PORTOFOLIO

TELECOM CHURN ANALYST

Presented by Rafli Fadilah Rahman

Today's Outline

Topic Highlights

The presentation of the analytical results of the teleco Churn Customer data set, which has been used in the company's **decision-making process** in relation to customer **churn**.

Dataset by : https://www.kaggle.com/datasets/blastchar/telco-customer-churn

WHEN YOU FOCUS ON YOUR CUSTOMERS, EVERYTHING ELSE WILL FOLLOW.

JEFF BEZOS

Table Of Content

Background

- The dataset in question is that of a telecommunications company, and it contains **behavioural** data from **two distinct customer groups: those who have churned and those who have not.**
- This dataset also contains customer **demographic** information and **service** data provided by **customers.**
- In order **to make informed decisions about future churn rates**, it is essential to **analyse** the data in question. This analysis should consider the **behaviour of customers** in order to support telecommunication company decisions.

Objectives

Given the aforementioned background, **it is clear that a machine learning model is required to predict customer churn**. This model must consider customer **demographic information, service usage by customers, and other customer behaviours.**

Description of data

- There are **21 columns** and **7043 rows** in the **Telco Churn Customer dataset**.
- For further details on the dataset and the description of each column, please refer to the following link : Gdrive DataSet
- This dataset contains several data types that were **converted from categorical to numerical** to facilitate **data cleaning and analysis** using Jupyter Notebook. The links are as follows: Data Cleaning & EDA: Data Cleaning & EDA

No	Name Column	Description	**Data Type**	Unique Value
1	Customer ID	Contains customer ID	Kategorial	ID Unique Customer
2	Gender	Contains the gender of the customer	Kategorial	Female = 0 ; Male = 1
3	SeniorCitizen	Customer is a senior citizen or not	Kategorial	Senior Citizen = 1 ; Not Senior Citizen = 0
4	Partners	Customer has a partner or not	Kategorial	Yes = 1 ; Not = 0
5	Dependent	Customer has dependents or not	Kategorial	Yes = 1 ; Not = 0
6	Tenure	Number of months subscribed	Numerik	Min = 0 ; Month max = 72
7	Phoneservice	Whether the customer has multiple lines or not	Kategorial	Yes = 1 ; Not = 0
8	Multiplies Line	Whether the customer has multiple lines or not	Kategorial	No phone service = 0 ; No = 1 ; Yes = 2
9	Internetservice	Customer's internet service provider	Kategorial	No =0 ; DLS: 1 ; Fiber optic ; 2

Dashboard Teleco Churn Rate Customer

customerID
Masukkan nilai

Filter

Churn

gender

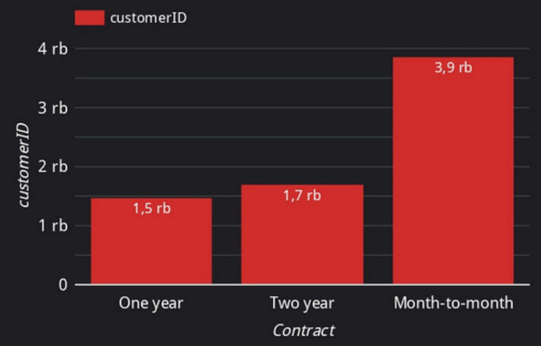
Contract

PaymentMethod

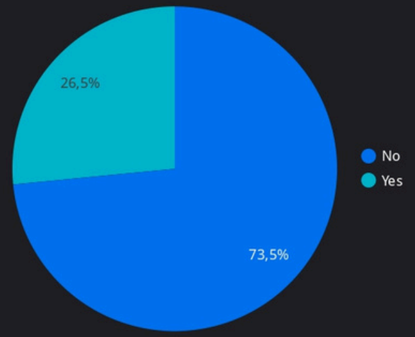
TechSupport

Total Customer	Tech Support	Average Monthly Charge's	Total Charges
7.043	7.043	$64,76	16.056.168,7

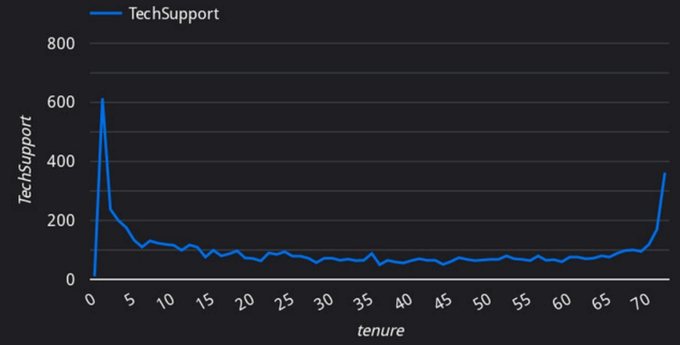
Duration Contract



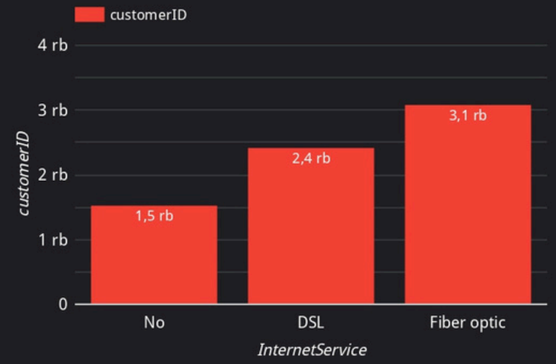
Count Of Customer by Churn



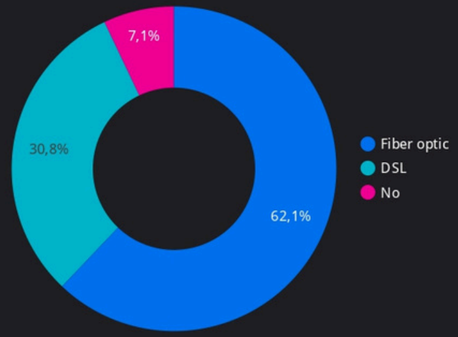
Total Tech Ticket by Tenure



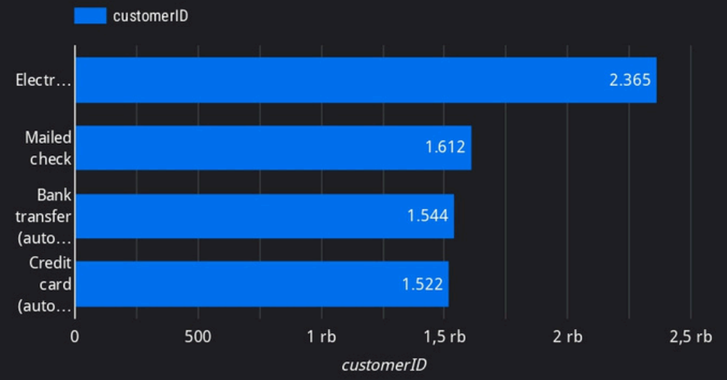
Internet Service



Monthly Charges by Internet Service



Payment Method Use By Customer



Dashboard data overview

To view the dashboard further can use the following link access: [Dashboard Teleco Customer]

Exploratory Data Analyst

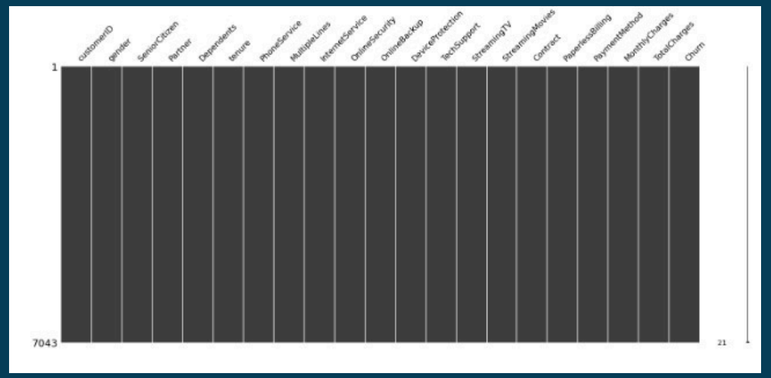
This dataset contains several data types that were converted from categorical to numerical to facilitate data cleaning and analysis using Jupyter Notebook. The links are as follows: Data Cleaning & EDA: Data Cleaning & EDA

This EDA focuses on factors in the data set that affect customer churn.

- Page 9
  Data cleansing process of the Telecochurn rate dataset.
- Page 10
  The objective is to determine the condition of the analysis target, namely customers and churn, using EDA dataset.
- Page 11 - 16
  Display visual results to show conditions for analysis purposes.

Exploratory Data Analyst
Data Cleaning

- At this stage, the dataset comprises 7,043 rows and 21 columns. Upon initial analysis, no null values were identified. However, subsequent investigation revealed the presence of **missing numerical values in the tenure column.** These values will be **removed to enhance data efficiency.**
- Furthermore, there **are 11 missing values in the Total Charges column**. In order to make improvements, I have **taken the average value of the column.**
- To facilitate more effective data processing, **I converted the categorical data to numerical values.**




```
gender              0
SeniorCitizen       0
Partner             0
Dependents          0
tenure              0
PhoneService        0
MultipleLines       0
InternetService     0
OnlineSecurity      0
OnlineBackup        0
DeviceProtection    0
TechSupport         0
StreamingTV         0
StreamingMovies     0
Contract            0
PaperlessBilling    0
PaymentMethod       0
MonthlyCharges      0
TotalCharges       11
Churn               0
dtype: int64
```

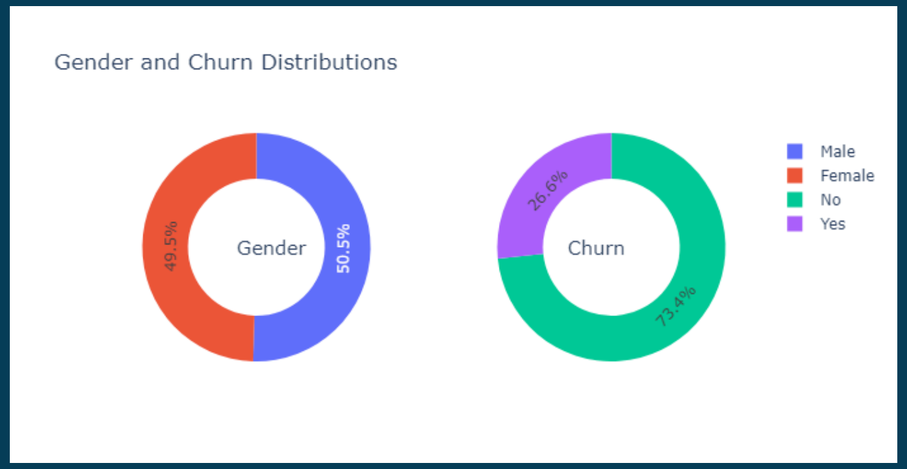
Exploratory Data Analyst

	tenure	MonthlyCharges	TotalCharges
count	7032.000000	7032.000000	7032.000000
mean	32.421786	64.798208	2283.300441
std	24.545260	30.085974	2266.771362
min	1.000000	18.250000	18.800000
25%	9.000000	35.587500	401.450000
50%	29.000000	70.350000	1397.475000
75%	55.000000	89.862500	3794.737500
max	72.000000	118.750000	8684.800000

We are currently undertaking a review of the distribution of numeric data, namely tenure, total charges and monthly charges.

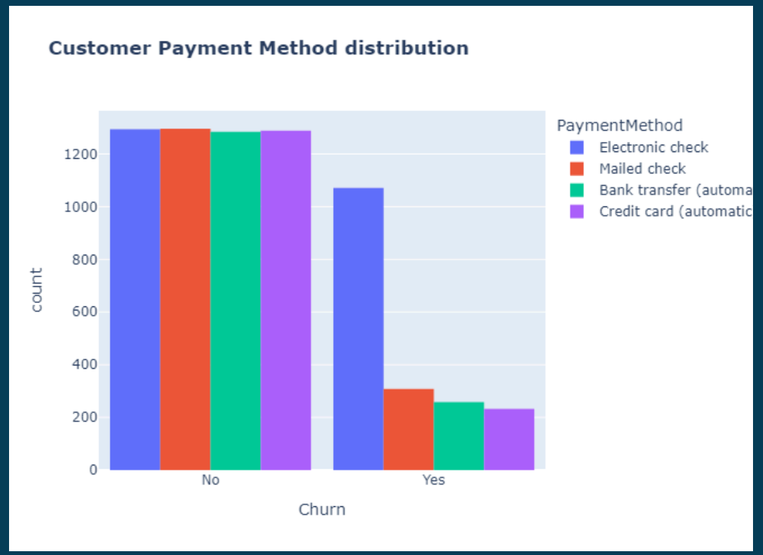
interpretation :

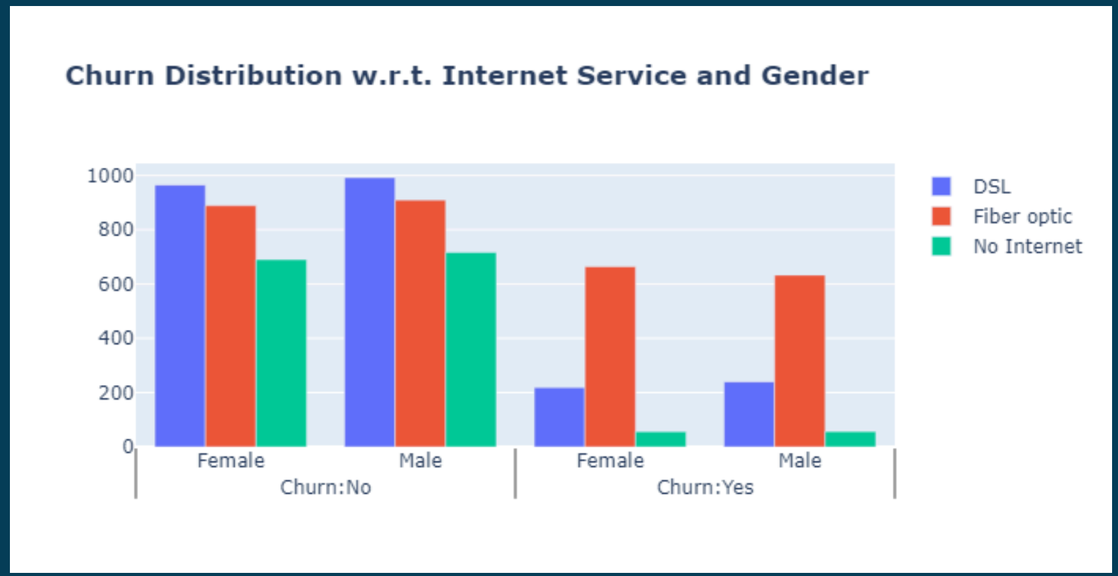
- **(Tenure) The average customer** uses internet services for **32.42 months.**
- The **average monthly cost of internet service** is **64.79.**
- The **total cost that the average customer has paid is 2283.30.**
- **Most customers (75%)** are still using the internet service.
- The **lowest monthly cost of internet service is 18.25**, and the highest is 118.75.
- The lowest total cost paid by customers was 18.80, and the highest was 8684.80.



The dataset shows the proportion of customers is 49.5% Female and 50.5% Male, and shows 73.4% No Churn and 26.6% Churn.

Data Visualisation

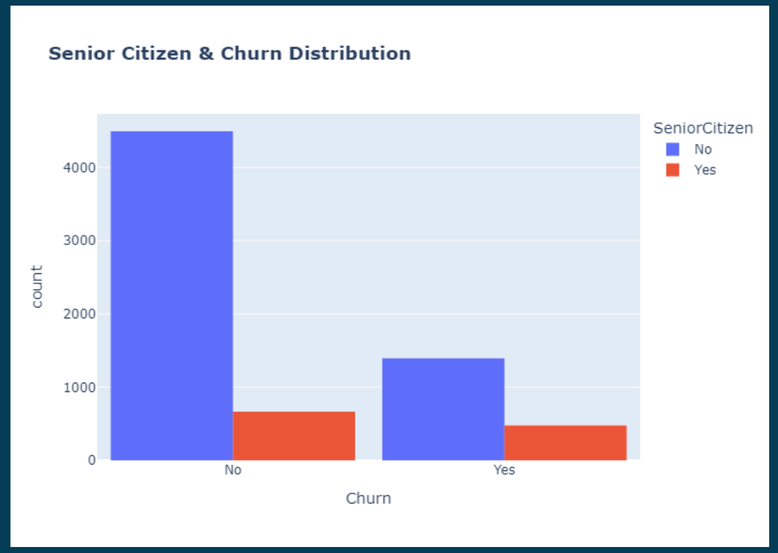




- In customer churn, **it shows that electronic check tends to churn**, which is in **the number 1 position in the users of payment methods that have the potential to churn.**
- The churn rate for customers who use **Electronic Check is 45.31%**.

