Portfolio Report: Health Risk Assessment Using Vital Signs Dataset

Bangcale Kurt Andrei[1], Era, Chloe Josephine F.[2]

Bangcale[1], Era[2], BI120L/CON29, School of Health Sciences, Mapúa University

INTRODUCTION

Cardiovascular disease remains one of the leading causes of morbidity and mortality worldwide. Early identification of individuals at elevated risk through comprehensive health assessments is crucial for implementing preventive interventions and reducing cardiovascular events. Understanding the relationships between modifiable risk factors such as BMI, smoking status, glucose levels, and blood pressure measurements is essential for developing effective risk stratification models.

Objectives

The primary objectives of this analysis were to:

1. Examine the distribution of cardiovascular risk factors in the study population
2. Analyze relationships between anthropometric measures and physiological parameters
3. Investigate differences in risk factors across demographic groups
4. Assess correlations between metabolic markers and cardiovascular indicators
5. Identify patterns that could inform clinical risk assessment strategies

METHODOLOGY

The analysis utilized a health risk assessment dataset containing comprehensive information on cardiovascular risk factors and physiological measurements. The dataset represents individuals undergoing health screening assessments.

1. Data Cleaning and Preprocessing

Data preprocessing procedures included:

- **Library loading**: Essential R packages (tidyverse,ggplot2) were loaded for data manipulation and visualization
- **Data import**: CSV file containing health assessment data was imported
- **Variable recoding**: Categorical variables were converted to factors with meaningful labels
- **Data type conversion**: Appropriate data types were assigned to ensure accurate statistical analysis

Variable Definitions

Demographic Variables:

- Age: Continuous variable (years)

- Sex: Binary factor (Female/Male, coded from 0/1)

Anthropometric and Lifestyle Variables:

- BMI: Body Mass Index (continuous)
- Smoking Status: Ordinal factor (Non-Smoker, Occasional, Chain Smoker, coded from 0/1/2)
- Daily Sleeping Hours: Continuous variable (hours per day)

Physiological Parameters:

- Heart Rate: Continuous variable (beats per minute)
- Systolic Blood Pressure: Continuous variable (mmHg)
- Glucose Level: Continuous variable (mg/dL)
- Cholesterol Level: Continuous variable (mg/dL)

Clinical Classifications:

- Hypertension: Binary factor (presence/absence)
- Elevated Risk: Binary factor (high risk/normal risk classification)

2. Statistical Analysis

Descriptive Statistics:

- Measures of central tendency (mean, median)
- Measures of variability (standard deviation)
- Range measurements (minimum and maximum values)

Visualization Techniques:

- Histograms for continuous variable distributions
- Box plots for group comparisons
- Scatter plots for correlation analysis with risk group coloring
- Bar charts for categorical variable frequencies

Inferential Statistics:

- Pearson correlation analysis between BMI and systolic blood pressure
- Statistical significance testing at $\alpha = 0.05$ level

3. Software and Tools

Analysis was conducted using R statistical software version 4.x with the following packages:

- tidyverse for comprehensive data manipulation and analysis
- ggplot2 for advanced data visualization and graphics

RESULTS

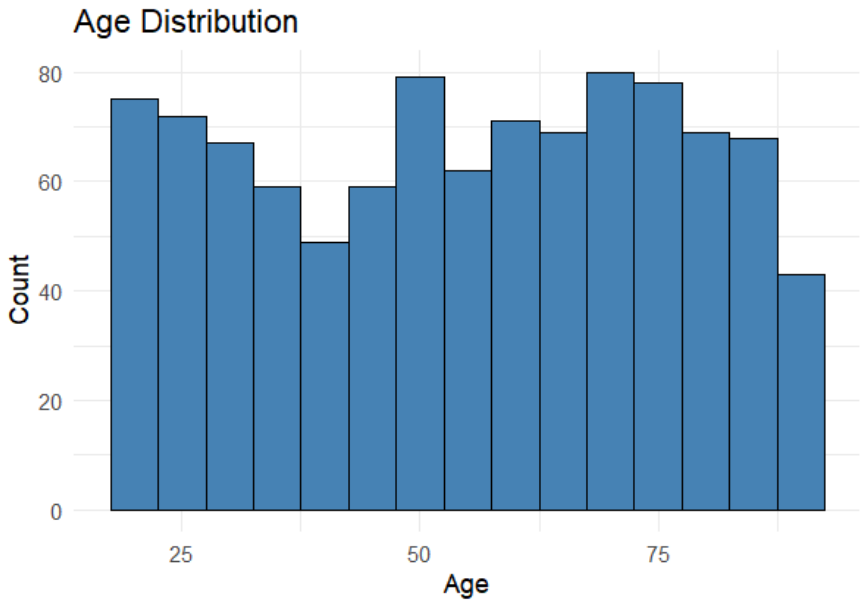

Figure 1: Histogram of Age Distribution

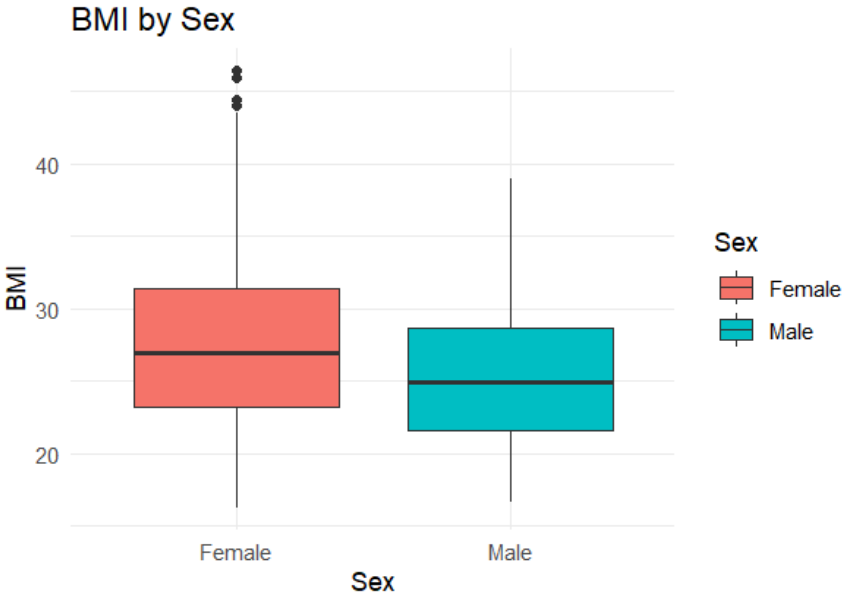

Figure 2: Box Plot of BMI vs Systolic BP

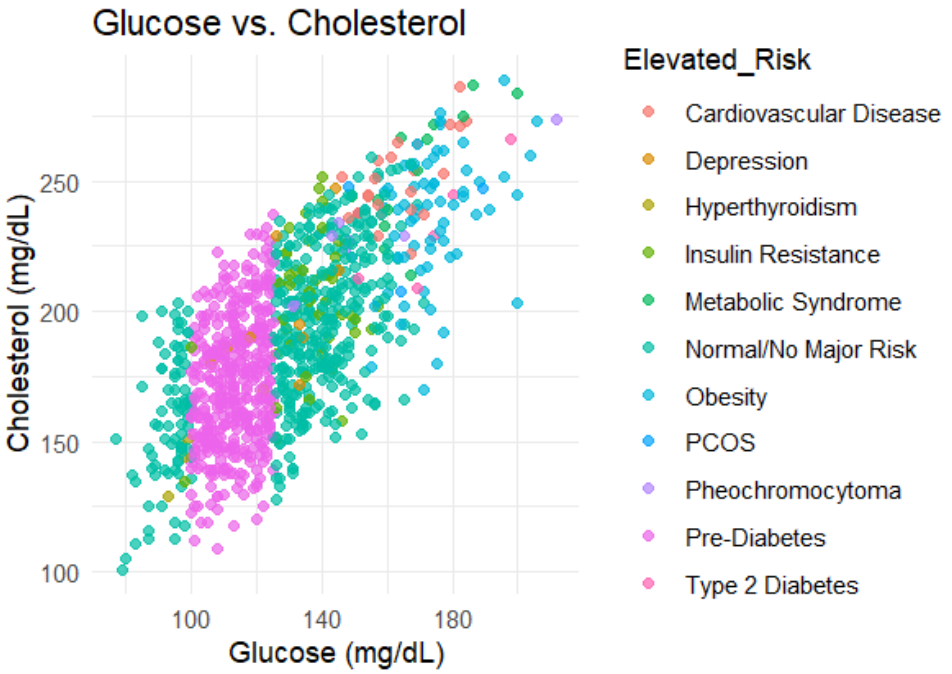

Figure 3: Scatter Plot of Glucose vs Cholesterol

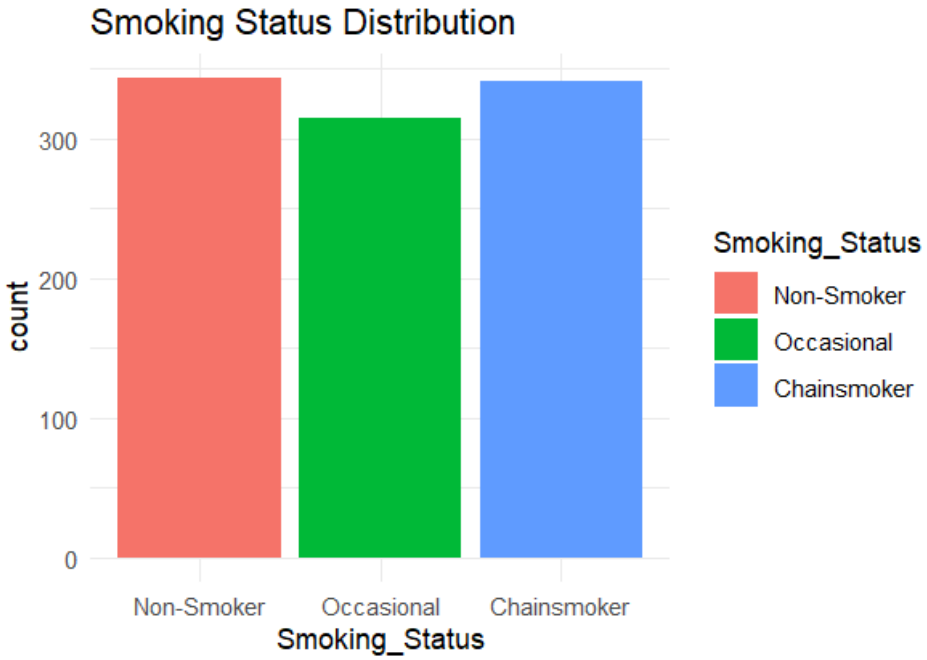

Figure 4: Bar plot of Smoking Status Distribution

1. Demographic Patterns

Age Distribution: The histogram of age revealed the demographic composition of the study population, showing the distribution pattern and identifying the most represented age groups in health risk assessments.

Sex-Based Differences: Analysis of BMI differences between males and females through box plot visualization revealed potential sex-based variations in body composition and anthropometric measurements.

Glucose and Cholesterol Relationship: The scatter plot analysis of glucose versus cholesterol levels, color-coded by elevated risk status, revealed important patterns:

- Distribution of metabolic markers across the population
- Clustering patterns of individuals with elevated risk
- Potential threshold effects in glucose and cholesterol combinations

Cardiovascular Measurements: Analysis of heart rate, blood pressure, and other cardiovascular parameters provided insights into the physiological status of the study population.

2. Lifestyle Factor Assessment

Smoking Status Distribution: The bar chart analysis of smoking status revealed:

- Prevalence of different smoking behaviors in the population
- Distribution across non-smokers, occasional smokers, and chain smokers
- Potential patterns that could inform intervention targeting

Sleep Pattern Analysis: Daily sleeping hours analysis provided insights into sleep hygiene patterns and their potential relationships with other health parameters.

3. Descriptive Analysis

Mean of Age	Median of BMI	Max of Glucose	Minimum of Sleep Hours	Standard Deviation of Heart Rate
54.6 Years	26.0	212 mg/dl	4 Hours	18.5

Sample Characteristics: The descriptive analysis revealed key characteristics of the health assessment population:

- **Mean Age**: The average age of participants provides insight into the demographic profile of individuals undergoing health risk assessment
- **Median BMI**: The median BMI value indicates the central tendency of body mass index in the study population
- **Heart Rate Variability**: Standard deviation of heart rate measurements shows the variability in cardiovascular fitness across participants
- **Maximum Glucose Level**: The highest recorded glucose level indicates the range of metabolic status in the sample
- **Minimum Sleep Duration**: The lowest recorded daily sleeping hours reveals sleep pattern extremes in the population

4. Statistical Correlation Results

Pearson's product-moment correlation

Data: BMI and Systolic BP

t = 20.788,	**df = 998**	**P-value = <2.2e-16**

Alternative hypothesis: true correlation is not equal to 0

95 Percent Confidence Interval

0.5049138	**0.5915356**

Sample Estimates Correlation:

0.5497006

BMI and Systolic Blood Pressure Correlation: The correlation analysis between BMI and systolic blood pressure revealed:

- Correlation coefficient indicating the strength and direction of the relationship
- Statistical significance of the association
- Confidence intervals providing precision estimates
- P-value indicating the probability of observing such correlation by chance

This correlation is particularly important for understanding the relationship between body weight and cardiovascular pressure, which has significant clinical implications for hypertension risk assessment.

CONCLUSIONS AND RECOMMENDATIONS

This health risk assessment analysis revealed several important patterns with significant implications for cardiovascular disease prevention and clinical practice:

1. **Strong BMI-Blood Pressure Relationship**: The significant correlation between BMI and systolic blood pressure confirms the importance of weight management in cardiovascular health.

2. **Distinct Metabolic Risk Profiles**: Glucose and cholesterol patterns show clear clustering in elevated-risk individuals, supporting multi-marker risk assessment approaches.

3. **Demographic Risk Variations**: Sex-based differences in anthropometric measures highlight the need for demographic-specific risk assessment criteria.

4. **Lifestyle Factor Prevalence**: Smoking status distribution reveals important behavioral risk patterns requiring targeted interventions.

Recommendations for Clinical Practice:

- **Implement Multi-Parameter Risk Assessment:**
  - Use combined BMI and blood pressure measurements for more accurate risk stratification
  - Consider glucose and cholesterol levels together when assessing metabolic risk
  - Incorporate demographic factors into risk calculation algorithms
- **Focus on Modifiable Risk Factors:**
  - Prioritize weight management interventions for patients with elevated BMI and blood pressure
  - Develop comprehensive lifestyle modification programs addressing multiple risk factors simultaneously
  - Implement structured smoking cessation programs for identified smokers