Capstone Project
TELECOM CHURN ANALYSIS

SUBMITTED BY
SAMPANNA MISHRA

CONTENTS

- **Business Problem Understanding**
- **Objectives**
- **Data Summary**
- **Exploratory Data Analysis**
- **Challenges**
- **Conclusion**
- **Recommendation**

BUSINESS PROBLEM UNDERSTANDING

❖ **<u>TELECOM CHURN:</u> Customer churn means shifting from one service provider to its competitor in the market. Customer churn is one of the biggest fears of any industry, particularly for the telecom industry.**

❖ **Retaining customers is one of the most critical challenges in the maturing mobile telecommunications service industry.Telecom research indicates that it is more expensive to gain a new customer than to retain an existing one. In order to retain existing customers, Telecom providers need to know the reasons of churn, which can be realized through the knowledge extracted from Telecom data.**

❖ **Its estimated that the cost for convincing a regular customer not to churn to the competitor is 16 times less than the cost of searching and establishing contact with a new customer and the cost of attracting new customers is 5 to 6 times more than that for retaining existing ones.**

❖ **This project aims to analyse the data to determine the cause of customer churn and what to do to retain the valuable customers.**

OBJECTIVE

❖ **To identify patterns that can yield to customers churn**

❖ **Retain churn customers by taking appropriate steps**

DATA SUMMARY

AI

TARGET VARIABLE
➔ Churn

CATEGORICAL DATA
➔ State
➔ International plan
➔ Voicemail plan
➔ Area Code

TELECOM DATASET

NUMERICAL DATA
➔ Number vmail messages
➔ Total day minutes
➔ Total day calls
➔ Total day charge
➔ Total eve minutes
➔ Total eve calls
➔ Total eve charge
➔ Total night minutes
➔ Total night calls
➔ Total night charge
➔ Total intl minutes
➔ Total intl calls
➔ Total intl charge
➔ Customer service calls
➔ Account length

DATA SUMMARY

Given below is the Orange Telecom Churn Dataset which further has been used for analysis .The first five rows have been displayed below:

	State	Account length	Area code	International plan	Voice mail plan	Number vmail messages	Total day minutes	Total day calls	Total day charge	Total eve minutes	Total eve calls	Total eve charge	Total night minutes	Total night calls	Total night charge	Total intl minutes	Total intl calls	Total intl charge	Customer service calls	Churn
0	KS	128	415	No	Yes	25	265.1	110	45.07	197.4	99	16.78	244.7	91	11.01	10.0	3	2.70	1	False
1	OH	107	415	No	Yes	26	161.6	123	27.47	195.5	103	16.62	254.4	103	11.45	13.7	3	3.70	1	False
2	NJ	137	415	No	No	0	243.4	114	41.38	121.2	110	10.30	162.6	104	7.32	12.2	5	3.29	0	False
3	OH	84	408	Yes	No	0	299.4	71	50.90	61.9	88	5.26	196.9	89	8.86	6.6	7	1.78	2	False
4	OK	75	415	Yes	No	0	166.7	113	28.34	148.3	122	12.61	186.9	121	8.41	10.1	3	2.73	3	False

# DEFINING INDEPENDENT VARIABLES

AI

- **State** → **51 unique states have been used**

- **Account Length** → **Length of the account**

- **Area code** → **Code number of areas of some states**

- **International plan** → **Yes indicates international plan is present and No indicates no subscription for international plan.**

- **Voicemail plan** → **Yes indicates subscription is present and No indicates no subscription for Voice mail.**

- **Number vmail messages** → **No.of voice mail messages ranging from 0 to 50.**

- **Total day minutes** → **Total number of minutes spent in the morning.**

- **Total day calls** → **Total charges to the customer making international calls.**

- **Total day charge** → **Total charges to the customer in the morning.**

- **Total eve minutes** → **Total no.of minutes spent in the evening.**

DEFINING INDEPENDENT VARIABLES

AI

- **Total eve calls** → Total no.of calls made in evening.

- **Total eve charge** → Total charges to the customer in the evening.

- **Total night minutes** → Total number of minutes spent in night.

- **Total night calls** → Total no.of calls made in night.

- **Total night charge** -> Total charges to the customer in the night.

- **Total intl minutes** → Total no.of minutes spent in international calls.

- **Total intl calls** → Total number of calls made to people outside the country.

- **Total intl charge** → Total charges to the customer making international calls.

- **Customer service calls** → Number of customer service calls made by customers.