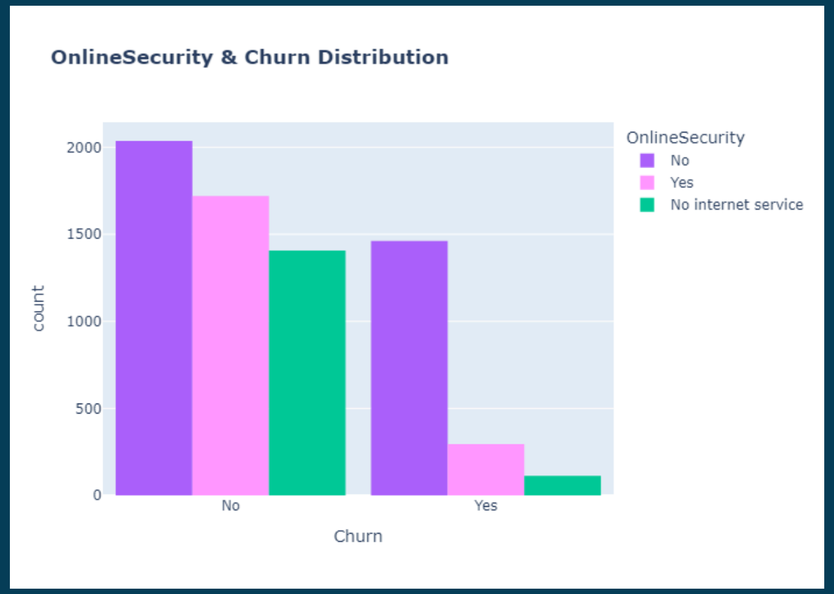
- Cumulatively Fiber Optic users are 1799 who do not experience Churn and as many as 1327, so a review of fiber optic services is needed.
- Fibre Optic **has the highest churn rate** in customers of **42.48%** in both female and female gender.

Data Visualisation



- **Customers categorised as elderly** have a **high churn rate** even though the total number of customers at that age is small.

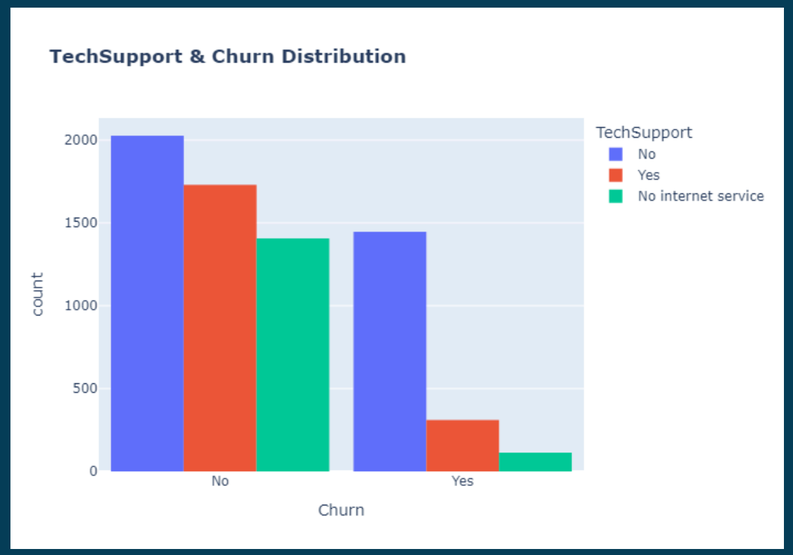
- **A churn rate of 41.66%** means that of those, 41.66 cancelled their **subscription.On this factor, elderly or SeniorCitizens are prone to churn rates.**



- **Customers who do not have Online Security have a churn rate of 41.6%**, so these customers are vulnerable to fraud.

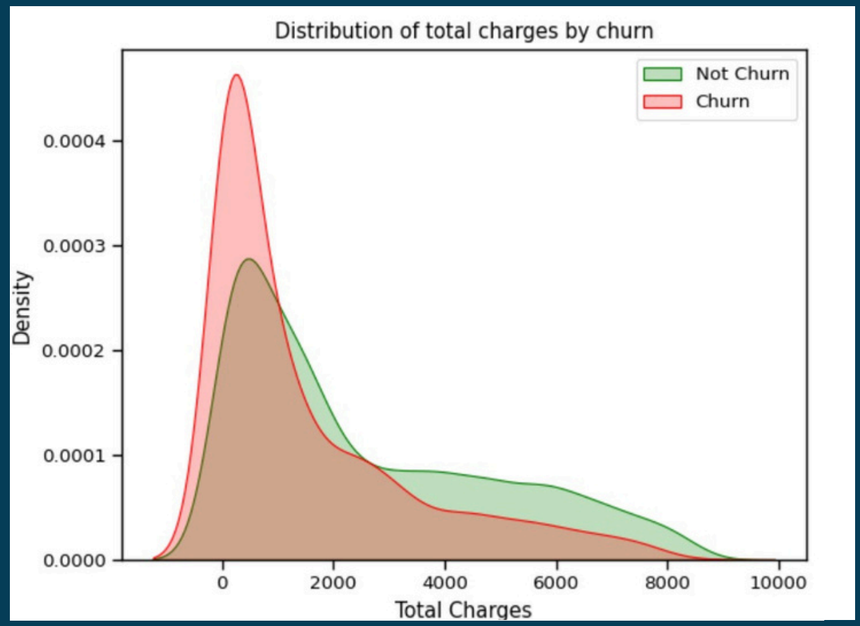
- **Not providing or not making it easy for them to utilise cyber security will impact the compan**y.

