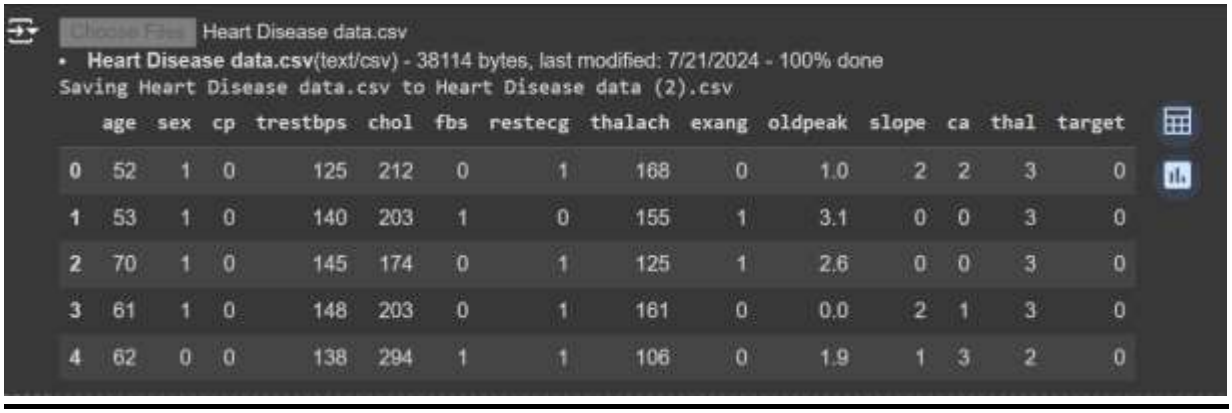
PYTHON CODE:

```
[9]  # Import necessary libraries
     import pandas as pd

     # Load the dataset
     from google.colab import files
     uploaded = files.upload()

     # Read the uploaded file
     data = pd.read_csv('Heart Disease data.csv')

     # Display the first few rows of the dataset
     data.head()
```



Choose Files Heart Disease data.csv

- **Heart Disease data.csv**(text/csv) - 38114 bytes, last modified: 7/21/2024 - 100% done

Saving Heart Disease data.csv to Heart Disease data (2).csv

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

```
     # Check for missing values
     missing_values = data.isnull().sum()
     print("Missing Values:\n", missing_values)
```

```
Missing Values:
 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
```

```python
# Fill missing values with the mean of the column
data.fillna(data.mean(), inplace=True)

missing_values_after = data.isnull().sum()
print("Missing Values after handling:\n", missing_values_after)
```

```
Missing Values after handling:
 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
```

```
data.dtypes

# Convert 'age' to integer if it is not already
data['age'] = data['age'].astype(int)

# Convert 'sex', 'cp', 'fbs', 'restecg', 'exang', 'slope', 'ca', 'thal', and 'target' to categorical if they are not already
categorical_columns = ['sex', 'cp', 'fbs', 'restecg', 'exang', 'slope', 'ca', 'thal', 'target']
for col in categorical_columns:
    data[col] = data[col].astype('category')

# Verify the data types again
data.dtypes
```

```
age              int64
sex           category
cp            category
trestbps         int64
chol             int64
fbs           category
restecg       category
thalach          int64
exang         category
oldpeak        float64
slope         category
ca            category
thal          category
target        category
dtype: object
```

```
data['age_category'] = pd.cut(data['age'], bins=[0, 30, 50, 70, 100], labels=['Young', 'Middle-aged', 'Senior', 'Old'])
```

```
[14] data.to_csv('Transformed_Heart_Disease_Data.csv', index=False)

     from google.colab import files
     files.download('Transformed_Heart_Disease_Data.csv')
```