COVID-19 Cases Analysis Using IBM Cognos

510421205025

PHASE-3

DEVELOPMENT PART 1

# Introduction

This document provides a structured guide for building a COVID-19 cases analysis using IBM Cognos. The analysis aims to understand the trend of COVID-19 cases and deaths over time and to compare COVID-19 statistics across different regions or countries.

# Step 1: Define Analysis Objectives

Clearly outline the objectives of your COVID-19 analysis. Some example objectives include:

* **Trend Analysis:** Understand the trend of COVID-19 cases and deaths over time.
* **Regional Comparison:** Compare COVID-19 statistics across different regions or countries.
* **Correlation Analysis:** Identify factors that may correlate with the spread of the virus.

**Step 2: Obtain COVID-19 Data**

Acquire a reliable and up-to-date COVID-19 dataset.

# Step 3: Data Cleaning and Processing

Ensure the data is clean, accurate, and well-structured before loading it into IBM Cognos.

* **Handle Missing Values:** Identify and decide how to handle missing values (remove, impute, etc.).
* **Data Types:** Ensure that data types are appropriate (e.g., date fields should be in datetime format).
* **Remove Duplicates:** Eliminate duplicate records from the dataset.
* **Data Transformation:** Perform necessary transformations (aggregations, filtering, create calculated fields).

In this session we have used Jupyter Notebook to clean the data.

Python code for cleaning dataset:

import pandas as pd

# Load the dataset

df = pd.read\_csv('dataset.csv')

# Print the first 5 rows of the dataset

print(df.head())

# Check for missing values

print(df.isnull().sum())

# Drop missing values

df = df.dropna()

# Print the first 5 rows of the dataset after dropping missing values

print(df.head())

# Check for duplicate rows

print(df.duplicated().sum())

# Drop duplicate rows

df = df.drop\_duplicates()

# Print the first 5 rows of the dataset after dropping duplicate rows

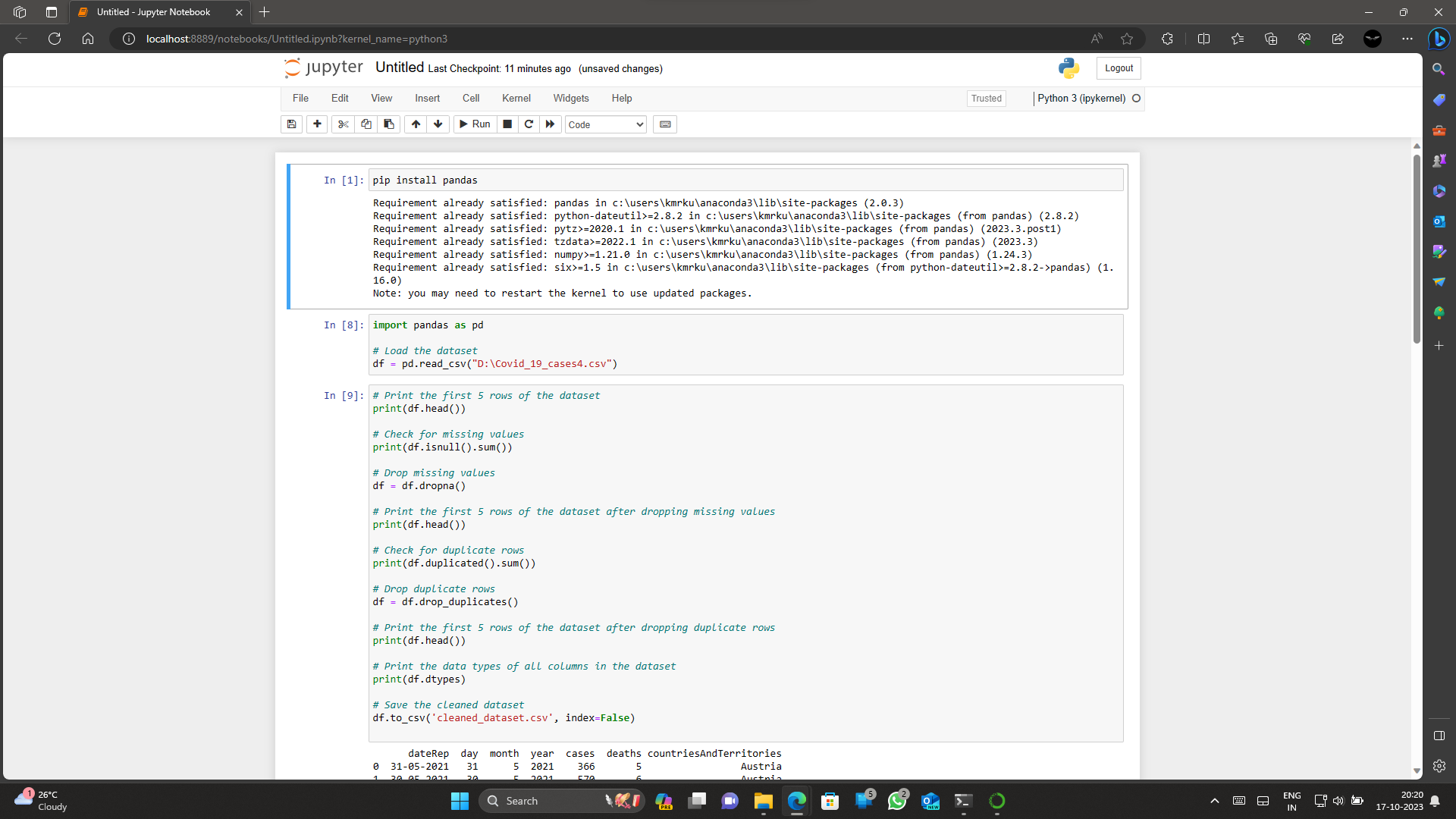
print(df.head())

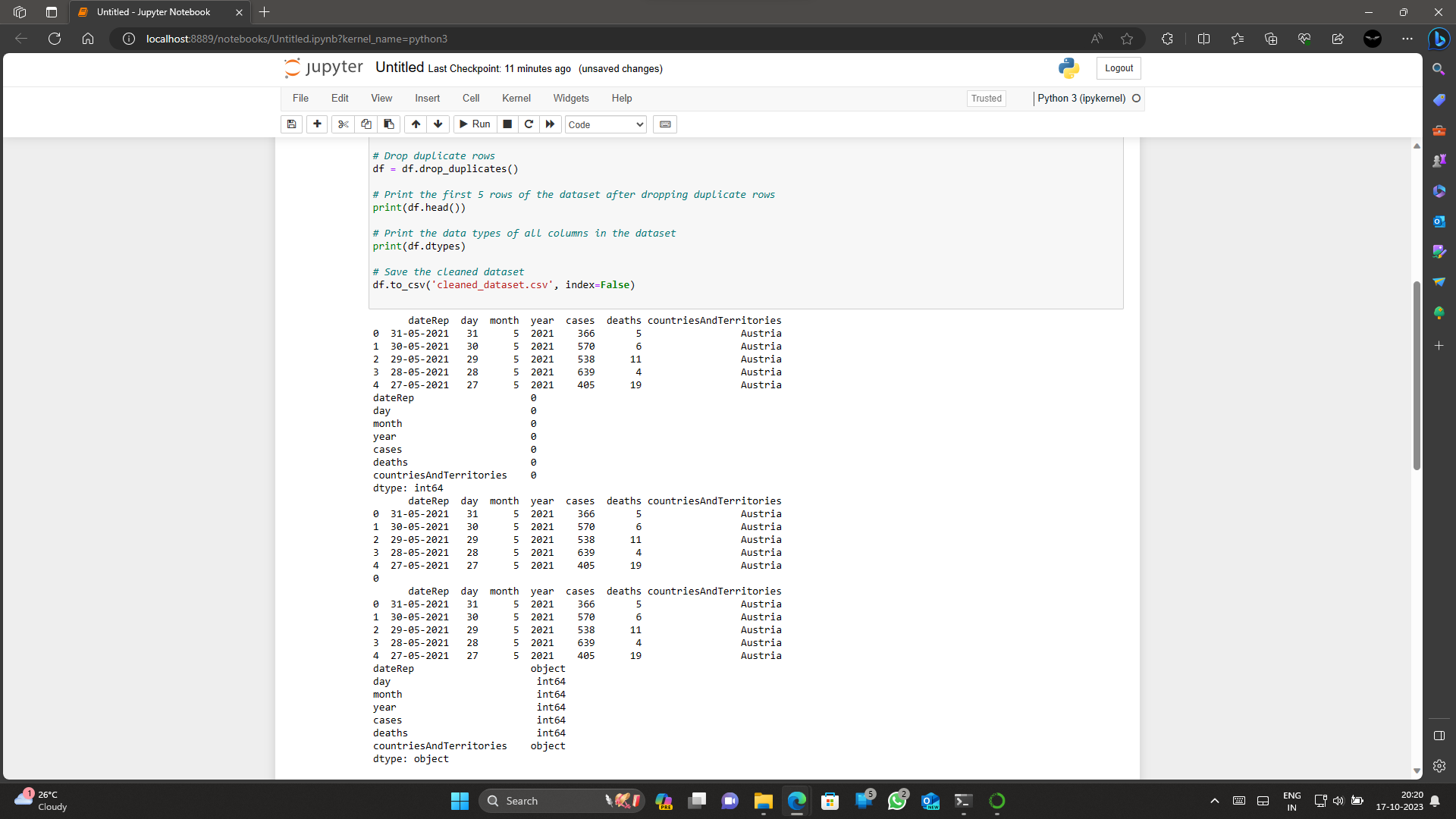
# Print the data types of all columns in the dataset

print(df.dtypes)

# Save the cleaned dataset

df.to\_csv('cleaned\_dataset.csv', index=False)





# Step 4: Load Data into IBM Cognos

* **Open IBM Cognos Analytics:** Log in to IBM Cognos Analytics.
* **Create a Data Module or Connect to Data Source:** Create a new data module. Connect to your data source directly.
* **Import Data:** Import the cleaned COVID-19 data into IBM Cognos.

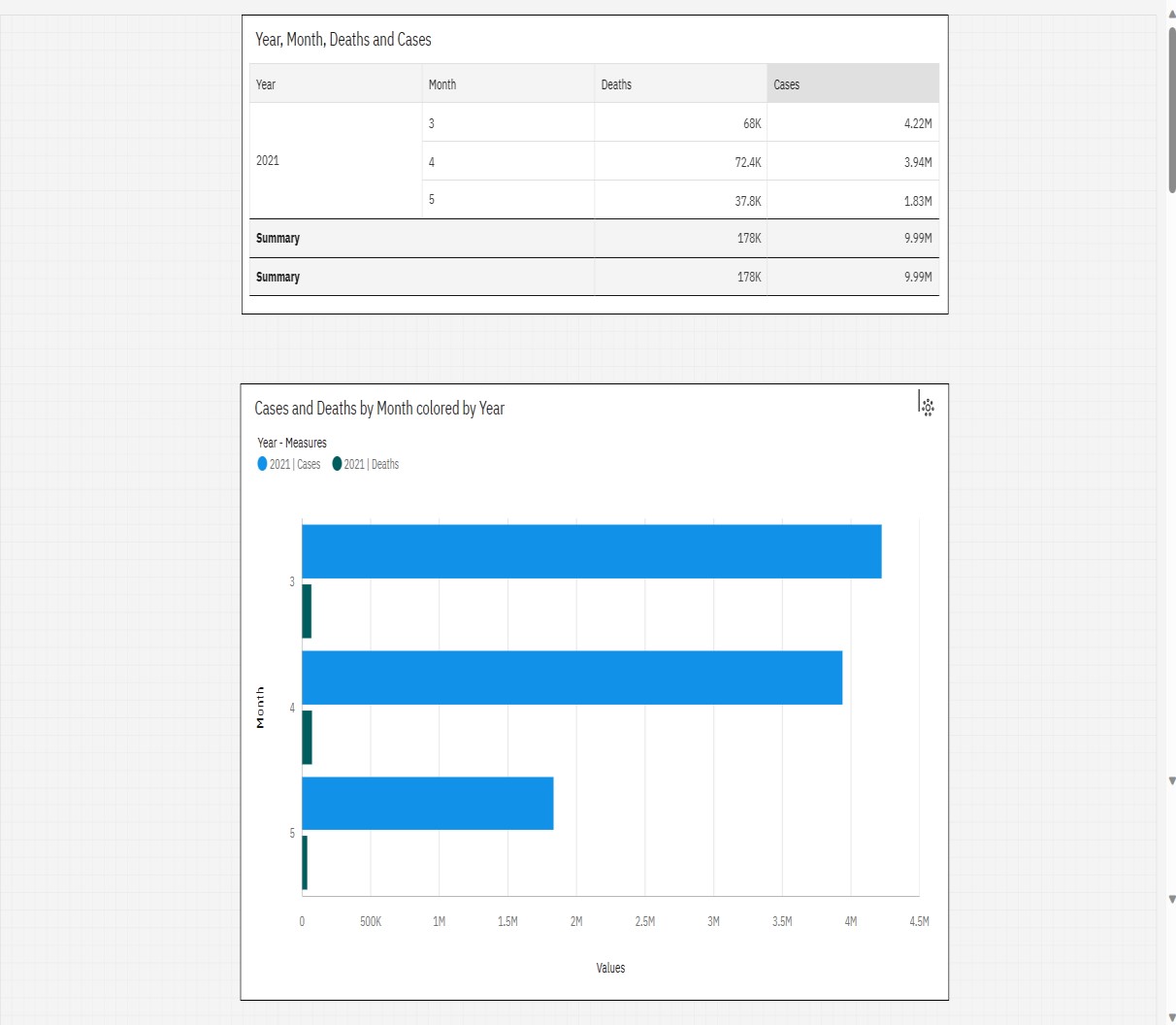
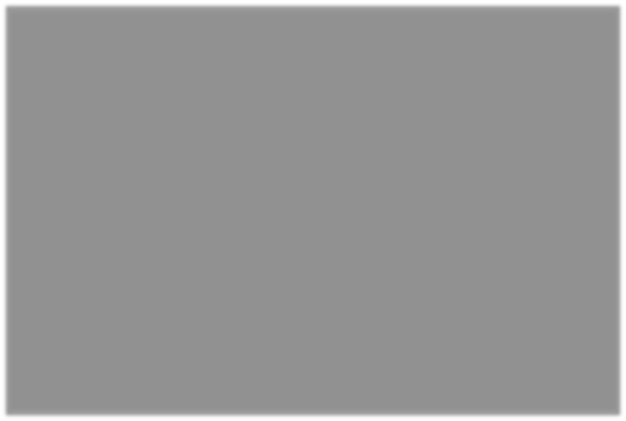
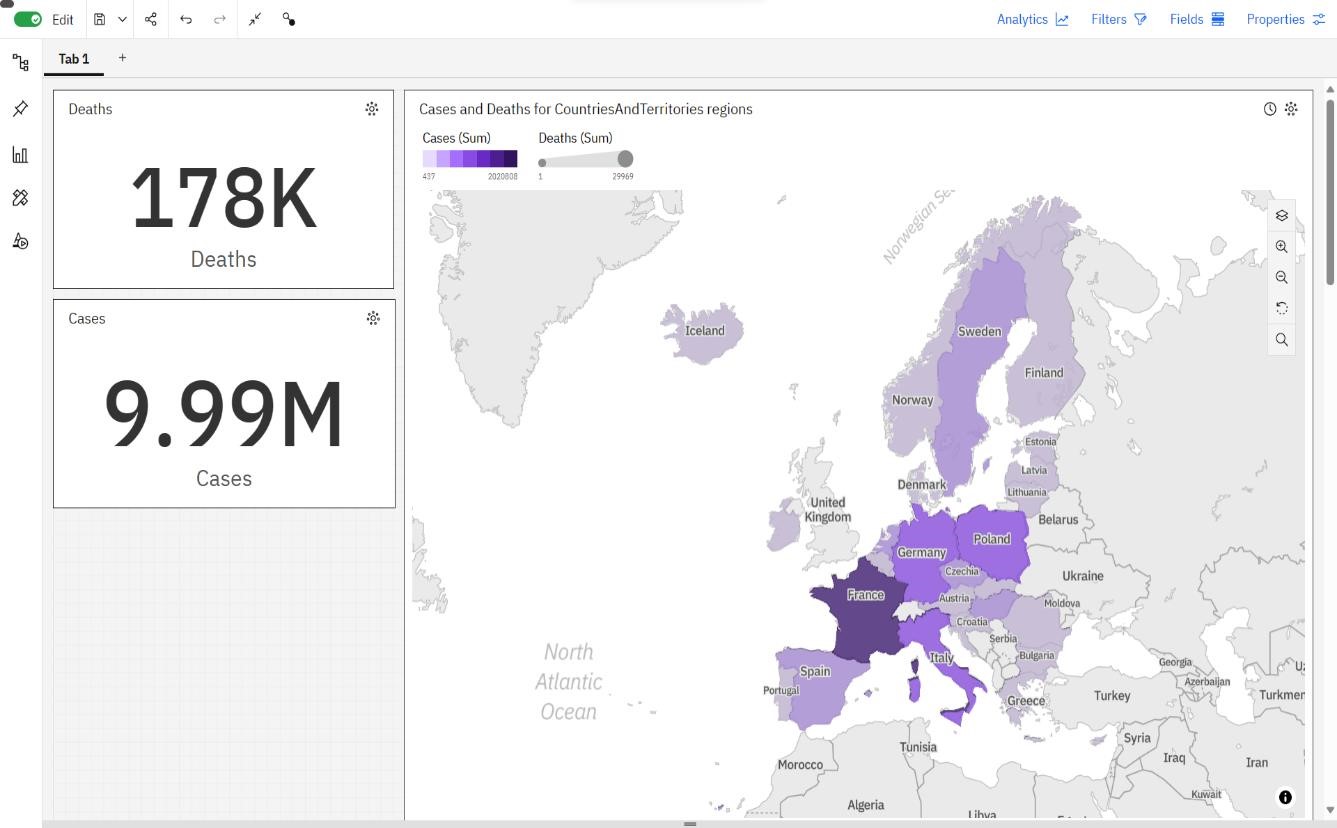
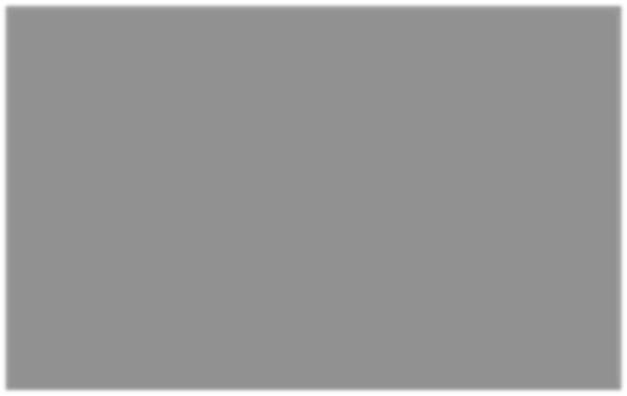
# Step 5: Build Visualizations

* **Create Reports and Dashboards:** Build reports and dashboards in IBM Cognos Analytics.
* **Choose Appropriate Chart Types:** Select chart types based on your analysis objectives (line charts, bar charts, maps, etc.).
* **Use Features like Filtering and Grouping:** Utilize features such as filtering, grouping, and drilling down for deeper insights.

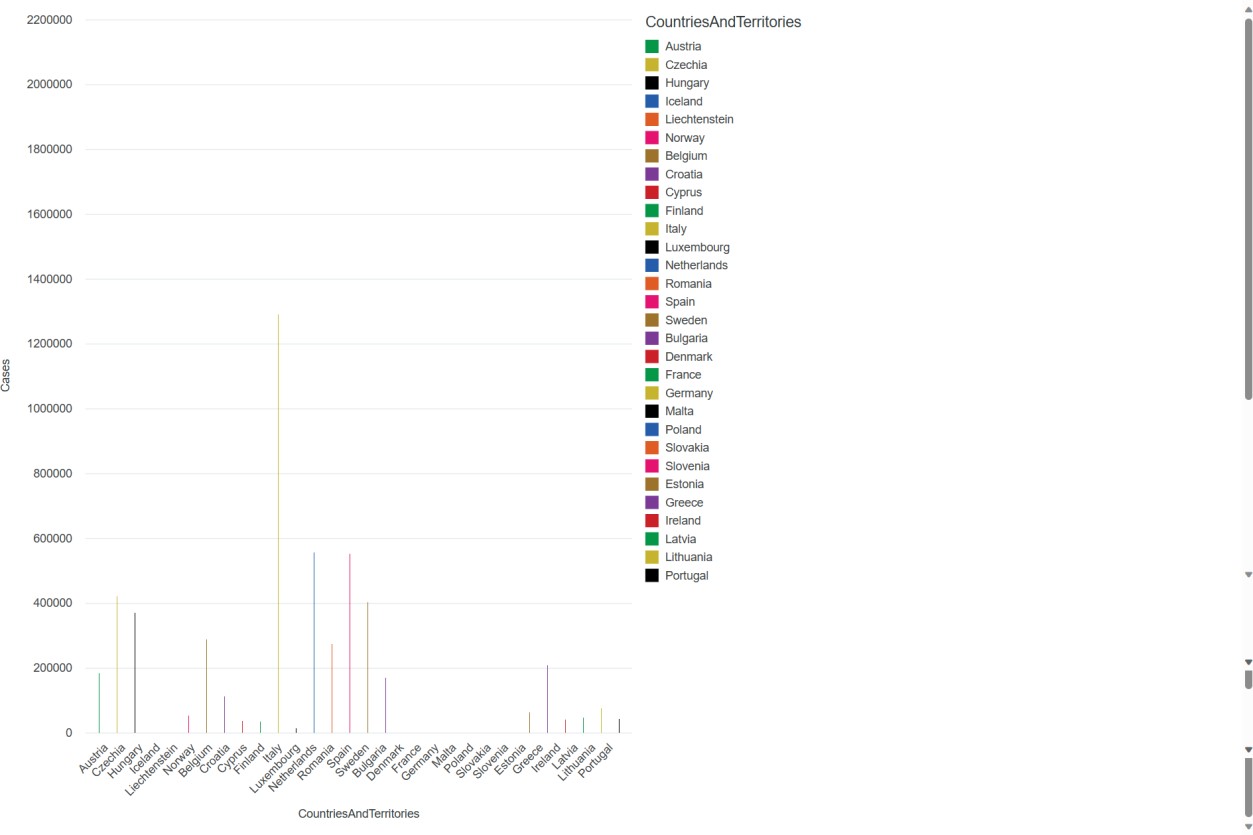
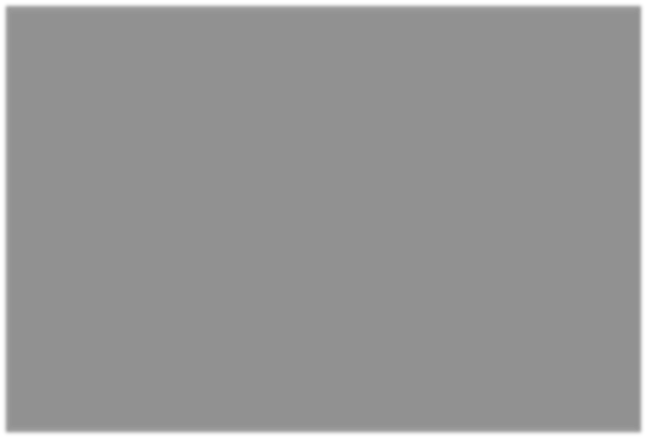
# Step 6: Analyse and Share

* **Conduct Analysis:** Analyse the visualizations to answer your defined objectives.
* **Share Findings:** Share your findings through interactive dashboards or static reports.
* **Schedule Data Refresh:** Consider scheduling or automating the data refresh process for up-to-date insights.

# Data Visualization and Analysis



o Drilling Down For Deeper Insights



## Conclusion

In conclusion, this guide provides a systematic approach to initiate a comprehensive COVID19 cases analysis using IBM Cognos for effective visualization. The outlined steps aim to facilitate the understanding of both temporal trends and regional variations in COVID-19 statistics.