```
import pandas as pd
# importing a data set
df = pd.read_csv("/content/Lab3.csv")
print(df)
```

```
    Sn Department     Name of student Program  \
0    1         BT              Ayushi  B.Tech
1    2         BT        Bharti Singh     NaN
2    3         BT       Muskan Bhatia     NaN
3    4         BT                 NaN   B.des
4    5         BT        Khushi Singh  B.Tech
5    6        NaN     Nikita Gajbhiye     NaN
6    7         BT           Pooja Rai  B.Tech
7    8         BT     Sanskriti Rawat  B.Tech
8    9         BT                 NaN  B.Arch
9   10        NaN        Deepak Kumar  B.Tech
10  11        ME                 NaN  B.Tech
11  11        ME        Biplop pandey  B.Tech
12  13        ME            Hariharan     NaN
13  14        NaN               Karan  B.Tech
14  15        ME           Kunal jain  B.Tech
15  16        NaN         Shaun joshi   B.des
16  17        ME                 NaN  B.Tech
17  18        ME         Augus pandey     NaN
18  19        CSE  Mr Ashim Chaudhary  B.Tech
19  20        NaN                NaN   B.des
20  21        CSE      Atamjeet Kohli  B.Tech
21  22        CSE    Miss Ashima Jain     NaN
22  23        ME                 NaN  B.Tech
23  24        ME         Augus pandey     NaN

    Paypackage at appointment in LPA PlacedBy
0                               4.50        D
1                               3.00        D
2                                NaN        D
3                               2.16        S
4                                NaN        S
5                               4.50      NaN
6                               4.50        S
7                               4.50        S
8                               4.50      NaN
9                                NaN        D
10                              3.90        S
11                               NaN      NaN
12                              4.40        S
13                              4.40      NaN
14                              4.40        S
15                              4.40      NaN
16                               NaN        S
17                              4.40        S
18                              7.00        S
19                              4.00      NaN
20                              7.50        D
21                               NaN        S
22                               NaN        S
23                              4.40        S
```

```
# Displaying the dimensions of data set
df.shape
```

```
(24, 6)
```

```
df.any()
```

```
Sn                                True
Department                        True
Name of student                   True
Program                           True
Paypackage at appointment in LPA  True
PlacedBy                          True
dtype: bool
```

```
print(df.isnull().any(axis=1))
```

```
0     False
1      True
2      True
3      True
4      True
5      True
6     False
7     False
8      True
9      True
10     True
11     True
12     True
13     True
14    False
15     True
16     True
17     True
18    False
19     True
20    False
21     True
22     True
23     True
dtype: bool
```

```
print(df.loc[:, df.isnull().any()])
```

```
   Department     Name of student Program  Paypackage at appointment in LPA  \
0          BT               Ayushi  B.Tech                              4.50
1          BT        Bharti Singh     NaN                              3.00
2          BT       Muskan Bhatia     NaN                               NaN
3          BT                 NaN   B.des                              2.16
4          BT        Khushi Singh  B.Tech                               NaN
5         NaN     Nikita Gajbhiye     NaN                              4.50
6          BT           Pooja Rai  B.Tech                              4.50
7          BT     Sanskriti Rawat  B.Tech                              4.50
8          BT                 NaN  B.Arch                              4.50
9         NaN        Deepak Kumar  B.Tech                               NaN
10        ME                  NaN  B.Tech                              3.90
11        ME        Biplop pandey  B.Tech                               NaN
