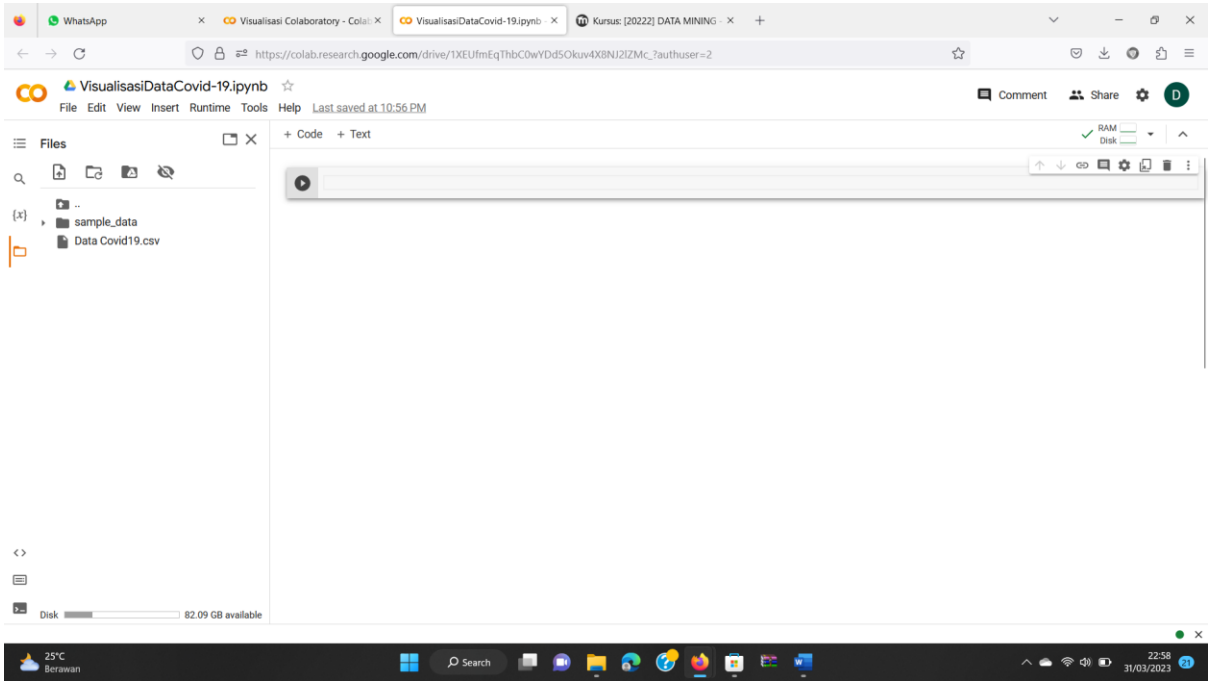
NAMA : DINA ETIKASARI

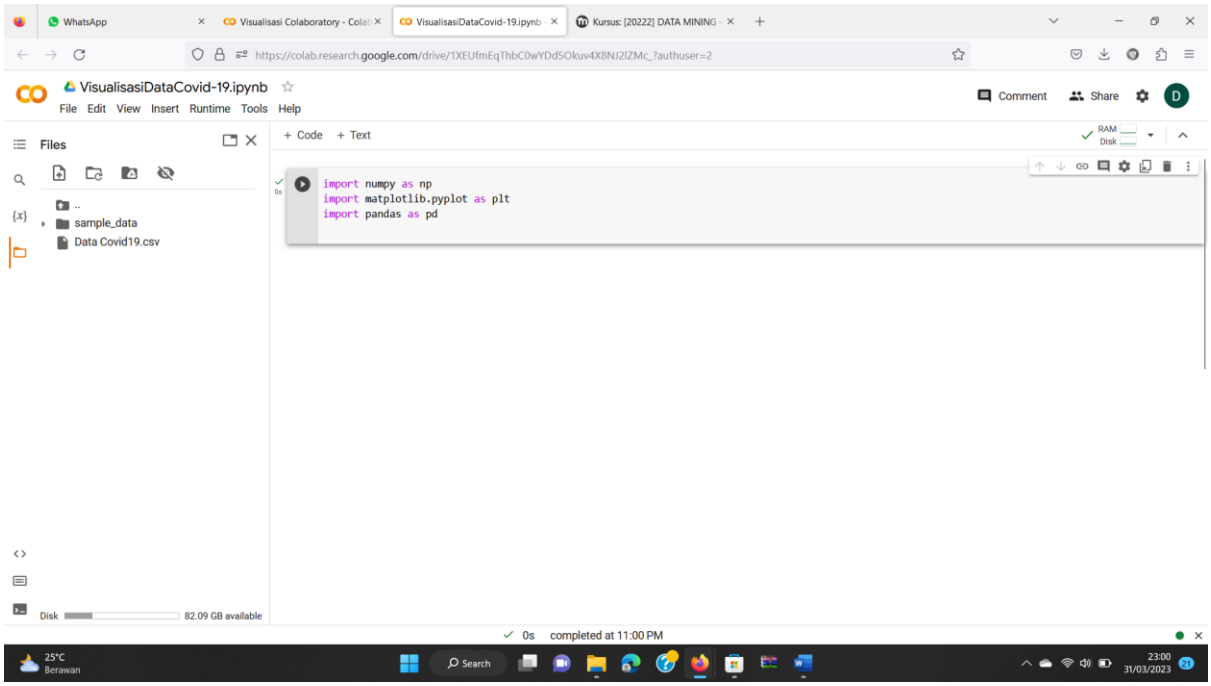
NIM : A11.2020.13191

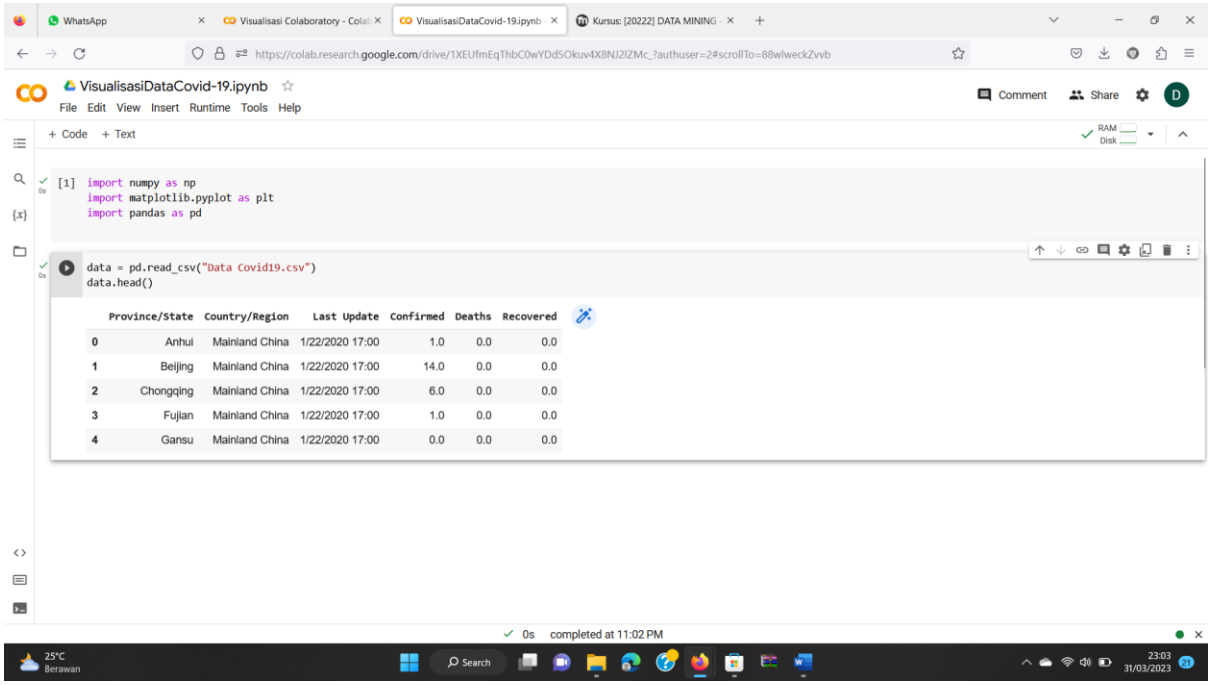
KELAS : A11.4616

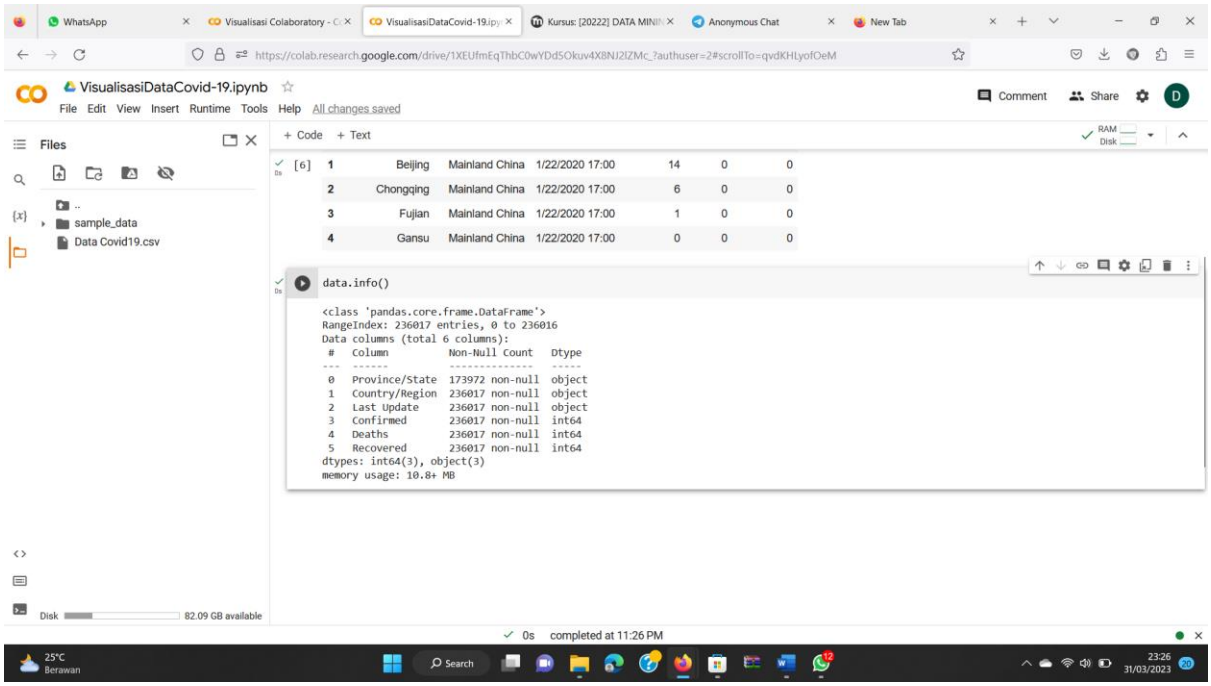
LINK GOOGLE COLAB TUGAS 4

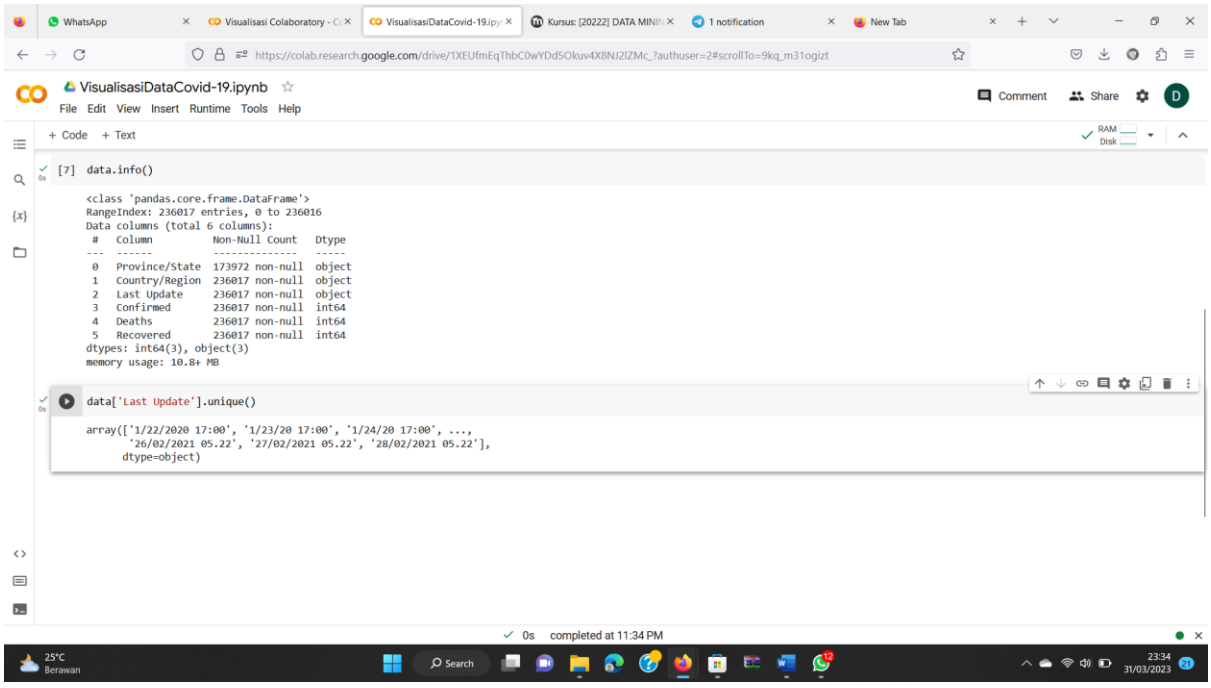
https://colab.research.google.com/drive/1XEUfmEqThbC0wYDd5Okuv4X8NJ2lZMc_?usp=sharing

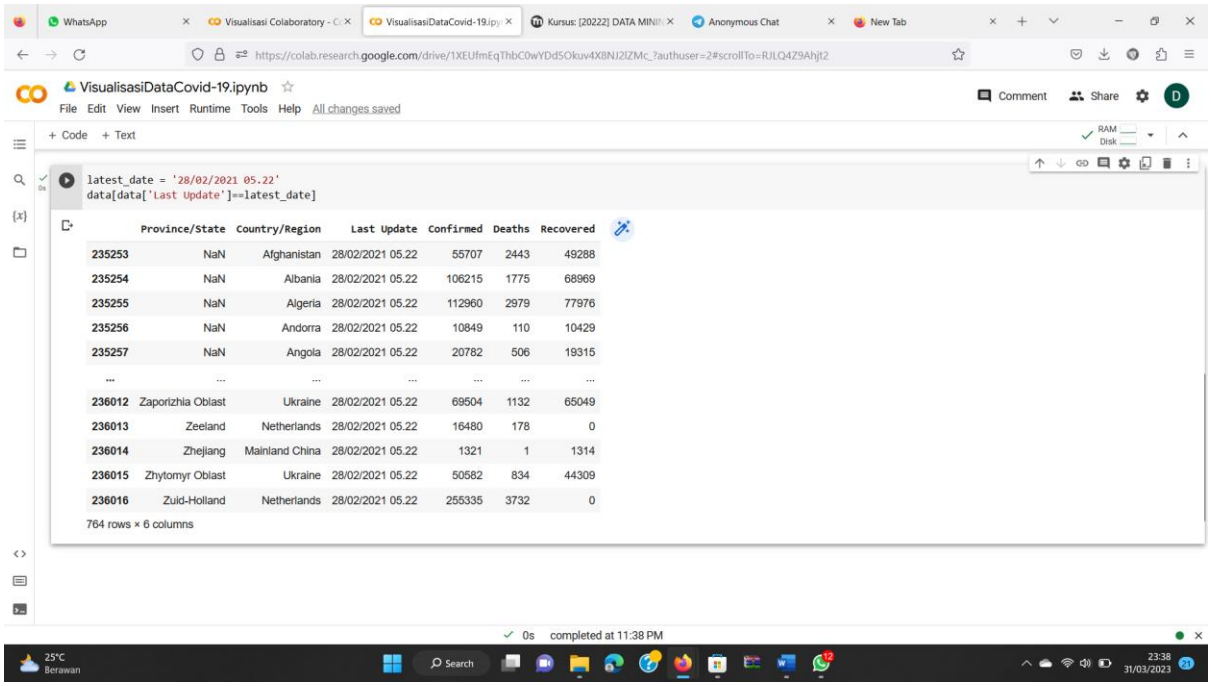



WhatsApp
Visualisasi Colaboratory - Colab
VisualisasiDataCovid-19.ipynb
Kursus: [20222] DATA MINING
https://colab.research.google.com/drive/1XEUfmEqThbC0wYDd5Okuv4X8NJ2lZMc_?authuser=2
VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Files
+ Code + Text
RAM
Disk
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
..
sample_data
Data Covid19.csv
Disk 82.09 GB available
0s completed at 11:00 PM
25°C
Berawan
Search
23:00
31/03/2023


WhatsApp
Visualisasi Colaboratory - Colab
VisualisasiDataCovid-19.ipynb
Kursus: [20222] DATA MINING
https://colab.research.google.com/drive/1XEUfmEqThbC0wYDd5Okuv4X8NJ2lZMc_?authuser=2#scrollTo=88wlweckZvvb
VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
+ Code + Text
RAM
Disk
[1] import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
data = pd.read_csv("Data Covid19.csv")
data.head()
Province/State Country/Region Last Update Confirmed Deaths Recovered
0 Anhui Mainland China 1/22/2020 17:00 1.0 0.0 0.0
1 Beijing Mainland China 1/22/2020 17:00 14.0 0.0 0.0
2 Chongqing Mainland China 1/22/2020 17:00 6.0 0.0 0.0
3 Fujian Mainland China 1/22/2020 17:00 1.0 0.0 0.0
4 Gansu Mainland China 1/22/2020 17:00 0.0 0.0 0.0
0s completed at 11:02 PM
25°C
Berawan
Search
23:03
31/03/2023


VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
Data Covid19.csv
+ Code + Text
[6]
1 Beijing Mainland China 1/22/2020 17:00 14 0 0
2 Chongqing Mainland China 1/22/2020 17:00 6 0 0
3 Fujian Mainland China 1/22/2020 17:00 1 0 0
4 Gansu Mainland China 1/22/2020 17:00 0 0 0
data.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 236017 entries, 0 to 236016
Data columns (total 6 columns):
Column Non-Null Count Dtype
0 Province/State 173972 non-null object
1 Country/Region 236017 non-null object
2 Last Update 236017 non-null object
3 Confirmed 236017 non-null int64
4 Deaths 236017 non-null int64
5 Recovered 236017 non-null int64
dtypes: int64(3), object(3)
memory usage: 10.8+ MB
Disk 82.09 GB available
0s completed at 11:26 PM


VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
+ Code + Text
[7] data.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 236017 entries, 0 to 236016
Data columns (total 6 columns):
Column Non-Null Count Dtype
0 Province/State 173972 non-null object
1 Country/Region 236017 non-null object
2 Last Update 236017 non-null object
3 Confirmed 236017 non-null int64
4 Deaths 236017 non-null int64
5 Recovered 236017 non-null int64
dtypes: int64(3), object(3)
memory usage: 10.8+ MB
data['Last Update'].unique()
array(['1/22/2020 17:00', '1/23/20 17:00', '1/24/20 17:00', ...,
'26/02/2021 05.22', '27/02/2021 05.22', '28/02/2021 05.22'],
dtype=object)
0s completed at 11:34 PM



VisualisasiDataCovid-19.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

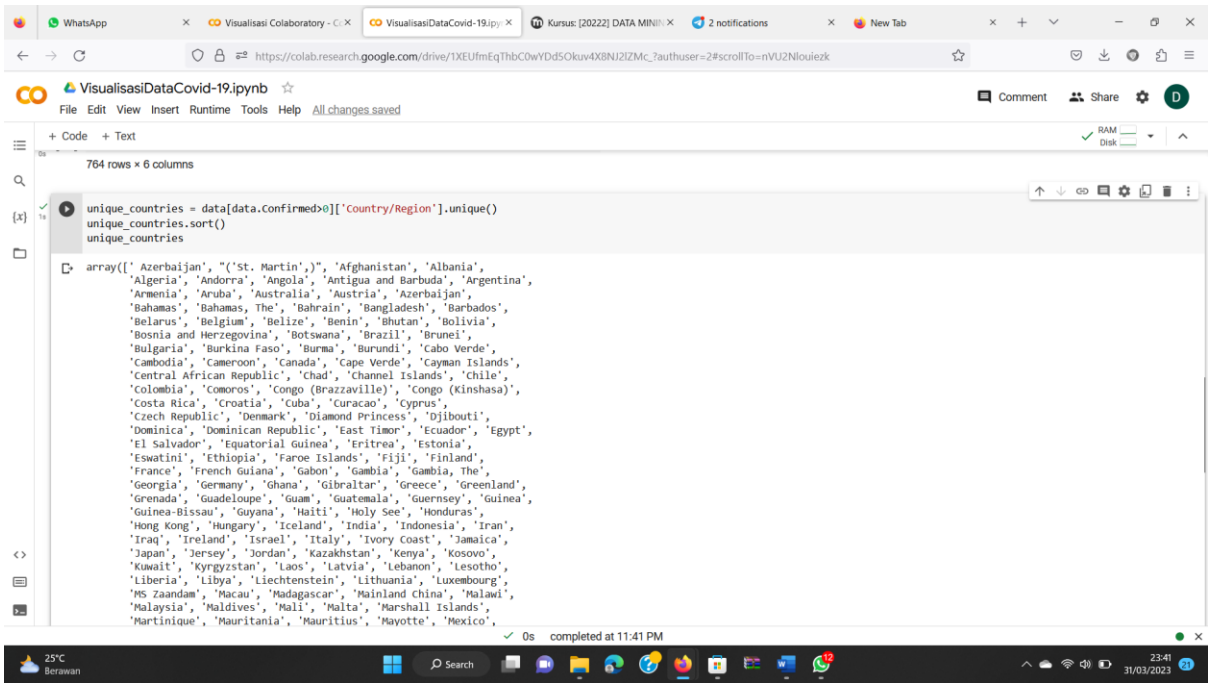
+ Code + Text

```
latest_date = '28/02/2021 05.22'
data[data['Last Update']==latest_date]
```

	Province/State	Country/Region	Last Update	Confirmed	Deaths	Recovered
235253	NaN	Afghanistan	28/02/2021 05.22	55707	2443	49288
235254	NaN	Albania	28/02/2021 05.22	106215	1775	68969
235255	NaN	Algeria	28/02/2021 05.22	112960	2979	77976
235256	NaN	Andorra	28/02/2021 05.22	10849	110	10429
235257	NaN	Angola	28/02/2021 05.22	20782	506	19315
...	...	...	...	...	...	...
236012	Zaporizhia Oblast	Ukraine	28/02/2021 05.22	69504	1132	65049
236013	Zeeland	Netherlands	28/02/2021 05.22	16480	178	0
236014	Zhejiang	Mainland China	28/02/2021 05.22	1321	1	1314
236015	Zhytomyr Oblast	Ukraine	28/02/2021 05.22	50582	834	44309
236016	Zuid-Holland	Netherlands	28/02/2021 05.22	255335	3732	0

764 rows × 6 columns

0s completed at 11:38 PM



VisualisasiDataCovid-19.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

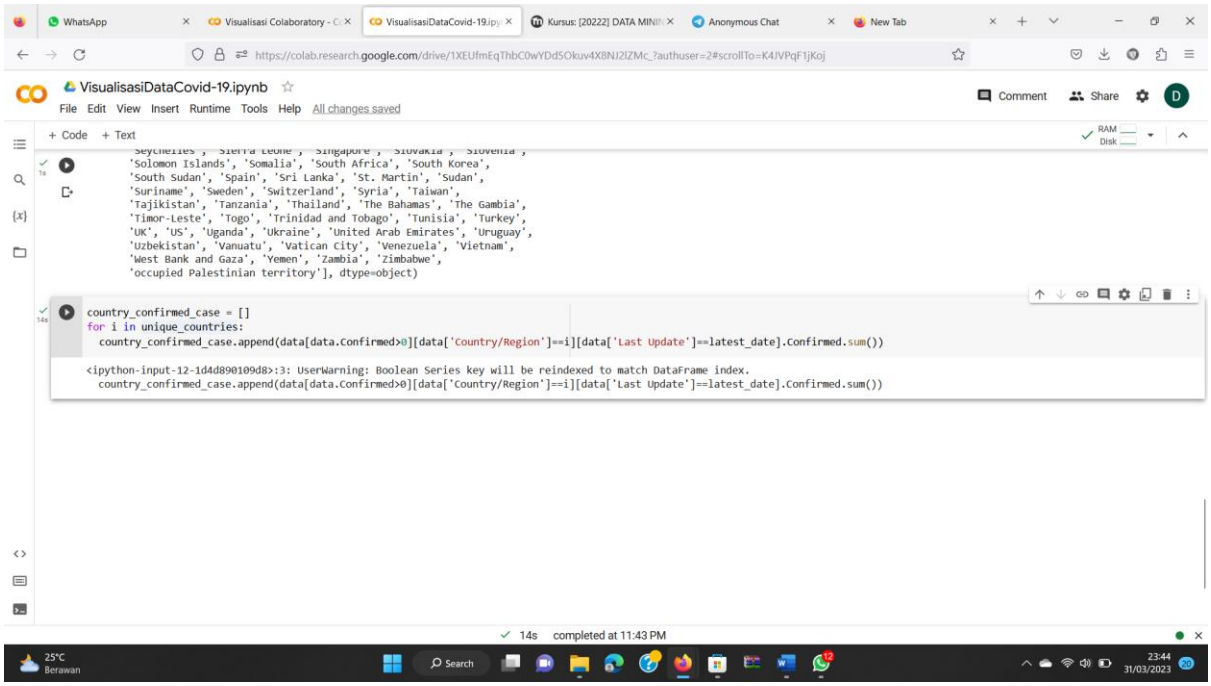
+ Code + Text

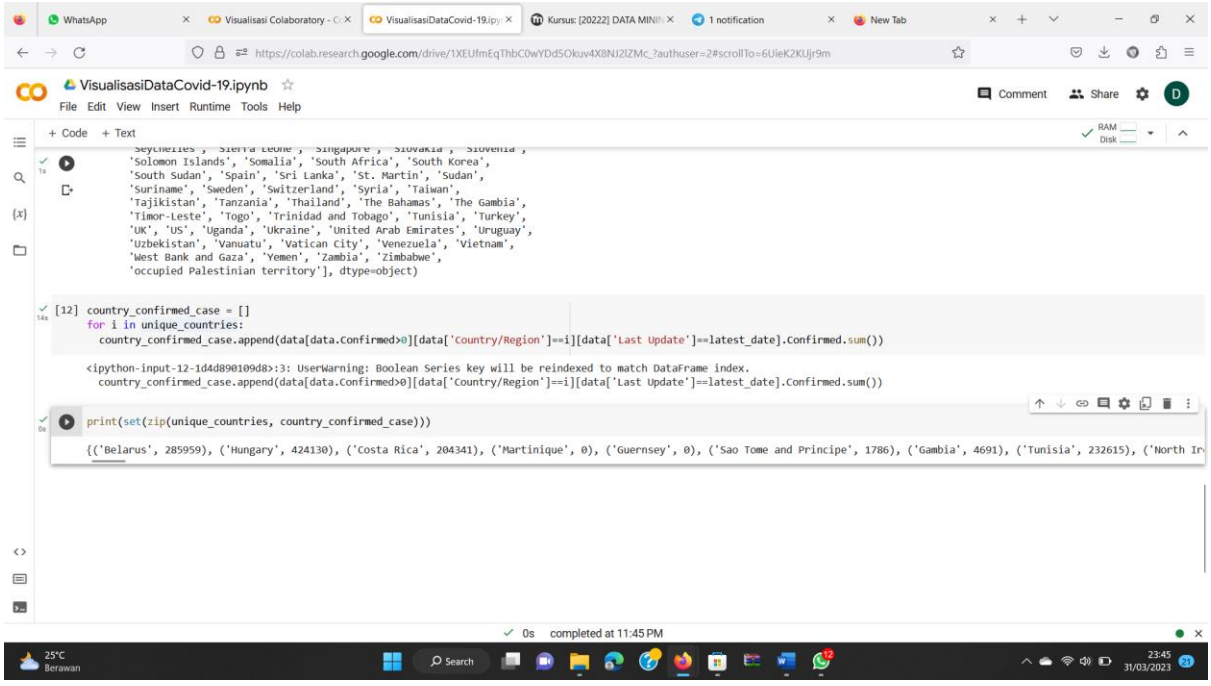
764 rows × 6 columns

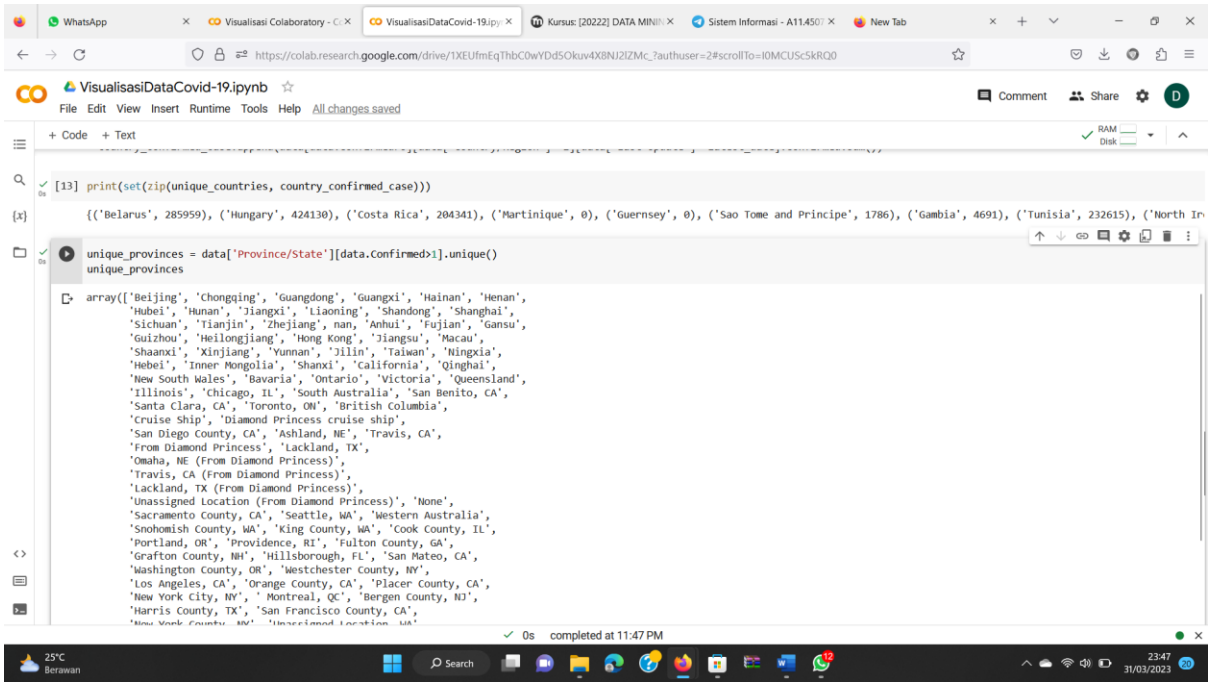
```
unique_countries = data[data.Confirmed>0]['Country/Region'].unique()
unique_countries.sort()
unique_countries
```

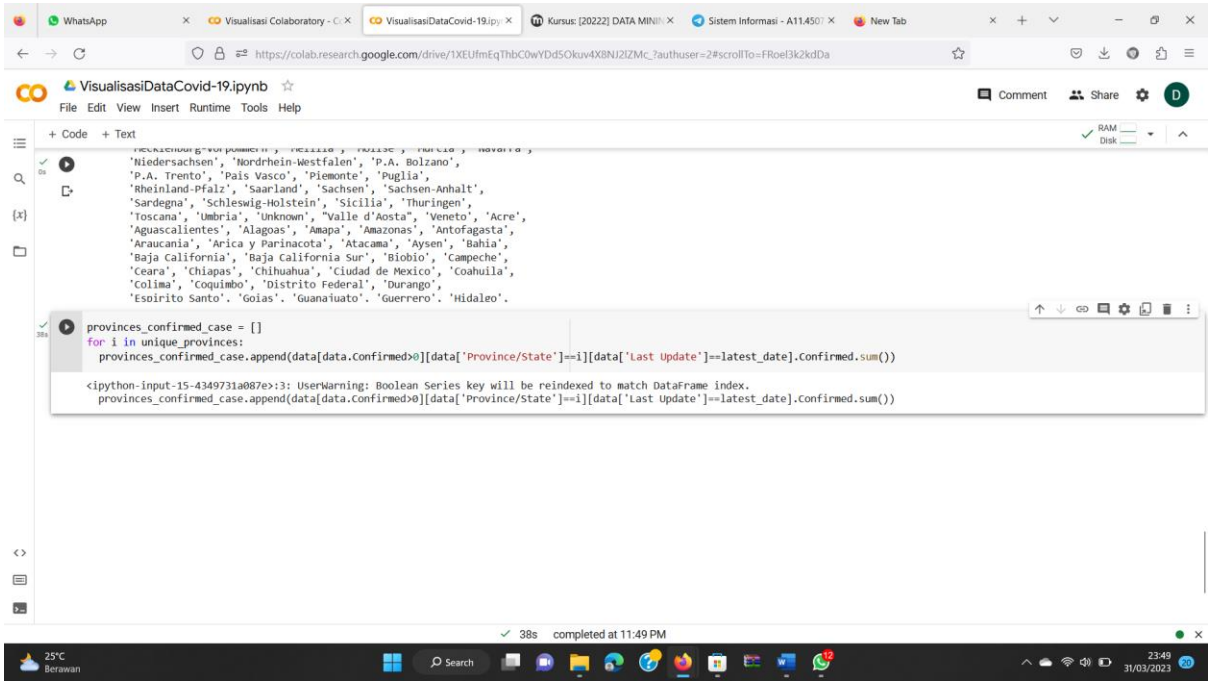
```
array([' Azerbaijan', "('St. Martin',)", 'Afghanistan', 'Albania',
       'Algeria', 'Andorra', 'Angola', 'Antigua and Barbuda', 'Argentina',
       'Armenia', 'Aruba', 'Australia', 'Austria', 'Azerbaijan',
       'Bahamas', 'Bahamas, The', 'Bahrain', 'Bangladesh', 'Barbados',
       'Belarus', 'Belgium', 'Belize', 'Benin', 'Bhutan', 'Bolivia',
       'Bosnia and Herzegovina', 'Botswana', 'Brazil', 'Brunei',
       'Bulgaria', 'Burkina Faso', 'Burma', 'Burundi', 'Cabo Verde',
       'Cambodia', 'Cameroon', 'Canada', 'Cape Verde', 'Cayman Islands',
       'Central African Republic', 'Chad', 'Channel Islands', 'Chile',
       'Colombia', 'Comoros', 'Congo (Brazzaville)', 'Congo (Kinshasa)',
       'Costa Rica', 'Croatia', 'Cuba', 'Curacao', 'Cyprus',
       'Czech Republic', 'Denmark', 'Diamond Princess', 'Djibouti',
       'Dominica', 'Dominican Republic', 'East Timor', 'Ecuador', 'Egypt',
       'El Salvador', 'Equatorial Guinea', 'Eritrea', 'Estonia',
       'Eswatini', 'Ethiopia', 'Faroe Islands', 'Fiji', 'Finland',
       'France', 'French Guiana', 'Gabon', 'Gambia', 'Gambia, The',
       'Georgia', 'Germany', 'Ghana', 'Gibraltar', 'Greece', 'Greenland',
       'Grenada', 'Guadeloupe', 'Guam', 'Guatemala', 'Guernsey', 'Guinea',
       'Guinea-Bissau', 'Guyana', 'Haiti', 'Holy See', 'Honduras',
       'Hong Kong', 'Hungary', 'Iceland', 'India', 'Indonesia', 'Iran',
       'Iraq', 'Ireland', 'Israel', 'Italy', 'Ivory Coast', 'Jamaica',
       'Japan', 'Jersey', 'Jordan', 'Kazakhstan', 'Kenya', 'Kosovo',
       'Kuwait', 'Kyrgyzstan', 'Laos', 'Latvia', 'Lebanon', 'Lesotho',
       'Liberia', 'Libya', 'Liechtenstein', 'Lithuania', 'Luxembourg',
       'MS Zaandam', 'Macau', 'Madagascar', 'Mainland China', 'Malawi',
       'Malaysia', 'Maldives', 'Mali', 'Malta', 'Marshall Islands',
       'Martinique', 'Mauritania', 'Mauritius', 'Mayotte', 'Mexico',
```

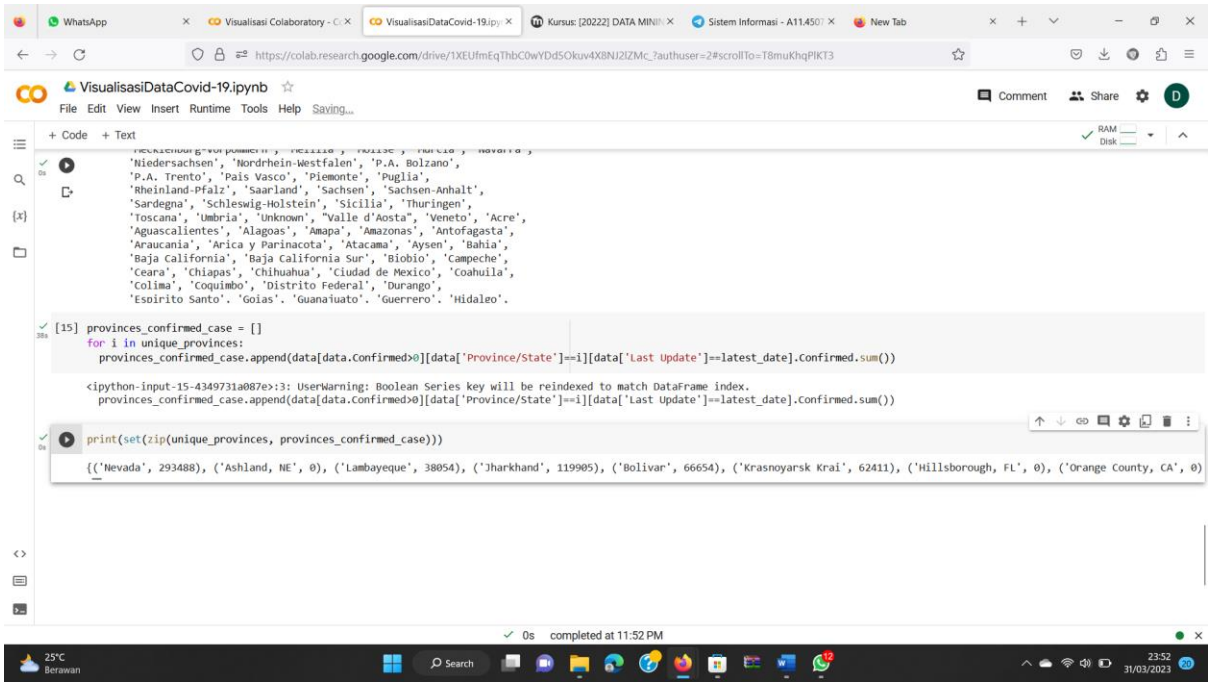
0s completed at 11:41 PM

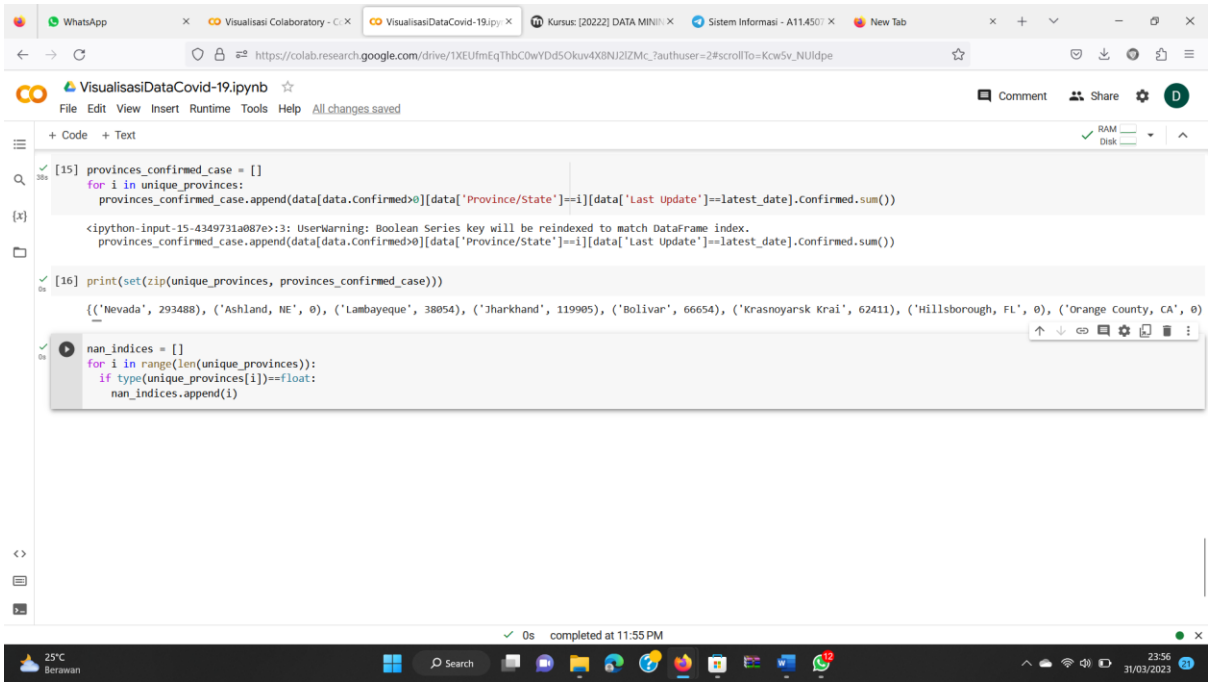

VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
'Solomon Islands', 'Somalia', 'South Africa', 'South Korea',
'South Sudan', 'Spain', 'Sri Lanka', 'St. Martin', 'Sudan',
'Suriname', 'Sweden', 'Switzerland', 'Syria', 'Taiwan',
'Tajikistan', 'Tanzania', 'Thailand', 'The Bahamas', 'The Gambia',
'Timor-Leste', 'Togo', 'Trinidad and Tobago', 'Tunisia', 'Turkey',
'UK', 'US', 'Uganda', 'Ukraine', 'United Arab Emirates', 'Uruguay',
'Uzbekistan', 'Vanuatu', 'Vatican City', 'Venezuela', 'Vietnam',
'West Bank and Gaza', 'Yemen', 'Zambia', 'Zimbabwe',
'occupied Palestinian territory'], dtype=object)
country_confirmed_case = []
for i in unique_countries:
country_confirmed_case.append(data[data.Confirmed>0][data['Country/Region']==i][data['Last Update']==latest_date].Confirmed.sum())
<ipython-input-12-1d4d890109d8>:3: UserWarning: Boolean Series key will be reindexed to match DataFrame index.
country_confirmed_case.append(data[data.Confirmed>0][data['Country/Region']==i][data['Last Update']==latest_date].Confirmed.sum())
14s completed at 11:43 PM
25°C Berawan
Search
23:44 31/03/2023

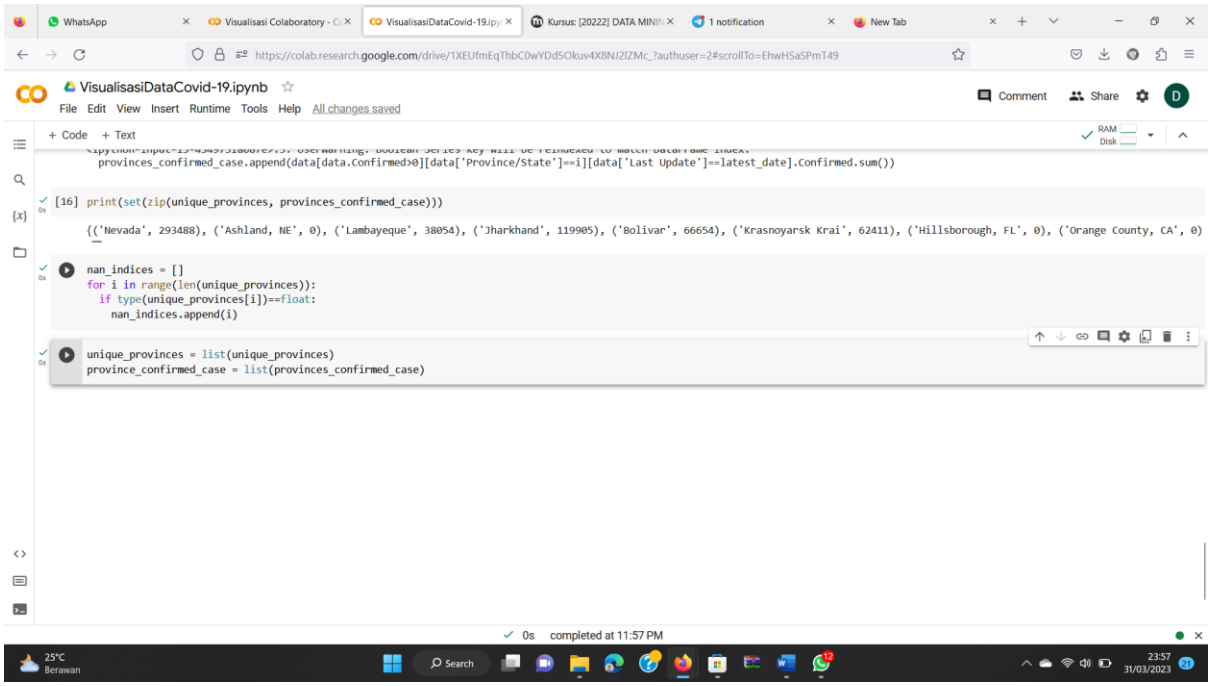

VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
+ Code + Text
'Solomon Islands', 'Somalia', 'South Africa', 'South Korea',
'South Sudan', 'Spain', 'Sri Lanka', 'St. Martin', 'Sudan',
'Suriname', 'Sweden', 'Switzerland', 'Syria', 'Taiwan',
'Tajikistan', 'Tanzania', 'Thailand', 'The Bahamas', 'The Gambia',
'Timor-Leste', 'Togo', 'Trinidad and Tobago', 'Tunisia', 'Turkey',
'UK', 'US', 'Uganda', 'Ukraine', 'United Arab Emirates', 'Uruguay',
'Uzbekistan', 'Vanuatu', 'Vatican City', 'Venezuela', 'Vietnam',
'West Bank and Gaza', 'Yemen', 'Zambia', 'Zimbabwe',
'occupied Palestinian territory'], dtype=object)
[12] country_confirmed_case = []
for i in unique_countries:
country_confirmed_case.append(data[data.Confirmed>0][data['Country/Region']==i][data['Last Update']==latest_date].Confirmed.sum())
<ipython-input-12-1d4d890109d8>:3: UserWarning: Boolean Series key will be reindexed to match DataFrame index.
country_confirmed_case.append(data[data.Confirmed>0][data['Country/Region']==i][data['Last Update']==latest_date].Confirmed.sum())
print(set(zip(unique_countries, country_confirmed_case)))
{('Belarus', 285959), ('Hungary', 424130), ('Costa Rica', 204341), ('Martinique', 0), ('Guernsey', 0), ('Sao Tome and Principe', 1786), ('Gambia', 4691), ('Tunisia', 232615), ('North Ir
0s completed at 11:45 PM
25°C Berawan
Search
23:45 31/03/2023

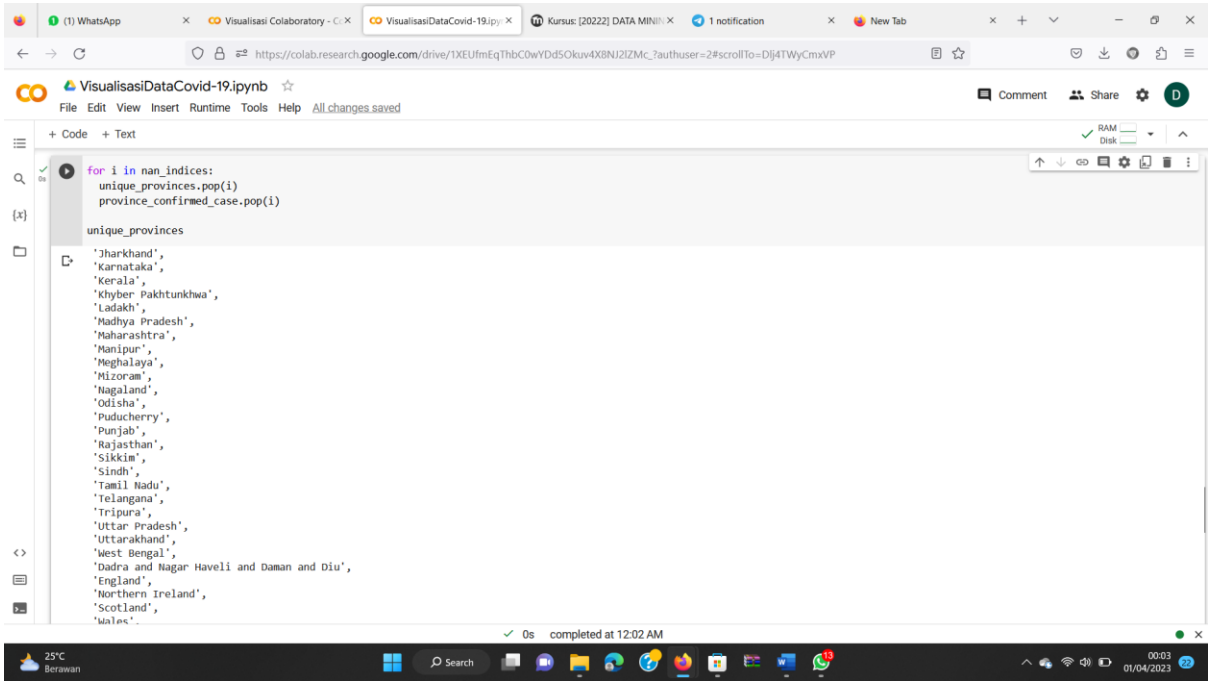

VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
[13] print(set(zip(unique_countries, country_confirmed_case)))
{('Belarus', 285959), ('Hungary', 424130), ('Costa Rica', 204341), ('Martinique', 0), ('Guernsey', 0), ('Sao Tome and Principe', 1786), ('Gambia', 4691), ('Tunisia', 232615), ('North Ir
unique_provinces = data['Province/State'][data.Confirmed>1].unique()
unique_provinces
array(['Beijing', 'Chongqing', 'Guangdong', 'Guangxi', 'Hainan', 'Henan',
'Hubei', 'Hunan', 'Jiangxi', 'Liaoning', 'Shandong', 'Shanghai',
'Sichuan', 'Tianjin', 'Zhejiang', nan, 'Anhui', 'Fujian', 'Gansu',
'Guizhou', 'Heilongjiang', 'Hong Kong', 'Jiangsu', 'Macau',
'Shaanxi', 'Xinjiang', 'Yunnan', 'Jilin', 'Taiwan', 'Ningxia',
'Hebei', 'Inner Mongolia', 'Shanxi', 'California', 'Qinghai',
'New South Wales', 'Bavaria', 'Ontario', 'Victoria', 'Queensland',
'Illinois', 'Chicago, IL', 'South Australia', 'San Benito, CA',
'Santa Clara, CA', 'Toronto, ON', 'British Columbia',
'Cruise Ship', 'Diamond Princess cruise ship',
'San Diego County, CA', 'Ashland, NE', 'Travis, CA',
'From Diamond Princess', 'Lackland, TX',
'Omaha, NE (From Diamond Princess)',
'Travis, CA (From Diamond Princess)',
'Lackland, TX (From Diamond Princess)',
'Unassigned Location (From Diamond Princess)', 'None',
'Sacramento County, CA', 'Seattle, WA', 'Western Australia',
'Snohomish County, WA', 'King County, WA', 'Cook County, IL',
'Portland, OR', 'Providence, RI', 'Fulton County, GA',
'Grafton County, NH', 'Hillsborough, FL', 'San Mateo, CA',
'Washington County, OR', 'Westchester County, NY',
'Los Angeles, CA', 'Orange County, CA', 'Placer County, CA',
'New York City, NY', ' Montreal, QC', 'Bergen County, NJ',
'Harris County, TX', 'San Francisco County, CA',
0s completed at 11:47 PM
25°C Berawan
Search
23:47 31/03/2023

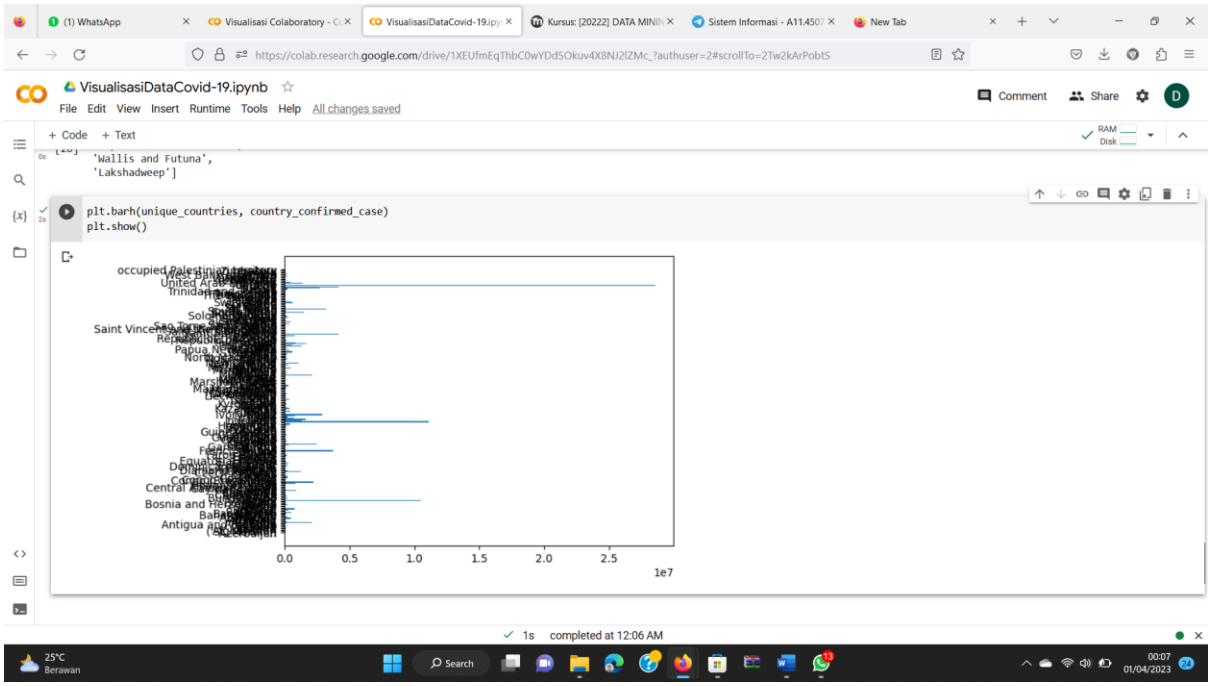

VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
+ Code + Text
'Niedersachsen', 'Nordrhein-Westfalen', 'P.A. Bolzano',
'P.A. Trento', 'Pais Vasco', 'Piemonte', 'Puglia',
'Rheinland-Pfalz', 'Saarland', 'Sachsen', 'Sachsen-Anhalt',
'Sardegna', 'Schleswig-Holstein', 'Sicilia', 'Thuringen',
'Toscana', 'Umbria', 'Unknown', "Valle d'Aosta", 'Veneto', 'Acre',
'Aguascalientes', 'Alagoas', 'Amapa', 'Amazonas', 'Antofagasta',
'Araucania', 'Arica y Parinacota', 'Atacama', 'Aysen', 'Bahia',
'Baja California', 'Baja California Sur', 'Biobio', 'Campeche',
'Ceara', 'Chiapas', 'Chihuahua', 'Ciudad de Mexico', 'Coahuila',
'Colima', 'Coquimbo', 'Distrito Federal', 'Durango',
'Esoirito Santo', 'Goias', 'Guanajuato', 'Guerrero', 'Hidalgo',
provinces_confirmed_case = []
for i in unique_provinces:
provinces_confirmed_case.append(data[data.Confirmed>0][data['Province/State']==i][data['Last Update']==latest_date].Confirmed.sum())
<ipython-input-15-4349731a087e>:3: UserWarning: Boolean Series key will be reindexed to match DataFrame index.
provinces_confirmed_case.append(data[data.Confirmed>0][data['Province/State']==i][data['Last Update']==latest_date].Confirmed.sum())
38s completed at 11:49 PM
25°C Berawan
Search
23:49 31/03/2023

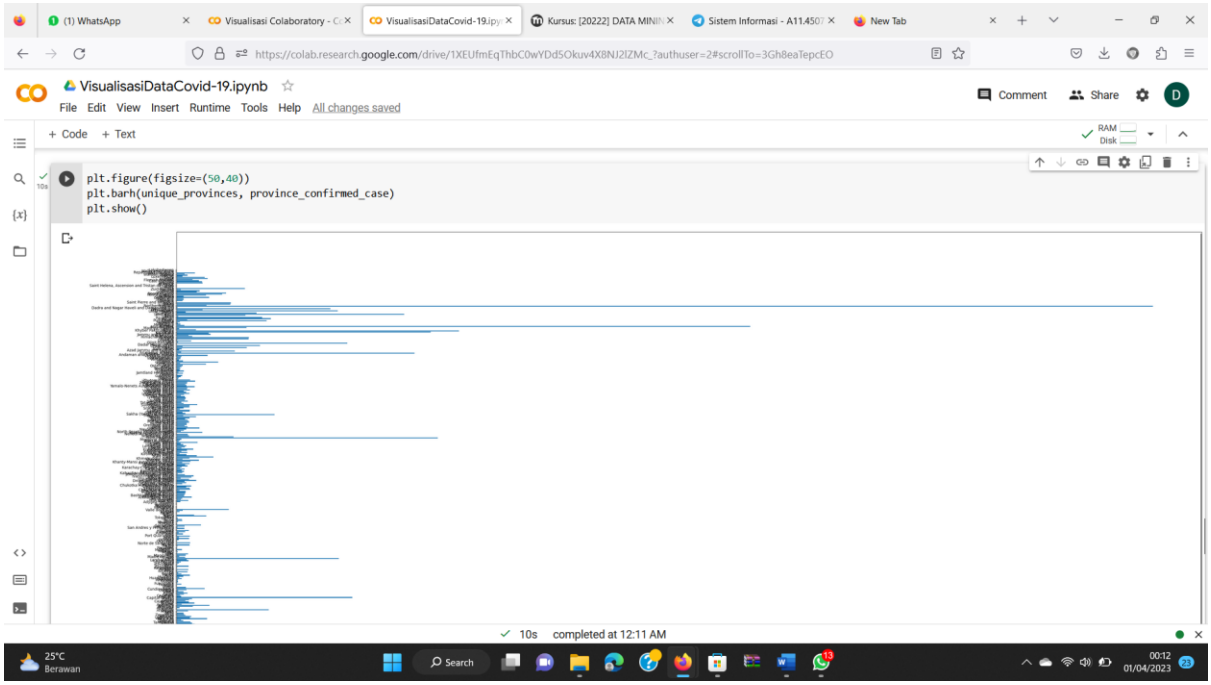

VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
+ Code + Text
'Niedersachsen', 'Nordrhein-Westfalen', 'P.A. Bolzano',
'P.A. Trento', 'Pais Vasco', 'Piemonte', 'Puglia',
'Rheinland-Pfalz', 'Saarland', 'Sachsen', 'Sachsen-Anhalt',
'Sardegna', 'Schleswig-Holstein', 'Sicilia', 'Thuringen',
'Toscana', 'Umbria', 'Unknown', 'Valle d'Aosta", 'Veneto', 'Acre',
'Aguascalientes', 'Alagoas', 'Amapa', 'Amazonas', 'Antofagasta',
'Araucania', 'Arica y Parinacota', 'Atacama', 'Aysen', 'Bahia',
'Baja California', 'Baja California Sur', 'Biobio', 'Campeche',
'Ceara', 'Chiapas', 'Chihuahua', 'Ciudad de Mexico', 'Coahuila',
'Colima', 'Coquimbo', 'Distrito Federal', 'Durango',
'Espirito Santo', 'Goias', 'Guanajuato', 'Guerrero', 'Hidalgo',
[15] provinces_confirmed_case = []
for i in unique_provinces:
provinces_confirmed_case.append(data[data.Confirmed>0][data['Province/State']==i][data['Last Update']==latest_date].Confirmed.sum())
<ipython-input-15-4349731a087e>:3: UserWarning: Boolean Series key will be reindexed to match DataFrame index.
provinces_confirmed_case.append(data[data.Confirmed>0][data['Province/State']==i][data['Last Update']==latest_date].Confirmed.sum())
print(set(zip(unique_provinces, provinces_confirmed_case)))
{('Nevada', 293488), ('Ashland, NE', 0), ('Lambayeque', 38054), ('Jharkhand', 119905), ('Bolivar', 66654), ('Krasnoyarsk Krai', 62411), ('Hillsborough, FL', 0), ('Orange County, CA', 0)
0s completed at 11:52 PM

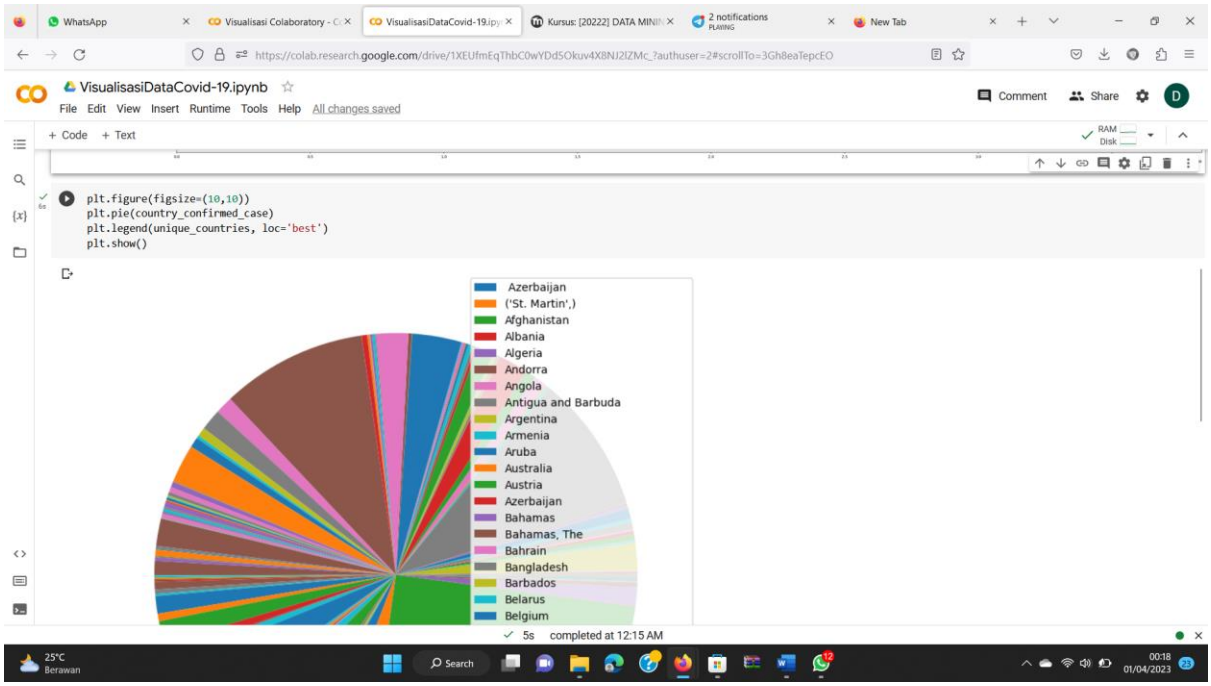

VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
[15] provinces_confirmed_case = []
for i in unique_provinces:
provinces_confirmed_case.append(data[data.Confirmed>0][data['Province/State']==i][data['Last Update']==latest_date].Confirmed.sum())
<ipython-input-15-4349731a087e>:3: UserWarning: Boolean Series key will be reindexed to match DataFrame index.
provinces_confirmed_case.append(data[data.Confirmed>0][data['Province/State']==i][data['Last Update']==latest_date].Confirmed.sum())
[16] print(set(zip(unique_provinces, provinces_confirmed_case)))
{('Nevada', 293488), ('Ashland, NE', 0), ('Lambayeque', 38054), ('Jharkhand', 119905), ('Bolivar', 66654), ('Krasnoyarsk Krai', 62411), ('Hillsborough, FL', 0), ('Orange County, CA', 0)
nan_indices = []
for i in range(len(unique_provinces)):
if type(unique_provinces[i])==float:
nan_indices.append(i)
0s completed at 11:55 PM

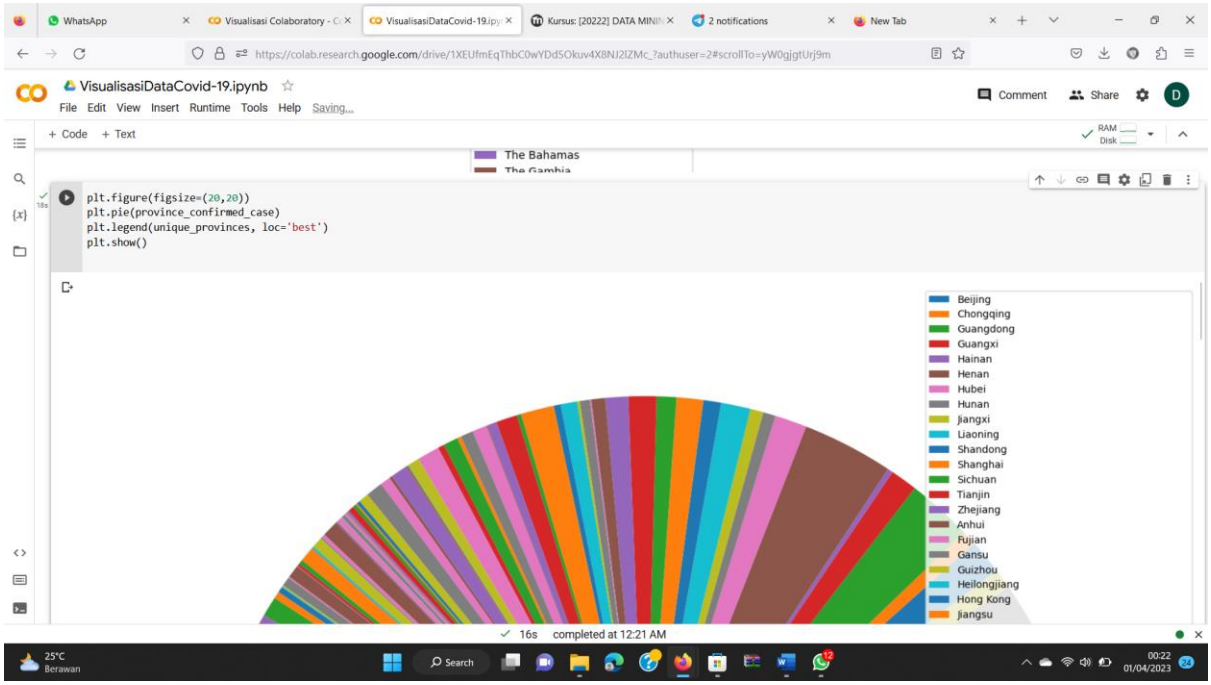

VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
provinces_confirmed_case.append(data[data.Confirmed>0][data['Province/State']==i][data['Last Update']==latest_date].Confirmed.sum())
[16] print(set(zip(unique_provinces, provinces_confirmed_case)))
{('Nevada', 293488), ('Ashland, NE', 0), ('Lambayeque', 38054), ('Jharkhand', 119905), ('Bolivar', 66654), ('Krasnoyarsk Krai', 62411), ('Hillsborough, FL', 0), ('Orange County, CA', 0)
nan_indices = []
for i in range(len(unique_provinces)):
if type(unique_provinces[i])==float:
nan_indices.append(i)
unique_provinces = list(unique_provinces)
province_confirmed_case = list(provinces_confirmed_case)
0s completed at 11:57 PM


VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
for i in nan_indices:
unique_provinces.pop(i)
province_confirmed_case.pop(i)
unique_provinces
'Jharkhand',
'Karnataka',
'Kerala',
'Khyber Pakhtunkhwa',
'Ladakh',
'Madhya Pradesh',
'Maharashtra',
'Manipur',
'Meghalaya',
'Mizoram',
'Nagaland',
'Odisha',
'Puducherry',
'Punjab',
'Rajasthan',
'Sikkim',
'Sindh',
'Tamil Nadu',
'Telangana',
'Tripura',
'Uttar Pradesh',
'Uttarakhand',
'West Bengal',
'Dadra and Nagar Haveli and Daman and Diu',
'England',
'Northern Ireland',
'Scotland',
'Wales'
0s completed at 12:02 AM


VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
'Wallis and Futuna',
'Lakshadweep']
plt.barh(unique_countries, country_confirmed_case)
plt.show()
occupied Palestinian territory
United Arab Emirates
Trinidad and Tobago
Saint Vincent and the Grenadines
Papua New Guinea
Marshall Islands
Guinea-Bissau
Equatorial Guinea
Dominican Republic
Central African Republic
Bosnia and Herzegovina
Antigua and Barbuda
0.0
0.5
1.0
1.5
2.0
2.5
1e7
1s completed at 12:06 AM
25°C
Berawan
Search
00:07
01/04/2023


VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
plt.figure(figsize=(50,40))
plt.barh(unique_provinces, province_confirmed_case)
plt.show()
10s completed at 12:11 AM
25°C
Berawan
Search
00:12
01/04/2023


VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
plt.figure(figsize=(10,10))
plt.pie(country_confirmed_case)
plt.legend(unique_countries, loc='best')
plt.show()
Azerbaijan
('St. Martin',)
Afghanistan
Albania
Algeria
Andorra
Angola
Antigua and Barbuda
Argentina
Armenia
Aruba
Australia
Austria
Azerbaijan
Bahamas
Bahamas, The
Bahrain
Bangladesh
Barbados
Belarus
Belgium
5s completed at 12:15 AM


VisualisasiDataCovid-19.ipynb
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
+ Code + Text
The Bahamas
plt.figure(figsize=(20,20))
plt.pie(province_confirmed_case)
plt.legend(unique_provinces, loc='best')
plt.show()
Beijing
Chongqing
Guangdong
Guangxi
Hainan
Henan
Hubei
Hunan
Jiangxi
Liaoning
Shandong
Shanghai
Sichuan
Tianjin
Zhejiang
Anhui
Fujian
Gansu
Guizhou
Heilongjiang
Hong Kong
Jiangsu
16s completed at 12:21 AM