CAPSTONE PROJECT

EDA on Telecom Churn Analysis

By

Harshith Gowda

Data Science Trainee, AlmaBetter

CONTENT

- Business Problem Understanding
- Objective
- Data Summary
- Exploratory Data Analysis
- Challenges
- Recommendation
- Conclusion
- Q & A

BUSINESS PROBLEM UNDERSTANDING

- Customer churn in the telecom industry poses one of the most significant risks to loss of revenue.

- The average churn rate in the telecom industry is approximately 1.9% per month, but could rise as high as 67% annually * as per survey conduct.

- Since the cost of acquiring new customers is up to 25 times higher than the cost of retaining them, reducing the churn rate of customer is key.

- To reduce customer churn, telecom companies need to predict which customers are at high risk of churn for this we taking advantage of the vast streams of rich telecom customer data.

- This project aims to analyse the data to determine the cause of customer churn customers who are most likely subject to churn, and what to do to retain the most valuable customer.

OBJECTIVE

- **Maximize:**
  Company's profit by retaining customer
- **Minimize:**
  Customer churn by identifying the key cause of the problem

Business Constraint:

- Provide offers and discount and improve the service quality without compromising with profit
- Maintain company’s brand value

DATA SUMMARY

Decision Variable
- ➔ Churn

Categorical Data
- ➔ Stata
- ➔ International plan
- ➔ Voicemail plan

Nominal Data
- ➔ 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

This is The Orange Telecom Churn Dataset.In the below table it's show the top and bottom 5 rows respectively

	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

	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
3328	AZ	192	415	No	Yes	36	156.2	77	26.55	215.5	126	18.32	279.1	83	12.56	9.9	6	2.67	2	False
3329	WV	68	415	No	No	0	231.1	57	39.29	153.4	55	13.04	191.3	123	8.61	9.6	4	2.59	3	False
3330	RI	28	510	No	No	0	180.8	109	30.74	288.8	58	24.55	191.9	91	8.64	14.1	6	3.81	2	False
3331	CT	184	510	Yes	No	0	213.8	105	36.35	159.6	84	13.57	139.2	137	6.26	5.0	10	1.35	2	False
3332	TN	74	415	No	Yes	25	234.4	113	39.85	265.9	82	22.60	241.4	77	10.86	13.7	4	3.70	0	False

FEATURES DESCRIPTION

STATE:There are 51 unique state present
ACCOUNT LENGTH:It is the length that the customer used their account
AREA CODE: There are 3 unique area code present

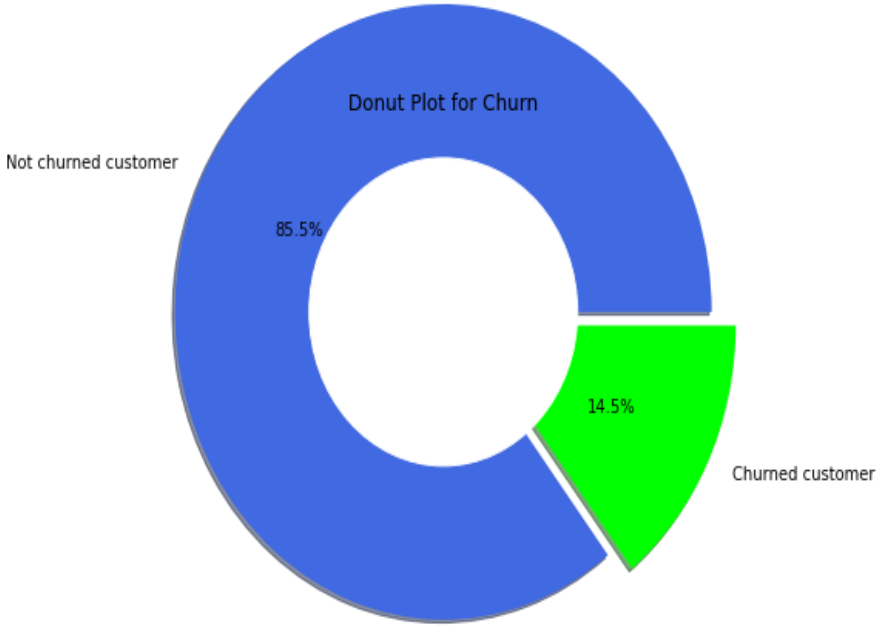
INTERNATIONAL PLAN & VOICEMAIL PLAN:
Both column are described as a categorical feature,yes means plan taken no means plan not taken
NO. OF VOICEMAIL MESSAGES: The number of voicemail make by the voicemail plan taken customer

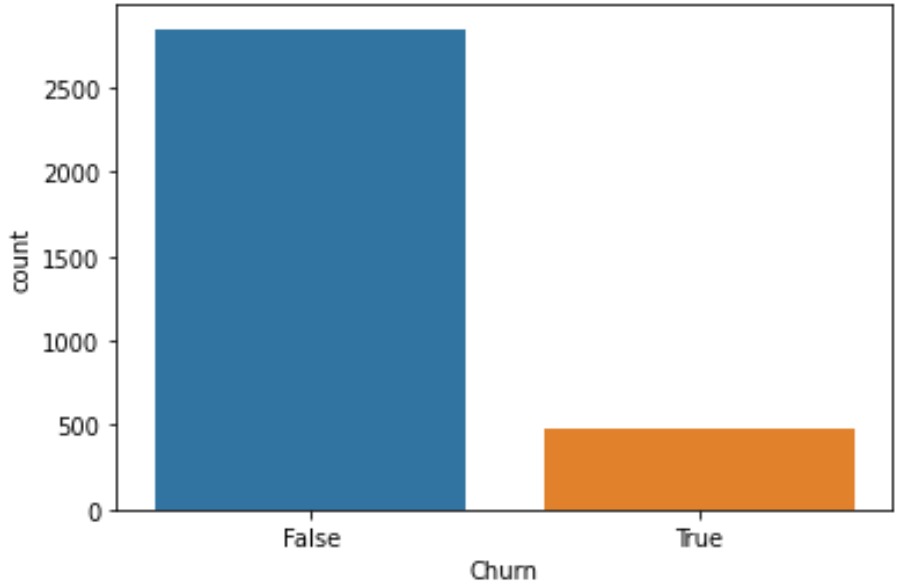
TOTAL (DAY/EVENING/NIGHT/INTERNATIONAL) (MINUTES/CALLS/CHARGES):
These are total 12 columns, and all are numerical data types.These contain the data of calls, minutes, charges of the customer with respective to the various time of the day and plan.

Customer service calls: It is the number of calls made by the customer to operator service centre
Churn:it is our target dependent variable having boolean data type of true and false

ANALYZING DEPENDENT VARIABLE “CHURN”

- Below plot on the left side is a donut plot shows the percentage of total churned and not churned customer
- And on the right side count plot shows the number of customer churned and not churned



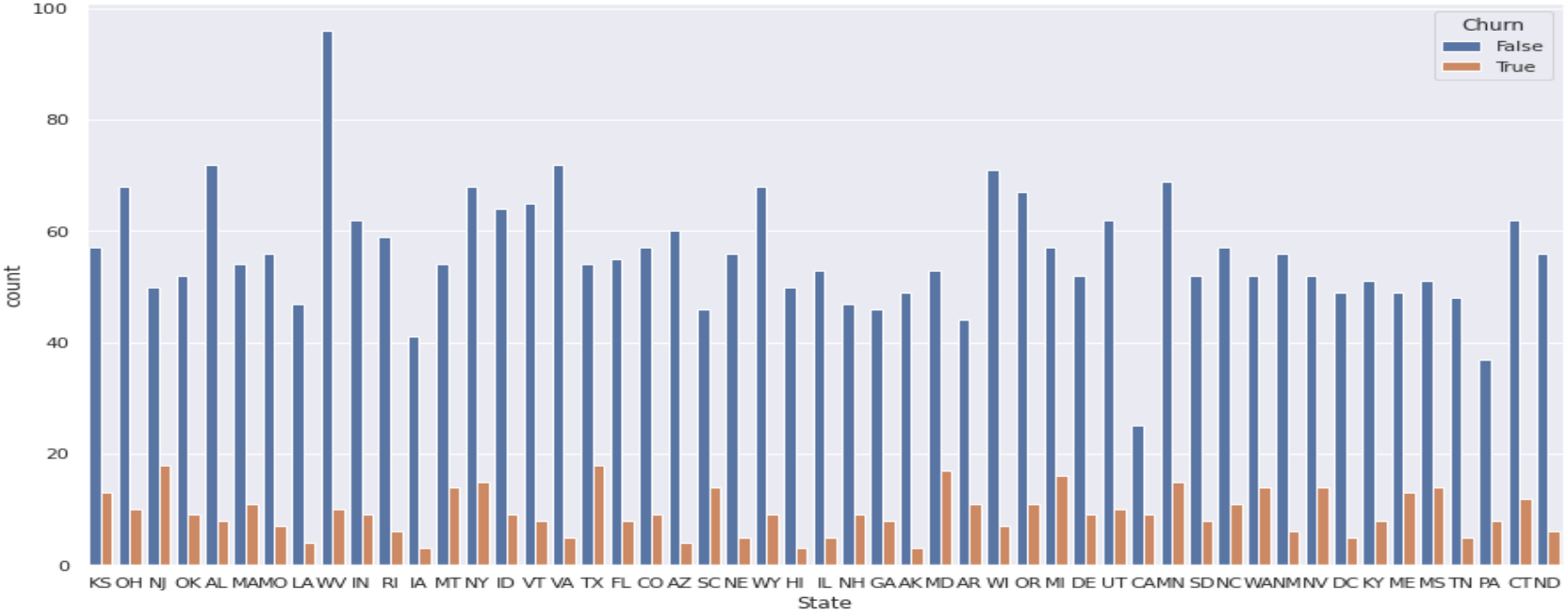


ANALYZING DEPENDENT VARIABLE “CHURN”

- Total Users number - 3333.
- 2850 - Non churn (85.5%)
- 483 - Churn (14.5%)
- From the above donut plot and count plot, It was found from this analysis that almost 14.5% of customers had churned.

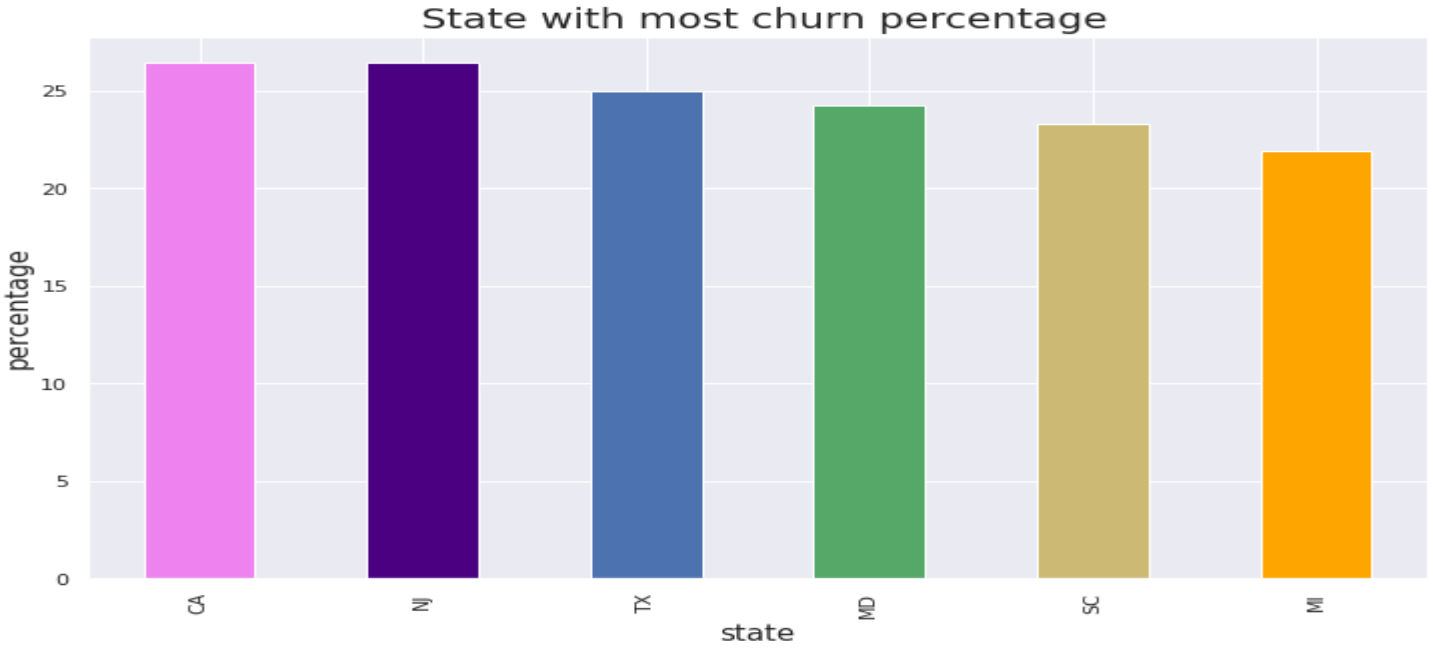
ANALYSIS STATE COLUMN

➢ **This Plot shows the Churn in each state**



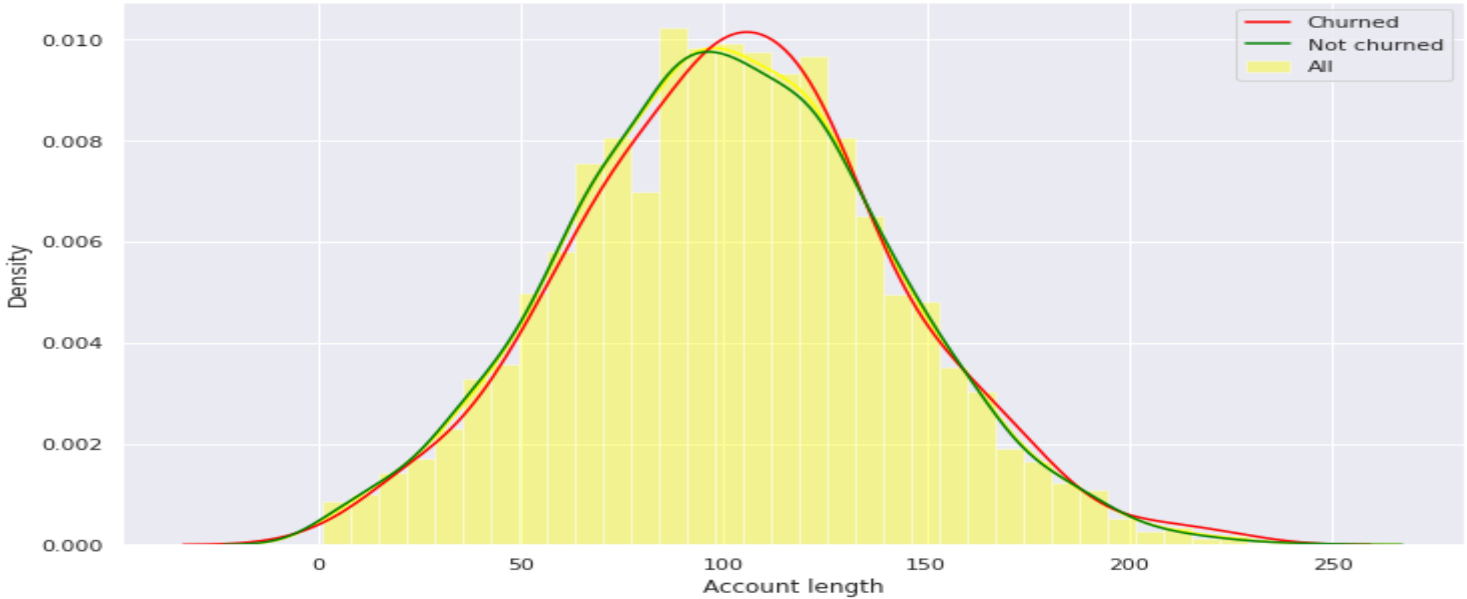
TOP STATE CHURN PERCENTAGE

- **CA, NJ ,TX , MD ,SC ,MI are the ones who have higher churn rate more than 21.74%**



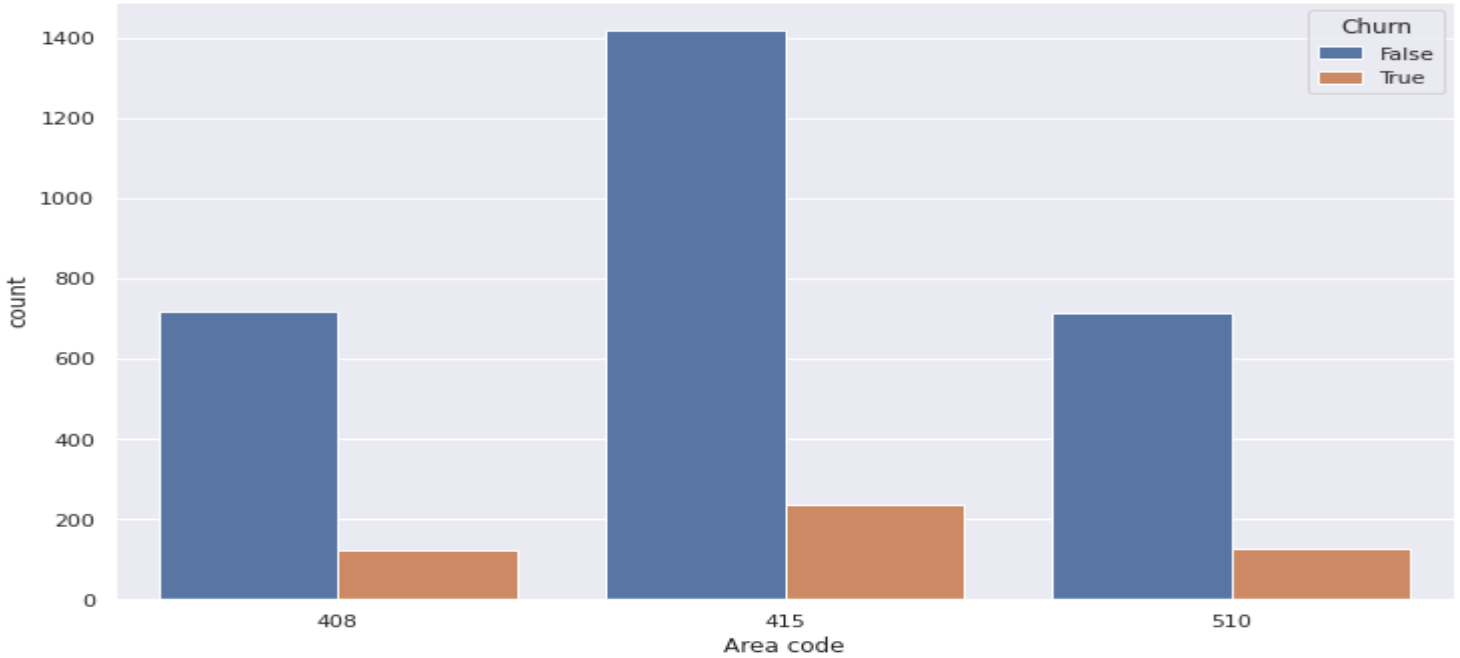
ACCOUNT LENGTH vs. CHURN

- This Plot show effect of Account Length on Churn.
- Here is no sign of customers leaving because of the length of usage of their account.



ANALYSIS OF AREA CODE

- This plot graph shows all the values with churn.
- Area code has only 3 unique values, and consider as a nominal data type and has equal number of churn.

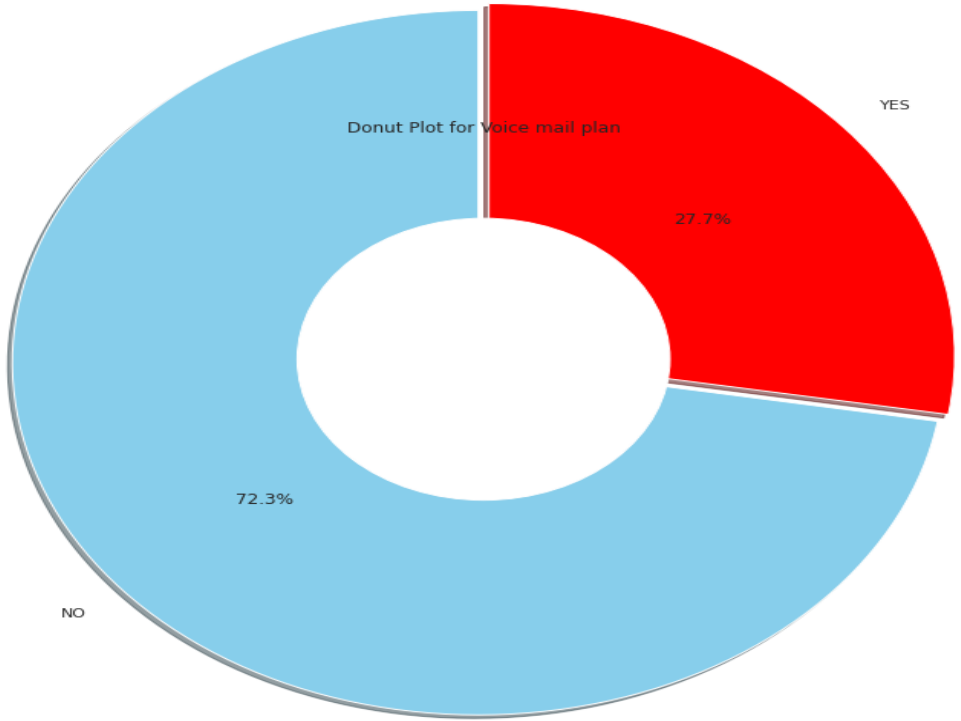


ANALYSIS OF VOICEMAIL PLAN

- There are 3333 people,

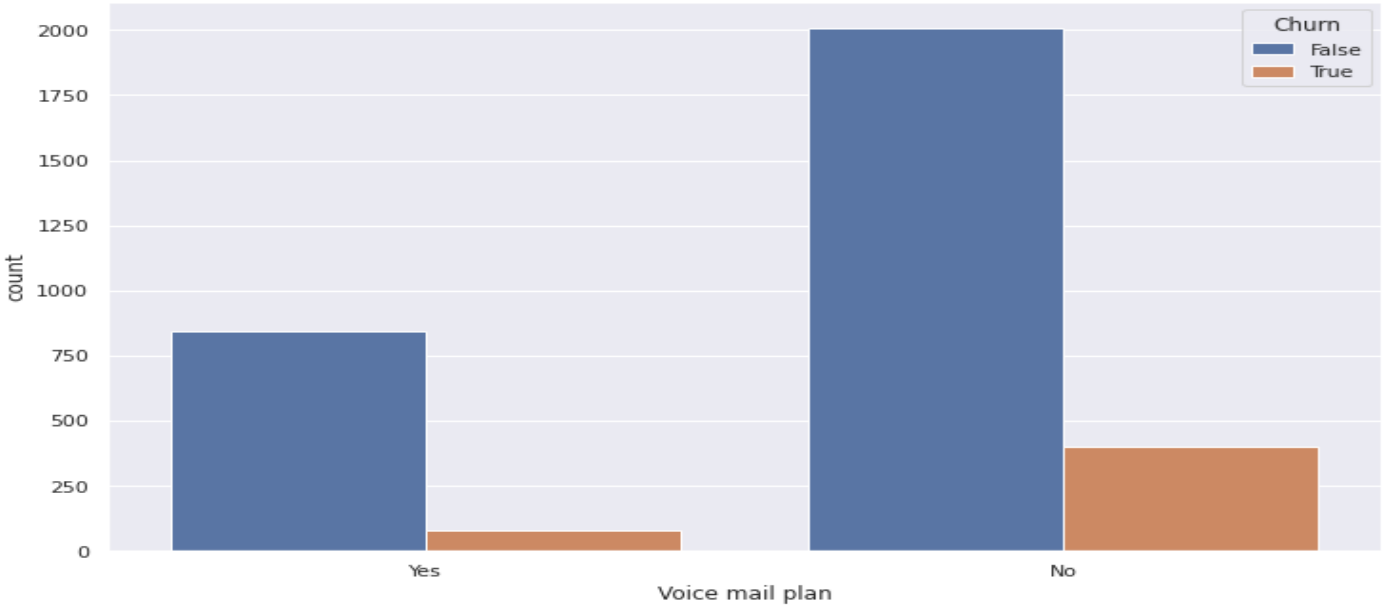
922 having Voicemail plan,

2411 do not have any Voicemail plan.



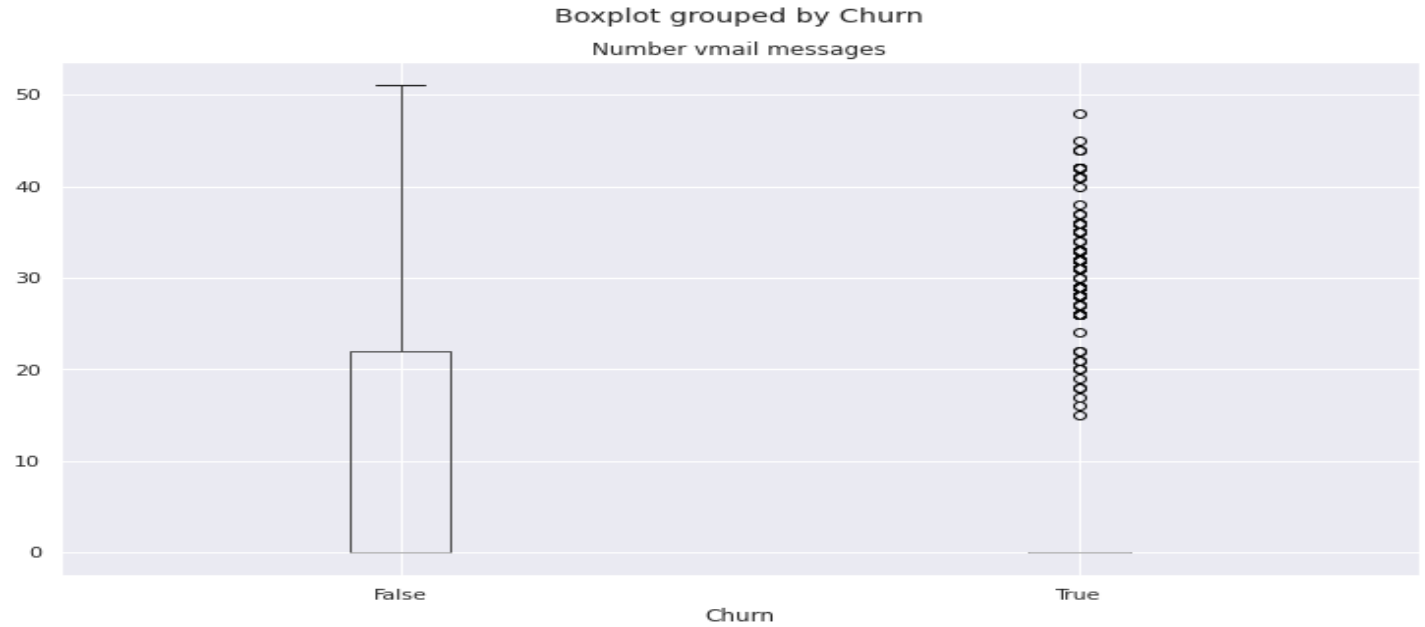
VOICEMAIL PLAN vs. CHURN

- This plot shows churn corresponding with the subscription of voicemail plan.
- Out of 922 people having Voicemail plan, 8.7% are Churn.



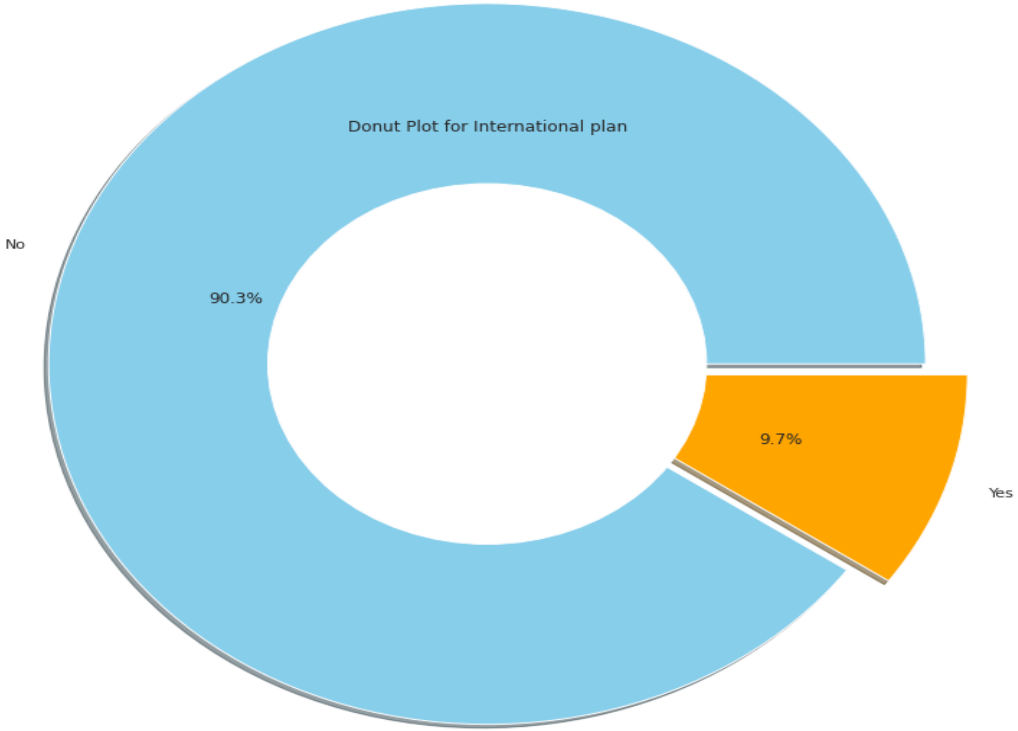
NO. OF VOICEMAIL vs. CHURN

- This box plot shows the relation between churn and no. of vmail.
- when there are more than 20 voice-mail messages then there is a churn.



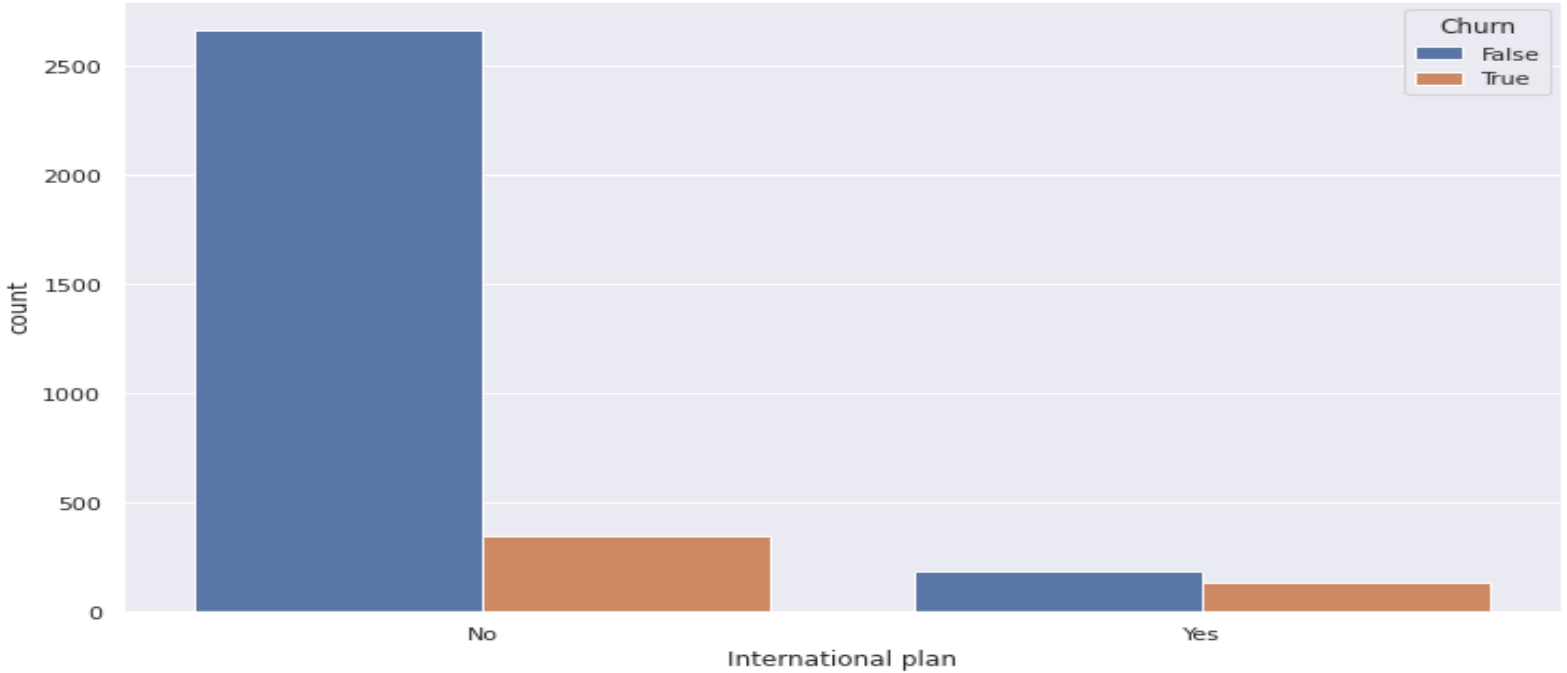
INTERNATIONAL PLAN

- There are 3333 people
- 323 have a International Plan
- 3010 do not have International Plan



INTERNATIONAL PLAN vs. CHURN

This is a count plot which shows the churned and not churned customer respective to their international plan



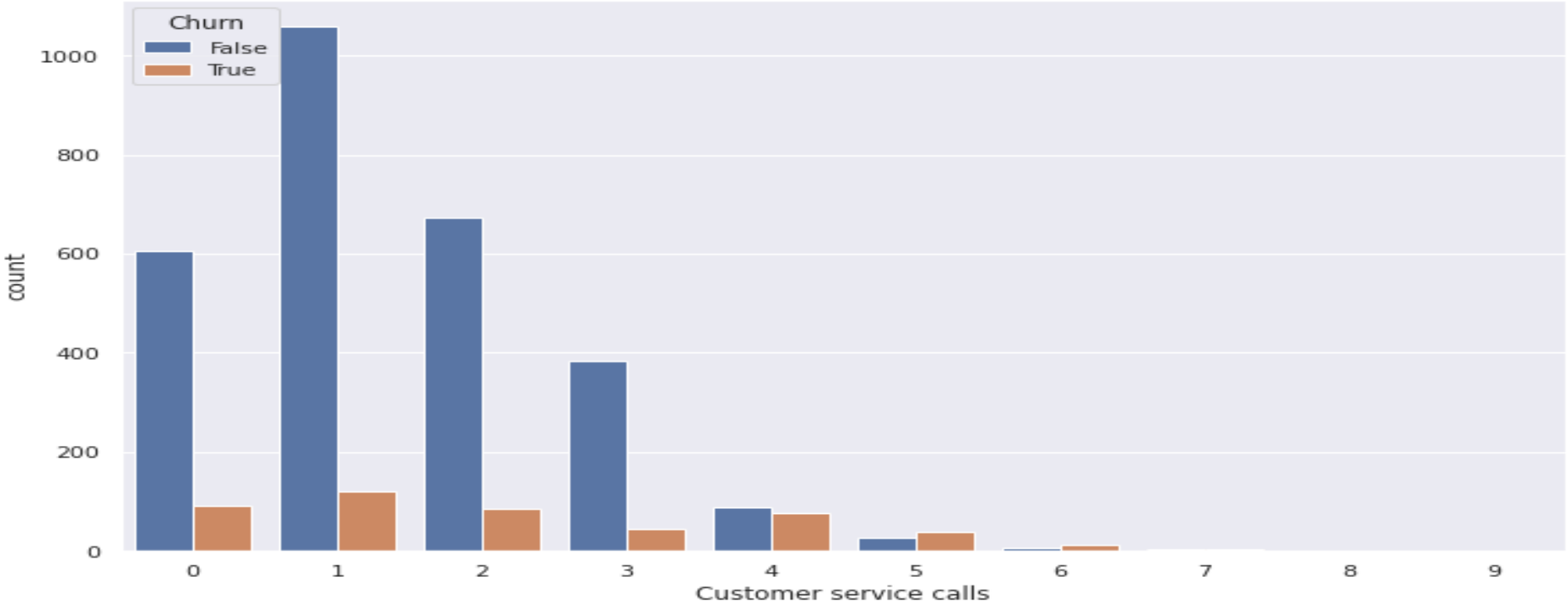
INTERNATIONAL PLAN vs. CHURN

- This side table map shows data about the percentage churn according to the international plan
- Which clearly shows that the churn percentage is 42.41% who takes the international plan
- That means there is some cause which effects the churn rate like call price or network issue

Churn	False	True	Percentage Churn
International plan			
No	2664	346	11.495017
Yes	186	137	42.414861

CUSTOMER SERVICE CALLS vs. CHURN

This plot shows Churn for number of customer service calls



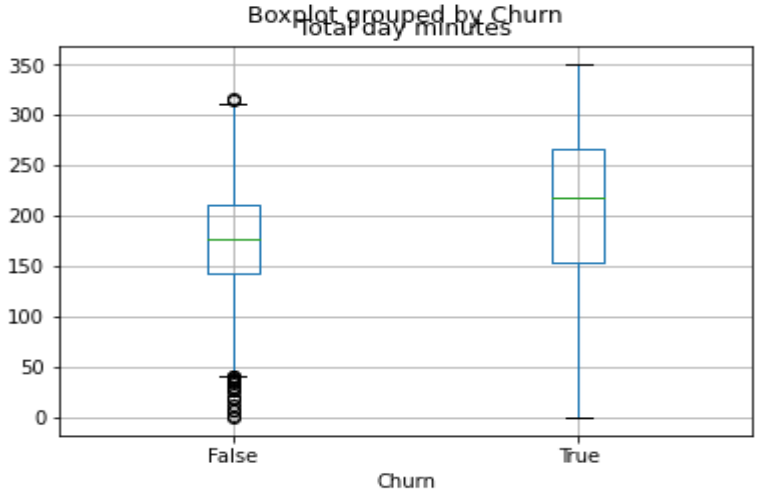
CUSTOMER SERVICE CALLS vs. CHURN

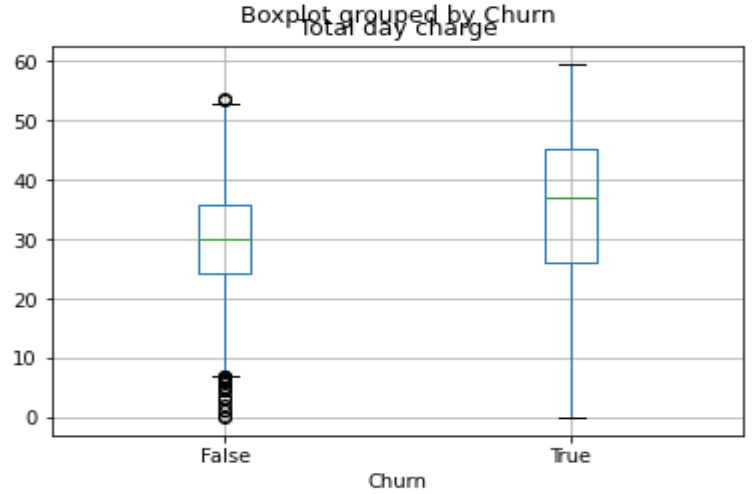
- This table mapping number of customer calls to the churn percentage
- It's clear that after 4 calls at least 45% of the subscribers churn.
- Customers with more than 4 service calls their probability of leaving is more

Churn	False	True	Percentage_Churn
Customer service calls			
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

DAY CALL MINUTES & DAY CALL CHARGE vs. CHURN

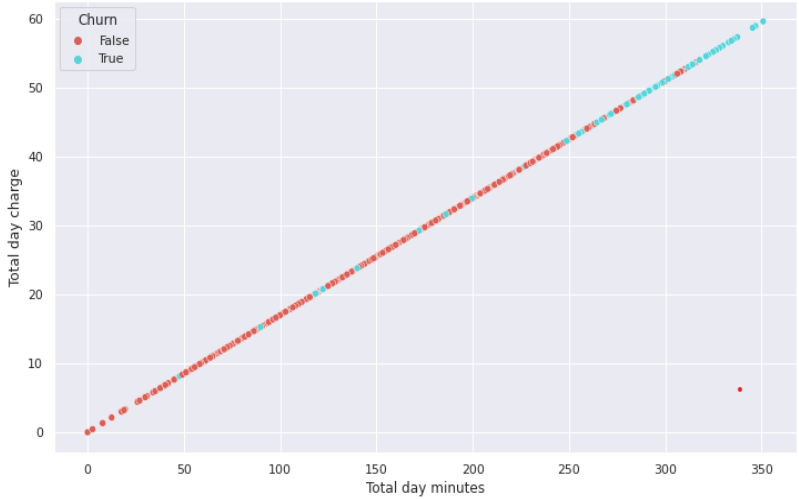
- Left side box-plot shows the relation between total day minutes with churn
- Right side box-plot shows the relation between total day call charges with churn
- Below box-plot shows that with users spending more 225 minutes or more tend to switch to other operator.
- The customer who have high call minutes also have high call price these tends to churn

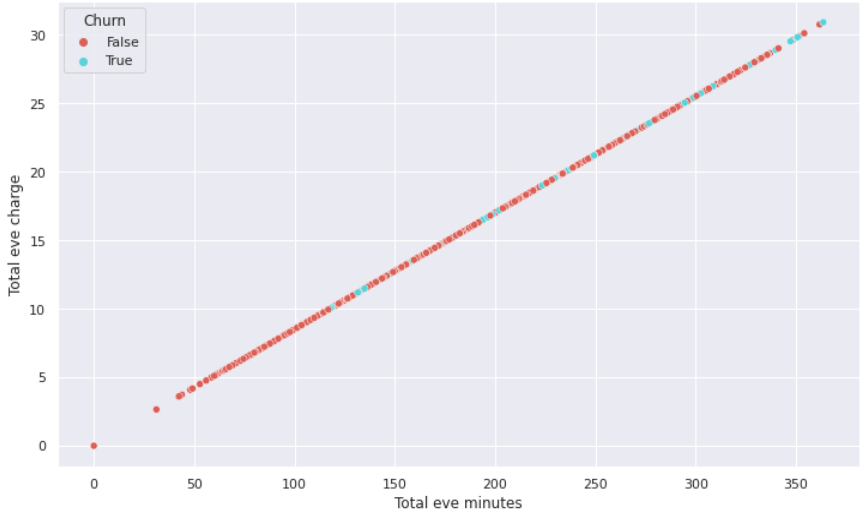




ANALYZING ALL CALLS MINUTES,ALL CALLS, ALL CALLS CHARGE

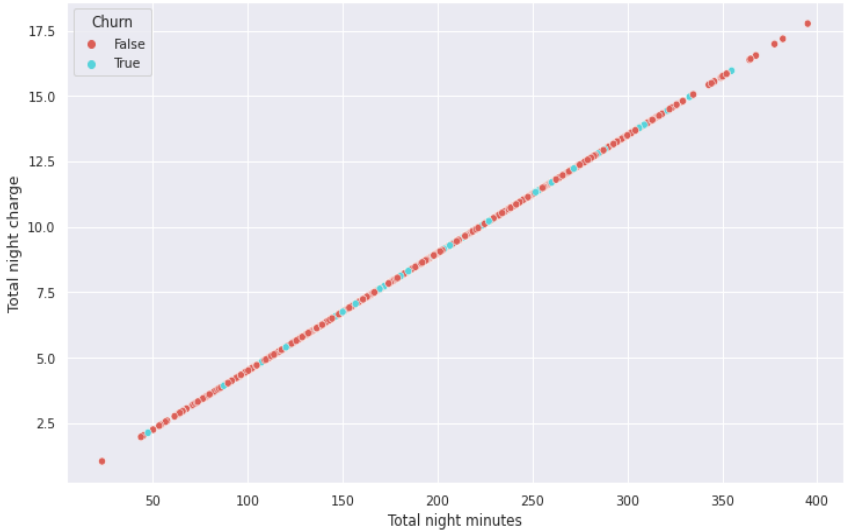
- Below plots are scatter plot which shows the relation between calls and churn
- Left side plot shows the Total day minutes, Total day charge With churn
- Right side plot shows the Total eve minutes, Total eve charge With churn

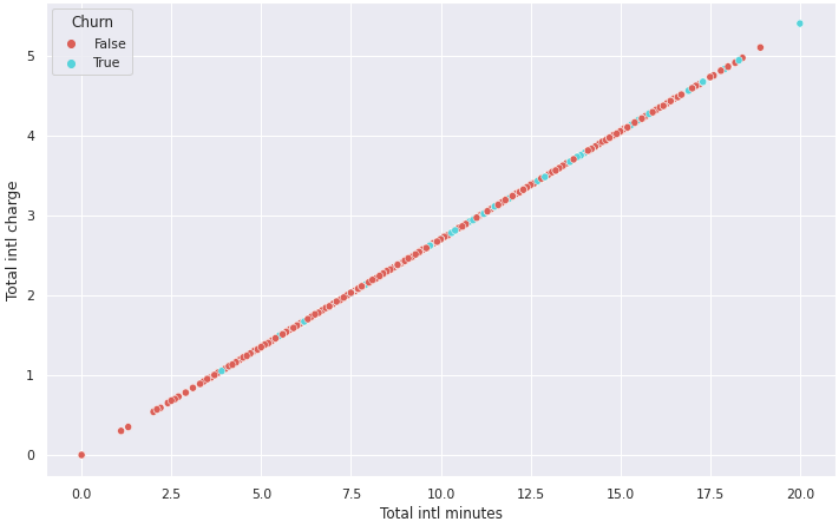




ANALYZING ALL CALLS MINUTES, ALL CALLS, ALL CALLS CHARGE

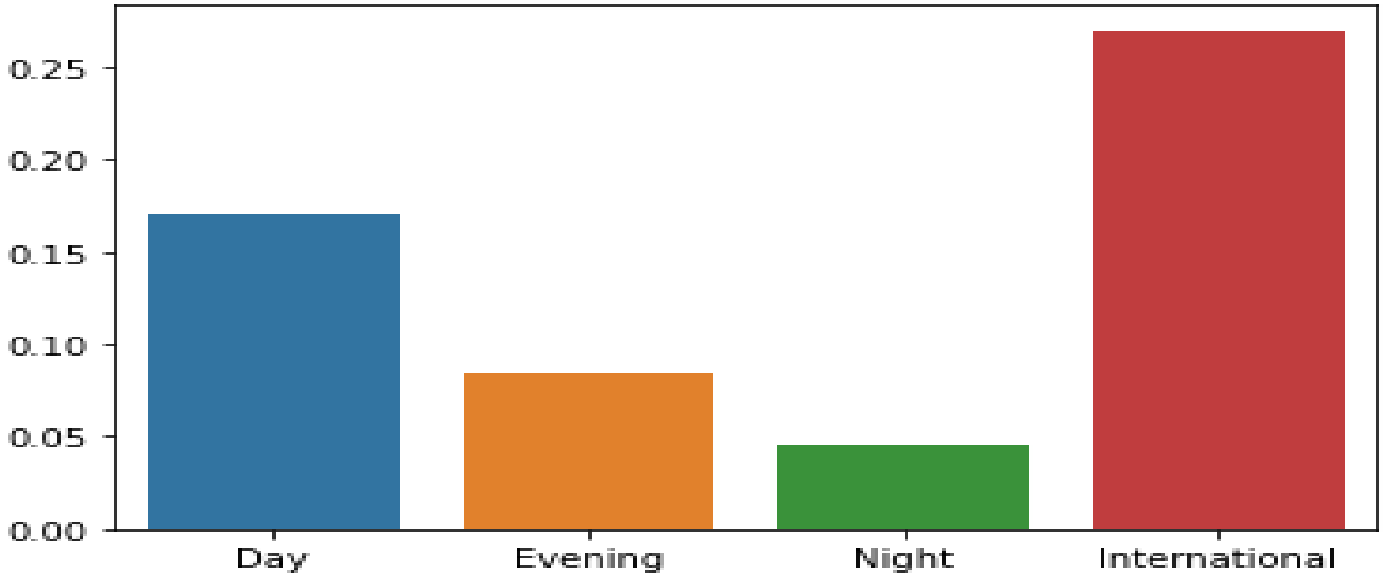
- **Left side scatter plot shows the Total night minutes, Total night charge With churn**
- **Right side plot shows the Total international minutes, Total international charge With churn**



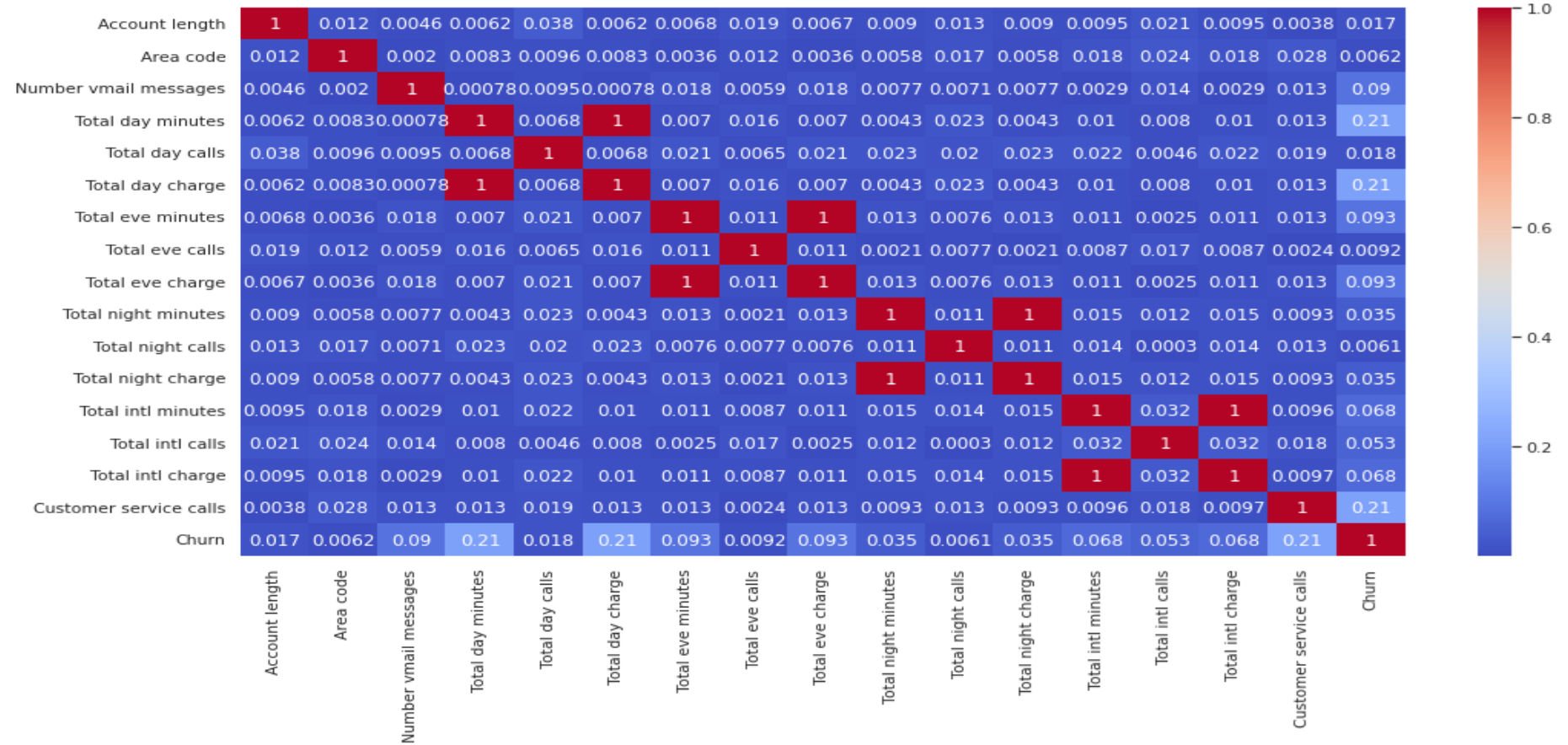


COMPARISON OF CALL CHARGES PER MINUTE

- Below this bar plot shows the comparison between all call charges per minute
- International call charges are high as compare to others it's an obvious thing but that may be a cause for international plan customers to churn out.



CORRELATION MATRIX



CHALLENGES

- Difficult to analyze columns like account length, area code.
- Need to plot lot of graph for columns as to understand the data.
- For calls data there is no direct relation to churn but related column has played some role.

CONCLUSION

- 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.
- In the international plan those customers who have this plan are churn more and also the international calling charges are also high so the customer who has the plan unsatisfied with network issues and high call charge.
- In the voicemail section when there are more than 20 voice-mail messages then there is a churn so it basically means that the quality of voice mail is not good.
- The customer who have high day call minutes also have high call price these customer tends to churn.
- In customer service calls data shows us that whenever an unsatisfied customer called the service centre the churn rate is high, which means the service centre didn't resolve the customer issue.

RECOMMENDATIONS

- Improve network coverage churned state.
- In international plan provide some discount plan to the customer.
- Improve the voicemail quality or take feedback from the customer.
- Provide discount to those customer who spent more minutes.
- Improve the service of call centre and take frequently feedback from the customer regarding their issue and try to solve it as soon as possible.

Q & A

THANK YOU