

```
In [3]: import pandas as pd
```

```
In [5]: df = pd.read_csv(r'C:\Users\arsha_4tjdyqj\Downloads\data.csv')
df
```

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 [7]: df.info
```

```
Out[7]: <bound method DataFrame.info of
InternetUsers \>

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

   IncomeGroup
0      High income
1      Low income
2 Upper middle income
3 Upper middle income
4      High income
..           ...
190 Lower middle income
191 Upper middle income
192      Low income
193 Lower middle income
194      Low income

[195 rows x 5 columns]>
```

In [9]: `df.info()`

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

In [11]: `len(df)`

Out[11]: 195

In [13]: `df.columns`

```
Out[13]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
               'IncomeGroup'],
              dtype='object')
```

In [15]: `type(df)`

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

In [17]: len(df.columns)

Out[17]: 5

In [19]: df.head(2)

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

In [21]: df.head()

	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 [23]: df.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 [25]: df.tail(1)

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
194	Zimbabwe	ZWE	35.715	18.5	Low income

In [27]: df[:]

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 [29]:

df[:::-1]

Out[29]:

	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 [31]:

df[:5]

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

In [33]:

df[0:200:10]

Out[33]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.900000	High income
10	Azerbaijan	AZE	18.300	58.700000	Upper middle income
20	Belarus	BLR	12.500	54.170000	Upper middle income
30	Canada	CAN	10.900	85.800000	High income
40	Costa Rica	CRI	15.022	45.960000	Upper middle income
50	Ecuador	ECU	21.070	40.353684	Upper middle income
60	Gabon	GAB	30.555	9.200000	Upper middle income
70	Greenland	GRL	14.500	65.800000	High income
80	India	IND	20.291	15.100000	Lower middle income
90	Kazakhstan	KAZ	22.730	54.000000	Upper middle income
100	Libya	LBY	21.425	16.500000	Upper middle income
110	Moldova	MDA	12.141	45.000000	Lower middle income
120	Mozambique	MOZ	39.705	5.400000	Low income
130	Netherlands	NLD	10.200	93.956400	High income
140	Poland	POL	9.600	62.849200	High income
150	Sudan	SDN	33.477	22.700000	Lower middle income
160	Suriname	SUR	18.455	37.400000	Upper middle income
170	Tajikistan	TJK	30.792	16.000000	Lower middle income
180	Uruguay	URY	14.374	57.690000	High income
190	Yemen, Rep.	YEM	32.947	20.000000	Lower middle income

In [35]:

df

Out[35]:

	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]:

df.describe()

Out[37]:

	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 [39]:

df[6:]

Out[39]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
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
10	Azerbaijan	AZE	18.300	58.7000	Upper middle income
...
190	Yemen, Rep.	YEM	32.947	20.0000	Lower middle income
191	South Africa	ZAF	20.850	46.5000	Upper middle income
192	Congo, Dem. Rep.	COD	42.394	2.2000	Low income
193	Zambia	ZMB	40.471	15.4000	Lower middle income
194	Zimbabwe	ZWE	35.715	18.5000	Low income

189 rows × 5 columns

In [41]: df.describe().transpose()

Out[41]:

	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 [43]: df.head(2)

Out[43]:

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

In [45]: df.columns=['a', 'b', 'c', 'd', 'e']

In [47]: df.head(1)

Out[47]:

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

In [49]: df.columns

Out[49]: Index(['a', 'b', 'c', 'd', 'e'], dtype='object')

```
In [51]: df.columns=['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
                  'IncomeGroup']
```

```
In [53]: df.head(1)
```

```
Out[53]:   CountryName  CountryCode  BirthRate  InternetUsers  IncomeGroup
0          Aruba        ABW      10.244       78.9  High income
```

```
In [55]: df[:]
```

```
Out[55]:   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 [57]: df[0:5]
```

```
Out[57]:   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 [59]: df[:10]
```

Out[59]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.9000	High income
1	Afghanistan	AFG	35.253	5.9000	Low income
2	Angola	AGO	45.985	19.1000	Upper middle income
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 [65]:

`df.isnull().sum()`

Out[65]:

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

In [67]:

`df.isnull()`

Out[67]:

	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 [69]: `df.dtypes`

```
Out[69]: CountryName    object
CountryCode     object
BirthRate      float64
InternetUsers  float64
IncomeGroup     object
dtype: object
```

In [71]: `df.columns`

```
Out[71]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
               'IncomeGroup'],
              dtype='object')
```

In [73]: `df_categorical=df[['CountryName', 'CountryCode', 'IncomeGroup']]`
`df_categorical.head()`

	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 [75]: `df.describe()`

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

In [77]: `df_categorical.describe()`

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

In [79]: df.InternetUsers<2

Out[79]: 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 [81]: Filter=df.InternetUsers<2

In [83]: 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 [85]: df

Out[85]:

	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 [87]: Filter2=df.BirthRate>40

In [89]: Filter2

Out[89]: 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 [91]: 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 [93]: import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

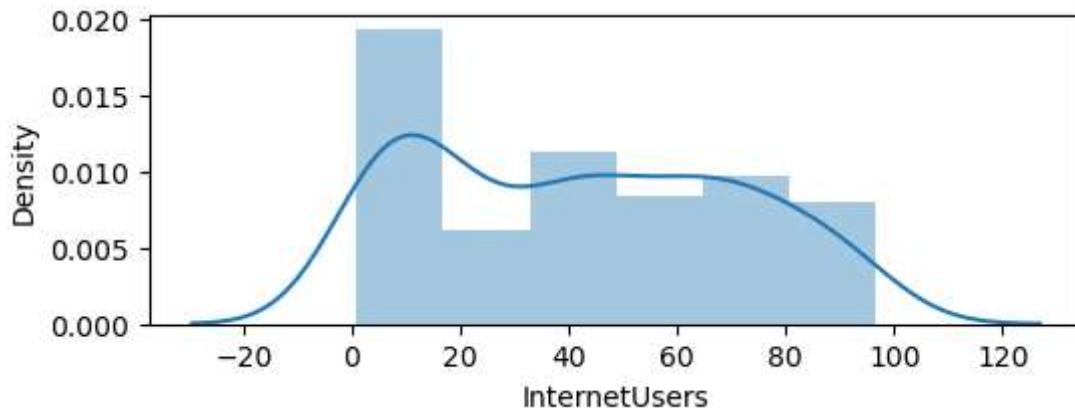
```
plt.rcParams['figure.figsize']=6,2
import warnings
warnings.filterwarnings('ignore')
```

In [94]: `df.head()`

Out[94]:

	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 [95]: `vis1=sns.distplot(df["InternetUsers"])`



In [96]: `vis2=sns.distplot(df["InternetUsers"])`

