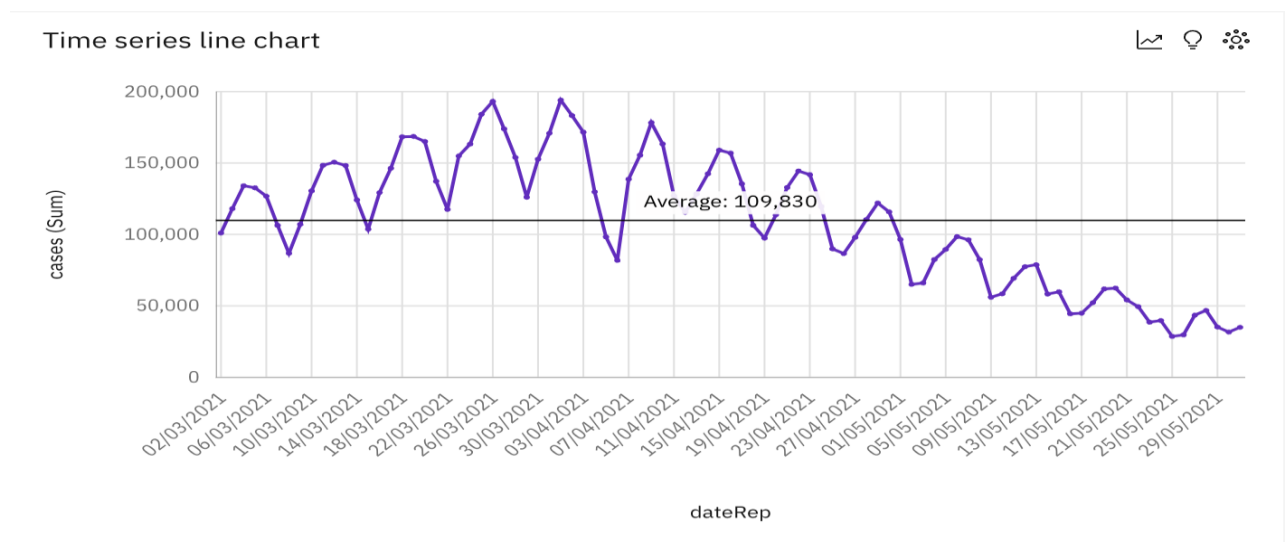


PHASE 4

Development Part-2 (Covid-19 Cases Analysis)

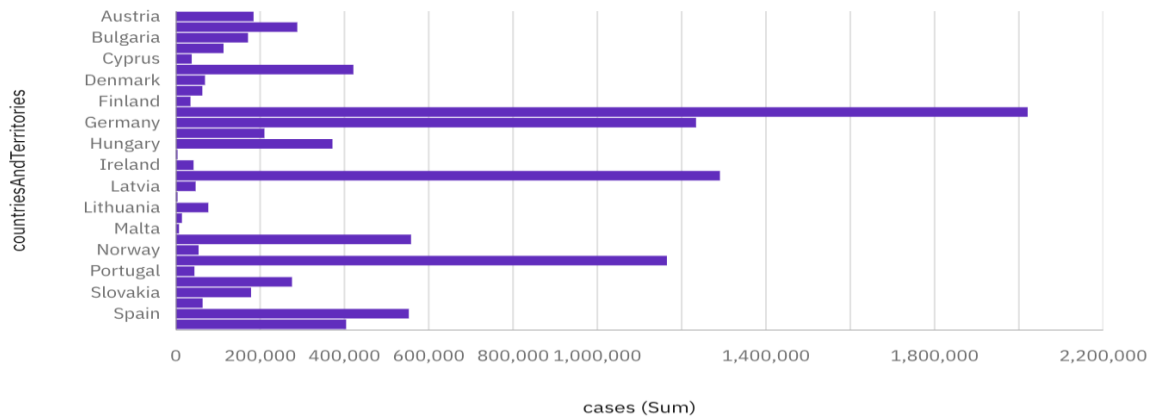
Continue building the analysis by creating visualizations using IBM Cognos and deriving insights from the data.

- **Building Data Models:** In IBM Cognos, create data models that represent the data structure. Define dimensions (e.g., time, geography) and measures (e.g., confirmed cases, deaths) that you want to analyse.
- **Creating Visualizations:** Use the IBM Cognos tools to create visualizations that effectively convey the insights. Some common types of visualizations for COVID-19 analysis include:
 - a. **Time Series Line Charts:** Show trends over time, such as daily new cases or cumulative cases.



- b. **Bar Charts:** Compare data across different categories, like cases by country.

Bar Charts

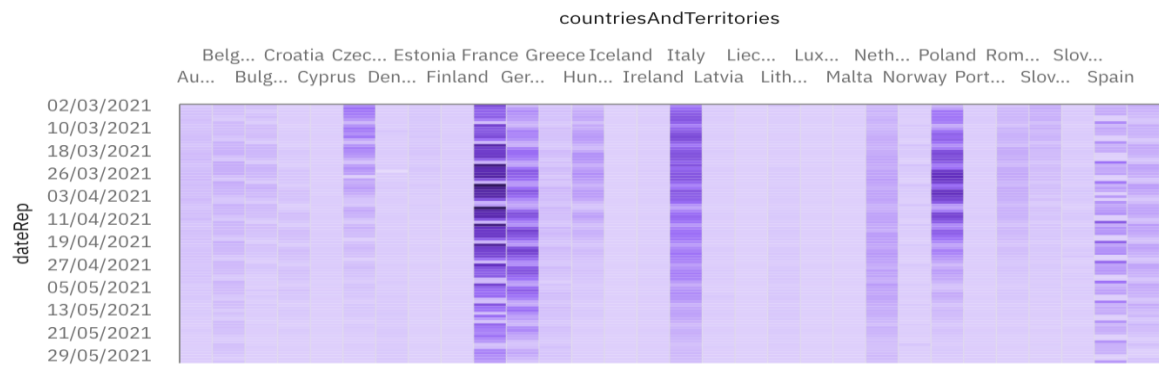


- c. **Heat Maps:** Visualize the spread of the virus geographically.

Heat Maps



cases (Sum)

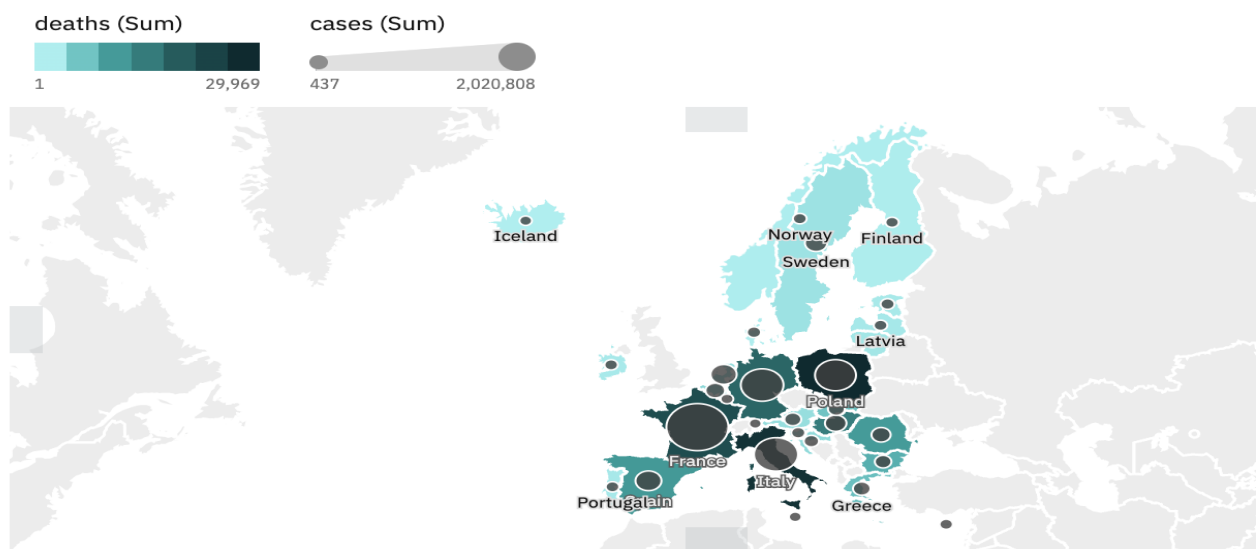


- d. **Pivot Tables:** Display detailed data with filtering and sorting options.

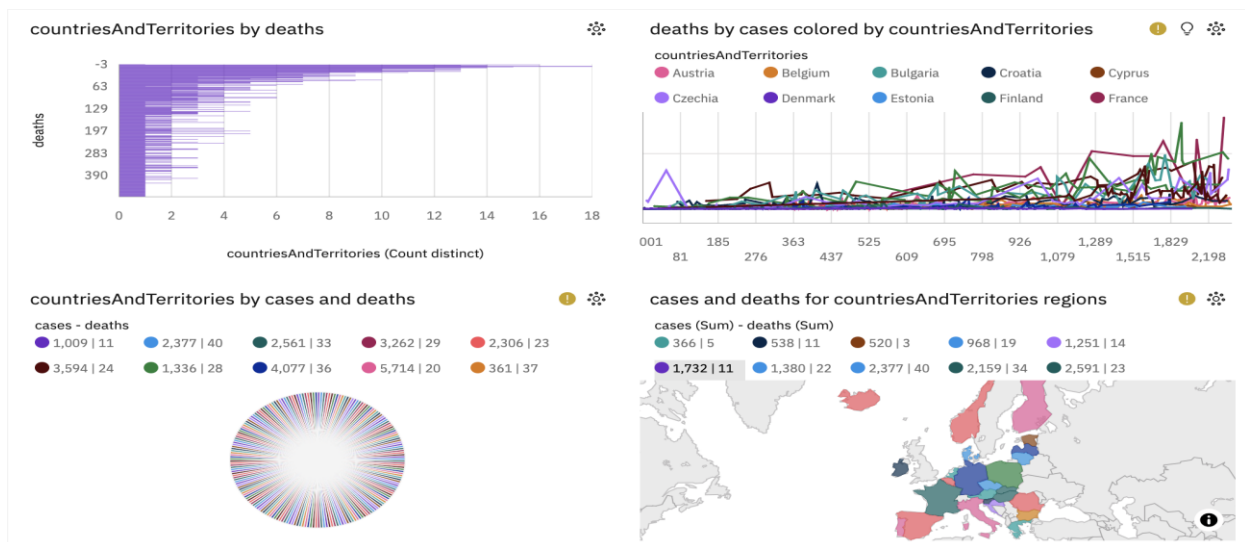
cases		Portugal	Romania	Slovakia	Slovenia	Spain	Sweden	Summary
17/05/2021	68	334	565	386	296	11,061	2,109	44,851
18/05/2021	11	199	392	552	114	3,988	4,646	52,261
19/05/2021	34	386	681	19	497	6,080	3,409	61,844
20/05/2021	37	511	707	512	532	5,733	3,165	62,409
21/05/2021	90	451	618	347	364	4,792	2,322	54,154
22/05/2021	80	559	484	354	331	0	1,909	49,452
23/05/2021	15	523	455	464	383	0	620	38,574
24/05/2021	76	413	307	376	159	11,067	1,195	39,728
25/05/2021	59	241	230	412	83	5,359	2,740	28,616
26/05/2021	06	375	381	132	325	5,007	1,419	29,670
27/05/2021	63	594	379	362	388	5,290	90	43,493
28/05/2021	02	572	307	335	362	5,482	297	46,796
29/05/2021	43	598	321	134	292	0	285	35,162
30/05/2021	76	609	265	355	353	0	1,949	31,680
31/05/2021	79	445	158	200	138	9,732	1,900	34,985
Summary	64	44,096	275,590	178,475	63,550	552,723	404,019	9,994,560

- **e. Geospatial Maps:** Represent data on a map to show regional variations.

Geospatial Maps



- **Creating Dashboards:** Build interactive dashboards that bring together multiple visualizations. This allows users to explore the data and gain insights in a user-friendly way.



- **Deriving Insights:** Use your visualizations and dashboards to answer specific questions, such as:

The summed values of **deaths** range from **-3** to **956**.

↓ ↑ ★

The total number of results for **countriesAndTerritories**, across all **deaths**, is over **2500**.

↓ ↑ ★

It is projected that by **2021-06-19**, **Germany** will exceed **Poland** in **deaths** by **45**.

↓ ↑ ★

From **2021-04-24** to **2021-04-25**, **1's countriesAndTerritories** dropped by **83%**.

↓ ↑ ★

cases 27890 has the highest total **deaths** due to **countriesAndTerritories Poland**.

↓ ↑ ★

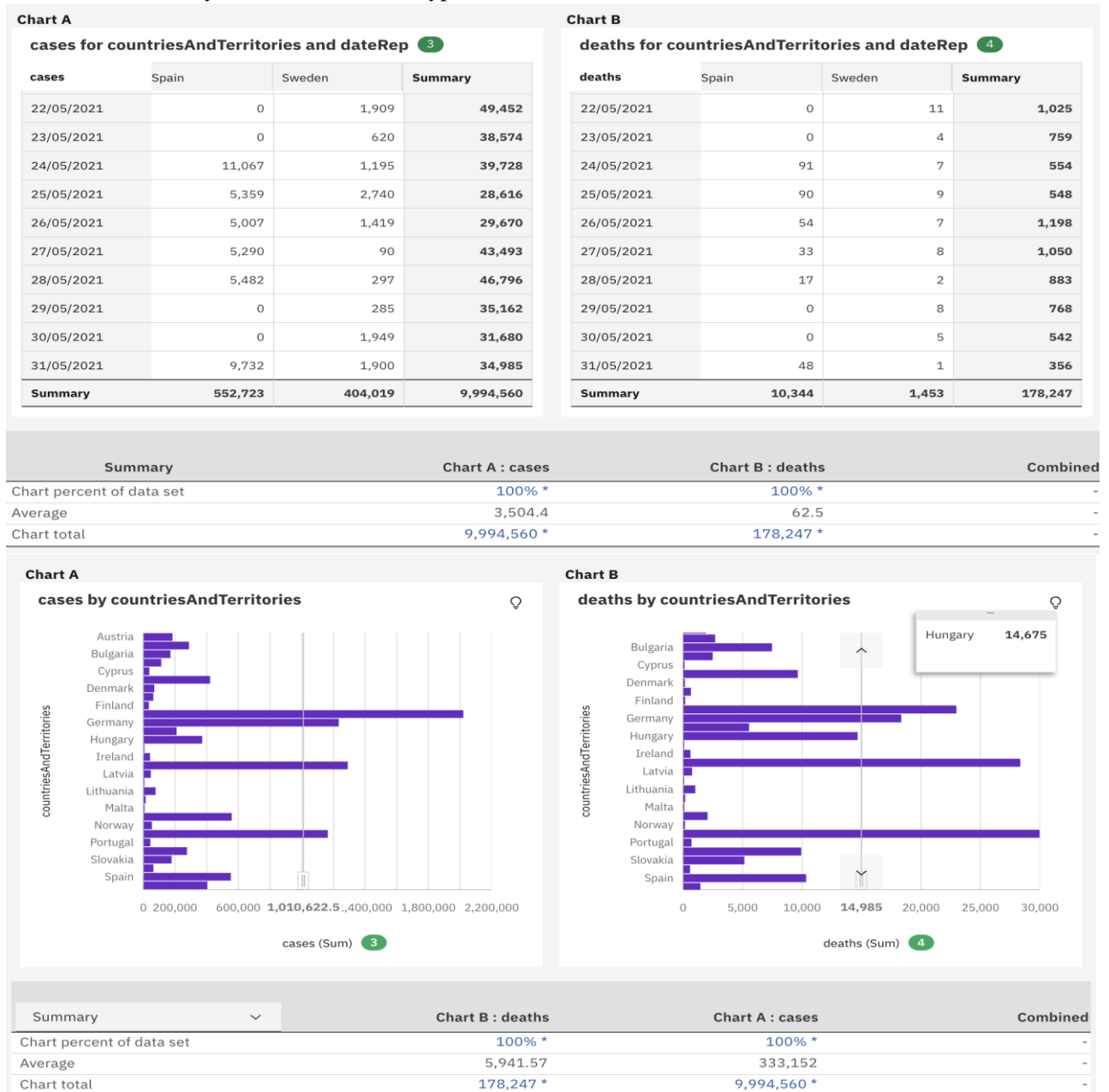
It is projected that by **2021-06-19**, **0** will exceed **1** in **countriesAndTerritories** by **5.04**.

↓ ↑ ★

Create charts and graphs in IBM Cognos to visualize and compare the mean values and standard deviations of COVID-19 cases and associated deaths.

- **Create a New Report:** In IBM Cognos, create a new report that will serve as the canvas for your charts and graphs.

- **Insert a Table:** To display mean values and standard deviations, you can start with a simple table. Drag and drop the measures "COVID-19 Cases" and "COVID-19 Deaths" into the table's data items.
- **Create Chart Visualizations:** To visualize and compare the mean values and standard deviations, you can use different types of charts:



Analyse the visualizations to identify trends, variations, and potential correlations between cases and deaths.

- **Import Libraries:** Start by importing the necessary libraries:
- **Load and Prepare Data:** Load your COVID-19 data into a Pandas DataFrame and ensure its properly cleaned and formatted.

```

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
data = pd.read_csv(r"/content/Covid_19_cases.csv")
print(data.head())

```

	dateRep	day	month	year	cases	deaths	countriesAndTerritories
0	31-05-2021	31	5	2021	366	5	Austria
1	30-05-2021	30	5	2021	570	6	Austria
2	29-05-2021	29	5	2021	538	11	Austria
3	28-05-2021	28	5	2021	639	4	Austria
4	27-05-2021	27	5	2021	405	19	Austria

- **Visualize Trends:** Create visualizations to identify trends and variations in COVID-19 cases and deaths over time. You can use line plots for this purpose.

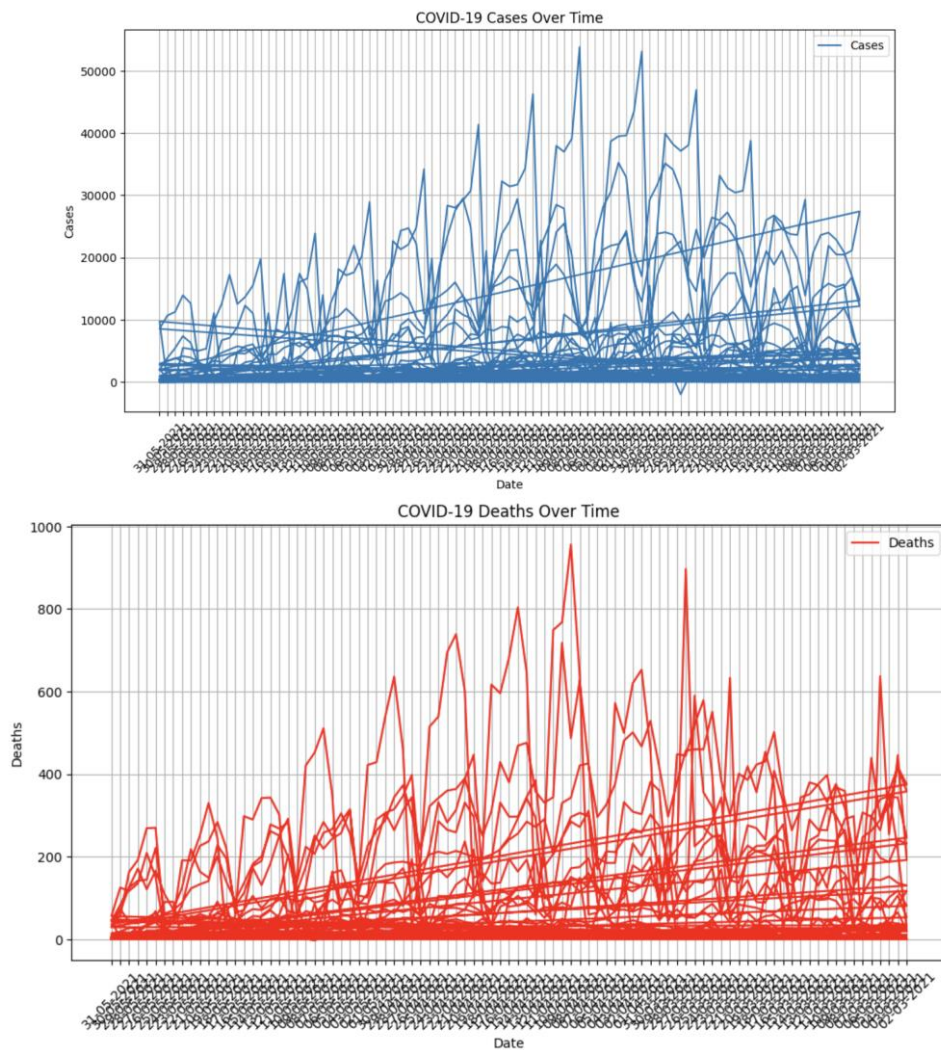
```

plt.figure(figsize=(12, 6))
plt.plot(data['dateRep'], data['cases'], label='Cases')
plt.title('COVID-19 Cases Over Time')
plt.xlabel('Date')
plt.ylabel('Cases')
plt.xticks(rotation=45)
plt.legend()
plt.grid(True)
plt.show()

plt.figure(figsize=(12, 6))
plt.plot(data['dateRep'], data['deaths'], label='Deaths', color='red')
plt.title('COVID-19 Deaths Over Time')
plt.xlabel('Date')
plt.ylabel('Deaths')
plt.xticks(rotation=45)
plt.legend()
plt.grid(True)
plt.show()

```

OUTPUT:



- **Identify Trends and Variations:** Examine the line plots to identify trends and variations in COVID-19 cases and deaths over time. Look for patterns such as spikes, plateaus, or significant increases or decreases in the data.
- **Correlation Analysis:** To identify potential correlations between COVID-19 cases and deaths, calculate the correlation coefficient between these two variables. A positive correlation suggests that as cases increase, deaths tend to increase, while a negative correlation indicates an inverse relationship.

```

correlation_coefficient = data['cases'].corr(data['deaths'])
print(f"Correlation Coefficient: {correlation_coefficient}")

Correlation Coefficient: 0.766308878657635

```

A correlation coefficient close to 1 indicates a strong positive correlation, close to -1 indicates a strong negative correlation, and close to 0 indicates a weak correlation.

This shows the development part-2 for the COVID-19 Cases Analysis were successfully implemented with the following steps which is followed to execute the Data Analytics with

IBM Cognos and Python with the given procedure.