

The background of the slide is an abstract, vibrant composition of numerous 3D rectangular blocks or cubes. These blocks are arranged in a way that creates a sense of depth and movement, with some blocks appearing to rise or fall. The color palette is diverse, featuring warm tones like yellows, oranges, and reds on the left side, transitioning through browns and purples in the center, and cooler tones like blues and greens on the right side. The lighting is dramatic, with strong highlights and shadows that emphasize the three-dimensional nature of the blocks.

# *Coronary Artery Disease (CAD)*

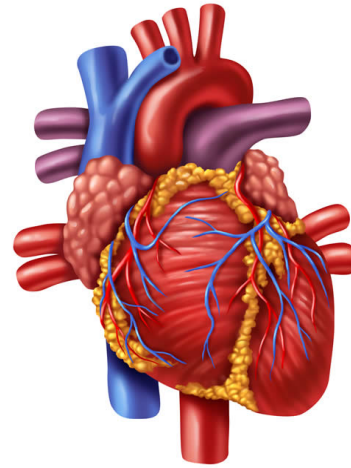
Henry Friederichs

# ***Content***

**Introduction**

**Analysis**


**Conclusion**




# ***INTRODUCTION***

# ER Diagram

EXERCISE	
Exercise_ID	
Exang	12
Oldpeak	13
Slope	14


 Dataset



4051

## Heart Disease UCI

<https://archive.ics.uci.edu/ml/datasets/Heart+Disease>


 ronit


updated 2 years ago (Version 1)

Data

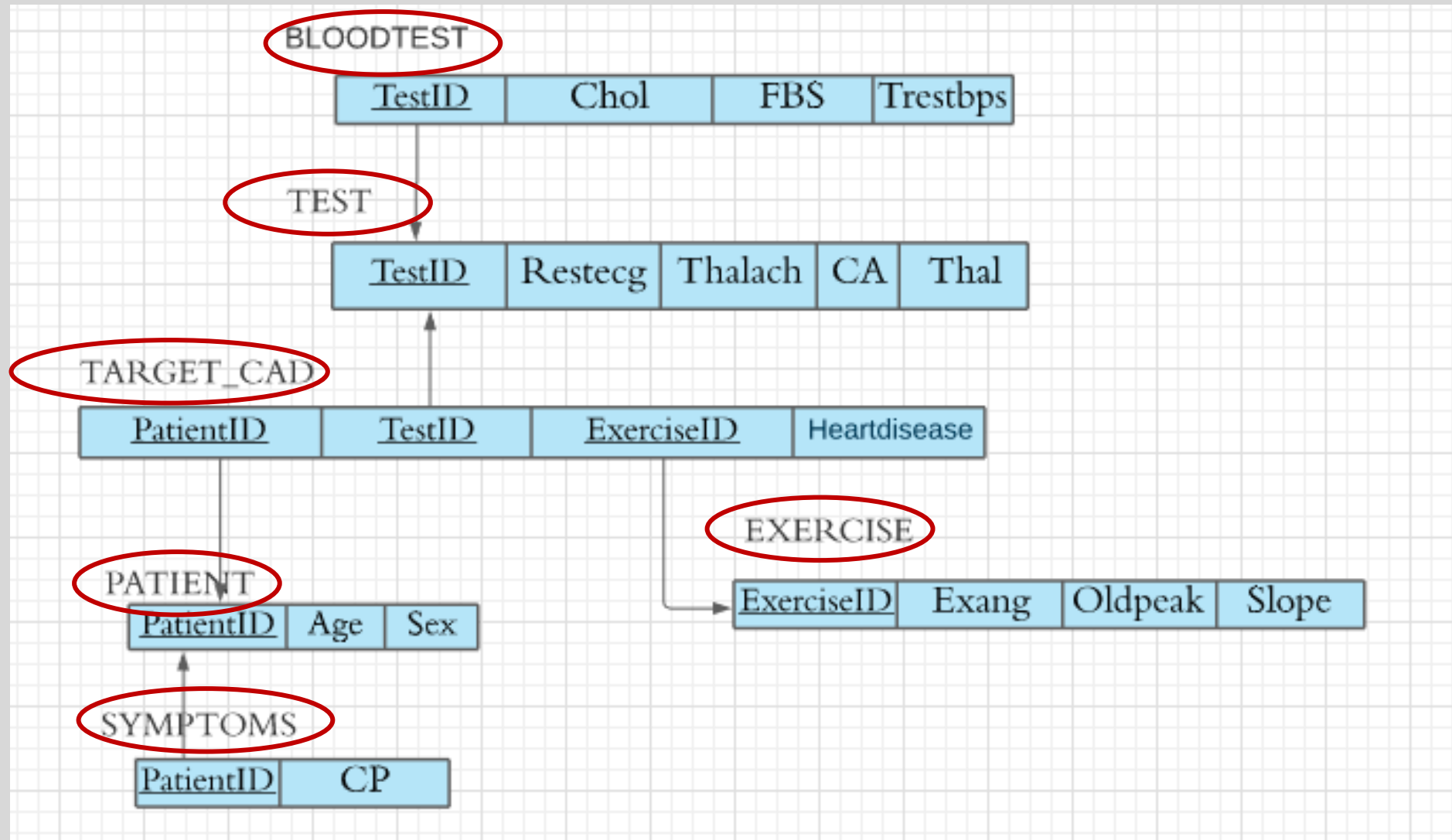
Tasks (11)Notebooks (1,360)Discussion (61)Activity

Usability 7.6

 License Reddit API Terms

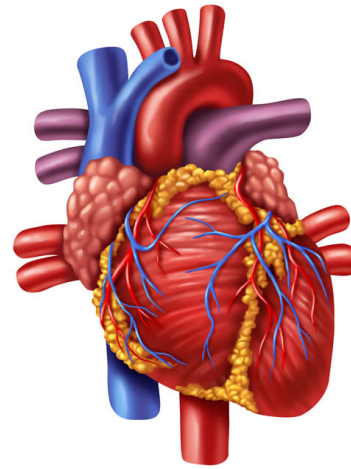
 Tags health, biology, classification, heart conditions, binary classification

# Relational Schema



# ***Questions of Interest***

- 1. What is the age/gender distribution of these patients?
- 2. What is the correlation between these test results?
- 3. what is the most common symptom that these patients have?
- 4. Which test results of each test table can be used as indicators for CAD patients?



***ANALYSIS***

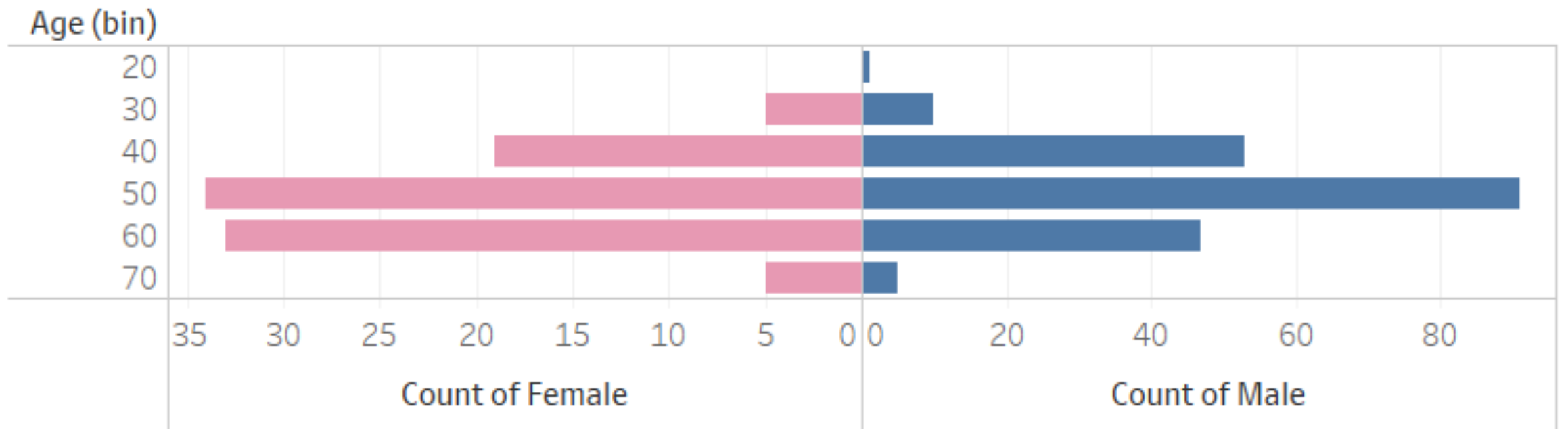
# ***The Proportion of CAD Patients and Non-CAD Patients***

```
SELECT If_Patients_Are_Diagnosed_CAD, COUNT(*) AS Number_of_Patients
FROM (SELECT CASE
        WHEN TA.HEARTDISEASE = 1 THEN 'CAD Patients'
        WHEN TA.HEARTDISEASE = 0 THEN 'Non-CAD Patients'
        END AS If_Patients_Are_Diagnosed_CAD
      FROM TARGET_CAD TA)
GROUP BY If_Patients_Are_Diagnosed_CAD
```



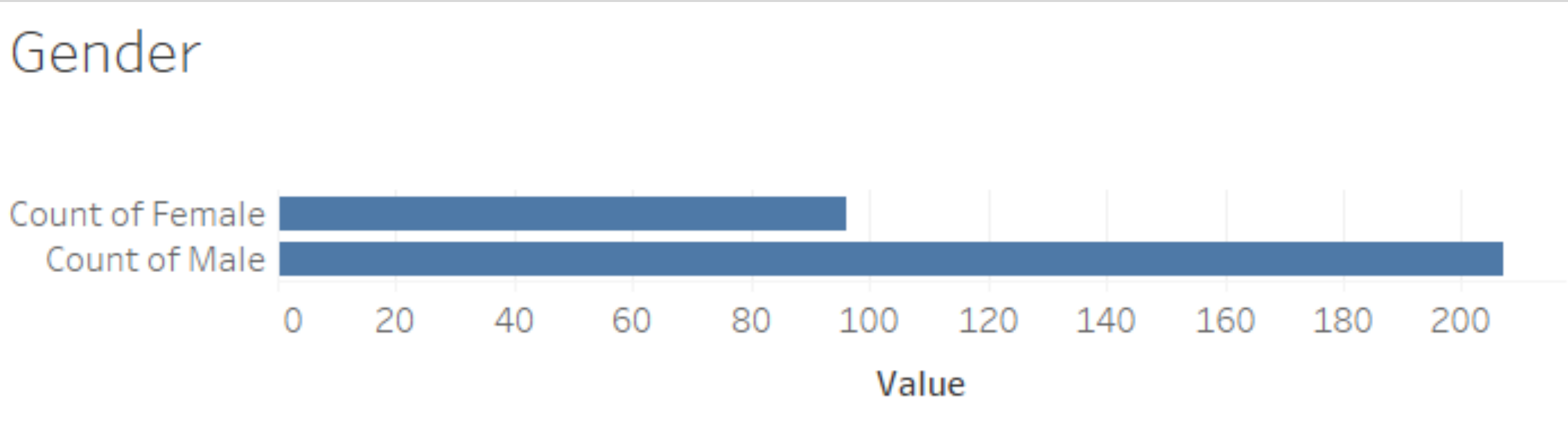
*1. What is the age/gender distribution of these patients?*

## Population Pyramid

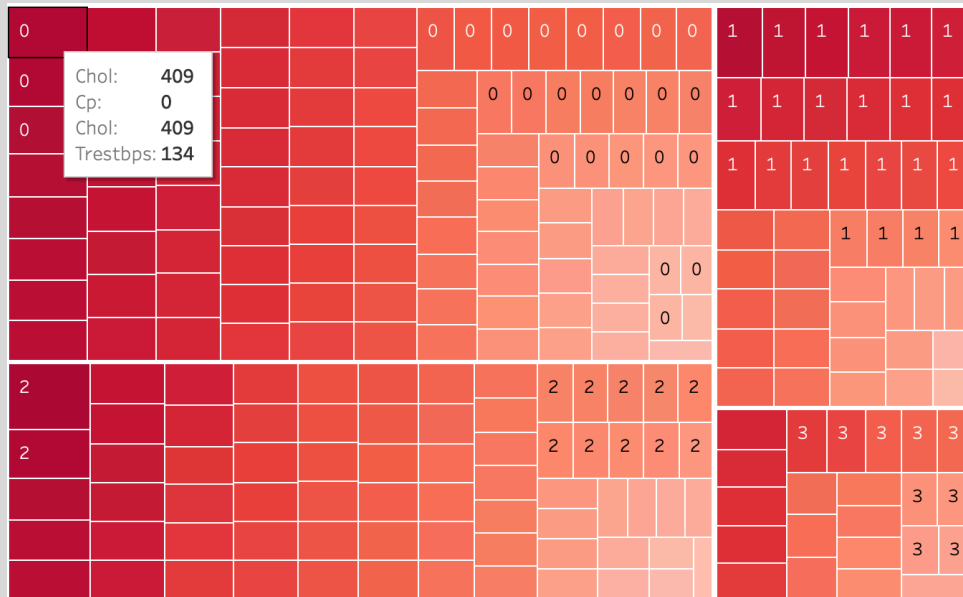


***Population Description***

# Gender



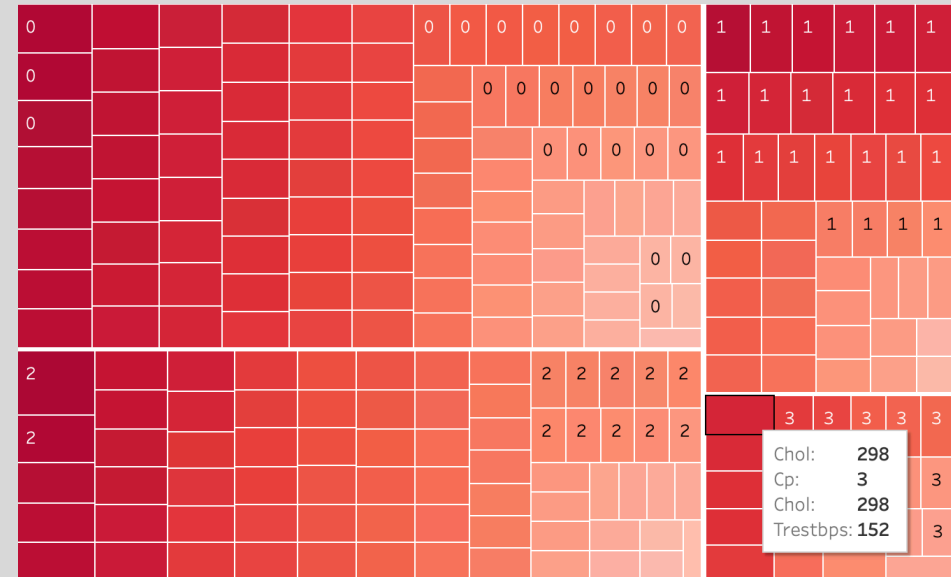
## 2. What is the correlation between these test results?



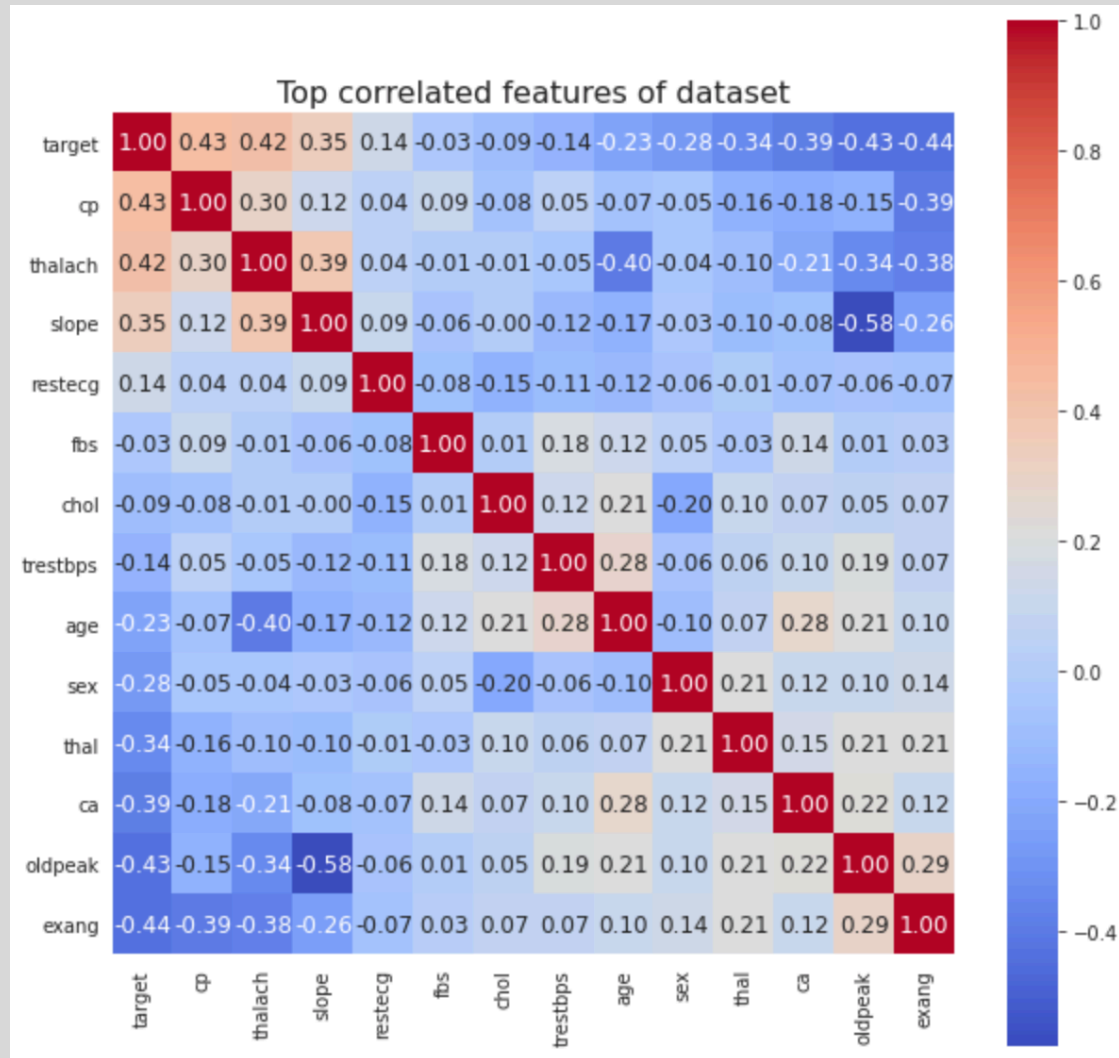
**Correlation**

### Chest Pain Type

- Value 0: typical angina
- Value 1: atypical angina
- Value 2: non-anginal pain
- Value 3: asymptomatic



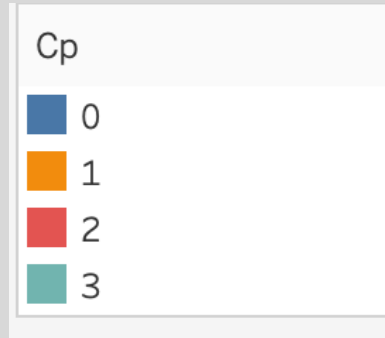
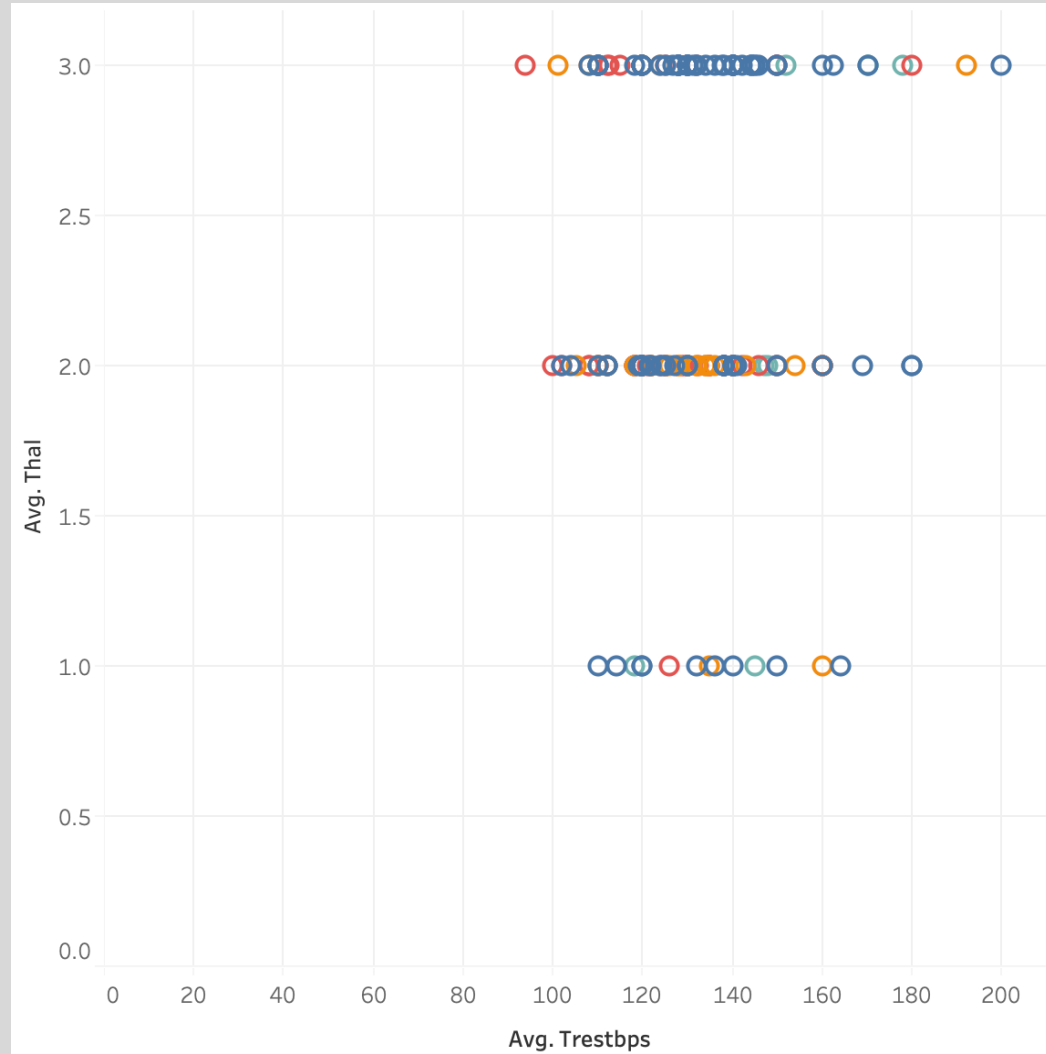
# Correlation Matrix



# ***Python code for heat map correlation***

```
top = 15
corr = data.corr()
top15 = corr.nlargest(top, 'target')['target'].index
corr_top15 = data[top15].corr()
f,ax = plt.subplots(figsize=(10,10))
sns.heatmap(corr_top15, square=True, ax=ax, annot=True, cmap='coolwarm', fmt='.2f',
annot_kws={'size':12})
plt.title('Top correlated features of dataset', size=16)
plt.show()
```

# *Maximum Heart Rate Correlation*



### 3. what is the most common symptom that these patients have?

```
SELECT Chest_Pain_Types, COUNT(*) AS Number_of_NonCAD_Patients
FROM (SELECT CASE
      WHEN CP = 1 THEN 'Typical angina'
      WHEN CP = 2 THEN 'Atypical angina'
      WHEN CP = 3 THEN 'Non-anginal pain'
      WHEN CP = 0 THEN 'Asymptomatic'
      END AS Chest_Pain_Types
      FROM SYMPTOMS S, TARGET_CAD T
      WHERE S.PATIENT_ID = T.PATIENT_ID
      AND T.HEARTDISEASE = 0)
GROUP BY Chest_Pain_Types
ORDER BY Number_of_NonCAD_Patients DESC
```

	Asymptomatic
<pre>SELECT Chest_Pain_Types, COUNT(*) AS Number_of_CAD_Patients FROM (SELECT CASE       WHEN CP = 1 THEN 'Typical angina'       WHEN CP = 2 THEN 'Atypical angina'       WHEN CP = 3 THEN 'Non-anginal pain'       WHEN CP = 0 THEN 'Asymptomatic'       END AS Chest_Pain_Types       FROM SYMPTOMS S, TARGET_CAD T       WHERE S.PATIENT_ID = T.PATIENT_ID       AND T.HEARTDISEASE = 1) GROUP BY Chest_Pain_Types ORDER BY Number_of_CAD_Patients DESC</pre>	

CAD patients:  
Atypical angina &  
Typical angina

Non-CAD patients:  
Asymptomatic

# Bloodtest Table of Asymptomatic CAD Patients

Cholesterol

```
SELECT cholesterol_Level, COUNT(*) AS NUM_Asymptomatic_CAD_Patients
FROM (SELECT CASE
      WHEN CHOL >= 240 THEN 'High: >= 240'
      WHEN CHOL >= 200 THEN 'Borderline high: >= 200'
      ELSE 'Normal:< 200'
      END AS cholesterol_Level
      FROM SYMPTOMS S, TARGET_CAD T, BLOODTEST B
      WHERE S.PATIENT_ID = T.PATIENT_ID
      AND B.TESTID = T.TESTID
      AND T.HEARTDISEASE = 1
      AND CP = 0)
GROUP BY cholesterol_Level
ORDER BY NUM_Asymptomatic_CAD_Patients DESC
```

Normal blood sugar: <= 120

97%

- People who are at high risk of cad regularly monitor their cholesterol and blood pressure level.



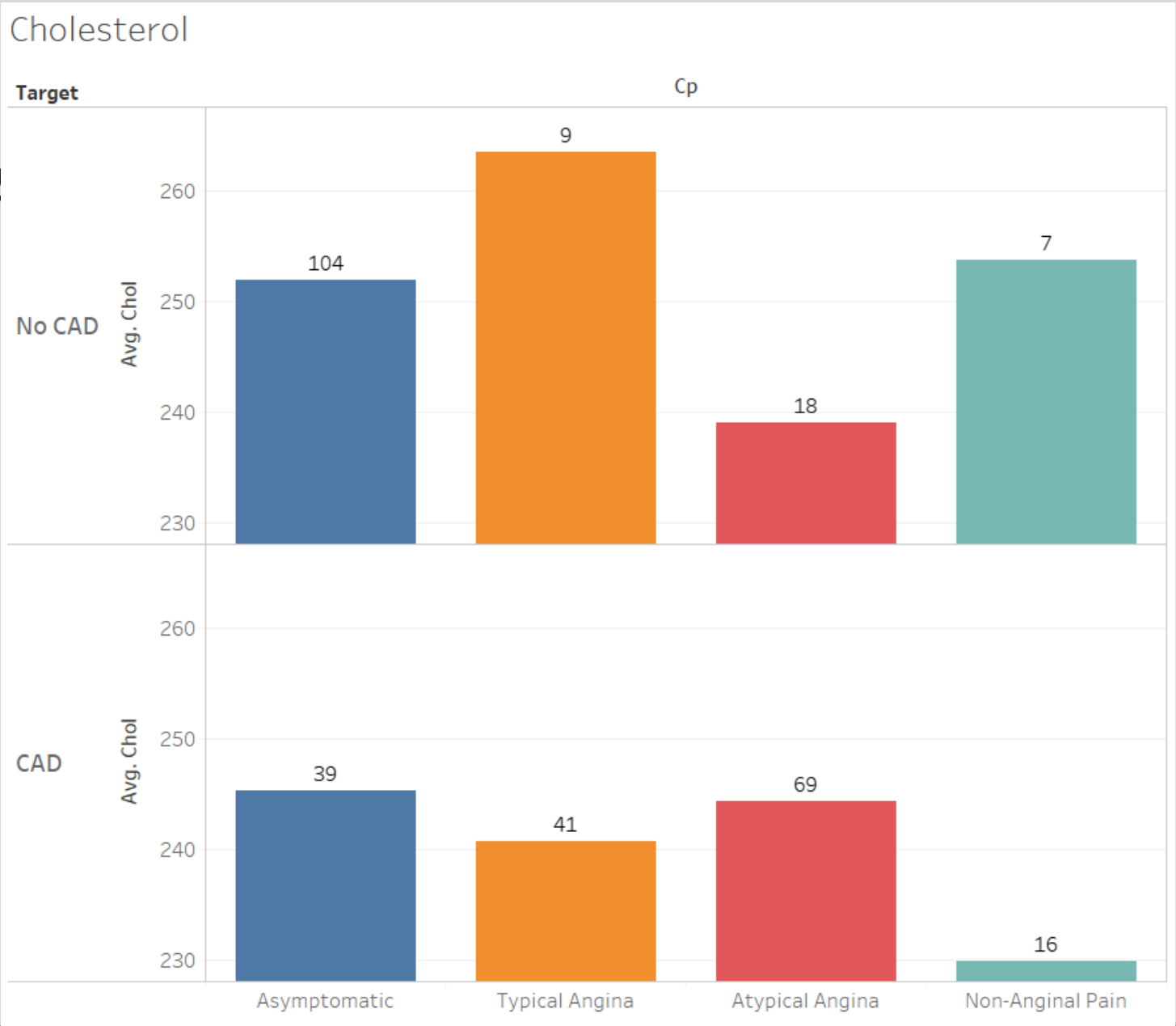
*4. Which test results of each test table can be used as indicators for CAD patients?*

## ***Blood test table***

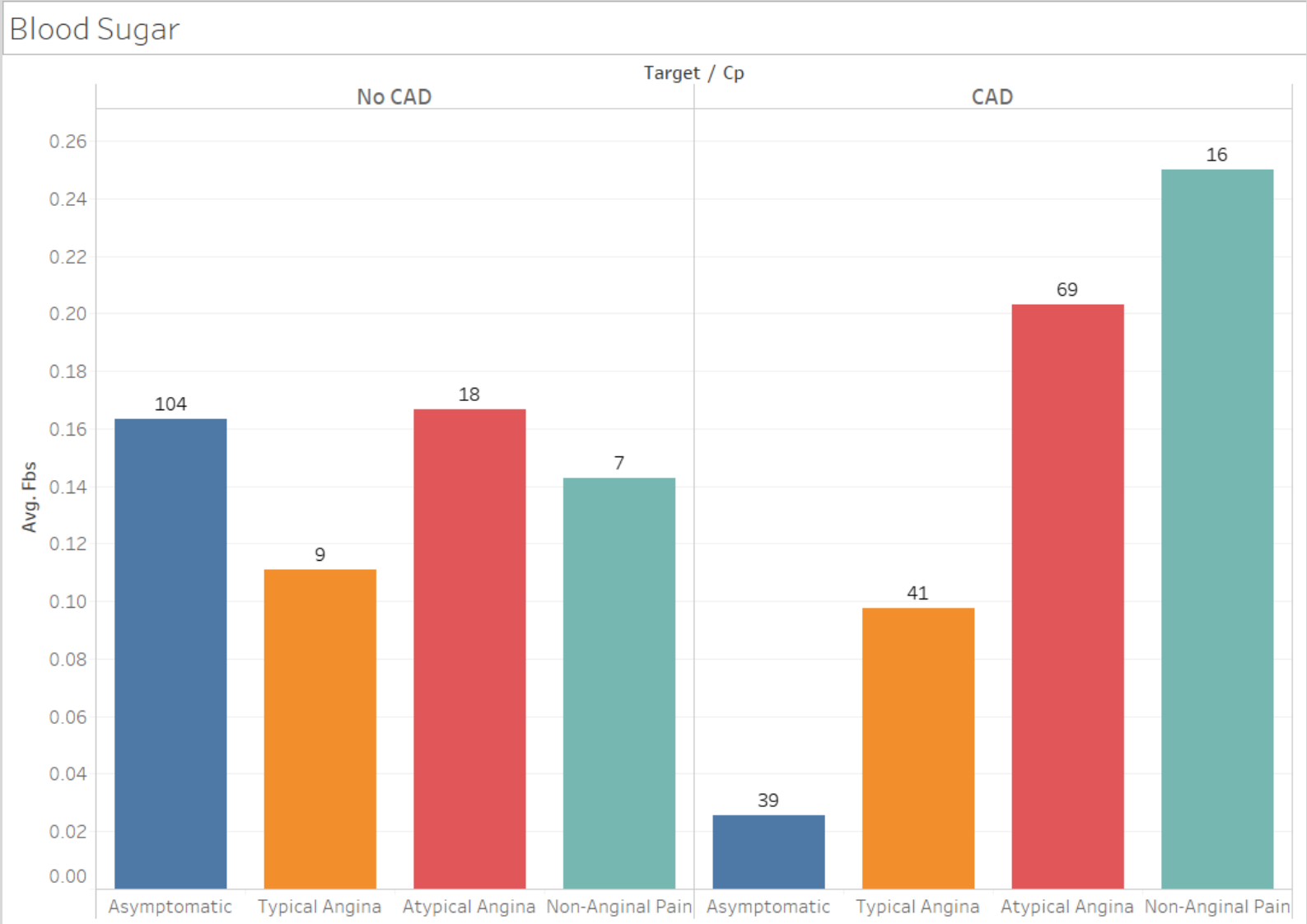
### ***Query***

- **Select** CP, BLOODTEST, chol, trestbps, fbs
- **From** target\_cad t, bloodtest e, symptoms s
- **Where** t.patient\_id = s.patient\_id
- and t.test\_id = e.test\_id

# Cholesterol and Chest Pain

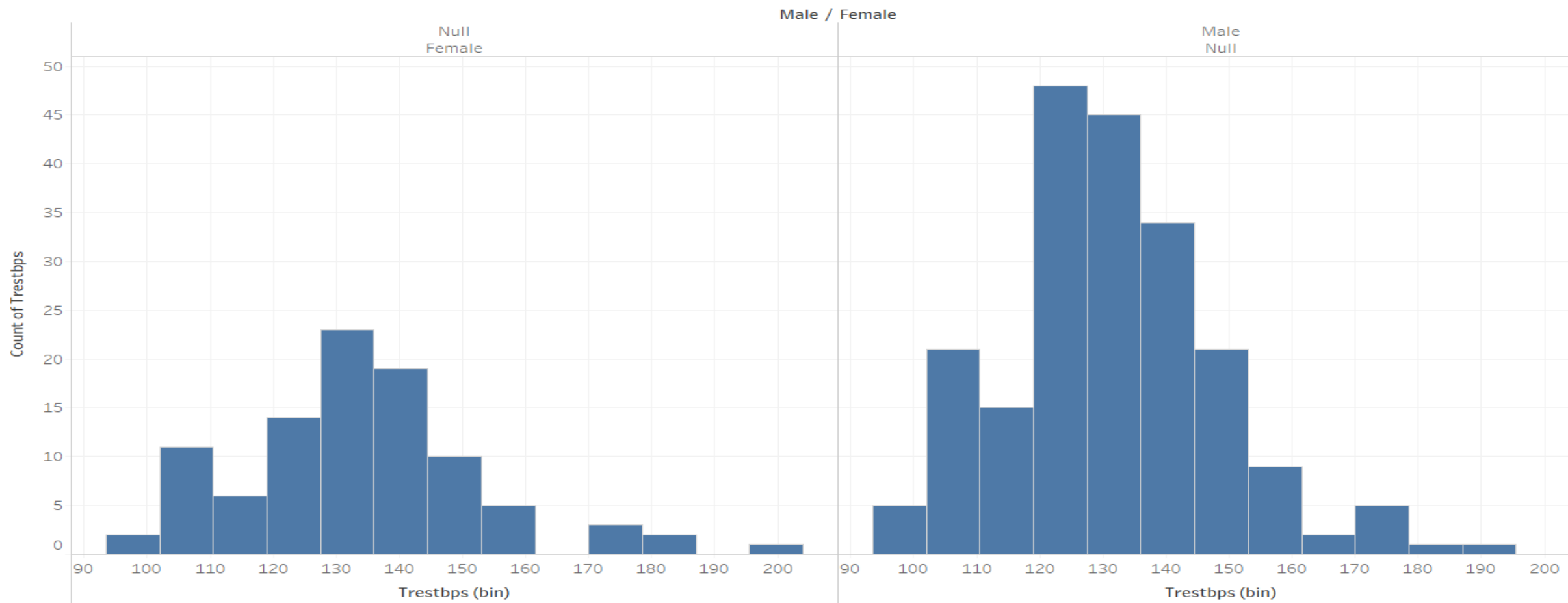


# Blood Sugar and Angina



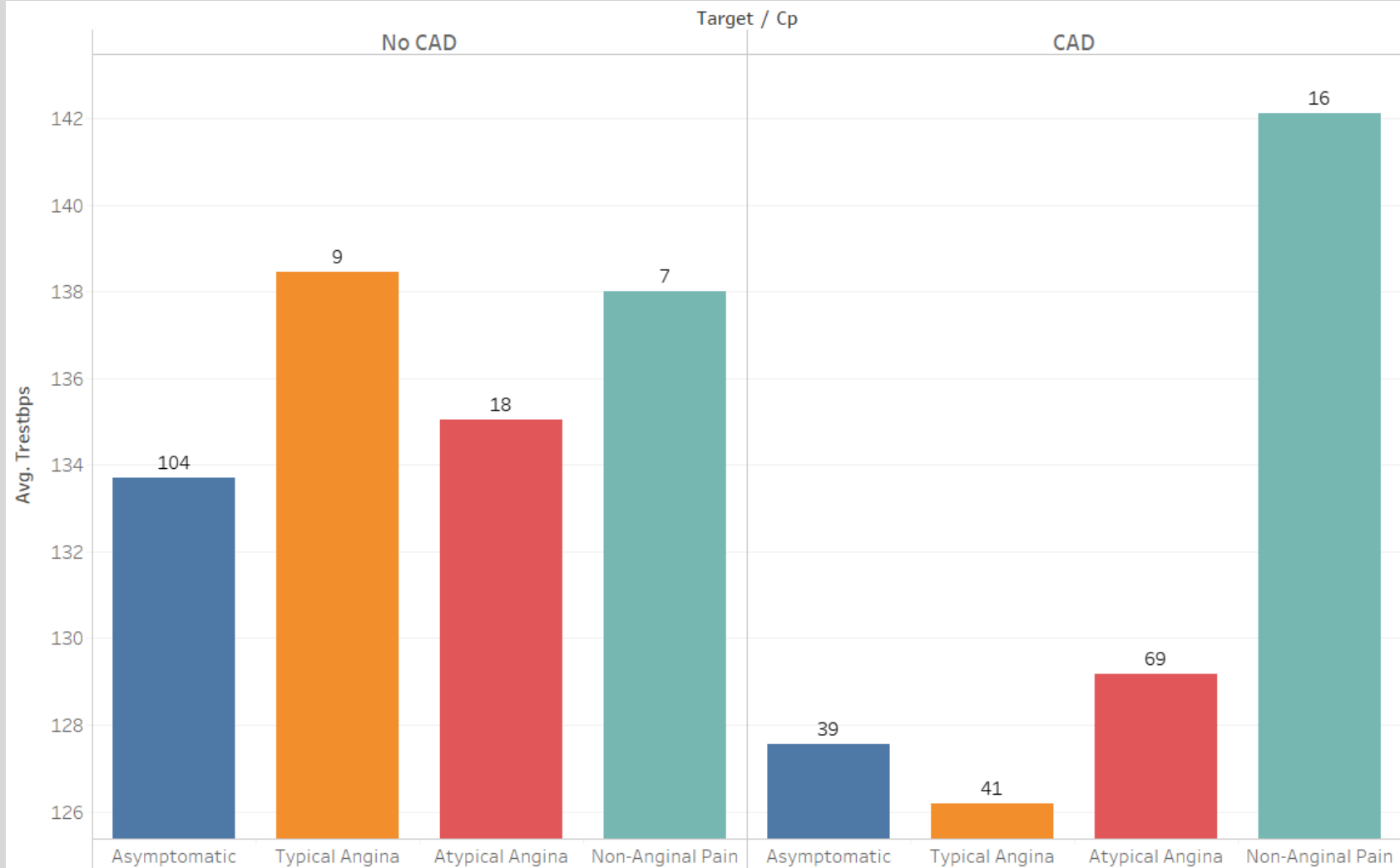
# ***Blood pressure and Gender***

Blood Pressure



# Blood Pressure and Angina

Bp and Angina



# ***Exercise table***

## ***Query***

**Select** CP, HEARTDISEASE, oldpeak, exang, slope, COUNT(\*)

**From** target\_cad t, exercise e, symptoms s

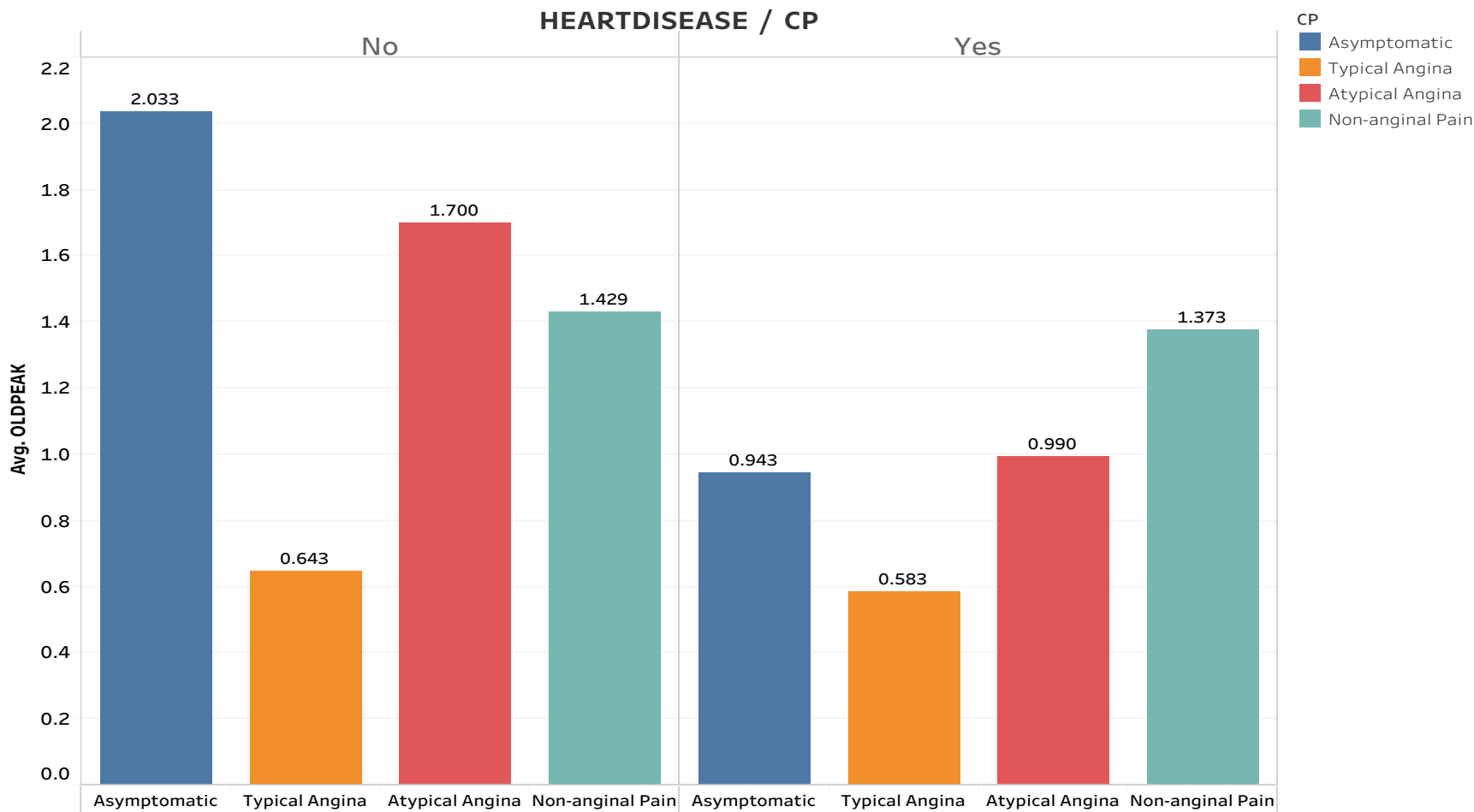
**Where** t.patient\_id = s.patient\_id

and t.exercise\_id = e.exercise\_id

**GROUP BY** CP, HEARTDISEASE, oldpeak, exang, slope

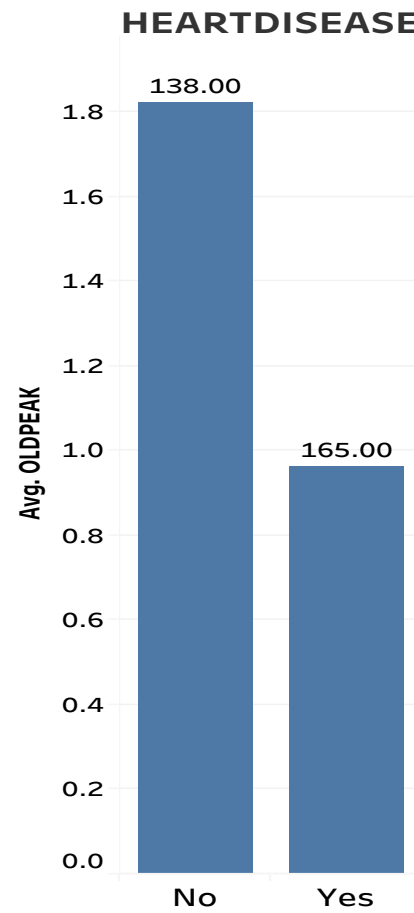
# Short Term Depression Induced by Exercise vs Heart Disease

HD/CP vs OLDPEAK



Average of OLDPEAK for each CP broken down by HEARTDISEASE. Color shows details about CP.

Average ST Depression vs Heart Disease



Average of OLDPEAK for each HEARTDISEASE. The marks are labeled by sum of COUNT(\*).

# Short-Term Peak Exercise vs Heart Disease

HD/CP vs SLOPE


		SLOPE			COUNT(*)
HEARTDISE..	CP	Upward	Flat Slope	Downward	
No	Atypical Angina		14.00	4.00	
	Non-anginal Pain		5.00	2.00	
	Typical Angina	1.00	4.00	4.00	
	Asymptomatic	11.00	68.00	25.00	
Yes	Asymptomatic		16.00	23.00	
	Typical Angina	1.00	8.00	32.00	
	Non-anginal Pain	3.00	6.00	7.00	
	Atypical Angina	5.00	19.00	45.00	

Sum of COUNT(\*) broken down by SLOPE vs. HEARTDISEASE and CP. Color shows sum of COUNT(\*). The marks are labeled by sum of COUNT(\*).



# Exercise Induced Angina vs Heart Disease

HD/CP vs EXANG

EXANG	CP	HEARTDISEASE		COUNT(*)
		No	Yes	
No	Asymptomatic	34.00	29.00	 1.00 70.00
	Typical Angina	8.00	38.00	
	Atypical Angina	14.00	62.00	
	Non-anginal Pain	6.00	13.00	
Yes	Asymptomatic	70.00	10.00	
	Typical Angina	1.00	3.00	
	Atypical Angina	4.00	7.00	
	Non-anginal Pain	1.00	3.00	

Sum of COUNT(\*) broken down by HEARTDISEASE vs. EXANG and CP. Color shows sum of COUNT(\*). The marks are labeled by sum of COUNT(\*).

Heart Disease vs Exang Totals

EXANG	HEARTDISEASE	
	0	1
0	62.0	142.0
1	76.0	23.0

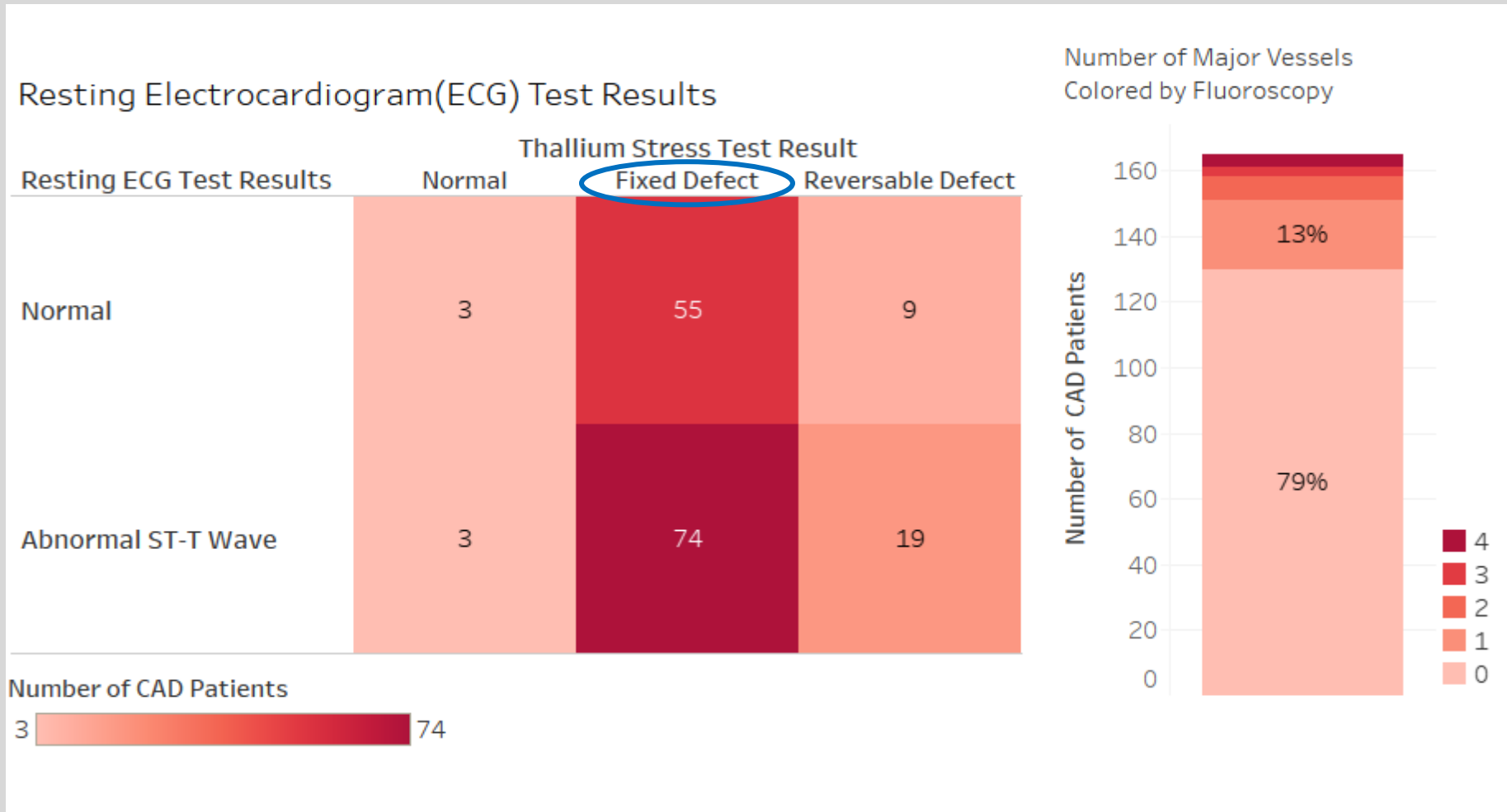
Sum of COUNT(\*) broken down by HEARTDISEASE vs. EXANG.

## ***Test table***

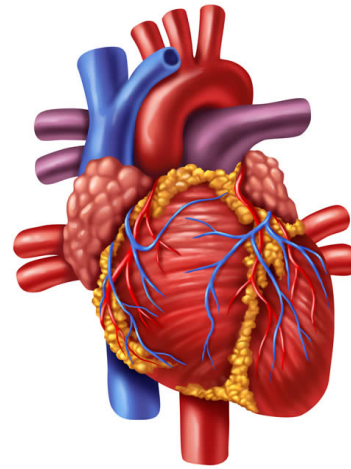
## ***Query***

```
SELECT t.Patient_ID, Restecg, Thalach, CA, Thal  
FROM TARGET_CAD t, TEST te, SYMPTOMS s  
Where t.Patient_ID = s.Patient_ID  
      and t.TestID = te.TestID  
      and Heartdisease = 1;
```

# Test Table



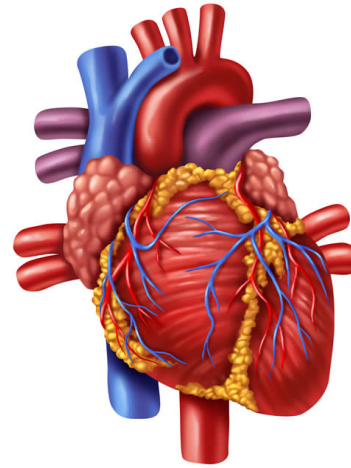
- The Thallium Stress Test of Test table appears to be the main indicator of CAD patients



***CONCLUSION***

# ***Conclusion***

- More males than females in this study
  - Females had a higher average age
- CAD patients showed mostly atypical or typical angina
  - Non-CAD patients were mostly asymptomatic as predicted
- Indicators:
  - Slope, oldpeak, max heart rate, chest pain type, thallium stress test, and blood sugar



***THANK YOU FOR LISTENING***