Churn reduction for a telecom company

1. What is churn?

Churn or customer attrition is a business metric that shows customers that stopped doing business or stopped using a particular service with a company. When a customer churns it means a company has lost that customer.

2. Why care about churn?

The goal of every business is to maximize profits. Research done by *Bain & Company* shows that for example in financial services, a 5% increase in customer retention (avoiding customer attrition) produces more than a 25% increase in profit. Many companies spend a lot of money on customer acquisition (attracting new customers), however, it has been shown that it can cost five times more to attract new customers than the cost spent in retaining them.

Customer retention and customer acquisition are both important for a company but focusing more on customer retention have more benefits such as building loyal relationships with customers, reducing customer acquisition costs, and even gain more insights on how to improve your business. The probability of selling a product or a service to an existing customer is between 60% and 70% while selling to new customers ranges between 5% and 20%.

By predicting which customer is going to churn can allow businesses to take appropriate actions to prevent losing a customer. So in this project, we are going to build a predictive churn model and see what kind of actions to take in order to reduce customer attrition.

3. Data

Data that was used in this study is of a telecom company, it has up to 20 features with a feature called churn which tells us if a customer has churned or not. This churn feature will be our target.

	state	account length	area code	phone number	international plan	voice mail plan	number vmail messages	total day minutes	total day calls	total day charge	...	total eve calls	total eve charge	total night minutes	total night calls	total night charge	total intl minutes	total intl calls	ch
0	KS	128	415	382-4657	no	yes	25	265.1	110	45.07	...	99	16.78	244.7	91	11.01	10.0	3	
1	OH	107	415	371-7191	no	yes	26	161.6	123	27.47	...	103	16.62	254.4	103	11.45	13.7	3	
2	NJ	137	415	358-1921	no	no	0	243.4	114	41.38	...	110	10.30	162.6	104	7.32	12.2	5	

3 rows × 21 columns

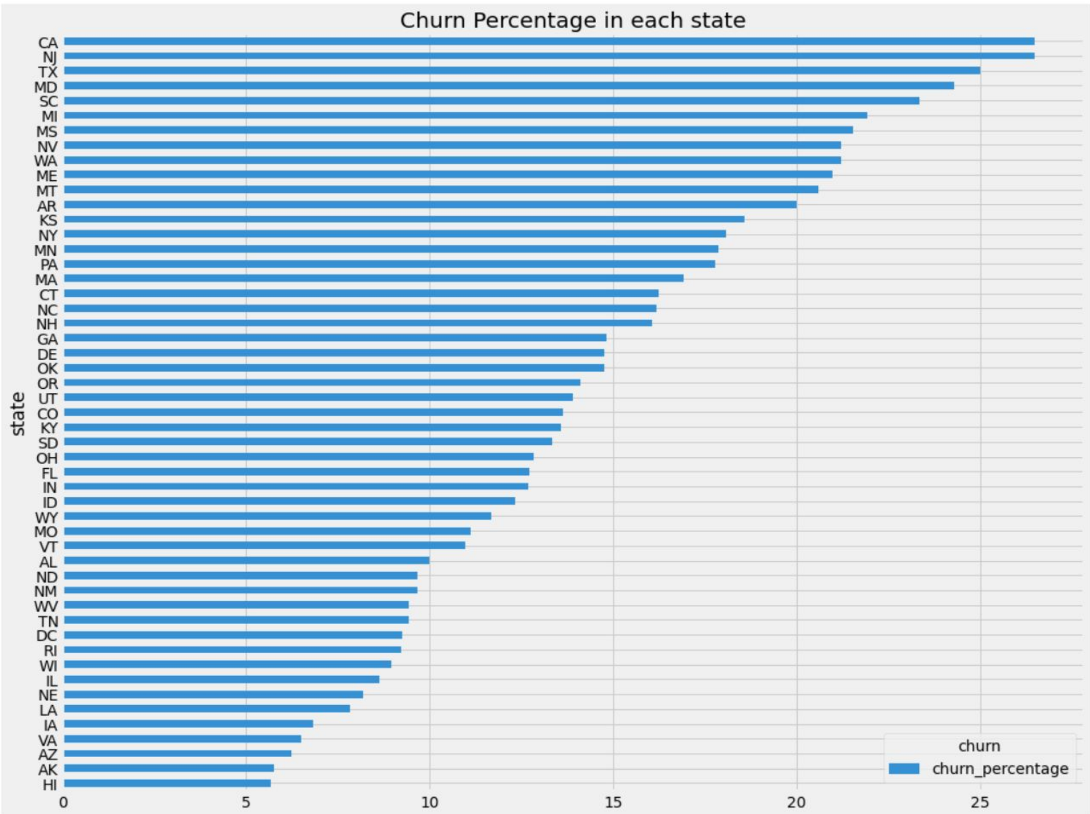
4. Exploratory Data Analysis

Before building a model, we have to understand our data and see if we can get some useful insights. Building a churn predictive model is not always the only solution of how to fight customers churning but simple insights from analyzing your data can help you to reduce customer attrition. Matter of fact many roles of how to reduce churn are held by product managers, content creators, or even account managers (not necessarily data scientists). All those people rarely use predictive models to come to their decisions.

4.1 Churn Percentage in Each State

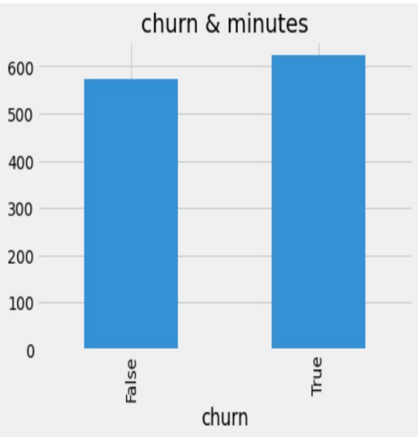
Many people will think that if a place has more phone lines then the churn rate will be higher, but according to this dataset, the churn rate does not depend on how many lines are there. The total number of lines from the top 10 states with a high churn rate is 636 which is almost the same number of lines (603) in the top 10 states with a low churn rate.

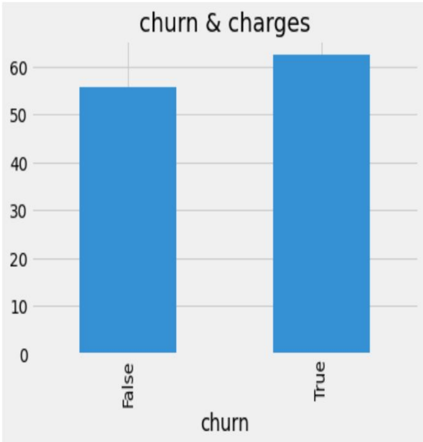
Another example is the state of Mississippi and Rhode Island, the two states have the same number of phone lines yet Mississippi has a 21.5% of churn rate while Rhode Island has 9.2%. So I don't think the number of phone lines has a significant impact on the churn rate in each state.

