

```
In [1]: #Experiment no.4
```

```
In [2]: #Aim:Missing value treatement
```

```
In [4]: #Name:Srushti Bawane  
#Roll no.:5  
#sec:B  
#sub:ET 1  
#date:20-08-2025
```

```
In [5]: import pandas as pd
```

```
In [1]: import os
```

```
In [7]: os.getcwd()
```

```
Out[7]: 'C:\\Users\\RH'
```

```
In [8]: os.chdir('C:\\Users\\RH\\Downloads')
```

```
In [13]: data=pd.read_csv("titanic - titanic.csv")
```

```
In [14]: data.head()
```


```
Out[14]:
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500

```
In [15]: data.describe()
```

Out[15]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204200
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693420
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200




```
In [16]: #check missing value by record
data.isna()
```

Out[16]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
0	False	False	False	False	False	False	False	False	False	False	True
1	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	True
3	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	True
...
886	False	False	False	False	False	False	False	False	False	False	True
887	False	False	False	False	False	False	False	False	False	False	False
888	False	False	False	False	False	True	False	False	False	False	True
889	False	False	False	False	False	False	False	False	False	False	False
890	False	False	False	False	False	False	False	False	False	False	True

891 rows × 12 columns



```
In [17]: ##check missing value by column
data.isna().any()
```

```
Out[17]: PassengerId    False
          Survived      False
          Pclass        False
          Name          False
          Sex           False
          Age           True
          SibSp         False
          Parch         False
          Ticket        False
          Fare          False
          Cabin         True
          Embarked      True
          dtype: bool
```

```
In [18]: data.isna().sum()
```

```
Out[18]: PassengerId    0
          Survived      0
          Pclass        0
          Name          0
          Sex           0
          Age          177
          SibSp         0
          Parch         0
          Ticket        0
          Fare          0
          Cabin        687
          Embarked      2
          dtype: int64
```

```
In [19]: #to drop element
          df1=data.dropna()
```

```
In [20]: df1
```

Out[20]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.283
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.100
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.862
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.700
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.550
...
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	female	47.0	1	1	11751	52.554
872	873	0	1	Carlsson, Mr. Frans Olof	male	33.0	0	0	695	5.000
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.158
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.000
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.000

183 rows × 12 columns



```
In [21]: df1.isna().sum()
```

```
Out[21]: PassengerId    0
         Survived      0
         Pclass       0
         Name         0
         Sex          0
         Age          0
         SibSp        0
         Parch        0
         Ticket       0
         Fare         0
         Cabin        0
         Embarked     0
         dtype: int64
```

```
In [22]: data["Age"].fillna(29.699118)
```

```
Out[22]: 0      22.000000
         1      38.000000
         2      26.000000
         3      35.000000
         4      35.000000
         ...
         886    27.000000
         887    19.000000
         888    29.699118
         889    26.000000
         890    32.000000
         Name: Age, Length: 891, dtype: float64
```

```
In [23]: data["Cabin"].fillna(29.699118)
```

```
Out[23]: 0      29.699118
         1           C85
         2      29.699118
         3           C123
         4      29.699118
         ...
         886    29.699118
         887           B42
         888    29.699118
         889           C148
         890    29.699118
         Name: Cabin, Length: 891, dtype: object
```

```
In [24]: import numpy as np
```

```
In [25]: #creating 1D array using numpy
```

```
In [26]: a1=np.array([10,20,30,40,50])
```

```
In [27]: a1
```

```
Out[27]: array([10, 20, 30, 40, 50])
```

```
In [28]: #creating 2D array using numpy
```

```
In [29]: a2=np.array([[10,20,30,40,50],[90,50,60,40,10]])
```

```
In [30]: a2
```

```
Out[30]: array([[10, 20, 30, 40, 50],  
               [90, 50, 60, 40, 10]])
```

```
In [31]: #creating multiD array using numpy
```

```
In [32]: a3=np.matrix('11,22;33,44;55,66')  
print(a3)
```

```
[[11 22]  
 [33 44]  
 [55 66]]
```

```
In [ ]:
```