

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

```
In [2]: import numpy as np
import matplotlib.pyplot as plt
```

```
In [3]: import seaborn as sns
```

```
In [4]: stats=pd.read_excel(r'C:\Users\nlnar\Downloads\data.xlsx')
```

```
In [5]: stats
```

```
Out[5]:
```

	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 [6]: len(stats) #no of rows
```

```
Out[6]: 195
```

```
In [7]: stats.columns #to call columns
```

```
Out[7]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
              'IncomeGroup'],
              dtype='object')
```

```
In [8]: len(stats.columns)
```

```
Out[8]: 5
```

```
In [9]: stats.shape #shows rows and col
```

Out[9]: (195, 5)

```
In [10]: stats.isnull()
```

Out[10]:

	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 [11]: stats.isna()
```

Out[11]:

	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 [12]: stats.isnull().sum()
```

```
Out[12]: CountryName      0
CountryCode      0
BirthRate        0
InternetUsers     0
IncomeGroup      0
dtype: int64
```

```
In [13]: stats.dtypes
```

```
Out[13]: CountryName      object
CountryCode      object
BirthRate        float64
InternetUsers     float64
IncomeGroup      object
dtype: object
```

```
In [14]: stats.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
```

```
In [15]: type(stats)
```

```
Out[15]: pandas.core.frame.DataFrame
```

```
In [16]: stats.head()
```

```
Out[16]:
```

	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 [17]: stats.tail()
```

```
Out[17]:
```

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

```
In [18]: stats.head(2)
```

```
Out[18]:
```

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

```
In [19]: stats.tail(2)
```

```
Out[19]:
```

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

```
In [20]: stats.columns
```

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

```
In [21]: stats['CountryName'] #to extract country name along with rows
```

```
Out[21]: 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
```

```
In [22]: stats[:]# to extract entire data
```

Out[22]:

	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 [23]: `stats['BirthRate']`

Out[23]:

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

Name: BirthRate, Length: 195, dtype: float64

In [24]: `stats[['InternetUsers','BirthRate']]` *#to call two columns*

Out[24]:

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

195 rows × 2 columns

## split the data set into numerical and categorical data

In [25]: `stats_numeric_data=stats[['BirthRate','InternetUsers']]`

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

Out[26]:

	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

In [27]: `stats_categorical_data=stats[['CountryCode','IncomeGroup']]`

In [28]: `stats_categorical_data.head()`

Out[28]:

	CountryCode	IncomeGroup
0	ABW	High income
1	AFG	Low income
2	AGO	Upper middle income
3	ALB	Upper middle income
4	ARE	High income

In [29]:

```
print(stats.shape)
print(stats_numeric_data.shape)
#print(stats_categorical_data.shape)
```

(195, 5)  
(195, 2)

## slicing in pandas

In [30]:

```
stats[:]
```

Out[30]:

	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 [31]:

```
stats[:6]
```

Out[31]:

	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
5	Argentina	ARG	17.716	59.9	High income

In [32]: `stats[:2]`

Out[32]:

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

In [33]: `stats[3:]`

Out[33]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
3	Albania	ALB	12.877	57.2	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0	High income
5	Argentina	ARG	17.716	59.9	High income
6	Armenia	ARM	13.308	41.9	Lower middle income
7	Antigua and Barbuda	ATG	16.447	63.4	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

192 rows × 5 columns

In [34]: `stats[3:9]`

Out[34]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
3	Albania	ALB	12.877	57.2	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0	High income
5	Argentina	ARG	17.716	59.9	High income
6	Armenia	ARM	13.308	41.9	Lower middle income
7	Antigua and Barbuda	ATG	16.447	63.4	High income
8	Australia	AUS	13.200	83.0	High income

In [35]:

```
stats[3:50:5]
```

Out[35]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
3	Albania	ALB	12.877	57.2000	Upper middle income
8	Australia	AUS	13.200	83.0000	High income
13	Benin	BEN	36.440	4.9000	Low income
18	Bahamas, The	BHS	15.339	72.0000	High income
23	Bolivia	BOL	24.236	36.9400	Lower middle income
28	Botswana	BWA	25.267	15.0000	Upper middle income
33	China	CHN	12.100	45.8000	Upper middle income
38	Comoros	COM	34.326	6.5000	Low income
43	Cyprus	CYP	11.436	65.4548	High income
48	Dominican Republic	DOM	21.198	45.9000	Upper middle income

In [36]:

```
stats[:, :-1]
```

Out[36]:

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

195 rows × 5 columns

In [37]: `stats[::-2] #reverse skipping two rows`

Out[37]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
<b>194</b>	Zimbabwe	ZWE	35.715	18.5	Low income
<b>192</b>	Congo, Dem. Rep.	COD	42.394	2.2	Low income
<b>190</b>	Yemen, Rep.	YEM	32.947	20.0	Lower middle income
<b>188</b>	West Bank and Gaza	PSE	30.394	46.6	Lower middle income
<b>186</b>	Vietnam	VNM	15.537	43.9	Lower middle income
...	...	...	...	...	...
<b>8</b>	Australia	AUS	13.200	83.0	High income
<b>6</b>	Armenia	ARM	13.308	41.9	Lower middle income
<b>4</b>	United Arab Emirates	ARE	11.044	88.0	High income
<b>2</b>	Angola	AGO	45.985	19.1	Upper middle income
<b>0</b>	Aruba	ABW	10.244	78.9	High income

98 rows × 5 columns

In [38]: `stats[50:100]`

Out[38]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
50	Ecuador	ECU	21.070	40.353684	Upper middle income
51	Egypt, Arab Rep.	EGY	28.032	29.400000	Lower middle income
52	Eritrea	ERI	34.800	0.900000	Low income
53	Spain	ESP	9.100	71.635000	High income
54	Estonia	EST	10.300	79.400000	High income
55	Ethiopia	ETH	32.925	1.900000	Low income
56	Finland	FIN	10.700	91.514400	High income
57	Fiji	FJI	20.463	37.100000	Upper middle income
58	France	FRA	12.300	81.919800	High income
59	Micronesia, Fed. Sts.	FSM	23.511	27.800000	Lower middle income
60	Gabon	GAB	30.555	9.200000	Upper middle income
61	United Kingdom	GBR	12.200	89.844100	High income
62	Georgia	GEO	13.332	43.300000	Lower middle income
63	Ghana	GHA	33.131	12.300000	Lower middle income
64	Guinea	GIN	37.337	1.600000	Low income
65	Gambia, The	GMB	42.525	14.000000	Low income
66	Guinea-Bissau	GNB	37.503	3.100000	Low income
67	Equatorial Guinea	GNQ	35.362	16.400000	High income
68	Greece	GRC	8.500	59.866300	High income
69	Grenada	GRD	19.334	35.000000	Upper middle income
70	Greenland	GRL	14.500	65.800000	High income
71	Guatemala	GTM	27.465	19.700000	Lower middle income
72	Guam	GUM	17.389	65.400000	High income
73	Guyana	GUY	18.885	35.000000	Lower middle income
74	Hong Kong SAR, China	HKG	7.900	74.200000	High income
75	Honduras	HND	21.593	17.800000	Lower middle income
76	Croatia	HRV	9.400	66.747600	High income
77	Haiti	HTI	25.345	10.600000	Low income
78	Hungary	HUN	9.200	72.643900	High income
79	Indonesia	IDN	20.297	14.940000	Lower middle income

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
80	India	IND	20.291	15.100000	Lower middle income
81	Ireland	IRL	15.000	78.247700	High income
82	Iran, Islamic Rep.	IRN	17.900	29.950000	Upper middle income
83	Iraq	IRQ	31.093	9.200000	Upper middle income
84	Iceland	ISL	13.400	96.546800	High income
85	Israel	ISR	21.300	70.800000	High income
86	Italy	ITA	8.500	58.459300	High income
87	Jamaica	JAM	13.540	37.100000	Upper middle income
88	Jordan	JOR	27.046	41.000000	Upper middle income
89	Japan	JPN	8.200	89.710000	High income
90	Kazakhstan	KAZ	22.730	54.000000	Upper middle income
91	Kenya	KEN	35.194	39.000000	Lower middle income
92	Kyrgyz Republic	KGZ	27.200	23.000000	Lower middle income
93	Cambodia	KHM	24.462	6.800000	Low income
94	Kiribati	KIR	29.044	11.500000	Lower middle income
95	Korea, Rep.	KOR	8.600	84.770000	High income
96	Kuwait	KWT	20.575	75.460000	High income
97	Lao PDR	LAO	27.051	12.500000	Lower middle income
98	Lebanon	LBN	13.426	70.500000	Upper middle income
99	Liberia	LBR	35.521	3.200000	Low income

In [39]: stats[0:200]

Out[39]:

	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 [40]: stats

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 [41]: stats

Out[41]:

	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

## 6th nov class

In [42]: `pd.__version__`

Out[42]: '2.2.3'

In [43]: `stats.columns`

Out[43]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',  
          'IncomeGroup'],  
          dtype='object')

In [44]: `stats.describe()` # summary of the data set ,only numerical data attributes are called

Out[44]:

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

In [45]: `stats.describe().transpose()`

Out[45]:

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

In [46]: `stats_numeric_data.describe().transpose()`

Out[46]:

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

In [47]: `stats_categorical_data.describe()` *#it describes the non numeric data of the data se*

Out[47]:

	CountryCode	IncomeGroup
<b>count</b>	195	195
<b>unique</b>	195	4
<b>top</b>	ABW	High income
<b>freq</b>	1	67

In [48]: `stats.head(2)`

Out[48]:

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

```
In [49]: stats['mycalc']=stats['BirthRate']*stats['InternetUsers']
stats.head(2) # to add the attribute or column syntax
```

```
Out[49]:
```

	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

```
In [50]: stats.head(4)
```

```
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

```
In [51]: stats.columns
```

```
Out[51]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
               'IncomeGroup', 'mycalc'],
              dtype='object')
```

```
In [52]: len(stats.columns)
```

```
Out[52]: 6
```

```
In [53]: stats=stats.drop('mycalc',axis=1) #axis 1 is column ;axis 0 is rows; drop fn drops
```

```
In [54]: stats
```

Out[54]:

	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 [55]: stats['InternetUsers']<2 #without filtering gives only boolean values
```

```
Out[55]: 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 [56]: stats[stats['InternetUsers']<2] # this is after filteration with filters
```

Out[56]:

	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 [57]:

```
stats[stats['BirthRate']>40]
```

Out[57]:

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

In [58]:

```
(stats[(stats.BirthRate>40)&(stats.InternetUsers<2)])
```

Out[58]:

	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 [59]: len(stats[(stats.BirthRate>40)&(stats.InternetUsers<2)])
```

```
Out[59]: 3
```

```
In [60]: stats[stats.IncomeGroup == 'High income'] # filtering the data
```

```
Out[60]:
```

	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 [61]: stats[stats.IncomeGroup == 'Low income']
```

Out[61]:

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

```
In [62]: stats.IncomeGroup.nunique()
```

```
Out[62]: 4
```

```
In [63]: stats.InternetUsers.nunique()
```

```
Out[63]: 181
```

```
In [64]: import matplotlib.pyplot as plt
```

```
In [65]: import seaborn as sns
```

```
In [66]: %matplotlib inline
```

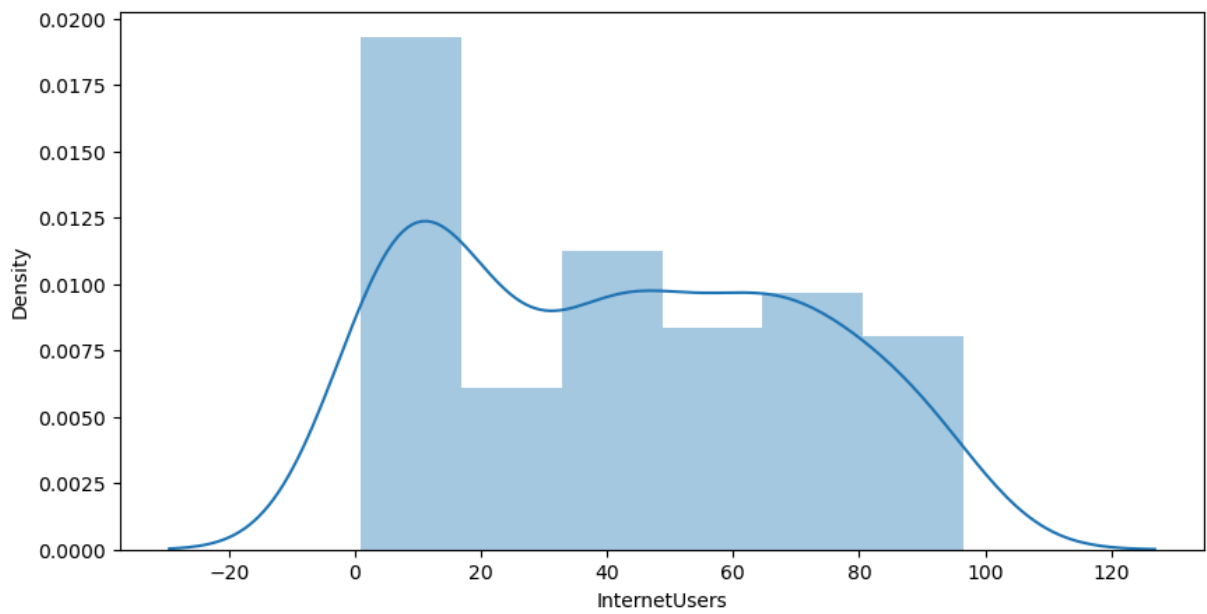
```
In [67]: plt.rcParams['figure.figsize']=10,5
```

```
In [68]: import warnings  
warnings.filterwarnings('ignore')
```

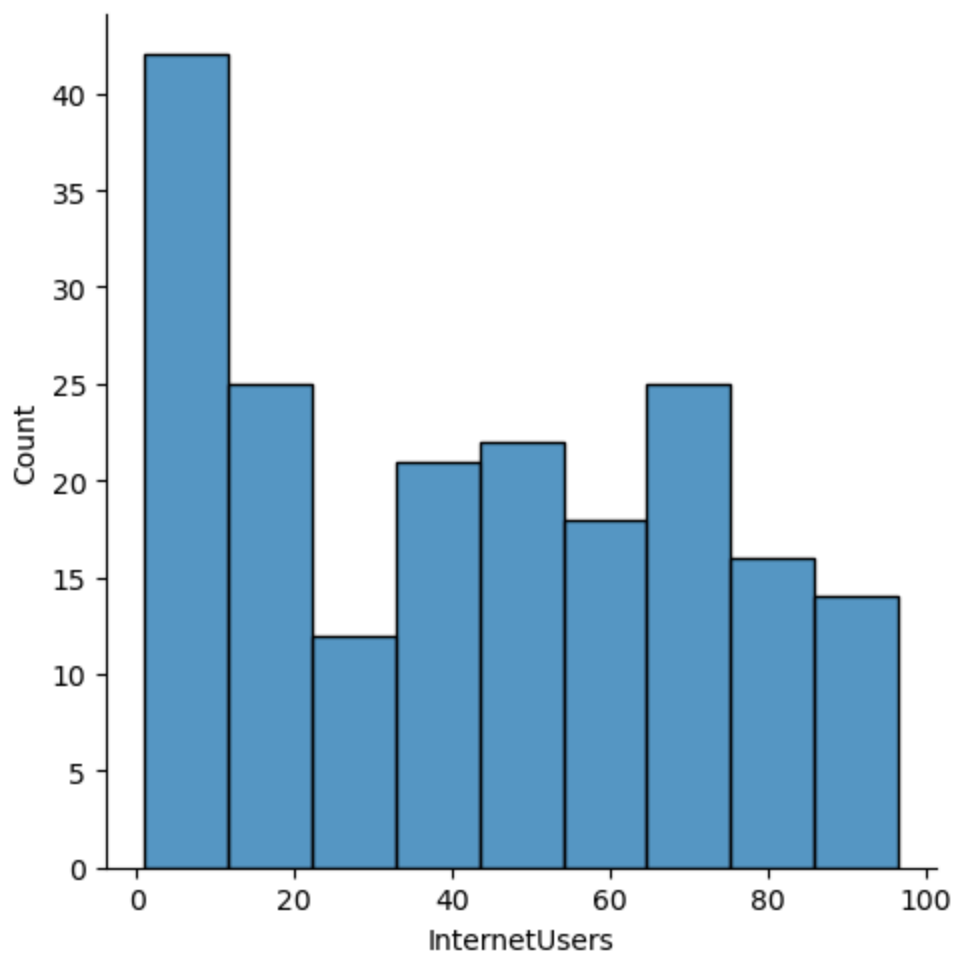
```
In [69]: stats['InternetUsers']
```

```
Out[69]: 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
```

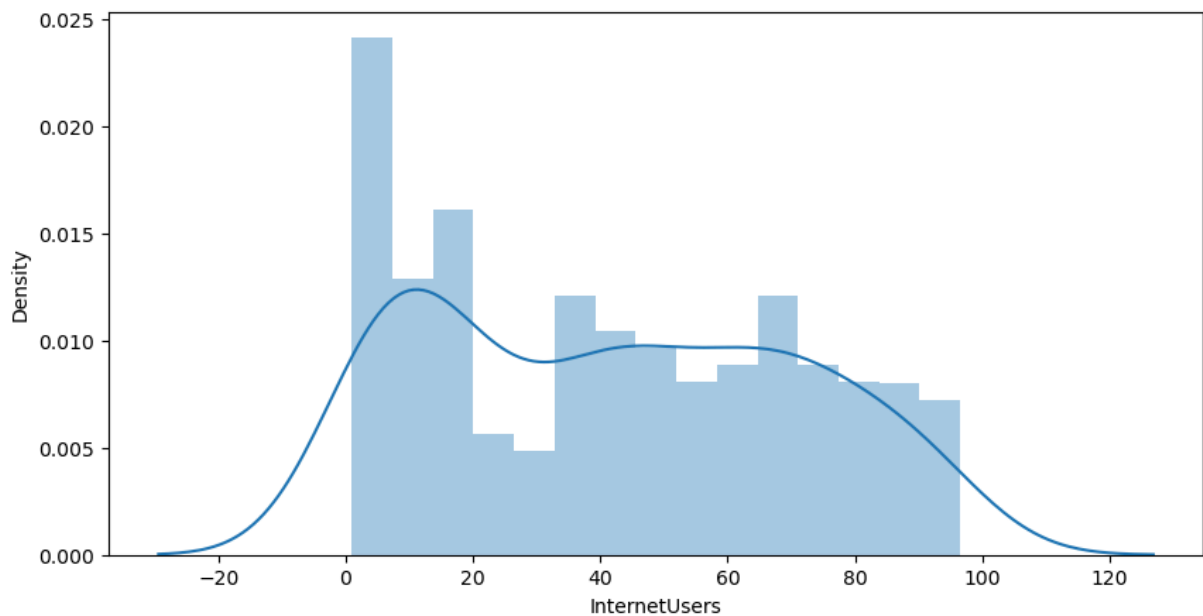
```
In [70]: vis1=sns.distplot(stats["InternetUsers"]) # data distribution  
plt.show()
```



```
In [71]: vis1=sns.displot(stats["InternetUsers"])  
plt.show()
```

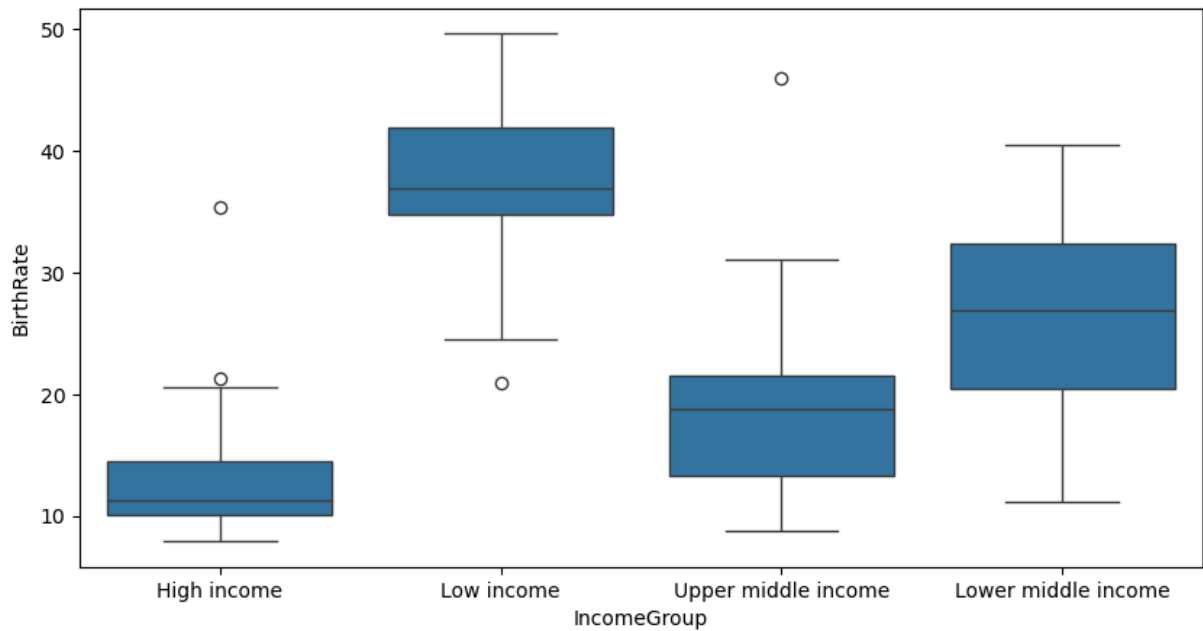


```
In [72]: vis1=sns.distplot(stats["InternetUsers"],bins=15) #bins are used for visualization,  
plt.show()
```

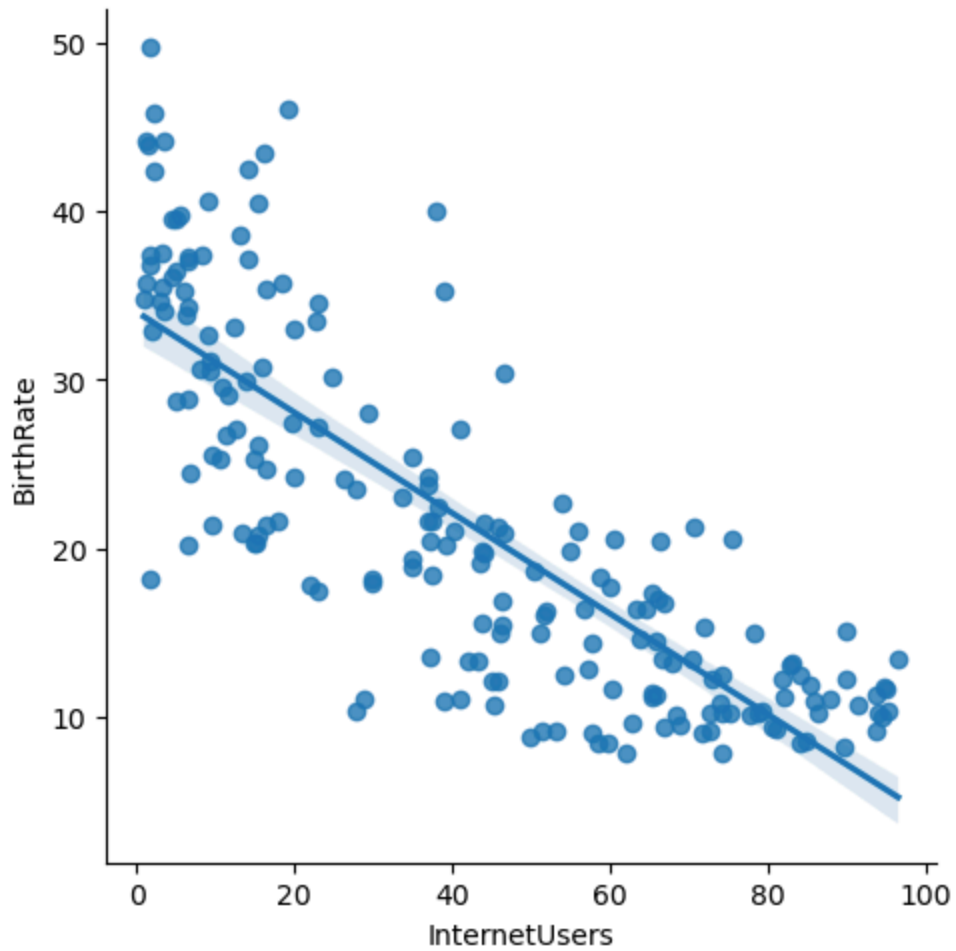


## box plot bi variat

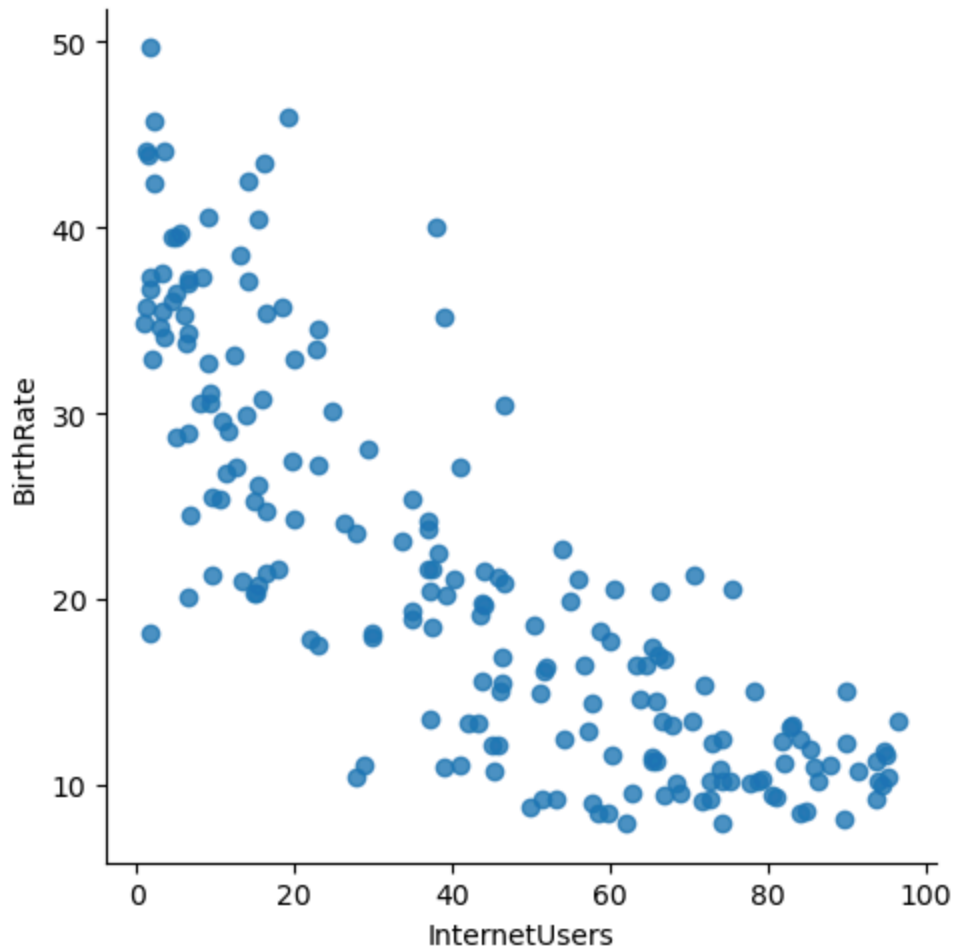
```
In [73]: vis4=sns.boxplot(data=stats,x="IncomeGroup",y='BirthRate')
