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

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

Out[2]:

	conversation_id	date	time	user_id	username	tweet	mentions	replies_count	retweets_count	likes_count
0	1258425982907637761	2020-05-07	23:58:29	1058474317	monologis_id	https://bit.ly/2L6CcbB | Seharusnya saat ini...	[]	0	3	0
1	1258320972198940675	2020-05-07	23:57:30	1179769476	its_dul	Klo kata gw Pemerintah tuh lagi menerapkan Her...	['mas__piyuuu']	0	0	0
2	1258356644427083777	2020-05-07	23:53:20	1012156669831229441	meonkbaong	Saat ini yang bisa saya lakukan hanya menyiapk...	['oiivert']	0	0	0
3	1258424368993931265	2020-05-07	23:52:04	1204303690061844481	rakyatdotnews	Satu Warga Positif Corona, Bupati: Kondisi ini...	[]	0	0	0
4	1258423545698246656	2020-05-07	23:48:48	227620381	annisathalib_	emosi bgt, lg kondisi begini gue disuruh liput...	[]	0	0	0

In [3]:
```
to_drop = ['conversation_id','time','mentions','tweet','username','hashtags','user_id']
```

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

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 52959 entries, 0 to 52958
Data columns (total 4 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   date            52959 non-null  object
 1   replies_count   52959 non-null  int64 
 2   retweets_count  52959 non-null  int64 
 3   likes_count     52959 non-null  int64 
dtypes: int64(3), object(1)
memory usage: 1.6+ MB
```

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

Out[6]:

	replies_count	retweets_count	likes_count
0	0	3	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	1
7	0	0	3
8	0	1	3
9	0	2	13
10	0	0	0

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

Out[7]:
```
0     2020-05-07
1     2020-05-07
2     2020-05-07
3     2020-05-07
4     2020-05-07
5     2020-05-07
6     2020-05-07
7     2020-05-07
8     2020-05-07
9     2020-05-07
10    2020-05-07
Name: date, dtype: object
```

In [8]:
```
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB

modelnb = GaussianNB()
```

In [9]:
```
nbtrain = modelnb.fit(x, y)
df.head(11)
```

Out[9]:

	date	replies_count	retweets_count	likes_count
0	2020-05-07	0	3	0
1	2020-05-07	0	0	0
2	2020-05-07	0	0	0
3	2020-05-07	0	0	0
4	2020-05-07	0	0	0
5	2020-05-07	0	0	0
6	2020-05-07	0	0	1
7	2020-05-07	0	0	3
8	2020-05-07	0	1	3
9	2020-05-07	0	2	13
10	2020-05-07	0	0	0

In [10]:
```
x_test = df.drop(["date"], axis=1)
x_test.head(11)
```

Out[10]:

	replies_count	retweets_count	likes_count
0	0	3	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	1
7	0	0	3
8	0	1	3
9	0	2	13
10	0	0	0

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

Out[11]:
```
0     2020-05-07
1     2020-05-07
2     2020-05-07
3     2020-05-07
4     2020-05-07
5     2020-05-07
6     2020-05-07
7     2020-05-07
8     2020-05-07
9     2020-05-07
10    2020-05-07
Name: date, dtype: object
```

In [12]:
```
Y_predict = nbtrain.predict(x_test)
print("Prediksi Naive Bayes : ",Y_predict)
```

```
Prediksi Naive Bayes :  ['2020-06-20' '2020-06-20' '2020-06-20' ... '2020-06-20' '2020-06-20'
 '2020-05-22']
```

In [13]:
```
from sklearn.metrics import accuracy_score
accuracy= accuracy_score(y_uji, Y_predict)
print("Akurasi Naive Bayes : ",accuracy)
```

```
Akurasi Naive Bayes :  0.0099510942427160 63
```

In [14]:
```
# Menghitung nilai akurasi dari klasifikasi naive bayes
from sklearn.metrics import classification_report
print(classification_report(y_uji, Y_predict))
```

```
C:\Users\ROG\AppData\Roaming\Python\Python39\site-packages\sklearn\metrics\_classification.py:1334: UndefinedMetricWarning: Pre
cision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero_division` parameter to c
ontrol this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
```

	precision	recall	f1-score	support
2020-05-01	0.00	0.00	0.00	1053
2020-05-02	0.00	0.00	0.00	858
2020-05-03	0.00	0.00	0.00	834
2020-05-04	0.00	0.00	0.00	1069
2020-05-05	0.00	0.00	0.00	986
2020-05-06	0.04	0.00	0.00	1290
2020-05-07	0.02	0.00	0.00	1120
2020-05-12	0.00	0.00	0.00	564
2020-05-13	0.02	0.00	0.01	3712
2020-05-14	0.04	0.00	0.01	1797
2020-05-15	0.00	0.00	0.00	1273
2020-05-18	0.00	0.00	0.00	146
2020-05-19	0.05	0.00	0.01	3506
2020-05-20	0.00	0.00	0.00	1680
2020-05-21	0.03	0.00	0.00	1269
2020-05-22	0.03	0.01	0.01	2182
2020-05-23	0.02	0.01	0.01	1116
2020-05-25	0.00	0.00	0.00	590
2020-05-26	0.02	0.00	0.00	1077
2020-05-27	0.00	0.00	0.00	1347
2020-05-28	0.00	0.00	0.00	1497
2020-05-29	0.00	0.00	0.00	1165
2020-05-30	0.07	0.00	0.01	772
2020-06-01	0.00	0.00	0.00	636
2020-06-02	0.00	0.00	0.00	850
2020-06-03	0.09	0.00	0.00	928
2020-06-04	0.00	0.00	0.00	882
2020-06-05	0.00	0.00	0.00	792
2020-06-06	0.00	0.00	0.00	540
2020-06-07	0.00	0.00	0.00	448
2020-06-13	0.00	0.00	0.00	417
2020-06-14	0.00	0.00	0.00	413
2020-06-15	0.17	0.00	0.00	765
2020-06-16	0.02	0.00	0.00	870
2020-06-17	0.00	0.00	0.00	597
2020-06-18	0.00	0.00	0.00	554
2020-06-19	0.00	0.00	0.00	569
2020-06-20	0.01	0.91	0.02	410
2020-06-21	0.00	0.00	0.00	394
2020-06-22	0.08	0.00	0.01	715
2020-06-23	0.00	0.00	0.00	639
2020-06-27	0.00	0.00	0.00	473
2020-06-28	0.00	0.00	0.00	334
2020-06-29	0.00	0.00	0.00	545
2020-07-02	0.00	0.00	0.00	706
2020-07-03	0.00	0.00	0.00	1174
2020-07-04	0.00	0.00	0.00	788
2020-07-05	0.00	0.00	0.00	761
2020-07-06	0.03	0.06	0.05	1171
2020-07-07	0.02	0.00	0.00	1789
2020-07-15	0.00	0.00	0.00	703
2020-07-20	0.00	0.00	0.00	745
2020-07-21	0.29	0.00	0.00	809
2020-07-22	0.00	0.00	0.00	639
accuracy			0.01	52959
macro avg	0.02	0.02	0.00	52959
weighted avg	0.02	0.01	0.00	52959

```
C:\Users\ROG\AppData\Roaming\Python\Python39\site-packages\sklearn\metrics\_classification.py:1334: UndefinedMetricWarning: Pre
cision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero_division` parameter to c
ontrol this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
C:\Users\ROG\AppData\Roaming\Python\Python39\site-packages\sklearn\metrics\_classification.py:1334: UndefinedMetricWarning: Pre
cision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero_division` parameter to c
ontrol this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
```