In [17]:
```
import numpy as np
import pandas as pd
```

In [19]:
```
store = pd.read_csv(r"C:\Users\ruchi\Downloads\Sample - Superstore_Orders.csv")
```

In [21]:
```
store
```

Out[21]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Or
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20 103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20 112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20 112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20 112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20 141
...	...	...	...	...	...	...	
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20 143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20 115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20 156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20 143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20 143

10194 rows × 19 columns

In [23]:
```
id(store)
```

Out[23]:
```
2185697356704
```

In [25]:
```
len(store)
```

Out[25]:
```
10194
```

In [27]:
```
store.columns
```

Out[27]:
```
Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer',
       'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
       'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',
       'Discount', 'Profit', 'Quantity', 'Sales'],
      dtype='object')
```

In [29]:
```
store.shape
```

Out[29]:
```
(10194, 19)
```

In [31]:
```
len(store.columns)
```

Out[31]:
```
19
```

In [33]:
```
store.isnull()
```

Out[33]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Posta Code
0	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...
10189	False	False	False	False	False	False	False	False
10190	False	False	False	False	False	False	False	False
10191	False	False	False	False	False	False	False	False
10192	False	False	False	False	False	False	False	False
10193	False	False	False	False	False	False	False	False

10194 rows × 19 columns

In [35]:
```
store.isna()
```

Out[35]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Posta Code
0	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...
10189	False	False	False	False	False	False	False	False
10190	False	False	False	False	False	False	False	False
10191	False	False	False	False	False	False	False	False
10192	False	False	False	False	False	False	False	False
10193	False	False	False	False	False	False	False	False

10194 rows × 19 columns

In [37]:
```
store.isnull().sum()
```

Out[37]:
```
Category          0
City              0
Country/Region    0
Customer Name     0
Manufacturer      0
Order Date        0
Order ID          0
Postal Code       0
Product Name      0
Region            0
Segment           0
Ship Date         0
Ship Mode         0
State/Province    0
Sub-Category      0
Discount          0
Profit            0
Quantity          0
Sales             0
dtype: int64
```

In [39]:
```
store.dtypes
```

Out[39]:
```
Category           object
City               object
Country/Region     object
Customer Name      object
Manufacturer       object
Order Date         object
Order ID           object
Postal Code        object
Product Name       object
Region             object
Segment            object
Ship Date          object
Ship Mode          object
State/Province     object
Sub-Category       object
Discount          float64
Profit            float64
Quantity            int64
Sales             float64
dtype: object
```

In [41]:
```
store.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 19 columns):
 #   Column          Non-Null Count  Dtype  
---  ------          --------------  -----  
 0   Category        10194 non-null  object 
 1   City            10194 non-null  object 
 2   Country/Region  10194 non-null  object 
 3   Customer Name   10194 non-null  object 
 4   Manufacturer    10194 non-null  object 
 5   Order Date      10194 non-null  object 
 6   Order ID        10194 non-null  object 
 7   Postal Code     10194 non-null  object 
 8   Product Name    10194 non-null  object 
 9   Region          10194 non-null  object 
 10  Segment         10194 non-null  object 
 11  Ship Date       10194 non-null  object 
 12  Ship Mode       10194 non-null  object 
 13  State/Province  10194 non-null  object 
 14  Sub-Category    10194 non-null  object 
 15  Discount        10194 non-null  float64
 16  Profit          10194 non-null  float64
 17  Quantity        10194 non-null  int64  
 18  Sales           10194 non-null  float64
dtypes: float64(3), int64(1), object(15)
memory usage: 1.5+ MB
```

In [43]:
```
pd.__version__
```

Out[43]:
```
'2.2.2'
```

In [45]:
```
store.head()
```

Out[45]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Po C
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326	60
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326	60
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	US-2020-112326	60
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817	19

In [47]:
```
store.head(3)
```

Out[47]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Post Coc
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	7709
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326	6054
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326	6054

In [49]:
```
store.tail()
```

Out[49]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Or
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20 143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20 115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20 156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20 143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20 143

In [51]:
```
store.tail(3)
```

Out[51]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Or
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20 156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20 143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20 143

In [53]:
```
store
```

Out[53]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Or
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20 103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20 112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20 112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20 112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20 141
...	...	...	...	...	...	...	
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20 143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20 115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20 156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20 143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20 143

10194 rows × 19 columns

In [55]:
```
store[:]
```

Out[55]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Or
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20 103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20 112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20 112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20 112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20 141
...	...	...	...	...	...	...	
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20 143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20 115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20 156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20 143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20 143

10194 rows × 19 columns

In [57]:
```
store[0:50:10]
```

Out[57]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Po C
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77
10	Office Supplies	Henderson	United States	Maria Etezadi	Southworth	06-01-2020	US-2020-167199	42
20	Furniture	Dover	United States	Seth Vernon	DAX	11-01-2020	US-2020-130092	19
30	Office Supplies	San Francisco	United States	Brian Dahlen	Tennsco	13-01-2020	US-2020-157147	94
40	Office Supplies	Scottsdale	United States	Toby Swindell	GBC	19-01-2020	US-2020-146591	85

In [59]:
```
store.head(1)
```

Out[59]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77095

In [61]:
```
store['Category']
```

Out[61]:
```
0        Office Supplies
1        Office Supplies
2        Office Supplies
3        Office Supplies
4        Office Supplies
             ...
10189    Office Supplies
10190    Office Supplies
10191    Office Supplies
10192         Technology
10193    Office Supplies
Name: Category, Length: 10194, dtype: object
```

In [63]:
```
store[['Category','City']]
```

Out[63]:

	Category	City
0	Office Supplies	Houston
1	Office Supplies	Naperville
2	Office Supplies	Naperville
3	Office Supplies	Naperville
4	Office Supplies	Philadelphia
...	...	...
10189	Office Supplies	New York City
10190	Office Supplies	Fairfield
10191	Office Supplies	Loveland
10192	Technology	New York City
10193	Office Supplies	Charlottetown

10194 rows × 2 columns

In [65]:
```
store[['Category','City','Customer Name']]
```

Out[65]:

	Category	City	Customer Name
0	Office Supplies	Houston	Darren Powers
1	Office Supplies	Naperville	Phillina Ober
2	Office Supplies	Naperville	Phillina Ober
3	Office Supplies	Naperville	Phillina Ober
4	Office Supplies	Philadelphia	Mick Brown
...	...	...	...
10189	Office Supplies	New York City	Patrick O'Donnell
10190	Office Supplies	Fairfield	Erica Bern
10191	Office Supplies	Loveland	Jill Matthias
10192	Technology	New York City	Patrick O'Donnell
10193	Office Supplies	Charlottetown	Harry Olson

10194 rows × 3 columns

In [67]:
```
store.columns
```

Out[67]:
```
Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer',
       'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
       'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',
       'Discount', 'Profit', 'Quantity', 'Sales'],
      dtype='object')
```

In [69]:
```
store.dtypes
```

Out[69]:
```
Category          object
City              object
Country/Region    object
Customer Name     object
Manufacturer      object
Order Date        object
Order ID          object
Postal Code       object
Product Name      object
Region            object
Segment           object
Ship Date         object
Ship Mode         object
State/Province    object
Sub-Category      object
Discount         float64
Profit           float64
Quantity           int64
Sales            float64
dtype: object
```

In [71]:
```
store_num = store[['Discount','Profit', 'Quantity', 'Sales']]
store_num
```

Out[71]:

	Discount	Profit	Quantity	Sales
0	0.2	5.5512	2	16.448
1	0.8	-5.4870	2	3.540
2	0.2	4.2717	3	11.784
3	0.2	-64.7748	3	272.736
4	0.2	4.8840	3	19.536
...	...	...	...	...
10189	0.2	19.7910	3	52.776
10190	0.2	6.4750	2	20.720
10191	0.2	-0.6048	3	3.024
10192	0.0	2.7279	7	90.930
10193	0.2	-0.6048	3	3.024

10194 rows × 4 columns

In [73]:
```
store.columns
```

Out[73]:
```
Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer',
       'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
       'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',
       'Discount', 'Profit', 'Quantity', 'Sales'],
      dtype='object')
```

In [75]:
```
store_cate = store[['Category', 'City', 'Country/Region', 'Customer Name', 'Manu
       'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
       'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category']]
```

In [77]:
```
store_cate.dtypes
```

Out[77]:
```
Category          object
City              object
Country/Region    object
Customer Name     object
Manufacturer      object
Order Date        object
Order ID          object
Postal Code       object
Product Name      object
Region            object
Segment           object
Ship Date         object
Ship Mode         object
State/Province    object
Sub-Category      object
dtype: object
```

In [79]:
```
store['Profit'].mean()
```

Out[79]:
```
28.673417166960963
```

In [81]:
```
store['Profit'].median()
```

Out[81]:
```
8.69
```

In [83]:
```
store['Profit'].mode()
```

Out[83]:
```
0    0.0
Name: Profit, dtype: float64
```

In [85]:
```
store['Profit'].var()
```

Out[85]:
```
54040.02971828826
```

In [87]:
```
store['Profit'].std()
```

Out[87]:
```
232.46511505662147
```

dataframe in python and how to import the dataset

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

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

In [7]:
```
df
```

Out[7]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
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 [9]:
```
len(df)
```

Out[9]:
```
195
```

In [11]:
```
df.shape
```

Out[11]:
```
(195, 6)
```

In [13]:
```
df.columns
```

Out[13]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed: 5'],
      dtype='object')
```

In [15]:
```
type(df)
```

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

In [17]:
```
df
```

Out[17]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
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 [19]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 195 entries, 0 to 194
Data columns (total 6 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 
 5   Unnamed: 5     195 non-null    float64
dtypes: float64(3), object(3)
memory usage: 9.3+ KB
```

In [21]:
```
df.columns
```

Out[21]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed: 5'],
      dtype='object')
```

In [23]:
```
len(df.columns)
```

Out[23]:
```
6
```

In [25]:
```
df.shape
```

Out[25]: (195, 6)

In [27]:
```
df.head()
```

Out[27]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
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 [29]:
```
df.tail()
```

Out[29]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
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

In [31]:
```
df.tail(2)
```

Out[31]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
193	Zambia	ZMB	40.471	15.4	Lower middle income	623.2534
194	Zimbabwe	ZWE	35.715	18.5	Low income	660.7275

In [33]:
```
df.columns
```

Out[33]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed: 5'],
      dtype='object')
```

In [35]:
```
df.head(2)
```

Out[35]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
0	Aruba	ABW	10.244	78.9	High income	808.2516
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927

In [37]:
```
df
```

Out[37]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
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 [39]:
```
df[::-1]
```

Out[39]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
194	Zimbabwe	ZWE	35.715	18.5	Low income	660.7275
193	Zambia	ZMB	40.471	15.4	Lower middle income	623.2534
192	Congo, Dem. Rep.	COD	42.394	2.2	Low income	93.2668
191	South Africa	ZAF	20.850	46.5	Upper middle income	969.5250
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income	658.9400
...	...	...	...	...	...	...
4	United Arab Emirates	ARE	11.044	88.0	High income	971.8720
3	Albania	ALB	12.877	57.2	Upper middle income	736.5644
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927
0	Aruba	ABW	10.244	78.9	High income	808.2516

195 rows × 6 columns

In [41]:
```
df[:5]
```

Out[41]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
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 [43]:
```
df[6:]
```

Out[43]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
6	Armenia	ARM	13.308	41.9000	Lower middle income	557.60520
7	Antigua and Barbuda	ATG	16.447	63.4000	High income	1042.73980
8	Australia	AUS	13.200	83.0000	High income	1095.60000
9	Austria	AUT	9.400	80.6188	High income	757.81672
10	Azerbaijan	AZE	18.300	58.7000	Upper middle income	1074.21000
...	...	...	...	...	...	...
190	Yemen, Rep.	YEM	32.947	20.0000	Lower middle income	658.94000
191	South Africa	ZAF	20.850	46.5000	Upper middle income	969.52500
192	Congo, Dem. Rep.	COD	42.394	2.2000	Low income	93.26680
193	Zambia	ZMB	40.471	15.4000	Lower middle income	623.25340
194	Zimbabwe	ZWE	35.715	18.5000	Low income	660.72750

189 rows × 6 columns

In [45]:
```
df[0:200:10]
```

Out[45]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
0	Aruba	ABW	10.244	78.900000	High income	808.251600
10	Azerbaijan	AZE	18.300	58.700000	Upper middle income	1074.210000
20	Belarus	BLR	12.500	54.170000	Upper middle income	677.125000
30	Canada	CAN	10.900	85.800000	High income	935.220000
40	Costa Rica	CRI	15.022	45.960000	Upper middle income	690.411120
50	Ecuador	ECU	21.070	40.353684	Upper middle income	850.252127
60	Gabon	GAB	30.555	9.200000	Upper middle income	281.106000
70	Greenland	GRL	14.500	65.800000	High income	954.100000
80	India	IND	20.291	15.100000	Lower middle income	306.394100
90	Kazakhstan	KAZ	22.730	54.000000	Upper middle income	1227.420000
100	Libya	LBY	21.425	16.500000	Upper middle income	353.512500
110	Moldova	MDA	12.141	45.000000	Lower middle income	546.345000
120	Mozambique	MOZ	39.705	5.400000	Low income	214.407000
130	Netherlands	NLD	10.200	93.956400	High income	958.355280
140	Poland	POL	9.600	62.849200	High income	603.352320
150	Sudan	SDN	33.477	22.700000	Lower middle income	759.927900
160	Suriname	SUR	18.455	37.400000	Upper middle income	690.217000
170	Tajikistan	TJK	30.792	16.000000	Lower middle income	492.672000
180	Uruguay	URY	14.374	57.690000	High income	829.236060
190	Yemen, Rep.	YEM	32.947	20.000000	Lower middle income	658.940000

In [47]:
```
df.describe()
```

Out[47]:

	BirthRate	InternetUsers	Unnamed: 5
count	195.000000	195.000000	195.000000
mean	21.469928	42.076471	653.559009
std	10.605467	29.030788	351.553521
min	7.900000	0.900000	28.990400
25%	12.120500	14.520000	361.263300
50%	19.680000	41.000000	682.074300
75%	29.759500	66.225000	892.690170
max	49.661000	96.546800	1552.589500

In [51]:
```
df.describe
```

Out[51]:
```
<bound method NDFrame.describe of              CountryName CountryCode  BirthR
ate  InternetUsers  \
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  Unnamed: 5
0            High income    808.2516
1             Low income    207.9927
2    Upper middle income    878.3135
3    Upper middle income    736.5644
4            High income    971.8720
..                   ...         ...
190  Lower middle income    658.9400
191  Upper middle income    969.5250
192           Low income     93.2668
193  Lower middle income    623.2534
194           Low income    660.7275

[195 rows x 6 columns]>
```

In [72]:
```
df.describe().transpose()
```

Out[72]:

	count	mean	std	min	25%	50%	75%	
BirthRate	195.0	21.469928	10.605467	7.9000	12.1205	19.6800	29.75950	
InternetUsers	195.0	42.076471	29.030788	0.9000	14.5200	41.0000	66.22500	
Unnamed: 5	195.0	653.559009	351.553521	28.9904	361.2633	682.0743	892.69017	1!

In [74]:
```
df.head(2)
```

Out[74]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
0	Aruba	ABW	10.244	78.9	High income	808.2516
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927

In [76]:
```
df.columns
```

Out[76]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed: 5'],
      dtype='object')
```

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

In [65]:
```
df.head(1)
```

Out[65]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed: 5
0	Aruba	ABW	10.244	78.9	High income	808.2516

In [67]:
```
df.columns
```

Out[67]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed: 5'],
      dtype='object')
```

In [69]:
```
df.columns = ['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed']
```

In [71]:
```
df.head(1)
```

Out[71]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
0	Aruba	ABW	10.244	78.9	High income	808.2516

In [73]:
```
df[:]
```

Out[73]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
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 [75]:
```
df[0:5]
```

Out[75]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
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 [77]:
```
df.isnull()
```

Out[77]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
...	...	...	...	...	...	...
190	False	False	False	False	False	False
191	False	False	False	False	False	False
192	False	False	False	False	False	False
193	False	False	False	False	False	False
194	False	False	False	False	False	False

195 rows × 6 columns

In [79]:
```
df.isnull().sum()
```

Out[79]:
```
CountryName      0
CountryCode      0
BirthRate        0
InternetUsers    0
IncomeGroup      0
Unnamed          0
dtype: int64
```

In [83]:
```
df.dtypes
```

Out[83]:
```
CountryName       object
CountryCode       object
BirthRate        float64
InternetUsers    float64
IncomeGroup       object
Unnamed          float64
dtype: object
```

In [85]:
```
df.columns
```

Out[85]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed'],
      dtype='object')
```

In [87]:
```
df_categoical = df[['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed']]
df_categoical.head()
```

Out[87]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
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 [89]:
```
df.describe()
```

Out[89]:

	BirthRate	InternetUsers	Unnamed
count	195.000000	195.000000	195.000000
mean	21.469928	42.076471	653.559009
std	10.605467	29.030788	351.553521
min	7.900000	0.900000	28.990400
25%	12.120500	14.520000	361.263300
50%	19.680000	41.000000	682.074300
75%	29.759500	66.225000	892.690170
max	49.661000	96.546800	1552.589500

In [91]:
```
df_categoical
```

Out[91]:

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
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 [93]:
```
df_categoical.describe()
```

Out[93]:

	BirthRate	InternetUsers	Unnamed
count	195.000000	195.000000	195.000000
mean	21.469928	42.076471	653.559009
std	10.605467	29.030788	351.553521
min	7.900000	0.900000	28.990400
25%	12.120500	14.520000	361.263300
50%	19.680000	41.000000	682.074300
75%	29.759500	66.225000	892.690170
max	49.661000	96.546800	1552.589500

In [101…
```
df_num = df[['BirthRate','InternetUsers']]
df_num.head()
```

Out[101…

	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 [105…
```
df_num.describe().transpose()
```

Out[105…

	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 [107…
```
df.head()
```

Out[107…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
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 [111…
```
df['IncomeGroup']
```

Out[111…
```
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 [113…
```
df.columns
```

Out[113…
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed'],
      dtype='object')
```

In [115…
```
['CountryName','BirthRate']
```

Out[115… `['CountryName', 'BirthRate']`

In [123…
```
df.head()
```

Out[123…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
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 [121…
```
df.BirthRate* df.InternetUsers
```

Out[121…
```
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 [125…
```
df.head(1)
```

Out[125…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
0	Aruba	ABW	10.244	78.9	High income	808.2516

In [127…
```
df['myCalc'] = df.BirthRate * df.InternetUsers
```

In [129…
```
df
```

Out[129…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed	
0	Aruba	ABW	10.244	78.9	High income	808.2516	8
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927	2
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135	8
3	Albania	ALB	12.877	57.2	Upper middle income	736.5644	7
4	United Arab Emirates	ARE	11.044	88.0	High income	971.8720	9
…	…	…	…	…	…	…	
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income	658.9400	6
191	South Africa	ZAF	20.850	46.5	Upper middle income	969.5250	9
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	6
194	Zimbabwe	ZWE	35.715	18.5	Low income	660.7275	6

195 rows × 7 columns

In [131…
```
df.columns
```

Out[131…
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed', 'myCalc'],
      dtype='object')
```

In [135…
```
len(df.columns)
```

Out[135…
```
7
```

In [137…
```
df
```

Out[137…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed	
0	Aruba	ABW	10.244	78.9	High income	808.2516	8
1	Afghanistan	AFG	35.253	5.9	Low income	207.9927	2
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135	8
3	Albania	ALB	12.877	57.2	Upper middle income	736.5644	7
4	United Arab Emirates	ARE	11.044	88.0	High income	971.8720	9
…	…	…	…	…	…	…	
190	Yemen, Rep.	YEM	32.947	20.0	Lower middle income	658.9400	6
191	South Africa	ZAF	20.850	46.5	Upper middle income	969.5250	9
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	6
194	Zimbabwe	ZWE	35.715	18.5	Low income	660.7275	6

195 rows × 7 columns

In [139…
```
df = df.drop('myCalc',axis = 1)
```

In [141…
```
df.columns
```

Out[141…
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup', 'Unnamed'],
      dtype='object')
```

In [143…
```
df.columns[3:4]
```

Out[143…
```
Index(['InternetUsers'], dtype='object')
```

In [145…
```
df.shape
```

Out[145…
```
(195, 6)
```

In [147…
```
df['InternetUsers']
```

Out[147…
```
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 [149…
```
df.InternetUsers<2
```

Out[149…
```
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 [151…
```
df
```

Out[151…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
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 [153…
```
Filter = df.InternetUsers < 2
```

In [155…
```
Filter
```

Out[155…
```
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 [157…
```
df[Filter]
```

Out[157…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
11	Burundi	BDI	44.151	1.3	Low income	57.3963
52	Eritrea	ERI	34.800	0.9	Low income	31.3200
55	Ethiopia	ETH	32.925	1.9	Low income	62.5575
64	Guinea	GIN	37.337	1.6	Low income	59.7392
117	Myanmar	MMR	18.119	1.6	Lower middle income	28.9904
127	Niger	NER	49.661	1.7	Low income	84.4237
154	Sierra Leone	SLE	36.729	1.7	Low income	62.4393
156	Somalia	SOM	43.891	1.5	Low income	65.8365
172	Timor-Leste	TLS	35.755	1.1	Lower middle income	39.3305

In [159…
```
len(df[Filter])
```

Out[159… 9

In [161…
```
df
```

Out[161…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
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 [163…
```
df.BirthRate>40
```

Out[163…
```
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 [165…
```
Filter2 = df.BirthRate>40
Filter2
```

Out[165…
```
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 [167…
```
df[Filter2]
```

Out[167…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135
11	Burundi	BDI	44.151	1.3	Low income	57.3963
14	Burkina Faso	BFA	40.551	9.1	Low income	369.0141
65	Gambia, The	GMB	42.525	14.0	Low income	595.3500
115	Mali	MLI	44.138	3.5	Low income	154.4830
127	Niger	NER	49.661	1.7	Low income	84.4237
128	Nigeria	NGA	40.045	38.0	Lower middle income	1521.7100
156	Somalia	SOM	43.891	1.5	Low income	65.8365
167	Chad	TCD	45.745	2.3	Low income	105.2135
178	Uganda	UGA	43.474	16.2	Low income	704.2788
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

In [169…
```
len(df[Filter2])
```

Out[169… 12

In [171…
```
df[Filter2]
```

Out[171…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
2	Angola	AGO	45.985	19.1	Upper middle income	878.3135
11	Burundi	BDI	44.151	1.3	Low income	57.3963
14	Burkina Faso	BFA	40.551	9.1	Low income	369.0141
65	Gambia, The	GMB	42.525	14.0	Low income	595.3500
115	Mali	MLI	44.138	3.5	Low income	154.4830
127	Niger	NER	49.661	1.7	Low income	84.4237
128	Nigeria	NGA	40.045	38.0	Lower middle income	1521.7100
156	Somalia	SOM	43.891	1.5	Low income	65.8365
167	Chad	TCD	45.745	2.3	Low income	105.2135
178	Uganda	UGA	43.474	16.2	Low income	704.2788
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

In [173…
```
Filter & Filter2
```

Out[173…
```
0      False
1      False
2      False
3      False
4      False
       ...
190    False
191    False
192    False
193    False
194    False
Length: 195, dtype: bool
```

In [175…
```
df[Filter & Filter ]
```

Out[175…

	CountryName	CountryCode	BirthRate	InternetUsers	IncomeGroup	Unnamed
11	Burundi	BDI	44.151	1.3	Low income	57.3963
52	Eritrea	ERI	34.800	0.9	Low income	31.3200
55	Ethiopia	ETH	32.925	1.9	Low income	62.5575
64	Guinea	GIN	37.337	1.6	Low income	59.7392
117	Myanmar	MMR	18.119	1.6	Lower middle income	28.9904
127	Niger	NER	49.661	1.7	Low income	84.4237
154	Sierra Leone	SLE	36.729	1.7	Low income	62.4393
156	Somalia	SOM	43.891	1.5	Low income	65.8365
172	Timor-Leste	TLS	35.755	1.1	Lower middle income	39.3305

In [ ]: