

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

```
In [2]: pd.__version__
```

```
Out[2]: '2.2.2'
```

```
In [3]: df = pd.read_csv(r"C:\Users\Lenovo\Downloads\dataset.csv")
```

```
In [4]: df
```

```
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
...	...	...	...	...	...
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 [5]: id(df)
```

```
Out[5]: 1232861841456
```

```
In [6]: len(df)
```

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

```
In [7]: df.columns
```

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

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

Out[8]: 5

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

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]: df.isna()
```

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]: df.isnull().sum()
```

```
Out[11]: CountryName      0
CountryCode      0
BirthRate        0
InternetUsers     0
IncomeGroup       0
dtype: int64
```

```
In [12]: df.isna().sum()
```

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

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

```
Out[13]:
```

	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 [14]: df.tail()
```

```
Out[14]:
```

	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 [15]: 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 [16]: df[:]
```

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]: df[1:10]

Out[17]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
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 [18]: df[100:195]

Out[18]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
<b>100</b>	Libya	LBY	21.425	16.5	Upper middle income
<b>101</b>	St. Lucia	LCA	15.430	46.2	Upper middle income
<b>102</b>	Liechtenstein	LIE	9.200	93.8	High income
<b>103</b>	Sri Lanka	LKA	17.863	21.9	Lower middle income
<b>104</b>	Lesotho	LSO	28.738	5.0	Lower middle income
...	...	...	...	...	...
<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

95 rows × 5 columns

In [19]:

df[1:190:10]

Out[19]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup
<b>1</b>	Afghanistan	AFG	35.253	5.9000	Low income
<b>11</b>	Burundi	BDI	44.151	1.3000	Low income
<b>21</b>	Belize	BLZ	23.092	33.6000	Upper middle income
<b>31</b>	Switzerland	CHE	10.200	86.3400	High income
<b>41</b>	Cuba	CUB	10.400	27.9300	Upper middle income
<b>51</b>	Egypt, Arab Rep.	EGY	28.032	29.4000	Lower middle income
<b>61</b>	United Kingdom	GBR	12.200	89.8441	High income
<b>71</b>	Guatemala	GTM	27.465	19.7000	Lower middle income
<b>81</b>	Ireland	IRL	15.000	78.2477	High income
<b>91</b>	Kenya	KEN	35.194	39.0000	Lower middle income
<b>101</b>	St. Lucia	LCA	15.430	46.2000	Upper middle income
<b>111</b>	Madagascar	MDG	34.686	3.0000	Low income
<b>121</b>	Mauritania	MRT	33.801	6.2000	Lower middle income
<b>131</b>	Norway	NOR	11.600	95.0534	High income
<b>141</b>	Puerto Rico	PRI	10.800	73.9000	High income
<b>151</b>	Senegal	SEN	38.533	13.1000	Lower middle income
<b>161</b>	Slovak Republic	SVK	10.100	77.8826	High income
<b>171</b>	Turkmenistan	TKM	21.322	9.6000	Upper middle income
<b>181</b>	United States	USA	12.500	84.2000	High income

In [20]: `df.head(2)`

Out[20]:

	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 [21]: `df.describe()`

Out[21]:

	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 [22]: `df['CountryName']`

```
Out[22]: 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 [23]: `df['IncomeGroup']`

```
Out[23]: 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 [24]: `df['BirthRate']`

```
Out[24]: 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 [25]: df[['CountryName', 'CountryCode', 'BirthRate']]
```

```
Out[25]:
```

	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

```
In [30]: df_cat = df[['CountryName', 'CountryCode', 'IncomeGroup']]
         df_cat
```



Out[30]:

	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 [31]: print(len(df.columns))
         print(len(df_cat.columns))
```

5

3

```
In [32]: print(df.columns)
```

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

```
In [33]: print(df_cat.columns)
```

```
Index(['CountryName', 'CountryCode', 'IncomeGroup'], dtype='object')
```

```
In [34]: df_cat.describe()
```

Out[34]:

	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 [35]: df_num = df[['BirthRate', 'InternetUsers']]
         df_num
```

Out[35]:

	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 [36]: `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 [37]: `df_cat.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 195 entries, 0 to 194
Data columns (total 3 columns):
#   Column          Non-Null Count  Dtype
---  -
0   CountryName     195 non-null   object
1   CountryCode     195 non-null   object
2   IncomeGroup     195 non-null   object
dtypes: object(3)
memory usage: 4.7+ KB
```

In [38]: `df_num.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 195 entries, 0 to 194
Data columns (total 2 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   BirthRate       195 non-null    float64
1   InternetUsers   195 non-null    float64
dtypes: float64(2)
memory usage: 3.2 KB
```

In [39]: `df.describe()`

Out[39]:

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

Out[40]:

	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 [41]: `df.describe().T`

Out[41]:

	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 [42]: `df.columns`

Out[42]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGroup'], dtype='object')

In [43]: `df.columns = ['A', 'B', 'C', 'D', 'E']`  
`df.columns`

Out[43]: Index(['A', 'B', 'C', 'D', 'E'], dtype='object')

In [45]: `df.columns = ['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers', 'IncomeGro`  
`df.columns`

```
Out[45]: Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
              'IncomeGroup'],
              dtype='object')
```

```
In [46]: df.head(1)
```

```
Out[46]:
```

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

```
In [48]: df[['CountryCode', 'BirthRate', 'InternetUsers']][4:8]
```

```
Out[48]:
```

	CountryCode	BirthRate	InternetUsers
4	ARE	11.044	88.0
5	ARG	17.716	59.9
6	ARM	13.308	41.9
7	ATG	16.447	63.4

```
In [49]: df[4:8][['CountryCode', 'BirthRate', 'InternetUsers']]
```

```
Out[49]:
```

	CountryCode	BirthRate	InternetUsers
4	ARE	11.044	88.0
5	ARG	17.716	59.9
6	ARM	13.308	41.9
7	ATG	16.447	63.4

```
In [50]: df.BirthRate * df.InternetUsers
```

```
Out[50]:
```

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 [51]: df['Prod'] = df.BirthRate * df.InternetUsers
```

```
In [52]: df.head()
```

Out[52]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Prod
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

In [54]: `df = df.drop('Prod', axis = 1)`

In [55]: `df.head()`

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

In [57]: `df[df.InternetUsers<2]`

Out[57]:

	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 [58]: `len(df[df.InternetUsers<2])`

Out[58]: 9

In [59]: `df.BirthRate>40`

```
Out[59]: 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 [60]: len(df.BirthRate>40)
```

```
Out[60]: 195
```

```
In [61]: Filter1 = df.InternetUsers<2
```

```
In [62]: Filter2 = df.BirthRate>40
```

```
In [63]: df[Filter1 & Filter2]
```

```
Out[63]:
```

	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 [ ]:
```