plt.show()
```



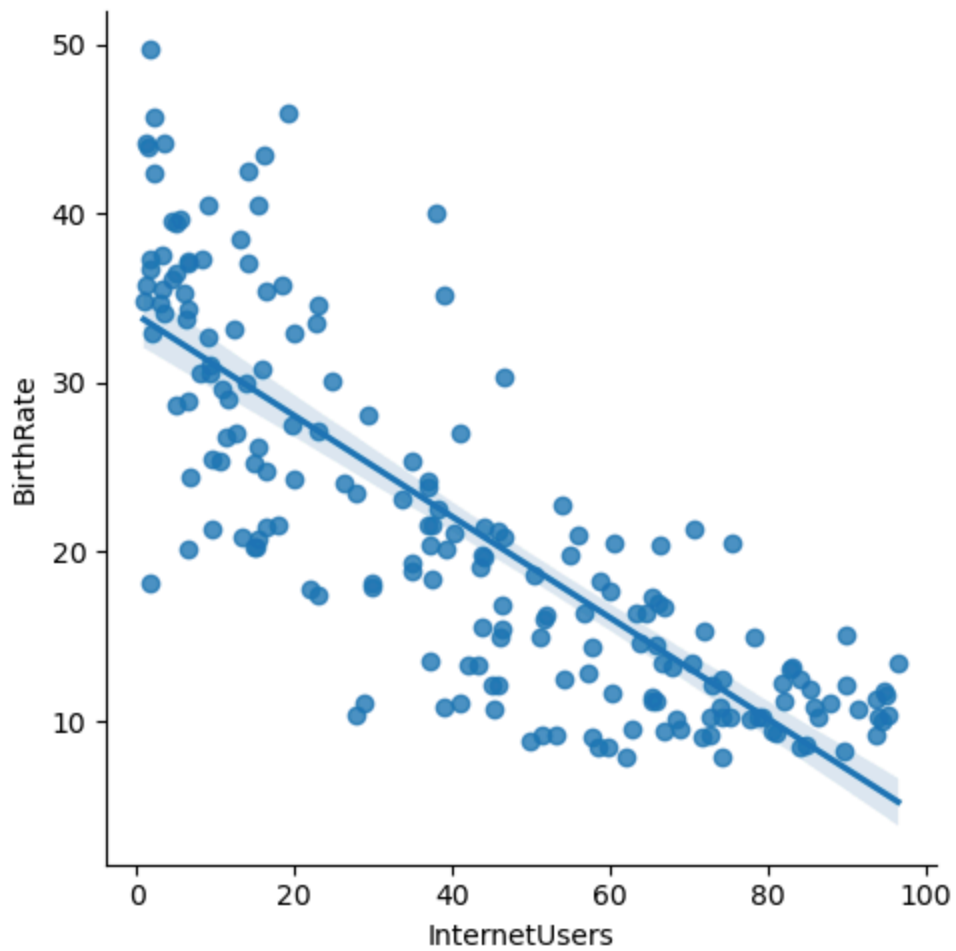
```
In [74]: vis5=sns.lmplot(data=stats,x="InternetUsers",y='BirthRate') # this is called as sca
plt.show()
```



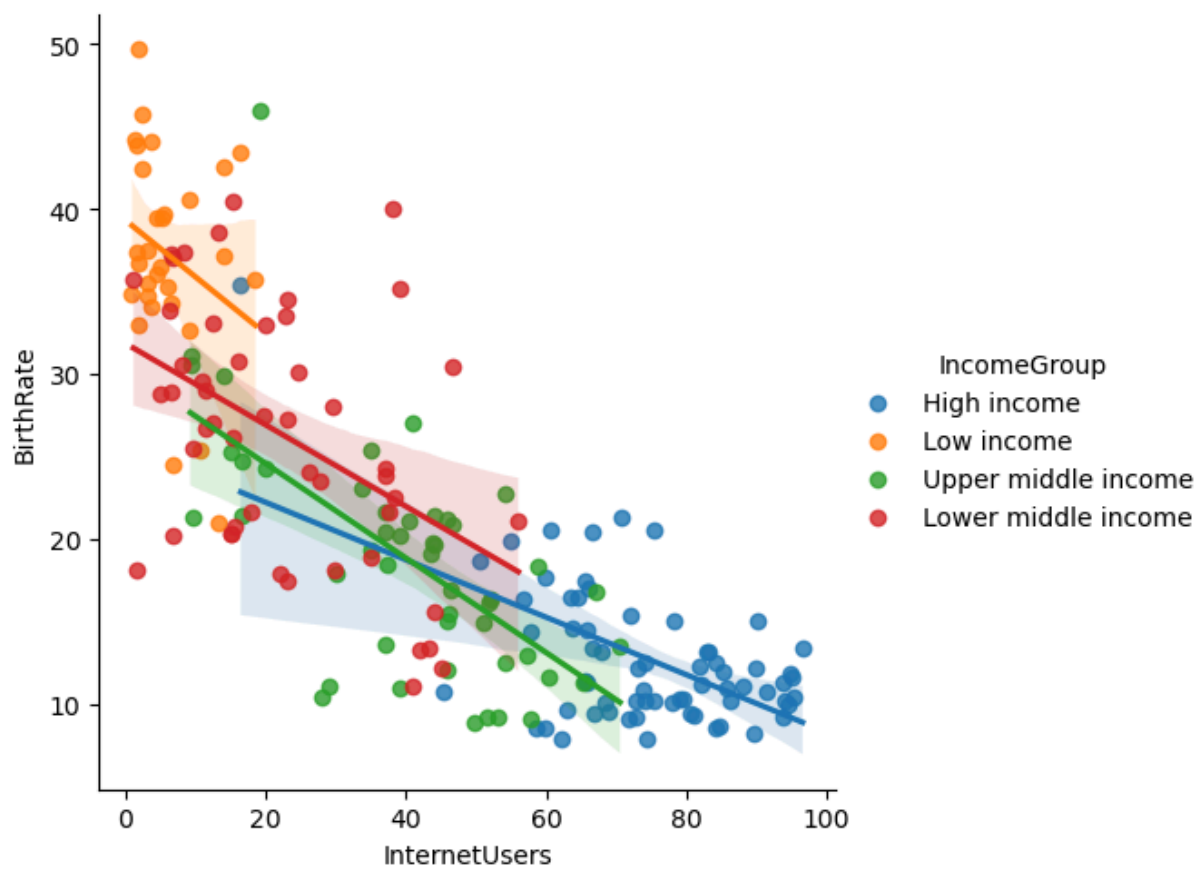
```
In [75]: vis5=sns.lmplot(data=stats,x="InternetUsers",y='BirthRate',fit_reg=False)
plt.show()
```



```
In [76]: vis5=sns.lmplot(data=stats,x="InternetUsers",y='BirthRate',fit_reg=True)
plt.show()
```



```
In [84]: vis5=sns.lmplot(data=stats,x="InternetUsers",y='BirthRate',fit_reg=True,hue='Income',
plt.show())
```



In [ ]: