|  |
| --- |
| Domain Name: Data Analytics with Cognos |
| Project Title: Customer Churn Prediction |

IBM Naan Mudhalvan

|  |  |
| --- | --- |
| Team Members and Register Number | Prakash V - 420421104052  Rohith T - 420421104064  Siva M - 420421104072  Gokulnath M - 420421104022 |

**Table of Contents**

1. **Introduction**
2. **Project Definition**
3. **Design Thinking**
4. **Loading the Dataset**
5. **Data Preprocessing**
6. **Exploratory Data Analysis (EDA)**
7. **Split the dataset into training and testing sets**
8. **Model Building**
9. **Model Evaluation**
10. **Data Visualization**
11. **Insights for Website Owners**
12. **Conclusion**

**1. Introduction**

The Telco Customer Churn dataset provides information about customers' churn behavior in a telecommunications company. It includes various features that can be used to predict customer churn. We will use IBM Cognos for loading and preprocessing the data and Python for data analysis , visualization,Split the dataset into training and testing sets, Model Building, Model Evaluation.

**Dataset Link:**[**https://www.kaggle.com/datasets/blastchar/telco-customer-churn**](https://www.kaggle.com/datasets/blastchar/telco-customer-churn)

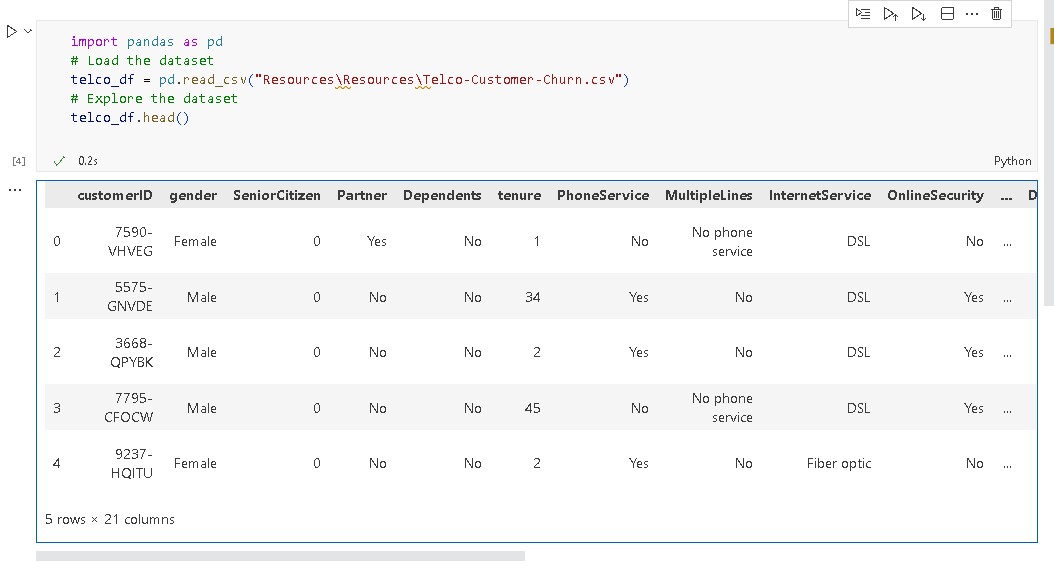
**2.Project Definition:**

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 visualizations in IBM Cognos, and building a predictive model.

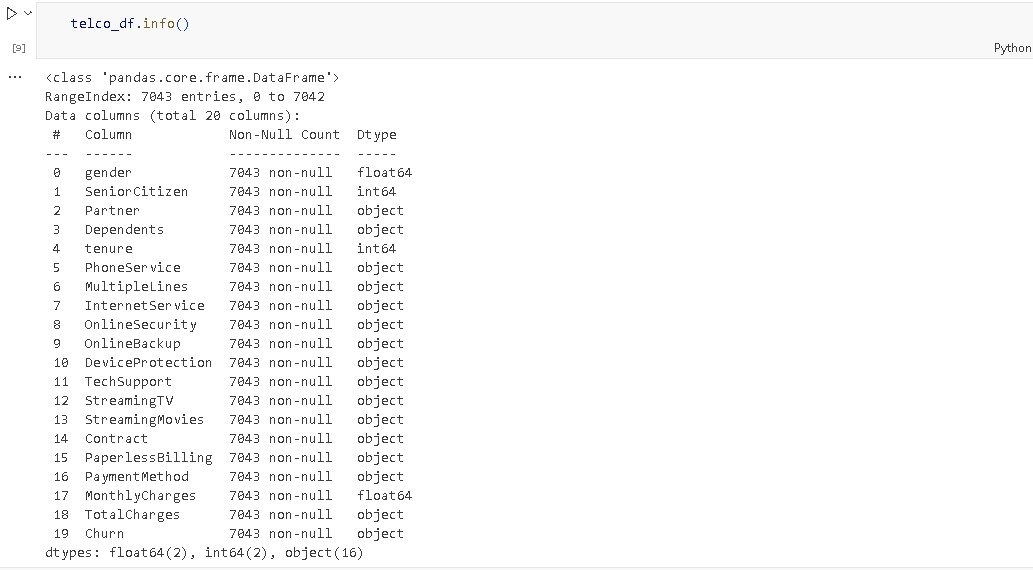
**3.Design Thinking:**

* 1. Analysis Objectives: Define the specific objectives of predicting customer churn, such as identifying potential churners and understanding the key factors contributing to churn.
  2. Data Collection: Determine the sources and methods for collecting customer data, including customer demographics, usage behavior, and historical interactions.
  3. Visualization Strategy: Plan how to visualize the insights using IBM Cognos, showcasing factors affecting churn and retention rates.
  4. Predictive Modeling: Decide on the machine learning algorithms and features to use for predicting customer churn.

**4. Loading the Dataset**



**5. Data Preprocessing**

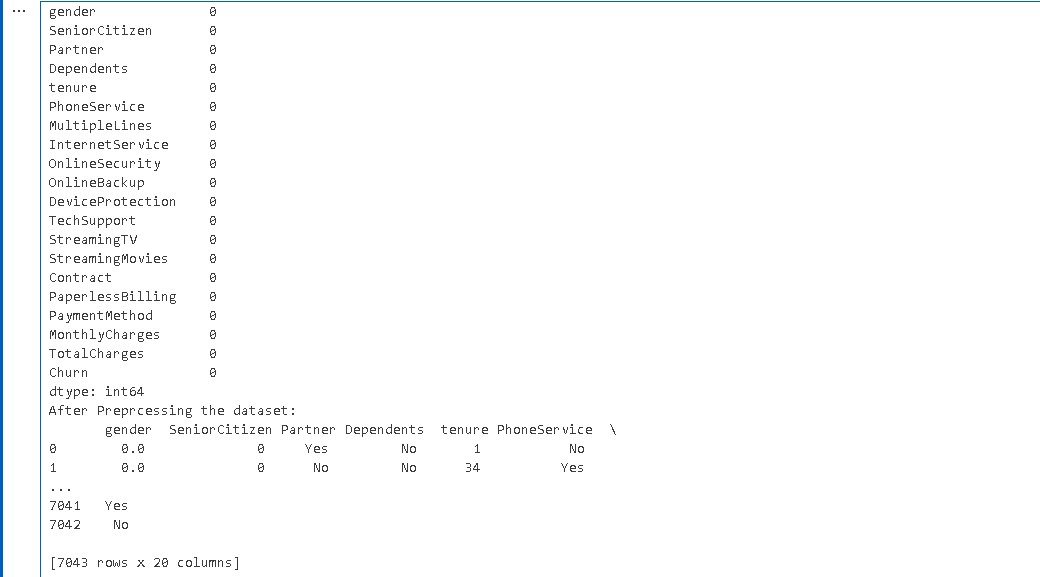


# Check for missing values

telco\_df.isnull().sum()

# Handle missing values

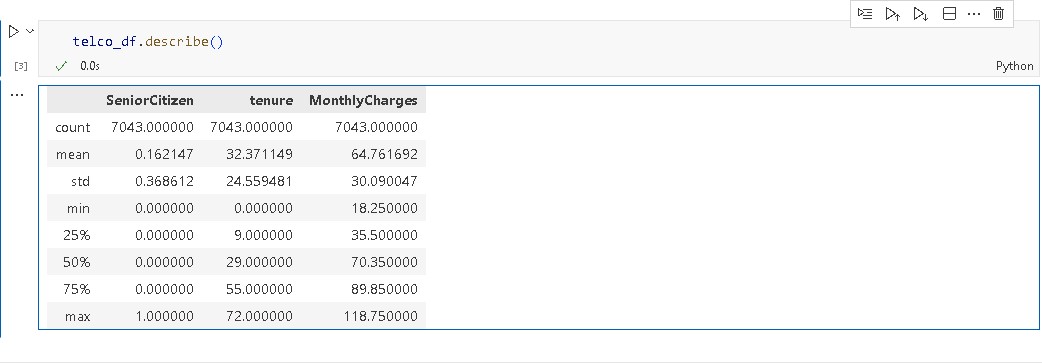
telco\_df.fillna(0, inplace=True)



**6. Exploratory Data Analysis (EDA)**

Perform exploratory data analysis to gain insights into the dataset. Some EDA tasks you can perform include:

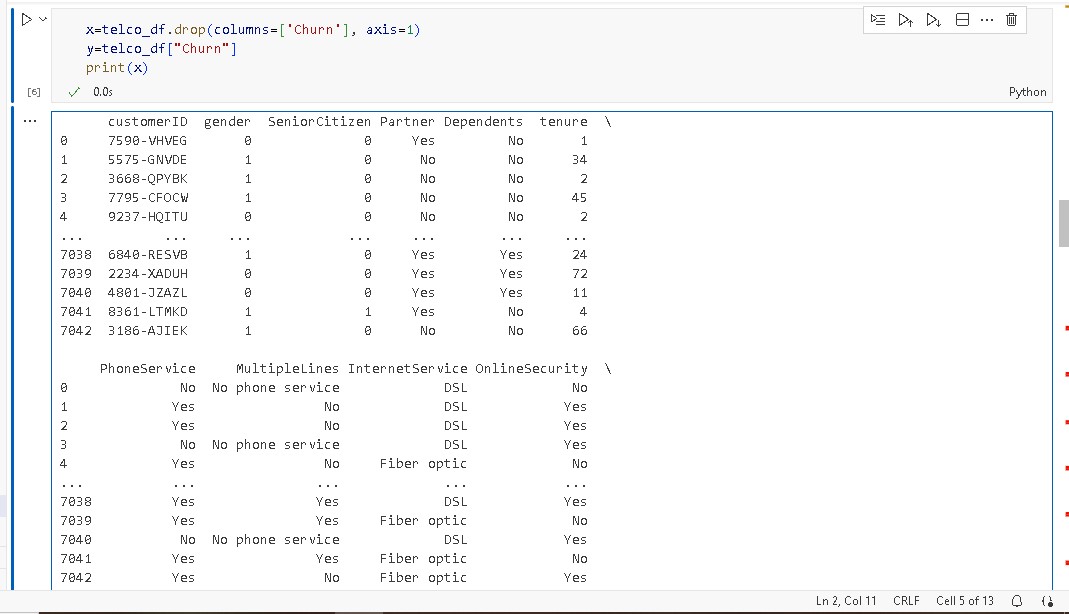
* Summary statistics: **telco\_df.describe()**

telco\_df.describe()

