Telecom Churn Analysis Capstone Project

By Matthew Smith

Problem Statement

A california telecom company is faced with customer churn of 26%.

What are some potential solutions to reduce churn?

Findings

- Customers indicated the top 3 reasons they churned were:
  - Competitor has better devices
  - Competitor making a better offer
  - Attitude of Customer Support
- The most predictive feature of a customer's likelihood to churn is tenure

Suggestions

- Add contract terms

- Offer lower priced plans
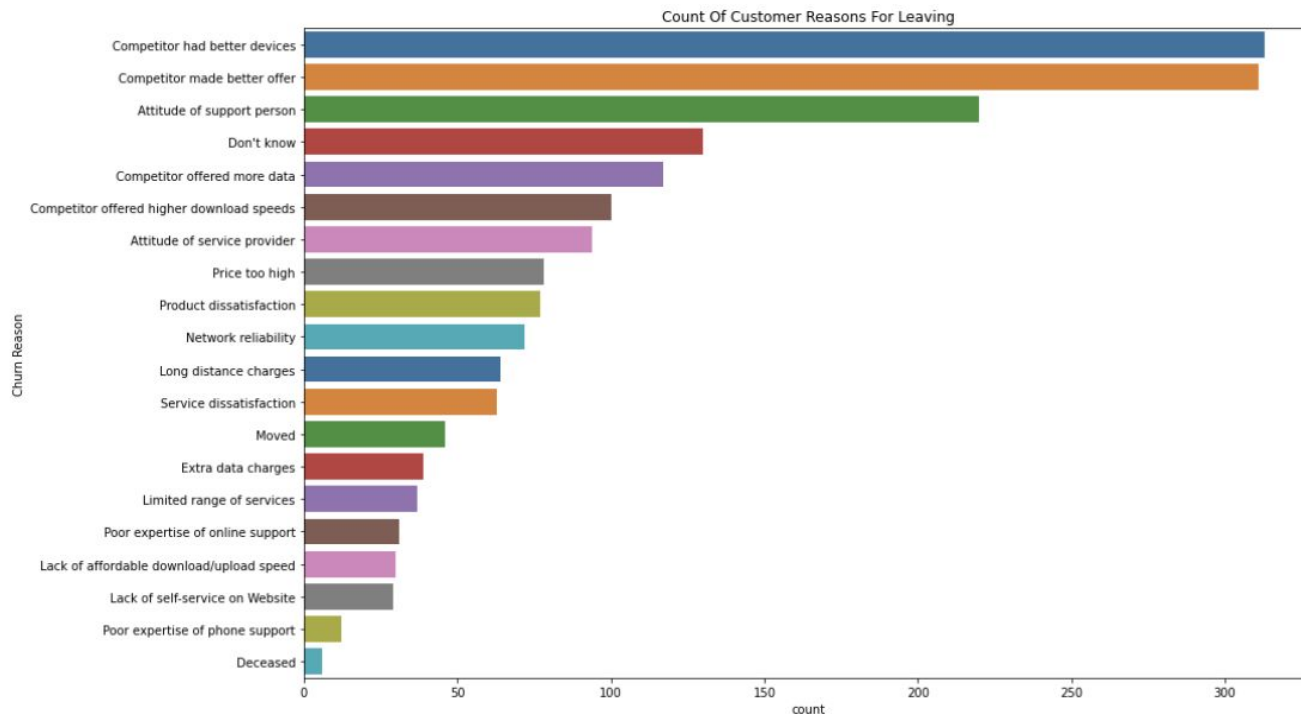
- Re-train customer support

Data Wrangling

Original Dataset:

https://www.kaggle.com/datasets/shilongzhuang/telecom-customer-churn-by-maven-analytics?resource=download
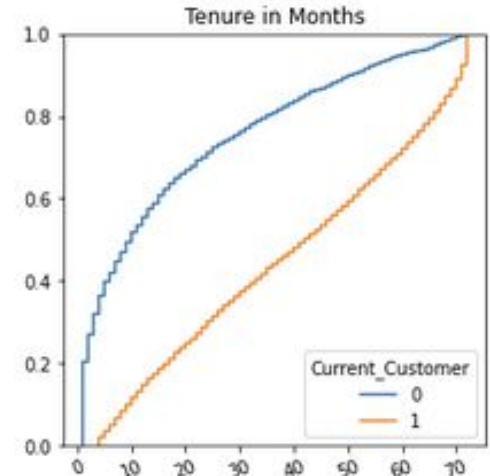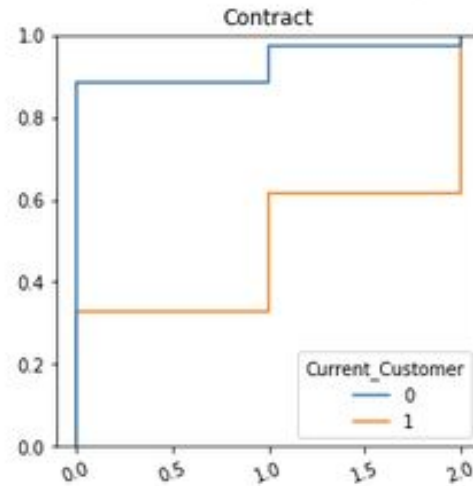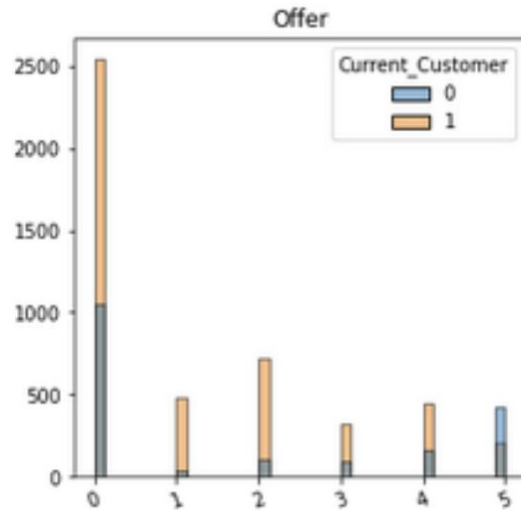
Key steps taken:

- Imputed missing values
- Created numeric categorical variables for appropriate categories

Reasons Given For Customer Churn

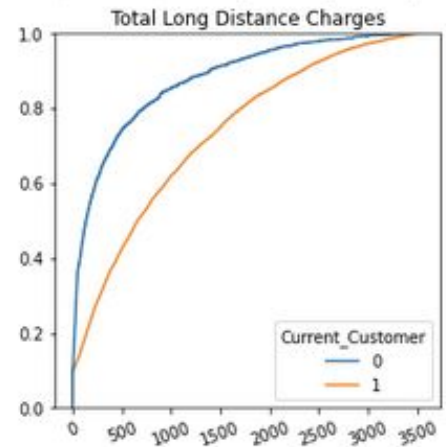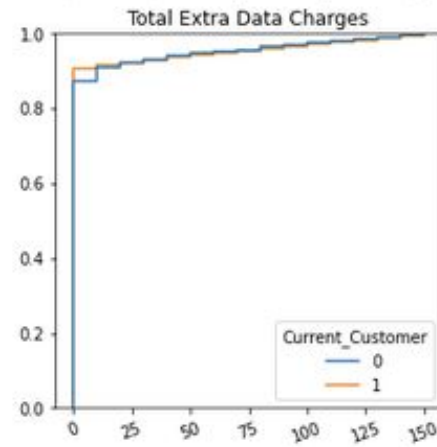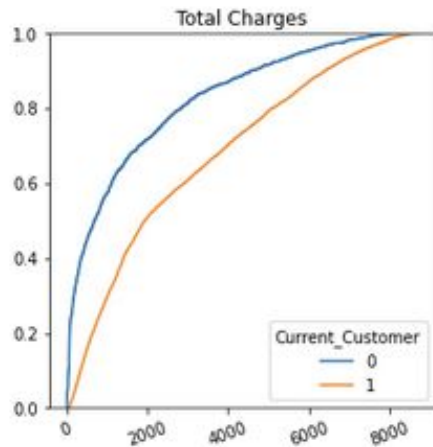

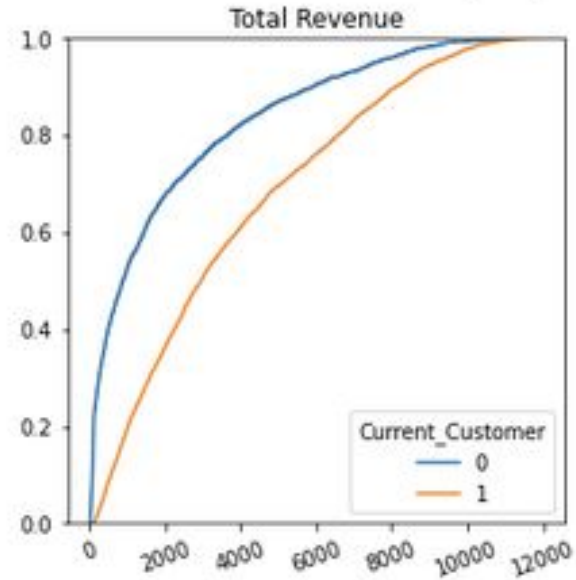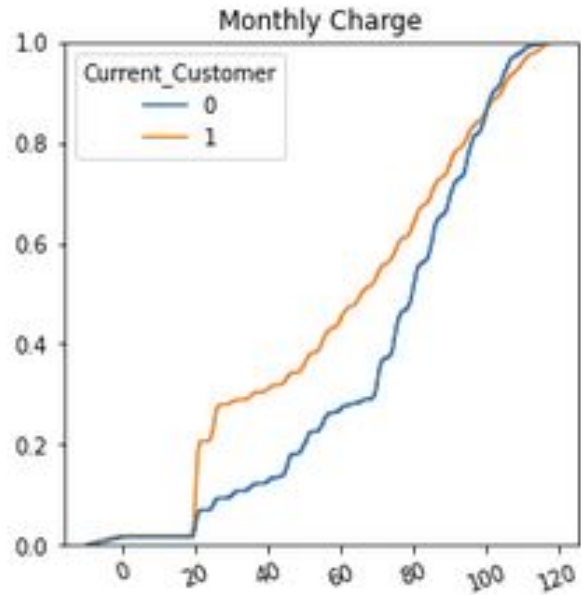Count Of Customer Reasons For Leaving
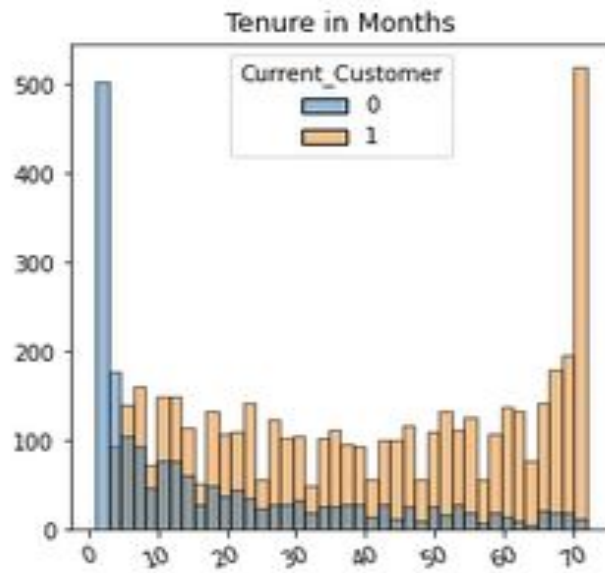
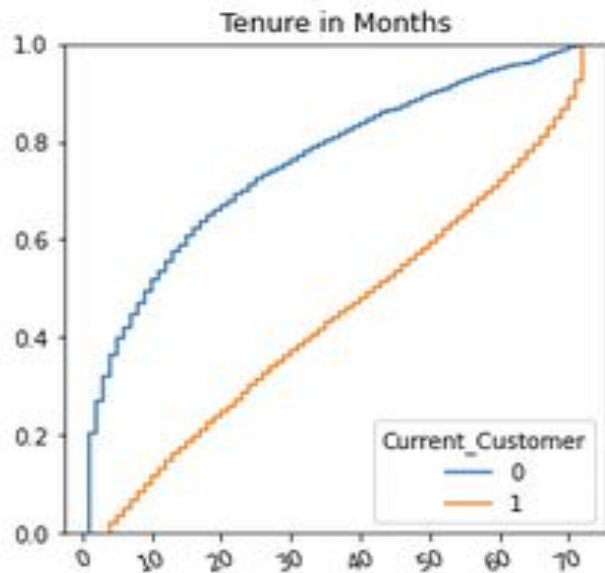# Exploratory Data Analysis

Exploratory Data Analysis (Continued)
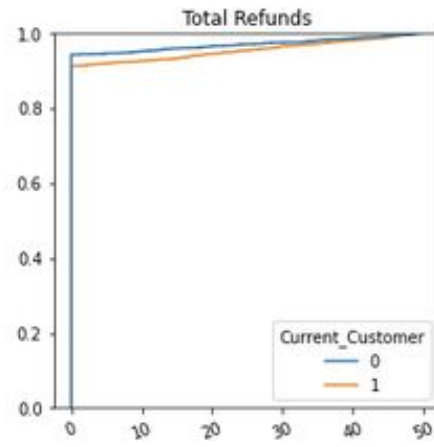
Exploratory Data Analysis (Continued)

Customer Tenure

Model Selection

|	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
gbc	Gradient Boosting Classifier	0.8774	0.9377	0.9464	0.8896	0.9171	0.6830	0.6885	0.2520
lightgbm	Light Gradient Boosting Machine	0.8742	0.9329	0.9356	0.8938	0.9141	0.6794	0.6823	0.0530
rf	Random Forest Classifier	0.8717	0.9252	0.9480	0.8819	0.9137	0.6652	0.6723	0.1220
ada	Ada Boost Classifier	0.8678	0.9265	0.9244	0.8946	0.9091	0.6668	0.6689	0.0800
et	Extra Trees Classifier	0.8637	0.9218	0.9298	0.8856	0.9071	0.6516	0.6547	0.1110
lda	Linear Discriminant Analysis	0.8439	0.9075	0.8900	0.8919	0.8908	0.6168	0.6176	0.0310
ridge	Ridge Classifier	0.8432	0.0000	0.9005	0.8830	0.8915	0.6085	0.6096	0.0220
dt	Decision Tree Classifier	0.8206	0.7800	0.8740	0.8752	0.8745	0.5597	0.5602	0.0230
knn	K Neighbors Classifier	0.7466	0.7117	0.8753	0.7927	0.8318	0.3229	0.3303	0.0320
nb	Naive Bayes	0.7188	0.8416	0.6735	0.9106	0.7738	0.4234	0.4584	0.0230
dummy	Dummy Classifier	0.7158	0.5000	1.0000	0.7158	0.8344	0.0000	0.0000	0.0170
svm	SVM - Linear Kernel	0.6743	0.0000	0.7467	0.8131	0.7435	0.2242	0.2674	0.0380
qda	Quadratic Discriminant Analysis	0.6423	0.6401	0.6456	0.8174	0.7054	0.2550	0.2713	0.0340

Model Performance
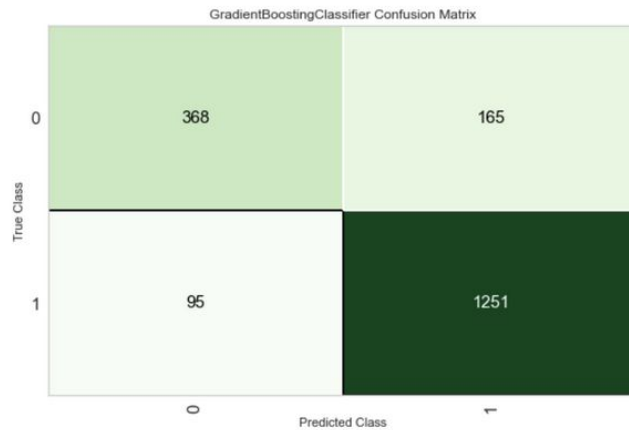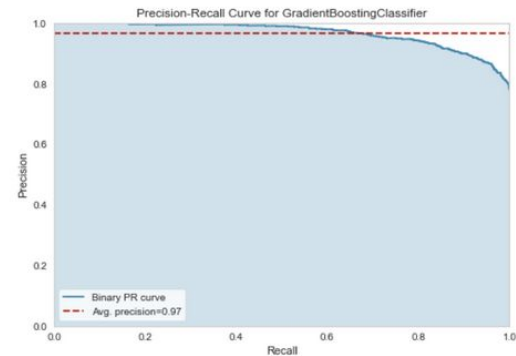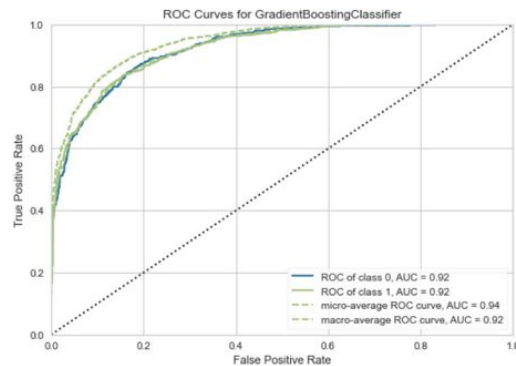
Takeaways

- Customer Tenure is the most important feature for predicting churn

- Customer stated reason for churn was competitor offer & device quality

Feature Importance Plot
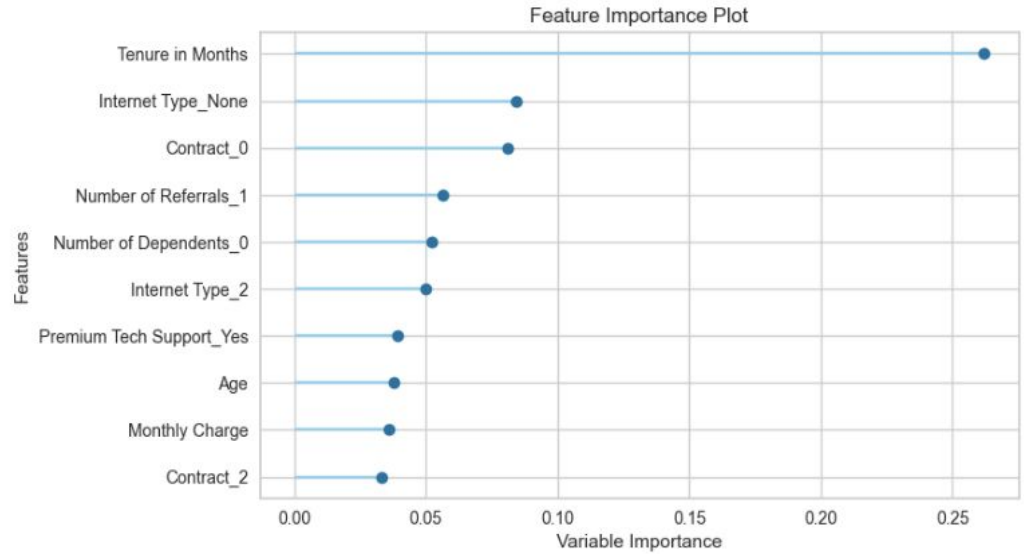
Recommendations

- Add contract terms to all plans
  - Reason: Customer tenure is a major factor in churn
    - Pros: Forces customer retention until end of contract, may make turnover more predictable
    - Cons: May be less attractive to potential customers

- Offer lower priced plans to beat competitors
  - Reason: Competitor made better offer is 2nd most common customer reason
  - Pros: More attractive to potential customers
  - Cons: Reduces profit per customer, may still result in churn

Recommendations (Continued)

- Improve customer support quality
  - Reason: Customer Support attitude was the 3rd most common reason given for customer churn.
  - Pros: More satisfied customers, potentially higher retention rates of customers
  - Cons: Cost of training & cost of monitoring customer support

Future Research

- Profile & features of high revenue customers
- Profile & features of new customers
- Market analysis of competitor offerings
- Impact of contract terms on customer retention

Questions & Answers