Title: Problem for Covid - 19 Data Analysis Project using Python

Dataset link :

Url = https://raw.githubusercontent.com/SR1608/Datasets/main/covid-data.csv

Perform following analysis on above dataset :

1. Import the dataset using Pandas from above mentioned url.

2. High Level Data Understanding:
- **a. Find no. of rows & columns in the dataset**
- **b. Data types of columns.**
- **c. Info & describe of data in dataframe.**

3. Low Level Data Understanding :
- **a. Find count of unique values in location column.**
- **b. Find which continent has maximum frequency using values counts.**
- **c. Find maximum & mean value in 'total_cases'.**
- **d. Find 25%,50% & 75% quartile value in 'total_deaths'.**
- **e. Find which continent has maximum 'human_development_index'.**
- **f. Find which continent has minimum 'gdp_per_capita'.**

4. Filter the dataframe with only this columns ['continent','location','date','total_cases','total_deaths','gdp_per_capita',' human_development_index'] and update the data frame.

5. Data Cleaning
- **a. Remove all duplicates observations**
- **b. Find missing values in all columns**
- **c. Remove all observations where continent column value is missing**
- **Tip : using subset parameter in dropna**
- **d. Fill all missing values with 0**

6. Date time format :
- **a. Convert date column in datetime format using pandas.to_datetime**

b. Create new column month after extracting month data from date column.

7. Data Aggregation:

a. Find max value in all columns using groupby function on 'continent' column

Tip: use reset_index() after applying groupby

b. Store the result in a new dataframe named 'df_groupby'. (Use df_groupby dataframe for all further analysis)

8. Feature Engineering :

a. Create a new feature 'total_deaths_to_total_cases' by ratio of 'total_deaths' column to 'total_cases'

9. Data Visualization :

a. Perform Univariate analysis on 'gdp_per_capita' column by plotting histogram using seaborn dist plot.

b. Plot a scatter plot of 'total_cases' & 'gdp_per_capita'

c. Plot Pairplot on df_groupby dataset.

d. Plot a bar plot of 'continent' column with 'total_cases' .

Tip : using kind='bar' in seaborn catplot

10.Save the df_groupby dataframe in your local drive using pandas.to_csv function .