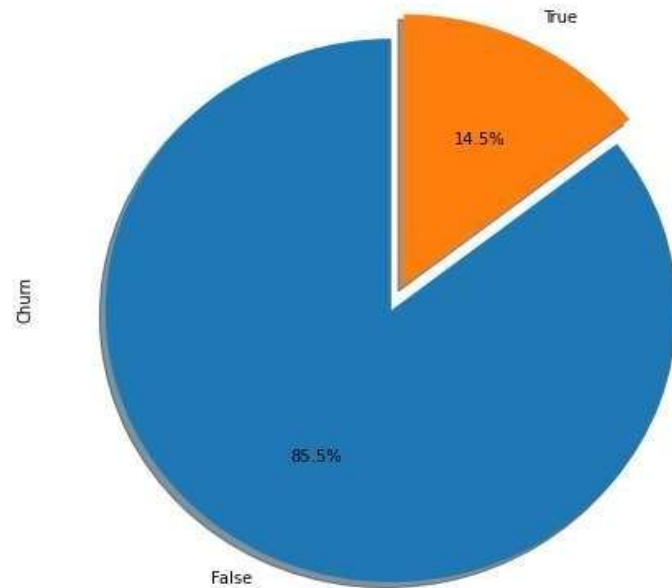
EDA on given data

Digging into data we understand that

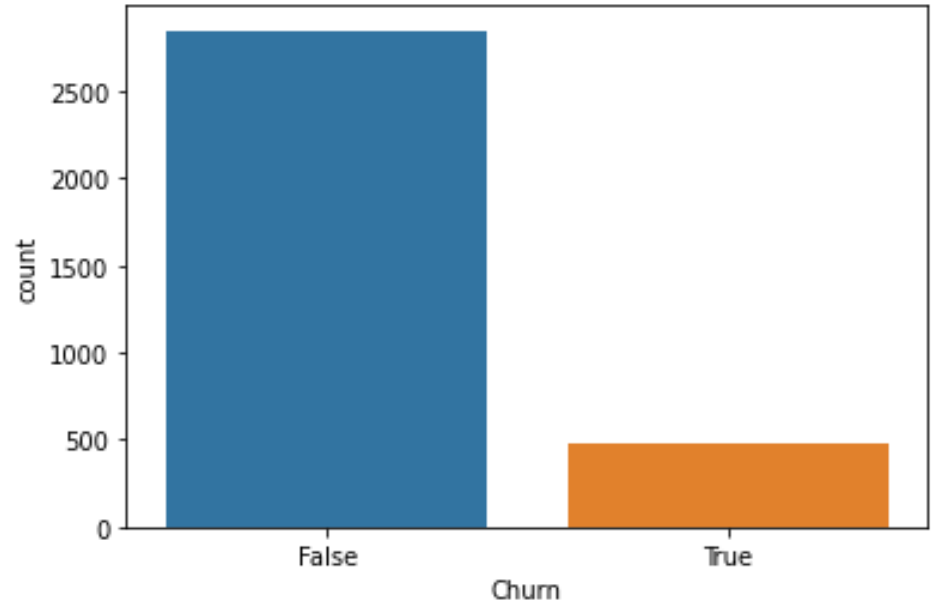
- There is no null value in the data set.
- In state column there are total 51 unique states.
- Total 20 columns with values such as float, integer, Boolean and object data type.
- Dependent variable should be considered as Churn.
- Graphical representation according to various columns and with manipulation of columns.

Churn percentage of states

Pie Chart for Churn

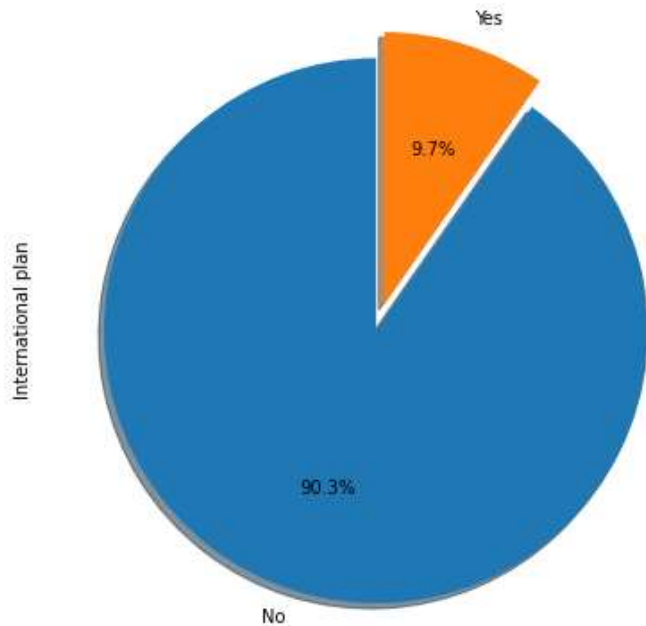


Count Plot for Churn

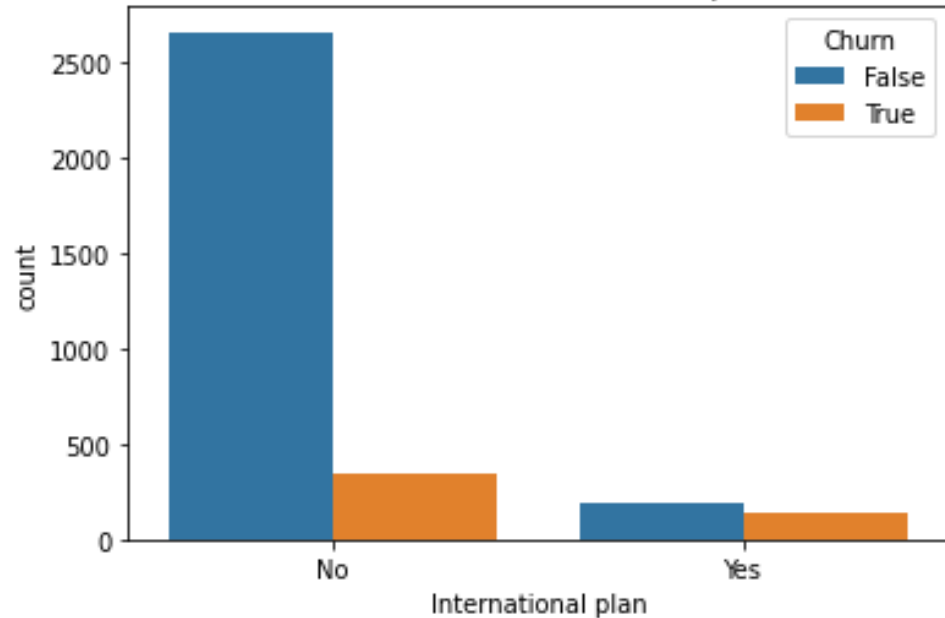


International Plan

Pie Chart for International plan

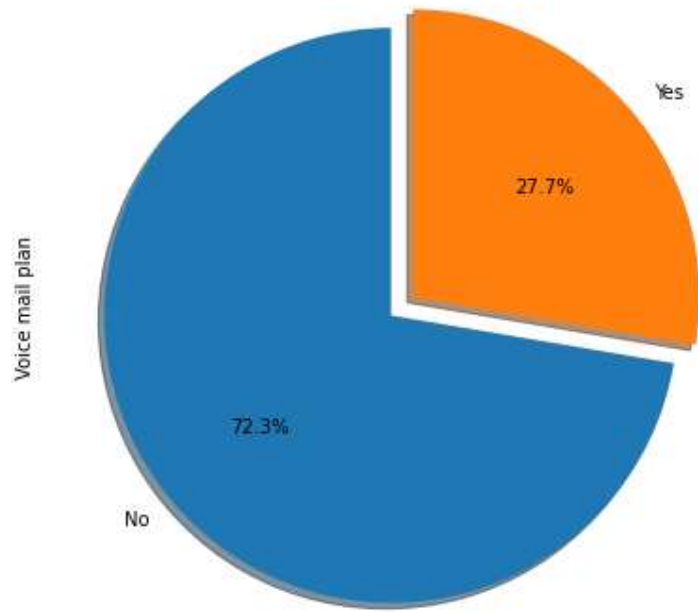


Count Plot for International plan

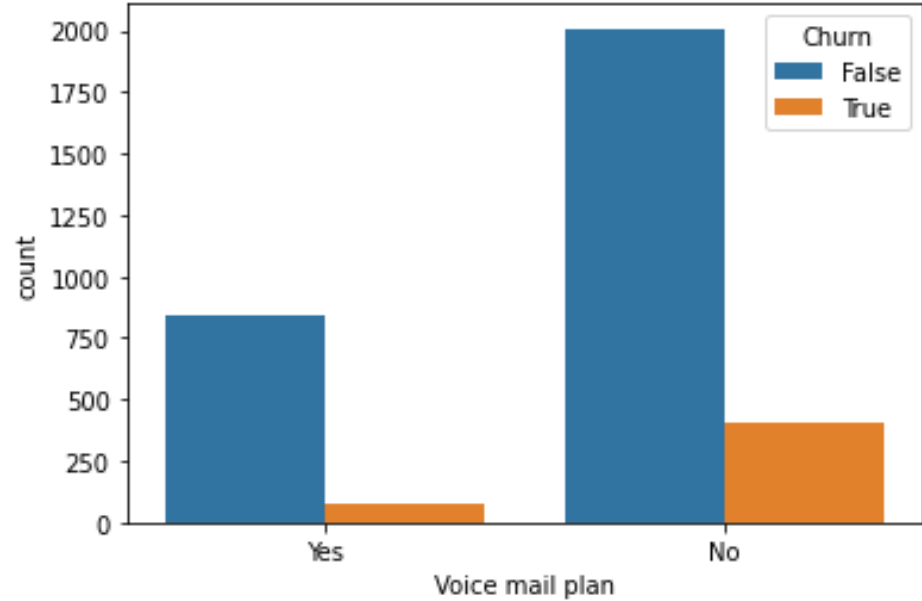


Voice Mail Plan

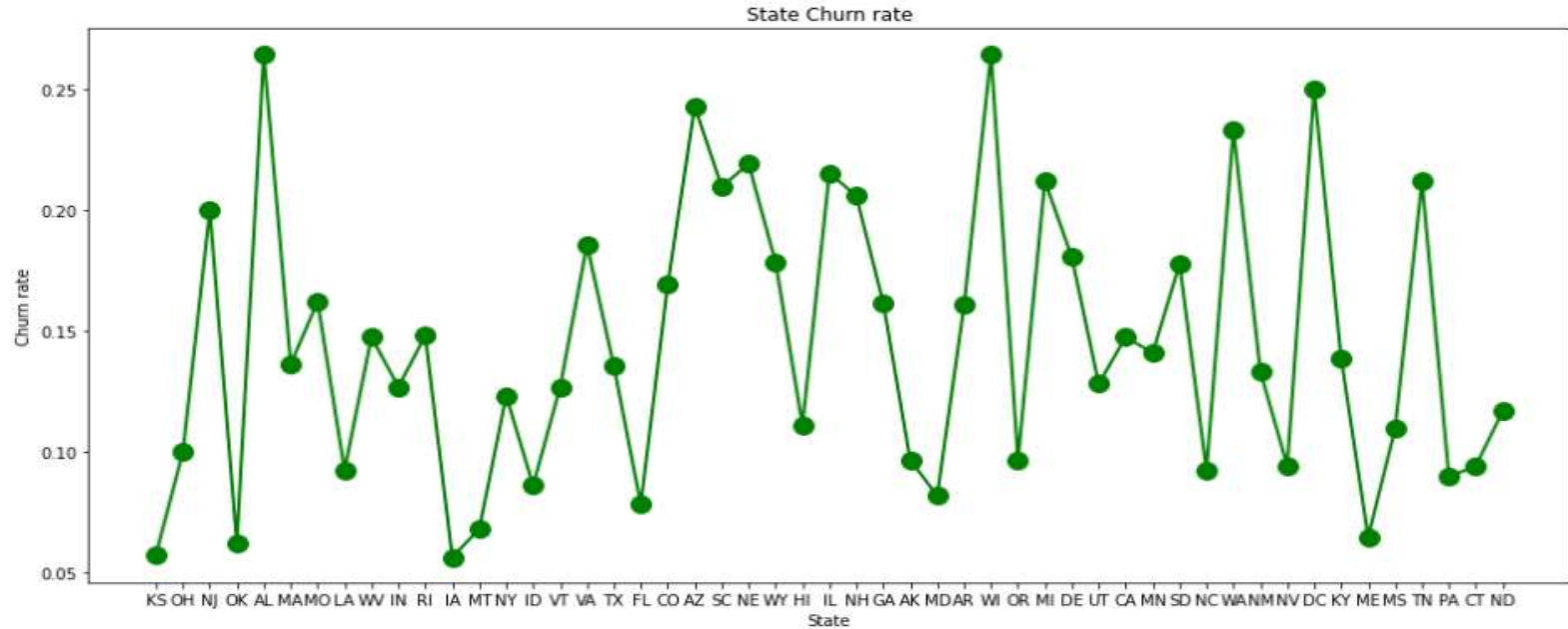
Pie Chart for Voice mail plan



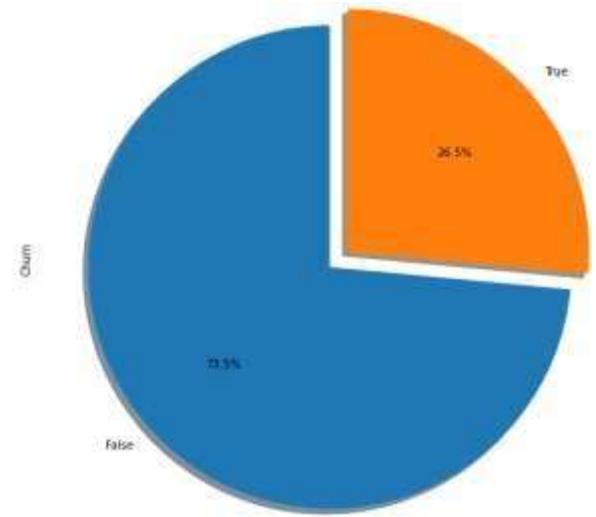
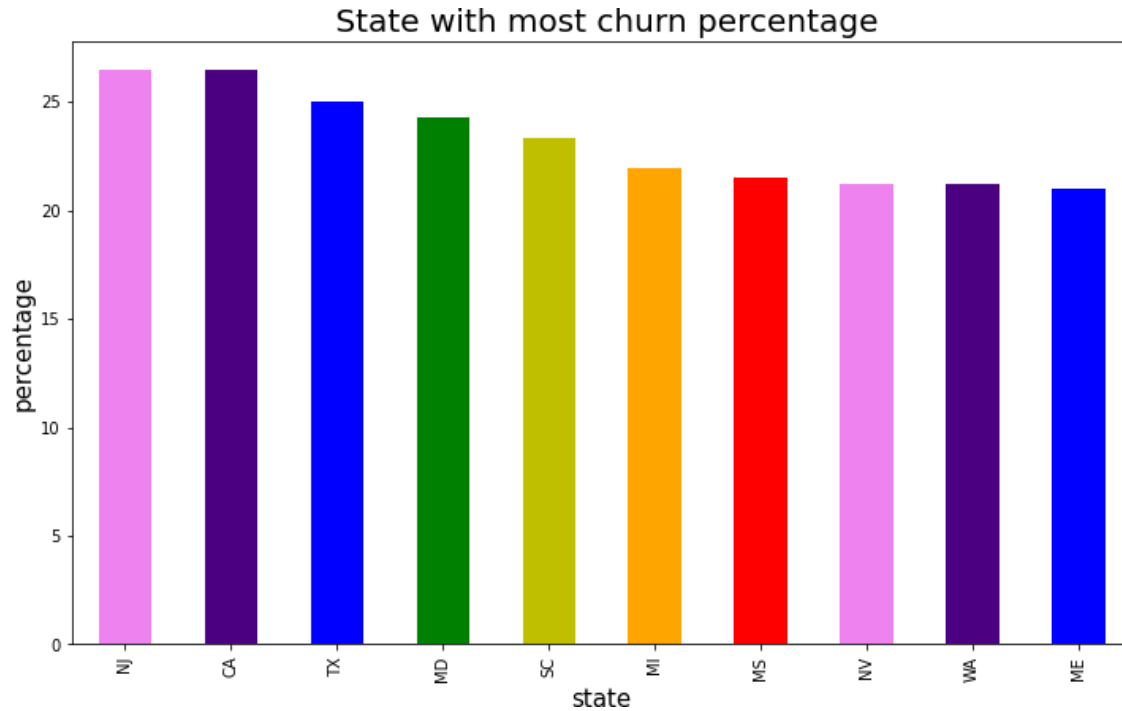
Count Plot for Voice mail plan



State Churn Rate

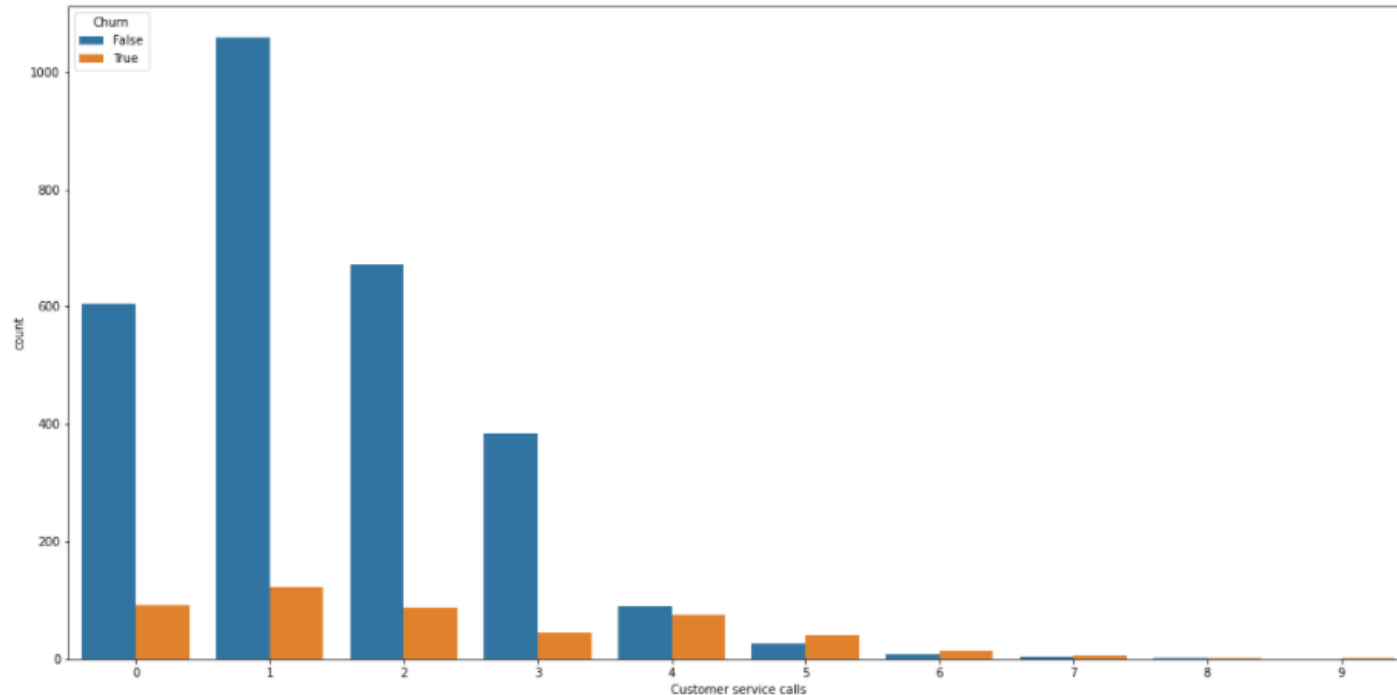


Churn percentage of states (Cont.)



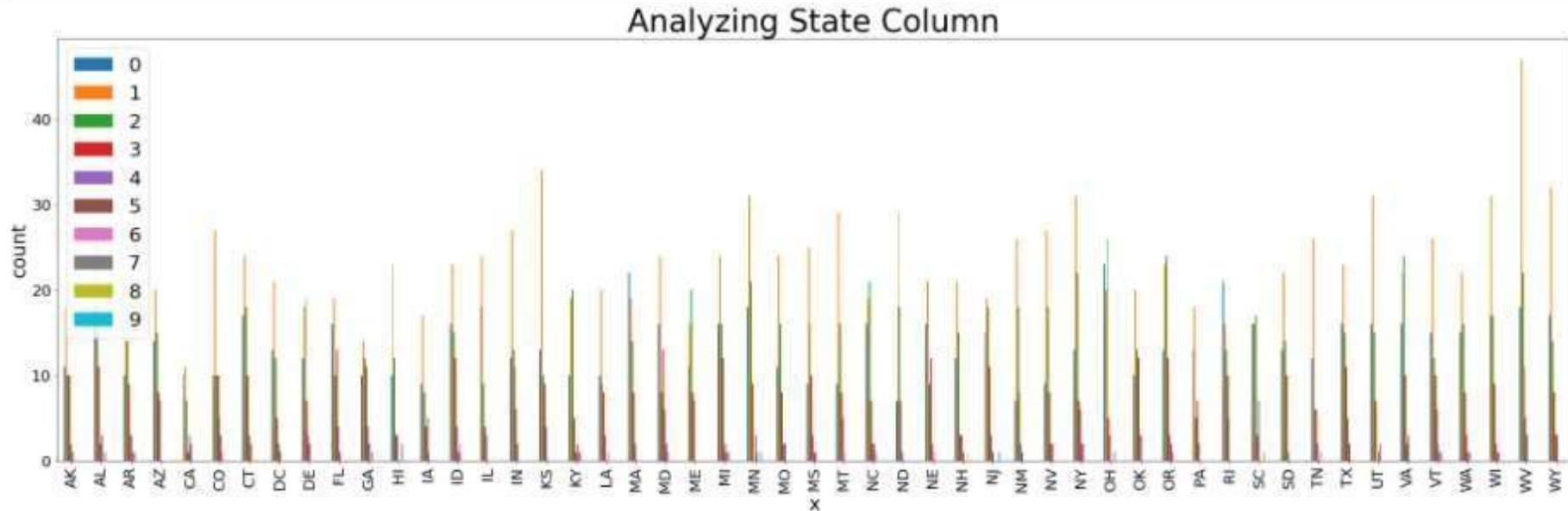
Churn percentage of states

- From graph below we can see there are very much less calls after 4. customers are only going for only 2 to 3 calls and then moving towards churn.



Churn percentage of states

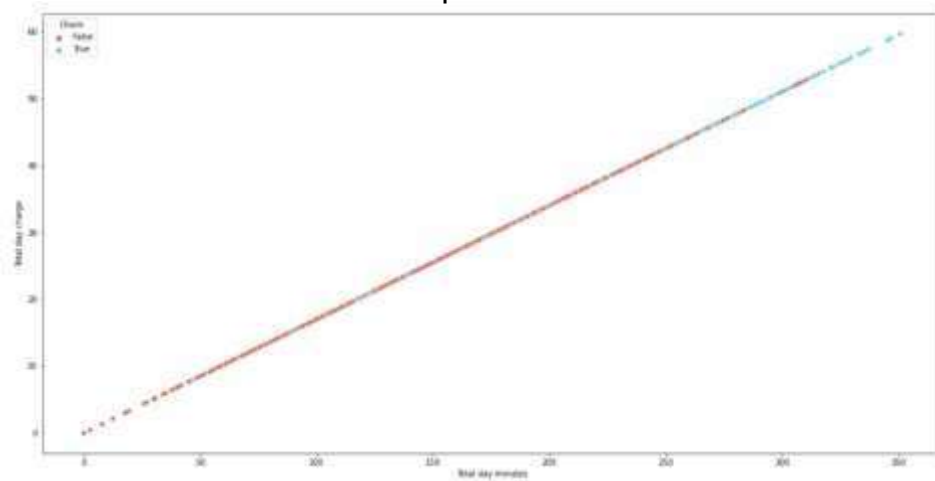
- We have seen top 10 and bottom 10 states with high and low percentage simultaneously. One more factor come into consideration before churn have customers called to customer care? Or not.
- From above graphs we know top state to churn is CA and NJ and bottom state is HI.



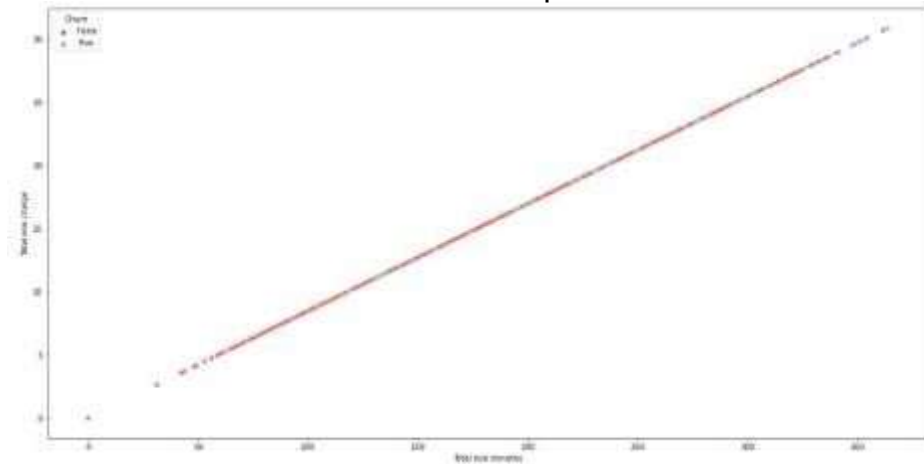
Churn percentage of states

- By scatter plot let us see which customers are moving towards churn
- Scatter plot 1 and scatter plot 2 shows relation between total day minutes and total day charge and total evening minutes and total evening charge.
- From scatter plot 1 as minutes increasing churn rate is increasing. It means customers who are churning using more minutes.
- From scatter plot 2 it is not same as day but there are some customers who are leaving with increasing minutes.

Scatter plot 1



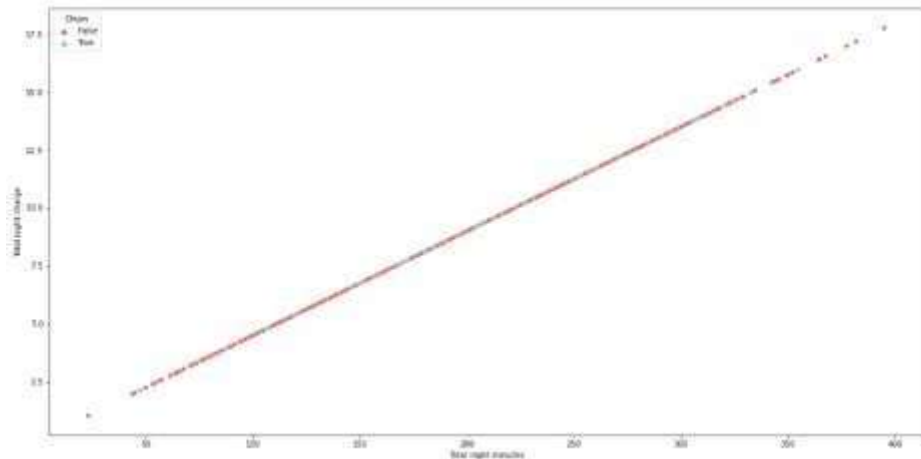
Scatter plot 2



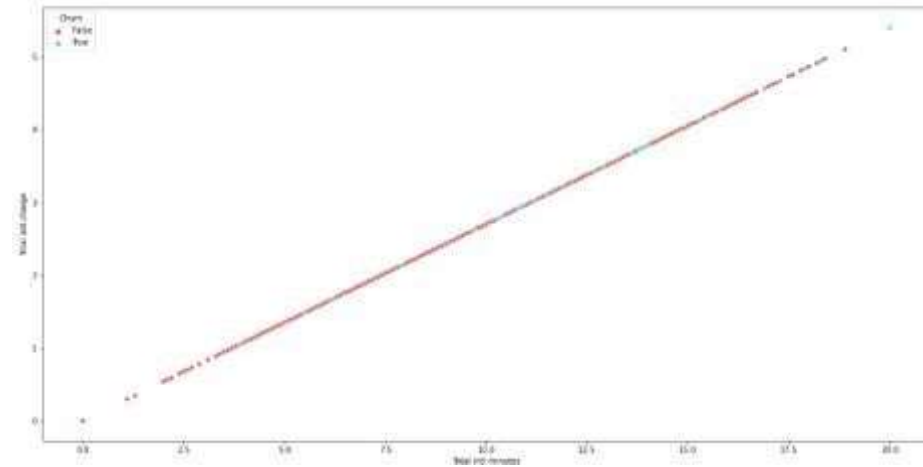
Churn percentage of states

- Scatter plot 3 shows relation between total night minutes and total night charges. In this customers are satisfied with increasing minutes there is less churn.
- Scatter plot 4 shows relation between total international minutes calls and total international charges. From graph we can say that customers above moderate use show tendency towards churn.

