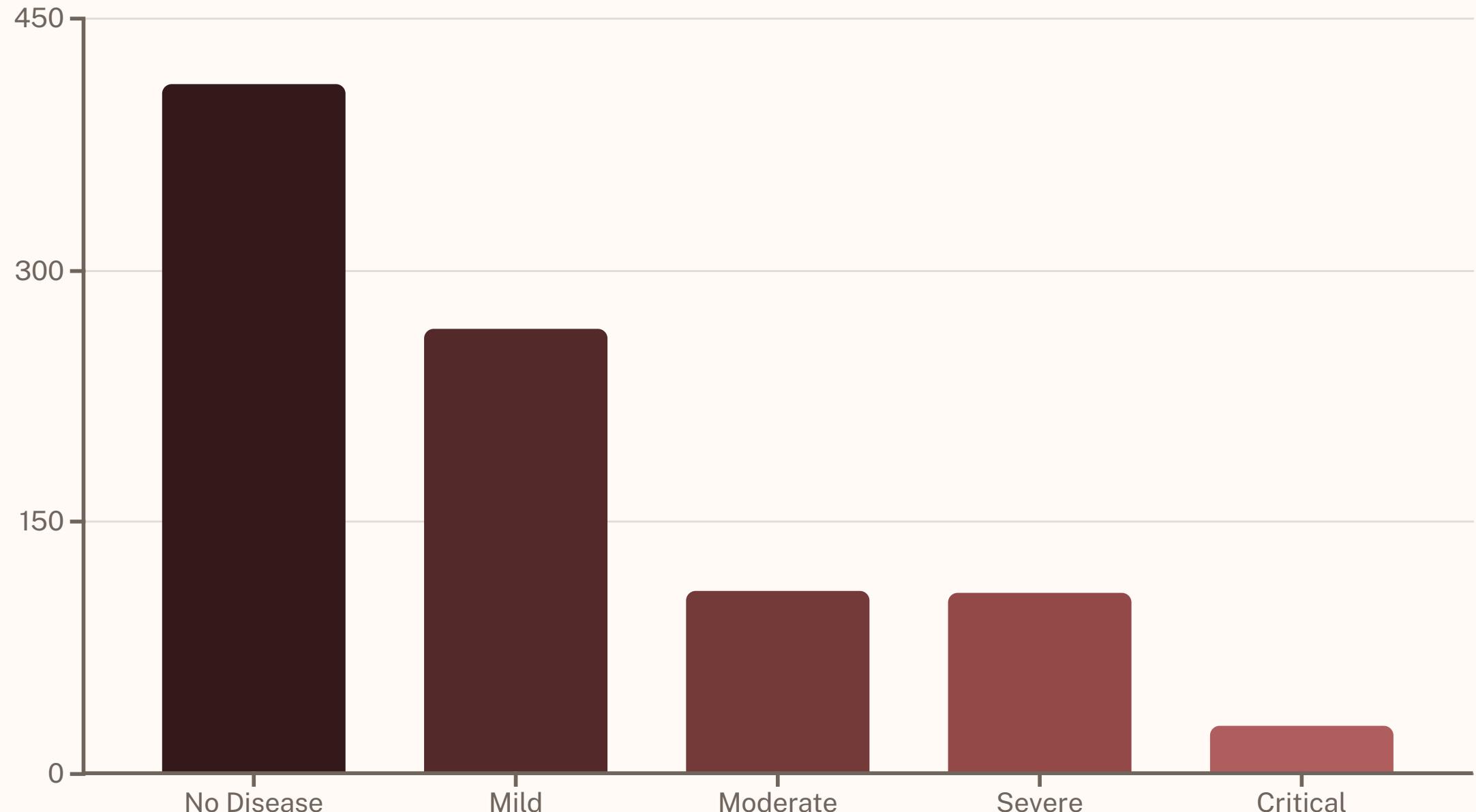




# Understanding Heart Disease Presence

This presentation explores the distribution and correlations of heart disease presence within a dataset, categorizing severity from no disease to critical. We will examine key physiological values, categorical data, and medical test outputs to identify significant trends and predictors.

# Heart Disease Severity Distribution



The dataset shows a highly imbalanced distribution of heart disease severity. Approximately 73.4% of patients have no or mild heart disease, while only about 20% exhibit moderate to critical conditions. As severity increases, the number of patients decreases significantly.

# Key Variables in Heart Disease Analysis

Trestbps

Resting blood pressure

Chol

Cholesterol levels

Thalach

Maximum heart rate achieved

Oldpeak

ST depression induced by exercise

Ca

Number of major vessels

Num

Heart disease severity (target variable)

These variables are crucial for understanding the factors influencing heart disease. Their interrelationships and individual impacts on severity will be further explored through correlation analysis.



## Correlation Matrix Overview

The heatmap illustrates the correlation matrix for various features and heart disease severity ('num'). No pair shows a very strong correlation (>0.7), with most variables exhibiting weak correlations (0.1-0.3). Notably, 'oldpeak' and 'ca' show moderate correlation with the target variable.

Some features, particularly 'thalach' (maximum heart rate), demonstrate negative relationships with other variables, indicating inverse trends.

# Age and Its Impact on Heart Health

Age shows several correlations with heart disease indicators:

- **Age vs. Thalach (-0.37):** Older patients tend to have a lower maximum heart rate.
- **Age vs. Ca (0.37):** Older individuals generally have more blocked vessels.
- **Age vs. Num (0.34):** Heart disease severity typically increases with age.
- **Age vs. Trestbps (0.24):** Older patients show slightly higher resting blood pressure.
- **Age vs. Oldpeak (0.26):** Older age is associated with more ST depression.
- **Age vs. Chol (-0.086):** Cholesterol levels do not significantly depend on age in this dataset.





# Blood Pressure and Cholesterol Insights



## Trestbps (Resting BP)

Weakly correlated with heart disease (0.12). Higher BP shows a slight, unreliable link to higher risk and slightly lower max heart rate (-0.1).



## Cholesterol

Almost no correlation with heart disease (-0.23) or age (-0.086). Not a reliable predictor in this dataset.

While commonly associated with heart health, resting blood pressure and cholesterol show surprisingly weak correlations in this specific dataset, suggesting other factors may be more influential.

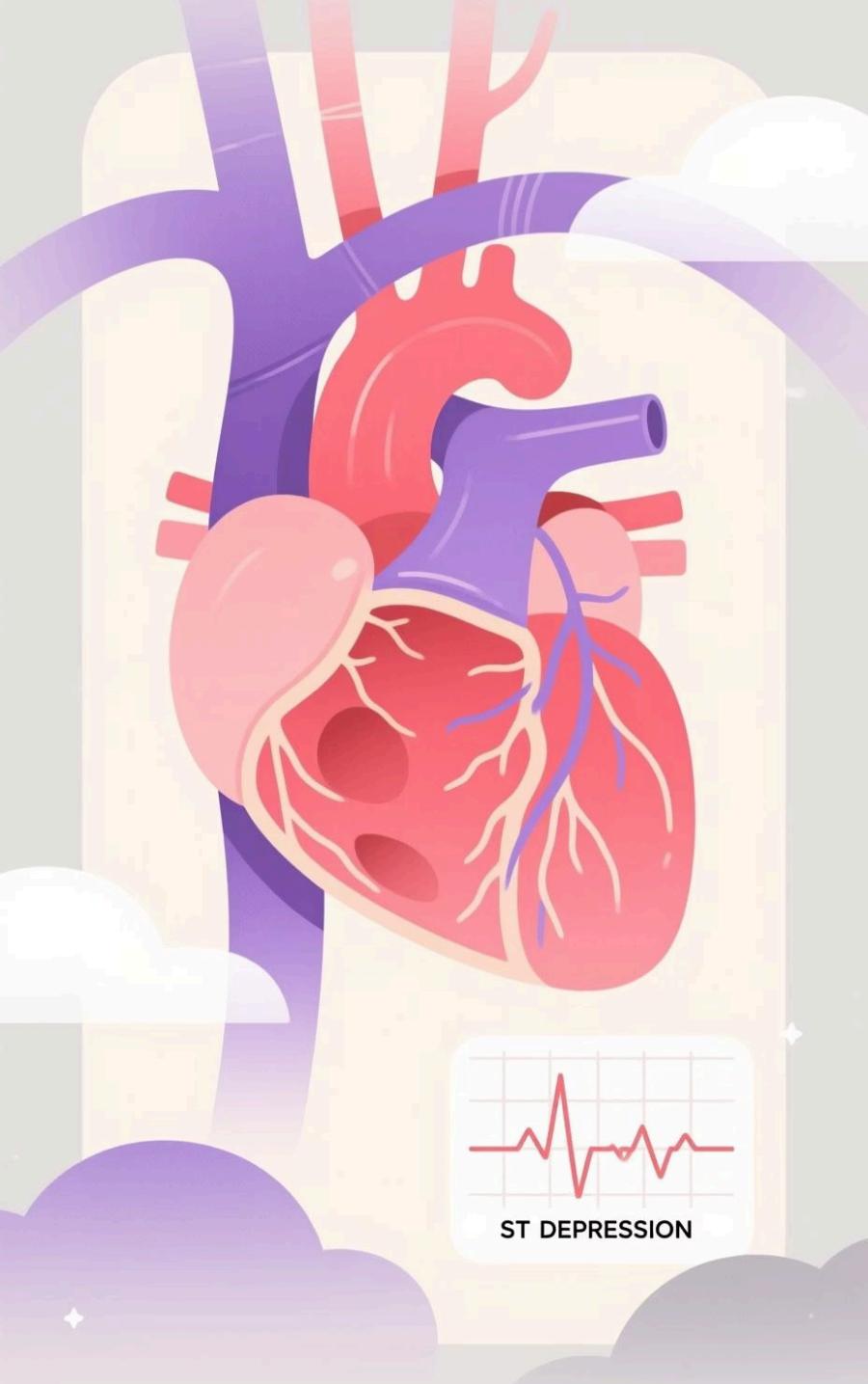
# Max Heart Rate (Thalach) as a Predictor



Maximum heart rate (thalach) exhibits significant inverse relationships:

- **Thalach vs. Num (-0.37):** A lower maximum heart rate is associated with a higher chance of heart disease.
- **Thalach vs. Age (-0.37):** Older age typically leads to a lower maximum heart rate.
- **Thalach vs. Oldpeak (-0.15):** Lower max heart rate can lead to slightly higher ST depression.
- **Thalach vs. Ca (-0.26):** Patients with more blocked vessels tend to have a lower maximum heart rate.

These findings highlight thalach as a clinically meaningful indicator, where a reduced capacity for maximum heart rate points towards increased heart disease severity.



# Oldpeak and Number of Vessels: Strong Predictors



## Oldpeak (ST Depression)

**Oldpeak vs. Num (0.44):** Higher ST depression strongly correlates with more severe heart disease, making it a clear positive linear relationship.

**Oldpeak vs. Age (0.26):** Older individuals tend to experience more ST segment depression.



## Number of Vessels (Ca)

**Ca vs. Num (0.52):** More blocked vessels are the most important predictor, leading to more severe heart disease.

**Ca vs. Age (0.37):** Older people are more likely to have increased blockages.

Both 'oldpeak' and 'ca' stand out as the strongest predictors of heart disease severity in this dataset, with 'ca' being the most influential.

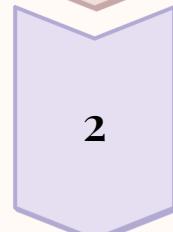
# Linear Trends in Feature Combinations

While no clear exponential relationships were found, several feature combinations show linear trends, particularly when considering medical logic and pairplots:



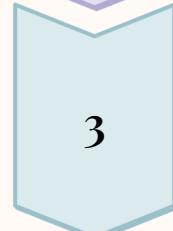
## 1 Oldpeak vs. Num

Strong Positive Linear: Higher ST depression leads to higher heart disease severity.



## 2 Thalach vs. Num

Strong Negative Linear: Lower maximum heart rate indicates higher heart disease severity.



## 3 Age vs. Num

Moderate Positive Linear: Older age slightly increases heart disease severity.

These linear relationships, though not always strong, provide valuable insights into the progression and indicators of heart disease.



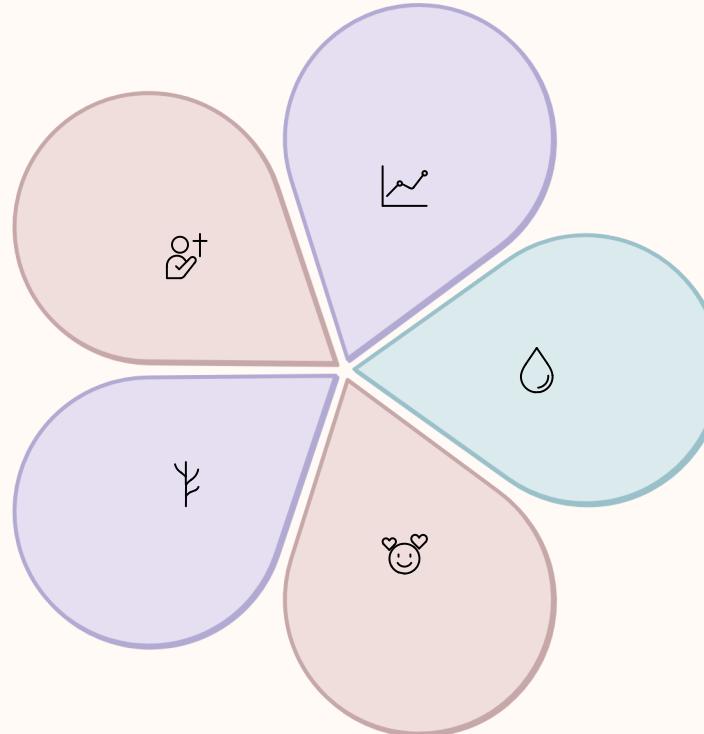
# Summary of Key Findings

## Imbalanced Data

Dataset dominated by non-disease or mild cases.

## Strongest Predictors

Oldpeak and number of vessels (ca) are highly correlated with severity.



## Age Impact

Increases severity, lowers max heart rate, more blocked vessels.

## Weak Predictors

Trestbps and cholesterol show minimal correlation.

## Thalach

Lower max heart rate indicates higher disease risk.

This analysis highlights the most influential factors in predicting heart disease severity, emphasizing the importance of 'oldpeak' and 'ca' for future diagnostic and prognostic models.