

# COVID -19 ANALAYSIS

## INNOVATION

<b>Date</b>	<b>10-10-2023</b>
<b>Team ID</b>	<b>720</b>
<b>Project Name</b>	<b>Covid-19 Analysis</b>

### Table of Contents

1	Introduction
2	Problem Statement
3	Data discovery
4	Data Preparation
5	Data modelling
6	Dashboarding Visualizing Insights
7	Communication of Results
8	Conclusion

## 1. Introduction

COVID-19, also known as Coronavirus Disease 2019, is a highly contagious respiratory illness caused by the novel coronavirus SARS-CoV-2. It emerged in late 2019 in Wuhan, China, and swiftly evolved into a global pandemic. The virus spreads primarily through respiratory droplets, leading to a wide range of symptoms, from mild respiratory issues to severe pneumonia and, in some cases, fatalities. The pandemic has had a profound impact on public health, economies, and daily life worldwide, prompting governments and healthcare systems to implement extensive measures such as lockdowns, mask mandates, and vaccination campaigns to control its spread. The ongoing battle against COVID-19 has underscored the importance of scientific research, international cooperation, and public health measures to combat emerging infectious diseases.

## 2. Problem Statement

The project involves analyzing COVID-19 cases and deaths data using IBM Cognos. The objective is to compare and contrast the mean values and standard deviations of cases and associated deaths per day and by country in the EU/EEA. This project encompasses defining analysis objectives, collecting COVID-19 data, designing relevant visualizations in IBM Cognos, and deriving insights from the data.

## **3. Data discovery**

### **3.1 Data Import and Initial Inspection:**

Begin by importing the COVID-19 dataset into your data analysis environment. Inspect the dataset's structure, including the number of rows and columns, data types, and the first few rows to get a sense of the data.

### **3.2 Data Exploration:**

Calculate summary statistics for cases and deaths to understand the overall scale of the pandemic.

Explore the distribution of cases and deaths by creating histograms, box plots, or density plots.

### **3.3 Temporal Analysis:**

Analyze how COVID-19 cases and deaths have evolved over time. Consider the day, month, and year columns.

Calculate daily, monthly, and yearly totals for cases and deaths.

Visualize trends using line plots or time series charts to identify peaks and patterns.

### **3.4 Geographical Analysis:**

Explore how COVID-19 cases and deaths are distributed across different countries and territories.

Calculate country-wise totals for cases and deaths.

Create maps, bar charts, or heatmaps to visualize the geographical spread.

## **4. Data preparation**

Data preparation for COVID-19 analysis involves cleaning, structuring, and organizing the dataset to ensure that it is ready for analysis. Below are the key steps involved in data preparation for COVID-19 data:

### **4.1 Data Collection:**

Obtain COVID-19 data from reliable sources, such as health agencies, government repositories, or research organizations. Ensure that the data is up-to-date and comprehensive.

### **4.2 Data Import:**

Import the raw data into your chosen data analysis tool or programming environment (e.g., Python, R, Excel).

### **4.3 Data Inspection:**

Examine the dataset's structure by checking the number of rows and columns, data types, and the first few rows of data to gain an initial understanding.

### **4.4 Handling Missing Values:**

Identify and handle missing values appropriately. Options include imputation (replacing missing values with estimated values), removal of rows or columns with missing data, or considering the missingness pattern.

#### **4.5 Data Cleaning:**

Address data inconsistencies and errors. This may involve correcting data entry mistakes, resolving duplicates, and ensuring uniform formats for dates and locations.

#### **4.6 Data Transformation:**

Convert data types as needed (e.g., date columns to date objects) and create new variables if necessary (e.g., calculating daily cases from cumulative data).

#### **4.7 Outlier Handling:**

Identify and handle outliers, which are extreme values that can skew analysis results. Options include capping values, transforming data, or conducting separate analyses with and without outliers.

## **5. Data modelling**

Modeling COVID-19 is a complex task that requires collaboration among epidemiologists, data scientists, and domain experts. It plays a critical role in understanding the pandemic's dynamics, guiding public health measures, and making informed decisions. However, models are simplifications of reality and should be used in conjunction with real-world data and expert judgment to inform decision-making.

## **6. Dashboarding Visualizing Insights**

### **6.1 Define Dashboard Objectives:**

Determine the main objectives of your COVID-19 dashboard. Are you aiming to track cases, deaths, vaccinations, or other key metrics? What insights do you want to provide to your audience?

### **6.2 Choose Visualization Tools:**

Select a suitable visualization tool for creating your dashboard. Common choices include Tableau, Power BI, Python libraries like Matplotlib and Plotly, R with Shiny, or even custom web development using JavaScript libraries like D3.js.

### **6.3 Data Integration:**

Integrate clean and up-to-date COVID-19 data into your chosen visualization tool. Ensure that data sources are connected and regularly updated.

### **6.4 Design Layout:**

Create a user-friendly and visually appealing layout for your dashboard. Consider the needs of different users, such as public health officials, policymakers, or the general public.

### **6.5 Interactive Elements:**

Incorporate interactive features, such as dropdown menus, filters, or sliders, to allow users to explore the data and customize their views.

### **6.6 Choose Visualization Types:**

Select appropriate chart types and visualizations for presenting COVID-19 insights. Examples include line charts for time series, bar charts for comparisons, maps for geographical data, and heatmaps for intensity visualization.

### **6.7 Time Series Analysis:**

Use time series charts to illustrate trends over time, such as daily or weekly changes in COVID-19 cases, deaths, or vaccination rates.

### **6.8 Geographical Mapping:**

If relevant, include maps to display the geographic distribution of COVID-19 cases, vaccination coverage, or other location-based data. Use color coding or symbols to represent data points.

### **6.9 Key Metrics and KPIs:**

Highlight key metrics and key performance indicators (KPIs) relevant to the COVID-19 situation. Display daily case counts, test positivity rates, vaccination percentages, or other critical data.

## **7. Communication of Results**

### **7.1 Finalize Report Layout:**

Review and finalize the layout of your report. Ensure that all elements, such as tables, charts, and text, are properly arranged and formatted.

### **7.2 Formatting:**

Apply formatting options to enhance the report's appearance. You can format fonts, colors, borders, and backgrounds to make the report visually appealing and easy to read.

### **7.3 Report Styling:**

Apply a consistent styling theme to your report to maintain a unified look and feel. Cognos Analytics provides predefined styles that you can choose from or customize.

### **7.4 Testing and Validation:**

Test your report to ensure that all data elements are displayed correctly. Verify that data calculations, aggregations, and filters are working as expected.

### **7.5 Interactivity:**

interactive prompts. This allows users to explore the data further.

### **7.6 Save and Publish:**

Save your report in Cognos Analytics and publish it to the appropriate location, such as a shared folder, report server, or email distribution list.

### **7.7 Security and Permissions:**

Set up security and permissions for the report to control who can access, view, or modify it. Define user roles and access levels as needed.

### **7.8 Documentation:**

Document the report creation process, including data sources, calculations, and any customizations. This documentation helps with report maintenance and troubleshooting.

## **8. Conclusion**

Summarize your key findings and insights from the data discovery process.

Highlight any notable trends or observations in COVID-19 cases and deaths.

Remember that the specific analysis you conduct will depend on your goals and research questions. This approach provides a foundation for exploring and understanding the COVID-19 dataset, ultimately helping you gain insights into the pandemic's trends and patterns

