

CUSTOMER CHURN PREDICTION

Data Analytics with cognos – Phase 5 DOCUMENTATION

Team Members:

- 1.ANANDHI V(au613021205002)
- 2.DEEPANA S(au613021205004)
- 3.HARITHA R(au613021205015)
- 4.KARTHIKA B(au613021205023)
- 5.SHARMILA C(au613021205049)

ABSTRACT:

The project involves using IBM Cognos to predict customer churn and identify factors influencing customer retention. The goal is to help businesses reduce customer attrition by understanding the patterns and reasons behind customers leaving. This project includes defining analysis objectives, collecting customer data, designing relevant visualization as in IBM Cognos, and building a predictive model.

1.OBJECTIVE

The objective of customer churn prediction is to proactively identify and predict which customers are likely to leave or discontinue their relationship with a business or service. By accurately forecasting customer churn, businesses can achieve several key goals:

1.1 Retention and Customer Loyalty:

- Retain valuable customers and build customer loyalty by addressing their concerns and needs before they decide to leave.

1.2 Revenue Protection and Growth:

- Protect existing revenue streams by reducing customer attrition and increasing the lifetime value of customers. It's often more cost-effective to keep existing customers than acquire new ones.

1.3 Resource Allocation:

- Optimize resource allocation by focusing efforts and resources on customers who are at the highest risk of churning, rather than applying a uniform approach to all customers.

1.4 Marketing and Communication Strategy:

- Tailor marketing and communication strategies to target specific customer segments with personalized retention offers and campaigns.

1.5 Enhanced Customer Experience:

- Improve the overall customer experience by addressing pain points and issues that may lead to churn, thereby ensuring a higher level of satisfaction.

1.6 Data-Driven Decision-Making:

- Leverage data analytics and machine learning to make informed decisions based on historical customer data rather than relying solely on intuition.

1.7 Proactive Problem Solving:

- Identify and address issues related to products, services, or customer support that are driving churn, helping to rectify problems and improve the business.

1.8 Competitive Advantage:

- Gain a competitive edge by outperforming competitors in customer retention, which can lead to long-term success in the market.

1.9 Customer Segmentation:

- Categorize customers into different segments based on their likelihood to churn, allowing for tailored strategies for each segment.

1.10 Cost Reduction:

- Minimize costs associated with customer acquisition, as it is often more expensive to acquire new customers than to retain existing ones.

1.11 Performance Measurement:

- Evaluate the effectiveness of retention strategies and campaigns by monitoring the actual churn rates and comparing them to predicted churn.

1.12 Real-Time Insights:

- Provide real-time or near-real-time insights into customer behavior, allowing for quick responses to changing customer dynamics.

1.13 Ethical Data Usage:

- Ensure that customer data is used in an ethical and responsible manner, respecting privacy and data protection regulations.

In summary, the primary objective of customer churn prediction is to leverage data and predictive models to anticipate customer departures, enabling businesses to take proactive measures to retain customers and enhance their overall success and sustainability.

2.DESIGN THINKING:

2.1 Analysis objective :

Conduct in-depth churn analysis to understand the patterns and trends associated with customers who have already churned. This could involve exploring when and why they left, which products or services they were using, and any commonalities among churned customers.

2.2 Data collection :

Customer churn data tells you how many customers have stopped using your product or service over a period of time.

Utilize data that the business already has in its internal records, such as customer registration information, account profiles, and purchase history.

This data may include age, gender, location, and contact information.

Conduct customer surveys to gather demographic information that may not be available in internal records.

2.3 Visualization Strategy :

Data visualization allows you to quickly detect patterns, trends, and outliers in churn, enabling you to make key business decisions that can address negative patterns before they get out of control.

Visualize customer segments derived from your analysis.

Use pie charts, bar charts, or tree maps to show the distribution of customers across segments.

Plan for regular updates of the visualizations as new data becomes available. Automation can help in keeping the visualizations up to date.

3.DEVELOPEMENT PHASE

3.1 Data Collection

Collect and gather relevant data, which may include customer demographics, transaction history, customer service interactions, and any other relevant information. Preprocess the data by handling missing values, encoding categorical variables, and scaling numerical features.

Dataset Link: <https://www.kaggle.com/datasets/blatchar/telco-customer-churn>

3.2 Data Preprocessing

- ❖ Clean and prepare the data for analysis.
- ❖ Handle missing values, outliers, and inconsistencies.
- ❖ Normalize or scale the data as needed

4.LOADING THE DATASET:

4.1.Importing libraries

Here, for preprocessing the dataset and manipulate the data, pandas is the library used to frame the data.

Code:

```
import pandas as pd
```

4.2.Loading the dataset

In this step, we are framing the data into the table using DataFrame in pandas, and display the head or 5 rows of the dataset.

Code:

```
# Replace with the actual filename
data=pd.read_csv("C:/Users/91962/Desktop/phase3_dataset.csv")
```

5.EXPLORING THE DATA SET:

After framing data, the first few or five rows of the data in displayed using the head() function.

Code:

```
data
```

OUTPUT:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	TechS
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	
1	5575-GNVOE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	
...
7038	8840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	...	Yes	
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	...	Yes	
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	...	No	
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...	No	
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	Yes	...	Yes	

7043 rows × 21 columns

6.DATA VISUALIZATION:

6.1 CORRELATION GRAPH:

```
import matplotlib.pyplot as plt

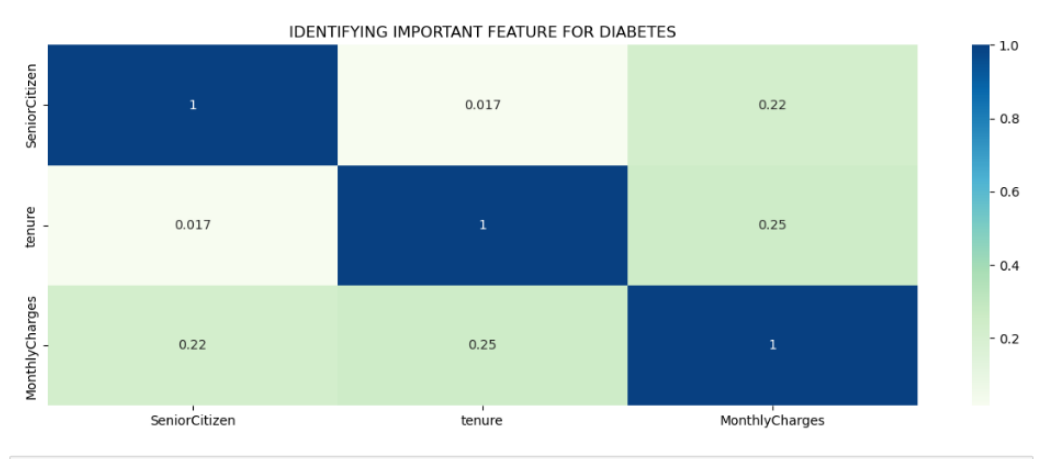
import seaborn as sns

plt.figure(figsize=(15,5))

sns.heatmap(data.corr(),cmap="GnBu",annot=True)

plt.title("IDENTIFYING IMPORTANT FEATURE FOR DIABETES")
```

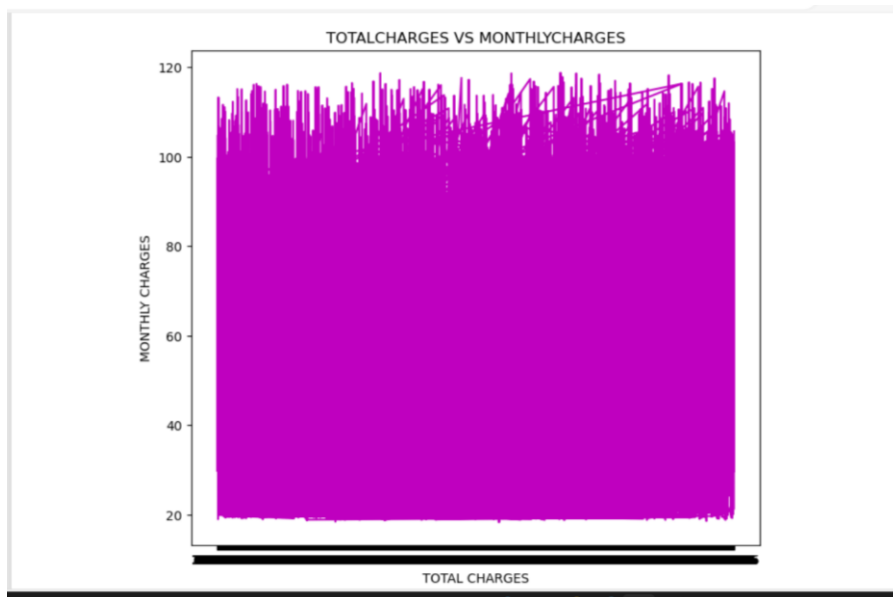
OUTPUT:



6.2 PLOT:

```
plt.figure(figsize=(8,7))  
plt.plot(data.TotalCharges,data.MonthlyCharges,'m')  
plt.title("TOTALCHARGES VS MONTHLYCHARGES")  
plt.xlabel("TOTAL CHARGES") plt.ylabel("MONTHLY CHARGES")
```

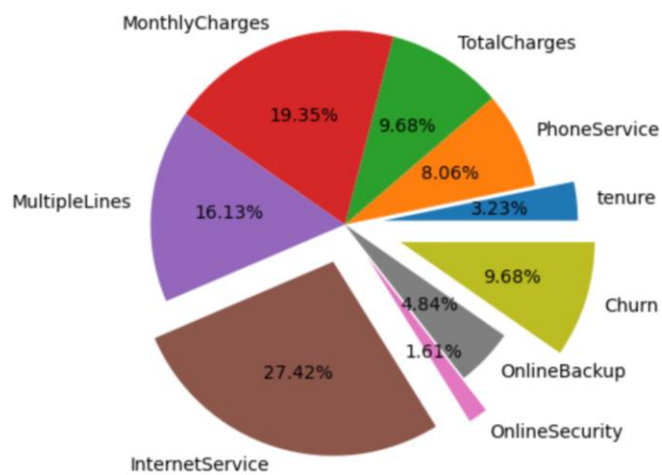
OUTPUT:



6.3 PIE PLOT:

```
x=[2000,5000,6000,12000,10000,17000,1000,3000,6000]  
y=['tenure','PhoneService','TotalCharges','MonthlyCharges','MultipleLines','InternetService','OnlineSecurity','OnlineBackup','Churn']  
plt.pie(x,labels=y,autopct='%0.2f%%',explode=[0.2,0,0,0,0,0.2,0.2,0,0.3])
```

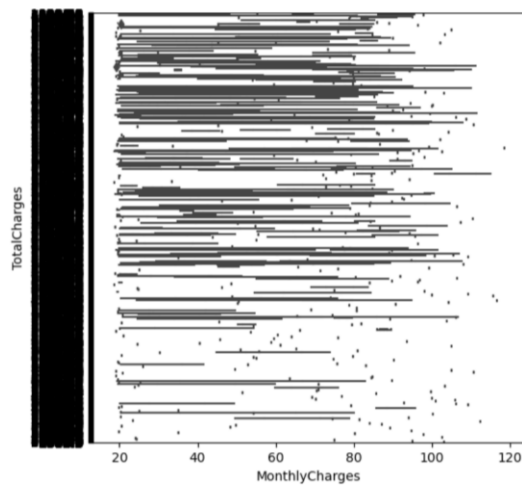
OUTPUT:



6.4 BOX PLOT

```
plt.figure(figsize=(6,6))  
sns.boxplot(x='MonthlyCharges',y='TotalCharges',data=data)
```

OUTPUT:



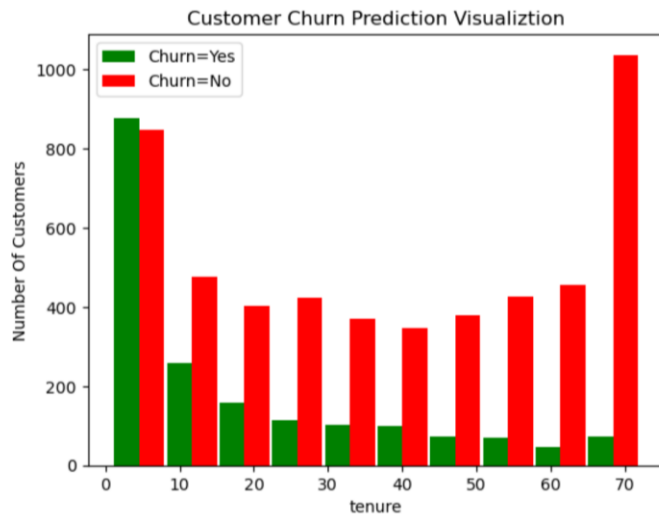
6.5 HISTOGRAM

```
tenure_churn_no = df1[df1.Churn=='No'].tenure  
tenure_churn_yes = df1[df1.Churn=='Yes'].tenure  
plt.xlabel("tenure")  
plt.ylabel("Number Of Customers")  
plt.title("Customer Churn Prediction Visualiztion")
```

```
plt.hist([tenure_churn_yes, tenure_churn_no], rwidth=0.95, color=['green',
'red'],label=['Churn=Yes','Churn=No'])

plt.legend();
```

OUTPUT:

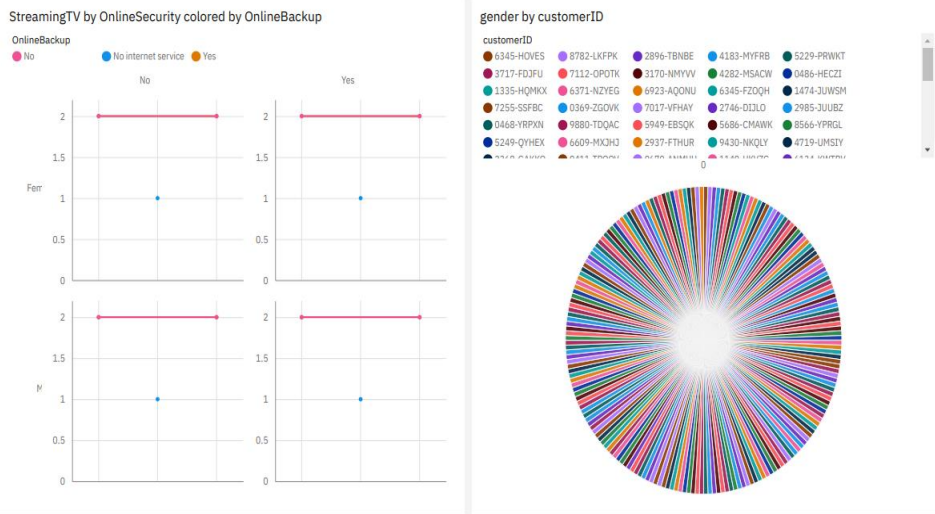


7. DATA VISUALIZATION

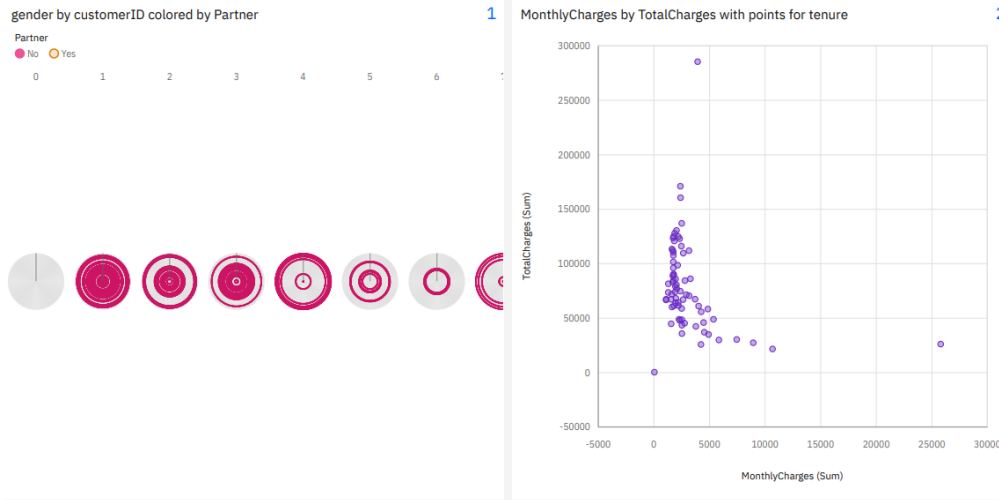
7.1 DATA ANALYTICS WITH IBM COGNOS

- ❖ IBM Cognos Introduction
 - ◆ Introduce IBM Cognos as a tool for data analytics.
- ❖ Data Exploration
 - ◆ Showcase how IBM Cognos aids in exploring and understanding the dataset.
- ❖ Visualization
 - ◆ Demonstrate the creation of visualizations in IBM Cognos.

TAB 1

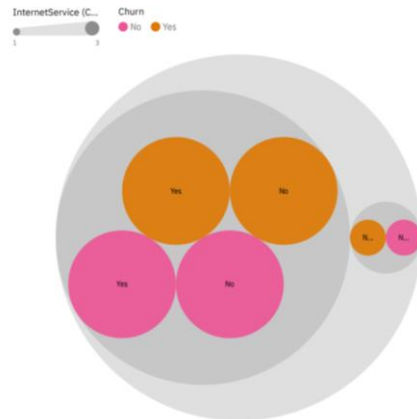


TAB 2

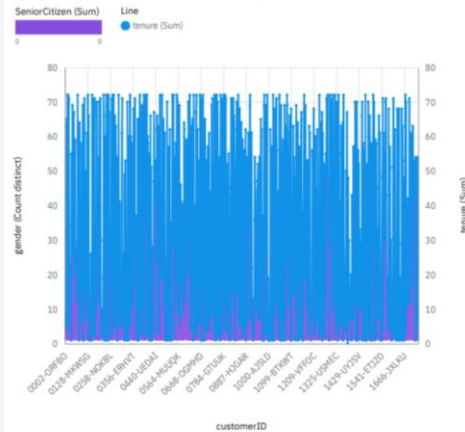


Tab 2

PhoneService and MultipleLines hierarchy colored by Churn and sized by InternetService

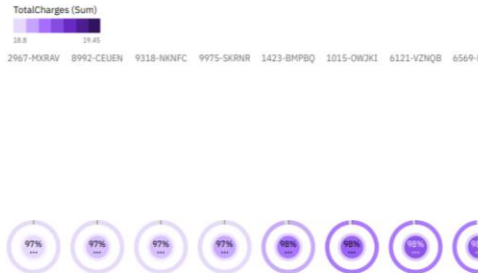


tenure and gender for customerID colored by SeniorCitizen

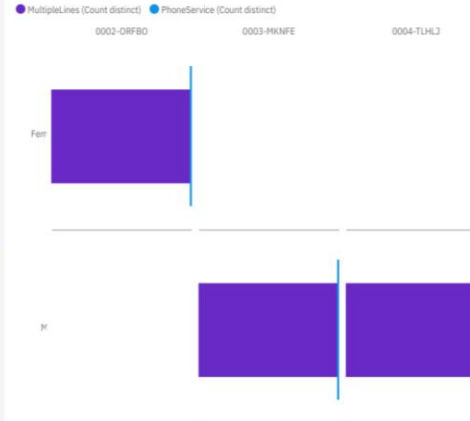


TAB 4

MonthlyCharges, TotalCharges, customerID

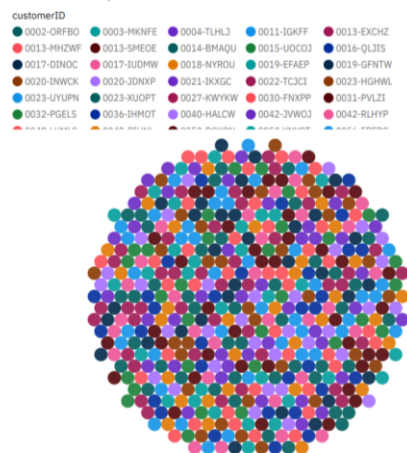


MultipleLines compared to PhoneService

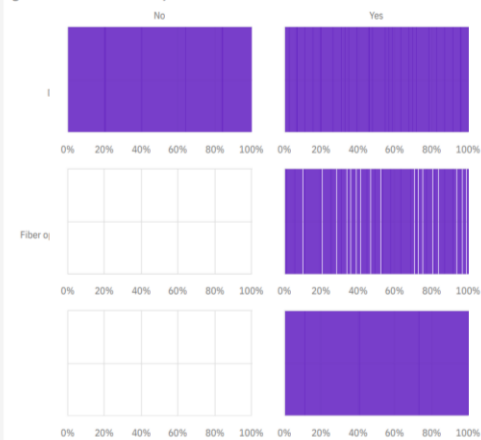


TAB 5

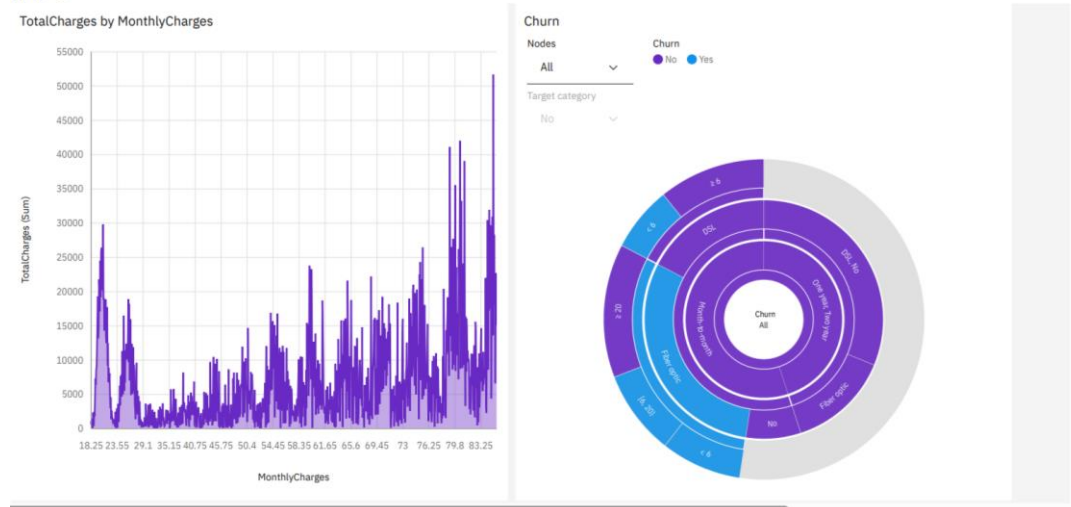
tenure colored by customerID



gender and OnlineSecurity for customerID



TAB 6

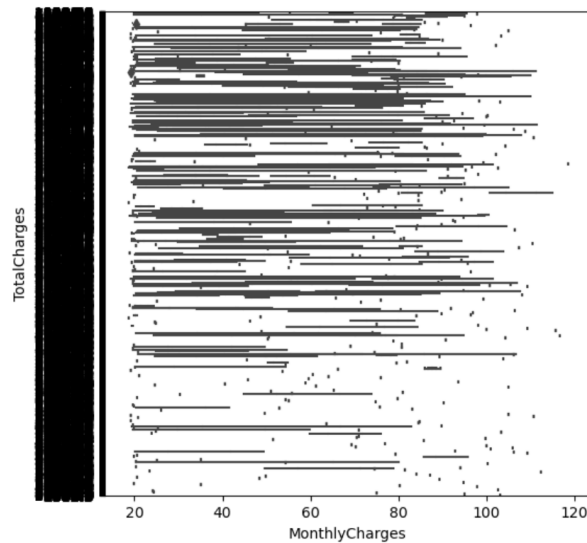


7.2 DATA VISUALIZATION WITH JUPYTER NOTEBOOK

- ❖ Jupyter Notebook
 - ◆ Introduction Present Jupyter Notebook as a tool for data analysis and visualization.
- ❖ Visualizing CUSTOMER CHURN PREDICTION
 - Use Jupyter Notebook to create visualizations of customer churn prediction.
- ❖ Geographic Mapping
 - ◆ Visualize customer by location using Jupyter Notebook.
- ❖ Time Series Analysis
 - Analyze temporal changes in customer using Jupyter Notebook

```
In [5]: plt.figure(figsize=(6,6))
sns.boxplot(x='MonthlyCharges',y='TotalCharges',data=data)

Out[5]: <Axes: xlabel='MonthlyCharges', ylabel='TotalCharges'>
```



8.PREDICTIVE MODELLING:

Building a predictive model for customer churn prediction involves using historical customer data to create a model that can forecast which customers are likely to churn in the future. Here are the key steps to develop a predictive model for customer churn:

8.1 Data Collection and Preparation:

Collect and clean historical customer data. This data should include features such as customer demographics, transaction history, customer support interactions, feedback, and, most importantly, a churn indicator (whether the customer churned or not).

8.2 Feature Selection and Engineering:

Identify relevant features that might impact customer churn, like customer tenure, purchase frequency, average transaction value, and customer feedback.

Create new features through feature engineering, such as churn-related scores or aggregations of historical data.

8.3 Data Split:

Split the data into training and testing sets. The training data is used to build the predictive model, while the testing data is used to evaluate the model's performance.

8.4 Model Selection:

Choose an appropriate machine learning model for your churn prediction task. Common models include logistic regression, decision trees, random forests, support vector machines, and neural networks. The choice depends on your dataset and goals.

8.5 Model Training:

Train the selected model using the training dataset. The model learns the patterns and relationships in the data that can help predict churn.

8.6 Model Evaluation:

Evaluate the model's performance on the testing dataset using appropriate metrics such as accuracy, precision, recall, F1-score, and ROC curves. The choice of metrics depends on your specific objectives.

8.7 Hyperparameter Tuning:

Fine-tune the model by adjusting hyperparameters to optimize its performance. Techniques like grid search or random search can be used.

8.8 Model Validation:

Validate the model's performance on a separate dataset, or through cross-validation, to ensure that the model generalizes well to new, unseen data.

8.9 Deployment:

Deploy the trained model within your business environment. This may involve integrating it into customer management systems, CRM software, or other relevant platforms.

8.10 Monitoring and Maintenance:

Implement monitoring systems to track the model's performance in real-time. Regularly retrain the model to adapt to changing customer behavior.

8.11 Actionable Insights:

Use the churn predictions and insights generated by the model to implement targeted retention strategies, such as personalized marketing campaigns, loyalty programs, or improving customer service.

8.12 Ethical Considerations:

Ensure that customer data is handled ethically and in compliance with data privacy regulations. Protect customer privacy and data security.

It's important to note that the success of your churn prediction model depends on the quality of data, the choice of features, and the selected model. Continuous improvement, regular updates, and a data-driven approach are key to achieving better accuracy and effectiveness in predicting and preventing customer churn.

9.INSIGHTS OF CUSTOMER CHURN PREDICTION:

Obtaining actionable insights from customer churn analysis is essential for developing effective strategies to retain customers and enhance business performance. Here are some valuable insights that can be derived from customer churn analysis:

9.1 Churn Rate Trends:

Identify trends in churn rates over time. Analyze whether churn is increasing, decreasing, or remaining steady. Seasonal patterns or external factors may influence these trends.

9.2 Segmentation:

Segment your customer base to understand which customer groups are more prone to churn. Segmentations could be based on demographics, behavior, or purchase history.

9.3 Churn Drivers:

Determine the factors that drive customer churn. Are there specific product or service issues, pricing concerns, or customer service problems that lead to higher churn rates?

9.4 Customer Lifetime Value (CLV):

Assess the CLV for both churned and retained customers. This helps in understanding the revenue impact of churn and in prioritizing high-value customers.

9.5 Churn Predictors:

Identify the key indicators that precede customer churn. These may include decreased purchase frequency, increased complaints, or extended periods of inactivity.

9.6 Customer Feedback Analysis:

Analyze customer feedback, surveys, and reviews to pinpoint issues and concerns that contribute to churn. Addressing these issues can help reduce churn.

9.7 Churn Timing:

Determine the timing of customer churn. Understanding when customers are most likely to churn can inform the timing of retention efforts or campaigns.

9.8 Retention Strategy Effectiveness:

Assess the effectiveness of your retention strategies and campaigns. Are they successful in reducing churn, and which strategies perform best?

9.9 Customer Satisfaction:

Measure customer satisfaction levels and their impact on churn. Identify the factors that influence satisfaction, and work to improve them.

9.10 Competitor Analysis:

Analyze the competitive landscape. Are customers leaving for competitors? Understanding the competition can help in crafting counter-strategies.

10.CONCLUSION:

In today's highly competitive business environment, customer churn prediction and retention strategies are integral to customer-focused business models. By understanding and proactively addressing the factors contributing to churn, businesses can improve customer satisfaction, maintain revenue, and establish a solid foundation for long-term growth and success.

LINK FOR JUPYTER NOTEBOOK (ipynb) :

<https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/PHASE%203%20JUPYTER%20NOTEBOOK.ipynb>

https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase4.ipynb

LINK FOR JUPYTER NOTEBOOK (pdf) :

[https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase4\(Part%201\).pdf](https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase4(Part%201).pdf)

LINK FOR IBM COGNOS VISUALIZATION (pdf):

[https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase4\(part%202\).pdf](https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase4(part%202).pdf)

GITHUB LINK:

PHASE 1: https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase1.pdf

PHASE 2: https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase2.pdf

PHASE 3: https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase3.pdf

PHASE 4: [https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase4\(part%202\).pdf](https://github.com/ANANDHI-VELU/CustomerChurnPrediction/blob/main/DAC_Phase4(part%202).pdf)

