CONNCETTEL CUSTOMER CHURN PREDICTION

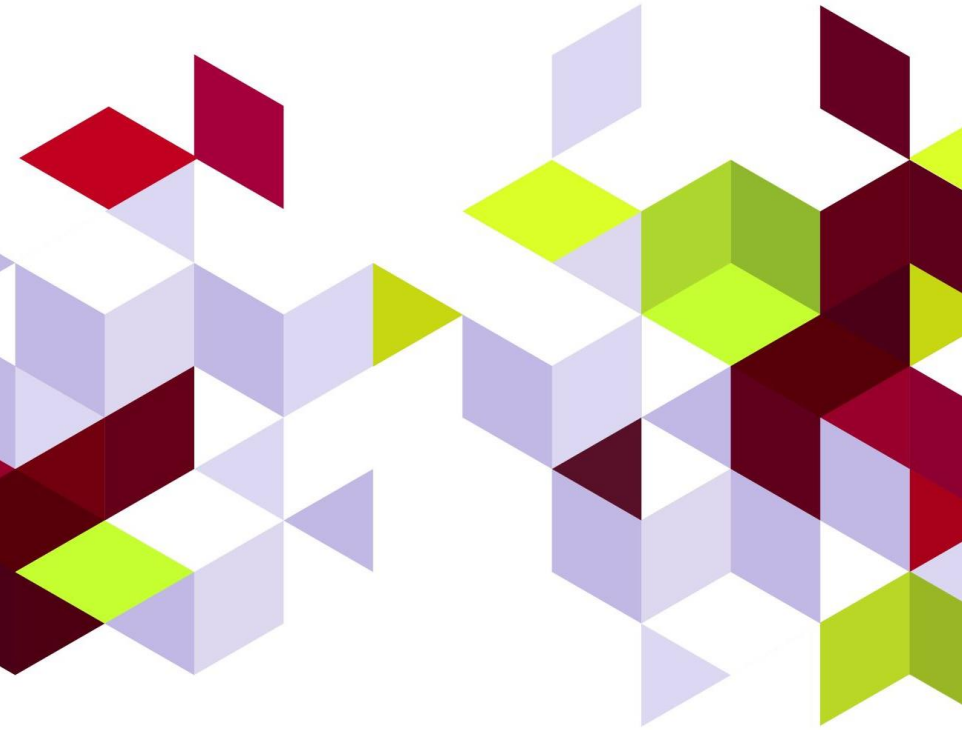
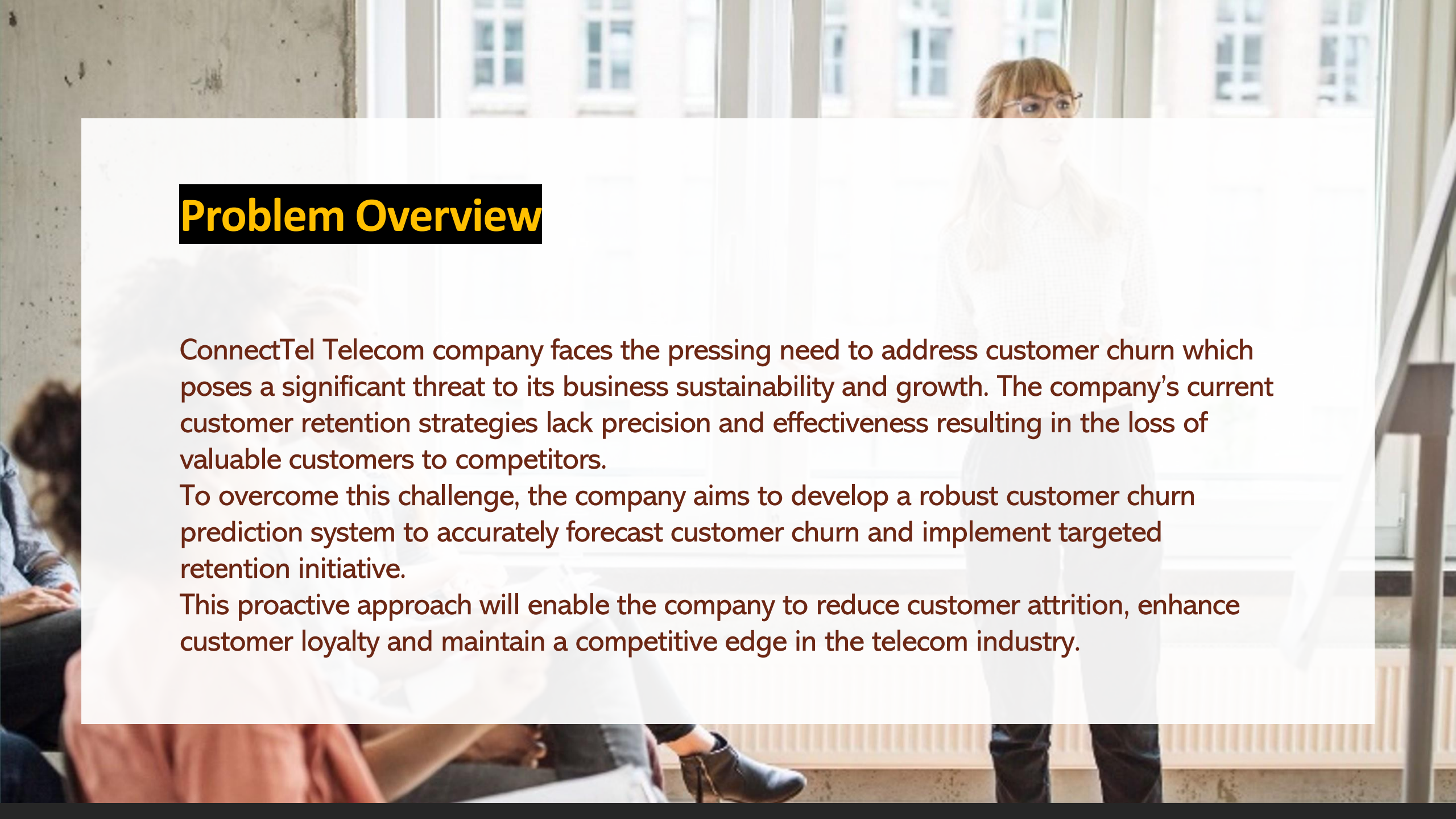


