In [3]:
```
# assignment 15
# Calculate the probability a student gets an A (80%+) in math, given they miss 10 or more classes.
```

In [4]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
file= pd.read_csv("student-mat.csv")
file
```

Out[4]:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
390	MS	M	20	U	LE3	A	2	2	services	services	...	5	5	4	4	5
391	MS	M	17	U	LE3	T	3	1	services	services	...	2	4	5	3	4
392	MS	M	21	R	GT3	T	1	1	other	other	...	5	5	3	3	3
393	MS	M	18	R	LE3	T	3	2	services	other	...	4	4	1	3	4
394	MS	M	19	U	LE3	T	1	1	other	at_home	...	3	2	3	3	3

395 rows × 33 columns

In [5]:
```
file.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 395 entries, 0 to 394
Data columns (total 33 columns):
 #   Column      Non-Null Count  Dtype 
---  ------      --------------  ----- 
 0   school      395 non-null    object
 1   sex         395 non-null    object
 2   age         395 non-null    int64 
 3   address     395 non-null    object
 4   famsize     395 non-null    object
 5   Pstatus     395 non-null    object
 6   Medu        395 non-null    int64 
 7   Fedu        395 non-null    int64 
 8   Mjob        395 non-null    object
 9   Fjob        395 non-null    object
 10  reason      395 non-null    object
 11  guardian    395 non-null    object
 12  traveltime  395 non-null    int64 
 13  studytime   395 non-null    int64 
 14  failures    395 non-null    int64 
 15  schoolsup   395 non-null    object
 16  famsup      395 non-null    object
 17  paid        395 non-null    object
 18  activities  395 non-null    object
 19  nursery     395 non-null    object
 20  higher      395 non-null    object
 21  internet    395 non-null    object
 22  romantic    395 non-null    object
 23  famrel      395 non-null    int64 
 24  freetime    395 non-null    int64 
 25  goout       395 non-null    int64 
 26  Dalc        395 non-null    int64 
 27  Walc        395 non-null    int64 
 28  health      395 non-null    int64 
 29  absences    395 non-null    int64 
 30  G1          395 non-null    int64 
 31  G2          395 non-null    int64 
 32  G3          395 non-null    int64 
dtypes: int64(16), object(17)
memory usage: 102.0+ KB
```

In [6]:
```
file.head()
```

Out[6]:

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	he
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	

5 rows × 33 columns

In [11]:
```

```

Out[11]: 75

In [ ]:
```

```

In [17]:
```
file['A']=np.where(file['G3']*5>=80,1,0)
file['ab']=np.where(file['absences']>=10,1,0)
file['A']
new=pd.DataFrame()
file['ab']
```

Out[17]:
```
0      0
1      0
2      1
3      0
4      0
      ..
390    1
391    0
392    0
393    0
394    0
Name: ab, Length: 395, dtype: int32
```

In [21]:
```
new['A']=file['A']
new['ab']=file['ab']
new['count']=1
new
```

Out[21]:

	A	ab	count
0	0	0	1
1	0	0	1
2	0	1	1
3	0	0	1
4	0	0	1
...	...	...	...
390	0	1	1
391	1	0	1
392	0	0	1
393	0	0	1
394	0	0	1

395 rows × 3 columns

pd.pivot_table(new, values='count', index=['A'], columns=['ab'], aggfunc=np.sum, fill_value=0)

calculation ¶

In [32]:
```
A=(35+5)/(35+5+277+78)
A
```

Out[32]:
```
0.10126582278481013
```

In [33]:
```
B=(78+5)/(35+5+277+78)
B
```

Out[33]:
```
0.21012658227848102
```

In [36]:
```
Z=5/(35+5+277+78)
P=Z/B
P
```

Out[36]:
```
0.060240963855421686
```

In [37]:
```
P=Z/B
P
```

Out[37]:
```
0.060240963855421686
```

In [38]:
```
P%100
```

Out[38]:
```
0.060240963855421686
```

In [39]:
```
# THE PROBABLITY OF A STUDENT GETTING A 80% MORE IN MATHS , GIVEN THEY MISS 10  OR MORE CLASSES IS 6%
```

In [ ]:
```

```