

```
!pip install pandas
```

```
Requirement already satisfied: pandas in  
/usr/local/lib/python3.10/dist-packages (2.2.2)  
Requirement already satisfied: numpy>=1.22.4 in  
/usr/local/lib/python3.10/dist-packages (from pandas) (1.26.4)  
Requirement already satisfied: python-dateutil>=2.8.2 in  
/usr/local/lib/python3.10/dist-packages (from pandas) (2.8.2)  
Requirement already satisfied: pytz>=2020.1 in  
/usr/local/lib/python3.10/dist-packages (from pandas) (2024.2)  
Requirement already satisfied: tzdata>=2022.7 in  
/usr/local/lib/python3.10/dist-packages (from pandas) (2024.2)  
Requirement already satisfied: six>=1.5 in  
/usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2-  
>pandas) (1.16.0)
```

```
import pandas as pd
```

```
df = pd.read_csv('heart_2020_cleaned(1).csv')
```

```
df
```

```
{"type": "dataframe", "variable_name": "df"}
```

What is the statistics of patients who is smoking, alcohol drinking all according to their age which is shown by a histogram graph?

```
import matplotlib.pyplot as plt
```

```
import pandas as pd
```

```
patients = pd.read_csv('heart_2020_cleaned(1).csv')
```

```
plt.hist(patients[patients['Smoking'] == 'Yes']['AgeCategory'])
```

```
plt.xlabel('Age Category')
```

```
plt.ylabel('Number of Patients')
```

```
plt.title('Smoking by Age Category')
```

```
plt.xticks(rotation=45)
```

```
plt.show()
```

```
plt.hist(patients[patients['AlcoholDrinking'] == 'Yes']  
['AgeCategory'])
```

```
plt.xlabel('Age Category')
```

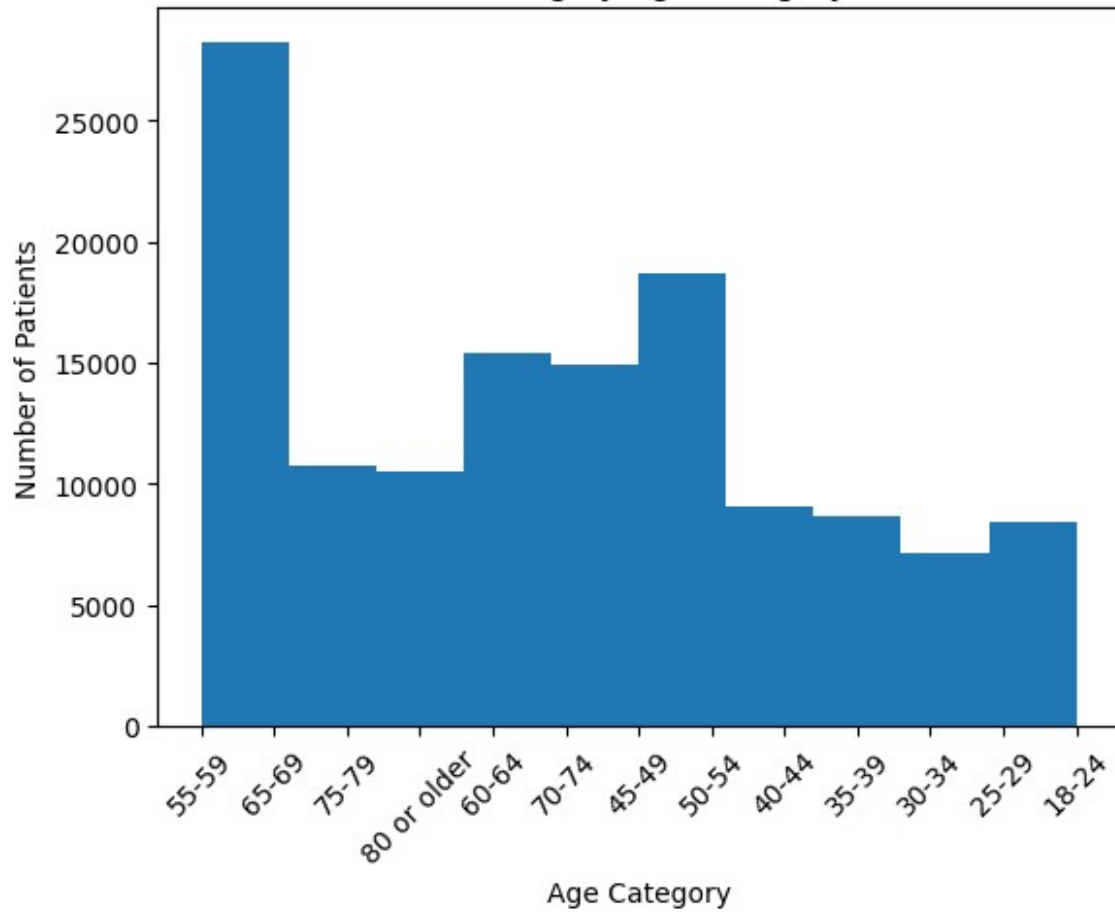
```
plt.ylabel('Number of Patients')
```

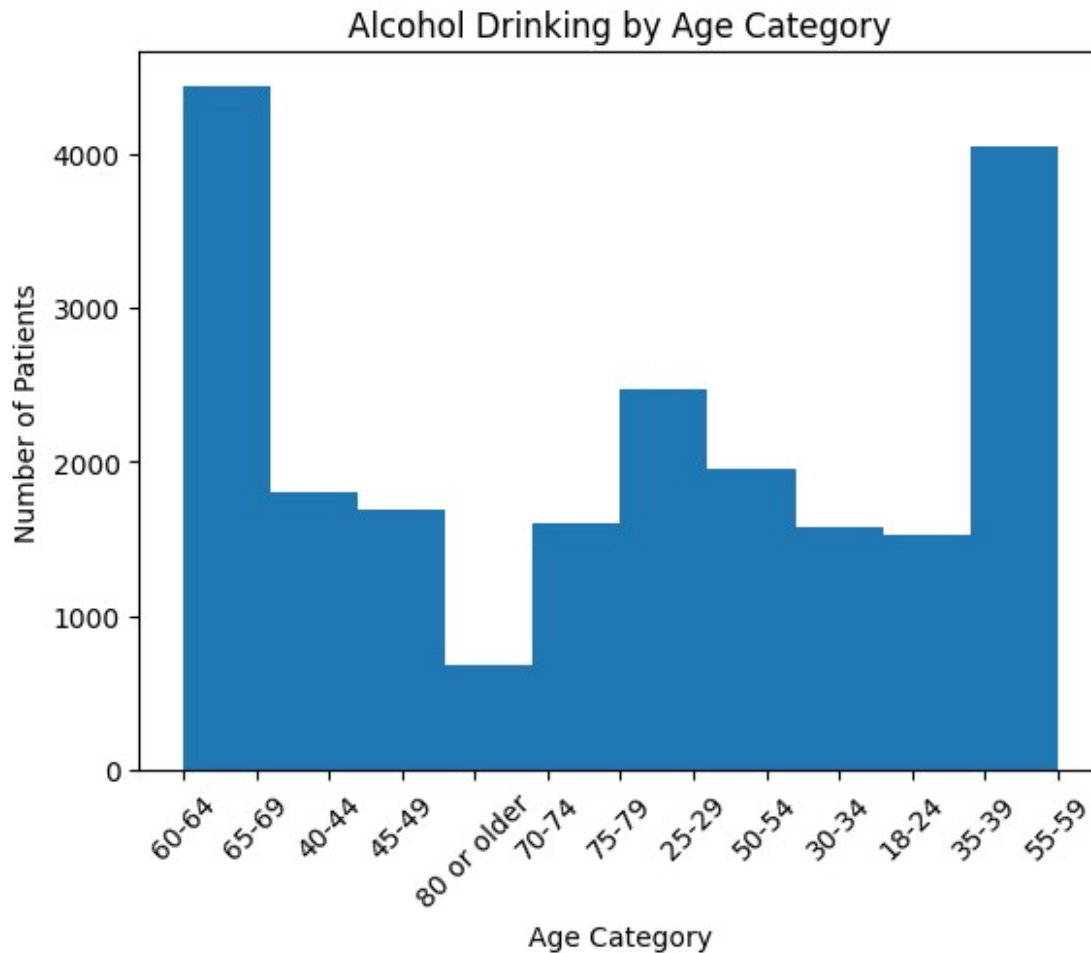
```
plt.title('Alcohol Drinking by Age Category')
```

```
plt.xticks(rotation=45)
```

```
plt.show()
```

Smoking by Age Category



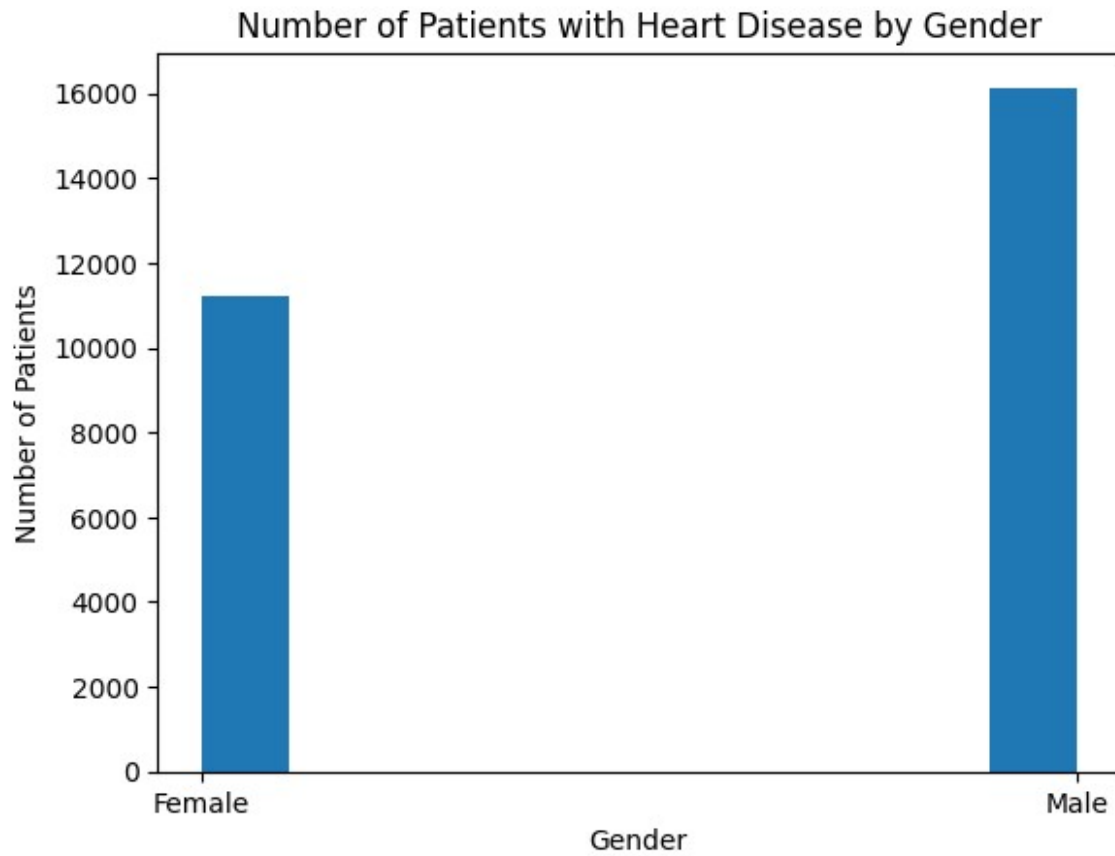


There are 2 separate graphs showing the number of patients based on their age who are drinking alcohol and smoking.

```
import matplotlib.pyplot as plt
import pandas as pd

patients = pd.read_csv('heart_2020_cleaned(1).csv')

plt.hist(patients[patients['HeartDisease'] == 'Yes']['Sex'])
plt.title('Number of Patients with Heart Disease by Gender')
plt.xlabel('Gender')
plt.ylabel('Number of Patients')
plt.show()
```



According to the graph that is shown, majority of the patients who are experiencing heart disease already are male which is around 16,000 patients. On the other hand, The number of patients for females are in between 10,000 - 12,000 patients.