**COVID VACCINES ANALYSIS**

**TEAM MEMBERS**

**310821243012: DHANUUSH.A**

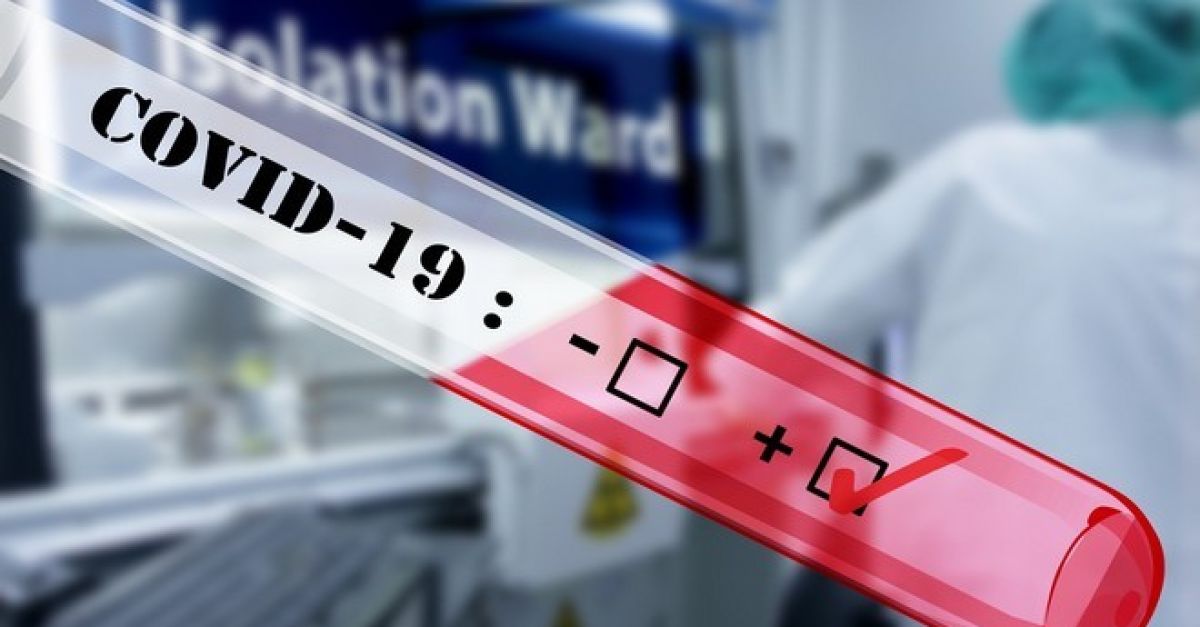
**310821243048: SAI NITHISH KUMAR.S**

**310821243034: MURALI KRISHNAN.M**

**310821243033: LOKESH.V**

**PHASE-I:** Problem Definition and Design Thinking

**PROJECT:** *DATA ANALYSIS ON COVID VACCINATION DATA*



**Objectives:**

The global response to the COVID-19 pandemic has been marked by the rapid development and deployment of vaccines. Data analysis plays a pivotal role in understanding vaccination trends, assessing vaccine effectiveness, and informing public health policies. This abstract provides an overview of key modules in the data analysis of COVID-19 vaccination.

**1. Data Collecting:**

A good sourced data will give out better analysis report so they should be more accurate, Complete, Covering geographic data, Accessible etc.

**DATASET LINK:** https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress

**2. Data Preprocessing:**

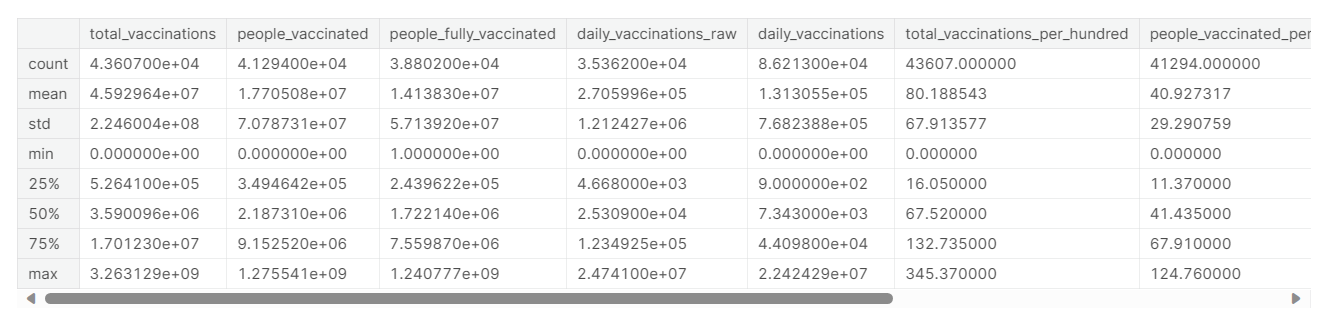
- Gathering vaccination data from various sources, including government records, healthcare providers, and research studies.

- Cleaning and standardizing data to remove inconsistencies, missing values, and errors.

- Integration of data from diverse formats and locations into a unified dataset.

**3. Descriptive Analysis:**

- Basic statistics such as vaccination rates, coverage, and demographic breakdowns.



- Visualizations, including heatmaps, time-series graphs, and geographical maps, to illustrate vaccination distribution and uptake.

**3. Vaccine Effectiveness Assessment:**

- Statistical methods to evaluate the real-world effectiveness of vaccines in preventing COVID-19 infection, hospitalization, and mortality.

- Adjustment for confounding factors such as age, comorbidities, and vaccination status.

**4. Temporal Analysis:**

- Examining vaccination trends over time to identify changes in uptake, hesitancy, and vaccine distribution.

- Forecasting future vaccination rates and potential challenges.

**5. Geospatial Analysis:**

- Mapping vaccine distribution, coverage, and hotspot areas to identify disparities in access.



- Geospatial clustering techniques to detect outbreaks and monitor their containment.

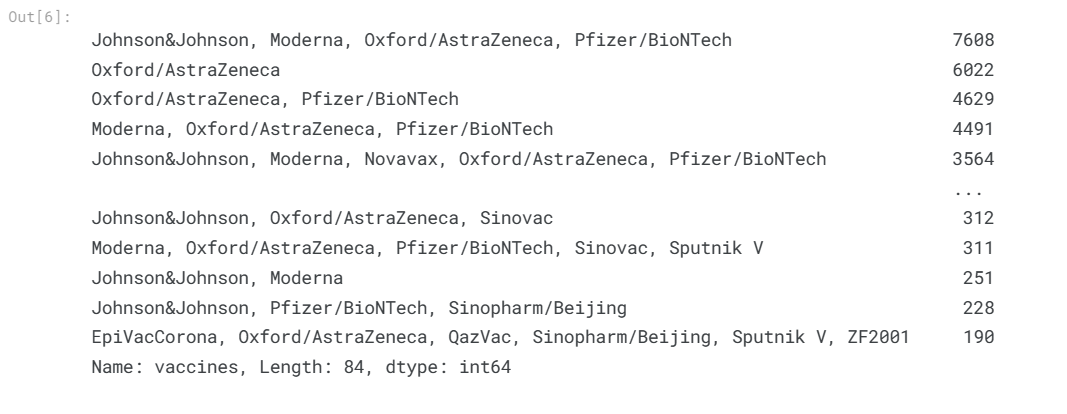
**6. Vaccine Hesitancy and Sentiment Analysis:**

- Natural language processing (NLP) to analyse public sentiment regarding COVID-19 vaccines from social media, news articles, and surveys.

- Identifying factors contributing to vaccine hesitancy and designing targeted interventions.

**7. Economic and Social Impact Assessment:**

- Analysing the economic consequences of vaccination campaigns, including healthcare cost savings and workforce productivity.



- Studying the social impact of vaccination on daily life, travel, and public gatherings.

**8. Predictive Modelling:**

- Developing machine learning models to predict future vaccination rates, vaccine supply requirements, and potential COVID-19 surges.

- Scenario analysis to inform policy decisions and resource allocation.

**9. Ethical Considerations and Privacy:**

- Addressing ethical challenges related to data privacy, informed consent, and transparency in data analysis.

- Ensuring responsible data handling practices.

**SOURCE CODE:**

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import plotly.express as px

import plotly.offline as py

import os

for dirname, \_, filenames in os.walk('/kaggle/input'):

for filename in filenames:

print(os.path.join(dirname, filename))

data = pd.read\_csv("/kaggle/input/covid-world-vaccination-progress/country\_vaccinations.csv")

data.head()

data.describe()

pd.to\_datetime(data.date)

data.country.value\_counts()

data = data[data.country.apply(lambda x: x not in ["England", "Scotland", "Wales", "Northern Ireland"])]

data.country.value\_counts()

data.vaccines.value\_counts()

df = data[["vaccines", "country"]]

df.head()

dict\_ = {}

for i in df.vaccines.unique():

dict\_[i] = [df["country"][j] for j in df[df["vaccines"]==i].index]

vaccines = {}

for key, value in dict\_.items():

vaccines[key] = set(value)

for i, j in vaccines.items():

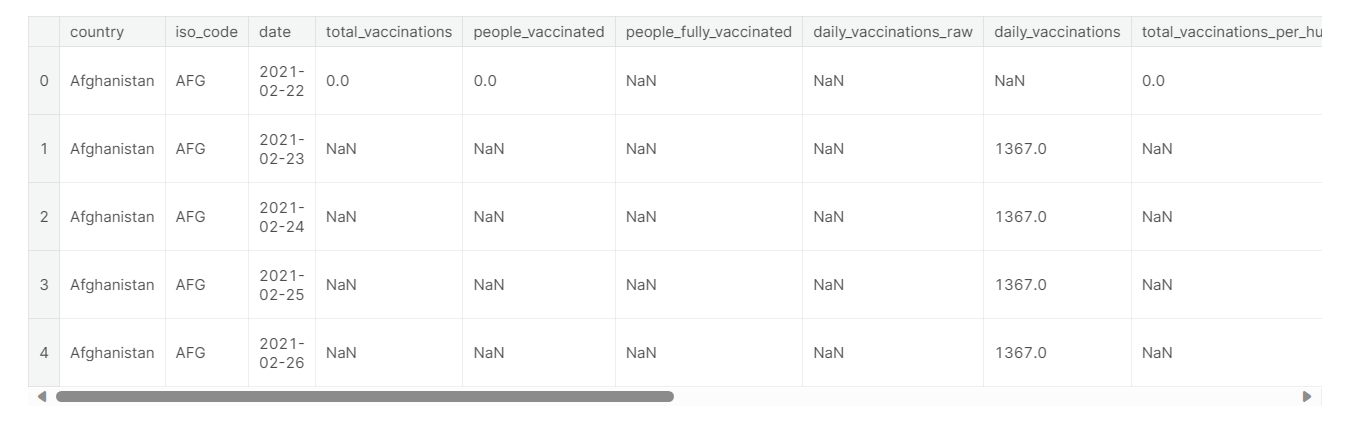
print(f"{i}:>>{j}")

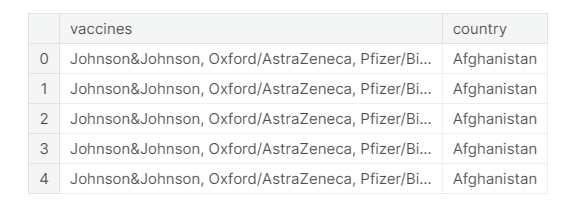
vaccine\_map = px.choropleth(data, locations = 'iso\_code', color = 'vaccines')

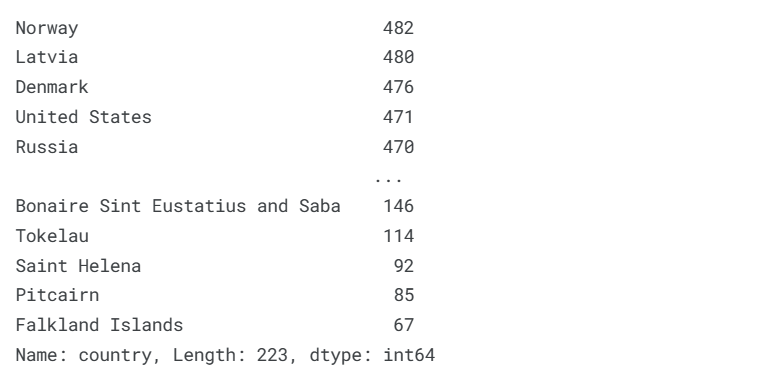
vaccine\_map.update\_layout(height=300, margin={"r":0,"t":0,"l":0,"b":0})

vaccine\_map.show()

**SAMPLE OUTPUT:**



****

****