Knowledge Discovery & Data Mining Lab-03

Name: Gurvinder Kaur Matharu
PRN: 20190802077

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

car.data
Attribute Information:
Class Values:
unacc, acc, good, vgood
Attributes:
buying: vhigh, high, med, low.
maint: vhigh, high, med, low.
doors: 2, 3, 4, 5more.
persons: 2, 4, more.
lug_boot: small, med, big.
safety: low, med, high.

In [2]:
```
with open('car.data') as f:
    data = f.read()
    data = data.split('\n')
print(data)
```

```
['vhigh,vhigh,2,2,small,low,unacc', 'vhigh,vhigh,2,2,small,med,unacc', 'vhigh,vhigh,2,2,small,hig
h,unacc', 'vhigh,vhigh,2,2,med,low,unacc', 'vhigh,vhigh,2,2,med,med,unacc', 'vhigh,vhigh,2,2,med,
high,unacc', 'vhigh,vhigh,2,2,big,low,unacc', 'vhigh,vhigh,2,2,big,med,unacc', 'vhigh,vhigh,2,2,b
ig,high,unacc', 'vhigh,vhigh,2,4,small,low,unacc', 'vhigh,vhigh,2,4,small,med,unacc', 'vhigh,vhig
h,2,4,small,high,unacc', 'vhigh,vhigh,2,4,med,low,unacc', 'vhigh,vhigh,2,4,med,med,unacc', 'vhig
h,vhigh,2,4,med,high,unacc', 'vhigh,vhigh,2,4,big,low,unacc', 'vhigh,vhigh,2,4,big,med,unacc', 'v
high,vhigh,2,4,big,high,unacc', 'vhigh,vhigh,2,more,small,low,unacc', 'vhigh,vhigh,2,more,small,m
ed,unacc', 'vhigh,vhigh,2,more,small,high,unacc', 'vhigh,vhigh,2,more,med,low,unacc', 'vhigh,vhig
h,2,more,med,med,unacc', 'vhigh,vhigh,2,more,med,high,unacc', 'vhigh,vhigh,2,more,big,low,unacc',
'vhigh,vhigh,2,more,big,med,unacc', 'vhigh,vhigh,2,more,big,high,unacc', 'vhigh,vhigh,3,2,small,l
ow,unacc', 'vhigh,vhigh,3,2,small,med,unacc', 'vhigh,vhigh,3,2,small,high,unacc', 'vhigh,vhigh,3,
2,med,low,unacc', 'vhigh,vhigh,3,2,med,med,unacc', 'vhigh,vhigh,3,2,med,high,unacc', 'vhigh,vhig
h,3,2,big,low,unacc', 'vhigh,vhigh,3,2,big,med,unacc', 'vhigh,vhigh,3,2,big,high,unacc', 'vhigh,v
high,3,4,small,low,unacc', 'vhigh,vhigh,3,4,small,med,unacc', 'vhigh,vhigh,3,4,small,high,unacc',
'vhigh,vhigh,3,4,med,low,unacc', 'vhigh,vhigh,3,4,med,med,unacc', 'vhigh,vhigh,3,4,med,high,unac
c', 'vhigh,vhigh,3,4,big,low,unacc', 'vhigh,vhigh,3,4,big,med,unacc', 'vhigh,vhigh,3,4,big,high,u
nacc', 'vhigh,vhigh,3,more,small,low,unacc', 'vhigh,vhigh,3,more,small,med,unacc', 'vhigh,vhigh,
3,more,small,high,unacc', 'vhigh,vhigh,3,more,med,low,unacc', 'vhigh,vhigh,3,more,med,med,unacc',
'vhigh,vhigh,3,more,med,high,unacc', 'vhigh,vhigh,3,more,big,low,unacc', 'vhigh,vhigh,3,more,big,
```

In [3]:
```
new_data = []
for line in data:
    new_data.append(line.split(','))
```

In [4]:
```
print(new_data)
```

```
'low', 'unacc'], ['vhigh', 'med', '5more', 'more', 'small', 'med', 'unacc'], ['vhigh', 'med', '5m
ore', 'more', 'small', 'high', 'acc'], ['vhigh', 'med', '5more', 'more', 'med', 'low', 'unacc'],
['vhigh', 'med', '5more', 'more', 'med', 'med', 'acc'], ['vhigh', 'med', '5more', 'more', 'med',
'high', 'acc'], ['vhigh', 'med', '5more', 'more', 'big', 'low', 'unacc'], ['vhigh', 'med', '5mor
e', 'more', 'big', 'med', 'acc'], ['vhigh', 'med', '5more', 'more', 'big', 'high', 'acc'], ['vhig
h', 'low', '2', '2', 'small', 'low', 'unacc'], ['vhigh', 'low', '2', '2', 'small', 'med', 'unac
c'], ['vhigh', 'low', '2', '2', 'small', 'high', 'unacc'], ['vhigh', 'low', '2', '2', 'med', 'lo
w', 'unacc'], ['vhigh', 'low', '2', '2', 'med', 'med', 'unacc'], ['vhigh', 'low', '2', '2', 'me
d', 'high', 'unacc'], ['vhigh', 'low', '2', '2', 'big', 'low', 'unacc'], ['vhigh', 'low', '2',
'2', 'big', 'med', 'unacc'], ['vhigh', 'low', '2', '2', 'big', 'high', 'unacc'], ['vhigh', 'low',
'2', '4', 'small', 'low', 'unacc'], ['vhigh', 'low', '2', '4', 'small', 'med', 'unacc'], ['vhig
h', 'low', '2', '4', 'small', 'high', 'acc'], ['vhigh', 'low', '2', '4', 'med', 'low', 'unacc'],
['vhigh', 'low', '2', '4', 'med', 'med', 'unacc'], ['vhigh', 'low', '2', '4', 'med', 'high', 'ac
c'], ['vhigh', 'low', '2', '4', 'big', 'low', 'unacc'], ['vhigh', 'low', '2', '4', 'big', 'med',
'acc'], ['vhigh', 'low', '2', '4', 'big', 'high', 'acc'], ['vhigh', 'low', '2', 'more', 'small',
'low', 'unacc'], ['vhigh', 'low', '2', 'more', 'small', 'med', 'unacc'], ['vhigh', 'low', '2', 'm
ore', 'small', 'high', 'unacc'], ['vhigh', 'low', '2', 'more', 'med', 'low', 'unacc'], ['vhigh',
'low', '2', 'more', 'med', 'med', 'unacc'], ['vhigh', 'low', '2', 'more', 'med', 'high', 'acc'],
['vhigh', 'low', '2', 'more', 'big', 'low', 'unacc'], ['vhigh', 'low', '2', 'more', 'big', 'med',
```

In [5]:
```
df = pd.DataFrame(new_data, columns=['buying', 'maint', 'doors', 'persons', 'lug_boot','safety','Labe
```

In [6]:
```
df
```

	buying	maint	doors	persons	lug_boot	safety	Label
0	vhigh	vhigh	2	2	small	low	unacc
1	vhigh	vhigh	2	2	small	med	unacc
2	vhigh	vhigh	2	2	small	high	unacc
3	vhigh	vhigh	2	2	med	low	unacc
4	vhigh	vhigh	2	2	med	med	unacc
...	...	...	...	...	...	...	...
1724	low	low	5more	more	med	high	vgood
1725	low	low	5more	more	big	low	unacc
1726	low	low	5more	more	big	med	good
1727	low	low	5more	more	big	high	vgood
1728		None	None	None	None	None	None

1729 rows × 7 columns

In [7]:
```
df.to_csv('car.csv', index=False) # converting into a csv file
```

In [8]:
```
df.to_excel('car_excel.xlsx', index=False) # converting to excel file
```

Haberman's Cancer Survival Dataset
Attribute Information:

1. Age of patient at time of operation (numerical)
2. Patient's year of operation (year - 1900, numerical)
3. Number of positive axillary nodes detected (numerical)
4. Survival status (class attribute)
   1 = the patient survived 5 years or longer
   2 = the patient died within 5 year

In [9]:
```
with open('haberman.data') as i:
    data1 = i.read()
    data1 = data1.split('\n')
```

In [10]:
```
print(data1)
```

```
['30,64,1,1', '30,62,3,1', '30,65,0,1', '31,59,2,1', '31,65,4,1', '33,58,10,1', '33,60,0,1', '34,59,
0,2', '34,66,9,2', '34,58,30,1', '34,60,1,1', '34,61,10,1', '34,67,7,1', '34,60,0,1', '35,64,13,1',
'35,63,0,1', '36,60,1,1', '36,69,0,1', '37,60,0,1', '37,63,0,1', '37,58,0,1', '37,59,6,1', '37,60,1
5,1', '37,63,0,1', '38,69,21,2', '38,59,2,1', '38,60,0,1', '38,60,0,1', '38,62,3,1', '38,64,1,1', '3
8,66,0,1', '38,66,11,1', '38,60,1,1', '38,67,5,1', '39,66,0,2', '39,63,0,1', '39,67,0,1', '39,58,0,
1', '39,59,2,1', '39,63,4,1', '40,58,2,1', '40,58,0,1', '40,65,0,1', '41,60,23,2', '41,64,0,2', '41,
67,0,2', '41,58,0,1', '41,59,8,1', '41,59,0,1', '41,64,0,1', '41,69,8,1', '41,65,0,1', '41,65,0,1',
'42,69,1,2', '42,59,0,2', '42,58,0,1', '42,60,1,1', '42,59,2,1', '42,61,4,1', '42,62,20,1', '42,65,
0,1', '42,63,1,1', '43,58,52,2', '43,59,2,2', '43,64,0,2', '43,64,0,2', '43,63,14,1', '43,64,2,1',
'43,64,3,1', '43,60,0,1', '43,63,2,1', '43,65,0,1', '43,66,4,1', '44,64,6,2', '44,58,9,2', '44,63,1
9,2', '44,61,0,1', '44,63,1,1', '44,61,0,1', '44,67,16,1', '45,65,6,2', '45,66,0,2', '45,67,1,2', '4
5,60,0,1', '45,67,0,1', '45,59,14,1', '45,64,0,1', '45,68,0,1', '45,67,1,1', '46,58,2,2', '46,69,3,
2', '46,62,5,2', '46,65,20,2', '46,62,0,1', '46,58,3,1', '46,63,0,1', '47,63,23,2', '47,62,0,2', '4
7,65,0,2', '47,61,0,1', '47,63,6,1', '47,66,0,1', '47,67,0,1', '47,58,3,1', '47,60,4,1', '47,68,4,
1', '47,66,12,1', '48,58,11,2', '48,58,11,2', '48,67,7,2', '48,61,8,1', '48,62,2,1', '48,64,0,1', '4
8,66,0,1', '49,63,0,2', '49,64,10,2', '49,61,1,1', '49,62,0,1', '49,66,0,1', '49,60,1,1', '49,62,1,
1', '49,63,3,1', '49,61,0,1', '49,67,1,1', '50,63,13,2', '50,64,0,2', '50,59,0,1', '50,61,6,1', '50,
61,0,1', '50,63,1,1', '50,58,1,1', '50,59,2,1', '50,61,0,1', '50,64,0,1', '50,65,4,1', '50,66,1,1',
'51,59,13,2', '51,59,3,2', '51,64,7,1', '51,59,1,1', '51,65,0,1', '51,66,1,1', '52,69,3,2', '52,59,
2,2', '52,62,3,2', '52,66,4,2', '52,61,0,1', '52,63,4,1', '52,69,0,1', '52,60,4,1', '52,60,5,1', '5
2,62,0,1', '52,62,1,1', '52,64,0,1', '52,65,0,1', '52,68,0,1', '53,58,4,2', '53,65,1,2', '53,59,3,
2', '53,60,9,2', '53,63,24,2', '53,65,12,2', '53,58,1,1', '53,60,1,1', '53,60,2,1', '53,61,1,1', '5
3,63,0,1', '54,60,11,2', '54,65,23,2', '54,65,5,2', '54,68,7,2', '54,59,7,1', '54,60,3,1', '54,66,0,
1', '54,67,46,1', '54,62,0,1', '54,69,7,1', '54,63,19,1', '54,58,1,1', '54,62,0,1', '55,63,6,2', '5
5,68,15,2', '55,58,1,1', '55,58,0,1', '55,58,1,1', '55,66,18,1', '55,66,0,1', '55,69,3,1', '55,69,2
2,1', '55,67,1,1', '56,65,9,2', '56,66,3,2', '56,60,0,1', '56,66,2,1', '56,66,1,1', '56,67,0,1', '5
6,60,0,1', '57,61,5,2', '57,62,14,2', '57,64,1,2', '57,64,9,1', '57,69,0,1', '57,61,0,1', '57,62,0,
1', '57,63,0,1', '57,64,0,1', '57,64,0,1', '57,67,0,1', '58,59,0,1', '58,60,3,1', '58,61,1,1', '58,6
7,0,1', '58,58,0,1', '58,58,3,1', '58,61,2,1', '59,62,35,2', '59,60,0,1', '59,63,0,1', '59,64,1,1',
'59,64,4,1', '59,64,0,1', '59,64,7,1', '59,67,3,1', '60,59,17,2', '60,65,0,2', '60,61,1,1', '60,67,
2,1', '60,61,25,1', '60,64,0,1', '61,62,5,2', '61,65,0,2', '61,68,1,2', '61,59,0,1', '61,59,0,1', '6
1,64,0,1', '61,65,8,1', '61,68,0,1', '61,59,0,1', '62,59,13,2', '62,58,0,2', '62,65,19,2', '62,62,6,
1', '62,66,0,1', '62,66,0,1', '62,58,0,1', '63,60,1,2', '63,61,0,1', '63,62,0,1', '63,63,0,1', '63,6
3,0,1', '63,66,0,1', '63,61,9,1', '63,61,28,1', '64,58,0,1', '64,65,22,1', '64,66,0,1', '64,61,0,1',
'64,68,0,1', '65,58,0,2', '65,61,2,2', '65,62,22,2', '65,66,15,2', '65,58,0,1', '65,64,0,1', '65,67,
0,1', '65,59,2,1', '65,64,0,1', '65,67,1,1', '66,58,0,2', '66,61,13,2', '66,58,0,1', '66,58,1,1', '6
6,68,0,1', '67,64,8,2', '67,63,1,2', '67,66,0,1', '67,66,0,1', '67,61,0,1', '67,65,0,1', '68,67,0,
1', '68,68,0,1', '69,67,8,2', '69,60,0,1', '69,65,0,1', '69,66,0,1', '70,58,0,2', '70,58,4,2', '70,6
6,14,1', '70,67,0,1', '70,68,0,1', '70,59,8,1', '70,63,0,1', '71,68,2,1', '72,63,0,2', '72,58,0,1',
'72,64,0,1', '72,67,3,1', '73,62,0,1', '73,68,0,1', '74,65,3,2', '74,63,0,1', '75,62,1,1', '76,67,0,
1', '77,65,3,1', '78,65,1,2', '83,58,2,2', '']
```

In [11]:
```
new_d = []
for l in data1:
    new_d.append(l.split(','))
```

In [12]:
```
print(new_d)
```

```
[['30', '64', '1', '1'], ['30', '62', '3', '1'], ['30', '65', '0', '1'], ['31', '59', '2', '1'], ['3
1', '65', '4', '1'], ['33', '58', '10', '1'], ['33', '60', '0', '1'], ['34', '59', '0', '2'], ['34',
'66', '9', '2'], ['34', '58', '30', '1'], ['34', '60', '1', '1'], ['34', '61', '10', '1'], ['34', '6
7', '7', '1'], ['34', '60', '0', '1'], ['35', '64', '13', '1'], ['35', '63', '0', '1'], ['36', '60',
'1', '1'], ['36', '69', '0', '1'], ['37', '60', '0', '1'], ['37', '63', '0', '1'], ['37', '58', '0',
'1'], ['37', '59', '6', '1'], ['37', '60', '15', '1'], ['37', '63', '0', '1'], ['38', '69', '21',
'2'], ['38', '59', '2', '1'], ['38', '60', '0', '1'], ['38', '60', '0', '1'], ['38', '62', '3',
'1'], ['38', '64', '1', '1'], ['38', '66', '0', '1'], ['38', '66', '11', '1'], ['38', '60', '1',
'1'], ['38', '67', '5', '1'], ['39', '66', '0', '2'], ['39', '63', '0', '1'], ['39', '67', '0',
'1'], ['39', '58', '0', '1'], ['39', '59', '2', '1'], ['39', '63', '4', '1'], ['40', '58', '2',
'1'], ['40', '58', '0', '1'], ['40', '65', '0', '1'], ['41', '60', '23', '2'], ['41', '64', '0',
'2'], ['41', '67', '0', '2'], ['41', '58', '0', '1'], ['41', '59', '8', '1'], ['41', '59', '0',
'1'], ['41', '64', '0', '1'], ['41', '69', '8', '1'], ['41', '65', '0', '1'], ['41', '65', '0',
'1'], ['42', '69', '1', '2'], ['42', '59', '0', '2'], ['42', '58', '0', '1'], ['42', '60', '1',
'1'], ['42', '59', '2', '1'], ['42', '61', '4', '1'], ['42', '62', '20', '1'], ['42', '65', '0',
'1'], ['42', '63', '1', '1'], ['43', '58', '52', '2'], ['43', '59', '2', '2'], ['43', '64', '0',
'2'], ['43', '64', '0', '2'], ['43', '63', '14', '1'], ['43', '64', '2', '1'], ['43', '64', '3',
'1'], ['43', '60', '0', '1'], ['43', '63', '2', '1'], ['43', '65', '0', '1'], ['43', '66', '4',
'1'], ['44', '64', '6', '2'], ['44', '58', '9', '2'], ['44', '63', '19', '2'], ['44', '61', '0',
'1'], ['44', '63', '1', '1'], ['44', '61', '0', '1'], ['44', '67', '16', '1'], ['45', '65', '6',
'2'], ['45', '66', '0', '2'], ['45', '67', '1', '2'], ['45', '60', '0', '1'], ['45', '67', '0',
'1'], ['45', '59', '14', '1'], ['45', '64', '0', '1'], ['45', '68', '0', '1'], ['45', '67', '1',
'1'], ['46', '58', '2', '2'], ['46', '69', '3', '2'], ['46', '62', '5', '2'], ['46', '65', '20',
'2'], ['46', '62', '0', '1'], ['46', '58', '3', '1'], ['46', '63', '0', '1'], ['47', '63', '23',
'2'], ['47', '62', '0', '2'], ['47', '65', '0', '2'], ['47', '61', '0', '1'], ['47', '63', '6',
'1'], ['47', '66', '0', '1'], ['47', '67', '0', '1'], ['47', '58', '3', '1'], ['47', '60', '4',
'1'], ['47', '68', '4', '1'], ['47', '66', '12', '1'], ['48', '58', '11', '2'], ['48', '58', '11',
'2'], ['48', '67', '7', '2'], ['48', '61', '8', '1'], ['48', '62', '2', '1'], ['48', '64', '0',
'1'], ['48', '66', '0', '1'], ['49', '63', '0', '2'], ['49', '64', '10', '2'], ['49', '61', '1',
'1'], ['49', '62', '0', '1'], ['49', '66', '0', '1'], ['49', '60', '1', '1'], ['49', '62', '1',
'1'], ['49', '63', '3', '1'], ['49', '61', '0', '1'], ['49', '67', '1', '1'], ['50', '63', '13',
'2'], ['50', '64', '0', '2'], ['50', '59', '0', '1'], ['50', '61', '6', '1'], ['50', '61', '0',
'1'], ['50', '63', '1', '1'], ['50', '58', '1', '1'], ['50', '59', '2', '1'], ['50', '61', '0',
'1'], ['50', '64', '0', '1'], ['50', '65', '4', '1'], ['50', '66', '1', '1'], ['51', '59', '13',
'2'], ['51', '59', '3', '2'], ['51', '64', '7', '1'], ['51', '59', '1', '1'], ['51', '65', '0',
'1'], ['51', '66', '1', '1'], ['52', '69', '3', '2'], ['52', '59', '2', '2'], ['52', '62', '3',
'2'], ['52', '66', '4', '2'], ['52', '61', '0', '1'], ['52', '63', '4', '1'], ['52', '69', '0',
'1'], ['52', '60', '4', '1'], ['52', '60', '5', '1'], ['52', '62', '0', '1'], ['52', '62', '1',
'1'], ['52', '64', '0', '1'], ['52', '65', '0', '1'], ['52', '68', '0', '1'], ['53', '58', '4',
'2'], ['53', '65', '1', '2'], ['53', '59', '3', '2'], ['53', '60', '9', '2'], ['53', '63', '24',
'2'], ['53', '65', '12', '2'], ['53', '58', '1', '1'], ['53', '60', '1', '1'], ['53', '60', '2',
'1'], ['53', '61', '1', '1'], ['53', '63', '0', '1'], ['54', '60', '11', '2'], ['54', '65', '23',
'2'], ['54', '65', '5', '2'], ['54', '68', '7', '2'], ['54', '59', '7', '1'], ['54', '60', '3',
'1'], ['54', '66', '0', '1'], ['54', '67', '46', '1'], ['54', '62', '0', '1'], ['54', '69', '7',
'1'], ['54', '63', '19', '1'], ['54', '58', '1', '1'], ['54', '62', '0', '1'], ['55', '63', '6',
'2'], ['55', '68', '15', '2'], ['55', '58', '1', '1'], ['55', '58', '0', '1'], ['55', '58', '1',
'1'], ['55', '66', '18', '1'], ['55', '66', '0', '1'], ['55', '69', '3', '1'], ['55', '69', '22',
'1'], ['55', '67', '1', '1'], ['56', '65', '9', '2'], ['56', '66', '3', '2'], ['56', '60', '0',
'1'], ['56', '66', '2', '1'], ['56', '66', '1', '1'], ['56', '67', '0', '1'], ['56', '60', '0',
'1'], ['57', '61', '5', '2'], ['57', '62', '14', '2'], ['57', '64', '1', '2'], ['57', '64', '9',
'1'], ['57', '69', '0', '1'], ['57', '61', '0', '1'], ['57', '62', '0', '1'], ['57', '63', '0',
'1'], ['57', '64', '0', '1'], ['57', '64', '0', '1'], ['57', '67', '0', '1'], ['58', '59', '0',
'1'], ['58', '60', '3', '1'], ['58', '61', '1', '1'], ['58', '67', '0', '1'], ['58', '58', '0',
'1'], ['58', '58', '3', '1'], ['58', '61', '2', '1'], ['59', '62', '35', '2'], ['59', '60', '0',
'1'], ['59', '63', '0', '1'], ['59', '64', '1', '1'], ['59', '64', '4', '1'], ['59', '64', '0',
'1'], ['59', '64', '7', '1'], ['59', '67', '3', '1'], ['60', '59', '17', '2'], ['60', '65', '0',
'2'], ['60', '61', '1', '1'], ['60', '67', '2', '1'], ['60', '61', '25', '1'], ['60', '64', '0',
'1'], ['61', '62', '5', '2'], ['61', '65', '0', '2'], ['61', '68', '1', '2'], ['61', '59', '0',
'1'], ['61', '59', '0', '1'], ['61', '64', '0', '1'], ['61', '65', '8', '1'], ['61', '68', '0',
'1'], ['61', '59', '0', '1'], ['62', '59', '13', '2'], ['62', '58', '0', '2'], ['62', '65', '19',
'2'], ['62', '62', '6', '1'], ['62', '66', '0', '1'], ['62', '66', '0', '1'], ['62', '58', '0',
'1'], ['63', '60', '1', '2'], ['63', '61', '0', '1'], ['63', '62', '0', '1'], ['63', '63', '0',
'1'], ['63', '63', '0', '1'], ['63', '66', '0', '1'], ['63', '61', '9', '1'], ['63', '61', '28',
'1'], ['64', '58', '0', '1'], ['64', '65', '22', '1'], ['64', '66', '0', '1'], ['64', '61', '0',
'1'], ['64', '68', '0', '1'], ['65', '58', '0', '2'], ['65', '61', '2', '2'], ['65', '62', '22',
'2'], ['65', '66', '15', '2'], ['65', '58', '0', '1'], ['65', '64', '0', '1'], ['65', '67', '0',
'1'], ['65', '59', '2', '1'], ['65', '64', '0', '1'], ['65', '67', '1', '1'], ['66', '58', '0',
'2'], ['66', '61', '13', '2'], ['66', '58', '0', '1'], ['66', '58', '1', '1'], ['66', '68', '0',
'1'], ['67', '64', '8', '2'], ['67', '63', '1', '2'], ['67', '66', '0', '1'], ['67', '66', '0',
'1'], ['67', '61', '0', '1'], ['67', '65', '0', '1'], ['68', '67', '0', '1'], ['68', '68', '0',
'1'], ['69', '67', '8', '2'], ['69', '60', '0', '1'], ['69', '65', '0', '1'], ['69', '66', '0',
'1'], ['70', '58', '0', '2'], ['70', '58', '4', '2'], ['70', '66', '14', '1'], ['70', '67', '0',
'1'], ['70', '68', '0', '1'], ['70', '59', '8', '1'], ['70', '63', '0', '1'], ['71', '68', '2',
'1'], ['72', '63', '0', '2'], ['72', '58', '0', '1'], ['72', '64', '0', '1'], ['72', '67', '3',
```

```
'1'], ['73', '62', '0', '1'], ['73', '68', '0', '1'], ['74', '65', '3', '2'], ['74', '63', '0',
'1'], ['75', '62', '1', '1'], ['76', '67', '0', '1'], ['77', '65', '3', '1'], ['78', '65', '1',
'2'], ['83', '58', '2', '2'], ['']]
```

In [13]:
```
df1 = pd.DataFrame(new_d, columns=['age', 'yr_of_operation', 'pos_axillary_nodes', 'survival_stat'])
```

In [14]:
```
df1
```

Out[14]:

	age	yr_of_operation	pos_axillary_nodes	survival_stat
0	30	64	1	1
1	30	62	3	1
2	30	65	0	1
3	31	59	2	1
4	31	65	4	1
...	...	...	...	...
302	76	67	0	1
303	77	65	3	1
304	78	65	1	2
305	83	58	2	2
306		None	None	None

307 rows × 4 columns

In [15]:
```
df1.to_csv('haberman_csv.csv', index=False) # converting into csv file
```

In [16]:
```
df1.to_excel('haberman_excel.xlsx', index=False) # converting into excel file
```