

ASSESSMENT -1

21BCE0516

ANUSHKA

anushka.2021a@vitstudent.ac.in (<mailto:anushka.2021a@vitstudent.ac.in>)

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

```
In [2]: data = {
    'Feature1': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],
    'Feature2': ['A', 'B', None, 'D', 'E', 'F', 'G', 'H', 'I', 'J'],
    'Feature3': [10.1, 20.2, 30.3, 40.4, 50.5, 60.6, 70.7, None, 90.9, 100.0],
    'Feature4': [True, False, True, False, True, False, True, False, True, False],
    'Feature5': ['X', 'Y', 'Z', 'X', 'Y', None, 'X', 'Y', 'Z', 'X']
}
```

```
In [3]: df=pd.DataFrame(data)
df
```

```
Out[3]:
```

	Feature1	Feature2	Feature3	Feature4	Feature5
0	1	A	10.1	True	X
1	2	B	20.2	False	Y
2	3	None	30.3	True	Z
3	4	D	40.4	False	X
4	5	E	50.5	True	Y
5	6	F	60.6	False	None
6	7	G	70.7	True	X
7	8	H	NaN	False	Y
8	9	I	90.9	True	Z
9	10	J	100.0	False	X

In [5]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   Feature1    10 non-null     int64
 1   Feature2     9 non-null     object
 2   Feature3     9 non-null     float64
 3   Feature4    10 non-null     bool
 4   Feature5     9 non-null     object
dtypes: bool(1), float64(1), int64(1), object(2)
memory usage: 462.0+ bytes
```

In [6]: `df.describe()`

Out[6]:

	Feature1	Feature3
count	10.00000	9.000000
mean	5.50000	52.633333
std	3.02765	30.845988
min	1.00000	10.100000
25%	3.25000	30.300000
50%	5.50000	50.500000
75%	7.75000	70.700000
max	10.00000	100.000000

In [7]: *#4th data will have index 3:*
`df.loc[3]`

Out[7]:

Feature1	4
Feature2	D
Feature3	40.4
Feature4	False
Feature5	X

Name: 3, dtype: object

```
In [8]: null_values = df.isnull()
null_values
```

```
Out[8]:
```

	Feature1	Feature2	Feature3	Feature4	Feature5
0	False	False	False	False	False
1	False	False	False	False	False
2	False	True	False	False	False
3	False	False	False	False	False
4	False	False	False	False	False
5	False	False	False	False	True
6	False	False	False	False	False
7	False	False	True	False	False
8	False	False	False	False	False
9	False	False	False	False	False

```
In [9]: null_count=null_values.sum()
null_count
```

```
Out[9]: Feature1    0
Feature2    1
Feature3    1
Feature4    0
Feature5    1
dtype: int64
```