CS661 - BIG DATA VISUAL ANALYTICS
PROJECT PROPOSAL REPORT
GROUP:23
COVID - 19: Trends and Impact Analysis

1. Introduction

The COVID-19 pandemic has significantly impacted global health, economies, and daily life. The large-scale availability of pandemic-related data allows for insightful visual analytics, helping researchers, policymakers, and the general public understand trends and patterns.

In this project, we aim to leverage the Our World in Data COVID-19 Dataset to create a comprehensive web-based dashboard. This dashboard will enable users to explore COVID-19 trends through interactive visualizations, gaining insights into infection rates, vaccination progress, and mortality trends worldwide.

2. Data Sources

Our project will use the following data sources:

- **Primary Source**: Our World in Data COVID-19 Dataset
- **URL**: OWID COVID-19 API

Data Description

The dataset includes country-wise and day-wise records of new cases, total cases, and deaths. It also contains vaccination-related data, such as new vaccinations, total vaccinations, and the vaccine manufacturers used, all categorized by country and date. Additionally, the dataset includes information on COVID-19 testing, hospitalizations, as well as Google Mobility data.

3. Specific Tasks

For this project, we will perform the following main tasks:

- **Data Aggregation**: Retrieve COVID-19 data from OWID and preprocess it for visualization.

- **Data Cleaning and Preprocessing**: Handle missing values, standardize formats, and ensure consistency in data for analysis.
- **Data Visualization**: Develop interactive visualizations to explore pandemic trends, vaccination impacts, and testing effectiveness.
- **User Interaction**: Create an intuitive dashboard to allow users to explore different aspects of the COVID-19 data.

4. Overall Solution

Our system will be a web-based visual analytics platform featuring interactive graphs and charts to present COVID-19 trends. The platform will be developed using modern web technologies to ensure accessibility and user-friendliness.

Visualization Tasks

The following tasks will be carried out in this project. Given that the dataset is structured by country and day, we can create visualizations at various temporal resolutions, including daily, monthly, and yearly.

- **Analysis of Disease Spread**:This task focuses on visualizing the spread of the disease across different countries and over various time periods, including days, months, and years. The goal is to identify disease hotspots and observe how they evolved over time. To achieve this, we will primarily utilize reported data on new and total cases.
- **Mortality Impact Analysis**:This task focuses on assessing the impact of the disease on human lives by analyzing and visualizing reported deaths over time. Additionally, we will consider incorporating the excess mortality rate to gain a more comprehensive understanding of the disease's true impact.
- **Analysis of Vaccination Data**: This task involves examining vaccination trends across various countries over time. The objective is to visualize vaccination distribution, identify patterns, and compare vaccination rates among different nations.
- **Testing Impact Analysis**: This task focuses on assessing the impact of COVID-19 testing by analyzing and visualizing reported testing data. We will explore potential relationships between testing rates, new cases, and deaths to understand how testing efforts influenced the detection, spread, and severity of the disease.
- **Google Mobility Analysis**: During the pandemic, Google released Community Mobility Reports, which showed how people's movements changed in response to lockdowns, social distancing measures, and reopening policies. This task involves analyzing and visualizing the same to assess their impact on the spread of the disease.
- **Interactive COVID-19 Snapshot**: This task allows users to select a specific date and view the COVID-19 data for that day. The visualization will present key statistics like new cases, deaths, vaccinations, and tests conducted on the chosen date. This feature helps users analyze daily fluctuations and compare country-wise responses on any given date.

- **Analyzing covid 19 impacts in india**: The task involves analyzing COVID-19's impact in India using statewise data on cases, testing, and vaccination. It includes assessing trends in infections, testing efficiency, and vaccination coverage to understand regional disparities and pandemic control effectiveness.

5. Tech Stack

- **Web UI Development** : For web UI development, we will explore different options, namely JavaScript-based frameworks (such as React, Vue, etc.) or Pythonic alternatives like Dash or Streamlit, which enable web UI development without requiring JavaScript.
- **Data Visualization** : We will explore different avenues for data visualization, namely D3.js if we are using JavaScript for UI development, and Plotly if we are using Streamlit or Dash. Both of these can be used to implement data filtering and to develop graphs, charts, and trend analysis dashboards.
- **Deployment**: Deployment will be done via Heroku or Streamlit Community Cloud (if Streamlit is used for web UI development).

6. Team Members and Responsibilities

The project will be divided among the following team members:

- **Web UI Development, Backend Development and Hosting**:
    - Rohan (241110057)
    - Sevak Shekokar (241110065)
    - Krishna Kumar Bais (241110038)
    - Sanjay Singh Shekhawat (241230012)
- **Visualization Module**:
    - Milan Roy (241110042)
    - Tsewang Chukey (241110092)
    - Tsewang Namgail (241110093)
    - Pushkar Bhardwaj (231110605)