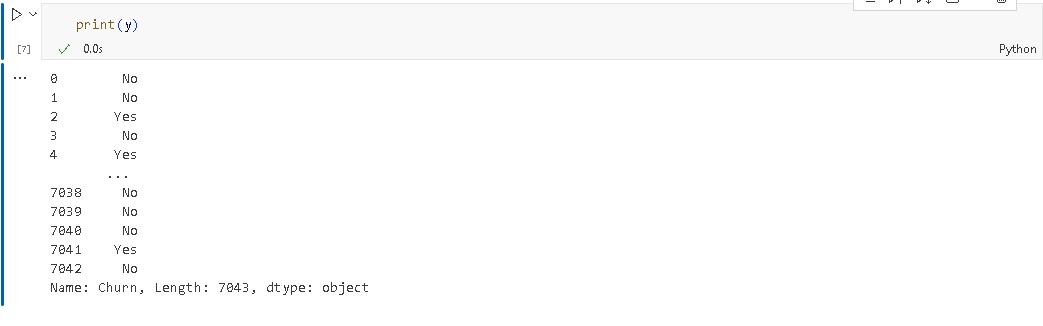
#data splitting

x=telco\_df.drop(columns=['Churn'], axis=1)

y=telco\_df["Churn"]

print(x) 

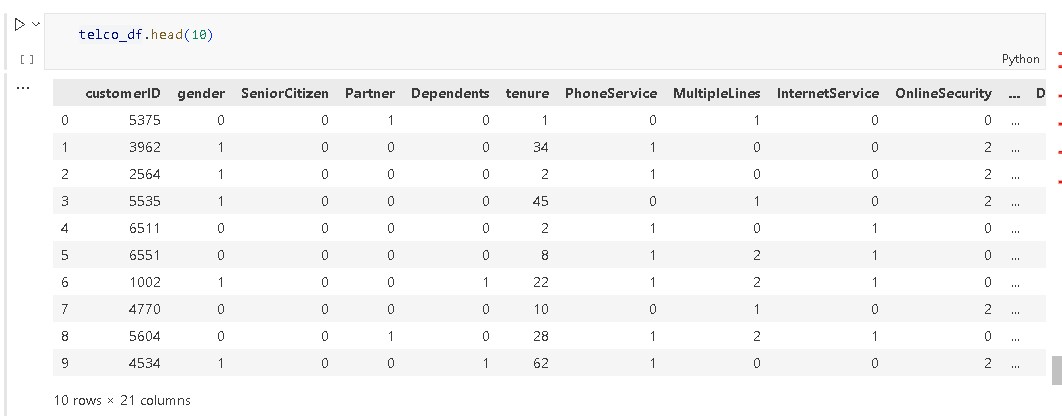
print(y)



