**Phase** **3** : **Development part 1**

**Coding:**

import pandas as pd

import matplotlib.pyplot as plt

# Load COVID-19 vaccine data (replace with your dataset)

vaccine\_data = pd.read\_csv('covid\_vaccine\_data.csv')

# Data Preprocessing (cleaning and exploration)

# You may need to handle missing data, filter relevant columns, and perform data transformations

# Example: Plot vaccine distribution by manufacturer

manufacturer\_counts = vaccine\_data['Manufacturer'].value\_counts()

manufacturer\_counts.plot(kind='bar')

plt.title('COVID-19 Vaccine Distribution by Manufacturer')

plt.xlabel('Manufacturer')

plt.ylabel('Count')

plt.show()

# Hypothesis Testing

# You can use statistical tests to test hypotheses related to vaccine effectiveness or other aspects.

# Machine Learning (Regression example)

from sklearn.linear\_model import LinearRegression

from sklearn.model\_selection import train\_test\_split

# Prepare the data for regression (X should be your features, y should be your target variable)

X = vaccine\_data[['Feature1', 'Feature2', ...]]

y = vaccine\_data['Target']

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2)

# Create a Linear Regression model

model = LinearRegression()

model.fit(X\_train, y\_train)

# Model Evaluation

# Evaluate the model's performance using appropriate metrics

predicted = model.predict(X\_test)

# Data Visualization (optional)

# You can create various visualizations to illustrate your findings and insights.

# Save or present your results in a report or presentation.