COVID-19 Data Analysis Project Report

Project Title

Analysis of COVID-19 Cases Worldwide with a Focus on India

Objective

The aim of this project was to perform data analysis on global COVID-19 data, visualize trends, and extract key statistics - especially focusing on India's current situation.

Dataset

Source: Preloaded CSV

Type: CSV

Fields used: Country, Confirmed, Recovered, Deaths

Tools and Libraries Used

- Python
- pandas (for data manipulation)
- matplotlib and seaborn (for data visualization)
- Jupyter Notebook

Steps Performed

1. Data Loading & Cleaning
   - Loaded the CSV dataset using pandas.
   - Checked for missing values and removed rows with missing critical data.

2. Data Sorting & Filtering

COVID-19 Data Analysis Project Report

- Sorted countries by number of confirmed cases.
- Selected the top 10 countries with the highest confirmed cases.

3. Visualization

- Created bar plots to visualize:
  - Top 10 countries by confirmed cases.
  - Top 10 countries by deaths.
  - Top 10 countries by recovered cases.

4. India-Specific Summary

- Extracted the latest COVID-19 statistics for India.
- Created a summary DataFrame with confirmed, recovered, and deaths.
- Saved the summary to 'india_covid_summary.csv'

Key Insights

- The top affected countries are USA, India, Brazil, etc.
- India's confirmed case count is [98.8%].
- Recovery and death trends vary significantly among the top 10 countries.

Output Files

- Visual graphs (stored in notebook)
- india_covid_summary.csv (summary of Indias COVID-19 data)

Conclusion

This mini-project helped in understanding how to clean, analyze, and visualize real-world data using Python.

It also demonstrated how to extract meaningful insights and summarize them effectively.