# Heart Disease Analysis Insights

## 1. General Demographics

Age Distribution & Risk:  
 - Majority of the dataset lies between 30–70 years, with an increasing trend of heart disease prevalence in the 50+ age group.

Gender & Heart Disease:  
 - Males tend to have a higher prevalence of heart disease compared to females.  
 - Gender seems to play a notable role when combined with other risk factors (e.g., smoking, cholesterol).

## 2. Clinical Indicators & Risk Factors

BMI (Body Mass Index)  
- The average BMI across the dataset is in the overweight range (25–30).  
- Higher BMI levels correlate with increased instances of heart disease.  
  
Blood Pressure  
- Individuals with high blood pressure (systolic ≥140 mmHg) showed a significantly higher risk of heart disease.  
- Average blood pressure of those with heart disease is markedly higher than those without.  
  
Cholesterol & Triglycerides  
- People with heart disease exhibit elevated total cholesterol and triglyceride levels, indicating poor lipid metabolism.  
- Both Low HDL (good cholesterol) and High LDL (bad cholesterol) are more common among heart disease positive individuals.  
  
CRP (C-Reactive Protein) & Homocysteine  
- High CRP levels, a marker of inflammation, are positively associated with heart disease cases.  
- Elevated Homocysteine levels, linked to vascular damage, are present in many individuals with heart disease.

Fasting Blood Sugar  
- A larger proportion of heart disease patients have fasting glucose levels > 126 mg/dL, which indicates undiagnosed or unmanaged diabetes.

## 3. Lifestyle Factors

Exercise Habits  
- Individuals with low or no physical activity have a notably higher incidence of heart disease.  
- Regular exercise appears to be a protective factor.  
  
Smoking  
- Smokers show a significantly increased heart disease rate.  
- Smoking is strongly correlated with high cholesterol, CRP, and low HDL levels.  
  
Alcohol Consumption  
- High alcohol consumption also trends toward higher heart disease risk, especially when combined with smoking and poor diet.  
  
Sugar Consumption  
- High sugar consumption correlates with higher triglycerides and BMI, indirectly elevating heart disease risk.  
  
Stress Level  
- A majority of heart disease patients report medium to high stress levels.  
- Chronic stress may exacerbate inflammatory markers (CRP), indirectly increasing heart disease risk.  
  
Sleep Hours  
- People sleeping less than 6 hours per day show increased markers of stress and inflammation.  
- Optimal sleep duration (7–8 hours) is more common among healthy individuals.

## 4. PCA Insights (Dimensionality Reduction)

Principal Component Analysis (PCA) revealed:  
 - Top features influencing variance: Age, CRP Level, Homocysteine, Triglycerides, BMI.  
 - These features can be used for risk prediction modeling or screening prioritization.

## 5. Statistical Trends

- A/B comparisons (e.g., smokers vs non-smokers, high vs low cholesterol) show statistically significant differences in heart disease rates using t-tests.  
- High Blood Pressure + Diabetes + High Cholesterol is the most common comorbidity cluster for individuals with heart disease.

## Key Takeaways

|  |  |
| --- | --- |
| Insight | Observation |
| Most At-Risk | Males, Age 55+, High BP, High BMI, Smoker, High CRP |
| Protective Factors | Regular exercise, low stress, normal cholesterol, no diabetes |
| Strong Predictors | CRP Level, Triglycerides, Homocysteine, Blood Pressure |
| Lifestyle Red Flags | Smoking, poor sleep, no exercise, high sugar/alcohol |