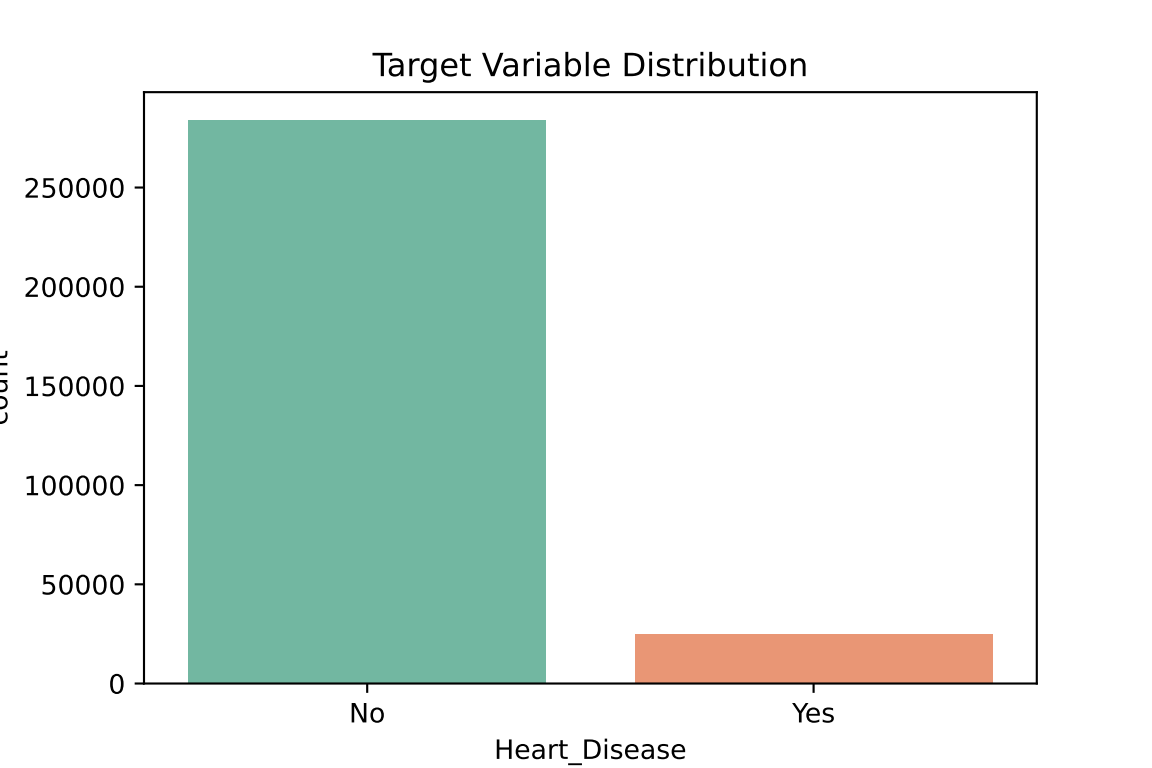
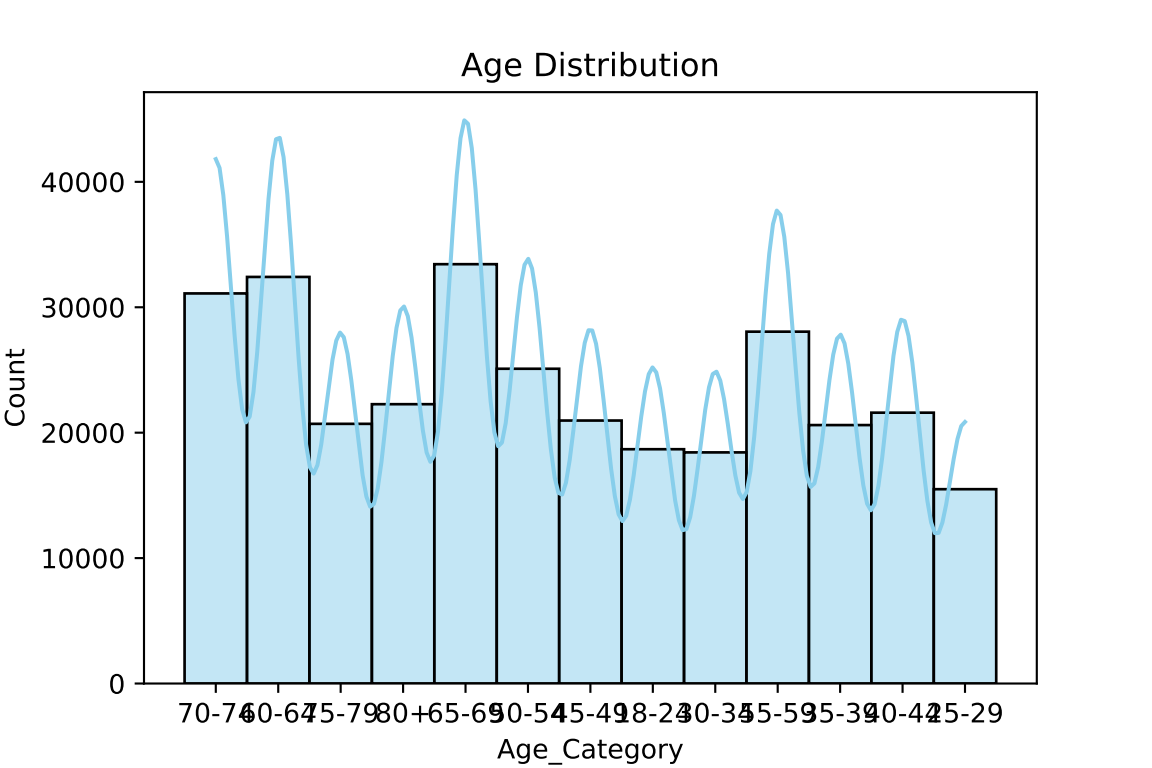
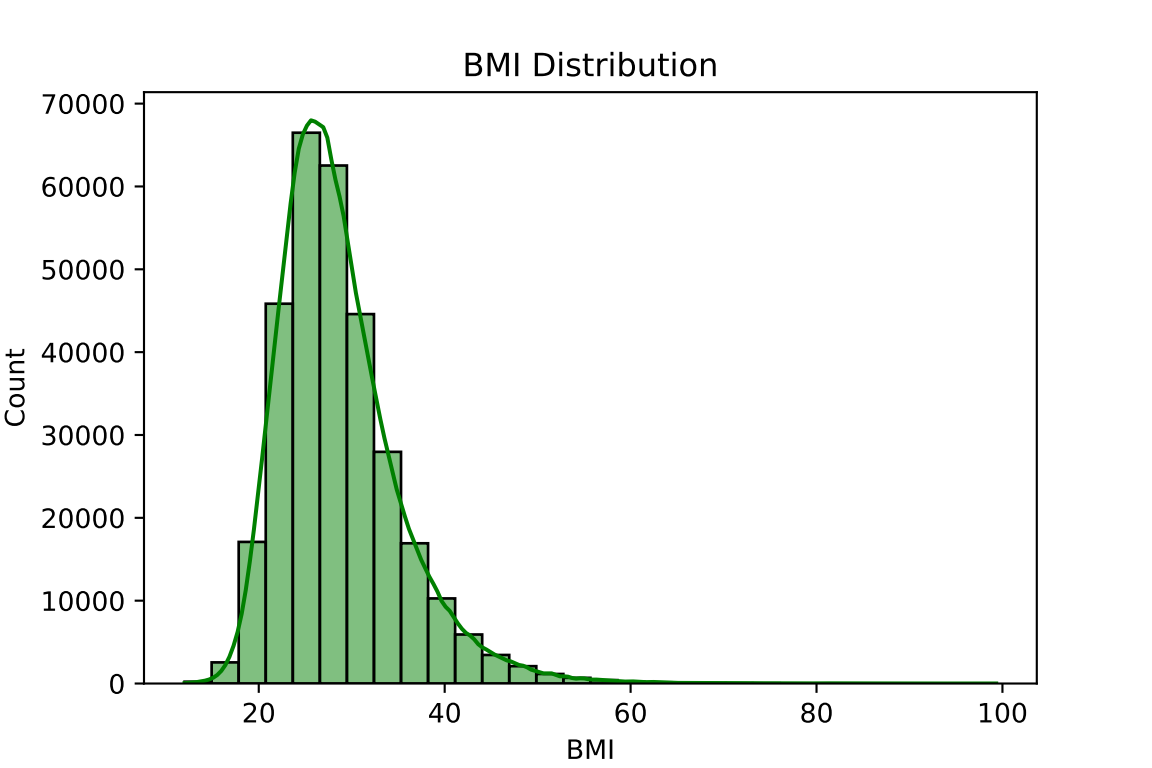
Dataset Overview

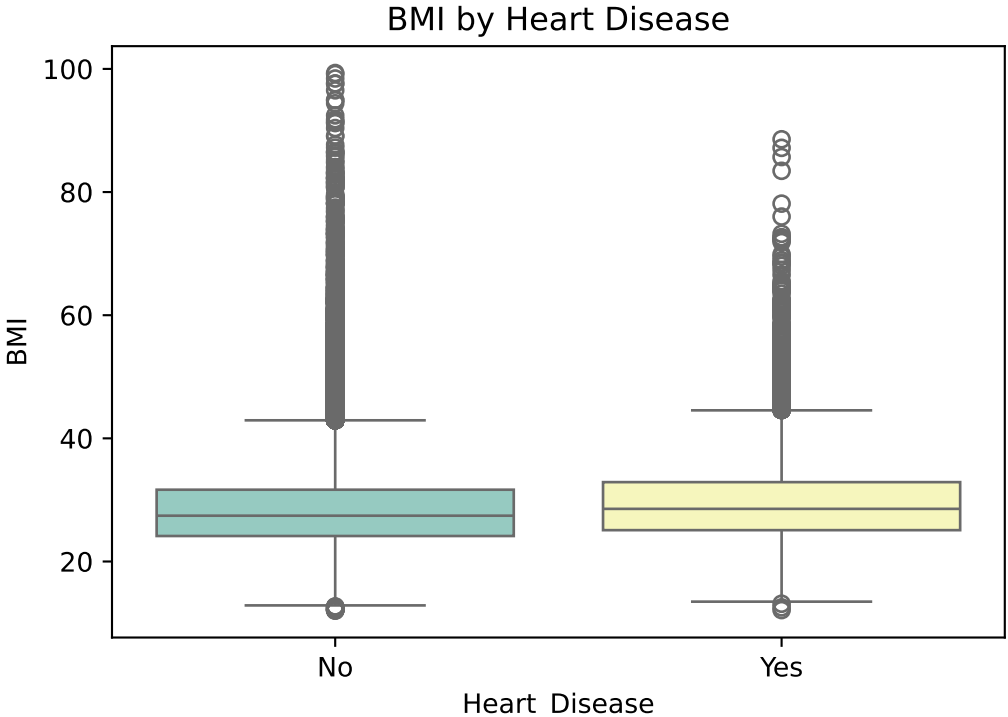
Shape: (308854, 19)

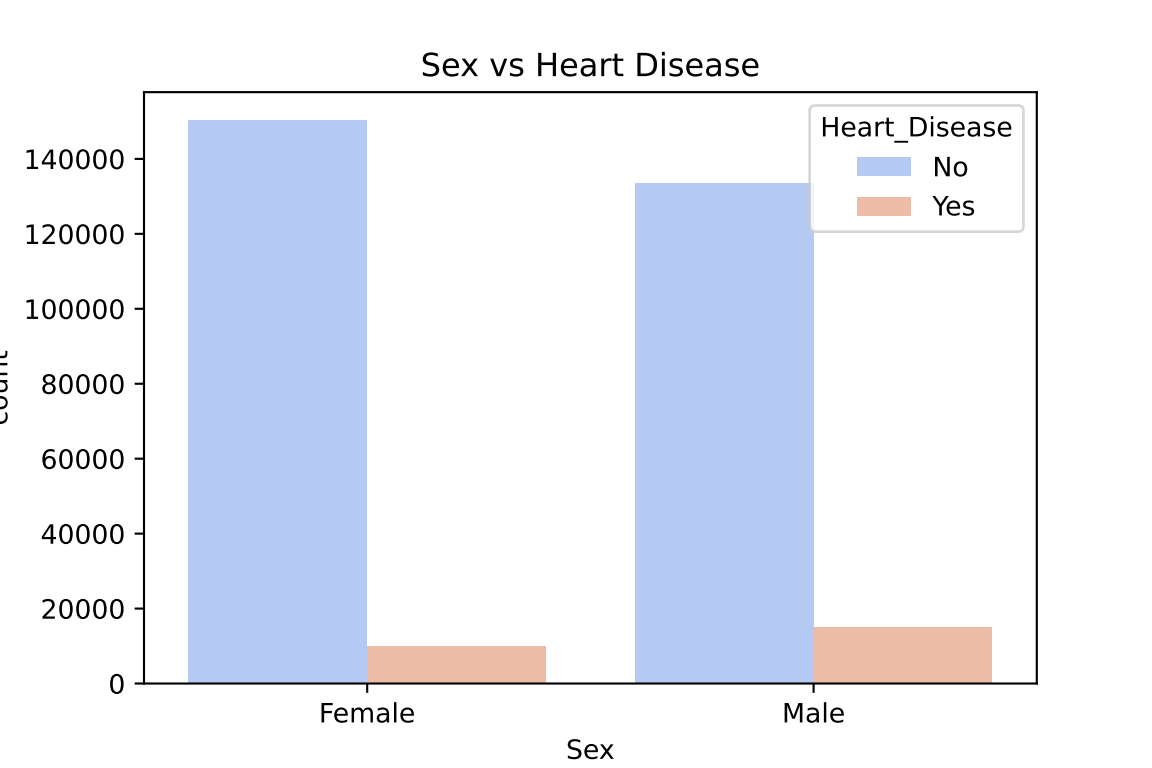
Missing values:
0
Duplicates: 80

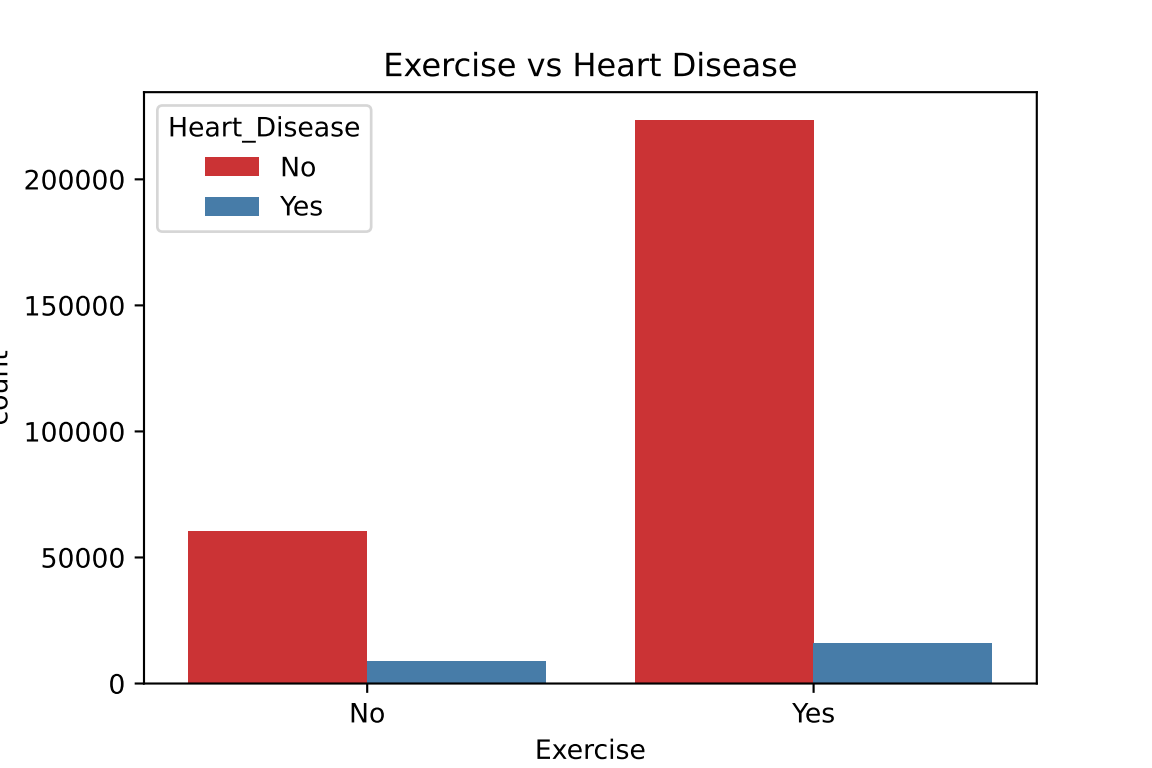

Target Variable Distribution
250000
200000
150000
100000
50000
0
No
Yes
Heart_Disease

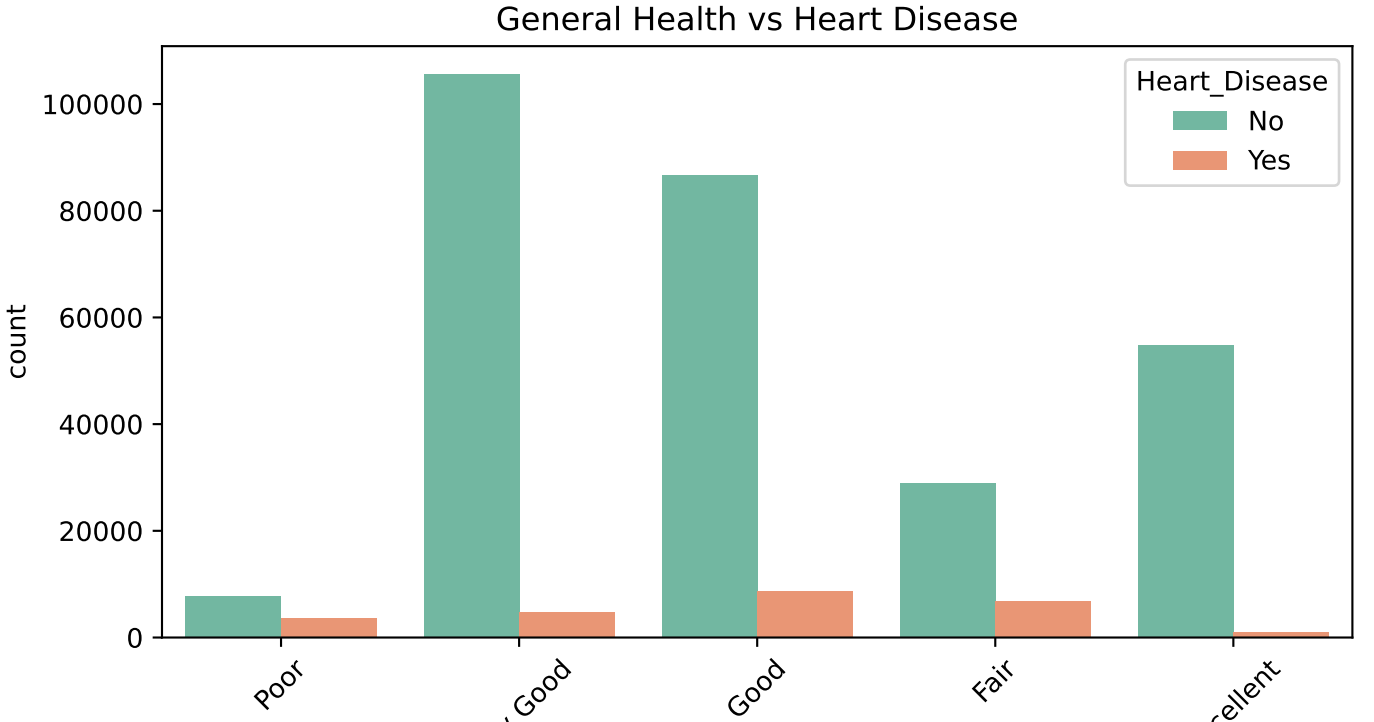

Age Distribution
Count
40000
30000
20000
10000
0
70-74
60-64
75-79
80+
65-69
50-54
45-49
18-24
30-34
55-59
35-39
40-44
25-29
Age_Category

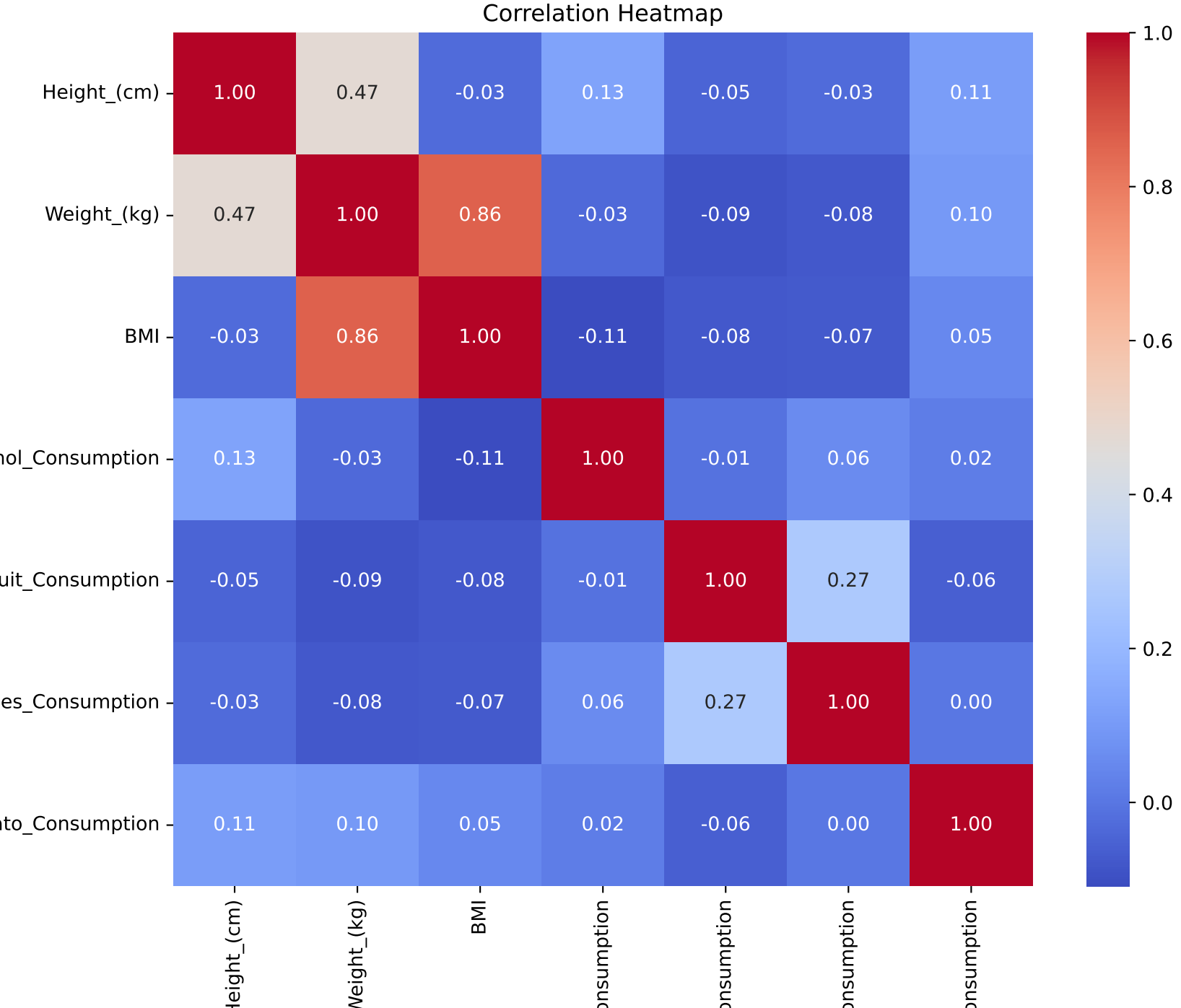

BMI Distribution
70000
60000
50000
40000
30000
20000
10000
0
Count
20
40
60
80
100
BMI

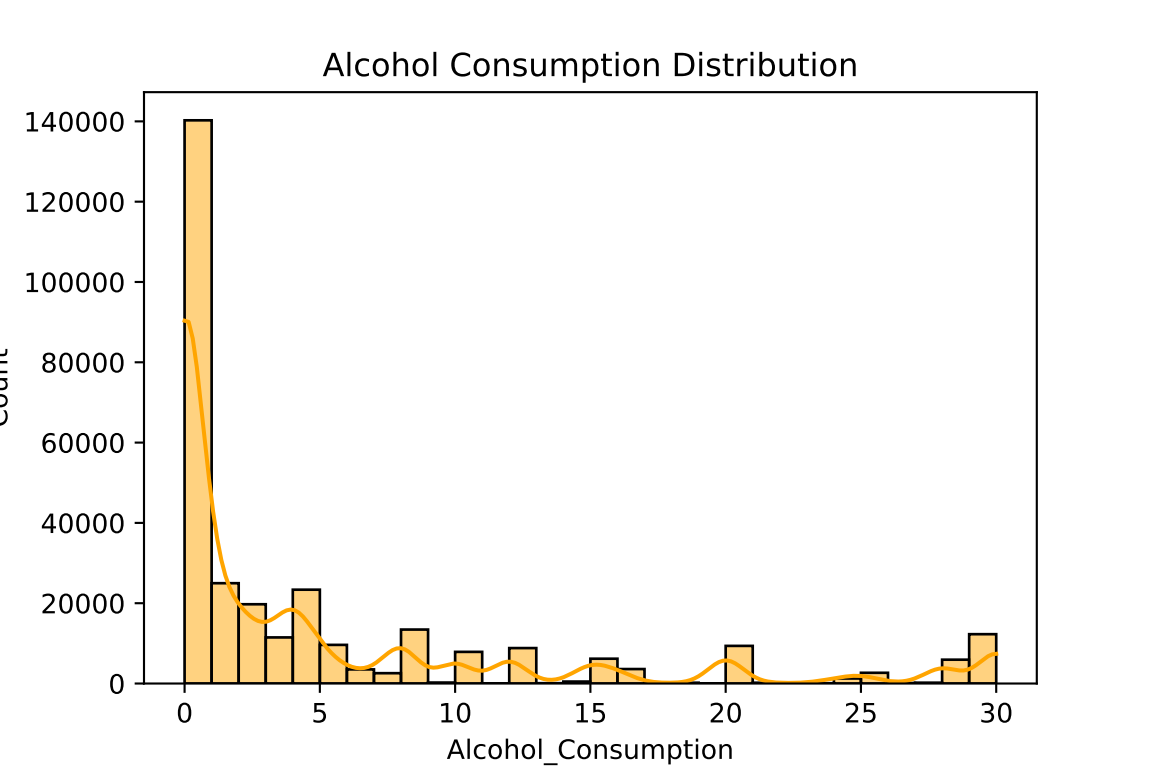

BMI by Heart Disease
BMI
100
80
60
40
20
No
Yes
Heart_Disease

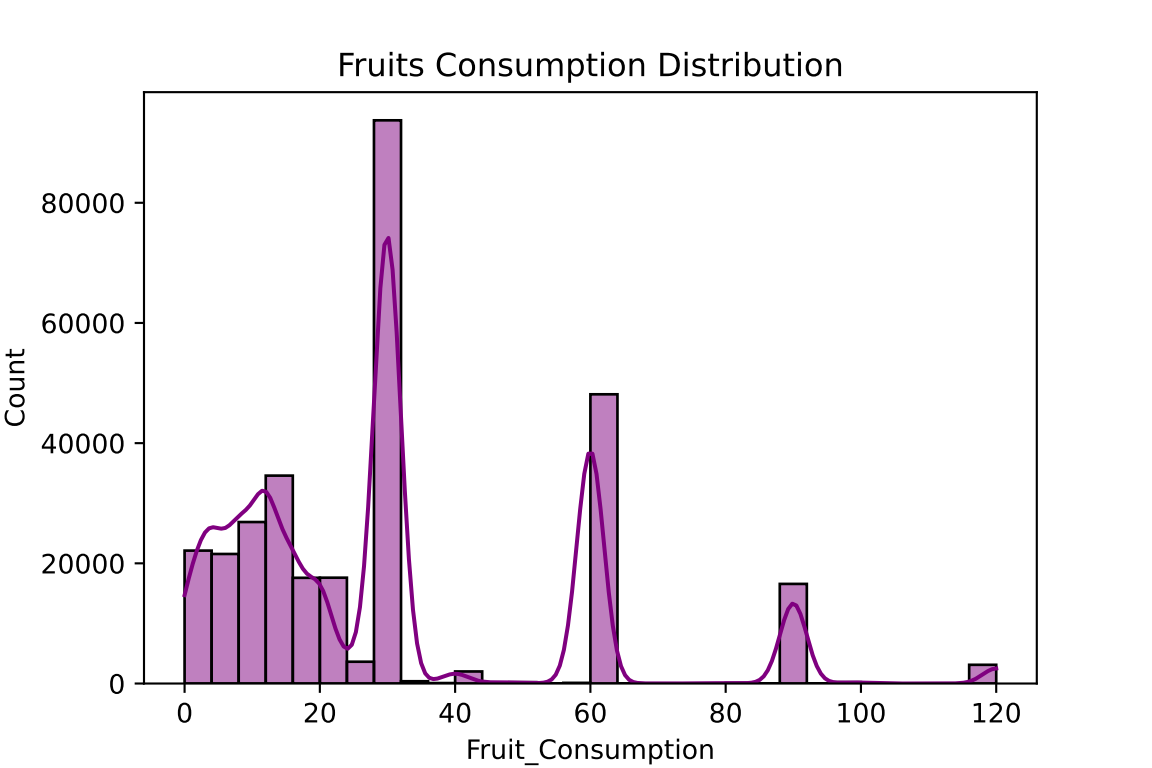

Sex vs Heart Disease
Heart_Disease
No
Yes
140000
120000
100000
80000
60000
40000
20000
0
Female
Male
Sex

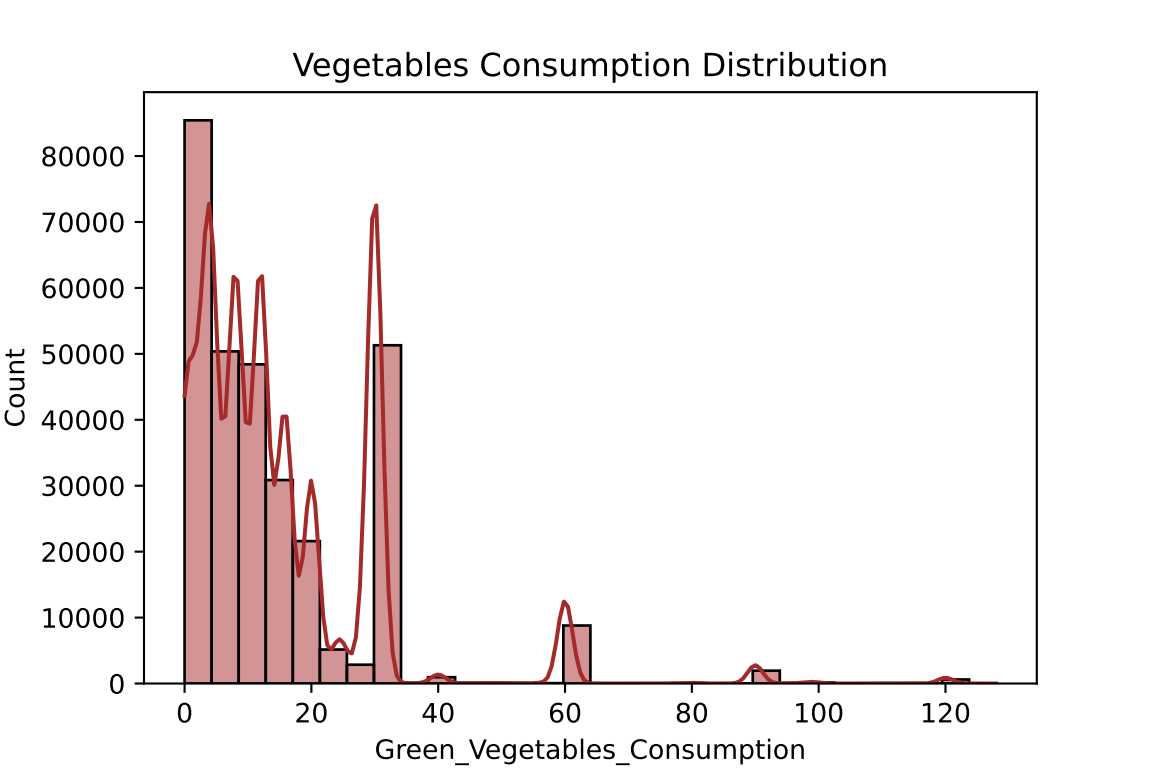

Exercise vs Heart Disease
Heart_Disease
No
Yes
200000
150000
100000
50000
0
No
Yes
Exercise

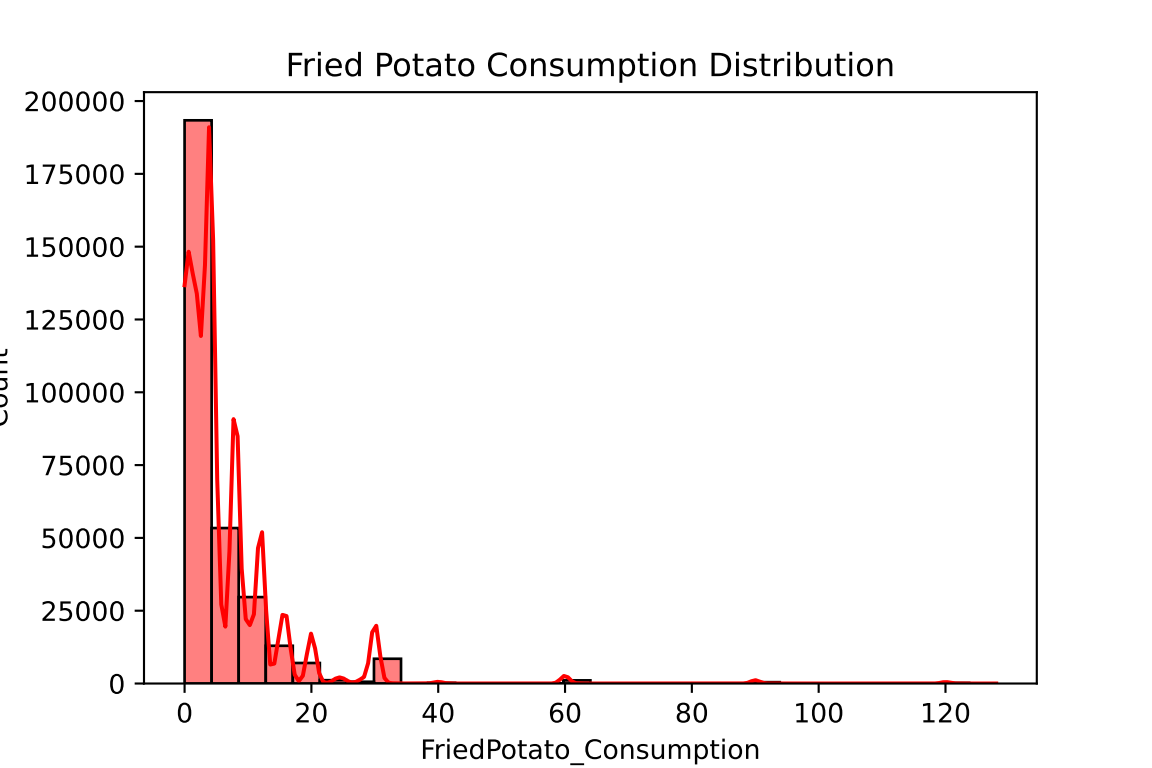

General Health vs Heart Disease
Heart_Disease
No
Yes
count
100000
80000
60000
40000
20000
0
Poor
Good
Good
Fair
ellent


Correlation Heatmap
Height_(cm)
Weight_(kg)
BMI
hol_Consumption
uit_Consumption
es_Consumption
to_Consumption
1.00 0.47 -0.03 0.13 -0.05 -0.03 0.11
0.47 1.00 0.86 -0.03 -0.09 -0.08 0.10
-0.03 0.86 1.00 -0.11 -0.08 -0.07 0.05
0.13 -0.03 -0.11 1.00 -0.01 0.06 0.02
-0.05 -0.09 -0.08 -0.01 1.00 0.27 -0.06
-0.03 -0.08 -0.07 0.06 0.27 1.00 0.00
0.11 0.10 0.05 0.02 -0.06 0.00 1.00
Height_(cm)
Weight_(kg)
BMI
onsumption
onsumption
onsumption
onsumption
1.0
0.8
0.6
0.4
0.2
0.0


Alcohol Consumption Distribution
140000
120000
100000
80000
60000
40000
20000
0
0
5
10
15
20
25
30
Alcohol_Consumption


Fruits Consumption Distribution
Count
80000
60000
40000
20000
0
0
20
40
60
80
100
120
Fruit_Consumption


Vegetables Consumption Distribution
Count
80000
70000
60000
50000
40000
30000
20000
10000
0
0
20
40
60
80
100
120
Green_Vegetables_Consumption


Fried Potato Consumption Distribution
200000
175000
150000
125000
100000
75000
50000
25000
0
Count
0
20
40
60
80
100
120
FriedPotato_Consumption