importing library pandas

```
import pandas as pd
```

importing into pandas dataframe

```
dataset = pd.read_csv("dirtydata.csv")
print(dataset)
```

prints the shape of dataset

```
dataset.shape
(32, 5)
```

shows datatypes dataset elements

```
Duration int64
Date object
Pulse int64
Maxpulse int64
Calories float64
dtype: object

dataset.columns

Index(['Duration', 'Date', 'Pulse', 'Maxpulse', 'Calories'],
dtype='object')
```

shows information of dataset

```
dataset.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 32 entries, 0 to 31
Data columns (total 5 columns):
              Non-Null Count Dtype
    Column
 0
    Duration 32 non-null
                              int64
 1
    Date
              31 non-null
                              object
    Pulse
 2
             32 non-null
                              int64
 3
    Maxpulse 32 non-null
                              int64
    Calories 30 non-null
 4
                              float64
dtypes: float64(1), int64(3), object(1)
memory usage: 1.4+ KB
```

dataset.describe() Duration Pulse Maxpulse Calories count 32.000000 32.000000 32.000000 30.000000 68.437500 103.500000 128,500000 266.013333 mean 7.832933 12,998759 164.876415 std 70.039591 min 30.000000 90.000000 101.000000 -300.000000 25% 60.000000 100.000000 120.000000 247.000000 60.000000 102.500000 127.500000 282.200000 50% 60.000000 132.250000 343.975000 75% 106.500000 max 450.000000 130.000000 175.000000 479.000000 dataset.isnull() Pulse Duration Date Maxpulse Calories 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 5 False False False False False 6 False False False False False 7 False False False False False 8 False False False False False 9 False False False False False 10 False False False False False 11 False False False False False 12 False False False False False False False False 13 False False 14 False False False False False 15 False False False False False 16 False False False False False 17 False False False False False 18 False False False False True 19 False False False False False 20 False False False False False 21 False False False False False 22 False True False False False 23 False False False False False 24 False False False False False 25 False False False False False 26 False False False False False 27 False False False False False False 28 False False False True 29 False False False False False 30 False False False False False 31 False False False False False

```
dataset.isnull().sum()

Duration 0
Date 1
Pulse 0
Maxpulse 0
Calories 2
dtype: int64
```

iloc shows the index location

```
dataset.iloc[:,1:3]
             Date
                   Pulse
0
    '2020/12/01'
                      110
1
    '2020/12/02'
                      117
2
    '2020/12/03'
                      103
3
    '2020/12/04'
                      109
4
    '2020/12/05'
                      117
5
    '2020/12/06'
                      102
6
    '2020/12/07'
                      110
7
    '2020/12/08'
                      104
8
    '2020/12/09'
                      109
9
                       98
    '2020/12/10'
10
    '2020/12/11'
                      103
11
    '2020/12/12'
                      100
12
    '2020/12/12'
                      100
13
    '2020/12/13'
                      106
14
                      104
    '2020/12/14'
15
                       98
    '2020/12/15'
16
                       98
    '2020/12/16'
17
    '2020/12/17'
                      100
18
    '2020/12/18'
                       90
19
    '2020/12/19'
                      103
20
    '2020/12/20'
                       97
21
    '2020/12/21'
                      108
22
              NaN
                      100
23
    '2020/12/23'
                      130
24
    '2020/12/24'
                      105
25
    '2020/12/25'
                      102
26
        20201226
                      100
27
    '2020/12/27'
                       92
28
    '2020/12/28'
                      103
29
    '2020/12/29'
                      100
30
    '2020/12/30'
                      102
31
    '2020/12/31'
                       92
dataset['Duration']
```

```
60
0
1
2
3
4
5
6
7
        60
        60
        45
        45
        60
        60
       450
8
        30
        60
10
        60
11
        60
12
        60
13
        60
14
        60
15
        60
16
        60
17
        60
        45
18
19
        60
20
        45
21
        60
22
        45
23
        60
24
        45
25
        60
26
        60
27
        60
        60
28
29
        60
30
        60
31
        60
Name: Duration, dtype: int64
print(dataset.Duration,dataset.Date)
        60
0
1
        60
2
3
4
5
6
        60
        45
        45
        60
        60
7
       450
8
        30
9
10
        60
        60
11
        60
12
        60
13
        60
```

```
14
       60
15
       60
16
       60
17
       60
18
       45
19
       60
20
       45
21
       60
22
       45
23
       60
24
       45
25
       60
26
       60
27
       60
28
       60
29
       60
30
       60
31
       60
Name: Duration, dtype: int64 0 '2020/12/01'
1
       '2020/12/02'
2
       '2020/12/03'
3
       '2020/12/04'
4
       '2020/12/05'
5
       '2020/12/06'
6
       '2020/12/07'
7
       '2020/12/08'
8
       '2020/12/09'
9
       '2020/12/10'
10
       '2020/12/11'
       '2020/12/12'
11
12
       '2020/12/12'
13
       '2020/12/13'
14
       '2020/12/14'
15
       '2020/12/15'
16
       '2020/12/16'
17
       '2020/12/17'
18
       '2020/12/18'
19
       '2020/12/19'
20
       '2020/12/20'
21
       '2020/12/21'
22
                NaN
23
       '2020/12/23'
24
       '2020/12/24'
25
       '2020/12/25'
26
           20201226
27
       '2020/12/27'
28
       '2020/12/28'
29
       '2020/12/29'
30
       '2020/12/30'
```

```
31 '2020/12/31'
Name: Date, dtype: object
```

mean() function provides average

```
X = dataset['Calories'].mean()
```

fillna() is function fill the values inside the table . inplace is used to provide specific location

```
dataset['Calories'].fillna(X,inplace = True)
print(dataset)
    Duration
                               Pulse
                                       Maxpulse
                                                     Calories
                         Date
0
           60
                '2020/12/01'
                                  110
                                             130
                                                   409.100000
1
           60
                '2020/12/02'
                                  117
                                             145
                                                   479.000000
2
           60
                '2020/12/03'
                                  103
                                             135
                                                   340.000000
3
           45
                '2020/12/04'
                                  109
                                             175
                                                   282.400000
4
           45
                '2020/12/05'
                                  117
                                             148
                                                   406.000000
5
           60
                '2020/12/06'
                                  102
                                             127
                                                  -300.000000
6
                '2020/12/07'
           60
                                  110
                                             136
                                                   374.000000
7
          450
                '2020/12/08'
                                  104
                                             134
                                                   253.300000
8
           30
                '2020/12/09'
                                  109
                                             133
                                                   195.100000
9
           60
                '2020/12/10'
                                   98
                                             124
                                                   269,000000
10
           60
                '2020/12/11'
                                  103
                                             147
                                                   329.300000
11
           60
                '2020/12/12'
                                  100
                                             120
                                                   250.700000
12
                                  100
                                             120
           60
                '2020/12/12'
                                                   250.700000
13
                '2020/12/13'
                                                   345.300000
           60
                                  106
                                             128
14
           60
                '2020/12/14'
                                             132
                                                   379.300000
                                  104
15
           60
                '2020/12/15'
                                   98
                                             123
                                                   275.000000
                                                   215.200000
16
           60
                '2020/12/16'
                                   98
                                             120
17
           60
                '2020/12/17'
                                  100
                                             120
                                                   300.000000
18
           45
                '2020/12/18'
                                   90
                                             112
                                                   266.013333
19
           60
                                             123
                                                   323,000000
                '2020/12/19'
                                  103
20
           45
                '2020/12/20'
                                   97
                                             125
                                                   243.000000
21
           60
                '2020/12/21'
                                  108
                                             131
                                                   364.200000
22
           45
                          NaN
                                  100
                                             119
                                                   282.000000
           60
23
                '2020/12/23'
                                  130
                                             101
                                                   300.000000
24
           45
                '2020/12/24'
                                  105
                                             132
                                                   246.000000
25
           60
                '2020/12/25'
                                             126
                                                   334.500000
                                  102
26
           60
                    20201226
                                  100
                                             120
                                                   250.000000
27
           60
                '2020/12/27'
                                   92
                                             118
                                                   241.000000
28
           60
                '2020/12/28'
                                  103
                                             132
                                                   266.013333
29
           60
                '2020/12/29'
                                  100
                                             132 -280.000000
30
                '2020/12/30'
                                                   380.300000
           60
                                  102
                                             129
31
           60
                '2020/12/31'
                                   92
                                             115
                                                   243.000000
```

abs() function convert negative values into positive values

```
dataset['Calories']=dataset['Calories'].abs()
dataset
    Duration
                                Pulse
                                       Maxpulse
                                                     Calories
                         Date
0
           60
                '2020/12/01'
                                  110
                                             130
                                                   409.100000
1
           60
                '2020/12/02'
                                  117
                                             145
                                                   479.000000
2
                '2020/12/03'
                                  103
                                             135
           60
                                                   340.000000
3
           45
                '2020/12/04'
                                  109
                                             175
                                                   282.400000
4
           45
                '2020/12/05'
                                  117
                                             148
                                                   406.000000
5
           60
                '2020/12/06'
                                  102
                                             127
                                                   300.000000
6
           60
                                             136
                '2020/12/07'
                                  110
                                                   374.000000
7
          450
                '2020/12/08'
                                  104
                                             134
                                                   253.300000
8
           30
                '2020/12/09'
                                  109
                                             133
                                                   195.100000
9
           60
                '2020/12/10'
                                             124
                                                   269.000000
                                   98
10
                '2020/12/11'
                                             147
                                                   329.300000
           60
                                  103
11
           60
                                  100
                                             120
                '2020/12/12'
                                                   250.700000
12
           60
                '2020/12/12'
                                  100
                                             120
                                                   250.700000
13
                                             128
           60
                '2020/12/13'
                                  106
                                                   345.300000
14
           60
                '2020/12/14'
                                  104
                                             132
                                                   379.300000
15
           60
                '2020/12/15'
                                   98
                                             123
                                                   275.000000
16
           60
                '2020/12/16'
                                   98
                                             120
                                                   215.200000
17
                '2020/12/17'
                                             120
                                                   300.000000
           60
                                  100
18
           45
                '2020/12/18'
                                   90
                                             112
                                                   266.013333
19
           60
                '2020/12/19'
                                             123
                                                   323.000000
                                  103
20
           45
                                   97
                                             125
                                                   243,000000
                '2020/12/20'
21
           60
                '2020/12/21'
                                  108
                                             131
                                                   364.200000
22
           45
                                             119
                          NaN
                                  100
                                                   282.000000
23
           60
                '2020/12/23'
                                  130
                                             101
                                                   300,000000
24
                '2020/12/24'
                                                   246.000000
           45
                                  105
                                             132
25
           60
                '2020/12/25'
                                  102
                                             126
                                                   334.500000
26
           60
                    20201226
                                  100
                                             120
                                                   250.000000
27
           60
                '2020/12/27'
                                   92
                                             118
                                                   241.000000
28
           60
                '2020/12/28'
                                  103
                                             132
                                                   266.013333
29
                                  100
                                             132
                                                   280.000000
           60
                '2020/12/29'
30
           60
                '2020/12/30'
                                  102
                                             129
                                                   380.300000
31
           60
                                   92
                                             115
                                                   243.000000
                '2020/12/31'
```

dropna() function to drop operation

```
dataset.dropna(subset=['Date'],inplace = True)
dataset
    Duration
                        Date
                               Pulse
                                       Maxpulse
                                                    Calories
0
                '2020/12/01'
                                 110
                                            130
                                                  409.100000
           60
1
                '2020/12/02'
                                 117
                                            145
                                                  479.000000
           60
2
           60
                '2020/12/03'
                                 103
                                            135
                                                  340.000000
3
                '2020/12/04'
           45
                                 109
                                            175
                                                  282.400000
4
           45
                '2020/12/05'
                                 117
                                            148
                                                  406.000000
```

```
5
           60
                '2020/12/06'
                                  102
                                             127
                                                   300.000000
6
                '2020/12/07'
                                             136
           60
                                  110
                                                   374.000000
7
          450
                '2020/12/08'
                                  104
                                             134
                                                   253.300000
8
           30
                '2020/12/09'
                                  109
                                             133
                                                   195.100000
9
           60
                '2020/12/10'
                                   98
                                             124
                                                   269.000000
10
           60
                '2020/12/11'
                                  103
                                             147
                                                   329.300000
11
           60
                '2020/12/12'
                                  100
                                             120
                                                   250.700000
12
           60
                '2020/12/12'
                                  100
                                             120
                                                   250.700000
13
                '2020/12/13'
                                             128
                                                   345.300000
           60
                                  106
14
           60
                '2020/12/14'
                                  104
                                             132
                                                   379.300000
15
                '2020/12/15'
                                   98
                                                   275.000000
           60
                                             123
16
           60
                '2020/12/16'
                                   98
                                             120
                                                   215.200000
17
                '2020/12/17'
                                             120
                                                   300.000000
           60
                                  100
18
           45
                '2020/12/18'
                                   90
                                             112
                                                   266.013333
19
           60
                '2020/12/19'
                                  103
                                             123
                                                   323.000000
20
           45
                '2020/12/20'
                                             125
                                                   243.000000
                                   97
21
           60
                '2020/12/21'
                                  108
                                             131
                                                   364.200000
23
           60
                '2020/12/23'
                                             101
                                                   300.000000
                                  130
           45
24
                '2020/12/24'
                                  105
                                             132
                                                   246.000000
25
           60
                '2020/12/25'
                                  102
                                             126
                                                   334.500000
26
                                             120
           60
                    20201226
                                  100
                                                   250.000000
27
           60
                '2020/12/27'
                                   92
                                             118
                                                   241.000000
28
           60
                '2020/12/28'
                                  103
                                             132
                                                   266.013333
29
           60
                '2020/12/29'
                                  100
                                             132
                                                   280.000000
30
           60
                '2020/12/30'
                                  102
                                             129
                                                   380.300000
31
           60
                '2020/12/31'
                                   92
                                             115
                                                   243.000000
dataset.dtypes
Duration
                int64
               object
Date
Pulse
                int64
Maxpulse
                int64
Calories
             float64
dtype: object
```

to update date in specified format use to date

```
dataset['Date'] = pd.to datetime(dataset['Date'], format = 'mixed')
dataset
                                   Maxpulse
    Duration
                    Date
                           Pulse
                                                Calories
0
           60 2020-12-01
                             110
                                              409.100000
                                        130
           60 2020-12-02
                                        145
1
                             117
                                              479.000000
2
           60 2020-12-03
                             103
                                        135
                                              340.000000
3
           45 2020-12-04
                             109
                                        175
                                              282.400000
4
           45 2020-12-05
                             117
                                        148
                                              406.000000
5
           60 2020-12-06
                             102
                                        127
                                              300.000000
6
           60 2020-12-07
                             110
                                        136
                                              374.000000
```

7	450 2020-12-08	104	134	253.300000	
8	30 2020-12-09	109	133	195.100000	
9	60 2020-12-10	98	124	269.000000	
10	60 2020-12-11	103	147	329.300000	
11	60 2020-12-12	100	120	250.700000	
12	60 2020-12-12	100	120	250.700000	
13	60 2020-12-13	106	128	345.300000	
14	60 2020-12-14	104	132	379.300000	
15	60 2020-12-15	98	123	275.000000	
16	60 2020-12-16	98	120	215.200000	
17	60 2020-12-17	100	120	300.000000	
18	45 2020-12-18	90	112	266.013333	
19	60 2020-12-19	103	123	323.000000	
20	45 2020-12-20	97	125	243.000000	
21	60 2020-12-21	108	131	364.200000	
23	60 2020-12-23	130	101	300.000000	
24	45 2020-12-24	105	132	246.000000	
25	60 2020-12-25	102	126	334.500000	
26	60 2020-12-26	100	120	250.000000	
27	60 2020-12-27	92	118	241.000000	
28	60 2020-12-28	103	132	266.013333	
29	60 2020-12-29	100	132	280.000000	
30	60 2020-12-30	102	129	380.300000	
31	60 2020-12-31	92	115	243.000000	

astype(int) convert data entries into integer

```
dataset['Calories'] = dataset['Calories'].astype(int)
dataset
    Duration
                     Date
                           Pulse
                                   Maxpulse
                                              Calories
0
           60 2020-12-01
                              110
                                         130
                                                    409
1
                                                    479
           60 2020-12-02
                              117
                                         145
2
           60 2020-12-03
                              103
                                         135
                                                    340
3
           45 2020-12-04
                                                    282
                              109
                                         175
4
           45 2020-12-05
                              117
                                         148
                                                    406
5
6
           60 2020-12-06
                              102
                                         127
                                                    300
           60 2020-12-07
                              110
                                         136
                                                    374
7
          450 2020-12-08
                              104
                                         134
                                                    253
8
           30 2020-12-09
                              109
                                         133
                                                    195
9
           60 2020-12-10
                               98
                                         124
                                                    269
10
           60 2020-12-11
                              103
                                         147
                                                    329
11
           60 2020-12-12
                              100
                                         120
                                                    250
12
           60 2020-12-12
                              100
                                         120
                                                    250
13
           60 2020-12-13
                              106
                                         128
                                                    345
14
           60 2020-12-14
                              104
                                                    379
                                         132
15
           60 2020-12-15
                               98
                                         123
                                                    275
           60 2020-12-16
                                         120
                                                    215
16
                               98
```

17	60 2020-12-17	100	120	300
18	45 2020-12-18	90	112	266
19	60 2020-12-19	103	123	323
20	45 2020-12-20	97	125	243
21	60 2020-12-21	108	131	364
23	60 2020-12-23	130	101	300
24	45 2020-12-24	105	132	246
25	60 2020-12-25	102	126	334
26	60 2020-12-26	100	120	250
27	60 2020-12-27	92	118	241
28	60 2020-12-28	103	132	266
29	60 2020-12-29	100	132	280
30	60 2020-12-30	102	129	380
31	60 2020-12-31	92	115	243

updating the value of specific location

```
dataset.loc[7, 'Duration']=45
dataset
                                                Calories
    Duration
                      Date
                            Pulse
                                    Maxpulse
           60 2020-12-01
0
                               110
                                          130
                                                     409
1
           60 2020-12-02
                                          145
                                                     479
                               117
2
                                          135
                                                     340
           60 2020-12-03
                               103
3
           45 2020-12-04
                               109
                                          175
                                                     282
4
           45 2020-12-05
                               117
                                          148
                                                     406
5
           60 2020-12-06
                               102
                                          127
                                                     300
6
           60 2020-12-07
                               110
                                          136
                                                     374
7
           45 2020-12-08
                                                     253
                               104
                                          134
8
           30 2020-12-09
                               109
                                          133
                                                     195
9
           60 2020-12-10
                                          124
                                98
                                                     269
10
           60 2020-12-11
                               103
                                          147
                                                     329
11
           60 2020-12-12
                               100
                                          120
                                                     250
12
           60 2020-12-12
                                                     250
                               100
                                          120
13
           60 2020-12-13
                                          128
                               106
                                                     345
14
           60 2020-12-14
                                                     379
                               104
                                          132
15
           60 2020-12-15
                                98
                                          123
                                                     275
16
           60 2020-12-16
                                98
                                          120
                                                     215
17
           60 2020-12-17
                               100
                                          120
                                                     300
18
           45 2020-12-18
                                          112
                                90
                                                     266
19
           60 2020-12-19
                               103
                                          123
                                                     323
20
                                97
           45 2020-12-20
                                          125
                                                     243
21
           60 2020-12-21
                               108
                                          131
                                                     364
23
           60 2020-12-23
                               130
                                          101
                                                     300
24
           45 2020-12-24
                                          132
                                                     246
                               105
25
           60 2020-12-25
                               102
                                          126
                                                     334
26
           60 2020-12-26
                               100
                                          120
                                                     250
           60 2020-12-27
27
                                92
                                          118
                                                     241
```

```
28
          60 2020-12-28
                             103
                                       132
                                                  266
29
          60 2020-12-29
                                       132
                                                  280
                             100
                                       129
30
          60 2020-12-30
                             102
                                                  380
31
          60 2020-12-31
                             92
                                       115
                                                  243
dataset.duplicated()
0
      False
1
      False
2
      False
3
      False
4
      False
5
6
      False
      False
7
      False
8
      False
9
      False
10
      False
11
      False
12
       True
13
      False
      False
14
15
      False
16
      False
17
      False
18
      False
19
      False
20
      False
21
      False
23
      False
24
      False
25
      False
26
      False
27
      False
28
      False
29
      False
30
      False
      False
31
dtype: bool
dataset.duplicated().sum()
1
dataset.drop duplicates(inplace = True)
dataset
                                  Maxpulse
                                             Calories
    Duration
                    Date
                           Pulse
0
          60 2020-12-01
                             110
                                       130
                                                  409
1
          60 2020-12-02
                             117
                                       145
                                                  479
```

```
2
           60 2020-12-03
                              103
                                          135
                                                     340
3
           45 2020-12-04
                              109
                                          175
                                                     282
4
           45 2020-12-05
                              117
                                          148
                                                     406
5
           60 2020-12-06
                                          127
                              102
                                                     300
6
           60 2020-12-07
                              110
                                          136
                                                     374
7
           45 2020-12-08
                              104
                                         134
                                                     253
8
           30 2020-12-09
                              109
                                         133
                                                     195
9
           60 2020-12-10
                               98
                                         124
                                                     269
10
           60 2020-12-11
                              103
                                          147
                                                     329
11
           60 2020-12-12
                              100
                                          120
                                                     250
           60 2020-12-13
                              106
                                                     345
13
                                         128
           60 2020-12-14
14
                              104
                                         132
                                                     379
15
           60 2020-12-15
                               98
                                         123
                                                     275
           60 2020-12-16
                               98
                                         120
                                                     215
16
17
           60 2020-12-17
                              100
                                          120
                                                     300
           45 2020-12-18
18
                               90
                                         112
                                                     266
19
           60 2020-12-19
                                         123
                              103
                                                     323
20
           45 2020-12-20
                               97
                                          125
                                                     243
21
           60 2020-12-21
                              108
                                         131
                                                     364
23
           60 2020-12-23
                              130
                                         101
                                                     300
24
           45 2020-12-24
                                         132
                                                     246
                              105
25
           60 2020-12-25
                              102
                                                     334
                                         126
26
           60 2020-12-26
                              100
                                          120
                                                     250
27
           60 2020-12-27
                               92
                                         118
                                                     241
28
           60 2020-12-28
                              103
                                         132
                                                     266
29
           60 2020-12-29
                              100
                                         132
                                                     280
30
           60 2020-12-30
                              102
                                          129
                                                     380
31
           60 2020-12-31
                               92
                                         115
                                                     243
dataset.duplicated().sum()
0
```

performing the or operation

```
dataset[(dataset['Duration']==45)|(dataset['Maxpulse']==120)]
    Duration
                     Date
                           Pulse
                                   Maxpulse
                                              Calories
3
           45 2020-12-04
                             109
                                        175
                                                   282
4
           45 2020-12-05
                             117
                                        148
                                                   406
7
          45 2020-12-08
                             104
                                        134
                                                   253
11
          60 2020-12-12
                             100
                                        120
                                                   250
           60 2020-12-16
16
                              98
                                        120
                                                   215
17
           60 2020-12-17
                             100
                                        120
                                                   300
18
           45 2020-12-18
                              90
                                        112
                                                   266
20
          45 2020-12-20
                              97
                                        125
                                                   243
24
          45 2020-12-24
                                        132
                                                   246
                             105
          60 2020-12-26
26
                             100
                                        120
                                                   250
df1 = dataset.loc[:,['Duration','Calories']]
```

```
df1
    Duration
                Calories
                      409
0
           60
1
           60
                      479
2
           60
                      340
3
4
           45
                      282
           45
                      406
5
           60
                      300
6
           60
                      374
7
           45
                      253
8
           30
                      195
9
           60
                      269
10
           60
                      329
11
           60
                      250
13
           60
                      345
14
           60
                      379
15
                      275
           60
16
                      215
           60
17
           60
                      300
18
           45
                      266
19
           60
                      323
20
           45
                      243
21
           60
                      364
23
           60
                      300
24
           45
                      246
25
           60
                      334
26
           60
                      250
27
           60
                      241
28
           60
                      266
29
           60
                      280
30
           60
                      380
31
           60
                      243
```

selecting all colums except maxpulse

```
df2 = dataset.loc[:,dataset.columns!= 'Maxpulse']
df2
    Duration
                                  Calories
                    Date
                          Pulse
0
          60 2020-12-01
                             110
                                       409
          60 2020-12-02
1
                             117
                                       479
2
          60 2020-12-03
                                       340
                             103
3
          45 2020-12-04
                             109
                                       282
4
          45 2020-12-05
                             117
                                       406
5
          60 2020-12-06
                             102
                                       300
6
          60 2020-12-07
                            110
                                       374
7
          45 2020-12-08
                             104
                                       253
8
          30 2020-12-09
                             109
                                       195
```

g)	60	2020-12-10	98	269
1	LO	60	2020-12-11	103	329
1	L1	60	2020-12-12	100	250
1	L3	60	2020-12-13	106	345
1	L4	60	2020-12-14	104	379
1	L5	60	2020-12-15	98	275
1	L6	60	2020-12-16	98	215
1	L7	60	2020-12-17	100	300
1	L8	45	2020-12-18	90	266
1	L9	60	2020-12-19	103	323
2	20	45	2020-12-20	97	243
2	21	60	2020-12-21	108	364
2	23	60	2020-12-23	130	300
2	24	45	2020-12-24	105	246
2	25	60	2020-12-25	102	334
2	26	60	2020-12-26	100	250
2	27	60	2020-12-27	92	241
2	28	60	2020-12-28	103	266
2	29	60	2020-12-29	100	280
3	30	60	2020-12-30	102	380
3	31	60	2020-12-31	92	243

updating the name of column maxpulse to new

```
df2['new']=dataset['Maxpulse']
df2
                                    Calories
    Duration
                            Pulse
                     Date
                                               new
0
           60 2020-12-01
                                               130
                              110
                                         409
1
           60 2020-12-02
                              117
                                         479
                                               145
2
           60 2020-12-03
                                         340
                                               135
                              103
3
           45 2020-12-04
                              109
                                         282
                                               175
4
           45 2020-12-05
                              117
                                         406
                                               148
5
           60 2020-12-06
                              102
                                         300
                                               127
6
           60 2020-12-07
                                               136
                              110
                                         374
7
           45 2020-12-08
                                         253
                                               134
                              104
8
           30 2020-12-09
                              109
                                         195
                                               133
9
           60 2020-12-10
                               98
                                         269
                                               124
10
           60 2020-12-11
                              103
                                         329
                                               147
11
           60 2020-12-12
                              100
                                         250
                                               120
13
           60 2020-12-13
                                         345
                              106
                                               128
14
                              104
                                         379
                                               132
           60 2020-12-14
15
           60 2020-12-15
                               98
                                         275
                                               123
16
           60 2020-12-16
                               98
                                         215
                                               120
17
           60 2020-12-17
                              100
                                         300
                                               120
18
           45 2020-12-18
                               90
                                         266
                                               112
19
           60 2020-12-19
                              103
                                         323
                                               123
20
           45 2020-12-20
                               97
                                         243
                                               125
```

```
21
           60 2020-12-21
                              108
                                         364
                                               131
23
           60 2020-12-23
                                               101
                              130
                                          300
24
           45 2020-12-24
                              105
                                         246
                                               132
25
           60 2020-12-25
                              102
                                         334
                                               126
26
           60 2020-12-26
                              100
                                         250
                                               120
27
           60 2020-12-27
                               92
                                         241
                                               118
28
           60 2020-12-28
                              103
                                         266
                                               132
29
           60 2020-12-29
                              100
                                         280
                                               132
30
           60 2020-12-30
                              102
                                         380
                                               129
31
           60 2020-12-31
                               92
                                         243
                                               115
dataset
    Duration
                     Date
                            Pulse
                                    Maxpulse
                                               Calories
           60 2020-12-01
0
                              110
                                          130
                                                     409
1
           60 2020-12-02
                              117
                                          145
                                                     479
2
           60 2020-12-03
                                         135
                                                     340
                              103
3
           45 2020-12-04
                                         175
                              109
                                                     282
4
           45 2020-12-05
                              117
                                         148
                                                     406
5
           60 2020-12-06
                              102
                                         127
                                                     300
6
           60 2020-12-07
                              110
                                         136
                                                     374
7
           45 2020-12-08
                              104
                                         134
                                                     253
8
           30 2020-12-09
                                         133
                              109
                                                     195
9
           60 2020-12-10
                               98
                                         124
                                                     269
10
           60 2020-12-11
                              103
                                         147
                                                     329
11
           60 2020-12-12
                              100
                                         120
                                                     250
13
           60 2020-12-13
                              106
                                         128
                                                     345
           60 2020-12-14
14
                              104
                                         132
                                                     379
15
           60 2020-12-15
                               98
                                         123
                                                     275
16
           60 2020-12-16
                               98
                                         120
                                                     215
17
           60 2020-12-17
                              100
                                         120
                                                     300
                               90
18
           45 2020-12-18
                                         112
                                                     266
19
           60 2020-12-19
                              103
                                         123
                                                     323
20
           45 2020-12-20
                               97
                                         125
                                                     243
21
           60 2020-12-21
                              108
                                          131
                                                     364
23
           60 2020-12-23
                              130
                                          101
                                                     300
24
           45 2020-12-24
                              105
                                         132
                                                     246
25
           60 2020-12-25
                              102
                                         126
                                                     334
26
           60 2020-12-26
                              100
                                         120
                                                     250
27
           60 2020-12-27
                               92
                                         118
                                                     241
28
           60 2020-12-28
                              103
                                         132
                                                     266
29
           60 2020-12-29
                              100
                                         132
                                                     280
30
           60 2020-12-30
                              102
                                         129
                                                     380
31
           60 2020-12-31
                               92
                                         115
                                                     243
```

various operations on nba dataset

```
import pandas as pd

df = pd.read_csv("nba.csv")
```

df									
Weight	. \	Name		Team	Number	Position	Age	Height	
_		Bradley	Boston	Celtics	0.0	PG	25.0	6-2	
1	Jae	Crowder	Boston	Celtics	99.0	SF	25.0	6-6	
	John	Holland	Boston	Celtics	30.0	SG	27.0	6-5	
205.0	R.J.	Hunter	Boston	Celtics	28.0	SG	22.0	6-5	
185.0 4 J 231.0	lonas	Jerebko	Boston	Celtics	8.0	PF	29.0	6-10	
	Shelv	/in Mack	U	tah Jazz	8.0	PG	26.0	6-3	
454	Ra	aul Neto	U	tah Jazz	25.0	PG	24.0	6-1	
	Tibor	Pleiss	U	tah Jazz	21.0	С	26.0	7-3	
256.0 456	Jeff	f Withey	U	tah Jazz	24.0	С	26.0	7 - 0	
231.0 457		NaN		NaN	NaN	NaN	NaN	NaN	
NaN									
3 4 453		Marque Dnivers Porgia St But	xas 77 tte 67 ity ate 11 NaN 50 	Salary 30337.0 96117.0 NaN 48640.0 00000.0					
454 455 456 457		Kan	NaN 29	00000.0 00000.0 47276.0 NaN					
[458 r	OWS >	c 9 colum	ns]						
df.dty	pes								
Name Team Number Positi Age Height	.on	object object int64 object int64 object							

```
Weight
              int64
College
             object
Salary
            float64
dtype: object
df['Position'].unique()
array(['PG', 'SF', 'SG', 'PF', 'C'], dtype=object)
dfn = df
dfn['Position'].replace(['PG','SF','SG','PF','C'],[0,1,2,3,4],inplace
= True)
/tmp/ipykernel 5112/2764792939.py:2: FutureWarning: A value is trying
to be set on a copy of a DataFrame or Series through chained
assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never
work because the intermediate object on which we are setting values
always behaves as a copy.
For example, when doing 'df[col].method(value, inplace=True)', try
using 'df.method({col: value}, inplace=True)' or df[col] =
df[col].method(value) instead, to perform the operation inplace on the
original object.
 dfn['Position'].replace(['PG','SF','SG','PF','C'],
[0,1,2,3,4], inplace = True)
/tmp/ipykernel 5112/2764792939.py:2: FutureWarning: Downcasting
behavior in `replace` is deprecated and will be removed in a future
version. To retain the old behavior, explicitly call
`result.infer objects(copy=False)`. To opt-in to the future behavior,
set `pd.set option('future.no silent downcasting', True)`
  dfn['Position'].replace(['PG','SF','SG','PF','C'],
[0,1,2,3,4], inplace = True)
dfn
                              Team Number Position Age Height
              Name
Weiaht \
     Avery Bradley Boston Celtics
                                         0
                                                   0
                                                       25
                                                            2-Jun
180
       Jae Crowder Boston Celtics
                                        99
                                                   1
                                                       25
1
                                                            6-Jun
235
2
      John Holland Boston Celtics
                                        30
                                                   2
                                                       27
                                                            5-Jun
205
3
       R.J. Hunter Boston Celtics
                                        28
                                                   2
                                                       22
                                                            5-Jun
185
     Jonas Jerebko Boston Celtics
                                         8
                                                   3
                                                       29
4
                                                           10-Jun
231
```

```
452
        Trey Lyles
                          Utah Jazz
                                          41
                                                      3
                                                          20
                                                              10-Jun
234
      Shelvin Mack
                          Utah Jazz
453
                                           8
                                                          26
                                                               3-Jun
203
454
         Raul Neto
                          Utah Jazz
                                          25
                                                      0
                                                          24
                                                               1-Jun
179
455
      Tibor Pleiss
                          Utah Jazz
                                          21
                                                          26
                                                               3-Jul
256
       Jeff Withey
                          Utah Jazz
456
                                          24
                                                          26
                                                              Jul-00
231
               College
                            Salary
0
                  Texas
                         7730337.0
1
             Marquette
                         6796117.0
2
     Boston University
                               NaN
3
         Georgia State
                         1148640.0
4
                    NaN
                         5000000.0
452
              Kentucky
                         2239800.0
453
                Butler
                         2433333.0
454
                    NaN
                          900000.0
455
                    NaN
                         2900000.0
456
                Kansas
                          947276.0
[457 rows x 9 columns]
from sklearn import preprocessing
lbt = preprocessing.LabelEncoder()
df['Position'] = lbt.fit transform(df['Position'])
df
                                              Position Age
              Name
                               Team
                                      Number
                                                              Height
Weight \
     Avery Bradley
                     Boston Celtics
                                                          25
                                           0
                                                               2-Jun
180
1
       Jae Crowder
                     Boston Celtics
                                          99
                                                      1
                                                          25
                                                               6-Jun
235
                                                      2
2
      John Holland Boston Celtics
                                          30
                                                          27
                                                               5-Jun
205
                                                      2
       R.J. Hunter Boston Celtics
                                          28
                                                          22
3
                                                               5-Jun
185
     Jonas Jerebko Boston Celtics
                                                      3
                                                          29
                                           8
                                                              10-Jun
231
. . .
```

Utah Jazz

41

20

3

10-Jun

Trey Lyles

452

234

```
453
      Shelvin Mack
                          Utah Jazz
                                           8
                                                     0
                                                          26
                                                               3-Jun
203
454
         Raul Neto
                          Utah Jazz
                                          25
                                                          24
                                                               1-Jun
179
      Tibor Pleiss
455
                          Utah Jazz
                                          21
                                                          26
                                                               3-Jul
256
       Jeff Withey
                          Utah Jazz
                                          24
                                                     4
                                                          26 Jul-00
456
231
               College
                            Salary
0
                 Texas
                         7730337.0
1
             Marquette
                         6796117.0
2
     Boston University
                               NaN
3
         Georgia State
                         1148640.0
4
                    NaN
                         5000000.0
. .
452
              Kentucky
                         2239800.0
453
                Butler
                         2433333.0
454
                    NaN
                          900000.0
455
                         2900000.0
                    NaN
456
                Kansas
                          947276.0
[457 rows x 9 columns]
df['Position'].unique()
array([0, 1, 2, 3, 4])
df['Position'].value_counts()
Position
2
     102
3
     100
      92
0
1
      85
4
      78
Name: count, dtype: int64
df['Age'].dropna(inplace = True)
df
              Name
                               Team
                                      Number
                                              Position
                                                        Age
                                                              Height
Weight
     Avery Bradley Boston Celtics
                                                     0
                                                          25
                                           0
                                                               2-Jun
180
       Jae Crowder Boston Celtics
                                          99
                                                     1
                                                          25
                                                               6-Jun
1
235
2
      John Holland Boston Celtics
                                          30
                                                     2
                                                          27
                                                               5-Jun
205
```

28

2

22

5-Jun

3

R.J. Hunter Boston Celtics

```
185
     Jonas Jerebko Boston Celtics
                                           8
                                                     3
                                                          29 10-Jun
4
231
. .
. . .
452
        Trey Lyles
                          Utah Jazz
                                          41
                                                     3
                                                          20
                                                              10-Jun
234
453
      Shelvin Mack
                          Utah Jazz
                                           8
                                                          26
                                                               3-Jun
203
                          Utah Jazz
454
         Raul Neto
                                          25
                                                     0
                                                          24
                                                               1-Jun
179
      Tibor Pleiss
455
                          Utah Jazz
                                          21
                                                          26
                                                               3-Jul
256
       Jeff Withey
                          Utah Jazz
                                          24
                                                          26 Jul-00
456
231
               College
                            Salary
0
                 Texas
                         7730337.0
1
             Marquette
                         6796117.0
2
     Boston University
                               NaN
3
         Georgia State
                         1148640.0
4
                    NaN
                         5000000.0
. .
                         2239800.0
452
              Kentucky
453
                Butler
                         2433333.0
454
                          900000.0
                    NaN
455
                    NaN
                         2900000.0
456
                Kansas
                          947276.0
[457 rows x 9 columns]
df['Age'].isnull().sum()
0
category =
pd.cut(df.Age,bins=[19,25,30,35,45],labels=['A','B','C','D'])
df.insert(5, 'Age1', category)
df
                                     Number
                                              Position Age Age1
                                                                   Height
              Name
                               Team
     Avery Bradley Boston Celtics
                                                     0
                                                          25
                                           0
                                                                Α
                                                                    2-Jun
                                          99
                                                     1
1
       Jae Crowder
                    Boston Celtics
                                                          25
                                                                Α
                                                                    6-Jun
      John Holland Boston Celtics
                                          30
                                                     2
                                                          27
                                                                В
                                                                    5-Jun
       R.J. Hunter
                    Boston Celtics
                                                     2
3
                                          28
                                                          22
                                                                Α
                                                                    5-Jun
```

4	Jonas Jerebko E	Boston Celtics	8	3	29	В	10-Jun
452	Trey Lyles	Utah Jazz	41	3	20	Α	10-Jun
453	Shelvin Mack	Utah Jazz	8	0	26	В	3-Jun
454	Raul Neto	Utah Jazz	25	0	24	Α	1-Jun
455	Tibor Pleiss	Utah Jazz	21	4	26	В	3-Jul
456	Jeff Withey	Utah Jazz	24	4	26	В	Jul-00
_		Texas 77303 Marquette 67961 University Tgia State 11486 NaN 50000 Kentucky 22398 Butler 24333 NaN 9000 NaN 29000 Kansas 9472	17.0 NaN 40.0 00.0 00.0 33.0 00.0				
	Name	Team N	umber	Posit	ion	Age	9
Heig 0 2		Boston Celtics	0.0		PG	25.0	0 6-

	Name	Team	Number	Position	Age	
Heig	ht \				_	
0	Avery Bradley	Boston Celtics	0.0	PG	25.0	6 -
2						
1	Jae Crowder	Boston Celtics	99.0	SF	25.0	6 -
6						
2	John Holland	Boston Celtics	30.0	SG	27.0	6 -
5						
3	R.J. Hunter	Boston Celtics	28.0	SG	22.0	6 -
5						
4	Jonas Jerebko	Boston Celtics	8.0	PF	29.0	6 -
10						
453	Shelvin Mack	Utah Jazz	8.0	PG	26.0	6 -
3						

```
454
         Raul Neto
                          Utah Jazz
                                        25.0
                                                           PG 24.0
                                                                       6-
1
455
      Tibor Pleiss
                          Utah Jazz
                                        21.0
                                                               26.0
                                                                       7 -
3
456
       Jeff Withey
                          Utah Jazz
                                        24.0
                                                               26.0
                                                                       7 -
457
                                              5077829.215909
               NaN
                                NaN
                                         NaN
                                                                NaN
NaN
     Weight
                    Salary
0
      180.0
             7.730337e+06
1
      235.0
             6.796117e+06
      205.0
2
             5.077829e+06
3
      185.0
             1.148640e+06
4
      231.0
             5.000000e+06
. .
      203.0
             2.433333e+06
453
454
      179.0
             9.000000e+05
455
      256.0
             2.900000e+06
      231.0
             9.472760e+05
456
             5.077829e+06
457
        NaN
[458 rows x 8 columns]
PG = df[(df['Position']=='PG')]
PG
              Name
                                        Team
                                             Number Position Age
Height \
     Avery Bradley
                             Boston Celtics
                                                 0.0
                                                            PG
                                                                25.0
6-2
                             Boston Celtics
                                                            PG
      Terry Rozier
                                                12.0
                                                                22.0
8
6-2
      Marcus Smart
                             Boston Celtics
                                                36.0
                                                            PG
                                                                22.0
9
6-4
11
     Isaiah Thomas
                             Boston Celtics
                                                 4.0
                                                            PG
                                                                27.0
5-9
19
      Jarrett Jack
                              Brooklyn Nets
                                                 2.0
                                                            PG
                                                                32.0
6-3
. .
                     Portland Trail Blazers
440
     Brian Roberts
                                                 2.0
                                                            PG
                                                                30.0
6-1
443
        Trey Burke
                                  Utah Jazz
                                                 3.0
                                                            PG
                                                                23.0
6-1
445
                                  Utah Jazz
                                                            PG
                                                                20.0
        Dante Exum
                                                11.0
6-6
453
      Shelvin Mack
                                  Utah Jazz
                                                 8.0
                                                            PG 26.0
6-3
```

```
454
         Raul Neto
                                 Utah Jazz
                                               25.0
                                                          PG 24.0
6 - 1
                    College
     Weight
                                Salary
      180.0
                             7730337.0
0
                      Texas
8
      190.0
                 Louisville
                             1824360.0
9
      220.0
             Oklahoma State
                             3431040.0
11
      185.0
                 Washington
                             6912869.0
19
      200.0
               Georgia Tech
                             6300000.0
440
      173.0
                     Dayton
                             2854940.0
443
      191.0
                   Michigan
                             2658240.0
445
      190.0
                        NaN
                             3777720.0
453
      203.0
                     Butler
                             2433333.0
454
      179.0
                        NaN
                              900000.0
[92 rows x 9 columns]
X = PG['Salary'].mean()
Χ
5077829.215909091
PG['Salary'].fillna(X,inplace = True)
/tmp/ipykernel 6266/2443542409.py:1: FutureWarning: A value is trying
to be set on a copy of a DataFrame or Series through chained
assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never
work because the intermediate object on which we are setting values
always behaves as a copy.
For example, when doing 'df[col].method(value, inplace=True)', try
using 'df.method({col: value}, inplace=True)' or df[col] =
df[col].method(value) instead, to perform the operation inplace on the
original object.
  PG['Salary'].fillna(X,inplace = True)
/tmp/ipykernel 6266/2443542409.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
See the caveats in the documentation:
https://pandas.pydata.org/pandas-docs/stable/user guide/indexing.html#
returning-a-view-versus-a-copy
  PG['Salary'].fillna(X,inplace = True)
PG
```

```
Number Position
              Name
                                        Team
                                                                 Age
Height \
     Avery Bradley
                             Boston Celtics
                                                 0.0
                                                            PG
                                                                25.0
6-2
8
      Terry Rozier
                             Boston Celtics
                                                12.0
                                                            PG
                                                                22.0
6-2
                             Boston Celtics
                                                            PG
                                                                22.0
9
      Marcus Smart
                                                36.0
6-4
     Isaiah Thomas
                             Boston Celtics
                                                 4.0
                                                            PG
11
                                                                27.0
5-9
19
      Jarrett Jack
                              Brooklyn Nets
                                                 2.0
                                                            PG
                                                                32.0
6-3
. .
440
     Brian Roberts
                     Portland Trail Blazers
                                                 2.0
                                                            PG
                                                                30.0
6-1
443
        Trey Burke
                                  Utah Jazz
                                                 3.0
                                                            PG
                                                                23.0
6-1
445
        Dante Exum
                                  Utah Jazz
                                                            PG
                                                                20.0
                                                11.0
6-6
      Shelvin Mack
                                  Utah Jazz
                                                            PG
453
                                                 8.0
                                                                26.0
6-3
454
         Raul Neto
                                  Utah Jazz
                                                25.0
                                                            PG 24.0
6-1
                     College
     Weight
                                 Salary
0
      180.0
                       Texas
                              7730337.0
8
      190.0
                 Louisville
                              1824360.0
9
      220.0
             Oklahoma State
                              3431040.0
11
      185.0
                 Washington
                              6912869.0
               Georgia Tech
19
      200.0
                              6300000.0
440
      173.0
                      Dayton
                              2854940.0
443
      191.0
                    Michigan
                              2658240.0
445
      190.0
                         NaN
                              3777720.0
453
      203.0
                      Butler
                              2433333.0
454
      179.0
                         NaN
                               900000.0
[92 rows x 9 columns]
C = df[(df['Position']=='C')]
SF = df[(df['Position']=='SF')]
SG = df[(df['Position']=='SG')]
y= C['Salary'].mean()
z = SF['Salary'].mean()
w=SG['Salary'].mean()
SG['Salary'].fillna(w,inplace = True)
SF['Salary'].fillna(z,inplace = True)
C['Salary'].fillna(y,inplace = True)
```

/tmp/ipykernel_6266/3803027798.py:1: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.

The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.

SG['Salary'].fillna(w,inplace = True)
/tmp/ipykernel_6266/3803027798.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation:

https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

SG['Salary'].fillna(w,inplace = True)

/tmp/ipykernel_6266/3803027798.py:2: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.

The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.

SF['Salary'].fillna(z,inplace = True)
/tmp/ipykernel_6266/3803027798.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation:

https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

SF['Salary'].fillna(z,inplace = True)

/tmp/ipykernel_6266/3803027798.py:3: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.

The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] =

df[col].method(value) instead, to perform the operation inplace on the original object.

C['Salary'].fillna(y,inplace = True)
/tmp/ipykernel_6266/3803027798.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation:

https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#
returning-a-view-versus-a-copy

C['Salary'].fillna(y,inplace = True)

SF SG C

		Name			Т	eam	Number	Position	Age
Heigl 7		ly Olynyk		Boston	Cel+	ics	41.0	С	25.0
, 7 - 0	ite c	cy ocyllyk		DOSCOII		103	41.0		23.0
10 6-9	Jared	Sullinger		Boston	Celt	ics	7.0	С	24.0
14	Tyl	er Zeller		Boston	Celt	ics	44.0	С	26.0
7-0									20.0
23 7-0	Br	ook Lopez		Brookl	.yn N	ets	11.0	С	28.0
27	Н	enry Sims		Brookl	vn N	ets	14.0	С	26.0
6-10	•••	CITT y 313		Dioon	. y		1110		2010
424	Ch	nic Vomen	Portland	Troil	D1	0.00	25.0	C	34.0
434 7-0	CII	ris Kaman	Porttand	ITALL	BLaZ	ers	35.0	С	34.0
439	Maso	n Plumlee	Portland	Trail	Blaz	ers	24.0	С	26.0
6-11									
447	Ru	dy Gobert		Ut	ah J	azz	27.0	С	23.0
7 - 1 455	Tib	or Pleiss		114	ah J	277	21.0	С	26.0
455 7-3	IID	or Preiss		U	.ali J	azz	21.0	C	20.0
456	Je	ff Withey		Ut	ah J	azz	24.0	С	26.0
7 - 0		-							
	الحامة أحمادا		C-11	C-	. 1				
7	Weight 238.0		College Gonzaga	21651	lary				
10	260.0	Ok	nio State	25692					
14	253.0		Carolina	26169					
23	275.0	_	Stanford	196890					
27	248.0		eorgetown		76.0				
434	265.0	Central	Michigan	50160	0.00				

439	235.0			Duke	1415520.0
447	245.0			NaN	1175880.0
455	256.0			NaN	2900000.0
456	231.0			Kansas	947276.0
	_	_	_		

[78 rows x 9 columns]

SG

	Name			Team	Number	Position	Age
Heig	ht \						
2	John Holland	[Boston	Celtics	30.0	SG	27.0
6-5		_					
3	R.J. Hunter	l	Boston	Celtics	28.0	SG	22.0
6-5 12	Evan Turner		Poston	Celtics	11.0	SG	27.0
6-7	Evali Turrier		505 (011	Cettics	11.0	30	27.0
13	James Young	ı	Boston	Celtics	13.0	SG	20.0
6-6							
15	Bojan Bogdanovic		Brook1	lyn Nets	44.0	SG	27.0
6-8							
422	Canald Handanaan	Daust Laurel	T 21	D1	0 0	CC	20.0
433 6-5	Gerald Henderson	Portland	irait	Blazers	9.0	SG	28.0
437	C.J. McCollum	Portland	Trail	Rlazers	3.0	SG	24.0
6-4	CISI HECOCCUIII	1 of ccana	IIGIC	DCUZCIS	3.0	30	24.0
438	Luis Montero	Portland	Trail	Blazers	44.0	SG	23.0
6-7							
444	Alec Burks		Ut	tah Jazz	10.0	SG	24.0
6-6							
449	Rodney Hood		U1	tah Jazz	5.0	SG	23.0
6-8							
	Weight	College		Salary			
2		iversity	5.0778	329e+06			
3		ia Staté		640e+06			

	Weight	College	Salary
2	205.0	Boston University	5.077829e+06
3	185.0	Georgia State	1.148640e+06
12	220.0	Ohio State	3.425510e+06
13	215.0	Kentucky	1.749840e+06
15	216.0	NaN	3.425510e+06
433	215.0	Duke	6.000000e+06
437	200.0	Lehigh	2.525160e+06
438	185.0	Westchester CC	5.250930e+05
444	214.0	Colorado	9.463484e+06
449	206.0	Duke	1.348440e+06

[102 rows x 9 columns]

SF

Name Team N	vuiibei	Position					
Jae Crowder Boston Celtics	99.0	SF					
25.032 Thanasis Antetokounmpo New York Knicks23.0	43.0	SF					
33 Carmelo Anthony New York Knicks	7.0	SF					
32.0 35 Cleanthony Early New York Knicks	11.0	SF					
25.0 42 Lance Thomas New York Knicks 28.0	42.0	SF					
428 Al-Farouq Aminu Portland Trail Blazers 25.0	8.0	SF					
432 Maurice Harkless Portland Trail Blazers 23.0	4.0	SF					
448 Gordon Hayward Utah Jazz	20.0	SF					
26.0 450 Joe Ingles Utah Jazz	2.0	SF					
28.0 451 Chris Johnson Utah Jazz	23.0	SF					
26.0							
Height Weight College Salary 1 6-6 235.0 Marquette 6796117.0 32 6-7 205.0 NaN 30888.0 33 6-8 240.0 Syracuse 22875000.0 35 6-8 210.0 Wichita State 845059.0 42 6-8 235.0 Duke 1636842.0							
428 6-9 215.0 Wake Forest 8042895.0 432 6-9 215.0 St. John's 2894059.0 448 6-8 226.0 Butler 15409570.0 450 6-8 226.0 NaN 2050000.0 451 6-6 206.0 Dayton 981348.0							
[85 rows x 9 columns]							
SF.isnull().sum()							
Name 0 Team 0 Number 0 Position 0 Age 0 Height 0 Weight 0 College 14							

```
Salary
dtype: int64
comb df = pd.concat([PG, SF,C,SG])
comb df
                                               Number Position Age
                 Name
                                         Team
Height
                               Boston Celtics
                                                   0.0
                                                             PG 25.0
        Avery Bradley
6-2
                               Boston Celtics
                                                             PG 22.0
         Terry Rozier
                                                  12.0
8
6-2
9
         Marcus Smart
                               Boston Celtics
                                                  36.0
                                                             PG 22.0
6-4
                               Boston Celtics
11
        Isaiah Thomas
                                                   4.0
                                                             PG 27.0
5-9
19
         Jarrett Jack
                                Brooklyn Nets
                                                   2.0
                                                             PG 32.0
6-3
. .
. . .
    Gerald Henderson Portland Trail Blazers
                                                   9.0
                                                             SG 28.0
433
6-5
        C.J. McCollum Portland Trail Blazers
437
                                                   3.0
                                                             SG 24.0
6-4
438
         Luis Montero Portland Trail Blazers
                                                  44.0
                                                             SG 23.0
6-7
444
           Alec Burks
                                    Utah Jazz
                                                  10.0
                                                             SG 24.0
6-6
449
          Rodney Hood
                                    Utah Jazz
                                                   5.0
                                                             SG 23.0
6-8
     Weight
                    College
                                Salary
      180.0
                      Texas
0
                             7730337.0
                 Louisville
8
      190.0
                             1824360.0
9
      220.0
             Oklahoma State
                             3431040.0
                 Washington
11
      185.0
                             6912869.0
19
      200.0
               Georgia Tech
                             6300000.0
433
      215.0
                       Duke
                             6000000.0
437
      200.0
                     Lehigh
                             2525160.0
438
      185.0
             Westchester CC
                              525093.0
444
      214.0
                   Colorado
                             9463484.0
449
      206.0
                       Duke
                             1348440.0
[357 rows x 9 columns]
comb_df.sort_index(axis = 0)
                              Team Number Position Age Height
              Name
Weight \
```

0 180.0	Avery Bradley	Boston Celt	ics 0.0) PG	25.0	6-2
1 235.0	Jae Crowder	Boston Celt	ics 99.0) SF	25.0	6-6
2 205.0	John Holland	Boston Celt	ics 30.0) SG	27.0	6-5
3	R.J. Hunter	Boston Celt	ics 28.0) SG	22.0	6-5
185.0 7 238.0	Kelly Olynyk	Boston Celt	cics 41.0) C	25.0	7-0
451 206.0	Chris Johnson	Utah J	lazz 23.0) SF	26.0	6-6
453 203.0	Shelvin Mack	Utah J	lazz 8.0	PG PG	26.0	6-3
454 179.0	Raul Neto	Utah J	lazz 25.0) PG	24.0	6-1
455 256.0	Tibor Pleiss	Utah J	Jazz 21.0) C	26.0	7-3
456 231.0	Jeff Withey	Utah J	Jazz 24.0) C	26.0	7 - 0
3 7 451 453 454 455 456	Marque Boston Univers Georgia St Gonz Day But	xas 7.73033 tte 6.79611 ity 5.07782 ate 1.14864 aga 2.16516 ton 9.81348 ler 2.43333 NaN 9.00000 NaN 2.90000 sas 9.47276	7e+06 9e+06 10e+06 50e+06 80e+05 33e+06 10e+05			
[35/	rows x 9 colum	ns J				