Dokumen Teknis

Capstone Project

Data And Artificial Intelligence

Dhea Amalia Ariantoputri

DAI 002

Mentor :
Noviyanti Tri Maretta Sagala, S.TI., M.Sc

Asisten Mentor :
Ainul Zakiy

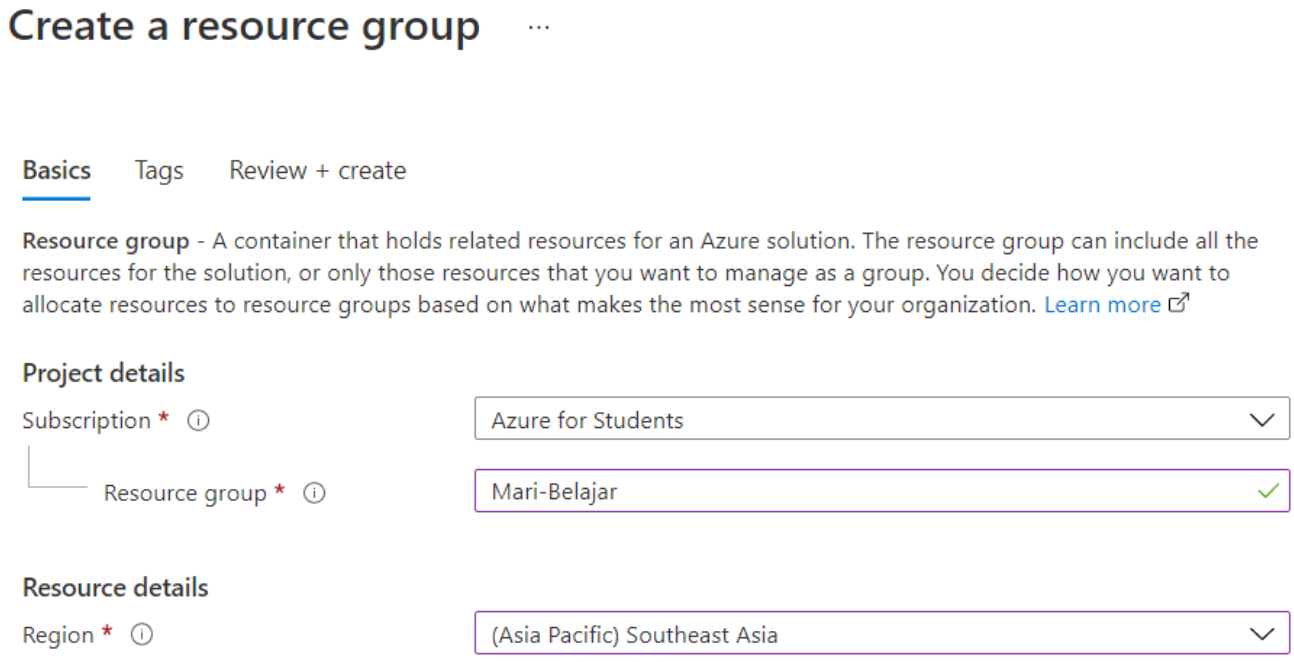
STUDI INDEPENDEN PT. MICROSOFT INDONESIA
NOVEMBER 2021

Latar Belakang

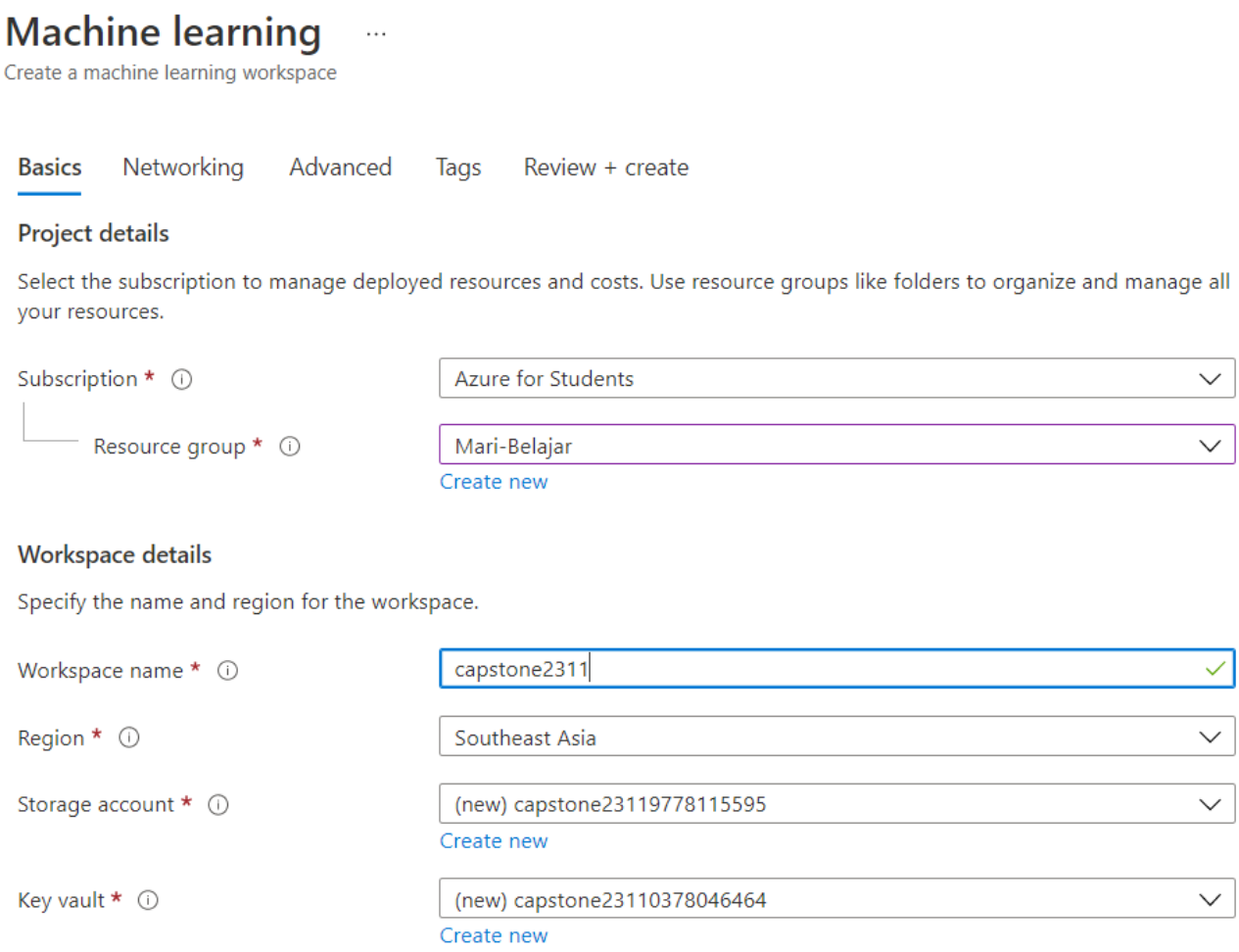
Kasus Covid-19 di Indonesia sudah memasuki tahun kedua. Klinik Mari Sehat ingin membuat sebuah dashboard yang akan ditampilkan di ruang depan klinik agar dapat dibaca oleh pasien. Dataset yang digunakan yaitu MariShat.

Pre-Processing Data

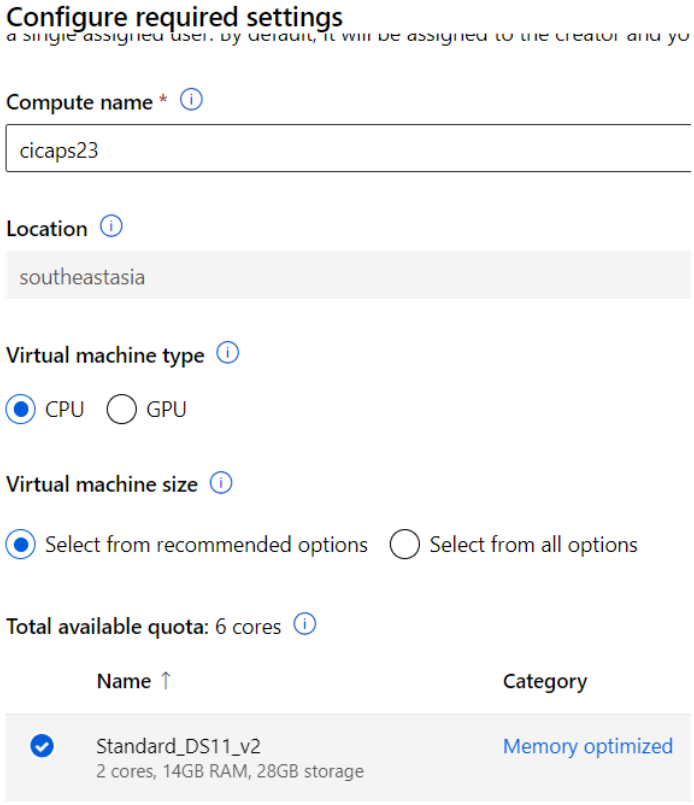
1. Pastikan memiliki akun Azure terlebih dahulu agar dapat membuat resource group.
2. Buat resource group



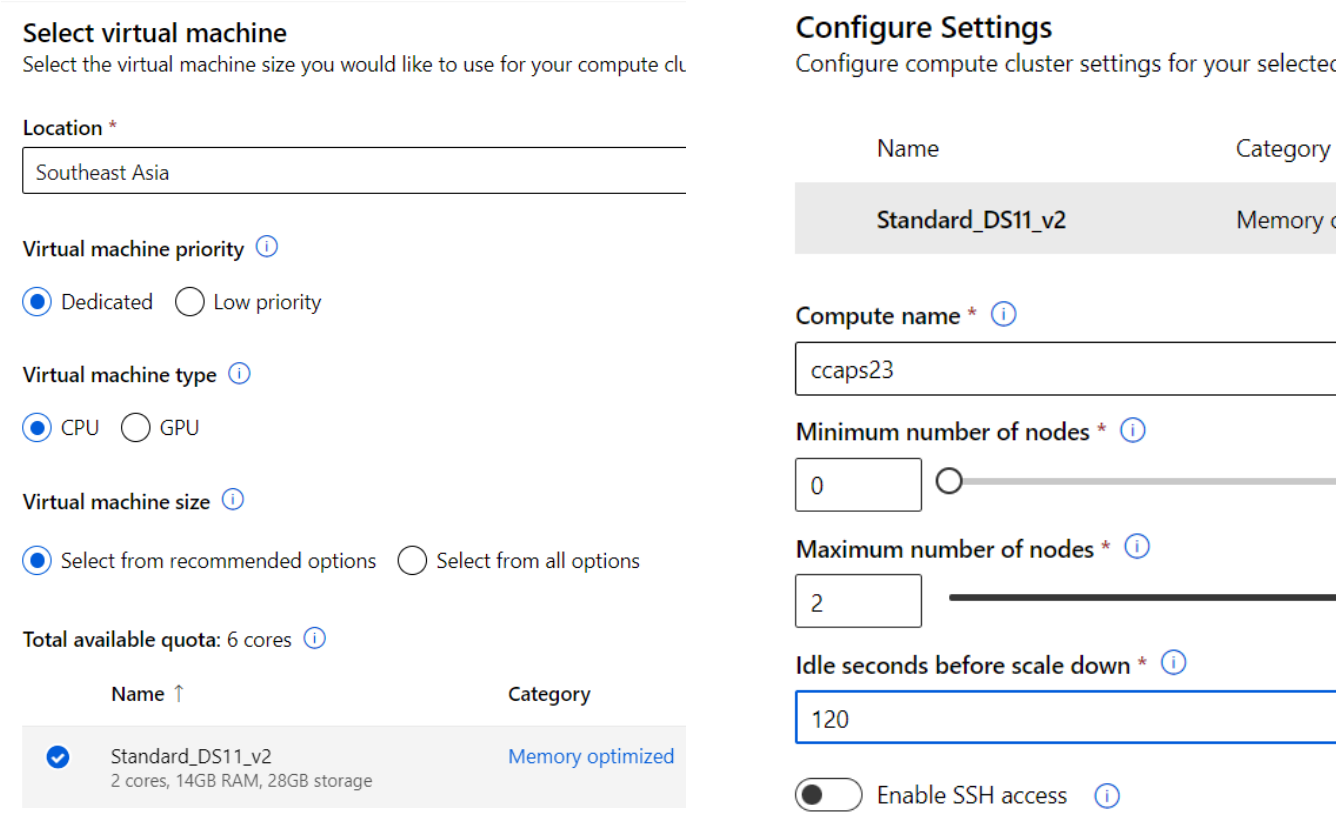
3. Membuat machine learning workspace. Setelah selesai, klik lauch studio



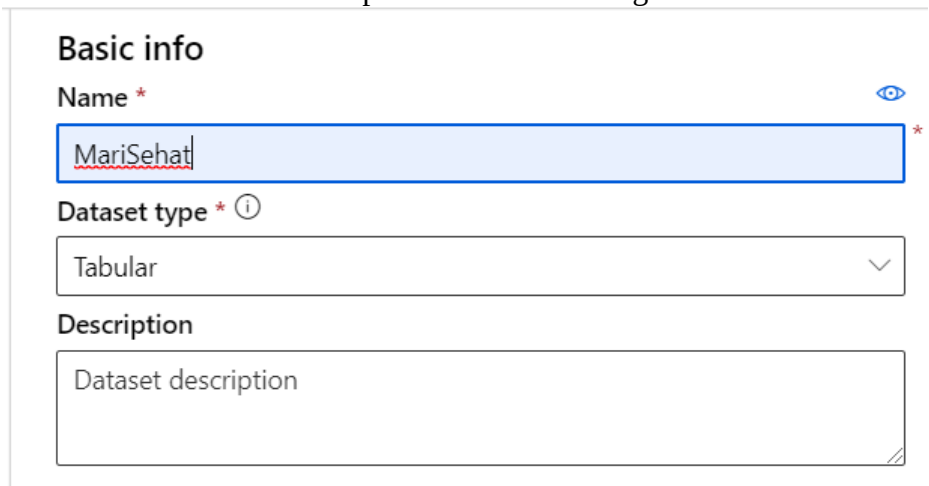
4. Pilih menu Compute pada sidebar. Buat Compute Instances

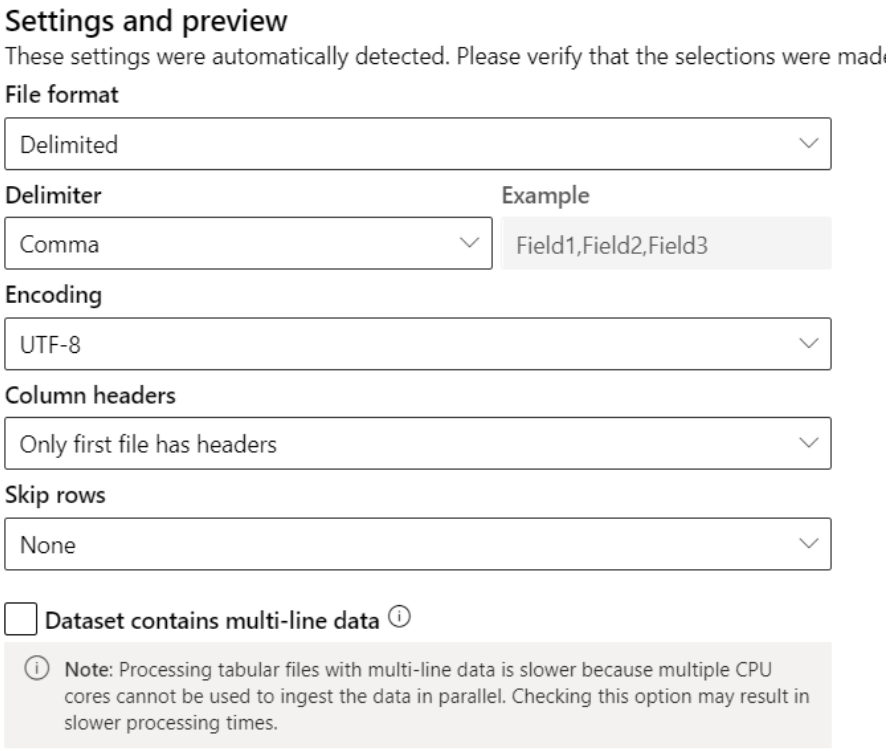


Buat Compute clusters

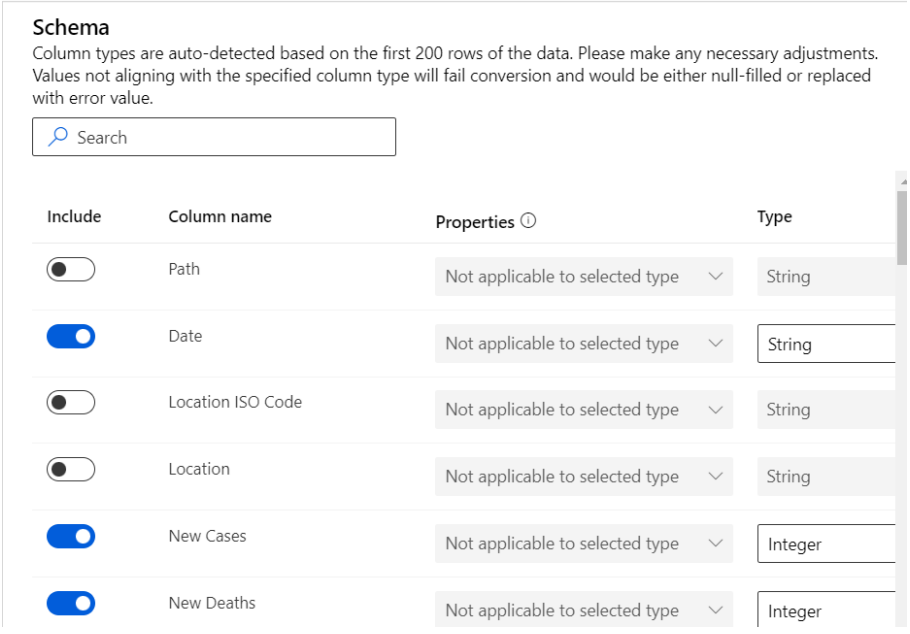


5. Pilih menu Datasets pada sidebar. Pilih register dataset from local files



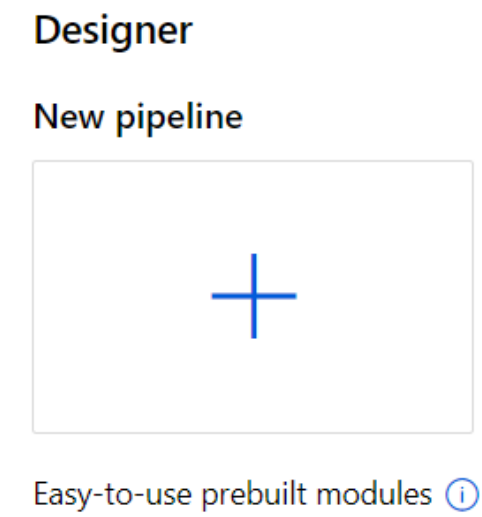


Pada schema, pilih kolom mana yang memang dibutuhkan untuk proses training dan testing nanti. Disini saya memilih Date, New Cases, New Deaths, New Recovered, New Active Cases, Total Cases, Total Deaths, Total Recovered, Total Active Cases, Province, Case Fatality Rate, Case Recovered Rate

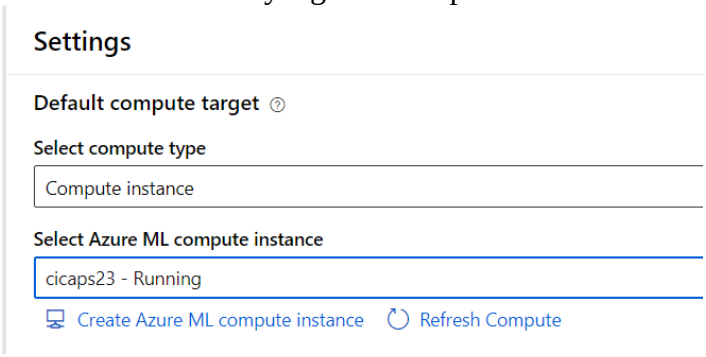


Setelah itu klik create.

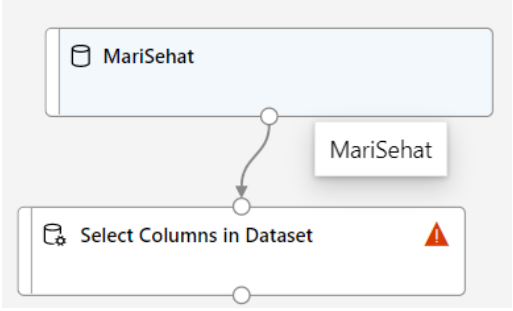
6. Pilih Designer pada sidebar. Pilih new pipeline



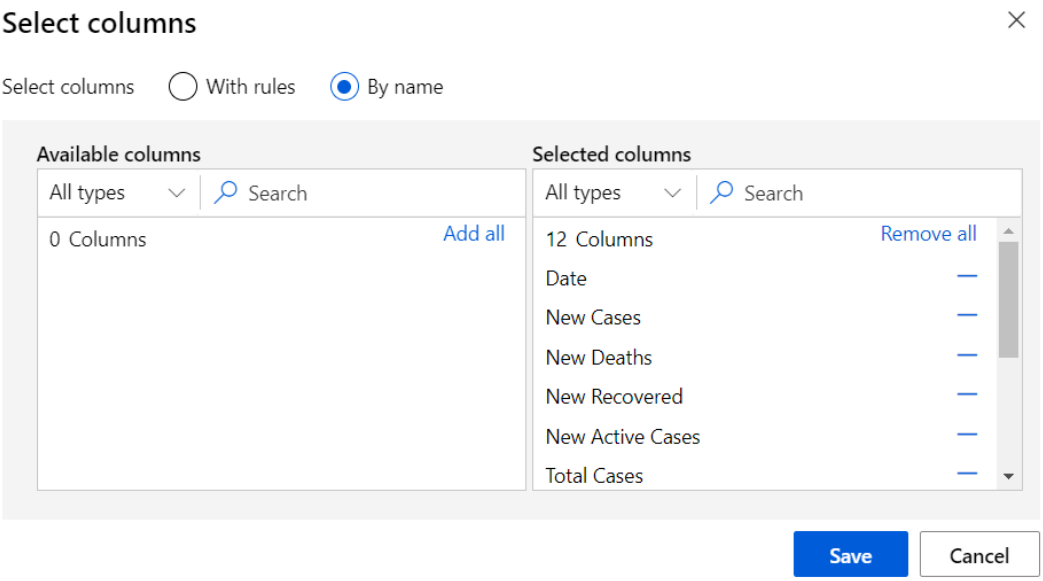
7. Pilih dataset yang sudah di proses sebelumna. Jangan lupa untuk memilih compute instance.



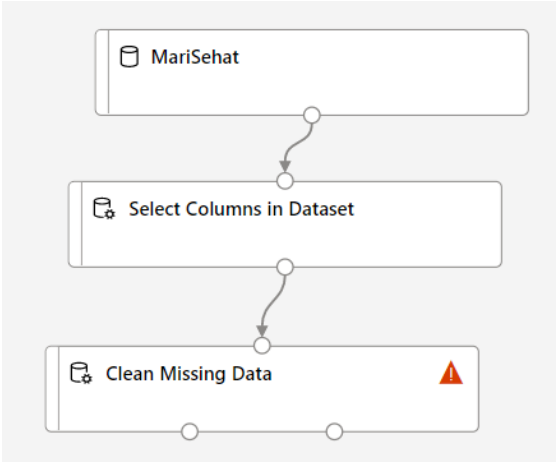
8. Pilih modul select columns in dataset lalu hubungkan dengan dataset.



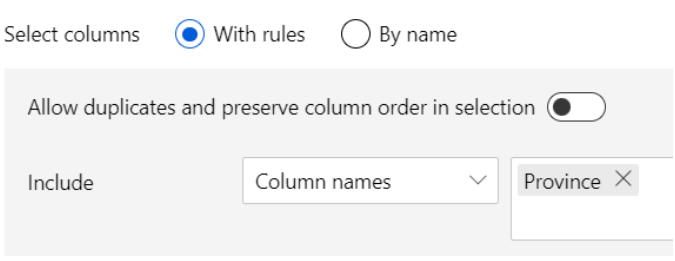
klik select columns in dataset, klik edit column



Tambahkan modul cleaning missing data



Karena pada kolom province ada yang kosong.



Untuk cleaning mode pilih remove entire row

Clean Missing Data

Columns to be cleaned *

Column names: Province

Minimum missing value ratio *

0.0

Maximum missing value ratio *

1.0

Cleaning mode *

Remove entire row

9. Lalu klik submit. Set pipeline terlebih dahulu.

Set up pipeline run

Experiment

Select existing Create new

New experiment name *

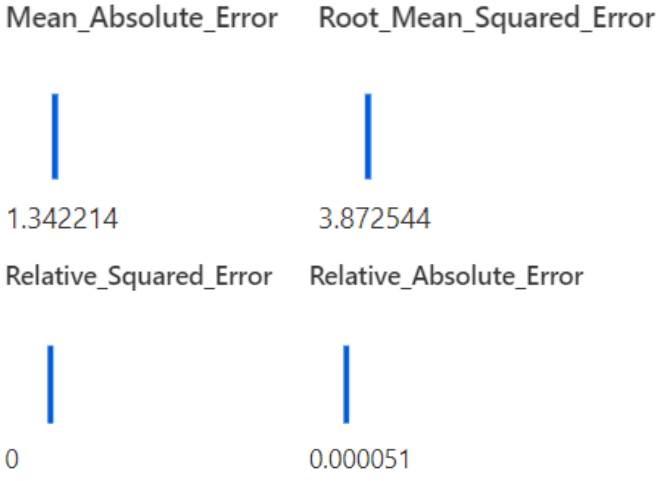
training-caps

Run description

Pipeline created on 20211123

Continue on failure step

Setelah berhasil di run, kita dapat melihat visualisasi data di score model dan evaluate model
Berikut hasil evaluate model

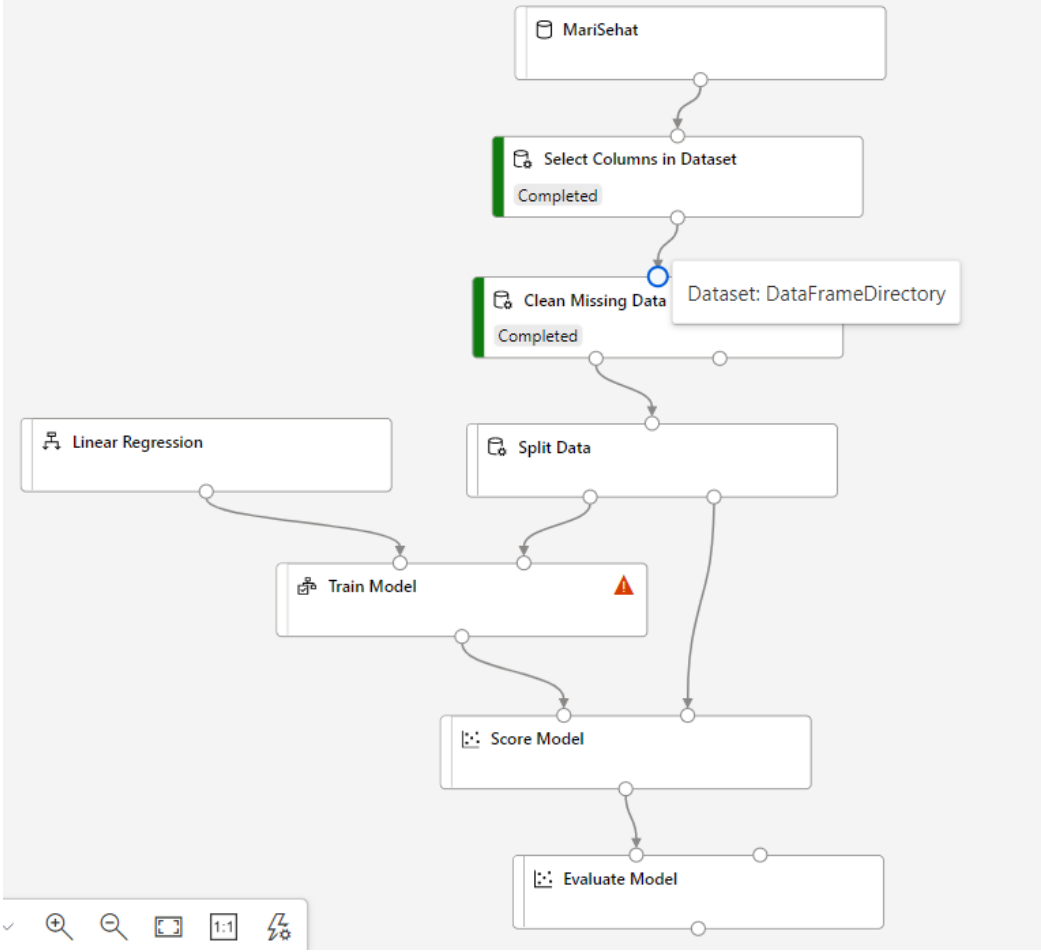


Coefficient_of_Determinatio
n

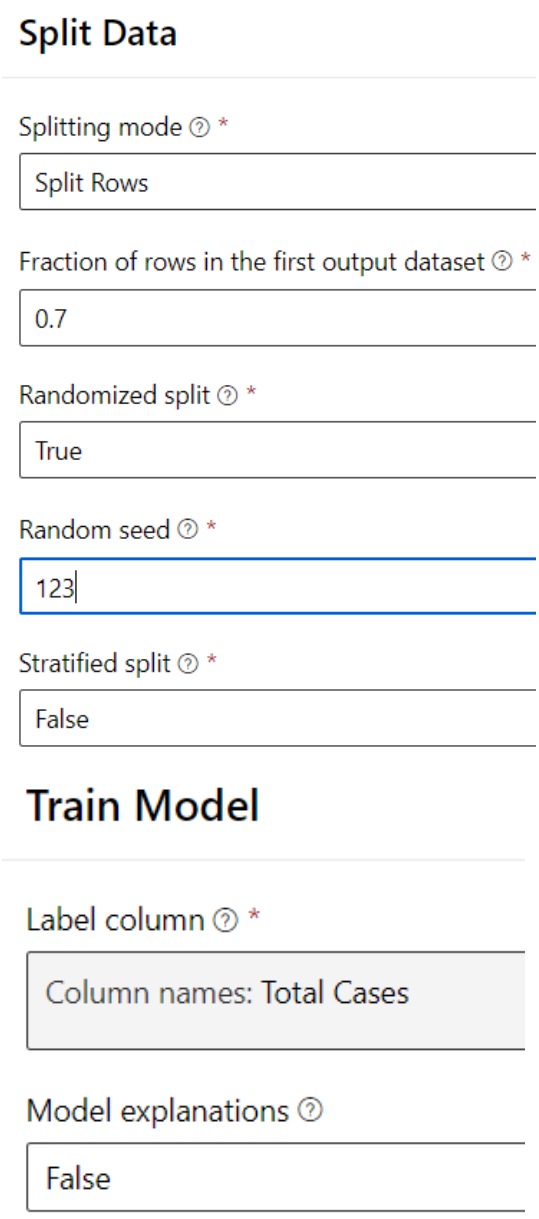
1

Output di atas menunjukkan bahwa coefficient of determination, atau R-squared adalah 100%, yang mana ini merupakan kinerja model yang baik.

10. Tambahkan beberapa modul seperti pada gambar

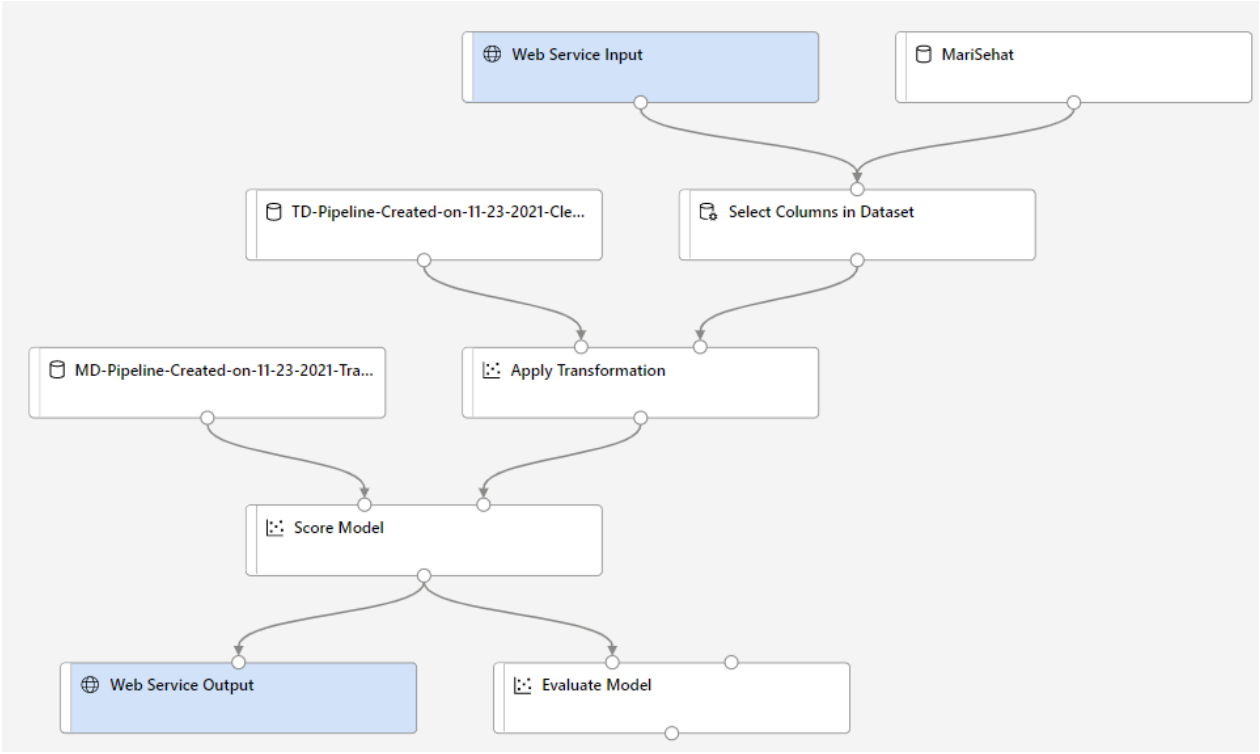


Berikut setting yang digunakan pada model ini



Setelah semua di setting dengan benar, klik submit dan pilih select existing, gunakan experiment yang sebelumnya dibuat.

11. Buat inference pipeline. Pilih real-time inference pipeline, maka akan otomatis terbentuk



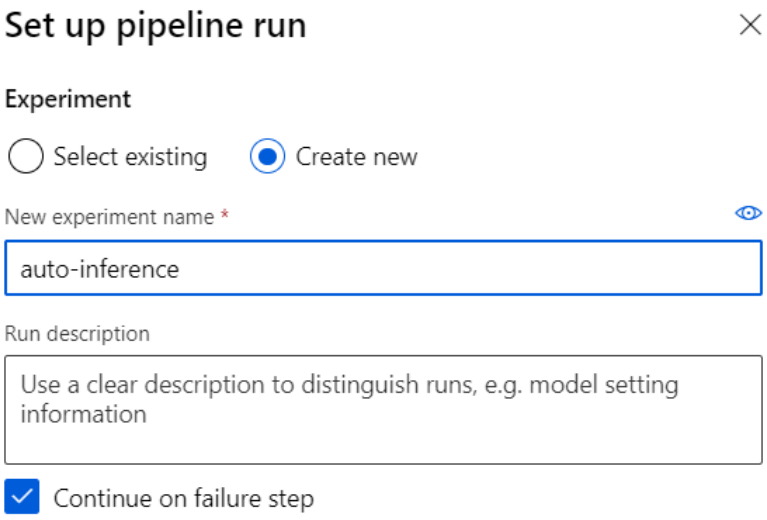
Dari gambar diatas, ada beberapa modul yang dihilangkan seperti, MariSehat dan evaluate model. Untuk MariSehat diganti dengan enter data manually dan pada modul select column in dataset diedit Total Cases dihilangkan karena disini akan memprediksi total kasus di waktu tertentu.

Di modul enter data manually, kita masukkan data baru untuk diprediksi
Date,New Cases,New Deaths,New Recovered,New Active Cases,Total Deaths,Total Recovered,Total Active Cases,Province,Case Fatality Rate,Case Recovered Rate
2021/08/01 00:00:00,140,7,95,39,1016,17631,4595,Aceh,0.04,0.76
2021/08/01 00:00:00,98,2,35,62,1662,46951,7644,Bali,0.03,0.95
2021/08/01 00:00:00,255,5,97,154,1567,52688,13066,Banten,0.02,0.92
2021/08/01 00:00:00,69,1,56,12,216,10762,2211,Bengkulu,0.02,0.89
2021/08/01 00:00:00,524,13,236,275,2178,59544,22366,Daerah Istimewa Yogyakarta,0.03,0.88
2021/08/01 00:00:00,3780,40,1753,1986,10150,565626,143745,DKI Jakarta,0.02,0.93
2021/08/01 00:00:00,108,2,106,1,328,14224,1781,Jambi,0.02,0.81
2021/08/01 00:00:00,2250,39,1485,726,6866,381364,101649,Jawa Barat,0.01,0.9
2021/08/01 00:00:00,1792,45,889,857,13552,243264,67934,Jawa Tengah,0.04,0.85
2021/08/01 00:00:00,608,41,329,236,14537,167534,20966,Jawa Timur,0.07,0.89
2021/08/01 00:00:00,127,5,109,11,523,16769,2269,Kalimantan Barat,0.01,0.9
2021/08/01 00:00:00,43,1,44,1,1113,35742,1333,Kalimantan Selatan,0.03,0.95
2021/08/01 00:00:00,105,1,43,61,578,22028,7512,Kalimantan Tengah,0.02,0.83
2021/08/01 00:00:00,371,8,137,225,2179,76878,12845,Kalimantan Timur,0.02,0.93
2021/08/01 00:00:00,60,1,21,39,232,12780,2604,Kalimantan Utara,0.02,0.92
2021/08/01 00:00:00,125,2,99,22,409,20959,2106,Kepulauan Bangka Belitung,0.02,0.93
2021/08/01 00:00:00,380,8,314,57,815,30841,7319,Kepulauan Riau,0.02,0.8
2021/08/01 00:00:00,176,6,99,71,1281,19382,5607,Lampung,0.05,0.85
2021/08/01 00:00:00,85,1,9,74,186,9336,4221,Maluku,0.02,0.89
2021/08/01 00:00:00,63,1,10,51,164,4870,2856,Maluku Utara,0.02,0.88
2021/08/01 00:00:00,31,1,65,-34,520,12888,1510,Nusa Tenggara Barat,0.04,0.81
2021/08/01 00:00:00,192,1,49,141,515,18466,8154,Nusa Tenggara Timur,0.02,0.89
2021/08/01 00:00:00,32,1,7,25,235,11850,10144,Papua,0.01,0.55
2021/08/01 00:00:00,118,1,21,96,221,10291,5171,Papua Barat,0.02,0.9
2021/08/01 00:00:00,357,10,348,-2,2239,75642,4894,Riau,0.03,0.92
2021/08/01 00:00:00,21,1,7,14,149,5749,837,Sulawesi Barat,0.02,0.94
2021/08/01 00:00:00,154,2,53,99,1065,63837,5755,Sulawesi Selatan,0.02,0.96

2021/08/01 00:00:00,52,1,17,33,445,13362,1984,Sulawesi Tengah,0.03,0.93
2021/08/01 00:00:00,61,1,9,50,274,10610,2883,Sulawesi Tenggara,0.02,0.91
2021/08/01 00:00:00,159,2,21,135,616,15987,4297,Sulawesi Utara,0.03,0.91
2021/08/01 00:00:00,509,8,206,294,1435,53372,12044,Sumatera Barat,0.02,0.89
2021/08/01 00:00:00,284,11,158,115,1811,30564,5020,Sumatera Selatan,0.05,0.87
2021/08/01 00:00:00,238,5,203,29,1342,38239,3673,Sumatera Utara,0.03,0.89

Angka diatas saya ambil berdasarkan keadaan kasus covid sebulan ke belakang lalu saya rata-rata dan saya jadikan sebagai kasus harian, lalu saya kalikan selama 22 hari kedepan, karena tanggal terakhir di dataset menunjukkan tanggal 9 Juli 2021. Saya ingin memprediksi bagaimana jumlah kasus saat tanggal 1 Agustus 2021 pukul 00:00:00

12. Submit inference pipeline yang sudah dibuat dan pilih create new, beri nama pipeline

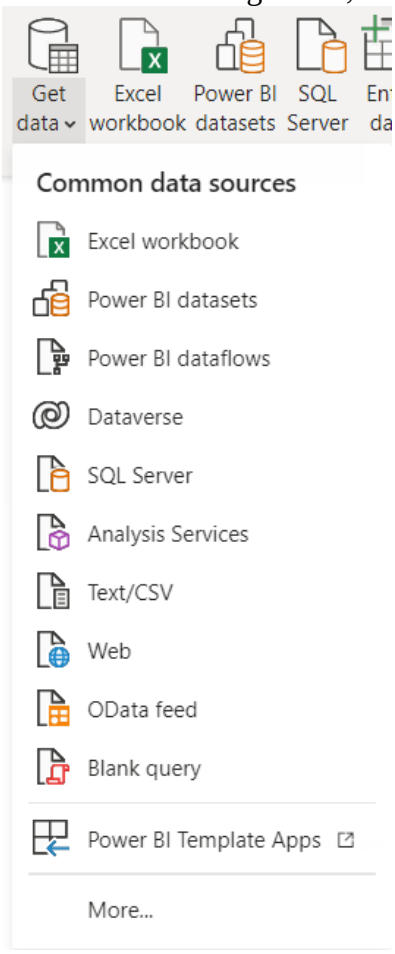


13. Tambahkan modul convert to CSV agar dataset bisa di download

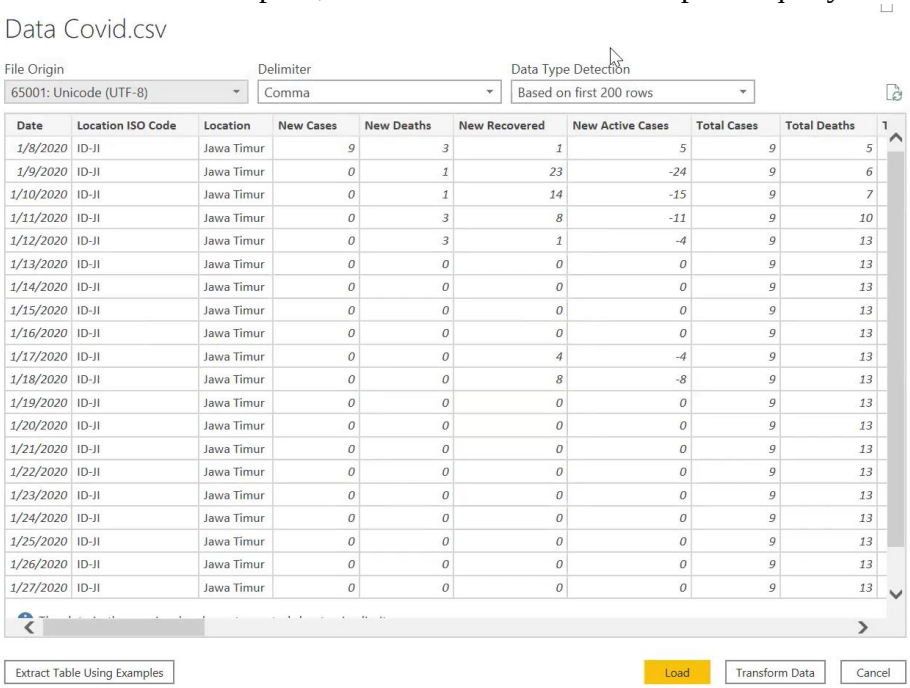
Hasil dari testing yang akan divisualisasikan menjadi dashboard di power BI

Getting Data di Power BI

1. Pastikan telah mendownload power BI desktop dan login terlebih dahulu.
2. Pilih menu get data, lalu pilih text/CSV

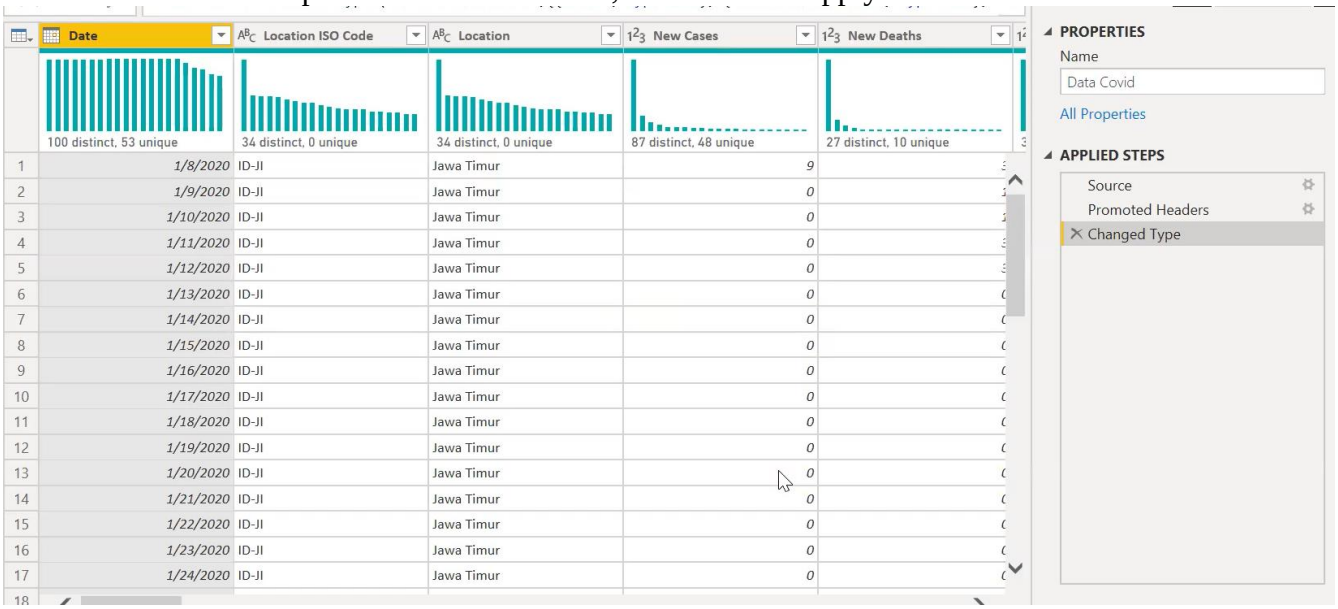


3. Setelah file dipilih, maka nanti akan masuk ke power query. Lalu klik transform



Date	Location ISO Code	Location	New Cases	New Deaths	New Recovered	New Active Cases	Total Cases	Total Deaths
1/8/2020	ID-JI	Jawa Timur	9	3	1	5	9	5
1/9/2020	ID-JI	Jawa Timur	0	1	23	-24	9	6
1/10/2020	ID-JI	Jawa Timur	0	1	14	-15	9	7
1/11/2020	ID-JI	Jawa Timur	0	3	8	-11	9	10
1/12/2020	ID-JI	Jawa Timur	0	3	1	-4	9	13
1/13/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/14/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/15/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/16/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/17/2020	ID-JI	Jawa Timur	0	0	4	-4	9	13
1/18/2020	ID-JI	Jawa Timur	0	0	8	-8	9	13
1/19/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/20/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/21/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/22/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/23/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/24/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/25/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/26/2020	ID-JI	Jawa Timur	0	0	0	0	9	13
1/27/2020	ID-JI	Jawa Timur	0	0	0	0	9	13

4. Transform data ini berguna memilah kolom mana yang akan divisualisasikan menjadi report dan dasboard. Disini juga dapat melihat perubahan apa saja yang terjadi pada datanya melalui applied steps. Kita juga dapat mengubah tipe data yang memang sesuai dengan datanya. Setelah dirasa sudah siap untuk divisualisasikan, klik Close & Apply

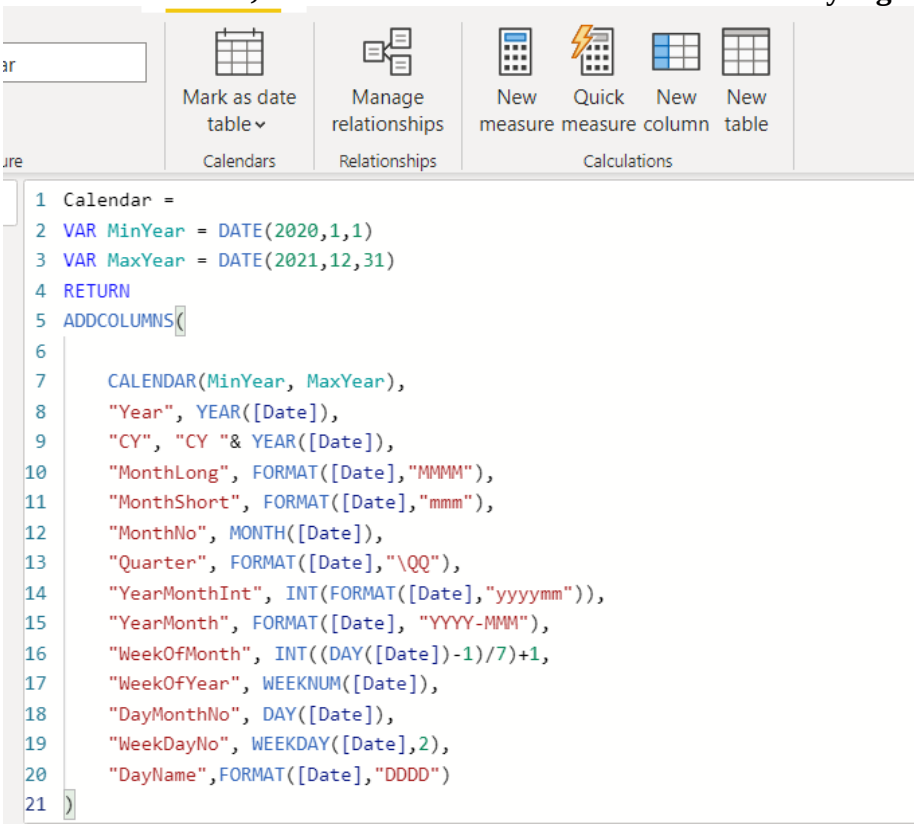


	Date	Location ISO Code	Location	New Cases	New Deaths
	100 distinct, 53 unique	34 distinct, 0 unique	34 distinct, 0 unique	87 distinct, 48 unique	27 distinct, 10 unique
1	1/8/2020	ID-JI	Jawa Timur	9	
2	1/9/2020	ID-JI	Jawa Timur	0	
3	1/10/2020	ID-JI	Jawa Timur	0	
4	1/11/2020	ID-JI	Jawa Timur	0	
5	1/12/2020	ID-JI	Jawa Timur	0	
6	1/13/2020	ID-JI	Jawa Timur	0	
7	1/14/2020	ID-JI	Jawa Timur	0	
8	1/15/2020	ID-JI	Jawa Timur	0	
9	1/16/2020	ID-JI	Jawa Timur	0	
10	1/17/2020	ID-JI	Jawa Timur	0	
11	1/18/2020	ID-JI	Jawa Timur	0	
12	1/19/2020	ID-JI	Jawa Timur	0	
13	1/20/2020	ID-JI	Jawa Timur	0	
14	1/21/2020	ID-JI	Jawa Timur	0	
15	1/22/2020	ID-JI	Jawa Timur	0	
16	1/23/2020	ID-JI	Jawa Timur	0	
17	1/24/2020	ID-JI	Jawa Timur	0	

PROPERTIES
Name
Data Covid
All Properties
APPLIED STEPS
Source
Promoted Headers
Changed Type

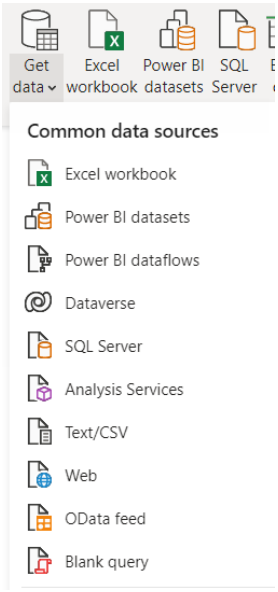
5. Di power BI juga dapat menambah,mengedit,menghapus query tertentu di DAX. Contoh calendar. Berikut caranya :

Klik new table, masukkan code tersebut ke dalam kotak yang tersedia




```
Calendar =
VAR MinYear = DATE(2020,1,1)
VAR MaxYear = DATE(2021,12,31)
RETURN
ADDCOLUMNS(

    CALENDAR(MinYear, MaxYear),
    "Year", YEAR([Date]),
    "CY", "CY "& YEAR([Date]),
    "MonthLong", FORMAT([Date],"MMMM"),
    "MonthShort", FORMAT([Date],"mmm"),
    "MonthNo", MONTH([Date]),
    "Quarter", FORMAT([Date],"\QQ"),
    "YearMonthInt", INT(FORMAT([Date],"yyyymm")),
    "YearMonth", FORMAT([Date], "YYYY-MMM"),
    "WeekOfMonth", INT((DAY([Date])-1)/7)+1,
    "WeekOfYear", WEEKNUM([Date]),
    "DayMonthNo", DAY([Date]),
    "WeekDayNo", WEEKDAY([Date],2),
    "DayName",FORMAT([Date],"DDDD")
)
```

Atau juga bisa dengan membuat blank query



Masukkan script ini kedalamnya, jangan lupa untuk mengganti StartYear dan EndYear dan lainnya. Source ini banyak terdapat di google

```
let
  StartDate = #date(StartYear,1,1),
  EndDate = #date(EndYear,12,31),
  NumberOfDays = Duration.Days( EndDate - StartDate ),
  Dates = List.Dates(StartDate, NumberOfDays+1, #duration(1,0,0,0)),
  #"Converted to Table" = Table.FromList(Dates, Splitter.SplitByNothing(), null, null, ExtraValues.Error),
  #"Renamed Columns" = Table.RenameColumns(#"Converted to Table",{{"Column1", "FullDateAlternateKey"}}),
  #"Changed Type" = Table.TransformColumnTypes(#"Renamed Columns",{{"FullDateAlternateKey", type date}}),
  #"Inserted Year" = Table.AddColumn(#"Changed Type", "Year", each Date.Year([FullDateAlternateKey]), type number),
  #"Inserted Month" = Table.AddColumn(#"Inserted Year", "Month", each Date.Month([FullDateAlternateKey]), type number),
  #"Inserted Month Name" = Table.AddColumn(#"Inserted Month", "Month Name", each Date.MonthName([FullDateAlternateKey]), type text),
  #"Inserted Quarter" = Table.AddColumn(#"Inserted Month Name", "Quarter", each Date.QuarterOfYear([FullDateAlternateKey]), type number),
  #"Inserted Week of Year" = Table.AddColumn(#"Inserted Quarter", "Week of Year", each Date.WeekOfYear([FullDateAlternateKey]), type number),
  #"Inserted Week of Month" = Table.AddColumn(#"Inserted Week of Year", "Week of Month", each Date.WeekOfMonth([FullDateAlternateKey]), type number),
  #"Inserted Day" = Table.AddColumn(#"Inserted Week of Month", "Day", each Date.Day([FullDateAlternateKey]), type number),
  #"Inserted Day of Week" = Table.AddColumn(#"Inserted Day", "Day of Week", each Date.DayOfWeek([FullDateAlternateKey]), type number),
  #"Inserted Day of Year" = Table.AddColumn(#"Inserted Day of Week", "Day of Year", each Date.DayOfYear([FullDateAlternateKey]), type number),
  #"Inserted Day Name" = Table.AddColumn(#"Inserted Day of Year", "Day Name", each Date.DayOfWeekName([FullDateAlternateKey]), type text),
  #"Added Custom" = Table.AddColumn(#"Inserted Day Name", "Fiscal Year", each if [Month]>=StartOfFiscalYear
then [Year]+1
else [Year]),
  #"Changed Type1" = Table.TransformColumnTypes(#"Added Custom",{{"Fiscal Year", Int64.Type}}),
  #"Added Custom1" = Table.AddColumn(#"Changed Type1", "Fiscal Period", each if [Month]>=StartOfFiscalYear
then [Month]-(StartOfFiscalYear-1)
else [Month]+(12-StartOfFiscalYear+1)),
  #"Changed Type2" = Table.TransformColumnTypes(#"Added Custom1",{{"Fiscal Period", Int64.Type}}),
  #"Inserted Division" = Table.AddColumn(#"Changed Type2", "Inserted Division", each [Fiscal Period] / 3, type number),
  #"Rounded Up" = Table.TransformColumns(#"Inserted Division",{{"Inserted Division", Number.RoundUp, Int64.Type}}),
  #"Renamed Columns1" = Table.RenameColumns(#"Rounded Up",{{"Inserted Division", "Fiscal Quarter"}})
in
  #"Renamed Columns1"
```

Selanjutnya di power BI juga bisa menambahkan query seperti dibawah ini, saya beri nama measure

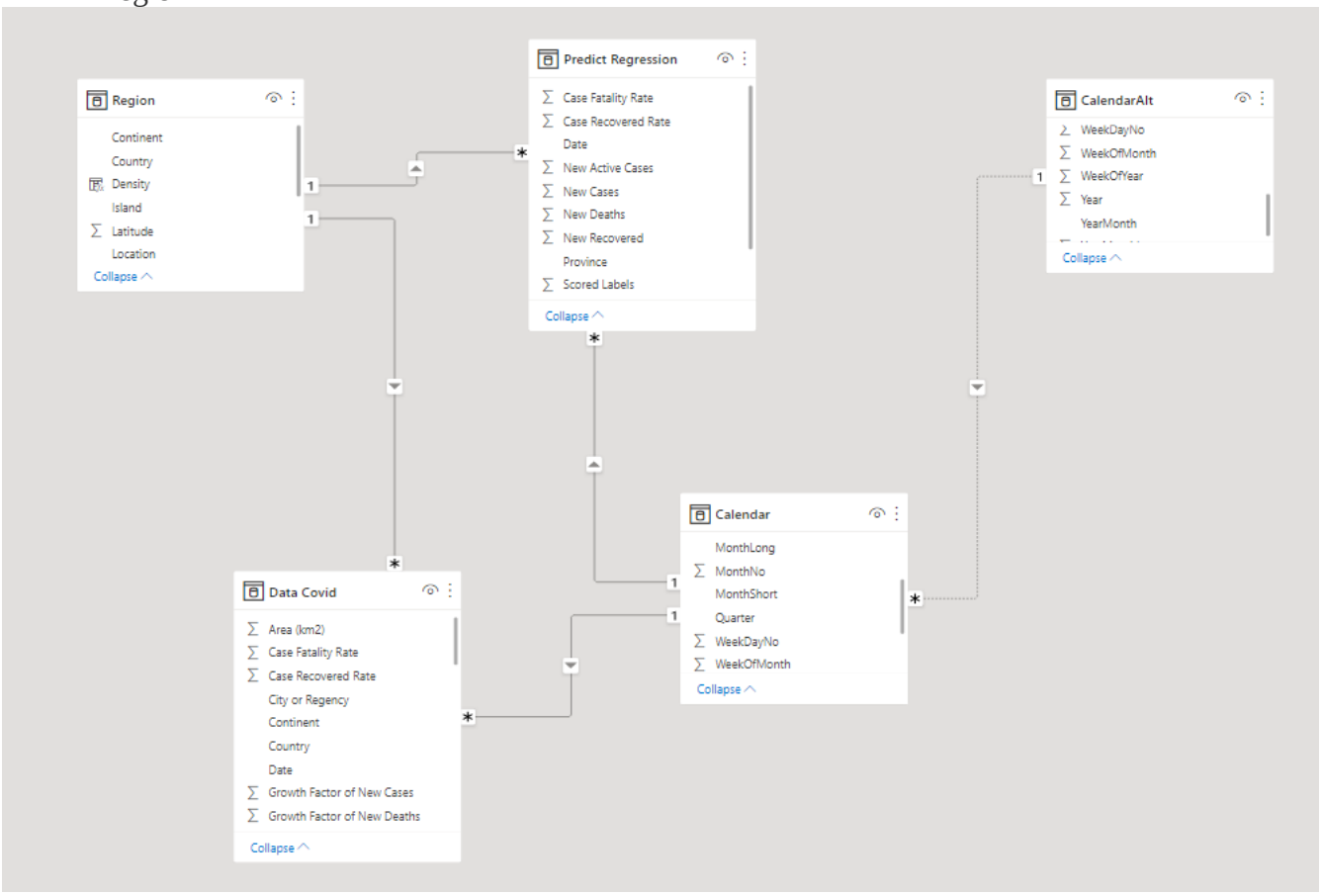
```
New Cases =
SUM('Data Covid'[New Cases])
```

```
Delta New Cases =
VAR _n = [New Cases Cumulative]
VAR _n_1 = [New Cases Cumulative H-1]
VAR _delta = _n - _n_1
VAR result = DIVIDE(_delta, _n_1)
RETURN
    _delta
```

```
Growth Act vs Predict =
VAR _Act = [Select Measure]
VAR _Pred = [Projection Cases]
VAR _delta = [Projection Cases] - [Select Measure]
RETURN
    DIVIDE(_Pred, _Act)-1
```

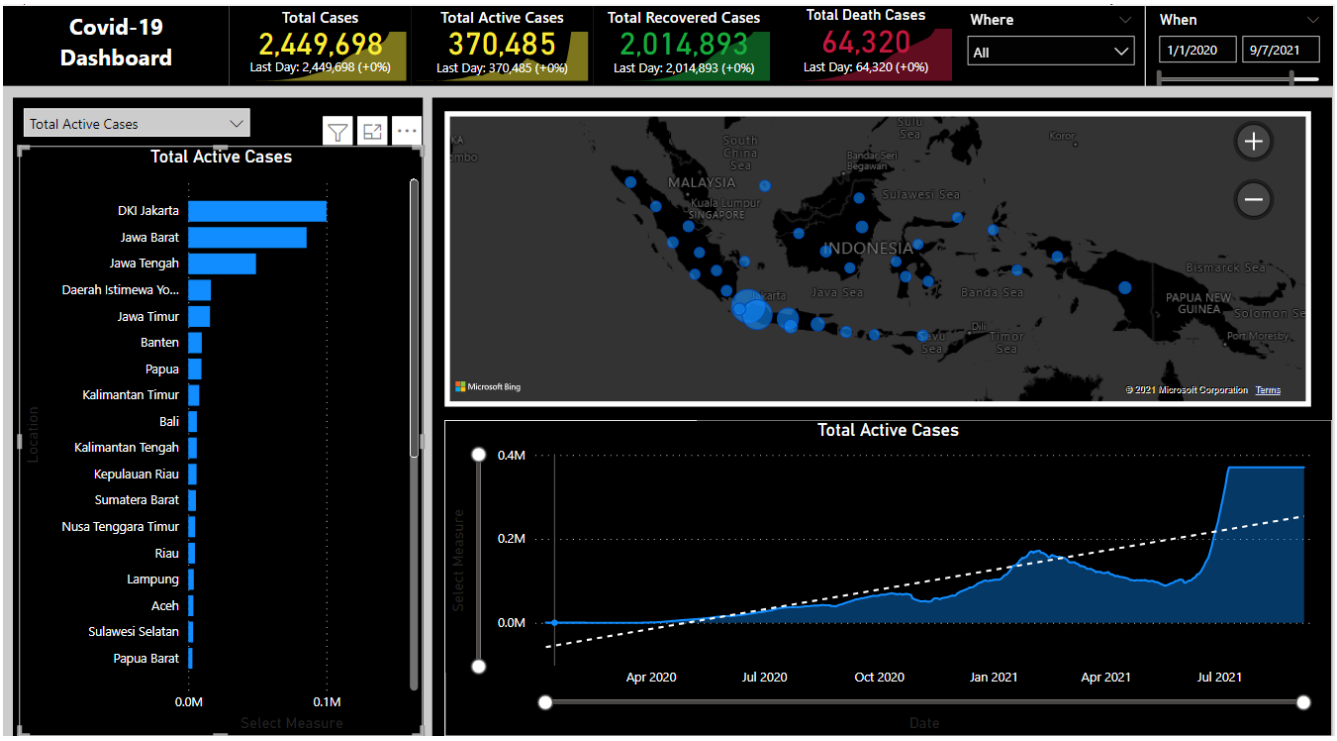
Dan lain-lain

6. Berikut hasil model data yang telah di relasikan dengan tabel lain. Di power BI terdapat dua jenis tabel, yaitu Fact tabel (tabel fakta) dan dimension table (tabel dimensi). Dimana tabel fakta berisi kunci utama yang merupakan gabungan dari kunci utama dari semua tabel dimensi dan setiap tabel dimensi berisi kunci utama. Yang termasuk fact tabel yaitu Calendar, Calender Alt, Region

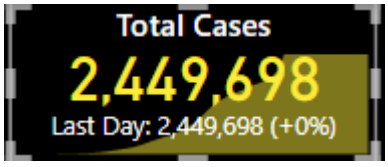


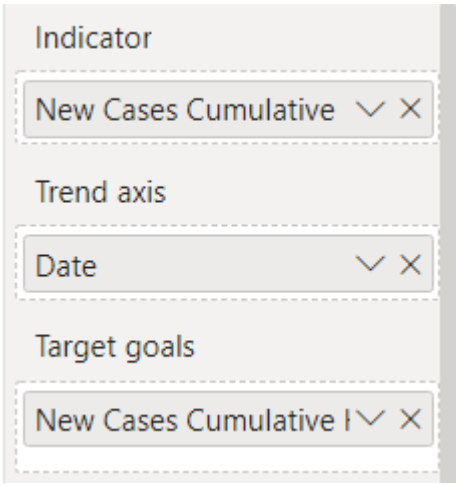
Visualisasi Data

1. Report Summary. Summary ini menggambarkan bagaimana keadaan covid pada tanggal yang ada di dataset. Report ini bersifat interaktif. Report ini bersifat analisis deskriptif



Untuk memvisualisasikan ini menggunakan KPI

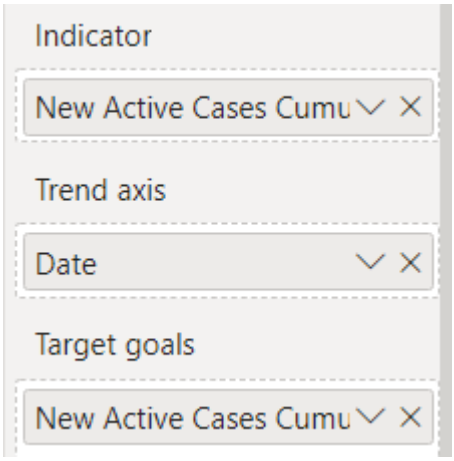





```
New Cases Cumulative H-1 =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Result =
    CALCULATE(
        [New Cases],
        REMOVEFILTERS('Calendar'),
        'Data Covid'[Date] < _MaxDate
        )

RETURN
    _Result
```

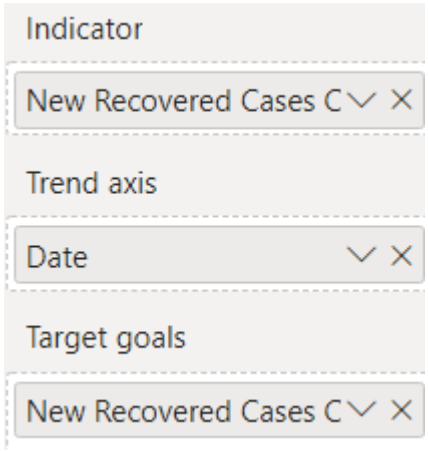
```
New Cases Cumulative =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Result =
    CALCULATE(
        [New Cases],
        FILTER(
            ALL('Calendar'),
            'Calendar'[Date] <= _MaxDate
        )
    )
RETURN
    _Result
```

Last day menunjukkan total kasus H-1 dan pertumbuhan yang terjadi



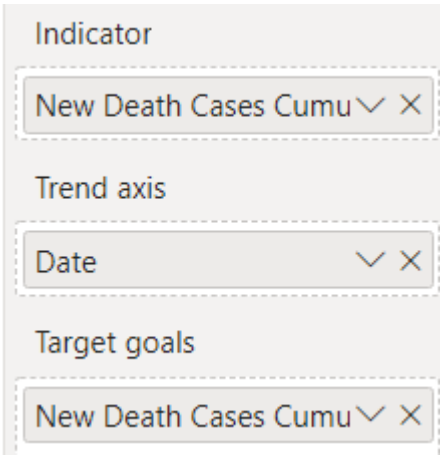

```
New Active Cases Cumulative =
VAR _MaxDate =
    MAX('Calendar'[Date])

VAR _filter =
    FILTER(
            ALL('Calendar'),
            'Calendar'[Date] <= _MaxDate)
VAR _Result =
    CALCULATE(
        [New Active Cases],
        FILTER(
            ALL('Calendar'),
            'Calendar'[Date] <= _MaxDate)
    )
RETURN
    _Result
```

```
New Active Cases Cumulative H-1 =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Result =
    CALCULATE(
        [New Active Cases],
        REMOVEFILTERS('Calendar'),
        'Data Covid'[Date] < _MaxDate
        )

RETURN
    _Result
```




```
1  New Recovered Cases Cumulative =
2  VAR _MaxDate = MAX('Calendar'[Date])
3  VAR _filter =
4      FILTER(
5              ALL('Calendar'),
6              'Calendar'[Date] <= _MaxDate)
7  VAR _Result =
8      CALCULATE(
9          [New Recovered],
10         FILTER(
11         ALL('Calendar'),
12         'Calendar'[Date] <= _MaxDate)
13     )
14 RETURN
15     _Result
```

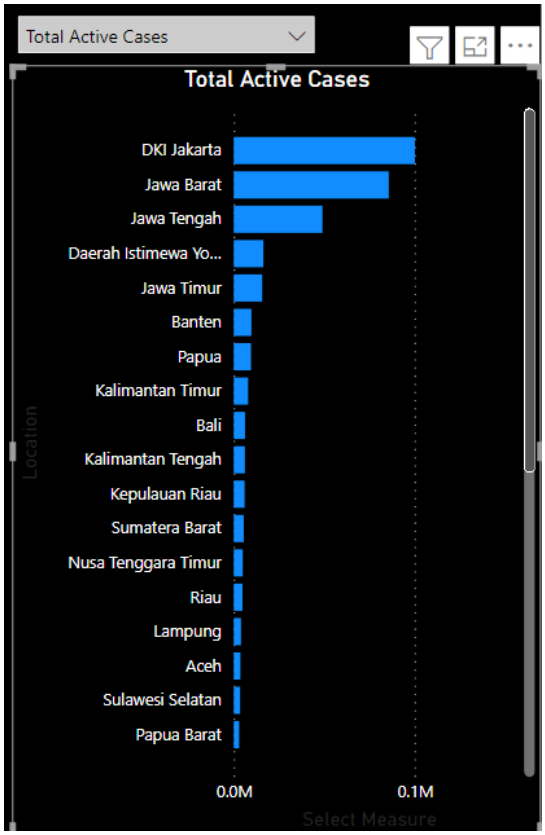
```
New Recovered Cases Cumulative H-1 =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Result =
    CALCULATE(
        [New Recovered],
        REMOVEFILTERS('Calendar'),
        'Data Covid'[Date] < _MaxDate
        )

RETURN
    _Result
```

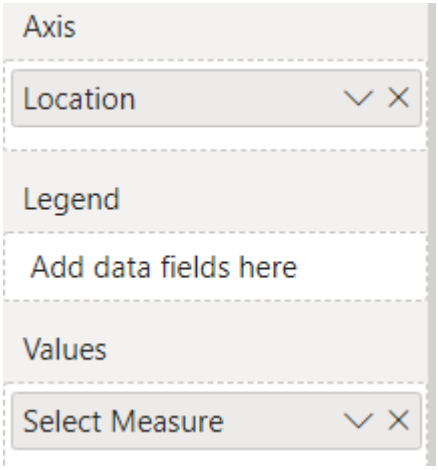



```
New Death Cases Cumulative =
VAR _MaxDate =
    MAX('Calendar'[Date])

VAR _filter =
    FILTER(
        ALL('Calendar'),
        'Calendar'[Date] <= _MaxDate)
VAR _Result =
    CALCULATE(
        [New Death],
        _filter
    )
RETURN
    _Result
```

```
New Death Cases Cumulative H-1 =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Result =
    CALCULATE(
        [New Death],
        REMOVEFILTERS('Calendar'),
        'Data Covid'[Date] < _MaxDate
        )

RETURN
    _Result
```

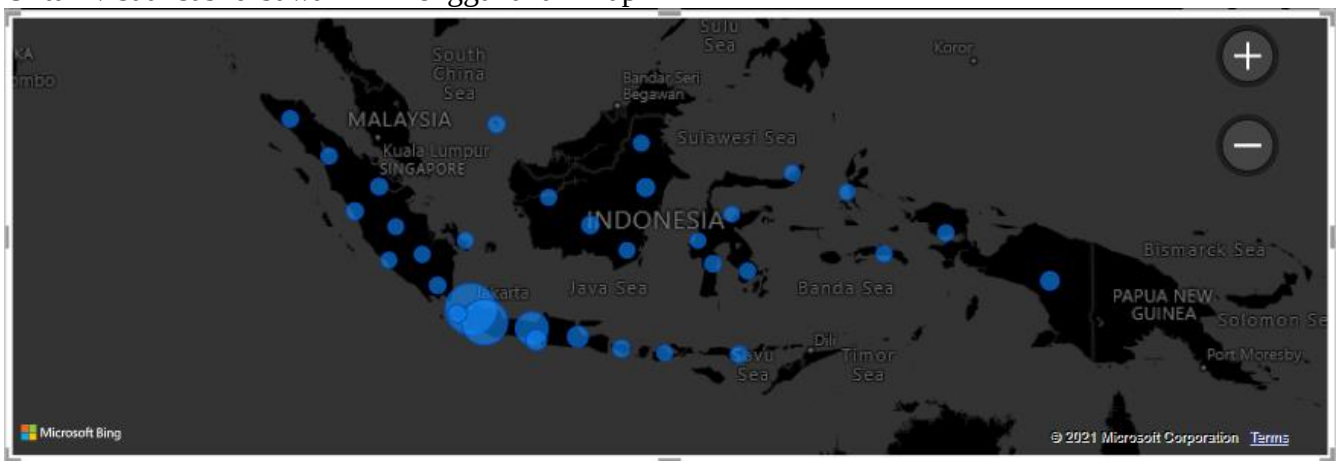
Untuk visualisasi selanjutnya menggunakan clustered bar chart

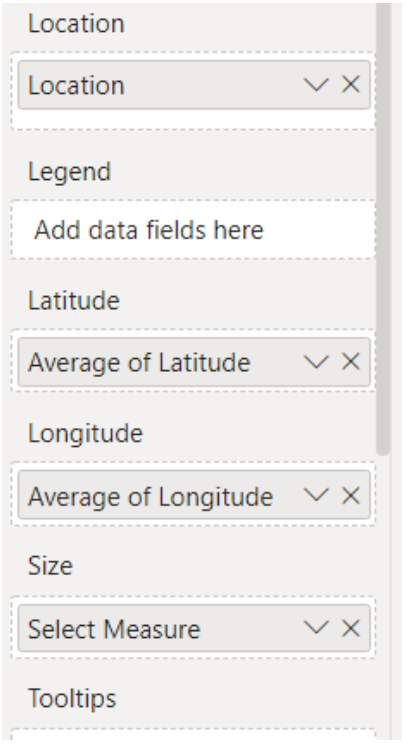





```
Select Measure =
VAR _Select = SELECTEDVALUE('Switcher Measure'[Measure])
VAR _New = [New Cases Cumulative]
VAR _Active = [New Active Cases Cumulative]
VAR _Recovered = [New Recovered Cases Cumulative]
VAR _Death = [New Death Cases Cumulative]
VAR _Result =
    SWITCH(
        TRUE(),
        _Select = "Total New Cases", _New,
        _Select = "Total Active Cases", _Active,
        _Select = "Total Recovered Cases", _Recovered,
        _Select = "Total Death Cases", _Death
    )
RETURN
    _Result
```

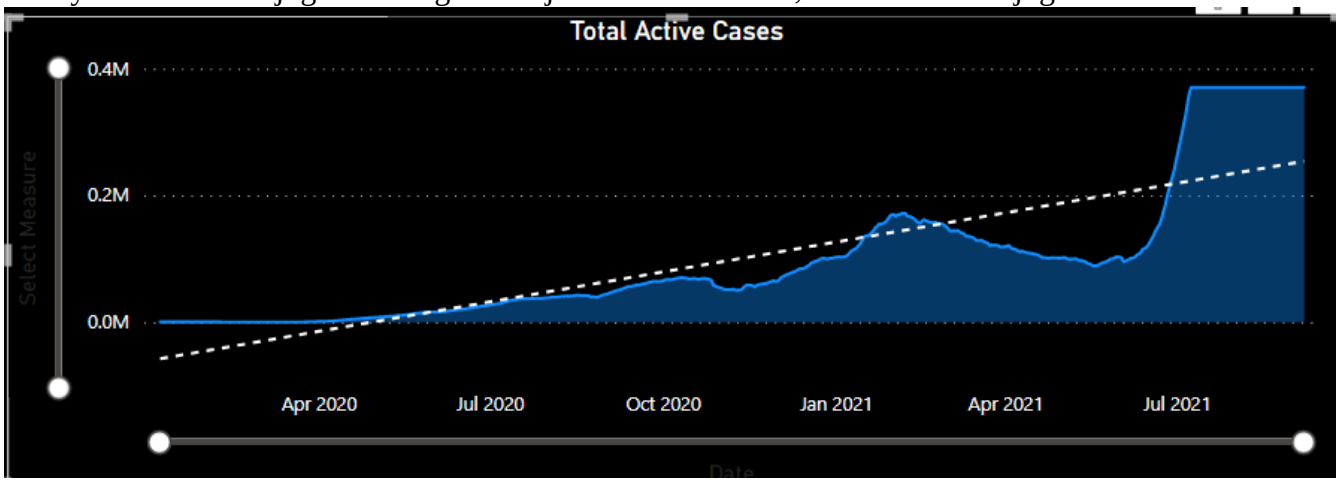
Untuk visualisasi dibawah ini menggunakan Map

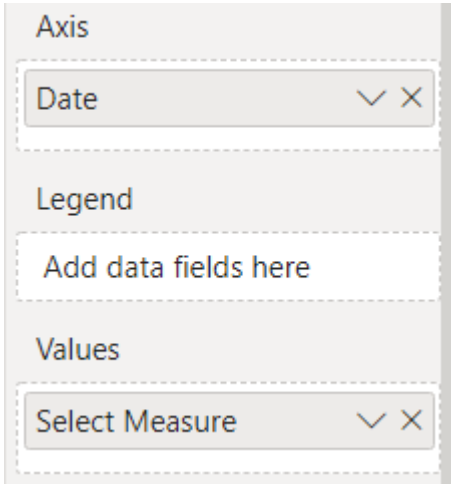





```
Select Measure =
VAR _Select = SELECTEDVALUE('Switcher Measure'[Measure])
VAR _New = [New Cases Cumulative]
VAR _Active = [New Active Cases Cumulative]
VAR _Recovered = [New Recovered Cases Cumulative]
VAR _Death = [New Death Cases Cumulative]
VAR _Result =
    SWITCH(
        TRUE(),
        _Select = "Total New Cases", _New,
        _Select = "Total Active Cases", _Active,
        _Select = "Total Recovered Cases", _Recovered,
        _Select = "Total Death Cases", _Death
    )
RETURN
    _Result
```

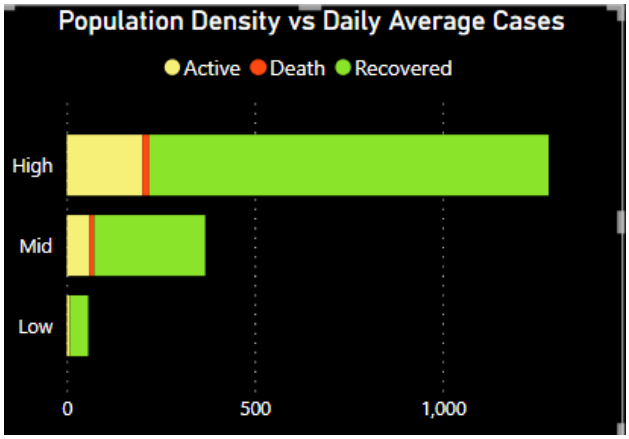
Untuk memvisualisasikannya menggunakan area chart. Garis putus-putus disini menunjukkan bahwa trennya naik. Hal ini juga memungkinkan jika kasus menurun, maka tren akan juga ikut turun.






```
Select Measure =
VAR _Select = SELECTEDVALUE('Switcher Measure'[Measure])
VAR _New = [New Cases Cumulative]
VAR _Active = [New Active Cases Cumulative]
VAR _Recovered = [New Recovered Cases Cumulative]
VAR _Death = [New Death Cases Cumulative]
VAR _Result =
    SWITCH(
        TRUE(),
        _Select = "Total New Cases", _New,
        _Select = "Total Active Cases", _Active,
        _Select = "Total Recovered Cases", _Recovered,
        _Select = "Total Death Cases", _Death
    )
RETURN
    _Result
```

2. Analytics page. Pada analytics page, saya menganalisa hubungan antara kepadatan penduduk dengan rata-rata kasus harian yang muncul. Saya membaginya menjadi tiga cluster, yaitu low, medium, dan high. Page ini dapat dianalisis menggunakan analisis diagnostic karena



Details

Location

Legend

Density

X Axis

Select Measure

Y Axis

Population Density

Size

Select Measure

```
Population Density =
MAXX(
    'Data Covid',
    DIVIDE('Data Covid'[Population], 'Data Covid'[Area (km2)])
)
```

```
Density =
SWITCH(
    TRUE(),
    Region[Location (clusters)] = "Cluster3", "Mid",
    Region[Location (clusters)] = "Cluster1", "Low",
    Region[Location (clusters)] = "Cluster2", "High"
)
```

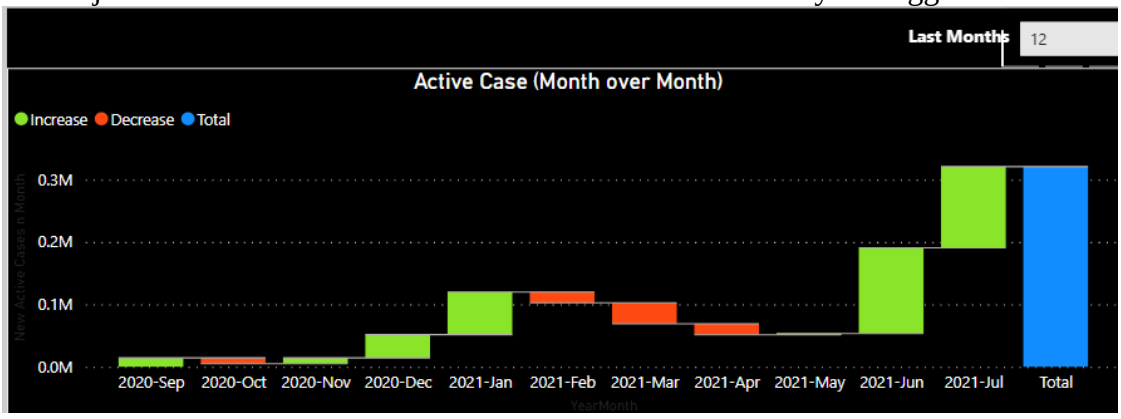
```
Select Measure =
VAR _Select = SELECTEDVALUE('Switcher Measure'[Measure])
VAR _New = [New Cases Cumulative]
VAR _Active = [New Active Cases Cumulative]
VAR _Recovered = [New Recovered Cases Cumulative]
VAR _Death = [New Death Cases Cumulative]
VAR _Result =
    SWITCH(
        TRUE(),
        _Select = "Total New Cases", _New,
        _Select = "Total Active Cases", _Active,
        _Select = "Total Recovered Cases", _Recovered,
        _Select = "Total Death Cases", _Death
    )
RETURN
    _Result
```

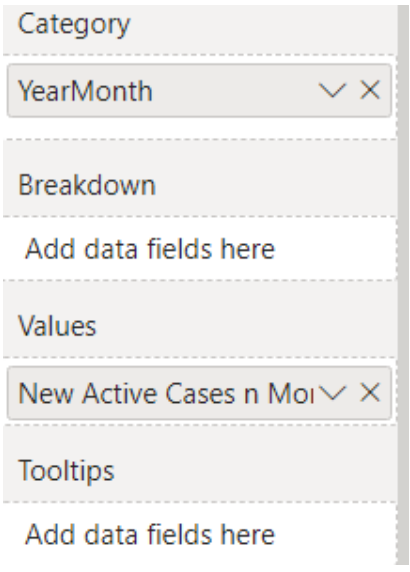
```
Density =
SWITCH(
    TRUE(),
    Region[Location (clusters)] = "Cluster3", "Mid",
    Region[Location (clusters)] = "Cluster1", "Low",
    Region[Location (clusters)] = "Cluster2", "High"
)
```

Location ISO Code	Location	Province	Country	Continent	Island	Time Zone	Longitude	Latitude	Location (clusters)	Density
ID-JI	Jawa Timur	Jawa Timur	Indonesia	Asia	Jawa	UTC+07:00	112.7329414	-7.723345579	Cluster3	Mid
ID-JK	DKI Jakarta	DKI Jakarta	Indonesia	Asia	Jawa	UTC+07:00	106.8361183	-6.204698991	Cluster2	High
ID-JB	Jawa Barat	Jawa Barat	Indonesia	Asia	Jawa	UTC+07:00	107.6037083	-6.920432083	Cluster3	Mid
ID-RI	Riau	Riau	Indonesia	Asia	Sumatera	UTC+07:00	101.8051092	0.511647851	Cluster1	Low
ID-BT	Banten	Banten	Indonesia	Asia	Jawa	UTC+07:00	106.1090043	-6.456736388	Cluster3	Mid
ID-JT	Jawa Tengah	Jawa Tengah	Indonesia	Asia	Jawa	UTC+07:00	110.2011149	-7.259097177	Cluster3	Mid
ID-SG	Sulawesi Tenggara	Sulawesi Tenggara	Indonesia	Asia	Sulawesi	UTC+08:00	122.070311	-4.124688793	Cluster1	Low
ID-BA	Bali	Bali	Indonesia	Asia	Nusa Tenggara	UTC+08:00	115.1317136	-8.369471688	Cluster3	Mid
ID-KI	Kalimantan Timur	Kalimantan Timur	Indonesia	Asia	Kalimantan	UTC+08:00	116.4684405	0.45385803	Cluster1	Low
ID-YO	Daerah Istimewa Yogyakarta	Daerah Istimewa Yogyakarta	Indonesia	Asia	Jawa	UTC+07:00	110.4448783	-7.89450185	Cluster3	Mid
ID-SU	Sumatera Utara	Sumatera Utara	Indonesia	Asia	Sumatera	UTC+07:00	99.05196442	2.191894453	Cluster1	Low
ID-KR	Kepulauan Riau	Kepulauan Riau	Indonesia	Asia	Sumatera	UTC+07:00	108.261746	3.916346	Cluster1	Low
ID-SN	Sulawesi Selatan	Sulawesi Selatan	Indonesia	Asia	Sulawesi	UTC+08:00	120.1620559	-3.731080714	Cluster1	Low
ID-JA	Jambi	Jambi	Indonesia	Asia	Sumatera	UTC+07:00	102.7236404	-1.69769766	Cluster1	Low
ID-MA	Maluku	Maluku	Indonesia	Asia	Maluku	UTC+09:00	129.576792	-3.192572	Cluster1	Low
ID-PA	Papua	Papua	Indonesia	Asia	Papua	UTC+09:00	138.69603	-4.66620953	Cluster1	Low
ID-MU	Maluku Utara	Maluku Utara	Indonesia	Asia	Maluku	UTC+09:00	127.5391072	0.212036949	Cluster1	Low
ID-SS	Sumatera Selatan	Sumatera Selatan	Indonesia	Asia	Sumatera	UTC+07:00	104.1694647	-3.216211808	Cluster1	Low
ID-AC	Aceh	Aceh	Indonesia	Asia	Sumatera	UTC+07:00	96.91052174	4.225614628	Cluster1	Low
ID-KT	Kalimantan Tengah	Kalimantan Tengah	Indonesia	Asia	Kalimantan	UTC+07:00	113.4176536	-1.602484653	Cluster1	Low
ID-LA	Lampung	Lampung	Indonesia	Asia	Sumatera	UTC+07:00	105.0214366	-4.916792975	Cluster1	Low
ID-ST	Sulawesi Tengah	Sulawesi Tengah	Indonesia	Asia	Sulawesi	UTC+08:00	121.2010927	-1.00413668	Cluster1	Low
ID-SA	Sulawesi Utara	Sulawesi Utara	Indonesia	Asia	Sulawesi	UTC+08:00	124.5212396	1.259638212	Cluster1	Low

Location	Location (clusters)	Population Density	Density
Aceh	Cluster1	90.54	Low
Bali	Cluster3	729.44	Mid
Banten	Cluster3	1,109.63	Mid
Bengkulu	Cluster1	100.38	Low
Daerah Istimewa Yogyakarta	Cluster3	1,158.96	Mid
DKI Jakarta	Cluster2	16,334.56	High
Jambi	Cluster1	69.79	Low
Jawa Barat	Cluster3	1,276.54	Mid
Jawa Tengah	Cluster3	1,108.63	Mid
Jawa Timur	Cluster3	846.79	Mid
Kalimantan Barat	Cluster1	36.81	Low
Kalimantan Selatan	Cluster1	103.84	Low
Kalimantan Tengah	Cluster1	16.74	Low
Kalimantan Timur	Cluster1	27.52	Low
Kalimantan Utara	Cluster1	8.59	Low
Kepulauan Bangka Belitung	Cluster1	84.01	Low
Kepulauan Riau	Cluster1	235.24	Low
Lampung	Cluster1	262.70	Low
Maluku	Cluster1	39.37	Low
Maluku Utara	Cluster1	40.89	Low
Nusa Tenggara Barat	Cluster1	283.77	Low
Nusa Tenggara Timur	Cluster1	111.07	Low
Papua	Cluster1	13.60	Low
Papua Barat	Cluster1	11.08	Low
Riau	Cluster1	69.80	Low
Sulawesi Barat	Cluster1	92.93	Low
Total		**16,334.56**	

Menunjukkan kasus aktif dari bulan ke bulan. Visualisasi datanya menggunakan waterfall chart






```
New Active Cases n Month =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Parameter = SELECTEDVALUE(Parameter[Parameter])
VAR _Period = DATESINPERIOD('Calendar'[Date], _MaxDate, -_Parameter, MONTH)
VAR _Result =
    CALCULATE(
        [New Active Cases],
        REMOVEFILTERS('Calendar'),
        KEEPFILTERS(_Period),
        USERELATIONSHIP(CalendarAlt[Date], 'Calendar'[Date])
    )
RETURN
    _Result
```

Location vs Daily Average Cases

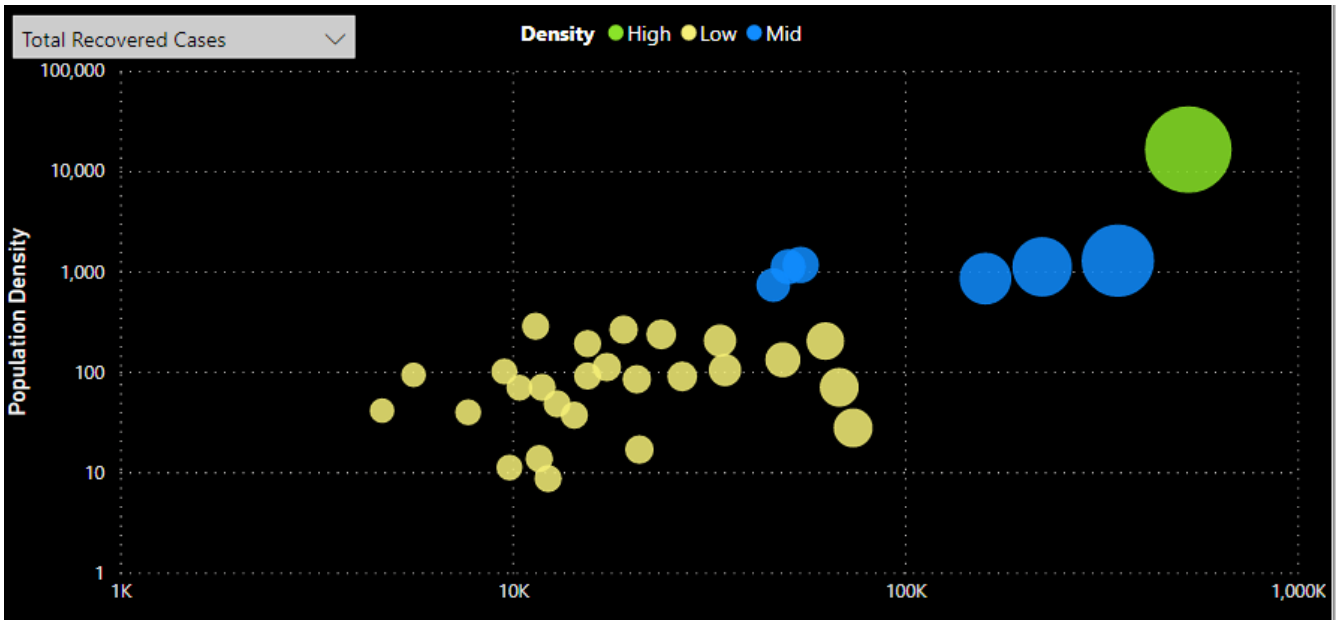
Location	Death	Recovered	Active
Jawa Tengah	26	457	100
Jawa Timur	25	292	29
DKI Jakarta	19	1,062	202
Jawa Barat	12	704	173
Kalimantan Timur	4	153	17
Riau	4	137	10
Daerah Istimewa Yogyakarta	4	113	34
Bali	3	95	13
Sumatera Selatan	3	57	5
Banten	3	103	20
Sumatera Barat	3	104	12

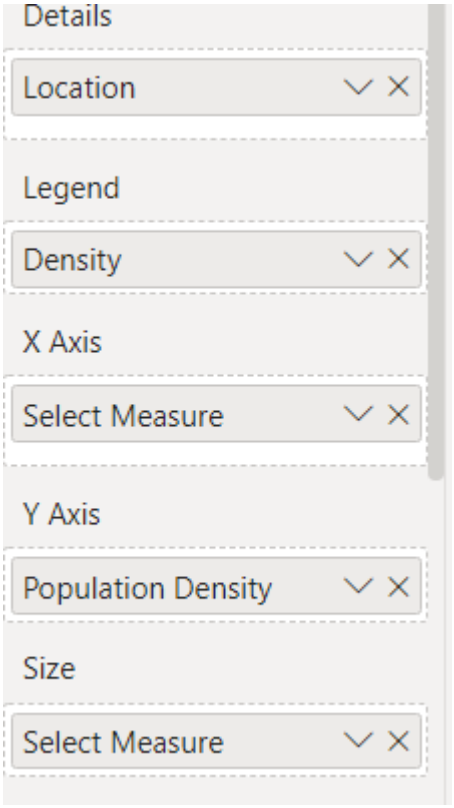

```
New Active Cases Cumulative Average =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Result =
    CALCULATE(
        AVERAGE('Data Covid'[New Active Cases]),
        FILTER(
            ALL('Calendar'),
            'Calendar'[Date] <= _MaxDate
        )
    )
RETURN
    _Result
```

```
New Recovered Cases Cumulative Average =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Result =
    CALCULATE(
        AVERAGE('Data Covid'[New Recovered]),
        FILTER(
            ALL('Calendar'),
            'Calendar'[Date] <= _MaxDate
        )
    )
RETURN
    _Result
```

```
New Death Cases Cumulative Average =
VAR _MaxDate = MAX('Calendar'[Date])
VAR _Result =
    CALCULATE(
        AVERAGE('Data Covid'[New Deaths]),
        FILTER(
            ALL('Calendar'),
            'Calendar'[Date] <= _MaxDate
        )
    )
RETURN
    _Result
```

Untuk memvisualisasikan menggunakan scatter chart



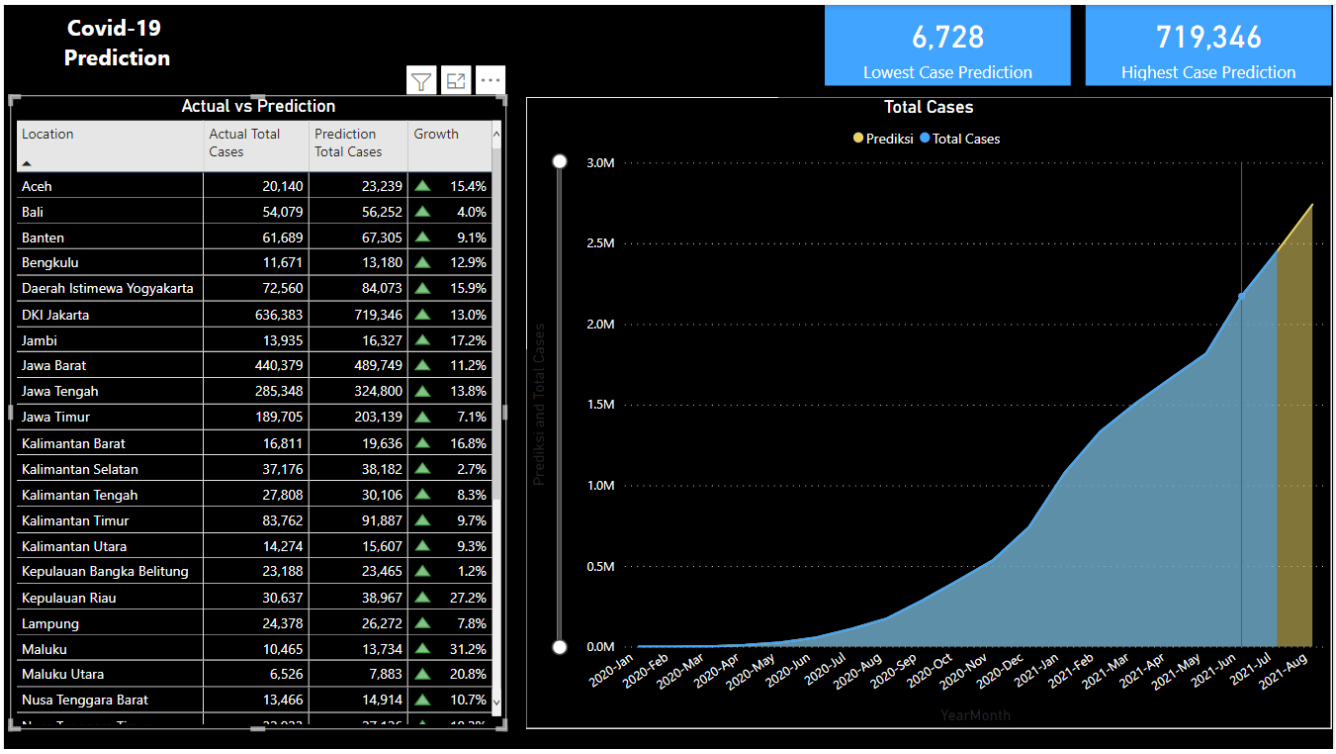



```
Select Measure =
VAR _Select = SELECTEDVALUE('Switcher Measure'[Measure])
VAR _New = [New Cases Cumulative]
VAR _Active = [New Active Cases Cumulative]
VAR _Recovered = [New Recovered Cases Cumulative]
VAR _Death = [New Death Cases Cumulative]
VAR _Result =
    SWITCH(
        TRUE(),
        _Select = "Total New Cases", _New,
        _Select = "Total Active Cases", _Active,
        _Select = "Total Recovered Cases", _Recovered,
        _Select = "Total Death Cases", _Death
    )
RETURN
    _Result
```

```
Population Density =
MAXX(
    'Data Covid',
    DIVIDE('Data Covid'[Population], 'Data Covid'[Area (km2)])
)
```

```
Density =
SWITCH(
    TRUE(),
    Region[Location (clusters)] = "Cluster3", "Mid",
    Region[Location (clusters)] = "Cluster1", "Low",
    Region[Location (clusters)] = "Cluster2", "High"
)
```

3. Predicition page. Pada page ini menggunakan analisis prediktif untuk menentukan pertambahan total kasus dari tanggal terakhir hingga memasuki awal bulan berikutnya. Di page ini saya menggunakan model machine learning regression yang sudah dibuat di azure

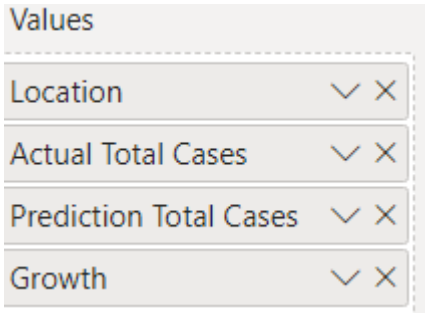


Location	Actual Total Cases	Prediction Total Cases	Growth
Aceh	20,140	23,239	15.4%
Bali	54,079	56,252	4.0%
Banten	61,689	67,305	9.1%
Bengkulu	11,671	13,180	12.9%
Daerah Istimewa Yogyakarta	72,560	84,073	15.9%
DKI Jakarta	636,383	719,346	13.0%
Jambi	13,935	16,327	17.2%
Jawa Barat	440,379	489,749	11.2%
Jawa Tengah	285,348	324,800	13.8%
Jawa Timur	189,705	203,139	7.1%
Kalimantan Barat	16,811	19,636	16.8%
Kalimantan Selatan	37,176	38,182	2.7%
Kalimantan Tengah	27,808	30,106	8.3%
Kalimantan Timur	83,762	91,887	9.7%
Kalimantan Utara	14,274	15,607	9.3%
Kepulauan Bangka Belitung	23,188	23,465	1.2%
Kepulauan Riau	30,637	38,967	27.2%
Lampung	24,378	26,272	7.8%
Maluku	10,465	13,734	31.2%
Maluku Utara	6,526	7,883	20.8%
Nusa Tenggara Barat	13,466	14,914	10.7%

Tabel ini menunjukkan adanya angka actual dan prediksi dimana angka actual disitu merupakan kasus bulan lalu

Actual vs Prediction

Location	Actual Total Cases	Prediction Total Cases	Growth
Aceh	20,140	23,239	15.4%
Bali	54,079	56,252	4.0%
Banten	61,689	67,305	9.1%
Bengkulu	11,671	13,180	12.9%
Daerah Istimewa Yogyakarta	72,560	84,073	15.9%
DKI Jakarta	636,383	719,346	13.0%
Jambi	13,935	16,327	17.2%
Jawa Barat	440,379	489,749	11.2%
Jawa Tengah	285,348	324,800	13.8%
Jawa Timur	189,705	203,139	7.1%
Kalimantan Barat	16,811	19,636	16.8%
Kalimantan Selatan	37,176	38,182	2.7%
Kalimantan Tengah	27,808	30,106	8.3%
Kalimantan Timur	83,762	91,887	9.7%
Kalimantan Utara	14,274	15,607	9.3%
Kepulauan Bangka Belitung	23,188	23,465	1.2%
Kepulauan Riau	30,637	38,967	27.2%
Lampung	24,378	26,272	7.8%
Maluku	10,465	13,734	31.2%
Maluku Utara	6,526	7,883	20.8%
Nusa Tenggara Barat	13,466	14,914	10.7%

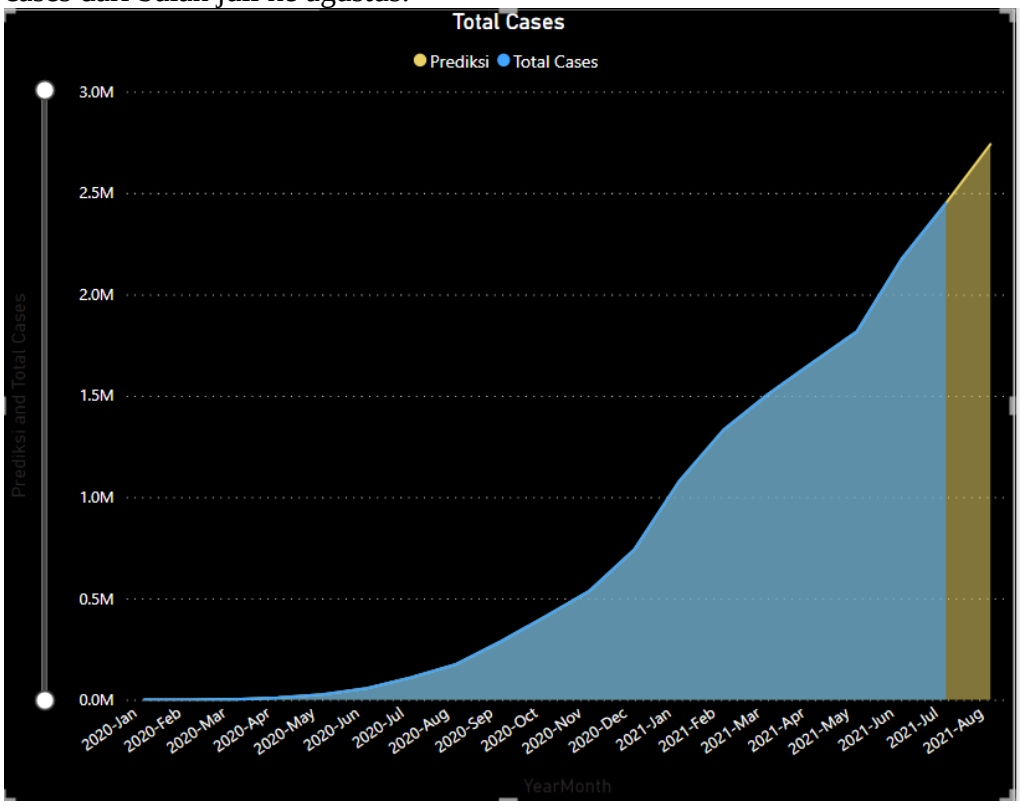


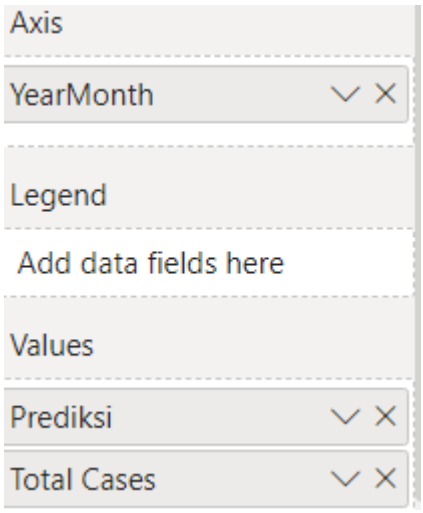

```
Total Cases 2 =
VAR _T =
SUMX(
    ADDCOLUMNS(
        VALUES(Region[Location]),
        "ABC", CALCULATE(MAX('Data Covid'[Total Cases]))),
        [ABC]
)
RETURN
    _T
```

```
Predict =
SUM('Predic Regression'[Scored Labels])
```

```
Growth Act vs Predict =
VAR _Act = [Total Cases 2]
VAR _Pred = [Predict]
VAR _delta = [Total Cases 2] - [Predict]
RETURN
    DIVIDE(_Pred, _Act)-1
```

Untuk menvisualisasikannya menggunakan area chart. Chart ini menunjukkan prediksi kenaikan total cases dari bulan juli ke agustus.






```
Projection Case 3 =
VAR _T =
    ADDCOLUMNS(
        VALUES('Calendar'[YearMonth]),
        "TOTALCASE", [Total Cases],
        "PREDICT", [Predict])
RETURN
CALCULATE(
    SUMX(
        _T,
        [TOTALCASE] + [PREDICT]
    ),
    VALUES('Calendar'[YearMonth]))
```

```
Total Cases =
SUMX(
ADDCOLUMNS(
    SUMMARIZE(
        'Data Covid', 'Data Covid'[Location], 'Calendar'[YearMonth]),
        "Abc", CALCULATE(
          MAX('Data Covid'[Total Cases]),
          VALUES(Region[Location])))
, [Abc])
```