CHRIST (Deemed to be) University Bangalore

Department of Computer Science

MCA171 Python Programming

Lab Exercises VII

**TOJIN VARKEY SIMSON**

**2447253 MCA-B**

Exploring the Relationship between COVID-19 and Economic Growth

Scenario Overview:

You have been provided with two datasets:

(1) COVID-19 Data: Monthly COVID-19 statistics (confirmed cases, deaths, and

recoveries) for five countries: USA, India, Brazil, Germany, and Japan, from

January 2020 to December 2021.

(2) GDP Growth Data: Quarterly GDP growth data for the same countries over

the same period.

(GDP: Gross Domestic Product)

Your task is to analyze these datasets using pandas DataFrame operations to explore

how the pandemic has affected the economic growth of each country.

Question

(1) Load the synthetic\_covid\_data.csv and synthetic\_gdp\_data.csv files into

two separate pandas DataFrames.

(2) Display the first few rows of each dataset to understand the structure.

(3) Ensure that the month column in the COVID-19 dataset is of type datetime64.

(4) Check for any duplicates in both datasets and remove them if necessary.

1

2

(5) Add a new column active\_cases to the COVID-19 dataset, calculated as:

active cases=confirmed cases recovered

(6) Aggregate the COVID-19 dataset by country and quarter to align it with the

GDP dataset. Calculate the total confirmed\_cases, deaths, and active\_cases

for each country and quarter.

(7) Merge the aggregated COVID-19 data with the GDP data on the country and

quarter columns.

(8) Display the resulting DataFrame after the merge.

(9) Identify which country experienced the highest overall GDP growth during

the pandemic.

(10) Determine which country had the highest number of total COVID-19 con-

firmed cases and deaths over the 2-year period.

Write a very short conclusion summarizing your findings, and observations between

COVID-19 cases and economic growth

**CODE:**

**import pandas as pd**

**covid\_data = pd.read\_csv('synthetic\_covid\_data.csv')**

**gdp\_data = pd.read\_csv('synthetic\_gdp\_data.csv')**

**print("COVID-19 Data:")**

**print(covid\_data.head())**

**print("\nGDP Data:")**

**print(gdp\_data.head())**

**covid\_data['month'] = pd.to\_datetime(covid\_data['month'])**

**print("\nDuplicate rows in COVID-19 Data: " + str(covid\_data.duplicated().sum()))**

**print("Duplicate rows in GDP Data: " + str(gdp\_data.duplicated().sum()))**

**covid\_data = covid\_data.drop\_duplicates()**

**gdp\_data = gdp\_data.drop\_duplicates()**

**covid\_data['active\_cases'] = covid\_data['confirmed\_cases'] - covid\_data['recovered']**

**covid\_data['quarter'] = covid\_data['month'].dt.to\_period('Q')**

**quarterly\_covid\_data = covid\_data.groupby(['country', 'quarter']).agg({**

**'confirmed\_cases': 'sum',**

**'deaths': 'sum',**

**'active\_cases': 'sum'**

**}).reset\_index()**

**def convert\_to\_period(quarter\_str):**

**try:**

**return pd.Period(quarter\_str, freq='Q')**

**except Exception as e:**

**print(f"Error converting {quarter\_str} to period: {e}")**

**return None**

**gdp\_data['quarter'] = gdp\_data['quarter'].astype(str).apply(convert\_to\_period)**

**gdp\_data = gdp\_data.dropna(subset=['quarter'])**

**merged\_data = pd.merge(quarterly\_covid\_data, gdp\_data, on=['country', 'quarter'])**

**print("\nMerged Data:")**

**print(merged\_data.head())**

**country\_gdp\_growth = merged\_data.groupby('country')['gdp\_growth'].sum()**

**highest\_gdp\_growth\_country = country\_gdp\_growth.idxmax()**

**print("\nThe country with the highest overall GDP growth is: " + highest\_gdp\_growth\_country)**

**country\_covid\_totals = merged\_data.groupby('country').agg({**

**'confirmed\_cases': 'sum',**

**'deaths': 'sum'**

**})**

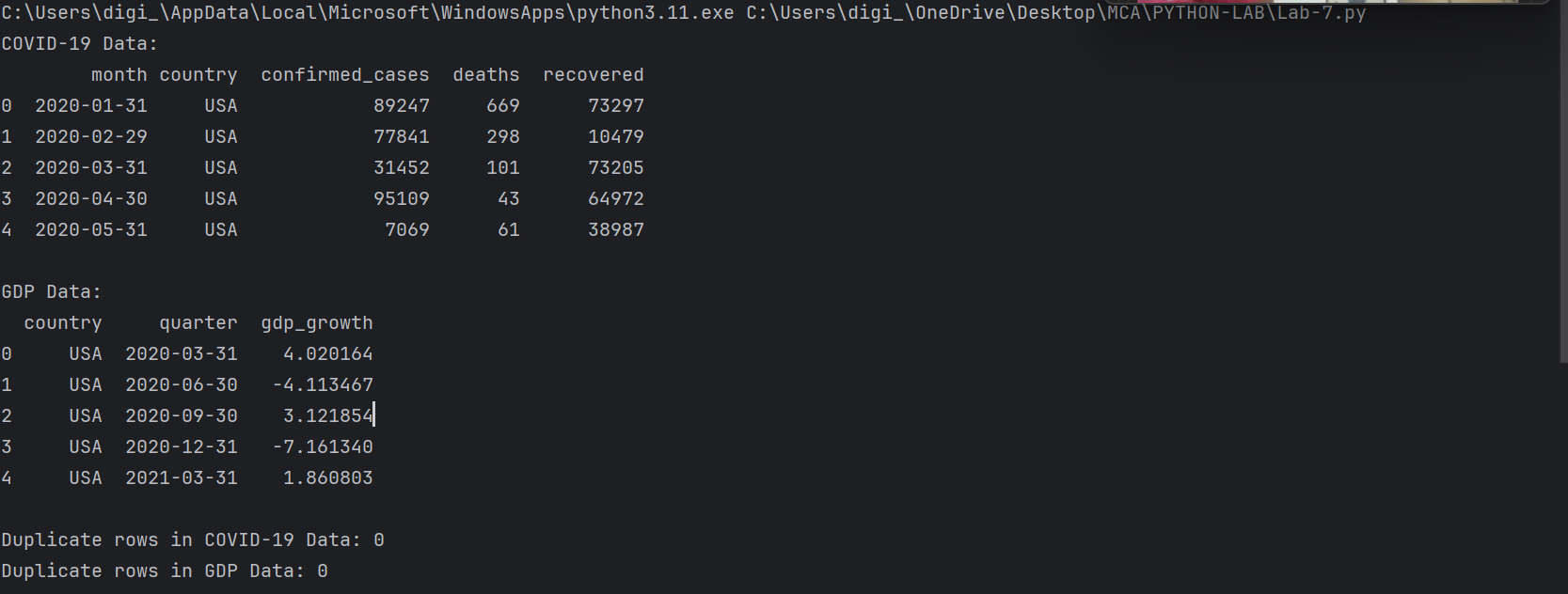
**highest\_cases\_country = country\_covid\_totals['confirmed\_cases'].idxmax()**

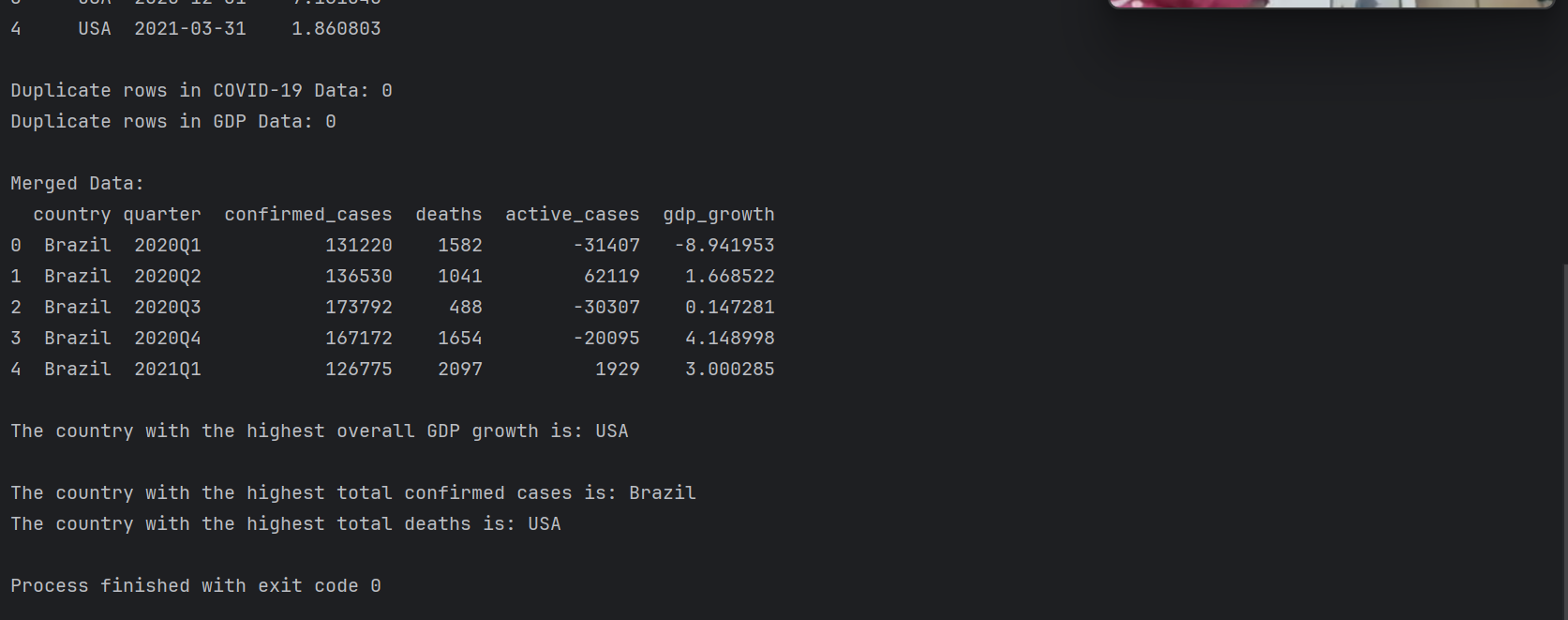
**highest\_deaths\_country = country\_covid\_totals['deaths'].idxmax()**

**print("\nThe country with the highest total confirmed cases is: " + highest\_cases\_country)**

**print("The country with the highest total deaths is: " + highest\_deaths\_country)**

**OUTPUT:**

****

****