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

```
In [2]:  df = pd.read_csv(r'data.csv')
```

```
In [4]:  df.head()
```

Out[4]:

	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 [5]:  df.columns
```

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

```
In [6]:  df['InternetUsers']
```

```
Out[6]:  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 [7]:  df_cat = df[['CountryName', 'CountryCode','IncomeGroup']]
         df_cat
```

Out[7]:

	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 [8]:
```python
df_num = df[['BirthRate', 'InternetUsers']]
df_num
```

Out[8]:

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

195 rows × 2 columns

In [9]:
```python
print(df.shape)
print(df_cat.shape)
print(df_num.shape)
```

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

In [10]: `df.describe()`

Out[10]:

	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 [11]: `df.describe().transpose()`

Out[11]:

	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 [12]: `df.columns = ['a','b','c','d','e']`

In [13]: `df.head(5)`

Out[13]:

	a	b	c	d	e
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 [14]: `df.columns = ['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers','IncomeGro`

In [15]: `df.head(5)`

Out[15]:

	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 [16]:
```
df.BirthRate * df.InternetUsers
```

Out[16]:
```
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 [17]:
```
df['myCalc'] = df.BirthRate * df.InternetUsers
```

In [18]:
```
df = df.drop('myCalc',axis = 1)
```

In [19]:
```
df['InternetUsers'] < 2
```

Out[19]:
```
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 [20]:
```
len(df[df['InternetUsers'] < 2])
```

Out[20]:
```
9
```

In [21]:
```
filter_1 = df[df['InternetUsers'] < 2]
filter_1
```

Out[21]:

	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 [22]:
```python
filter_2 = df[df['BirthRate'] > 40]
filter_2
```

Out[22]:

	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 [23]:
```python
df[(df.BirthRate > 40) & (df.InternetUsers < 2)]
```

Out[23]:

	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 [24]: `df[df.IncomeGroup == 'High income'].head()`

Out[24]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
0	Aruba	ABW	10.244	78.9	High income
4	United Arab Emirates	ARE	11.044	88.0	High income
5	Argentina	ARG	17.716	59.9	High income
7	Antigua and Barbuda	ATG	16.447	63.4	High income
8	Australia	AUS	13.200	83.0	High income

In [25]: `df.IncomeGroup.unique()`

Out[25]:
```
array(['High income', 'Low income', 'Upper middle income',
       'Lower middle income'], dtype=object)
```

In [26]: `df.IncomeGroup.nunique()`

Out[26]: 4

In [27]:
```python
import matplotlib.pyplot as plt
import seaborn as sns

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

import warnings
warnings.filterwarnings('ignore')
```
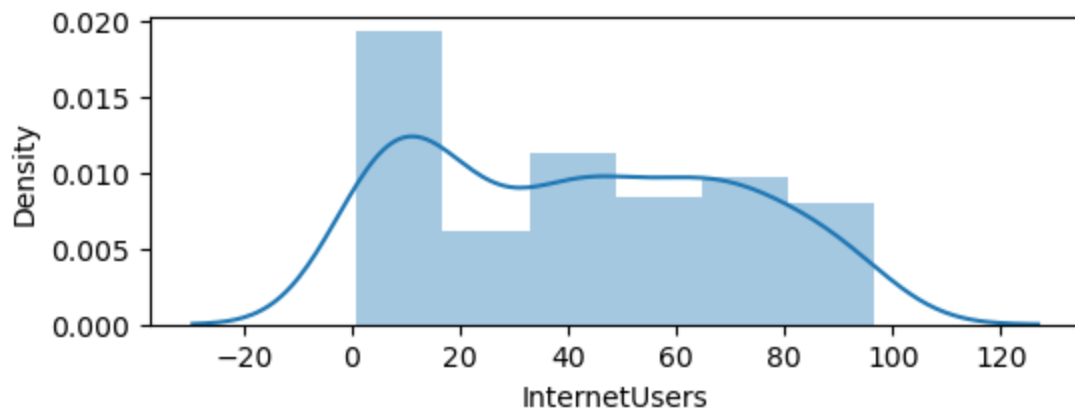
In [28]: `df["InternetUsers"]`

Out[28]:
```
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 [29]:
```python
vis1 = sns.distplot(df["InternetUsers"])
plt.show()
```


In [ ]: