

```
In [2]: import pandas as pd
import numpy as np
import seaborn as sns
```

```
In [3]: data = pd.read_csv("heart.csv")
```

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

Out[4]:	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	tar
0	52	1	0	125	212	0	1	168	0	1.0	2	2	3	
1	53	1	0	140	203	1	0	155	1	3.1	0	0	3	
2	70	1	0	145	174	0	1	125	1	2.6	0	0	3	
3	61	1	0	148	203	0	1	161	0	0.0	2	1	3	
4	62	0	0	138	294	1	1	106	0	1.9	1	3	2	

```
In [6]: data.tail()
```

```
Out[6]:
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
1020	59	1	1	140	221	0	1	164	1	0.0	2	0	2
1021	60	1	0	125	258	0	0	141	1	2.8	1	1	3
1022	47	1	0	110	275	0	0	118	1	1.0	1	1	2
1023	50	0	0	110	254	0	0	159	0	0.0	2	0	2
1024	54	1	0	120	188	0	1	113	0	1.4	1	1	3

```
In [11]: data.describe()
```

Out[11]:								
	age	sex	cp	trestbps	chol	fbs	restecg	thca
count	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000
mean	54.434146	0.695610	0.942439	131.611707	246.000000	0.149268	0.500000	0.500000
std	9.072290	0.460373	1.029641	17.516718	51.59251	0.356527	0.500000	0.500000
min	29.000000	0.000000	0.000000	94.000000	126.000000	0.000000	0.000000	0.000000
25%	48.000000	0.000000	0.000000	120.000000	211.000000	0.000000	0.000000	0.000000

	age	sex	cp	trestbps	chol	fbs	re
50%	56.000000	1.000000	1.000000	130.000000	240.000000	0.000000	1.0
75%	61.000000	1.000000	2.000000	140.000000	275.000000	0.000000	1.0
max	77.000000	1.000000	3.000000	200.000000	564.000000	1.000000	2.0

In [9]: `data.shape`

Out[9]: (1025, 14)

In [12]: `data.columns`

Out[12]: Index(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach', 'exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'], dtype='object')

In [13]: `data.nunique()`

Out[13]: age 41
sex 2
cp 4
trestbps 49
chol 152
fbs 2
restecg 3
thalach 91
exang 2
oldpeak 40
slope 3
ca 5
thal 4
target 2
dtype: int64

In [17]: `data['target'].unique()`

Out[17]: array([0, 1], dtype=int64)

In [41]: `data.isnull().sum()`

Out[41]: age 0
sex 0
cp 0
trestbps 0
chol 0
fbs 0
restecg 0
thalach 0

```
exang      0
oldpeak    0
slope      0
ca         0
thal       0
target     0
dtype: int64
```

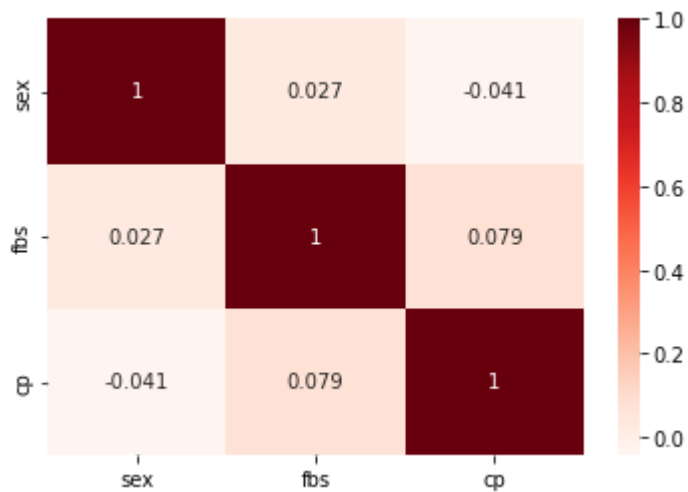
```
In [64]: import matplotlib.pyplot as plt
```

```
In [65]: data[['sex', 'fbs', 'cp']].corr()
```

```
Out[65]:
```

	sex	fbs	cp
sex	1.000000	0.027200	-0.041119
fbs	0.027200	1.000000	0.079294
cp	-0.041119	0.079294	1.000000

```
In [66]: sns.heatmap(data[['sex', 'fbs', 'cp']].corr(), annot=True, cmap = 'Reds')
plt.show()
```



```
In [70]: data.groupby('age')['sex'].mean()
```

```
Out[70]:
```

age	
29	1.000000
34	0.500000
35	0.733333
37	0.500000
38	1.000000
39	0.500000
40	1.000000
41	0.625000
42	0.769231
43	0.730769
44	0.833333

```

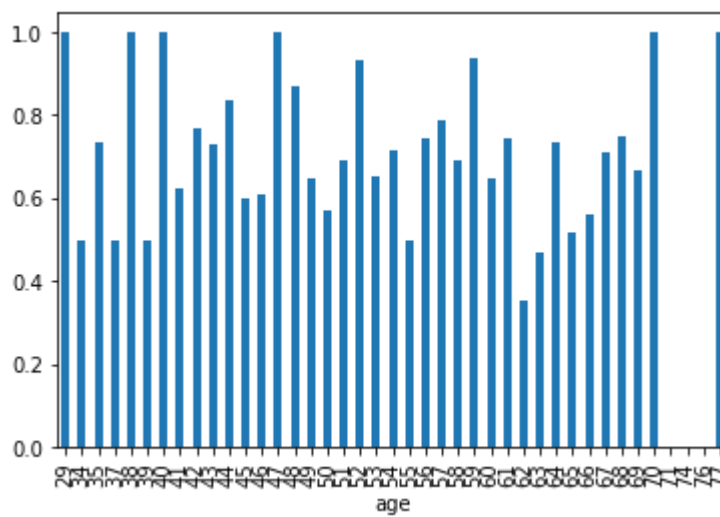
45    0.600000
46    0.608696
47    1.000000
48    0.869565
49    0.647059
50    0.571429
51    0.692308
52    0.930233
53    0.653846
54    0.716981
55    0.500000
56    0.743590
57    0.789474
58    0.691176
59    0.934783
60    0.648649
61    0.741935
62    0.351351
63    0.468750
64    0.735294
65    0.518519
66    0.560000
67    0.709677
68    0.750000
69    0.666667
70    1.000000
71    0.000000
74    0.000000
76    0.000000
77    1.000000
Name: sex, dtype: float64

```

```

In [72]: data.groupby('age')['sex'].mean().plot.bar()
          plt.show()

```



```

In [ ]:

```