Content

- PROBLEM STATEMENT
- DATA SUMMARY
- ANALYSIS
- CHALLENGES
- CONCLUSION
- Q & A

Problem Statement

What is Customer Churn?

Customer Churn, also known as **Customer Attrition** is the loss of clients or customers from a particular company or service provider.

This project aims to analyse the data to determine what variables are correlated with customer churn. Additionally a prediction model is built to identify the people that might churn.

Data Summary

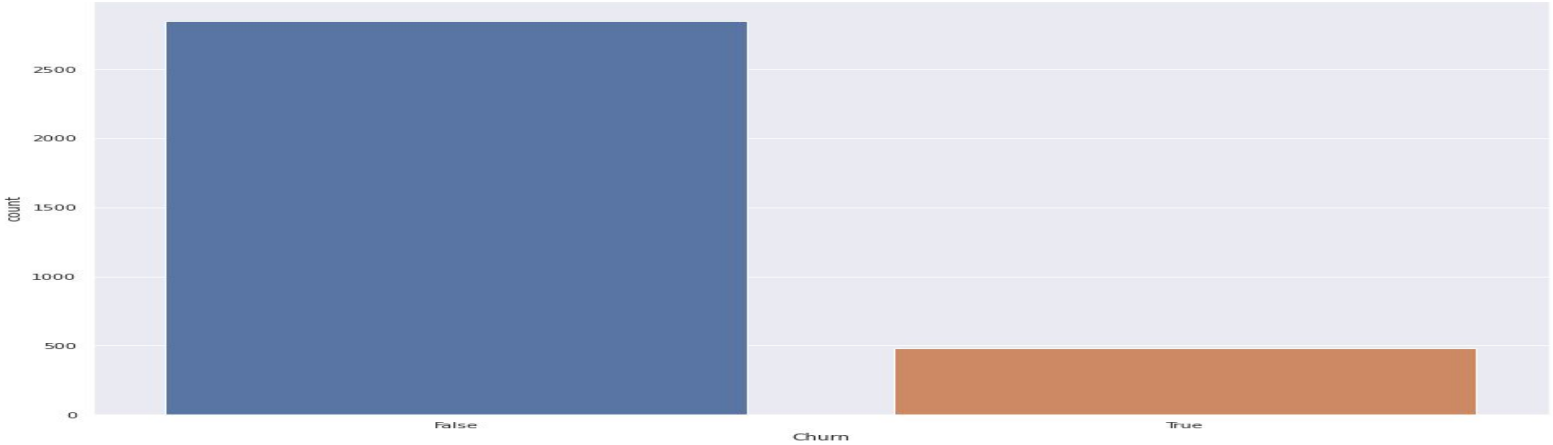
The data below is obtained from Orange S.A, a French multinational telecommunications corporation.

State	Account_length	Area_code	International_plan	Voice_mail_plan	Number_vmail_messages	Total_day_minutes	Total_day_calls	Total_day_charge
KS	128	415	No	Yes	25	265.1	110	45.07
OH	107	415	No	Yes	26	161.6	123	27.47
NJ	137	415	No	No	0	243.4	114	41.38
OH	84	408	Yes	No	0	299.4	71	50.90
OK	75	415	Yes	No	0	166.7	113	28.34

Total_eve_minutes	Total_eve_calls	Total_eve_charge	Total_night_minutes	Total_night_calls	Total_night_charge	Total_intl_minutes	Total_intl_charge	Customer_service_calls	Churn
197.4	99	16.78	244.7	91	11.01	10.0	2.70	1	False
195.5	103	16.62	254.4	103	11.45	13.7	3.70	1	False
121.2	110	10.30	162.6	104	7.32	12.2	3.29	0	False
61.9	88	5.26	196.9	89	8.86	6.6	1.78	2	False
148.3	122	12.61	186.9	121	8.41	10.1	2.73	3	False

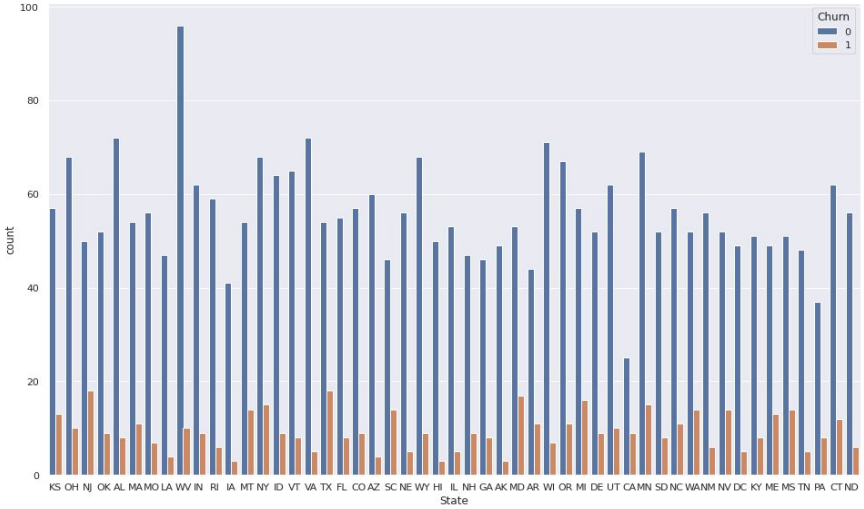
Churn

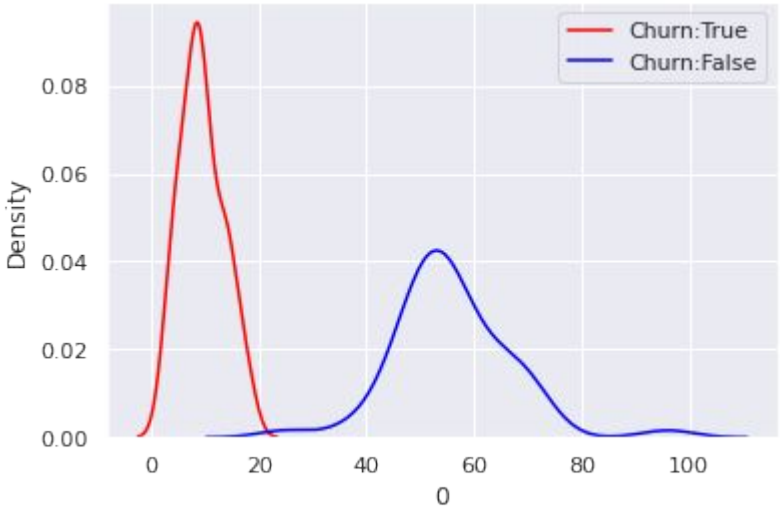
- A countplot of Churn would give us a brief idea of the number of customers who left the service.
- It was found from this analysis that almost 14.5% of customers had churned .



State vs Churn

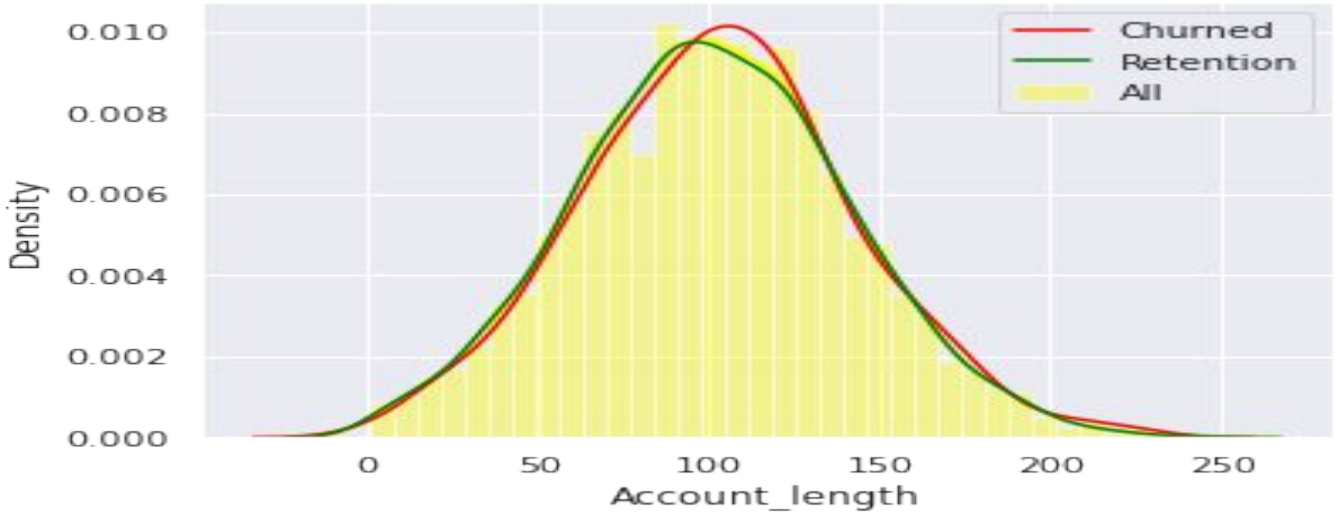
- Plot to check if a State contributed to the churn rate
- Plot on the left shows Churn in each state whereas plot on the right shows the average Churn in total.





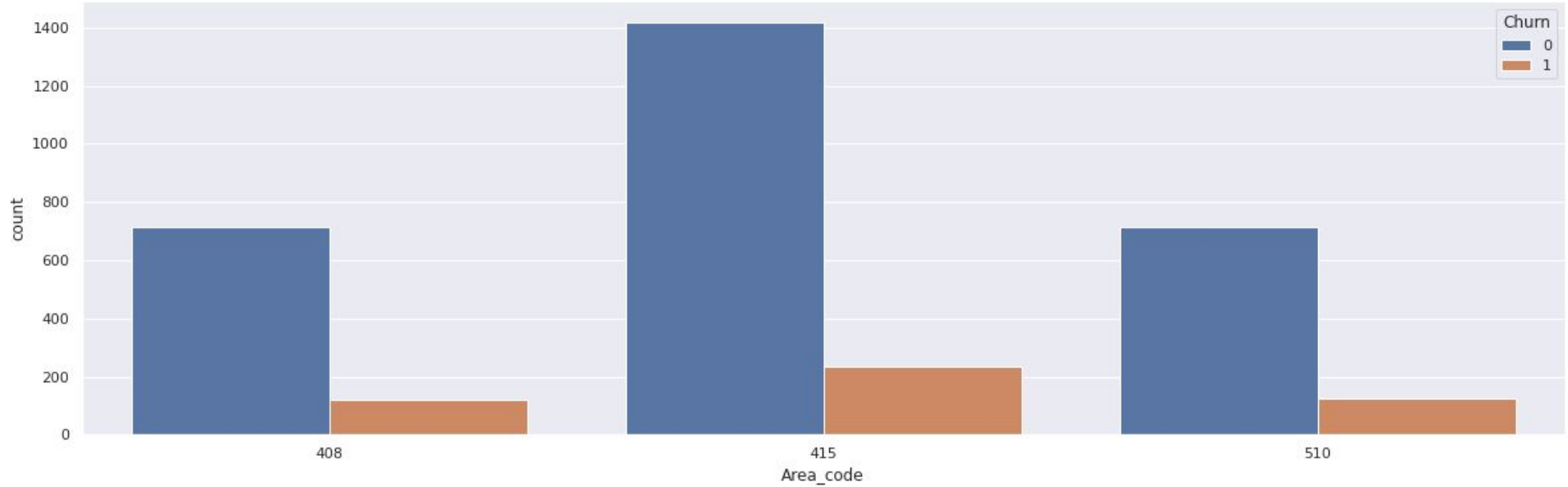
Account Length vs Churn

- Plot to check if the *Account Length* is a factor which affects *Churn*
- In general there is no evidence of customers leaving because of the length of usage of their account.



Area Code

- Area code tells us which area is a subscriber from in a particular state.
- Area code has only 3 unique values, hence a nominal data type.
- Plotting all these values with churn will give us the following plot graph.



Area Code

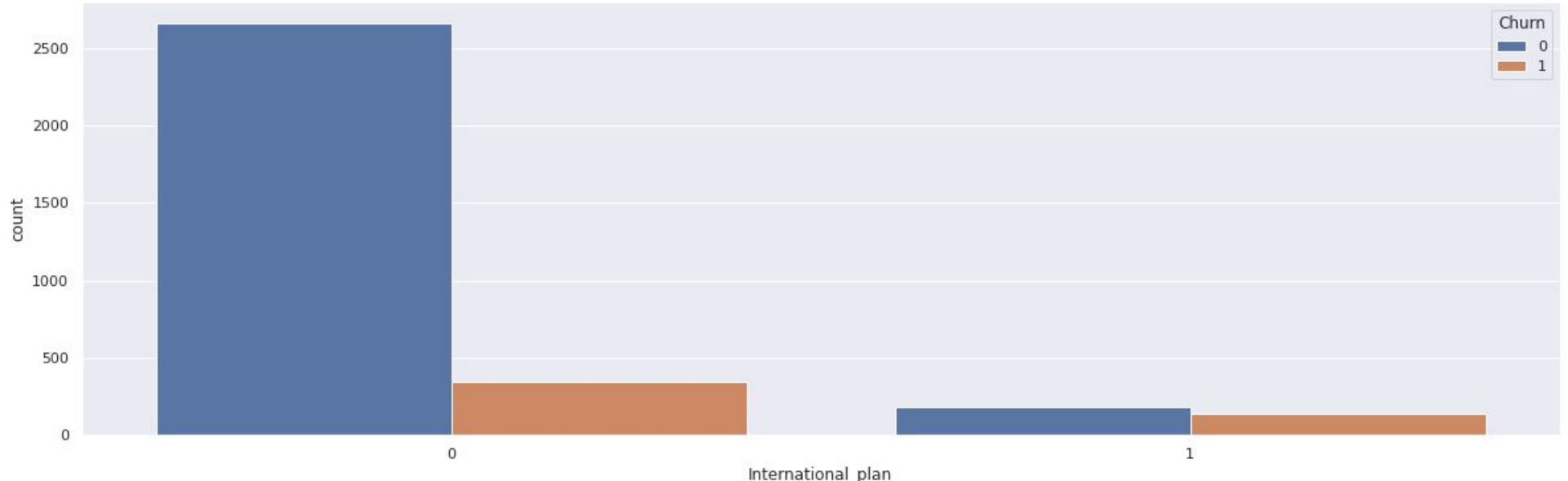
- The Area Code feature does not give us interesting insights.
- The rate of churn for any of the two states is close to each other.

Churn	0	1	Percentage_Churn
Area_code			
408	716	122	14.558473
415	1419	236	14.259819
510	715	125	14.880952

- It is clearly visible that the churn rate for each area code is close to 15%.
- Area code is not a good feature to correlate with our churn.

International Plan

- International plan is an attribute data type with unique values of 0 and 1 corresponding to false and true respectively.
- A plot for churn corresponding the subscription of International plan:



International Plan

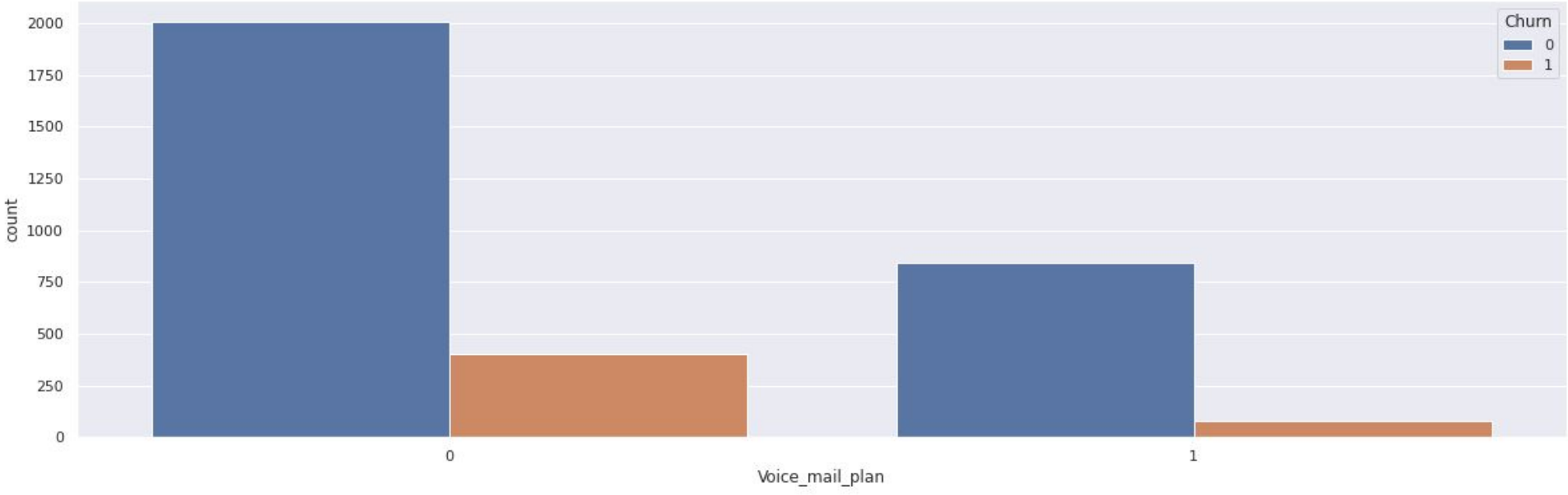
- There was a sharp raise in rate of churn if the subscriber has an International plan.
- Let's jump into numbers and see what percentage of subscribers churn if they have an International Plan.

Churn	0	1	Percentage_Churn
International_plan			
0	2664	346	11.495017
1	186	137	42.414861

- There is something wrong with the international plan, due to which subscribers churn.
- It can be anything, high prices or bad network.
- We will save this information for now and use it later on.

Voicemail Plan

- International plan is an attribute data type with unique values of 0 and 1 corresponding to false and true respectively.
- A plot for churn corresponding the subscription of voicemail plan:



Voicemail Plan

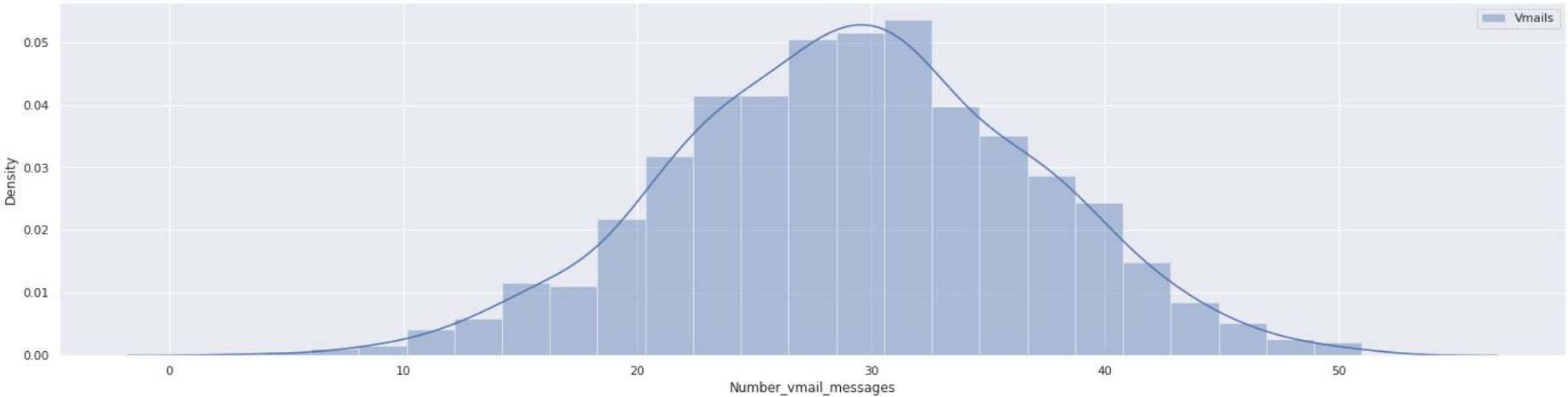
- Although, Voicemail plan subscribers churn too, but as compared to subscribers who do not have voicemail plan, those who have voicemail plan tend to stay.

Churn	0	1	Percentage_Churn
Voice_mail_plan			
0	2008	403	16.715056
1	842	80	8.676790

- The churn rate is half for subscribers with a voicemail plan.

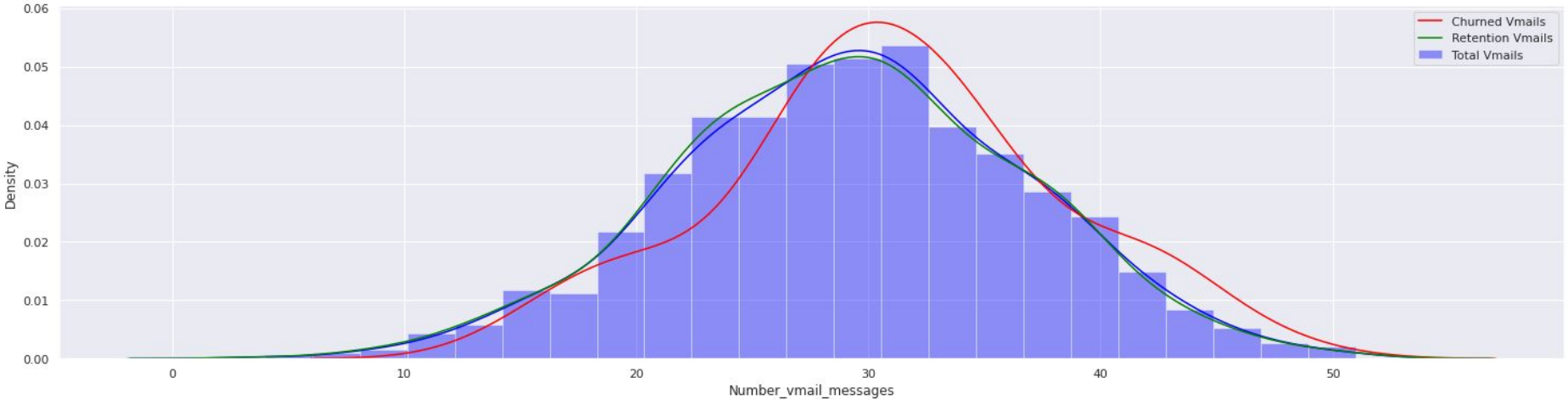
Number of Voicemail Messages

- Number of voicemail messages is defined only for the subscribers with voicemail plans.
- Below is a distribution plot for number of v-mails for all subscribers.



Number of Voicemail Messages

- Judging by number of voicemails everyone has made is difficult.
- Separate graphs for subscribers churned and retended subscribers can be plotted.



- Both, Churned and Retended subscribers are almost distributed equally.

Calls

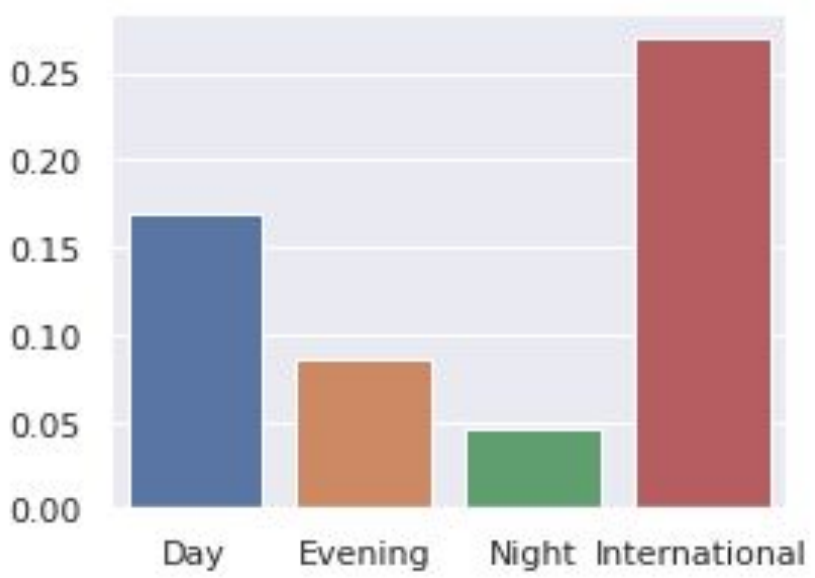
Total	Day	Evening	Night	International
Calls	Total Number of Day Calls	Total Number of Evening Calls	Total Number of Night Calls	Total Number of International Calls
Minutes	Total Day Minutes	Total Evening Minutes	Total Night Minutes	Total International Minutes
Charges	Total Charges charged for Day Calls	Total Charges charged for Evening Calls	Total Charges charged for Night Calls	Total Charges charged for International Calls

Calls

- All these columns may have a lot of data but they provide limited information.

Price Per Minute

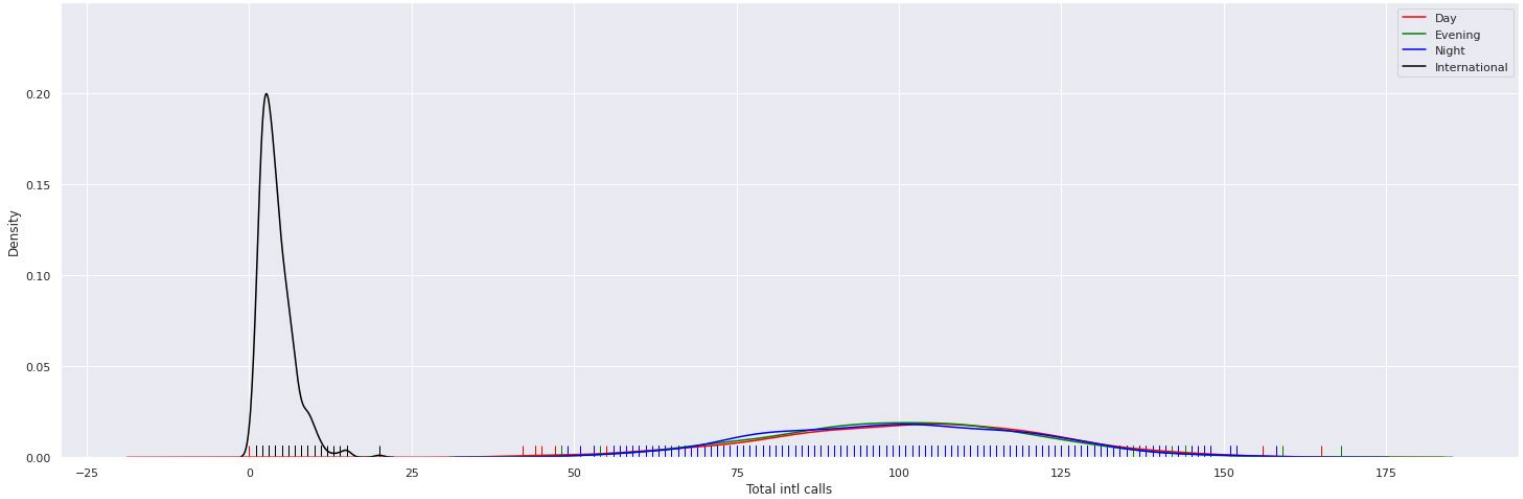
- The price per minute for each call is calculated by taking the ratio of total charges for the subscriber and total number of minutes for the subscriber.
- We calculated the mean price per minute for all categories of calls.



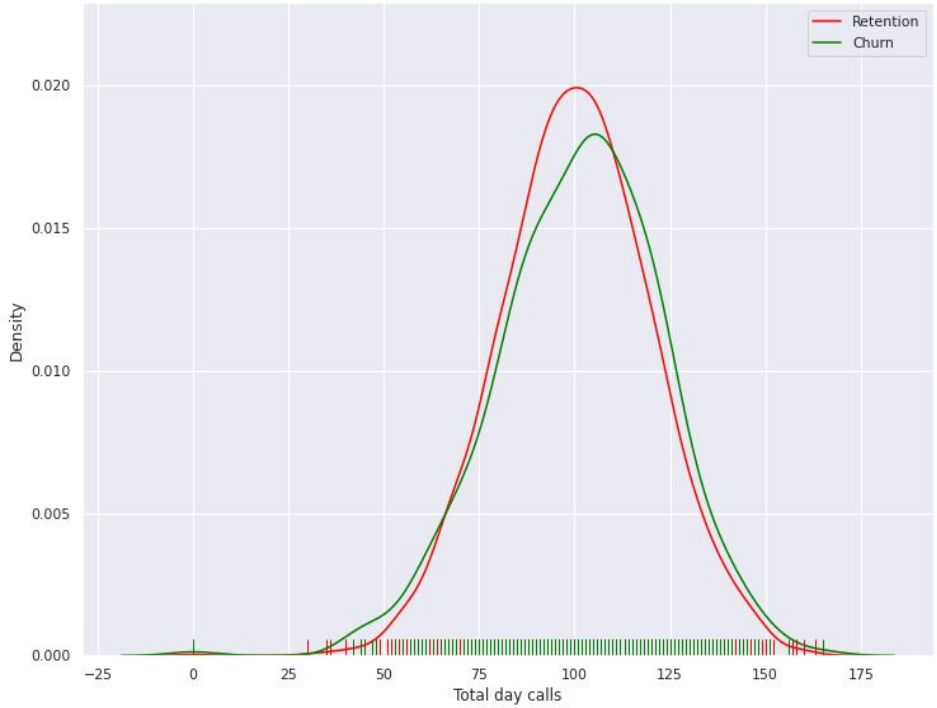
Calls

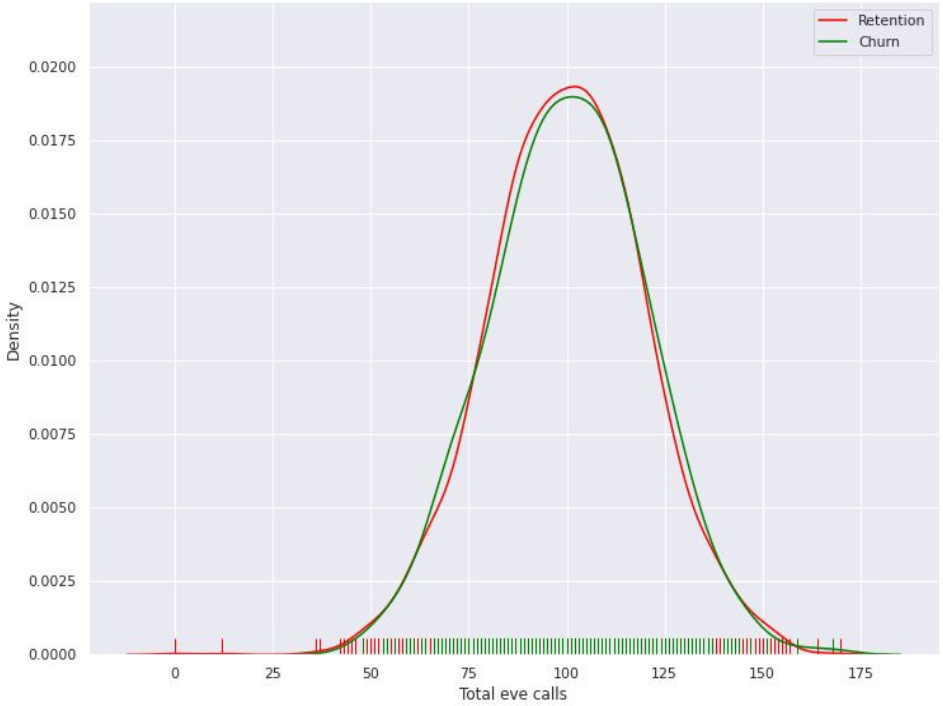
- The price for international calls is one and a half times as compared for day and almost six times the price of night call charges.
- That can be a reason why subscribers with international plan tend to churn.

Distribution of total calls.

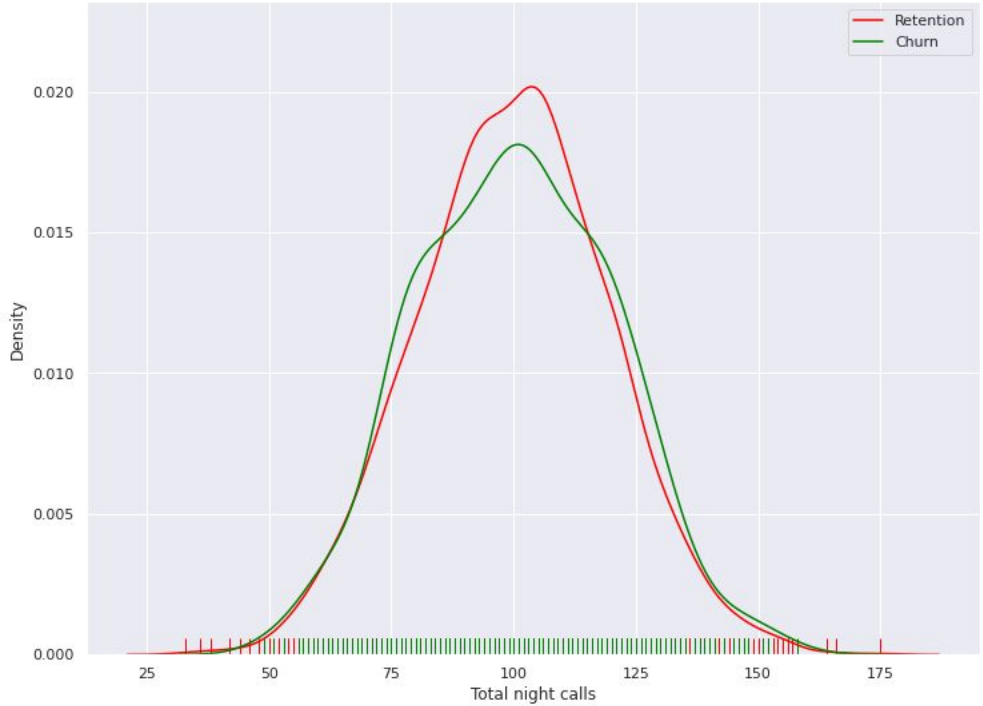


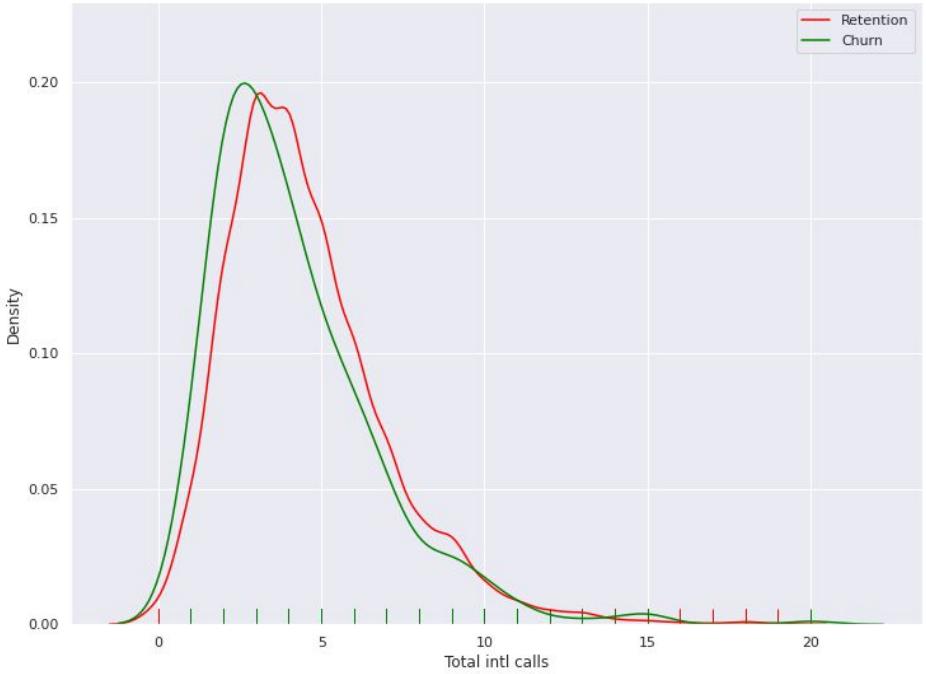
Calls


Retention
Churn
0.020
0.015
0.010
0.005
0.000
Density
−25
0
25
50
75
100
125
150
175
Total day calls


Retention
Churn
0.0200
0.0175
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
Density
0
25
50
75
100
125
150
175
Total eve calls

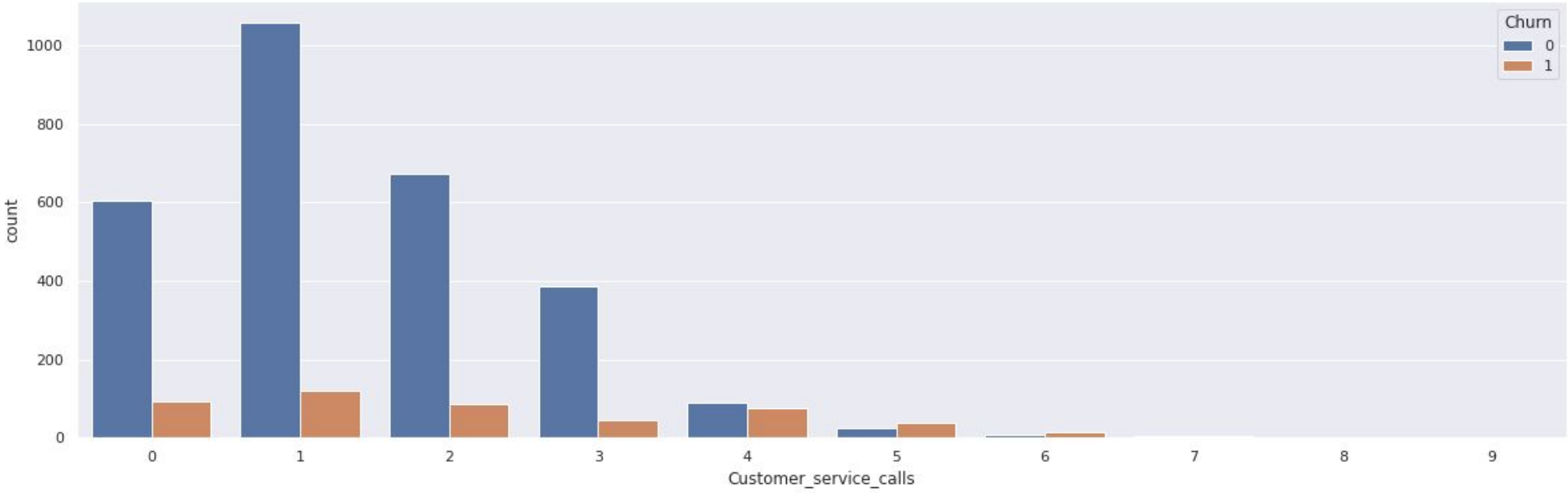
Calls


Retention
Churn
Density
0.020
0.015
0.010
0.005
0.000
25
50
75
100
125
150
175
Total night calls


Retention
Churn
Density
0.20
0.15
0.10
0.05
0.00
0
5
10
15
20
Total intl calls

Number of Customer Service Calls

- The number of total customer calls can reveal a lot about the customer service at the company.
- Here is a plot of Churn for discrete number of customer calls.



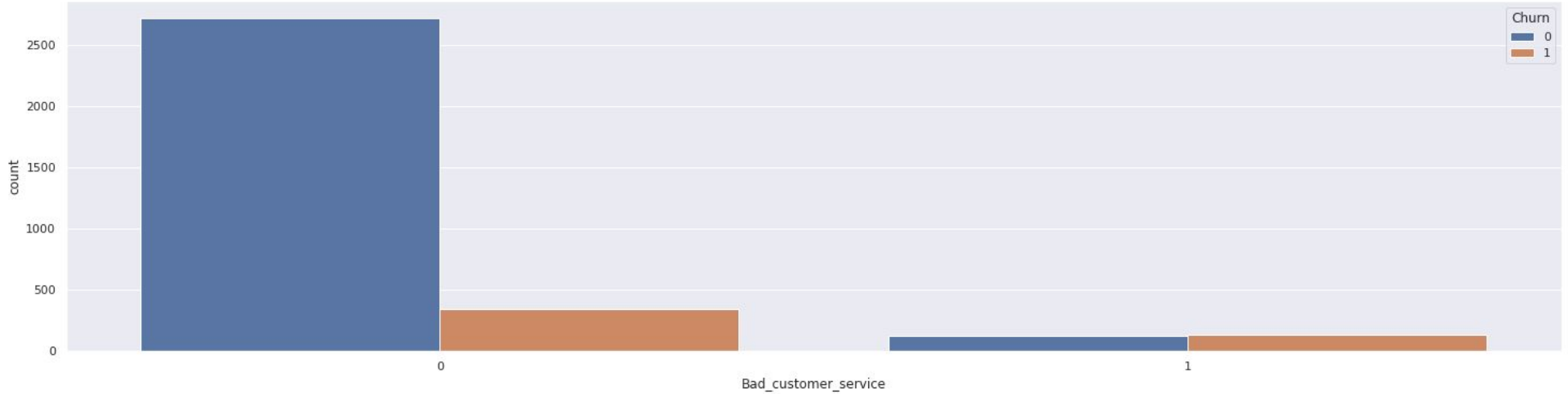
Number of Customer Service Calls

- The churn rate sharply increases after customer 3 number of customer calls.
- This information can be useful and can be used later to categorise the service.
- On the right is a table mapping number of customer calls to the churn percentage.
- It's clear that after 3 calls at least 45% of the subscribers churn.

Churn	0	1	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

Bad Customer Services

- To analyse customer service calls further we divided it into two sub categories.
- Here is a plot, it shows the two categories, one is good service where calls are less than 3, and bad service, where calls are greater than 3.



Bad Customer Services

- Below are finding from above plot:
  If customer calls are less than 3 the churn is 11.25%.
  If customer calls are greater than 3 the churn is 51.68 %.

Churn	0	1	Percentage_Churn
Bad_customer_service			
0	2721	345	11.252446
1	129	138	51.685393

Multivariate Analysis

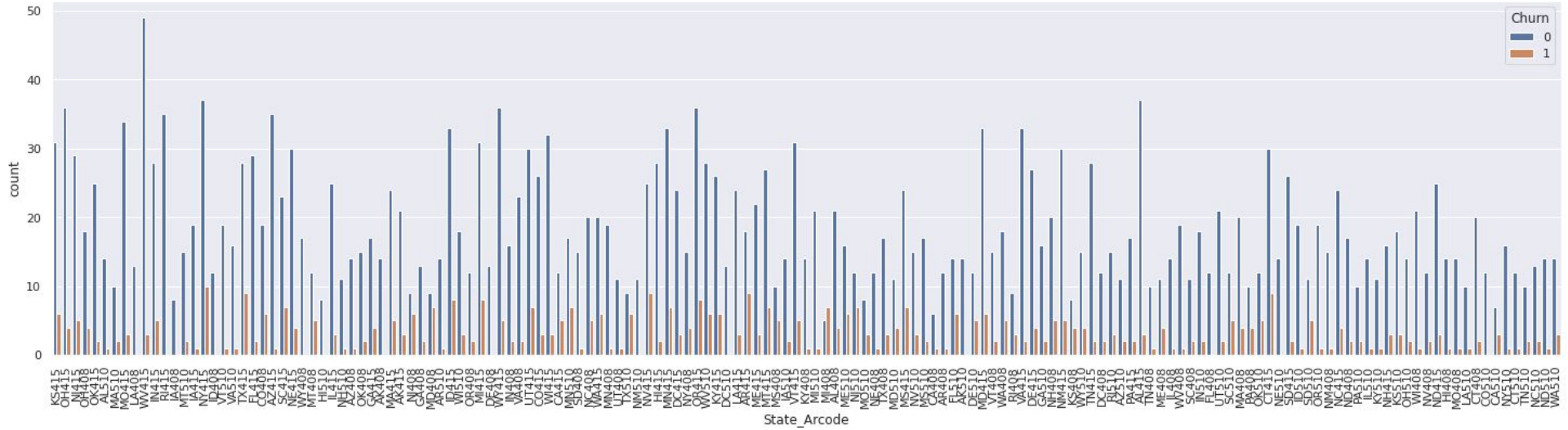
- After analysing every column of dataset individually against Churn, we decided to plot the combination of variable to see if they can provide us more insight.

- As when we plotted area code and state separately, we haven’t got much information. So we combined state and area code, to see which of them is having high churn rate..

- Also, will discuss how bad customer service and international plans combinely affect customer churn.

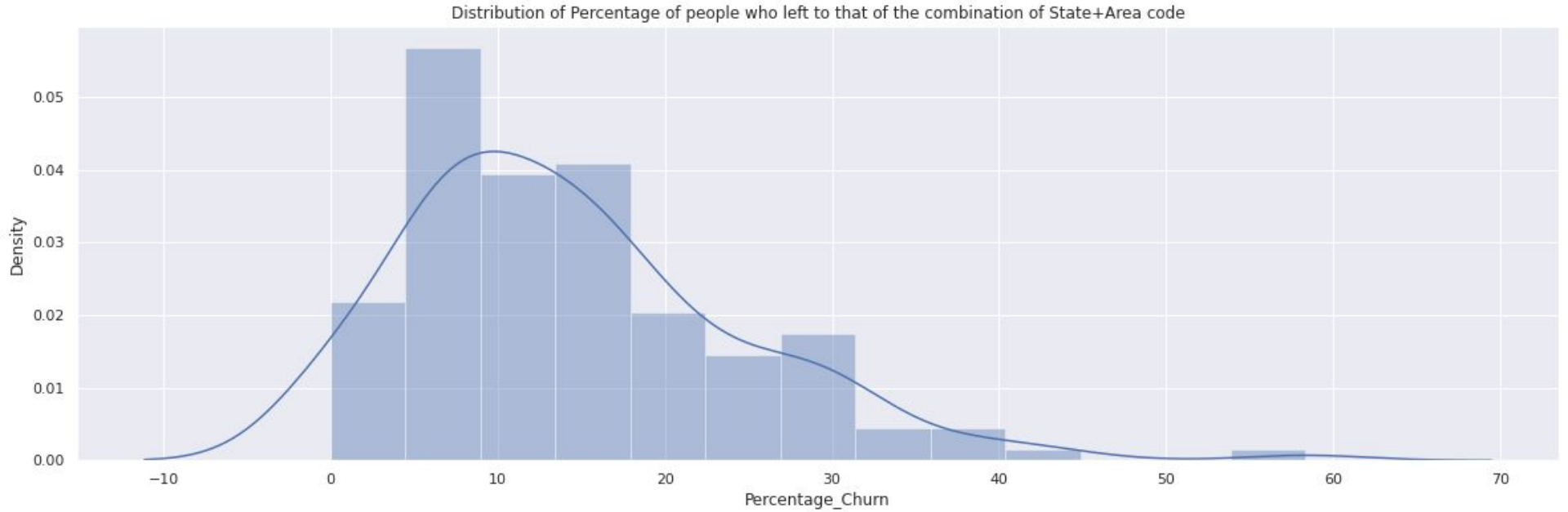
State and Area Code

- As we know , we have in total 51 States and 3 unique area codes.
- Below graph represents all possible combination of state and area code against Churn



State and Area Code

- Below graph represents how churn percentage is distributed among all states.
- It also depicts that most of states are having churn as 10 %.



State and Area Code

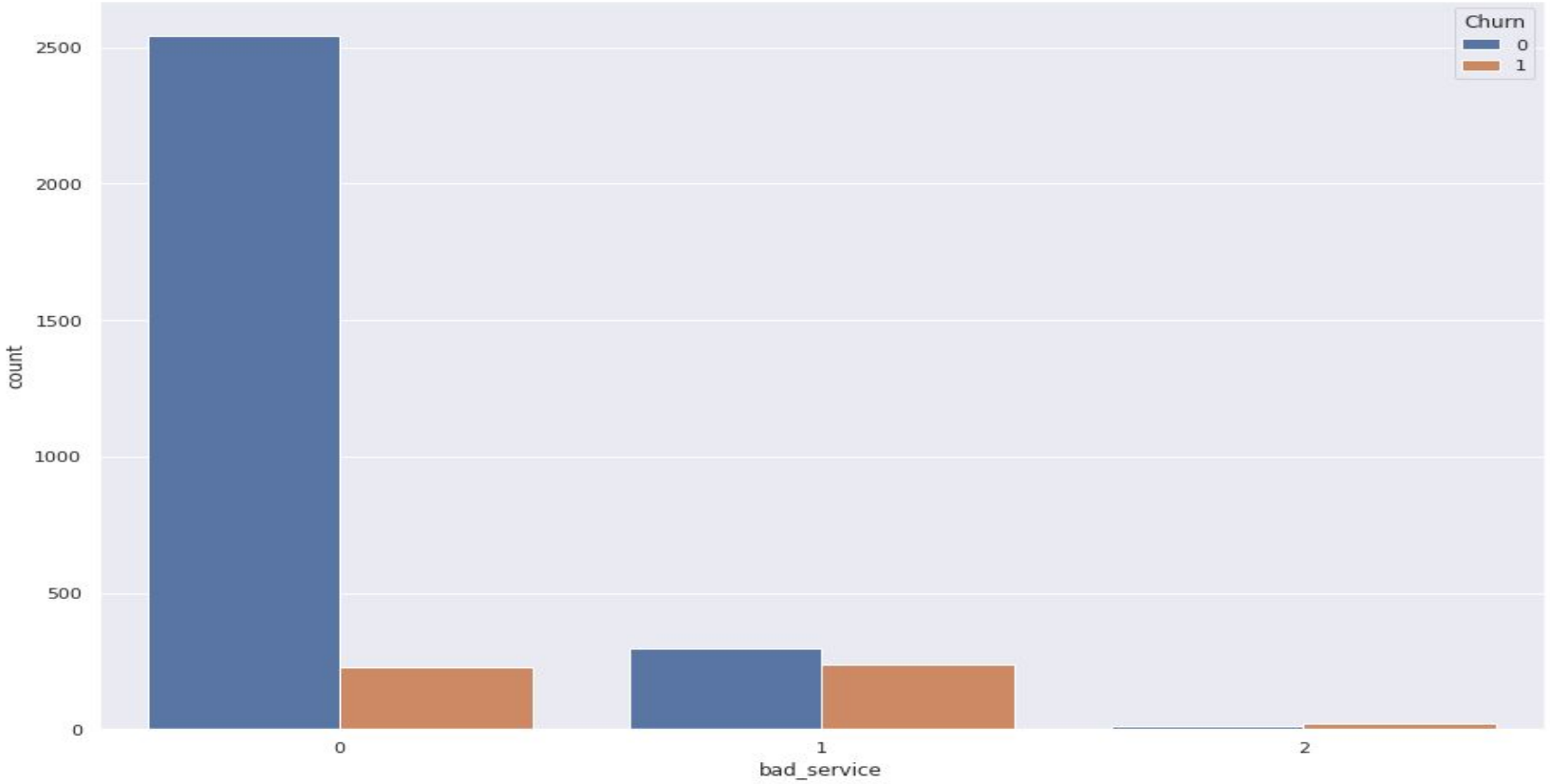
- If we filter states_areacode having higher churn percentage keeping filter limit as 40 %, from above mentioned graph, we will get this results.
- This results shows the state and area code which are highly affected by Churn

Churn	False	True	Percentage_Churn
State_AreaCode			
MD408	9	7	43.750000
MI408	5	7	58.333333
NJ408	9	6	40.000000
TX510	9	6	40.000000

Composite Factor

- So to dig in deep we combined factors like international plan and customer service quality(because of their contribution to churn), and tried to plot a new Composite factor containing a mix of both.
- Lets see how it goes..

Composite Factor



Composite Factor

From above graph following are 3 scenarios:

- If 0: then customer is not using international plans and also having either no or less than 3 customer calls. So by looking at graph we can say more people are tend to stay.

- If 1: then customer is either having international plan or more than 3 customer calls. So more people are leaving.

- If 2: then people are having international plan and having so many customer service call, which shows us that they are might not happy with services or facing various issues.

Challenges

Throughout this journey from analysing data to finding conclusions , below are few challenges that we faced:

- When we started with data we were not sure which column to consider, so we tried to map everything.
- Later we got so many insights , but it was being difficult to reach to proper conclusions.
- We have also faced timing issues among us , so we tried solve this issue by discussing and by finding suitable time frame when we all can discuss and contribute.
- Atlast we were not able to conclude our analysis, so we tried using composite factors , which really helped us to find our answers.

Conclusion

So after analyzing the dataset, below are some key findings:

- Customers opted for international plan, are more likely to leave.
- Voice mail plan service is used by many existing customers.
- International call charges are higher than , any other call charges.
- No. of customer service call is directly proportional to churn rate.
- There is set of state and area code combination, where churn rate is 40%.

Recommendations

Here are few recommendations on the basis of our basic analysis:

- Telecom service can consider revising there international plans , like providing free incoming calls or reducing price.
- Improvement is required in customer service support. We should work on solving issues related to every customer , and try to gain customer satisfaction .
- Few state and area code having higher churn rate, we should try giving better offers to existing customer, and better support so they will not churn.

Q & A