

Assignment2

March 25, 2022

```
[2]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```
[3]: data = pd.read_csv('AcademicPerformance.csv')
```

```
[4]: data.head()
```

```
[4]:  gender NationalITY PlaceofBirth      StageID GradeID SectionID Topic \
0      M           KW      KuwaIT  lowerlevel    G-04          A    IT
1      M           KW      KuwaIT  lowerlevel    G-04          A    IT
2      M           KW      KuwaIT  lowerlevel    G-04          A    IT
3      M           KW      KuwaIT  lowerlevel    G-04          A    IT
4      M           KW      KuwaIT  lowerlevel    G-04          A    IT
```

```
      Semester Relation  raisedhands  VisITedResources  AnnouncementsView \
0            F  Father           15             16             2
1            F  Father           20             20             3
2            F  Father           10              7             0
3            F  Father           30             25             5
4            F  Father           40             50            12
```

```
      Discussion ParentAnsweringSurvey ParentschoolSatisfaction \
0            20                Yes                Good
1            25                Yes                Good
2            30                No                 Bad
3            35                No                 Bad
4            50                No                 Bad
```

```
      StudentAbsenceDays Class
0            Under-7      M
1            Under-7      M
2            Above-7      L
3            Above-7      L
4            Above-7      M
```

```
[5]: data.describe()
```

```
[5]:      raisedhands  VisITedResources  AnnouncementsView  Discussion
count    480.000000          480.000000          480.000000  480.000000
mean     46.775000           54.797917           37.918750   43.283333
std      30.779223           33.080007           26.611244   27.637735
min       0.000000           0.000000           0.000000    1.000000
25%      15.750000           20.000000           14.000000   20.000000
50%      50.000000           65.000000           33.000000   39.000000
75%      75.000000           84.000000           58.000000   70.000000
max     100.000000          99.000000           98.000000   99.000000
```

```
[6]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 480 entries, 0 to 479
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype
---  -
0   gender                480 non-null   object
1   NationalITy           480 non-null   object
2   PlaceofBirth          480 non-null   object
3   StageID               480 non-null   object
4   GradeID               480 non-null   object
5   SectionID             480 non-null   object
6   Topic                 480 non-null   object
7   Semester              480 non-null   object
8   Relation              480 non-null   object
9   raisedhands           480 non-null   int64
10  VisITedResources       480 non-null   int64
11  AnnouncementsView      480 non-null   int64
12  Discussion             480 non-null   int64
13  ParentAnsweringSurvey  480 non-null   object
14  ParentschoolSatisfaction 480 non-null   object
15  StudentAbsenceDays     480 non-null   object
16  Class                 480 non-null   object
dtypes: int64(4), object(13)
memory usage: 63.9+ KB
```

```
[7]: data.dropna()
```

```
[7]:      gender NationalITy PlaceofBirth      StageID GradeID SectionID  \
0         M           KW      KuwaIT  lowerlevel    G-04         A
1         M           KW      KuwaIT  lowerlevel    G-04         A
2         M           KW      KuwaIT  lowerlevel    G-04         A
3         M           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	16	
1	IT	F	Father	20	20	
2	IT	F	Father	10	7	
3	IT	F	Father	30	25	
4	IT	F	Father	40	50	
..	
475	Chemistry	S	Father	5	4	
476	Geology	F	Father	50	77	
477	Geology	S	Father	55	74	
478	History	F	Father	30	17	
479	History	S	Father	35	14	

	AnnouncementsView	Discussion	ParentAnsweringSurvey	\
0	2	20	Yes	
1	3	25	Yes	
2	0	30	No	
3	5	35	No	
4	12	50	No	
..	
475	5	8	No	
476	14	28	No	
477	25	29	No	
478	14	57	No	
479	23	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]

```
[8]: data.shape
```

```
[8]: (480, 17)
```

```
[9]: data['ParentschoolSatisfaction'].replace(['Bad', 'Good'],[0,1], inplace=True)
```

```
[11]: data['ParentAnsweringSurvey'].replace(['No', 'Yes'],[0,1], inplace=True)
```

```
[12]: data.head()
```

```
[12]:  gender NationalITy PlaceofBirth      StageID GradeID SectionID Topic \
0      M      KW      KuwaIT  lowerlevel    G-04      A      IT
1      M      KW      KuwaIT  lowerlevel    G-04      A      IT
2      M      KW      KuwaIT  lowerlevel    G-04      A      IT
3      M      KW      KuwaIT  lowerlevel    G-04      A      IT
4      M      KW      KuwaIT  lowerlevel    G-04      A      IT

      Semester Relation  raisedhands  VisITedResources  AnnouncementsView \
0      F  Father      15      16      2
1      F  Father      20      20      3
2      F  Father      10      7      0
3      F  Father      30      25      5
4      F  Father      40      50      12

      Discussion  ParentAnsweringSurvey  ParentschoolSatisfaction \
0      20      1      1
1      25      1      1
2      30      0      0
3      35      0      0
4      50      0      0

      StudentAbsenceDays Class
0      Under-7      M
1      Under-7      M
2      Above-7      L
3      Above-7      L
4      Above-7      M
```

```
[ ]:
```