

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

```
In [2]: df=pd.read_csv("https://raw.githubusercontent.com/dsrscientist/dataset4/ref
```

```
In [3]: df
```

Out[3]:

	Seat No.	PH- 121	HS- 101	CY- 105	HS- 105/12	MT- 111	CS- 105	CS- 106	EL- 102	EE- 119	...	CS- 312	CS- 317	CS- 403	CS- 421	CS- 406
0	CS- 97001	B-	D+	C-	C	C-	D+	D	C-	B-	...	C-	C-	C-	C-	A-
1	CS- 97002	A	D	D+	D	B-	C	D	A	D+	...	D+	D	C	D	A-
2	CS- 97003	A	B	A	B-	B+	A	B-	B+	A-	...	B	B	A	C	A
3	CS- 97004	D	C+	D+	D	D	A-	D+	C-	D	...	D+	C	D+	C-	B-
4	CS- 97005	A-	A-	A-	B+	A	A	A-	B+	A	...	B-	B+	B+	B-	A-
...
566	CS- 97567	B	A	A	A-	A+	A	A-	A-	A+	...	A-	A-	A	A	A
567	CS- 97568	A+	A	A	A	A	A	A	A-	A	...	B+	B+	A	A	A-
568	CS- 97569	B	A	A-	B+	A	A	A	A	A	...	A-	B	A	B+	A
569	CS- 97570	A	B+	D	A	D	D+	B-	C-	B-	...	D	B	B	C-	D
570	CS- 97571	C	D	D	C	C	D+	B	C+	C	...	C+	C	B-	D	F

571 rows × 43 columns



In [4]: `df.head()`

Out[4]:

	Seat No.	PH-121	HS-101	CY-105	HS-105/12	MT-111	CS-105	CS-106	EL-102	EE-119	...	CS-312	CS-317	CS-403	CS-421	CS-406	C 4'
0	CS-97001	B-	D+	C-	C	C-	D+	D	C-	B-	...	C-	C-	C-	C-	A-	
1	CS-97002	A	D	D+	D	B-	C	D	A	D+	...	D+	D	C	D	A-	
2	CS-97003	A	B	A	B-	B+	A	B-	B+	A-	...	B	B	A	C	A	
3	CS-97004	D	C+	D+	D	D	A-	D+	C-	D	...	D+	C	D+	C-	B-	
4	CS-97005	A-	A-	A-	B+	A	A	A-	B+	A	...	B-	B+	B+	B-	A-	

5 rows × 43 columns



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

Out[5]:

	Seat No.	PH-121	HS-101	CY-105	HS-105/12	MT-111	CS-105	CS-106	EL-102	EE-119	...	CS-312	CS-317	CS-403	CS-421	CS-406	
566	CS-97567	B	A	A	A-	A+	A	A-	A-	A+	...	A-	A-	A	A	A	
567	CS-97568	A+	A	A	A	A	A	A	A-	A	...	B+	B+	A	A	A-	
568	CS-97569	B	A	A-	B+	A	A	A	A	A	...	A-	B	A	B+	A	
569	CS-97570	A	B+	D	A	D	D+	B-	C-	B-	...	D	B	B	C-	D	
570	CS-97571	C	D	D	C	C	D+	B	C+	C	...	C+	C	B-	D	F	

5 rows × 43 columns



In [6]: `df.shape`

Out[6]: (571, 43)

```
In [7]: df.columns
```

```
Out[7]: Index(['Seat No.', 'PH-121', 'HS-101', 'CY-105', 'HS-105/12', 'MT-111',  
              'CS-105', 'CS-106', 'EL-102', 'EE-119', 'ME-107', 'CS-107', 'HS-20  
5/20',  
              'MT-222', 'EE-222', 'MT-224', 'CS-210', 'CS-211', 'CS-203', 'CS-21  
4',  
              'EE-217', 'CS-212', 'CS-215', 'MT-331', 'EF-303', 'HS-304', 'CS-30  
1',  
              'CS-302', 'TC-383', 'MT-442', 'EL-332', 'CS-318', 'CS-306', 'CS-31  
2',  
              'CS-317', 'CS-403', 'CS-421', 'CS-406', 'CS-414', 'CS-419', 'CS-42  
3',  
              'CS-412', 'CGPA'],  
              dtype='object')
```

```
In [8]: df.columns.tolist()
```

```
Out[8]: ['Seat No.',  
        'PH-121',  
        'HS-101',  
        'CY-105',  
        'HS-105/12',  
        'MT-111',  
        'CS-105',  
        'CS-106',  
        'EL-102',  
        'EE-119',  
        'ME-107',  
        'CS-107',  
        'HS-205/20',  
        'MT-222',  
        'EE-222',  
        'MT-224',  
        'CS-210',  
        'CS-211',  
        'CS-203',  
        'CS-214',  
        'EE-217',  
        'CS-212',  
        'CS-215',  
        'MT-331',  
        'EF-303',  
        'HS-304',  
        'CS-301',  
        'CS-302',  
        'TC-383',  
        'MT-442',  
        'EL-332',  
        'CS-318',  
        'CS-306',  
        'CS-312',  
        'CS-317',  
        'CS-403',  
        'CS-421',  
        'CS-406',  
        'CS-414',  
        'CS-419',  
        'CS-423',  
        'CS-412',  
        'CGPA']
```

In [9]: df.dtypes

```
Out[9]: Seat No.      object
PH-121      object
HS-101      object
CY-105      object
HS-105/12   object
MT-111      object
CS-105      object
CS-106      object
EL-102      object
EE-119      object
ME-107      object
CS-107      object
HS-205/20   object
MT-222      object
EE-222      object
MT-224      object
CS-210      object
CS-211      object
CS-203      object
CS-214      object
EE-217      object
CS-212      object
CS-215      object
MT-331      object
EF-303      object
HS-304      object
CS-301      object
CS-302      object
TC-383      object
MT-442      object
EL-332      object
CS-318      object
CS-306      object
CS-312      object
CS-317      object
CS-403      object
CS-421      object
CS-406      object
CS-414      object
CS-419      object
CS-423      object
CS-412      object
CGPA        float64
dtype: object
```

```
In [11]: df.isnull().sum()
```

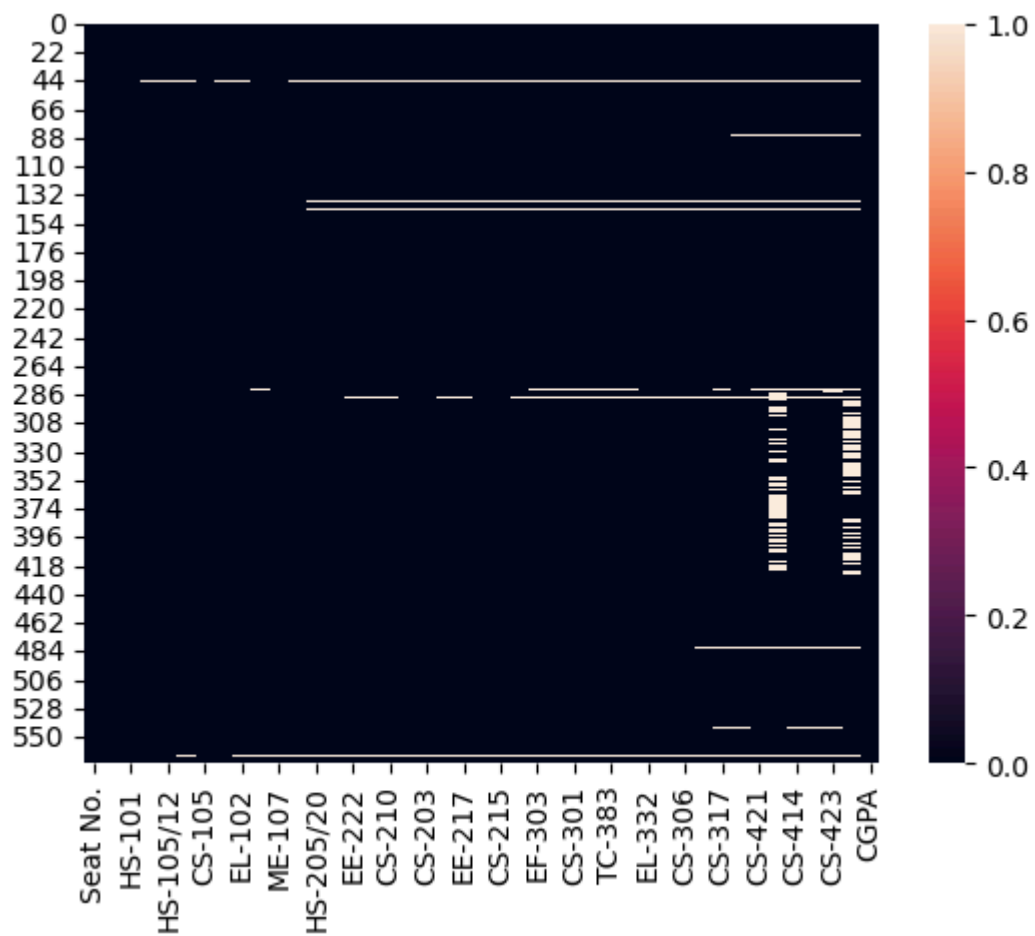
```
Out[11]: Seat No.      0  
PH-121      0  
HS-101      0  
CY-105      1  
HS-105/12   1  
MT-111      2  
CS-105      0  
CS-106      2  
EL-102      2  
EE-119      2  
ME-107      2  
CS-107      2  
HS-205/20   5  
MT-222      5  
EE-222      7  
MT-224      7  
CS-210      7  
CS-211      5  
CS-203      5  
CS-214      6  
EE-217      6  
CS-212      6  
CS-215      6  
MT-331      9  
EF-303     10  
HS-304     10  
CS-301     10  
CS-302     10  
TC-383     10  
MT-442     10  
EL-332      9  
CS-318      9  
CS-306      9  
CS-312     10  
CS-317     12  
CS-403     12  
CS-421     12  
CS-406     85  
CS-414     13  
CS-419     13  
CS-423     14  
CS-412     79  
CGPA        0  
dtype: int64
```

In [12]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 571 entries, 0 to 570
Data columns (total 43 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Seat No.        571 non-null   object
1   PH-121          571 non-null   object
2   HS-101          571 non-null   object
3   CY-105          570 non-null   object
4   HS-105/12       570 non-null   object
5   MT-111          569 non-null   object
6   CS-105          571 non-null   object
7   CS-106          569 non-null   object
8   EL-102          569 non-null   object
9   EE-119          569 non-null   object
10  ME-107          569 non-null   object
11  CS-107          569 non-null   object
12  HS-205/20       566 non-null   object
13  MT-222          566 non-null   object
14  EE-222          564 non-null   object
15  MT-224          564 non-null   object
16  CS-210          564 non-null   object
17  CS-211          566 non-null   object
18  CS-203          566 non-null   object
19  CS-214          565 non-null   object
20  EE-217          565 non-null   object
21  CS-212          565 non-null   object
22  CS-215          565 non-null   object
23  MT-331          562 non-null   object
24  EF-303          561 non-null   object
25  HS-304          561 non-null   object
26  CS-301          561 non-null   object
27  CS-302          561 non-null   object
28  TC-383          561 non-null   object
29  MT-442          561 non-null   object
30  EL-332          562 non-null   object
31  CS-318          562 non-null   object
32  CS-306          562 non-null   object
33  CS-312          561 non-null   object
34  CS-317          559 non-null   object
35  CS-403          559 non-null   object
36  CS-421          559 non-null   object
37  CS-406          486 non-null   object
38  CS-414          558 non-null   object
39  CS-419          558 non-null   object
40  CS-423          557 non-null   object
41  CS-412          492 non-null   object
42  CGPA            571 non-null   float64
dtypes: float64(1), object(42)
memory usage: 191.9+ KB
```

```
In [13]: sns.heatmap(df.isnull())
```

```
Out[13]: <Axes: >
```



In [14]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 571 entries, 0 to 570
Data columns (total 43 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Seat No.        571 non-null    object
1   PH-121          571 non-null    object
2   HS-101          571 non-null    object
3   CY-105          570 non-null    object
4   HS-105/12       570 non-null    object
5   MT-111          569 non-null    object
6   CS-105          571 non-null    object
7   CS-106          569 non-null    object
8   EL-102          569 non-null    object
9   EE-119          569 non-null    object
10  ME-107          569 non-null    object
11  CS-107          569 non-null    object
12  HS-205/20       566 non-null    object
13  MT-222          566 non-null    object
14  EE-222          564 non-null    object
15  MT-224          564 non-null    object
16  CS-210          564 non-null    object
17  CS-211          566 non-null    object
18  CS-203          566 non-null    object
19  CS-214          565 non-null    object
20  EE-217          565 non-null    object
21  CS-212          565 non-null    object
22  CS-215          565 non-null    object
23  MT-331          562 non-null    object
24  EF-303          561 non-null    object
25  HS-304          561 non-null    object
26  CS-301          561 non-null    object
27  CS-302          561 non-null    object
28  TC-383          561 non-null    object
29  MT-442          561 non-null    object
30  EL-332          562 non-null    object
31  CS-318          562 non-null    object
32  CS-306          562 non-null    object
33  CS-312          561 non-null    object
34  CS-317          559 non-null    object
35  CS-403          559 non-null    object
36  CS-421          559 non-null    object
37  CS-406          486 non-null    object
38  CS-414          558 non-null    object
39  CS-419          558 non-null    object
40  CS-423          557 non-null    object
41  CS-412          492 non-null    object
42  CGPA            571 non-null    float64
dtypes: float64(1), object(42)
memory usage: 191.9+ KB
```

```
In [15]: for i in df.columns:
          print(df[i].value_counts())
          print("\n")
```

```
WU      2
F        2
Name: count, dtype: int64
```

CS-107

```
A        105
A-        81
B+        57
B         55
C-        49
B-        43
A+        42
D         38
C+        34
D+        31
C         30
WU         2
I          1
F          1
```

```
In [ ]:
```

```
In [ ]:
```