12        ME            Hariharan     NaN                              4.40
13        NaN               Karan  B.Tech                              4.40
14        ME           Kunal jain  B.Tech                              4.40
15        NaN         Shaun joshi   B.des                              4.40
16        ME                  NaN  B.Tech                               NaN
17        ME         Augus pandey     NaN                              4.40
```

```
18        CSE  Mr Ashim Chaudhary  B.Tech                          7.00
19        NaN                 NaN   B.des                          4.00
20        CSE      Atamjeet Kohli  B.Tech                          7.50
21        CSE    Miss Ashima Jain     NaN                           NaN
22        ME                  NaN  B.Tech                           NaN
23        ME         Augus pandey     NaN                          4.40

   PlacedBy
0         D
1         D
2         D
3         S
4         S
5       NaN
6         S
7         S
8       NaN
9         D
10        S
11      NaN
12        S
13      NaN
14        S
15      NaN
16        S
17        S
18        S
19      NaN
20        D
21        S
22        S
23        S
```

```
print(df['Paypackage at appointment in LPA'])
```

```
0     4.50
1     3.00
2      NaN
3     2.16
4      NaN
5     4.50
6     4.50
7     4.50
8     4.50
9      NaN
10    3.90
11     NaN
12    4.40
13    4.40
14    4.40
15    4.40
16     NaN
17    4.40
18    7.00
19    4.00
20    7.50
21     NaN
22     NaN
23    4.40
Name: Paypackage at appointment in LPA, dtype: float64
```

```
s=0
for i in df['Sn']:
    s+=i
s=int(s)
a=s/len(df['Sn'])
print("The average Rating of the datas in our data set is : ",round(a,3))
```

```
The average Rating of the datas in our data set is :  12.458
```

```
# Displaying the first 10 rows of the data set
df.head(10)
```

	Sn	Department	Name of student	Program	Paypackage at appointment in LPA	PlacedBy
0	1	BT	Ayushi	B.Tech	4.50	D
1	2	BT	Bharti Singh	NaN	3.00	D
2	3	BT	Muskan Bhatia	NaN	NaN	D
3	4	BT	NaN	B.des	2.16	S
4	5	BT	Khushi Singh	B.Tech	NaN	S
5	6	NaN	Nikita Gajbhiye	NaN	4.50	NaN
6	7	BT	Pooja Rai	B.Tech	4.50	S
7	8	BT	Sanskriti Rawat	B.Tech	4.50	S
8	9	BT	NaN	B.Arch	4.50	NaN
9	10	NaN	Deepak Kumar	B.Tech	NaN	D

```
# Dtypes is used for what is data type for each column
df.dtypes
```

```
Sn                                int64
Department                       object
Name of student                  object
Program                          object
Paypackage at appointment in LPA    float64
PlacedBy                         object
dtype: object
```

```
# How many null values are present in each column can done by isnull().sum()
print(df.isnull().sum())
```

```
Sn                                0
Department                        5
Name of student                   4
Program                           7
Paypackage at appointment in LPA    7
PlacedBy                          6
dtype: int64
```

```
# For calculating the average of each column dataset
df.isna().sum()/len(df)*100
```

```
Sn                                 0.000000
Department                        20.833333
Name of student                   16.666667
Program                           29.166667
Paypackage at appointment in LPA  29.166667
PlacedBy                          25.000000
dtype: float64
```

```
# dropna used to drop rows and columns with missing values
"""Threshold value will be setup for the no.of non-null values will be retained.This means that any columns with less than 0.5% of non-null values will be dropped. The purpose of setting a threshold like thi
# axis will specifies that the operation should be applied to columns rather tham rows. axis=0 should be used for drop rows.
# This means that any columns with less than 0.5% of non-null values will be dropped. The purpose of setting a threshold like this is to remove columns that have too many missing values, which can make the d
# If inplace=True, the operation will modify the original DataFrame and return None. In other words, the changes are applied directly to the DataFrame object, and no new object is created.
# If inplace=False (which is the default), the operation will return a new DataFrame with the specified changes applied, leaving the original DataFrame unmodified.
df.dropna(thresh=0.5/100*len(df),axis=1,inplace=False)
```

	Sn	Department	Name of student	Program	Paypackage at appointment in LPA	PlacedBy
0	1	BT	Ayushi	B.Tech	4.50	D
1	2	BT	Bharti Singh	NaN	3.00	D
2	3	BT	Muskan Bhatia	NaN	NaN	D
3	4	BT	NaN	B.des	2.16	S
4	5	BT	Khushi Singh	B.Tech	NaN	S
5	6	NaN	Nikita Gajbhiye	NaN	4.50	NaN
6	7	BT	Pooja Rai	B.Tech	4.50	S
7	8	BT	Sanskriti Rawat	B.Tech	4.50	S
8	9	BT	NaN	B.Arch	4.50	NaN
9	10	NaN	Deepak Kumar	B.Tech	NaN	D
10	11	ME	NaN	B.Tech	3.90	S
11	11	ME	Biplop pandey	B.Tech	NaN	NaN
12	13	ME	Hariharan	NaN	4.40	S
13	14	NaN	Karan	B.Tech	4.40	NaN
14	15	ME	Kunal jain	B.Tech	4.40	S
15	16	NaN	Shaun joshi	B.des	4.40	NaN
16	17	ME	NaN	B.Tech	NaN	S
17	18	ME	Augus pandey	NaN	4.40	S
18	19	CSE	Mr Ashim Chaudhary	B.Tech	7.00	S
19	20	NaN	NaN	B.des	4.00	NaN
20	21	CSE	Atamjeet Kohli	B.Tech	7.50	D
21	22	CSE	Miss Ashima Jain	NaN	NaN	S
22	23	ME	NaN	B.Tech	NaN	S

```
# By using drop duplicates method we can rempve redudant columns
df.drop_duplicates
```

```
<bound method DataFrame.drop_duplicates of     Sn Department     Name of student Program  \
0    1         BT              Ayushi  B.Tech
1    2         BT        Bharti Singh     NaN
2    3         BT       Muskan Bhatia     NaN
3    4         BT                 NaN   B.des
4    5         BT        Khushi Singh  B.Tech
5    6        NaN     Nikita Gajbhiye     NaN
6    7         BT           Pooja Rai  B.Tech
7    8         BT     Sanskriti Rawat  B.Tech
8    9         BT                 NaN  B.Arch
9   10        NaN        Deepak Kumar  B.Tech
10  11        ME                 NaN  B.Tech
11  11        ME        Biplop pandey  B.Tech
12  13        ME            Hariharan     NaN
13  14        NaN               Karan  B.Tech
14  15        ME           Kunal jain  B.Tech
15  16        NaN         Shaun joshi   B.des
16  17        ME                 NaN  B.Tech
17  18        ME         Augus pandey     NaN
18  19        CSE  Mr Ashim Chaudhary  B.Tech
19  20        NaN                NaN   B.des
20  21        CSE      Atamjeet Kohli  B.Tech
21  22        CSE    Miss Ashima Jain     NaN
22  23        ME                 NaN  B.Tech
23  24        ME         Augus pandey     NaN

    Paypackage at appointment in LPA PlacedBy
0                               4.50        D
1                               3.00        D
2                                NaN        D
3                               2.16        S
4                                NaN        S
5                               4.50      NaN
6                               4.50        S
7                               4.50        S
8                               4.50      NaN
9                                NaN        D
10                              3.90        S
11                               NaN      NaN
12                              4.40        S
13                              4.40      NaN
14                              4.40        S
15                              4.40      NaN
16                               NaN        S
17                              4.40        S
18                              7.00        S
19                              4.00      NaN
20                              7.50        D
21                               NaN        S
22                               NaN        S
23                              4.40        S  >
```

```
#Remove the null values from the dataset by dropping the rows
df.dropna
```

```
<bound method DataFrame.dropna of     Sn Department     Name of student Program  \
0    1         BT              Ayushi  B.Tech
1    2         BT        Bharti Singh     NaN
```

```
2    3         BT       Muskan Bhatia     NaN
3    4         BT                 NaN   B.des
4    5         BT        Khushi Singh  B.Tech
5    6        NaN     Nikita Gajbhiye     NaN
6    7         BT           Pooja Rai  B.Tech
7    8         BT     Sanskriti Rawat  B.Tech
8    9         BT                 NaN  B.Arch
9   10        NaN        Deepak Kumar  B.Tech
10  11        ME                 NaN  B.Tech
11  11        ME        Biplop pandey  B.Tech
12  13        ME            Hariharan     NaN
13  14        NaN               Karan  B.Tech
14  15        ME           Kunal jain  B.Tech
15  16        NaN         Shaun joshi   B.des
16  17        ME                 NaN  B.Tech
17  18        ME         Augus pandey     NaN
18  19        CSE  Mr Ashim Chaudhary  B.Tech
19  20        NaN                NaN   B.des
20  21        CSE      Atamjeet Kohli  B.Tech
21  22        CSE    Miss Ashima Jain     NaN
22  23        ME                 NaN  B.Tech
23  24        ME         Augus pandey     NaN

    Paypackage at appointment in LPA PlacedBy
0                               4.50        D
1                               3.00        D
2                                NaN        D
3                               2.16        S
4                                NaN        S
5                               4.50      NaN
6                               4.50        S
7                               4.50        S
8                               4.50      NaN
9                                NaN        D
10                              3.90        S
11                               NaN      NaN
12                              4.40        S
13                              4.40      NaN
14                              4.40        S
15                              4.40      NaN
16                               NaN        S
17                              4.40        S
18                              7.00        S
19                              4.00      NaN
20                              7.50        D
21                               NaN        S
22                               NaN        S
23                              4.40        S  >
```

```
# Remove the null values from the dataset by dropping the columns
df=df.dropna(axis=1)
df
```

	Sn	Department	Name of student	Program	Paypackage at appointment in LPA	PlacedBy
0	1	BT	Ayushi	B.Tech	4.5	D

```
# Remove the null values from the dataset by dropping the rows
df=df.dropna(axis=0)
df
```

	Sn	Department	Name of student	Program	Paypackage at appointment in LPA	PlacedBy
0	1	BT	Ayushi	B.Tech	4.5	D
6	7	BT	Pooja Rai	B.Tech	4.5	S
7	8	BT	Sanskriti Rawat	B.Tech	4.5	S
14	15	ME	Kunal jain	B.Tech	4.4	S
18	19	CSE	Mr Ashim Chaudhary	B.Tech	7.0	S

```
#  Calculate the average of each numeric columns of the dataset
df.mean()
```

```
<ipython-input-6-6c9a7d4a3ac7>:2: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError.  Select onl
  df.mean()
Sn                              11.833333
Paypackage at appointment in LPA     5.400000
dtype: float64
```

```
df.isna().sum()/len(df)*100
```

```
Sn    0.0
dtype: float64
```

```
#Display all the unique values of each column
unique_values = [df[col].unique() for col in df.columns]
for i, col in enumerate(df.columns):
    print(f"Unique values of {col}: {unique_values[i]}")
```

```
Unique values of Sn: [ 1  2  3  4  5  6  7  8  9 10 11 13 14 15 16 17 18 19 20 21 22 23 24]
Unique values of Department: ['BT' nan 'ME ' 'CSE']
Unique values of Name of student: ['Ayushi' 'Bharti Singh' 'Muskan Bhatia' nan 'Khushi Singh'
 'Nikita Gajbhiye' 'Pooja Rai' 'Sanskriti Rawat' 'Deepak Kumar'
 'Biplop pandey' 'Hariharan' 'Karan' 'Kunal jain' 'Shaun joshi' ' NaN'
 'Augus pandey' 'Mr Ashim Chaudhary' 'Atamjeet Kohli' 'Miss Ashima Jain']
Unique values of Program: ['B.Tech' nan 'B.des' 'B.Arch']
Unique values of Paypackage at appointment in LPA: [4.5  3.    nan 2.16 3.9  4.4  7.   4.   7.5 ]
Unique values of PlacedBy: ['D' 'S' nan]
```

```
#Display all the unique values of each column
unique_values = [df[col].unique() for col in df.columns]
for i, col in enumerate(df.columns):
    print(f"Unique values of {col}: {unique_values[i]}")
```

```
Unique values of Sn: [ 1  2  3  4  5  6  7  8  9 10 11 13 14 15 16 17 18 19 20 21 22 23 24]
Unique values of Department: ['BT' nan 'ME ' 'CSE']
Unique values of Name of student: ['Ayushi' 'Bharti Singh' 'Muskan Bhatia' nan 'Khushi Singh'
 'Nikita Gajbhiye' 'Pooja Rai' 'Sanskriti Rawat' 'Deepak Kumar'
 'Biplop pandey' 'Hariharan' 'Karan' 'Kunal jain' 'Shaun joshi' ' NaN'
 'Augus pandey' 'Mr Ashim Chaudhary' 'Atamjeet Kohli' 'Miss Ashima Jain']
Unique values of Program: ['B.Tech' nan 'B.des' 'B.Arch']
Unique values of Paypackage at appointment in LPA: [4.5  3.    nan 2.16 3.9  4.4  7.   4.   7.5 ]
Unique values of PlacedBy: ['D' 'S' nan]
```

```
#Fill the null values of numeric columns with the average value(mean) without using imputer method
mean_values = df.mean().to_dict()

for col in df.columns:
    if df[col].dtype == 'float64':
        mean = mean_values[col]
        df[col].fillna(mean, inplace=True)

print(df)
```

```
    Sn Department     Name of student Program  \
0    1         BT              Ayushi  B.Tech
1    2         BT        Bharti Singh     NaN
2    3         BT       Muskan Bhatia     NaN
3    4         BT                 NaN   B.des
4    5         BT        Khushi Singh  B.Tech
5    6        NaN     Nikita Gajbhiye     NaN
6    7         BT           Pooja Rai  B.Tech
7    8         BT     Sanskriti Rawat  B.Tech
8    9         BT                 NaN  B.Arch
9   10        NaN        Deepak Kumar  B.Tech
10  11        ME                 NaN  B.Tech
11  11        ME        Biplop pandey  B.Tech
12  13        ME            Hariharan     NaN
13  14        NaN               Karan  B.Tech
14  15        ME           Kunal jain  B.Tech
15  16        NaN         Shaun joshi   B.des
16  17        ME                 NaN  B.Tech
17  18        ME         Augus pandey     NaN
18  19        CSE  Mr Ashim Chaudhary  B.Tech
19  20        NaN                NaN   B.des
20  21        CSE      Atamjeet Kohli  B.Tech
21  22        CSE    Miss Ashima Jain     NaN
22  23        ME                 NaN  B.Tech
23  24        ME         Augus pandey     NaN

    Paypackage at appointment in LPA PlacedBy
0                           4.500000        D
1                           3.000000        D
2                           4.497647        D
3                           2.160000        S
4                           4.497647        S
5                           4.500000      NaN
6                           4.500000        S
7                           4.500000        S
8                           4.500000      NaN
9                           4.497647        D
10                          3.900000        S
11                          4.497647      NaN
12                          4.400000        S
13                          4.400000      NaN
```

```
14                          4.400000        S
15                          4.400000      NaN
16                          4.497647        S
17                          4.400000        S
18                          7.000000        S
19                          4.000000      NaN
20                          7.500000        D
21                          4.497647        S
22                          4.497647        S
23                          4.400000        S
<ipython-input-4-0888a6c2e796>:2: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError.  Select onl
  mean_values = df.mean().to_dict()
```

```
# Fill the null values of categorical/text columns with the most frequent value(mode) without using imputer method
mode_values = df.mode().to_dict(orient='records')[0]

# Fill the null values with the mode
for col in df.columns:
    if df[col].dtype == 'object':
        mode = mode_values[col]
        df[col].fillna(mode, inplace=True)

# Print the DataFrame
print(df)
```

```
    Sn Department     Name of student Program  \
0    1         BT              Ayushi  B.Tech
1    2         BT        Bharti Singh  B.Tech
2    3         BT       Muskan Bhatia  B.Tech
3    4         BT                 NaN   B.des
4    5         BT        Khushi Singh  B.Tech
5    6         BT     Nikita Gajbhiye  B.Tech
6    7         BT           Pooja Rai  B.Tech
7    8         BT     Sanskriti Rawat  B.Tech
8    9         BT                 NaN  B.Arch
9   10         BT        Deepak Kumar  B.Tech
10  11        ME                 NaN  B.Tech
11  11        ME        Biplop pandey  B.Tech
12  13        ME            Hariharan  B.Tech
13  14         BT               Karan  B.Tech
14  15        ME           Kunal jain  B.Tech
15  16         BT         Shaun joshi   B.des
16  17        ME                 NaN  B.Tech
17  18        ME         Augus pandey  B.Tech