Scatter plot 3

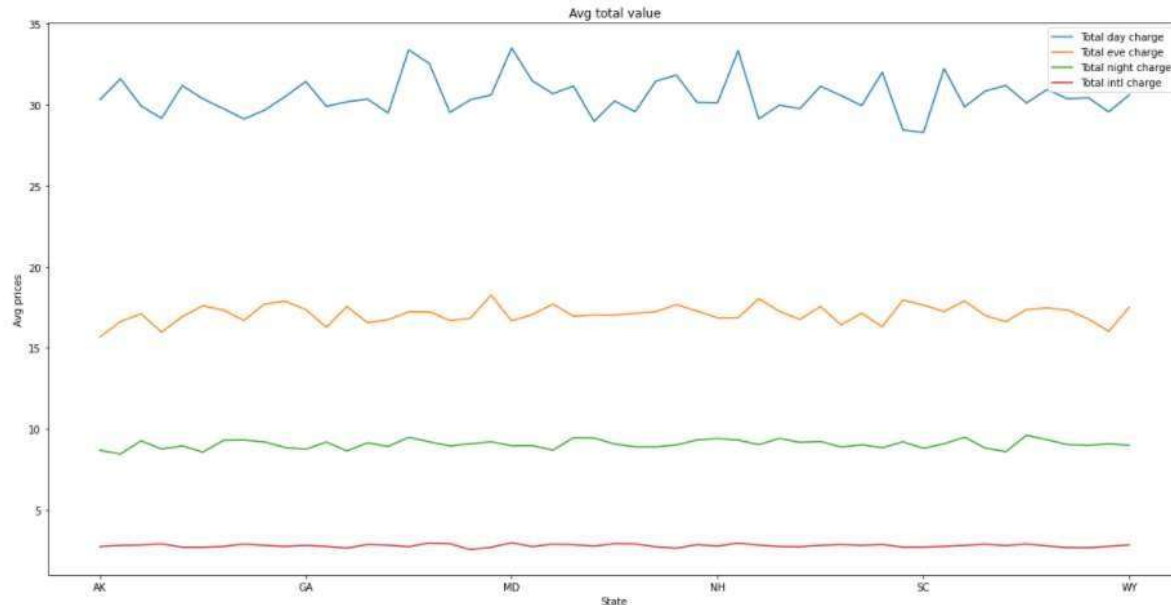


Scatter plot 4



Company revenue from customers

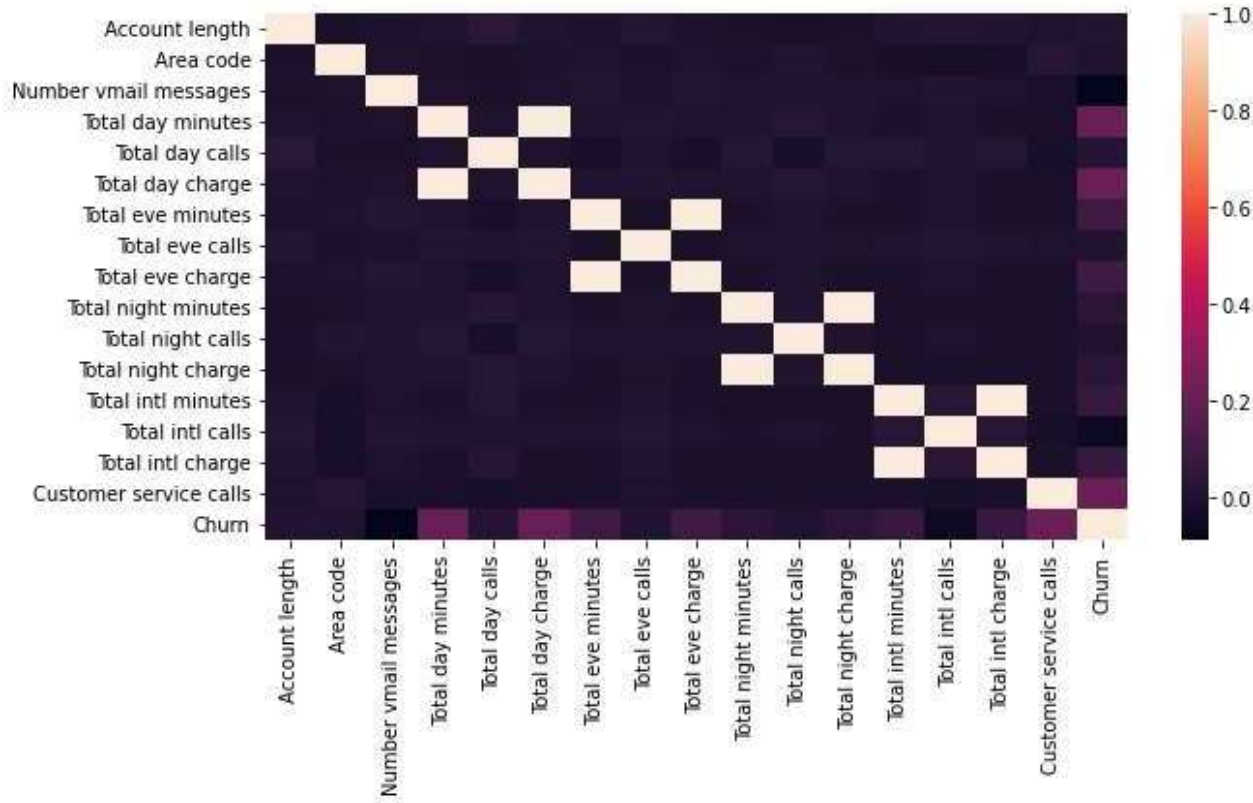
- From above all the graphs we have seen customers their tendency towards churn rates etc.
- Let us see revenue of company from customers. The graph shows relation between average prices and states.
- Maximum revenue is generated by customers who use day calls.



- Evening calls and night calls generated revenue but not at large scale.
- From prices we have seen international calls have more rates but revenue generated is small.

Relation between Churn and Others

- From heat map below we can find direct and in direct co-relation between various factors.



Relation between Churn and Others

- From heat map we can conclude that Total day minutes directly co-related to Total day charge, Total evening minutes directly co-related to Total evening charge, Total night minutes directly co-related to Total night charge, Total international minutes directly co-related to Total international charge.
- There is not more but some what co-relation between Total day minutes and Churn, Total day charge and churn, Customer service calls and churn.

Conclusion

From the given data and after performing EAD and comparison with the all the elements we say that there are some factors which company should taken care in consideration.

- States with high percentage of Churn are not approaching towards costumer service center. Instead when port request is put by some customers customer service should approach them.
- There is mix match churn rate for voice mail plans and voice mail messages.
- People with international plan who use more international minutes are moving towards churn.

Conclusion

- In Customer service calls data shows us that whenever an unsatisfied customer called the service center the churn rate is high, which means the service center didn't resolve the customer issue.
- For a telecom company it is necessary to approach towards customers on ground level and within certain period of time launch new schemes.
- States where customers churn rate is high increase advertisement in that area and increase customer service centers.

References

- 1“Heterogeneous ensemble stacking with minority upliftment (HESMU) for churn prediction on imbalanced telecom data” by K. Sivasaankar Karuppaiah, N.P. Gopalan Palanisamy Materials today :Proceedings xxx (xxxx) xxx.
- 2 “Churn prediction on huge telecom data using hybrid firefly based classification” By Ammar A. Q. Ahmed ↑ , Maheswari D. Egyptian Informatics Journal xxx (2017) xxx-xxx
- 3“Customer Churn Prediction in Telecommunication Sector using Rough Set Approach” By Adnan Amina , Sajid Anwara , Awais Adnana , Muhammad Nawaza , Khalid Alawfib , Amir Hussainc , Kaizhu Huangd Neurocomputing, <http://dx.doi.org/10.1016/j.neucom.2016.12.009>
- 4“Applying data mining to telecom churn management” By Shin-Yuan Hung a , David C. Yen b, Hsiu-Yu Wang Expert Systems with Applications 31 (2006) 515–524

Thank You...