TABLE OF CONTENT

Introduction

ConnectTel is a leading telecommunication company at the forefront of innovation and connectivity solutions.
With a strong presence in the global market, ConnectTel has established itself as a trusted provider of reliable voice, data and internet services. The company offer a range of telecommunication solutions, including mobile networks, broadband connections and enterprise solutions.
ConnectTel caters to both individual and corporate customers and they are committed to providing exceptional customer service and cutting-edge technology.

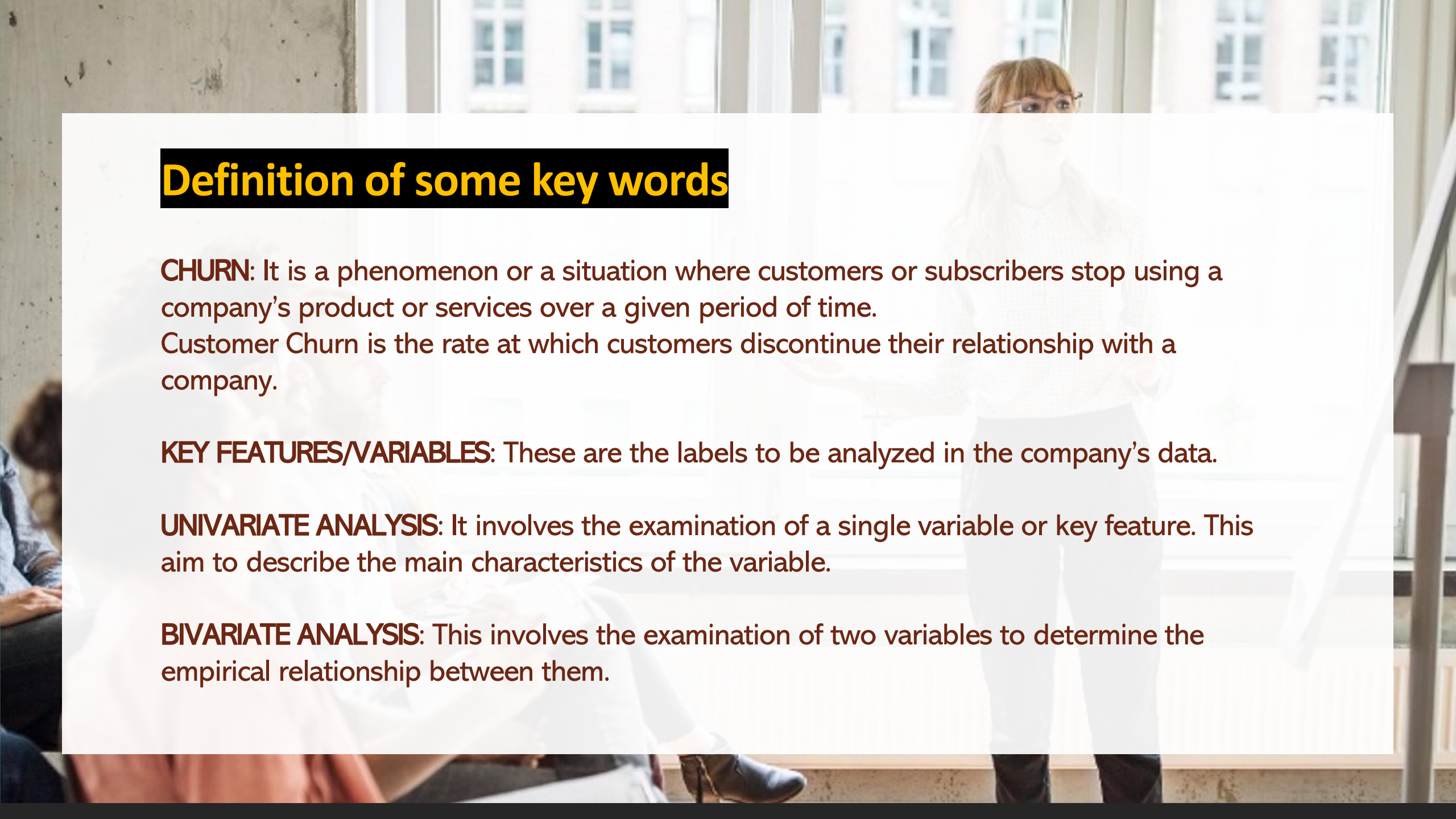


Problem Overview

ConnectTel Telecom company faces the pressing need to address customer churn which poses a significant threat to its business sustainability and growth. The company's current customer retention strategies lack precision and effectiveness resulting in the loss of valuable customers to competitors.

To overcome this challenge, the company aims to develop a robust customer churn prediction system to accurately forecast customer churn and implement targeted retention initiative.

This proactive approach will enable the company to reduce customer attrition, enhance customer loyalty and maintain a competitive edge in the telecom industry.



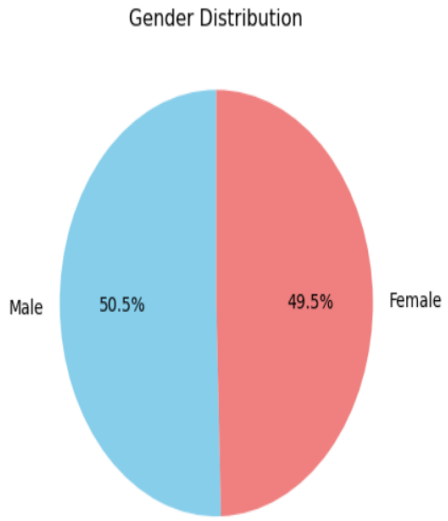
Definition of some key words

CHURN: It is a phenomenon or a situation where customers or subscribers stop using a company's product or services over a given period of time.
Customer Churn is the rate at which customers discontinue their relationship with a company.

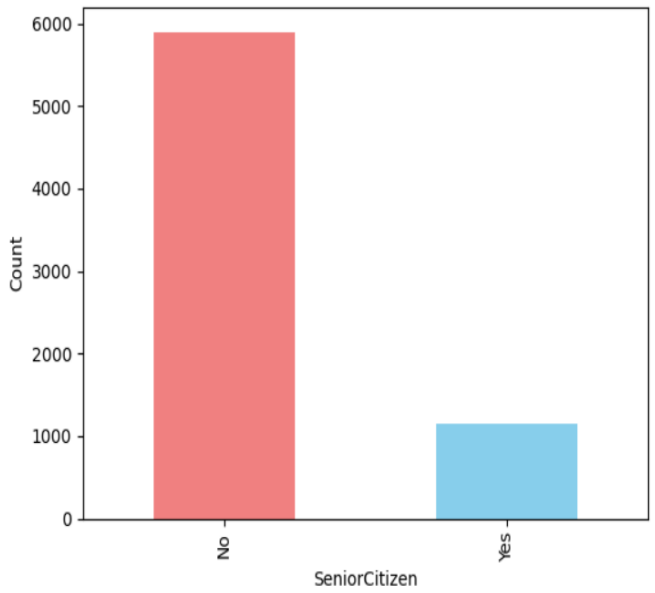
KEY FEATURES/VARIABLES: These are the labels to be analyzed in the company's data.

UNIVARIATE ANALYSIS: It involves the examination of a single variable or key feature. This aim to describe the main characteristics of the variable.

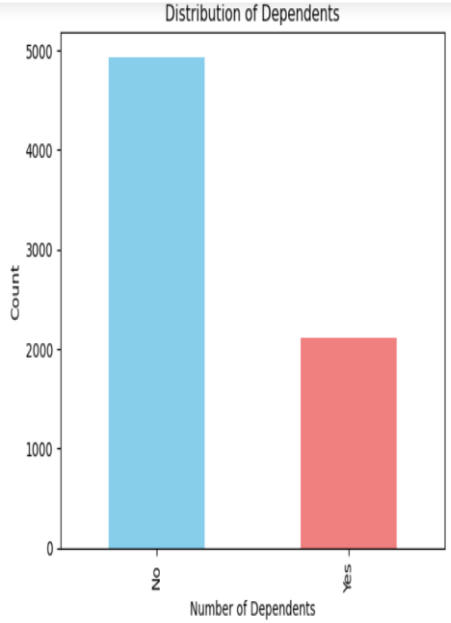
BIVARIATE ANALYSIS: This involves the examination of two variables to determine the empirical relationship between them.



Observations - The difference in the percentage of male customers to female customers are slim therefore both the male and female customers are almost equal subscribers.



Observations - This visualizes that 1,142 customers are categorized as Senior Citizen while 5,901 customers are not in the category of Senior Citizen.



Observation - There is a high number of customers that are independent about 4,900 customers compared to those that are dependent about 2,100

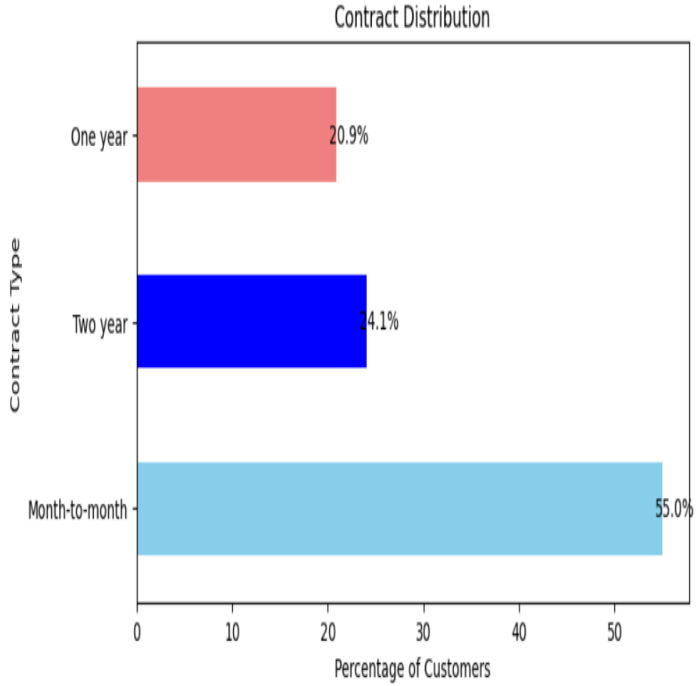
Exploratory Data Analysis between labels and some key features.

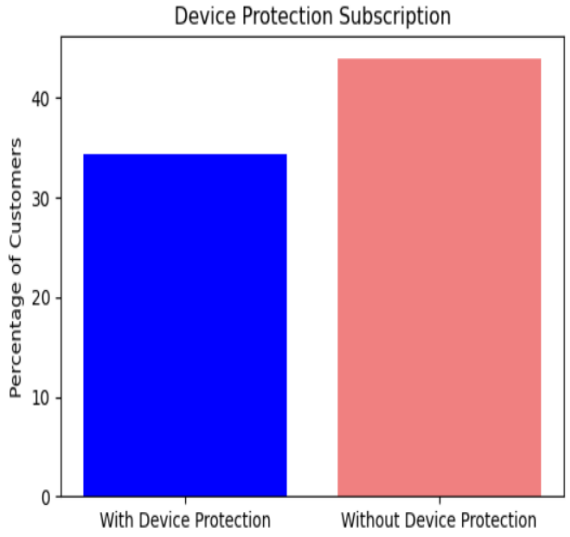
These visualization that shows the distribution of Gender, Senior Citizen and Dependents, will help in understanding the customers information.

Gender: This shows an almost equal distribution between both gender in the company's data.

Senior Citizen: This shows that a large percentage of company's customers are not senior citizen, about 5,901 are not identified as Senior Citizen while 1,142 customers are Senior Citizen. This helps in understanding any churn patterns among senior and non-senior customers.

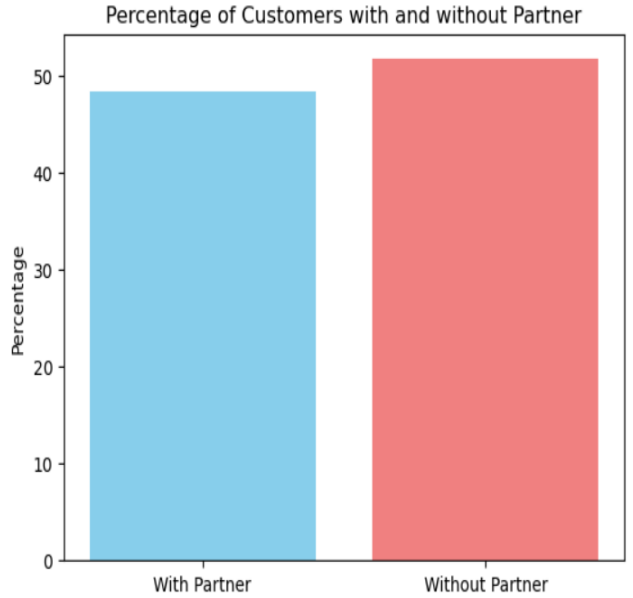
Dependents: This shows a high number of customers about 4,900 are independent while about 2,100 customers are labeled as dependents. This helps in assessing the influence of having dependents on customer churn.





This visualization shows that there is a high percentage of customers without

Device Protection about 44% in comparison to customers with Device Protection about 35%



Observation - About 48% of customers have partners and about 52% are without partners

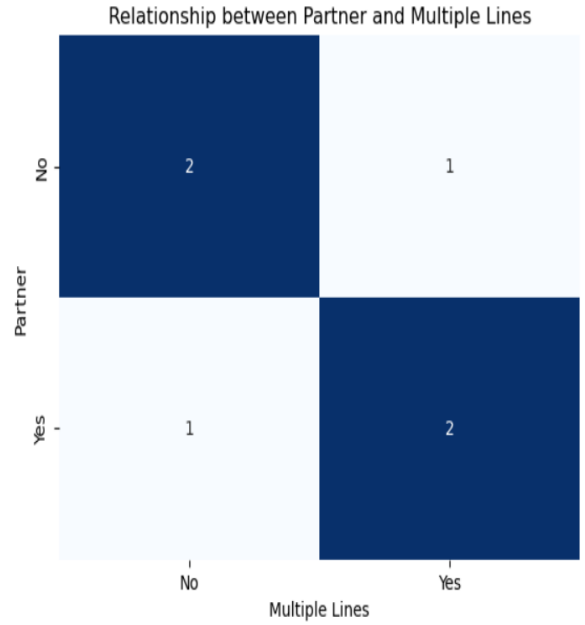
More EDA on some key features.

These are showing the percentage distribution on Contract Type, Customer Device Protection and Customers with or without partner. These visualization will help us understand the customer churn of the company.

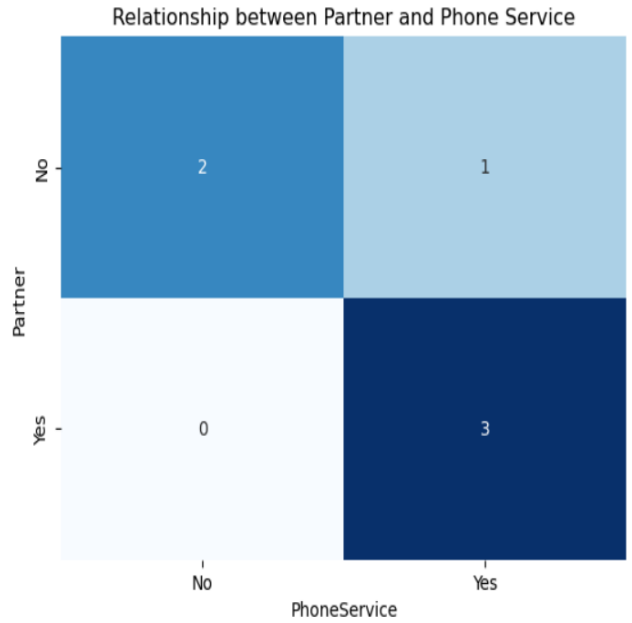
Contract Type: 55% of customers subscribe for a monthly contract, 24% opt for a 2year contract subscription while 21% subscribe for a yearly contract.

Device Protection: This shows that about 56% of customers are without device protection, while 44% have their device protected.

Partner: This visualization shows that 52% of the company customers are without partner, while 48% are with partner.



Observation - This visualization shows that there is equal percentage of customers who has partners aand use multiple lines and those who do not have partners and dont use multiple lines.



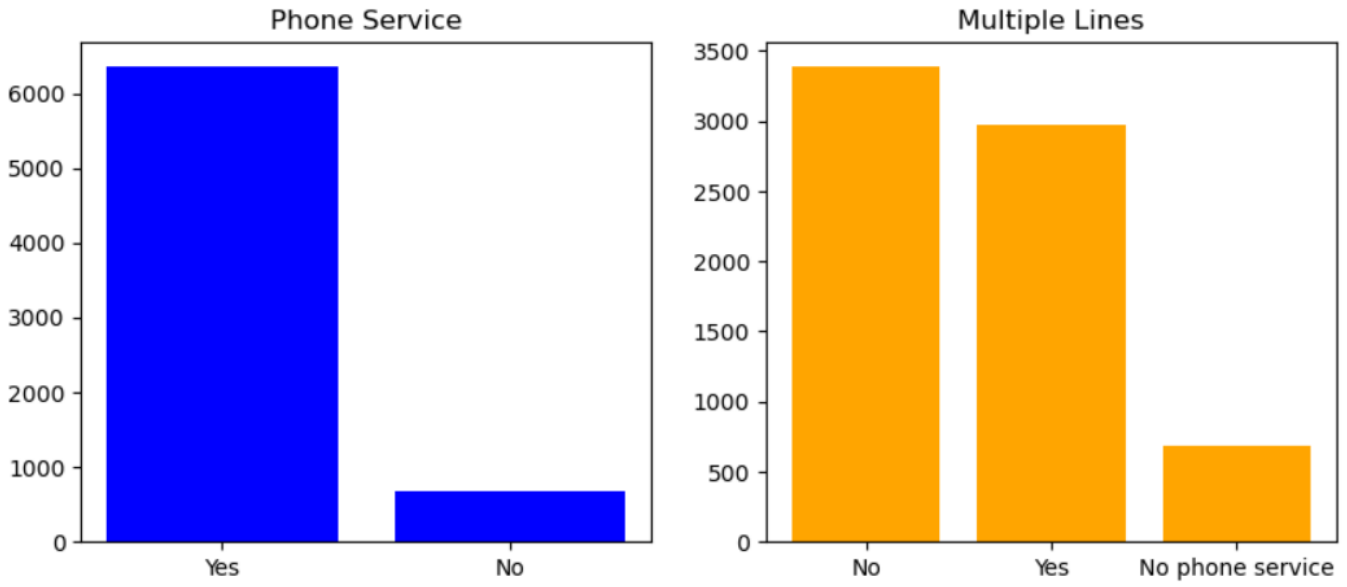
This visualization indicates that a high percentage of customers who have partners use Phone Service

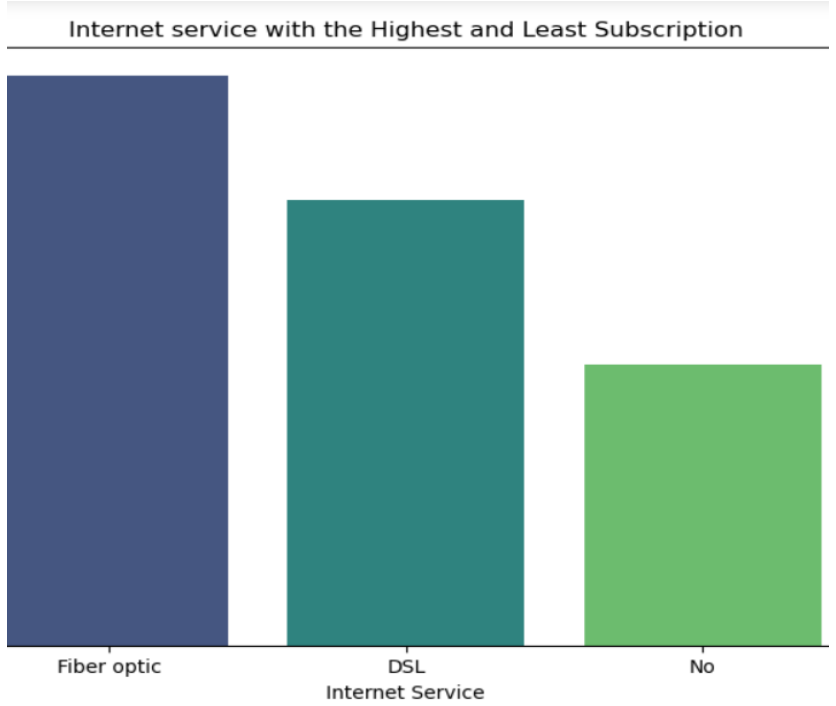
Correlations on some Variables.

These visualizations shows the positive and negative correlation between customers who has partners and have connection to a phone service or use multiple lines and customers without partners and their preferred use.

The first visualization to the left shows a positive correlation which is an equal percentage of customers who has partners and use multiple lines and those who don't.

The second visualization shows a positive correlation of customers who has partners and use phone service and a negative correlation of customers who use phone service and are without partners.



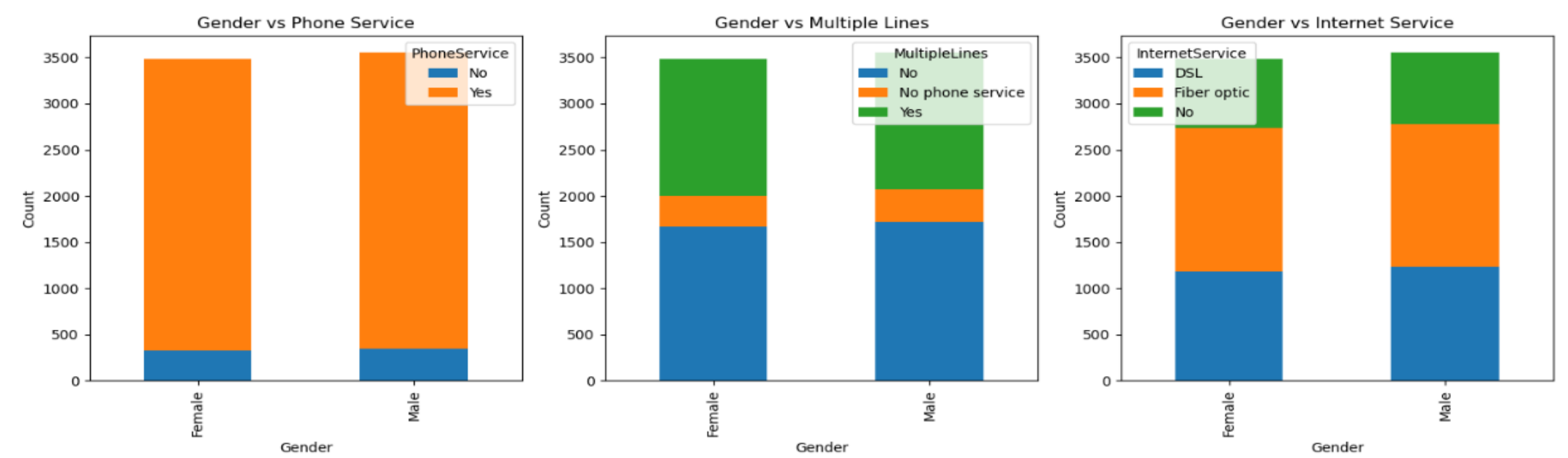


Observation - Over 6,400 customers use phone service and about 2,900 out of the 6,400 use multiple lines

Univariate Analysis on some Variables.

These visualization provide a clear understanding of customer preferences for phone and internet services. Most customers subscribe to phone service, with an almost balanced split between those using multiple lines and those not.

Fiber Optic is the most popular preferred internet service, followed by DSL, with a considerable number of customers not subscribing to internet services.



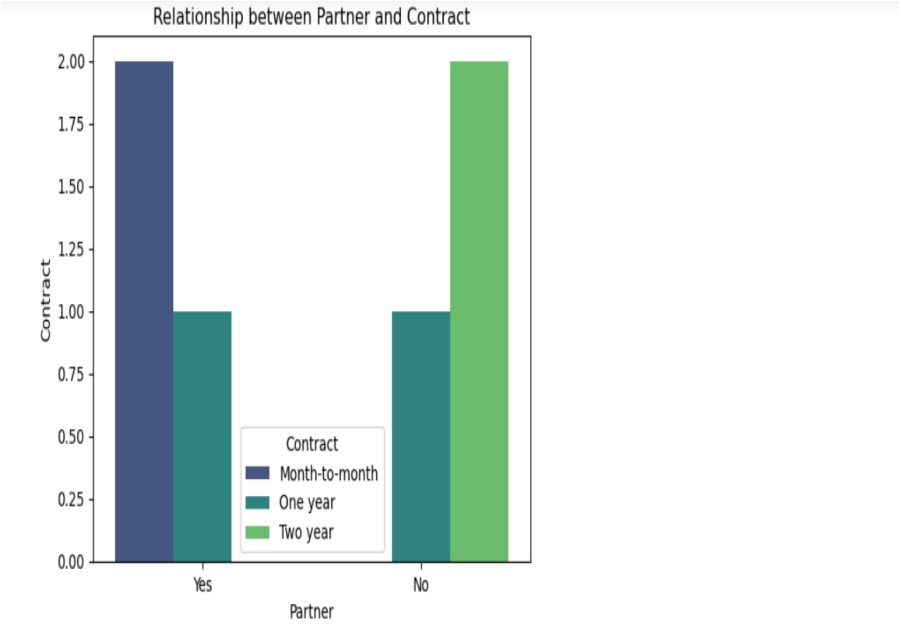
Observations - This visualization shows the comparative view of how different genders interact with these services (Phone Service, Multiple lines, Internet Service)

Bivariate Analysis on some variables.

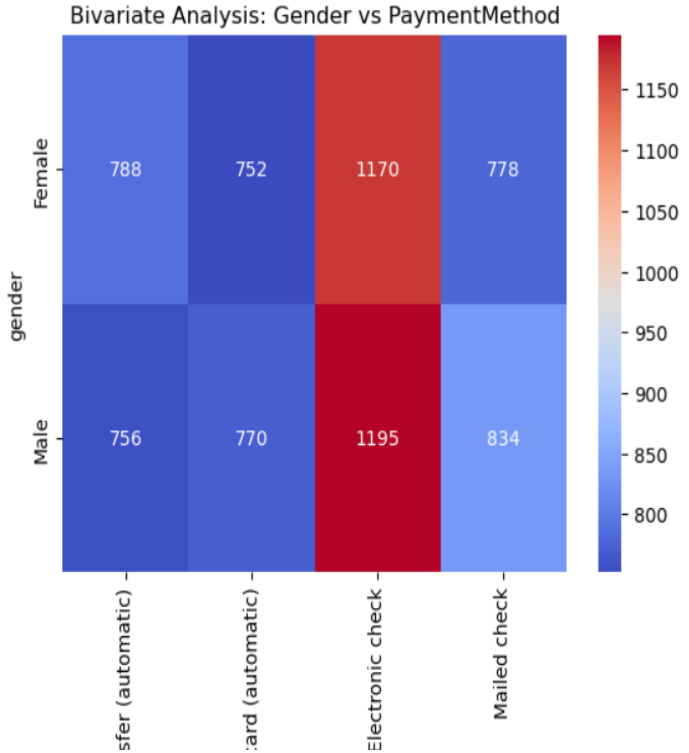
This visualization shows the relationship between certain variables. It aims to understand how each gender in the company's data interact and react to using phone and internet services.

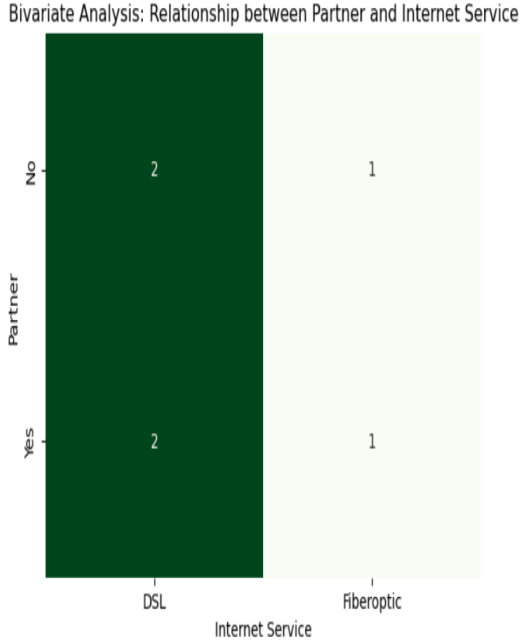
Both genders have almost equal access to phone service and use multiple lines with few number lacking access to these services and an equal number not using multiple lines.

For the internet service, Fiber Optic being the most preferred among both genders, closely followed by DSL while a substantial number of customers are not subscribed to the internet.



This visualization shows that customers without partners tend to subscribe to a longer contract period (Two Year) while those with partners subscribe more of month to month contract.





Observation - This shows that the most preferred Internet Service for customers is DSL

More Bivariate Analysis on some key features.

These visualization shows the correlation between customers who has partners and their preferred contract tenure and customers who do not. Customers who has partners tend to subscribe to a monthly contract tenure while customer without partners subscribe more to a 2-year contract period.

The next visualization shows the correlation on the most and least preferred payment method between both genders. For both male and female gender, electronic check is the most preferred payment method while bank transfer is the least preferred for the male gender and credit card is the least preferred for the female gender.

For the last visualization, it shows an equal distribution between partner and internet service. Both customers with and without partners prefer the DSL service.

Accuracy Score

	XGB Classifier	Random Forest	K-Nearest Neighbors	SGD Classifier	SVC	Naive Bayes	Decision Tree	Logistic Regression
0	79.28%	78.92%	76.79%	81.12%	80.77%	75.8%	72.82%	81.55%

Precision List

	XGB Classifier	Random Forest	K-Nearest Neighbors	SGD Classifier	SVC	Naive Bayes	Decision Tree	Logistic Regression
0	62.54%	65.93%	56.97%	74.68%	69.03%	52.94%	48.57%	68.77%

Recall List

	XGB Classifier	Random Forest	K-Nearest Neighbors	SGD Classifier	SVC	Naive Bayes	Decision Tree	Logistic Regression
0	54.16%	47.72%	50.4%	30.83%	49.6%	77.21%	45.58%	55.5%

ROC List

	XGB Classifier	Random Forest	K-Nearest Neighbors	SGD Classifier	SVC	Naive Bayes	Decision Tree	Logistic Regression
0	71.24%	69.42%	68.35%	63.53%	70.79%	76.25%	64.1%	73.21%

Machine Learning Model Development used and best performing model.

This customer churn prediction was trained on several supervised machine learning models, which includes but not limited to these; XGB Classifier, Random Forest, K-Nearest Neighbors, SGD Classifier, SVC, Naïve Bayes, Decision Tree, Logistic Regression.

Based on performance, Logistic Regression performed the best with an
- Accuracy Score: 81.55%
- Precision: 68.77%
- Recall: 55.5%
- ROC AUC: 73.21%.

Business Impact

Running a customer churn prediction model can impact the business in various positive ways:

1. **Improved Customer Retention/Relationship:** Taking proactive measures to retain customers who are likely to churn, by offering discounts, addressing specific complaints or providing personalized services. Retention campaigns could be made more efficient and effective for high-risk customers. Also, engaging with customers at risk of churning could provide opportunity to gather valuable feedback and address issues before they escalate.

2. **Increased Revenue:** Retaining customers directly translates to sustained revenue. By reducing churn, the average lifetime value of a customer increases boosting long-term profitability.

3. **Saves Cost:** Retaining existing customers is generally cheaper than acquiring new ones. Churn prediction helps in minimizing marketing and sales costs and the resources can be allocated more efficiently.

4. **Competitive Advantage:** By addressing issues leading to churn, the business can enhance overall customer experience, making them more competitive in the market, customer centric and gain positive reputation.

5. **Strategic Decision Making:** Insights gained from churn prediction model can guide strategic decisions such as product development, pricing strategies and market segmentation. It can also aid in better financial management.

6. **Enhanced Marketing Effectiveness:** With churn predictions, marketing efforts can be more personalized and targeted, increasing their effectiveness and ROI. By identifying patterns in churn, a better loyalty program can be designed.

Conclusion and Recommendation

In summary, running a customer churn prediction model enables businesses to retain more customers, increase revenue, reduce cost and gain valuable insights into customer behavior. These benefits collectively lead to improved profitability, competitive advantage and sustainable growth.

Recommendation: Using the best trained Machine Learning Model Logistic Regression effectively, can help the company proactively address churn by:

- Input customer data which includes the customer demographic details, usage patterns, service plans and payment history.
- The model can predict the probability of churn for each customer using the customer's data.
- Customers with predicted probability above a set threshold (e.g, 1) are flagged as likely to churn and necessary actions taken immediately.
- The business can reach out to this flagged customers with retention campaigns.

Thank you

Stella Enwerem

www.linkedin.com/in/stellamaris-enwerem

stellae15@outlook.com