ANA-522-OL1 Spring 2022
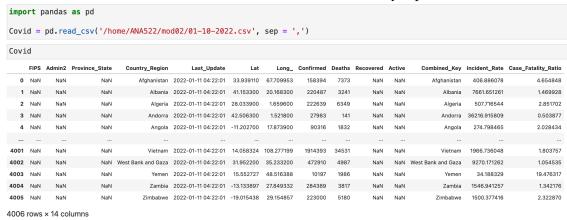# Mod02 HW: pandas
Due: Friday February 4th at midnight

Explore Covid-19 Dataset COVID-19 Data repository for the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) is available on https://github.com/CSSEGISandData/COVID-19. Within the repository, there are daily reports for Covid-19 cases around the world. The field description for daily reports dataset can be found here and copied in the following:

- FIPS: US only. Federal Information Processing Standards code that uniquely identifies counties within the USA.

- Admin2: County name. US only.

- Province_State: Province, state or dependency name.

- Country_Region: Country, region or sovereignty name. The names of locations included on the Website correspond with the official designations used by the U.S. Department of State.

- Last Update: MM/DD/YYYY HH:mm:ss (24 hour format, in UTC).

- Lat and Long_: Dot locations on the dashboard. All points (except for Australia) shown on the map are based on geographic centroids, and are not representative of a specific address, building or any location at a spatial scale finer than a province/state. Australian dots are located at the centroid of the largest city in each state.

- Confirmed: Counts include confirmed and probable (where reported).

- Deaths: Counts include confirmed and probable (where reported).

- Recovered: Recovered cases are estimates based on local media reports, and state and local reporting when available, and therefore may be substantially lower than the true number. US state-level recovered cases are from COVID Tracking Project.

- Active: Active cases = total cases - total recovered - total deaths.

- Incident_Rate: Incidence Rate = cases per 100,000 persons.

- Case_Fatality_Ratio (%): Case-Fatality Ratio (%) = Number recorded deaths / Number cases.

In this assignment, we are using only one daily report of January 10th, 2022. The dataset file of comma-separated values format is 01-10-2022.csv. You can create a new Jupyter Notebook for the assignment and add the first cell to load the dataset with statements in the following:

```
import pandas as pd
Covid = pd.read_csv('/home/ANA522/mod02/01-10-2022.csv', sep = ',')
```

You should be able to see the DataFrame of the Covid-19 daily report now and can start to explore.

```
import pandas as pd

Covid = pd.read_csv('/home/ANA522/mod02/01-10-2022.csv', sep = ',')

Covid
```

	FIPS	Admin2	Province_State	Country_Region	Last_Update	Lat	Long_	Confirmed	Deaths	Recovered	Active	Combined_Key	Incident_Rate	Case_Fatality_Ratio
0	NaN	NaN	NaN	Afghanistan	2022-01-11 04:22:01	33.939110	67.709953	158394	7373	NaN	NaN	Afghanistan	406.886078	4.654848
1	NaN	NaN	NaN	Albania	2022-01-11 04:22:01	41.153300	20.168300	220487	3241	NaN	NaN	Albania	7661.651261	1.469928
2	NaN	NaN	NaN	Algeria	2022-01-11 04:22:01	28.033900	1.659600	222639	6349	NaN	NaN	Algeria	507.716544	2.851702
3	NaN	NaN	NaN	Andorra	2022-01-11 04:22:01	42.506300	1.521800	27983	141	NaN	NaN	Andorra	36216.915809	0.503877
4	NaN	NaN	NaN	Angola	2022-01-11 04:22:01	-11.202700	17.873900	90316	1832	NaN	NaN	Angola	274.798465	2.028434
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4001	NaN	NaN	NaN	Vietnam	2022-01-11 04:22:01	14.058324	108.277199	1914393	34531	NaN	NaN	Vietnam	1966.736048	1.803757
4002	NaN	NaN	NaN	West Bank and Gaza	2022-01-11 04:22:01	31.952200	35.233200	472910	4987	NaN	NaN	West Bank and Gaza	9270.171262	1.054535
4003	NaN	NaN	NaN	Yemen	2022-01-11 04:22:01	15.552727	48.516388	10197	1986	NaN	NaN	Yemen	34.188329	19.476317
4004	NaN	NaN	NaN	Zambia	2022-01-11 04:22:01	-13.133897	27.849332	284389	3817	NaN	NaN	Zambia	1546.941257	1.342176
4005	NaN	NaN	NaN	Zimbabwe	2022-01-11 04:22:01	-19.015438	29.154857	223000	5180	NaN	NaN	Zimbabwe	1500.377416	2.322870

4006 rows × 14 columns

Please come up with Two (reasonably good) questions that you are interested in, and manipulate the data with the help of pandas to answer proposed questions. The format is similar to the pandas Lab exercise questions.

Turn in your Jupyter Notebook file (.ipynb) to upload it on the Blackboard.

Your two question statements (but not the answers) will be shared with the whole class in the Class Cafe Forum.