# Data Science – Pandas – DataFrame – Rename column, index

## 10. Pandas – DataFrame – Rename column & index

## Contents

1. rename(p) method - Changing single column name	2
2. rename(p) method - Changing multiple column names	
3. columns attribute - Changing multiple column names	8
4. rename(p) method - Changing index in DataFrame	10
5. index attribute - Changing multiple column names	12
5. Converting column names from lower case to upper case	15

#### 10. Pandas – DataFrame – Rename column & index

- ✓ Sometimes we need to process column names and index.
- ✓ Based on requirement we can rename/change/modify/update column names and index.

### 1. rename(p) method - Changing single column name

- ✓ rename(p) is a predefined method in DataFrame
- ✓ We can access rename() method by using DataFrame object.
- ✓ By using this method we can change single column name in DataFrame.

```
Program
              Creating DataFrame and checking column names
Name
              demo1.py
Input file
              sales3.csv
              import pandas as pd
              df = pd.read csv("sales3.csv")
              print(df.head())
              print()
              print(df.columns)
Output
                 192837
                           Veeru
                                                        LG Mobile
                                                                     65999
                         Neelima
                                              Apple iPad 10.2-inch
                 192838
                          Balaji
                                             34in Ultrawide Monitor
                 192839
                                                                     75999
                 192840
                          Shahid
                                     20
                                                        iPhone 11
                                                                     60000
                                     10 Bose SoundSport Headphones
                           Vinay
                                                                     69999
               Index(['ord id', 'cust name', 'cust id', 'prod name', 'prod cost'], dtype='object'
```

#### Note

- ✓ We need to use python dictionary to specify old column as key and new. column as value.
- ✓ Whenever changing column then key should be match with column name in DataFrame otherwise changes will not reflects.

```
Program
             Renaming single column in DataFrame
Name
             demo2.py
Input file
             sales3.csv
             import pandas as pd
             df1 = pd.read csv("sales3.csv")
             d = {
                  "ord id": "Order Id"
             }
             df2 = df1.rename(columns = d)
             print(df1.head())
             print()
             print(df2.head())
Output
                ord id cust name
                                                          prod name
                                                                     prod cost
                192837
                          Veeru
                                                          LG Mobile
                                                                         65999
                        Neelima
                                      19
                                                Apple iPad 10.2-inch
                192838
                                                                         63000
                                      12
                                              34in Ultrawide Monitor
                192839
                         Balaii
                                                                         75999
                         Shahid
                                      20
                                                          iPhone 11
                192840
                                                                         60000
                                      10 Bose SoundSport Headphones
                192841
                          Vinay
                                                                         69999
                Order Id cust name cust id
                                                            prod name
                                                                       prod cost
                  192837
                            Veeru
                                        3
                                                            LG Mobile
                                                                           65999
                                                                           63000
                          Neelima
                                        19
                                                  Apple iPad 10.2-inch
```

12

20

192838

192839

192840

192841

Balaji

Shahid

Vinay

iPhone 11

75999

60000

69999

34in Ultrawide Monitor

10 Bose SoundSport Headphones

### Note

✓ Whenever changing column then key should be match with column name in DataFrame otherwise changes will not reflects.

```
Program
            Renaming single column in DataFrame
Name
            demo3.py
Input file
            sales3.csv
            import pandas as pd
            df1 = pd.read_csv("sales3.csv")
            d = {
                "orddddd id": "Order Id"
            }
            df2 = df1.rename(columns = d)
            print(df1.head())
            print()
            print(df2.head())
Output
```

	ord id	cust name	cust id	prod name	prod cost
0	192837	Veeru	3	LG Mobile	65999
1	192838	Neelima	19	Apple iPad 10.2-inch	63000
2	192839	Balaji	12	34in Ultrawide Monitor	75999
3	192840	Shahid	20	iPhone 11	60000
4	192841	Vinay	10	Bose SoundSport Headphones	69999
	ord id	cust name	cust id	prod name	prod cost
0	192837	Veeru	3	LG Mobile	65999
1	192838	Neelima	19	Apple iPad 10.2-inch	63000
2	192839	Balaji	12	34in Ultrawide Monitor	75999
3	192840	Shahid	20	iPhone 11	60000
4	192841	Vinay	10	Bose SoundSport Headphones	69999

### 2. rename(p) method - Changing multiple column names

- √ rename(p) is a predefined method in DataFrame
- ✓ We can access rename() method by using DataFrame object.
- ✓ We can even change multiple column names by using rename(p) method.
- ✓ We need to use python dictionary to specify old column as key and new column as value.

```
Renaming multiple column in DataFrame
Program
Name
            demo4.py
            sales3.csv
Input file
            import pandas as pd
            df1 = pd.read_csv("sales3.csv")
            d = {
                 'ord id': 'Order Id',
                 'cust name': 'Customer Name',
                 'cust id': 'Customer Id',
                 'prod name': 'Product Name',
                 'prod cost': 'Product Cost'
            }
            df2 = df1.rename(columns = d)
            print(df1.head())
            print()
            print(df2.head())
Output
```

## Data Science – Pandas – DataFrame – Rename column, index

```
ord id cust name
                   cust id
                                                          prod cost
                                              LG Mobile
192837
           Veeru
                                                              65999
192838
         Neelima
                        19
                                  Apple iPad 10.2-inch
                                                              63000
                                 34in Ultrawide Monitor
192839
          Balaji
                                                              75999
           Shahid
                                              iPhone 11
192840
                                                              60000
192841
           Vinay
                        10 Bose SoundSport Headphones
                                                              69999
Order Id Customer Name Customer Id
                                                      Product Name Product Cost
  192837
                 Veeru
                                                         LG Mobile
                                                                            65999
                                           Apple iPad 10.2-inch
34in Ultrawide Monitor
               Neelima
                                   19
  192838
                                                                            63000
  192839
                Balaji
                                                                            75999
  192840
                 Shahid
                                   20
                                                         iPhone 11
                                                                            60000
  192841
                  Vinay
                                       Bose SoundSport Headphones
                                   10
                                                                            69999
```

```
Program
             Renaming multiple column in DataFrame
Name
             demo5.py
Input file
             sales3.csv
             import pandas as pd
             df1 = pd.read csv("sales3.csv")
             d = {
                  "ord id": "order_id",
                  "cust name": "customer_name",
                  "cust id": "customer id",
                  "prod name": "product name",
                  "prod cost": "product_cost"
             }
             df2 = df1.rename(columns = d)
             print(df1.head())
             print()
             print(df2.head())
Output
                ord id cust name
                                cust id
                                                        prod name
                                                                 prod cost
                                                        LG Mobile
                192837
                          Veeru
                                                                     65999
                192838
                        Neelima
                                              Apple iPad 10.2-inch
                                                                     63000
                                     12
                                            34in Ultrawide Monitor
                                                                     75999
                192839
                         Balaji
                192840
                                     20
                                                        iPhone 11
                         Shahid
                                                                     60000
                                     10 Bose SoundSport Headphones
                192841
                                                                     69999
                          Vinay
                order_id customer_name customer_id
                                                              product_name product_cost
                  192837
                               Veeru
                                                                LG Mobile
                                                                                65999
                  192838
                              Neelima
                                              19
                                                       Apple iPad 10.2-inch
                                                                                63000
```

192839

192840

192841

Balaji

Shahid

Vinay

12

20

10

34in Ultrawide Monitor

Bose SoundSport Headphones

iPhone 11

75999

60000

69999

### 3. columns attribute - Changing multiple column names

- ✓ columns is predefined attribute in DataFrame class.
- ✓ We can access columns attribute by using DataFrame object.
- ✓ By using columns attribute we can even change multiple column names.

```
Program
              Renaming multiple column in DataFrame
Name
              demo6.py
Input file
              sales3.csv
              import pandas as pd
              df1 = pd.read_csv("sales3.csv")
              print(df1.head())
              df1.columns = [
                   "order id",
                   "customer name",
                   "customer id",
                   "product name",
                   "product cost"
              1
              print()
              print(df1.head())
Output
                                                                   prod cost
                 ord id cust name
                 192837
                           Veeru
                                                         LG Mobile
                                                                       65999
                         Neelima
                                               Apple iPad 10.2-inch
                 192838
                                      19
                                                                       63000
                          Balaji
                                      12
                                             34in Ultrawide Monitor
                 192839
                          Shahid
                                      20
                 192840
                                                         iPhone 11
                                                                       60000
                 192841
                                      10 Bose SoundSport Headphones
                           Vinay
                 order_id customer_name customer_id
                                                               product_name product_cost
                   192837
                                Veeru
                                                                  LG Mobile
                                                                                   65999
                   192838
                              Neelima
                                               19
                                                        Apple iPad 10.2-inch
                                                                                   63000
                               Balaji
                   192839
                                               12
                                                      34in Ultrawide Monitor
                                                                                   75999
                   192840
                               Shahid
                                               20
                                                                  iPhone 11
                                                                                   60000
                                               10 Bose SoundSport Headphones
                   192841
                                Vinay
                                                                                   69999
```

### Note

✓ While using columns attributes to change column names then Number
of DataFrame column names should be match with existing columns
otherwise we will get an error.

```
Renaming multiple column in DataFrame
Program
Name
            demo7.py
            sales3.csv
Input file
            import pandas as pd
            df1 = pd.read_csv("sales3.csv")
            print(df1.head())
            df1.columns = [
                "order id",
                "customer_name",
                "customer id",
                "product_name",
                "product cost",
                "total"
            ]
            print()
            print(df1.head())
Output
            ValueError: Length mismatch: Expected axis has 5 elements, new
            values have 6 elements
```

### 4. rename(p) method - Changing index in DataFrame

- ✓ rename(p) is a predefined method in DataFrame
- ✓ We can access rename() method by using DataFrame object.
- ✓ Based on requirement we can change the index in DataFrame.
- ✓ By using rename(p) we can change the DataFrame index

```
Program
           Creating a sample Dataframe
Name
           demo8.py
           import pandas as pd
           d = {
               "order id": [11, 21, 31],
               "customer_name": ["Prasad", "Daniel", "Jeswanth"],
               "product": ["iPhone", "hTC", "macbook"]
           }
           df1 = pd.DataFrame(d)
           print(df1)
Output
              order id customer name
                                              product
                                           iPhone 11
                      11
                                  Prasad
                                  Daniel
                      21
                                                   hTC
                                Jeswanth
                                              macbook
                      31
```

```
Program
           Changing index in DataFrame
           demo9.py
Name
           import pandas as pd
           d = {
               "order_id": [11, 21, 31],
               "customer_name": ["Prasad", "Daniel", "Jeswanth"],
               "product": ["iPhone", "hTC", "macbook"]
           }
           i = \{0: 77, 1: 88, 2: 99\}
           df1 = pd.DataFrame(d)
           df2 = df1.rename(index = i)
           print(df1)
           print()
           print(df2)
Output
               order id customer name
                                           product
                      11
                                 Prasad
                                           iPhone
                      21
                                 Daniel
                                                hTC
                               Jeswanth macbook
                      31
               order_id customer_name
                                            product
                                             iPhone
           77
                       11
                                  Prasad
                                  Daniel
                                                 hTC
           88
                       21
           99
                       31
                                Jeswanth
                                           macbook
```

### 5. index attribute - Changing multiple column names

- ✓ index is predefined attribute in DataFrame class.
- ✓ We can access index attribute by using DataFrame object.
- ✓ By using index attribute we can even change index in DataFrame

```
Program
           Changing index in DataFrame, using axis parameter
Name
           demo10.py
           import pandas as pd
           d = {
               "order_id": [11, 21, 31],
               "customer_name": ["Prasad", "Daniel", "Jeswanth"],
               "product": ["iPhone", "hTC", "macbook"]
           }
           df1 = pd.DataFrame(d)
           print(df1)
           df1.index = [77, 88, 99]
           print()
           print(df1)
Output
               order id customer_name
                                           product
                                             iPhone
                      11
                                  Prasad
                      21
                                  Daniel
                                                hTC
                      31
                               Jeswanth
                                           macbook
                order_id customer_name
                                             product
                                             iPhone
           77
                       11
                                   Prasad
```

Daniel

Jeswanth

88

99

21

31

hTC

macbook

Program Changing index in DataFrame, using index attribute.

Name demo11.py
Input file sales31.csv

import pandas as pd

df1 = pd.read\_csv("sales31.csv")
print(df1)

df1.index = range(10, 20)

print()
print(df1)

## Output

_					
	ord id	cust name	cust id	prod name	prod cost
0	192837	Veeru	3	LG Mobile	65999
1	192838	Neelima	19	Apple iPad 10.2-inch	63000
2	192839	Balaji	12	34in Ultrawide Monitor	75999
3	192840	Shahid	20	iPhone 11	60000
4	192841	Vinay	10	Bose SoundSport Headphones	69999
5	192842	Vinay	10	20in Monitor	65999
6	192843	Balaji	12	iPhone 8	55999
7	192844	Nireekshan	1	27in 4K Gaming Monitor	55999
8	192845	Venu	23	iPhone 9	55999
9	192846	Veeru	3	20in Monitor	63000
	ord id	cust name	cust id	prod name	prod cost
10	192837	Veeru	3	LG Mobile	65999
11	192838	Neelima	19	Apple iPad 10.2-inch	63000
12	192839	Balaji	12	34in Ultrawide Monitor	75999
13	192840	Shahid	20	iPhone 11	60000
14	192841	Vinay	10	Bose SoundSport Headphones	69999
15	192842	Vinay	10	20in Monitor	65999
16	192843	Balaji	12	iPhone 8	55999
17	192844	Nireekshan	1	27in 4K Gaming Monitor	55999
18	192845	Venu	23	iPhone 9	55999
19	192846	Veeru	3	20in Monitor	63000

```
Program
           Changing columns and index in a Dataframe
           demo12.py
Name
           import pandas as pd
           d = {
               "Ord Id": [11, 21, 31],
               "Customer Name": ["Prasad", "Daniel", "Jeswanth"],
               "Product": ["iPhone", "hTC", "macbook"]
           }
           df1 = pd.DataFrame(d)
           print(df1)
           df1.index = [333, 444, 555]
           df1.columns = ["order_id", "customer_name", "product"]
           print()
           print(df1)
Output
              Ord Id Customer Name
                                       Product
                                        iPhone
                   11
                              Prasad
                   21
                              Daniel
                                            hTC
                            Jeswanth macbook
                   31
                 order_id customer_name
                                            product
                                   Prasad iPhone
           333
                       11
                                   Daniel
           444
                                                 hTC
                       21
           555
                       31
                                 Jeswanth macbook
```

### 6. Converting column names from lower case to upper case

✓ Based on requirement we can convert DataFrame column names from lower case to upper case.

```
Program
            Converting all columns names into upper case
Name
            demo13.py
Input file
            sales3.csv
            import pandas as pd
            df1 = pd.read csv("sales3.csv")
            print(df1.head())
            df1.columns = df1.columns.str.upper()
            print()
            print(df1.head())
Output
               ord id cust name
                                 cust id
                                                                     prod cost
                                                          prod name
               192837
                          Veeru
                                                          LG Mobile
                                                                         65999
               192838
                                      19
                                               Apple iPad 10.2-inch
                                                                         63000
                        Neelima
                                      12
               192839
                         Balaji
                                              34in Ultrawide Monitor
                                                                         75999
               192840
                         Shahid
                                      20
                                                          iPhone 11
                                                                         60000
               192841
                         Vinay
                                      10 Bose SoundSport Headphones
                                                                         69999
               ORD ID CUST NAME CUST ID
                                                          PROD NAME
                                                                     PROD COST
               192837
                          Veeru
                                      3
                                                          LG Mobile
                                                                         65999
                                      19
                        Neelima
                                               Apple iPad 10.2-inch
                                                                         63000
               192838
                                      12
                                              34in Ultrawide Monitor
               192839
                         Balaji
                                                                         75999
               192840
                         Shahid
                                      20
                                                          iPhone 11
                                                                         60000
               192841
                          Vinay
                                      10 Bose SoundSport Headphones
                                                                         69999
```