18  19        CSE  Mr Ashim Chaudhary  B.Tech
19  20         BT                NaN   B.des
20  21        CSE      Atamjeet Kohli  B.Tech
21  22        CSE    Miss Ashima Jain  B.Tech
22  23        ME                 NaN  B.Tech
23  24        ME         Augus pandey  B.Tech

    Paypackage at appointment in LPA PlacedBy
0                               4.50        D
1                               3.00        D
2                                NaN        D
3                               2.16        S
4                                NaN        S
5                               4.50        S
6                               4.50        S
7                               4.50        S
```

```
8                              4.50        S
9                               NaN        D
10                             3.90        S
11                              NaN        S
12                             4.40        S
13                             4.40        S
14                             4.40        S
15                             4.40        S
16                              NaN        S
17                             4.40        S
18                             7.00        S
19                             4.00        S
20                             7.50        D
21                              NaN        S
22                              NaN        S
23                             4.40        S
```

```
# Fill the null values of both categorical/text and numeric columns with the most frequent value(mode) by using sklearn library's imputer method
from sklearn.impute import SimpleImputer
import numpy as np
imputer = SimpleImputer(strategy='most_frequent')

df = pd.DataFrame(imputer.fit_transform(df), columns=df.columns)

print(df)
```

```
    Sn Department     Name of student Program  \
0    1         BT              Ayushi  B.Tech
1    2         BT        Bharti Singh  B.Tech
2    3         BT       Muskan Bhatia  B.Tech
3    4         BT                 NaN   B.des
4    5         BT       Khushi Singh  B.Tech
5    6         BT     Nikita Gajbhiye  B.Tech
6    7         BT           Pooja Rai  B.Tech
7    8         BT     Sanskriti Rawat  B.Tech
8    9         BT                 NaN  B.Arch
9   10         BT        Deepak Kumar  B.Tech
10  11        ME                 NaN  B.Tech
11  11        ME        Biplop pandey  B.Tech
12  13        ME            Hariharan  B.Tech
13  14         BT              Karan  B.Tech
14  15        ME           Kunal jain  B.Tech
15  16         BT         Shaun joshi   B.des
16  17        ME                 NaN  B.Tech
17  18        ME         Augus pandey  B.Tech
18  19        CSE  Mr Ashim Chaudhary  B.Tech
19  20         BT                NaN   B.des
20  21        CSE      Atamjeet Kohli  B.Tech
21  22        CSE    Miss Ashima Jain  B.Tech
22  23        ME                 NaN  B.Tech
23  24        ME         Augus pandey  B.Tech

   Paypackage at appointment in LPA PlacedBy
0                              4.5        D
1                              3.0        D
2                         4.497647        D
3                             2.16        S
4                         4.497647        S
5                              4.5        S
6                              4.5        S
7                              4.5        S
8                              4.5        S
```

```
9                          4.497647        D
10                              3.9        S
11                         4.497647        S
12                              4.4        S
13                              4.4        S
14                              4.4        S
15                              4.4        S
16                         4.497647        S
17                              4.4        S
18                              7.0        S
19                              4.0        S
20                              7.5        D
21                         4.497647        S
22                         4.497647        S
23                              4.4        S
```

```
# Fill the null values of both categorical/text and numeric columns with the most frequent value(mode) by using sklearn library's imputer method on only specific rows and columns using iloc method
imputer = SimpleImputer(strategy='most_frequent')
df_subset = df.iloc[1:4, 1:3]
df_subset = pd.DataFrame(imputer.fit_transform(df_subset), columns=df_subset.columns)
df.iloc[1:4, 1:3] = df_subset
print(df)
```

```
    Sn Department     Name of student Program  \
0    1         BT              Ayushi  B.Tech
1    2         BT        Bharti Singh  B.Tech
2    3         BT       Muskan Bhatia  B.Tech
3    4         BT                 NaN   B.des
4    5         BT        Khushi Singh  B.Tech
5    6         BT     Nikita Gajbhiye  B.Tech
6    7         BT           Pooja Rai  B.Tech
7    8         BT     Sanskriti Rawat  B.Tech
8    9         BT                 NaN  B.Arch
9   10         BT        Deepak Kumar  B.Tech
10  11        ME                 NaN  B.Tech
11  11        ME        Biplop pandey  B.Tech
12  13        ME            Hariharan  B.Tech
13  14         BT              Karan  B.Tech
14  15        ME           Kunal jain  B.Tech
15  16         BT         Shaun joshi   B.des
16  17        ME                 NaN  B.Tech
17  18        ME         Augus pandey  B.Tech
18  19        CSE  Mr Ashim Chaudhary  B.Tech
19  20         BT                NaN   B.des
20  21        CSE      Atamjeet Kohli  B.Tech
21  22        CSE    Miss Ashima Jain  B.Tech
22  23        ME                 NaN  B.Tech
23  24        ME         Augus pandey  B.Tech

   Paypackage at appointment in LPA PlacedBy
0                               4.5        D
1                               3.0        D
2                          4.497647        D
3                              2.16        S
4                          4.497647        S
5                               4.5        S
6                               4.5        S
7                               4.5        S
8                               4.5        S
9                          4.497647        D
10                              3.9        S
```

```
11                      4.497647       S
12                           4.4       S
13                           4.4       S
14                           4.4       S
15                           4.4       S
16                      4.497647       S
17                           4.4       S
18                           7.0       S
19                           4.0       S
20                           7.5       D
21                      4.497647       S
22                      4.497647       S
23                           4.4       S
```

```
# Use the describe() function to display the mean, standard deviation, and Inter Quartile Range (IQR) values of the numeric columns of the dataset.
df.describe()
```

	Sn	Department	Name of student	Program	Paypackage at appointment in LPA	PlacedBy
count	24	24	24	24	24.0	24
unique	23	3	18	3	8.0	2
top	11	BT	NaN	B.Tech	4.4	S
freq	2	13	6	20	13.0	19

```
# Show the count of null values and missing values seperately in each column of the dataset

null_counts = df.isnull().sum()
print("Null values in each column:")
print(null_counts)
missing_counts = df.isna().sum()
print("\nMissing values in each column:")
print(missing_counts)
```

```
Null values in each column:
Sn                              0
Department                      0
Name of student                 0
Program                         0
Paypackage at appointment in LPA    0
PlacedBy                        0
dtype: int64

Missing values in each column:
Sn                              0
Department                      0
Name of student                 0
Program                         0
Paypackage at appointment in LPA    0
PlacedBy                        0
dtype: int64
```

```
# Export the cleaned dataset to a new csv file without null or missing or duplicate value.
# Remove missing values
df = df.dropna()
# Remove duplicates
```

```
df = df.drop_duplicates()

# Save the cleaned data to a new CSV file
df.to_csv('cleaned_data.csv', index=False)
```

```
def function(a):
  return a*a
x=map(function, (1,2,3,4))
print(set(x))
```

```
{16, 1, 4, 9}
```

```
tup=(5,27,46,48,92,17,2,84)
newtuple = tuple(map(lamda x:x+3))
```

```
from functools import reduce
reduce(lambda a,b: a+b,[52,6,45,3])
```

```
106106
```