**PROJECT TITLE**

**Analyzing COVID-19 Cases and Deaths Data using IBM Cognos**

**TEAM MEMBERS**

**YUVARAJ T 2021115124**

**ABINAYA S 2021115004**

**AARTHI R 2021115001**

**TAMIZHARASAN D 2021115325**

**HARISH M 2021115326**

**PHASE 2**

**INTRODUCTION:**

The project focuses on analyzing and understanding the dynamics of COVID-19 cases. It aims to leverage data analysis techniques to gain insights into the spread and impact of the virus.

**DATA SOURCES:**

The project relies on data from various sources, including official health organizations, government databases, and publicly available datasets.

**MACHINE LEARNING ALGORITHMS:**

The following machine learning algorithms are utilized in this project

* Logistic Regression: Used for binary classification tasks, such as predicting the likelihood of COVID-19 infection.
* Random Forest: Applied for both classification (e.g., identifying high-risk areas) and regression (e.g., predicting case counts) tasks.
* Time Series Analysis: Techniques like ARIMA and Exponential Smoothing are used to model and forecast COVID-19 case trends over time.
* SIR Models: Epidemiological models like Susceptible-Infectious-Recovered (SIR) are employed to simulate the virus's spread.

**DATA VISUALIZATION:**

Data visualization tools and libraries, such as Matplotlib and Tableau, are used to create informative charts, graphs, and maps to communicate insights effectively.

**SENTIMENT ANALYSIS:**

Natural language processing techniques are applied to analyse public sentiment, helping to understand the social impact and public response to the pandemic.

**CONCLUSION:**

This project provides a data-driven approach to better understand the COVID-19 pandemic. By utilizing various machine learning algorithms, time series analysis, and sentiment analysis, it offers valuable insights for health authorities, policymakers, and the public. These insights can aid in making informed decisions, implementing effective measures, and responding to the pandemic's challenges.