**Dataframe in python and how to import dataset**

**17july**

In [1]:

**import** pandas **as** pd

In [2]: stats**=** pd**.**read\_csv(r'C:\Users\nithi\OneDrive\Desktop\full stack data science & ai with chatgpt\JULY\17th,18th\data.csv')

# In [3]:

stats

Out[3]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [4]:

*# Explore data in python*

*#1. Full dataframe*

*#2. How many rows & columns. you have to chk the row becuase the no. of rows should matched with client data*

len(stats) *#195 rows imported (this is for tracking later part )*

Out[4]:195

In [5]:

stats**.**shape *#dimension of data*

Out[5]:(195, 5)

In [7]:

stats**.**columns *#columns of dataset*

Out[7]:Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',

'IncomeGroup'], dtype='object')

In [8]:

type(stats)

Out[8]:pandas.core.frame.DataFrame

In [11]: stats**.**info() *#info() is used to find missing values and datatypes of attribute & information of datset*

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 195 entries, 0 to 194

Data columns (total 5 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

1. CountryName 195 non-null object
2. CountryCode 195 non-null object
3. BirthRate 195 non-null float64
4. InternetUsers 195 non-null float64 4 IncomeGroup 195 non-null object dtypes: float64(2), object(3) memory usage: 7.7+ KB

In [12]:

stats**.**head()

Out[12]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

* 1. Aruba ABW 10.244 78.9 High income
  2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

## **2** Angola AGO 45.985 19.1

income

### Upper middle

**3** Albania ALB 12.877 57.2

income

United Arab

## **4** ARE 11.044 88.0 High income

Emirates

In [13]:

stats**.**head(10)

Out[13]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9000 High income
2. Afghanistan AFG 35.253 5.9000 Low income

Upper middle

1. Angola AGO 45.985 19.1000

income

Upper middle

1. Albania ALB 12.877 57.2000

income

United Arab

1. ARE 11.044 88.0000 High income

Emirates

1. Argentina ARG 17.716 59.9000 High income

Lower middle

1. Armenia ARM 13.308 41.9000

income

1. Antigua and Barbuda ATG 16.447 63.4000 High income
2. Australia AUS 13.200 83.0000 High income
3. Austria AUT 9.400 80.6188 High income

In [14]:

stats**.**tail()

Out[14]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

### Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

Congo, Dem.

1. COD 42.394 2.2 Low income

Rep.

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income

In [15]:

stats**.**tail(10)

Out[15]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Virgin Islands (U.S.) VIR 10.700 45.3 High income

Lower middle

1. Vietnam VNM 15.537 43.9

income

Lower middle

1. Vanuatu VUT 26.739 11.3

income

West Bank and Lower middle

1. PSE 30.394 46.6

Gaza income

Lower middle

1. Samoa WSM 26.172 15.3

income

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income

In [17]:

stats**.**describe() *# descriptive statistical information of dataset*

*#stats.describe . DESCRIBE WILL GIVE ONLY NUMERICAL INFORMATION*

|  |  |  |
| --- | --- | --- |
|  | **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 |

Out[17]:

In [18]:

stats**.**describe()**.**transpose() *#converts col to rows by transpose*

Out[18]: **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

**Renaming columns of data**

In [19]:

stats**.**head(6)

Out[19]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

1. Argentina ARG 17.716 59.9 High income

In [21]:

stats**.**info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 195 entries, 0 to 194

Data columns (total 5 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

1. CountryName 195 non-null object
2. CountryCode 195 non-null object
3. BirthRate 195 non-null float64
4. InternetUsers 195 non-null float64 4 IncomeGroup 195 non-null object dtypes: float64(2), object(3) memory usage: 7.7+ KB

In [22]:

stats**.**columns

Out[22]:Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',

'IncomeGroup'], dtype='object')

In [23]: stats**.**columns **=** ['a','b','c','d','e']

In [24]:

stats**.**head()

Out[24]: **a b c d e**

* 1. Aruba ABW 10.244 78.9 High income
  2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

* 1. Angola AGO 45.985 19.1

income

Upper middle

* 1. Albania ALB 12.877 57.2

income

United Arab

* 1. ARE 11.044 88.0 High income

Emirates

In [25]:

stats**.**columns

Out[25]:Index(['a', 'b', 'c', 'd', 'e'], dtype='object')

In [27]:

stats**.**info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 195 entries, 0 to 194 Data columns (total 5 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

1. a 195 non-null object
2. b 195 non-null object
3. c 195 non-null float64
4. d 195 non-null float64 4 e 195 non-null object dtypes: float64(2), object(3) memory usage: 7.7+ KB

In [28]: stats**.**columns **=** ['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers','IncomeGroup']

In [29]:

stats**.**head()

Out[29]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

* 1. Aruba ABW 10.244 78.9 High income
  2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

* 1. Angola AGO 45.985 19.1

income

Upper middle

* 1. Albania ALB 12.877 57.2

income

United Arab

* 1. ARE 11.044 88.0 High income

Emirates

In [30]:

stats**.**columns

Out[30]:Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',

'IncomeGroup'], dtype='object')

In [31]:

*# subsetting a dataframes in pandas*

*#1. Rows*

*#2. Columns*

*#3. combine the two*

ROWS based subsetting python knows with index

In [32]:

stats[:]

Out[32]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [33]:

stats[2:100]

Out[33]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

Upper middle

1. Angola AGO 45.985 19.10

income

Upper middle

1. Albania ALB 12.877 57.20

income

United Arab

1. ARE 11.044 88.00 High income

Emirates

1. Argentina ARG 17.716 59.90 High income

Lower middle

1. Armenia ARM 13.308 41.90

income

**...** ... ... ... ... ...

1. Korea, Rep. KOR 8.600 84.77 High income
2. Kuwait KWT 20.575 75.46 High income

Lower middle

1. Lao PDR LAO 27.051 12.50

income

Upper middle

1. Lebanon LBN 13.426 70.50

income

1. Liberia LBR 35.521 3.20 Low income

98 rows × 5 columns

In [34]:

stats[::**-**1]

Out[34]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

**194** Zimbabwe ZWE 35.715 18.5 Low income

Lower middle

**193** Zambia ZMB 40.471 15.4

income

**192** Congo, Dem. Rep. COD 42.394 2.2 Low income

Upper middle

**191** South Africa ZAF 20.850 46.5

income

Lower middle

**190** Yemen, Rep. YEM 32.947 20.0

income

**...** ... ... ... ... ...

United Arab

**4** ARE 11.044 88.0 High income

Emirates

Upper middle

**3** Albania ALB 12.877 57.2

income

Upper middle

**2** Angola AGO 45.985 19.1

income

**1** Afghanistan AFG 35.253 5.9 Low income **0** Aruba ABW 10.244 78.9 High income

195 rows × 5 columns

In [35]:

stats[::**-**2]

Out[35]: **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

Lower middle

**190** Yemen, Rep. YEM 32.947 20.0

income

Lower middle

**188** West Bank and Gaza PSE 30.394 46.6

income

Lower middle

**186** Vietnam VNM 15.537 43.9

income

**...** ... ... ... ... ...

**8** Australia AUS 13.200 83.0 High income

Lower middle

**6** Armenia ARM 13.308 41.9

income

United Arab

**4** ARE 11.044 88.0 High income

Emirates

Upper middle

**2** Angola AGO 45.985 19.1

income

**0** Aruba ABW 10.244 78.9 High income

98 rows × 5 columns

In [36]: stats[::10]

Out[36]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

**0** Aruba ABW 10.244 78.900000 High income

Upper middle

**10** Azerbaijan AZE 18.300 58.700000

income

Upper middle

**20** Belarus BLR 12.500 54.170000

income

**30** Canada CAN 10.900 85.800000 High income

Upper middle

**40** Costa Rica CRI 15.022 45.960000

income

Upper middle

**50** Ecuador ECU 21.070 40.353684

income

Upper middle

**60** Gabon GAB 30.555 9.200000

income

**70** Greenland GRL 14.500 65.800000 High income

Lower middle

**80** India IND 20.291 15.100000

income

Upper middle

**90** Kazakhstan KAZ 22.730 54.000000

income

Upper middle

**100** Libya LBY 21.425 16.500000

income

Lower middle

**110** Moldova MDA 12.141 45.000000

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

Lower middle

**150** Sudan SDN 33.477 22.700000

income

Upper middle

**160** Suriname SUR 18.455 37.400000

income

Lower middle

**170** Tajikistan TJK 30.792 16.000000

income

**180** Uruguay URY 14.374 57.690000 High income

Lower middle

**190** Yemen, Rep. YEM 32.947 20.000000

income

COLUMNS

In [37]:

stats**.**columns

Out[37]:Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',

'IncomeGroup'], dtype='object')

In [38]:

stats

Out[38]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [39]:

stats**.**head()

Out[39]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

In [40]:

stats['IncomeGroup']

Out[40]:0 High income

1. Low income
2. Upper middle income
3. Upper middle income
4. High income

...

1. Lower middle income
2. Upper middle income
3. Low income
4. Lower middle income
5. Low income

Name: IncomeGroup, Length: 195, dtype: object

In [41]:

['CountryName', 'BirthRate']

Out[41]:['CountryName', 'BirthRate']

In [42]:

stats[['CountryName', 'BirthRate']]

Out[42]: **CountryName BirthRate**

* 1. Aruba 10.244
  2. Afghanistan 35.253
  3. Angola 45.985
  4. Albania 12.877

United Arab

* 1. 11.044

Emirates

**...** ... ...

1. Yemen, Rep. 32.947
2. South Africa 20.850
3. Congo, Dem. Rep. 42.394
4. Zambia 40.471
5. Zimbabwe 35.715
6. rows × 2 columns

In [43]:

stats['CountryName']**.**head()

Out[43]:0 Aruba

1. Afghanistan
2. Angola
3. Albania
4. United Arab Emirates

Name: CountryName, dtype: object

In [44]:

stats[4:8] *#rows subsetting*

Out[44]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

United Arab

1. ARE 11.044 88.0 High income

Emirates

1. Argentina ARG 17.716 59.9 High income

Lower middle

1. Armenia ARM 13.308 41.9

income

1. Antigua and Barbuda ATG 16.447 63.4 High income

In [45]:

stats[['CountryName','BirthRate']][4:8]

Out[45]: **CountryName BirthRate**

United Arab

1. 11.044

Emirates

1. Argentina 17.716
2. Armenia 13.308
3. Antigua and Barbuda 16.447

In [46]:

*#2nd variable for holding dataframe*

In [47]: df **=** stats[['CountryName','BirthRate']]

In [48]:

df

Out[48]: **CountryName BirthRate**

1. Aruba 10.244
2. Afghanistan 35.253
3. Angola 45.985
4. Albania 12.877

United Arab

1. 11.044

Emirates

**...** ... ...

1. Yemen, Rep. 32.947
2. South Africa 20.850
3. Congo, Dem. Rep. 42.394
4. Zambia 40.471
5. Zimbabwe 35.715
6. rows × 2 columns

In [49]:

df**.**head()

Out[49]: **CountryName BirthRate**

1. Aruba 10.244
2. Afghanistan 35.253
3. Angola 45.985
4. Albania 12.877

United Arab

1. 11.044

Emirates

In [51]:

*#3rd variable* df1 **=** stats[4:8] df1

Out[51]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

United Arab

1. ARE 11.044 88.0 High income

Emirates

1. Argentina ARG 17.716 59.9 High income

Lower middle

1. Armenia ARM 13.308 41.9

income

1. Antigua and Barbuda ATG 16.447 63.4 High income

In [52]: stats

Out[52]:

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [53]:

stats**.**head()

Out[53]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

In [54]:

*#subset dataframe*

In [55]:

stats[['CountryName','BirthRate','InternetUsers']][4:9]

Out[55]: **CountryName BirthRate InternetUsers**

United Arab

1. 11.044 88.0

Emirates

1. Argentina 17.716 59.9
2. Armenia 13.308 41.9
3. Antigua and Barbuda 16.447 63.4
4. Australia 13.200 83.0

In [56]:

stats**.**head()

Out[56]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

In [57]:

stats**.**CountryName

Out[57]:0 Aruba

1. Afghanistan
2. Angola
3. Albania4 United Arab Emirates

...

1. Yemen, Rep.
2. South Africa
3. Congo, Dem. Rep.
4. Zambia
5. Zimbabwe

Name: CountryName, Length: 195, dtype: object

In [58]:stats**.**BirthRate

Out[58]:0 10.244

1. 35.253
2. 45.985
3. 12.877
4. 11.044

...

1. 32.947
2. 20.850
3. 42.394
4. 40.471
5. 35.715

Name: BirthRate, Length: 195, dtype: float64

In [60]:

stats**.**BirthRate**\***stats**.**InternetUsers

Out[60]:0 808.2516

1. 207.9927
2. 878.3135
3. 736.5644
4. 971.8720

...

1. 658.9400
2. 969.5250
3. 93.2668
4. 623.2534
5. 660.7275

Length: 195, dtype: float64

In [61]:

stats

Out[61]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

* 1. Aruba ABW 10.244 78.9 High income
  2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

* 1. Angola AGO 45.985 19.1

income

Upper middle

* 1. Albania ALB 12.877 57.2

income

United Arab

* 1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [62]: stats['myCalc'] **=** stats**.**BirthRate**\***stats**.**InternetUsers

In [63]:

stats**.**head()

Out[63]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup myCalc**

1. Aruba ABW 10.244 78.9 High income 808.2516
2. Afghanistan AFG 35.253 5.9 Low income 207.9927

Upper middle

1. Angola AGO 45.985 19.1 878.3135

income

Upper middle

1. Albania ALB 12.877 57.2 736.5644

income

United Arab

1. ARE 11.044 88.0 High income 971.8720

Emirates

In [64]:

stats**.**columns

Out[64]:Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',

'IncomeGroup', 'myCalc'], dtype='object')

In [65]: len(stats**.**columns)

Out[65]:6

In [86]: stats**=**stats**.**drop('myCalc', axis**=**1)

In [87]:

stats

Out[87]:

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [88]:

stats**.**columns

Out[88]:Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',

'IncomeGroup'], dtype='object')

In [89]:

stats**.**columns[3:4]

Out[89]:Index(['InternetUsers'], dtype='object')

In [90]:

stats

Out[90]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [91]:

stats**.**InternetUsers**<**20

Out[91]:0 False

1. True
2. True
3. False
4. False

...

1. False
2. False
3. True
4. True
5. True

Name: InternetUsers, Length: 195, dtype: bool

In [92]:

Filter **=** stats**.**InternetUsers**<**20

Filter

Out[92]:0 False

1. True
2. True
3. False
4. False

...

1. False
2. False
3. True
4. True
5. True

Name: InternetUsers, Length: 195, dtype: bool

In [93]:

type(Filter)

Out[93]:pandas.core.series.Series

In [94]:

stats

Out[94]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

* 1. Aruba ABW 10.244 78.9 High income
  2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

* 1. Angola AGO 45.985 19.1

income

Upper middle

* 1. Albania ALB 12.877 57.2

income

United Arab

* 1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [95]:

stats[10:101]

Out[95]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

Upper middle

1. Azerbaijan AZE 18.300 58.7000

income

1. Burundi BDI 44.151 1.3000 Low income
2. Belgium BEL 11.200 82.1702 High income
3. Benin BEN 36.440 4.9000 Low income
4. Burkina Faso BFA 40.551 9.1000 Low income

**...** ... ... ... ... ...

1. Kuwait KWT 20.575 75.4600 High income

Lower middle

1. Lao PDR LAO 27.051 12.5000

income

Upper middle

1. Lebanon LBN 13.426 70.5000

income

1. Liberia LBR 35.521 3.2000 Low income

Upper middle

1. Libya LBY 21.425 16.5000

income

91 rows × 5 columns

In [96]:

stats[10:90:4]

Out[96]:

Upper middle

**10** Azerbaijan AZE 18.300 58.700000

income

**14** Burkina Faso BFA 40.551 9.100000 Low income

**18** Bahamas, The BHS 15.339 72.000000 High income

**22** Bermuda BMU 10.400 95.300000 High income

**26** Brunei Darussalam BRN 16.405 64.500000 High income

**30** Canada CAN 10.900 85.800000 High income

Lower middle

**34** Cote d'Ivoire CIV 37.320 8.400000

income

**38** Comoros COM 34.326 6.500000 Low income

**42** Cayman Islands CYM 12.500 74.100000 High income

Lower middle

**46** Djibouti DJI 25.486 9.500000

income

Upper middle

**50** Ecuador ECU 21.070 40.353684

income

**54** Estonia EST 10.300 79.400000 High income

**58** France FRA 12.300 81.919800 High income

Lower middle

**62** Georgia GEO 13.332 43.300000

income

**66** Guinea-Bissau GNB 37.503 3.100000 Low income

**70** Greenland GRL 14.500 65.800000 High income

Hong Kong SAR,

**74** HKG 7.900 74.200000 High income

China

**78** Hungary HUN 9.200 72.643900 High income

Upper middle

**82** Iran, Islamic Rep. IRN 17.900 29.950000

income

**86** Italy ITA 8.500 58.459300 High income

In [97]: Filter

Out[97]:0 False

1. True
2. True
3. False
4. False

...

1. False
2. False
3. True
4. True
5. True

Name: InternetUsers, Length: 195, dtype: bool

In [100]: len(Filter)

Out[100]:195

In [99]:

stats[Filter]

Out[99]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

* 1. Afghanistan AFG 35.253 5.9 Low income

Upper middle

* 1. Angola AGO 45.985 19.1

income

**11** Burundi BDI 44.151 1.3 Low income

1. Benin BEN 36.440 4.9 Low income
2. Burkina Faso BFA 40.551 9.1 Low income

**...** ... ... ... ... ...

Lower middle

**187** Vanuatu VUT 26.739 11.3

income

Lower middle

**189** Samoa WSM 26.172 15.3

income

Congo, Dem.

1. COD 42.394 2.2 Low income

Rep.

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income

64 rows × 5 columns

In [101]:len(stats[Filter])

Out[101]:64

In [102]:

stats

Out[102]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

**...** ... ... ... ... ...

Lower middle

1. Yemen, Rep. YEM 32.947 20.0

income

Upper middle

1. South Africa ZAF 20.850 46.5

income

1. Congo, Dem. Rep. COD 42.394 2.2 Low income

Lower middle

1. Zambia ZMB 40.471 15.4

income

1. Zimbabwe ZWE 35.715 18.5 Low income
2. rows × 5 columns

In [106]:

stats**.**BirthRate**>**30

Out[106]:0 False

1. True
2. True
3. False
4. False

...

1. True
2. False
3. True
4. True
5. True

Name: BirthRate, Length: 195, dtype: bool

In [110]:

Filter2 **=** stats**.**BirthRate**>**30

Filter2

Out[110]:0 False

1. True
2. True
3. False
4. False

...

1. True
2. False
3. True
4. True
5. True

Name: BirthRate, Length: 195, dtype: bool

In [111]:

stats[Filter2]

Out[111]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

* 1. Afghanistan AFG 35.253 5.90 Low income

Upper middle

* 1. Angola AGO 45.985 19.10

income

**11** Burundi BDI 44.151 1.30 Low income

1. Benin BEN 36.440 4.90 Low income
2. Burkina Faso BFA 40.551 9.10 Low income

Central African

**29** CAF 34.076 3.50 Low income

Republic

Lower middle

1. Cote d'Ivoire CIV 37.320 8.40

income

Lower middle

1. Cameroon CMR 37.236 6.40

income

Lower middle

1. Congo, Rep. COG 37.011 6.60

income

**38** Comoros COM 34.326 6.50 Low income

**52** Eritrea ERI 34.800 0.90 Low income

**CountryName CountryCode BirthRate InternetUsers IncomeGroup**

**55** Ethiopia ETH 32.925 1.90 Low income

Upper middle

**60** Gabon GAB 30.555 9.20

income

Lower middle

1. Ghana GHA 33.131 12.30

income

1. Guinea GIN 37.337 1.60 Low income
2. Gambia, The GMB 42.525 14.00 Low income
3. Guinea-Bissau GNB 37.503 3.10 Low income
4. Equatorial Guinea GNQ 35.362 16.40 High income

Upper middle

**83** Iraq IRQ 31.093 9.20

income

Lower middle

**91** Kenya KEN 35.194 39.00

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

1. Mozambique MOZ 39.705 5.40 Low income

Lower middle

1. Mauritania MRT 33.801 6.20

income

**123** Malawi MWI 39.459 5.05 Low income

1. Niger NER 49.661 1.70 Low income

Lower middle

1. Nigeria NGA 40.045 38.00

income

**148** Rwanda RWA 32.689 9.00 Low income

Lower middle

1. Sudan SDN 33.477 22.70

income

Lower middle

1. Senegal SEN 38.533 13.10

income

Lower middle

1. Solomon Islands SLB 30.578 8.00

income

1. Sierra Leone SLE 36.729 1.70 Low income

**156** Somalia SOM 43.891 1.50 Low income

1. South Sudan SSD 37.126 14.10 Low income

Lower middle

1. Sao Tome and Principe STP 34.537 23.00

income

Lower middle

**164** Swaziland SWZ 30.093 24.70

income

1. Chad TCD 45.745 2.30 Low income
2. Togo TGO 36.080 4.50 Low income

Lower middle

**170** Tajikistan TJK 30.792 16.00

income

Lower middle

**172** Timor-Leste TLS 35.755 1.10

income

1. Tanzania TZA 39.518 4.40 Low income
2. Uganda UGA 43.474 16.20 Low income

Lower middle

**188** West Bank and Gaza PSE 30.394 46.60

income

Lower middle

**190** Yemen, Rep. YEM 32.947 20.00

income

1. Congo, Dem. Rep. COD 42.394 2.20 Low income

Lower middle

1. Zambia ZMB 40.471 15.40

income

1. Zimbabwe ZWE 35.715 18.50 Low income

In [113]: len(stats[Filter2])

Out[113]:48

In [114]: len(Filter2)

Out[114]:195

In [115]:

Filter**&**Filter2

1. False
2. True
3. True
4. False
5. False

...

1. False
2. False
3. True
4. True
5. True

Length: 195, dtype: bool

In [116]:

stats[Filter **&** Filter2]

**CountryName CountryCode BirthRate InternetUsers IncomeGroup**

* 1. Afghanistan AFG 35.253 5.90 Low income

Upper middle

* 1. Angola AGO 45.985 19.10

income

**11** Burundi BDI 44.151 1.30 Low income

1. Benin BEN 36.440 4.90 Low income
2. Burkina Faso BFA 40.551 9.10 Low income

Central African

**29** CAF 34.076 3.50 Low income

Republic

Lower middle

1. Cote d'Ivoire CIV 37.320 8.40

income

Lower middle

1. Cameroon CMR 37.236 6.40

income

Lower middle

1. Congo, Rep. COG 37.011 6.60

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

Upper middle

**60** Gabon GAB 30.555 9.20

income

Lower middle

1. Ghana GHA 33.131 12.30

income

1. Guinea GIN 37.337 1.60 Low income
2. Gambia, The GMB 42.525 14.00 Low income
3. Guinea-Bissau GNB 37.503 3.10 Low income
4. Equatorial Guinea GNQ 35.362 16.40 High income

Upper middle

**83** Iraq IRQ 31.093 9.20

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

1. Mozambique MOZ 39.705 5.40 Low income

Lower middle

1. Mauritania MRT 33.801 6.20

income

**123** Malawi MWI 39.459 5.05 Low income

**127** Niger NER 49.661 1.70 Low income

**148** Rwanda RWA 32.689 9.00 Low income

Lower middle

**151** Senegal SEN 38.533 13.10

income

Lower middle

1. Solomon Islands SLB 30.578 8.00

income

1. 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

1. Chad TCD 45.745 2.30 Low income
2. Togo TGO 36.080 4.50 Low income

Lower middle

**170** Tajikistan TJK 30.792 16.00

income

Lower middle

**172** Timor-Leste TLS 35.755 1.10

income

1. Tanzania TZA 39.518 4.40 Low income
2. Uganda UGA 43.474 16.20 Low income
3. Congo, Dem. Rep. COD 42.394 2.20 Low income

Lower middle

1. Zambia ZMB 40.471 15.40

income

1. Zimbabwe ZWE 35.715 18.50 Low income

In [117]:

stats[(stats**.**InternetUsers**<**20)**&**(stats**.**BirthRate**>**30)]

**CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Afghanistan AFG 35.253 5.90 Low income

Upper middle

1. Angola AGO 45.985 19.10

income

**11** Burundi BDI 44.151 1.30 Low income

1. Benin BEN 36.440 4.90 Low income
2. Burkina Faso BFA 40.551 9.10 Low income

Central African

**29** CAF 34.076 3.50 Low income

Republic

Lower middle

1. Cote d'Ivoire CIV 37.320 8.40

income

Lower middle

1. Cameroon CMR 37.236 6.40

income

Lower middle

1. Congo, Rep. COG 37.011 6.60

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

Upper middle

**60** Gabon GAB 30.555 9.20

income

Lower middle

1. Ghana GHA 33.131 12.30

income

1. Guinea GIN 37.337 1.60 Low income
2. Gambia, The GMB 42.525 14.00 Low income
3. Guinea-Bissau GNB 37.503 3.10 Low income
4. Equatorial Guinea GNQ 35.362 16.40 High income

Upper middle

**83** Iraq IRQ 31.093 9.20

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

1. Mozambique MOZ 39.705 5.40 Low income

Lower middle

1. Mauritania MRT 33.801 6.20

income

**123** Malawi MWI 39.459 5.05 Low income

**127** Niger NER 49.661 1.70 Low income

**148** Rwanda RWA 32.689 9.00 Low income

Lower middle

**151** Senegal SEN 38.533 13.10

income

Lower middle

1. Solomon Islands SLB 30.578 8.00

income

1. 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

1. Chad TCD 45.745 2.30 Low income
2. Togo TGO 36.080 4.50 Low income

Lower middle

**170** Tajikistan TJK 30.792 16.00

income

Lower middle

**172** Timor-Leste TLS 35.755 1.10

income

1. Tanzania TZA 39.518 4.40 Low income
2. Uganda UGA 43.474 16.20 Low income
3. Congo, Dem. Rep. COD 42.394 2.20 Low income

Lower middle

1. Zambia ZMB 40.471 15.40

income

1. Zimbabwe ZWE 35.715 18.50 Low income

In [118]:

stats**.**head()

**CountryName CountryCode BirthRate InternetUsers IncomeGroup**

1. Aruba ABW 10.244 78.9 High income
2. Afghanistan AFG 35.253 5.9 Low income

Upper middle

1. Angola AGO 45.985 19.1

income

Upper middle

1. Albania ALB 12.877 57.2

income

United Arab

1. ARE 11.044 88.0 High income

Emirates

In [119]:

stats**.**IncomeGroup **==** 'Low income'

Out[119]:0 False

1. True
2. False
3. False
4. False

...

1. False
2. False
3. True
4. False
5. True

Name: IncomeGroup, Length: 195, dtype: bool

In [120]:

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

Out[120]: **CountryName CountryCode BirthRate InternetUsers IncomeGroup**

**1** Afghanistan AFG 35.253 5.90 Low income

**11** Burundi BDI 44.151 1.30 Low income

1. Benin BEN 36.440 4.90 Low income
2. Burkina Faso BFA 40.551 9.10 Low income

Central African

**29** CAF 34.076 3.50 Low income

Republic

**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

1. Guinea GIN 37.337 1.60 Low income
2. Gambia, The GMB 42.525 14.00 Low income
3. 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

1. Chad TCD 45.745 2.30 Low income
2. Togo TGO 36.080 4.50 Low income
3. Tanzania TZA 39.518 4.40 Low income
4. 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 [121]: len(stats[stats**.**IncomeGroup **==** 'Low income'])

Out[121]:30

In [123]:

*#unique categories* stats**.**IncomeGroup**.**unique()

Out[123]:array(['High income', 'Low income', 'Upper middle income',

'Lower middle income'], dtype=object)

**1> importing data into python**

**2> Dataframe via panda**

**3> exploring datasets: head()tail()info()describe()**

**4> Renaming columns**

**5> subsetting dataframes**

**6> Basic operations with dataframe**

**7> filtering data frames**

In [ ]: