```
In [1]: import pandas as pd
```

```
In [2]: df=pd.read_csv(r"C:\Users\Admin\Downloads\data.csv")
```

```
In [3]: df
```

Out[3]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.9	High income
1	Afghanistan	AFG	35.253	5.9	Low income
2	Angola	AGO	45.985	19.1	Upper middle income
3	Albania	ALB	12.877	57.2	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0	High income
...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income
191	South Africa	ZAF	20.850	46.5	Upper middle income
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income
193	Zambia	ZMB	40.471	15.4	Lower middle income
194	Zimbabwe	ZWE	35.715	18.5	Low income

195 rows × 5 columns

```
In [4]: len(df)
```

Out[4]: 195

```
In [5]: df.shape
```

Out[5]: (195, 5)

```
In [6]: df.columns
```

Out[6]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
                'IncomeGroup'],
               dtype='object')

```
In [7]: type(df)
```

Out[7]: pandas.core.frame.DataFrame

```
In [8]: df
```

Out[8]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.9	High income
1	Afghanistan	AFG	35.253	5.9	Low income
2	Angola	AGO	45.985	19.1	Upper middle income
3	Albania	ALB	12.877	57.2	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0	High income
...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income
191	South Africa	ZAF	20.850	46.5	Upper middle income
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income
193	Zambia	ZMB	40.471	15.4	Lower middle income
194	Zimbabwe	ZWE	35.715	18.5	Low income

195 rows × 5 columns

In [9]: ```df.info```

Out[9]: <bound method DataFrame.info of   CountryName CountryCode BirthRat
         e InternetUsers \
         0   Aruba   ABW   10.244   78.9
         1   Afghanistan   AFG   35.253   5.9
         2   Angola   AGO   45.985   19.1
         3   Albania   ALB   12.877   57.2
         4   United Arab Emirates   ARE   11.044   88.0
         ..   ..
         190   Yemen, Rep.   YEM   32.947   20.0
         191   South Africa   ZAF   20.850   46.5
         192   Congo, Dem. Rep.   COD   42.394   2.2
         193   Zambia   ZMB   40.471   15.4
         194   Zimbabwe   ZWE   35.715   18.5

                      IncomeGroup
         0   High income
         1   Low income
         2   Upper middle income
         3   Upper middle income
         4   High income
         ..   ...
         190 Lower middle income
         191 Upper middle income
         192   Low income
         193 Lower middle income
         194   Low income

         [195 rows x 5 columns]>

In [10]:
```python
df_num=df[['BirthRate', 'InternetUsers']]
df_num
```

Out[10]:

	BirthRate	InternetUsers
0	10.244	78.9
1	35.253	5.9
2	45.985	19.1
3	12.877	57.2
4	11.044	88.0
...	...	...
190	32.947	20.0
191	20.850	46.5
192	42.394	2.2
193	40.471	15.4
194	35.715	18.5

195 rows × 2 columns

In [11]:
```python
df.tail()
```

Out[11]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income
191	South Africa	ZAF	20.850	46.5	Upper middle income
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income
193	Zambia	ZMB	40.471	15.4	Lower middle income
194	Zimbabwe	ZWE	35.715	18.5	Low income

In [12]: `df.head()`

Out[12]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.9	High income
1	Afghanistan	AFG	35.253	5.9	Low income
2	Angola	AGO	45.985	19.1	Upper middle income
3	Albania	ALB	12.877	57.2	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0	High income

In [13]: `df.tail(2)`

Out[13]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
193	Zambia	ZMB	40.471	15.4	Lower middle income
194	Zimbabwe	ZWE	35.715	18.5	Low income

In [14]: `df.head(2)`

Out[14]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.9	High income
1	Afghanistan	AFG	35.253	5.9	Low income

