

Pandas

1. Pandas is the library/class in python that will help you deal with dataframes and data series.
2. It has functions for analyzing, cleaning, exploring, and manipulating data.
3. Pandas can clean messy data sets, and make them readable and relevant
4. What can pandas do, Pandas gives you answers about the data. Like: a. Is there a correlation between two or more columns b. What is average value? c. If there are any na values in the data d. Range of values of column

```
In [1]: 1 import pandas as pd  
        2  
        3 print(pd.__version__)
```

1.1.5

Reading a dataframe

```
In [2]: 1 df = pd.read_csv('gapminder.csv')
```

In [3]:

```
1 df
```

Out[3]:

	population	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific
...
134	3350832.0	2.11	0.5	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America
135	26952719.0	2.46	0.1	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia
136	86589342.0	1.86	0.4	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific
137	13114579.0	5.88	13.6	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa
138	13495462.0	3.85	15.1	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa

139 rows × 10 columns

In [8]:

```
1 df.head()
```

Out[8]:

	population	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific

In [10]: 1 df.tail()

Out[10]:

	population	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
134	3350832.0	2.11	0.5	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America
135	26952719.0	2.46	0.1	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia
136	86589342.0	1.86	0.4	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific
137	13114579.0	5.88	13.6	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa
138	13495462.0	3.85	15.1	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa

In [11]: 1 df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 139 entries, 0 to 138
Data columns (total 10 columns):
#   Column                Non-Null Count  Dtype
---  -
0   population            139 non-null    float64
1   fertility             139 non-null    float64
2   HIV                  139 non-null    float64
3   CO2                  139 non-null    float64
4   BMI_male             139 non-null    float64
5   GDP                  139 non-null    float64
6   BMI_female           139 non-null    float64
7   life                 139 non-null    float64
8   child_mortality       139 non-null    float64
9   Region               139 non-null    object
dtypes: float64(9), object(1)
memory usage: 11.0+ KB
```

In [12]: 1 df.shape

Out[12]: (139, 10)

In [13]: 1 df.columns

Out[13]: Index(['population', 'fertility', 'HIV', 'CO2', 'BMI_male', 'GDP',
 'BMI_female', 'life', 'child_mortality', 'Region'],
 dtype='object')

In [14]: 1 df.describe()

Out[14]:

	population	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality
count	1.390000e+02	139.000000	139.000000	139.000000	139.000000	139.000000	139.000000	139.000000	139.000000
mean	3.549977e+07	3.005108	1.915612	4.459874	24.623054	16638.784173	126.701914	69.602878	45.097122
std	1.095121e+08	1.615354	4.408974	6.268349	2.209368	19207.299083	4.471997	9.122189	45.724667
min	2.773150e+05	1.280000	0.060000	0.008618	20.397420	588.000000	117.375500	45.200000	2.700000
25%	3.752776e+06	1.810000	0.100000	0.496190	22.448135	2899.000000	123.232200	62.200000	8.100000
50%	9.705130e+06	2.410000	0.400000	2.223796	25.156990	9938.000000	126.519600	72.000000	24.000000
75%	2.791973e+07	4.095000	1.300000	6.589156	26.497575	23278.500000	130.275900	76.850000	74.200000
max	1.197070e+09	7.590000	25.900000	48.702062	28.456980	126076.000000	135.492000	82.600000	192.000000

In [18]: 1 df.rename(columns = {
 2 'population' : 'pop'
 3 },inplace = True) #renaming columns

In [19]:

1 df

Out[19]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific
...
134	3350832.0	2.11	0.5	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America
135	26952719.0	2.46	0.1	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia
136	86589342.0	1.86	0.4	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific
137	13114579.0	5.88	13.6	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa
138	13495462.0	3.85	15.1	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa

139 rows × 10 columns

In [34]:

```

1 import numpy as np
2 df.replace('Middle East & North Africa',np.NaN,inplace = True)#No reassignment to variable required since : inplace

```

```
In [35]: 1 df.isnull().sum()# Total number of na/null values
```

```
Out[35]: pop                0
fertility                 0
HIV                      0
CO2                      0
BMI_male                 0
GDP                      0
BMI_female               0
life                     0
child_mortality          0
Region                   10
dtype: int64
```

```
In [36]: 1 df
```

```
Out[36]:
```

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	NaN
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific
...
134	3350832.0	2.11	0.5	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America
135	26952719.0	2.46	0.1	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia
136	86589342.0	1.86	0.4	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific
137	13114579.0	5.88	13.6	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa
138	13495462.0	3.85	15.1	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa

139 rows × 10 columns

In [37]: 1 df.dropna()

Out[37]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific
5	8331465.0	1.41	0.3	8.183160	26.46741	43952.0	124.1394	80.4	4.6	Europe & Central Asia
...
134	3350832.0	2.11	0.5	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America
135	26952719.0	2.46	0.1	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia
136	86589342.0	1.86	0.4	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific
137	13114579.0	5.88	13.6	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa
138	13495462.0	3.85	15.1	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa

129 rows × 10 columns

In [38]: 1 df = df.fillna('Middle East & North Africa')*#fill all na values in the dataframe with Middle East & North Africa*

```
In [39]: 1 df.isna().sum()
```

```
Out[39]: pop                0
fertility                 0
HIV                      0
CO2                      0
BMI_male                 0
GDP                      0
BMI_female               0
life                    0
child_mortality          0
Region                   0
dtype: int64
```

```
In [41]: 1 df['Region']
```

```
Out[41]: 0    Middle East & North Africa
1         Sub-Saharan Africa
2                America
3    Europe & Central Asia
4    East Asia & Pacific
      ...
134                America
135    Europe & Central Asia
136    East Asia & Pacific
137    Sub-Saharan Africa
138    Sub-Saharan Africa
Name: Region, Length: 139, dtype: object
```


In [42]:

1 df

Out[42]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific
...
134	3350832.0	2.11	0.5	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America
135	26952719.0	2.46	0.1	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia
136	86589342.0	1.86	0.4	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific
137	13114579.0	5.88	13.6	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa
138	13495462.0	3.85	15.1	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa

139 rows × 10 columns

```
In [55]: 1 df[['pop', 'CO2', 'Region']]
```

```
Out[55]:
```

	pop	CO2	Region
0	34811059.0	3.328945	Middle East & North Africa
1	19842251.0	1.474353	Sub-Saharan Africa
2	40381860.0	4.785170	America
3	2975029.0	1.804106	Europe & Central Asia
4	21370348.0	18.016313	East Asia & Pacific
...
134	3350832.0	2.489764	America
135	26952719.0	4.476669	Europe & Central Asia
136	86589342.0	1.479347	East Asia & Pacific
137	13114579.0	0.148982	Sub-Saharan Africa
138	13495462.0	0.654323	Sub-Saharan Africa

139 rows × 3 columns

```
In [45]: 1 type(df['Region'])
```

```
Out[45]: pandas.core.series.Series
```

```
In [46]: 1 type(df[['pop', 'CO2', 'Region']])
```

```
Out[46]: pandas.core.frame.DataFrame
```

In [48]: 1 df.head()

Out[48]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific

In [50]: 1 df.iloc[1,2]

Out[50]: 2.0

In [52]: 1 df.iloc[0:10,0:3] # iloc[0:10,0:3] -> Columns 0,1,2 and rows 0,1,2...9

Out[52]:

	pop	fertility	HIV
0	34811059.0	2.73	0.10
1	19842251.0	6.43	2.00
2	40381860.0	2.24	0.50
3	2975029.0	1.40	0.10
4	21370348.0	1.96	0.10
5	8331465.0	1.41	0.30
6	8868713.0	1.99	0.10
7	348587.0	1.89	3.10
8	148252473.0	2.38	0.06
9	277315.0	1.83	1.30

```
In [62]: 1 df[df['Region'] == 'Middle East & North Africa']
          2
```

Out[62]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.10	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa
34	809639.0	3.76	2.60	0.612800	23.38403	2502.0	129.3376	61.0	81.0	Middle East & North Africa
36	78976122.0	2.95	0.06	2.512394	26.73243	9974.0	125.0931	70.1	31.4	Middle East & North Africa
60	72530693.0	1.88	0.20	7.892211	25.31003	15955.0	125.1859	76.9	21.4	Middle East & North Africa
62	7093808.0	2.92	0.20	10.001188	27.13151	28562.0	121.0838	80.9	4.9	Middle East & North Africa
69	4109389.0	1.57	0.10	3.996722	27.20117	14158.0	127.5037	77.6	11.3	Middle East & North Africa
85	31350544.0	2.44	0.10	1.594083	25.63182	6091.0	126.5284	73.3	35.8	Middle East & North Africa
95	2652281.0	2.89	0.10	15.572080	26.24109	47799.0	126.8870	74.6	11.9	Middle East & North Africa
104	1388962.0	2.20	0.06	48.702062	28.13138	126076.0	126.3153	80.4	9.5	Middle East & North Africa
128	10408091.0	2.04	0.06	2.440669	25.15699	9938.0	128.6291	76.5	19.4	Middle East & North Africa

```
In [70]: 1 df[df['fertility'] < 1.3]
```

Out[70]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
111	4849641.0	1.28	0.1	4.114441	23.83996	65991.0	121.1736	80.9	2.8	East Asia & Pacific

```
In [77]: 1 df['Size'] = 10
```

In [79]:

```
1 df.head()
```

Out[79]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region	Size
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa	10
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa	10
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America	10
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia	10
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific	10

In [80]:

```
1 df = df.drop('Size',axis=1)
```

In [81]:

```
1 df
```

Out[81]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific
...
134	3350832.0	2.11	0.5	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America
135	26952719.0	2.46	0.1	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia
136	86589342.0	1.86	0.4	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific
137	13114579.0	5.88	13.6	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa
138	13495462.0	3.85	15.1	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa

139 rows × 10 columns

In [82]:

```
1 df['Region'].value_counts()
```

Out[82]:

Europe & Central Asia	41
Sub-Saharan Africa	40
America	27
East Asia & Pacific	14
Middle East & North Africa	10
South Asia	7

Name: Region, dtype: int64

In [83]:

1 df

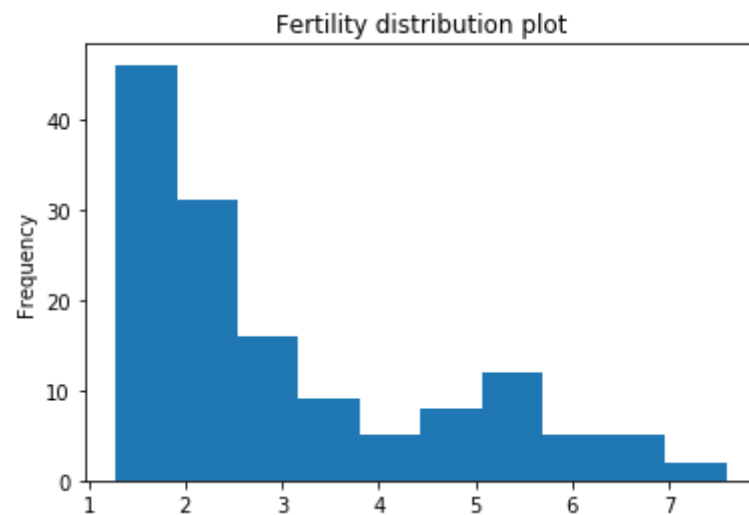
Out[83]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region
0	34811059.0	2.73	0.1	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa
1	19842251.0	6.43	2.0	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa
2	40381860.0	2.24	0.5	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America
3	2975029.0	1.40	0.1	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia
4	21370348.0	1.96	0.1	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific
...
134	3350832.0	2.11	0.5	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America
135	26952719.0	2.46	0.1	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia
136	86589342.0	1.86	0.4	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific
137	13114579.0	5.88	13.6	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa
138	13495462.0	3.85	15.1	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa

139 rows × 10 columns

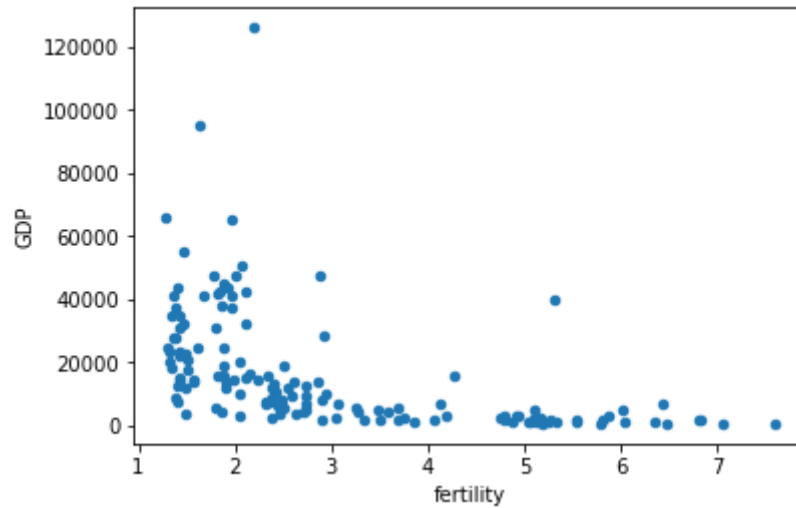
```
In [86]: 1 import matplotlib.pyplot as plt  
2 df['fertility'].plot(kind = 'hist',title = 'Fertility distribution plot')
```

Out[86]: <matplotlib.axes._subplots.AxesSubplot at 0x1e859ea3d48>




```
In [88]: 1 df.plot(kind = 'scatter',x= 'fertility',y = 'GDP')
```

```
Out[88]: <matplotlib.axes._subplots.AxesSubplot at 0x1e85a0b0188>
```



```
In [89]: 1 df.to_csv('updated_data.csv')
```

```
In [90]: 1 data = df
```

```
In [91]: 1 data.to_csv('updated_data.csv')
```

```
In [94]: 1 df['GDP'].describe()
```

```
Out[94]: count      139.000000  
mean      16638.784173  
std       19207.299083  
min        588.000000  
25%       2899.000000  
50%       9938.000000  
75%      23278.500000  
max      126076.000000  
Name: GDP, dtype: float64
```

```
In [105]: 1 def rich_or_not(gdp):  
2     if gdp<2899:  
3         return 'Poor'  
4     elif gdp>23278:  
5         return 'Rich'  
6     else:  
7         return 'Middle Income'  
8
```

```
In [106]: 1 df['Rich_or_Not'] = df['GDP'].apply(rich_or_not)
```

In [108]: 1 df.head(10)

Out[108]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region	Rich_or_Not
0	34811059.0	2.73	0.10	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa	Middle Income
1	19842251.0	6.43	2.00	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa	Middle Income
2	40381860.0	2.24	0.50	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America	Middle Income
3	2975029.0	1.40	0.10	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia	Middle Income
4	21370348.0	1.96	0.10	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific	Rich
5	8331465.0	1.41	0.30	8.183160	26.46741	43952.0	124.1394	80.4	4.6	Europe & Central Asia	Rich
6	8868713.0	1.99	0.10	5.109538	25.65117	14365.0	128.6024	70.6	43.3	Europe & Central Asia	Middle Income
7	348587.0	1.89	3.10	3.131921	27.24594	24373.0	124.3862	72.2	14.5	America	Rich
8	148252473.0	2.38	0.06	0.319161	20.39742	2265.0	125.0307	68.4	55.9	South Asia	Poor
9	277315.0	1.83	1.30	6.008279	26.38439	16075.0	126.3940	75.3	15.4	America	Middle Income

In [109]: 1 df.tail(10)

Out[109]:

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region	Rich_or_Not
129	70344357.0	2.15	0.06	4.021903	26.70371	16454.0	124.0675	75.1	22.2	Europe & Central Asia	Middle Income
130	31014427.0	6.34	6.40	0.100853	22.35833	1437.0	134.5204	57.2	89.3	Sub-Saharan Africa	Poor
131	46028476.0	1.38	1.10	7.032359	25.42379	8762.0	131.4962	68.2	12.9	Europe & Central Asia	Middle Income
132	61689620.0	1.87	0.20	8.526467	27.39249	37739.0	124.0845	79.5	5.6	Europe & Central Asia	Rich
133	304473143.0	2.07	0.60	18.545992	28.45698	50384.0	118.4777	78.2	7.7	America	Rich
134	3350832.0	2.11	0.50	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America	Middle Income
135	26952719.0	2.46	0.10	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia	Middle Income
136	86589342.0	1.86	0.40	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific	Middle Income
137	13114579.0	5.88	13.60	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa	Middle Income
138	13495462.0	3.85	15.10	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa	Poor

```
In [110]: 1 pd.concat([df.head(10),df.tail(10)],axis = 0)
```

```
Out[110]:
```

	pop	fertility	HIV	CO2	BMI_male	GDP	BMI_female	life	child_mortality	Region	Rich_or_Not
0	34811059.0	2.73	0.10	3.328945	24.59620	12314.0	129.9049	75.3	29.5	Middle East & North Africa	Middle Income
1	19842251.0	6.43	2.00	1.474353	22.25083	7103.0	130.1247	58.3	192.0	Sub-Saharan Africa	Middle Income
2	40381860.0	2.24	0.50	4.785170	27.50170	14646.0	118.8915	75.5	15.4	America	Middle Income
3	2975029.0	1.40	0.10	1.804106	25.35542	7383.0	132.8108	72.5	20.0	Europe & Central Asia	Middle Income
4	21370348.0	1.96	0.10	18.016313	27.56373	41312.0	117.3755	81.5	5.2	East Asia & Pacific	Rich
5	8331465.0	1.41	0.30	8.183160	26.46741	43952.0	124.1394	80.4	4.6	Europe & Central Asia	Rich
6	8868713.0	1.99	0.10	5.109538	25.65117	14365.0	128.6024	70.6	43.3	Europe & Central Asia	Middle Income
7	348587.0	1.89	3.10	3.131921	27.24594	24373.0	124.3862	72.2	14.5	America	Rich
8	148252473.0	2.38	0.06	0.319161	20.39742	2265.0	125.0307	68.4	55.9	South Asia	Poor
9	277315.0	1.83	1.30	6.008279	26.38439	16075.0	126.3940	75.3	15.4	America	Middle Income
129	70344357.0	2.15	0.06	4.021903	26.70371	16454.0	124.0675	75.1	22.2	Europe & Central Asia	Middle Income
130	31014427.0	6.34	6.40	0.100853	22.35833	1437.0	134.5204	57.2	89.3	Sub-Saharan Africa	Poor
131	46028476.0	1.38	1.10	7.032359	25.42379	8762.0	131.4962	68.2	12.9	Europe & Central Asia	Middle Income
132	61689620.0	1.87	0.20	8.526467	27.39249	37739.0	124.0845	79.5	5.6	Europe & Central Asia	Rich
133	304473143.0	2.07	0.60	18.545992	28.45698	50384.0	118.4777	78.2	7.7	America	Rich
134	3350832.0	2.11	0.50	2.489764	26.39123	15317.0	124.2604	76.0	13.0	America	Middle Income
135	26952719.0	2.46	0.10	4.476669	25.32054	3733.0	124.3462	68.7	49.2	Europe & Central Asia	Middle Income
136	86589342.0	1.86	0.40	1.479347	20.91630	4085.0	121.9367	75.4	26.2	East Asia & Pacific	Middle Income
137	13114579.0	5.88	13.60	0.148982	20.68321	3039.0	132.4493	52.0	94.9	Sub-Saharan Africa	Middle Income
138	13495462.0	3.85	15.10	0.654323	22.02660	1286.0	131.9745	49.0	98.3	Sub-Saharan Africa	Poor

```
In [111]: 1 from urllib.request import urlopen
          2
          3 urlopen('https://gist.githubusercontent.com/aakashns/28b2e504b3350afd9bdb157893f9725c/raw/994b65665757f4f8887db1
          4 'countries.csv')
          5 countries_df = pd.read_csv('countries.csv')
```

```
In [112]: 1 countries_df
```

Out[112]:

	location	continent	population	life_expectancy	hospital_beds_per_thousand	gdp_per_capita
0	Afghanistan	Asia	38928341.0	64.83	0.50	1803.987
1	Albania	Europe	2877800.0	78.57	2.89	11803.431
2	Algeria	Africa	43851043.0	76.88	1.90	13913.839
3	Andorra	Europe	77265.0	83.73	NaN	NaN
4	Angola	Africa	32866268.0	61.15	NaN	5819.495
...
205	Vietnam	Asia	97338583.0	75.40	2.60	6171.884
206	Western Sahara	Africa	597330.0	70.26	NaN	NaN
207	Yemen	Asia	29825968.0	66.12	0.70	1479.147
208	Zambia	Africa	18383956.0	63.89	2.00	3689.251
209	Zimbabwe	Africa	14862927.0	61.49	1.70	1899.775

210 rows × 6 columns

```
In [ ]: 1
```

