Handling Null values

August 17, 2024

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

```
C:\Users\Vikas\anaconda3\lib\site-packages\pandas\core\arrays\masked.py:60:
UserWarning: Pandas requires version '1.3.6' or newer of 'bottleneck' (version
'1.3.5' currently installed).
  from pandas.core import (
```

```python
[2]: ## import the dataset
     data = pd.read_csv("Bengaluru_House_Data.csv")
```

```python
[3]: data.head()
```

```
[3]:             area_type    availability                       location        size  \
     0  Super built-up  Area          19-Dec  Electronic City Phase II       2 BHK
     1            Plot  Area  Ready To Move          Chikka Tirupathi  4 Bedroom
     2        Built-up  Area  Ready To Move               Uttarahalli      3 BHK
     3  Super built-up  Area  Ready To Move         Lingadheeranahalli      3 BHK
     4  Super built-up  Area  Ready To Move                   Kothanur      2 BHK

          society total_sqft  bath  balcony   price
     0   Coomee        1056   2.0      1.0   39.07
     1  Theanmp        2600   5.0      3.0  120.00
     2      NaN        1440   2.0      3.0   62.00
     3  Soiewre        1521   3.0      1.0   95.00
     4      NaN        1200   2.0      1.0   51.00
```

```python
[5]: data.shape
```

```
[5]: (13320, 9)
```

```python
[6]: data.isnull().sum()
```

```
[6]: area_type          0
     availability       0
     location           1
     size              16
     society         5502
     total_sqft         0
     bath              73
```

```
balcony         609
price             0
dtype: int64
```

[7]: `data.isnull().sum().sum()`

[7]: 6201

0.1 Filling null values

[8]: `df = data.fillna(value = 0)`

[9]: `df`

[9]:
```
              area_type     availability                  location  \
0       Super built-up  Area          19-Dec  Electronic City Phase II
1                Plot  Area  Ready To Move          Chikka Tirupathi
2             Built-up  Area  Ready To Move                Uttarahalli
3       Super built-up  Area  Ready To Move         Lingadheeranahalli
4       Super built-up  Area  Ready To Move                   Kothanur
...                ...            ...                       ...
13315         Built-up  Area  Ready To Move                 Whitefield
13316   Super built-up  Area  Ready To Move              Richards Town
13317         Built-up  Area  Ready To Move      Raja Rajeshwari Nagar
13318   Super built-up  Area          18-Jun            Padmanabhanagar
13319   Super built-up  Area  Ready To Move                Doddathoguru

           size  society total_sqft  bath  balcony   price
0         2 BHK   Coomee       1056   2.0      1.0   39.07
1     4 Bedroom  Theanmp       2600   5.0      3.0  120.00
2         3 BHK        0       1440   2.0      3.0   62.00
3         3 BHK  Soiewre       1521   3.0      1.0   95.00
4         2 BHK        0       1200   2.0      1.0   51.00
...         ...      ...        ...   ...      ...     ...
13315 5 Bedroom  ArsiaEx       3453   4.0      0.0  231.00
13316     4 BHK        0       3600   5.0      0.0  400.00
13317     2 BHK  Mahla T       1141   2.0      1.0   60.00
13318     4 BHK  SollyCl       4689   4.0      1.0  488.00
13319     1 BHK        0        550   1.0      1.0   17.00

[13320 rows x 9 columns]
```

[10]: `df.isnull().sum().sum()`

[10]: 0

```
[11]: ## Now want to fill the null value previous value
      df1 =data.fillna(method = 'pad')
      df1
```

C:\Users\Vikas\AppData\Local\Temp\ipykernel_17760\955742772.py:2: FutureWarning:
DataFrame.fillna with 'method' is deprecated and will raise in a future version.
Use obj.ffill() or obj.bfill() instead.
  df1 =data.fillna(method = 'pad')

```
[11]:                 area_type   availability                     location  \
      0       Super built-up  Area         19-Dec  Electronic City Phase II
      1                 Plot  Area  Ready To Move           Chikka Tirupathi
      2             Built-up  Area  Ready To Move                Uttarahalli
      3       Super built-up  Area  Ready To Move        Lingadheeranahalli
      4       Super built-up  Area  Ready To Move                   Kothanur
      ...                ...            ...                          ...
      13315         Built-up  Area  Ready To Move                Whitefield
      13316  Super built-up  Area  Ready To Move              Richards Town
      13317         Built-up  Area  Ready To Move      Raja Rajeshwari Nagar
      13318  Super built-up  Area         18-Jun            Padmanabhanagar
      13319  Super built-up  Area  Ready To Move                Doddathoguru

                  size   society total_sqft  bath  balcony   price
      0          2 BHK   Coomee        1056   2.0      1.0   39.07
      1      4 Bedroom  Theanmp        2600   5.0      3.0  120.00
      2          3 BHK  Theanmp        1440   2.0      3.0   62.00
      3          3 BHK  Soiewre        1521   3.0      1.0   95.00
      4          2 BHK  Soiewre        1200   2.0      1.0   51.00
      ...          ...      ...         ...   ...      ...     ...
      13315  5 Bedroom  ArsiaEx        3453   4.0      0.0  231.00
      13316      4 BHK  ArsiaEx        3600   5.0      0.0  400.00
      13317      2 BHK  Mahla T        1141   2.0      1.0   60.00
      13318      4 BHK  SollyCl        4689   4.0      1.0  488.00
      13319      1 BHK  SollyCl         550   1.0      1.0   17.00

      [13320 rows x 9 columns]
```

```
[13]: df1.isnull().sum()
```

```
[13]: area_type     0
      availability  0
      location      0
      size          0
      society       0
      total_sqft    0
      bath          0
      balcony       0
      price         0
```

```
dtype: int64
```

[14]: ```
Now want to fill the null value upcomming value
data.head()
```

[14]:
```
             area_type   availability   location   size \
0 Super built-up Area   19-Dec Electronic City Phase II   2 BHK
1   Plot Area Ready To Move   Chikka Tirupathi 4 Bedroom
2   Built-up Area Ready To Move   Uttarahalli   3 BHK
3 Super built-up Area Ready To Move   Lingadheeranahalli   3 BHK
4 Super built-up Area Ready To Move   Kothanur   2 BHK

     society total_sqft bath balcony   price
0   Coomee   1056   2.0   1.0   39.07
1 Theanmp   2600   5.0   3.0 120.00
2   NaN   1440   2.0   3.0   62.00
3 Soiewre   1521   3.0   1.0   95.00
4   NaN   1200   2.0   1.0   51.00
```

[15]: ```
df2=data.fillna(method = 'bfill')
df2
```

```
C:\Users\Vikas\AppData\Local\Temp\ipykernel_17760\3819208495.py:1:
FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a
future version. Use obj.ffill() or obj.bfill() instead.
  df2=data.fillna(method = 'bfill')
```

[15]:
```
             area_type     availability                       location  \
0      Super built-up  Area              19-Dec  Electronic City Phase II
1                Plot  Area  Ready To Move            Chikka Tirupathi
2            Built-up  Area  Ready To Move                 Uttarahalli
3      Super built-up  Area  Ready To Move       Lingadheeranahalli
4      Super built-up  Area  Ready To Move                    Kothanur
...               ...          ...                    ...
13315        Built-up  Area  Ready To Move                  Whitefield
13316  Super built-up  Area  Ready To Move               Richards Town
13317        Built-up  Area  Ready To Move       Raja Rajeshwari Nagar
13318  Super built-up  Area              18-Jun          Padmanabhanagar
13319  Super built-up  Area  Ready To Move                 Doddathoguru

           size   society total_sqft  bath  balcony   price
0         2 BHK    Coomee       1056   2.0      1.0   39.07
1     4 Bedroom   Theanmp       2600   5.0      3.0  120.00
2         3 BHK   Soiewre       1440   2.0      3.0   62.00
3         3 BHK   Soiewre       1521   3.0      1.0   95.00
4         2 BHK   DuenaTa       1200   2.0      1.0   51.00
...         ...       ...          ...   ...      ...     ...
13315  5 Bedroom  ArsiaEx       3453   4.0      0.0  231.00
```

```
13316      4 BHK   Mahla T       3600    5.0      1.0  400.00
13317      2 BHK   Mahla T       1141    2.0      1.0   60.00
13318      4 BHK   SollyCl       4689    4.0      1.0  488.00
13319      1 BHK       NaN        550    1.0      1.0   17.00

[13320 rows x 9 columns]
```

[16]: ```
fill the null values with column wise previous value
df3 = data.fillna(method = 'pad',axis=1)
df3
```

```
C:\Users\Vikas\AppData\Local\Temp\ipykernel_17760\3996429455.py:2:
FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a
future version. Use obj.ffill() or obj.bfill() instead.
  df3 = data.fillna(method = 'pad',axis=1)
```

[16]:
```
                   area_type    availability                    location  \
0         Super built-up  Area          19-Dec  Electronic City Phase II
1                  Plot  Area  Ready To Move          Chikka Tirupathi
2              Built-up  Area  Ready To Move               Uttarahalli
3         Super built-up  Area  Ready To Move       Lingadheeranahalli
4         Super built-up  Area  Ready To Move                  Kothanur
...                   ...            ...                      ...
13315         Built-up  Area  Ready To Move               Whitefield
13316  Super built-up  Area  Ready To Move            Richards Town
13317         Built-up  Area  Ready To Move  Raja Rajeshwari Nagar
13318  Super built-up  Area          18-Jun          Padmanabhanagar
13319  Super built-up  Area  Ready To Move              Doddathoguru

             size  society total_sqft bath balcony   price
0          2 BHK   Coomee      1056  2.0     1.0   39.07
1      4 Bedroom  Theanmp      2600  5.0     3.0   120.0
2          3 BHK    3 BHK      1440  2.0     3.0    62.0
3          3 BHK  Soiewre      1521  3.0     1.0    95.0
4          2 BHK    2 BHK      1200  2.0     1.0    51.0
...           ...      ...       ...  ...     ...      ...
13315  5 Bedroom  ArsiaEx      3453  4.0     0.0   231.0
13316      4 BHK    4 BHK      3600  5.0     5.0   400.0
13317      2 BHK   Mahla T      1141  2.0     1.0    60.0
13318      4 BHK  SollyCl      4689  4.0     1.0   488.0
13319      1 BHK    1 BHK       550  1.0     1.0    17.0

[13320 rows x 9 columns]
```

[17]: ```
fill the null values with column wise next value
df4 = data.fillna(method = 'bfill',axis=1)
df4
```

```
C:\Users\Vikas\AppData\Local\Temp\ipykernel_17760\932812257.py:2: FutureWarning:
DataFrame.fillna with 'method' is deprecated and will raise in a future version.
Use obj.ffill() or obj.bfill() instead.
  df4 = data.fillna(method = 'bfill',axis=1)
```

[17]:
```
              area_type   availability                 location  \
0       Super built-up  Area         19-Dec  Electronic City Phase II
1                 Plot  Area  Ready To Move          Chikka Tirupathi
2              Built-up  Area  Ready To Move               Uttarahalli
3       Super built-up  Area  Ready To Move        Lingadheeranahalli
4       Super built-up  Area  Ready To Move                  Kothanur
...                 ...            ...                          ...
13315          Built-up  Area  Ready To Move               Whitefield
13316  Super built-up  Area  Ready To Move             Richards Town
13317          Built-up  Area  Ready To Move      Raja Rajeshwari Nagar
13318  Super built-up  Area         18-Jun          Padmanabhanagar
13319  Super built-up  Area  Ready To Move             Doddathoguru

             size   society total_sqft bath balcony   price
0           2 BHK   Coomee       1056  2.0     1.0   39.07
1       4 Bedroom  Theanmp       2600  5.0     3.0   120.0
2           3 BHK     1440       1440  2.0     3.0    62.0
3           3 BHK  Soiewre       1521  3.0     1.0    95.0
4           2 BHK     1200       1200  2.0     1.0    51.0
...           ...      ...        ...  ...     ...     ...
13315   5 Bedroom  ArsiaEx       3453  4.0     0.0   231.0
13316       4 BHK     3600       3600  5.0   400.0   400.0
13317       2 BHK  Mahla T       1141  2.0     1.0    60.0
13318       4 BHK  SollyCl       4689  4.0     1.0   488.0
13319       1 BHK      550        550  1.0     1.0    17.0

[13320 rows x 9 columns]
```

[18]:
```
## filling the different value in null different coloumns
df6 = data.fillna({'society':'abcd','balcony':'defg'})
df6
```

[18]:
```
              area_type   availability                 location  \
0       Super built-up  Area         19-Dec  Electronic City Phase II
1                 Plot  Area  Ready To Move          Chikka Tirupathi
2              Built-up  Area  Ready To Move               Uttarahalli
3       Super built-up  Area  Ready To Move        Lingadheeranahalli
4       Super built-up  Area  Ready To Move                  Kothanur
...                 ...            ...                          ...
13315          Built-up  Area  Ready To Move               Whitefield
13316  Super built-up  Area  Ready To Move             Richards Town
13317          Built-up  Area  Ready To Move      Raja Rajeshwari Nagar
13318  Super built-up  Area         18-Jun          Padmanabhanagar
```

```
13319  Super built-up  Area  Ready To Move                Doddathoguru
```

```
              size   society  total_sqft  bath  balcony    price
0            2 BHK    Coomee        1056   2.0      1.0    39.07
1        4 Bedroom   Theanmp        2600   5.0      3.0   120.00
2            3 BHK      abcd        1440   2.0      3.0    62.00
3            3 BHK   Soiewre        1521   3.0      1.0    95.00
4            2 BHK      abcd        1200   2.0      1.0    51.00
...            ...       ...         ...   ...      ...      ...
13315    5 Bedroom   ArsiaEx        3453   4.0      0.0   231.00
13316        4 BHK      abcd        3600   5.0     defg   400.00
13317        2 BHK   Mahla T        1141   2.0      1.0    60.00
13318        4 BHK   SollyCl        4689   4.0      1.0   488.00
13319        1 BHK      abcd         550   1.0      1.0    17.00

[13320 rows x 9 columns]
```

[20]:
```python
## filling null values mean of a coloumns    ## min ,max
df7 = data.fillna(value=data['balcony'].max())
df7
```

[20]:
```
              area_type     availability                   location  \
0        Super built-up  Area       19-Dec   Electronic City Phase II
1                 Plot  Area  Ready To Move          Chikka Tirupathi
2              Built-up  Area  Ready To Move                Uttarahalli
3        Super built-up  Area  Ready To Move        Lingadheeranahalli
4        Super built-up  Area  Ready To Move                   Kothanur
...                 ...       ...            ...                       ...
13315          Built-up  Area  Ready To Move                 Whitefield
13316  Super built-up  Area  Ready To Move             Richards Town
13317          Built-up  Area  Ready To Move    Raja Rajeshwari Nagar
13318  Super built-up  Area       18-Jun            Padmanabhanagar
13319  Super built-up  Area  Ready To Move               Doddathoguru

              size   society  total_sqft  bath  balcony    price
0            2 BHK    Coomee        1056   2.0      1.0    39.07
1        4 Bedroom   Theanmp        2600   5.0      3.0   120.00
2            3 BHK       3.0        1440   2.0      3.0    62.00
3            3 BHK   Soiewre        1521   3.0      1.0    95.00
4            2 BHK       3.0        1200   2.0      1.0    51.00
...            ...       ...         ...   ...      ...      ...
13315    5 Bedroom   ArsiaEx        3453   4.0      0.0   231.00
13316        4 BHK       3.0        3600   5.0      3.0   400.00
13317        2 BHK   Mahla T        1141   2.0      1.0    60.00
13318        4 BHK   SollyCl        4689   4.0      1.0   488.00
13319        1 BHK       3.0         550   1.0      1.0    17.00
```

```
[13320 rows x 9 columns]
```

1 dropna()

```
[21]: df8 = data.dropna()
      df8
```

```
[21]:                area_type     availability                  location  \
      0      Super built-up  Area           19-Dec  Electronic City Phase II
      1                Plot  Area   Ready To Move            Chikka Tirupathi
      3      Super built-up  Area   Ready To Move           Lingadheeranahalli
      5      Super built-up  Area   Ready To Move                  Whitefield
      11               Plot  Area   Ready To Move                  Whitefield
      …                 …             …                         …
      13313  Super built-up  Area   Ready To Move                  Uttarahalli
      13314  Super built-up  Area   Ready To Move            Green Glen Layout
      13315          Built-up  Area   Ready To Move                Whitefield
      13317          Built-up  Area   Ready To Move       Raja Rajeshwari Nagar
      13318  Super built-up  Area           18-Jun             Padmanabhanagar

                   size  society total_sqft  bath  balcony   price
      0           2 BHK  Coomee         1056   2.0      1.0   39.07
      1       4 Bedroom  Theanmp        2600   5.0      3.0  120.00
      3           3 BHK  Soiewre        1521   3.0      1.0   95.00
      5           2 BHK  DuenaTa        1170   2.0      1.0   38.00
      11      4 Bedroom  Prrry M        2785   5.0      3.0  295.00
      …             …        …          …      …        …        …
      13313       3 BHK  Aklia R        1345   2.0      1.0   57.00
      13314       3 BHK  SoosePr        1715   3.0      3.0  112.00
      13315   5 Bedroom  ArsiaEx        3453   4.0      0.0  231.00
      13317       2 BHK  Mahla T        1141   2.0      1.0   60.00
      13318       4 BHK  SollyCl        4689   4.0      1.0  488.00

      [7496 rows x 9 columns]
```

```
[22]: df9 = data.dropna(how='any')
      df9
```

```
[22]:                area_type     availability                  location  \
      0      Super built-up  Area           19-Dec  Electronic City Phase II
      1                Plot  Area   Ready To Move            Chikka Tirupathi
      3      Super built-up  Area   Ready To Move           Lingadheeranahalli
      5      Super built-up  Area   Ready To Move                  Whitefield
      11               Plot  Area   Ready To Move                  Whitefield
      …                 …             …                         …
      13313  Super built-up  Area   Ready To Move                  Uttarahalli
```

```
13314  Super built-up  Area  Ready To Move         Green Glen Layout
13315       Built-up  Area  Ready To Move                Whitefield
13317       Built-up  Area  Ready To Move    Raja Rajeshwari Nagar
13318  Super built-up  Area         18-Jun       Padmanabhanagar


              size   society  total_sqft  bath  balcony   price
0           2 BHK    Coomee        1056   2.0      1.0   39.07
1       4 Bedroom   Theanmp        2600   5.0      3.0  120.00
3           3 BHK   Soiewre        1521   3.0      1.0   95.00
5           2 BHK   DuenaTa        1170   2.0      1.0   38.00
11      4 Bedroom    Prrry M       2785   5.0      3.0  295.00
...           ...       ...         ...   ...      ...     ...
13313       3 BHK   Aklia R        1345   2.0      1.0   57.00
13314       3 BHK   SoosePr        1715   3.0      3.0  112.00
13315   5 Bedroom   ArsiaEx        3453   4.0      0.0  231.00
13317       2 BHK   Mahla T        1141   2.0      1.0   60.00
13318       4 BHK   SollyCl        4689   4.0      1.0  488.00

[7496 rows x 9 columns]
```

2 repalce()

```python
import numpy as np
df10 = data.replace(to_replace=np.nan, value=87698)
df10
```

```
[23]:               area_type   availability                 location  \
0      Super built-up  Area         19-Dec  Electronic City Phase II
1                Plot  Area  Ready To Move           Chikka Tirupathi
2            Built-up  Area  Ready To Move               Uttarahalli
3      Super built-up  Area  Ready To Move       Lingadheeranahalli
4      Super built-up  Area  Ready To Move                  Kothanur
...             ...           ...                        ...
13315       Built-up  Area  Ready To Move                Whitefield
13316  Super built-up  Area  Ready To Move             Richards Town
13317       Built-up  Area  Ready To Move    Raja Rajeshwari Nagar
13318  Super built-up  Area         18-Jun       Padmanabhanagar
13319  Super built-up  Area  Ready To Move              Doddathoguru


              size   society  total_sqft  bath  balcony   price
0           2 BHK    Coomee        1056   2.0      1.0   39.07
1       4 Bedroom   Theanmp        2600   5.0      3.0  120.00
2           3 BHK     87698        1440   2.0      3.0   62.00
3           3 BHK   Soiewre        1521   3.0      1.0   95.00
4           2 BHK     87698        1200   2.0      1.0   51.00
...           ...       ...         ...   ...      ...     ...
```

```
13315  5 Bedroom  ArsiaEx       3453   4.0      0.0  231.00
13316      4 BHK    87698       3600   5.0  87698.0  400.00
13317      2 BHK  Mahla T       1141   2.0      1.0   60.00
13318      4 BHK  SollyCl       4689   4.0      1.0  488.00
13319      1 BHK    87698        550   1.0      1.0   17.00

[13320 rows x 9 columns]
```

[24]:
```python
df11 = data.replace(to_replace=3.0, value=5.0)
df11
```

[24]:
```
                 area_type   availability                   location  \
0        Super built-up  Area         19-Dec  Electronic City Phase II
1                 Plot  Area  Ready To Move           Chikka Tirupathi
2             Built-up  Area  Ready To Move                 Uttarahalli
3        Super built-up  Area  Ready To Move        Lingadheeranahalli
4        Super built-up  Area  Ready To Move                   Kothanur
...                   ...            ...                        ...
13315         Built-up  Area  Ready To Move                 Whitefield
13316  Super built-up  Area  Ready To Move              Richards Town
13317         Built-up  Area  Ready To Move      Raja Rajeshwari Nagar
13318  Super built-up  Area         18-Jun            Padmanabhanagar
13319  Super built-up  Area  Ready To Move                Doddathoguru

            size   society total_sqft  bath  balcony   price
0          2 BHK   Coomee        1056   2.0      1.0   39.07
1      4 Bedroom  Theanmp        2600   5.0      5.0  120.00
2          3 BHK      NaN        1440   2.0      5.0   62.00
3          3 BHK  Soiewre        1521   5.0      1.0   95.00
4          2 BHK      NaN        1200   2.0      1.0   51.00
...          ...      ...         ...   ...      ...     ...
13315  5 Bedroom  ArsiaEx       3453   4.0      0.0  231.00
13316      4 BHK      NaN        3600   5.0      NaN  400.00
13317      2 BHK  Mahla T       1141   2.0      1.0   60.00
13318      4 BHK  SollyCl       4689   4.0      1.0  488.00
13319      1 BHK      NaN        550   1.0      1.0   17.00

[13320 rows x 9 columns]
```

3   interpolate()

1 ## value is missing 5 automatimically fill the value 2 3 4 nan 6 7 8

[25]:
```python
data['balcony']=data['balcony'].interpolate(method='linear')
data
```

```
[25]:              area_type    availability                   location  \
      0      Super built-up  Area           19-Dec  Electronic City Phase II
      1                Plot  Area   Ready To Move           Chikka Tirupathi
      2             Built-up  Area   Ready To Move                Uttarahalli
      3      Super built-up  Area   Ready To Move        Lingadheeranahalli
      4      Super built-up  Area   Ready To Move                   Kothanur
      ...                ...             ...                         ...
      13315         Built-up  Area   Ready To Move                 Whitefield
      13316  Super built-up  Area   Ready To Move              Richards Town
      13317         Built-up  Area   Ready To Move      Raja Rajeshwari Nagar
      13318  Super built-up  Area           18-Jun            Padmanabhanagar
      13319  Super built-up  Area   Ready To Move                Doddathoguru

                  size  society total_sqft  bath  balcony   price
      0          2 BHK   Coomee       1056   2.0      1.0   39.07
      1      4 Bedroom  Theanmp       2600   5.0      3.0  120.00
      2          3 BHK      NaN       1440   2.0      3.0   62.00
      3          3 BHK  Soiewre       1521   3.0      1.0   95.00
      4          2 BHK      NaN       1200   2.0      1.0   51.00
      ...          ...      ...        ...   ...      ...     ...
      13315  5 Bedroom  ArsiaEx       3453   4.0      0.0  231.00
      13316      4 BHK      NaN       3600   5.0      0.5  400.00
      13317      2 BHK  Mahla T       1141   2.0      1.0   60.00
      13318      4 BHK  SollyCl       4689   4.0      1.0  488.00
      13319      1 BHK      NaN        550   1.0      1.0   17.00

      [13320 rows x 9 columns]

[ ]:
```