In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:
```
df = pd.read_csv("Indonesia_coronavirus_daily_data.csv")
df.head()
```

Out[2]:

	Date	Province	Daily_Case	Daily_Death	Daily_Recovered	Active_Case	Cumulative_Case	Cumulative_Recovered	Cumulative_Death	Cumulative_Active_Cas
0	2020-03-01	DKI JAKARTA	2	0	0	2	2	0	0	
1	2020-03-02	DKI JAKARTA	2	0	0	2	4	0	0	
2	2020-03-03	DKI JAKARTA	2	0	0	2	6	0	0	
3	2020-03-04	DKI JAKARTA	2	0	0	2	8	0	0	
4	2020-03-05	DKI JAKARTA	0	1	0	-1	8	0	1	

In [3]:
```
to_drop = ['Date']
```

In [4]:
```
df.drop(to_drop, inplace=True, axis=1)
```

In [5]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17613 entries, 0 to 17612
Data columns (total 9 columns):
 #   Column                 Non-Null Count  Dtype 
---  ------                 --------------  ----- 
 0   Province               17613 non-null  object
 1   Daily_Case             17613 non-null  int64 
 2   Daily_Death            17613 non-null  int64 
 3   Daily_Recovered        17613 non-null  int64 
 4   Active_Case            17613 non-null  int64 
 5   Cumulative_Case        17613 non-null  int64 
 6   Cumulative_Recovered   17613 non-null  int64 
 7   Cumulative_Death       17613 non-null  int64 
 8   Cumulative_Active_Case 17613 non-null  int64 
dtypes: int64(8), object(1)
memory usage: 1.2+ MB
```

In [6]:
```
x = df.drop(["Province"], axis=1)
x.head(11)
```

Out[6]:

	Daily_Case	Daily_Death	Daily_Recovered	Active_Case	Cumulative_Case	Cumulative_Recovered	Cumulative_Death	Cumulative_Active_Case
0	2	0	0	2	2	0	0	2
1	2	0	0	2	4	0	0	4
2	2	0	0	2	6	0	0	6
3	2	0	0	2	8	0	0	8
4	0	1	0	-1	8	0	1	7
5	0	0	0	0	8	0	1	7
6	0	2	0	-2	8	0	3	5
7	0	0	0	0	8	0	3	5
8	0	1	0	-1	8	0	4	4
9	0	0	0	0	8	0	4	4
10	0	0	0	0	8	0	4	4

In [7]:
```
y = df["Province"]
y.head(11)
```

Out[7]:
```
0     DKI JAKARTA
1     DKI JAKARTA
2     DKI JAKARTA
3     DKI JAKARTA
4     DKI JAKARTA
5     DKI JAKARTA
6     DKI JAKARTA
7     DKI JAKARTA
8     DKI JAKARTA
9     DKI JAKARTA
10    DKI JAKARTA
Name: Province, dtype: object
```

In [9]:
```
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB

modelnb = GaussianNB()
```

In [10]:
```
nbtrain = modelnb.fit(x, y)
df.head(11)
```

Out[10]:

	Province	Daily_Case	Daily_Death	Daily_Recovered	Active_Case	Cumulative_Case	Cumulative_Recovered	Cumulative_Death	Cumulative_Active_Case
0	DKI JAKARTA	2	0	0	2	2	0	0	2
1	DKI JAKARTA	2	0	0	2	4	0	0	4
2	DKI JAKARTA	2	0	0	2	6	0	0	6
3	DKI JAKARTA	2	0	0	2	8	0	0	8
4	DKI JAKARTA	0	1	0	-1	8	0	1	7
5	DKI JAKARTA	0	0	0	0	8	0	1	7
6	DKI JAKARTA	0	2	0	-2	8	0	3	5
7	DKI JAKARTA	0	0	0	0	8	0	3	5
8	DKI JAKARTA	0	1	0	-1	8	0	4	4
9	DKI JAKARTA	0	0	0	0	8	0	4	4
10	DKI JAKARTA	0	0	0	0	8	0	4	4

In [12]:
```
x_test = df.drop(["Province"], axis=1)
x_test.head(11)
```

Out[12]:

	Daily_Case	Daily_Death	Daily_Recovered	Active_Case	Cumulative_Case	Cumulative_Recovered	Cumulative_Death	Cumulative_Active_Case
0	2	0	0	2	2	0	0	2
1	2	0	0	2	4	0	0	4
2	2	0	0	2	6	0	0	6
3	2	0	0	2	8	0	0	8
4	0	1	0	-1	8	0	1	7
5	0	0	0	0	8	0	1	7
6	0	2	0	-2	8	0	3	5
7	0	0	0	0	8	0	3	5
8	0	1	0	-1	8	0	4	4
9	0	0	0	0	8	0	4	4
10	0	0	0	0	8	0	4	4

In [13]:
```
y_uji = df["Province"]
y_uji.head(11)
```

Out[13]:
```
0     DKI JAKARTA
1     DKI JAKARTA
2     DKI JAKARTA
3     DKI JAKARTA
4     DKI JAKARTA
5     DKI JAKARTA
6     DKI JAKARTA
7     DKI JAKARTA
8     DKI JAKARTA
9     DKI JAKARTA
10    DKI JAKARTA
Name: Province, dtype: object
```

In [14]:
```
Y_predict = nbtrain.predict(x_test)
print("Prediksi Naive Bayes : ",Y_predict)
```

```
Prediksi Naive Bayes :  ['SULAWESI BARAT' 'SULAWESI BARAT' 'SULAWESI BARAT' ... 'BANTEN'
 'KALIMANTAN TIMUR' 'BANTEN']
```

In [15]:
```
from sklearn.metrics import accuracy_score
accuracy= accuracy_score(y_uji, Y_predict)
print("Akurasi Naive Bayes : ",accuracy)
```

```
Akurasi Naive Bayes :  0.1538068472151252
```

In [16]:
```
# Menghitung nilai akurasi dari klasifikasi naive bayes
from sklearn.metrics import classification_report
print(classification_report(y_uji, Y_predict))
```

```
                           precision    recall  f1-score   support

                     ACEH       0.01      0.01      0.01       511
                     BALI       0.00      0.00      0.00       526
                   BANTEN       0.05      0.01      0.02       533
                 BENGKULU       0.00      0.00      0.00       506
DAERAH ISTIMEWA YOGYAKARTA       0.03      0.01      0.01       521
              DKI JAKARTA        0.74      0.39      0.51       536
                GORONTALO        0.18      0.63      0.28       496
                    JAMBI        0.00      0.00      0.00       517
               JAWA BARAT        0.69      0.26      0.38       535
              JAWA TENGAH        0.56      0.31      0.40       529
               JAWA TIMUR        0.47      0.40      0.43       589
         KALIMANTAN BARAT        0.05      0.00      0.00       508
        KALIMANTAN SELATAN       0.13      0.28      0.18       507
         KALIMANTAN TENGAH       0.05      0.02      0.03       511
          KALIMANTAN TIMUR       0.09      0.08      0.08       523
          KALIMANTAN UTARA       0.11      0.03      0.05       509
 KEPULAUAN BANGKA BELITUNG       0.71      0.01      0.02       507
           KEPULAUAN RIAU        0.03      0.00      0.00       518
                  LAMPUNG        0.04      0.01      0.01       511
                   MALUKU        0.17      0.16      0.17       517
             MALUKU UTARA        0.01      0.02      0.02       515
       NUSA TENGGARA BARAT       0.12      0.23      0.16       506
       NUSA TENGGARA TIMUR       0.05      0.01      0.02       497
                    PAPUA        0.58      0.54      0.56       515
              PAPUA BARAT        0.08      0.02      0.03       510
                    RIAU         0.11      0.06      0.07       535
          SULAWESI BARAT         0.05      0.57      0.09       509
         SULAWESI SELATAN        0.39      0.37      0.38       520
          SULAWESI TENGAH        0.05      0.01      0.01       511
        SULAWESI TENGGARA        0.22      0.34      0.27       527
          SULAWESI UTARA         0.16      0.34      0.22       511
          SUMATERA BARAT         0.04      0.02      0.03       511
         SUMATERA SELATAN        0.24      0.05      0.08       514
          SUMATERA UTARA         0.00      0.00      0.00       522

                accuracy                           0.15     17613
               macro avg        0.18      0.15      0.13     17613
            weighted avg        0.19      0.15      0.14     17613
```