**Comprehensive Analysis of COVID-19 Vaccination Data**

**Readme**

**Table of Contents**

* Overview
* Project Description
* Data Sources
* Prerequisites
* Getting Started
* Usage
* Data Analysis
* Contributing
* License

**Overview**

This project aims to provide a comprehensive analysis of COVID-19 vaccination data. It utilizes real-world data to uncover insights into vaccination rates, demographics, distribution, and more. The analysis is conducted using various data analysis tools, including Python libraries like pandas, matplotlib, and seaborn. The results can be useful for researchers, policymakers, and anyone interested in understanding the progress of COVID-19 vaccination efforts.

**Project Description**

Data Exploration: This project explores various aspects of COVID-19 vaccination data, such as the number of vaccine doses administered, vaccination rates, vaccination coverage, and the distribution of vaccines by region.

Visualization: The project includes data visualization using plots and charts to help users understand the trends and patterns in vaccination data. This includes bar charts, line graphs, heatmaps, and more.

Statistical Analysis: Statistical analysis is performed to identify correlations and trends in the data. This can include regression analysis, hypothesis testing, and more.

Data Cleaning: Data from multiple sources is cleaned and transformed to create a cohesive dataset for analysis.

Data Sources: Information on the data sources used in the project is provided.

**Data Sources**

[**https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress**](https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress)

Data is collected daily from [**Our World in Data**](https://ourworldindata.org/) GitHub repository for [covid-19](https://github.com/owid/covid-19-data), merged and uploaded. Country level vaccination data is gathered and assembled in one single file. Then, this data file is merged with locations data file to include vaccination sources information. A second file, with manufacturers information, is included.

**Prerequisites**

Before you begin, ensure you have met the following requirements:

* Python 3.x installed on your local machine.
* Jupyter Notebook or an integrated development environment (IDE) for running Python scripts.
* Required Python libraries listed in the requirements.txt file.

**Getting Started**

* Clone the repository to your local machine
* Install the required Python libraries
* Open the Jupyter Notebook or your preferred Python IDE.
* Load and run the project's notebooks to start analyzing the COVID-19 vaccination data.

**Usage**

To use this project, follow these steps:

* Open the Jupyter Notebook or your Python IDE.
* Run the notebooks in the project directory. The notebooks are organized based on the aspects of COVID-19 vaccination data you want to analyze.
* Explore the data, visualize trends, and perform statistical analysis as needed.
* Customize the analysis to focus on specific regions, timeframes, or other parameters of interest.
* Interpret the results and draw conclusions based on the analysis.

**Data Analysis**

The project notebooks cover various aspects of COVID-19 vaccination data, including but not limited to:

* Daily vaccination rates
* Vaccination coverage by age group
* Regional vaccination disparities
* Vaccine distribution over time
* Impact of vaccination on case and mortality rate.