

# Challenge 1 Binar Academy

## Data Scientist

Nama : Yusuf Pradnya Darpita  
No. Registrasi: 149368779100-293

### 1. Google Data Studio Study Case

#### - Visualisasi Dashboard

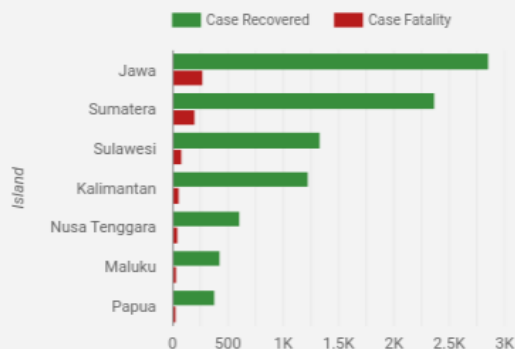
Dashboard yang dibuat dengan Google Data Studio dapat dilihat pada link di bawah ini:

<https://datastudio.google.com/reporting/c2713989-b61e-42cb-88e9-34cc07c71c10>

Screenshot Dashboard Covid 19 tahun 2020 - 2022 per provinsi:



Case Recovered and Fatality Rate By Island



Growth Factor of New Case and New Death

Island	New Case	New Death
1. Sumatera	4,279.23	2,990.36
2. Sulawesi	3,317.82	1,616.53
3. Jawa	2,803.43	2,408.4
4. Kalimantan	2,006.79	1,583.65
5. Nusa Tenggara	1,337.94	1,034.84
6. Papua	1,139.69	478.83

Location	New Case
1. DKI Jakarta	341,744
2. Jawa Barat	215,287
3. Jawa Tengah	154,555
4. Jawa Timur	129,778
5. Sulawesi Selatan	56,337
6. Kalimantan Timur	55,831
7. Bali	34,380
8. Riau	31,453
9. Banten	29,493
10. Sumatera Barat	29,204
11. Daerah Istimewa Yogyakarta	27,967
12. Sumatera Utara	24,666
13. Kalimantan Selatan	22,043
14. Papua	17,083
15. Sumatera Selatan	15,970
16. Sulawesi Utara	15,020
17. Kalimantan Tengah	13,924
18. Lampung	12,608
19. Sulawesi Tengah	10,099
20. Sulawesi Tenggara	10,054

## - Tujuan

Dashboard yang dibuat memuat data tentang Covid 19 di Indonesia pada tahun 2020 - 2022. Dashboard dibuat agar pengguna dapat memilih rentang tanggal dan provinsi yang ingin diteliti lebih lanjut. Pada bagian atas dashboard ditunjukkan total kasus, total kasus aktif, total kematian, dan total recovered. Pada bagian bawah, dibagi menjadi 2 yaitu dimensi pulau dan dimensi provinsi. Dashboard ini dibuat dengan tujuan untuk memantau kasus dan dampak yang telah ditimbulkan pada area dan waktu tertentu.

## 2. SQL Study Case

Dokumentasi kode dapat diakses pada link:

[https://github.com/jonywony/Binar\\_Study\\_Case/tree/main/Challenge%201](https://github.com/jonywony/Binar_Study_Case/tree/main/Challenge%201)

### 1. Pertanyaan 1

/\*

Pertanyaan 1 : Hitung jumlah total kasus covid aktif yang baru di setiap provinsi. Urutkan berdasarkan jumlah kasus yang paling besar.

Kolom : Province, New\_Active\_Cases

\*/

-- Memilih kolom sesuai soal

SELECT

province,

-- Menggunakan fungsi sum untuk menjumlahkan new\_active\_cases

SUM(new\_active\_cases) AS sum\_new\_active\_cases

-- Mengambil data dari tabel covid19\_indonesia

FROM

`data-scientist-academy.challenge\_1.covid19\_indonesia`

-- Memberikan batasan jika kolom province null maka tidak dihitung

WHERE

province IS NOT NULL

-- Mengelompokkan berdasarkan setiap provinsi

GROUP BY

province

-- Mengurutkan hasil query berdasarkan jumlah kasus aktif terbesar ke terkecil

ORDER BY

sum\_new\_active\_cases DESC

/\*

Hasil yang diperoleh: Bisa dilihat pada gambar di bawah ini:

\*/

Query results [SAVE RESULTS](#) [EXPLORE DATA](#)

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS
Row	province	sum_new_a...		
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		

Results per page: 50 1 - 34 of 34 [REFRESH](#)

PERSONAL HISTORY PROJECT HISTORY

Gambar 1.Pertanyaan1\_hasil.png

## 2. Pertanyaan 2

/\*

Pertanyaan 2 : Ambil 2 location iso code yang memiliki jumlah total kematian karena covid paling sedikit.

Kolom : Location\_ISO\_Code, Total\_deaths

\*/

-- Memilih kolom sesuai soal

SELECT

location\_iso\_code,

SUM(total\_deaths) AS sum\_total\_deaths

-- Mengambil data dari tabel covid19\_indonesia

FROM

`data-scientist-academy.challenge\_1.covid19\_indonesia`

-- Mengelompokkan data berdasarkan kolom location\_iso\_code

GROUP BY

location\_iso\_code

-- mengurutkan data jumlah kasus kematian berdasarkan yang terkecil ke terbesar

ORDER BY

sum\_total\_deaths ASC

-- Membatasi keluaran hanya 2 hasil saja

LIMIT 2

/\*

Hasil yang diperoleh, location\_iso\_code dengan jumlah kasus total kematian paling sedikit adalah:

1. ID-MA : 141796 kasus

2. ID-MU : 167511 kasus

\*/

Query results				<a href="#">SAVE RESULTS</a>	<a href="#">EXPLORE DATA</a>	<a href="#">↕</a>
JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		
Row	location_iso_code	sum_total_d...				
1	ID-MA	141796				
2	ID-MU	167511				
PERSONAL HISTORY		PROJECT HISTORY		<a href="#">REFRESH</a> ^		

Gambar 2.Pertanyaan2\_hasil.png

## 3. Pertanyaan 3

/\*

Pertanyaan 3 : Kapan (tanggal) Indonesia memiliki rate kasus recovered paling tinggi? Berapa ratenya?

Kolom : Date, Case\_Recovered\_Rate

\*/

-- Memilih kolom sesuai soal

SELECT

date,

MAX(case\_recovered\_rate) AS max\_case\_recovered\_rate\_IDN

-- Mengambil data dari tabel covid19\_indonesia

FROM

`data-scientist-academy.challenge\_1.covid19\_indonesia`

-- Membatasi hasil dengan hanya menampilkan kasus dimana location\_iso\_code adalah "IDN" atau hanya kasus di Indonesia saja (Negara bukan Provinsi)

WHERE

location\_iso\_code = 'IDN'

-- Mengelompokkan data berdasarkan tanggal

GROUP BY

date

-- Mengurutkan berdasarkan data terbesar ke terkecil

ORDER BY

max\_case\_recovered\_rate\_IDN DESC

-- Membatasi hasil hanya 5 karena ada 5 tanggal yang memiliki nilai tertinggi

LIMIT 5

/\*

Hasil yang diperoleh, tanggal dan rate kasus recovered paling tinggi adalah:

1. 2022-05-30 dengan rate 0.9737

2. 2022-05-29 dengan rate 0.9737

3. 2022-05-27 dengan rate 0.9737

4. 2022-05-24 dengan rate 0.9737

5. 2022-05-23 dengan rate 0.9737

\*/

Query results SAVE RESULTS EXPLORE DATA

JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS
Row	date	max_case_f...		
1	2022-05-30	0.9737		
2	2022-05-29	0.9737		
3	2022-05-27	0.9737		
4	2022-05-24	0.9737		
5	2022-05-23	0.9737		

PERSONAL HISTORY PROJECT HISTORY REFRESH

Gambar 3.Pertanyaan3\_hasil.png

#### 4. Pertanyaan 4

/\*

Pertanyaan 4 : Hitung total case fatality rate dan case recovered rate dari masing-masing location iso code. Urutkan dari yang paling rendah.

Kolom : Location\_ISO\_Code, Case\_Fatality\_Rate, Case\_Recovered\_Rate  
\*/

-- Memilih kolom sesuai soal

SELECT

location\_iso\_code,

-- Menggunakan fungsi ROUND untuk merapikan hasil

ROUND(SUM(case\_fatality\_rate), 3) AS sum\_case\_fatality\_rate,

ROUND(SUM(case\_recovered\_rate), 3) AS sum\_case\_recovered\_rate

-- Mengambil data dari tabel covid19\_indonesia

FROM

`data-scientist-academy.challenge\_1.covid19\_indonesia`

-- Mengelompokkan data berdasarkan location\_iso\_code

GROUP BY

location\_iso\_code

-- Mengurutkan berdasarkan data sum\_case\_fatality\_rate terkecil ke terbesar lalu sum\_case\_recovered\_rate terkecil ke terbesar

ORDER BY

sum\_case\_fatality\_rate ASC,

sum\_case\_recovered\_rate ASC

/\*

Hasil yang diperoleh: Dapat dilihat pada Pertanyaan\_4\_Hasil.png

\*/

Query results				SAVE RESULTS	EXPLORE DATA	
JOB INFORMATION		RESULTS	JSON	EXECUTION DETAILS		
Row	location_iso_code	sum_case_fatality_rate	sum_case_recovered_rate			
1	ID-KU	14.285	733.727			
2	ID-NT	15.935	700.821			
3	ID-PA	16.895	608.233			
4	ID-JA	17.327	760.529			
5	ID-SG	19.669	741.664			
6	ID-KB	20.561	771.574			
7	ID-SR	21.756	732.872			
8	ID-SN	22.457	775.297			

Results per page: 50 1 - 35 of 35 |< < > >|

PERSONAL HISTORY PROJECT HISTORY REFRESH ^

Gambar 4.Pertanyaan4\_hasil.png

## 5. Pertanyaan 5

/\*

Pertanyaan 5 : Pada tanggal berapa total kasus covid mulai menyentuh angka 30.000-an? Hitung ada berapa data yang tercatat ketika kasus covid lebih dari atau sama dengan 30.000?

Untuk menjawab pertanyaan ini, diperlukan 2 query karen aitu saya membaginya menjadi 5.1 dan 5.2

\*/

/\*

Pertanyaan 5.1 : Pada tanggal berapa total kasus covid mulai menyentuh angka 30.000-an?

Disini saya berasumsi bahwa pertanyaan mengacu pada seluruh lokasi

Kolom : Date

\*/

-- Memilih kolom sesuai soal

SELECT

date

-- Mengambil data dari tabel covid19\_indonesia

FROM

`data-scientist-academy.challenge\_1.covid19\_indonesia`

-- Hanya tanggal yang memiliki total\_cases lebih dari 30000 yang ditampilkan

WHERE

total\_cases >= 30000

-- Mengurutkan tanggal dari yang terkecil ke terbesar

ORDER BY

date ASC

-- Tanggal pertama kali total kasus mencapai 30000 atau lebih adalah dengan mengambil nilai pertama dari sorting yang dilakukan  
LIMIT 1

/\*

Hasil yang diperoleh:

Tanggal pertama saat kasus covid mencapai 30000 adalah tanggal 2020-06-06

\*/

/\*

Pertanyaan 5.2 : Hitung ada berapa data yang tercatat ketika kasus covid lebih dari atau sama dengan 30.000?

Disini saya berasumsi bahwa pertanyaan mengacu pada seluruh lokasi

Kolom : Total\_Cases

\*/

-- Memilih kolom sesuai soal

SELECT

COUNT(total\_cases) AS count\_total\_cases\_more\_than\_equal\_30000

-- Mengambil data dari tabel covid19\_indonesia

FROM

`data-scientist-academy.challenge\_1.covid19\_indonesia`

-- Hanya tanggal yang memiliki total\_cases lebih dari 30000 yang ditampilkan

WHERE

total\_cases >= 30000

/\*

Hasil yang diperoleh:

data yang tercatat ketika kasus covid lebih dari atau sama dengan 30.000 adalah 14399

\*/

Query results

SAVE RESULTS

EXPLORE DATA

JOB INFORMATION

RESULTS

JSON

EXECUTION DETAILS

Row	date	
1	2020-06-06	

PERSONAL HISTORY


PROJECT HISTORY


REFRESH


Gambar 5.Pertanyaan5.1\_hasil.png



Query results

 SAVE RESULTS ▾

 EXPLORE DATA ▾



JOB INFORMATION

RESULTS


JSON


EXECUTION DETAILS

Row	count_total_...
1	14399

PERSONAL HISTORY

PROJECT HISTORY

 REFRESH



Gambar 6.Pertanyaan5.2\_hasil.png