|  |
| --- |
| 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. **Prerequisites**
3. **Project Setup**
4. **Loading the Dataset**
5. **Data Preprocessing**
6. **Exploratory Data Analysis (EDA)**
7. **Data Visualization**
8. **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 and visualization.

**2. Prerequisites**

* IBM Cognos setup and access
* Python environment (Jupyter Notebook or any other IDE)
* Python libraries: Pandas, NumPy, Matplotlib, Seaborn

**3. Project Setup**

1. Download the Telco Customer Churn dataset from the Kaggle link provided and save it in a directory of your choice.
2. Open IBM Cognos and create a new project or workspace.

**4. Loading the Dataset**

We'll start by loading the Telco Customer Churn dataset using Python. You can use the Pandas library for this task. Make sure you've installed Pandas before proceeding.Let's start by loading the Telco Customer Churn dataset into a Jupyter Notebook.

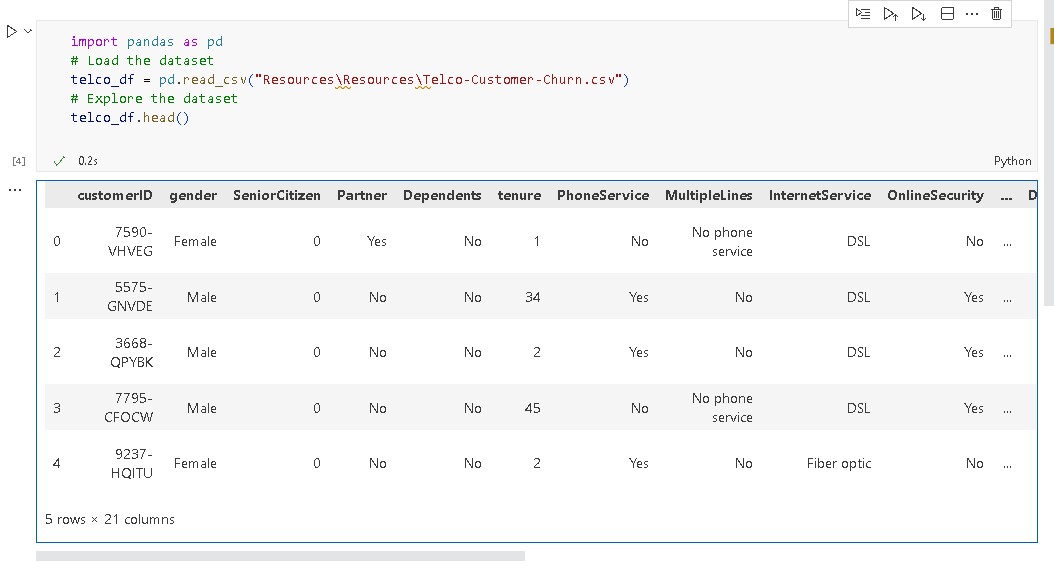
import pandas as pd

# Load the dataset

telco\_df = pd.read\_csv("telco-customer-churn.csv")

# Explore the dataset

telco\_df.head()

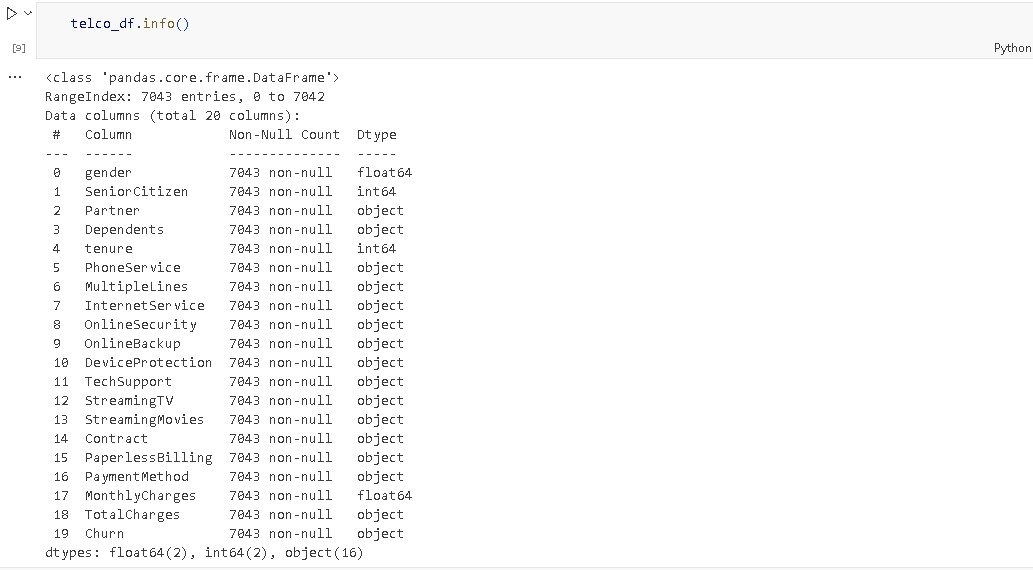


**5. Data Preprocessing**

Data preprocessing is crucial for analysis. We will handle missing values, convert data types, and clean the dataset.

#information about dataset

telco\_df.info()

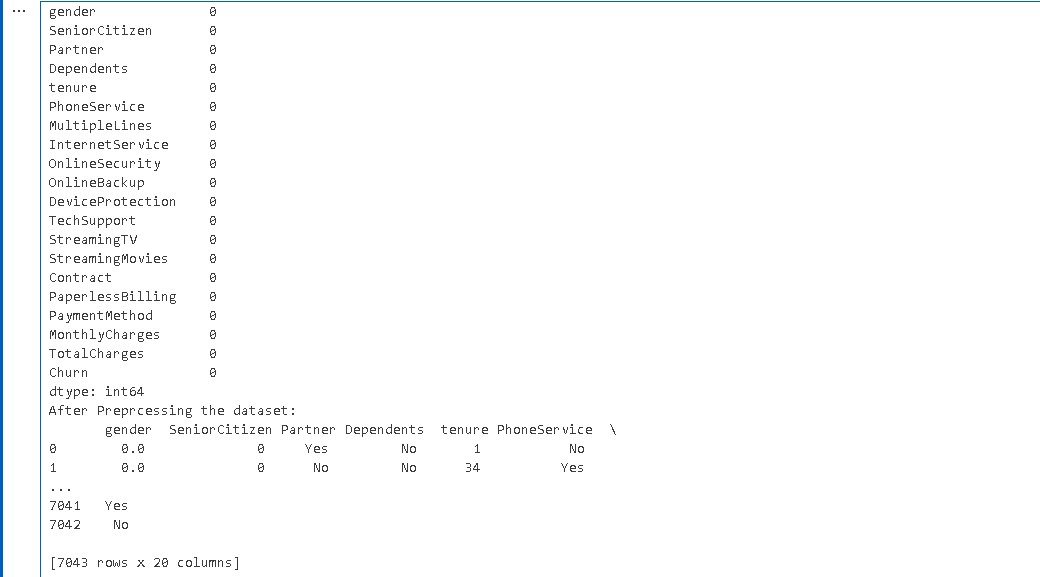


# Check for missing values

telco\_df.isnull().sum()

# Handle missing values

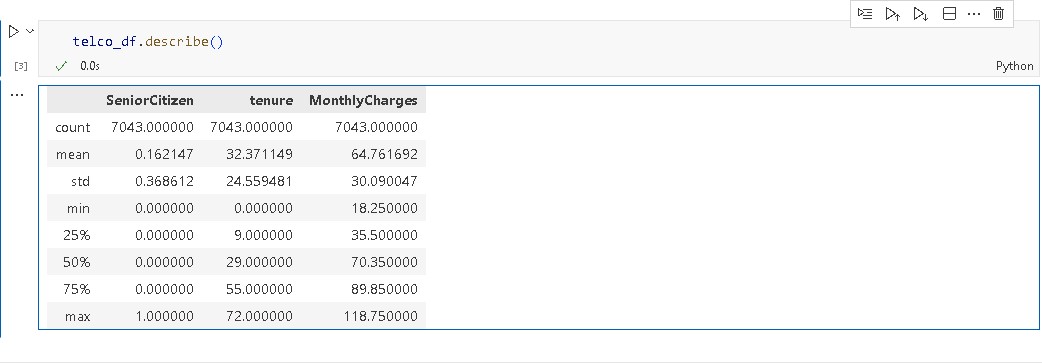
telco\_df.fillna(0, inplace=True)



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

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

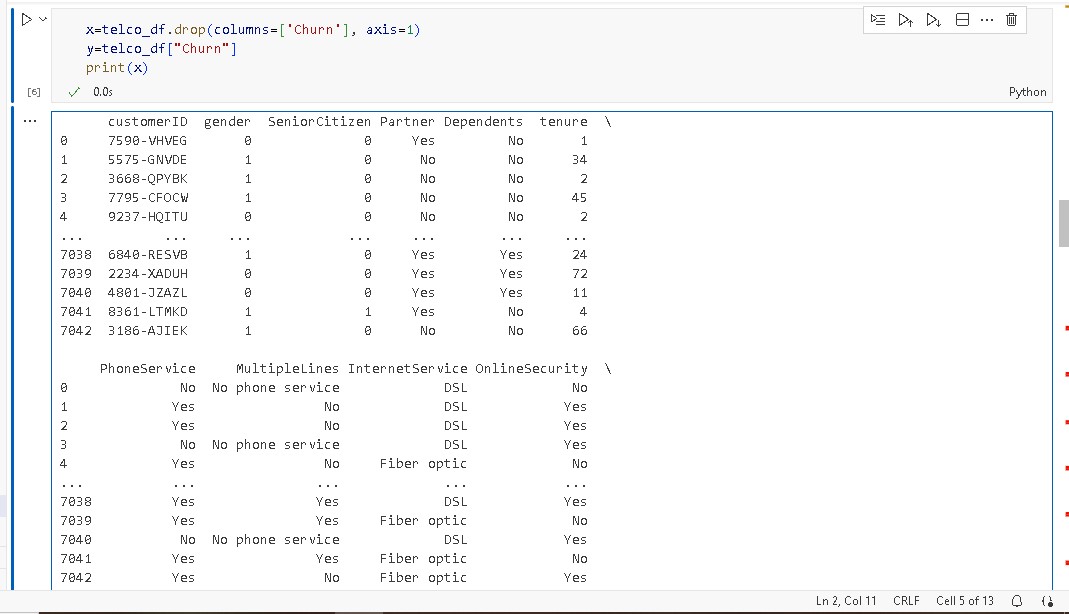
* Summary statistics: **telco\_df.describe()**

telco\_df.describe()

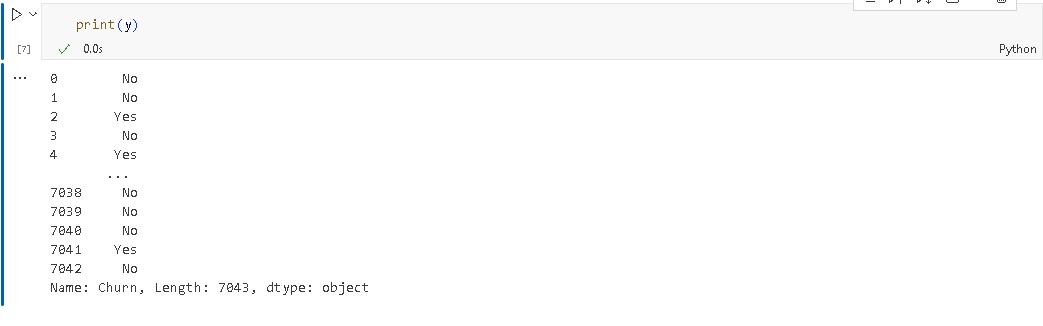
#data splitting

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

y=telco\_df["Churn"]

print(x) 

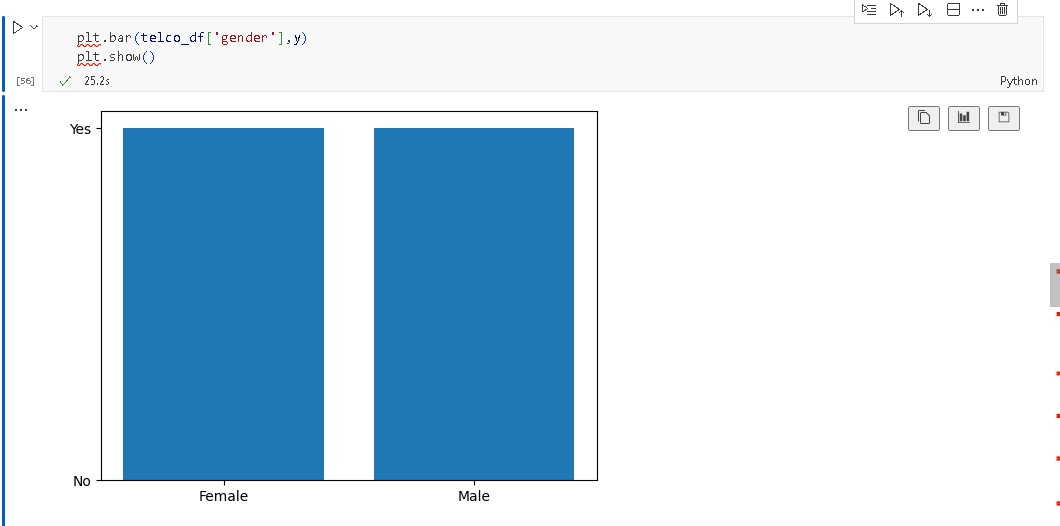
print(y)

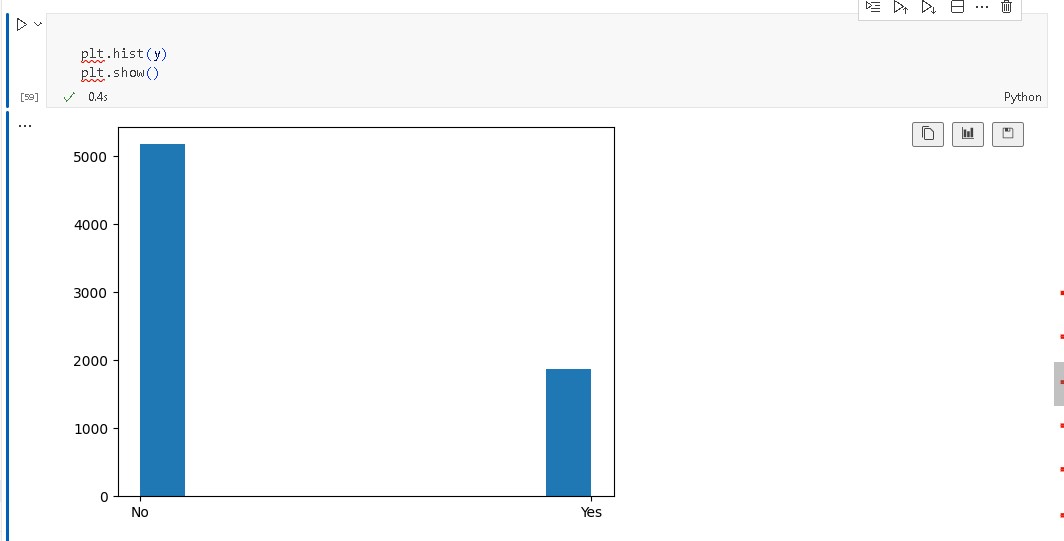


**7. Data Visualization**

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**



**8.Conclusion**

In this project, we successfully loaded and preprocessed 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.