```
import numpy as np
import pandas as pd
df=pd.read csv("Hotel Dataset.csv")
df
    CustomerID Age Group Rating(1-5)
                                              Hotel FoodPreference
                                                                       Bill
0
                    20-25
                                       4
                                                Ibis
                                                                 veg
                                                                       1300
1
              2
                    30-35
                                          LemonTree
                                                             Non-Veg
                                                                       2000
                    25-30
2
              3
                                       6
                                              RedFox
                                                                 Veg
                                                                       1322
3
                    20-25
                                      - 1
                                          LemonTree
                                                                 Veg
                                                                       1234
                       35+
                                       3
                                                Ibis
                                                          Vegetarian
                                                                        989
                       35+
                                       3
                                                Ibys
5
                                                             Non-Veg
                                                                       1909
                                       4
6
                       35+
                                              RedFox
                                                          Vegetarian
                                                                       1000
7
                                          LemonTree
                    20-25
                                                                 Veg
                                                                       2999
8
                    25-30
                                       2
                                                Ibis
                                                             Non-Veg
                                                                       3456
9
              9
                    25-30
                                       2
                                                Ibis
                                                             Non-Veg
                                                                       3456
10
             10
                    30-35
                                              RedFox
                                                             non-Veg -6755
    No0fPax
              EstimatedSalary Age Group.1
          2
3
0
                         40000
                                      20-25
1
                         59000
                                      30-35
           2
2
                         30000
                                      25 - 30
3
           2
                                      20-25
                        120000
4
           2
                         45000
                                        35+
5
           2
                        122220
                                        35+
6
          - 1
                         21122
                                        35+
7
                                      20-25
         - 10
                        345673
8
           3
                        -99999
                                      25 - 30
9
           3
                        -99999
                                      25 - 30
10
           4
                         87777
                                      30-35
df.duplicated()
0
      False
1
      False
2
      False
3
      False
4
      False
5
      False
```

```
6
      False
7
      False
8
      False
9
       True
10
      False
dtype: bool
df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11 entries, 0 to 10
Data columns (total 9 columns):
#
     Column
                       Non-Null Count
                                        Dtype
 0
     CustomerID
                       11 non-null
                                        int64
1
     Age Group
                       11 non-null
                                        object
 2
     Rating(1-5)
                       11 non-null
                                        int64
 3
     Hotel
                       11 non-null
                                        object
 4
     FoodPreference
                       11 non-null
                                        object
 5
     Bill
                       11 non-null
                                        int64
 6
     No0fPax
                       11 non-null
                                        int64
 7
     EstimatedSalary
                      11 non-null
                                        int64
8
                       11 non-null
     Age Group.1
                                        object
dtypes: int64(5), object(4)
memory usage: 920.0+ bytes
df.drop duplicates(inplace=True)
df
    CustomerID Age_Group Rating(1-5)
                                             Hotel FoodPreference
                                                                     Bill
/
0
                    20-25
                                              Ibis
                                                               veg
                                                                     1300
                                         LemonTree
1
             2
                    30-35
                                      5
                                                           Non-Veg
                                                                     2000
2
                    25-30
                                      6
                                            RedFox
                                                               Veg
                                                                     1322
3
                    20-25
                                     - 1
                                         LemonTree
                                                                     1234
                                                               Veg
                      35+
                                      3
                                              Ibis
                                                        Vegetarian
                                                                      989
                                      3
5
                      35+
                                              Ibys
                                                           Non-Veg
                                                                     1909
6
                      35+
                                            RedFox
                                                        Vegetarian
                                                                     1000
                    20-25
                                         LemonTree
7
             8
                                      7
                                                               Veg
                                                                     2999
8
             9
                    25-30
                                      2
                                              Ibis
                                                           Non-Veg
                                                                    3456
                                      5
10
            10
                    30-35
                                            RedFox
                                                           non-Veg -6755
```

```
No0fPax
              EstimatedSalary Age Group.1
0
           2
                         40000
                                       20-25
1
           3
                         59000
                                       30-35
           2
2
                         30000
                                       25 - 30
3
           2
                        120000
                                       20-25
4
           2
                         45000
                                         35+
5
           2
                        122220
                                         35 +
6
          - 1
                                         35+
                         21122
7
         - 10
                        345673
                                       20-25
8
           3
                        -99999
                                       25 - 30
           4
10
                         87777
                                       30-35
len(df)
10
index=np.array(list(range(0,len(df))))
df.set index(index,inplace=True)
index
df
                                              Hotel FoodPreference
   CustomerID Age Group Rating(1-5)
                                                                       Bill
NoOfPax \
             1
                    20-25
                                                Ibis
                                                                 veg
                                                                       1300
2
1
             2
                    30-35
                                          LemonTree
                                                             Non-Veg
                                                                       2000
3
2
                    25-30
                                             RedFox
                                                                 Veg
                                                                       1322
2
3
                    20-25
                                          LemonTree
                                                                       1234
                                                                 Veg
2
4
             5
                      35+
                                       3
                                                          Vegetarian
                                                                        989
                                                Ibis
2
5
             6
                      35+
                                                Ibys
                                                             Non-Veg
                                                                       1909
2
6
                      35+
                                             RedFox
                                                          Vegetarian
                                                                       1000
- 1
7
                    20-25
                                          LemonTree
                                                                       2999
             8
                                       7
                                                                 Veg
- 10
                                       2
8
                    25-30
                                                Ibis
                                                             Non-Veg
                                                                      3456
3
9
                                       5
                                             RedFox
                                                             non-Veg -6755
            10
                    30-35
4
   EstimatedSalary Age_Group.1
0
              40000
                            20-25
1
              59000
                            30 - 35
2
              30000
                            25 - 30
```

```
3
             120000
                           20-25
4
              45000
                             35+
5
             122220
                             35+
6
              21122
                             35+
7
             345673
                           20-25
8
                           25-30
             -99999
9
                           30-35
              87777
df.drop(['Age_Group.1'],axis=1,inplace=True)
df
   CustomerID Age Group Rating(1-5) Hotel FoodPreference
                                                                     Bill
NoOfPax \
0
             1
                   20-25
                                               Ibis
                                                                      1300
                                                                veg
2
1
                   30-35
                                         LemonTree
                                                            Non-Veg
                                                                      2000
3
2
                                      6
                                            RedFox
                                                                      1322
                   25-30
                                                                Veg
2
3
                                                                      1234
                   20-25
                                     - 1
                                         LemonTree
                                                                Veg
2
4
                      35+
                                      3
                                                         Vegetarian
                                               Ibis
                                                                     989
2
5
                     35+
                                      3
                                                                      1909
                                               Ibys
                                                            Non-Veg
2
6
                     35+
                                            RedFox
                                                         Vegetarian
                                                                      1000
- 1
7
             8
                   20-25
                                         LemonTree
                                                                Veg
                                                                     2999
- 10
                                      2
8
                   25 - 30
                                               Ibis
                                                            Non-Veg 3456
3
9
            10
                   30-35
                                      5
                                            RedFox
                                                            non-Veg -6755
4
   EstimatedSalary
0
              40000
              59000
1
2
              30000
3
             120000
4
              45000
5
             122220
6
              21122
7
             345673
8
             -99999
9
              87777
df.CustomerID.loc[df.CustomerID<0]=np.nan</pre>
df.Bill.loc[df.Bill<0]=np.nan</pre>
df.EstimatedSalary.loc[df.EstimatedSalary<0]=np.nan</pre>
df
```

c:\users\asus\appdata\local\programs\python\python37\lib\sitepackages\pandas\core\indexing.py:1732: 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

self. setitem single block(indexer, value, name)

	CustomerID	Age_Group	Rating(1-5)	Hotel	FoodPreference	Bill
0	1.0	20-25	4	Ibis	veg	1300.0
1	2.0	30-35	5	LemonTree	Non-Veg	2000.0
2	3.0	25-30	6	RedFox	Veg	1322.0
3	4.0	20-25	-1	LemonTree	Veg	1234.0
4	5.0	35+	3	Ibis	Vegetarian	989.0
5	6.0	35+	3	Ibys	Non-Veg	1909.0
6	7.0	35+	4	RedFox	Vegetarian	1000.0
7	8.0	20-25	7	LemonTree	Veg	2999.0
8	9.0	25-30	2	Ibis	Non-Veg	3456.0
9	10.0	30-35	5	RedFox	non-Veg	NaN

	NoOfPax	EstimatedSalary
0	2	40000.0
1	3	59000.0
2	2	30000.0
3	2	120000.0
4	2	45000.0
5	2	122220.0
6	-1	21122.0
7	- 10	345673.0
8	3	NaN
9	4	87777.0

c:\users\asus\appdata\local\programs\python\python37\lib\sitepackages\pandas\core\indexing.py:1732: 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
  self._setitem_single_block(indexer, value, name)
                                                                     Bill
   CustomerID Age Group Rating(1-5)
                                            Hotel FoodPreference
\
0
          1.0
                  20-25
                                    4
                                             Ibis
                                                                   1300.0
                                                              veg
          2.0
1
                   30-35
                                        LemonTree
                                                         Non-Veg
                                                                   2000.0
          3.0
2
                  25-30
                                           RedFox
                                                              Veg
                                                                   1322.0
          4.0
                                        LemonTree
3
                  20-25
                                                              Veg
                                                                  1234.0
          5.0
                     35+
                                                      Vegetarian
                                                                    989.0
                                    3
                                             Ibis
5
          6.0
                     35+
                                    3
                                             Ibys
                                                         Non-Veg 1909.0
6
          7.0
                     35+
                                           RedFox
                                                      Vegetarian
                                                                   1000.0
7
          8.0
                   20-25
                                        LemonTree
                                                              Veg
                                                                   2999.0
          9.0
8
                   25 - 30
                                    2
                                             Ibis
                                                         Non-Veg 3456.0
9
         10.0
                  30-35
                                    5
                                           RedFox
                                                         non-Veg
                                                                      NaN
   No0fPax
            EstimatedSalary
0
       2.0
                     40000.0
1
       3.0
                     59000.0
2
       2.0
                     30000.0
3
       2.0
                    120000.0
4
       2.0
                     45000.0
5
       2.0
                    122220.0
6
       NaN
                     21122.0
7
       NaN
                    345673.0
8
       3.0
                         NaN
9
       4.0
                     87777.0
df.Age Group.unique()
array(['20-25', '30-35', '25-30', '35+'], dtype=object)
df.Hotel.unique()
array(['Ibis', 'LemonTree', 'RedFox', 'Ibys'], dtype=object)
df.Hotel.replace(['Ibys'],'Ibis',inplace=True)
df.FoodPreference.unique
```

```
<bound method Series.unique of 0</pre>
                                              veq
        Non-Veg
1
2
            Veg
3
            Vea
4
     Vegetarian
5
        Non-Veg
6
     Vegetarian
7
            Veg
8
        Non-Veg
9
        non-Veg
Name: FoodPreference, dtype: object>
df.FoodPreference.replace(['Vegetarian','veg'],'Veg',inplace=True)
df.FoodPreference.replace(['non-Veg'],'Non-Veg',inplace=True)
df.EstimatedSalary.fillna(round(df.EstimatedSalary.mean()),inplace=Tru
df.NoOfPax.fillna(round(df.NoOfPax.median()),inplace=True)
df['Rating(1-5)'].fillna(round(df['Rating(1-5)'].median()),
inplace=True)
df.Bill.fillna(round(df.Bill.mean()),inplace=True)
df
                                            Hotel FoodPreference
   CustomerID Age Group Rating(1-5)
                                                                      Bill
/
0
          1.0
                   20-25
                                              Ibis
                                                               Veg
                                                                    1300.0
          2.0
                                     5
                                         LemonTree
1
                   30 - 35
                                                           Non-Veg
                                                                    2000.0
2
                                     6
          3.0
                   25-30
                                            RedFox
                                                               Veg
                                                                    1322.0
3
          4.0
                   20-25
                                    - 1
                                         LemonTree
                                                               Veg
                                                                    1234.0
          5.0
                     35+
                                     3
                                                                     989.0
                                              Ibis
                                                               Veg
          6.0
                                     3
                                              Ibis
                     35+
                                                           Non-Veg
                                                                    1909.0
          7.0
                                            RedFox
6
                     35 +
                                                               Veg
                                                                    1000.0
7
          8.0
                   20-25
                                     7
                                         LemonTree
                                                               Veg
                                                                    2999.0
          9.0
                                     2
8
                   25 - 30
                                              Ibis
                                                                    3456.0
                                                           Non-Veg
9
         10.0
                   30-35
                                     5
                                            RedFox
                                                           Non-Veg
                                                                    1801.0
   NoOfPax
            EstimatedSalary
0
       2.0
                     40000.0
       3.0
                     59000.0
1
2
       2.0
                     30000.0
3
                    120000.0
       2.0
```

4	2.0	45000.0
5	2.0	122220.0
6	2.0	21122.0
7	2.0	345673.0
8	3.0	96755.0
9	4.0	87777.0