Data Visualisation



- **41.63% Customers who experience churn are customers who do not use / do not have technical support tend to experience churn.**

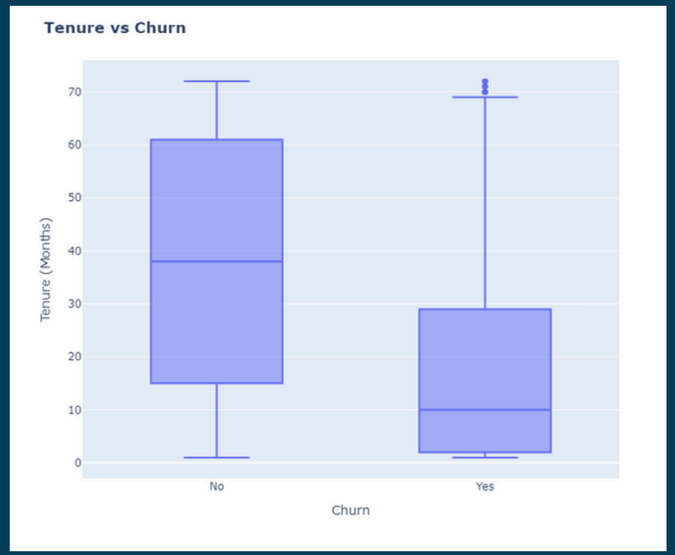
- **A churn rate of 41.66%** means that of those, 41.66 cancelled their **subscription.On this factor, elderly or SeniorCitizens are prone to churn rates.**



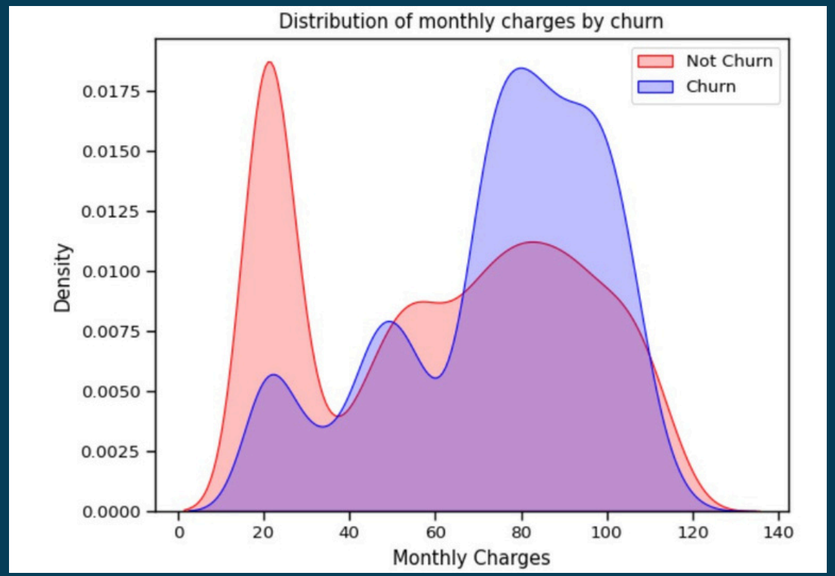
- This graph shows that customers who **spend more money with the company are less likely to Churn.**

- **Customers who experience churn tend to have lower total costs**. that customers lack confidence in the products provided by the company in the short term

Data Visualisation

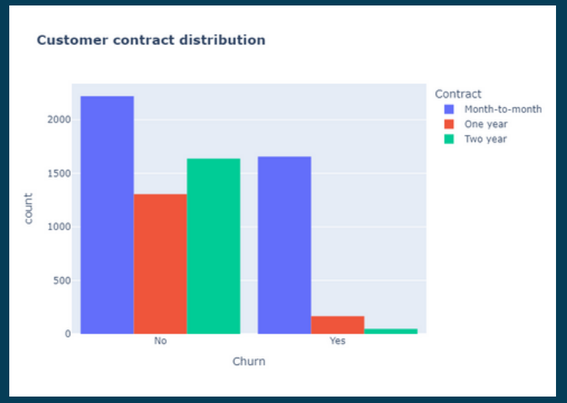


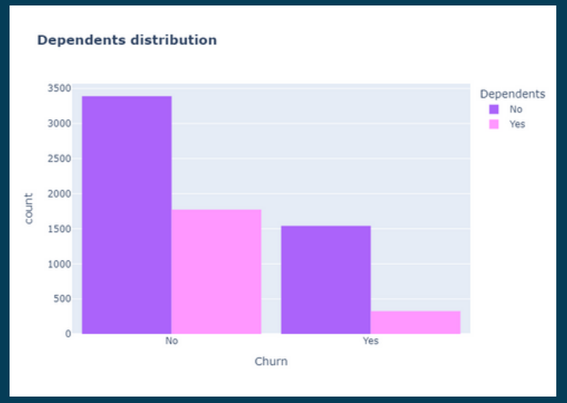
- It is evident that customers who **subscribe for longer periods** of time are **less likely** to churn
- The whiskers for churn customers extend further to the left than the whiskers for customers who do not churn, **suggesting that there is a higher proportion of churn customers with very short durations.**

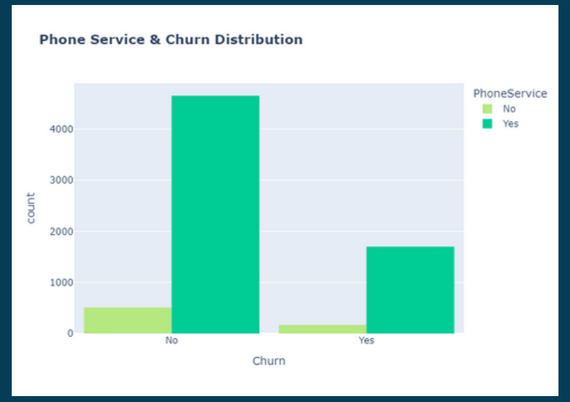


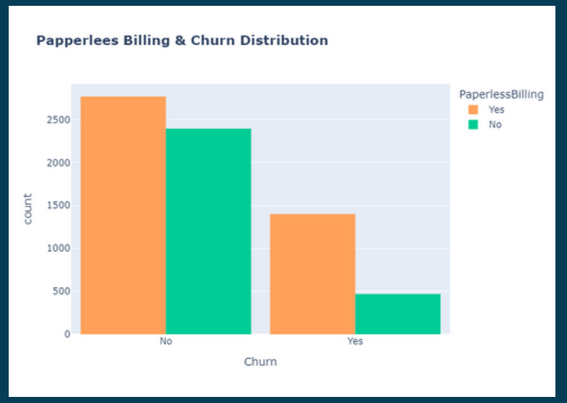
- Customers **with high monthly fees are more likely to churn** (blue curve) than those with lower fees (red curve).
- The blue curve (Churn) looks wider. **this shows that customers are more sensitive to price changes.**

Data Visualisation

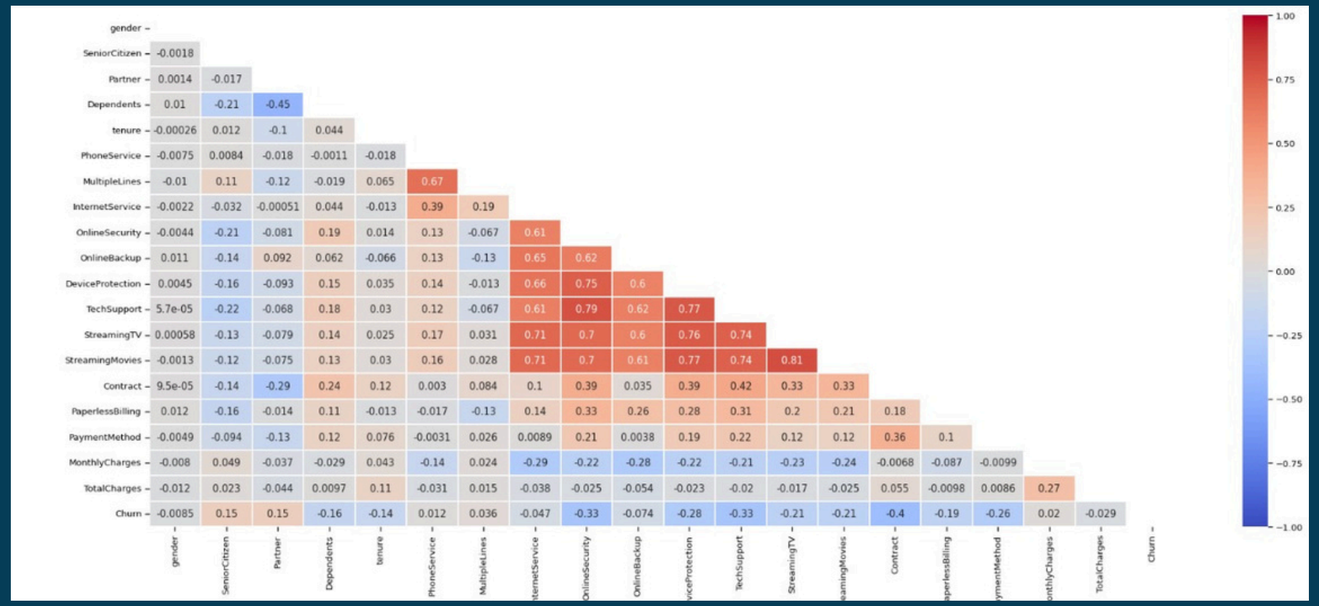








Data Visualisation Correlation Matrix



Factors that correlate with Churn :

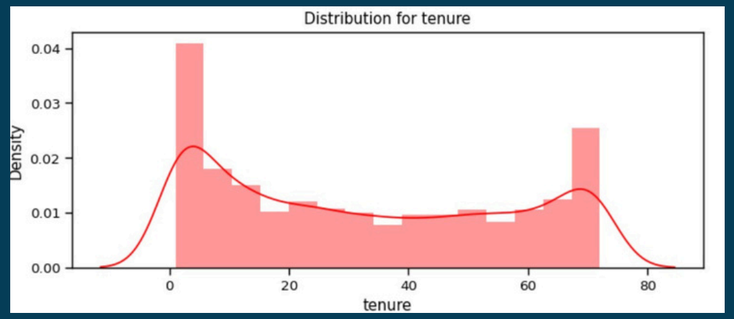
- SeniorCitizen moves in the same direction as Churn with a value of (0.15)
- Partner moves positively with churn by (0.15)
- MonthlyCharges moves in the same direction but in a weak correlation with a value of (0.02)
- PhoneService moves in the same direction but in a weak correlation with a value of (0.012)

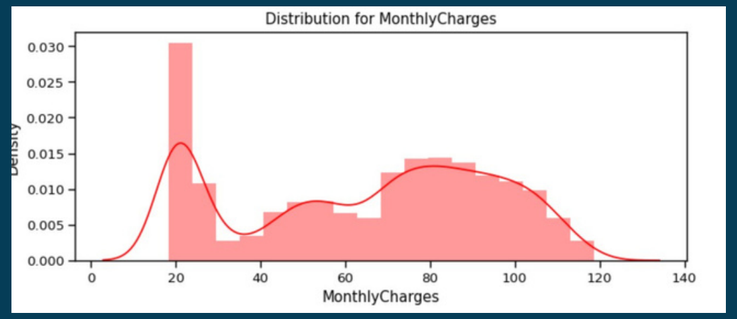
further factor analysis :

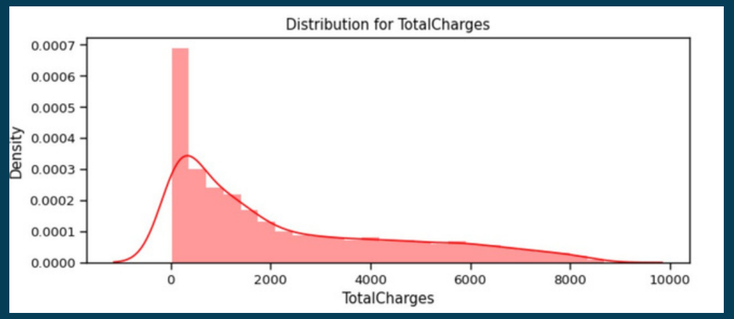
- Customers who have been with the company for a long time are much **less likely to churn than new customers.**
- Customers with **higher monthly bills are less likely to churn** than customers with lower monthly bills lower monthly bill
- Customers with **higher total bills are slightly less likely to churn** than customers with lower total bills.