EXPLORATORY DATA ANALYSIS

The plot shows the churn in each state:
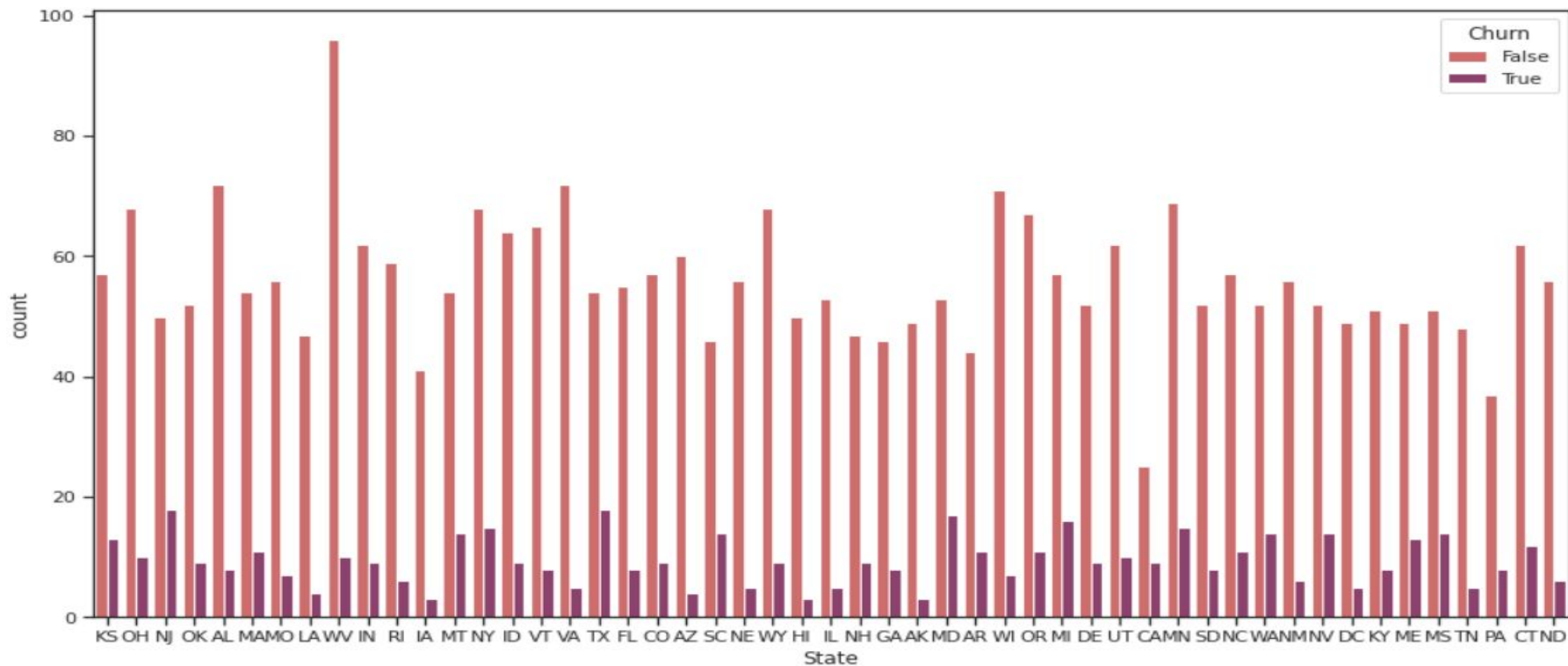
EXPLORATORY DATA ANALYSIS

