Heart Disease Risk Analysis

Overview

Understanding heart disease risk factors across different age groups is crucial for effective prevention. This study employs a structured dataset with demographic, clinical, and lifestyle variables to investigate the role of age, gender, and BMI in heart disease prevalence. With a focus on individuals aged 50 and above, the analysis aims to uncover trends that contribute to increased susceptibility, particularly in the presence of hypertension and metabolic conditions. These insights will guide data-driven healthcare policies and interventions to achieve a 10% reduction in heart disease cases by the end of 2025.

Problem Statement

By the end of 2025, determine the impact of age, gender, and BMI on the likelihood of developing heart disease, focusing on individuals within the age brackets of 20-29, 30-39, 40-49, 50-59, and 60-69. The goal is to identify key risk factors across these age groups and reduce heart disease incidence by 10%, particularly in individuals aged 50 and above.

Goals and Objectives

- Identify key risk factors influencing heart disease across different age groups.

- Examine the impact of gender differences on heart disease prevalence.

- Assess the influence of BMI on heart disease risk.

- Evaluate the role of smoking history and other lifestyle factors.

- Provide data-driven recommendations to reduce heart disease incidence by 10%.

Tools and Technologies Used

- **Programming Languages:** Python (Pandas, NumPy, Matplotlib, Seaborn)

- **Data Visualization:** Tableau

- **Data Storage and Management:** SQL, CSV format

- **Others:** PowerPoint, Excel

Methodology: HDEIP Framework

HDEIP stand for Hypothesis formation, Data Sourcing, Exploratory Data Analysis, Insight and Presentation. This analysis was structured using the **HDEIP Framework**, which consists of:

1. **Hypothesis Formation:** Defining key research questions and problem statements.

2. **Data Sourcing:** Collecting structured datasets with demographic, clinical, and lifestyle factors.

3. **Exploratory Data Analysis (EDA):** Identifying patterns, trends, and correlations.

4. **Insight Generation:** Deriving actionable insights from the data.

5. **Presentation:** Communicating findings to stakeholders via reports and visualizations.

Analytical Approach

To ensure a comprehensive analysis, both **Macro-Level** and **Micro-Level** approaches were applied.

Macro-Level Analysis (Identifying Broad Patterns and Trends)

- **Data Aggregation:** Summarizing key metrics (averages, medians, totals).

- **Trend Identification:** Observing overarching patterns in age, BMI, and gender distributions.

- **Outlier Detection:** Identifying extreme values that required further investigation.

- **Correlation Analysis:** Exploring variable relationships to detect strong associations.

Micro-Level Analysis (Granular Investigation of Key Drivers)

- **Detailed Segmentation:** Breaking down data by subgroups (age, gender, BMI, health conditions).

- **Outlier Investigation:** Examining specific cases of extreme values.

- **Hypothesis Testing:** Applying statistical tests to validate insights.

- **Feature Impact Analysis:** Assessing how specific factors influenced heart disease prevalence.

Key Insights and Their Influence on Strategy

1. Age-Specific Risk Factors

- Heart disease prevalence rises sharply with age, particularly for individuals aged **50+**.

- Overweight and obese individuals aged **50-69** have the highest incidence of heart disease.

- **Recommendation:** Implement weight management and early hypertension detection programs.

2. Gender Differences

- Males show a slightly higher heart disease prevalence, particularly among smokers.

- Smoking exacerbates risks across BMI categories, with underweight smokers being most vulnerable.

- **Recommendation:** Gender-focused preventive programs, emphasizing smoking cessation and weight control.

3. Smoking and Heart Disease

- Former smokers exhibit the highest prevalence of heart disease (~10%), highlighting long-term risks.

- Current smokers still face significant risks, while non-smokers have lower rates.

- **Recommendation:** Enhance smoking cessation programs and long-term cardiovascular monitoring.

4. BMI and Heart Disease Risk

- Obesity is strongly correlated with heart disease in individuals aged **40-59**.

- Even overweight individuals have elevated risks, necessitating early intervention.

- **Recommendation:** Implement targeted weight management and lifestyle modification programs.

5. Diabetes and Heart Disease

- The presence of both **diabetes and hypertension** significantly increases heart disease risk in individuals aged **50+**.
- **Recommendation:** Targeted health campaigns to promote diabetes and blood pressure management.

Challenges Encountered

- **Data Imbalance:** Certain age groups had fewer records, requiring adjustments for bias.

- **Missing Data:** Some clinical variables were incomplete and needed imputation techniques.

- **Correlation vs. Causation:** Establishing causal relationships required robust statistical validation.

- **Stakeholder Engagement:** Communicating findings effectively to both technical and non-technical audiences.

Communication of Findings

The results were presented through a structured **Executive Presentation**, which included:

- **Slide 1: Executive Summary** (High-level overview of findings and strategic impact)

- **Slide 2: Age-Specific Risk Factors** (Visualizing trends and targeted interventions)

- **Slide 3: Gender Differences in Heart Disease** (Insights into male and female risk factors)

- **Slide 4: Smoking History and Heart Disease** (Impact analysis of smoking on heart disease)

- **Slide 5: BMI and Heart Disease Risk** (BMI thresholds and associated risks)

- **Slide 6: Impact of Diabetes on Heart Disease** (Diabetes and hypertension correlations)

Conclusion

This analysis provides a data-driven foundation for reducing heart disease incidence by 10% by the end of 2025. By leveraging **structured analytical methods**, **macro- and micro-level insights**, and **strategic recommendations**, the study equips healthcare policymakers and practitioners with the tools needed to implement effective preventive measures.