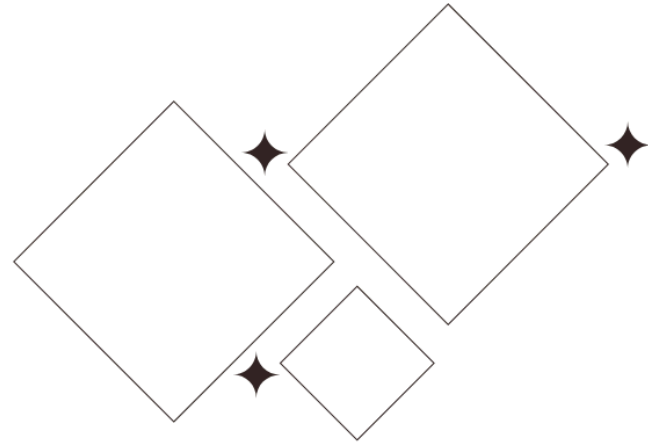# Statistical Exploration of Health Indicators

Vital Signs Diagnosis

Prepared By

Brylle Matthew A. Lupac

Naomi Christienne Tiama

Introduction to the Dataset

The dataset titled "Vital Signs and Diagnosis Data – Group 005" contains health-related data of 1,000 patients. Variables include age, weight, height, blood pressure, cholesterol, glucose levels, heart rate, daily sleep, physical activity hours, and diagnosis labels such as hypertension and smoking status.

This dataset aims to explore correlations and descriptive patterns in vital signs and assess relationships between variables to inform public health interventions.

Methods Used for Analysis

- Software: R with libraries: tidyverse, psych, ggplot2, ggpubr

- Data Preprocessing:

    - Converted numeric variables from character to numeric type.

    - Converted categorical variables (like Sex) into factors.

- Descriptive Statistics: Used psych::describe() for mean, median, SD, etc.

- Visualization:

    - Boxplot of BMI by sex

- Statistical Testing:

    - Pearson Correlation between Systolic and Diastolic Blood Pressure

    - Chi-square test between Hypertension and Smoking Status

Key Results and Figures

Descriptive Statistics Summary

Variable	Mean	Median	Std. Dev	Min	Max
Age	55.22	56.00	20.50	18	90
Weight (kg)	62.93	63.00	9.15	42	87
Height (cm)	152.95	152.50	11.77	130	175
BMI	27.25	26.58	5.39	0	46.75
Systolic BP	134.02	134.00	10.28	97	161
Diastolic BP	82.32	83.00	7.91	56	102
Heart Rate	101.77	101.00	19.03	50	154
Glucose (mg/dL)	130.11	129.00	23.55	16	205
Cholesterol (mg/dL)	191.85	192.00	32.45	111	287
Daily Sleeping Hours	5.31	5.00	1.20	4	9

Boxplot: BMI by Sex
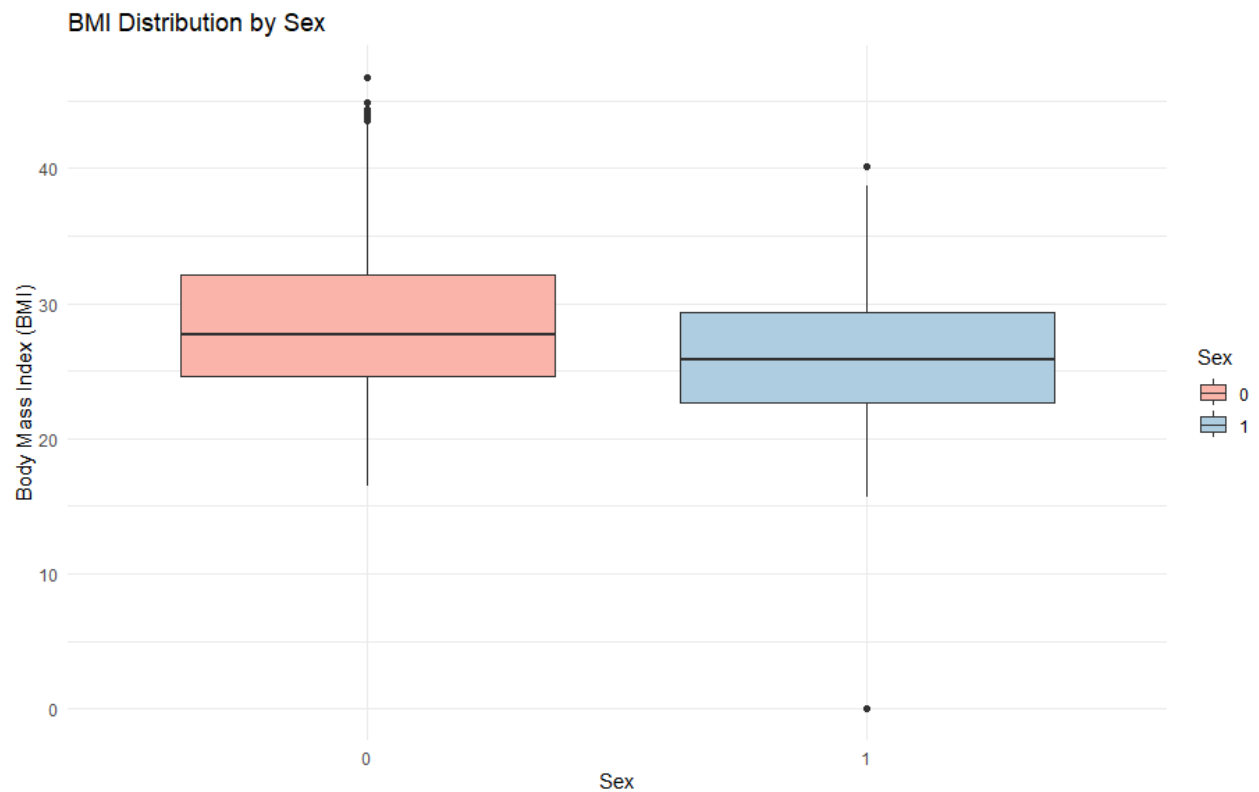
A boxplot showing BMI distribution by sex reveals variability and slight skewness in male participants.

Pearson Correlation: Systolic vs Diastolic Blood Pressure

- Correlation coefficient: 0.77

- $t = 37.95$, df = 992

- p-value $< 2.2e-16$

- 95% CI: [0.743, 0.794]
- There is a strong, statistically significant positive correlation between systolic and diastolic blood pressure.

D. Chi-Square Test: Hypertension × Smoking Status

- $\chi^2 = 134.01$

- df = 8

- p-value < 2.2e-16
- There is a significant association between a patient's smoking status and the presence of hypertension.

Interpretation and Brief Conclusion

This exploratory analysis revealed key health trends in the dataset:

- Blood pressure values are strongly interrelated.

- BMI varies significantly between sexes.

- Hypertension is significantly associated with smoking behavior, indicating a possible risk group.

Despite these findings, no single variable alone explains all health outcomes. Further multivariate modeling (e.g., logistic regression) is recommended for predicting specific diagnoses.