

Project 7: COVID-19 using Cognos

Phase 1: Problem Definition and Design Thinking

Project Definition: 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.

Design Thinking:

- 1. Analysis Objectives:** Define the specific objectives of analyzing COVID-19 cases and deaths data, such as comparing mean values and standard deviations.
- 2. Data Collection:** Obtain the provided data file containing COVID-19 cases and deaths information per day and by country in the EU/EEA.
- 3. Visualization Strategy:** Plan how to visualize the mean values and standard deviations using IBM Cognos to create informative charts and graphs.
- 4. Insights Generation:** Identify potential insights from the comparison of mean values and standard deviation of cases and deaths.

Analysis Objectives:

- **Analysis of COVID-19 pandemic on 12 most affected countries including China.**
- **Study of the performance curves of various countries in fighting this pandemic.**
- **Analyses the effectiveness of lockdown and to know the further implementations that can be done accordingly.**
- **Learning the right steps to contain the infection by not repeating the mistakes of other countries.**

Data Collection:

- **Dataset consist of weekly confirmed cases and weekly cumulative confirmed cases for 35 weeks.**

- Then the distribution of the data was examined using the most up-to-date Covid-19 weekly case data and its parameters were obtained according to the statistical distributions.

Visualization:

1.Data Preparation:

- Ensure your data source is properly formatted and loaded into IBM Cognos.

2.Select the Data:

- Identify the dataset or data source that contains the relevant data for which you want to calculate mean values and standard deviations.

3.Create Queries:

- Build a query or report that selects the necessary fields for your analysis. You will typically need a numeric measure to calculate means and standard deviations and one or more categorical dimensions for grouping.

4.Aggregate Data:

- In your query, use aggregation functions to calculate mean values and standard deviations. For mean, use the AVG() function, and for standard deviation, use the STDEV() function.

5.Choose Visualization Types:

- Decide which types of charts or graphs will best represent your data. Common choices include:
- Bar Charts: Useful for comparing means across categories.
- Line Charts: Useful for showing trends in mean values over time.
- Error Bars: Can be added to bar or line charts to represent standard deviations.
- Create Charts and Graphs:
- Use the chart creation tools in IBM Cognos to create visualizations based on your data and selected chart types. Ensure that the mean values and standard deviations are displayed clearly.

6.Customize Visualizations:

- Customize the appearance of your charts and graphs by adjusting colors, labels, legends, and axes to make them more informative and visually appealing.

Insights Generation:

1.Variability in Cases and Deaths:

- A higher standard deviation for cases compared to deaths indicates greater variability in the number of reported cases. This suggests that the number of cases fluctuates significantly over time or across regions.

2.Central Tendency:

- If the mean number of cases is significantly higher than the mean number of deaths, it suggests that, on average, more people are getting infected than succumbing to the disease. This can be an indicator of the disease's severity.

3.Outliers:

- Large standard deviations may be indicative of outliers in the data. Outliers could represent significant spikes in cases or deaths that warrant further investigation.

4.Trends Over Time:

- Comparing mean values and standard deviations over time can reveal trends. For example, if the mean cases are decreasing while the mean deaths are stable or decreasing, it may indicate improved healthcare and management of cases.

5.Regional Variations:

- If you have data for different regions or locations, comparing means and standard deviations can highlight regional disparities. Higher standard deviations in certain regions may signify outbreaks or variations in healthcare quality.