In [44]:
```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
```

In [13]:
```
data = pd.read_csv('healdata2.csv')
```

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

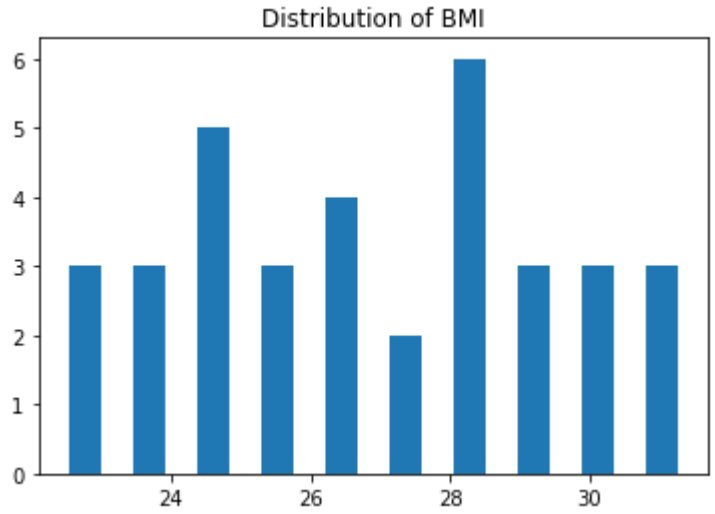
```
15           89           78
16           92           71
17          110           60
18           88           84
19          105           63
20           95           80
21          110           67
22          120           56
23           95           75
24          100           72
25           82           76
26          100           68
27          110           55
28           90           80
29           95           62
30          105           69
31           80           85
32          115           63
33           85           92
34           95           74
```

In [48]:
```
#No1

np.mean(data.age)
```

Out[48]: 42.51428571428571

In [57]:
```
#No2
plt.title('Distribution of BMI')
plt.hist(data['bmi'], rwidth= 0.5)
plt.show()
```



In [67]:
```
df=pd.DataFrame(data)
```

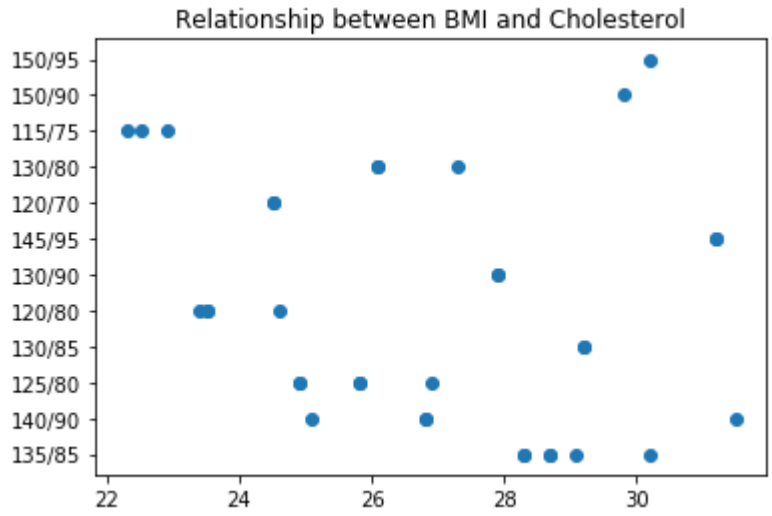
In [70]:
```
#No3

smokers = len(df[df['smoker'] == 'Yes'])
print(smokers)
```

```
17
```

In [74]:
```
#No4
bmi =df.bmi
cholesterol = df.cholesterol
plt.scatter(bmi, cholesterol)
plt.title('Relationship between BMI and Cholesterol')
```

Out[74]: Text(0.5, 1.0, 'Relationship between BMI and Cholesterol')



In [ ]:

In [ ]: