MINI PROJECT
ANALYZING DATA

Irman Maulana Pamungkas

INDONESIA'S COVID-19 DISTRIBUTION IN 2020-2021

From this dataset, I'm defining the problems from dataset by looking for which one of the data are genuinely important to be discussed or described.

Field name	Type	Mode
Date	DATE	NULLABLE
Province	STRING	NULLABLE
Daily_Case	INTEGER	NULLABLE
Daily_Death	INTEGER	NULLABLE
Daily_Recovered	INTEGER	NULLABLE
Active_Case	INTEGER	NULLABLE
Cumulative_Case	INTEGER	NULLABLE
Cumulative_Recovered	INTEGER	NULLABLE
Cumulative_Death	INTEGER	NULLABLE
Cumulative_Active_Case	INTEGER	NULLABLE

ABOUT DATASET

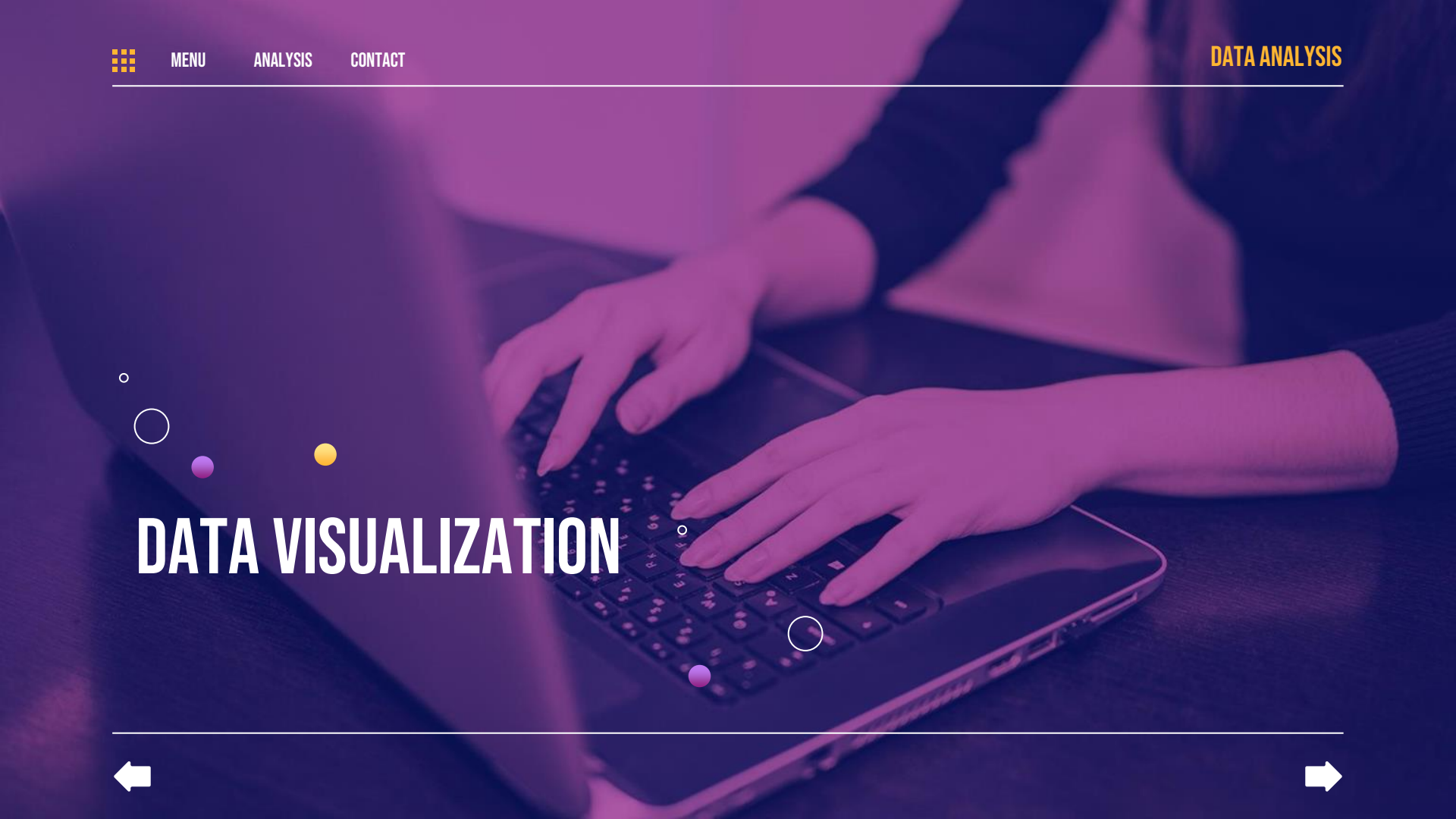
Indonesia Coronavirus daily data Columns Description:

- Date: Means the Date of Observation
- Province: Means Location/ Province which the data was observed
- Daily_Case: Means Daily new number of confirmed case in observed province
- Daily_Death: Means Daily new number of confirmed death in observed province
- Daily_Recovered: Means Daily new number of confirmed recover in observed province
- Active_Case: Means Daily new number of active case such as isolated or threated in hospital (still didnt recover nor die)
- Cumulative_Case: Means total for each day number number of confirmed case of the row's date, for the row's province
- Cumulative_Recovered: Means total for each day number number of confirmed Recover of the row's date, for the row's province
- Cumulative_Death: Means total for each day number of confirmed death of the row's date, for the row's province
- Cumulative_Active_Case: Means total for each day number of active case of the row's date, for the row's province

THE PROBLEM QUESTIONS:

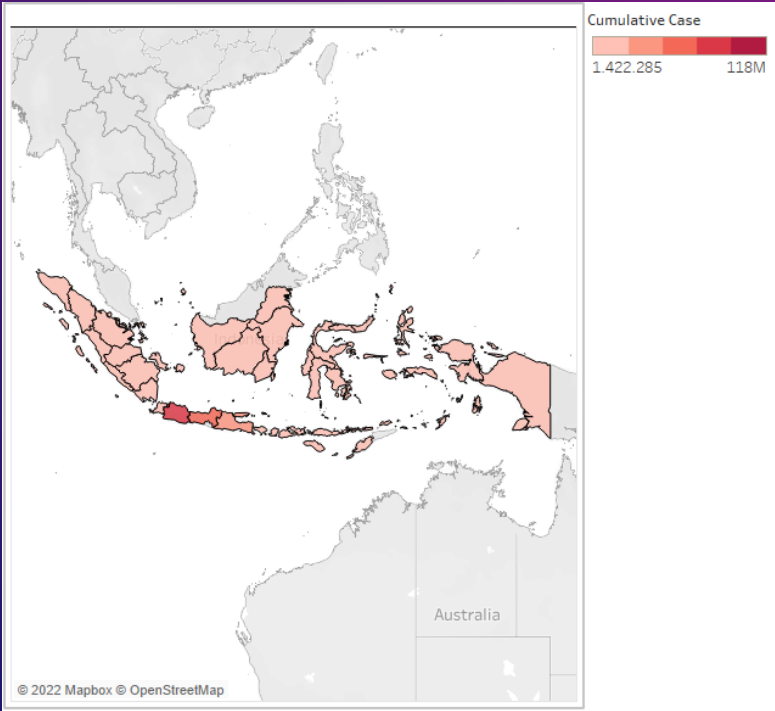
After looking for which data that are genuinely interesting to be discussed, then I breakdown the problem questions. So, here's the problem questions I've found:

1. How the distribution of Covid-19 in Indonesia from 2020-2021?
2. How the description of Covid-19 cases in Indonesia day by day from 2020-2021?
3. How the image of Covid-19 daily case in every Indonesia's provinces?
4. How the data present about Covid-19 case in Indonesia in every aspect such as case, death, and recovered from 2020-2021?
5. How the relations between death and recovered of Covid-19 cases in Indonesia?



DATA VISUALIZATION

1. THE DISTRIBUTION OF COVID-19 IN INDONESIA IN 2020-2021

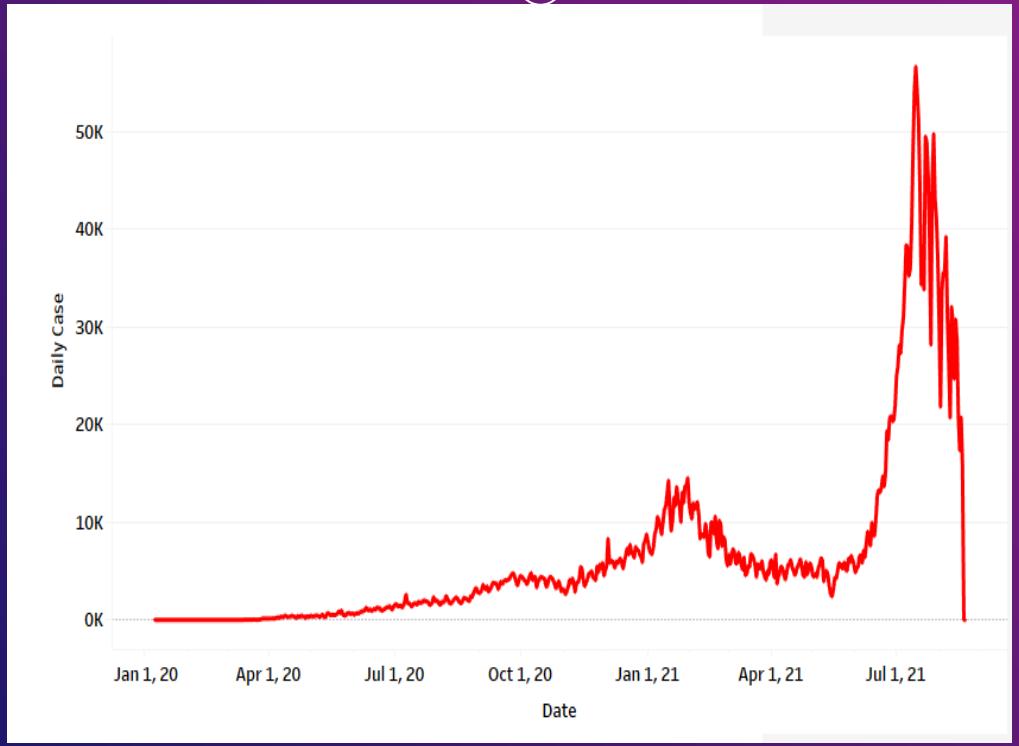


Here's the SQL Query to find the data from dataset:

```sql
SELECT
    Province,
    SUM(Cumulative_Case),
    SUM(Cumulative_Death)
FROM `irman-latihan-1.project_data_analyst10021206.covid19_indo`
GROUP BY Province;
```

Here's the finding present of how the distribution of Covid-19 in Indonesia throughout 2020-2021. I'm using the map chart because I used Province data, Cumulative Case, and Cumulative Death to looking for how the distribution and it makes clear how the distribution is it in every Indonesia's provinces. This finding presents that every provinces in Indonesia has Covid-19 cases, and DKI Jakarta Province has the most Cumulative Covid-19 cases throughout 2020-2021 with 117.633.477 cases and 1.978.802 death cases.

2. THE DESCRIPTION OF COVID-19 CASES IN INDONESIA DAY BY DAY FROM 2020-2021

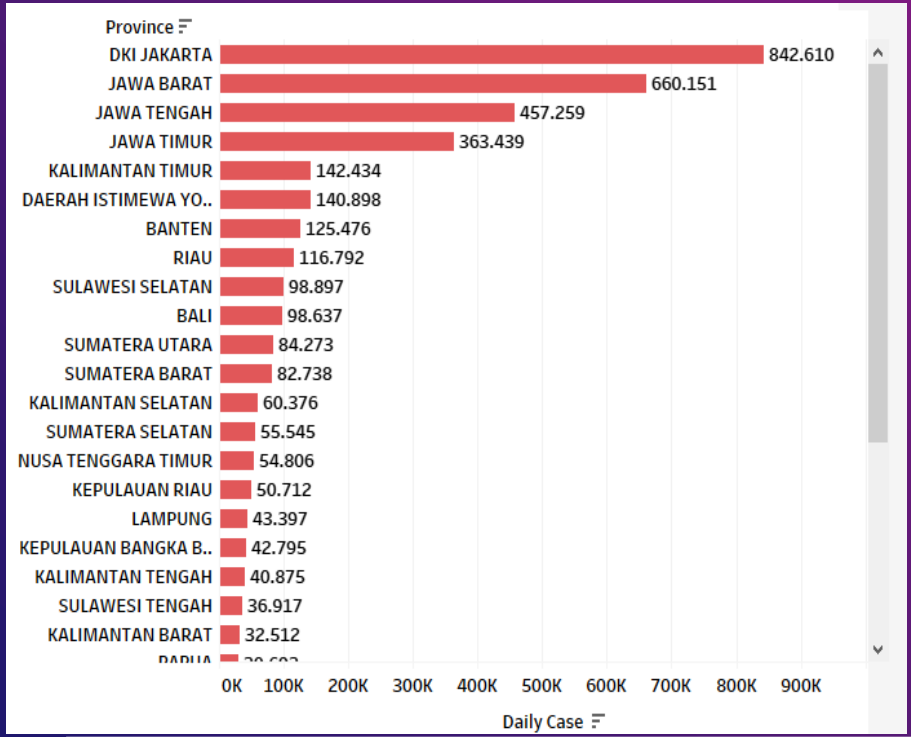


Here's the SQL Query to find the required data from dataset:

```sql
SELECT
    Date,
    SUM(Daily_Case)
FROM `irman-latihan-1.project_data_analyst10021206.covid19_indo`
GROUP BY Date;
```

Here's the second one. I'm using the lines chart because I would like to find the daily case of Covid-19 in Indonesia day by day and it requires time spans variable to make this chart. This finding show that in early 2020 were not many case and start to increase in early 2021 when there was a Nataru's holiday. The most daily cases of Covid-19 in Indonesia happened on 15th of July 2021 with 56.694 cases.

3. THE IMAGE OF COVID-19 DAILY CASE IN EVERY INDONESIA'S PROVINCES



Here's the SQL Query to find the required data from dataset:

```sql
SELECT
    Province,
    SUM(Daily_Case)
FROM `irman-latihan-1.project_data_analyst10021206.covid19_indo`
GROUP BY Province;
```

The third one of finding present shows about the total of daily cases of Covid-19 in every provinces. I'm using bar chart to find the total of daily cases of Covid-19 in Indonesia based on provinces. This chart order by the total daily case to find daily cases in provinces from highest to lowest. So it shows that DKI Jakarta Province has the most total daily cases, West Java as the second, Central Java follows as third, and so on.

4. THE DATA PRESENT ABOUT COVID-19 CASE IN INDONESIA ON CUMULATIVE OF CASE, DEATH, AND RECOVERED FROM 2020-2021

481.467.054
Cumulative Case

13.939.996
Cumulative Death

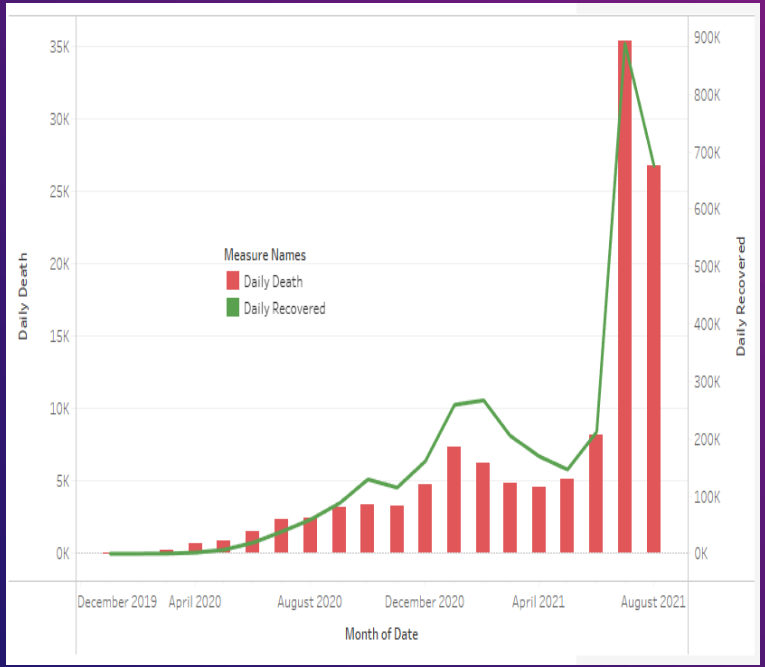
410.461.830
Cumulative Recovered

This data shows how Covid-19 cases in Indonesia on Cumulative Case, Death, and Recovered throughout 2020-2021.

Here's the SQL Query to find the required data from dataset:

```sql
SELECT
    SUM(Cumulative_Case),
    SUM(Cumulative_Death),
    SUM(Cumulative_Recovered)
FROM `irman-latihan-1.project_data_analyst10021206.covid19_indo`
```

3. THE RELATIONS BETWEEN DEATH AND RECOVERED OF COVID-19 CASES IN INDONESIA



Here's the SQL Query to find the required data from dataset:

```sql
SELECT
    FORMAT_DATE('%b %y', Date) AS Month,
    SUM(Daily_Death),
    SUM(Daily_Recovered)
FROM `irman-latihan-1.project_data_analyst10021206.covid19_indo`
GROUP BY Month;
```

Here's the last one of the finding presents. In this question, I'm using the combo chart to find the relation between total cases of Daily Death and Daily Recovered in Indonesia month by month throughout 2020-2021. The bar represent Daily Death, and the line represent Daily Recovered. This chart shows that the Daily Recovered directly proportional to Daily Death.

TOOLS AND RESOURCES

01

SQL QUERY

Google Big Query

02

DATA VIZUALIZATION AND DASHBOARD

tableau

01

DATASET

kaggle

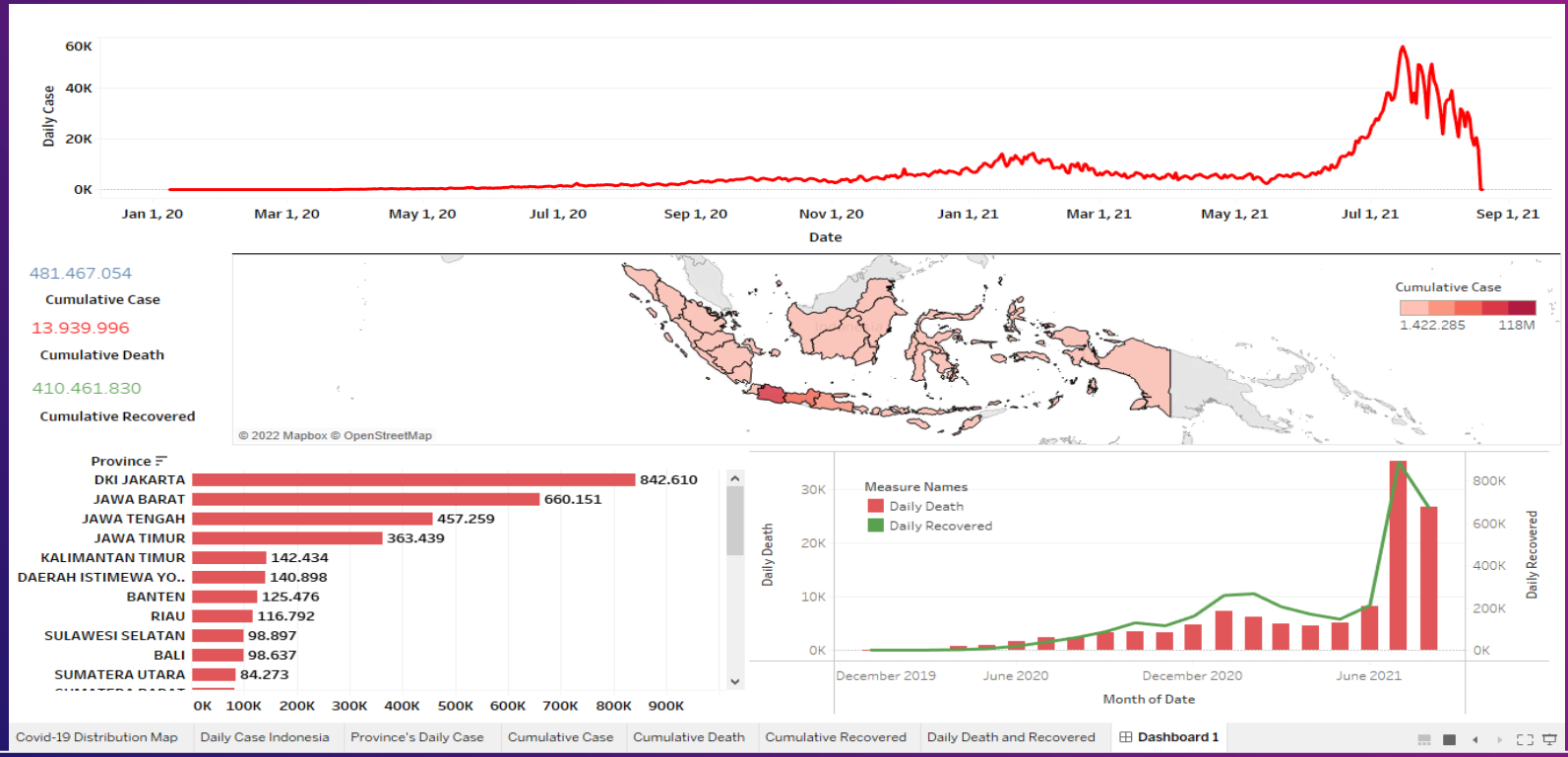
COVID-19 Indonesia Dataset (Case and Vaccination) by Thoriqul Aziz

04

POWERPOINT TEMPLATE

slidesgo

DASHBOARD


60K
40K
20K
0K
Daily Case
Jan 1, 20
Mar 1, 20
May 1, 20
Jul 1, 20
Sep 1, 20
Nov 1, 20
Jan 1, 21
Mar 1, 21
May 1, 21
Jul 1, 21
Sep 1, 21
Date
481.467.054
Cumulative Case
13.939.996
Cumulative Death
410.461.830
Cumulative Recovered
Cumulative Case
1.422.285
118M
© 2022 Mapbox © OpenStreetMap
Province
DKI JAKARTA 842.610
JAWA BARAT 660.151
JAWA TENGAH 457.259
JAWA TIMUR 363.439
KALIMANTAN TIMUR 142.434
DAERAH ISTIMEWA YO.. 140.898
BANTEN 125.476
RIAU 116.792
SULAWESI SELATAN 98.897
BALI 98.637
SUMATERA UTARA 84.273
0K 100K 200K 300K 400K 500K 600K 700K 800K 900K
Measure Names
Daily Death
Daily Recovered
30K
20K
10K
0K
Daily Death
800K
600K
400K
200K
0K
Daily Recovered
December 2019
June 2020
December 2020
June 2021
Month of Date
Covid-19 Distribution Map
Daily Case Indonesia
Province's Daily Case
Cumulative Case
Cumulative Death
Cumulative Recovered
Daily Death and Recovered
Dashboard 1

CERTIFICATION OF COMPLETION BY REVOU

RevoU Mini Course

Issued 27 May 2022

IRMAN MAULANA

has been awarded a certificate of completion for the

Intro to Data Analytics

a 2-weeks certified online course offered by RevoU

Matteo Sutto
CEO and Co-Founder
PT Revolusi Cita Edukasi

Verify at https://certificates.revou.co/irman-maulana-certificate-completion-damc22.pdf
Revou has confirmed the identity of this individual and their participation in the course

CERTIFICATE OF COMPLETION