```
import pandas as pd
```

```
pd.__version__
```

```
'2.2.3'
```

```
df = pd.read_csv(r"C:\Users\ttwrd\OneDrive\Attachments\Desktop\
data.csv")
```

```
df
```

```
              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]
```

```
id(df)
```

```
2855610182352
```

```
len(df)
```

```
195
```

```
df.columns
```

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

```
len(df.columns)
```

```
df.isnull()
```

```
     CountryName  CountryCode  BirthRate  InternetUsers  IncomeGroup
0          False        False      False          False        False
1          False        False      False          False        False
2          False        False      False          False        False
3          False        False      False          False        False
4          False        False      False          False        False
..           ...          ...        ...            ...          ...
190        False        False      False          False        False
191        False        False      False          False        False
192        False        False      False          False        False
193        False        False      False          False        False
194        False        False      False          False        False

[195 rows x 5 columns]
```

```
df.isna()
```

```
     CountryName  CountryCode  BirthRate  InternetUsers  IncomeGroup
0          False        False      False          False        False
1          False        False      False          False        False
2          False        False      False          False        False
3          False        False      False          False        False
4          False        False      False          False        False
..           ...          ...        ...            ...          ...
190        False        False      False          False        False
191        False        False      False          False        False
192        False        False      False          False        False
193        False        False      False          False        False
194        False        False      False          False        False

[195 rows x 5 columns]
```

```
df.isnull().sum()
```

```
CountryName      0
CountryCode      0
BirthRate        0
InternetUsers    0
IncomeGroup      0
dtype: int64
```

```
df.isna().sum()
```

```
df.head()
```

```
df.tail()
```

```
df.info()
```

```
df[:]
```

```
df[1:11]
```

```
df[::-1]
```

```
df[1:100:10]
```

```
df[10:21]
```

```
df
```

```
df.head(2)
```

```
df.describe()
```

```
        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
```

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

```
df['CountryName']
```

```
0                     Aruba
1               Afghanistan
2                    Angola
3                   Albania
4      United Arab Emirates
               ...
190             Yemen, Rep.
191            South Africa
192         Congo, Dem. Rep.
193                   Zambia
194                 Zimbabwe
Name: CountryName, Length: 195, dtype: object
```

```
df['CountryCode']
```

```
df[['CountryName','CountryCode','IncomeGroup']]
```

```
              CountryName CountryCode          IncomeGroup
0                   Aruba         ABW          High income
1             Afghanistan         AFG           Low income
2                  Angola         AGO  Upper middle income
3                 Albania         ALB  Upper middle income
```

```
4    United Arab Emirates          ARE           High income
..                    ...          ...                   ...
190           Yemen, Rep.          YEM  Lower middle income
191          South Africa          ZAF  Upper middle income
192      Congo, Dem. Rep.          COD           Low income
193                Zambia          ZMB  Lower middle income
194              Zimbabwe          ZWE           Low income

[195 rows x 3 columns]
```

```
df_cat = df[['CountryName','CountryCode','IncomeGroup']]
df_cat
```

```
              CountryName CountryCode          IncomeGroup
0                   Aruba          ABW          High income
1             Afghanistan          AFG           Low income
2                  Angola          AGO  Upper middle income
3                 Albania          ALB  Upper middle income
4    United Arab Emirates          ARE          High income
..                    ...          ...                   ...
190           Yemen, Rep.          YEM  Lower middle income
191          South Africa          ZAF  Upper middle income
192      Congo, Dem. Rep.          COD           Low income
193                Zambia          ZMB  Lower middle income
194              Zimbabwe          ZWE           Low income

[195 rows x 3 columns]
```

```
print(len(df.columns))
print(len(df_cat.columns))
```

```
5
3
```

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

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

```
print((df_cat.columns))
```

```
Index(['CountryName', 'CountryCode', 'IncomeGroup'], dtype='object')
```

```
df_cat.describe()
```

```
       CountryName CountryCode  IncomeGroup
count          195         195          195
unique         195         195            4
top          Aruba         ABW  High income
freq             1           1           67
```

```
df_num = df[['BirthRate', 'InternetUsers']]
df_num
```

```
     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]
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 195 entries, 0 to 194
Data columns (total 5 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   CountryName    195 non-null    object 
 1   CountryCode    195 non-null    object 
 2   BirthRate      195 non-null    float64
 3   InternetUsers  195 non-null    float64
 4   IncomeGroup    195 non-null    object 
dtypes: float64(2), object(3)
memory usage: 7.7+ KB
```

```
df_cat.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 195 entries, 0 to 194
Data columns (total 3 columns):
 #   Column       Non-Null Count  Dtype 
---  ------       --------------  ----- 
 0   CountryName  195 non-null    object
 1   CountryCode  195 non-null    object
 2   IncomeGroup  195 non-null    object
dtypes: object(3)
memory usage: 4.7+ KB
```

```
df_num.info()
```

```
df.describe()
```

```
        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
```

```
df.describe().transpose()
```

```
                count       mean        std  min      25%    50%  \
BirthRate       195.0  21.469928  10.605467  7.9  12.1205  19.68
29.7595
InternetUsers   195.0  42.076471  29.030788  0.9  14.5200  41.00
66.2250

                    max
BirthRate       49.6610
InternetUsers   96.5468
```

```
df.describe().T
```

```
                count       mean        std  min      25%    50%  \
BirthRate       195.0  21.469928  10.605467  7.9  12.1205  19.68
29.7595
InternetUsers   195.0  42.076471  29.030788  0.9  14.5200  41.00
66.2250

                    max
BirthRate       49.6610
InternetUsers   96.5468
```

```
df.columns
```

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

```
df.columns = ['a','b','c','d','e']
```

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

```
       a    b       c     d            e
0  Aruba  ABW  10.244  78.9  High income
```

```
df.columns = ['CountryName', 'CountryCode', 'BirthRate',
'InternetUsers', 'IncomeGroup']
df.head(1)
```

```
  CountryName CountryCode  BirthRate  InternetUsers  IncomeGroup
0       Aruba         ABW     10.244           78.9  High income
```

```
df[['CountryCode','BirthRate','InternetUsers']][4:8] #subset
```

```
  CountryCode  BirthRate  InternetUsers
4         ARE     11.044           88.0
5         ARG     17.716           59.9
6         ARM     13.308           41.9
7         ATG     16.447           63.4
```

```
df[4:8][['CountryCode','BirthRate','InternetUsers']]
```

```
  CountryCode  BirthRate  InternetUsers
4         ARE     11.044           88.0
5         ARG     17.716           59.9
6         ARM     13.308           41.9
7         ATG     16.447           63.4
```

```
df.columns
```

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

```
df.BirthRate * df.InternetUsers
```

```
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
```

```
df.head(2)
```

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

```
df['newcolumn'] = df.BirthRate * df.InternetUsers
```

```
df.head(5)
```

```
            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
```

```
           IncomeGroup  newcolumn
0          High income   808.2516
1           Low income   207.9927
2  Upper middle income   878.3135
3  Upper middle income   736.5644
4          High income   971.8720
```

```
len(df.columns)
```

```
6
```

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

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

```
  CountryName CountryCode  BirthRate  InternetUsers  IncomeGroup
0       Aruba         ABW     10.244           78.9  High income
```

```
df
```

```
               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]
```

```
df.InternetUsers<2
```

```
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
```

```
df[df.InternetUsers<2]
```

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

```
len(df[df.InternetUsers<2])
```

```
9
```

```
df.BirthRate>40
```

```
0      False
1      False
2       True
3      False
```

```
4      False
       ...
190    False
191    False
192     True
193     True
194    False
Name: BirthRate, Length: 195, dtype: bool
```

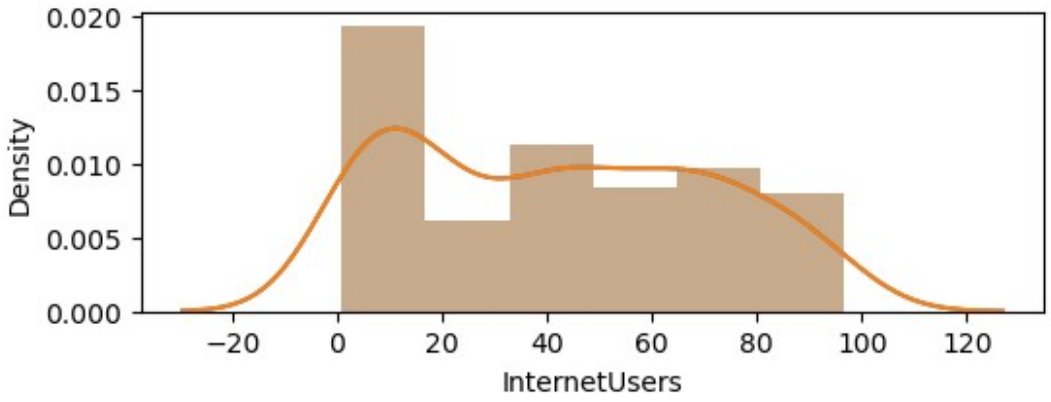
```
df[df.BirthRate>40]
```

```
          CountryName CountryCode  BirthRate  InternetUsers  \
2              Angola         AGO     45.985           19.1
11            Burundi         BDI     44.151            1.3
14       Burkina Faso         BFA     40.551            9.1
65        Gambia, The         GMB     42.525           14.0
115             Mali         MLI     44.138            3.5
127            Niger         NER     49.661            1.7
128          Nigeria         NGA     40.045           38.0
156          Somalia         SOM     43.891            1.5
167             Chad         TCD     45.745            2.3
178           Uganda         UGA     43.474           16.2
192  Congo, Dem. Rep.         COD     42.394            2.2
193           Zambia         ZMB     40.471           15.4

             IncomeGroup
2    Upper middle income
11            Low income
14            Low income
65            Low income
115           Low income
127           Low income
128  Lower middle income
156           Low income
167           Low income
178           Low income
192           Low income
193  Lower middle income
```

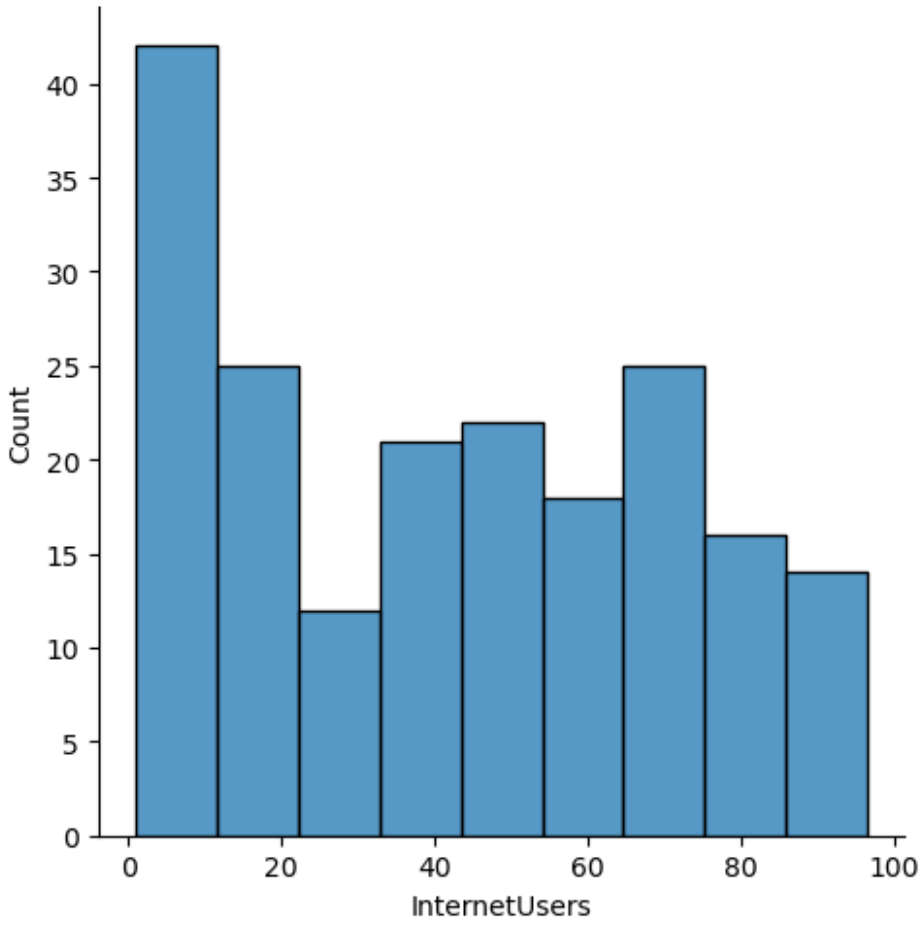
```
Filter = df.InternetUsers < 2

Filter2 = df.BirthRate >40

df[Filter & Filter2]
```

```
    CountryName CountryCode  BirthRate  InternetUsers IncomeGroup
11      Burundi         BDI     44.151            1.3  Low income
127       Niger         NER     49.661            1.7  Low income
156     Somalia         SOM     43.891            1.5  Low income
```

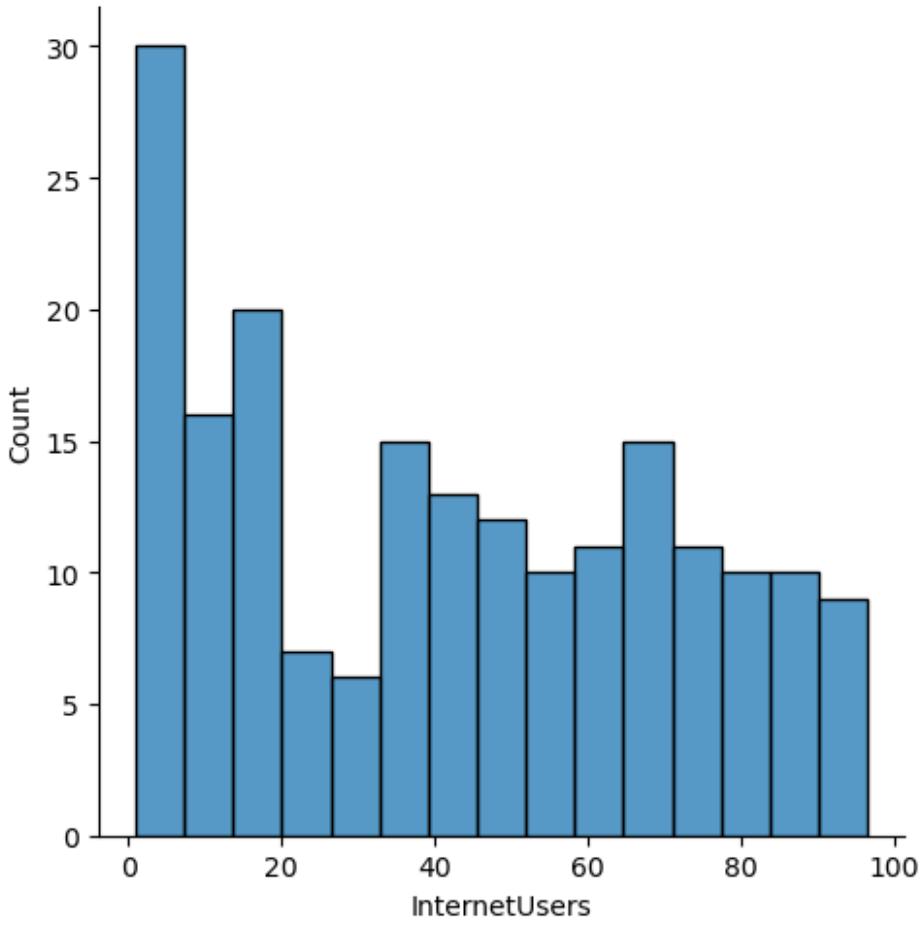
```
import matplotlib.pyplot as plt
import seaborn as sns

%matplotlib  inline
plt.rcParams['figure.figsize']=6,2
import warnings
warnings.filterwarnings('ignore')

vis1=sns.distplot(df['InternetUsers'])

plt.show(vis1)
```



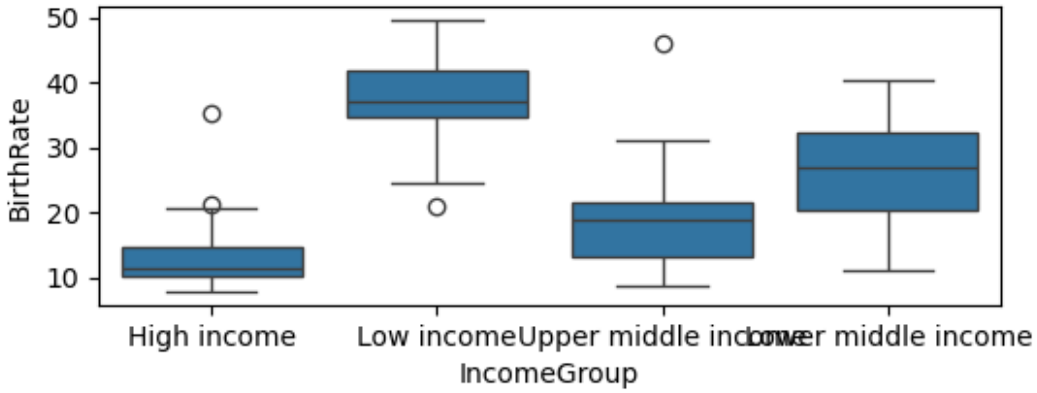

```
vis2=sns.displot(df['InternetUsers'])

plt.show(vis2)
```



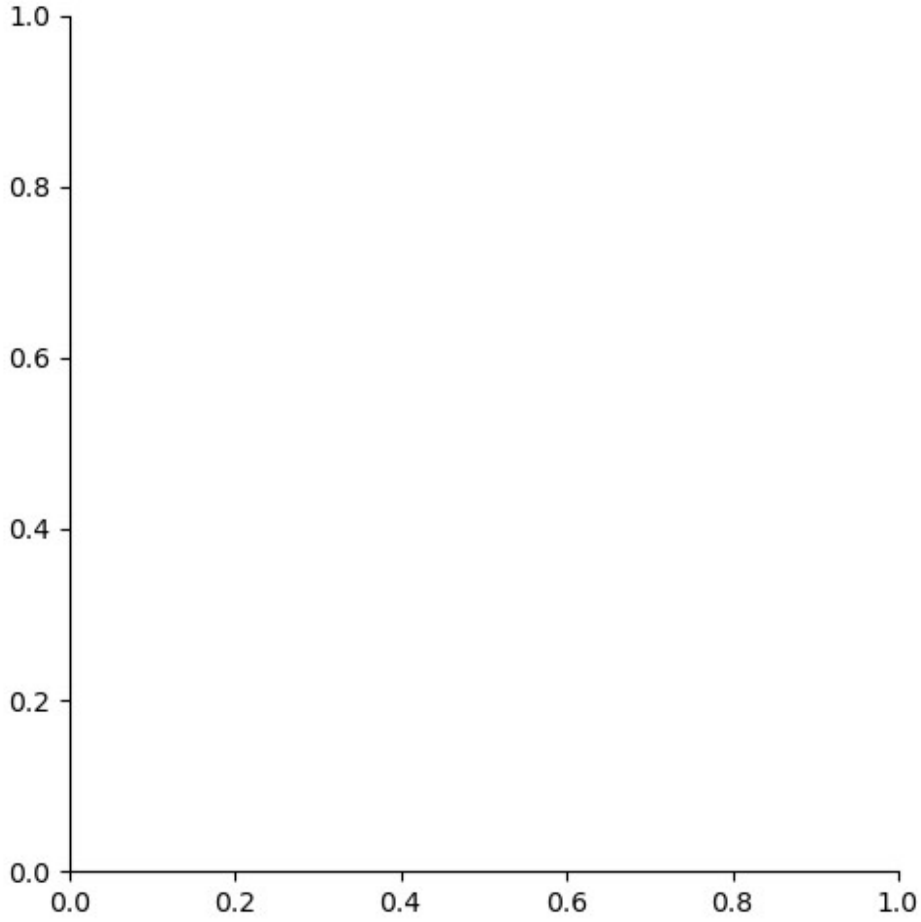

```
vis3=sns.displot(df['InternetUsers'],bins=15)
plt.show(vis3)
```

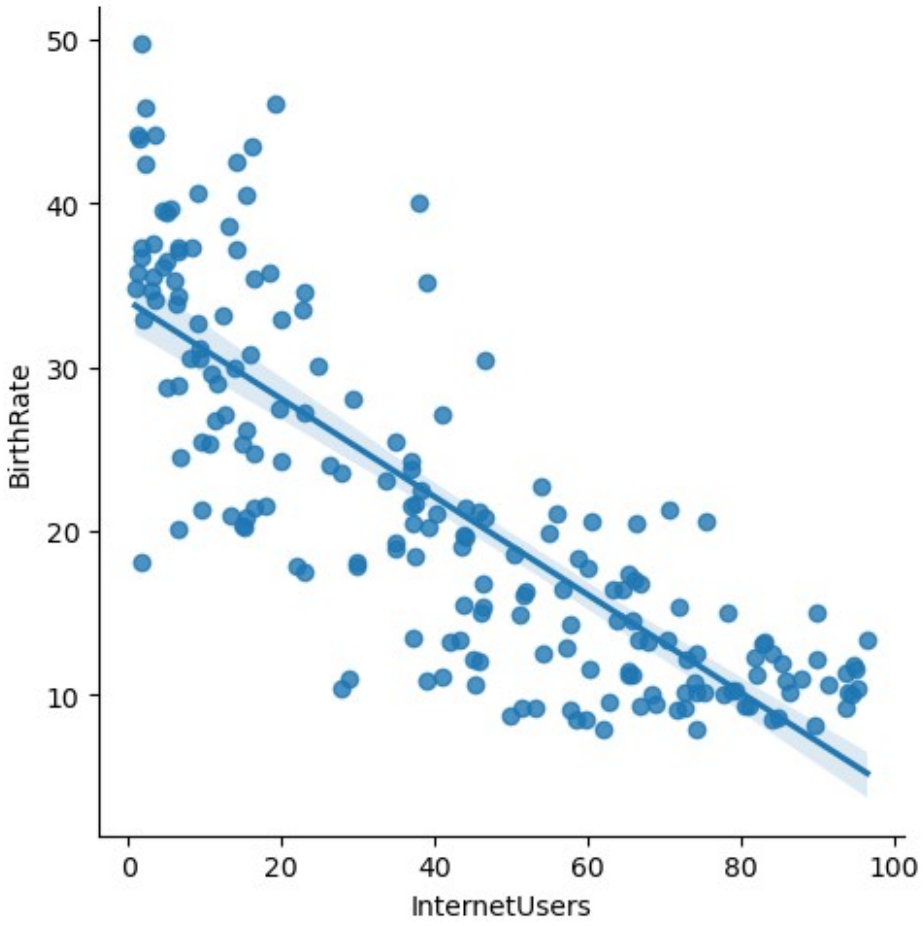


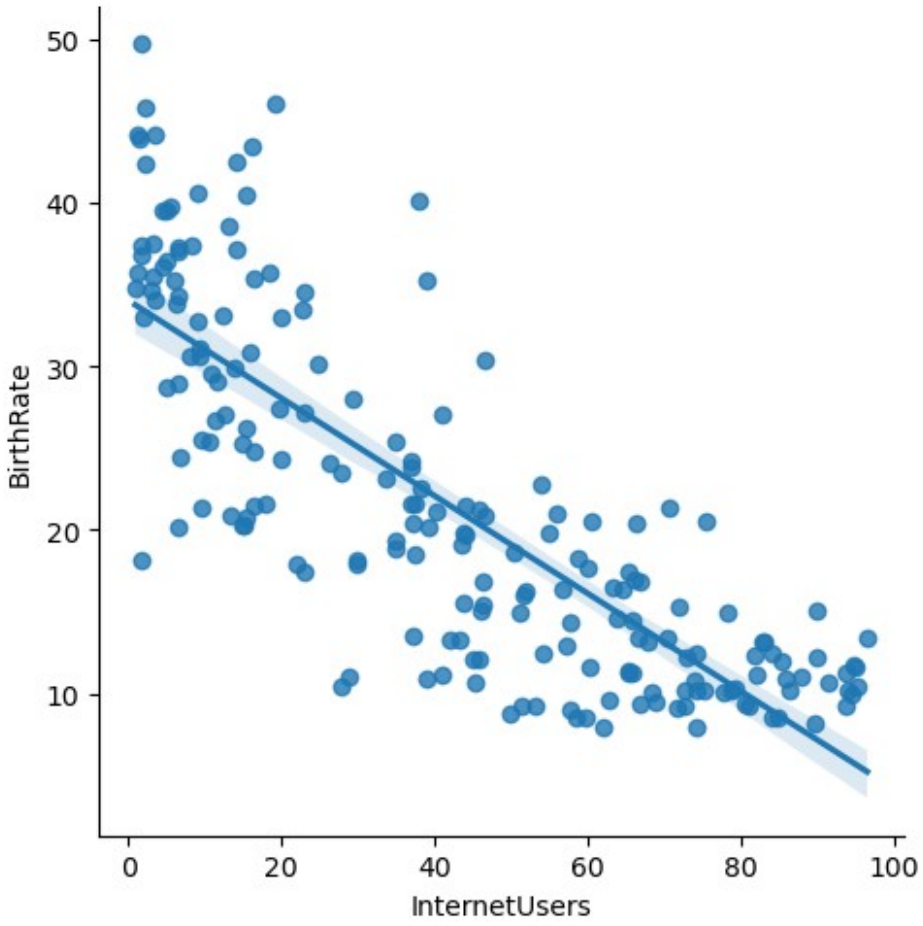

```
vis3=sns.boxplot(data=df,x='IncomeGroup',y='BirthRate')
plt.show(vis3)
```



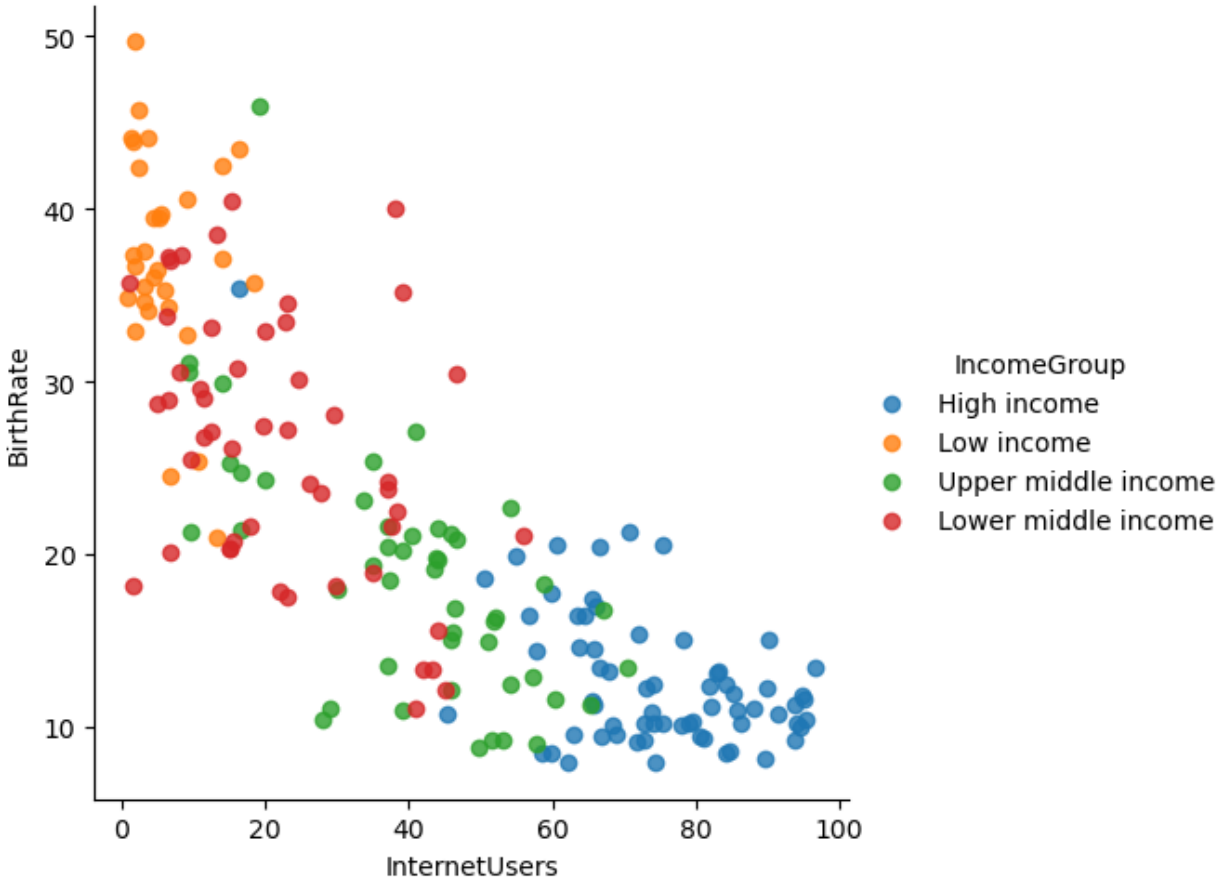

```
vis4=sns.lmplot(data=df,x='InternetUsers',y='BirthRate')
plt.show(vis4)
```


1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


50
40
30
20
10
BirthRate
0
20
40
60
80
100
InternetUsers




```
vis4=sns.lmplot(data=df,x='InternetUsers',y='BirthRate',fit_reg=False,
hue='IncomeGroup')
plt.show(vis4)
```


50
40
30
20
10
BirthRate
0
20
40
60
80
100
InternetUsers
IncomeGroup
High income
Low income
Upper middle income
Lower middle income