



COVID-19 (NOVEMBER 2023)

Data visualisation

ABSTRACT

Summarized analysis with pictures, charts, and a dashboard about Coronavirus data from WHO. Only with 6 illustrations, it is clearly and fast understandable the effects and progress by the novel virus in the planet Earth.

YIT 19488399 (Tony)

Data Visualization - CMP020X302A

Module Tutor: Mohammad Ahmad

BSc Computer science, year 3rd (autumn term).

Task 1: questions that a visualization will answer.

- ❖ From “vaccination data” dataset:
 1. Number of different vaccines used by region and ALL countries that used 11 and 12 different types of vaccines.
 2. Double bars plot about the speed and amount of vaccines administration to people in different countries (11 different countries in each bar plot).

- ❖ From “latest reported counts of cases and deaths” dataset:
 3. Total number of deaths by country with top100 countries. Looks like BUBLES but code is SCATTER.
 4. Total cases cumulative death by continent (region). Percentage information in each piece of the chart. PIE.
 5. Confirmed cases per country and region by days (24 hours timeframe). MAP/INTERACTIVE MAP.

NOTES: each python code has at the beginning a commented link. This is the reference of the source of my code.

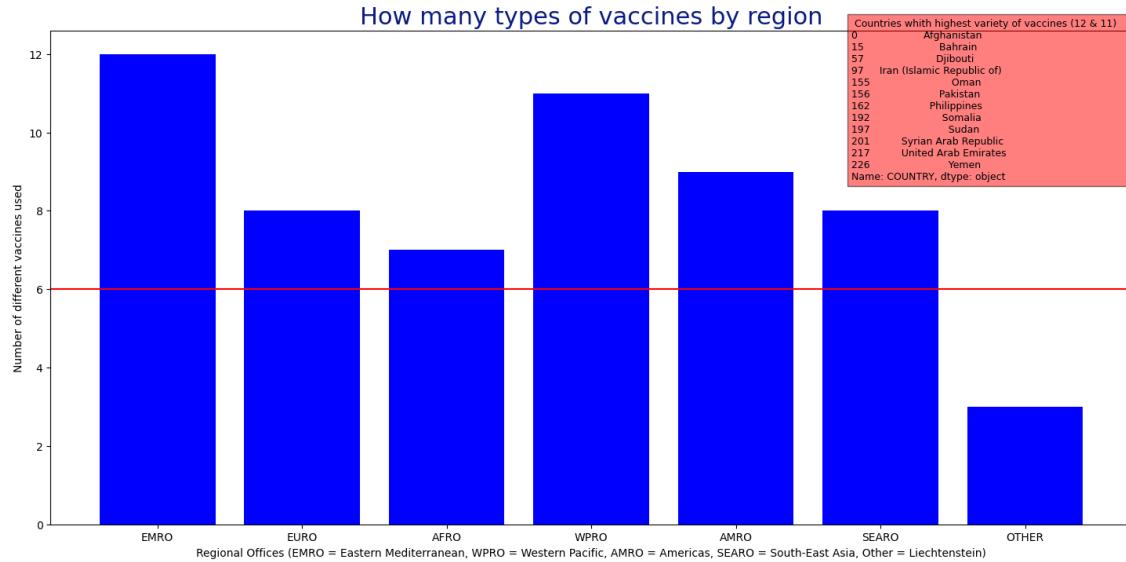
Task 2: TABLEAU or POWER BI and follow instructions (one dashboard with five dynamic visualizations to analyze data).

A dashboard with 8 worksheets, 1 picture and short text (2 references) to explain coronavirus data from WHO. Additional dataset used from official WHO website, plus dataset provided from module in Moodle. Datasets are recent, but not live updated.

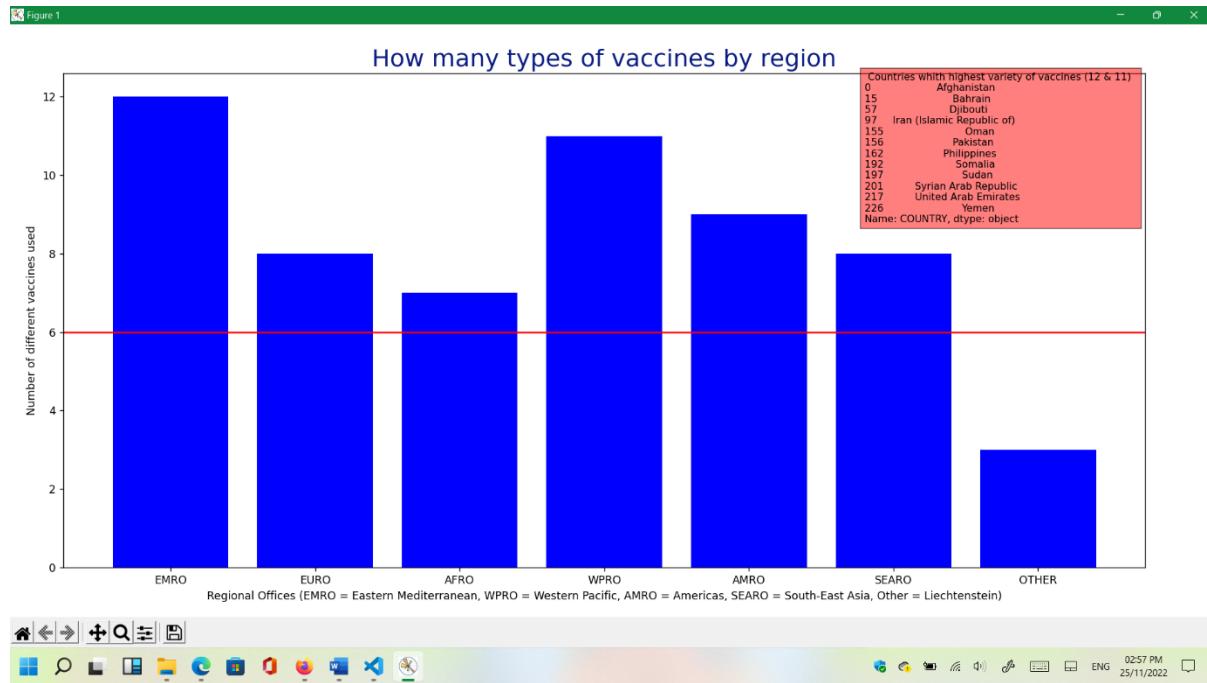
All dashboard is interactive by region and country selection in the “Highlight” area.

Selecting country anywhere in the dashboard will have same effects. Analysis and data visualization implemented in Tableau Public software and online because it is free to use and save the work can be done in the cloud (online).

First question: Number of different vaccines used by country or region



Output from running python code from VSC and screenshot. Both images are same data visualization, but different formats.



Python code follows with screenshot included. In this occasion, code output gives some information on terminal for familiarize with data (.csv):

```

import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

data = pd.read_csv("vaccination-data.csv") #reading the database
print(data.head(), data.columns)
integer_vaccines_type = data["NUMBER_VACCINES_TYPES_USED"]
integer_vaccines_type.fillna(0)
print(integer_vaccines_type, len(integer_vaccines_type), integer_vaccines_type.unique())
region = data["WHO_REGION"]
print(region, len(region), region.unique())
# print(max(data["NUMBER_VACCINES_TYPES_USED"]))
print(data['NUMBER_VACCINES_TYPES_USED'].nlargest(n=12)) #untill 11 inclusive
print(data.loc[data['NUMBER_VACCINES_TYPES_USED']>10, 'COUNTRY'])
top_variety = data.loc[data['NUMBER_VACCINES_TYPES_USED']>10, 'COUNTRY']
print(top_variety, len(top_variety))
top_varietylist = np.array(top_variety)
fig, axe = plt.subplots(figsize=(14, 7))
plt.bar(data["WHO_REGION"], data["NUMBER_VACCINES_TYPES_USED"], color = "blue")
plt.xlabel('Regional Offices (EMRO = Eastern Mediterranean, WPRO = Western Pacific, AMRO = Americas, SEARO = South-East Asia, Other = Liechtenstein)')
plt.ylabel('Number of different vaccines used')
plt.title("How many types of vaccines by region", color = "#091b82", fontsize = 22)
plt.axhline(y = 6, color = 'r', linestyle = '-')
axe.text(4.8, 8.8, f" Countries whith highest variety of vaccines (12 & 11) \n{top_variety}", bbox=dict(facecolor='red', alpha=0.5), fontsize = 9)
plt.tight_layout(pad=1.1, w_pad=0.5, h_pad=0.2)
plt.show()
fig.savefig("output1.png")
plt.close(fig)

```

The screenshot shows the Visual Studio Code interface with the following details:

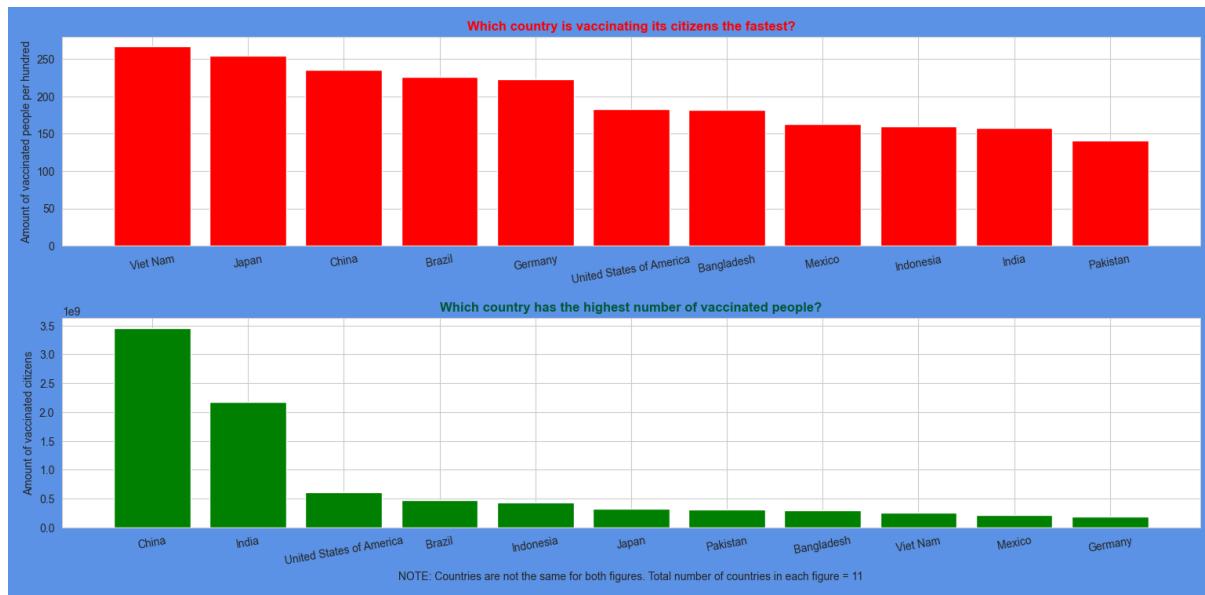
- File Explorer:** Shows files in the DV_COURSEWORK folder, including q1DV.py, q2DV.py, q3DV.py, q4DV.py, q5DV.py, output1.png, output2.png, output3.png, output4.png, Task1and2.docx, vaccination-data.csv, and WHO-COVID-19-global-...
- Code Editor:** Displays the Python script q1DV.py.
- Terminal:** Shows the output of the script, listing countries and their vaccination data. The output includes:

```

97  Iran (Islamic Republic of)
155      Oman
156      Pakistan
162      Philippines
192      Somalia
197      Sudan
201  Syrian Arab Republic
217      United Arab Emirates
226      Yemen
Name: COUNTRY, dtype: object 12
PS C:\Users\Aldin\OneDrive\Desktop\DV_Coursework>

```

Second question: Which countries started vaccinations first (speed) and which countries has the highest amount of vaccinated people?



```
# https://towardsdatascience.com/covid-19-vaccination-progress-analysis-around-the-world-736d7e57f198
import matplotlib.pyplot as plt
import seaborn as sns
```

```

import pandas as pd

sns.set_style('whitegrid')
df = pd.read_csv("vaccination-data.csv")
plt.figure(1, figsize=(15, 7), facecolor="#5b92e5")
plt.subplot(211)
cols = ['COUNTRY', 'TOTAL_VACCINATIONS', 'ISO3',
'VACCINES_USED', 'TOTAL_VACCINATIONS_PER100']
vacc_amount = df[cols].groupby('COUNTRY').max().sort_values('TOTAL_VACCINATIONS',
ascending=False).dropna(subset=['TOTAL_VACCINATIONS'])
vacc_amount = vacc_amount.iloc[:11]
vacc_amount = vacc_amount.sort_values('TOTAL_VACCINATIONS_PER100', ascending=False)
plt.bar(vacc_amount.index, vacc_amount.TOTAL_VACCINATIONS_PER100, color = 'r')
plt.xticks(rotation = 10)
plt.title("Which country is vaccinating its citizens the fastest?", color = "r",
fontweight="bold")
plt.ylabel('Amount of vaccinated people per hundred')
plt.subplot(212)
cols = ['COUNTRY', 'TOTAL_VACCINATIONS', 'ISO3', 'VACCINES_USED']
vacc_amount = df[cols].groupby('COUNTRY').max().sort_values('TOTAL_VACCINATIONS',
ascending=False).dropna(subset=['TOTAL_VACCINATIONS'])
vacc_amount = vacc_amount.iloc[:11]
plt.bar(vacc_amount.index, vacc_amount.TOTAL_VACCINATIONS, color = 'g')
plt.title('Which country has the highest number of vaccinated people?', color ="#025839",
fontweight="bold")
plt.ylabel('Amount of vaccinated citizens')
plt.xlabel('NOTE: Countries are not the same for both figures. Total number of countries in
each figure = 11')
plt.xticks(rotation = 10)
plt.tight_layout()
plt.show()
plt.savefig("output2.png")

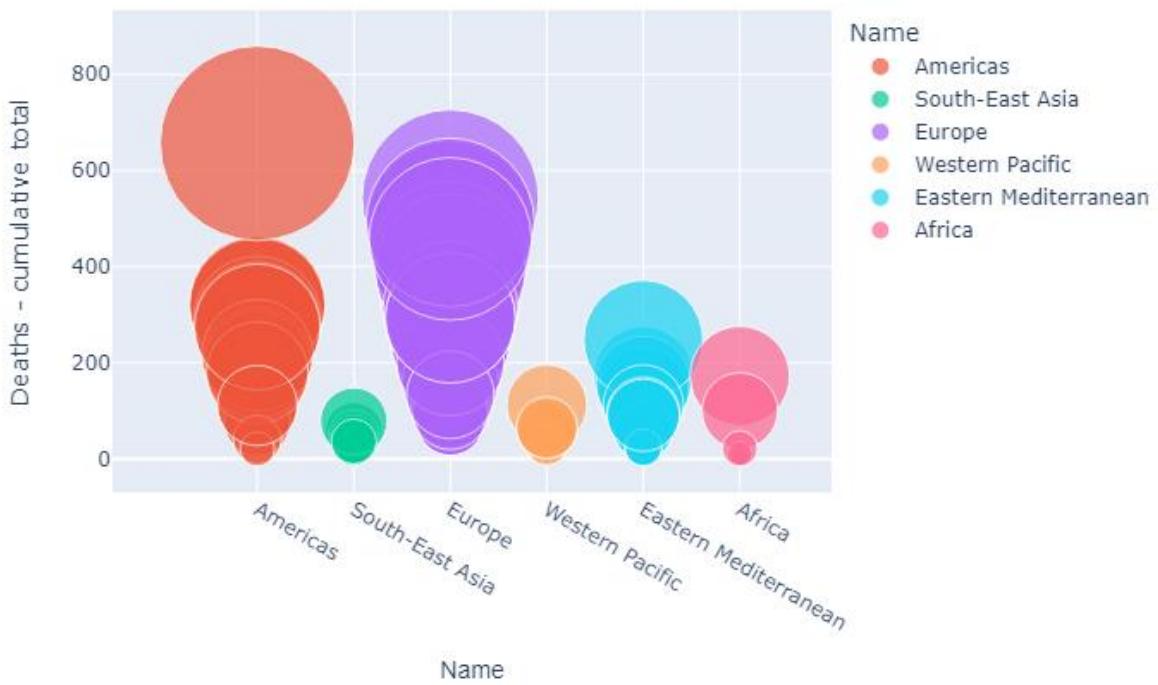
```

The screenshot shows the Visual Studio Code interface with the following details:

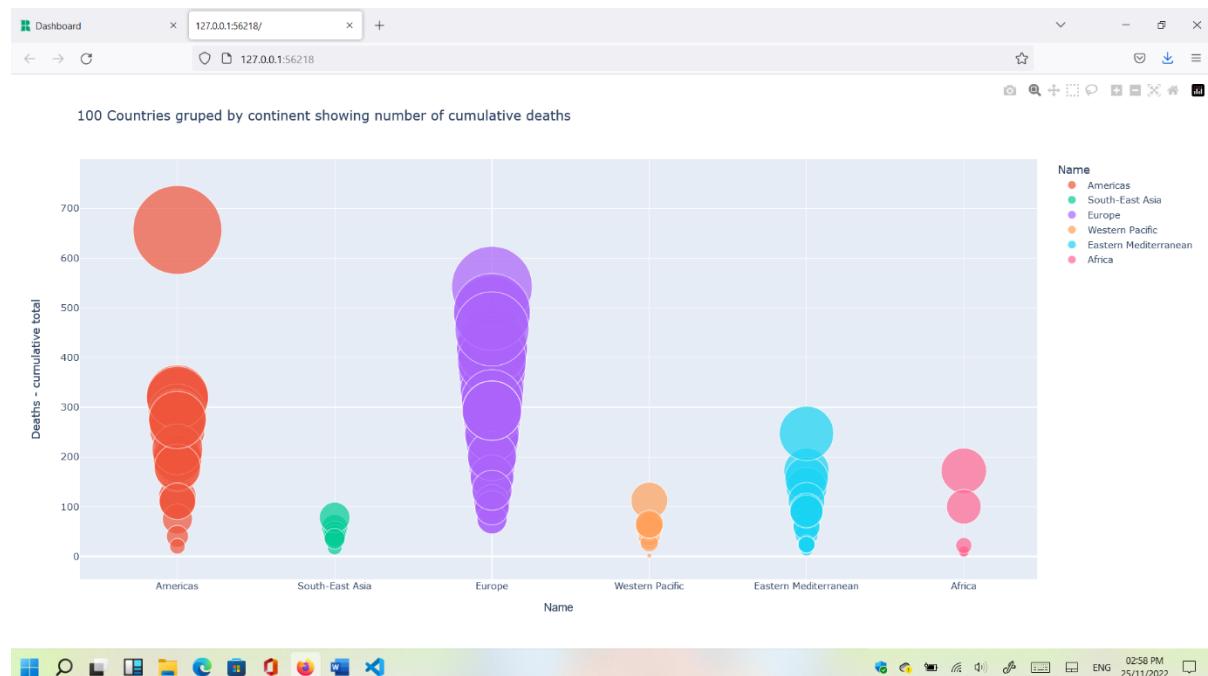
- File Explorer:** Shows files in the DV_COURSEWORK folder, including q1DV.py, q2DV.py, q3DV.py, q4DV.py, q5DV.py, output1.png, output2.png, output3.png, output4.png, TaskList2.docx, vaccination-data.csv, and WHO-COVID-19-global-t...
- Code Editor:** Displays the Python script q2DV.py.
- Status Bar:** Shows the current file is q2DV.py, the commit is master, and the status bar includes information like In 6 Col 1, Spaces: 4, UTT: 8, CRLF, Python 3.9.13 64-bit, and the date/time 03:05 PM 25/11/2022.

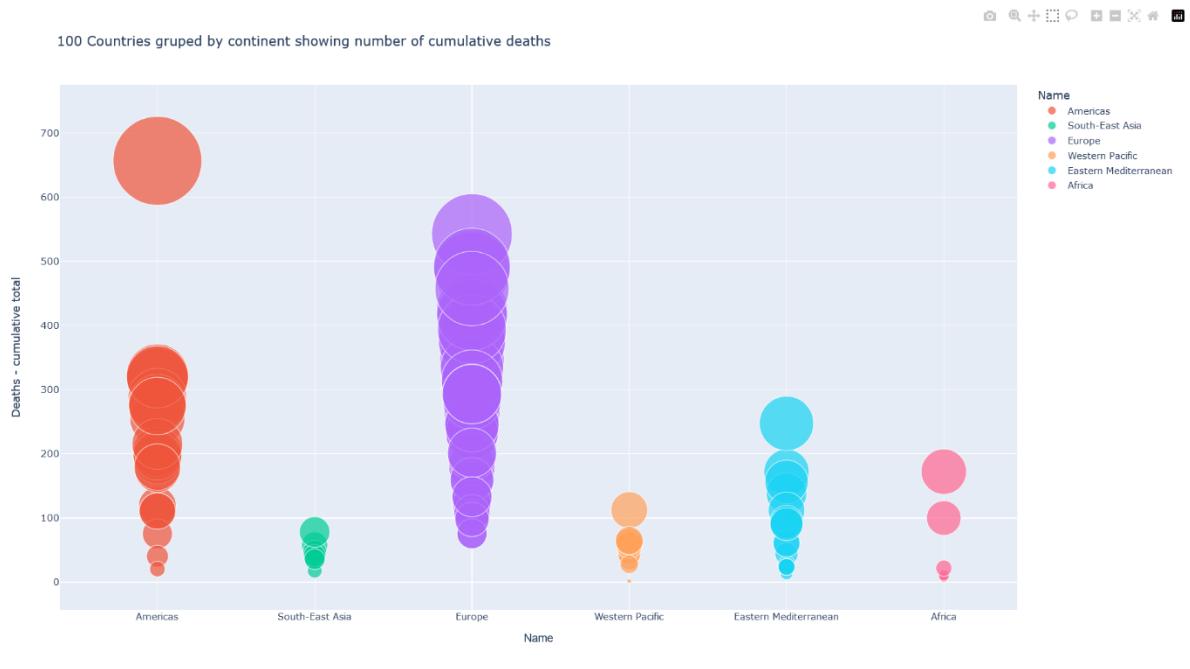
Third question: Cumulative deaths grouped by countries and showing only top 100 countries (total countries are 237):

100 Countries grouped by continent showing number of cumulative deaths



First image is the output that generates python code from VSC editor and saved using plotly buttons (top right). Second and third are screenshots of same picture.





```
#https://www.geeksforgeeks.org/covid-19-analysis-and-visualization-using-plotly-express/
import plotly.express as px
import pandas as pd
import plotly.io as pio

df = pd.read_csv("WHO-COVID-19-global-table-data.csv")
figure3 = px.scatter(df.head(100), x='Name', y= 'Deaths - cumulative total',
    hover_data=['WHO Region'],
    color='Name', size= 'Deaths - cumulative total', size_max=80, hover_name="Name",
    title=("100 Countries grouped by continent showing number of cumulative deaths"))
figure3.update_xaxes(title_font_family="Helvetica")
figure3.show()
# export as static image
pio.write_image(figure3, "output3.png")
```

The screenshot shows a Visual Studio Code interface with the following details:

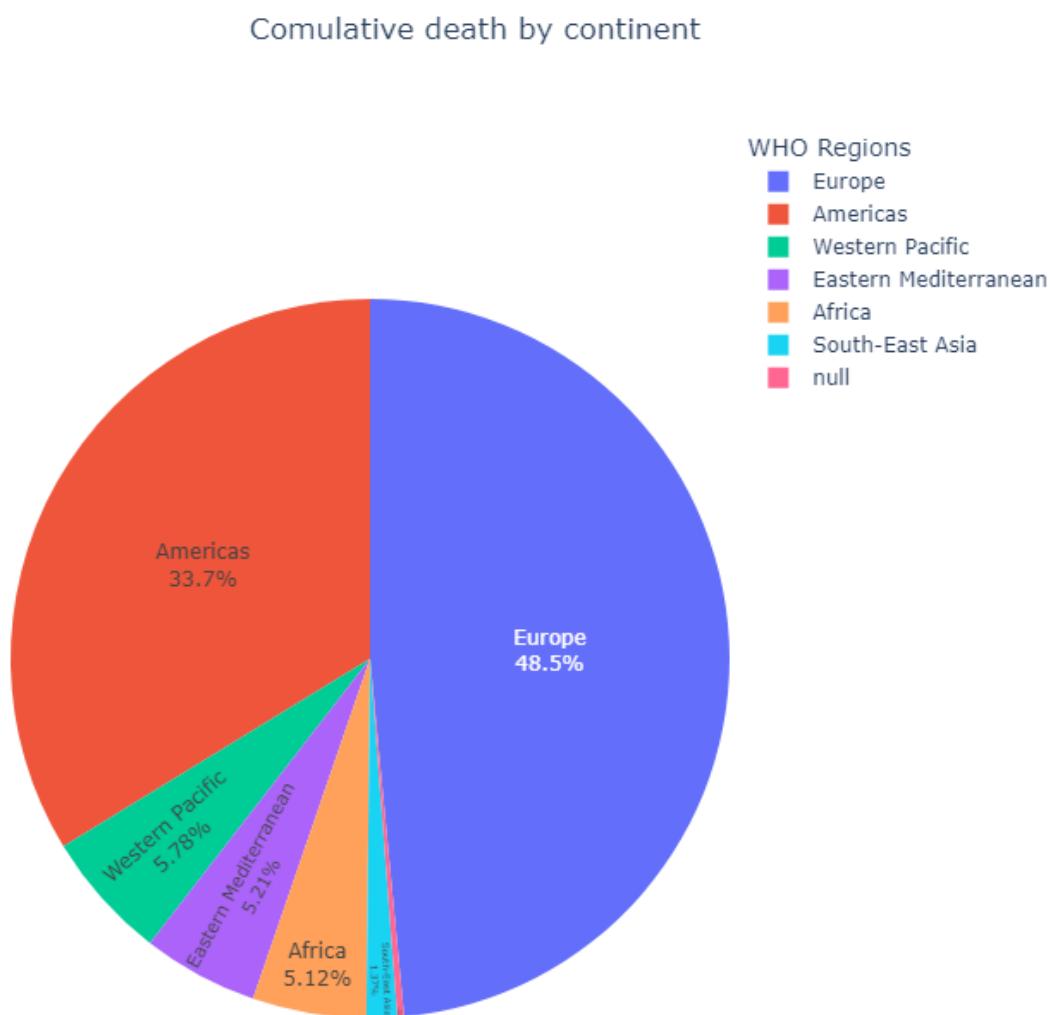
- File Menu:** File, Edit, Selection, View, Go, Run, Terminal, Help.
- Editor Area:** Five tabs are open: q1DV.py, q2DV.py, q3DV.py (active), q4DV.py, and q5DV.py. The q3DV.py tab contains the following Python code:

```
q3DV.py > ...
1 #https://www.geeksforgeeks.org/covid-19-analysis-and-visualization-using-plotly-express/
2 import plotly.express as px
3 import pandas as pd
4 import plotly.io as pio
5
6
7 df = pd.read_csv("WHO-COVID-19-global-table-data.csv")
8 figure1 = px.scatter(df.head(100), x="Name", y="Deaths - cumulative total",
9                      color="WHO_Region")
10
11 figure1.update_traces(size=100, hover_name="Name",
12 title="100 Countries grouped by continent showing number of cumulative deaths")
13 figure1.show()
14 # export as static image
15 pio.write_image(figure1, "output1.png")
16
```

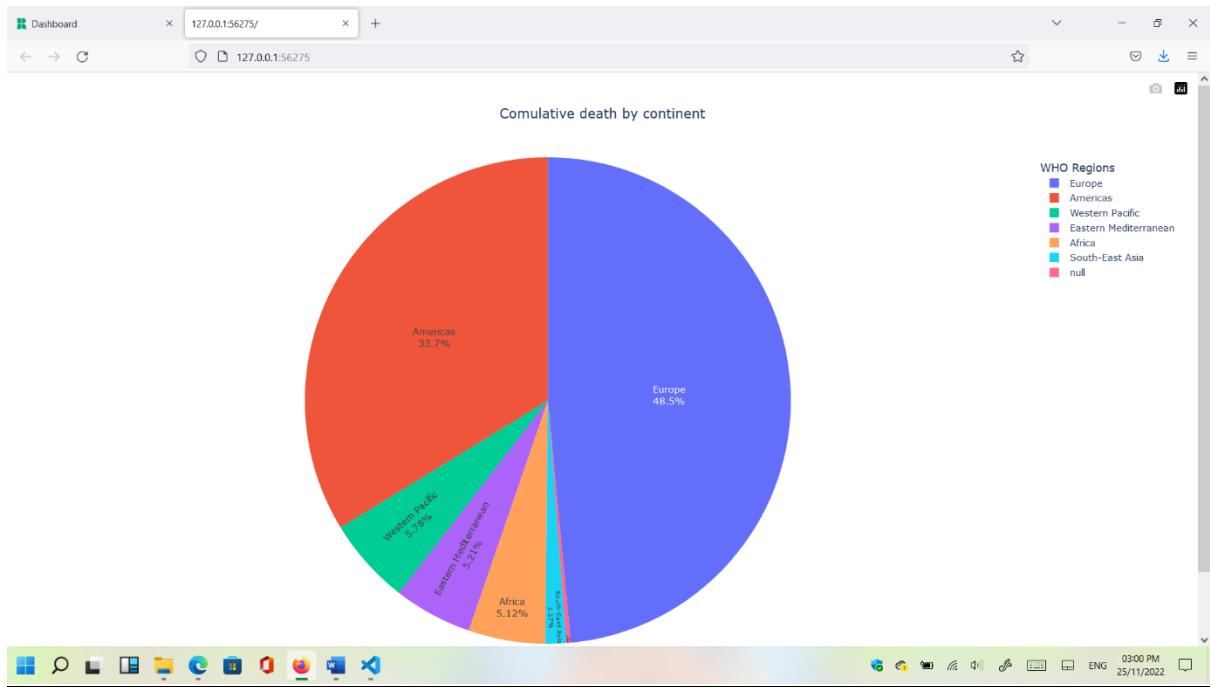
- Explorer Panel:** Shows files in the DV_COURSEWORK folder: output1.png, output2.png, output3.png, output4.png, q1DV.py, q2DV.py, q3DV.py (selected), q4DV.py, q5DV.py, Taskand2.docx, vaccination-data.csv, and WHO-COVID-19-global-t..
- Bottom Status Bar:** master*, Line 7, Col 55, Tab Size: 4, UFT-8, CRLF, Python 3.9.13 64-bit, ENG, 03:06 PM, 25/11/2022.

Fourth question: Total cumulative deaths grouped by continent using ALL countries.

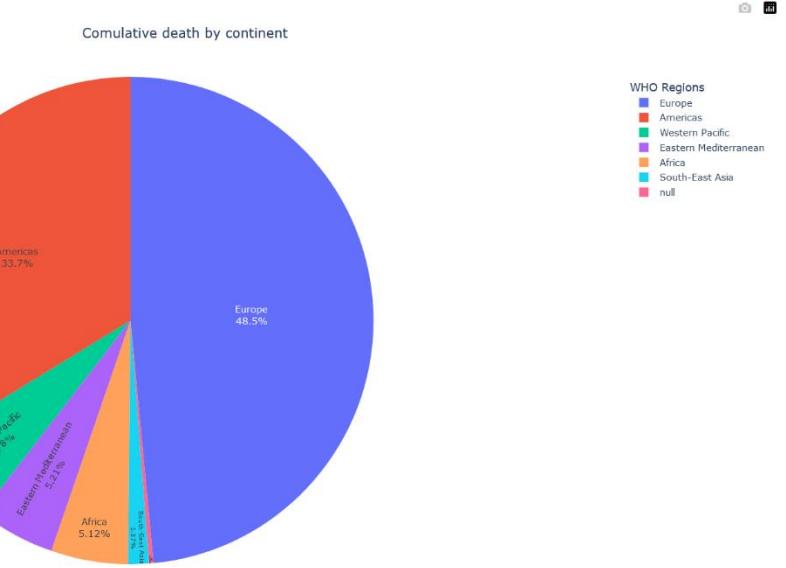
Pie chart.



First picture is the output saved using top right button implemented in plotly. This picture is from VSC editor, using python programming language.



Both images are screenshots from previous initial image. Browser is used (Firefox).
The third and last image is a full screen on the image on the browser (2nd).



```

# https://www.kaggle.com/code/siddheshmahajan/covid-19-data-visualization
import pandas as pd
import plotly.io as pio
import plotly.express as px

df = pd.read_csv("WHO-COVID-19-global-table-data.csv")
df_countries = df.groupby(['WHO Region']).max().reset_index()
df_countries = df_countries.drop_duplicates(subset = ['WHO Region'])
#the classic pie chart
fig = px.pie(df_countries, values = 'Deaths - cumulative total', names='Name', height=800,
title='Cumulative death by continent')
fig.update_traces(textposition='inside', textinfo='percent+label')
fig.update_layout(
    title_x = 0.5,
    autosize = True,
    geo=dict(
        showframe = False,
        showcoastlines = False,
    ))
fig.update_layout(legend=dict(
    title="WHO Regions"
)) # write a title for the legend on the right
fig.show()
pio.write_image(fig, "output4.png")

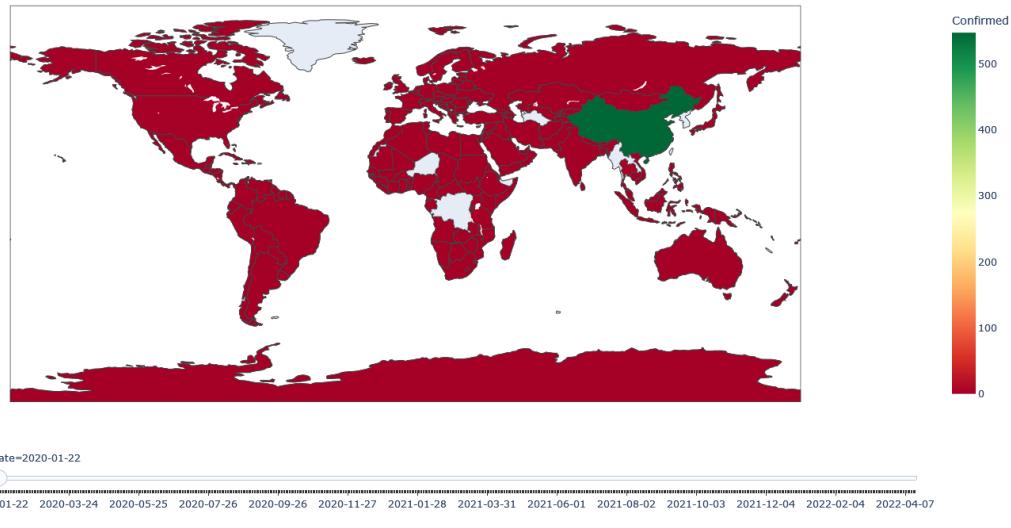
```

The screenshot shows the Visual Studio Code interface with the following details:

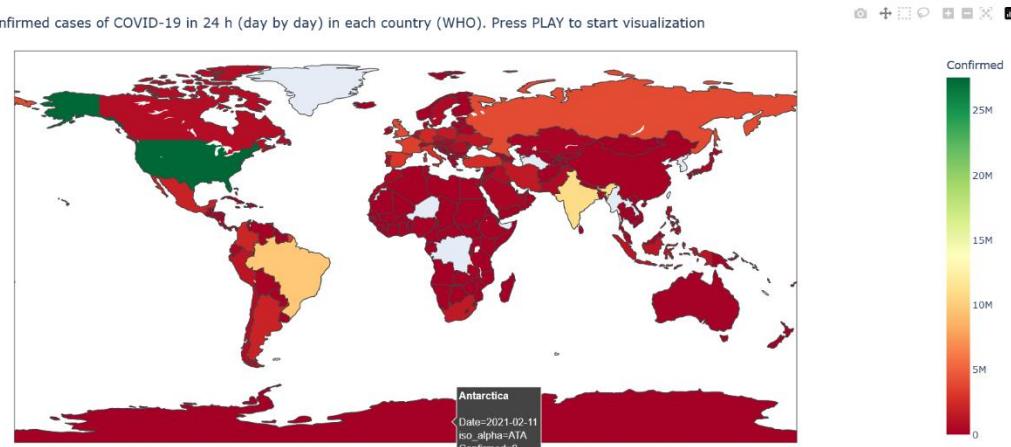
- File Explorer:** Shows files in the DV_COURSEWORK folder, including q1DV.py, q2DV.py, q3DV.py, q4DV.py (the active file), and q5DV.py.
- Code Editor:** Displays the Python script for generating a pie chart from a WHO COVID-19 global table CSV file.
- Terminal:** Shows the command-line output of the Python code execution, which prints country names and their corresponding death counts.
- Status Bar:** Provides information about the current file (q4DV.py), the Python interpreter (Python 3.9.13 64-bit), and the date/time (25/11/2022 03:07 PM).

Fifth question: Animated visualization of the spread of the coronavirus. Updates by day (24 hours).

Picture about confirmed cases of COVID-19 in 24 h (day by day) in each country (WHO). Press PLAY to start visualization

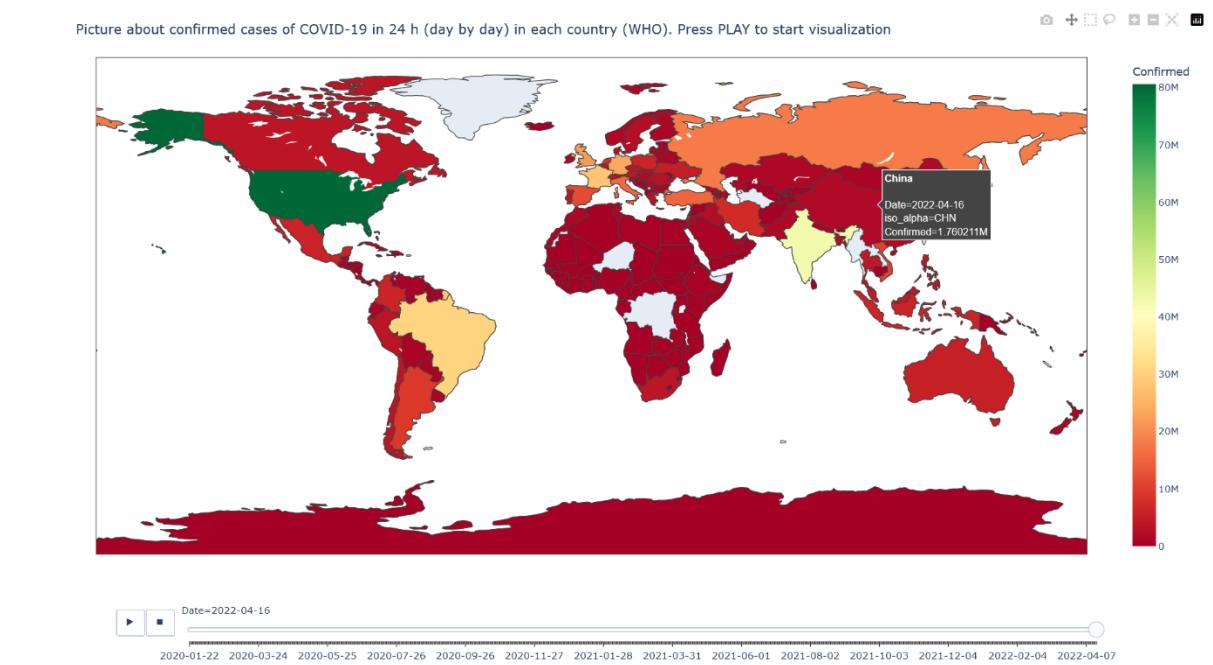
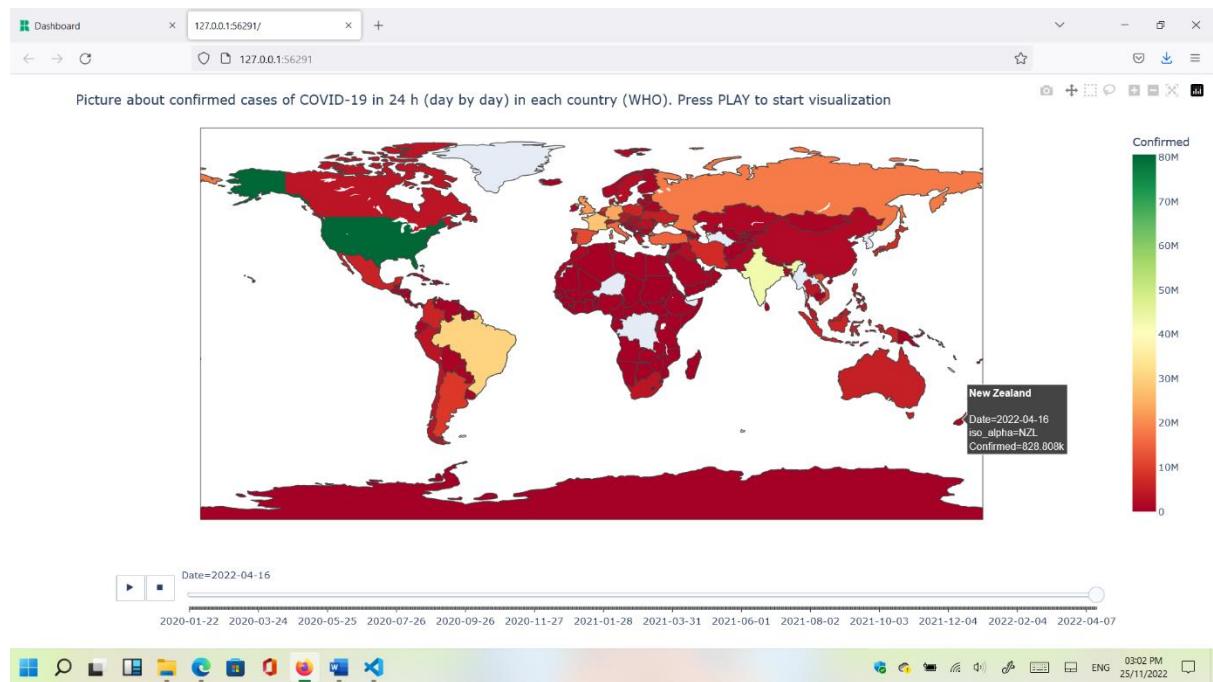


Picture about confirmed cases of COVID-19 in 24 h (day by day) in each country (WHO). Press PLAY to start visualization



Date=2021-02-11
2020-01-22 2020-03-24 2020-05-25 2020-07-26 2020-09-26 2020-11-27 2021-01-28 2021-03-31 2021-06-01 2021-08-02 2021-10-03 2021-12-04 2022-02-04 2022-04-07





```

# https://www.kaggle.com/code/lokeshrth4617/covid-19-world-wise-visualization
# https://opensource.com/article/20/4/python-map-covid-19
import pycountry
import plotly.express as px
import pandas as pd

# ----- Step 1 -----
URL_DATASET = r'https://raw.githubusercontent.com/datasets/covid-19/master/data/countries-aggregated.csv'
df1 = pd.read_csv(URL_DATASET)
# print(df1.head) # Uncomment to see what the dataframe is like
# ----- Step 2 -----
list_countries = df1['Country'].unique().tolist()
# print(list_countries) # Uncomment to see list of countries
d_country_code = {} # To hold the country names and their ISO
for country in list_countries:
    try:
        country_data = pycountry.countries.search_fuzzy(country)
        # country_data is a list of objects of class pycountry.db.Country
        # The first item ie at index 0 of list is best fit object of class Country have an
        alpha_3 attribute
        country_code = country_data[0].alpha_3
        d_country_code.update({country: country_code})
    except:
        print('could not add ISO 3 code for ->', country)
        d_country_code.update({country: ''}) # If could not find country, make ISO code ''
    #
# print(d_country_code) # Uncomment to check dictionary

# create a new column iso_alpha in the df and fill it with appropriate iso 3 code
for k, v in d_country_code.items():
    df1.loc[(df1.Country == k), 'iso_alpha'] = v
# print(df1.head) # Uncomment to confirm that ISO codes added
# ----- Step 3 -----
fig = px.choropleth(data_frame = df1,
                     locations= "iso_alpha",
                     color= "Confirmed", # value in column 'Confirmed' determines color
                     hover_name= "Country",
                     color_continuous_scale= 'RdYlGn', # color scale red, yellow green
                     animation_frame= "Date")
fig.update_layout(title="Picture about confirmed cases of COVID-19 in 24 h (day by day) in
each country (WHO). Press PLAY to start visualization")
fig.show()

```

The screenshot shows a Visual Studio Code window with the following details:

- File Explorer:** Shows files in the 'DV_COURSEWORK' folder, including q1DV.py, q2DV.py, q3DV.py, q4DV.py, q5DV.py, output1.png, output2.png, output3.png, output4.png, Task1and2.docx, vaccination-data.csv, and WHO-COVID-19-global-t...
- Code Editor:** Displays a Python script named q5DV.py. The code performs the following steps:
 - Imports necessary libraries: requests, pandas, pycountry, and plotly.express.
 - Downloads a dataset from Kaggle and GitHub.
 - Prints the first row of the dataset.
 - Prints unique country names.
 - Creates a dictionary to hold country names and ISO codes.
 - Tries to add ISO 3 codes for each country using pycountry's search_fuzzy method. If successful, it updates the ISO code in the dataset; if not found, it creates a new column 'iso_alpha' and sets its value to the country name.
 - Creates a choropleth map using plotly.express with 'Confirmed' cases as the value and 'iso_alpha' as the location.
 - Updates the map's layout with a title and animation frames.
 - Shows the map.
- Terminal:** Shows command-line output indicating the script ran successfully.
- Bottom Status Bar:** Shows the file path (PS :C:\Users\Admin\OneDrive\Desktop\DV_Coursework), line count (Ln 39, Col 112), spaces (Spaces: 4), encoding (UTF-8), CRLF, Python version (3.9.13 64-bit), and system information (03:07 PM, 25/11/2022).

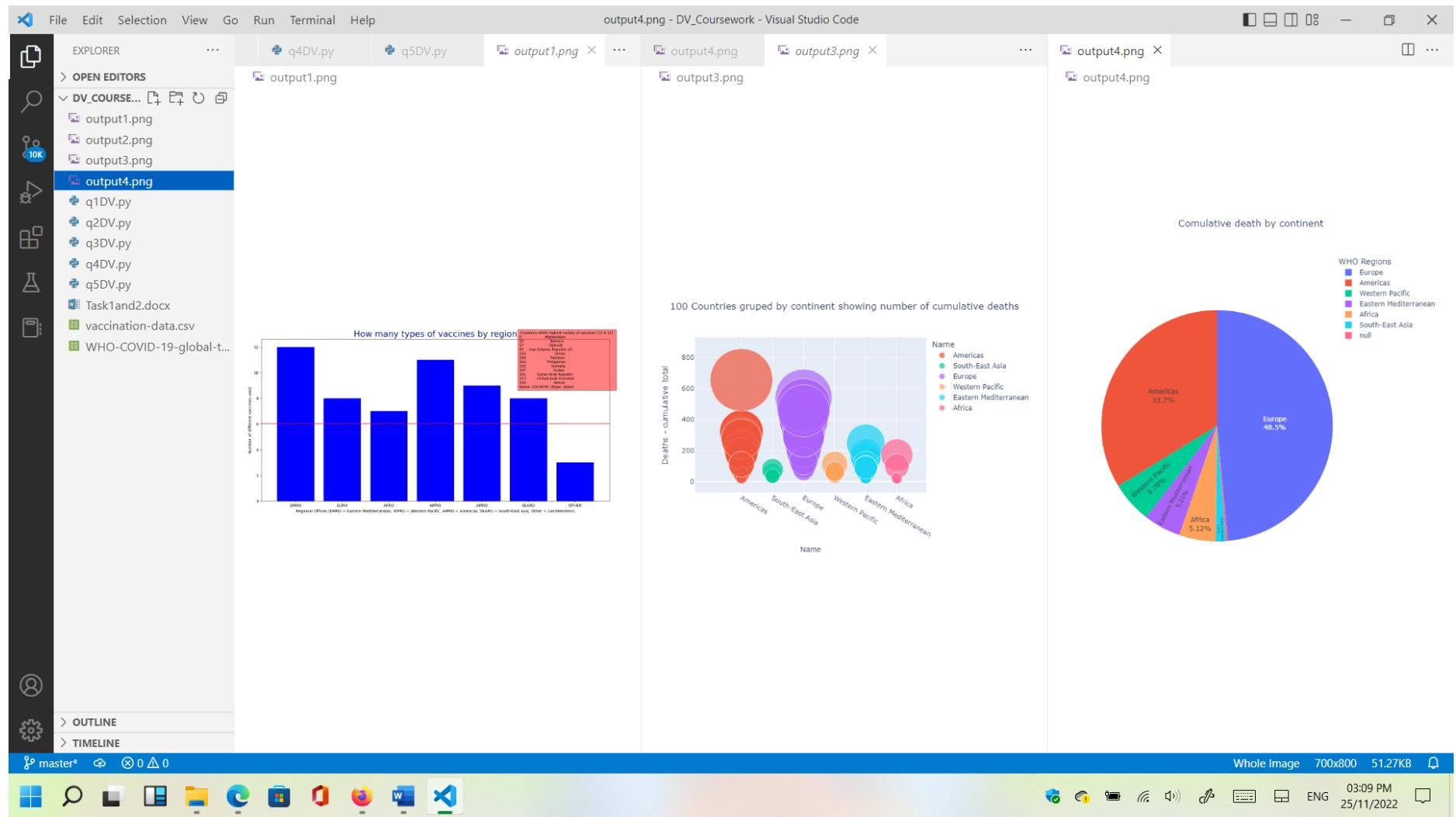
This screenshot is nearly identical to the one above, showing the same Visual Studio Code interface and Python script (q5DV.py). The main difference is in the terminal output, which now includes the following message at the end:

```
# print(df1.head) # Uncomment to confirm that ISO codes added
```

This indicates that the script has successfully added ISO 3 codes to the dataset and printed the updated head of the DataFrame.

Code on my GitHub:

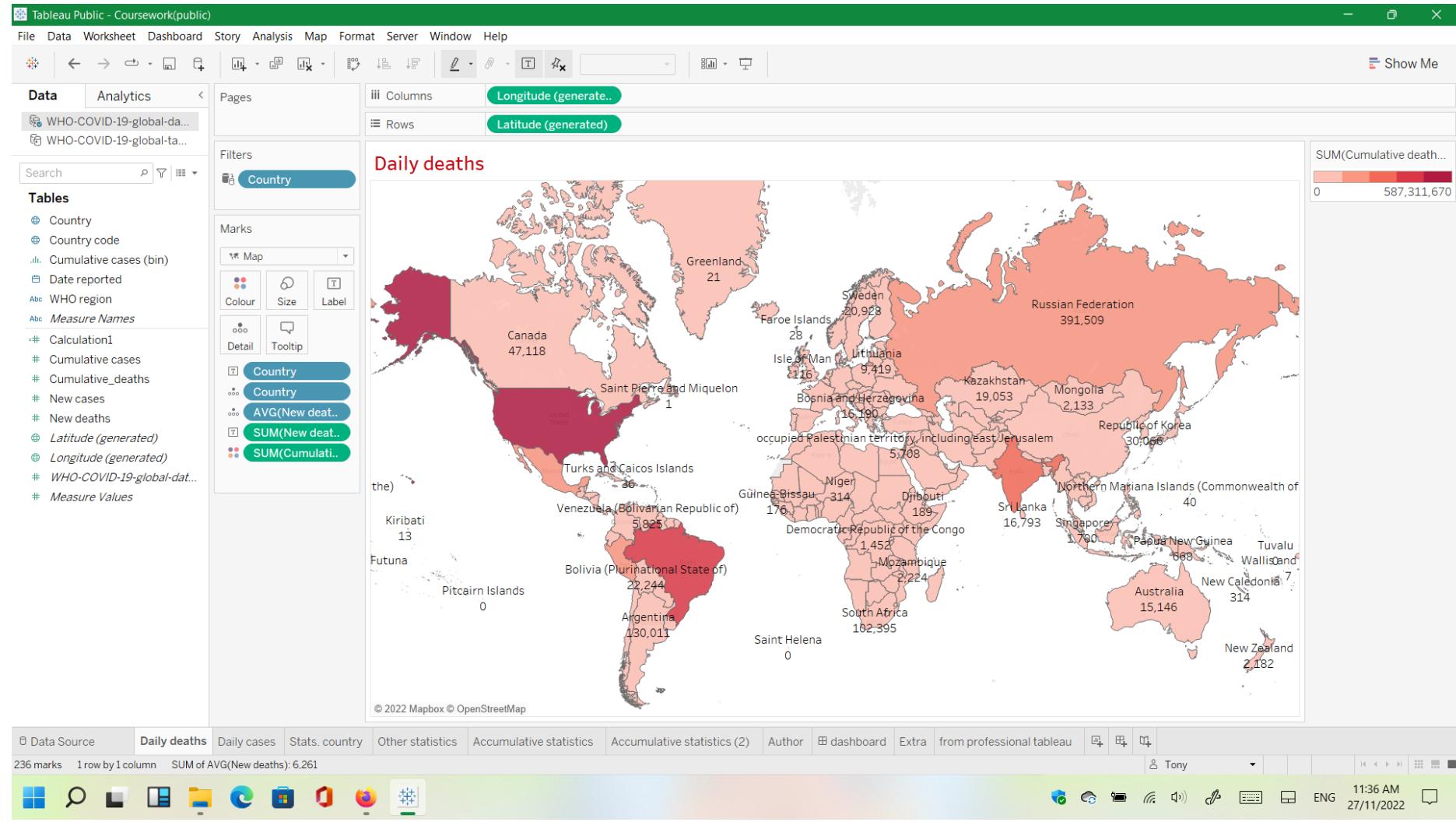
https://github.com/ittonycash/data_visualisation/tree/main/coursework

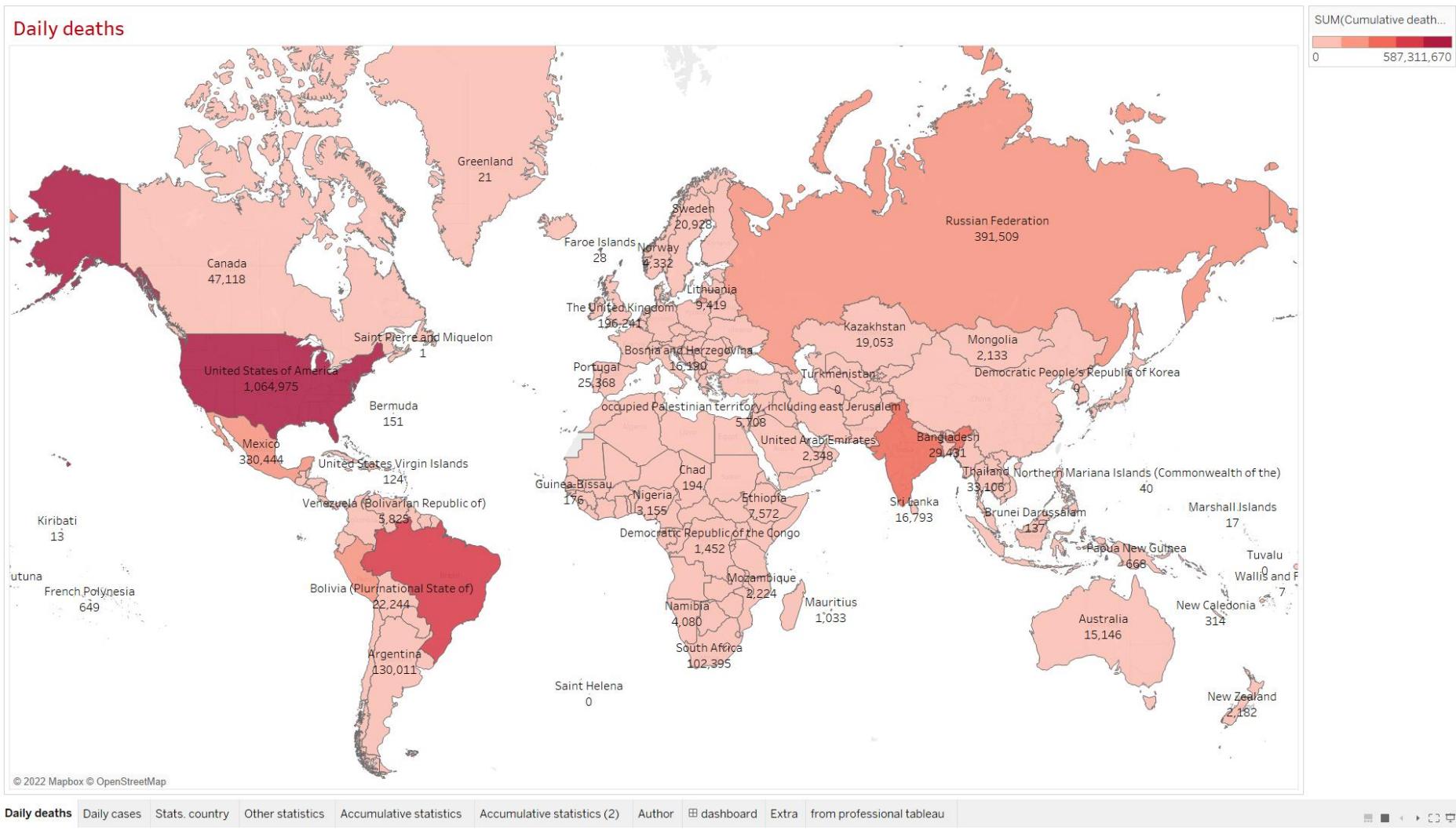


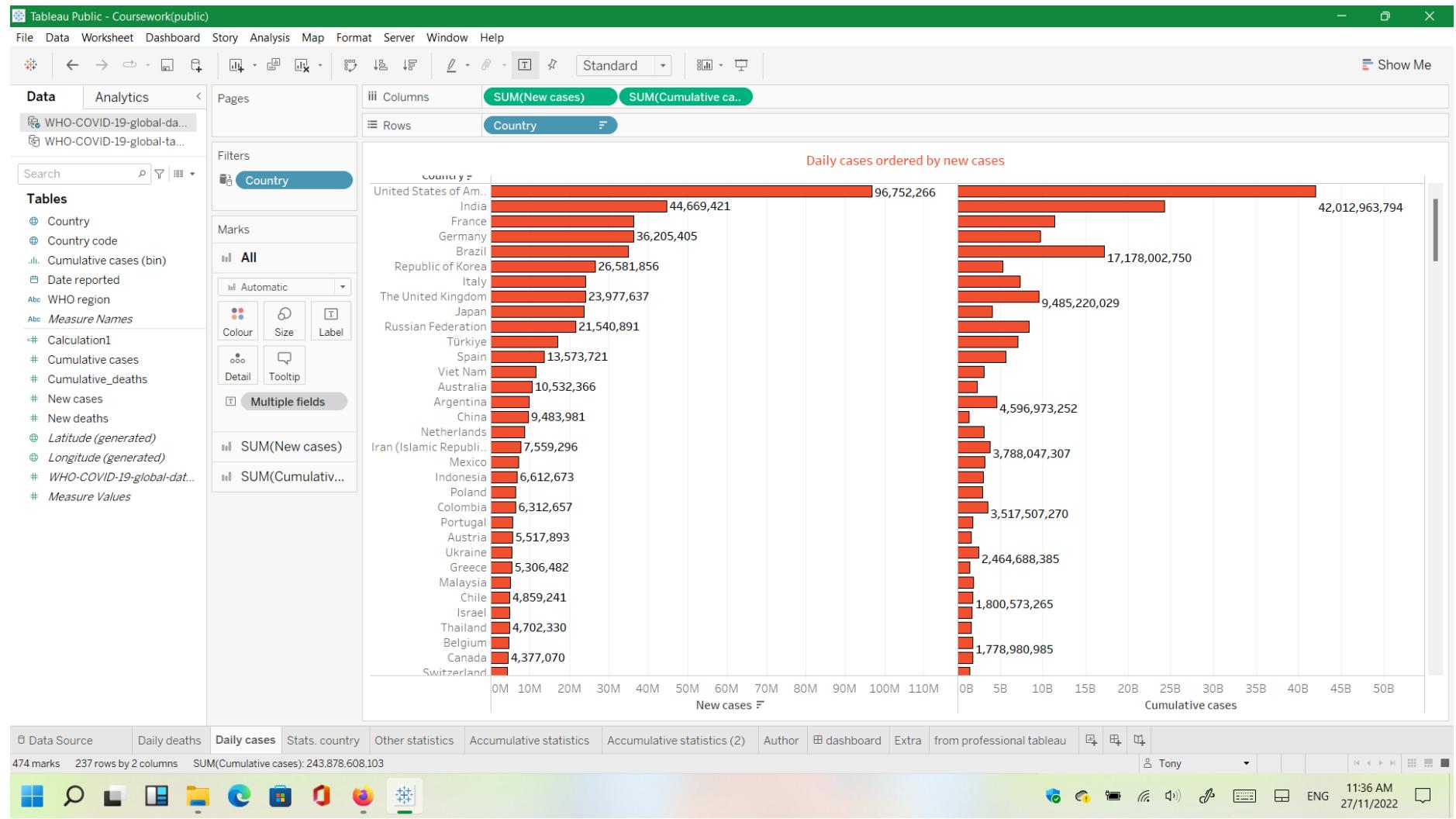
My audience:



Task 2 screenshots. Individual sheets and final dashboard:







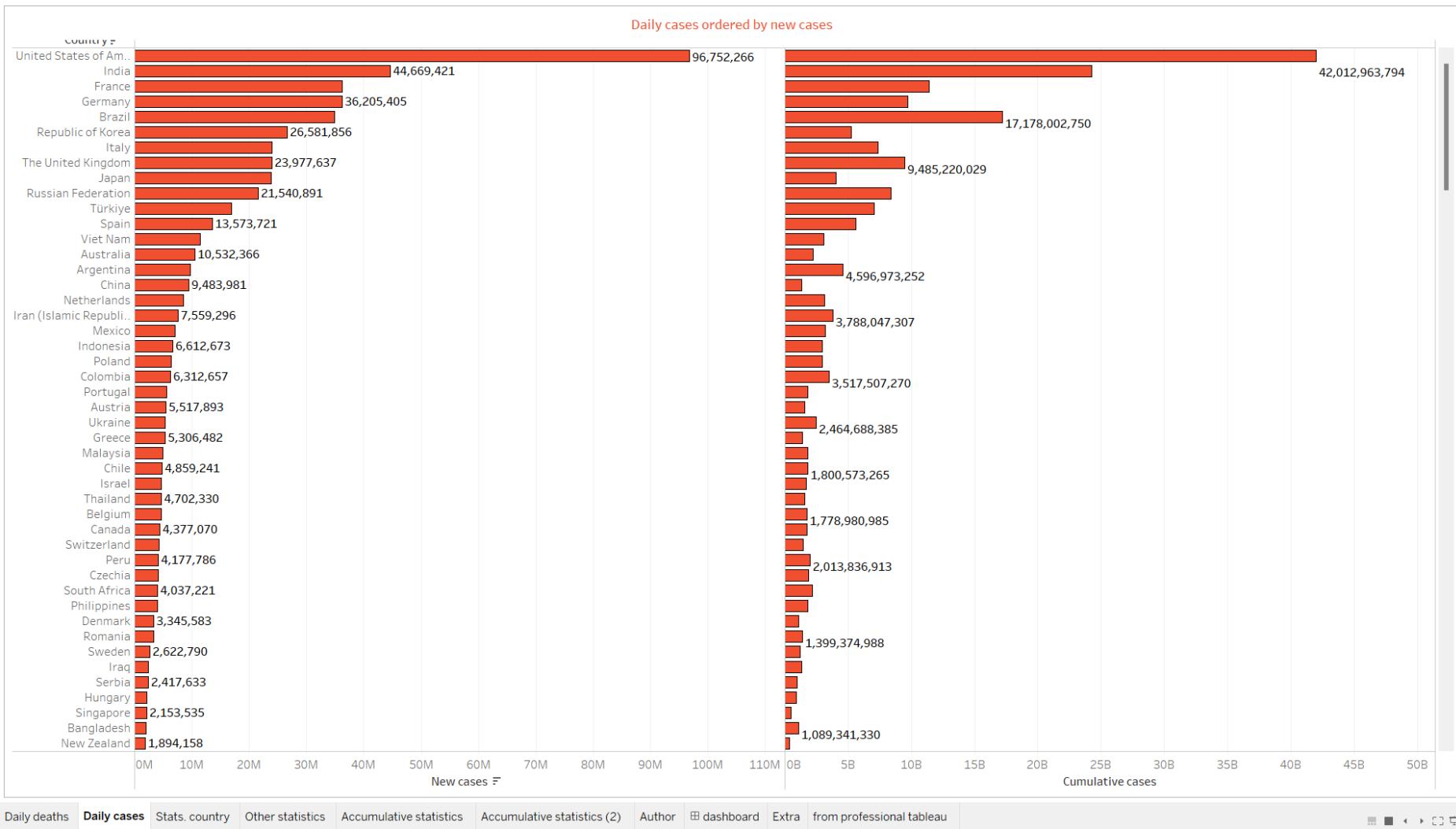


Tableau Public - Coursework(public)

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Show Me

Data Analytics

WHO-COVID-19-global-data WHO-COVID-19-global-ta...

Pages Measure Names

Rows Country YEAR(Date report..)

Filters

Measure Names

Country

Tables

- Country
- Country code
- Cumulative cases (bin)
- Date reported
- WHO region
- Measure Names
- # Calculation1
- # Cumulative cases
- # Cumulative_deaths
- # New cases
- # New deaths
- # Latitude (generated)
- # Longitude (generated)
- # WHO-COVID-19-global-dat...
- # Measure Values

Marks

Automatic

Colour Size Text

Detail Tooltip

YEAR(Date r..) Measure Values

Measure Values

- AGG(Calculation1)
- CNT(WHO-COVID-19..)
- SUM(Cumulative cas..)
- AVG(Cumulative_de..)
- SUM(New cases)
- SUM(New deaths)

Country Year of D.. Calculation1 Count of WHO-CO.. Cumulative cases Cumulative_deat.. New cases New deaths

Country	Year of D..	Calculation1	Count of WHO-CO..	Cumulative cases	Cumulative_deat..	New cases	New deaths
Afghanistan	2020	814	364	8,501,745	814	52,330	2,189
	2021	4,874	365	39,510,471	4,874	105,729	5,167
	2022	7,681	325	59,479,666	7,681	47,008	475
Albania	2020	240	364	3,669,242	240	57,727	1,174
	2021	2,382	365	50,956,601	2,382	151,789	2,041
	2022	3,499	325	94,417,392	3,499	122,590	379
Algeria	2020	1,027	364	10,482,230	1,027	99,311	2,751
	2021	4,275	365	57,054,531	4,275	118,726	3,520
	2022	6,818	325	85,633,669	6,818	52,915	610
American Samoa	2020	0	364	0	0	0	0
	2021	0	365	573	0	11	0
	2022	22	325	1,723,827	22	8,251	34
Andorra	2020	42	364	675,517	42	7,983	84
	2021	122	365	5,070,326	122	15,139	56
	2022	152	325	13,618,367	152	23,702	16
Angola	2020	100	364	1,326,113	100	17,433	405
	2021	1,006	365	14,729,111	1,006	59,354	1,352
	2022	1,901	325	32,559,843	1,901	26,344	160
Anguilla	2020	0	364	1,049	0	13	0
	2021	1	365	129,088	1	1,661	5
	2022	10	325	1,027,273	10	2,192	7
Antigua and Barbuda	2020	2	364	22,547	2	158	5
	2021	51	365	684,001	51	4,137	114
	2022	138	325	2,609,380	138	4,811	27
Argentina	2020	14,195	364	156,512,045	14,195	1,674,319	48,271
	2021	92,505	365	1,471,656,467	92,505	4,106,203	69,905
	2022	128,299	325	2,968,804,740	128,299	3,941,196	11,835
Armenia	2020	683	364	14,552,901	683	159,409	2,823
	2021	4,762	365	86,241,547	4,762	185,521	5,149
	2022	8,546	325	136,518,652	8,546	100,701	737
Aruba	2020	13	364	558,482	13	5,373	49
	2021	117	365	4,325,845	117	13,648	132
	2022	216	325	12,300,848	216	24,473	54

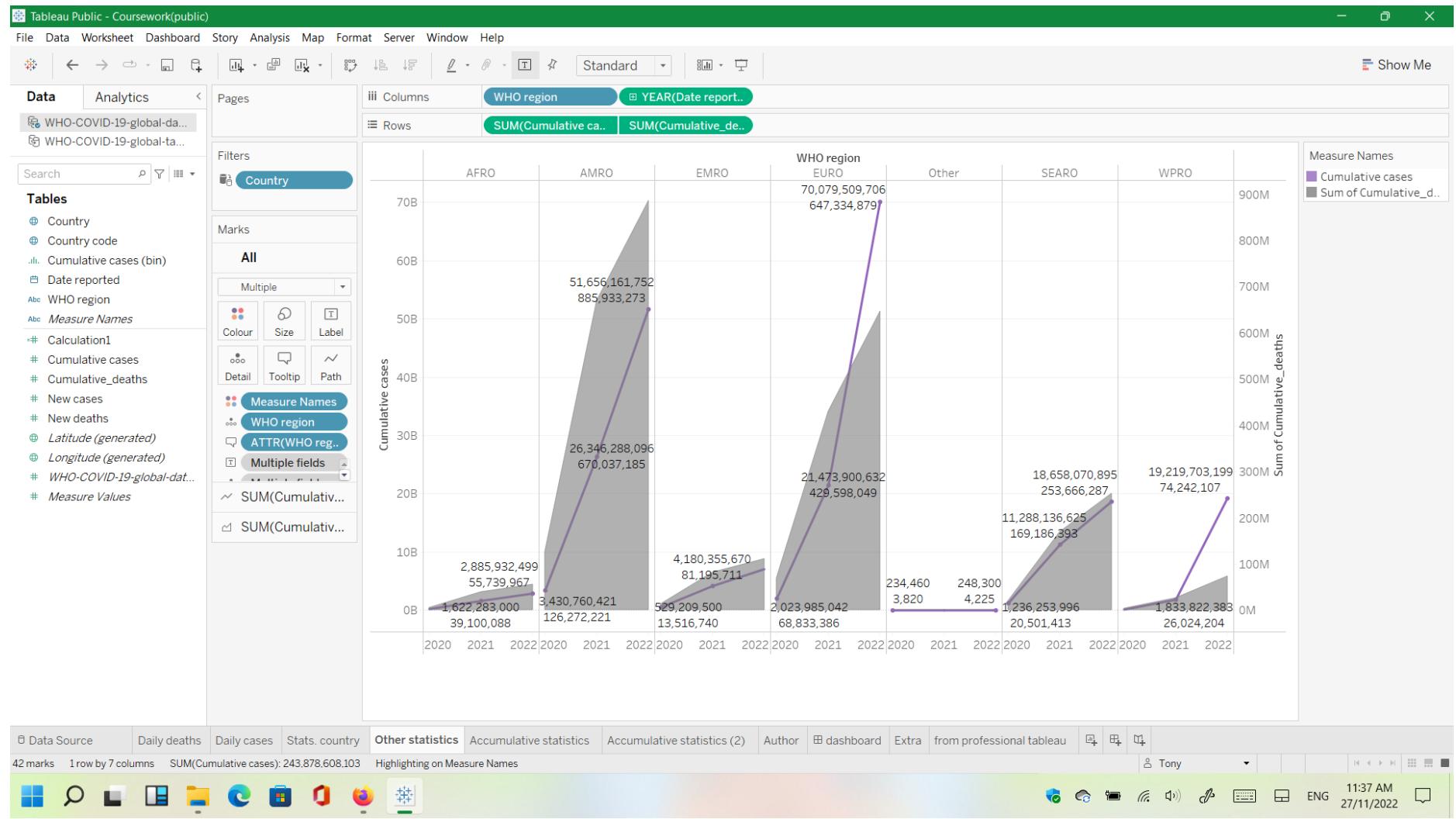
Data Source Daily deaths Daily cases Stats. country Other statistics Accumulative statistics Accumulative statistics (2) Author dashboard Extra from professional tableau

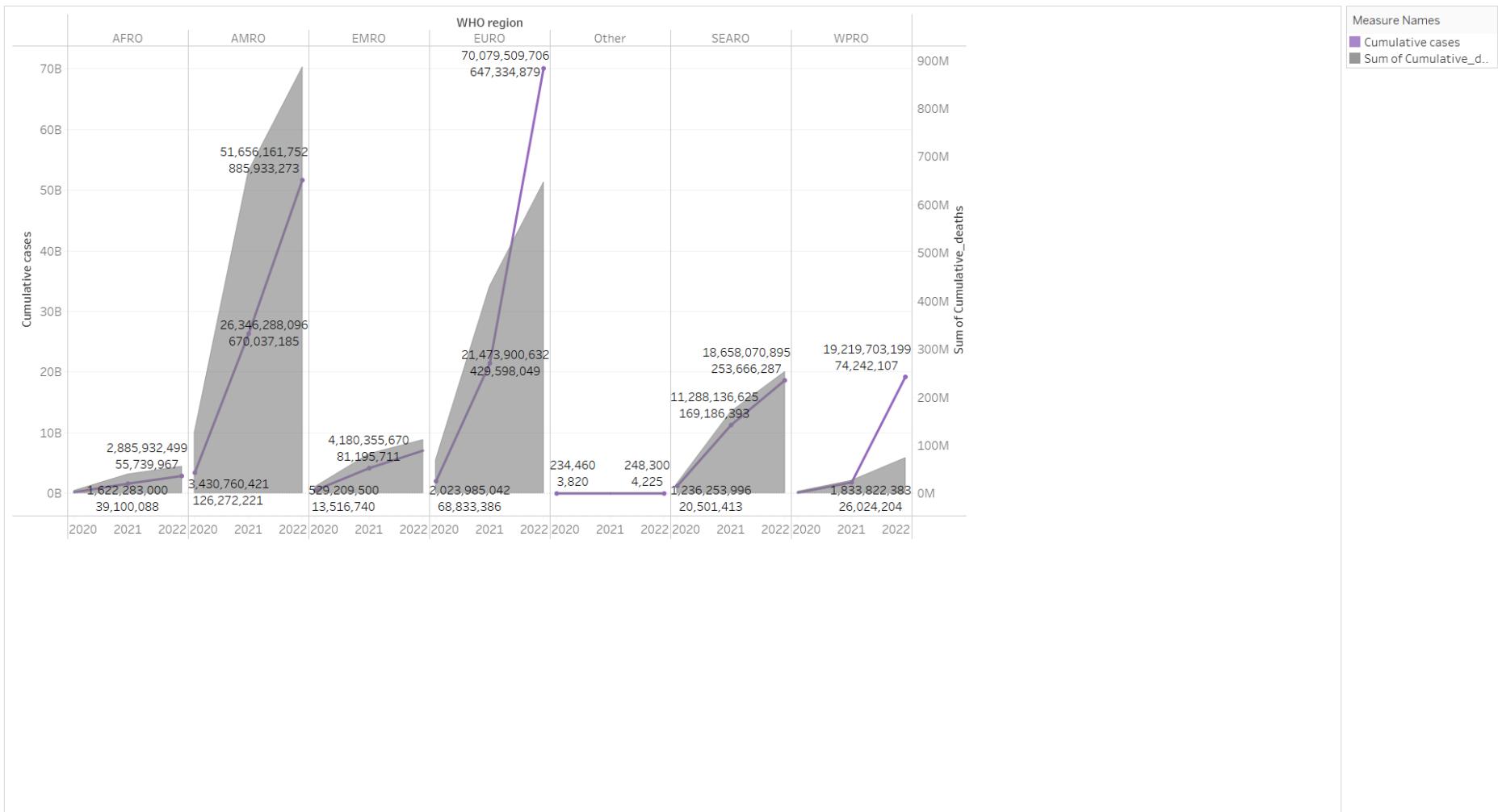
4266 marks 711 rows by 6 columns SUM of Measure Values: 244,541,515,568

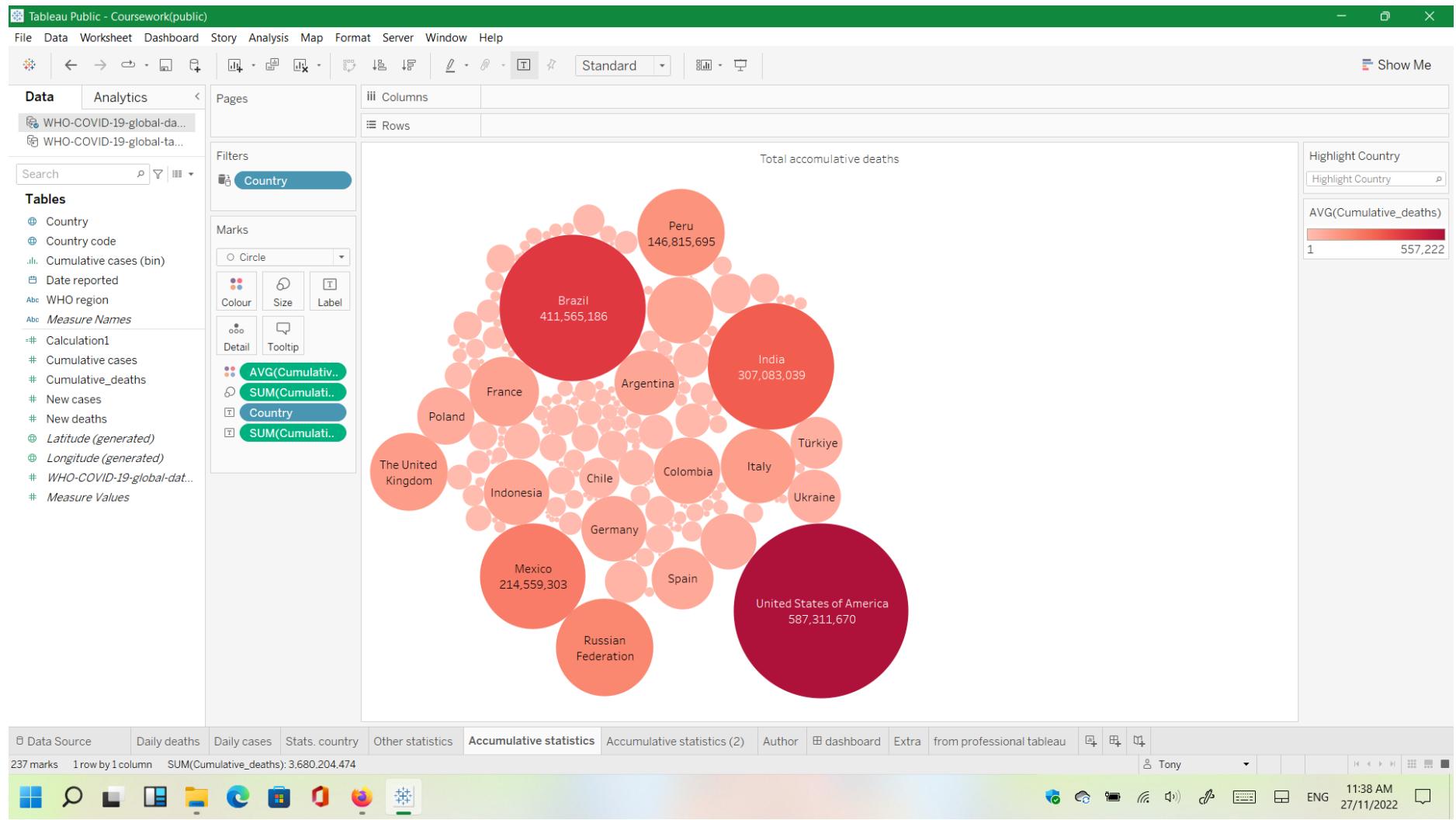
Tony 11:37 AM 27/11/2022 ENG

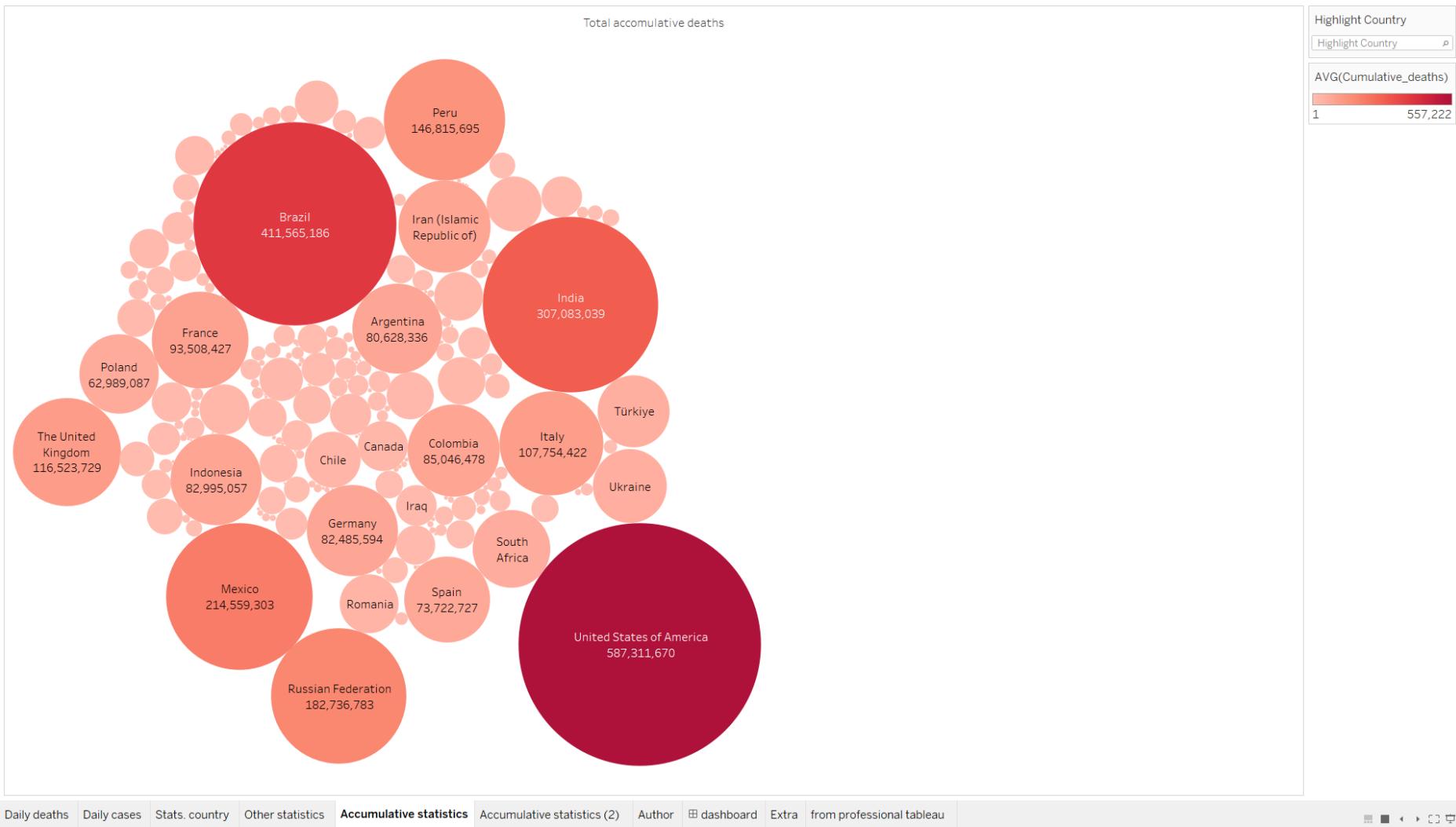
Windows Taskbar icons

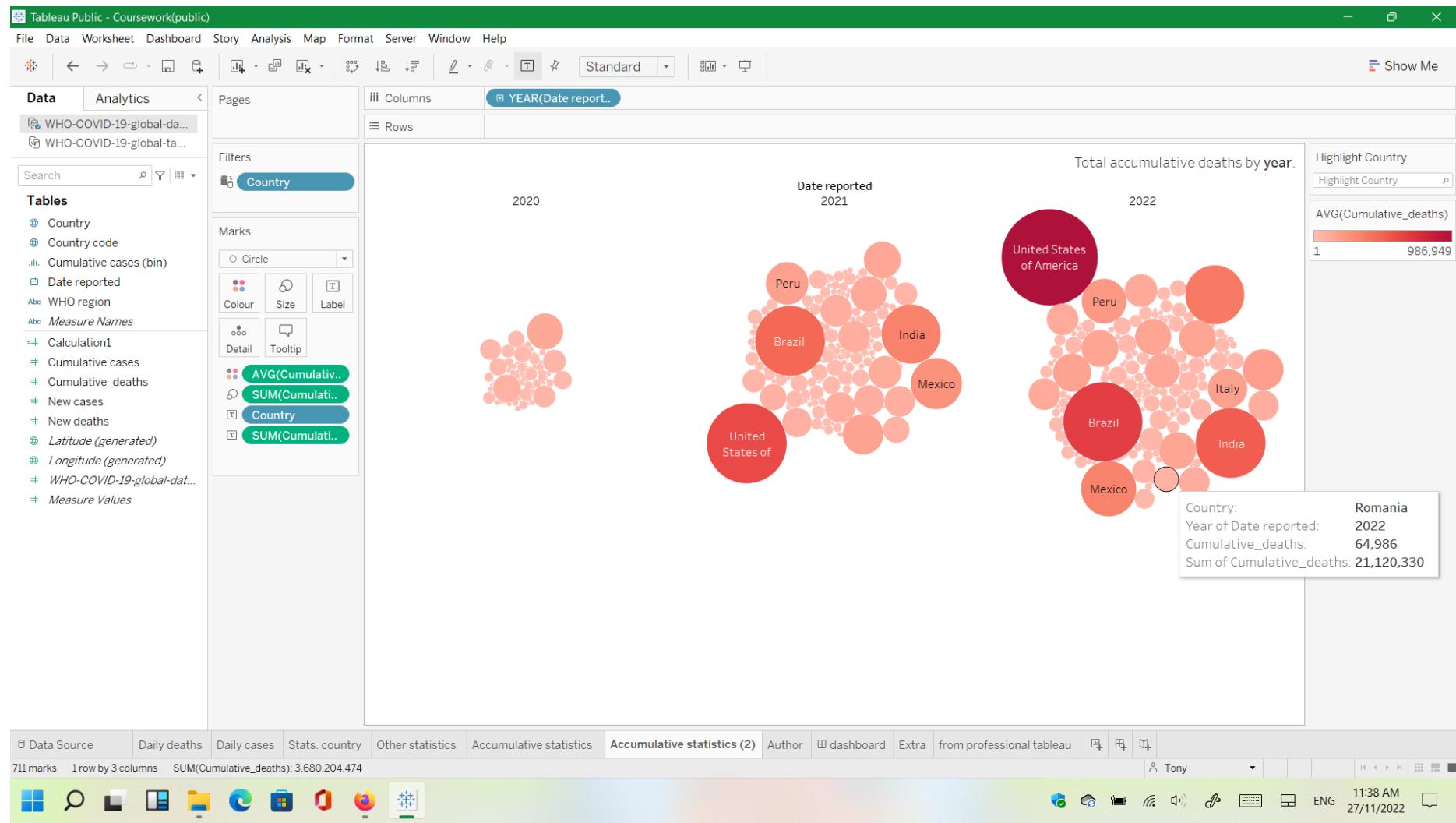
							Stats. country. Year of ...
Country	Year of D..	Calculation1	Count of WHO-CO..	Cumulative cases	Cumulative_deat..	New cases	New deaths
Afghanistan	2020	814	364	8,501,745	814	52,330	2,189
	2021	4,874	365	39,510,471	4,874	105,729	5,167
	2022	7,681	325	59,479,666	7,681	47,008	475
Albania	2020	240	364	3,669,242	240	57,727	1,174
	2021	2,382	365	50,956,601	2,382	151,789	2,041
	2022	3,499	325	94,417,392	3,499	122,590	379
Algeria	2020	1,027	364	10,482,230	1,027	99,311	2,751
	2021	4,275	365	57,054,531	4,275	118,726	3,520
	2022	6,818	325	85,633,669	6,818	52,915	610
American Samoa	2020	0	364	0	0	0	0
	2021	0	365	573	0	11	0
	2022	22	325	1,723,827	22	8,251	34
Andorra	2020	42	364	675,517	42	7,983	84
	2021	122	365	5,070,326	122	15,139	56
	2022	152	325	13,618,367	152	23,702	16
Angola	2020	100	364	1,326,113	100	17,433	405
	2021	1,006	365	14,729,111	1,006	59,354	1,352
	2022	1,901	325	32,559,843	1,901	26,344	160
Anguilla	2020	0	364	1,049	0	13	0
	2021	1	365	129,088	1	1,661	5
	2022	10	325	1,027,273	10	2,192	7
Antigua and Barbuda	2020	2	364	22,547	2	158	5
	2021	51	365	684,001	51	4,137	114
	2022	138	325	2,609,380	138	4,811	27
Argentina	2020	14,195	364	156,512,045	14,195	1,674,319	48,271
	2021	92,505	365	1,471,656,467	92,505	4,106,203	69,905
	2022	128,299	325	2,968,804,740	128,299	3,941,196	11,835
Armenia	2020	683	364	14,552,901	683	159,409	2,823
	2021	4,762	365	86,241,547	4,762	185,521	5,149
	2022	8,546	325	136,518,652	8,546	100,701	737
Aruba	2020	13	364	558,482	13	5,373	49
	2021	117	365	4,325,845	117	13,648	132
	2022	216	325	12,300,848	216	24,473	54
Australia	2020	361	364	5,043,709	361	28,381	909
	2021	1,165	365	27,675,084	1,165	334,177	1,317
	2022	9,368	325	2,218,877,282	9,368	10,169,808	12,920
Austria	2020	1,138	364	20,854,203	1,138	352,707	7,384
	2021	12,503	365	247,012,631	12,503	911,990	9,304
	2022	19,465	325	1,315,431,366	19,465	4,253,196	4,456
Azerbaijan	2020	430	364	12,512,631	430	216,584	2,575
	2021	5,163	365	137,039,996	5,163	399,768	5,771
	2022	9,590	325	253,677,738	9,590	207,602	1,626
Bahamas	2020	48	364	762,513	48	7,857	170
	2021	353	365	5,352,732	353	16,619	547
	2022	800	325	11,321,833	800	12,973	116
Bahrain	2020	129	364	12,840,326	129	92,425	352











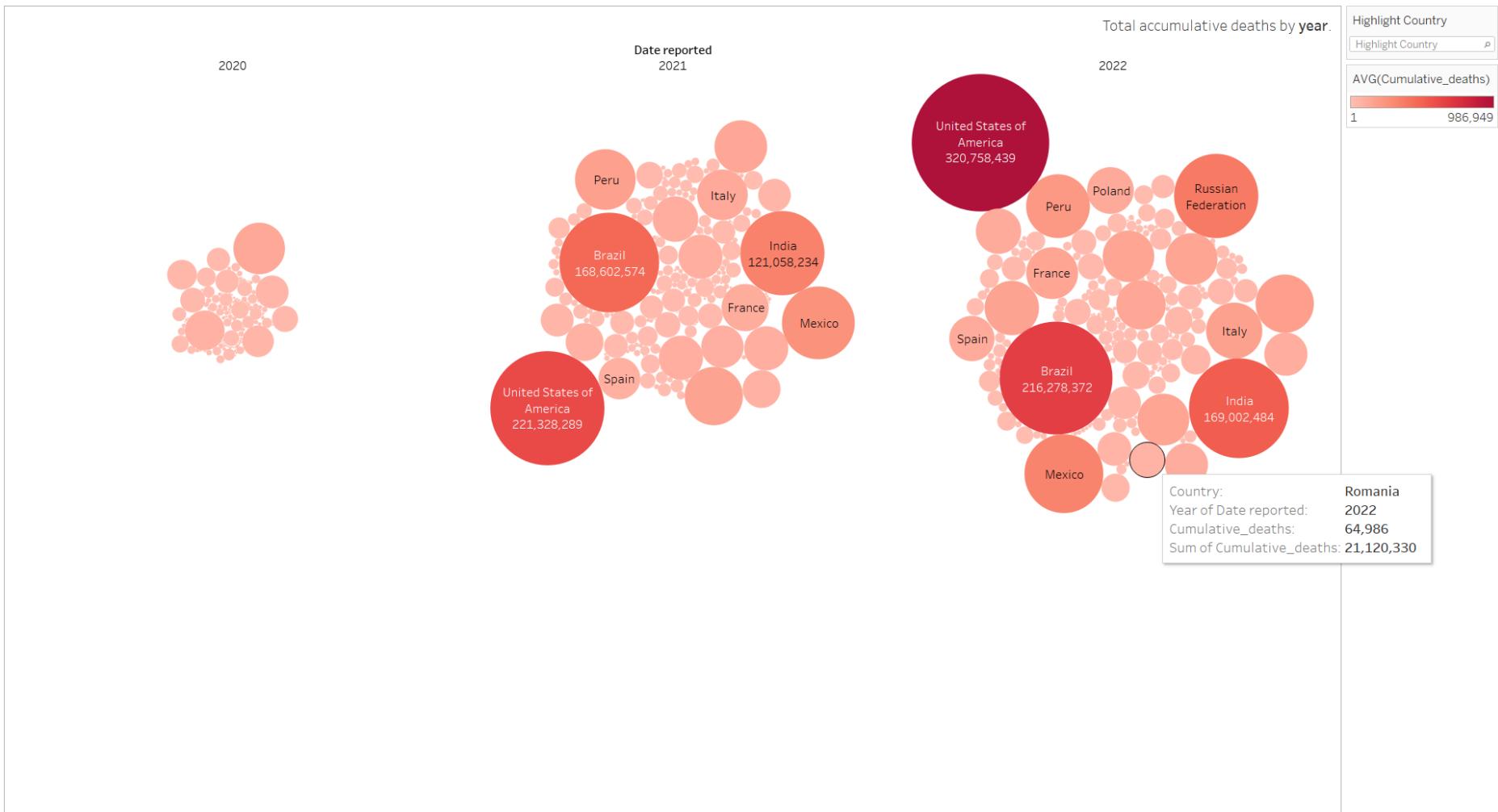


Tableau Public - Coursework(public)

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Show Me

Data Analytics

WHO-COVID-19-global-data...

WHO-COVID-19-global-ta...

Search

Tables

- Country
- Country code
- Cumulative cases (bin)
- Date reported
- WHO region
- Measure Names
- Calculation1
- Cumulative cases
- Cumulative_deaths
- New cases
- New deaths
- Latitude (generated)
- Longitude (generated)
- WHO-COVID-19-global-dat...
- Measure Values

Pages

Columns

Rows

Filters

Tony (YIT19488399), Data Visualization, autumn module.
UoR, year 3, BSc Computer Science

Marks

Automatic

Colour Size Text

Detail Tooltip

Drop field here

Drop field here

Drop field here

Data Source Daily deaths Daily cases Stats. country Other statistics Accumulative statistics Accumulative statistics (2) Author dashboard Extra from professional tableau

Tony

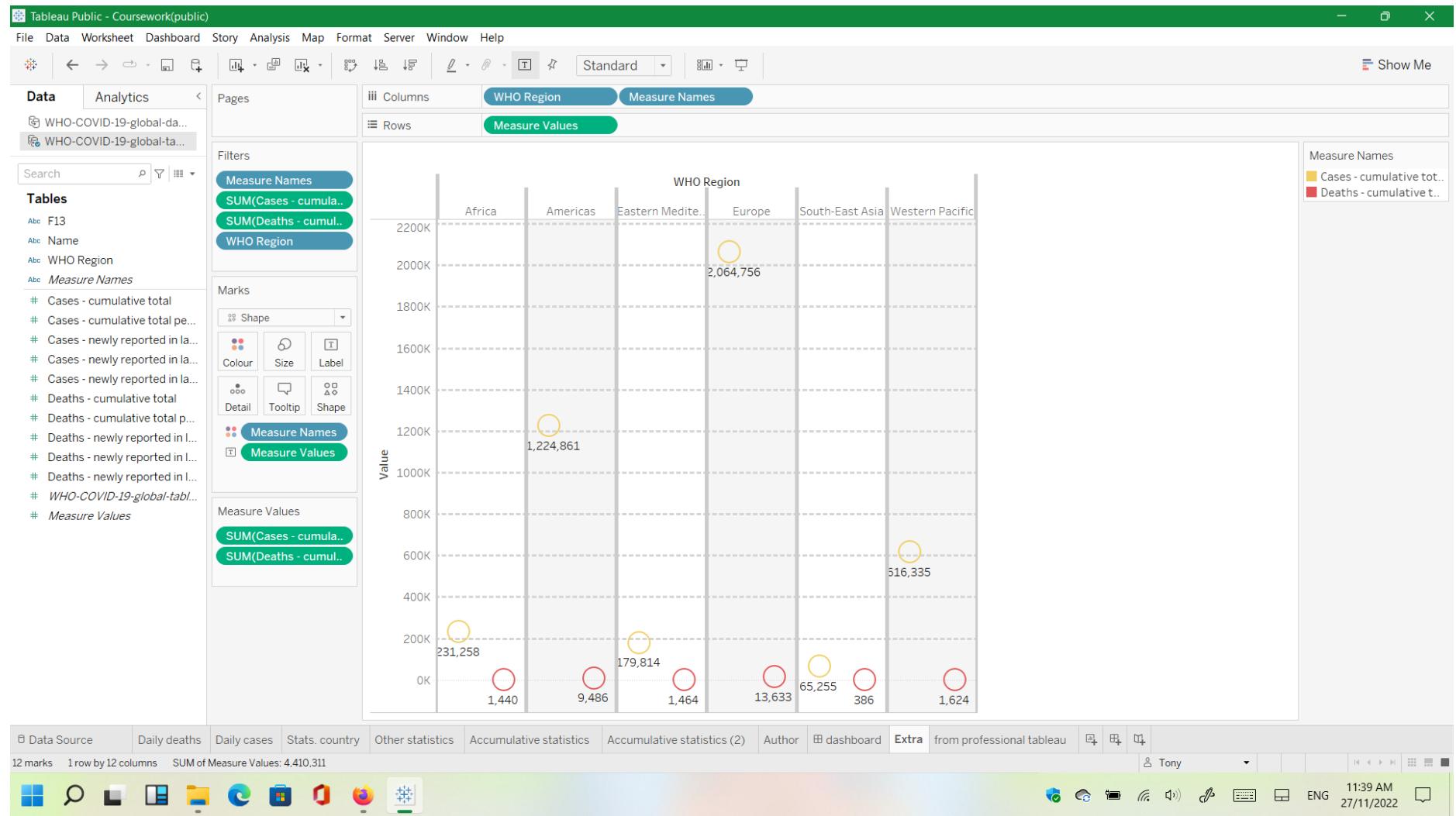
11:38 AM
27/11/2022

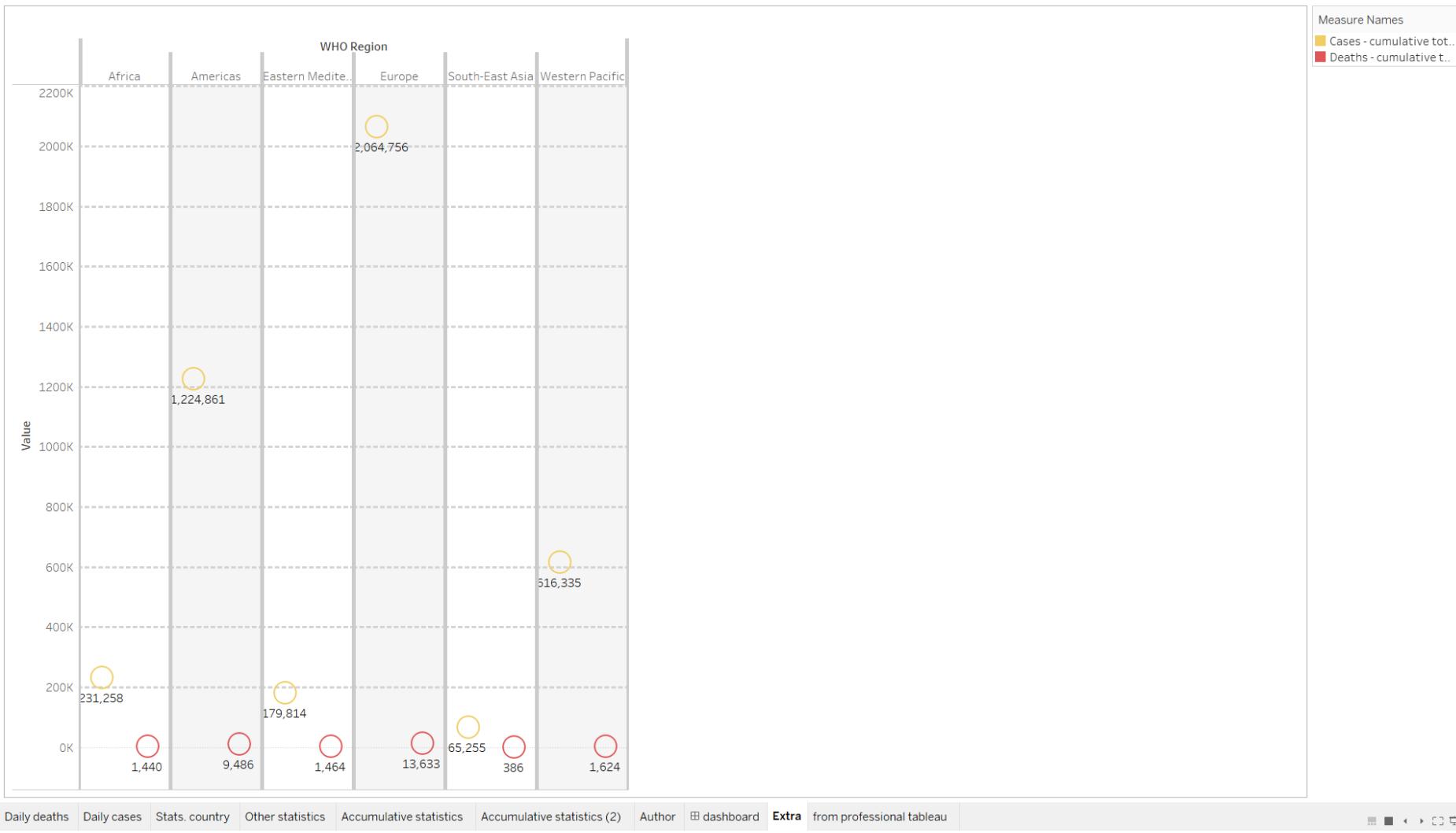
Tony (YIT19488399), Data Visualization, autumn module.
[UoR, year 3, BSc Computer Science](#)

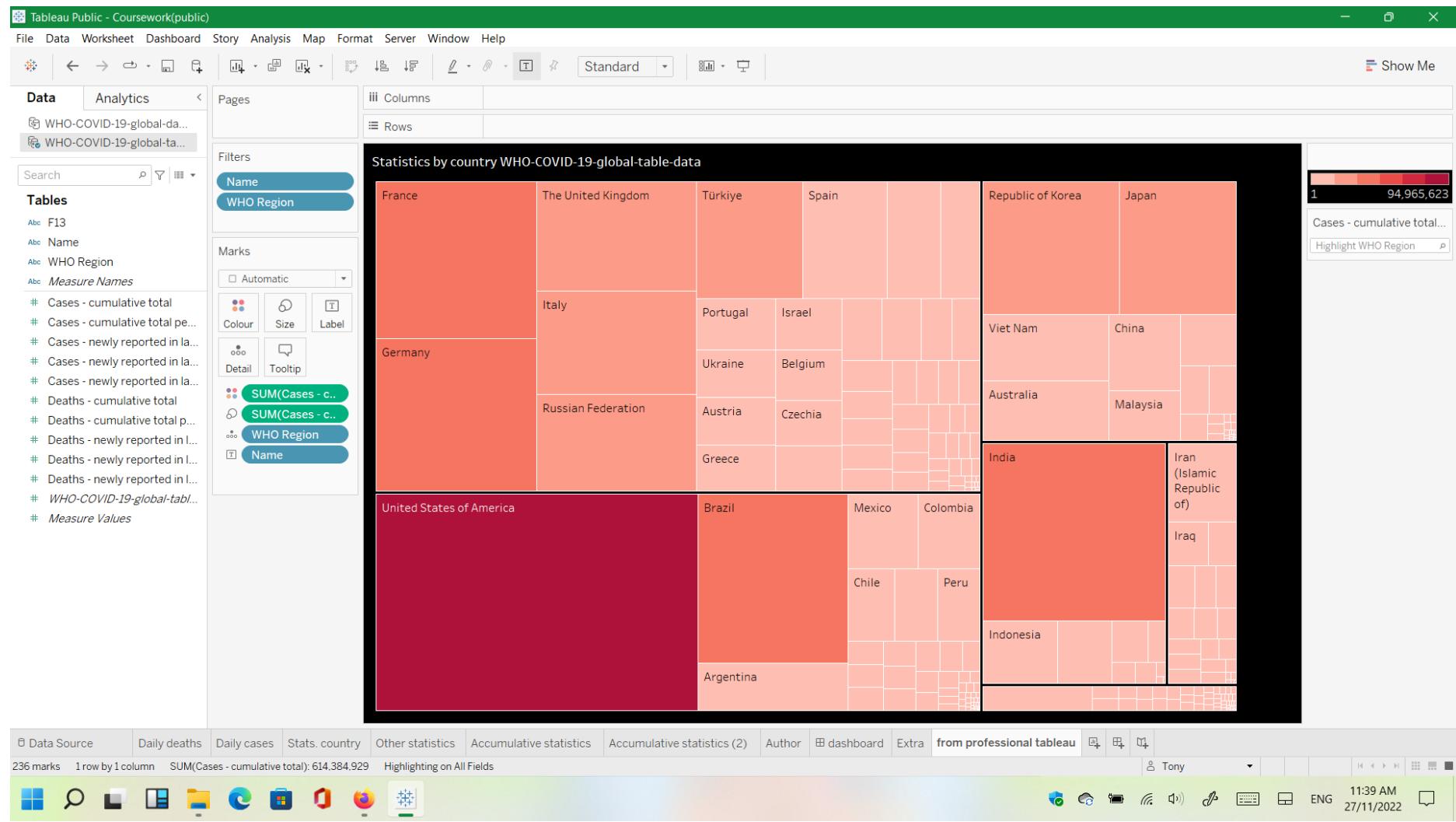
Drop field here

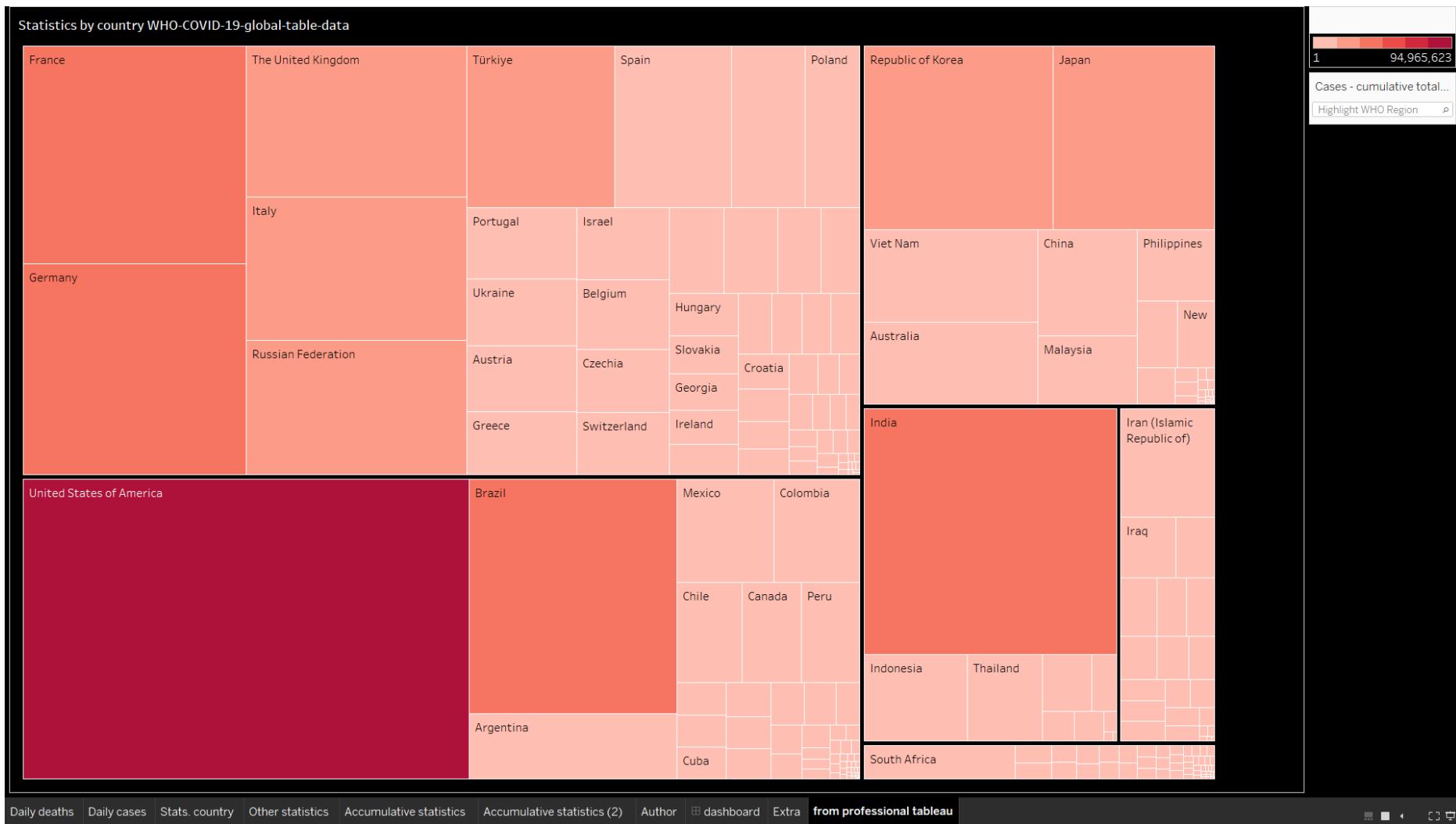
Drop
field
here

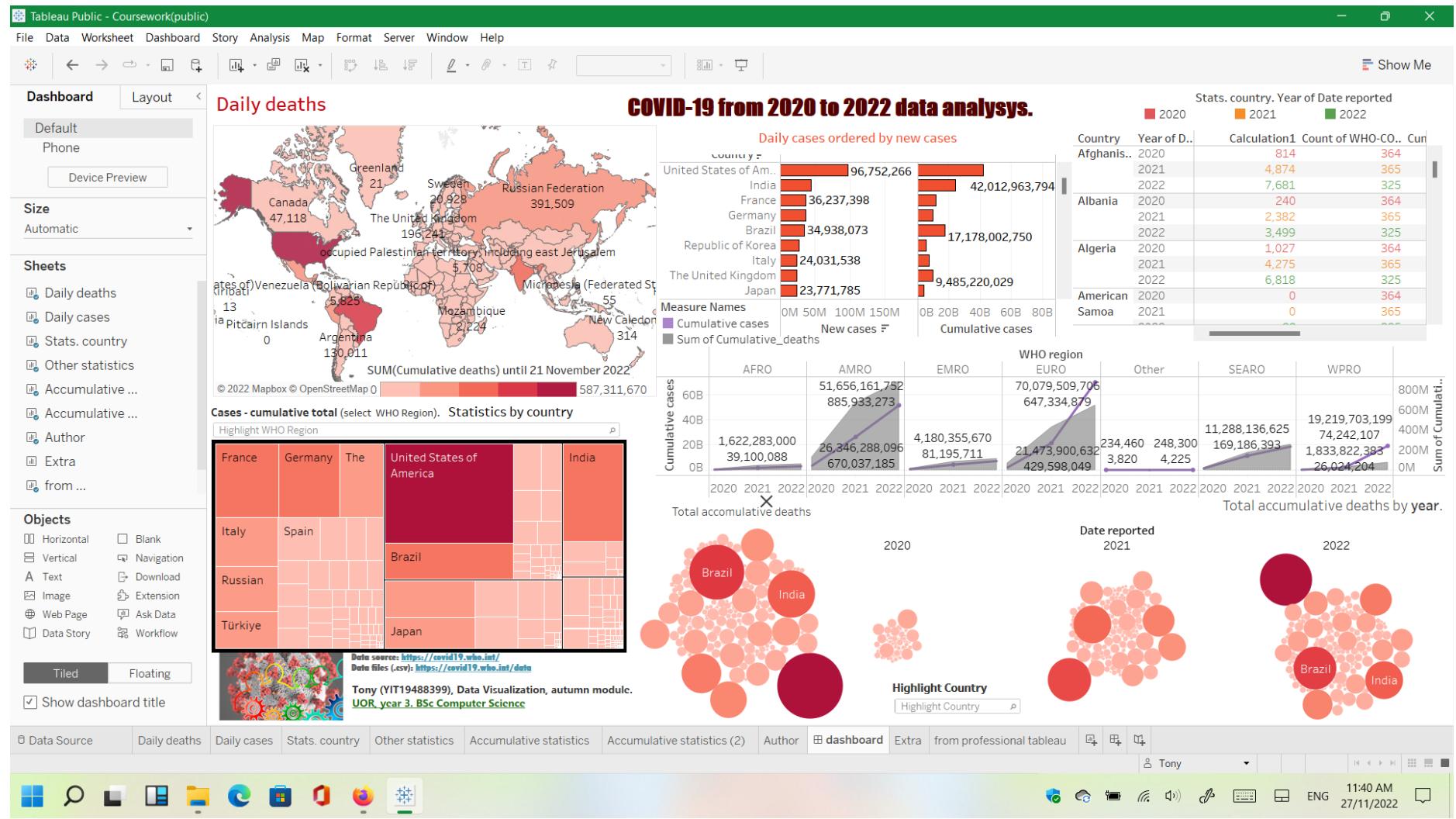
Drop field here



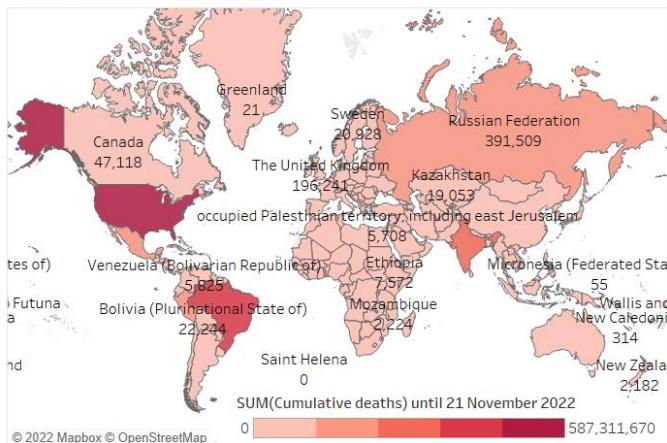




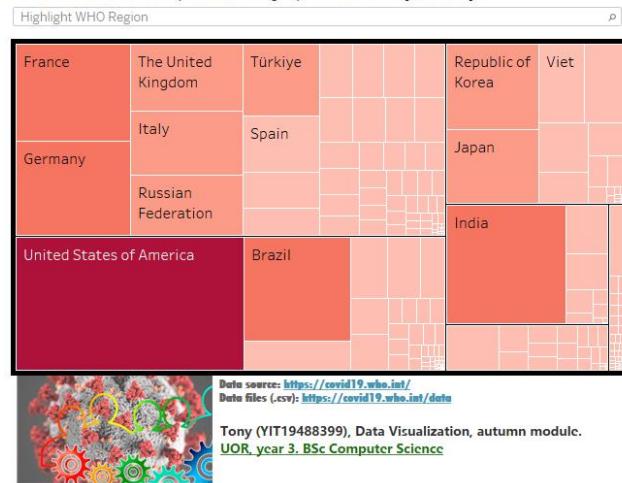




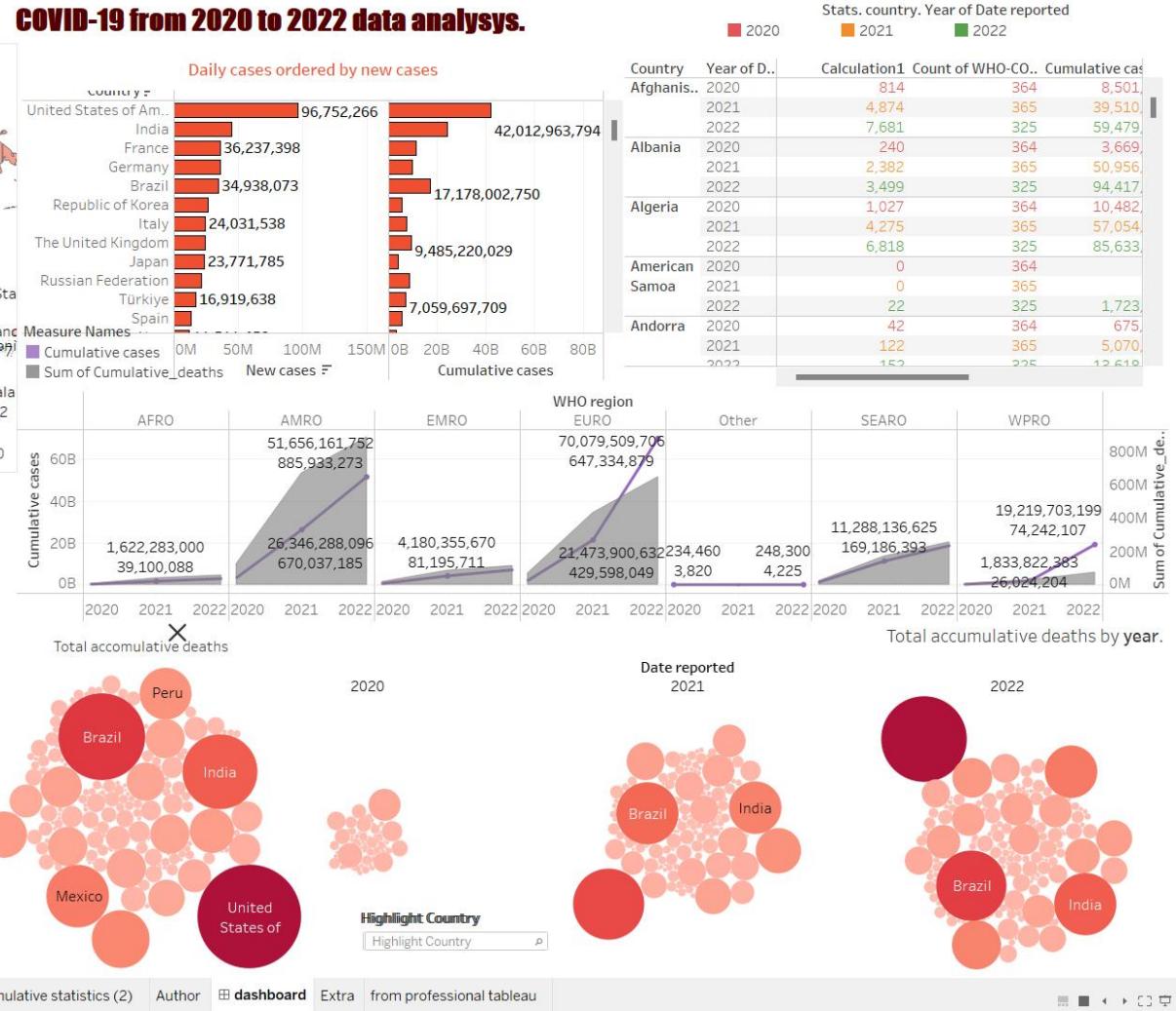
Daily deaths



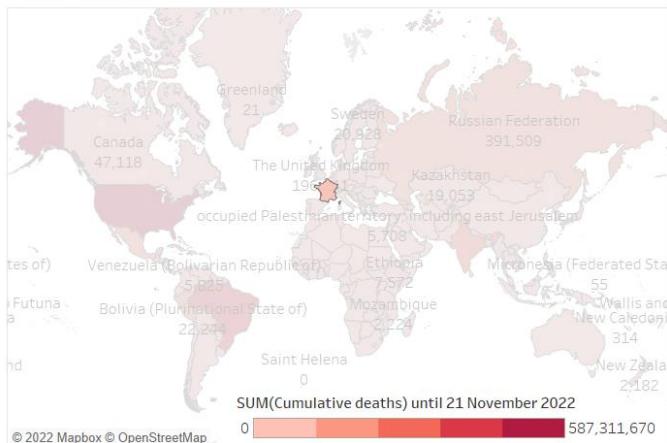
Cases - cumulative total (select WHO Region). Statistics by country



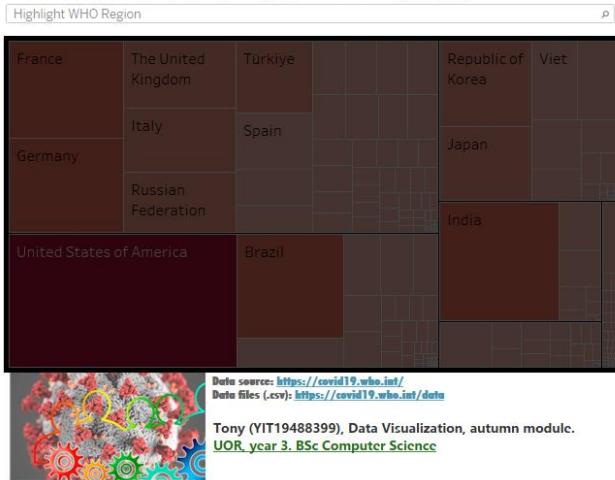
COVID-19 from 2020 to 2022 data analysis.



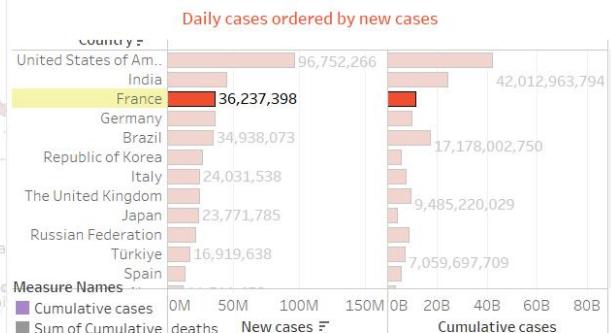
Daily deaths



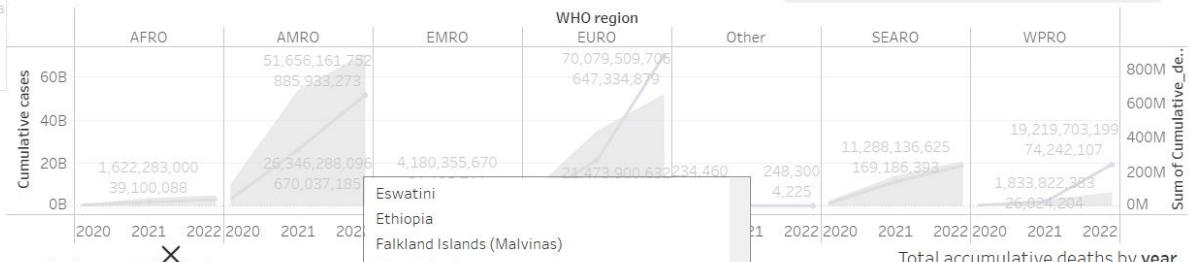
Cases - cumulative total (select WHO Region). Statistics by country



COVID-19 from 2020 to 2022 data analysys.



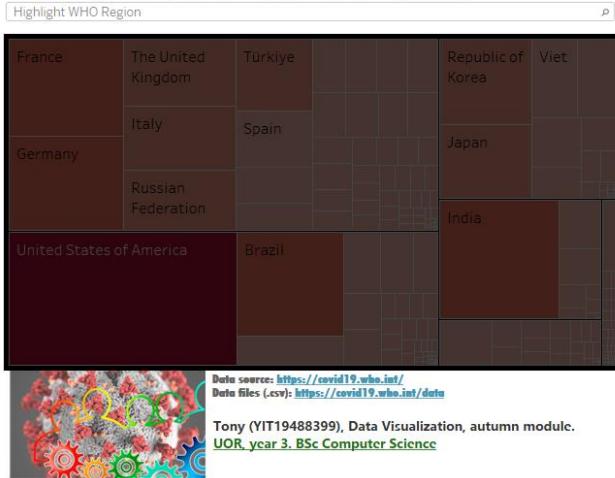
		Stats. country.	Year of Date reported	
Country	Year of D..	Calculation1	Count of WHO-CO..	Cumulative cas..
Afghanis..	2020	814	364	8,501
	2021	4,374	365	39,316
	2022	7,681	325	59,479
Albania	2020	240	364	3,669
	2021	2,382	365	50,956
	2022	3,499	325	94,417
Algeria	2020	1,027	364	10,482
	2021	4,275	365	57,054
	2022	6,818	325	85,633
American	2020	0	364	
	2021	0	365	
	2022	22	325	1,723
Andorra	2020	42	364	675
	2021	122	365	5,070
	2022	162	325	12,810



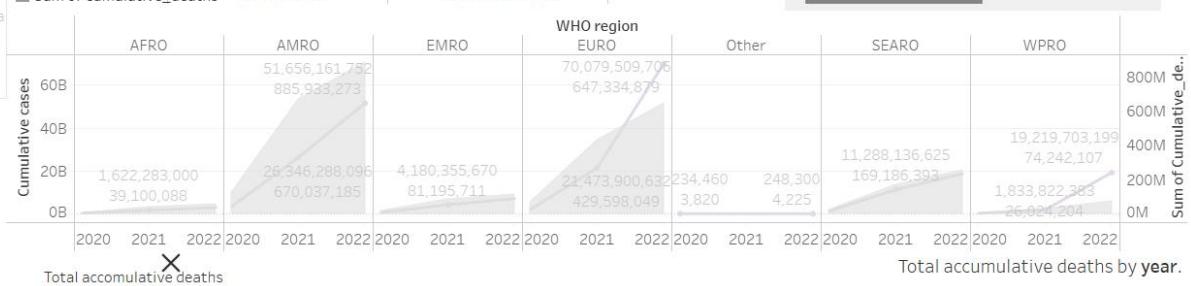
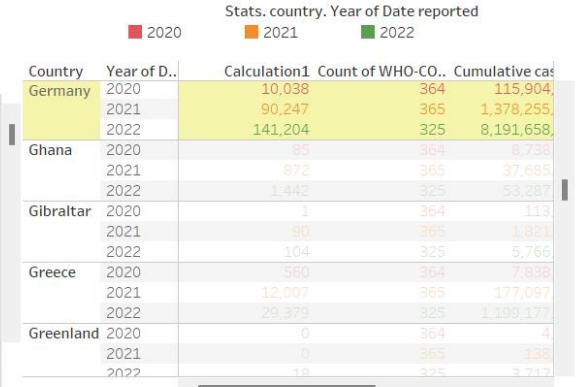
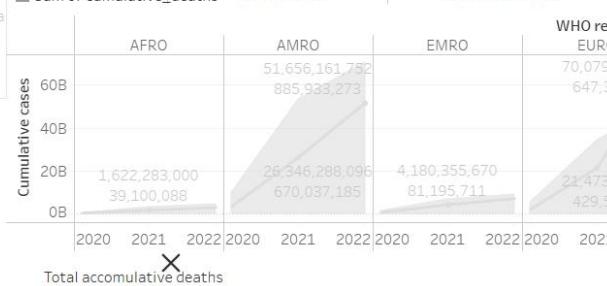
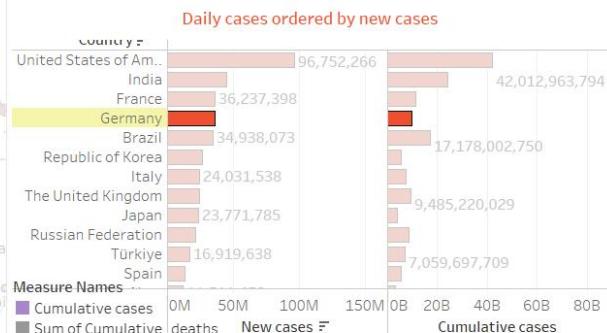
Daily deaths



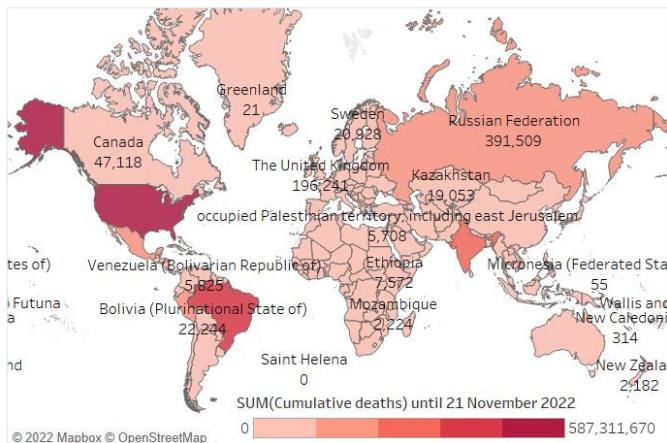
Cases - cumulative total (select WHO Region). Statistics by country



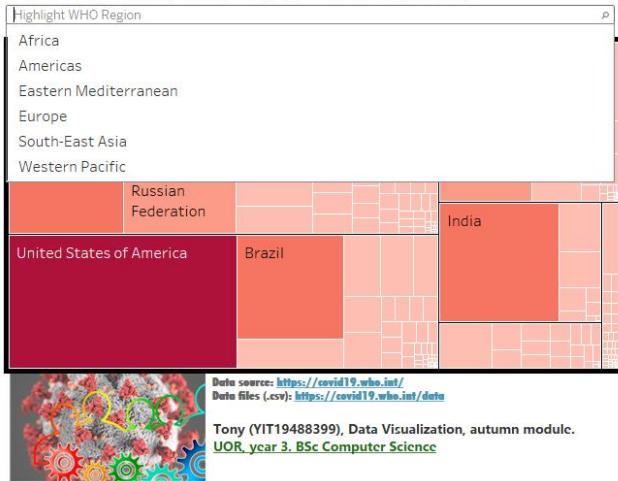
COVID-19 from 2020 to 2022 data analysys.



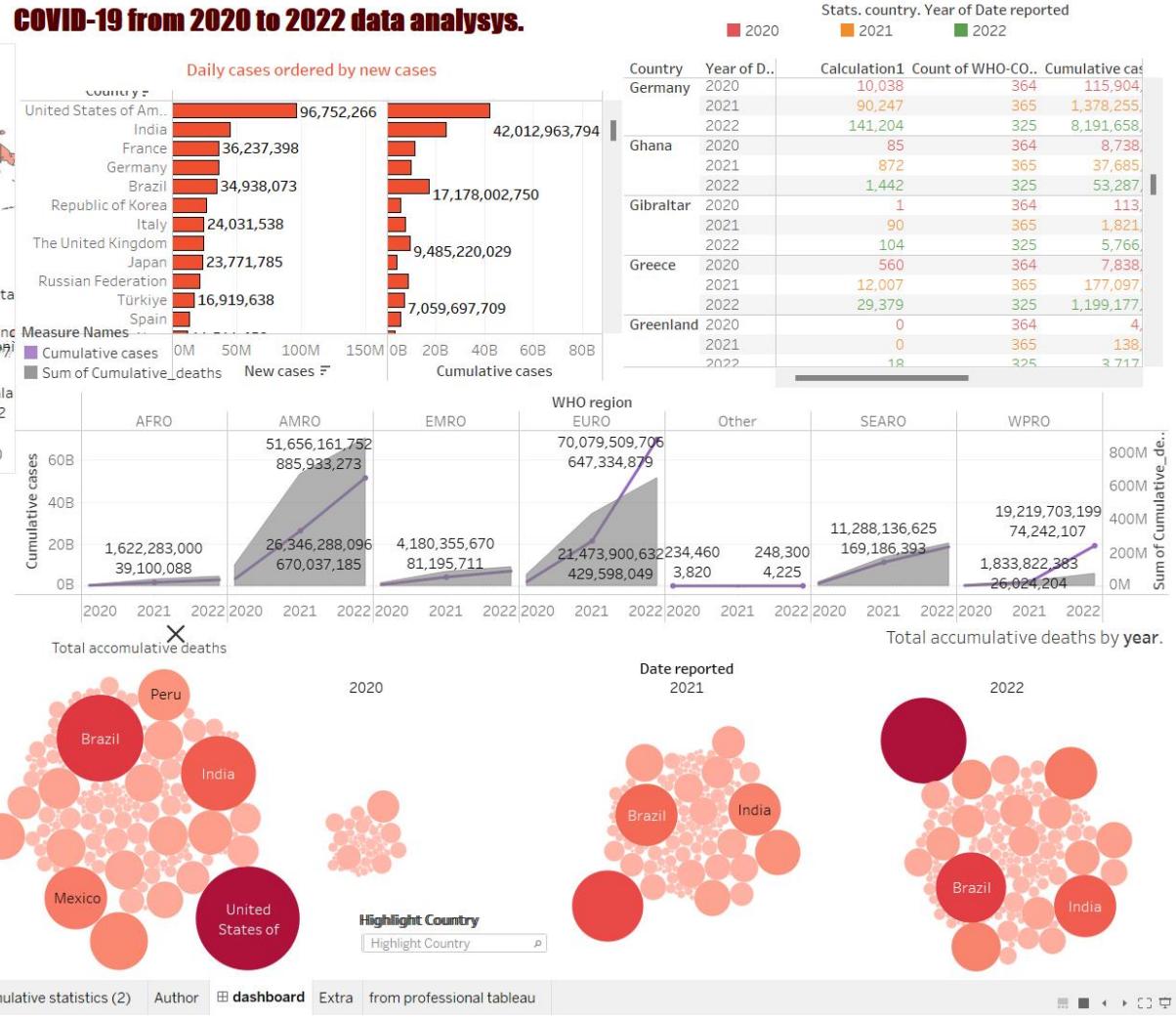
Daily deaths



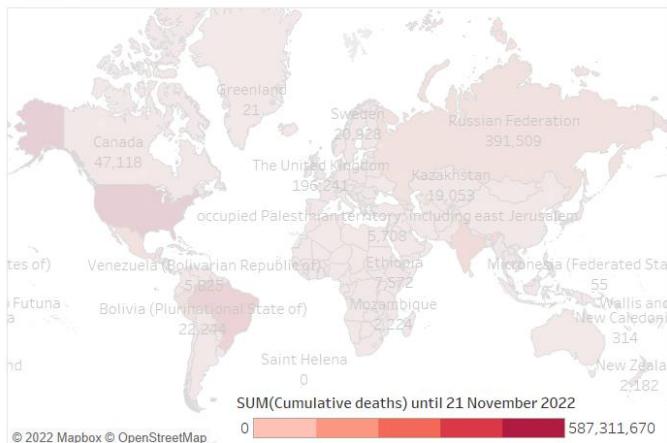
Cases - cumulative total (select WHO Region). Statistics by country



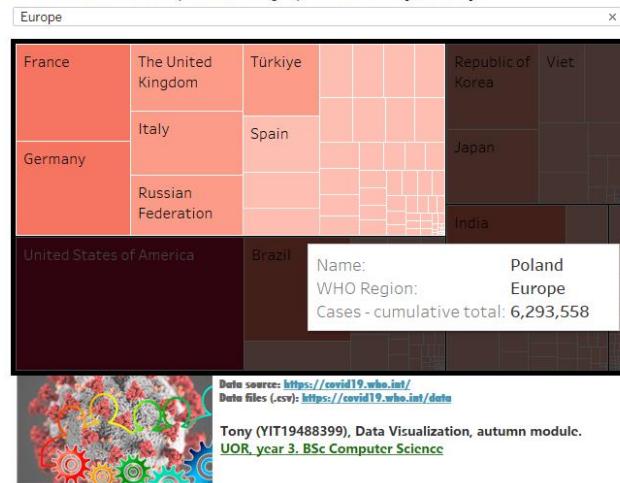
COVID-19 from 2020 to 2022 data analysis.



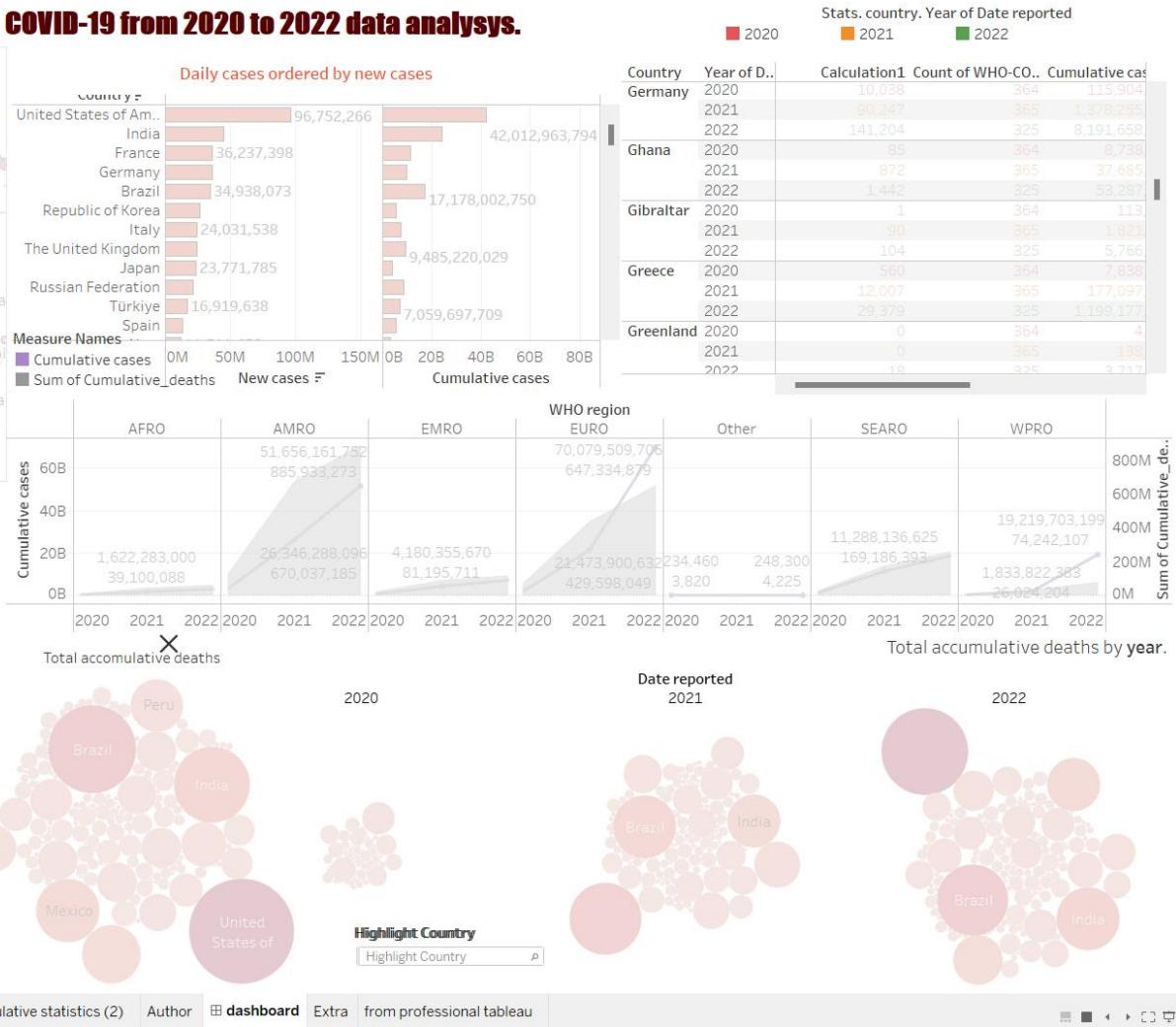
Daily deaths



Cases - cumulative total (select WHO Region). Statistics by country



COVID-19 from 2020 to 2022 data analysis.



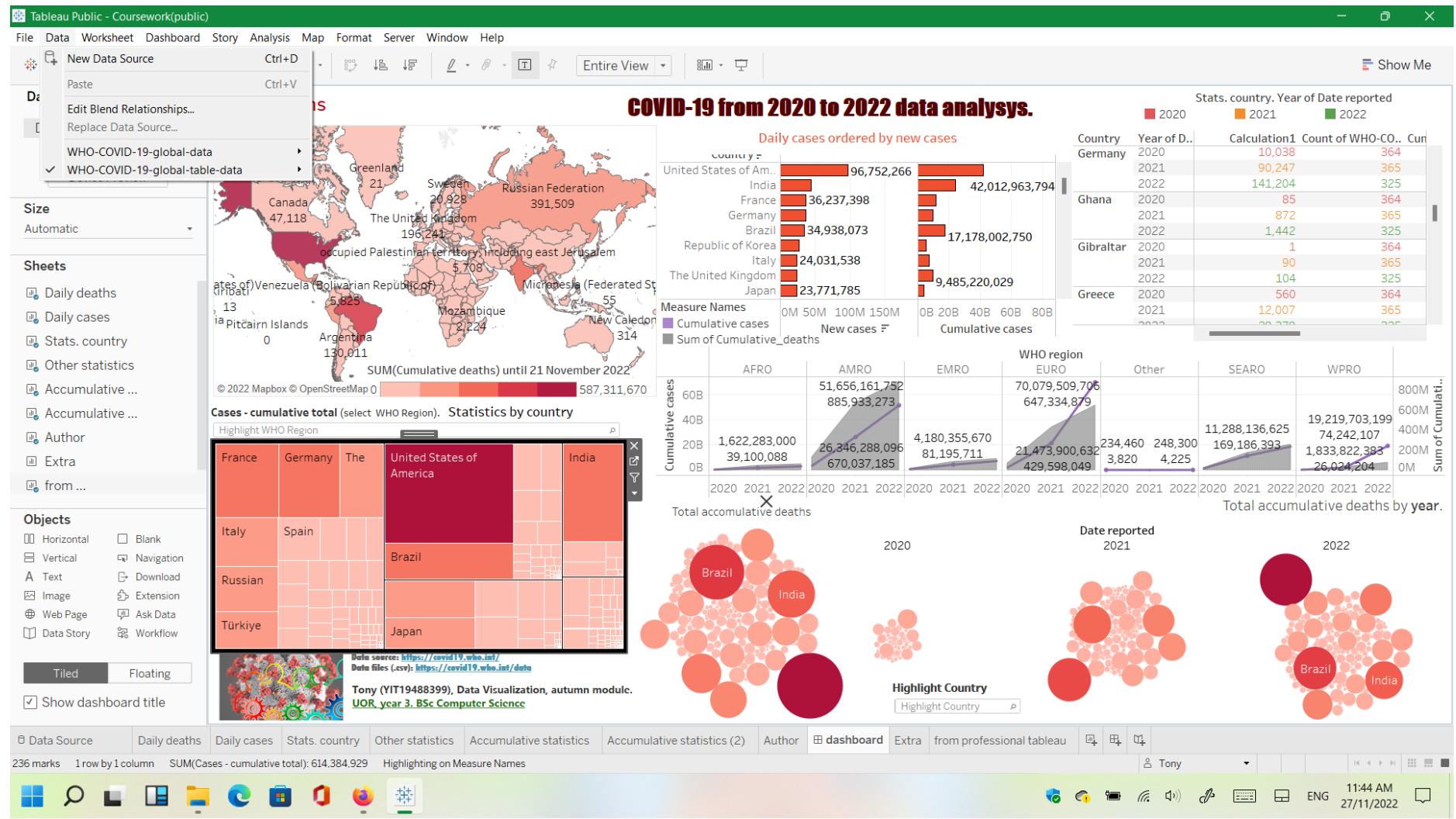


Tableau Public - Coursework(public)

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Dashboard Layout

Daily deaths

Optimise Workbook

Check Best Practices

Select an item to see the best practices guidelines and information on resolving issues.

• 14/17 Passed

Take action Updating these items to follow best practices won't impact workbook functionality.

Needs review Updating these items to follow best practices may require a trade-off in the workbook.

Dashboard size not fixed The dashboard *dashboard* is not fixed-size.

14 Passed Passed items follow best practice guidelines.

Rerun Optimiser Last run at 11:48

Published by Tony (YIT19488399), Data Visualization, autumn module. UoR, year 3. BSc Computer Science

Data source: <https://covid19.who.int/>
Data files (.csv): <https://covid19.who.int/WHO-COVID-19-global-data.csv>

Highlight Country

Highlight Country

Stats. country, Year of Date reported

Year	Count of WHO-COVID-19 cases	Count of WHO-COVID-19 deaths
2020	10,038	364
2021	90,247	365
2022	141,204	325
	85	364
	872	365
	1,442	325
	1	364
	90	365
	104	325
	560	364
	12,007	365
	20,220	325

Calculation1 Count of WHO-COVID-19 cases

Category	Value
SEARO	11,288,136,625
WPRO	19,219,703,199
Euro	18,300
AFRO	169,186,393
Middle East & Africa	74,242,107
Eastern Mediterranean	1,225
Sum of Cumulative	1,833,822,393
Total accumulative deaths by year	26,024,204

2022

Bubble chart showing total accumulative deaths by year for 2022. The chart includes data for Brazil and India.

Objects

- Horizontal
- Blank
- Vertical
- Navigation
- Text
- Download
- Image
- Extension
- Web Page
- Ask Data
- Data Story
- Workflow

Tiled Floating

Show dashboard title

Data Source Daily deaths Daily cases Stats. country Other statistics Accumulative statistics Accumulative statistics (2) Author dashboard Extra from professional tableau

Tony

11:48 AM 27/11/2022

Windows Taskbar icons

Tableau Public - Coursework(public)

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Show Me

Dashboard Layout

Default Phone Device Preview

Size Automatic

Sheets

- Daily deaths
- Daily cases
- Stats. country
- Other statistics
- Accumulative ...
- Accumulative ...
- Author
- Extra
- from ...

Objects

- Horizontal
- Blank
- Vertical
- Navigation
- A Text
- Download
- Image
- Extension
- Ask Data
- Web Page
- Data Story
- Workflow

Tiled Floating

Show dashboard title

Daily deaths Daily cases Stats. country Other statistics Accumulative statistics Accumulative statistics (2) Author dashboard Extra from professional tableau

Optimise Workbook

Check Best Practices

14/17 Passed

2 Take action Updating these items to follow best practices won't impact workbook functionality.

- Unused fields Multiple data sources have unused fields.
- Non-materialised calculations The data source WHO-COVID-19-global-data has not been pre-computed.

1 Needs review Updating these items to follow best practices may require a trade-off in the workbook.

14 Passed Passed items follow best practice guidelines.

Rerun Optimiser Last run at 11:48

Publish Close

Stats. country. Year of Date reported

Year	Count of WHO-COVID-19-global-data	Current
2020	10,038	364
2021	90,247	365
2022	141,204	325
	85	364
	872	365
	1,442	325
	1	364
	90	365
	104	325
	560	364
	12,007	365
	20,270	325

Calculation1 Count of WHO-COVID-19-global-data

SEARO WPRO

Year	SEARO	WPRO
2020	11,288,136,625	19,219,703,199
2021	169,186,393	74,242,107
2022	1,833,822,383	26,024,204

Total accumulative deaths by year.

2022

2021

2020

Brazil India

Tony (YIT1948399), Data Visualization, autumn module.
UoR year 3. BSc Computer Science

Highlight Country

Highlight Country

11:48 AM 27/11/2022

Notes:

Each worksheet has double screenshot, and the dashboard has meaningful visualizations to explain it. More screenshots with features are provided previously. And a video on YouTube explains better my dashboard. If necessary, I can comment my coursework in person with Lecturer (Mohammad Ahmad).

Refinement task 1 ("Provide a short write-up (no more than 2 paragraphs) describing your plot design and display."):

1. The simplest possible plot. Because is the first plot, I preprocessed, analyzed, and cleaned the data. It has 2 pictures and the smallest one is very informative, but not aesthetically nicely presented because python was designed for numerical and statistical data tasks at the beginning (data analysis not same data visualization) [1].

I use blue color for vaccination and red to put attention to top countries with more brands. Text inside is black to show seriousness and standardization.

2. More advance bar chart because has 2 plots and uses 2 different colors. 11 countries I think is okay but can be more if like. Countries are NOT the same for both plots, nevertheless coincidences is/are possible. Blue background because plot is about vaccines and WHO used a lot of blue

color. Red and green is to show difference between analyses in respective plots (data visualization).

3. This plot uses bubbles, and it opens in a browser. Legend clarifies continents and these colors are by default. Each bubble represents one country and as bigger is the bubble/circle, more deaths accumulate the country in question. It opens on a browser and this is the default for each user/machine (can be different browser and maybe slightly modifications are visible). I used Firefox.

This is the first plot with global data table dataset (NOT vaccinations dataset).

4. Same information as in figure 3 but analyzing ALL the countries in the planet in one single figure in one plot. The pie chart shows 2 information but only percentages are visible in the plot. Put mouse in the pie chart to see hidden information that will pop up.

Colors are by default and the plot opens again in a web browser.

Legend is very useful to speed understanding.

5. This is dynamic plot because it changes day by day. Even the data source is different from provided in coursework. The period is some years, and the animation shows the advance of the coronavirus in the planet. Plot opens in web browser and only requires press play bottom at the bottom left by the public. Visualizations take about a minute to be completed.

Red is adequate, but green is not. Although this green color shows clear and obvious difference (contrast). The substitute of the green color could be black, but this color is better to avoid when health is the theme (black = death).

Flaws in task 2 (Tableau Public):

Tree map about continents is not synchronized with the dashboard. When selecting a country from this tree map, dashboard doesn't work as expected. The same behavior happens when selecting country or continent in other part of the dashboard (NOT tree map), the tree map on the dashboard is not updating.

The conclusion is that Tableau is Business oriented and is not for "rocket science" analysis/studies/investigations. This behavior could be avoided if dashboard was designed with code like python, but this task will be much more tedious and complicated.

Summary: Tableau is good but can be improved and developed to analyze two datasets at the same time.

References:

- [1] ‘best programming language for data visualization’. best programming language for data visualization

Tableau Public own workspace:

<https://public.tableau.com/app/profile/tony6278>

Tableau Public dashboard direct link for task 2:

<https://public.tableau.com/app/profile/tony6278/viz/Courseworkpublic/dashboard>

YouTube video for Dashboard: <https://youtu.be/ceYehmEPF8k>

#END#