Phase 3

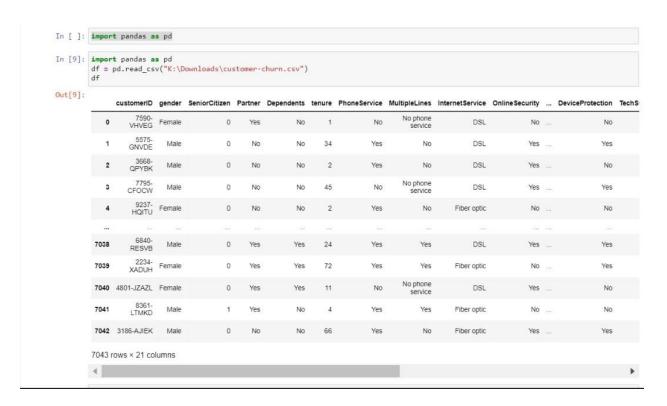
Customer Churn Prediction

To load and preprocess a Customer Churn dataset for analysis, you can follow these general steps using Python and Pandas. Make sure you have a Customer Churn dataset in a suitable format available.

1. **Import Libraries**: Start by importing the necessary Python libraries, including Pandas, to load and preprocess the dataset.

> Import pandas as pd

2. **Load the Customer Churn Dataset**: Load the COVID-19 dataset into a Pandas DataFrame. You can use **pd.read_csv()** for CSV files, but the method may vary depending on the file format.



3. **Data Inspection:** Before preprocessing, inspect the data to understand its structure and identify any potential issues.

```
In [10]: print(df.head())
             customerID gender SeniorCitizen Partner Dependents tenure PhoneService \
             7590-VHVEG Female
5575-GNVDE Male
                                               9
                                                      Yes
                                                                  No
                                                                                         No
                                                                  No
                                                       No
                                                                           34
                                                                                        Yes
             3668-QPYBK
                            Male
             7795-CF0CW
                           Male
                                                       No
                                                                  No
                                                                                         No
          4 9237-HQITU Female
                                                      No
                                                                  No
                                                                                        Yes
                {\it MultipleLines\ InternetService\ OnlineSecurity\ \dots\ DeviceProtection\ } \\
            No phone service
                                            DSL
                                                             No ...
Yes ...
                                            DSL
                            No
                                            DSL
                                                            Yes ...
                                                                                    No
             No phone service
                                                            Yes ...
                                   Fiber optic
                                                             No ...
                                                                                    No
            TechSupport StreamingTV StreamingMovies
                                                              Contract PaperlessBilling \
                     No
                                                   No Month-to-month
                     No
                                  No
                                                   No
                                                              One year
                                                                                      No
                                                   No Month-to-month
                    Yes
                                  No
                                                  No
                                                              One year
                                                                                      No
                                                  No Month-to-month
                     No
                                  No
                                                                                      Yes
                      PaymentMethod MonthlyCharges TotalCharges Churn
Electronic check 29.85 29.85 No
                                                                1889.5
                           Mailed check
                                                             108.15
1840.75
                                                  53.85
                                                                        Yes
             Bank transfer (automatic)
                                                  42.30
                                                                           No
                      Electronic check
                                                  70.70
                                                               151.65
          [5 rows x 21 columns]
In [11]: print(df.isnull().sum())
          customerID
          gender
SeniorCitizen
          Partner
          Dependents
          tenure
          MultipleLines
          InternetService
          OnlineSecurity
          OnlineBackup
          DeviceProtection
          TechSupport
          StreamingTV
          StreamingMovies
Contract
          PaperlessBilling
PaymentMethod
          MonthlyCharges
          TotalCharges
          Churn
          dtype: int64
In [13]: print(df.dtypes)
          customerID
                                 object
          gender
SeniorCitizen
                                 object
                                  int64
          Partner
                                 object
          Dependents
          tenure
PhoneService
                                  int64
                                 object
          MultipleLines
                                 object
```

4. Data Preprocessing:

InternetService

DeviceProtection

OnlineSecurity

OnlineBackup

TechSupport

StreamingTV StreamingMovies

Contract

object

object

object

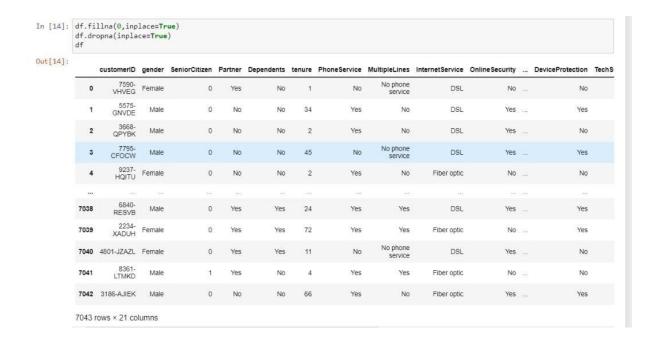
object

object

object

a. Data Cleaning:

Handle missing values by either imputing them or removing rows with missing data



b. Data Transformation: If necessary, transform the data to suit your analysis objectives. For instance, you may want to aggregate data by date or region.



c. Data Filtering:

Filter the data to focus on a specific time frame or specific regions of interest.



5. Save the Preprocessed Data (Optional): If you want to save the preprocessed data for future analysis, you can use Pandas to save it to a new CSV file.

```
In [23]: df.to_csv('preprocessed_customer_churn.csv',index=False)
```

These are the general steps to load and preprocess a Customer Churn dataset using Python and Pandas. Remember that the specific preprocessing steps and operations may vary depending on the structure of your dataset and your analysis objectives