In [1]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
df=pd.read_csv('C:/Users/adiva/Downloads/gapminder.csv')
df
```

Out[1]:

	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 [2]:

```
print(df.head())
```

```
   population  fertility  HIV        CO2  BMI_male      GDP  BMI_female  lif
e  \
0  34811059.0       2.73  0.1   3.328945  24.59620  12314.0    129.9049  75.
3
1  19842251.0       6.43  2.0   1.474353  22.25083   7103.0    130.1247  58.
3
2  40381860.0       2.24  0.5   4.785170  27.50170  14646.0    118.8915  75.
5
3   2975029.0       1.40  0.1   1.804106  25.35542   7383.0    132.8108  72.
5
4  21370348.0       1.96  0.1  18.016313  27.56373  41312.0    117.3755  81.
5

   child_mortality                      Region
0             29.5  Middle East & North Africa
1            192.0           Sub-Saharan Africa
2             15.4                      America
3             20.0        Europe & Central Asia
4              5.2          East Asia & Pacific
```

In [5]:

```
print(df.shape)
```

```
(139, 10)
```

In [6]:

```
print(df.columns)
```

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

In [7]:

```
print(df.dtypes)
```

```
population         float64
fertility          float64
HIV                float64
CO2                float64
BMI_male           float64
GDP                float64
BMI_female         float64
life               float64
child_mortality    float64
Region              object
dtype: object
```

In [8]:

```
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 139 entries, 0 to 138
Data columns (total 10 columns):
population         139 non-null float64
fertility          139 non-null float64
HIV                139 non-null float64
CO2                139 non-null float64
BMI_male           139 non-null float64
GDP                139 non-null float64
BMI_female         139 non-null float64
life               139 non-null float64
child_mortality    139 non-null float64
Region             139 non-null object
dtypes: float64(9), object(1)
memory usage: 11.0+ KB
None
```

In [17]:

```
df_fertility=df.['fertility']
```

```
  File "<ipython-input-17-70f58bf4c2f9>", line 1
    df_fertility=df.['fertility']
                    ^
SyntaxError: invalid syntax
```

In [19]:

```
df_fertility=df['fertility']
```

In [20]:

```
print(df_fertility)
```

```
0      2.73
1      6.43
2      2.24
3      1.40
4      1.96
       ...
134    2.11
135    2.46
136    1.86
137    5.88
138    3.85
Name: fertility, Length: 139, dtype: float64
```

In [22]:

```
print(df_fertility.head())
```

```
0    2.73
1    6.43
2    2.24
3    1.40
4    1.96
Name: fertility, dtype: float64
```

In [23]:

```
print(df_fertility.tail())
```

```
134    2.11
135    2.46
136    1.86
137    5.88
138    3.85
Name: fertility, dtype: float64
```

In [24]:

```
subset=df[['population', 'fertility', 'HIV']]
```

In [25]:

```
print(subset.head())
```

```
   population  fertility  HIV
0  34811059.0       2.73  0.1
1  19842251.0       6.43  2.0
2  40381860.0       2.24  0.5
3   2975029.0       1.40  0.1
4  21370348.0       1.96  0.1
```

In [26]:

```
print(subset.tail())
```

```
     population  fertility   HIV
134   3350832.0       2.11   0.5
135  26952719.0       2.46   0.1
136  86589342.0       1.86   0.4
137  13114579.0       5.88  13.6
138  13495462.0       3.85  15.1
```

In [27]:

```
print(df.loc[0])
```

```
population                       3.48111e+07
fertility                               2.73
HIV                                      0.1
CO2                                  3.32894
BMI_male                             24.5962
GDP                                    12314
BMI_female                           129.905
life                                    75.3
child_mortality                         29.5
Region             Middle East & North Africa
Name: 0, dtype: object
```

In [28]:

```
print(df.loc[99])
```

```
population         6.04713e+06
fertility                 3.06
HIV                        0.3
CO2                   0.698582
BMI_male               25.5422
GDP                       6684
BMI_female             123.615
life                      73.6
child_mortality           25.7
Region                 America
Name: 99, dtype: object
```

In [30]:

```
print(df.tail(n=1))
```

```
     population  fertility   HIV       CO2  BMI_male     GDP  BMI_female  \
138  13495462.0       3.85  15.1  0.654323   22.0266  1286.0    131.9745

     life  child_mortality             Region
138  49.0             98.3  Sub-Saharan Africa
```

In [31]:

```
print(df.loc[[0,10,20]])
```

```
    population  fertility  HIV       CO2  BMI_male      GDP  BMI_female  lif
e  \
0   34811059.0       2.73  0.1  3.328945  24.59620  12314.0    129.9049  75.
3
10   9526453.0       1.42  0.2  6.488174  26.16443  14488.0    129.7968  70.
1
20   8821795.0       6.48  3.5  0.031389  21.50291    723.0    134.1955  57.
4

    child_mortality                      Region
0              29.5  Middle East & North Africa
10              7.2       Europe & Central Asia
20            108.6          Sub-Saharan Africa
```

In [34]:

```
print(df.iloc[0])
```

```
population                         3.48111e+07
fertility                                 2.73
HIV                                        0.1
CO2                                    3.32894
BMI_male                               24.5962
GDP                                      12314
BMI_female                             129.905
life                                      75.3
child_mortality                           29.5
Region             Middle East & North Africa
Name: 0, dtype: object
```

In [35]:

```
print(df.iloc[99])
```

```
population         6.04713e+06
fertility                 3.06
HIV                        0.3
CO2                   0.698582
BMI_male               25.5422
GDP                       6684
BMI_female             123.615
life                      73.6
child_mortality           25.7
Region                 America
Name: 99, dtype: object
```

In [36]:

```
print(df.iloc[-1])
```

```
population                1.34955e+07
fertility                        3.85
HIV                              15.1
CO2                          0.654323
BMI_male                      22.0266
GDP                              1286
BMI_female                    131.975
life                               49
child_mortality                  98.3
Region              Sub-Saharan Africa
Name: 138, dtype: object
```

In [39]:

```
print(df.iloc[[0,10,100]])
```

```
     population  fertility  HIV       CO2  BMI_male      GDP  BMI_female  \
0    34811059.0       2.73  0.1  3.328945  24.59620  12314.0    129.9049
10    9526453.0       1.42  0.2  6.488174  26.16443  14488.0    129.7968
100  28642048.0       2.58  0.4  1.450134  24.77041   9249.0    119.6368

     life  child_mortality                      Region
0    75.3             29.5  Middle East & North Africa
10   70.1              7.2       Europe & Central Asia
100  76.8             23.2                     America
```

In [44]:

```
subset=df.loc[:,['population','HIV','BMI_male']]
print(subset.head())
```

```
   population  HIV  BMI_male
0  34811059.0  0.1  24.59620
1  19842251.0  2.0  22.25083
2  40381860.0  0.5  27.50170
3   2975029.0  0.1  25.35542
4  21370348.0  0.1  27.56373
```

In [46]:

```
subset=df.iloc[:,[1,3,6]]
print(subset.head())
```

```
   fertility        CO2  BMI_female
0       2.73   3.328945    129.9049
1       6.43   1.474353    130.1247
2       2.24   4.785170    118.8915
3       1.40   1.804106    132.8108
4       1.96  18.016313    117.3755
```

In [50]:

```
print(df.loc[42,'BMI_male'])
```

```
26.733390000000004
```

In [51]:

```
print(df.iloc[[0,1,2,3],[0,1,2,3,4]])
```

```
   population  fertility  HIV       CO2  BMI_male
0  34811059.0       2.73  0.1  3.328945  24.59620
1  19842251.0       6.43  2.0  1.474353  22.25083
2  40381860.0       2.24  0.5  4.785170  27.50170
3   2975029.0       1.40  0.1  1.804106  25.35542
```

In [54]:

```
print(df.iloc[0:4,[0,1,2,3]])
```

```
   population  fertility  HIV       CO2
0  34811059.0       2.73  0.1  3.328945
1  19842251.0       6.43  2.0  1.474353
2  40381860.0       2.24  0.5  4.785170
3   2975029.0       1.40  0.1  1.804106
```

In [56]:

```
print(df.iloc[0:4,[0,1,2,3]])
```

```
   population  fertility  HIV       CO2
0  34811059.0       2.73  0.1  3.328945
1  19842251.0       6.43  2.0  1.474353
2  40381860.0       2.24  0.5  4.785170
3   2975029.0       1.40  0.1  1.804106
```

In [60]:

```
print(df.groupby('Region')['fertility'].mean())
```

```
Region
America                       2.460741
East Asia & Pacific           2.293571
Europe & Central Asia         1.692439
Middle East & North Africa    2.538000
South Asia                    2.672857
Sub-Saharan Africa            5.142000
Name: fertility, dtype: float64
```

In [61]:

```
flat=df.groupby('Region')['fertility'].mean().reset_index()
```

In [62]:

```
print(flat.head())
```

```
                      Region  fertility
0                    America   2.460741
1         East Asia & Pacific   2.293571
2       Europe & Central Asia   1.692439
3  Middle East & North Africa   2.538000
4                  South Asia   2.672857
```

In [63]:

```
ana=df.groupby('Region')['fertility'].mean()
```

In [64]:

```
print(ana.head())
```

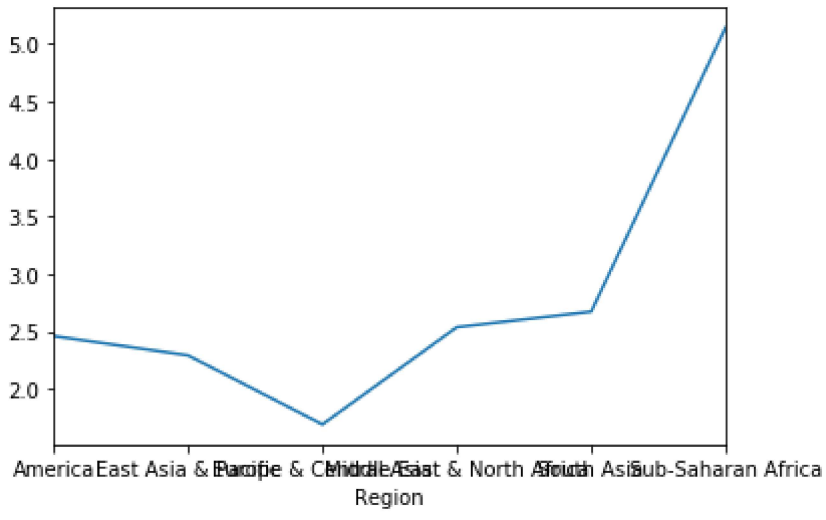
```
Region
America                        2.460741
East Asia & Pacific            2.293571
Europe & Central Asia          1.692439
Middle East & North Africa     2.538000
South Asia                     2.672857
Name: fertility, dtype: float64
```

In [65]:

```
ana.plot()
```

Out[65]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1c706407208>
```



In [ ]:

```

```