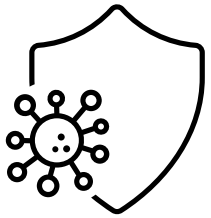


Covid – 19

Vaccination Progress Analysis

Overview

ABOUT

For this case study, one should select their own data to analyze, with the goal of conducting an exploratory visual analysis in Python and finding connections between variables that seem worth exploring.

OBJECTIVE

To build an interactive dashboard that will visually showcase well-curated results of an advanced exploratory analysis conducted in Python.

DATASET

Data set is supposed to be sourced by one's own choice and COVID vaccination dataset is chosen due to it's current importance and it meets the conditions for sourcing open data.

Analysis

Data Understanding and Preparation

- Importing Python libraries and data.
- Data wrangling and subsetting.
- Data consistency check for duplicate and missing values.
- Hypothesis formation based on different variables.

Analysis

- Exploring relations between different variables using scatterplots
- Geospatial Analysis
- Regression and Cluster Analysis
- Time series analysis (on different dataset)

Visualizations

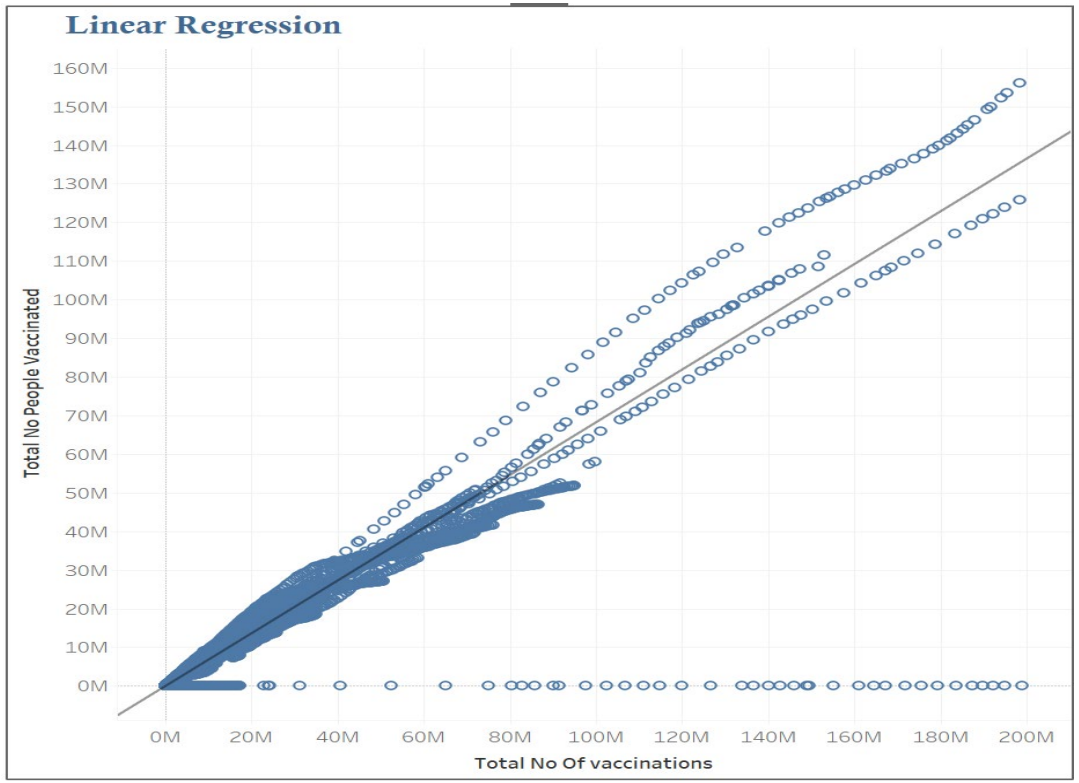
- Using Tableau, an interactive dashboard was created using all the findings from the analysis.
- Link to dashboard
- Link to GitHub repository

Findings

Hypothesis :

If a person gets at least one dose of vaccine(that is total number of vaccinated people) then the country will be more immune (As the total number of vaccinations will also increase).

- To Test the above hypothesis, I conducted linear regression analysis using python for two variables 'Total no of vaccinations' and 'Total no of people vaccinated'



The line covers most of the data points and has a upward trend.
We have a R-squared value which explains the variance in the data.

The **positive trend line** and **R-squared** value from plot which is **0.97** proves that the **correlation** between total vaccinations and total number of people vaccinated is **strong**.

Hence we can say that as total number of vaccinated people increases the total number of vaccinations also increase.