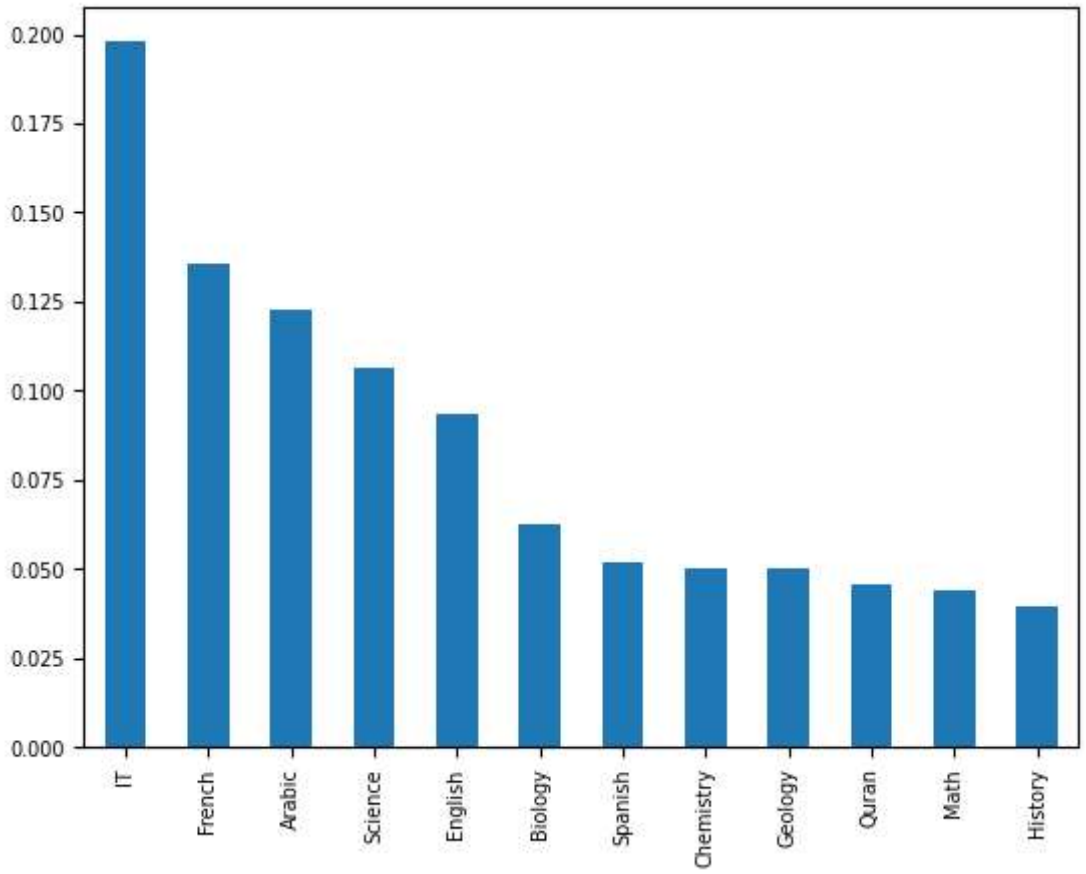
In [5]:
```
data.columns
```

Out[5]:
```
Index(['gender', 'NationalITy', 'PlaceofBirth', 'StageID', 'GradeID',
       'SectionID', 'Topic', 'Semester', 'Relation', 'raisedhands',
       'VisITedResources', 'AnnouncementsView', 'Discussion',
       'ParentAnsweringSurvey', 'ParentschoolSatisfaction',
       'StudentAbsenceDays', 'Class'],
      dtype='object')
```

In [4]:
```
print("percentage",data.Topic.value_counts(normalize=True))
data.Topic.value_counts(normalize=True).plot(kind='bar',fontsize=7)
```

```
percentage IT           0.197917
French       0.135417
Arabic       0.122917
Science      0.106250
English      0.093750
Biology      0.062500
Spanish      0.052083
Chemistry    0.050000
Geology      0.050000
Quran        0.045833
Math         0.043750
History      0.039583
Name: Topic, dtype: float64
```

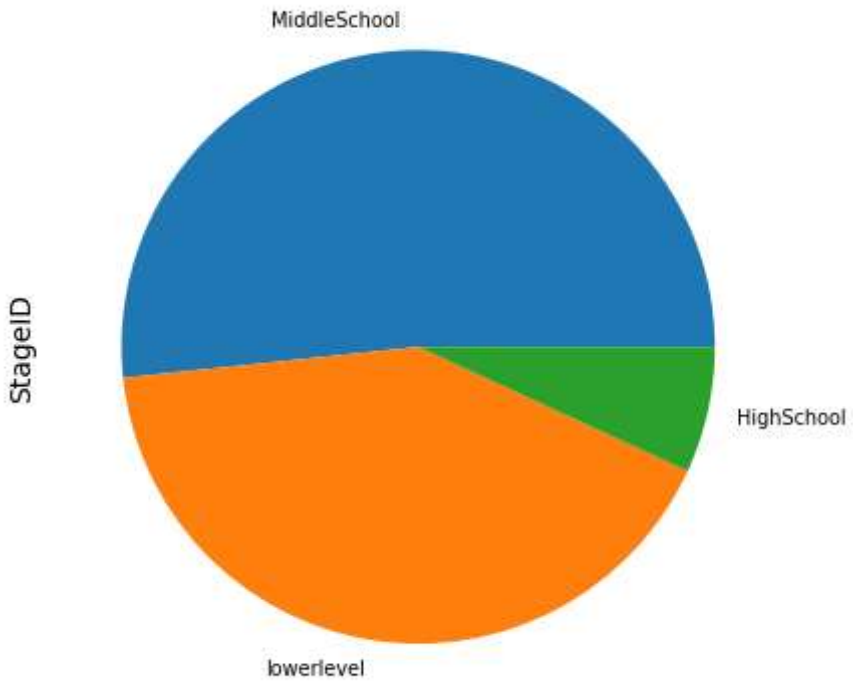
Out[4]: <Axes: >



In [3]:
```
print("percentage",data.StageID.value_counts(normalize=True))
data.StageID.value_counts(normalize=True).plot(kind='pie',fontsize=7)
```

```
percentage MiddleSchool    0.516667
lowerlevel      0.414583
HighSchool      0.068750
Name: StageID, dtype: float64
```

Out[3]: <Axes: ylabel='StageID'>




In [3]:
```
import numpy as np #load and read data
import pandas as pd
data=pd.read_csv(r'E:\NUST-dairy\Module\Project\AI-Data.csv') #r is use to tak
data.head() #use to extract data
```

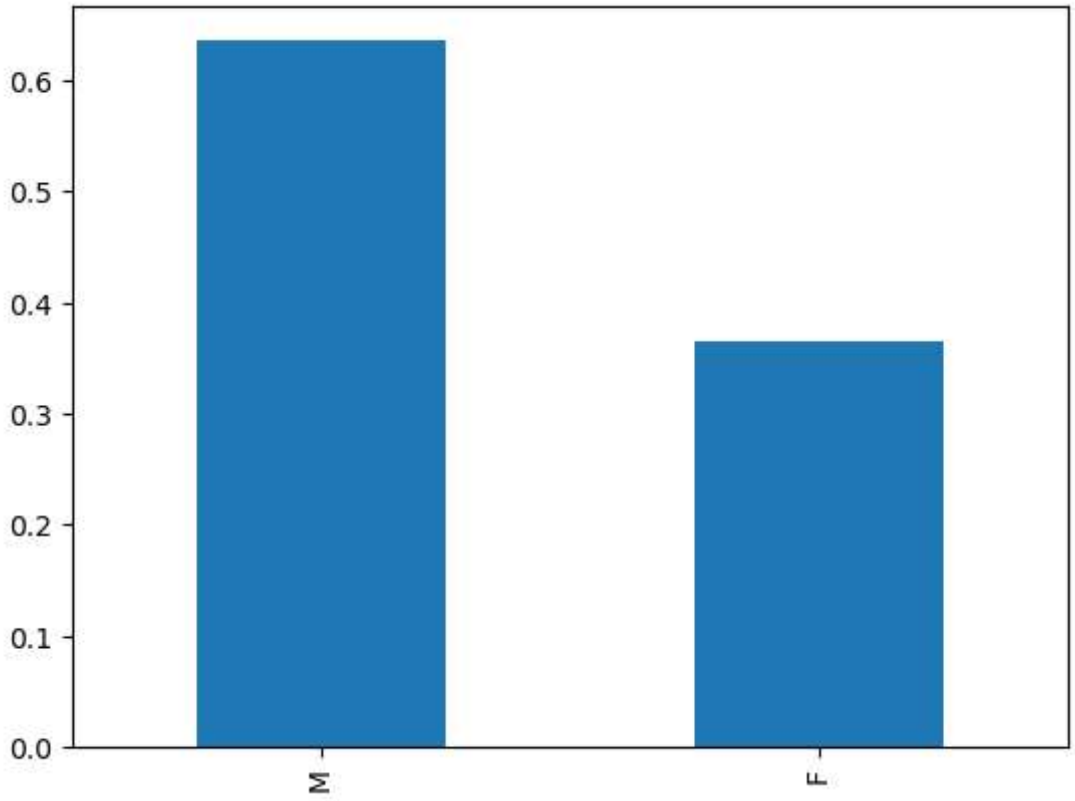
Out[3]:

| | gender | NationalITy | PlaceofBirth | StageID | GradeID | SectionID | Topic | Semester | Relation | ra |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | M | KW | KuwaIT | lowerlevel | G-04 | A | IT | F | Father | |
| 1 | M | KW | KuwaIT | lowerlevel | G-04 | A | IT | F | Father | |
| 2 | M | KW | KuwaIT | lowerlevel | G-04 | A | IT | F | Father | |
| 3 | M | KW | KuwaIT | lowerlevel | G-04 | A | IT | F | Father | |
| 4 | M | KW | KuwaIT | lowerlevel | G-04 | A | IT | F | Father | |

In [32]:
```
print("percentage",data.gender.value_counts(normalize=True))
data.gender.value_counts(normalize=True).plot(kind='bar')
```

```
percentage M    0.635417
F    0.364583
Name: gender, dtype: float64
```

