```
   Age Sex ChestPainType  RestingBP  Cholesterol  FastingBS RestingECG  MaxHR  \
0   40   M           ATA        140          289          0     Normal    172
1   49   F           NAP        160          180          0     Normal    156
2   37   M           ATA        130          283          0         ST     98
3   48   F           ASY        138          214          0     Normal    108
4   54   M           NAP        150          195          0     Normal    122

  ExerciseAngina  Oldpeak ST_Slope  HeartDisease
0              N      0.0       Up             0
1              N      1.0     Flat             1
2              N      0.0       Up             0
3              Y      1.5     Flat             1
4              N      0.0       Up             0
Index(['Age', 'Sex', 'ChestPainType', 'RestingBP', 'Cholesterol', 'FastingBS',
       'RestingECG', 'MaxHR', 'ExerciseAngina', 'Oldpeak', 'ST_Slope',
       'HeartDisease'],
      dtype='object')
```

```
Age               0
Sex               0
ChestPainType     0
RestingBP         0
Cholesterol       0
FastingBS         0
RestingECG        0
MaxHR             0
ExerciseAngina    0
Oldpeak           0
ST_Slope          0
HeartDisease      0
dtype: int64
```

	Age	Sex	RestingBP	Cholesterol	FastingBS	MaxHR	ExerciseAngina	Oldpeak	HeartDisease	ChestPainType_ASY	ChestPainType_ATA	ChestPainType_NAP	ChestPainType_TA	RestingECG_LVH	RestingECG_Normal	RestingECG_ST	ST_Slope_Down	ST_Slope_Flat	ST_Slope_Up
0	40	1	140	289	0	172	0	0.0	0	False	True	False	False	False	True	False	False	False	True
1	49	0	160	180	0	156	0	1.0	1	False	False	True	False	False	True	False	False	True	False
2	37	1	130	283	0	98	0	0.0	0	False	True	False	False	False	False	True	False	False	True
3	48	0	138	214	0	108	1	1.5	1	True	False	False	False	False	True	False	False	True	False
4	54	1	150	195	0	122	0	0.0	0	False	False	True	False	False	True	False	False	False	True

Chrome File Edit View History Bookmarks Profiles Tab Window Help Fri Jun 28 2:05 PM

HeartDiseasePrediction.ipynb | Connect Job | (7) WhatsApp

colab.research.google.com/drive/1Th1xgZUJLJWNgTogxaQFI3pA8WMjXLxw#scrollTo=DFrB1INLxMxR

Relaunch to update

+ Code + Text

```
# Display the cleaned dataset
df.head()
```

```
Age                 0
Sex                 0
RestingBP           0
Cholesterol         0
FastingBS           0
MaxHR               0
ExerciseAngina      0
Oldpeak             0
HeartDisease        0
ChestPainType_ASY   0
ChestPainType_ATA   0
ChestPainType_NAP   0
ChestPainType_TA    0
RestingECG_LVH      0
RestingECG_Normal   0
RestingECG_ST       0
ST_Slope_Down       0
ST_Slope_Flat       0
ST_Slope_Up         0
dtype: int64
```

	Age	Sex	RestingBP	Cholesterol	FastingBS	MaxHR	ExerciseAngina	Oldpeak	HeartDisease	ChestPainType_ASY	ChestPainType_ATA	ChestPainType_NAP	ChestPainType_TA
0	40	1	140	289	0	172	0	0.0	0	False	True	False	False
1	49	0	160	180	0	156	0	1.0	1	False	False	True	False
2	37	1	130	283	0	98	0	0.0	0	False	True	False	False
3	48	0	138	214	0	108	1	1.5	1	True	False	False	False
4	54	1	150	195	0	122	0	0.0	0	False	False	True	False

```
[ ] # Summary statistics
    print(df.describe())
```

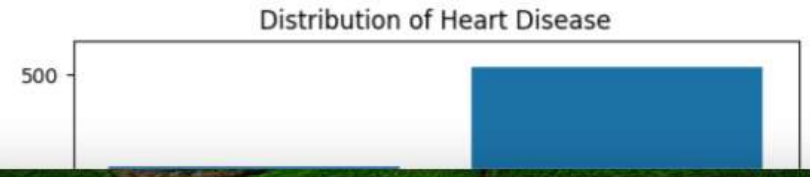
```
plt.show()

# Correlation matrix
plt.figure(figsize=(12, 8))
sns.heatmap(df.corr(), annot=True, fmt='.2f', cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()

# Pairplot
sns.pairplot(df, hue='HeartDisease')
plt.show()
```

```
              Age         Sex   RestingBP  Cholesterol   FastingBS  \
count  918.000000  918.000000  918.000000   918.000000  918.000000
mean    53.510893    0.789760  132.396514   198.799564    0.233115
std      9.432617    0.407701   18.514154   109.384145    0.423046
min     28.000000    0.000000    0.000000     0.000000    0.000000
25%     47.000000    1.000000  120.000000   173.250000    0.000000
50%     54.000000    1.000000  130.000000   223.000000    0.000000
75%     60.000000    1.000000  140.000000   267.000000    0.000000
max     77.000000    1.000000  200.000000   603.000000    1.000000

            MaxHR  ExerciseAngina     Oldpeak  HeartDisease
count  918.000000      918.000000  918.000000    918.000000
mean   136.809368        0.404139    0.887364      0.553377
std     25.460334        0.490992    1.066570      0.497414
min     60.000000        0.000000   -2.600000      0.000000
25%    120.000000        0.000000    0.000000      0.000000
50%    138.000000        0.000000    0.600000      1.000000
75%    156.000000        1.000000    1.500000      1.000000
max    202.000000        1.000000    6.200000      1.000000
```



Chrome File Edit View History Bookmarks Profiles Tab Window Help Fri Jun 28 2:05 PM

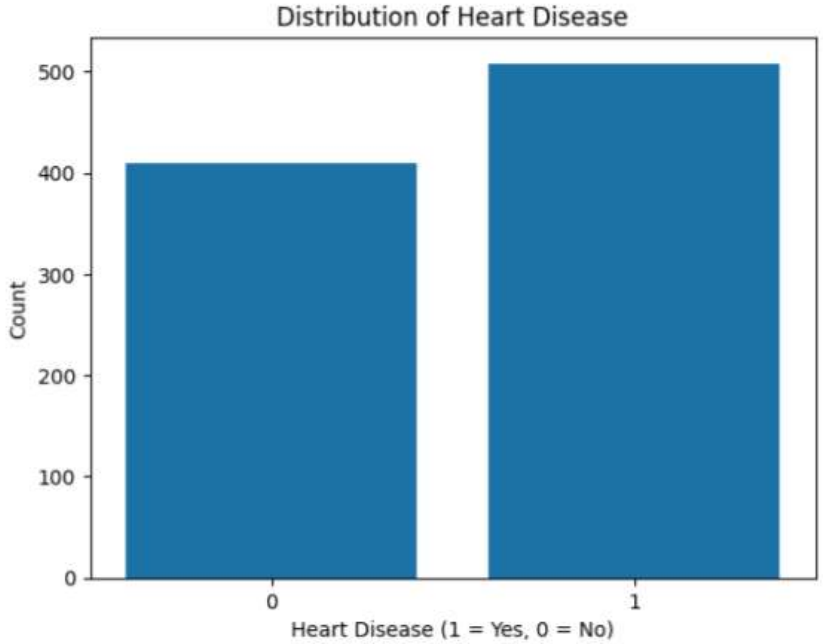
HeartDiseasePrediction.ipynb | Connect Job | (7) WhatsApp

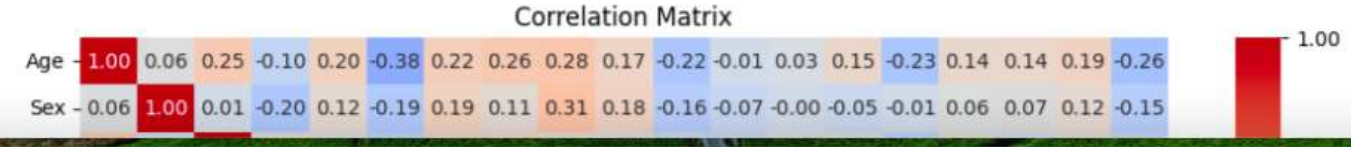
colab.research.google.com/drive/1Th1xgZUJLJWNgTogxaQFI3pA8WMjXLxw#scrollTo=DFrB1INLxMxR

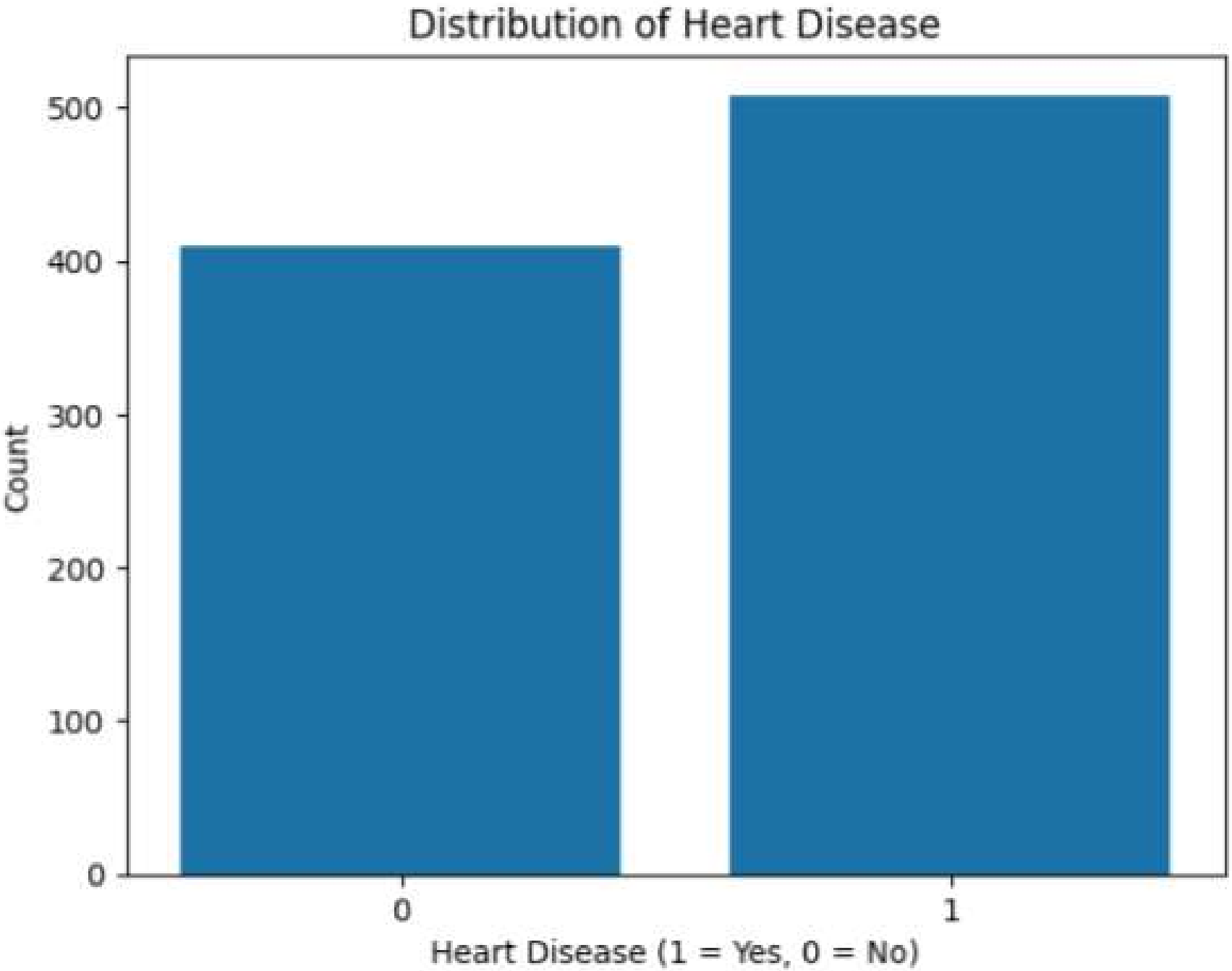
Relaunch to update

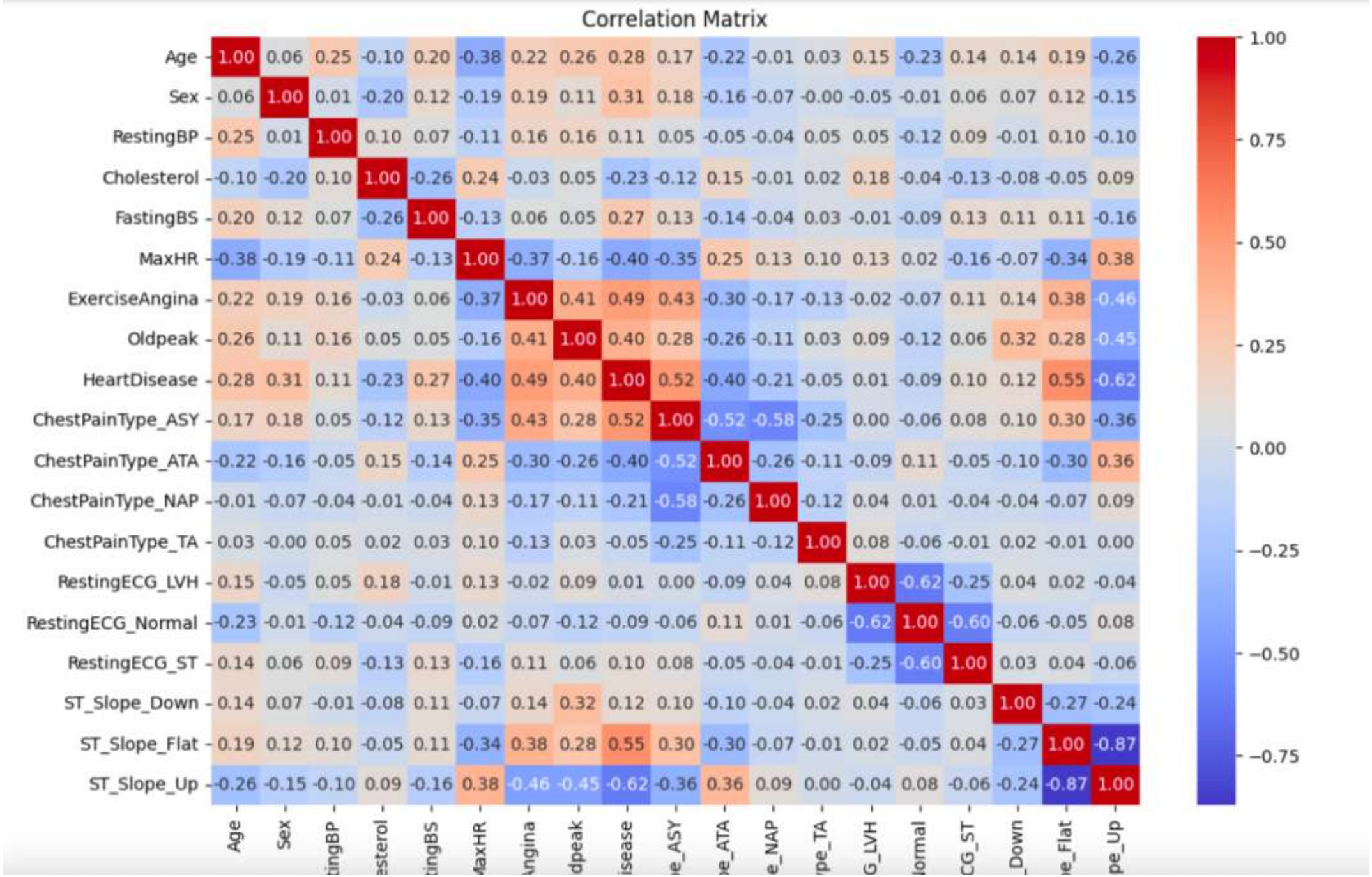
+ Code + Text

	MaxHR	ExerciseAngina	Oldpeak	HeartDisease
count	918.000000	918.000000	918.000000	918.000000
mean	136.809368	0.404139	0.887364	0.553377
std	25.460334	0.490992	1.066570	0.497414
min	60.000000	0.000000	-2.600000	0.000000
25%	120.000000	0.000000	0.000000	0.000000
50%	138.000000	0.000000	0.600000	1.000000
75%	156.000000	1.000000	1.500000	1.000000
max	202.000000	1.000000	6.200000	1.000000






Distribution of Heart Disease
500
400
300
200
100
0
Count
0
1
Heart Disease (1 = Yes, 0 = No)


Correlation Matrix
Age 1.00 0.06 0.25 -0.10 0.20 -0.38 0.22 0.26 0.28 0.17 -0.22 -0.01 0.03 0.15 -0.23 0.14 0.14 0.19 -0.26
Sex 0.06 1.00 0.01 -0.20 0.12 -0.19 0.19 0.11 0.31 0.18 -0.16 -0.07 -0.00 -0.05 -0.01 0.06 0.07 0.12 -0.15
RestingBP 0.25 0.01 1.00 0.10 0.07 -0.11 0.16 0.16 0.11 0.05 -0.05 -0.04 0.05 0.05 -0.12 0.09 -0.01 0.10 -0.10
Cholesterol -0.10 -0.20 0.10 1.00 -0.26 0.24 -0.03 0.05 -0.23 -0.12 0.15 -0.01 0.02 0.18 -0.04 -0.13 -0.08 -0.05 0.09
FastingBS 0.20 0.12 0.07 -0.26 1.00 -0.13 0.06 0.05 0.27 0.13 -0.14 -0.04 0.03 -0.01 -0.09 0.13 0.11 0.11 -0.16
MaxHR -0.38 -0.19 -0.11 0.24 -0.13 1.00 -0.37 -0.16 -0.40 -0.35 0.25 0.13 0.10 0.13 0.02 -0.16 -0.07 -0.34 0.38
ExerciseAngina 0.22 0.19 0.16 -0.03 0.06 -0.37 1.00 0.41 0.49 0.43 -0.30 -0.17 -0.13 -0.02 -0.07 0.11 0.14 0.38 -0.46
Oldpeak 0.26 0.11 0.16 0.05 0.05 -0.16 0.41 1.00 0.40 0.28 -0.26 -0.11 0.03 0.09 -0.12 0.06 0.32 0.28 -0.45
HeartDisease 0.28 0.31 0.11 -0.23 0.27 -0.40 0.49 0.40 1.00 0.52 -0.40 -0.21 -0.05 0.01 -0.09 0.10 0.12 0.55 -0.62
ChestPainType_ASY 0.17 0.18 0.05 -0.12 0.13 -0.35 0.43 0.28 0.52 1.00 -0.52 -0.58 -0.25 0.00 -0.06 0.08 0.10 0.30 -0.36
ChestPainType_ATA -0.22 -0.16 -0.05 0.15 -0.14 0.25 -0.30 -0.26 -0.40 -0.52 1.00 -0.26 -0.11 -0.09 0.11 -0.05 -0.10 -0.30 0.36
ChestPainType_NAP -0.01 -0.07 -0.04 -0.01 -0.04 0.13 -0.17 -0.11 -0.21 -0.58 -0.26 1.00 -0.12 0.04 0.01 -0.04 -0.04 -0.07 0.09
ChestPainType_TA 0.03 -0.00 0.05 0.02 0.03 0.10 -0.13 0.03 -0.05 -0.25 -0.11 -0.12 1.00 0.08 -0.06 -0.01 0.02 -0.01 0.00
RestingECG_LVH 0.15 -0.05 0.05 0.18 -0.01 0.13 -0.02 0.09 0.01 0.00 -0.09 0.04 0.08 1.00 -0.62 -0.25 0.04 0.02 -0.04
RestingECG_Normal -0.23 -0.01 -0.12 -0.04 -0.09 0.02 -0.07 -0.12 -0.09 -0.06 0.11 0.01 -0.06 -0.62 1.00 -0.60 -0.06 -0.05 0.08
RestingECG_ST 0.14 0.06 0.09 -0.13 0.13 -0.16 0.11 0.06 0.10 0.08 -0.05 -0.04 -0.01 -0.25 -0.60 1.00 0.03 0.04 -0.06
ST_Slope_Down 0.14 0.07 -0.01 -0.08 0.11 -0.07 0.14 0.32 0.12 0.10 -0.10 -0.04 0.02 0.04 -0.06 0.03 1.00 -0.27 -0.24
ST_Slope_Flat 0.19 0.12 0.10 -0.05 0.11 -0.34 0.38 0.28 0.55 0.30 -0.30 -0.07 -0.01 0.02 -0.05 0.04 -0.27 1.00 -0.87
ST_Slope_Up -0.26 -0.15 -0.10 0.09 -0.16 0.38 -0.46 -0.45 -0.62 -0.36 0.36 0.09 0.00 -0.04 0.08 -0.06 -0.24 -0.87 1.00
Age
Sex
RestingBP
Cholesterol
FastingBS
MaxHR
ExerciseAngina
Oldpeak
HeartDisease
ChestPainType_ASY
ChestPainType_ATA
ChestPainType_NAP
ChestPainType_TA
RestingECG_LVH
RestingECG_Normal
RestingECG_ST
ST_Slope_Down
ST_Slope_Flat
ST_Slope_Up
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75

```
Accuracy: 0.85
Confusion Matrix:
[[67 10]
 [17 90]]
Classification Report:
              precision    recall  f1-score   support

           0       0.80      0.87      0.83        77
           1       0.90      0.84      0.87       107

    accuracy                           0.85       184
   macro avg       0.85      0.86      0.85       184
weighted avg       0.86      0.85      0.85       184
```

Chrome File Edit View History Bookmarks Profiles Tab Window Help Fri Jun 28 2:03 PM

HeartDiseasePrediction.ipynb | Connect Job | (7) WhatsApp

colab.research.google.com/drive/1Th1xgZUJLJWNgTogxaQFI3pA8WMjXLxw#scrollTo=Zn8Q2_v_7iYF

Relaunch to update

HeartDiseasePrediction.ipynb

File Edit View Insert Runtime Tools Help

Comment Share

+ Code + Text Connect Gemini

```
[ ] # Display the first few rows of the dataset
    print(df.head())

    # Print out the column names
    print(df.columns)
```

```
   Age Sex ChestPainType  RestingBP  Cholesterol  FastingBS RestingECG  MaxHR  \
0   40   M           ATA        140          289          0     Normal    172
1   49   F           NAP        160          180          0     Normal    156
2   37   M           ATA        130          283          0         ST     98
3   48   F           ASY        138          214          0     Normal    108
4   54   M           NAP        150          195          0     Normal    122

  ExerciseAngina  Oldpeak ST_Slope  HeartDisease
0              N      0.0       Up             0
1              N      1.0     Flat             1
2              N      0.0       Up             0
3              Y      1.5     Flat             1
4              N      0.0       Up             0
Index(['Age', 'Sex', 'ChestPainType', 'RestingBP', 'Cholesterol', 'FastingBS',
       'RestingECG', 'MaxHR', 'ExerciseAngina', 'Oldpeak', 'ST_Slope',
       'HeartDisease'],
      dtype='object')
```

ETL (Extract, Transform, Load)

```
[ ] # Checking for missing values
    missing_values = df.isnull().sum()
    print(missing_values)
```