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

In [2]:
```
Data1 = pd.read_csv('AcademicPerformance.csv')
print(Data1)
```

```
    gender NationalITy PlaceofBirth       StageID GradeID SectionID  \
0        M          KW       KuwaIT    lowerlevel    G-04         A
1        M          KW       KuwaIT    lowerlevel    G-04       NaN
2        M          KW       KuwaIT    lowerlevel    G-04         A
3      NaN          KW       KuwaIT    lowerlevel    G-04         A
4        M          KW       KuwaIT    lowerlevel    G-04         A
..     ...         ...          ...           ...     ...       ...
475      F      Jordan       Jordan  MiddleSchool    G-08         A
476      F      Jordan       Jordan  MiddleSchool    G-08         A
477      F      Jordan       Jordan  MiddleSchool    G-08         A
478      F      Jordan       Jordan  MiddleSchool    G-08         A
479      F      Jordan       Jordan  MiddleSchool    G-08         A

         Topic Semester Relation  raisedhands  VisITedResources  \
0           IT        F   Father         15.0              16.0
1           IT        F   Father         20.0              20.0
2           IT      NaN   Father         10.0               7.0
3           IT        F   Father         30.0              25.0
4           IT        F   Father         40.0              50.0
..         ...      ...      ...          ...               ...
475  Chemistry        S   Father          5.0               4.0
476    Geology        F   Father         50.0              77.0
477    Geology        S   Father         55.0              74.0
478    History        F   Father         30.0              17.0
479    History        S   Father         35.0              14.0

     AnnouncementsView  Discussion ParentAnsweringSurvey  \
0                  2.0          20                   Yes
1                  3.0          25                   Yes
2                  0.0          30                    15
3                  5.0          35                    No
4                 12.0          50                    No
..                 ...         ...                   ...
475                5.0           8                    No
476               14.0          28                    No
477               25.0          29                    No
478               14.0          57                    No
479               23.0          62                    No

    ParentschoolSatisfaction StudentAbsenceDays Class
0                       Good            Under-7     M
1                       Good            Under-7     M
2                        Bad            Above-7     L
3                        Bad            Above-7     L
4                        Bad            Above-7     M
..                       ...                ...   ...
475                      Bad            Above-7     L
476                      Bad            Under-7     M
477                      Bad            Under-7     M
478                      Bad            Above-7     L
479                      Bad            Above-7     L

[480 rows x 17 columns]
```

In [3]:
```
Data1.shape
```

Out[3]:
```
(480, 17)
```

In [4]:
```
Data1.isnull().sum()
```

Out[4]:
```
gender                    10
NationalITy                0
PlaceofBirth               0
StageID                    0
GradeID                    0
SectionID                  6
Topic                      0
Semester                   9
Relation                   0
raisedhands               10
VisITedResources           5
AnnouncementsView          4
Discussion                 0
ParentAnsweringSurvey      0
ParentschoolSatisfaction   0
StudentAbsenceDays         0
Class                      0
dtype: int64
```

In [5]:
```
Data1.dropna(inplace=True)
print(Data1.isnull().sum())
```

```
gender                    0
NationalITy               0
PlaceofBirth              0
StageID                   0
GradeID                   0
SectionID                 0
Topic                     0
Semester                  0
Relation                  0
raisedhands               0
VisITedResources          0
AnnouncementsView         0
Discussion                0
ParentAnsweringSurvey     0
ParentschoolSatisfaction  0
StudentAbsenceDays        0
Class                     0
dtype: int64
```

In [6]:
```
import pandas as pd
import numpy as np
```

In [7]:
```
Data2 = pd.read_csv('AcademicPerformance.csv')
print(Data2)
```

```
    gender NationalITy PlaceofBirth       StageID GradeID SectionID  \
0        M          KW       KuwaIT    lowerlevel    G-04         A
1        M          KW       KuwaIT    lowerlevel    G-04       NaN
2        M          KW       KuwaIT    lowerlevel    G-04         A
3      NaN          KW       KuwaIT    lowerlevel    G-04         A
4        M          KW       KuwaIT    lowerlevel    G-04         A
..     ...         ...          ...           ...     ...       ...
475      F      Jordan       Jordan  MiddleSchool    G-08         A
```

```
476      F      Jordan       Jordan  MiddleSchool    G-08         A
477      F      Jordan       Jordan  MiddleSchool    G-08         A
478      F      Jordan       Jordan  MiddleSchool    G-08         A
479      F      Jordan       Jordan  MiddleSchool    G-08         A

         Topic Semester Relation  raisedhands  VisITedResources  \
0           IT        F   Father         15.0              16.0
1           IT        F   Father         20.0              20.0
2           IT      NaN   Father         10.0               7.0
3           IT        F   Father         30.0              25.0
4           IT        F   Father         40.0              50.0
..         ...      ...      ...          ...               ...
475  Chemistry        S   Father          5.0               4.0
476    Geology        F   Father         50.0              77.0
477    Geology        S   Father         55.0              74.0
478    History        F   Father         30.0              17.0
479    History        S   Father         35.0              14.0

     AnnouncementsView  Discussion ParentAnsweringSurvey  \
0                  2.0          20                   Yes
1                  3.0          25                   Yes
2                  0.0          30                    15
3                  5.0          35                    No
4                 12.0          50                    No
..                 ...         ...                   ...
475                5.0           8                    No
476               14.0          28                    No
477               25.0          29                    No
478               14.0          57                    No
479               23.0          62                    No

    ParentschoolSatisfaction StudentAbsenceDays Class
0                       Good            Under-7     M
1                       Good            Under-7     M
2                        Bad            Above-7     L
3                        Bad            Above-7     L
4                        Bad            Above-7     M
..                       ...                ...   ...
475                      Bad            Above-7     L
476                      Bad            Under-7     M
477                      Bad            Under-7     M
478                      Bad            Above-7     L
479                      Bad            Above-7     L

[480 rows x 17 columns]
```

In [8]:
```
Data2.notnull()
```

Out[8]:

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Relation	rais
0	True	True	True	True	True	True	True	True	True	
1	True	True	True	True	True	False	True	True	True	
2	True	True	True	True	True	True	True	False	True	
3	False	True	True	True	True	True	True	True	True	
4	True	True	True	True	True	True	True	True	True	
...	...	...	...	...	...	...	...	...	...	
475	True	True	True	True	True	True	True	True	True	
476	True	True	True	True	True	True	True	True	True	

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Relation	rai
477	True	True	True	True	True	True	True	True	True	
478	True	True	True	True	True	True	True	True	True	
479	True	True	True	True	True	True	True	True	True	

480 rows × 17 columns

In [9]:
```
series1 = pd.notnull(Data2["raisedhands"])
Data2[series1]
```

Out[9]:

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Rela
0	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
1	M	KW	KuwaIT	lowerlevel	G-04	NaN	IT	F	Fa
2	M	KW	KuwaIT	lowerlevel	G-04	A	IT	NaN	Fa
3	NaN	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
4	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
...	...	...	...	...	...	...	...	...	
475	F	Jordan	Jordan	MiddleSchool	G-08	A	Chemistry	S	Fa
476	F	Jordan	Jordan	MiddleSchool	G-08	A	Geology	F	Fa
477	F	Jordan	Jordan	MiddleSchool	G-08	A	Geology	S	Fa
478	F	Jordan	Jordan	MiddleSchool	G-08	A	History	F	Fa
479	F	Jordan	Jordan	MiddleSchool	G-08	A	History	S	Fa

470 rows × 17 columns

In [10]:
```
df=Data2
df.fillna(0)
```

Out[10]:

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Rela
0	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
1	M	KW	KuwaIT	lowerlevel	G-04	0	IT	F	Fa
2	M	KW	KuwaIT	lowerlevel	G-04	A	IT	0	Fa
3	0	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
4	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
...	...	...	...	...	...	...	...	...	
475	F	Jordan	Jordan	MiddleSchool	G-08	A	Chemistry	S	Fa
476	F	Jordan	Jordan	MiddleSchool	G-08	A	Geology	F	Fa
477	F	Jordan	Jordan	MiddleSchool	G-08	A	Geology	S	Fa

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Rela
478	F	Jordan	Jordan	MiddleSchool	G-08	A	History	F	Fa
479	F	Jordan	Jordan	MiddleSchool	G-08	A	History	S	Fa

480 rows × 17 columns

In [11]:
```
df['GradeID'].fillna('Grade')
```

Out[11]:
```
0      G-04
1      G-04
2      G-04
3      G-04
4      G-04
       ...
475    G-08
476    G-08
477    G-08
478    G-08
479    G-08
Name: GradeID, Length: 480, dtype: object
```

In [12]:
```
df[ 'VisITedResources'].fillna(df[ 'VisITedResources'].mean())
```

Out[12]:
```
0      16.0
1      20.0
2       7.0
3      25.0
4      50.0
       ...
475     4.0
476    77.0
477    74.0
478    17.0
479    14.0
Name: VisITedResources, Length: 480, dtype: float64
```

In [13]:
```
df[ 'VisITedResources'].fillna(df[ 'VisITedResources'].median())
```

Out[13]:
```
0      16.0
1      20.0
2       7.0
3      25.0
4      50.0
       ...
475     4.0
476    77.0
477    74.0
478    17.0
479    14.0
Name: VisITedResources, Length: 480, dtype: float64
```

In [15]:
```
df[ 'Class'].value_counts()
```

Out[15]:
```
M    211
H    142
L    127
Name: Class, dtype: int64
```

In [16]:
```
df['Class'].fillna(df['Class'].mode()[0])
```

Out[16]:
```
0      M
1      M
2      L
3      L
4      M
      ..
475    L
476    M
477    M
478    L
479    L
Name: Class, Length: 480, dtype: object
```

In [17]:
```
df.fillna(method="backfill")
```

Out[17]:

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Rela
0	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
1	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
2	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
3	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
4	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
...	...	...	...	...	...	...	...	...	
475	F	Jordan	Jordan	MiddleSchool	G-08	A	Chemistry	S	Fa
476	F	Jordan	Jordan	MiddleSchool	G-08	A	Geology	F	Fa
477	F	Jordan	Jordan	MiddleSchool	G-08	A	Geology	S	Fa
478	F	Jordan	Jordan	MiddleSchool	G-08	A	History	F	Fa
479	F	Jordan	Jordan	MiddleSchool	G-08	A	History	S	Fa

480 rows × 17 columns

In [18]:
```
df.fillna(method="pad")
```

Out[18]:

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Rela
0	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
1	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
2	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
3	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
4	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Fa
...	...	...	...	...	...	...	...	...	

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester	Rela
475	F	Jordan	Jordan	MiddleSchool	G-08	A	Chemistry	S	Fa
476	F	Jordan	Jordan	MiddleSchool	G-08	A	Geology	F	Fa
477	F	Jordan	Jordan	MiddleSchool	G-08	A	Geology	S	Fa
478	F	Jordan	Jordan	MiddleSchool	G-08	A	History	F	Fa
479	F	Jordan	Jordan	MiddleSchool	G-08	A	History	S	Fa

480 rows × 17 columns

In [19]:
```
df.describe()
```

Out[19]:

	raisedhands	VisITedResources	AnnouncementsView	Discussion
count	470.000000	475.000000	476.000000	480.000000
mean	47.506383	55.374737	38.237395	43.283333
std	30.483226	32.769213	26.493621	27.637735
min	0.000000	0.000000	0.000000	1.000000
25%	17.000000	20.000000	15.000000	20.000000
50%	50.000000	66.000000	33.000000	39.000000
75%	75.000000	84.000000	58.000000	70.000000
max	100.000000	99.000000	98.000000	99.000000

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: