In [18]:

```
import pandas as pd
import numpy as np
```

In [19]:

```
data = pd.read_excel("D:\KSR\Python\ExcelFile\Cust_Data.xlsx")
```

In [20]:

```
data
```

Out[20]:

	CustID	CustName	CustLoc	CustAge
0	1	Kiran	Pune	29
1	1	Kiran	Pune	29
2	2	Sathish	Maldives	30
3	3	Eshwar	Mumbai	34
4	5	Santhosh	Maldives	30
5	6	Mahesh	Tirupati	32
6	NaN	NaN	NaN	NaN
7	7	NaN	Bangalore	NaN

In [21]:

```
data.shape
```

Out[21]:

```
(8, 4)
```

In [22]:

```
list(data.columns)
```

Out[22]:

```
['CustID', 'CustName', 'CustLoc', 'CustAge']
```

In [23]:

```
print("The no of rows in given data:", data.shape[0])
```

```
The no of rows in given data: 8
```

In [24]:

```
print("The no of columns in given data:", data.shape[1])
```

The no of columns in given data: 4

In [25]:

```
data.isna().sum()
```

Out[25]:

```
CustID      1
CustName    2
CustLoc     1
CustAge     2
dtype: int64
```

In [26]:

```
(data.isna().sum() / data.shape[0]) * 100
```

Out[26]:

```
CustID      12.5
CustName    25.0
CustLoc     12.5
CustAge     25.0
dtype: float64
```

In [27]:

```
data.head(10)
```

Out[27]:

	CustID	CustName	CustLoc	CustAge
0	1	Kiran	Pune	29
1	1	Kiran	Pune	29
2	2	Sathish	Maldives	30
3	3	Eshwar	Mumbai	34
4	5	Santhosh	Maldives	30
5	6	Mahesh	Tirupati	32
6	NaN	NaN	NaN	NaN
7	7	NaN	Bangalore	NaN

In [29]:

```
data = data.dropna(subset = ['CustID'])
```

In [30]:

```
data.head(10)
```

Out[30]:

	CustID	CustName	CustLoc	CustAge
0	1	Kiran	Pune	29
1	1	Kiran	Pune	29
2	2	Sathish	Maldives	30
3	3	Eshwar	Mumbai	34
4	5	Santhosh	Maldives	30
5	6	Mahesh	Tirupati	32
7	7	NaN	Bangalore	NaN

In [36]:

```
data['CustAge'] = round(data['CustAge'].fillna(data['CustAge'].mea
```

```
C:\Users\ragav\AppData\Local\Temp\ipykernel_22504\393619598.py:1: 
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pa
  data['CustAge'] = round(data['CustAge'].fillna(data['CustAge'].m
```

In [37]:

```
data.head(10)
```

Out[37]:

	CustID	CustName	CustLoc	CustAge
0	1	Kiran	Pune	29
1	1	Kiran	Pune	29
2	2	Sathish	Maldives	30
3	3	Eshwar	Mumbai	34
4	5	Santhosh	Maldives	30
5	6	Mahesh	Tirupati	32
7	7	NaN	Bangalore	31

In [38]:

```
data['CustName'] = data['CustName'].fillna("Guest/NA")
```

```
C:\Users\ragav\AppData\Local\Temp\ipykernel_22504\2610292362.py:1:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pa
  data['CustName'] = data['CustName'].fillna("Guest/NA")
```

In [39]:

```
data.head(10)
```

Out[39]:

	CustID	CustName	CustLoc	CustAge

0	1	Kiran	Pune	29
1	1	Kiran	Pune	29
2	2	Sathish	Maldives	30
3	3	Eshwar	Mumbai	34
4	5	Santhosh	Maldives	30
5	6	Mahesh	Tirupati	32
7	7	Guest/NA	Bangalore	31

In [43]:

```
(data.isna().sum() / data.shape[0])*100
```

Out[43]:

```
CustID      0.0
CustName    0.0
CustLoc     0.0
CustAge     0.0
dtype: float64
```

In [ ]: