Heart Disease Risk Factors

January 2026

Clinical Insights & Data-Driven Results

codecademy

Bootcamp • Applied Data Science with Python

By: Blair Bateman

Data Source

Data set is a sample patients who were evaluated for heart disease at the Cleveland Clinic Foundation.

The principal investigators responsible for data collection were:

Hungarian Institute of Cardiology. Budapest: Andras Janosi, M.D.

University Hospital, Zurich, Switzerland: William Steinbrunn, M.D.

University Hospital, Basel, Switzerland: Matthias Pfisterer, M.D.

V.A. Medical Center, Long Beach and Cleveland Clinic Foundation: Robert Detrano, M.D., Ph.D.

Key Research Questions

1 Cholesterol Analysis: Do patients with heart disease have mean cholesterol levels significantly higher than 240 mg/dl?

2 Fasting Blood Sugar: Is the prevalence of elevated fasting blood sugar (>120 mg/dl) in this clinical sample significantly higher than the 1988 U.S. national average of 8%?
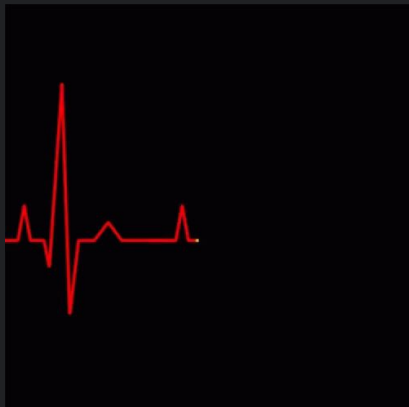
3 Comparative Analysis: How do cardiovascular metrics differ between patients with and without heart disease?

Overall View Of Patients

Heart Disease Distribution (n=303 patients)

Without Heart Disease — 54.1%

With Heart Disease — 45.9%
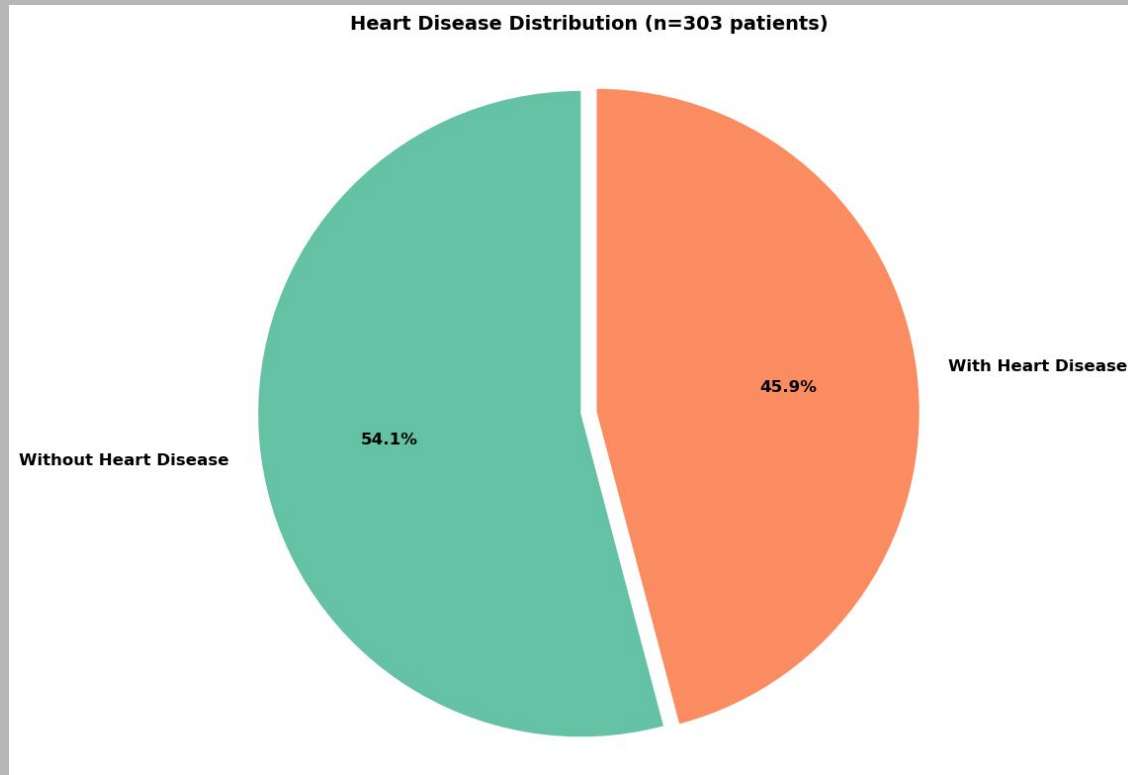
🔴 CRITICAL: Maximum Heart Rate

WITH Heart Disease

139 bpm

WITHOUT Heart Disease

158 bpm

Difference: 19.12 bpm ↓ | Effect Size: LARGE (d = -0.918)

✓ Strongest predictor identified | p < 0.0001
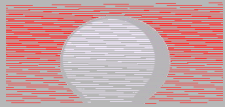
Notable Findings

Resting Blood Pressure

Heart Disease: 134.6 mmHg vs No-Heart Disease: 129.3 mmHg
Small effect (d=0.305)

Cholesterol

Heart Disease: 251.5 mg/dL vs No-Heart Disease: 242.6 mg/dL
Negligible effect (d=0.171)

Key Insight

Multiple factors matter—no single marker is deterministic. Risk is
multifactorial.

BREAKTHROUGH:
Fasting Blood Sugar

Study Prevalence (FBS > 120 mg/dL)

14.85%

vs. 1988 Baseline: 8%

85% HIGHER than expected

Statistical Significance: p = 0.0000469 ✓

SUMMARY TABLE: Effect Sizes and Statistical Significance

Comparison	Mean Difference	Cohen's d	Effect Size	p-value
Max Heart Rate (HD vs No HD)	-19.12 bpm	-0.918	Large	0.0000 ***
Age (HD vs No HD)	4.04 years	0.458	Small	0.0001 ***
Resting BP (HD vs No HD)	5.32 mmHg	0.305	Small	0.0085 **
Cholesterol (HD vs No HD)	8.83 mg/dL	0.171	Negligible	0.1391 ns

Effect Size Interpretation: |d| < 0.2 = negligible, 0.2-0.5 = small, 0.5-0.8 = medium, > 0.8 = large
Significance: *** p<0.001, ** p<0.01, * p<0.05, ns = not significant

Key Research Question 1

Cholesterol Analysis: Do patients with heart disease have mean cholesterol levels significantly higher than 240 mg/dl?

Result: YES, but with important context

Patients with heart disease have a mean cholesterol level of 251.47 mg/dl, which is significantly higher than the 240 mg/dl threshold (p = 0.0035). However, the effect size is small (Cohen's d = 0.232), meaning the clinical significance is modest.

Interestingly, patients without heart disease have a mean cholesterol of 242.64 mg/dl, which is NOT significantly different from 240 mg/dl (p = 0.2640, effect size = negligible).

Key Research Question **2**

Fasting Blood Sugar: Is the prevalence of elevated fasting blood sugar (>120 mg/dl) in this clinical sample significantly higher than the 1988 U.S. national average of 8%?

Result: YES, statistically significant

The prevalence of elevated fasting blood sugar in your clinical sample is 14.85%, which is significantly higher than the 1988 baseline of 8% (p = 0.0000469). This represents nearly double the expected rate.

Important note: Your project acknowledges this 8% baseline is from 1988, so results should be interpreted with caution given that modern rates have likely changed over the past three decades.

Key Research Question 3

Comparative Analysis: How do cardiovascular metrics differ between patients with and without heart disease?

Metric	Patients WITH HD	Patients WITHOUT HD	Mean Difference	Effect Size	p-value
Maximum Heart Rate	139.26 bpm	158.38 bpm	-19.12 bpm	Large (d = -0.918)	< 0.0001 ***
Age	56.63 years	52.59 years	+4.04 years	Small (d = 0.458)	< 0.0001 ***
Resting BP	134.57 mmHg	129.25 mmHg	+5.32 mmHg	Small (d = 0.305)	0.0085 **
Cholesterol	251.47 mg/dl	242.64 mg/dl	+8.83 mg/dl	Negligible (d = 0.171)	0.1391 ns

Effect Size Interpretation: |d| < 0.2 = negligible, 0.2-0.5 = small, 0.5-0.8 = medium, > 0.8 = large
Significance: *** p<0.001, ** p<0.01, * p<0.05, ns = not significant

The Bottom Line

Many Factors Causes Heart disease. Source New Baseline Data.

Focus on Fitness • Manage Multiple Risk Factors • Screen Proactively