```python
! pip install pandas numpy matplotlib seaborn  plotly

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('Heart Disease data.csv')
df.head(10)
```

```
    age  sex  cp  trestbps  chol  fbs  restecg  thalach  exang  oldpeak
slope  \
0    52    1   0       125   212    0        1      168      0      1.0
2
1    53    1   0       140   203    1        0      155      1      3.1
0
2    70    1   0       145   174    0        1      125      1      2.6
0
3    61    1   0       148   203    0        1      161      0      0.0
2
4    62    0   0       138   294    1        1      106      0      1.9
1
5    58    0   0       100   248    0        0      122      0      1.0
1
6    58    1   0       114   318    0        2      140      0      4.4
0
7    55    1   0       160   289    0        0      145      1      0.8
1
8    46    1   0       120   249    0        0      144      0      0.8
2
9    54    1   0       122   286    0        0      116      1      3.2
1

   ca  thal  target
0   2     3       0
1   0     3       0
2   0     3       0
3   1     3       0
4   3     2       0
5   0     2       1
6   3     1       0
7   1     3       0
8   0     3       0
9   2     2       0
```

```python
df.columns
```

```
Index(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg',
'thalach',
```

```
        'exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'],
      dtype='object')

print(df.isnull().sum())

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

plt.figure(figsize=(8, 6))
sns.histplot(df['age'], bins=30, kde=True)
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.show()
```
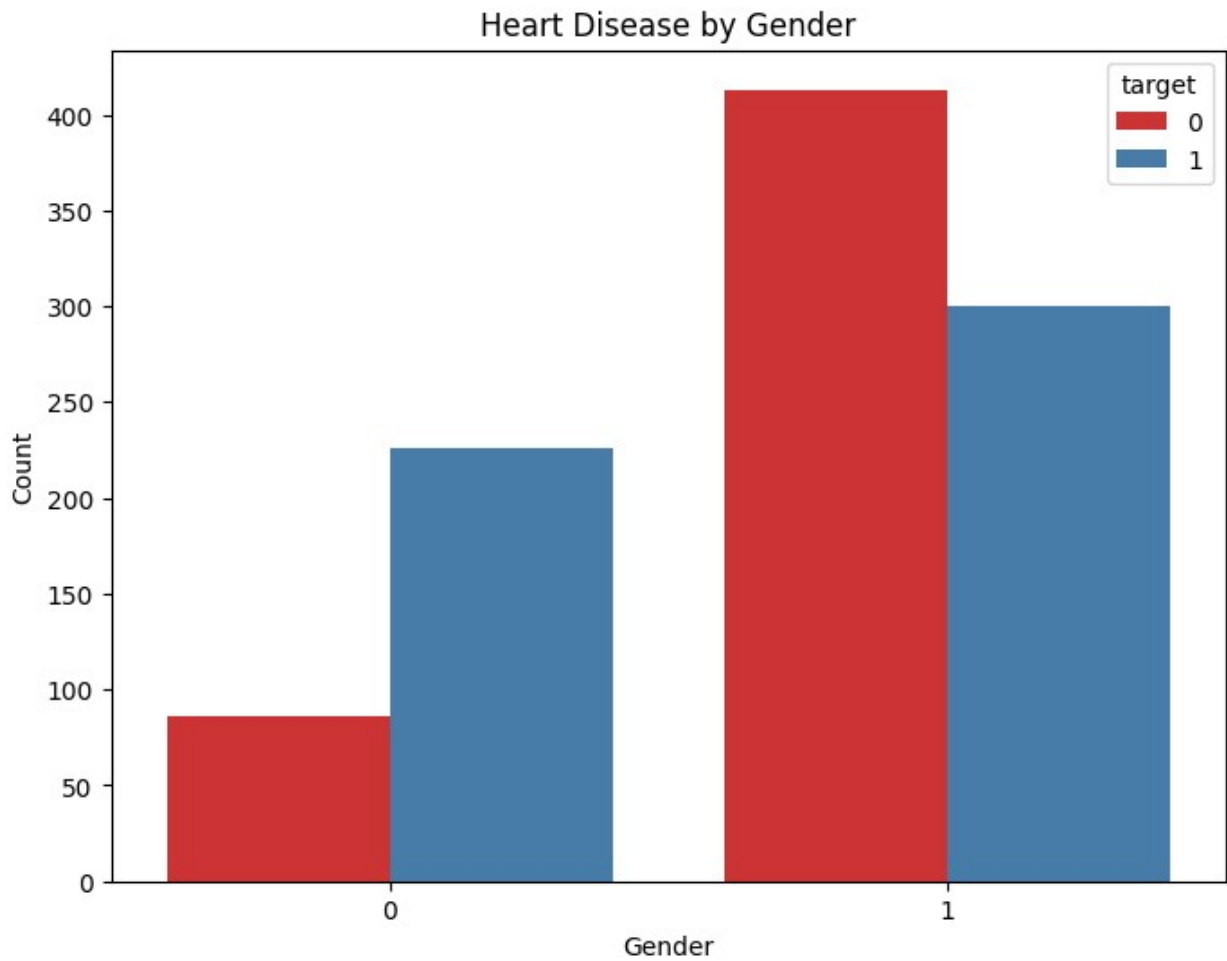
Age Distribution
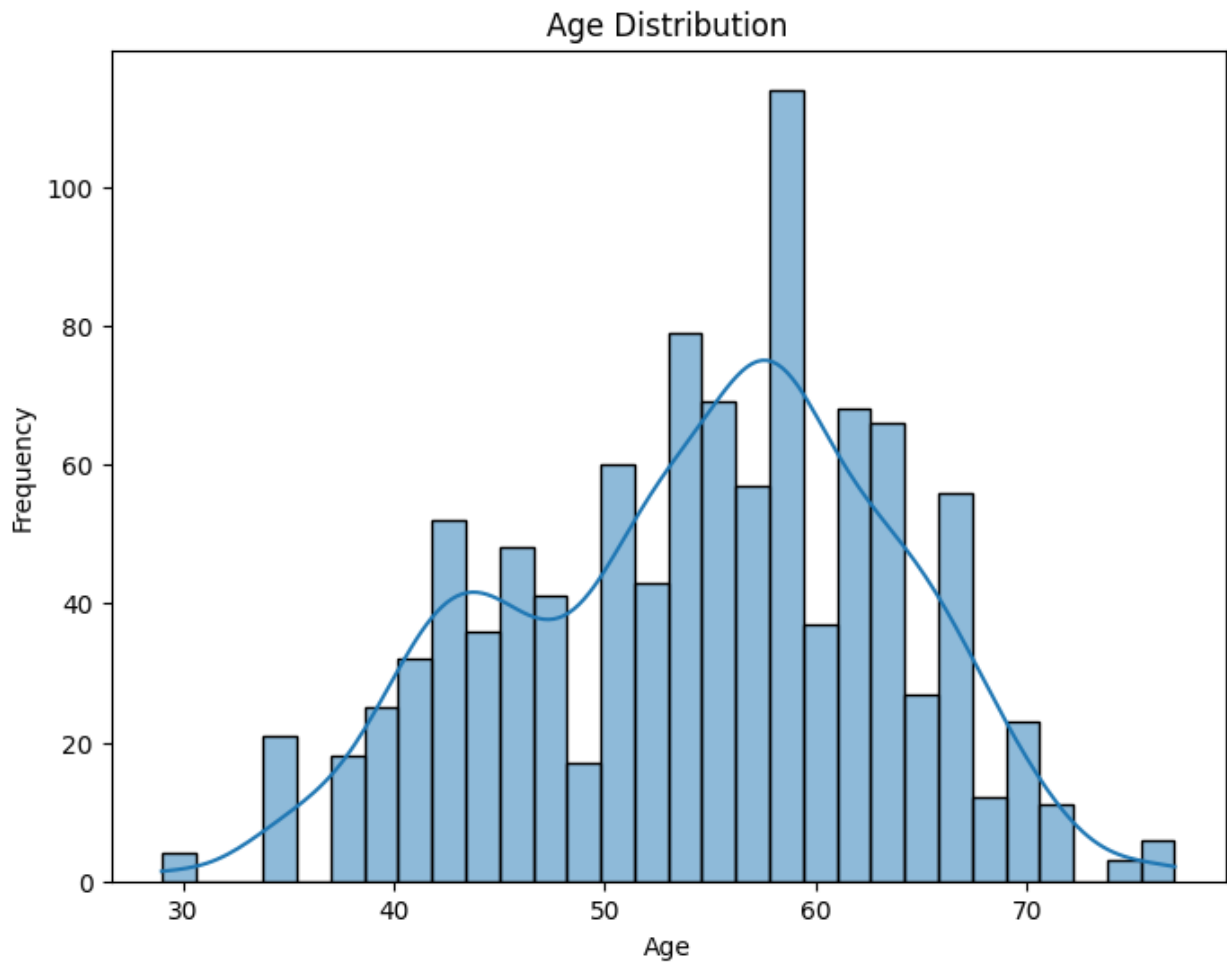
```
plt.figure(figsize=(8, 6))
sns.countplot(data=df, x='sex', hue='target', palette='Set1')
plt.title('Heart Disease by Gender')
plt.xlabel('Gender')
plt.ylabel('Count')
plt.show()
```

Heart Disease by Gender

```
plt.figure(figsize=(10, 8))
sns.countplot(data=df, x='age', hue='target', palette='Set2')
plt.title('Heart Disease by Age Group')
plt.xlabel('Age Group')
plt.ylabel('Count')
plt.show()
```

Heart Disease by Age Group
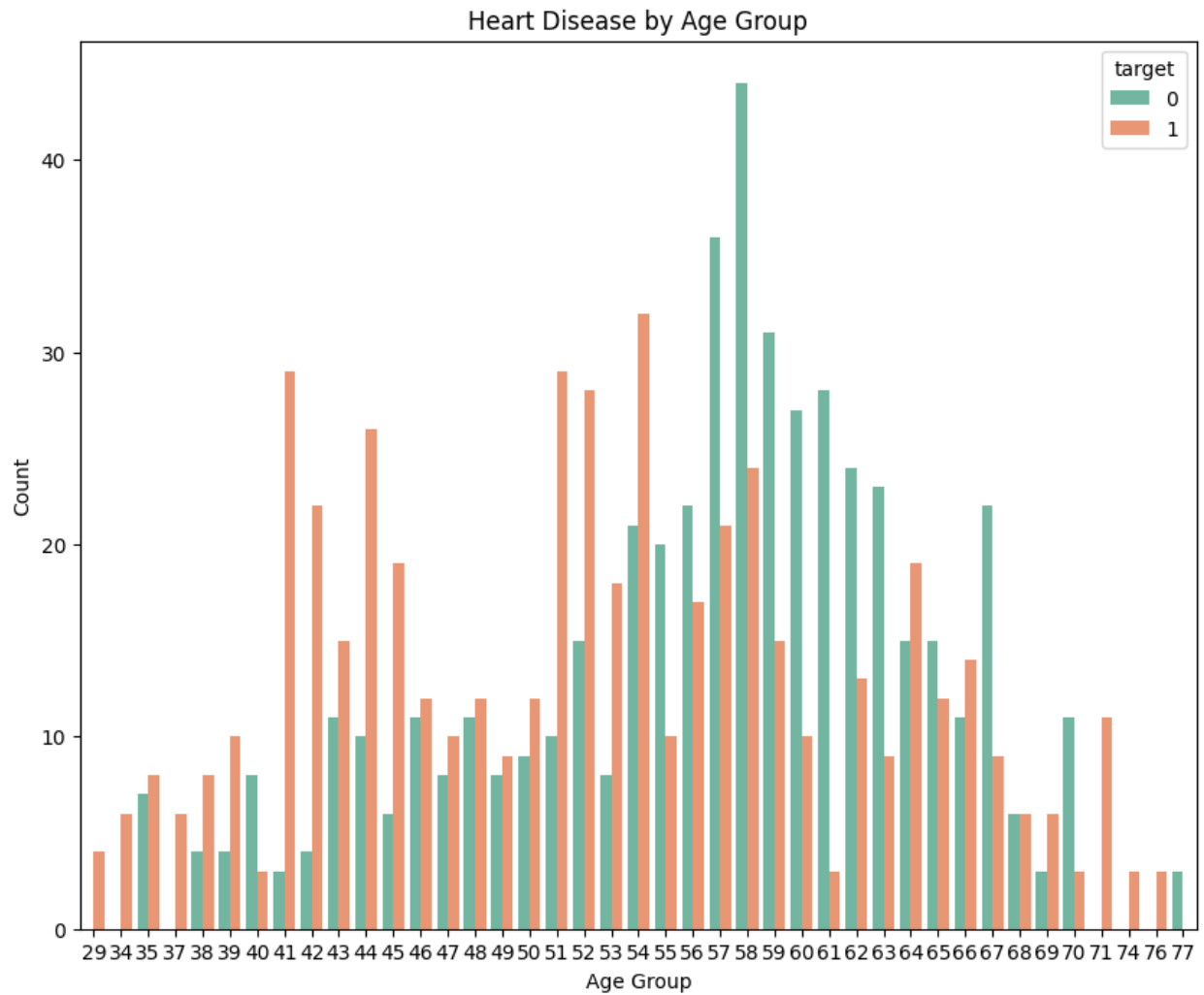
```
plt.figure(figsize=(10, 8))
corr_matrix = df.corr()
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt='.2f')
plt.title('Correlation Heatmap')
plt.show()
```

Correlation Heatmap
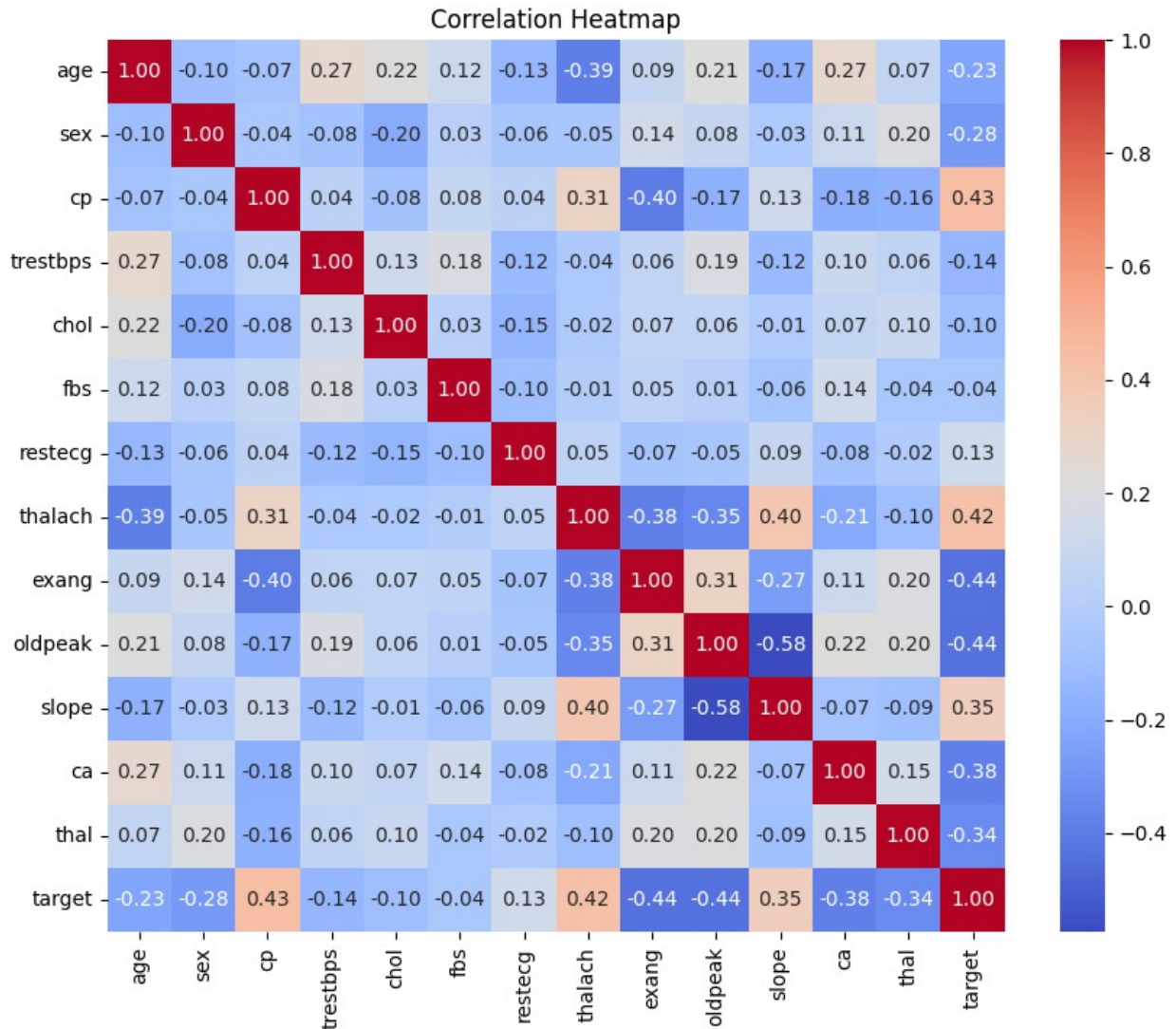

```
df['age_target'] = df['age'].apply(lambda x: 'Young' if x <= 50 else
'Old')
plt.figure(figsize=(8, 10))
sns.countplot(data=df, x='age_target', hue='target', palette='Set3')
plt.title('Heart Disease by Age (Young vs Old)')
plt.xlabel('Age Category')
plt.ylabel('Count')
plt.show()
```

Heart Disease by Age (Young vs Old)