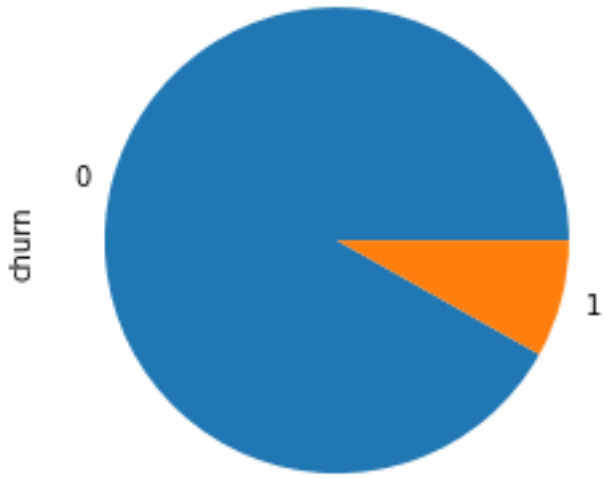
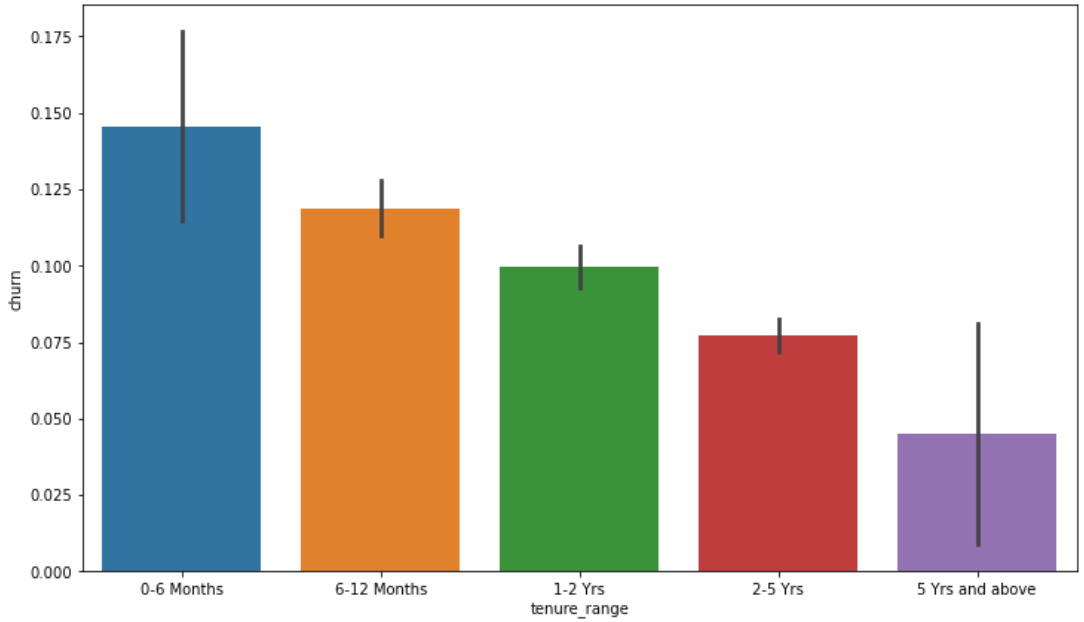
Telecom churn

Case study





As we can see that 91% of the customers do not churn, there is a possibility of class imbalance

It can be seen that the maximum churn rate happens within 0-6 month, but it gradually decreases as the customer retains in the network.

- Avg Outgoing Calls & calls on romaning for 6 & 7th months are positively correlated with churn.
- Avg Revenue, No. Of Recharge for 8th month has negative correlation with churn.
- As the number of recharge rate increases, the churn rate decreases clearly.

- Accuracy of the logistic regression model with PCA: 0.7548050216642596

•Customers who churn show lower average monthly local incoming calls from fixed line in the action period by 1.27 standard deviations , compared to users who don't churn , when all other factors are held constant. This is the strongest indicator of churn.

•Customers who churn show lower number of recharges done in action period by 1.20 standard deviations, when all other factors are held constant. This is the second strongest indicator of churn.

•Further customers who churn have done 0.6 standard deviations higher recharge than non-churn customers. This factor when coupled with above factors is a good indicator of churn.

•Customers who churn are more likely to be users of 'monthly 2g package-0 / monthly 3g package-0' in action period (approximately 0.3 std deviations higher than other packages), when all other factors are held constant.

Based on the indicators the recommendations to the telecom company are :
•Concentrate on users with 1.27 std devations lower than average incoming calls from fixed line. They are most likely to churn.
•Concentrate on users who recharge less number of times (less than 1.2 std deviations compared to avg) in the 8th month. They are second most likely to churn.
•Models with high sensitivity are the best for predicting churn. Use the PCA + Logistic Regression model to predict churn. It has an ROC score of 0.87, test sensitivity of 100%