In [47]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [48]:
```
import pandas as pd
```

In [49]:
```
pip install openpyxl
```

```
Requirement already satisfied: openpyxl in c:\users\santo\anaconda3\lib\site-packages (3.1.5)
Requirement already satisfied: et-xmlfile in c:\users\santo\anaconda3\lib\site-packages (from openpyxl) (1.1.0)
Note: you may need to restart the kernel to use updated packages.
```

In [50]:
```
import pandas as pd

stats = pd.read_excel(r"C:\Users\santo\OneDrive\Desktop\naresh it classroom\Project python\santosh data.xlsx")

stats
```

Out[50]:

	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 [51]:
```
len(stats)
```

Out[51]:
```
195
```

In [52]:
```
stats.columns
```

Out[52]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

In [53]:
```
len(stats.columns)
```

Out[53]:
```
5
```

In [54]:
```
stats.shape
```

Out[54]: (195, 5)

In [55]:
```
stats.isnull() #To check the missing Value
```

Out[55]:

	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.isna()
```

Out[12]:

	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 [13]:
```
stats.isnull().sum()
```

Out[13]:
```
CountryName      0
CountryCode      0
BirthRate        0
InternetUsers    0
IncomeGroup      0
dtype: int64
```

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

Out[14]:
```
CountryName       object
CountryCode       object
BirthRate        float64
InternetUsers    float64
IncomeGroup       object
dtype: object
```

In [15]:
```
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 [16]:
```
type(stats)
```

Out[16]:
```
pandas.core.frame.DataFrame
```

In [17]:
```
stats.head()
```

Out[17]:

	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 [18]:
```
stats.tail()
```

Out[18]:

	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 [19]:
```
stats.head(2)
```

Out[19]:

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

In [20]:
```
stats.tail(2)
```

Out[20]:

	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 [21]:
```
# can you split
```

In [22]:
```
stats.columns
```

Out[22]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

In [23]:
```
stats['CountryName']
```

Out[23]:
```
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 [24]:
```
stats['CountryCode']
```

Out[24]:
```
0      ABW
1      AFG
2      AGO
3      ALB
4      ARE
      ...
190    YEM
191    ZAF
192    COD
193    ZMB
194    ZWE
Name: CountryCode, Length: 195, dtype: object
```

In [25]:
```
stats['BirthRate']
```

Out[25]:
```
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 [26]:
```
stats['InternetUsers']
```

Out[26]:
```
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 [27]:
```
stats[:]
```

Out[27]:

	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 [28]:
```
stats['IncomeGroup']
```

Out[28]:
```
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
Name: IncomeGroup, Length: 195, dtype: object
```

In [29]:
```
stats[['BirthRate','InternetUsers']]
```

Out[29]:

	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

lets split the dataset to numerical & categories a

In [30]:
```
stats.columns
```

Out[30]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

In [31]:
```
stats_numeric_data=stats[['BirthRate','InternetUsers']]
```

In [32]:
```
stats_numeric_data.head()
```

Out[32]:

	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 [33]:
```
stats_numeric_data=stats[['CountryName', 'CountryCode','IncomeGroup']]
```

In [34]:
```
stats_numeric_data
```

Out[34]:

	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 [35]:
```
import pandas as pd

stats = pd.read_excel(r"C:\Users\santo\OneDrive\Desktop\naresh it classroom\Project python\santosh data.xlsx")

# Numeric columns
stats_numeric_data = stats.select_dtypes(include=['number'])

# Categorical columns
stats_categorical_data = stats.select_dtypes(exclude=['number'])

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

```
(195, 5)
(195, 2)
(195, 3)
```

slicing in pandas

In [36]:
```
stats[:]
```

Out[36]:

	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 [37]:
```
stats[:6]
```

Out[37]:

	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 [38]:
```
stats[:3]
```

Out[38]:

	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

In [39]:
```
stats[3:]
```

Out[39]:

	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 [40]:
```
stats[3:10]
```

Out[40]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
3	Albania	ALB	12.877	57.2000	Upper middle income
4	United Arab Emirates	ARE	11.044	88.0000	High income
5	Argentina	ARG	17.716	59.9000	High income
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

In [41]:
```
stats[3:50:5]
```

Out[41]:

	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 [42]:
```
stats[::-1]
```

Out[42]:

	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 [43]:
```
stats[::-2]
```

Out[43]:

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

98 rows × 5 columns

In [44]:
```
stats[50:100]
```

Out[44]:

	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

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
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
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

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
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 [45]:
```
stats[0:200]
```

Out[45]:

	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 [46]:
```
import pandas as pd
```

In [57]:
```
(pd.__version__)
```

Out[57]:
```
'2.2.3'
```

In [58]:
```
stats
```

Out[58]:

	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 [60]:
```
stats.columns
```

Out[60]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

In [61]:
```
stats_numeric_data.head(1)
```

Out[61]:

	BirthRate	InternetUsers
0	10.244	78.9

In [62]:
```
stats.head()
```

Out[62]:

	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 [63]:
```
stats.describe()
```

Out[63]:

	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 [65]:
```
stats.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 [66]:
```
stats.describe().transpose()
```

Out[66]:

	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 [67]:
```
stats_numeric_data.describe().transpose()
```

Out[67]:

	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 [68]:
```
stats_categorical_data.describe()
```

Out[68]:

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

In [69]:
```
stats.head(2)
```

Out[69]:

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

In [70]:
```
stats['BirthRate'] *stats['InternetUsers']
```

Out[70]:
```
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 [73]:
```
stats['myCalc'] = stats['BirthRate'] * stats['InternetUsers']
```

In [74]:
```
stats.head(2)
```

Out[74]:

	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 [75]:
```
stats.columns
```

Out[75]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'myCalc'],
      dtype='object')
```

In [76]:
```
len(stats.columns)
```

Out[76]:
```
6
```

In [8]:
```
import pandas as pd

stats = pd.read_excel(r"C:\Users\santo\OneDrive\Desktop\naresh it classroom\Project python\santosh data.xlsx")

stats
```

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]:

```
import pandas as pd

stats = pd.read_excel(r"C:\Users\santo\OneDrive\Desktop\naresh it classroom\Project python\santosh data.xlsx")

stats

print(stats.columns)
if 'myCalc' in stats.columns:
    stats = stats.drop('myCalc', axis=1)
    print("Column 'myCalc' dropped successfully.")
else:
    print("'myCalc' column not found in DataFrame.")
```

```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
'myCalc' column not found in DataFrame.
```

In [11]:
```
stats['myCalc'] = stats['BirthRate'] * stats['InternetUsers']
```

In [12]:
```
stats.columns
```

Out[12]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'myCalc'],
      dtype='object')
```

In [13]:
```
stats
```

Out[13]:

	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 [14]:
```
len(stats.columns)
```

Out[14]: 6

In [15]:
```
stats=stats.drop('myCalc', axis = 1)
```

In [16]:
```
stats
```

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
...	...	...	...	...	...
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 [17]:
```
stats.head(1)
```

Out[17]:

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

In [88]:
```
stats['InternetUsers']<2
```

Out[88]:
```
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 [87]:
```
stats[stats['InternetUsers']<2]
```

Out[87]:

	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 [90]:
```
len(stats[stats['InternetUsers']<2])
```

Out[90]:
```
9
```

In [91]:
```
stats['BirthRate'] >40
```

Out[91]:
```
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
```

In [92]:
```
stats[stats['BirthRate']>40]
```

Out[92]:

	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 [93]:
```
len(stats[stats['BirthRate']>40])
```

Out[93]: 12

In [94]:
```
stats[(stats.BirthRate>40)&(stats.InternetUsers<2)]
```

Out[94]:

	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 [95]:
```
stats
```

Out[95]:

	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 [99]:
```
stats[stats['IncomeGroup'] == 'High income']
```

Out[99]:

	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 [100…

```
stats[stats['IncomeGroup'] == 'Low income']
```

Out[100...

	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

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
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 [101…
```
stats.IncomeGroup.unique()
```

Out[101…
```
array(['High income', 'Low income', 'Upper middle income',
       'Lower middle income'], dtype=object)
```

In [103…
```
stats.IncomeGroup.nunique()
```

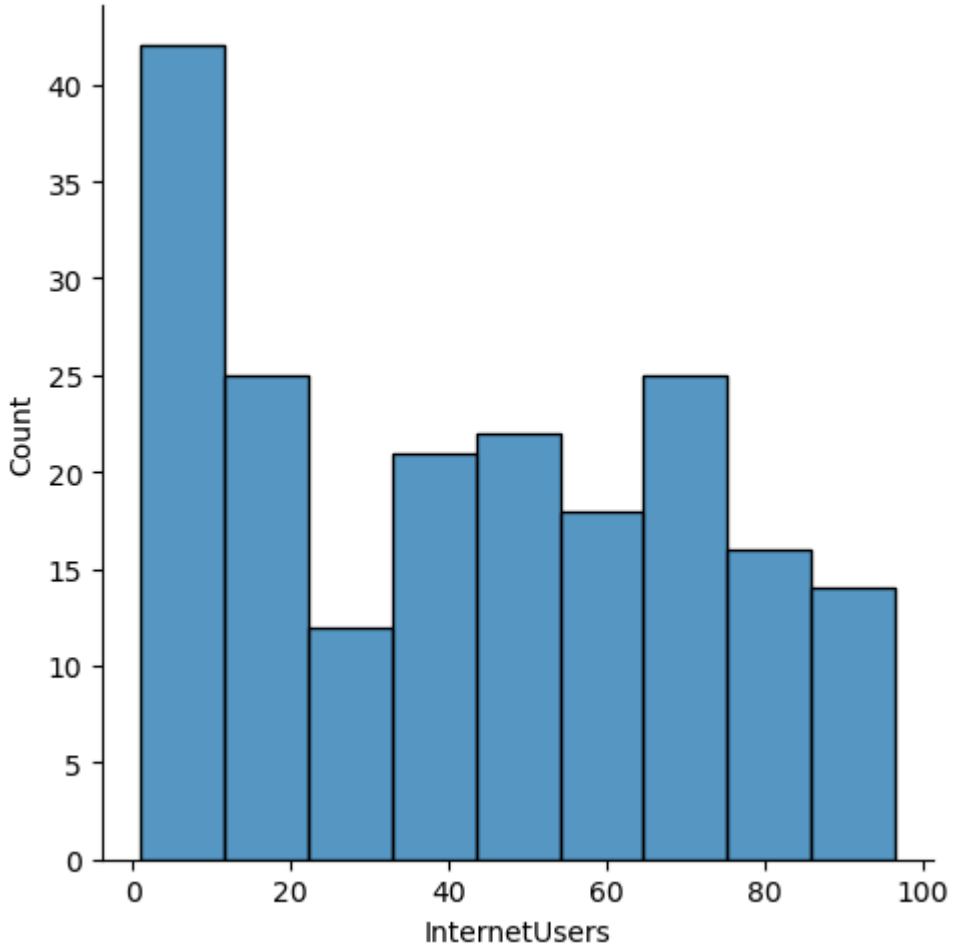
Out[103…
```
4
```

In [105…
```
import matplotlib.pyplot as plt
import seaborn as sns
import warnings

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

warnings.filterwarnings('ignore')  #os error
```

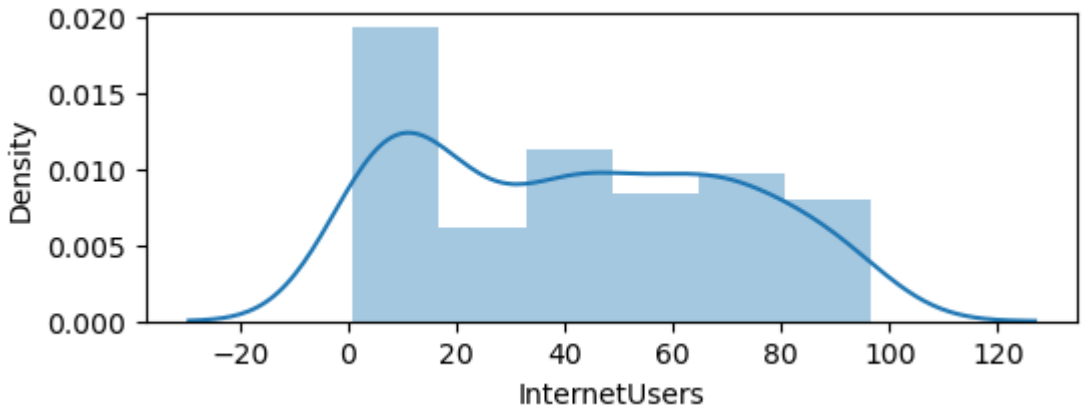
In [106…
```
stats["InternetUsers"]
```

Out[106…
```
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
```

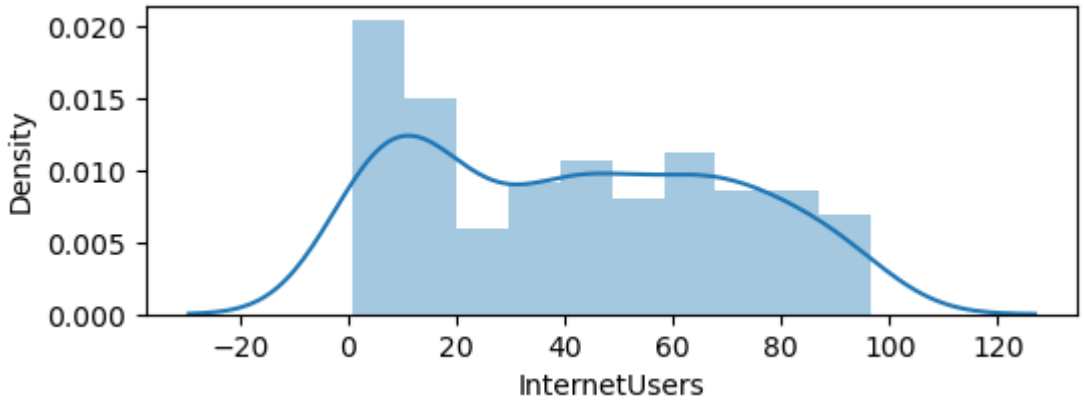
In [110…
```
vis1= sns.displot(stats["InternetUsers"])
