

Vaccination Data Analysis Report

1. Project Overview

This project analyzes global COVID-19 vaccination data to provide insights into vaccination trends, adoption of different vaccines, and progress by countries. It includes preprocessing for missing values, data cleaning, and visual analytics.

2. Dataset Overview

Source File: country_vaccinations.csv

Total Records: 86,512

Columns: 15

Key Features:

- country: Country name
- iso_code: ISO 3166 country code
- date: Date of record
- total_vaccinations: Cumulative doses administered
- people_vaccinated: Individuals with at least one dose
- people_fully_vaccinated: Fully vaccinated individuals
- daily_vaccinations_raw: Raw daily doses
- daily_vaccinations: Smoothed daily doses
- *_per_hundred / *_per_million: Normalized metrics
- vaccines: Vaccine types used
- source_name & source_website: Data sources

3. Data Cleaning & Preprocessing

Null Handling Summary:

Vaccination Data Analysis Report

- total_vaccinations: 42,905 missing
- people_vaccinated: 45,218 missing
- people_fully_vaccinated: 47,710 missing
- daily_vaccinations_raw: 51,150 missing

Strategy:

- Imputed with Mean: daily_vaccinations, *_per_hundred, *_per_million
- Imputed with 0: total_vaccinations, people_vaccinated, people_fully_vaccinated, daily_vaccinations_raw
- Dropped rows still missing critical data

Result: No missing values remain. Dataset cleaned and typecast for analysis.

4. Data Analysis & Visualizations

4.1 Global Leaders:

Top countries by vaccination count: USA, India, Brazil

4.2 Daily Trends:

Time-series shows rise in global daily vaccinations post-2020

4.3 Vaccine Diversity:

Countries use a mix of Pfizer, Moderna, AstraZeneca, etc.

4.4 Country Focus - India:

India shows consistent growth in vaccine deployment

Vaccination Data Analysis Report

5. Insights Summary

- USA, India, Brazil lead in administration volume
- Clear global acceleration in vaccine rollouts post-2020
- Multiple vaccines adopted across countries
- Outliers suggest possible data entry issues (e.g., over 1B daily doses)
- India demonstrated a structured roll-out plan

6. Conclusion

The cleaned vaccination dataset reveals actionable trends in how nations approached COVID-19 vaccine distribution. The study supports strategic planning and provides transparency into global public health progress.

7. Future Enhancements

- Vaccination rate prediction models
- Dashboards for real-time tracking
- Studies on booster shots and variant coverage
- Equity studies across income regions