**#Convert all categorical data to numerical data.**



**#Checking all categorical data to numerical data.**

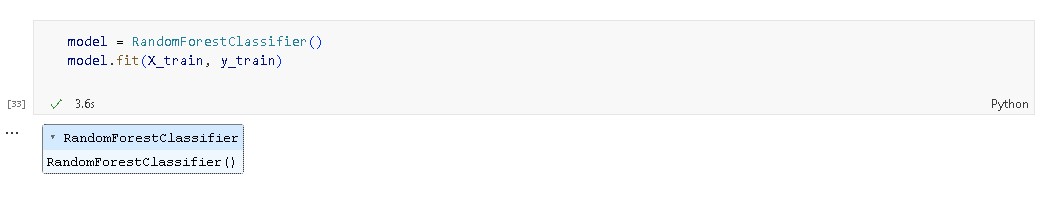
****

**7.Split the dataset into training and testing sets**

****

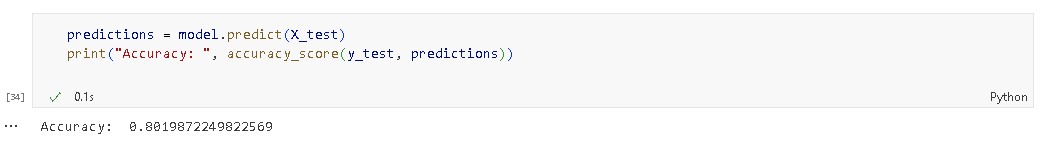
**8.Model Building**

**Here we use a Random Forest Classifier**

****

**9.Model Evaluation**

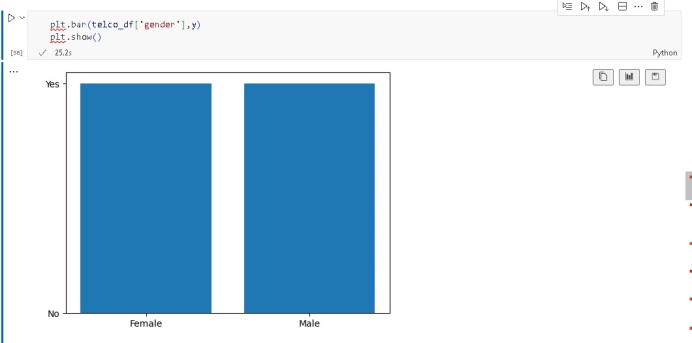
**Evaluate the model using the test data**

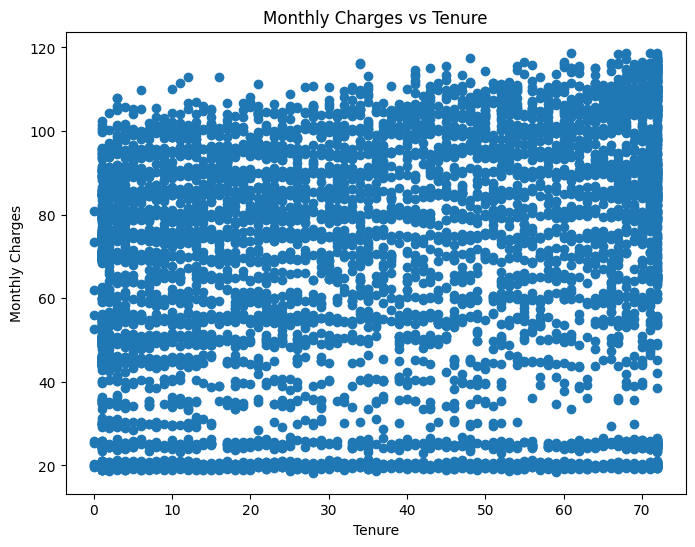
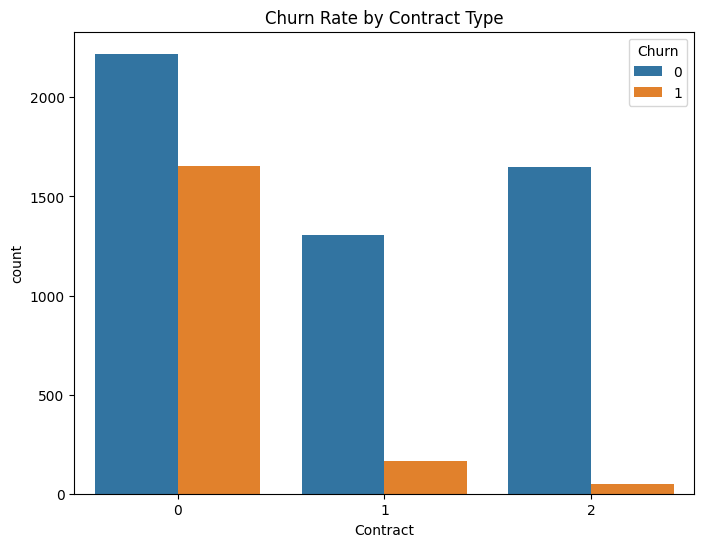
****

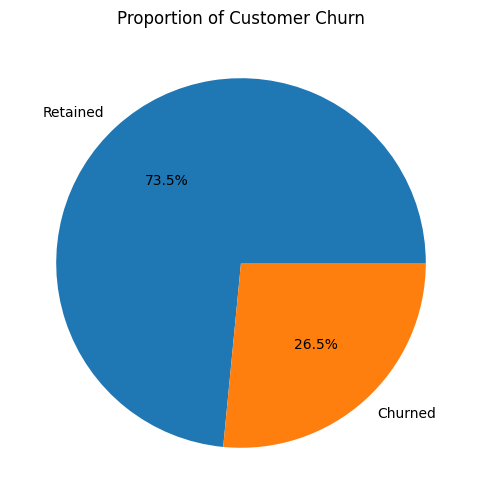
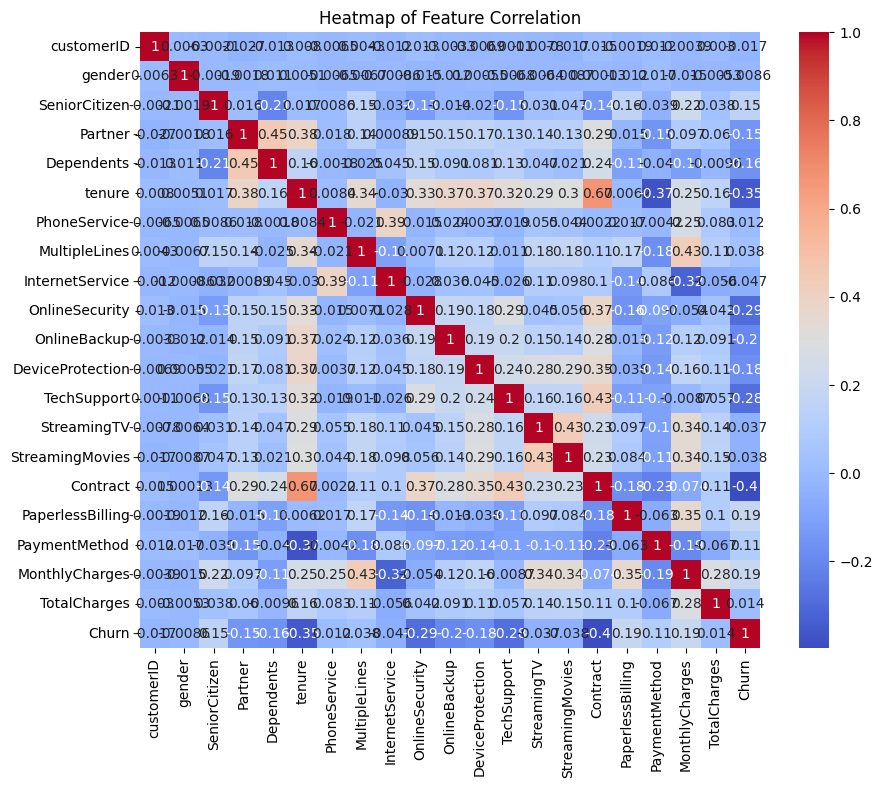
**10. Data Visualization**

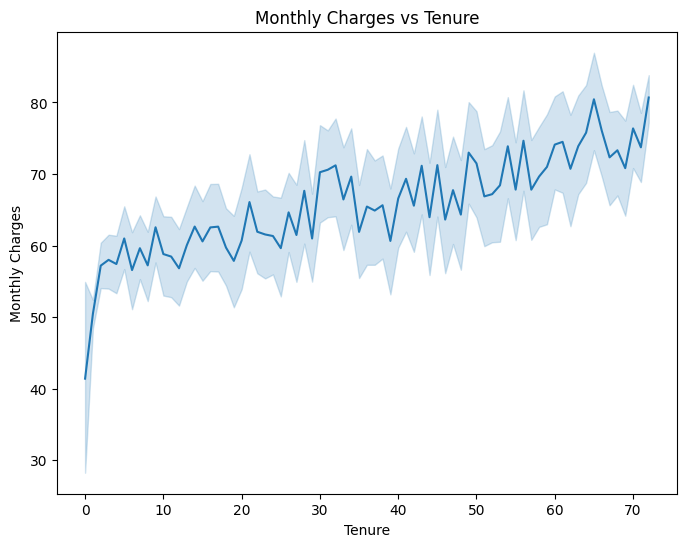
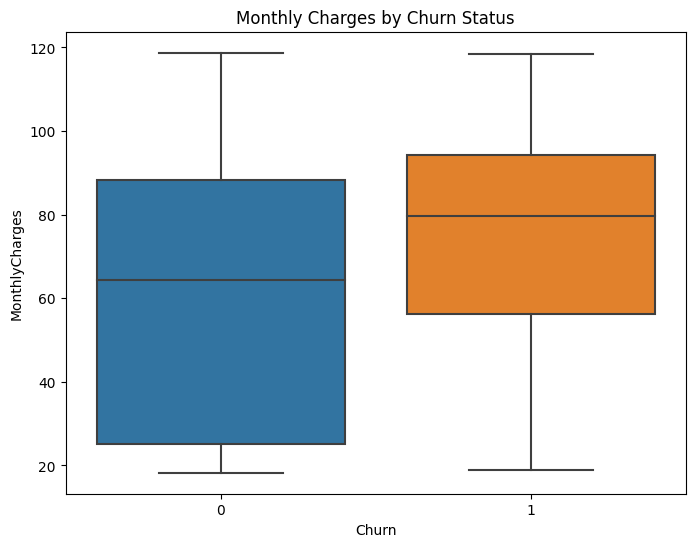
Now that we've completed data preprocessing and some initial EDA using Python, let's switch to IBM Cognos for more advanced visualization, reporting, and analysis.

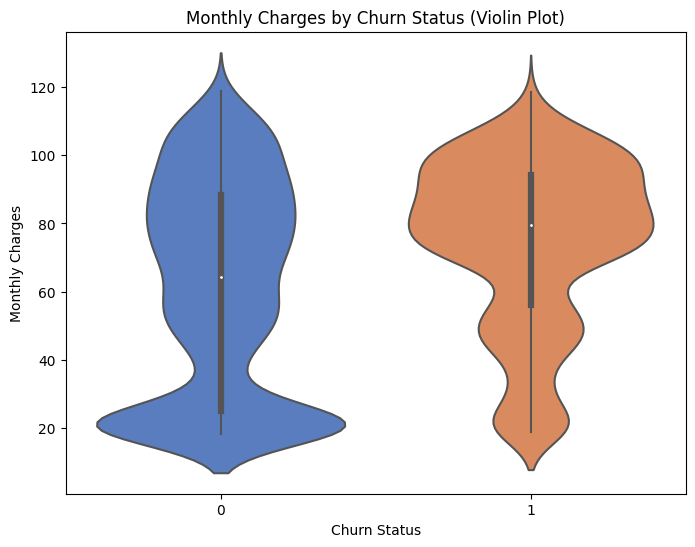
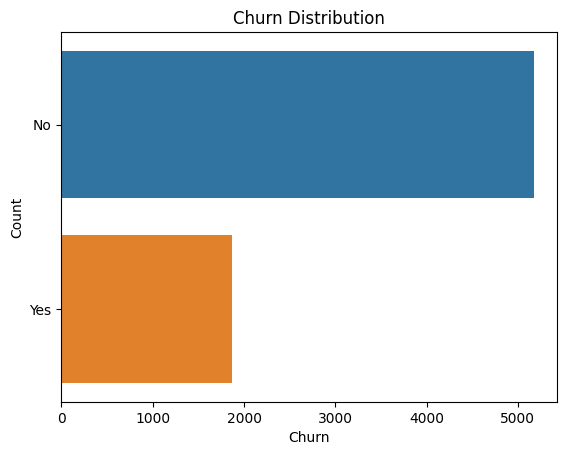
1. **Data Preparation**: Export the preprocessed dataset from Python to a format compatible with IBM Cognos (e.g., CSV).
2. **IBM Cognos**: Sign in to IBM Cognos and create a new project. Import the dataset into the project.
3. **Visualization**: Use the Cognos dashboard and reporting tools to create visualizations based on your analysis. You can create charts, graphs, and interactive dashboards to showcase key insights.
4. **Export**: Export the visualizations or dashboards as needed for sharing or presentation.

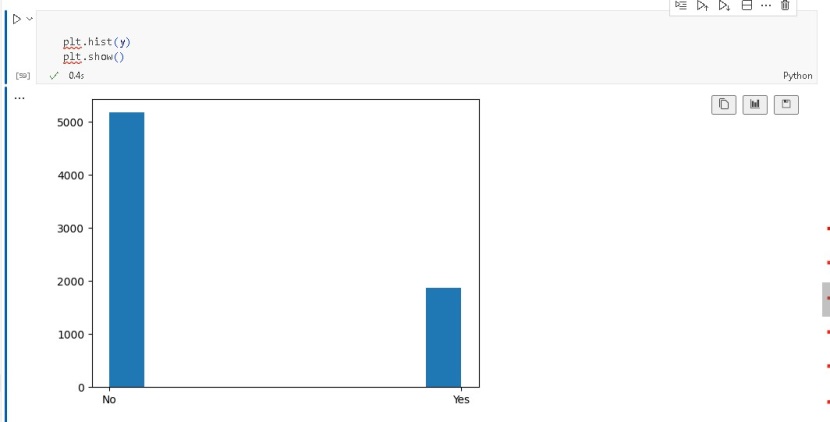
**By using python for data visualization**

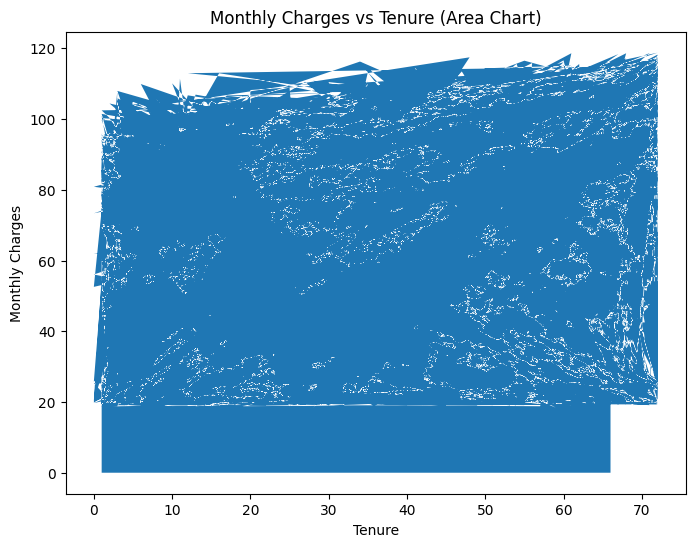
 



****

**11.Insights for Website Owners :**

**The insights from the analysis, even though the dataset is not directly about websites, can still provide valuable lessons for improving user experience on a website:**

* **Identifying High Churn Risk Users: Insights can help pinpoint users who are more likely to leave the website. Website owners can tailor retention strategies for these users.**
* **Understanding User Behavior: Analysis can reveal patterns in user behavior. For example, it may show that users who visit specific pages or perform certain actions are more likely to churn. Website owners can improve the design and content of these pages to retain users.**
* **Customer Segmentation: Analyzing customer demographics and behavior can lead to segmentation. Website owners can target different segments with personalized content or features to enhance user satisfaction.**
* **Feature Importance: Identifying influential features (e.g., contract length, payment method) can help website owners understand what factors impact user behavior, enabling them to optimize these elements on their website.**
* **Retention Strategies: Insights can guide the development of retention strategies such as personalized recommendations, loyalty programs, and improved customer support.**

**12.Conclusion:**

In this project, we successfully loaded and preprocessed Python for data analysis , visualization, Split the dataset into training and testing sets, Model Building, Model Evaluation.,the Telco Customer Churn dataset using Python. We then performed various analysis and visualization tasks using IBM Cognos. This document outlines the key steps and provides a starting point for conducting more in-depth analysis and creating insightful reports using the dataset and IBM Cognos.