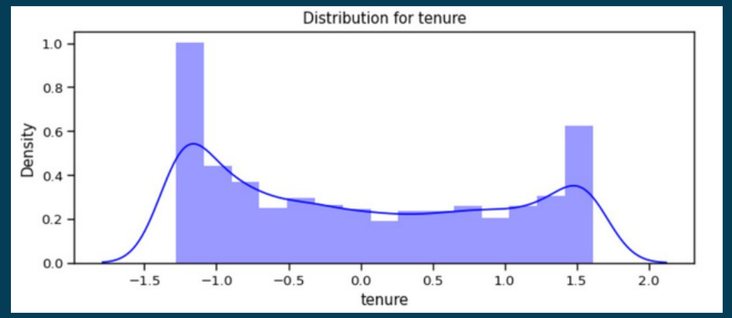
Data Prepocessing

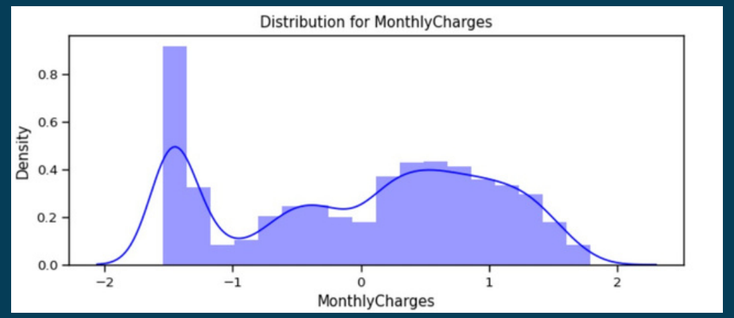
- At this stage, it is advisable to convert the categorical data to numerical format in order to achieve more optimal results in machine learning.
- The 3 factors of MonthlyCharges, TotalCharges, and Tenure will be converted to the same scale.

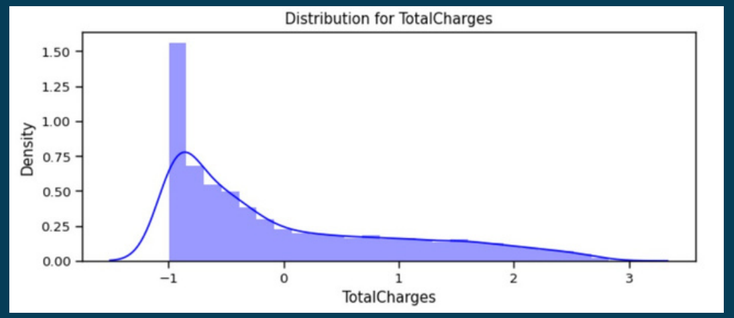










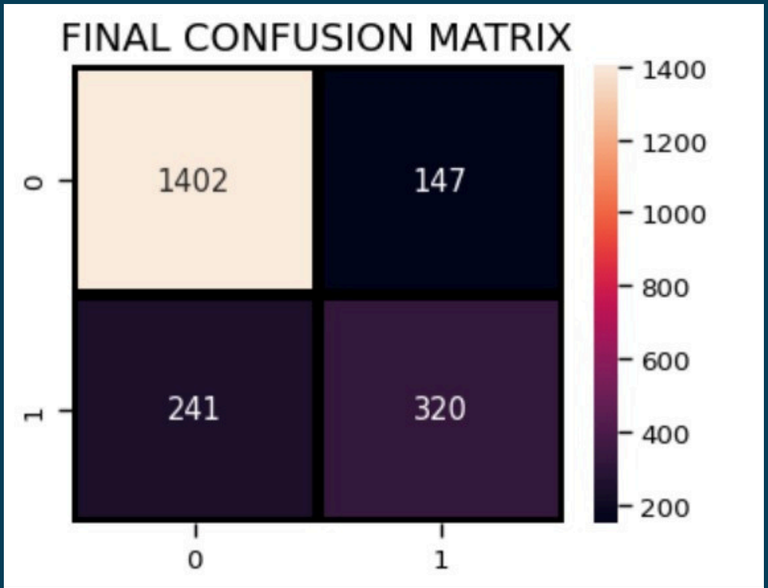


Model Evaluation & Prediction

	Accuracy	Precision	Recall	f1-score
AdaBoost Clasiffer	0.8037	0.81	0.81	0.80
Gradient Boosting Classifier	0.8075	0.80	0.81	0.80
Logistic Regression	0.8037	0.80	0.80	0.80
RandomForest	0.7791	0.80	0.80	0.80
Desicion Tree Clasifier	0.7298	0.81	0.81	0.73
K-Neighbors Classifier	0.7819	0.77	0.78	0.77
SVC	0.7341	0.54	0.73	0.62
Voting Classifier	0.8161	0.81	0.82	0.81

Model Evaluation & Prediction

The modelling evaluation results show that VotingClassifer runs well on the dataset with an accuracy of **0.8161**



Of the 1549 customers who did not actually churn, the model correctly predicted 1402 (ideal). But, the model incorrectly predicted 147 of them would churn (not ideal). There were 561 customers who actually churned, and the model correctly identified 324. However, the model missed 241 customers who churned by predicting they would stay subscribed (needs improvement).

```
print("Presentase Prediksi Churn:", presentase_prediksi_churn, "%")
```

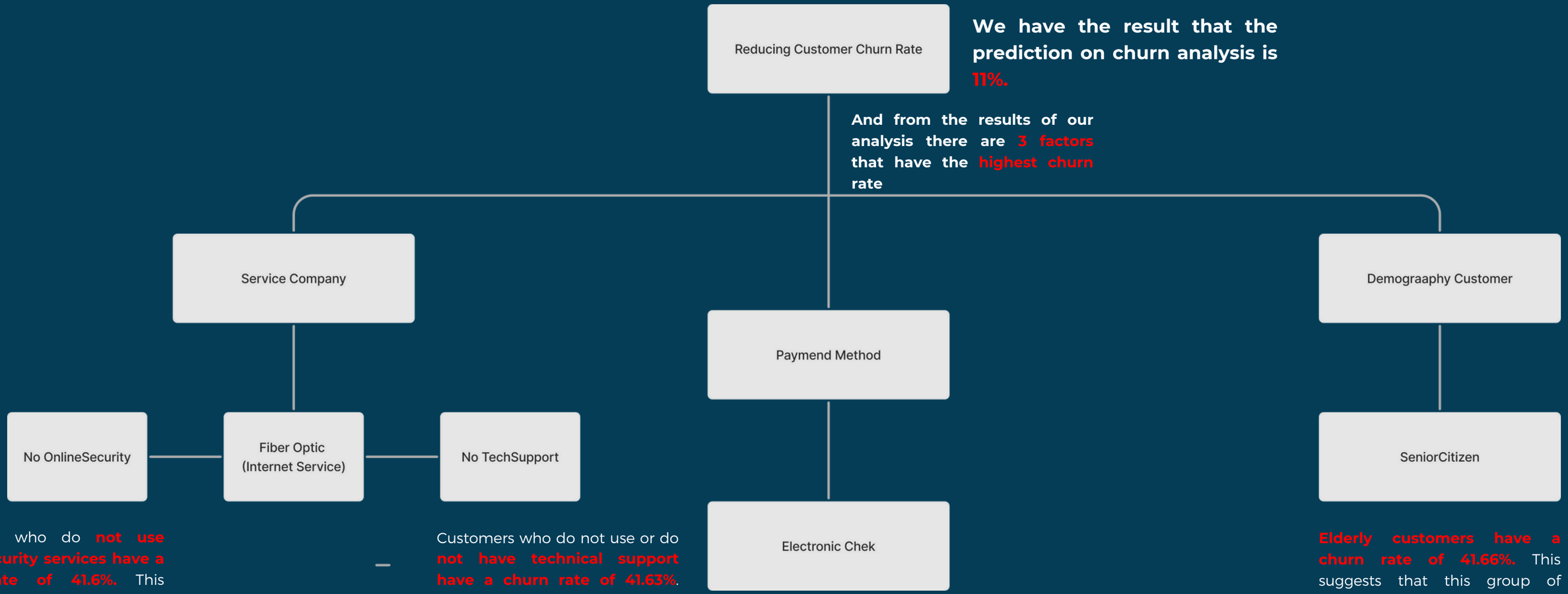
```
Presentase Prediksi Churn: 11.421800947867299 %
```

Conclusion

Based on the analysis results, the predicted churn rate shows a decrease of 15.5% to 11%. Despite the decrease, new customer acquisition is much more expensive than loyal customer retention. Therefore, it is necessary to further evaluate several components and factors of churn to improve service quality from various factors that contribute to the increase in churn.

From the analysis, it can be seen that the churn factors are only divided into two major scopes, namely internal customers and company services. Therefore, further data is needed such as surveys to customers to find out subjectively from the customer's side, what kind of service quality makes them remain a customer of the company's services.

Recommendation



Recommendation

No Online Security

Telecoms companies that experience customer churn due to a large number of customers not having online security can address this with education and awareness, easy-to-use security solutions, good customer support, promotions and incentives, and collaboration. Education and awareness raising can be done through educational campaigns, provision of educational materials, and consultancy services. Easy-to-use security solutions can include security packages, built-in security features, and security applications. Good customer support can include easily accessible customer service, security experts, and clear protocols. Promotions and incentives can include discounts, loyalty programmes, and competitions. Collaborations can be made with security providers, non-profit organisations, and online security initiatives.

Fiber Optic

To overcome the churn rate of Fibre Optic customers, further data is needed on complaints from customers about the service. The biggest possibility is the quality of service and prices that are not competitive compared to competitors so that it can provide solutions to customers. What can be done after identifying is to improve service quality and customer experience, offer loyalty programmes and retention solutions, collect feedback and improve customer satisfaction, and strengthen communication and marketing.

Recommendation

No Tech Support

To overcome high customer churn due to the absence of technical support, telecommunication companies should improve the availability and quality of technical support, focus on customer retention with loyalty programmes and churn solutions, as well as evaluate customer accessibility in obtaining technical support, and implement technical service SOP to facilitate customers.

Payment Method (Elecktronic Check)

increase payment channels that can facilitate customers. This allows customers to make payments with ease and low transfer fees. In addition, cooperate with banks that provide M-banking / E-Wallet with member or customer loyalty programmes as a customer retention strategy.

Senior Citizen

in the correlation metrix, it can be seen that Total Charges also affect customers with advanced age levels. the possible cause is that customers with advanced age are reluctant to spend too much / sensitive to product prices, this is because at an advanced age customers prefer to spend their money on cheaper providers and give their money to hobbies or family. To overcome this, the company must invite elderly customers to a customer retention or loyalty programme such as a family subscription programme for a certain period of time to retain customers.

Get in Touch

GIT HUB

https://github.com/Raflur-soejdamiko

EMAIL ADDRESS

raflifadilahrahman15@gmail.com

PHONE NUMBER

085895979448