

Out[98]:

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

In [108...]

```
df_num = df[['BirthRate', 'InternetUsers']]  
df_num.head()
```

Out[108...]

	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

New code

In [219...]

```
df[:]
```

Out[219...]

	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 []:

In [36]: df.columns = ['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup']

In [38]: df.head(1)

Out[38]:

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

In [40]: df[:]

Out[40]:

	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 [42]:

df[0:5]

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

In [44]:

df.columns

Out[44]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup'], dtype='object')

In [46]:

df[['CountryName', 'CountryCode']]

Out[46]:

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

195 rows × 2 columns

In [48]:

`df.isnull()`

Out[48]:

	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 × 5 columns

In [50]:

`df.isnull().sum()`

Out[50]:

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

In [52]: `df.dtypes`

Out[52]:

	CountryName	object
CountryCode	object	
BirthRate	float64	
InternetUsers	float64	
IncomeGroup	object	
dtype:	object	

In [54]: `df.columns`

Out[54]:

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

In [59]:

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

Out[59]:

	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 × 3 columns

In [61]: `df_categorical.describe()`

Out[61]:

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

In [65]: `df.describe()`

Out[65]:

	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 [67]: df_numerical = df[['BirthRate']]

In [69]: df_numerical

Out[69]:

	BirthRate
0	10.244
1	35.253
2	45.985
3	12.877
4	11.044
...	...
190	32.947
191	20.850
192	42.394
193	40.471
194	35.715

195 rows × 1 columns

In [71]: df_num = df[['BirthRate', 'InternetUsers']]
df_num

Out[71]:

	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 [73]:

`df_num.describe()`

Out[73]:

	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 [75]:

`df.columns`

Out[75]:

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

In [77]:

`['CountryName', 'BirthRate']`

Out[77]:

`['CountryName', 'BirthRate']`

In [79]:

`df.head()`

Out[79]:

	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 [81]:

`df.columns`

```
Out[81]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
       dtype='object')
```

In [83]:

`len(df.columns)`

Out[83]:

5

In [85]:

`df.BirthRate * df.InternetUsers`

```
Out[85]: 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 [101...]

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

In [103...]

`df`

Out[103...]

	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 [91]:

df

Out[91]:

	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 [93]: `df.columns`Out[93]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup', 'myCalc'], dtype='object')`In [95]: `len(df.columns)`

Out[95]: 6

In [105...]: `df = df.drop('myCalc', axis = 1)`In [107...]: `df`

Out[107...]

	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 [109...]

df['InternetUsers'] < 2

Out[109...]

```
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 [113...]

df.InternetUsers < 2

Out[113...]

```
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 [115... Filter = df.InternetUsers < 2
```

```
In [117... Filter
```

```
Out[117... 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 [119... df[Filter]
```

	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

```
In [123... 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

```
In [125...]: len(df[Filter])
```

```
Out[125...]: 9
```

```
In [129...]: Filter2 = df.BirthRate > 40
```

```
In [131...]: Filter & Filter2
```

```
Out[131...]: 0      False
 1      False
 2      False
 3      False
 4      False
 ...
190     False
191     False
192     False
193     False
194     False
Length: 195, dtype: bool
```

```
In [133...]: 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

```
In [135...]: df[(df.BirthRate > 40) & (df.InternetUsers < 2)]
```

	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

```
In [137...]: df[df.IncomeGroup == 'Low income']
```

Out[137...]

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
1	Afghanistan	AFG	35.253	5.90	Low income
11	Burundi	BDI	44.151	1.30	Low income
13	Benin	BEN	36.440	4.90	Low income
14	Burkina Faso	BFA	40.551	9.10	Low income
29	Central African Republic	CAF	34.076	3.50	Low income
38	Comoros	COM	34.326	6.50	Low income
52	Eritrea	ERI	34.800	0.90	Low income
55	Ethiopia	ETH	32.925	1.90	Low income
64	Guinea	GIN	37.337	1.60	Low income
65	Gambia, The	GMB	42.525	14.00	Low income
66	Guinea-Bissau	GNB	37.503	3.10	Low income
77	Haiti	HTI	25.345	10.60	Low income
93	Cambodia	KHM	24.462	6.80	Low income
99	Liberia	LBR	35.521	3.20	Low income
111	Madagascar	MDG	34.686	3.00	Low income
115	Mali	MLI	44.138	3.50	Low income
120	Mozambique	MOZ	39.705	5.40	Low income
123	Malawi	MWI	39.459	5.05	Low income
127	Niger	NER	49.661	1.70	Low income
132	Nepal	NPL	20.923	13.30	Low income
148	Rwanda	RWA	32.689	9.00	Low income
154	Sierra Leone	SLE	36.729	1.70	Low income
156	Somalia	SOM	43.891	1.50	Low income
158	South Sudan	SSD	37.126	14.10	Low income
167	Chad	TCD	45.745	2.30	Low income
168	Togo	TGO	36.080	4.50	Low income
177	Tanzania	TZA	39.518	4.40	Low income
178	Uganda	UGA	43.474	16.20	Low income
192	Congo, Dem. Rep.	COD	42.394	2.20	Low income
194	Zimbabwe	ZWE	35.715	18.50	Low income

In [141...]

df[df.IncomeGroup == 'High income']

Out[141...]

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.90	High income
4	United Arab Emirates	ARE	11.044	88.00	High income
5	Argentina	ARG	17.716	59.90	High income
7	Antigua and Barbuda	ATG	16.447	63.40	High income
8	Australia	AUS	13.200	83.00	High income
...
174	Trinidad and Tobago	TTO	14.590	63.80	High income
180	Uruguay	URY	14.374	57.69	High income
181	United States	USA	12.500	84.20	High income
184	Venezuela, RB	VEN	19.842	54.90	High income
185	Virgin Islands (U.S.)	VIR	10.700	45.30	High income

67 rows × 5 columns

In [143...]

df.IncomeGroup.unique()

Out[143...]

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

In [147...]

df.IncomeGroup.nunique()

Out[147...]

4

In [150...]

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

%matplotlib inline
plt.rcParams['figure.figsize'] = 6,2

import warnings
warnings.filterwarnings('ignore')
```

In [152...]

df.head()

Out[152...]

	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

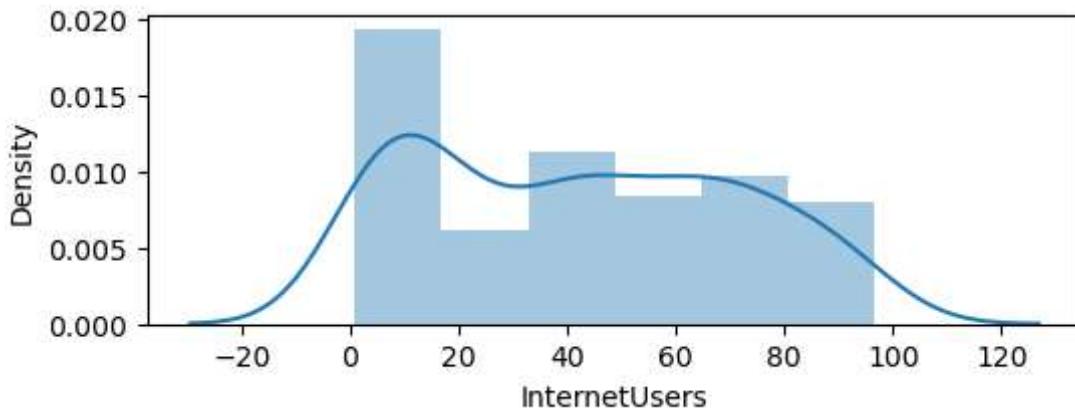
In [154...]

df["InternetUsers"]

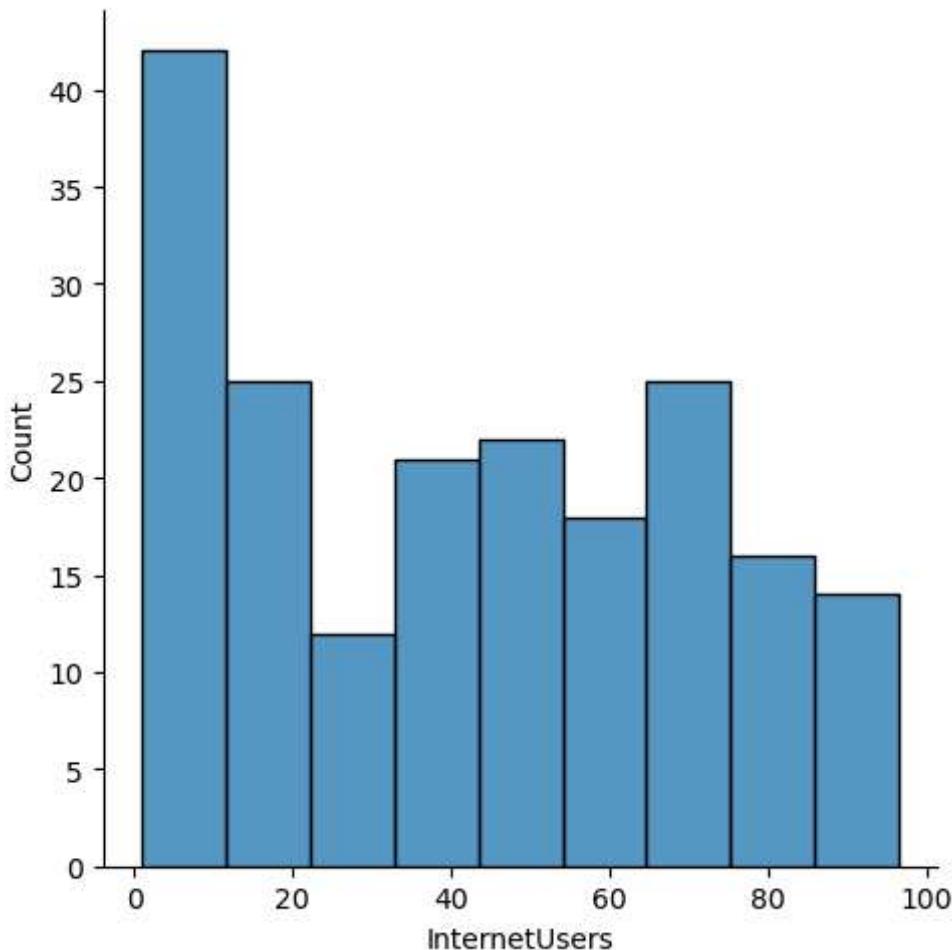
```
Out[154... 0    78.9
          1    5.9
          2   19.1
          3   57.2
          4   88.0
          ...
         190  20.0
         191  46.5
         192  2.2
         193  15.4
         194  18.5
Name: InternetUsers, Length: 195, dtype: float64
```

```
In [176... vis1 = sns.distplot(df["InternetUsers"])
```

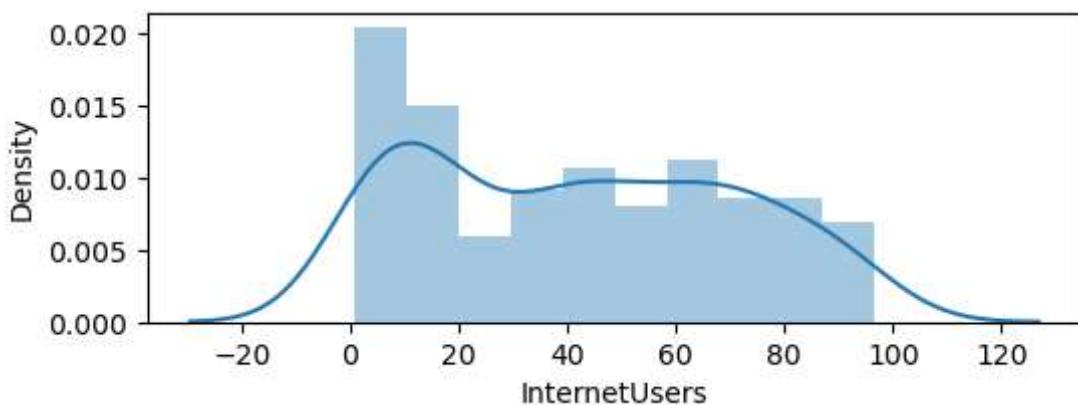
```
In [178... plt.show(vis1)
```



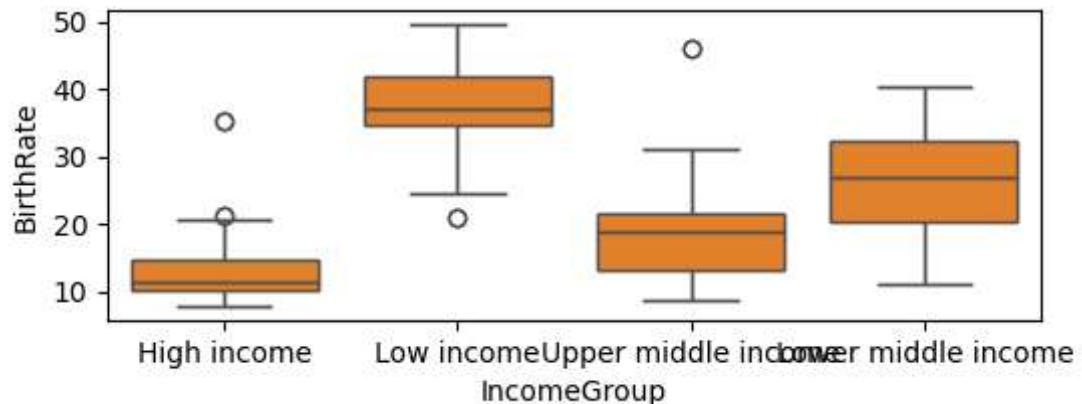
```
In [184... vis2 = sns.distplot(df["InternetUsers"])
plt.show(vis2)
```



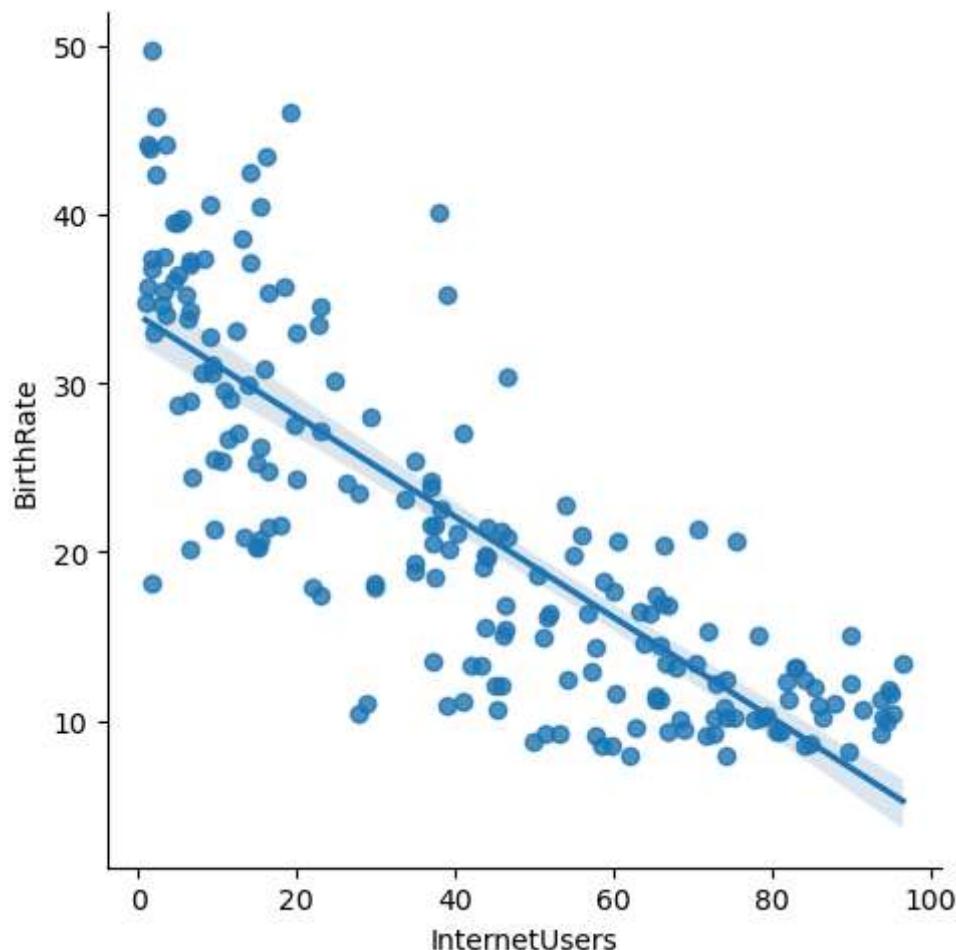
```
In [190]: vis3 = sns.distplot(df["InternetUsers"], bins = 10)  
plt.show(vis3)
```



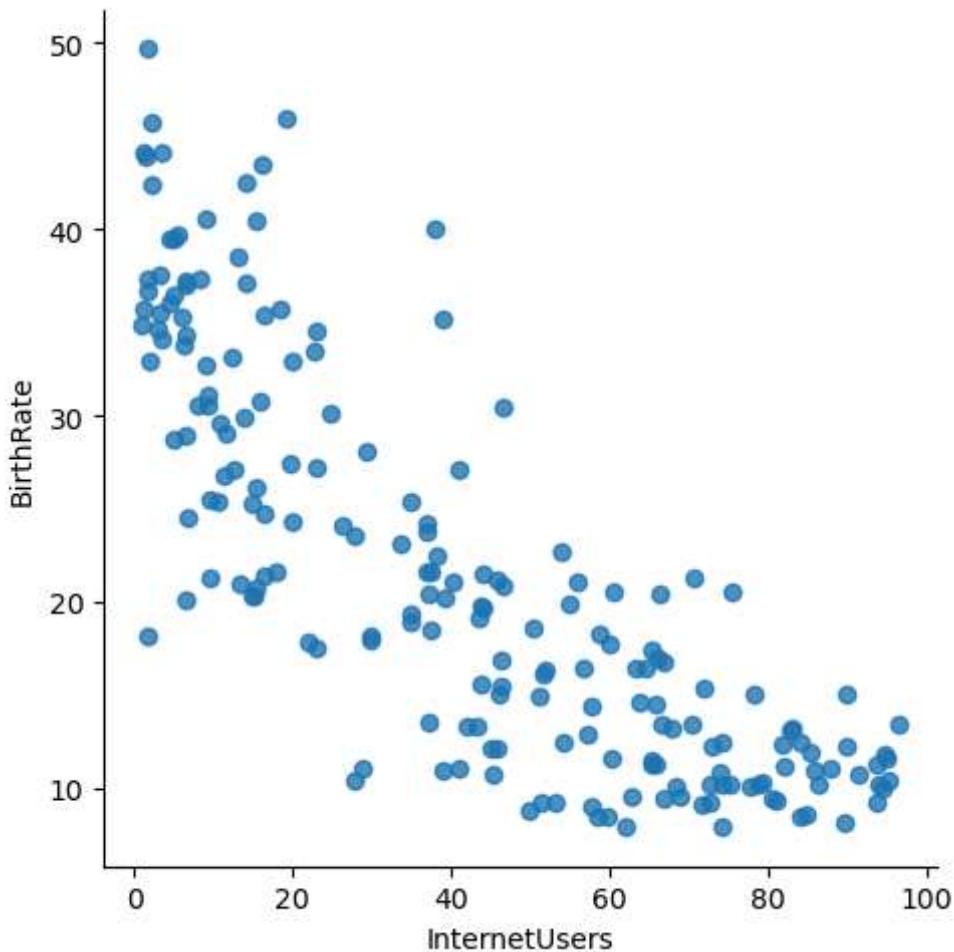
```
In [194]: vis4 = sns.boxplot(data = df, x = 'IncomeGroup', y = 'BirthRate')  
plt.show(vis4)
```



```
In [208]: vis5 = sns.lmplot(data = df, x = 'InternetUsers', y = 'BirthRate')
plt.show(vis5)
```



```
In [210]: vis5 = sns.lmplot(data = df, x = 'InternetUsers', y = 'BirthRate', fit_reg = False)
plt.show(vis5)
```



```
In [216]: vis5 = sns.lmplot(data = df, x = 'InternetUsers', y = 'BirthRate', fit_reg = True)
plt.show(vis5)
```

