

Data Science 4-R

G.E Group | Binar Academy

FINAL PRESENTATION BINAR ACADEMY DATA SCIENCE



digitalent



BINAR



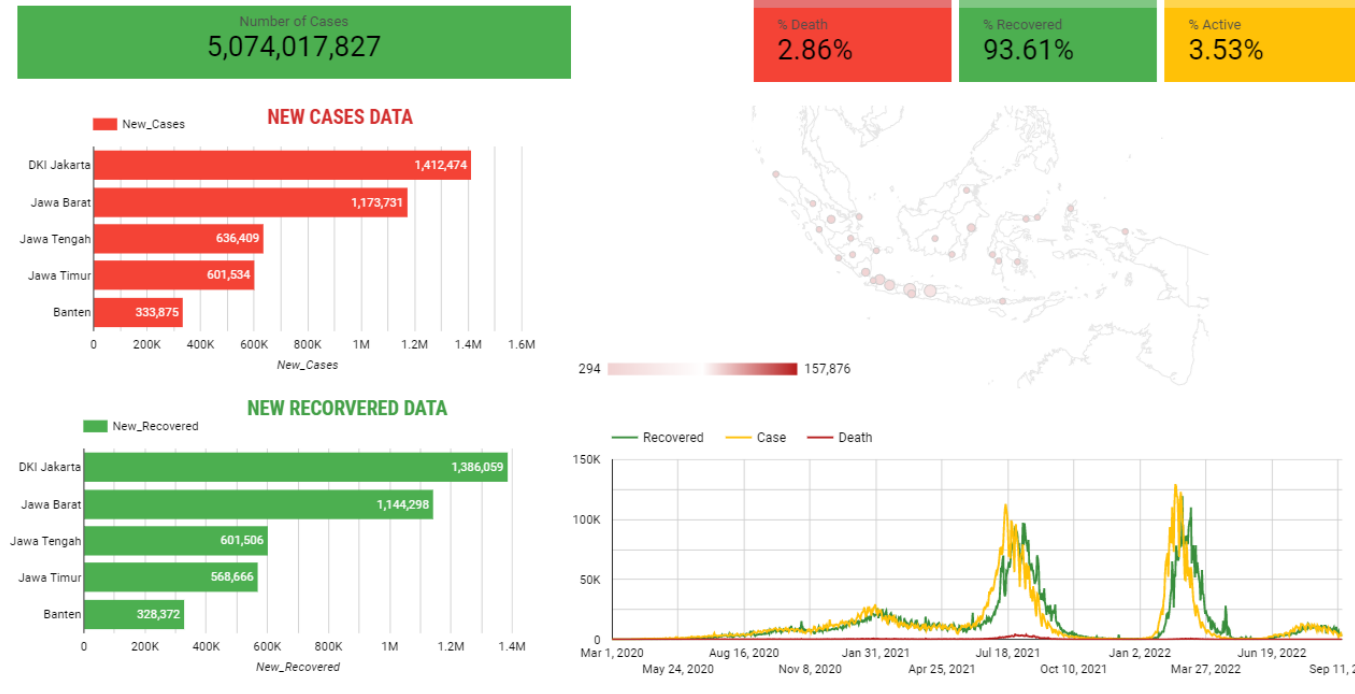
Muhammad Ghifari Kusuma

<https://github.com/MasGhiff/>

Google Data Studio

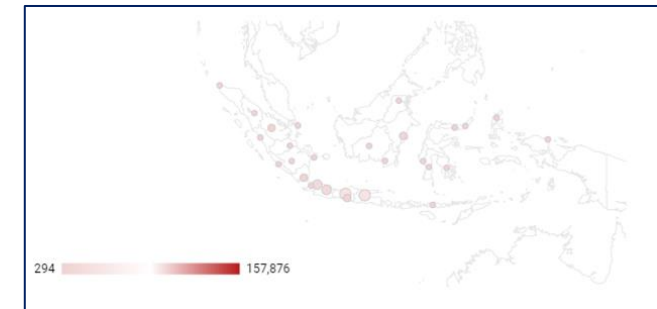
Dashboard Covid 19 in Indonesia

INDONESIA COVID SPREAD 2019-2022



We're visualize a data about covid spread in Indonesia from 2019-2022. The visualization consist of % of death, recovered and active case. Based on time series visualization concluded, the highest spread of covid-19 was in July 2021, and March 2022. In addition, Jakarta shown the highest of New Recovered Data outperform another city, but Jakarta also shown the highest of new cases data found.

Number of Deaths 145,259,577	$\frac{SUM(Total\ Death)}{SUM(Total\ Cases)}$
% Death 2.86%	
Total_Recovered 4,749,798,246	$\frac{SUM(Total\ Recovered)}{SUM(Total\ Cases)}$
% Recovered 93.61%	
Total_Active_Cases 178,960,004	$\frac{SUM(Total\ Active\ Cases)}{SUM(Total\ Cases)}$
% Active 3.53%	



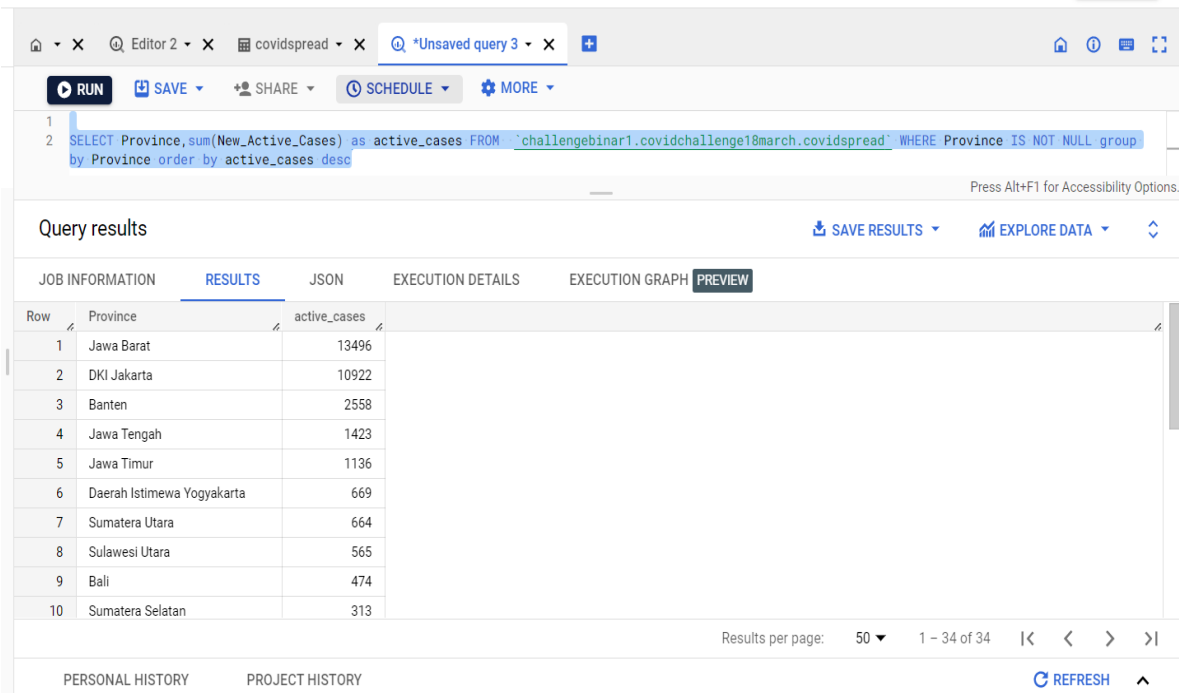
The image shown the spread of covid in each of Province. The highest number, indicates with strong red color.

SQL

Total Number New Active Covid 19

The total number of new active Covid-19 cases in each province is then sorted by the largest number of cases

```
SELECT Province,sum(New_Active_Cases) as  
active_cases FROM  
`challengebinar1.covidchallenge18march.covidspr  
ead` WHERE Province IS NOT NULL group by  
Province order by active_cases desc
```



The screenshot shows a SQL query editor interface. At the top, there's a toolbar with buttons for RUN, SAVE, SHARE, SCHEDULE, and MORE. Below the toolbar, the SQL query is displayed in a text area. The query is: `SELECT Province, sum(New_Active_Cases) as active_cases FROM `challengebinar1.covidchallenge18march.covidspread` WHERE Province IS NOT NULL group by Province order by active_cases desc`. Below the query, there's a section for 'Query results' with tabs for JOB INFORMATION, RESULTS, JSON, EXECUTION DETAILS, and EXECUTION GRAPH. The 'RESULTS' tab is selected, showing a table with 10 rows of data. The table has columns for Row, Province, and active_cases. The data is sorted by active_cases in descending order. At the bottom right, there's a 'Results per page' dropdown set to 50, and a 'REFRESH' button.

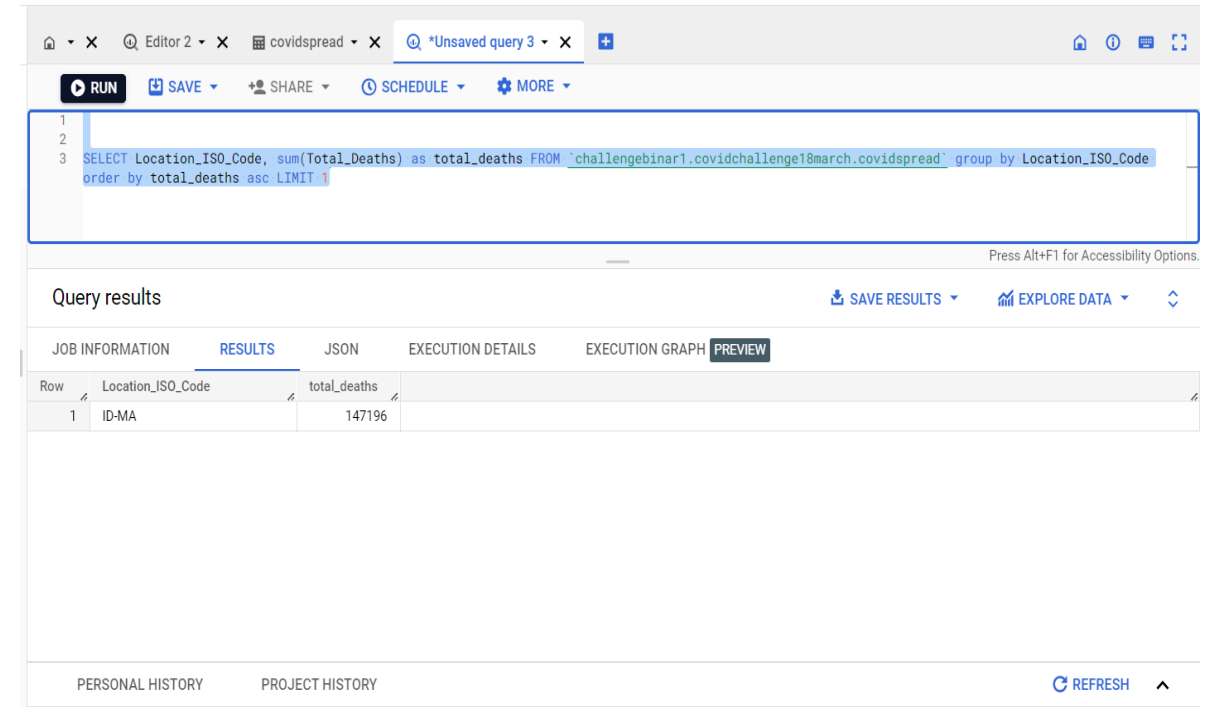
Row	Province	active_cases
1	Jawa Barat	13496
2	DKI Jakarta	10922
3	Banten	2558
4	Jawa Tengah	1423
5	Jawa Timur	1136
6	Daerah Istimewa Yogyakarta	669
7	Sumatera Utara	664
8	Sulawesi Utara	565
9	Bali	474
10	Sumatera Selatan	313

SQL

Total Number of Deaths

Take 2 (two) location iso codes that have the least total number of deaths due to Covid-19

```
SELECT Location_ISO_Code, sum(Total_Deaths)
as total_deaths FROM
`challengebinar1.covidchallenge18march.covidspr
ead` group by Location_ISO_Code order by
total_deaths asc LIMIT 1
```



The screenshot shows a web-based SQL editor interface. At the top, there's a toolbar with buttons for RUN, SAVE, SHARE, SCHEDULE, and MORE. Below the toolbar is a text area containing the SQL query: `SELECT Location_ISO_Code, sum(Total_Deaths) as total_deaths FROM `challengebinar1.covidchallenge18march.covidspread` group by Location_ISO_Code order by total_deaths asc LIMIT 1`. Below the query editor, there's a section titled "Query results" with tabs for JOB INFORMATION, RESULTS (selected), JSON, EXECUTION DETAILS, EXECUTION GRAPH, and PREVIEW. The RESULTS tab shows a table with two columns: "Location_ISO_Code" and "total_deaths". The first row shows "ID-MA" with a value of "147196". At the bottom of the interface, there are tabs for PERSONAL HISTORY and PROJECT HISTORY, and a REFRESH button.

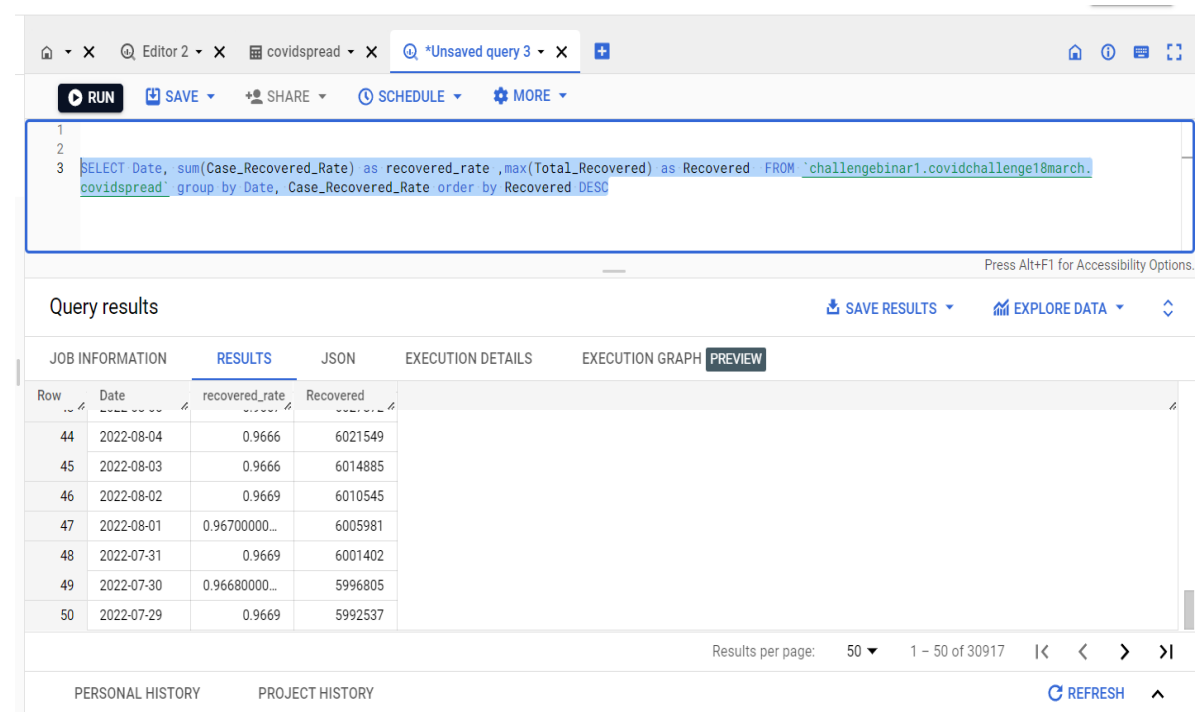
Row	Location_ISO_Code	total_deaths
1	ID-MA	147196

SQL

Dates Recovered Case is Highest

Data on the dates when the rate of recovered cases in Indonesia is highest and the number of rates

```
SELECT Date, sum(Case_Recovered_Rate) as
recovered_rate ,max(Total_Recovered) as
Recovered FROM
`challengebinar1.covidchallenge18march.covidspr
ead` group by Date, Case_Recovered_Rate order
by Recovered DESC
```



The screenshot shows a SQL query editor with the following query:

```
SELECT Date, sum(Case_Recovered_Rate) as recovered_rate ,max(Total_Recovered) as Recovered FROM `challengebinar1.covidchallenge18march.covidspread` group by Date, Case_Recovered_Rate order by Recovered DESC
```

Below the query editor, the "Query results" section displays a table with the following data:

Row	Date	recovered_rate	Recovered
44	2022-08-04	0.9666	6021549
45	2022-08-03	0.9666	6014885
46	2022-08-02	0.9669	6010545
47	2022-08-01	0.96700000...	6005981
48	2022-07-31	0.9669	6001402
49	2022-07-30	0.96680000...	5996805
50	2022-07-29	0.9669	5992537

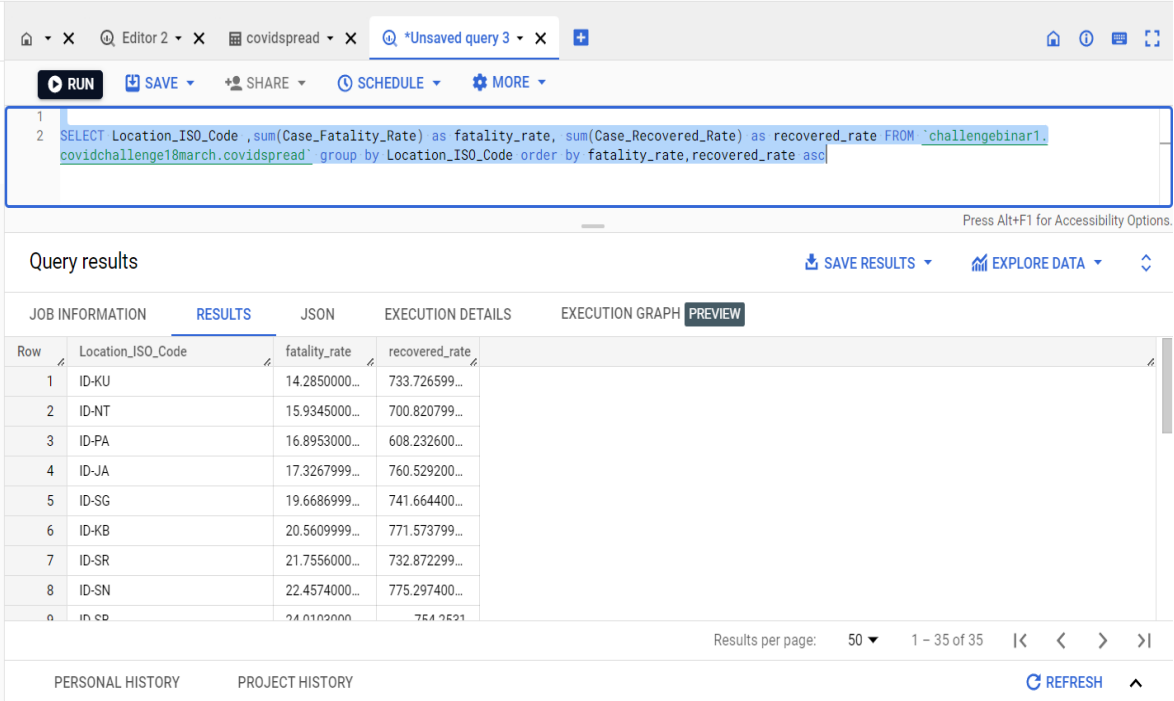
The table is part of a web application interface with tabs for "JOB INFORMATION", "RESULTS", "JSON", "EXECUTION DETAILS", "EXECUTION GRAPH", and "PREVIEW". The "RESULTS" tab is currently selected. At the bottom, there are links for "PERSONAL HISTORY", "PROJECT HISTORY", and a "REFRESH" button.

SQL

Total Case Fatality Rate

The total case fatality rate and case recovered rate of each location iso code sorted from the lowest data

```
SELECT Location_ISO_Code
,sum(Case_Fatality_Rate) as fatality_rate,
sum(Case_Recovered_Rate) as recovered_rate
FROM
`challengebinar1.covidchallenge18march.covidspr
ead` group by Location_ISO_Code order by
fatality_rate,recovered_rate asc
```



The screenshot shows a SQL query editor with the following query:

```
SELECT Location_ISO_Code ,sum(Case_Fatality_Rate) as fatality_rate, sum(Case_Recovered_Rate) as recovered_rate FROM `challengebinar1.covidchallenge18march.covidspread` group by Location_ISO_Code order by fatality_rate,recovered_rate asc
```

The query results are displayed in a table with the following columns: Row, Location_ISO_Code, fatality_rate, and recovered_rate. The results are sorted by fatality_rate and recovered_rate in ascending order.

Row	Location_ISO_Code	fatality_rate	recovered_rate
1	ID-KU	14.2850000...	733.726599...
2	ID-NT	15.9345000...	700.820799...
3	ID-PA	16.8953000...	608.232600...
4	ID-JA	17.3267999...	760.529200...
5	ID-SG	19.6686999...	741.664400...
6	ID-KB	20.5609999...	771.573799...
7	ID-SR	21.7556000...	732.872299...
8	ID-SN	22.4574000...	775.297400...

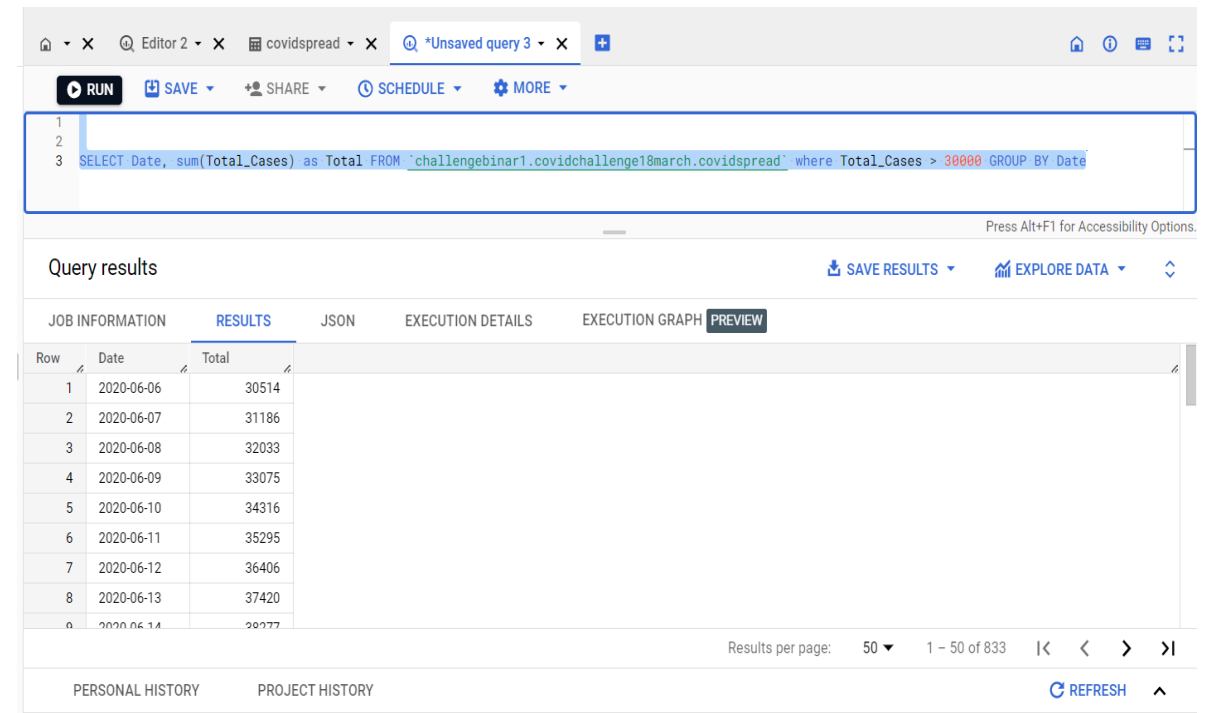
The interface also includes a 'Query results' section with options to 'SAVE RESULTS' and 'EXPLORE DATA'. At the bottom, there are tabs for 'PERSONAL HISTORY' and 'PROJECT HISTORY', and a 'REFRESH' button.

SQL

Data on Date Case Covid is Highest

Data on the dates when total Covid-19 cases began to touch the 30,000s

```
SELECT Date, sum(Total_Cases) as Total FROM
`challengebinar1.covidchallenge18march.covidspr
ead` where Total_Cases > 30000 GROUP BY
Date
```



The screenshot shows a SQL query editor with the following query:

```
SELECT Date, sum(Total_Cases) as Total FROM `challengebinar1.covidchallenge18march.covidspread` where Total_Cases > 30000 GROUP BY Date
```

The query results are displayed in a table with the following columns: Row, Date, and Total. The results show the total number of COVID-19 cases for each date from 2020-06-06 to 2020-06-14, where the total cases are greater than 30,000.

Row	Date	Total
1	2020-06-06	30514
2	2020-06-07	31186
3	2020-06-08	32033
4	2020-06-09	33075
5	2020-06-10	34316
6	2020-06-11	35295
7	2020-06-12	36406
8	2020-06-13	37420
9	2020-06-14	38977

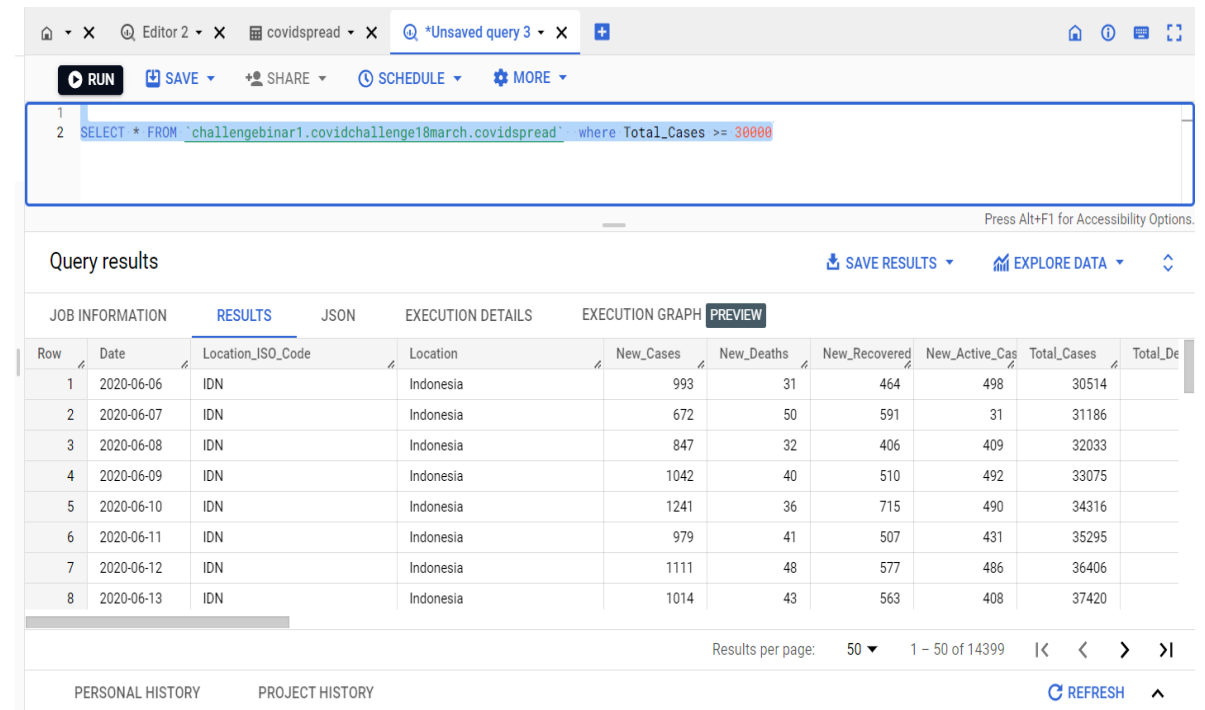
The interface also includes tabs for JOB INFORMATION, RESULTS, JSON, EXECUTION DETAILS, EXECUTION GRAPH, and PREVIEW. The RESULTS tab is currently selected. At the bottom, there are links for PERSONAL HISTORY, PROJECT HISTORY, and a REFRESH button.

SQL

Total Data Recovered ≥ 30.000

The amount of data recorded when Covid-19 cases are greater than or equal to 30,000

```
SELECT * FROM
`challengebinar1.covidchallenge18march.covidspr
ead` where Total_Cases  $\geq$  30000
```



The screenshot shows a SQL query editor with the following query:

```
SELECT * FROM `challengebinar1.covidchallenge18march.covidspread` where Total_Cases  $\geq$  30000
```

The query results are displayed in a table with the following columns: Row, Date, Location_ISO_Code, Location, New_Cases, New_Deaths, New_Recovered, New_Active_Cas, Total_Cases, and Total_Deaths. The results show data for Indonesia from June 6 to June 13, 2020.

Row	Date	Location_ISO_Code	Location	New_Cases	New_Deaths	New_Recovered	New_Active_Cas	Total_Cases	Total_Deaths
1	2020-06-06	IDN	Indonesia	993	31	464	498	30514	
2	2020-06-07	IDN	Indonesia	672	50	591	31	31186	
3	2020-06-08	IDN	Indonesia	847	32	406	409	32033	
4	2020-06-09	IDN	Indonesia	1042	40	510	492	33075	
5	2020-06-10	IDN	Indonesia	1241	36	715	490	34316	
6	2020-06-11	IDN	Indonesia	979	41	507	431	35295	
7	2020-06-12	IDN	Indonesia	1111	48	577	486	36406	
8	2020-06-13	IDN	Indonesia	1014	43	563	408	37420	

The interface also includes options to RUN, SAVE, SHARE, SCHEDULE, and MORE. The query results can be saved, explored, or previewed. The results show 14399 rows in total, with the first 8 rows displayed.

Conclusion

- The highest total number of new cases of Covid 19 in Indonesia is in West Java Province with a total of 13.496 cases
- The highest total deaths in covid 19 are located in iso code ID-MA with a total of 147,196 deaths
- As many as 30,917 data on the highest recovery rate date were recorded
- ISO Code ID-KU is recorded as the ISO Code with the lowest Total Fatality Rate and ISO Code ID-PA is recorded as the lowest recovery rate
- On Date 2020-06-06 is the first total death date to touch above 30,000 deaths
- A total of 14399 data recorded Covid 19 cases in Indonesia that exceeded the total of 30,000 cases