

# 05/11/25

In [2]: `import pandas as pd`

In [3]: `stats=pd.read_excel(r'C:\Users\pandu\OneDrive\Documents\Desktop\NIT\data.xlsx')`  
`stats`

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]: `# to get the rows`  
`len(stats)`

Out[4]: 195

In [5]: `# to get the columns`  
`stats.columns`

Out[5]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup'], dtype='object')`

In [6]: `len(stats.columns)`

Out[6]: 5

In [7]: `# to get rows and columns`  
`stats.shape`

Out[7]: (195, 5)

In [8]: `# to know is there any null values are there or not  
stats.isnull()`

Out[8]:

	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 [9]: `stats.isna()`

Out[9]:

	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 [10]: # to get how many null values are there in dataset  
stats.isnull().sum()
```

```
Out[10]: CountryName      0  
CountryCode       0  
BirthRate        0  
InternetUsers    0  
IncomeGroup       0  
dtype: int64
```

```
In [11]: stats.dtypes
```

```
Out[11]: CountryName      object  
CountryCode       object  
BirthRate        float64  
InternetUsers    float64  
IncomeGroup       object  
dtype: object
```

```
In [12]: 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 [13]: type(stats)
```

```
Out[13]: pandas.core.frame.DataFrame
```

```
In [14]: stats.head()
```

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

	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 [16]: `stats.head(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

In [17]: `stats.tail(3)`

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
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 [18]: `stats.columns`

Out[18]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup'],  
 dtype='object')`

In [19]: `# to get 1 column  
stats['CountryName']`

Out[19]: 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 [20]: stats['BirthRate']
```

```
Out[20]: 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 [21]: # to get 2 columns
stats[['CountryName', 'CountryCode']]
```

```
Out[21]:      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 [22]: # to get 3 columns
stats[['CountryName', 'CountryCode', 'BirthRate']]
```

Out[22]:

	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

## lets split the data to numerical data and categorical data

In [24]: `stats.columns`Out[24]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup'], dtype='object')`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[['CountryName', 'CountryCode', 'IncomeGroup']]`

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

Out[28]:

	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

In [29]: `print(stats.shape)`  
`print(stats_numeric_data.shape)`  
`print(stats_categorical_data.shape)`

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

## slicing in pandas

In [31]: `stats`

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
...	...	...	...	...	...
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 [32]: `stats[:6]`

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
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 [33]:

stats[:3]

Out[33]:

	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 [34]:

stats[3:]

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
...	...	...	...	...	...
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 [35]:

stats[3:10]

Out[35]:

	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 [36]:

stats[3:55]

Out[36]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
3	Albania	ALB	12.877	57.200000	Upper middle income
4	United Arab Emirates	ARE	11.044	88.000000	High income
5	Argentina	ARG	17.716	59.900000	High income
6	Armenia	ARM	13.308	41.900000	Lower middle income
7	Antigua and Barbuda	ATG	16.447	63.400000	High income
8	Australia	AUS	13.200	83.000000	High income
9	Austria	AUT	9.400	80.618800	High income
10	Azerbaijan	AZE	18.300	58.700000	Upper middle income
11	Burundi	BDI	44.151	1.300000	Low income
12	Belgium	BEL	11.200	82.170200	High income
13	Benin	BEN	36.440	4.900000	Low income
14	Burkina Faso	BFA	40.551	9.100000	Low income
15	Bangladesh	BGD	20.142	6.630000	Lower middle income
16	Bulgaria	BGR	9.200	53.061500	Upper middle income
17	Bahrain	BHR	15.040	90.000040	High income
18	Bahamas, The	BHS	15.339	72.000000	High income
19	Bosnia and Herzegovina	BIH	9.062	57.790000	Upper middle income
20	Belarus	BLR	12.500	54.170000	Upper middle income
21	Belize	BLZ	23.092	33.600000	Upper middle income
22	Bermuda	BMU	10.400	95.300000	High income
23	Bolivia	BOL	24.236	36.940000	Lower middle income
24	Brazil	BRA	14.931	51.040000	Upper middle income
25	Barbados	BRB	12.188	73.000000	High income
26	Brunei Darussalam	BRN	16.405	64.500000	High income
27	Bhutan	BTN	18.134	29.900000	Lower middle income
28	Botswana	BWA	25.267	15.000000	Upper middle income
29	Central African Republic	CAF	34.076	3.500000	Low income
30	Canada	CAN	10.900	85.800000	High income
31	Switzerland	CHE	10.200	86.340000	High income
32	Chile	CHL	13.385	66.500000	High income

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
33	China	CHN	12.100	45.800000	Upper middle income
34	Cote d'Ivoire	CIV	37.320	8.400000	Lower middle income
35	Cameroon	CMR	37.236	6.400000	Lower middle income
36	Congo, Rep.	COG	37.011	6.600000	Lower middle income
37	Colombia	COL	16.076	51.700000	Upper middle income
38	Comoros	COM	34.326	6.500000	Low income
39	Cabo Verde	CPV	21.625	37.500000	Lower middle income
40	Costa Rica	CRI	15.022	45.960000	Upper middle income
41	Cuba	CUB	10.400	27.930000	Upper middle income
42	Cayman Islands	CYM	12.500	74.100000	High income
43	Cyprus	CYP	11.436	65.454800	High income
44	Czech Republic	CZE	10.200	74.110400	High income
45	Germany	DEU	8.500	84.170000	High income
46	Djibouti	DJI	25.486	9.500000	Lower middle income
47	Denmark	DNK	10.000	94.629700	High income
48	Dominican Republic	DOM	21.198	45.900000	Upper middle income
49	Algeria	DZA	24.738	16.500000	Upper middle income
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

In [37]: `stats[3:55:5]`

Out[37]:

	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
53	Spain	ESP	9.100	71.6350	High income

In [38]:

stats[:::-1]

Out[38]:

	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 [39]:

stats[:::-2]

Out[39]:

	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 [40]:

stats[50:100]

Out[40]:

	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 [41]: stats[0:200]

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

## 06/11/25

In [43]: `pd.__version__`

Out[43]: '2.2.2'

In [44]: `stats.columns`Out[44]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup'], dtype='object')`In [46]: `stats_numeric_data.head()`Out[46]: `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 [48]: `stats_categorical_data.head()`

Out[48]:

	CountryName	CountryCode	IncomeGroup
<b>0</b>	Aruba	ABW	High income
<b>1</b>	Afghanistan	AFG	Low income
<b>2</b>	Angola	AGO	Upper middle income
<b>3</b>	Albania	ALB	Upper middle income
<b>4</b>	United Arab Emirates	ARE	High income

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

Out[50]:

	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
<b>2</b>	Angola	AGO	45.985	19.1	Upper middle income
<b>3</b>	Albania	ALB	12.877	57.2	Upper middle income
<b>4</b>	United Arab Emirates	ARE	11.044	88.0	High income

In [52]: `stats.describe()`

Out[52]:

	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 [54]: `stats.describe().transpose()`

Out[54]:

	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 [60]: `stats_numeric_data.describe()`

Out[60]:

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

Out[62]:

	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 [64]: `stats_categorical_data.describe()`

Out[64]:

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

In [66]: `stats.columns`Out[66]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup'], dtype='object')`In [76]: `# adding a new column``stats['myCalc']=stats['BirthRate'] * stats['InternetUsers']`In [78]: `stats`

Out[78]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	muCalc	myCalc
0	Aruba	ABW	10.244	78.9	High income	808.2516	808.25
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927	207.99
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135	878.31
3	Albania	ALB	12.877	57.2	Upper middle income	736.5644	736.56
4	United Arab Emirates	ARE	11.044	88.0	High income	971.8720	971.87
...	...	...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income	658.9400	658.94
191	South Africa	ZAF	20.850	46.5	Upper middle income	969.5250	969.52
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income	93.2668	93.26
193	Zambia	ZMB	40.471	15.4	Lower middle income	623.2534	623.25
194	Zimbabwe	ZWE	35.715	18.5	Low income	660.7275	660.72

195 rows × 7 columns



In [80]: `stats.columns`

Out[80]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup', 'muCalc', 'myCalc'], dtype='object')`

In [82]: `len(stats.columns)`

Out[82]: 7

In [88]: `# deleting a column`  
`stats=stats.drop('myCalc', axis=1)`

In [90]: `stats`

Out[90]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	muCalc
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 [92]:

```
stats=stats.drop('muCalc',axis=1)
stats
```

Out[92]:

	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 [94]: `stats.columns`Out[94]: `Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup'], dtype='object')`In [96]: `len(stats.columns)`

Out[96]: 5

In [98]: `stats['InternetUsers'] < 2`Out[98]: 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

## filtering in pandas

In [104...]: `stats[stats['InternetUsers'] < 2]`

Out[104...]

	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 [106...]

len(stats[stats['InternetUsers'] &lt; 2])

Out[106...]

9

In [112...]

stats[stats['BirthRate'] &gt; 40]

Out[112...]

	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 [116...]

len(stats[stats['BirthRate'] &gt; 40])

Out[116...]

12

In [122...]

stats[(stats.BirthRate &gt; 40) &amp; (stats.InternetUsers &lt; 2)]

Out[122...]

	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 [132...]

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

Out[132...]

	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 [130...]

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

Out[130...]

	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 [134... stats.IncomeGroup.unique()
```

```
Out[134... array(['High income', 'Low income', 'Upper middle income',
       'Lower middle income'], dtype=object)
```

```
In [136... stats.IncomeGroup.nunique()
```

```
Out[136... 4
```

```
In [146... import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
plt.rcParams['figure.figsize']=6,2
import warnings
warnings.filterwarnings('ignore') # os error
```

```
In [148... stats.columns
```

```
Out[148... Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
       dtype='object')
```

```
In [150... stats['InternetUsers']
```

```
Out[150... 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 [170... min(stats['InternetUsers'])
```

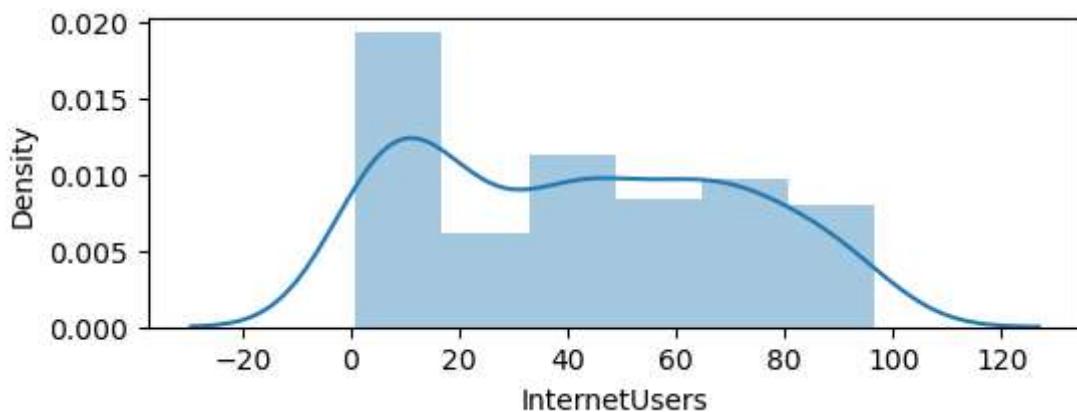
```
Out[170... 0.9
```

```
In [172... max(stats['InternetUsers'])
```

```
Out[172... 96.5468
```

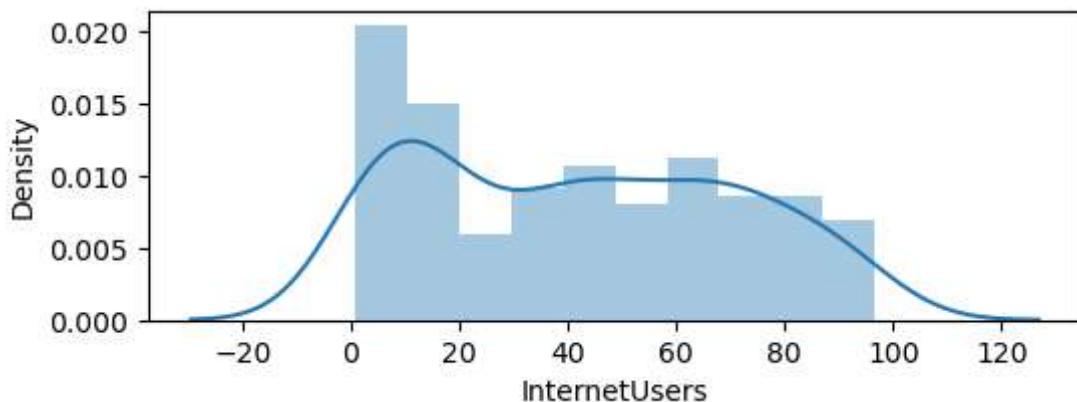
## univariate analysis (plot the graph using 1 variable)

```
In [188... vis1=sns.distplot(stats['InternetUsers'])
plt.show()
```



In [ ]:

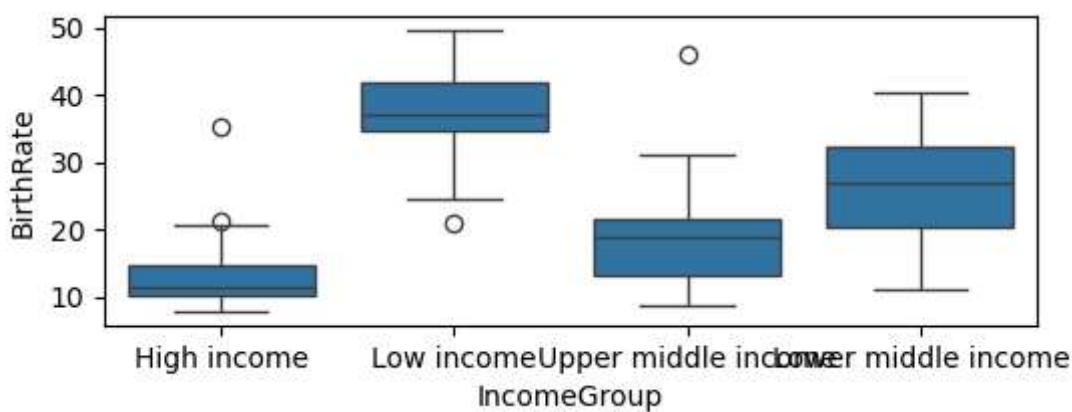
```
vis3=sns.distplot(stats['InternetUsers'],bins=10)
plt.show()
```



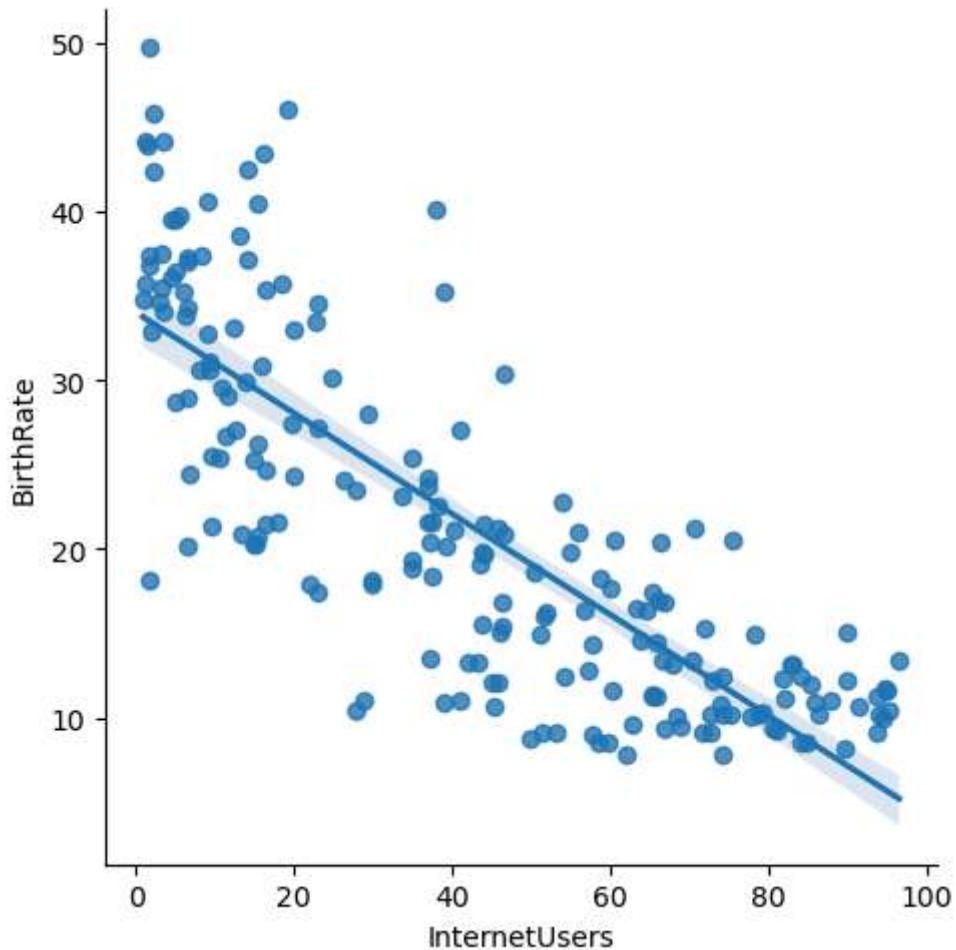
## Bivariate analysis (plot the graph using 2 variable)

In [194...]

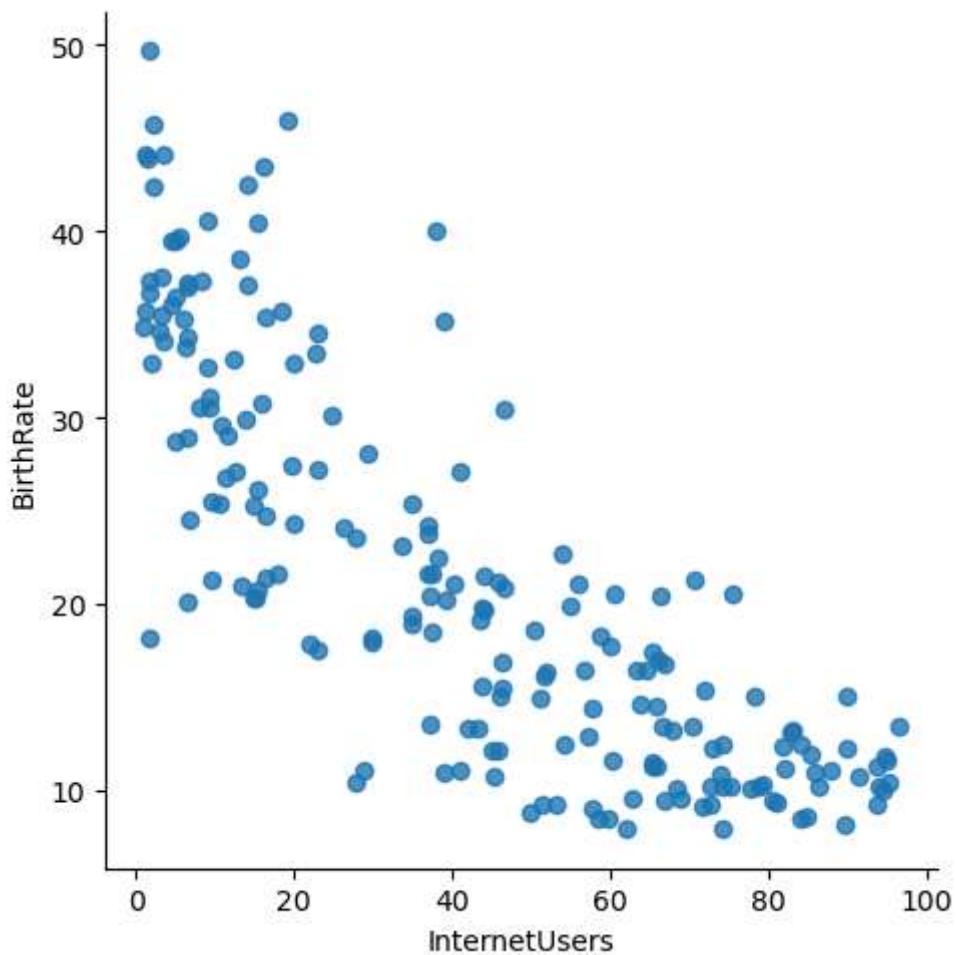
```
# Box Plot
vis4=sns.boxplot(data=stats,x='IncomeGroup',y='BirthRate')
plt.show()
```



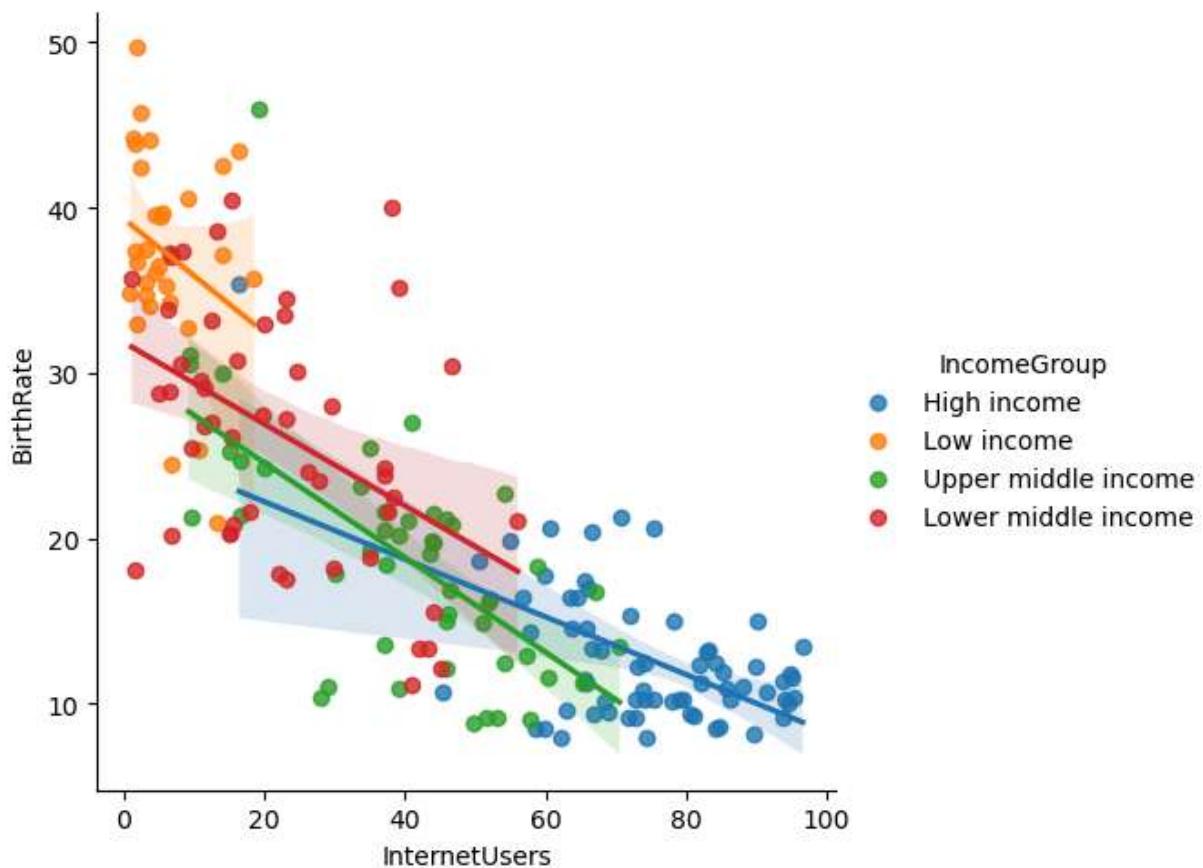
```
In [200... vis5=sns.lmplot(data=stats,x='InternetUsers',y='BirthRate')  
plt.show()
```



```
In [202... vis5=sns.lmplot(data=stats,x='InternetUsers',y='BirthRate',fit_reg=False)  
plt.show()
```



```
In [206]: vis5=sns.lmplot(data=stats,x='InternetUsers',y='BirthRate',fit_reg=True,hue='Income',  
plt.show()
```



In [ ]: