**Steps for Aggregation and Visualization of Data for Disease Outbreak Surveillance**

**1. Data Aggregation with Python**

**Objective:** Combine FHIR .json files from multiple hospitals into a unified dataset.

**Steps:**

1. **Set Up Python Environment:**
   * Install required libraries:

pip install pandas json numpy

1. **Load and Parse JSON Files:**
   * Use a Python script to read FHIR .json files (e.g., aspirus\_data.json, portage\_data.json).
   * Example script snippet:

import json

import pandas as pd

def load\_fhir\_data(file\_path):

with open(file\_path, 'r') as f:

data = json.load(f)

return data

# Load data from each hospital

hospitals = ['aspirus', 'portage', 'bcmh', 'mgh']

all\_data = []

for hospital in hospitals:

data = load\_fhir\_data(f'{hospital}\_data.json')

all\_data.extend(data['entry']) # Extract patient entries

1. **Extract Key Data Points:**
   * Filter for COVID-19 cases using FHIR Condition resources with ICD-10 code U07.1.

covid\_cases = []

for entry in all\_data:

if entry['resource']['resourceType'] == 'Condition':

if any(code['coding'][0]['code'] == 'U07.1' for code in entry['resource']['code']['coding']):

covid\_cases.append({

'city': entry['resource']['subject']['reference'].split('/')[-1],

'date': entry['resource']['onsetDateTime']

})

1. **Aggregate and Export to CSV:**
   * Convert to a Pandas DataFrame and save:

df = pd.DataFrame(covid\_cases)

df.to\_csv('aggregated\_covid\_cases.csv', index=False)

**2. Data Visualization in Google Looker Studio**

**Objective:** Create an interactive dashboard to visualize COVID-19 outbreaks.

**Steps:**

1. **Upload Data to Looker Studio:**
   * Go to [Google Looker Studio](https://lookerstudio.google.com/).
   * Click **Create** → **Report** → **Import CSV** (upload aggregated\_covid\_cases.csv).
2. **Create a Geo Map Visualization:**
   * Click **Add Chart** → **Geo Map**.
   * Set **Location Dimension** to city.
   * Set **Metric** to Record Count (or SUM of cases).
3. **Create a Bar Chart:**
   * Click **Add Chart** → **Bar Chart**.
   * Set **Dimension** to city.
   * Set **Metric** to Record Count.
4. **Configure Dashboard:**
   * Add a title (e.g., "COVID-19 Outbreak Surveillance").
   * Set **Auto-Refresh** to **15 minutes** (under *Report Settings* → *Data Freshness*).
5. **Share the Dashboard:**
   * Click **Share** → Enable **"Anyone with the link can view"**.