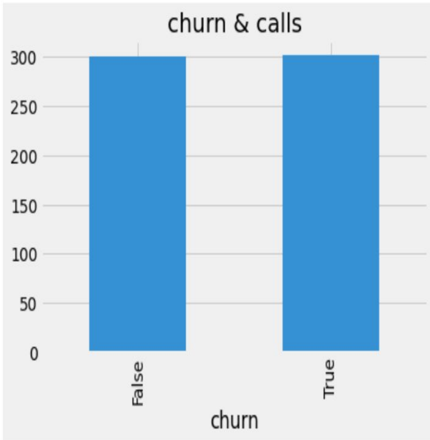


4.2 Churn vs total number minutes, the total number of calls, charges, and international services

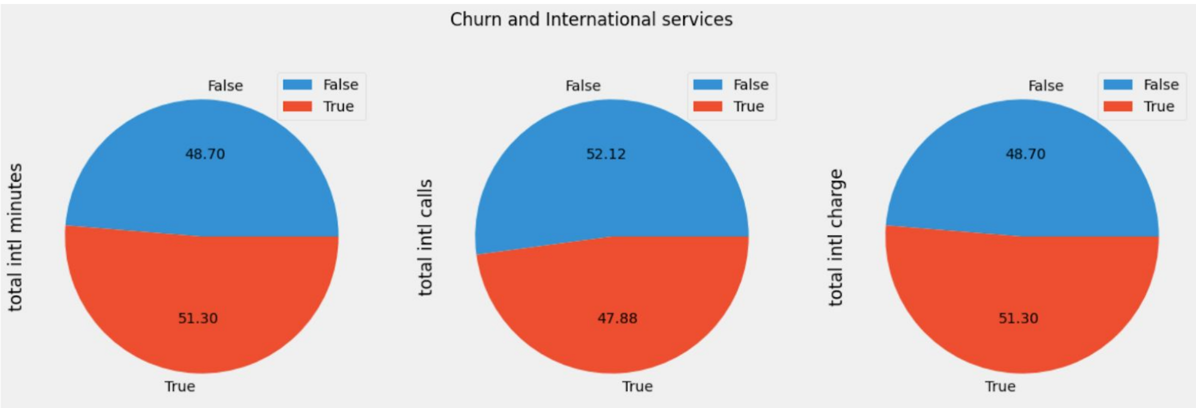
Looking at the graphs below, we see that the numbers are almost the same but the numbers of people who churn are a little bit higher than those who do not churn. People who churn spent more minutes on average more than those who did not churn, they had more calls and even more international minutes.





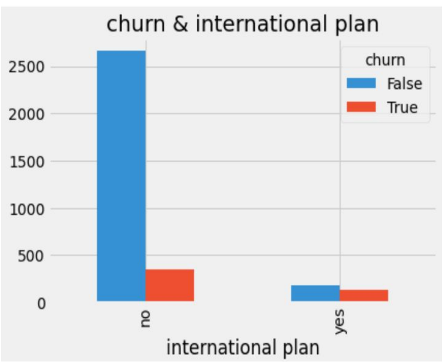


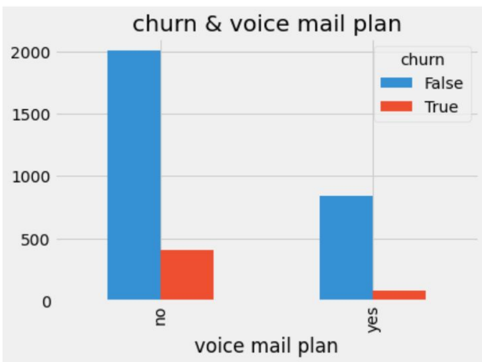
The only time where people who do not churn has a bigger number is the total number of international calls by 52% of all calls.



4.2 Churn vs International Plan and Voice mail plan

Almost 15% of all lines churn, this affects people who did not churn to have more numbers in all these categories. For people that have International plans and voice mail plans, many of them do not churn.





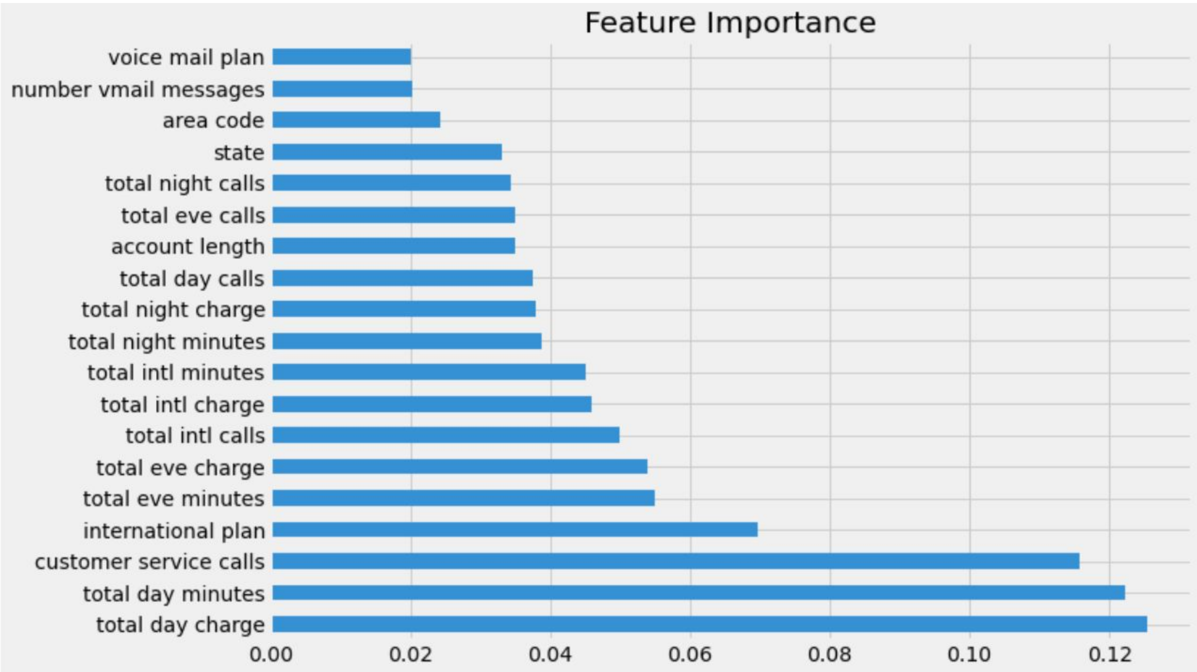
5. Feature Engineering, Pre-processing, and Processing

Before building a model we have to replace categorical features namely *area code, state, international plan, voice mail plan* and *churn* into numerical ones. After we have to set our target which the feature *churn* and then split the data into the training and testing data.

```
X_train
```

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	to nig char
5	72	2	0	1	39	92.8	98	15.78	271.2	115	23.05	167.1	83	7.
14	97	2	0	0	0	239.8	125	40.77	214.8	111	18.26	143.3	81	6.
7	147	1	0	0	0	124.4	74	21.15	320.9	78	27.28	157.2	126	7.
4	66	3	0	0	0	154.0	133	26.18	198.9	121	16.91	151.9	100	6.
48	117	2	0	0	0	191.1	93	32.49	282.8	56	24.04	84.8	118	3.

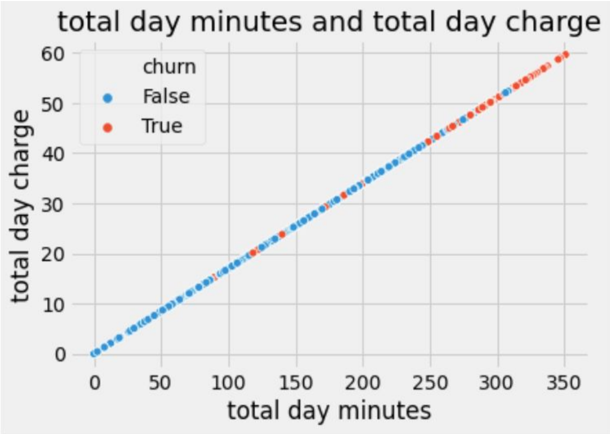
Trying, testing, and selecting models can be time-consuming, for processing, with the use of Tpot library, the best pipeline is selected, in our case, it is the Extra Trees Classifier with an internal CV score of 0.95, and when tested to testing data it had a score 0.93 out of 1.
After training the model we discover important features that have significant effects on our target (Churn).



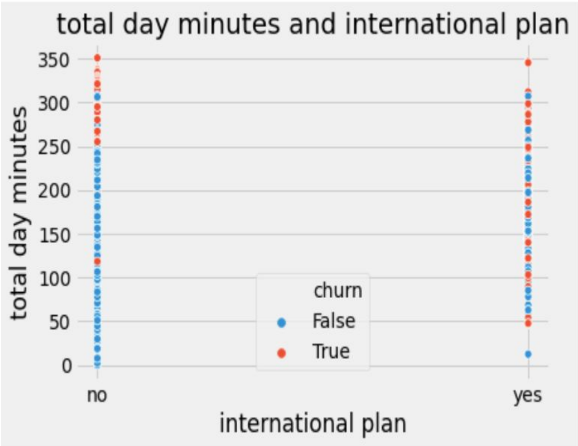
To derive the best insights from this model, we will focus on the most significant features namely '*today day charge', 'today day minutes', 'customer service calls',* and '*international plans'.*

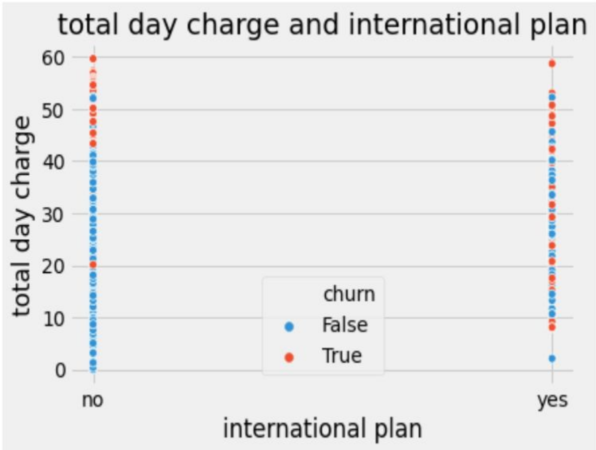
6. Final Insights and Action recommendation

Conclusion 1: People with more ‘ *today day charge’* and *‘today day minutes’* will churn more.

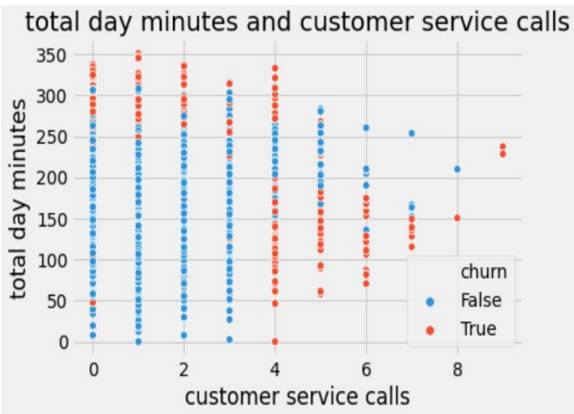


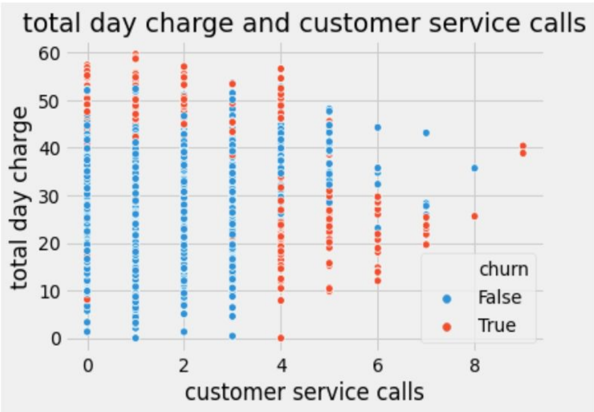
Conclusion 2: People with international plans will churn more despite the number of minutes or charges they have.





Conclusion 3: People with a lot of customer calls will churn more than those with fewer customer calls





Actions recommended:

- Avoid more than 3 customer service calls per customer
- Resolve the root cause of people with international plans to churn or remove it
- Focus more on local plans and on states with a low churn rate.

References:

1. What is churn: wikipedia.org/wiki/Customer_attrition
2. Bain & Company Research: https://media.bain.com/Images/BB_Prescription_cutting_costs.pdf
3. Customer Acquisition vs Customer Retention: https://www.invespcro.com/blog/customer-acquisition-retention/
4. Data: bigml.com/user/francisco/gallery/dataset/5163ad540c0b5e5b22000383
5. How to reduce churn with data science: fightchurnwithdata.com/how-to-reduce-churn-with-data-science/
6. Tpot Library: epistasislab.github.io/tpot/
7. Extra Trees Classifier:
8. scikit-learn.org/stable/modules/generated/sklearn.ensemble.ExtraTreesClassifier.html