In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [2]:
```
data  =pd.read_csv('F:\Kuliah/Semester 6/Data Mining/UAS/dataset1.csv',delimiter=";")
```

In [3]:
```
data.head()
```

Out[3]:

	Age	Income	Student	Credit_rating	Class (buy_computer)
0	<=30	High	No	Fair	No
1	<=30	High	No	Excellent	No
2	31..40	High	No	Fair	Yes
3	> 40	Medium	No	Fair	Yes
4	> 40	Low	Yes	Fair	Yes

In [4]:
```
data.tail(10)
```

Out[4]:

	Age	Income	Student	Credit_rating	Class (buy_computer)
41	> 40	Low	Yes	Fair	No
42	31..40	Low	Yes	Fair	Yes
43	31..40	Low	Yes	Excellent	No
44	<= 30	High	No	Excellent	No
45	<= 30	Medium	Yes	Excellent	Yes
46	> 40	Low	Yes	Fair	Yes
47	<= 30	Low	Yes	Fair	Yes
48	31..40	Medium	No	Fair	No
49	31..40	High	Yes	Excellent	Yes
50	> 40	Medium	No	Excellent	No

In [5]:
```
data['Age'].value_counts()
```

Out[5]:
```
> 40      17
<= 30     15
31..40    14
<=30       5
Name: Age, dtype: int64
```

In [6]:
```
data['Income'].value_counts()
```

Out[6]:
```
Low       21
Medium    19
High      11
Name: Income, dtype: int64
```

In [7]:
```
data['Student'].value_counts()
```

Out[7]:
```
Yes    27
No     24
Name: Student, dtype: int64
```

In [8]:
```
data['Credit_rating'].value_counts()
```

Out[8]:
```
Fair         31
Excellent    20
Name: Credit_rating, dtype: int64
```

In [9]:
```
data['Class (buy_computer)'].value_counts()
```

Out[9]:
```
Yes    29
No     22
Name: Class (buy_computer), dtype: int64
```

In [10]:
```
data.shape
```

Out[10]:
```
(51, 5)
```

In [11]:
```
PYes = 27/51
PNo = 24/51
```

In [12]:
```
pd.crosstab(data['Age'],data['Income'])
```

Out[12]:

Income	High	Low	Medium
Age			
31..40	5	5	4
<= 30	1	6	8
<=30	5	0	0
> 40	0	10	7

In [13]:
```
pd.crosstab(data['Age'],data['Student'])
```

Out[13]:

Student	No	Yes
Age		
31..40	7	7
<= 30	6	9
<=30	5	0
> 40	6	11

In [14]:
```
pd.crosstab(data['Age'],data['Credit_rating'])
```

Out[14]:

Credit_rating	Excellent	Fair
Age		
31..40	7	7
<= 30	4	11
<=30	3	2
> 40	6	11

In [15]:
```
pd.crosstab(data['Income'],data['Class (buy_computer)'])
```

Out[15]:

Class (buy_computer)	No	Yes
Income		
High	6	5
Low	11	10
Medium	5	14

In [16]:
```
pd.crosstab(data['Income'],data['Credit_rating'])
```

Out[16]:

Credit_rating	Excellent	Fair
Income		
High	5	6
Low	8	13
Medium	7	12

In [17]:
```
pd.crosstab(data['Income'],data['Age'])
```

Out[17]:

Age	31..40	<= 30	<=30	> 40
Income				
High	5	1	5	0
Low	5	6	0	10
Medium	4	8	0	7

In [18]:
```
PHighNo = 6/22
PLowNo = 11/22
PMediumNo = 5/22

PHighYes = 5/29
PLowYes = 10/29
PMediumYes = 5/29

PHigh = 11/52
PLow = 21/51
PMedium = 19/51


print (PHighNo)
```

```
0.2727272727272727
```

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

```
0.1724137931034483
```

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

```
0.21153846153846154
```

In [21]:
```
print (PLowNo)
```

```
0.5
```

In [22]:
```
print (PLowYes)
```

```
0.3448275862068966
```

In [23]:
```
print (PLow)
```

```
0.4117647058823529
```

In [24]:
```
print (PMediumYes)
```

```
0.1724137931034483
```

In [25]:
```
print (PMediumNo)
```

```
0.22727272727272727
```

In [26]:
```
print (PMedium)
```

```
0.37254901960784315
```

In [27]:
```
data.describe()
```

Out[27]:

	Age	Income	Student	Credit_rating	Class (buy_computer)
count	51	51	51	51	51
unique	4	3	2	2	2
top	> 40	Low	Yes	Fair	Yes
freq	17	21	27	31	29

In [29]:
```
data.to_excel('F:\Kuliah/Semester 6/Data Mining/UAS/jawaban/JawabanNo1.xls')
```

In [ ]:
```
```