

GITHUB PORTFOLIO

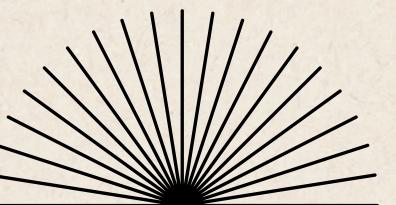
DEMOGRAPHIC BEHAVIORAL DATA

E REPORT

NAME OF PROJECT:
GITHUB PORTFOLIO

PRESENTED BY:
Morales, Enrico James E. &
Pinzon Jazreel S.

MAPUA UNIVERSITY
BI120L_CON29



Introduction

This report explores demographic and behavioral health data from 1,000 individuals to identify patterns and relationships among variables such as age, sex, BMI, smoking habits, physical activity, and health literacy. Using R programming, the data was cleaned, summarized, and visualized to uncover trends that may inform public health insights. Key objectives include examining BMI distribution by sex, the impact of physical activity on health metrics, and correlations among lifestyle factors.



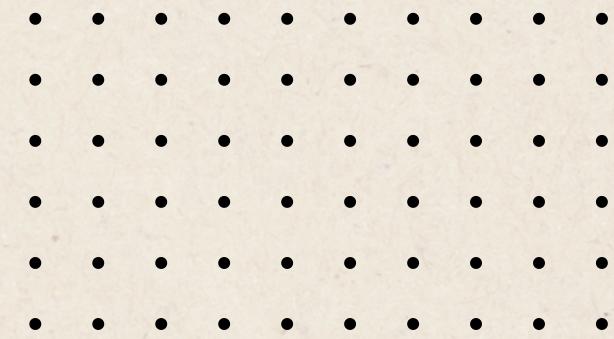
METHODS USED FOR ANALYSIS

The analysis was conducted using R, employing packages such as tidyverse, janitor, psych, ggplot2, and corrplot. Data cleaning involved renaming columns, removing variables with excessive missing values, and converting key variables to factors. Descriptive statistics were generated using summary() and describe(). Visualizations including histograms, bar plots, box plots, and scatter plots were created with ggplot2. Statistical analysis included a t-test for BMI differences by sex and Pearson correlation to assess relationships, with correlation matrices visualized using corrplot().

Statistical tools and tests used include:

- Descriptive statistics (summary(), describe()) to explore distributions of variables
- Pearson correlation test (cor.test()) to assess the linear relationship between BMI and physical activity
- Welch Two-Sample t-test (t.test()) to evaluate BMI differences between male and female participants
- Correlation matrix (cor()) to examine relationships among numeric variables
- Visual tools such as histograms, bar plots, box plots, scatter plots, and correlation heatmaps to illustrate findings

KEY RESULTS AND PLOTS



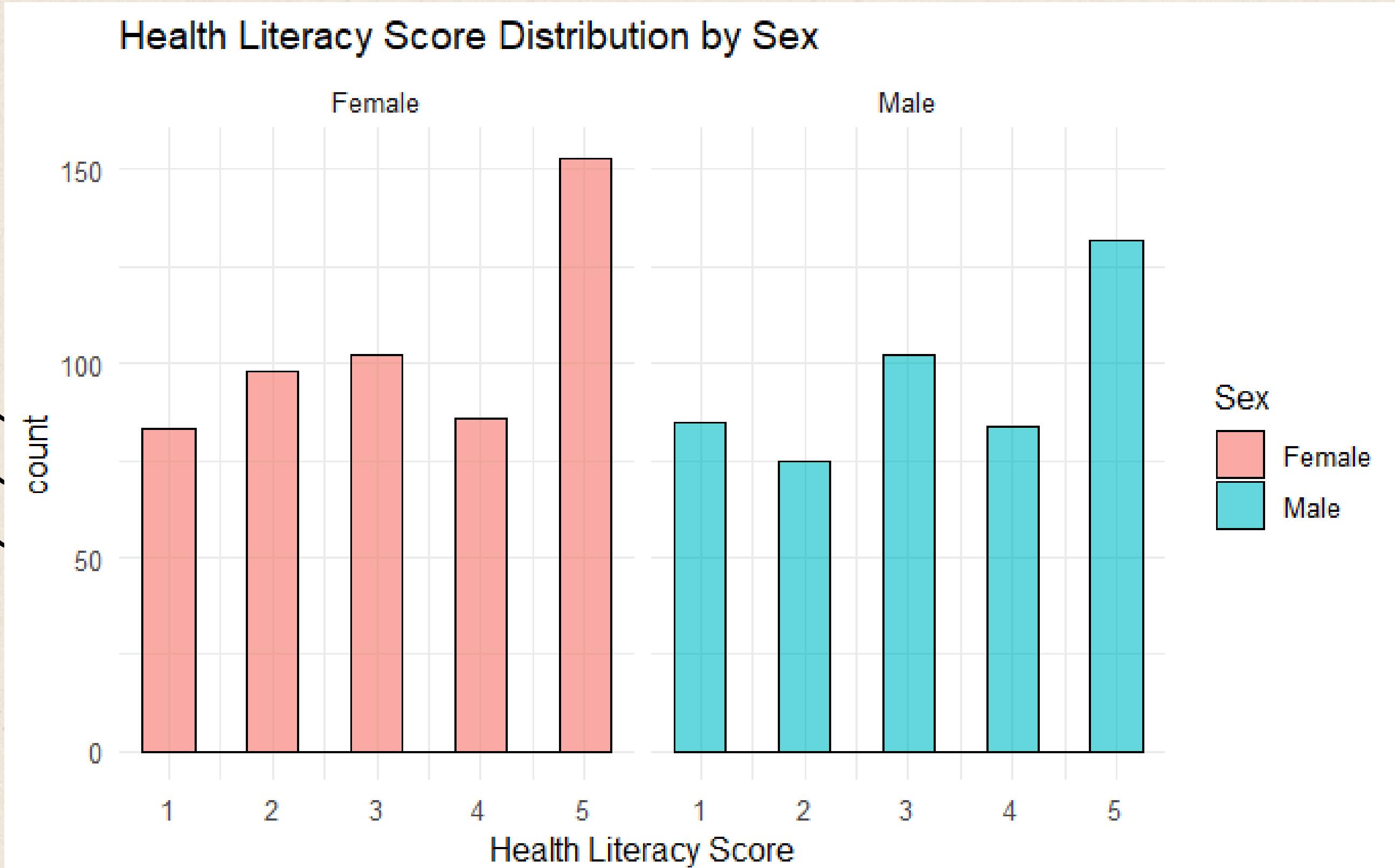
08/10

- Histogram of Health Literacy Score by Sex showed a relatively even distribution, with males slightly more concentrated at higher scores.
- Bar Plot revealed that Chainsmokers were the most common smoking group, followed by Occasional and Non-Smokers.
- Boxplot comparing BMI by sex showed that females had a higher median BMI and a wider spread of values than males.
- Scatter Plot between BMI and Physical Activity indicated no clear trend; a fitted regression line confirmed a weak, non-significant correlation.
- Correlation Test found a very weak, non-significant relationship between BMI and physical activity ($r = 0.027$, $p = 0.402$).
- T-test showed a significant difference in BMI between sexes ($p < 0.001$), with females having a higher mean BMI.
- Correlation Matrix & Plot highlighted strong negative correlation between height and BMI, and moderate positive correlations between BMI, weight, and socioeconomic status.

PLOTS

05/10

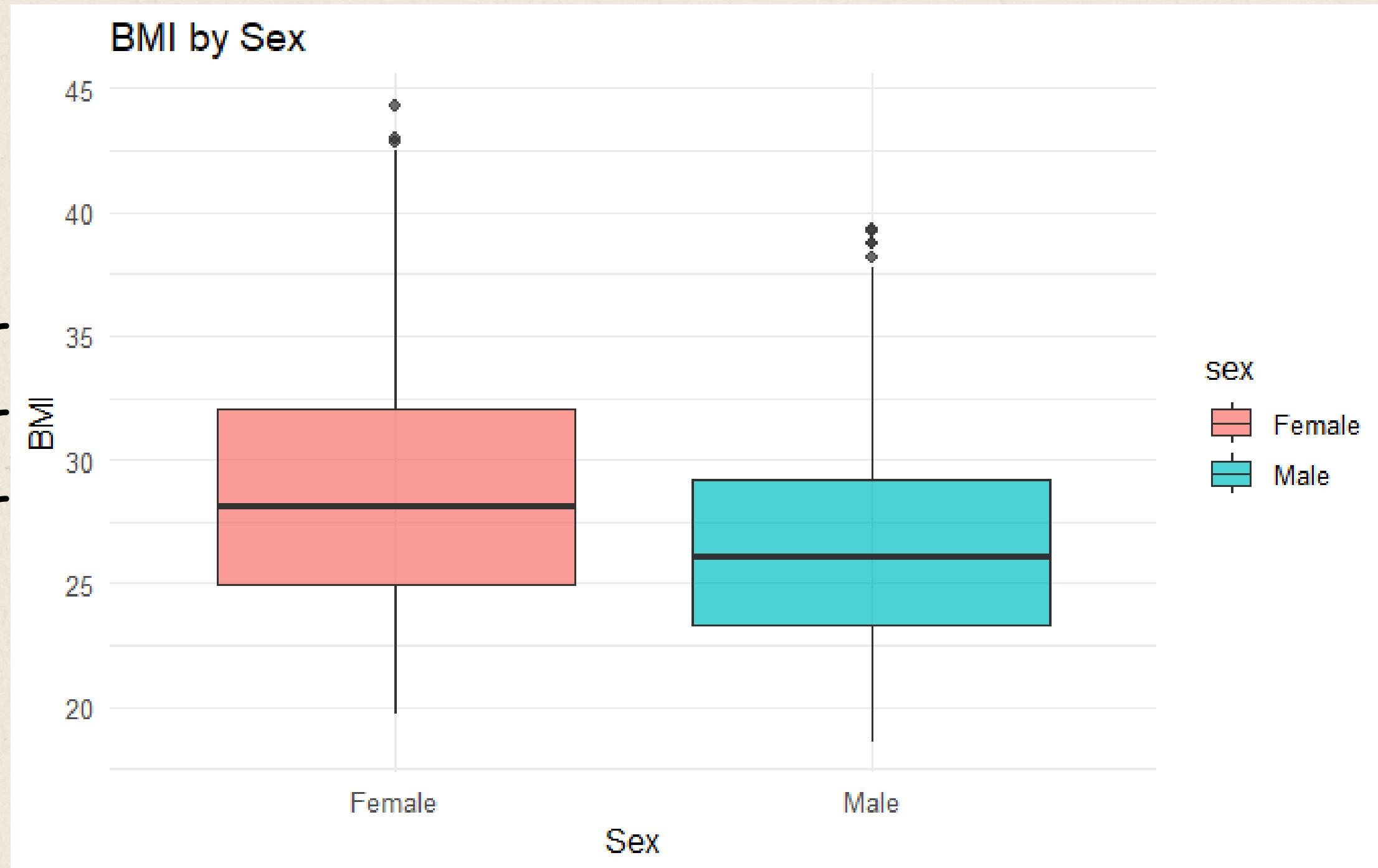
Health Literacy Score Distribution by Sex



(Figure 1)

The distribution shows that both males and females vary in health literacy, but males tend to cluster slightly toward higher scores.

PLOTS

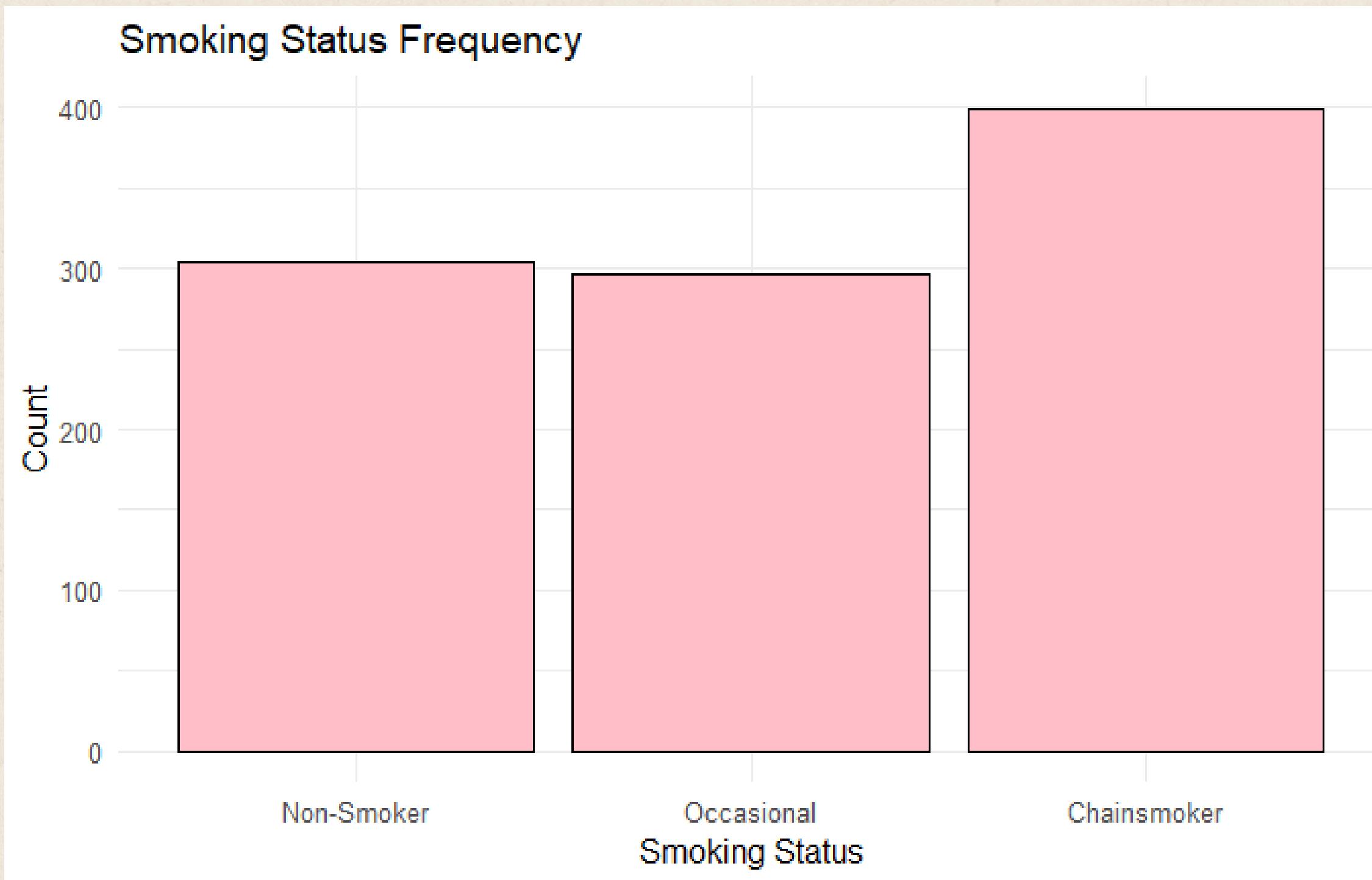


(Figure 2)

Females displayed a higher median BMI and a broader range compared to males, indicating sex-related differences in body composition.

PLOTS

05/10

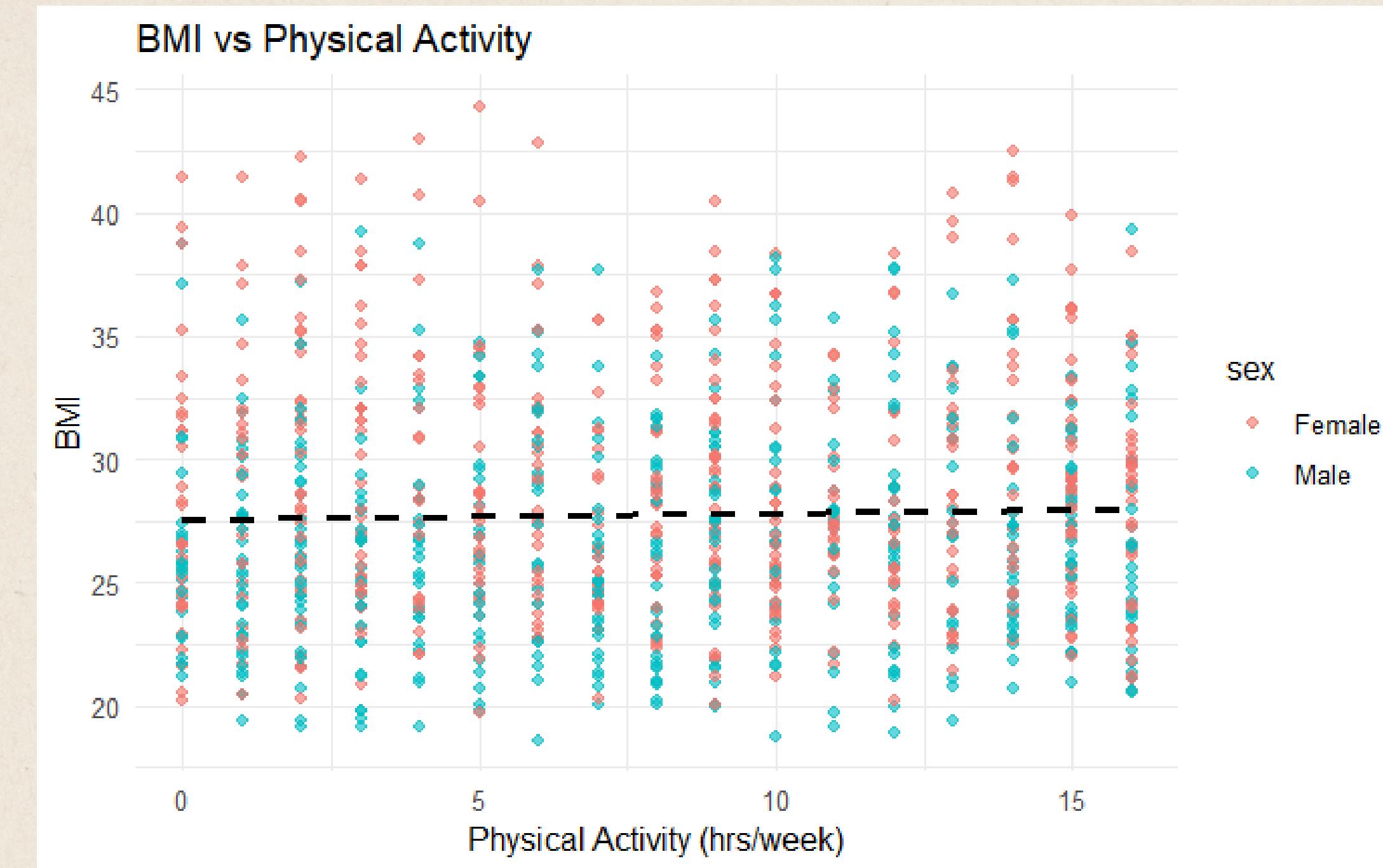


(FIGURE 3)

Chainsmokers were the most prevalent group, suggesting a concerning trend in high-frequency smoking within the population.

PLOTS

05/10



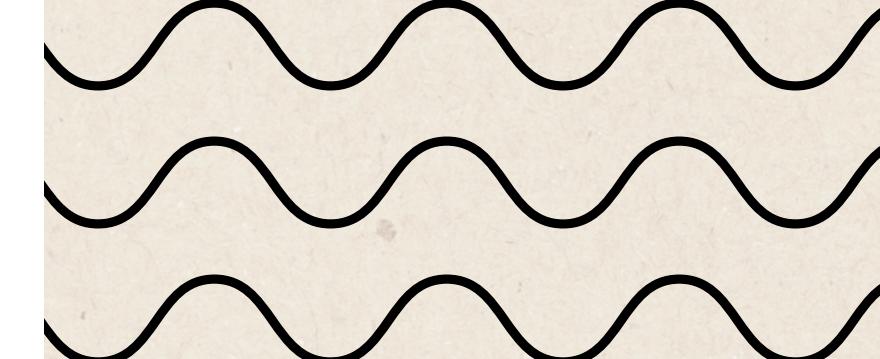
Scatter Plot (Figure 4)

BMI vs Physical Activity

No clear relationship is observed; physical activity hours do not strongly predict BMI. The regression line confirms this weak association.

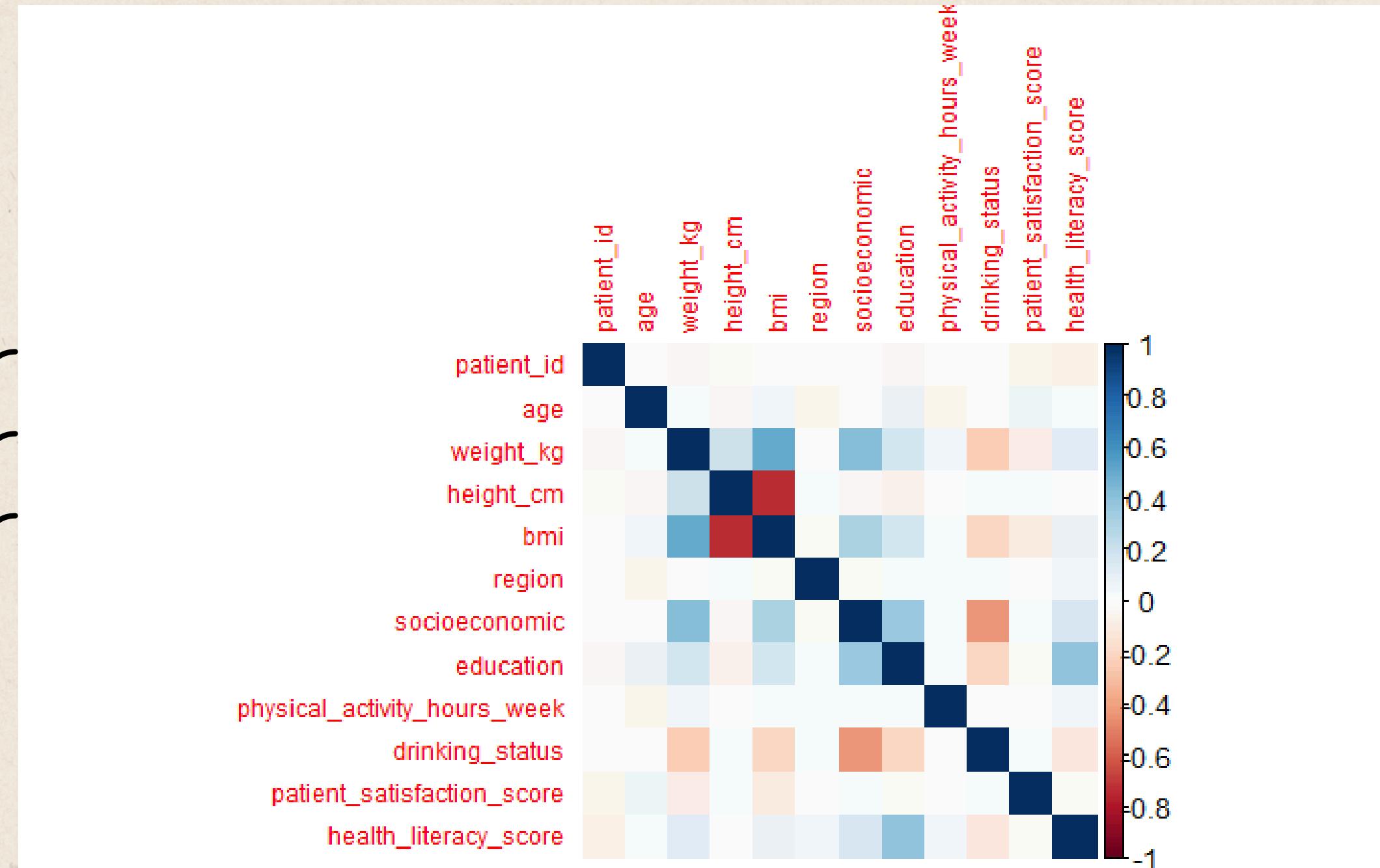
Correlation Test (BMI & Physical Activity)

A very weak, non-significant positive correlation ($r = 0.027$, $p = 0.402$) suggests that increased physical activity does not clearly affect BMI in this sample.



PLOTS

05/10



Corr PPlot (Figure 5)

Height and BMI were strongly negatively correlated ($r = -0.73$), while BMI showed moderate positive links with weight and socioeconomic status. Education and health literacy were also positively associated.

INTERPRETATION AND CONCLUSION

THE ANALYSIS REVEALED NOTABLE BEHAVIORAL AND HEALTH-RELATED TRENDS ACROSS DEMOGRAPHIC GROUPS. FEMALES EXHIBITED HIGHER BMI LEVELS THAN MALES, SUPPORTED BY A STATISTICALLY SIGNIFICANT T-TEST RESULT. HOWEVER, PHYSICAL ACTIVITY SHOWED NO SIGNIFICANT CORRELATION WITH BMI, SUGGESTING OTHER FACTORS MAY INFLUENCE BODY COMPOSITION MORE STRONGLY. THE CORRELATION MATRIX HIGHLIGHTED MEANINGFUL ASSOCIATIONS SUCH AS THE STRONG INVERSE RELATIONSHIP BETWEEN HEIGHT AND BMI, AND A POSITIVE LINK BETWEEN EDUCATION AND HEALTH LITERACY. THESE FINDINGS underscore THE COMPLEX INTERPLAY BETWEEN DEMOGRAPHIC, BEHAVIORAL, AND HEALTH INDICATORS, EMPHASIZING THE NEED FOR MULTIFACTORIAL APPROACHES WHEN ADDRESSING PUBLIC HEALTH CONCERNS.