Phase 1 : Problem Definition and Design Thinking

Problem Definition:

The problem is to conduct an in-depth analysis of Covid-19 vaccine data, focusing on vaccine efficacy, distribution, and adverse effects. The goal is to provide insights that aid policymakers and health organizations in optimizing vaccine deployment strategies. This project involves data collection, data preprocessing, exploratory data analysis, statistical analysis, and visualization.

Design Thinking:

Data collection:

*Data on how many doses have been distributed to different regions or countries

* The percentage of the population that has received one or both vaccine doses.

Data Preprocessing:

*Remove duplicate records

*Correct any data entry errors or inconsistencies

*Split the dataset into training, validation, and test sets for machine learning models.

Elaboraty Data Analytics:

*Creat visualization such as barcharts, line plots to illustrate trends and pattern.

*Conduct a detailed analysis of vaccine efficacy by age groups,genders and other demographics.

*Using ML algorithm like regression, decision trees, deep learning to estimate vaccine coverage.

Statistics Analysis:

*Use regression models to explore relationship between variables such as

Linear regression: predict vaccine coverage based on demographic factors

Logistic regression: Model the likelihood of adverse events occurring.

Visualisation:

Create a map and bar charts showing the geographic distribution of vaccines to visualize where doses have been allocated.

Insights And Recommendations:

Develop vaccination strategies that consider the unique needs and preferences of different demographic

groups, such as age ,gender,and ethnicity