Pandas_may 6th,25

May 7, 2025

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

```python
[5]: print(pd.__version__)
```

```
2.2.2
```

```python
[11]: store=pd.read_csv(r'/Users/shashi/Downloads/data.csv')
```

```python
[29]: store
```

```
[29]:          CountryName CountryCode  BirthRate  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
      0           High income
      1            Low income
      2    Upper middle income
      3    Upper middle income
      4           High income
      ..                  ...
      190  Lower middle income
      191  Upper middle income
      192           Low income
      193  Lower middle income
      194           Low income

      [195 rows x 5 columns]
```

```
[33]: len(store)
```

```
[33]: 195
```

```
[35]: len(store.columns)
```

```
[35]: 5
```

```
[37]: type(store)
```

```
[37]: pandas.core.frame.DataFrame
```

```
[39]: store
```

```
[39]:          CountryName CountryCode   BirthRate   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
      0         High income
      1          Low income
      2    Upper middle income
      3    Upper middle income
      4         High income
      ..                  …
      190  Lower middle income
      191  Upper middle income
      192          Low income
      193  Lower middle income
      194          Low income

      [195 rows x 5 columns]
```

```
[41]: store.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
```

[ ]: `# Info shows the range between 0-194, if data set is large memory will increase`

[43]: `store.columns`

[43]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

[49]: `store.head()`   #top rows

[49]:
```
          CountryName CountryCode  BirthRate  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

          IncomeGroup
0         High income
1          Low income
2  Upper middle income
3  Upper middle income
4         High income
```

[51]: `store.tail()`   #bottom rows

[51]:
```
          CountryName CountryCode  BirthRate  InternetUsers  \
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
190  Lower middle income
191  Upper middle income
192          Low income
193  Lower middle income
194          Low income
```

```
[53]: store.tail(4)    # bottom 4 rows
```

```
[53]:           CountryName CountryCode  BirthRate  InternetUsers  \
      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
      191  Upper middle income
      192           Low income
      193  Lower middle income
      194           Low income
```

```
[55]: store.head(2)    #top 2 rows
```

```
[55]:    CountryName CountryCode  BirthRate  InternetUsers  IncomeGroup
      0       Aruba         ABW     10.244           78.9  High income
      1  Afghanistan         AFG     35.253            5.9   Low income
```

```
[57]: store[::-1]      #we get reverse table
```

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

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

```
[195 rows x 5 columns]
```

```
[59]: store[:5]
```

```
[59]:          CountryName CountryCode  BirthRate  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

                    IncomeGroup
      0             High income
      1              Low income
      2    Upper middle income
      3    Upper middle income
      4             High income
```

```
[61]: store[5:]
```

```
[61]:          CountryName CountryCode  BirthRate  InternetUsers  \
      5             Argentina       ARG     17.716        59.9000
      6               Armenia       ARM     13.308        41.9000
      7    Antigua and Barbuda      ATG     16.447        63.4000
      8             Australia       AUS     13.200        83.0000
      9               Austria       AUT      9.400        80.6188
      ..                  ...       ...        ...            ...
      190          Yemen, Rep.      YEM     32.947        20.0000
      191         South Africa      ZAF     20.850        46.5000
      192     Congo, Dem. Rep.     COD     42.394         2.2000
      193              Zambia       ZMB     40.471        15.4000
      194            Zimbabwe       ZWE     35.715        18.5000

                     IncomeGroup
      5              High income
      6       Lower middle income
      7              High income
      8              High income
      9              High income
      ..                     ...
      190     Lower middle income
      191     Upper middle income
      192              Low income
      193     Lower middle income
      194              Low income
```

```
[190 rows x 5 columns]
```

```
[63]: store[0:200:10]
```

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

```
[67]: store.describe()    # descriptive stats - will give only numerical data of
       ↪birth data & interner users such as
       #mean,std deviation, 25 % percentil, 75% percentile, mac value
```

```
[67]:           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
```

```
[69]: store.describe().transpose()    #transpose converts rows into columns
```

```
[69]:                 count        mean         std  min      25%    50%      75% \
      BirthRate       195.0   21.469928   10.605467  7.9  12.1205  19.68  29.7595
      InternetUsers   195.0   42.076471   29.030788  0.9  14.5200  41.00  66.2250

                        max
```

```
          BirthRate        49.6610
          InternetUsers    96.5468
```

[77]: `store.columns`

[77]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

[79]: `store.columns=['a','b','c','d','e']`

[81]: `store.columns`

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

[91]:
```
store.columns=['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup']
```

[93]: `store.columns`

[93]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

[89]: `store.head()`

[89]:
```
                      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
```

[95]: `store[:]`

[95]:
```
             CountryName CountryCode  BirthRate  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
0           High income
1            Low income
2    Upper middle income
3    Upper middle income
4           High income
..                  …
190  Lower middle income
191  Upper middle income
192          Low income
193  Lower middle income
194          Low income

[195 rows x 5 columns]
```

[99]: `store.isnull()`

[99]:
	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 x 5 columns]
```

[101]: `store.isnull().sum()   # same as store.isna()`

[101]:
```
CountryName      0
CountryCode      0
BirthRate        0
InternetUsers    0
IncomeGroup      0
dtype: int64
```

[133]: `store.isna()`

[133]:
	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 x 5 columns]
```

[103]: `store.dtypes`

[103]:
```
CountryName      object
CountryCode      object
BirthRate       float64
InternetUsers   float64
IncomeGroup      object
dtype: object
```

[105]: `store.columns`

[105]:
```
Index(['CountryName', 'CountryCode', 'BirthRate', 'InternetUsers',
       'IncomeGroup'],
      dtype='object')
```

[107]: `store.head(2)`

[107]:
```
   CountryName CountryCode  BirthRate  InternetUsers  IncomeGroup
0        Aruba         ABW     10.244           78.9  High income
1  Afghanistan         AFG     35.253            5.9   Low income
```

[109]: `store`

[109]:
```
              CountryName CountryCode  BirthRate  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
```

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

[195 rows x 5 columns]
```

[ ]: *#divide the data into two part – categorical data into one part & numerical*␣
    ↪data into 2nd type.

[113]: `store['CountryName']`

[113]:
```
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
```

[117]: `store[['CountryName','CountryCode']]`

[117]:
```
         CountryName CountryCode
0              Aruba         ABW
1        Afghanistan         AFG
2             Angola         AGO
3            Albania         ALB
4   United Arab Emirates       ARE
..               …           …
190      Yemen, Rep.         YEM
191     South Africa         ZAF
192  Congo, Dem. Rep.        COD
193           Zambia         ZMB
194         Zimbabwe         ZWE
```

```
[195 rows x 2 columns]
```

```
[131]: store_categorical = store[['CountryName', 'CountryCode', 'IncomeGroup']]
       store_categorical.head()
```

```
[131]:             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
```

```
[139]: store.describe()
```

```
[139]:        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
```

```
[141]: store_categorical.describe()
```

```
[141]:         CountryName CountryCode   IncomeGroup
       count           195         195           195
       unique          195         195             4
       top           Aruba         ABW   High income
       freq              1           1            67
```

```
[147]: num = store[['BirthRate', 'InternetUsers']]
```

```
[149]: store.num.head()
```

```
[149]:    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
```

```
[151]: store
```

```
[151]:             CountryName CountryCode   BirthRate   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
0          High income
1           Low income
2   Upper middle income
3   Upper middle income
4          High income
..                   …
190  Lower middle income
191  Upper middle income
192           Low income
193  Lower middle income
194           Low income

[195 rows x 5 columns]
```

[ ]: