Unveiling The Dynamics of Covid-19- Examination of Testing, Vaccination and Travel Patterns

1. **Project Description:** “Unveiling the Dynamics of COVID-19: Examination of Testing, Vaccination, and Travel Patterns"
   This project aims to analyze archived COVID-19 data from Los Angeles and San Francisco, incorporating metrics such as cases, new fatalities, testing rates, hospitalization, and passenger flow patterns. The analysis will uncover trends, evaluate the impact on healthcare systems, and assess the decision-making factors influencing health officials, travel flow patterns and individuals.

2. **Objective:** The purpose of utilizing COVID-19 data for visualization is to gain a comprehensive understanding of the counties, offering inclusive insights into the local impact of the pandemic. This data facilitates the analysis of yearly recorded cases, hospital admissions, and date-wise occurrences across the entire county, categorized by new cases, deaths, and testing, providing a clear overview of the pandemic's spread and severity. Additionally, the project involves comparing passenger traffic at San Francisco International Airport (SFO) and Los Angeles International Airport (LAX) during three significant phases (pre-COVID, during-COVID, and post-COVID), and examining and visualizing the impact of COVID-19 on monthly passenger flow at SFO.

3. **Hypothesis:**
   a) Throughout the pandemic in Florida, Individuals aged 75 to 84 and older are more vulnerable to fatalities from COVID and related infections, a trend that can be linked to the decline in immune function and other health complications commonly associated with advanced age.
   b) In December 2020, LA counties conducted 11,094,477 daily tests, identifying 990,965 positive cases. In December 2021, 21,085,395 daily tests were conducted with 777,135 positive cases, despite ongoing vaccination. Similarly, in December 2022 daily test counting was 8,140,589 with a decline of 501,793 positive cases. This prompts the exploration of vaccine efficacy and potential factors contributing to sustained positivity rates.

c) Throughout the pandemic in the United States, higher mortality rates were recorded, suggesting a seasonal pattern in COVID-19 deaths, with higher rates observed in colder months (November-March) relative to warmer months(April - October).
d) Similar passenger traffic patterns are expected between San Francisco International Airport (SFO) and Los Angeles International Airport (LAX). The anticipation is for LAX to consistently show higher overall passenger counts, with domestic flights outweighing international flights across pre-COVID (Jan 2018-Dec 2019), pandemic (Jan 2020-Dec 2021), and post-pandemic (Jan 2022-Dec 2022) periods.
e) The monthly passenger flow at San Francisco International Airport was significantly affected by COVID-19, leading to a 50% reduction during the pandemic (January 2020 to December 2021). A subsequent recovery in passenger numbers is anticipated post-pandemic (January 2022 to December 2022), aligning with the levels observed in the pre-COVID period (January 2018 to December 2019).

4. **Data Sources:**
Primary Data Sets:
COVID-19 Testing and Case Counts in LA County | Los Angeles - Open Data Portal (lacity.org)
Pandemic Recovery Dashboard (lacity.org)
LA County COVID Cases | Los Angeles - Open Data Portal (lacity.org)

Column Name	Description
county	Country Name
State	State Name
Fips	Federal Information Processing standards (FIPS)
date	Date of the data
Lat	Latitude of the County
Lon	Longitude of the county
Cases	Total no of cases in county
Deaths	Total no of Deaths in the county till date
state_cases	Total no of Cases in the State
state_deaths	Total no of Deaths in the state till date
new_cases	No of New Cases in the county on date
new_deaths	No of New DEATHS in the county on date
new_state_cases	No of New Cases in the country on date
new_state_deaths	No of New DEATHS in the country on date

County_Performed	No of testings Performed in County as on Date
County_positive	No of Positives in the Testing in County as on Date

Secondary Data Sets:

https://data.lacity.org/Transportation/Los-Angeles-International-Airport- Passenger-Traffi/g3qu-7q2u
https://data.sfgov.org/Transportation/Air-Traffic- PassengerStatistics/rkru-6vcg

Column Name	Description
reportperiod	Time Period- Month and Year
Terminal	Terminal number of the flight in the Airport
Arrival_Departure	Gives the action performed by the passenger- Arrival or Departure
Domestic International	Gives if the Flight is Domestic or International
Passenger_Count	Gives the total count of passengers

5. **Number of Records:**
   There are about 34690 records for the LA County Covid Cases, including covid testing and pandemic recovery data.
   There are about 7936 records in the San Francisco dataset and 2210 records in the Los Angeles dataset.

6. **Data Cleansing Tool:** Python and MS Excel are the data cleansing tools we will be using for the visualization as they have widely used open-source data manipulation and analysis libraries that provide powerful tools for cleaning, transforming, and analyzing tabular data. We will use Python for calculations and creation of custom data frames because it allows us to export Excel files for use in tableau visualization. Excel can be used to make sure the dataset is accurate and ready for visualization by identifying and eliminating duplicate entries, fixing formatting errors, and filling in missing values.

7. **Visualization Tool:** In our project, we'll use Tableau as our Primary tool for the main Visualization as it can handle Large Datasets and provide with a variety of visualization options to effectively present insights. As we are dealing with the dataset related to COVID-19 Los Angeles County and travel data of San Francisco and Los Angeles, tableau helps in visualizing the data in different types like Line charts, Bar charts and Maps where we can visualize the states wise data in the US map and various other types of visualizations. Additionally, we will be using MS Excel to perform some of the Pre-Visualization Tasks like small scale visualizations and Basic Analysis.