plt.show()
```



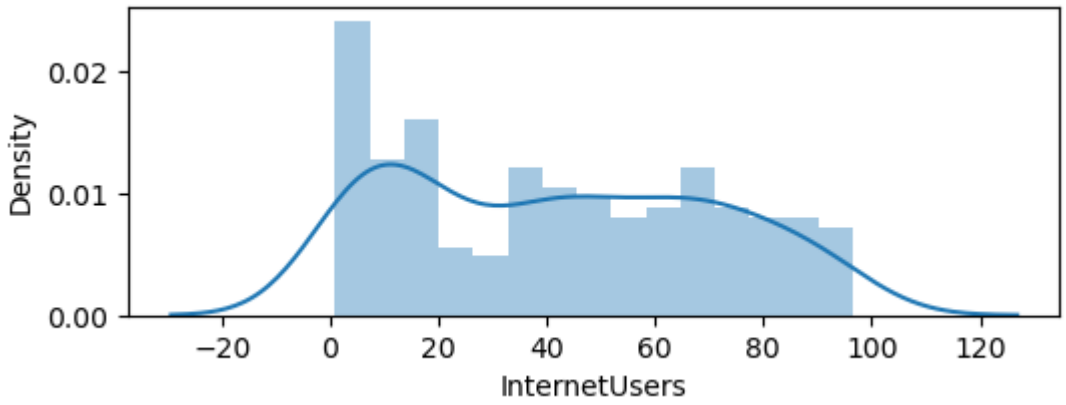
In [111…
```
vis1= sns.distplot(stats["InternetUsers"])
plt.show()
```



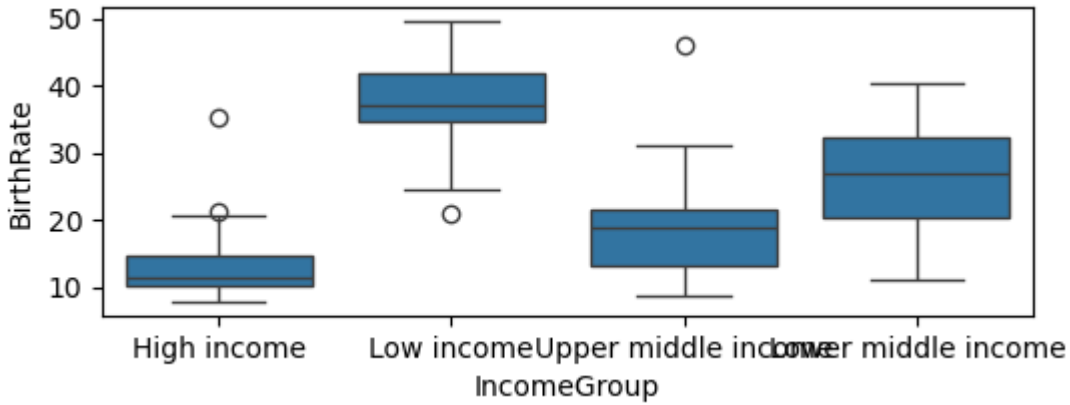
In [113…
```
vis2= sns.distplot(stats["InternetUsers"],bins=10)
plt.show()
```



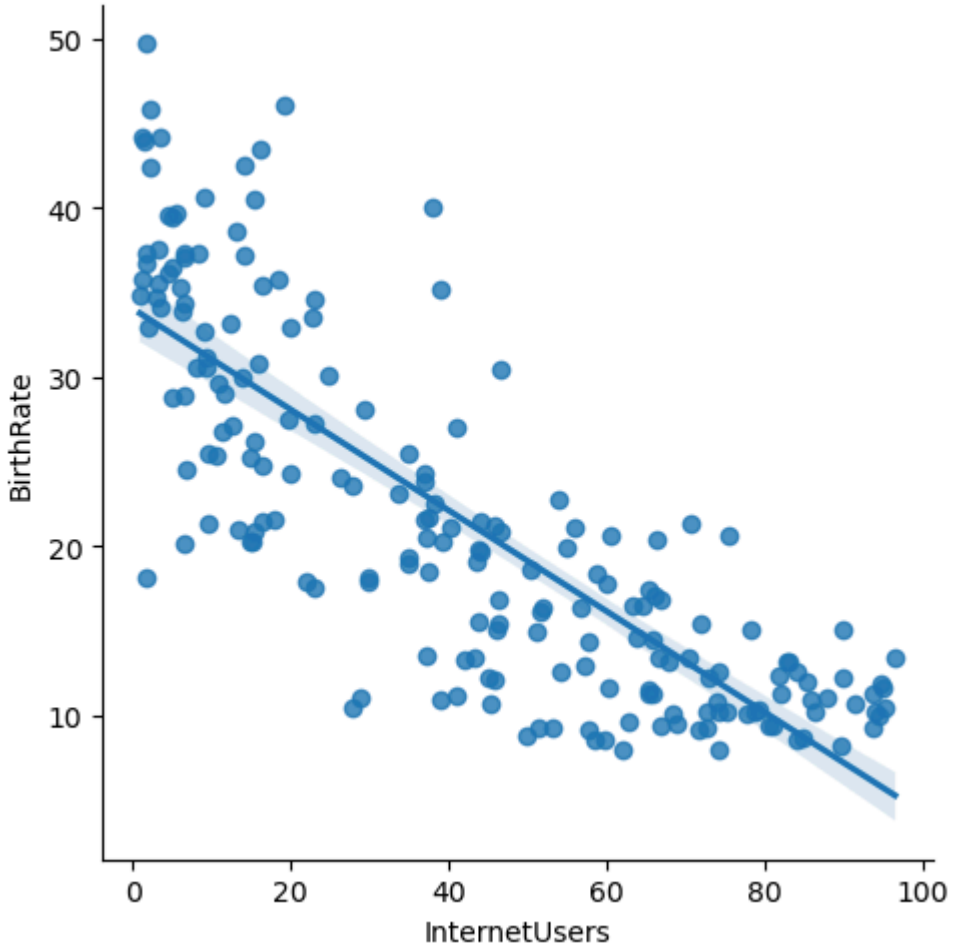
In [114…
```
vis2= sns.distplot(stats["InternetUsers"],bins=15)
plt.show()
```



In [118…
```
#Box Plots:
vis4=sns.boxplot(data = stats, x="IncomeGroup",y="BirthRate") #BI -Variable Analysis
plt.show()
```

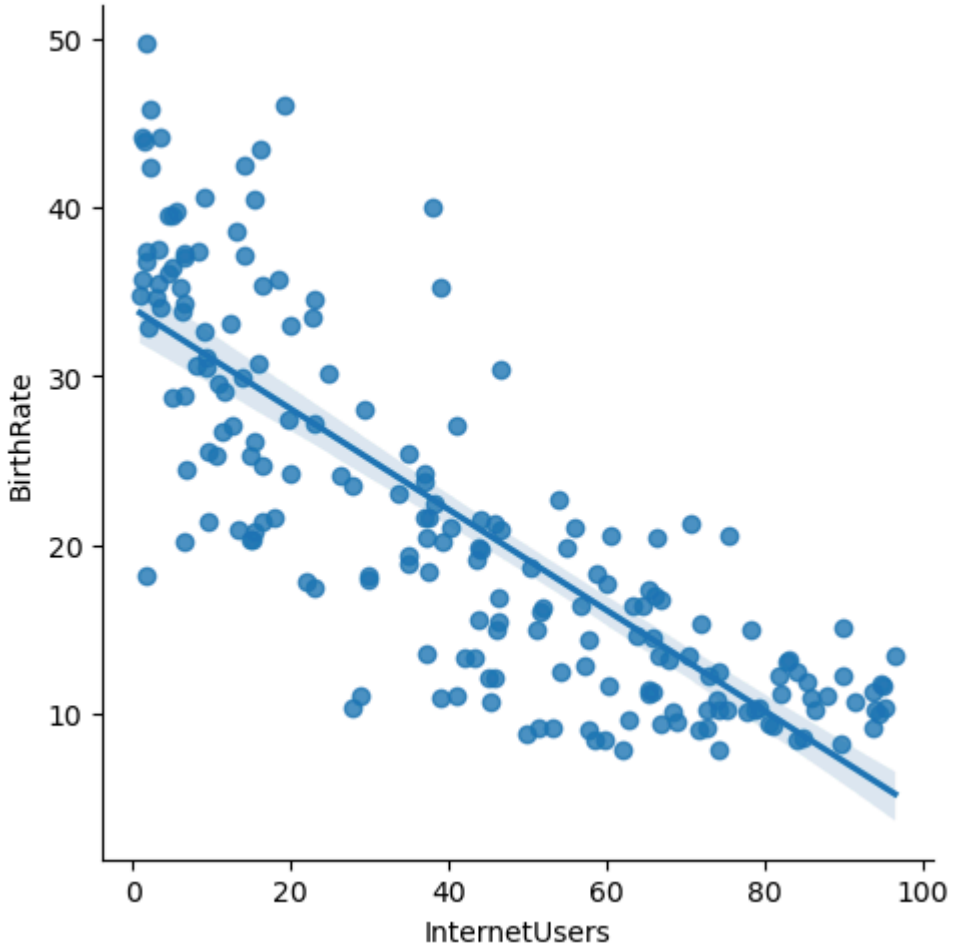


In [122…
```
vis5 = sns.lmplot(data=stats, x='InternetUsers', y='BirthRate')
plt.show()
```



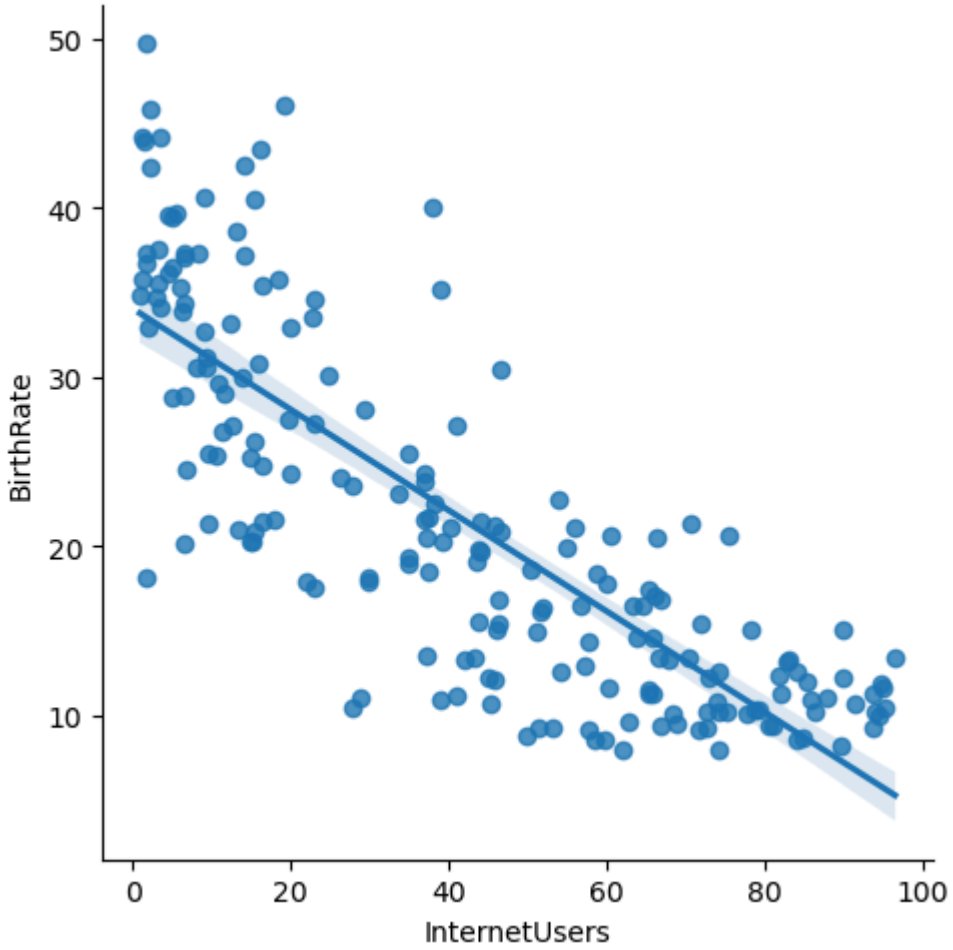
In [125…

```
vis5 = sns.lmplot(data=stats, x='InternetUsers', y='BirthRate',fit_reg=10)
plt.show()
```



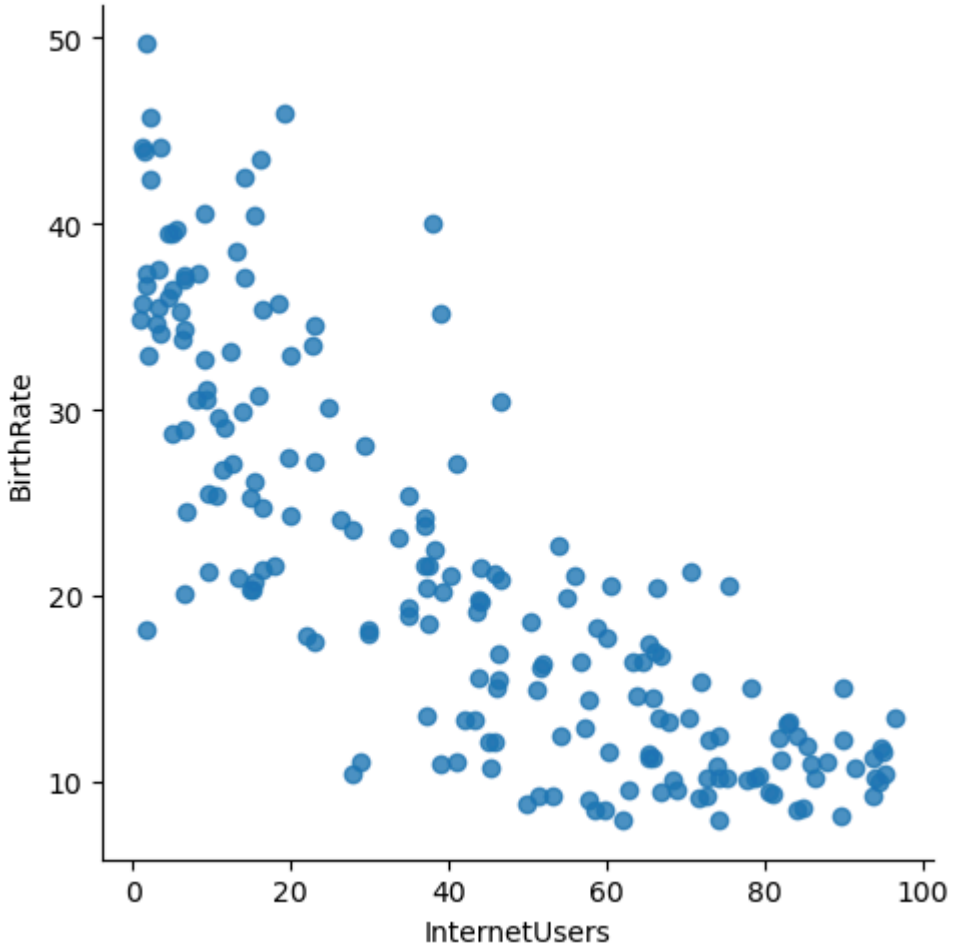
In [126…

```
vis5 = sns.lmplot(data=stats, x='InternetUsers', y='BirthRate',fit_reg=15)
plt.show()
```



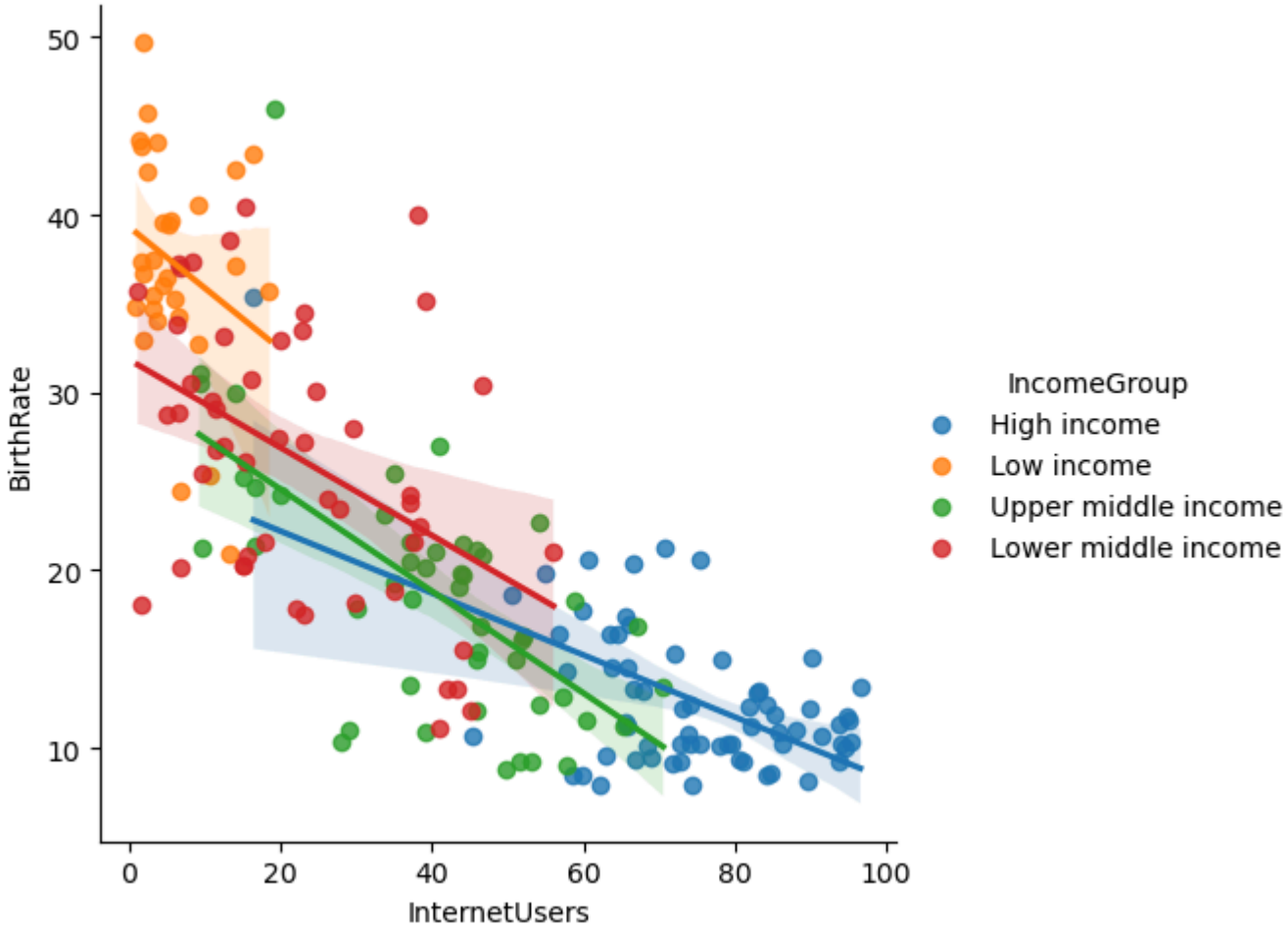
In [123…

```
vis5 = sns.lmplot(data=stats, x='InternetUsers', y='BirthRate',fit_reg=False)
plt.show()
```



In [124…

```
vis5 = sns.lmplot(data=stats, x='InternetUsers', y='BirthRate',fit_reg=True,hue='IncomeGroup')
plt.show()
```


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

In [ ]: