Course 7

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The disease was first identified in 2019 in Wuhan, China, and has since spread globally, resulting in the 2019–20 coronavirus pandemic. Common symptoms include fever, cough and shortness of breath. Muscle pain, sputum production and sore throat are less common. The rate of deaths per number of diagnosed cases is on average 3.4%, ranging from 0.2% in those less than 20 to approximately 15% in those over 80 years old.

Data Source (Date wise) : 2019 Novel Coronavirus COVID-19 (2019-nCoV) Data Repository by Johns Hopkins CSSE

Data Source: https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_daily_reports/08-19-2021.csv

In case it gives you error while reading this file from the specified path , please try the solutions indicated below: You can add them as an argument
error_bad_lines=False
1st Solution:
data = pd.read_csv(yourfileName', error_bad_lines=False)
2nd Solution
df = pandas.read_csv(yourfileName, sep='delimiter', header=None)
3rd Solution:

encoding="ISO-8859-1"

Field description

- Province/State: China - province name; US/Canada/Australia/ - city name, state/province name; Others - name of the event (e.g., "Diamond Princess" cruise ship); other countries - blank.
- Country/Region: country/region name conforming to WHO (will be updated).
- Last Update: MM/DD/YYYY HH:mm (24 hour format, in UTC).
- Confirmed: the number of confirmed cases. For Hubei Province: from Feb 13 (GMT +8), we report both clinically diagnosed and lab-confirmed cases. For lab-confirmed cases only (Before Feb 17), please refer to who_covid_19_situation_reports. For Italy, the diagnosis standard might be changed since Feb 27 to "slow the growth of new case numbers." (Source)
- Deaths: the number of deaths.
- Recovered: the number of recovered cases.

Using above dataset we have created some exercises on COVID-19 (Spread of the novel coronavirus, Analysis, Visualization, Prediction & Comparisons.

Please use the jupyter lab on your local computer to solve all the questions written below. Each group needs to share one ipynb file.

1. Write a Python program to display the first 5 rows from COVID-19 dataset. Also print the dataset information and check the missing values.

2. Write a python program to replace all the missing values with "Tobefound" for province column.
3. Normalize the values for "Recovered" column as per the z-score.
4. Create bin values for "Recovered" column in the data set.
5. Write a Python program to get the maximum number of confirmed, deaths, recovered and active cases of Novel Coronavirus (COVID-19) Country wise.
6. Write a Python program to get the total number of confirmed deaths and recovered people of Novel Coronavirus (COVID-19) cases Country/Region - Province/State wise
7. Write a Python program to get the country wise (country name taken as input from user) cases of confirmed, deaths and recovered cases of Novel Coronavirus (COVID-19).
8. Write a Python program to get the country wise deaths cases of Novel Coronavirus (COVID-19).
9. Write a Python program to list countries with no cases of Novel Coronavirus (COVID-19) recovered.
10. Write a Python program to list countries where all cases of Novel Coronavirus (COVID-19) died.
11. Write a Python program to list countries with all cases of Novel Coronavirus (COVID-19) recovered.
12. Write a Python program to get the top 10 countries data (Last Update, Country/Region, Confirmed, Deaths, Recovered) of Novel Coronavirus (COVID-19).
13. Write a Python program to create a plot (lines) of total deaths, confirmed, recovered and active cases Country wise where deaths greater than 150.