The plot below shows the top 5 state churn percentage. NJ, CA, TX, MD,SC,MJ are the ones who have higher churn rate
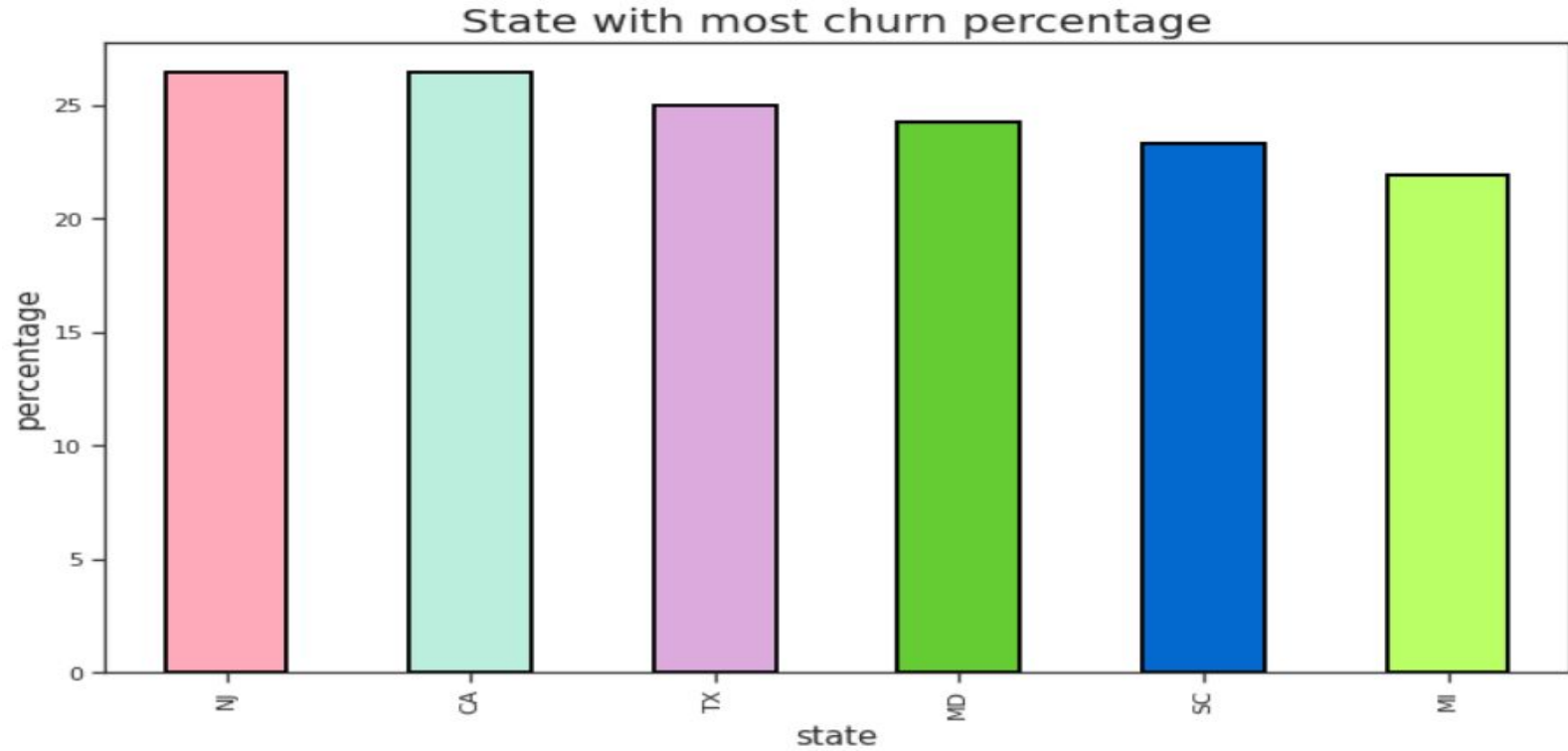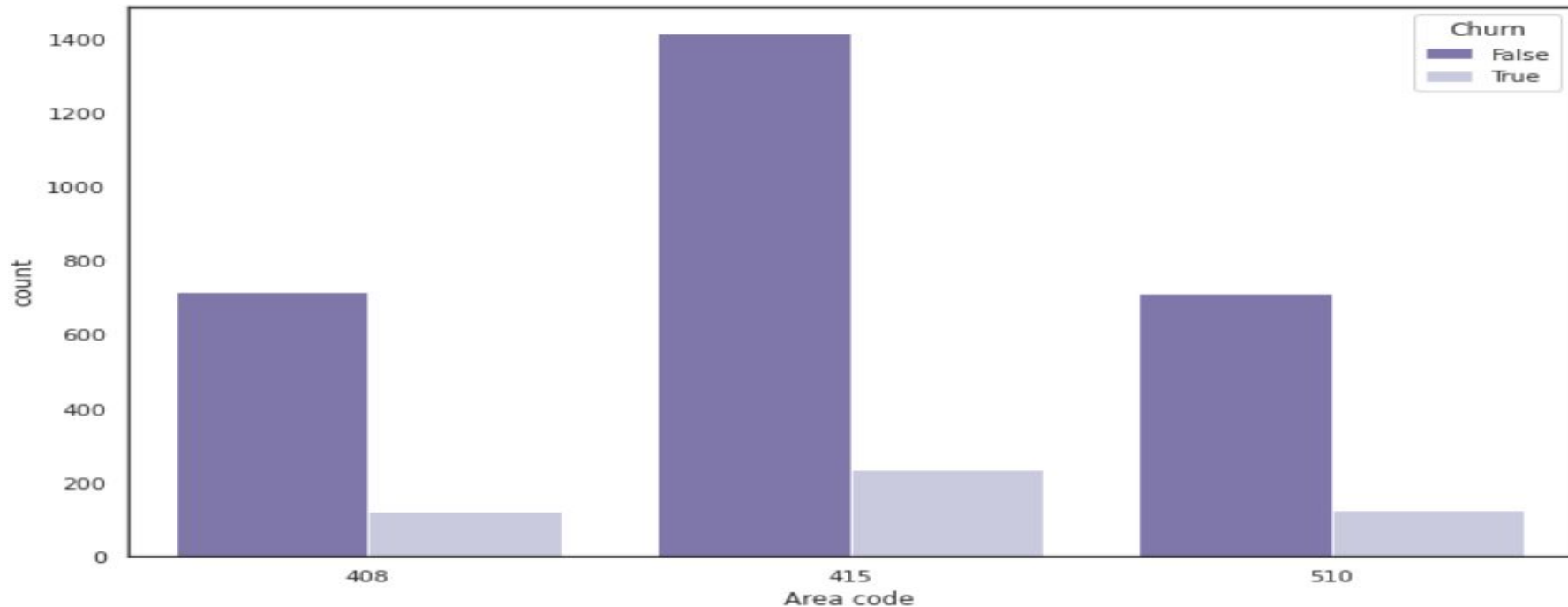
EXPLORATORY DATA ANALYSIS

The plot shows effect of Account length on churn.Evident that it does not play significant role in customer churn.
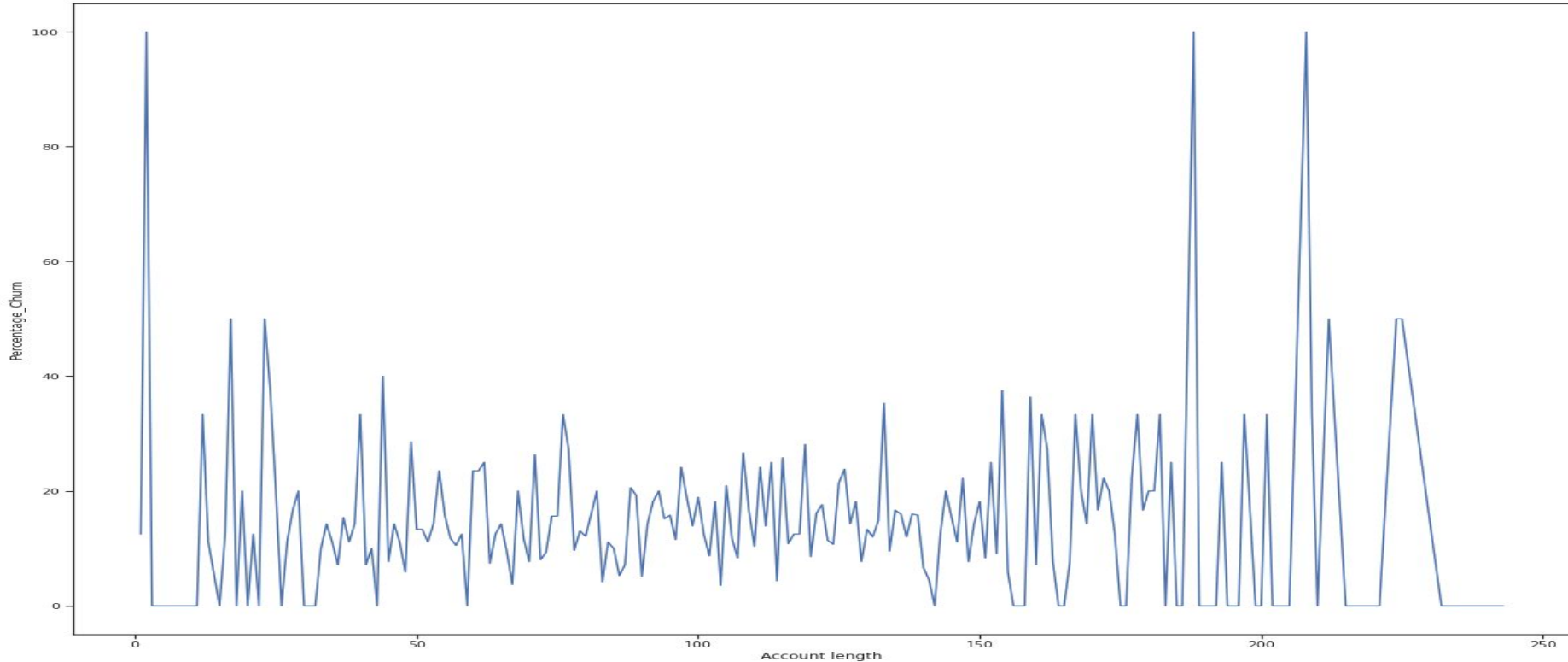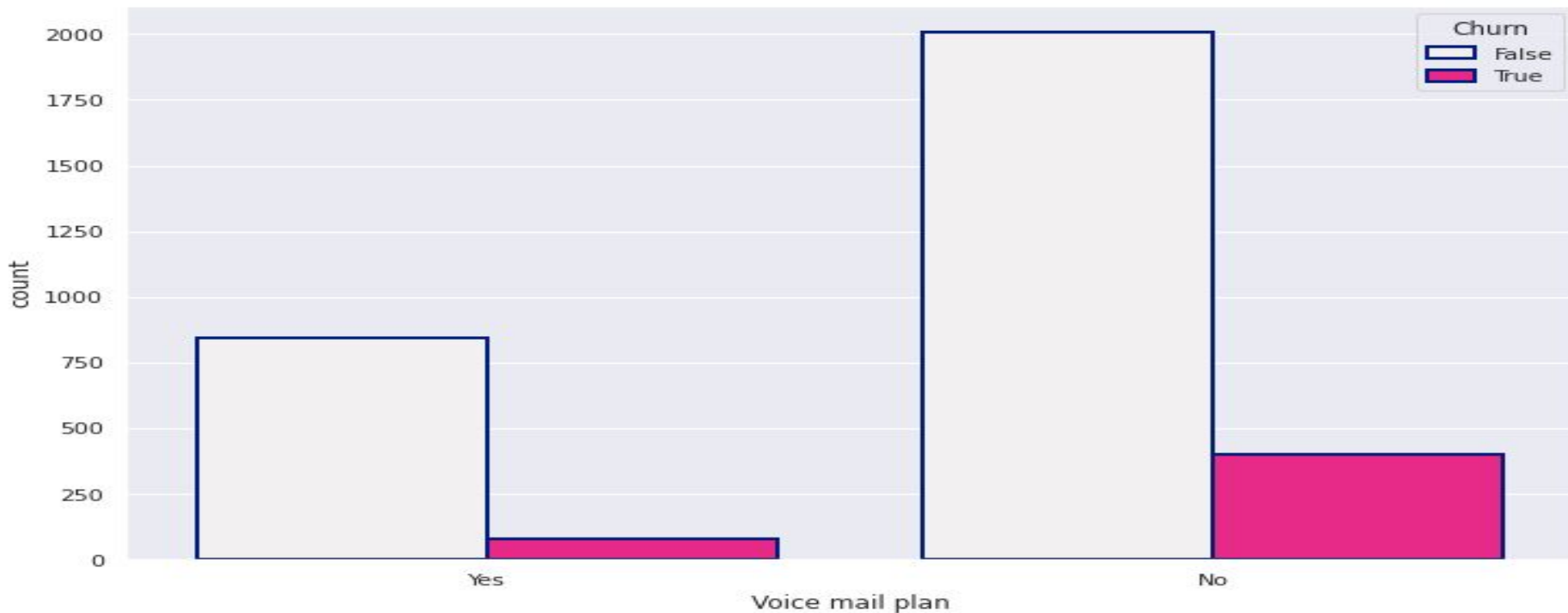
EXPLORATORY DATA ANALYSIS

We observe that there are only three unique Area Codes and the churn rate of these area codes are almost same.Most probably there is not any relationship between area code and churn.

EXPLORATORY DATA ANALYSIS

In the dataset 2411 don't have subscription for voice mail plan while 922 people have. The churn rate among the people who subscription is 8.7%.
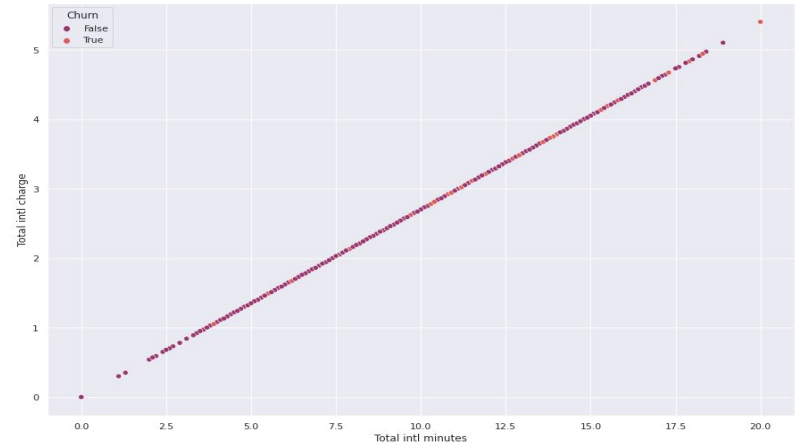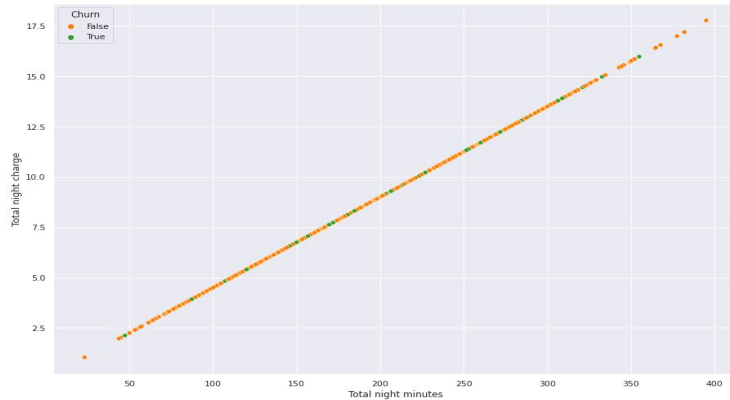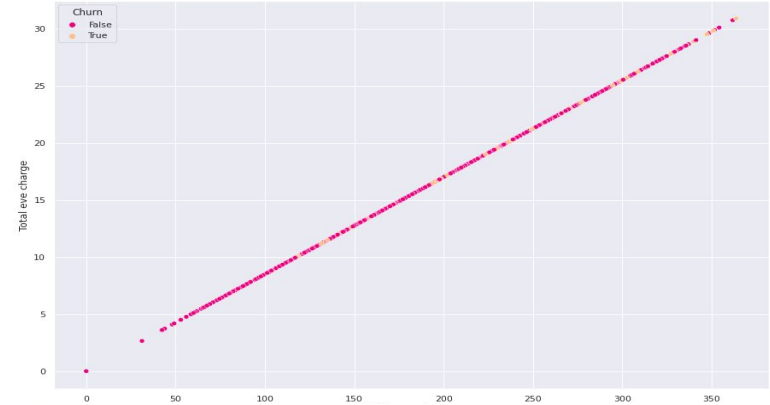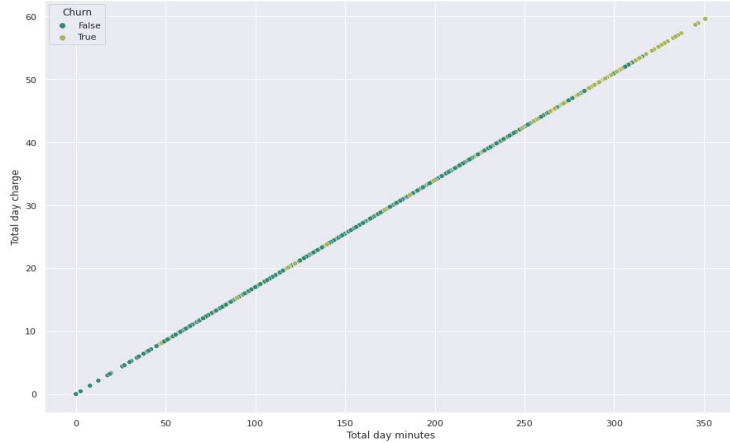
EXPLORATORY DATA ANALYSIS

The plot shows 323 people have subscribed for international plan while 3010 people have not subscribed. The churn percentage for subscribed people is 42.4%. For whose who don't have subscription is 11.4%.
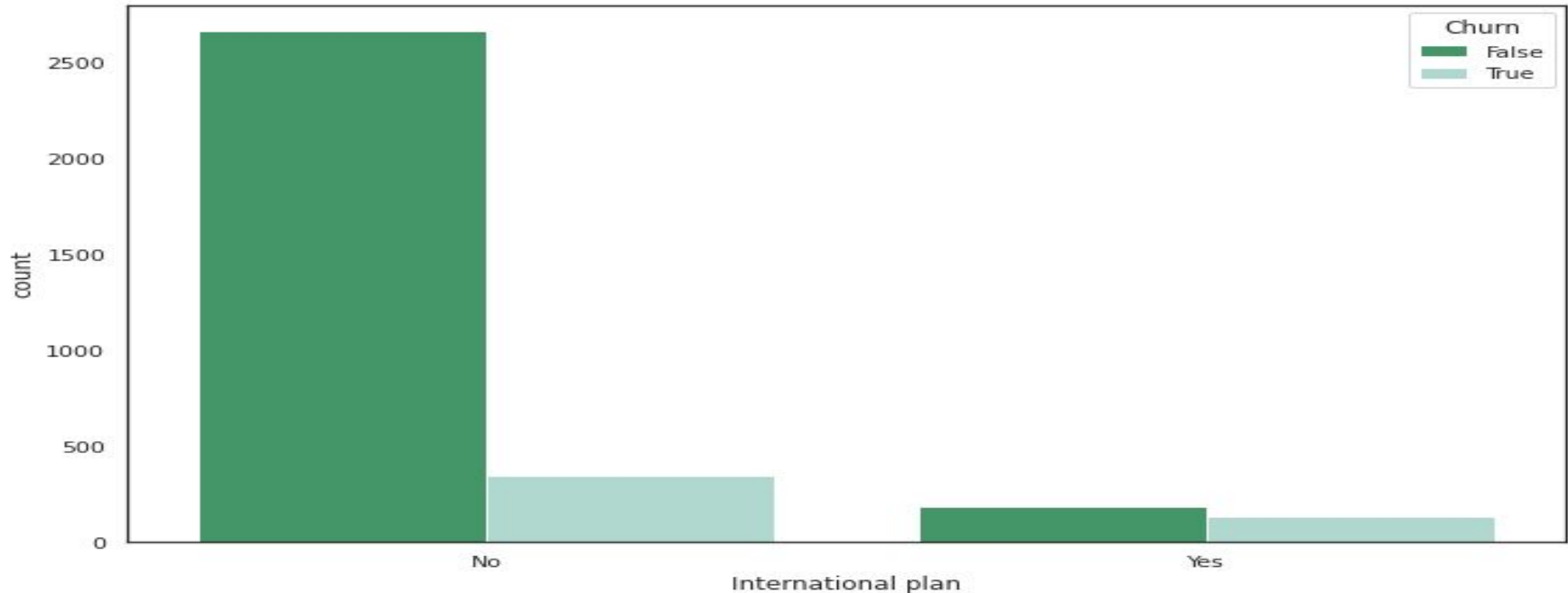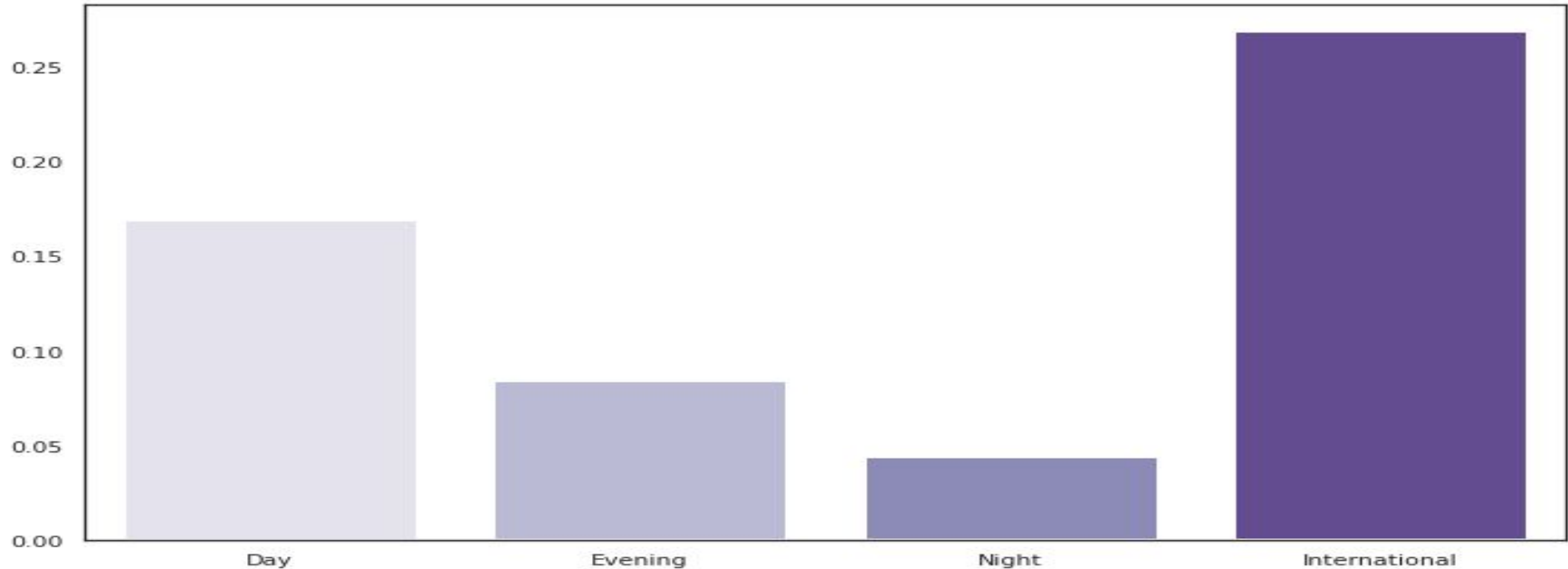
EXPLORATORY DATA ANALYSIS

Analysing relationship of all call minutes and all call charges. It includes day, evening, night and international calls. They share a linear relationship.

EXPLORATORY DATA ANALYSIS

The plot below shows the comparison between all call and charges per minute. International call charges are high as compared to others.Also customer churn most having international plans.
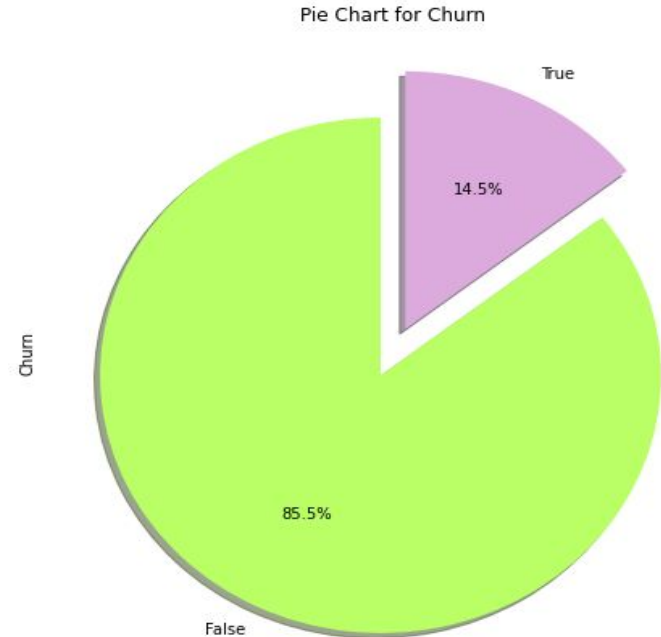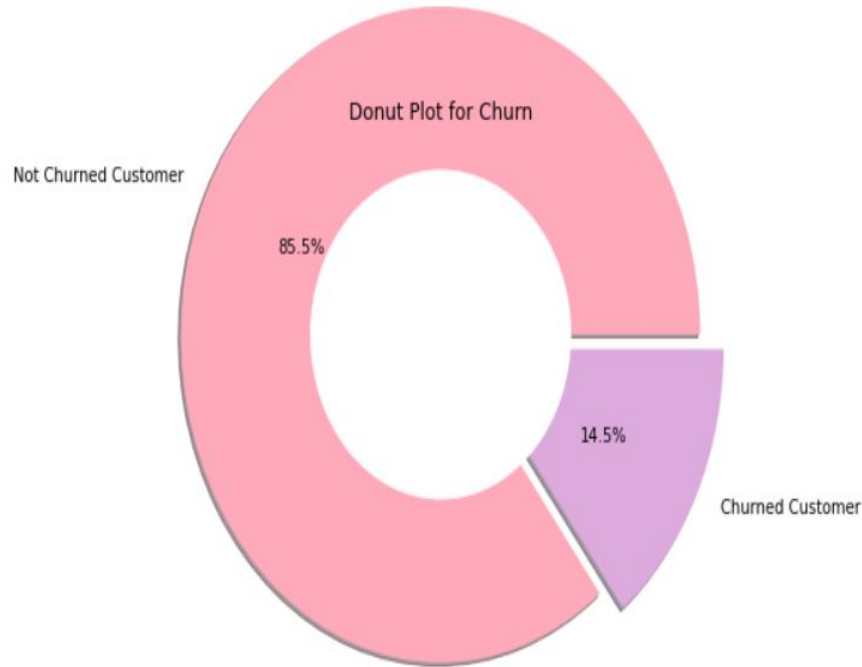
EXPLORATORY DATA ANALYSIS

AI

The figures below shows the mapping between number of customer service calls and percentage churn.It can be seen that after 4 calls at least 45% of customers churn.The trend depicts that more than 4 customer calls the the probability of person churning increases.
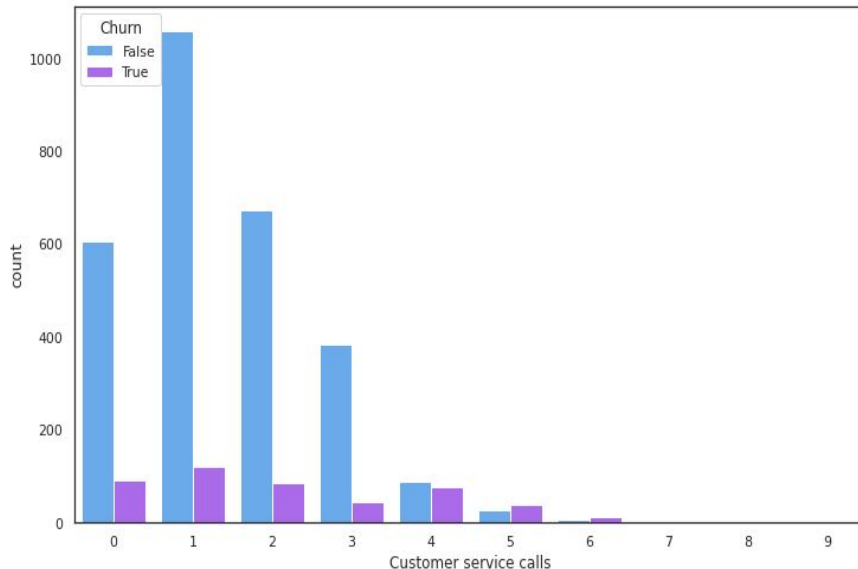
Churn Customer service calls	False	True	Percebtage_Churn
0	605	92	13.199426
1	1059	122	10.330229
2	672	87	11.462451
3	385	44	10.256410
4	90	76	45.783133
5	26	40	60.606061
6	8	14	63.636364
7	4	5	55.555556
8	1	1	50.000000
9	0	2	100.000000

EXPLORATORY DATA ANALYSIS

The donut plot and pie plot of customer churn shows almost 15% of customers churned from the telecom company.

EXPLORATORY DATA ANALYSIS

CORRELATION MATRIX

CHALLENGES

➢ **Analysing the call charges among day, evening , night international type vs total number of minutes for each category.**

➢ **Difficult to analyse columns like account length and area code.**

CONCLUSION

➤ Those customers who have International plan churn more and also the international calling charges are also high so the customer who has the plan unsatisfied with network high call charge.

➤ There are some states where the churn rate is high as compared to others may be due to low network coverage.

➤ Area code and Account length do not play any kind of role regarding the churn rate so,it's redundant data columns.

➤ Total day call minutes, total day calls, Total day charge, Total eve minutes, Total eve calls, Total eve charge, Total night minutes, Total night calls, Total night charge, these columns didn't play any kind of role regarding the churn rate.

➤ 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.

RECOMMENDATION

AI

➢ **Improve the service of call center and take frequently feedback from the customer regarding their issue and try to solve it as soon as possible.**

➢ **Improve network coverage in the top five states with highest churning rate.**

➢ **In international plan provide some discount plan to the customer.**