

# Healthcare Analytics: Heart Disease Analysis

## Menternship Project Report

---

### ◆ Introduction

This analysis is based on the **Behavioral Risk Factor Surveillance System (BRFSS 2015)** dataset, containing information on ~250,000 respondents. The primary objective was to explore health and demographic factors associated with **heart disease and heart attacks**, using exploratory data analysis (EDA) and visualization techniques.

---

### ◆ Key Findings

#### 1. Heart Disease Prevalence

- Approximately **9% of respondents** reported heart disease or a heart attack.
- The dataset is **imbalanced**, with many more “No” than “Yes” cases.
- This highlights the need for balanced approaches in predictive modeling.

#### 2. Factors Associated with Heart Disease

- **High Blood Pressure (Hypertension):**
  - Prevalence of heart disease is significantly higher among individuals with HighBP.
  - Chi-square tests confirm a **strong statistical association** ( $p < 0.001$ ).
- **High Cholesterol:**
  - Similarly, individuals with high cholesterol show higher heart disease prevalence.
  - Indicates the importance of cholesterol screening and control.
- **Age:**
  - Heart disease prevalence rises sharply with age.
  - Highest rates observed in respondents aged **65 years and above**.
- **Body Mass Index (BMI):**
  - Average BMI of individuals with heart disease is **slightly higher** than those without.
  - Obesity emerges as a contributing risk factor.
- **Lifestyle Factors:**
  - **Smoking** and **lack of physical activity** show a moderate association with heart disease.

- Respondents reporting poor mental/physical health days also show slightly higher prevalence.
  - **Demographics (Socioeconomic Factors):**
    - Prevalence is somewhat higher among **males** compared to females.
    - Respondents with **lower income and education levels** show higher prevalence, suggesting socioeconomic disparities.
- 

#### ◆ **Insights & Interpretation**

- Heart disease is influenced by a **combination of biological, lifestyle, and socioeconomic factors**.
  - Preventable risk factors such as **hypertension, high cholesterol, obesity, and smoking** play a crucial role.
  - Older adults and individuals from disadvantaged backgrounds are disproportionately affected, highlighting the importance of **targeted interventions**.
- 

#### ◆ **Recommendations**

##### 1. **Strengthen Preventive Healthcare:**

- Encourage routine screening for **blood pressure** and **cholesterol**.
- Promote awareness campaigns on healthy diet, exercise, and weight management.

##### 2. **Target High-Risk Groups:**

- Focus health interventions on **older adults, hypertensive patients, and those with high cholesterol**.
- Design programs specifically for **low-income and low-education groups**.

##### 3. **Promote Lifestyle Modifications:**

- Implement smoking cessation programs.
- Encourage physical activity campaigns at the community level.
- Provide resources for stress and mental health management.

##### 4. **Further Research:**

- Explore multivariate predictive modeling using this dataset.
  - Assess the impact of combined risk factors (e.g., obesity + hypertension).
-

## ◆ Conclusion

This analysis demonstrates that **heart disease prevalence** is strongly associated with modifiable health risks (blood pressure, cholesterol, BMI, smoking) as well as demographic factors (age, income, education).

By targeting these factors through **preventive healthcare strategies, awareness programs, and tailored interventions**, policymakers and healthcare providers can **reduce the burden of heart disease** and improve population health outcomes.