In [15]: `df[::-1]`

Out[15]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
194	Zimbabwe	ZWE	35.715	18.5	Low income
193	Zambia	ZMB	40.471	15.4	Lower middle income
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income
191	South Africa	ZAF	20.850	46.5	Upper middle income
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income
...	...	...	...	...	...
4	United Arab Emirates	ARE	11.044	88.0	High income
3	Albania	ALB	12.877	57.2	Upper middle income
2	Angola	AGO	45.985	19.1	Upper middle income
1	Afghanistan	AFG	35.253	5.9	Low income
0	Aruba	ABW	10.244	78.9	High income

195 rows × 5 columns

In [16]:
```python
df[6:]
```

Out[16]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
6	Armenia	ARM	13.308	41.9000	Lower middle income
7	Antigua and Barbuda	ATG	16.447	63.4000	High income
8	Australia	AUS	13.200	83.0000	High income
9	Austria	AUT	9.400	80.6188	High income
10	Azerbaijan	AZE	18.300	58.7000	Upper middle income
...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0000	Lower middle income
191	South Africa	ZAF	20.850	46.5000	Upper middle income
192	Congo, Dem. Rep.	COD	42.394	2.2000	Low income
193	Zambia	ZMB	40.471	15.4000	Lower middle income
194	Zimbabwe	ZWE	35.715	18.5000	Low income

189 rows × 5 columns

In [17]:
```python
df[0:200:10]
```

Out[17]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.900000	High income
10	Azerbaijan	AZE	18.300	58.700000	Upper middle income
20	Belarus	BLR	12.500	54.170000	Upper middle income
30	Canada	CAN	10.900	85.800000	High income
40	Costa Rica	CRI	15.022	45.960000	Upper middle income
50	Ecuador	ECU	21.070	40.353684	Upper middle income
60	Gabon	GAB	30.555	9.200000	Upper middle income
70	Greenland	GRL	14.500	65.800000	High income
80	India	IND	20.291	15.100000	Lower middle income
90	Kazakhstan	KAZ	22.730	54.000000	Upper middle income
100	Libya	LBY	21.425	16.500000	Upper middle income
110	Moldova	MDA	12.141	45.000000	Lower middle income
120	Mozambique	MOZ	39.705	5.400000	Low income
130	Netherlands	NLD	10.200	93.956400	High income
140	Poland	POL	9.600	62.849200	High income
150	Sudan	SDN	33.477	22.700000	Lower middle income
160	Suriname	SUR	18.455	37.400000	Upper middle income
170	Tajikistan	TJK	30.792	16.000000	Lower middle income
180	Uruguay	URY	14.374	57.690000	High income
190	Yemen, Rep.	YEM	32.947	20.000000	Lower middle income

In [18]:
```python
df.describe() #describes the statistical data of the data set
             #this function always print numerical data
```

Out[18]:

	BirthRate	InternetUsers
count	195.000000	195.000000
mean	21.469928	42.076471
std	10.605467	29.030788
min	7.900000	0.900000
25%	12.120500	14.520000
50%	19.680000	41.000000
75%	29.759500	66.225000
max	49.661000	96.546800

In [19]: `df.describe().transpose() #this function will convert rows to columns`

Out[19]:

	count	mean	std	min	25%	50%	75%	max
BirthRate	195.0	21.469928	10.605467	7.9	12.1205	19.68	29.7595	49.6610
InternetUsers	195.0	42.076471	29.030788	0.9	14.5200	41.00	66.2250	96.5468

In [20]: `df.columns`

Out[20]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',`
         `'IncomeGroup'],`
         `dtype='object')`

In [21]: `df.columns=['a','b','c','d','e'] #as list is mutable we  can change the columns`
         `#renaming the atributes`

In [22]: `df.columns`

Out[22]: `Index(['a', 'b', 'c', 'd', 'e'], dtype='object')`

In [23]: `df.head()`

Out[23]:

	a	b	c	d	e
0	Aruba	ABW	10.244	78.9	High income
1	Afghanistan	AFG	35.253	5.9	Low income
2	Angola	AGO	45.985	19.1	Upper middle income
3	Albania	ALB	12.877	57.2	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0	High income

In [24]: `df.columns=['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',`
         `'IncomeGroup']`

In [25]: `df.columns`

Out[25]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',`
         `'IncomeGroup'],`
         `dtype='object')`

In [26]: `df.head()`

Out[26]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.9	High income
1	Afghanistan	AFG	35.253	5.9	Low income
2	Angola	AGO	45.985	19.1	Upper middle income
3	Albania	ALB	12.877	57.2	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0	High income

In [27]: `df.columns`

Out[27]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

In [28]: `df[['CountryName', 'CountryCode', 'BirthRate']]`

Out[28]:

	CountryName	CountryCode	BirthRate
0	Aruba	ABW	10.244
1	Afghanistan	AFG	35.253
2	Angola	AGO	45.985
3	Albania	ALB	12.877
4	United Arab Emirates	ARE	11.044
...	...	...	...
190	Yemen, Rep.	YEM	32.947
191	South Africa	ZAF	20.850
192	Congo, Dem. Rep.	COD	42.394
193	Zambia	ZMB	40.471
194	Zimbabwe	ZWE	35.715

195 rows × 3 columns

In [29]: `df.isnull`

```
Out[29]: <bound method DataFrame.isnull of                    CountryName CountryCode  BirthR
         ate  InternetUsers  \
         0                  Aruba        ABW      10.244            78.9
         1            Afghanistan        AFG      35.253             5.9
         2                 Angola        AGO      45.985            19.1
         3                Albania        ALB      12.877            57.2
         4     United Arab Emirates        ARE      11.044            88.0
         ..                   ...        ...         ...             ...
         190          Yemen, Rep.        YEM      32.947            20.0
         191         South Africa        ZAF      20.850            46.5
         192     Congo, Dem. Rep.        COD      42.394             2.2
         193               Zambia        ZMB      40.471            15.4
         194             Zimbabwe        ZWE      35.715            18.5

                  IncomeGroup
         0        High income
         1         Low income
         2      Upper middle income
         3      Upper middle income
         4        High income
         ..               ...
         190  Lower middle income
         191  Upper middle income
         192          Low income
         193  Lower middle income
         194          Low income

         [195 rows x 5 columns]>
```

```python
In [30]: df_cate=df[['CountryName', 'CountryCode']]
         df_cate.head()
```

Out[30]:

	CountryName	CountryCode
0	Aruba	ABW
1	Afghanistan	AFG
2	Angola	AGO
3	Albania	ALB
4	United Arab Emirates	ARE

```python
In [31]: df.describe()
```

Out[31]:

	BirthRate	InternetUsers
count	195.000000	195.000000
mean	21.469928	42.076471
std	10.605467	29.030788
min	7.900000	0.900000
25%	12.120500	14.520000
50%	19.680000	41.000000
75%	29.759500	66.225000
max	49.661000	96.546800

In [32]: ```df_cate.describe()```

Out[32]:

	CountryName	CountryCode
count	195	195
unique	195	195
top	Aruba	ABW
freq	1	1

In [33]: ```df_num=df[['BirthRate', 'InternetUsers']]```

In [34]: ```df_num```

Out[34]:

	BirthRate	InternetUsers
0	10.244	78.9
1	35.253	5.9
2	45.985	19.1
3	12.877	57.2
4	11.044	88.0
...	...	...
190	32.947	20.0
191	20.850	46.5
192	42.394	2.2
193	40.471	15.4
194	35.715	18.5

195 rows × 2 columns

In [35]: ```df_num.transpose```

```
Out[35]:  <bound method DataFrame.transpose of        BirthRate   InternetUsers
          0         10.244           78.9
          1         35.253            5.9
          2         45.985           19.1
          3         12.877           57.2
          4         11.044           88.0
          ..          ...            ...
          190       32.947           20.0
          191       20.850           46.5
          192       42.394            2.2
          193       40.471           15.4
          194       35.715           18.5

          [195 rows x 2 columns]>
```

In [36]: `df.head`

```
Out[36]:  <bound method NDFrame.head of            CountryName CountryCode   BirthRate
          InternetUsers  \
          0                    Aruba      ABW      10.244           78.9
          1               Afghanistan    AFG      35.253            5.9
          2                   Angola     AGO      45.985           19.1
          3                   Albania    ALB      12.877           57.2
          4       United Arab Emirates   ARE      11.044           88.0
          ..                     ...     ...        ...            ...
          190            Yemen, Rep.    YEM      32.947           20.0
          191            South Africa   ZAF      20.850           46.5
          192        Congo, Dem. Rep.  COD      42.394            2.2
          193                Zambia     ZMB      40.471           15.4
          194               Zimbabwe    ZWE      35.715           18.5

                        IncomeGroup
          0             High income
          1              Low income
          2       Upper middle income
          3       Upper middle income
          4             High income
          ..                    ...
          190   Lower middle income
          191   Upper middle income
          192            Low income
          193   Lower middle income
          194            Low income

          [195 rows x 5 columns]>
```

In [37]: `df.BirthRate*df.InternetUsers`

```
Out[37]:  0        808.2516
          1        207.9927
          2        878.3135
          3        736.5644
          4        971.8720
                    ...
          190      658.9400
          191      969.5250
          192       93.2668
          193      623.2534
          194      660.7275
          Length: 195, dtype: float64
```

```
In [38]:  df['mycalc']=df.BirthRate*df.InternetUsers #adding a column
```

```
In [39]:  df
```

Out[39]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	mycalc
0	Aruba	ABW	10.244	78.9	High income	808.2516
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135
3	Albania	ALB	12.877	57.2	Upper middle income	736.5644
4	United Arab Emirates	ARE	11.044	88.0	High income	971.8720
...	...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income	658.9400
191	South Africa	ZAF	20.850	46.5	Upper middle income	969.5250
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income	93.2668
193	Zambia	ZMB	40.471	15.4	Lower middle income	623.2534
194	Zimbabwe	ZWE	35.715	18.5	Low income	660.7275

195 rows × 6 columns

```
In [40]:  df.columns
```

```
Out[40]:  Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
                 'IncomeGroup', 'mycalc'],
                dtype='object')
```

```
In [41]:  len(df.columns)
```

```
Out[41]:  6
```

In [42]: `df.drop('mycalc',axis=1)`

Out[42]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.9	High income
1	Afghanistan	AFG	35.253	5.9	Low income
2	Angola	AGO	45.985	19.1	Upper middle income
3	Albania	ALB	12.877	57.2	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0	High income
...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income
191	South Africa	ZAF	20.850	46.5	Upper middle income
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income
193	Zambia	ZMB	40.471	15.4	Lower middle income
194	Zimbabwe	ZWE	35.715	18.5	Low income

195 rows × 5 columns

In [43]: `df['InternetUsers']`

Out[43]:
```
0      78.9
1       5.9
2      19.1
3      57.2
4      88.0
       ...
190    20.0
191    46.5
192     2.2
193    15.4
194    18.5
Name: InternetUsers, Length: 195, dtype: float64
```

In [44]: `df['InternetUsers']<2`

Out[44]: 0   False
          1   False
          2   False
          3   False
          4   False
                   ...
          190   False
          191   False
          192   False
          193   False
          194   False
          Name: InternetUsers, Length: 195, dtype: bool

In [45]: ```
          filter=df['InternetUsers']<2
          filter
          ```

Out[45]:  0        False
          1        False
          2        False
          3        False
          4        False
                   ...
          190      False
          191      False
          192      False
          193      False
          194      False
          Name: InternetUsers, Length: 195, dtype: bool

In [46]:  ```
          df[filter] #filtering the data
          ```

Out[46]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	mycalc
11	Burundi	BDI	44.151	1.3	Low income	57.3963
52	Eritrea	ERI	34.800	0.9	Low income	31.3200
55	Ethiopia	ETH	32.925	1.9	Low income	62.5575
64	Guinea	GIN	37.337	1.6	Low income	59.7392
117	Myanmar	MMR	18.119	1.6	Lower middle income	28.9904
127	Niger	NER	49.661	1.7	Low income	84.4237
154	Sierra Leone	SLE	36.729	1.7	Low income	62.4393
156	Somalia	SOM	43.891	1.5	Low income	65.8365
172	Timor-Leste	TLS	35.755	1.1	Lower middle income	39.3305

In [47]: ```
          len(df[filter])
          ```

Out[47]:  9

In [48]:  ```
          df
          ```

Out[48]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	mycalc
0	Aruba	ABW	10.244	78.9	High income	808.2516
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135
3	Albania	ALB	12.877	57.2	Upper middle income	736.5644
4	United Arab Emirates	ARE	11.044	88.0	High income	971.8720
...	...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income	658.9400
191	South Africa	ZAF	20.850	46.5	Upper middle income	969.5250
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income	93.2668
193	Zambia	ZMB	40.471	15.4	Lower middle income	623.2534
194	Zimbabwe	ZWE	35.715	18.5	Low income	660.7275

195 rows × 6 columns

In [49]: `df[filter]`

Out[49]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	mycalc
11	Burundi	BDI	44.151	1.3	Low income	57.3963
52	Eritrea	ERI	34.800	0.9	Low income	31.3200
55	Ethiopia	ETH	32.925	1.9	Low income	62.5575
64	Guinea	GIN	37.337	1.6	Low income	59.7392
117	Myanmar	MMR	18.119	1.6	Lower middle income	28.9904
127	Niger	NER	49.661	1.7	Low income	84.4237
154	Sierra Leone	SLE	36.729	1.7	Low income	62.4393
156	Somalia	SOM	43.891	1.5	Low income	65.8365
172	Timor-Leste	TLS	35.755	1.1	Lower middle income	39.3305

In [50]: `df`

Out[50]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	mycalc
0	Aruba	ABW	10.244	78.9	High income	808.2516
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135
3	Albania	ALB	12.877	57.2	Upper middle income	736.5644
4	United Arab Emirates	ARE	11.044	88.0	High income	971.8720
...	...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income	658.9400
191	South Africa	ZAF	20.850	46.5	Upper middle income	969.5250
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income	93.2668
193	Zambia	ZMB	40.471	15.4	Lower middle income	623.2534
194	Zimbabwe	ZWE	35.715	18.5	Low income	660.7275

195 rows × 6 columns

```python
In [51]:  filter1=df['BirthRate']>40
```

```python
In [52]:  df.IncomeGroup.unique()
```

Out[52]: array(['High income', 'Low income', 'Upper middle income',
          'Lower middle income'], dtype=object)

```python
In [53]:  df.IncomeGroup.nunique()
```

Out[53]: 4

```python
In [54]:  df.CountryName.nunique() #we have taken data of 195 countries to analze the data
```

Out[54]: 195

```python
In [55]:  df.columns
```

Out[55]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
          'IncomeGroup', 'mycalc'],
          dtype='object')

In [56]:
```python
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
plt.rcParams['figure.figsize']=10,2
import warnings
warnings.filterwarnings('ignore') #ignoring os errors
```
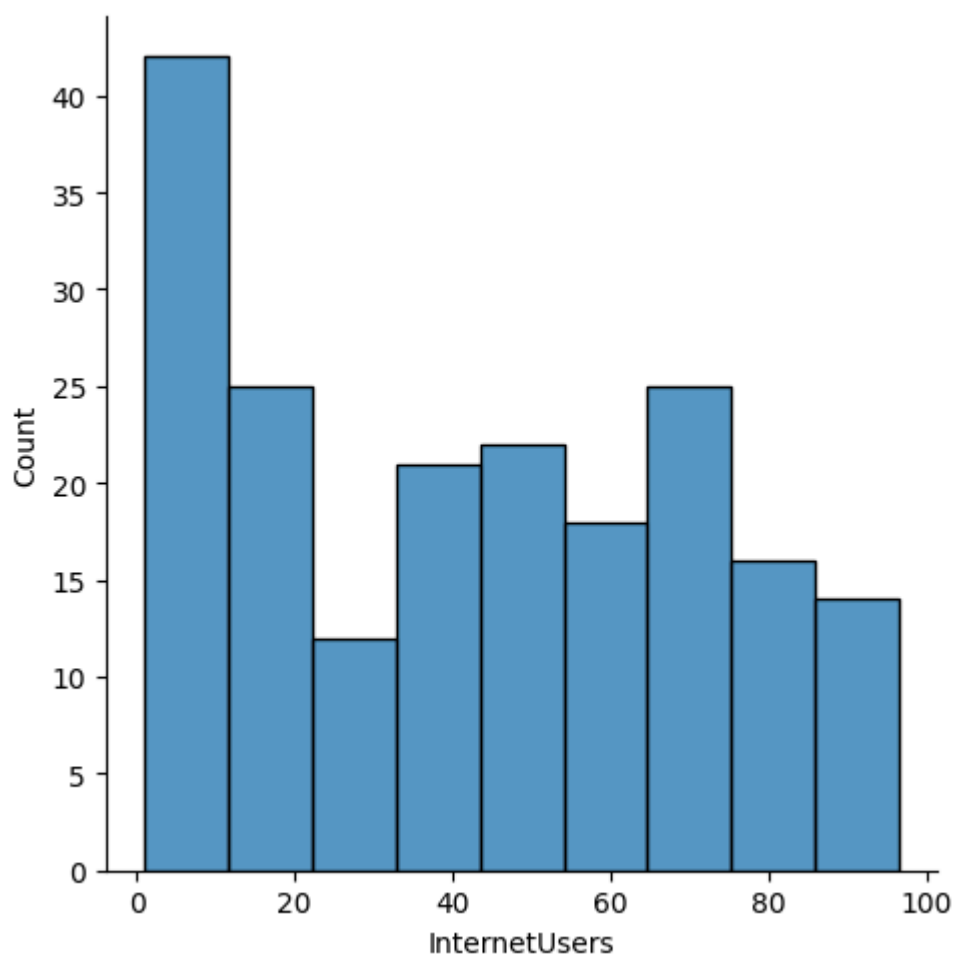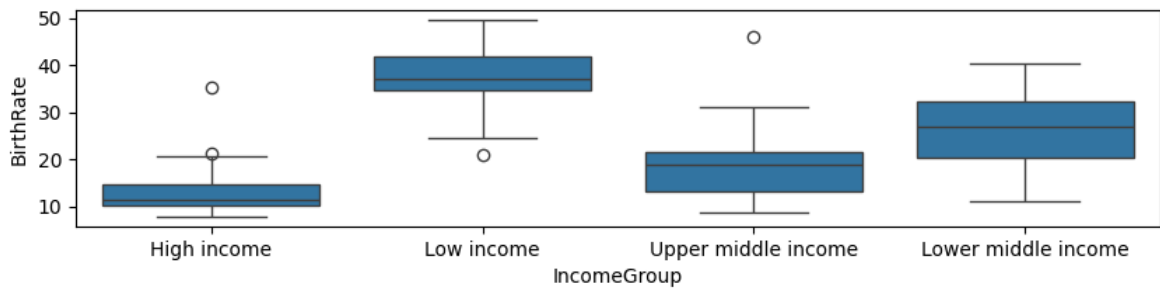
In [57]:
```python
df.columns
```

Out[57]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'mycalc'],
      dtype='object')
```

In [58]:
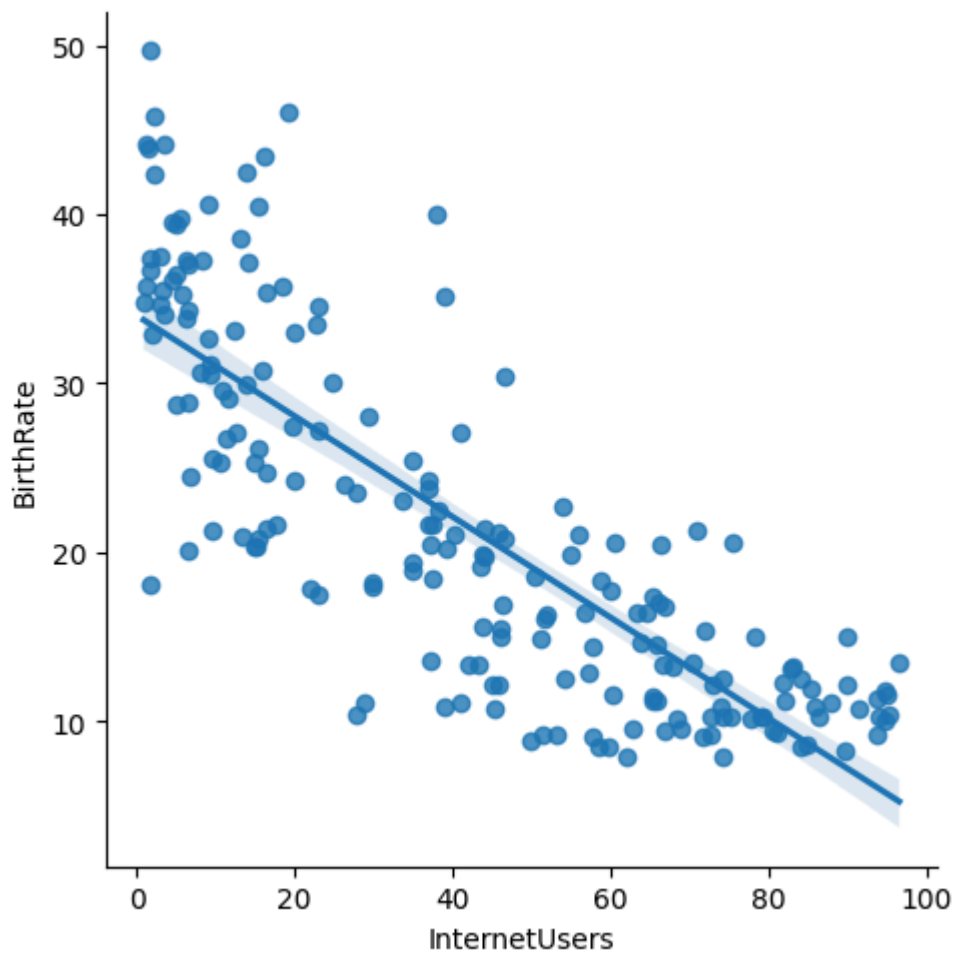```python
vis1=sns.distplot(df['InternetUsers'])
plt.show()
```
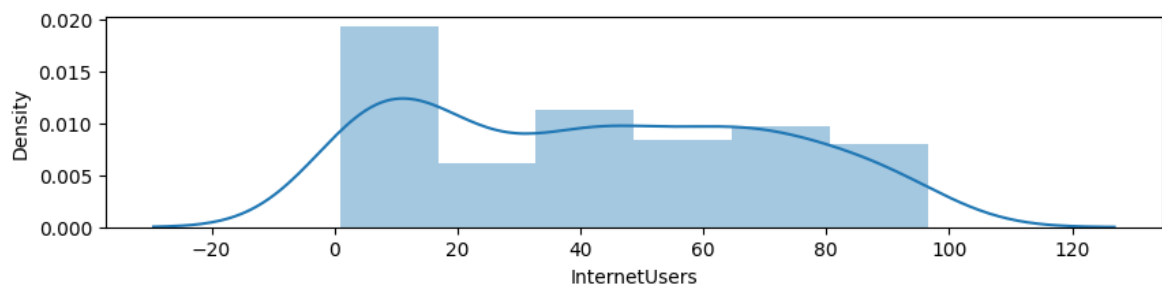


In [59]:
```python
vis2=sns.displot(df['InternetUsers'])
plt.show()
```
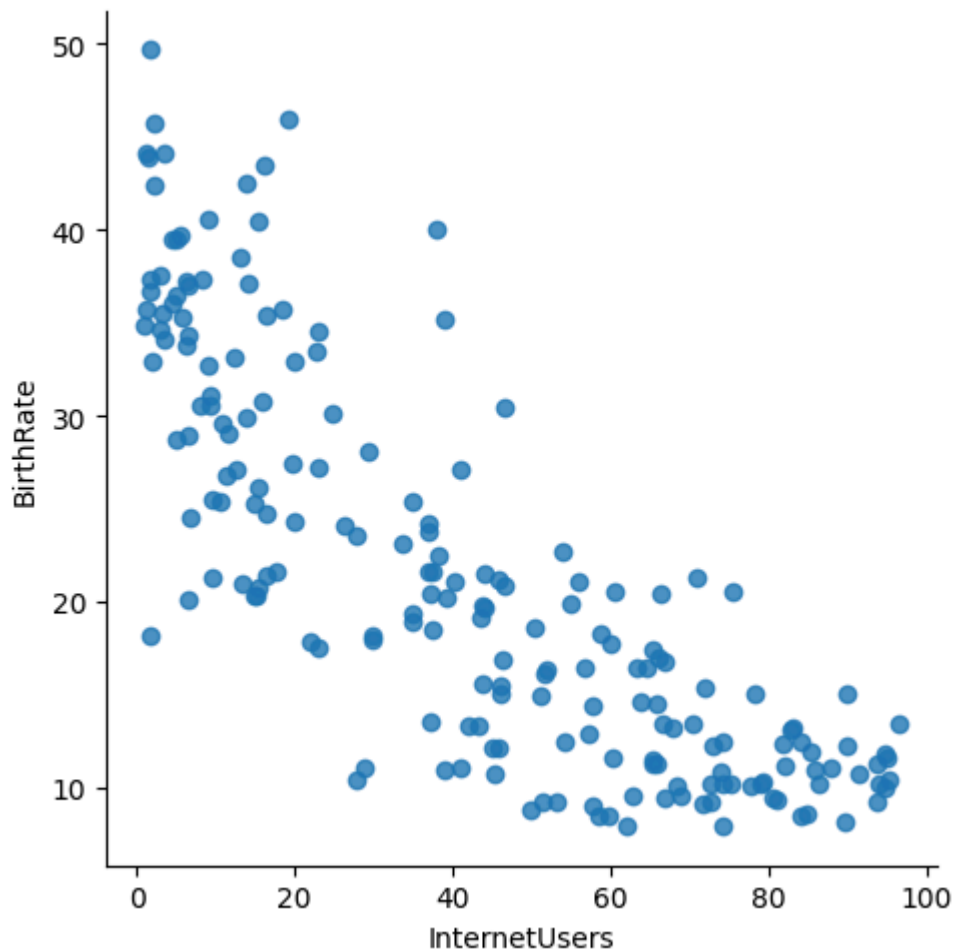
In [60]:
```python
vis3=sns.boxplot(data=df,x='IncomeGroup',y='BirthRate')
plt.show()
```


In [61]:
```python
vis4=sns.lmplot(data=df,x='InternetUsers',y='BirthRate')
plt.show()
```


In [62]:
```python
vis5=sns.lmplot(data=df,x='InternetUsers',y='BirthRate',fit_reg=False)
plt.show()
```

```
In [63]:  vis6=sns.lmplot(data=df,x='InternetUsers',y='BirthRate',hue='IncomeGroup',fit_re
          plt.show()
```