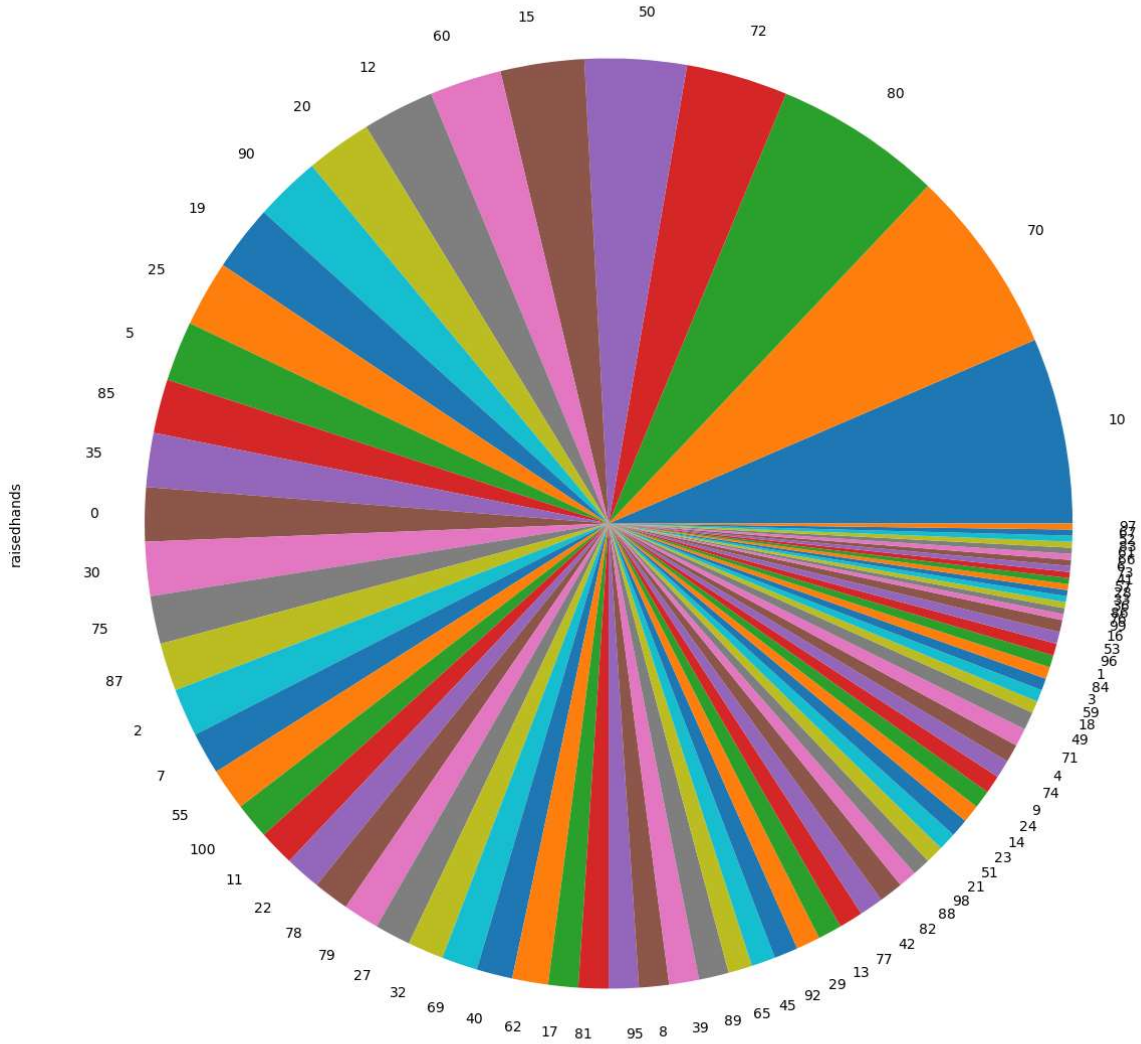
Out[32]: <Axes: >



In [33]:
```
print("percentage",data.raisedhands.value_counts(normalize=True))
data.raisedhands.value_counts(normalize=True).plot(kind='pie',figsize=(15,15))
```

```
percentage 10    0.064583
70    0.064583
80    0.058333
72    0.035417
50    0.035417
        ...
61    0.002083
83    0.002083
52    0.002083
67    0.002083
97    0.002083
Name: raisedhands, Length: 82, dtype: float64
```

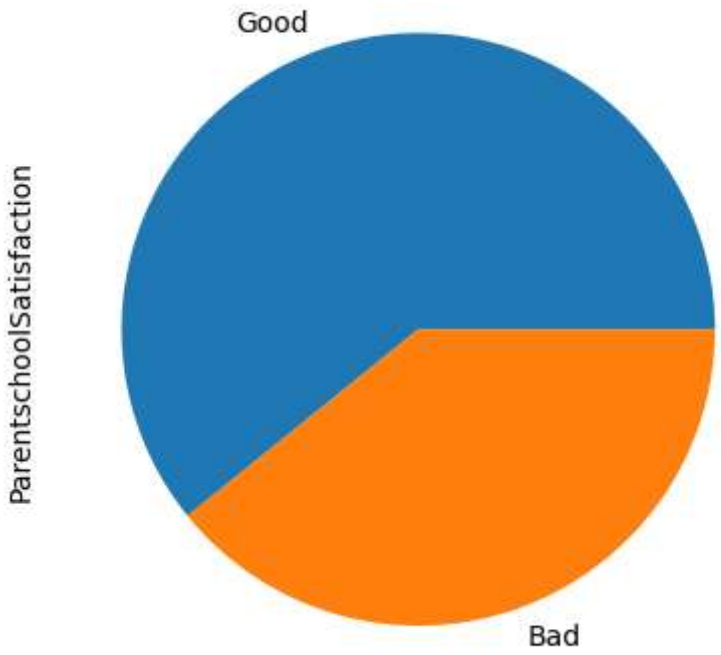
Out[33]: <Axes: ylabel='raisedhands'>



In [34]:
```
print("percentage",data.ParentschoolSatisfaction.value_counts(normalize=True))
data.ParentschoolSatisfaction.value_counts(normalize=True).plot(kind='pie')
```

```
percentage Good    0.608333
Bad     0.391667
Name: ParentschoolSatisfaction, dtype: float64
```

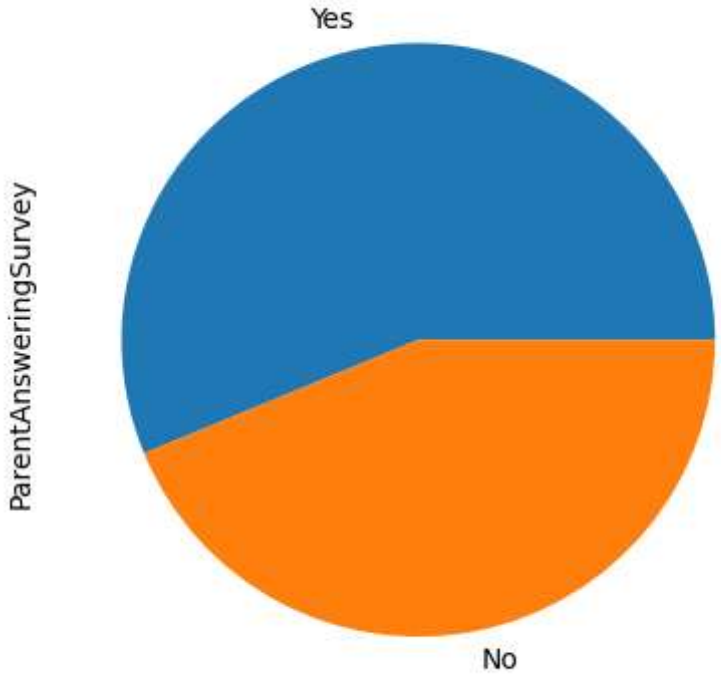
Out[34]:
```
<Axes: ylabel='ParentschoolSatisfaction'>
```



In [35]:
```
print("percentage",data.ParentAnsweringSurvey.value_counts(normalize=True))
data.ParentAnsweringSurvey.value_counts(normalize=True).plot(kind='pie')
```

```
percentage Yes    0.5625
No     0.4375
Name: ParentAnsweringSurvey, dtype: float64
```

Out[35]:
```
<Axes: ylabel='ParentAnsweringSurvey'>
```



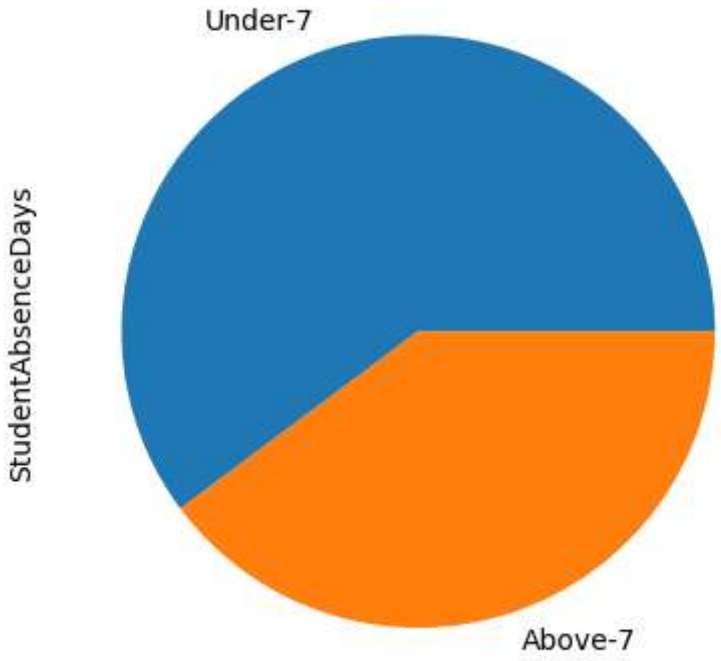

In [36]:
```
data["StudentAbsenceDays"].value_counts() #289,191 is the number of students
```

Out[36]:
```
Under-7    289
Above-7    191
Name: StudentAbsenceDays, dtype: int64
```

In [37]:
```
print("percentage",data.StudentAbsenceDays.value_counts(normalize=True)) #dnor
data.StudentAbsenceDays.value_counts(normalize=True).plot(kind='pie')
```

```
percentage Under-7    0.602083
Above-7    0.397917
Name: StudentAbsenceDays, dtype: float64
```

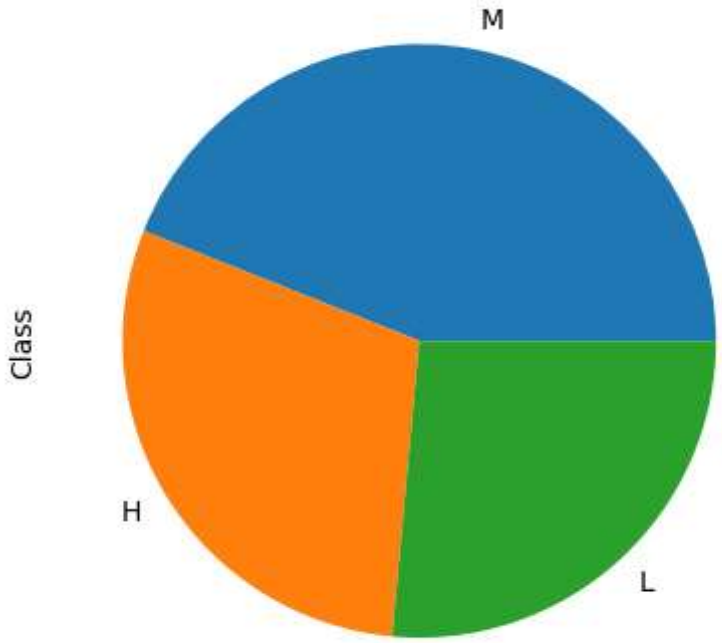
Out[37]:
```
<Axes: ylabel='StudentAbsenceDays'>
```



In [38]:
```
print("percentage",data.Class.value_counts(normalize=True)) #distibute number
data.Class.value_counts(normalize=True).plot(kind='pie')
```

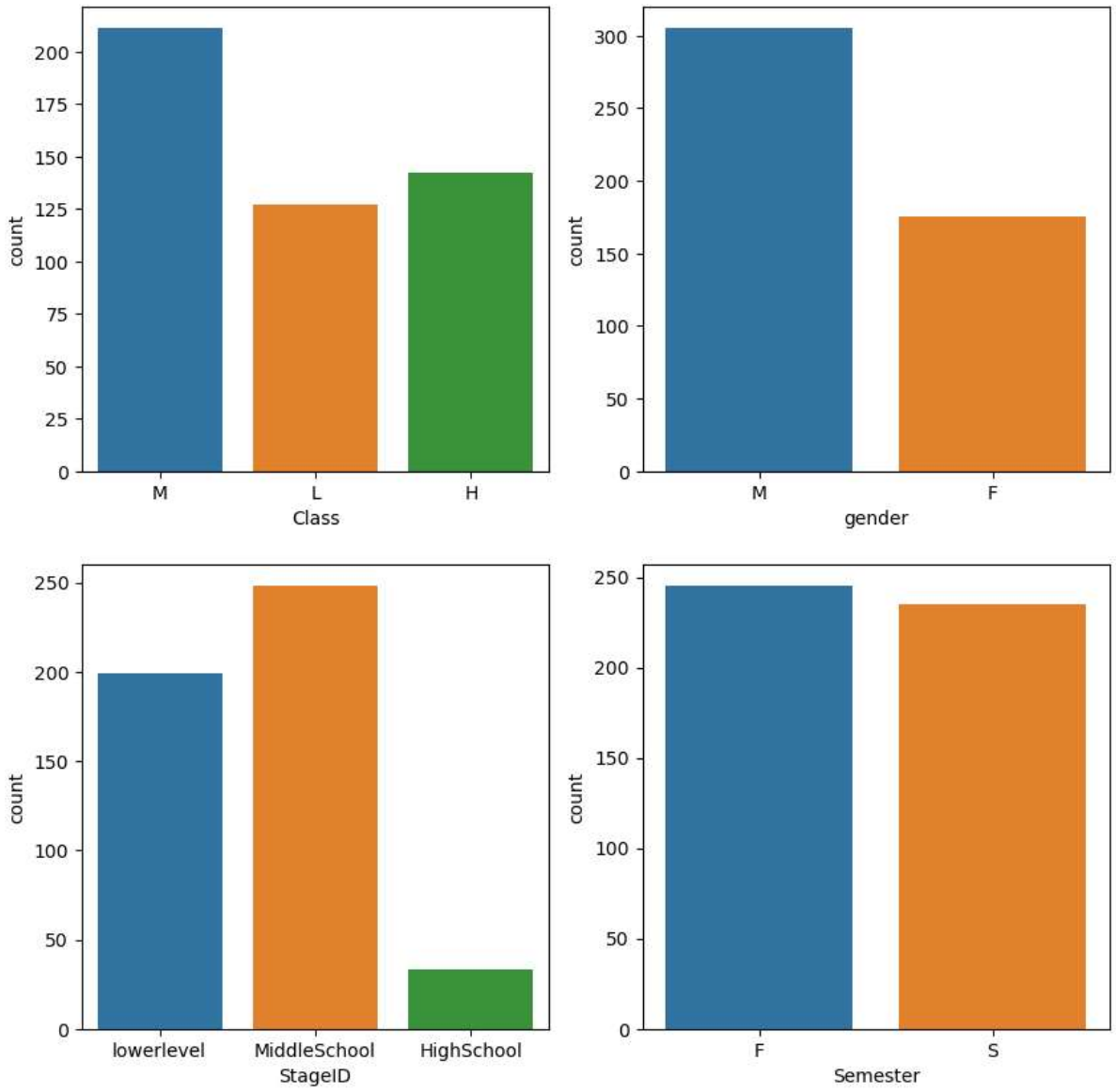
```
percentage M    0.439583
H    0.295833
L    0.264583
Name: Class, dtype: float64
```

Out[38]: <Axes: ylabel='Class'>



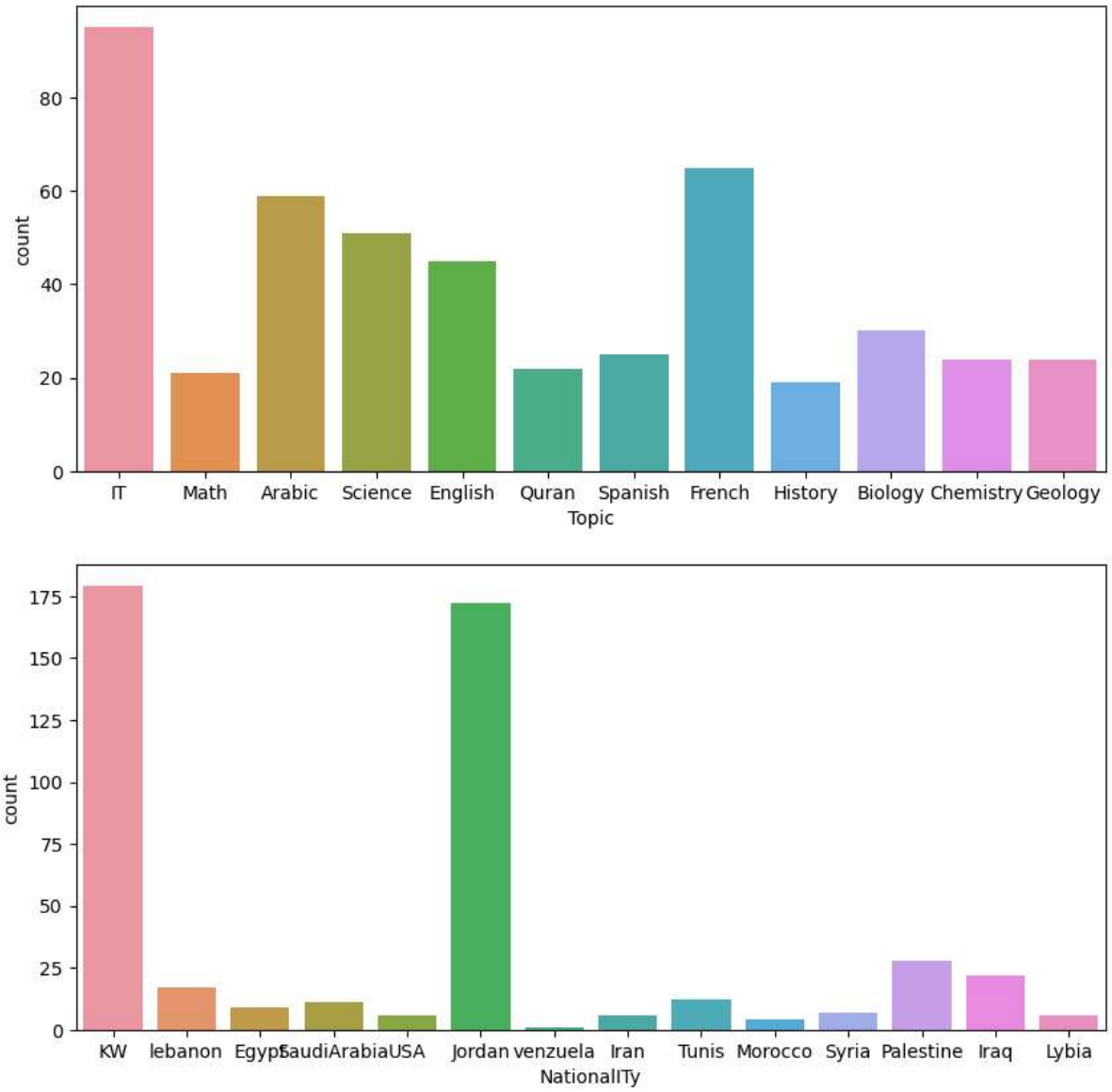
In [4]:
```
#!install seaborn
import seaborn as sns
import matplotlib.pyplot as plt
fig,axarr=plt.subplots(2,2,figsize=(10,10)) #ax, pick and label colmn,axarr is
sns.countplot(x="Class",data=data, ax=axarr[0,0])
sns.countplot(x="gender",data=data, ax=axarr[0,1])
sns.countplot(x="StageID",data=data, ax=axarr[1,0])
sns.countplot(x="Semester",data=data, ax=axarr[1,1])
```

Out[4]: <Axes: xlabel='Semester', ylabel='count'>



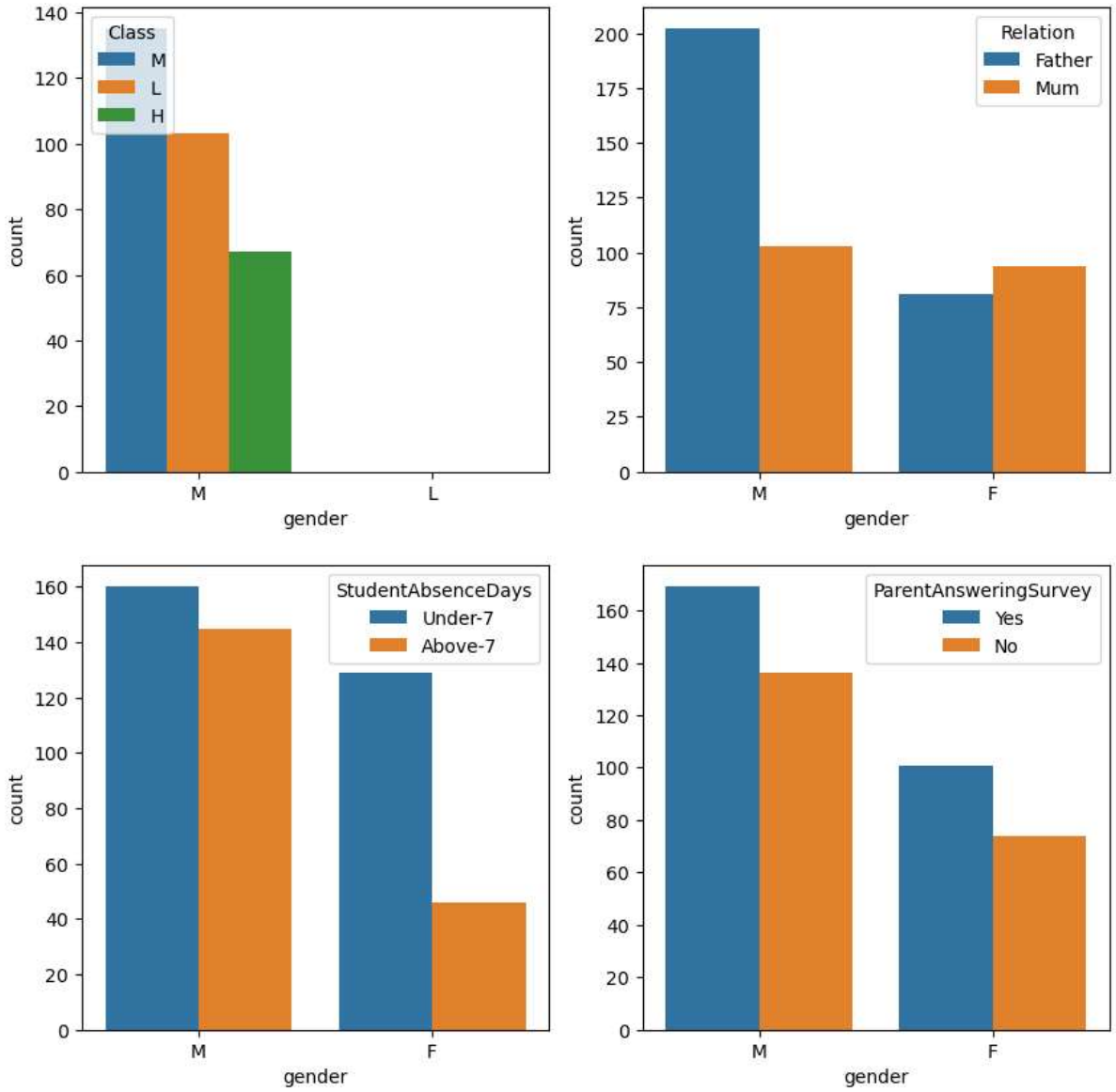
In [22]:
```
import seaborn as sns
import matplotlib.pyplot as plt
fig,axarr=plt.subplots(2,1,figsize=(10,10))
sns.countplot(x="Topic",data=data,ax=axarr[0])
sns.countplot(x="NationalITy",data=data,ax=axarr[1])
```

Out[22]: <Axes: xlabel='NationalITy', ylabel='count'>



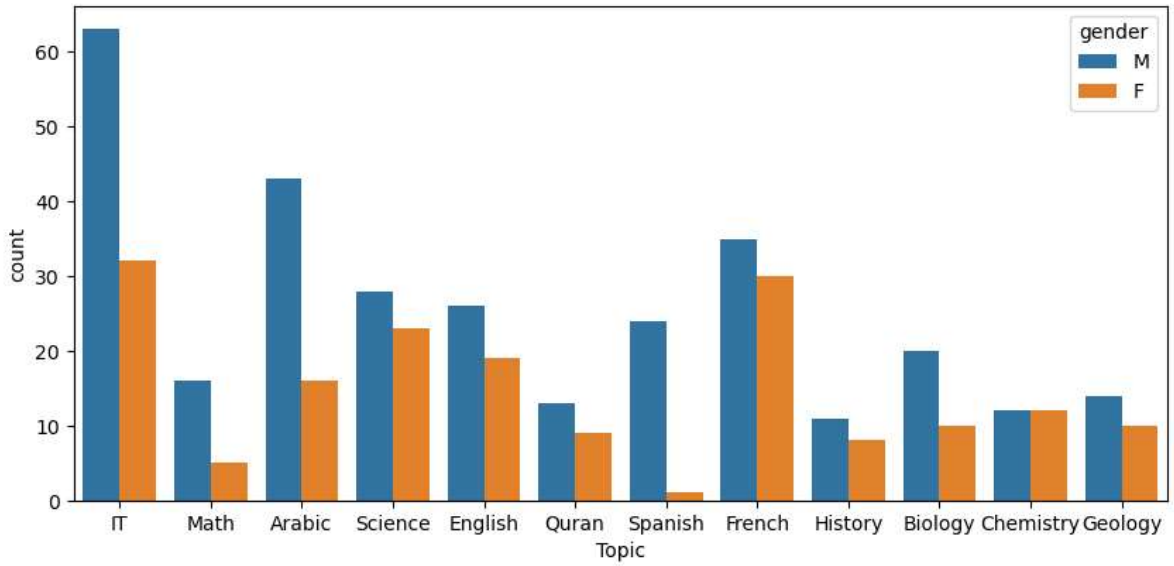
In [41]:
```
fig,axarr=plt.subplots(2,2,figsize=(10,10))
sns.countplot(x="gender",hue="Class",data=data,ax=axarr[0,0],order=["M","L"])
sns.countplot(x="gender",hue="Relation",data=data,ax=axarr[0,1])
sns.countplot(x="gender",hue="StudentAbsenceDays",data=data,ax=axarr[1,0])
sns.countplot(x="gender",hue="ParentAnsweringSurvey",data=data,ax=axarr[1,1])
```

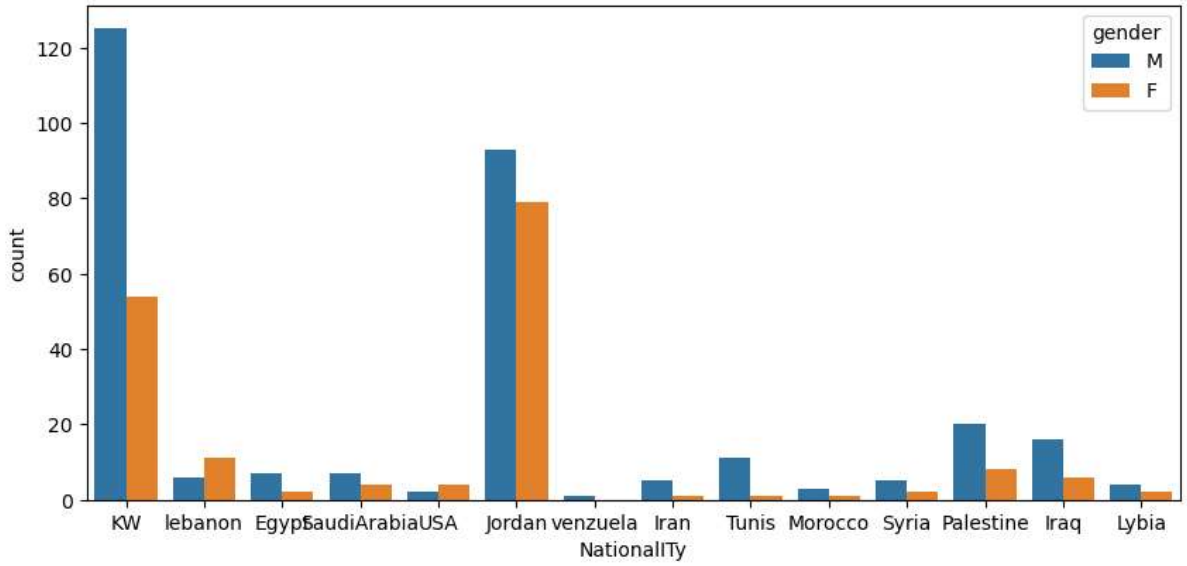
Out[41]: <Axes: xlabel='gender', ylabel='count'>



In [42]:
```
fig,axarr=plt.subplots(2,1,figsize=(10,10))
sns.countplot(x="Topic",hue="gender",data=data,ax=axarr[0])
sns.countplot(x="NationalITy",hue="gender",data=data,ax=axarr[1])
```

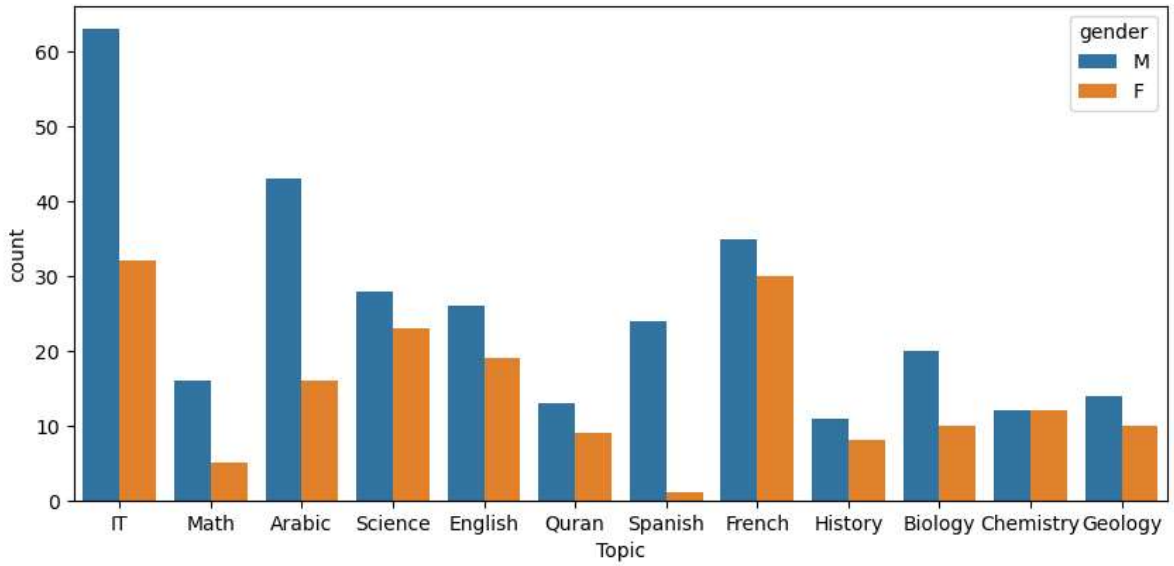
Out[42]: <Axes: xlabel='NationalITy', ylabel='count'>

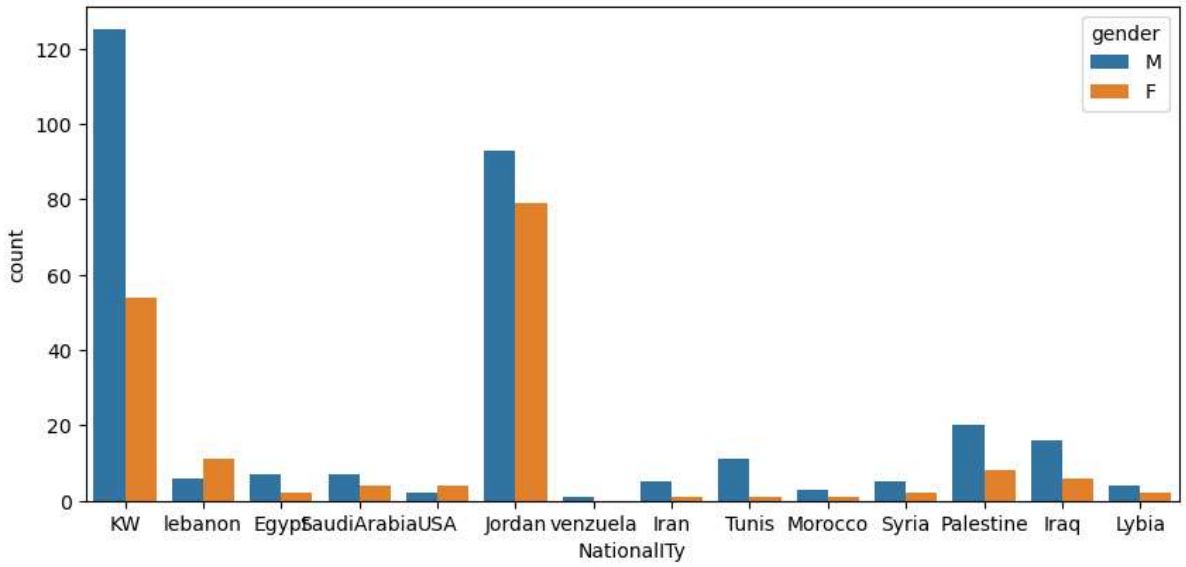


In [43]:
```
import seaborn as sns
import matplotlib.pyplot as plt
fig,axarr=plt.subplots(2,1,figsize=(10,10))
sns.countplot(x='Topic', hue="gender",data=data,ax=axarr[0])
sns.countplot(x='NationalITy', hue="gender",data=data,ax=axarr[1])
```

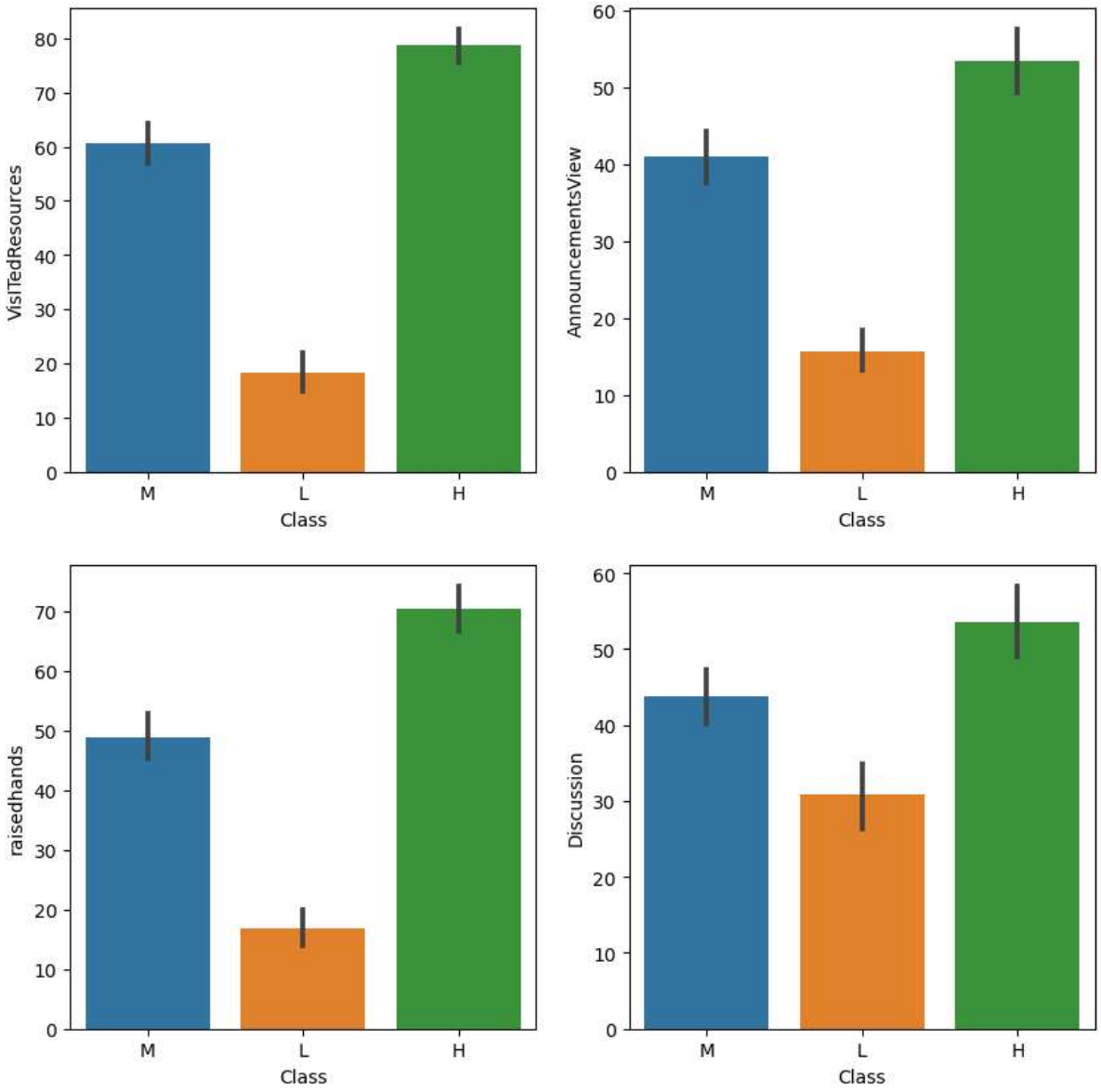
Out[43]: <Axes: xlabel='NationalITy', ylabel='count'>



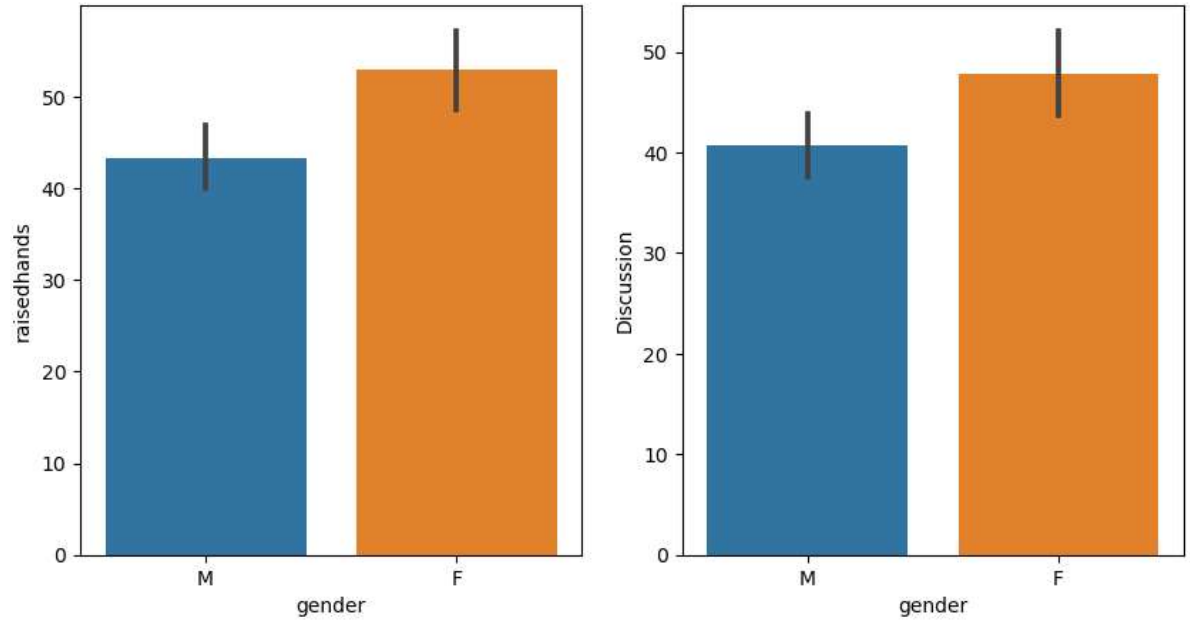
In [44]:
```
fig,axarr=plt.subplots(2,2,figsize=(10,10))
sns.barplot(x="Class",y="VisITedResources",data=data, ax=axarr[0,0])
sns.barplot(x="Class",y="AnnouncementsView",data=data,  ax=axarr[0,1])
sns.barplot(x="Class",y="raisedhands",data=data, ax=axarr[1,0])
sns.barplot(x="Class",y="Discussion",data=data, ax=axarr[1,1])
```

Out[44]: <Axes: xlabel='Class', ylabel='Discussion'>



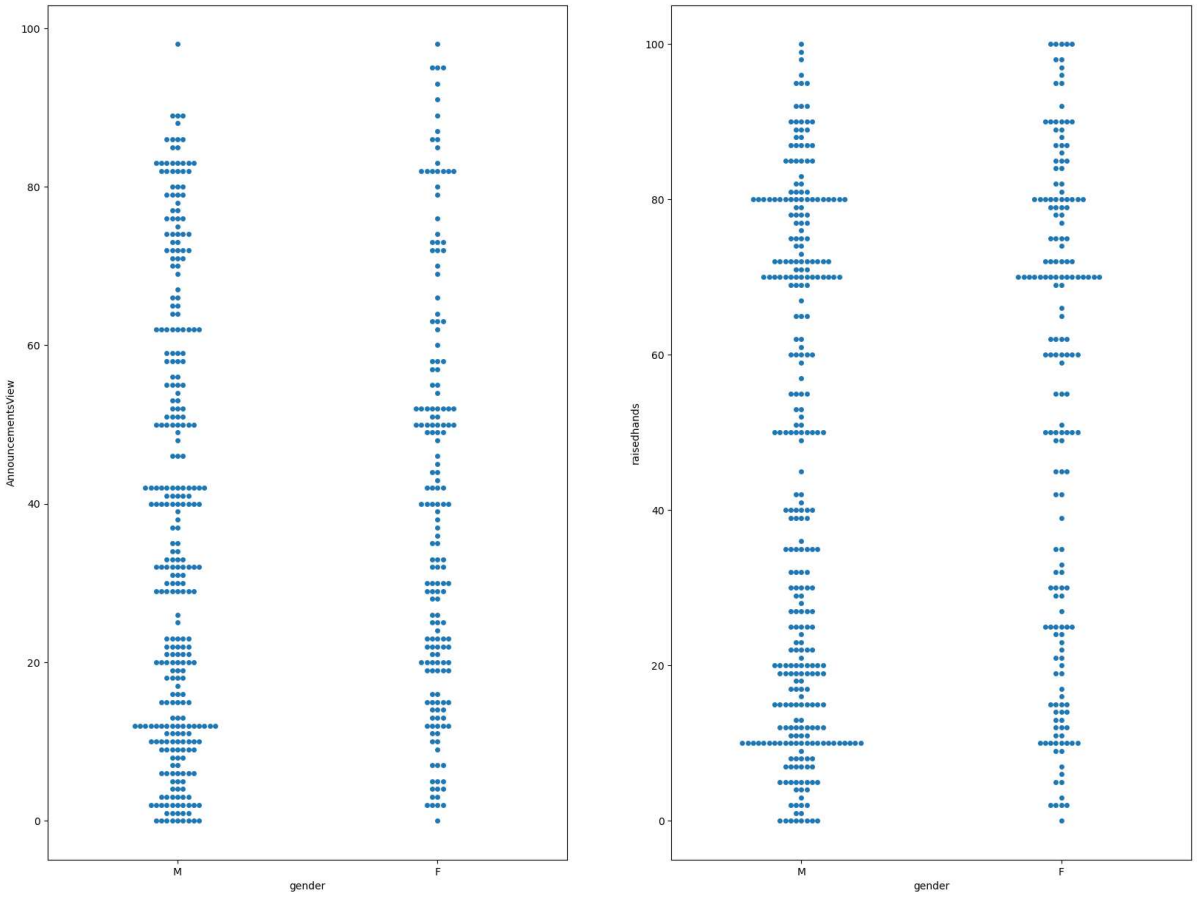
In [45]:
```
fig,(axis1,axis2) =plt.subplots(1,2,figsize=(10,5))
sns.barplot(x='gender',y='raisedhands', data=data,ax=axis1)
sns.barplot(x='gender',y='Discussion', data=data,ax=axis2)
```

Out[45]: <Axes: xlabel='gender', ylabel='Discussion'>



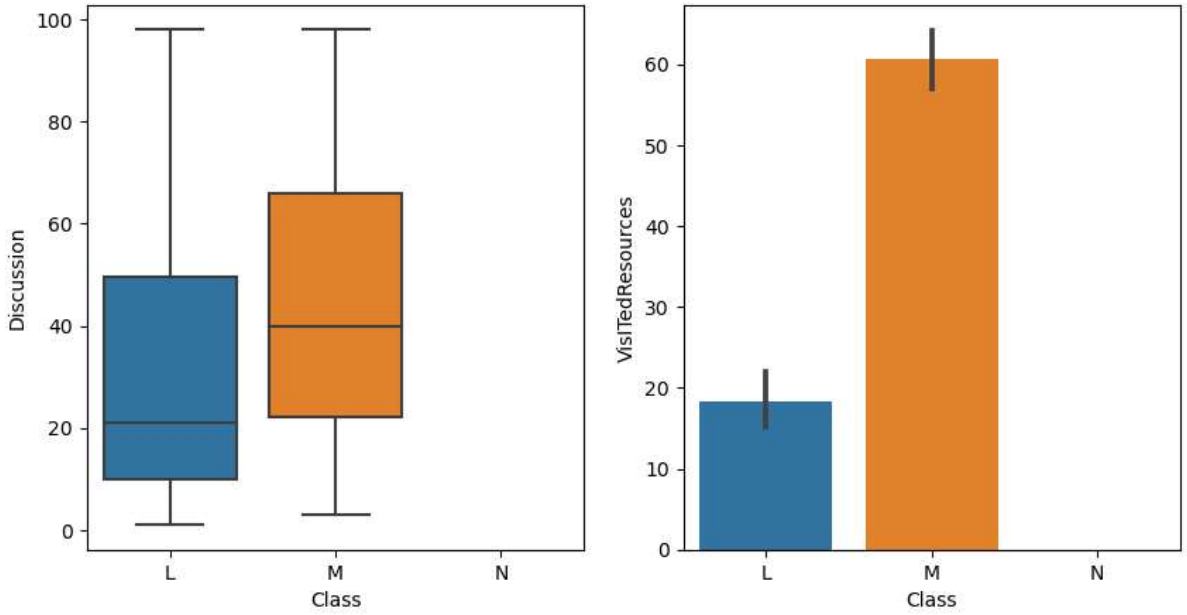
In [5]:
```
import matplotlib.pyplot as plt
import seaborn as sns
fig,(axis1,axis2) =plt.subplots(1,2,figsize=(20,15))
sns.swarmplot(x='gender',y='AnnouncementsView', data=data,ax=axis1)
sns.swarmplot(x='gender',y='raisedhands', data=data,ax=axis2)
```

Out[5]:
```
<Axes: xlabel='gender', ylabel='raisedhands'>
```



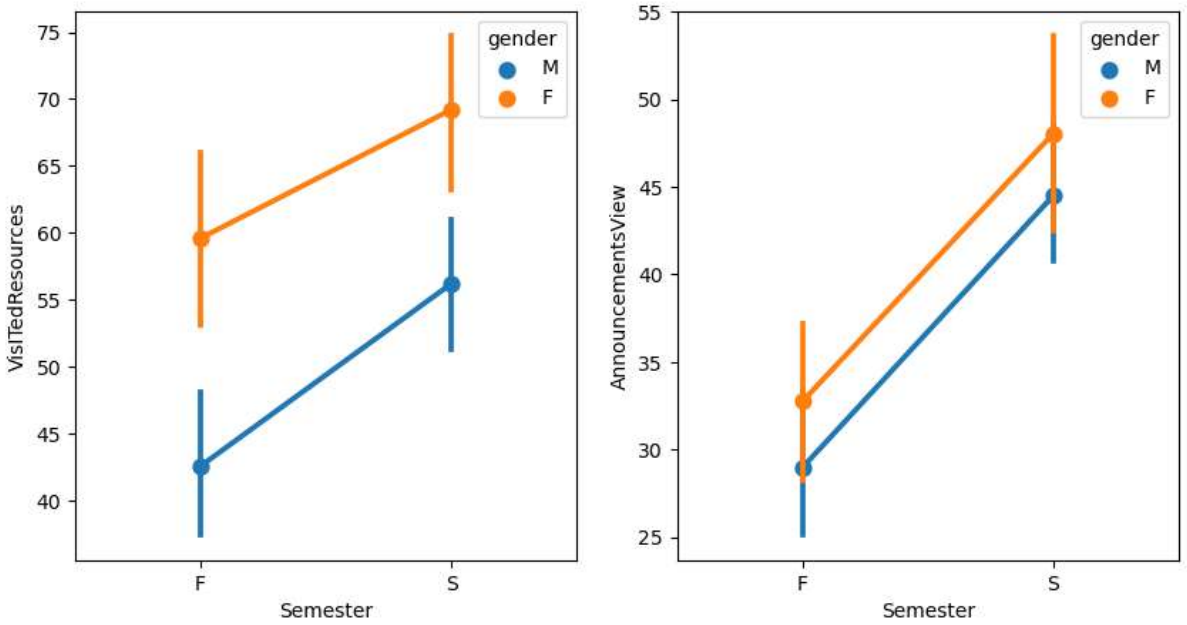
In [47]:
```
fig,(axis1,axis2) =plt.subplots(1,2,figsize=(10,5))
sns.boxplot(x='Class',y='Discussion', data=data, order=['L','M','N'],ax=axis1)
sns.barplot(x='Class',y='VisITedResources', data=data,order=['L','M','N'],ax=ax
```

Out[47]: <Axes: xlabel='Class', ylabel='VisITedResources'>



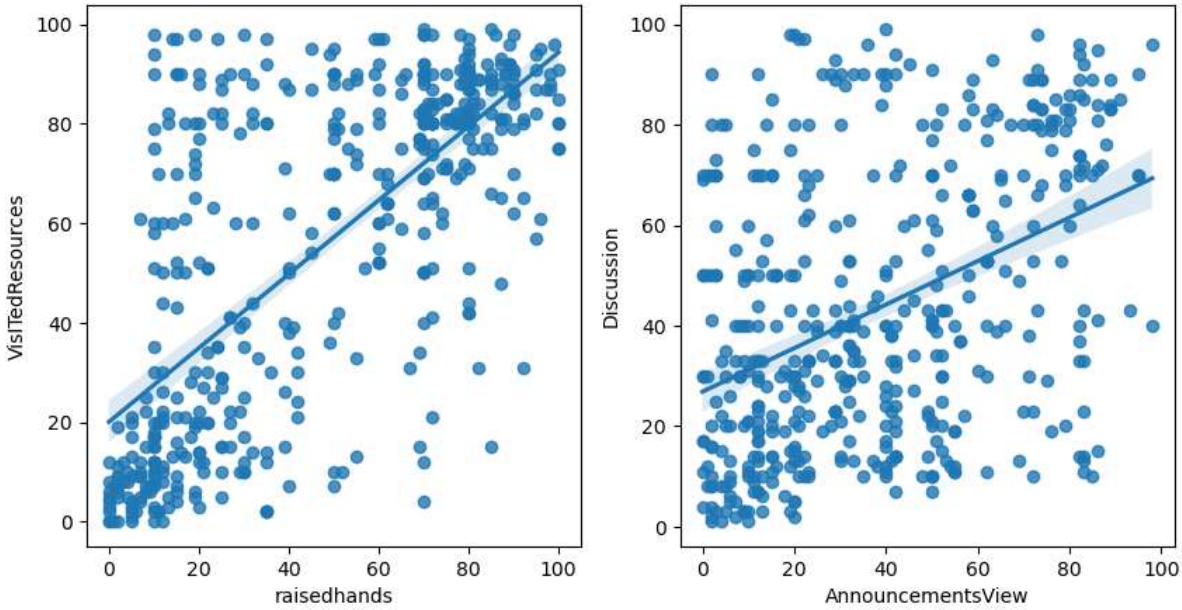

In [48]:
```
fig,(axis1,axis2) =plt.subplots(1,2,figsize=(10,5))
sns.pointplot(x='Semester',y='VisITedResources',hue='gender', data=data, ax=axi
sns.pointplot(x='Semester',y='AnnouncementsView',hue='gender', data=data,ax=axi
```

Out[48]: <Axes: xlabel='Semester', ylabel='AnnouncementsView'>

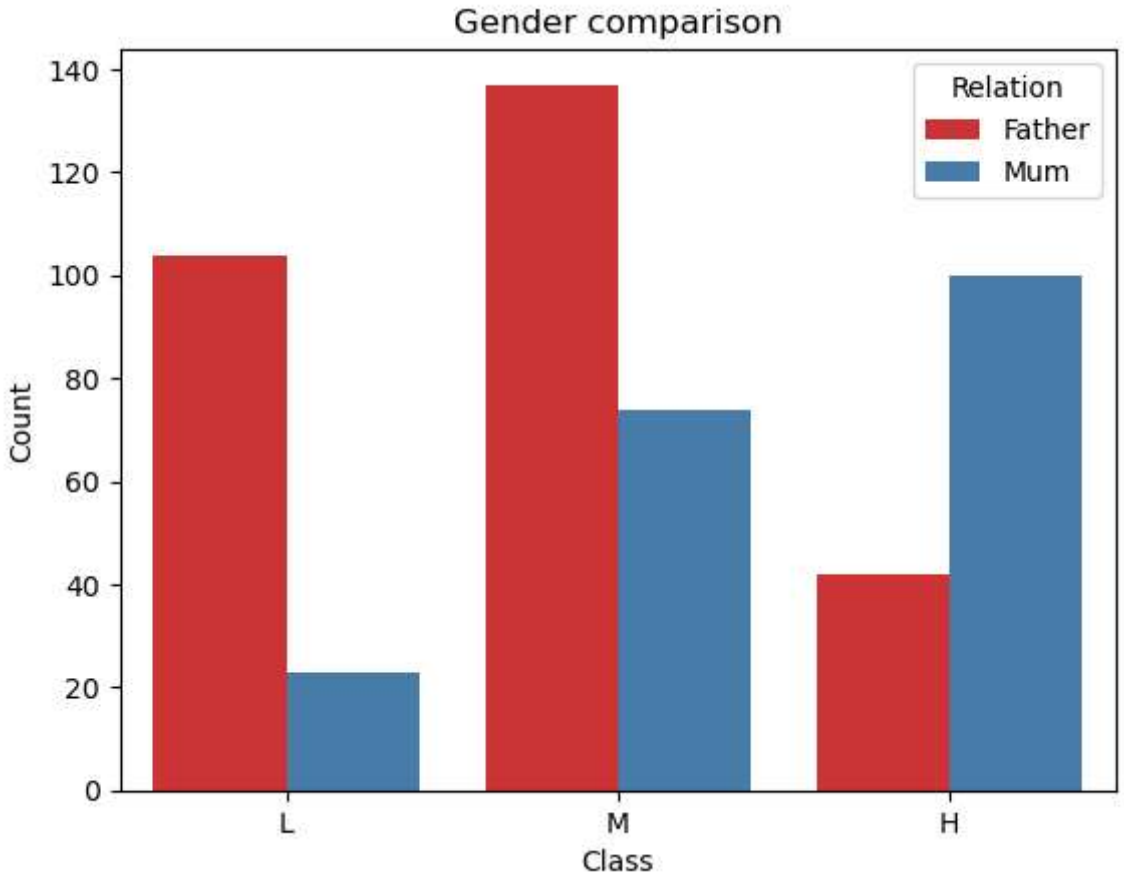


In [49]:
```
fig,(axis1,axis2) =plt.subplots(1,2,figsize=(10,5))
sns.regplot(x='raisedhands',y='VisITedResources', data=data, ax=axis1)
sns.regplot(x='AnnouncementsView',y='Discussion', data=data,ax=axis2)
```

Out[49]: <Axes: xlabel='AnnouncementsView', ylabel='Discussion'>

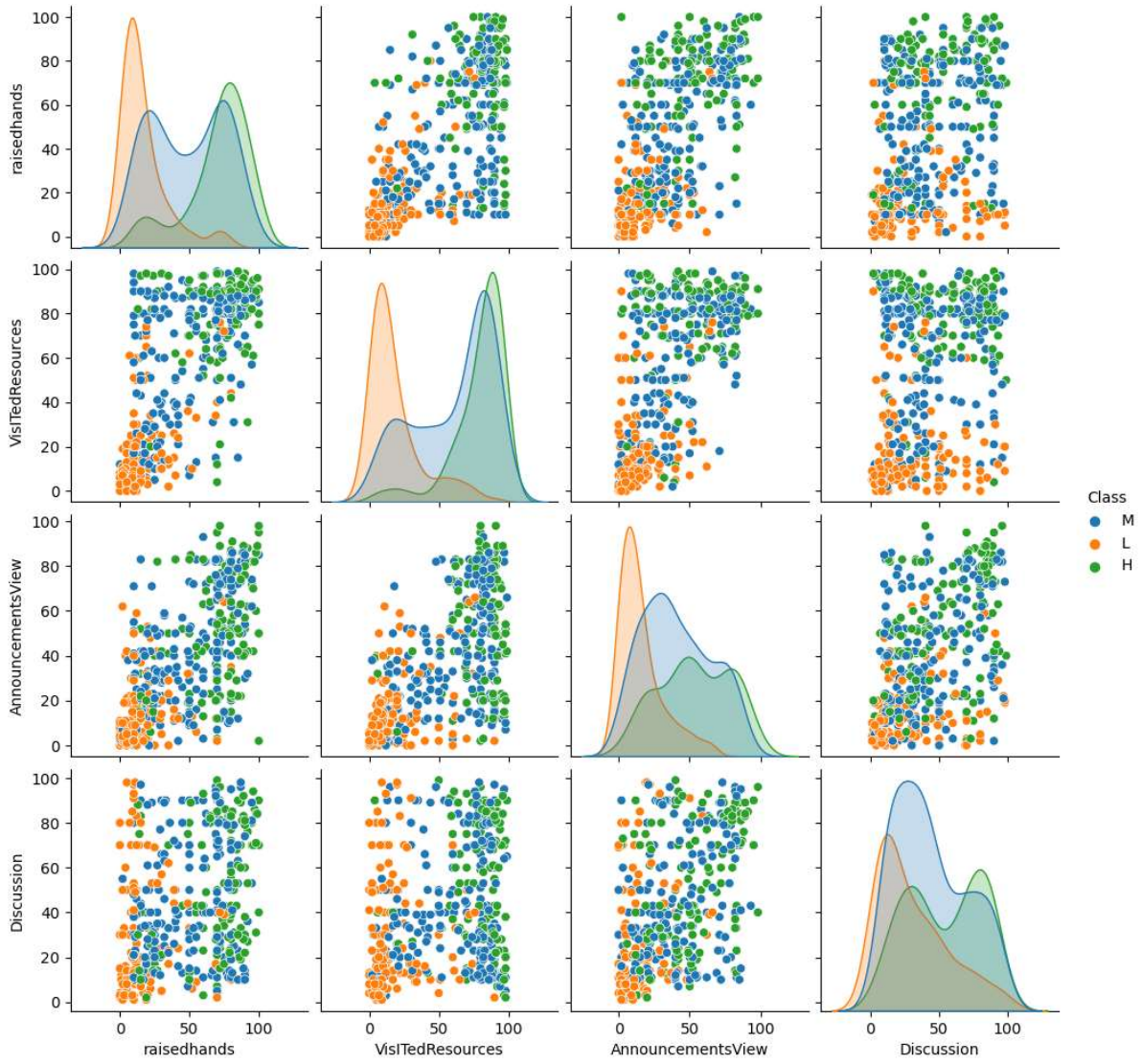



In [50]:
```
plot= sns.countplot(x='Class',hue='Relation',data=data,order=['L','M','H'],pale
plot.set(xlabel='Class',ylabel='Count',title='Gender comparison')
plt.show()
```



In [51]:
```
sns.pairplot(data,hue='Class')
```

Out[51]:
```
<seaborn.axisgrid.PairGrid at 0x1c26963f850>
```




# ENCODING

## Gender

In [6]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('gender',axis=1)
Target=data['gender']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  Semester
\
0              4             4        2        1          0      7         0
1              4             4        2        1          0      7         0
2              4             4        2        1          0      7         0
3              4             4        2        1          0      7         0
4              4             4        2        1          0      7         0
..           ...           ...      ...      ...        ...    ...       ...
475            3             3        1        5          0      2         1
476            3             3        1        5          0      5         0
477            3             3        1        5          0      5         1
478            3             3        1        5          0      6         0
479            3             3        1        5          0      6         1

     Relation  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         0            5                 4                  5           8
476         0           50                77                 14          28
477         0           55                74                 25          29
478         0           30                17                 14          57
479         0           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

In [ ]:

# Semester Encoding

In [7]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('Semester',axis=1)
Target=data['Semester']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Relation  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         0            5                 4                  5           8
476         0           50                77                 14          28
477         0           55                74                 25          29
478         0           30                17                 14          57
479         0           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# PARENTANSWERINGSURVEY

In [54]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('ParentAnsweringSurvey',axis=1)
Target=data['ParentAnsweringSurvey']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  Relation  raisedhands  VisITedResources  AnnouncementsView  \
0           0         0           15                16                  2
1           0         0           20                20                  3
2           0         0           10                 7                  0
3           0         0           30                25                  5
4           0         0           40                50                 12
..        ...       ...          ...               ...                ...
475         1         0            5                 4                  5
476         0         0           50                77                 14
477         1         0           55                74                 25
478         0         0           30                17                 14
479         1         0           35                14                 23

     Discussion  ParentschoolSatisfaction  StudentAbsenceDays  Class
0            20                         1                   1      2
1            25                         1                   1      2
2            30                         0                   0      1
3            35                         0                   0      1
4            50                         0                   0      2
..          ...                       ...                 ...    ...
475           8                         0                   0      1
476          28                         0                   1      2
477          29                         0                   1      2
478          57                         0                   0      1
479          62                         0                   0      1

[480 rows x 16 columns]
```

# Topic

In [55]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('Topic',axis=1)
Target=data['Topic']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Semester
\
0         1            4             4        2        1          0         0
1         1            4             4        2        1          0         0
2         1            4             4        2        1          0         0
3         1            4             4        2        1          0         0
4         1            4             4        2        1          0         0
..      ...          ...           ...      ...      ...        ...       ...
475       0            3             3        1        5          0         1
476       0            3             3        1        5          0         0
477       0            3             3        1        5          0         1
478       0            3             3        1        5          0         0
479       0            3             3        1        5          0         1

     Relation  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         0            5                 4                  5           8
476         0           50                77                 14          28
477         0           55                74                 25          29
478         0           30                17                 14          57
479         0           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# SectionID'

In [56]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('SectionID',axis=1)
Target=data['SectionID']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  Topic  Semester  \
0         1            4             4        2        1      7         0
1         1            4             4        2        1      7         0
2         1            4             4        2        1      7         0
3         1            4             4        2        1      7         0
4         1            4             4        2        1      7         0
..      ...          ...           ...      ...      ...    ...       ...
475       0            3             3        1        5      2         1
476       0            3             3        1        5      5         0
477       0            3             3        1        5      5         1
478       0            3             3        1        5      6         0
479       0            3             3        1        5      6         1

     Relation  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         0            5                 4                  5           8
476         0           50                77                 14          28
477         0           55                74                 25          29
478         0           30                17                 14          57
479         0           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# StageID'

In [57]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('StageID',axis=1)
Target=data['StageID']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  GradeID  SectionID  Topic  Semester
\
0         1            4             4        1          0      7         0
1         1            4             4        1          0      7         0
2         1            4             4        1          0      7         0
3         1            4             4        1          0      7         0
4         1            4             4        1          0      7         0
..      ...          ...           ...      ...        ...    ...       ...
475       0            3             3        5          0      2         1
476       0            3             3        5          0      5         0
477       0            3             3        5          0      5         1
478       0            3             3        5          0      6         0
479       0            3             3        5          0      6         1

     Relation  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         0            5                 4                  5           8
476         0           50                77                 14          28
477         0           55                74                 25          29
478         0           30                17                 14          57
479         0           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# GradeID

In [58]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('GradeID',axis=1)
Target=data['GradeID']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  SectionID  Topic  Semester
\
0         1            4             4        2          0      7         0
1         1            4             4        2          0      7         0
2         1            4             4        2          0      7         0
3         1            4             4        2          0      7         0
4         1            4             4        2          0      7         0
..      ...          ...           ...      ...        ...    ...       ...
475       0            3             3        1          0      2         1
476       0            3             3        1          0      5         0
477       0            3             3        1          0      5         1
478       0            3             3        1          0      6         0
479       0            3             3        1          0      6         1

     Relation  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         0            5                 4                  5           8
476         0           50                77                 14          28
477         0           55                74                 25          29
478         0           30                17                 14          57
479         0           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# NationalIT

In [59]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('NationalITy',axis=1)
Target=data['NationalITy']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  PlaceofBirth  StageID  GradeID  SectionID  Topic  Semester  \
0         1             4        2        1          0      7         0
1         1             4        2        1          0      7         0
2         1             4        2        1          0      7         0
3         1             4        2        1          0      7         0
4         1             4        2        1          0      7         0
..      ...           ...      ...      ...        ...    ...       ...
475       0             3        1        5          0      2         1
476       0             3        1        5          0      5         0
477       0             3        1        5          0      5         1
478       0             3        1        5          0      6         0
479       0             3        1        5          0      6         1

     Relation  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         0            5                 4                  5           8
476         0           50                77                 14          28
477         0           55                74                 25          29
478         0           30                17                 14          57
479         0           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# placeofbirth

In [9]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('PlaceofBirth',axis=1)
Target=data['PlaceofBirth']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  StageID  GradeID  SectionID  Topic  Semester  \
0         1            4        2        1          0      7         0
1         1            4        2        1          0      7         0
2         1            4        2        1          0      7         0
3         1            4        2        1          0      7         0
4         1            4        2        1          0      7         0
..      ...          ...      ...      ...        ...    ...       ...
475       0            3        1        5          0      2         1
476       0            3        1        5          0      5         0
477       0            3        1        5          0      5         1
478       0            3        1        5          0      6         0
479       0            3        1        5          0      6         1

     Relation  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         0            5                 4                  5           8
476         0           50                77                 14          28
477         0           55                74                 25          29
478         0           30                17                 14          57
479         0           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# Relation

In [10]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('Relation',axis=1)
Target=data['Relation']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  raisedhands  VisITedResources  AnnouncementsView  Discussion
\
0           0           15                16                  2          20
1           0           20                20                  3          25
2           0           10                 7                  0          30
3           0           30                25                  5          35
4           0           40                50                 12          50
..        ...          ...               ...                ...         ...
475         1            5                 4                  5           8
476         0           50                77                 14          28
477         1           55                74                 25          29
478         0           30                17                 14          57
479         1           35                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# Raisedhands

In [11]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('raisedhands',axis=1)
Target=data['raisedhands']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  Relation  VisITedResources  AnnouncementsView  Discussion  \
0           0         0                16                  2          20
1           0         0                20                  3          25
2           0         0                 7                  0          30
3           0         0                25                  5          35
4           0         0                50                 12          50
..        ...       ...               ...                ...         ...
475         1         0                 4                  5           8
476         0         0                77                 14          28
477         1         0                74                 25          29
478         0         0                17                 14          57
479         1         0                14                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# VisITedResources

In [8]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('VisITedResources',axis=1)
Target=data['VisITedResources']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  Relation  raisedhands  AnnouncementsView  Discussion  \
0           0         0           15                  2          20
1           0         0           20                  3          25
2           0         0           10                  0          30
3           0         0           30                  5          35
4           0         0           40                 12          50
..        ...       ...          ...                ...         ...
475         1         0            5                  5           8
476         0         0           50                 14          28
477         1         0           55                 25          29
478         0         0           30                 14          57
479         1         0           35                 23          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# Announcement

In [62]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('AnnouncementsView',axis=1)
Target=data['AnnouncementsView']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  Relation  raisedhands  VisITedResources  Discussion  \
0           0         0           15                16          20
1           0         0           20                20          25
2           0         0           10                 7          30
3           0         0           30                25          35
4           0         0           40                50          50
..        ...       ...          ...               ...         ...
475         1         0            5                 4           8
476         0         0           50                77          28
477         1         0           55                74          29
478         0         0           30                17          57
479         1         0           35                14          62

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# Discussion

In [9]:
```
 from sklearn.preprocessing import LabelEncoder
Features=data.drop('Discussion',axis=1)
Target=data['Discussion']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  Relation  raisedhands  VisITedResources  AnnouncementsView  \
0           0         0           15                16                  2
1           0         0           20                20                  3
2           0         0           10                 7                  0
3           0         0           30                25                  5
4           0         0           40                50                 12
..        ...       ...          ...               ...                ...
475         1         0            5                 4                  5
476         0         0           50                77                 14
477         1         0           55                74                 25
478         0         0           30                17                 14
479         1         0           35                14                 23

     ParentAnsweringSurvey  ParentschoolSatisfaction  StudentAbsenceDays  \
0                        1                         1                   1
1                        1                         1                   1
2                        0                         0                   0
3                        0                         0                   0
4                        0                         0                   0
..                     ...                       ...                 ...
475                      0                         0                   0
476                      0                         0                   1
477                      0                         0                   1
478                      0                         0                   0
479                      0                         0                   0

     Class
0        2
1        2
2        1
3        1
4        2
..     ...
475      1
476      2
477      2
478      1
479      1

[480 rows x 16 columns]
```

# ParentschoolSatisfaction

In [10]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('ParentschoolSatisfaction',axis=1)
Target=data['ParentschoolSatisfaction']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  Relation  raisedhands  VisITedResources  AnnouncementsView  \
0           0         0           15                16                  2
1           0         0           20                20                  3
2           0         0           10                 7                  0
3           0         0           30                25                  5
4           0         0           40                50                 12
..        ...       ...          ...               ...                ...
475         1         0            5                 4                  5
476         0         0           50                77                 14
477         1         0           55                74                 25
478         0         0           30                17                 14
479         1         0           35                14                 23

     Discussion  ParentAnsweringSurvey  StudentAbsenceDays  Class
0            20                      1                   1      2
1            25                      1                   1      2
2            30                      0                   0      1
3            35                      0                   0      1
4            50                      0                   0      2
..          ...                    ...                 ...    ...
475           8                      0                   0      1
476          28                      0                   1      2
477          29                      0                   1      2
478          57                      0                   0      1
479          62                      0                   0      1

[480 rows x 16 columns]
```

# #StudentAbsenceDays

In [11]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('StudentAbsenceDays',axis=1)
Target=data['StudentAbsenceDays']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  Relation  raisedhands  VisITedResources  AnnouncementsView  \
0           0         0           15                16                  2
1           0         0           20                20                  3
2           0         0           10                 7                  0
3           0         0           30                25                  5
4           0         0           40                50                 12
..        ...       ...          ...               ...                ...
475         1         0            5                 4                  5
476         0         0           50                77                 14
477         1         0           55                74                 25
478         0         0           30                17                 14
479         1         0           35                14                 23

     Discussion  ParentAnsweringSurvey  ParentschoolSatisfaction  Class
0            20                      1                         1      2
1            25                      1                         1      2
2            30                      0                         0      1
3            35                      0                         0      1
4            50                      0                         0      2
..          ...                    ...                       ...    ...
475           8                      0                         0      1
476          28                      0                         0      2
477          29                      0                         0      2
478          57                      0                         0      1
479          62                      0                         0      1

[480 rows x 16 columns]
```

# Class

In [12]:
```
from sklearn.preprocessing import LabelEncoder
Features=data.drop('Class',axis=1)
Target=data['Class']
label=LabelEncoder()
Cat_Colums=Features.dtypes.pipe(lambda Features:Features[Features=='object']).i
for col in Cat_Colums:
    Features[col]=label.fit_transform(Features[col])
print(Features)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic  \
0         1            4             4        2        1          0      7
1         1            4             4        2        1          0      7
2         1            4             4        2        1          0      7
3         1            4             4        2        1          0      7
4         1            4             4        2        1          0      7
..      ...          ...           ...      ...      ...        ...    ...
475       0            3             3        1        5          0      2
476       0            3             3        1        5          0      5
477       0            3             3        1        5          0      5
478       0            3             3        1        5          0      6
479       0            3             3        1        5          0      6

     Semester  Relation  raisedhands  VisITedResources  AnnouncementsView  \
0           0         0           15                16                  2
1           0         0           20                20                  3
2           0         0           10                 7                  0
3           0         0           30                25                  5
4           0         0           40                50                 12
..        ...       ...          ...               ...                ...
475         1         0            5                 4                  5
476         0         0           50                77                 14
477         1         0           55                74                 25
478         0         0           30                17                 14
479         1         0           35                14                 23

     Discussion  ParentAnsweringSurvey  ParentschoolSatisfaction  \
0            20                      1                         1
1            25                      1                         1
2            30                      0                         0
3            35                      0                         0
4            50                      0                         0
..          ...                    ...                       ...
475           8                      0                         0
476          28                      0                         0
477          29                      0                         0
478          57                      0                         0
479          62                      0                         0

     StudentAbsenceDays
0                     1
1                     1
2                     0
3                     0
4                     0
..                  ...
475                   0
476                   1
477                   1
478                   0
479                   0

[480 rows x 16 columns]
```

# Test and Train data split

In [13]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(Features,Target ,test_size=0.2,
print(x_test)
print(y_train)
print(y_test)
print(x_train)
```

```
     gender  NationalITy  PlaceofBirth  StageID  GradeID  SectionID  Topic
\
103       1            4             4        2        0          1      7
431       1            3             8        1        5          0      5
475       0            3             3        1        5          0      2
231       1            9             3        1        5          0     11
400       1            3             3        1        4          0      1
..      ...          ...           ...      ...      ...        ...    ...
1         1            4             4        2        1          0      7
43        0            4             4        0        6          0      7
149       0            3            12        1        4          0      9
437       1            3             3        1        5          0      5
393       0            2             2        1        4          0      1

     Semester  Relation  raisedhands  VisITedResources  AnnouncementsView
\
103         0         0            1                 7                  6
431         1         0           80                89                 23
475         1         0            5                 4                  5
```

In [31]:
```
#Logis
```

In [14]:
```
from sklearn.linear_model import LogisticRegression
Logit_Model=LogisticRegression()
Logit_Model.fit(x_train,y_train)
```

```
C:\Users\sumaira\anaconda2\lib\site-packages\sklearn\linear_model\_logistic.p
y:458: ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html (https://sciki
t-learn.org/stable/modules/preprocessing.html)
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion (https://scikit-learn.org/stable/modules/linear_model.html#logistic-regr
ession)
  n_iter_i = _check_optimize_result(
```

Out[14]:
```
LogisticRegression()
```

**In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.**
**On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.**

# logisticRegression's prediction,score & Report

In [17]:
```
from sklearn.metrics import classification_report,accuracy_score
prediction=Logit_Model.predict(x_test)
Score=accuracy_score(y_test,prediction)
Report=classification_report(y_test,prediction)
```

In [18]:
```
print(prediction)
```

```
['L' 'M' 'L' 'L' 'M' 'L' 'M' 'H' 'M' 'M' 'M' 'L' 'H' 'L' 'M' 'L' 'H' 'M'
 'M' 'M' 'H' 'H' 'L' 'L' 'H' 'L' 'M' 'H' 'H' 'M' 'M' 'M' 'L' 'M' 'M' 'L'
 'L' 'L' 'H' 'L' 'M' 'M' 'M' 'M' 'H' 'L' 'M' 'L' 'M' 'L' 'H' 'H' 'M' 'M'
 'H' 'M' 'H' 'L' 'L' 'H' 'L' 'H' 'M' 'M' 'H' 'H' 'L' 'L' 'H' 'M' 'H' 'H'
 'H' 'H' 'H' 'H' 'M' 'H' 'M' 'H' 'L' 'H' 'M' 'M' 'L' 'M' 'M' 'M' 'M' 'M'
 'L' 'M' 'M' 'H' 'M' 'H']
```

In [19]:
```
print(Score)
```

```
0.75
```

In [20]:
```
print(Report)
```

```
              precision    recall  f1-score   support

           H       0.80      0.69      0.74        35
           L       0.77      0.91      0.83        22
           M       0.70      0.72      0.71        39

    accuracy                           0.75        96
   macro avg       0.76      0.77      0.76        96
weighted avg       0.75      0.75      0.75        96
```

In [ ]:

In [ ]:

In [ ]: