Question:

Specific: COVID-19 cases in Klang Valley area

Measurable: less than 3000 cases per day

Action-focused: Booster dose (3rd dose)

Relevant: Reduce COVID-19 cases

Time-bound: January to April

Can the booster dose (3rd dose) reduce COVID-19 cases in Klang Valley area to less than 3000 cases per day from January to April?

Aim:

Creating a dashboard to show the effectiveness of booster dose towards reducing COVID-19 cases in Klang Valley

Type of data:

csv

Data:

Secondary data collection. The reason to use secondary data over primary data is because it saves time to create code to gather and extract information from internet.

Methods:

The method used is Government records because the GitHub repository was maintained and updated daily by Ministry of Health Malaysia, which is a reliable source of COVID-19 Malaysia data.

Metadata:

Dataset containing COVID positive cases information on different state of Malaysia.

The data was collected by Goh Jia Ean on 14/07/2022.

Data was collected via Government records from Ministry of Health official GitHub repository: https://github.com/MoH-Malaysia/covid19-public

The file named "cases_state.csv" contain information as shown below:

Column	Data Type	Description
date	date	yyyy-mm-dd format; data correct as of 1200hrs on that date
state	string - categorical	name of the Malaysia state
cases_new	integer	cases reported in the 24hrs since the last report
cases_import	integer	imported cases reported in the 24hrs since the last report
cases_recovered	integer	recovered cases reported in the 24hrs since the last report
cases_active	integer	individuals who have not yet recovered or died from COVID
cases_cluster	integer	number of cases attributable to clusters
cases_unvax	integer	Number of individuals who has not yet vaccinated who tested positive for COVID
cases_pvax	integer	number of partially-vaccinated (1-dose) individuals who tested positive for COVID
cases_fvax	integer	number of fully-vaccinated (2-doses) individuals who tested positive for COVID
cases_boost	integer	number of fully-vaccinated with booster (3rd dose) individuals who tested positive for COVID
cases_child	integer	COVID positive individuals between ages 0-11
cases_adolescent	integer	COVID positive individuals between ages 12-17
cases_adult	integer	COVID positive individuals between ages 18-59
cases_elderly	integer	COVID positive individuals between ages 60+
cases_0_4	integer	COVID positive individuals between ages 0-4
cases_5_11	integer	COVID positive individuals between ages 5-11
cases_12_17	integer	COVID positive individuals between ages 12-17
cases_18_29	integer	COVID positive individuals between ages 18-29
cases_30_39	integer	COVID positive individuals between ages 30-39
cases_40_49	integer	COVID positive individuals between ages 40-49
cases_50_59	integer	COVID positive individuals between ages 50-59
cases_60_69	integer	COVID positive individuals between ages 60-69
cases_70_79	integer	COVID positive individuals between ages 70-79
cases_80	integer	COVID positive individuals ages 80 and above

Decision making steps:

1. Explore the files in the repositories which contain data that answers the question
2. Copy and paste the data into a notepad and save as csv file
3. Open the csv file in Microsoft Excel for data wrangling
4. Save the cleaned data as a separate file
5. Data analysis in Excel – Pinpoint the date of booster dose rollout with the COVID cases (Jan-Apr)
6. Data visualisation in Tableau – Geographical data, line graph, etc.

Tools used:

- GitHub (data source)
- Notepad (csv file)
- Microsoft Excel (read csv file, data wrangling and data analysis)
- Tableau (data visualization)

Dates and timelines:

Activity	14/7	15/7	16/7	17/7
Understand and asking question				
Data collection				
Data wrangling				
Data analysis and visualisation				
Document the process				
Communicate insights				

Contact information

Data collector: MoH-Malaysia (https://github.com/MoH-Malaysia)