

**Princess Sumaya University for Technology**

King Hussein School for Computing Sciences

**Studying Factors Affecting Obesity in Jordan**

**Analytical Report**

**Data Visualization**

Summer - 2023/2024

**Supervised by:**

Dr. Amera Al-Amery

**Prepared by:**

Dalia Tamimi 20200288

Saja Salameh 20210562

Yumna Al-Natsheh 20221019

**TABLE OF CONTENT**

[TABLE OF FIGURES 3](#_Toc174226736)

[Team contribution: 4](#_Toc174226737)

[1. Executive Summary 5](#_Toc174226738)

[2. Introduction 6](#_Toc174226739)

[3. Discussion 7](#_Toc174226740)

[**3.1 Main statistical features of the variables in the dataset** 7](#_Toc174226741)

[**3.2 Central Statistical Relationships and Correlations** 7](#_Toc174226742)

[**3.3 Data Visuals** 8](#_Toc174226743)

[**3.3.1 Data cleaning** 8](#_Toc174226744)

[**3.3.2 Visualization** 8](#_Toc174226745)

[4. Recommendations and Further Work 12](#_Toc174226746)

# TABLE OF FIGURES

[Figure 1: Gender vs Obesity 8](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226643)

[Figure 2: Gender vs Obesity 8](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226644)

[Figure 3: Smoke vs Obesity 9](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226645)

[Figure 4: Smoke vs Obesity 9](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226646)

[Figure 5: Age vs Obesity Indicator 9](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226647)

[Figure 6: Age vs Obesity Indicator 9](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226648)

[Figure 7: Mental Distress Level vs Obesity Indicator 10](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226649)

[Figure 8: Mental Distress Level vs Obesity Indicator 10](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226650)

[Figure 9: Resilience Level vs Obesity Indicator 10](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226651)

[Figure 10: Resilience Level vs Obesity Indicator 10](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226652)

[Figure 11: PSQI vs Obesity Indicator 11](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226653)

[Figure 12: PSQI vs Obesity Indicator 11](file:///C:\Users\billo89\Downloads\Analytical%20Report(00000000).docx#_Toc174226654)

# Team contribution:

1. Dalia Tamimi

* Explored and cleaned the Resilience Scale (RS) and Demographic and Lifestyle Factors and did the code for it
* Analytical report

1. Saja Salameh

* Cleaned and explored the Pittsburgh Sleep Quality Index (PSQI) and did the code for it.
* Creation of Visualizations for Data Representation
* Presentation

1. Yumna Al Natsheh

* Explored and cleaned Kessler Psychological Distress Scale (K10+) and did the code for it.
* Technical report.

# 1. Executive Summary

The primary objective of this study is to investigate and uncover the factors influencing obesity in Jordan using data collected from over 900 participants in Jordan. Both exploratory data analysis (EDA) and explanatory analysis are utilized to understand correlations and statistical features, using data visualizations. This report focuses on the factors affecting obesity, uncovering key insights and highlighting important trends and patterns. It also emphasizes the need for more study in this area. The paper ends with recommendations for taking proactive measures and makes recommendations for future research for the analysis.

# 2. Introduction

Obesity is an increasing medical issue in Jordan, as the incidence of obesity grows it becoming increasingly important to understand the underlying causes Jordan's Ministry of Health recognized the critical need to address this issue and started searching and analyzing this study to know the factors of obesity in the country.

this study aims to achieve the goal by focusing on key factors such as The Pittsburgh Sleep Quality Index (PSQI), Resilience Scale (RS), Kessler Psychological Distress Scale (K10+), Demographic and Lifestyle Factors such as gender, age, family history with overweight, smoking status, transportation and city of residence. by examining the complex relationships between these factors and obesity the study seeks to uncover correlations, statistical relations, and underlying trends and the analysis employs exploratory data analysis (EDA) techniques and several statistical tests such as chi-Squared and Fisher's exact tests also help in identifying key drivers of obesity, understand their interactions, and provide recommendations.

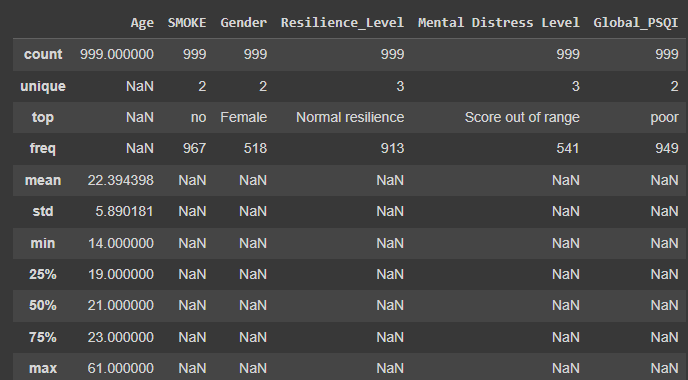
# 3. Discussion

Let’s take a look at the distribution of the data in each column:

We assume all the main columns are normally distributed.

## **3.1 Main statistical features of the variables in the dataset**

Table 1: Statistical calculation



### **3.2 Central Statistical Relationships and Correlations**

In our analysis of the dataset, we aimed to determine the significance of relationships between various variables and obesity. We employed two statistical tests for this purpose:

* Chi-Square Test: Applied to all columns except the 'smoke' variable. The Chi-Square test assesses whether there is a significant association between categorical variables.
* Fisher's Exact Test: Used for the 'smoke' variable due to its highly imbalanced counts, which could affect the reliability of the Chi-Square test.

**If the calculated p-value is less than 0.05, the relationship between the variables is considered significant. Conversely, if the p-value exceeds 0.05, the relationship is not considered significant. For the age variable, we first converted it into categories to apply the Chi-Square tests and visualize its distribution effectively*.***

## **3.3 Data Visuals**

### **3.3.1 Data cleaning**

A number of data preparation procedures were carried out in the early phases of the analysis using Python as our main language for this part of our work. One way to improve the readability and usability of the columns was to rename them. Additionally, we imputed missing values using the mode for categorical variables and mean for numerical(normally distributed) variables, To maintain a representative and clean dataset for further analysis, We add a column for the global range of Pittsburgh Sleep Quality Index (PSQI) that 0 to 21, with higher scores indicating worse sleep quality with 0–4 indicating “good” sleep and 5–21 indicating “poor” sleep and column for Resilience Scale (RS) for the score range 6 - 13 Low resilience ,14 - 21 Normal resilience ,22 - 30 High resilience, also for Kessler Psychological Distress Scale (K10+) 10 - 19 Likely to be well, 20 - 29 Likely to have a moderate disorder,30 - 50 Likely to have a severe disorder.

This will help us to make a good visualization and assumptions.

### **3.3.2 Visualization**

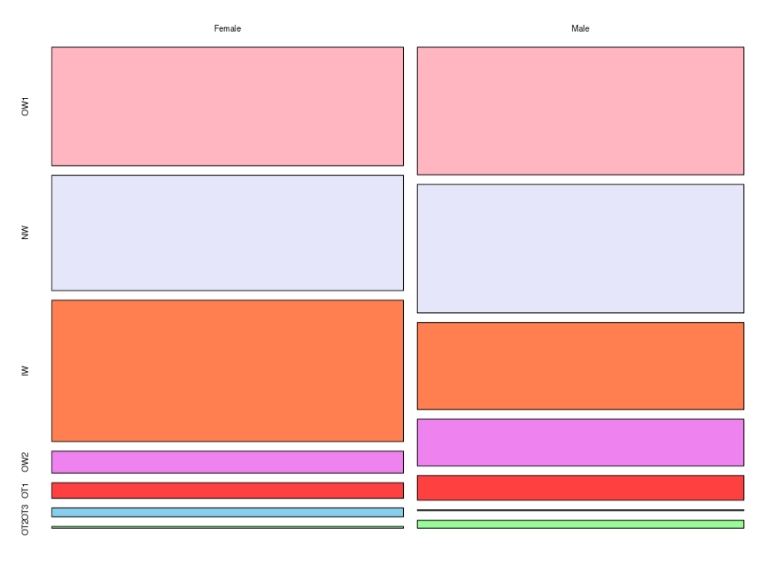
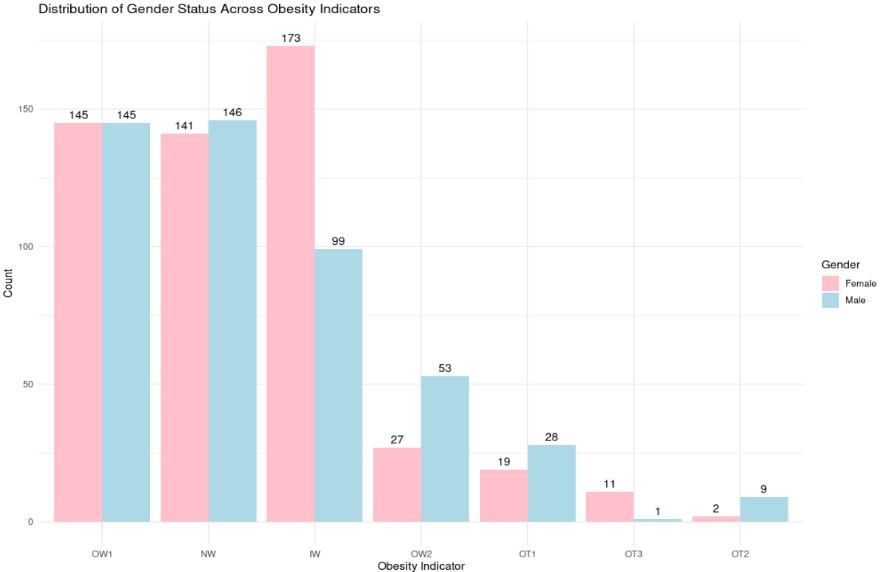
****

Figure 1: Gender vs Obesity

Figure 2: Gender vs Obesity

The visual representations did not provide a very clear association between genders.

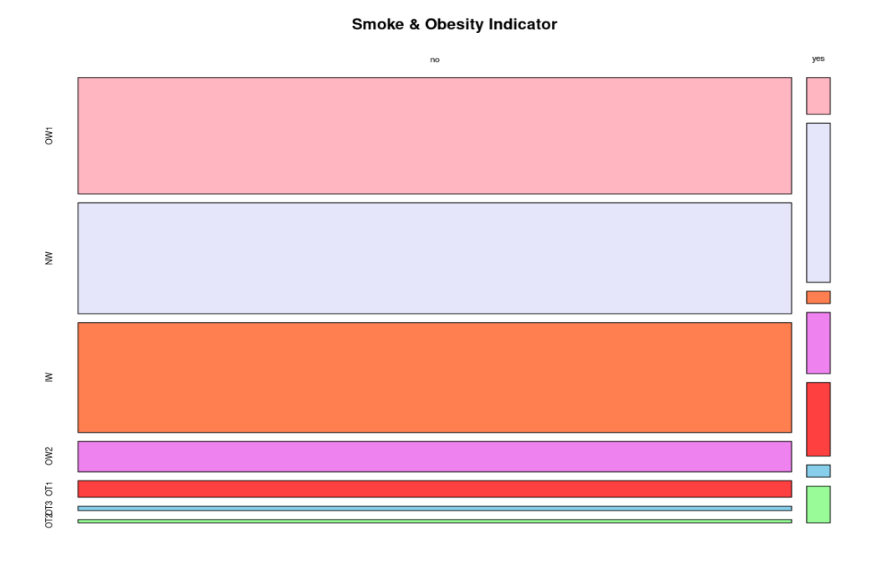
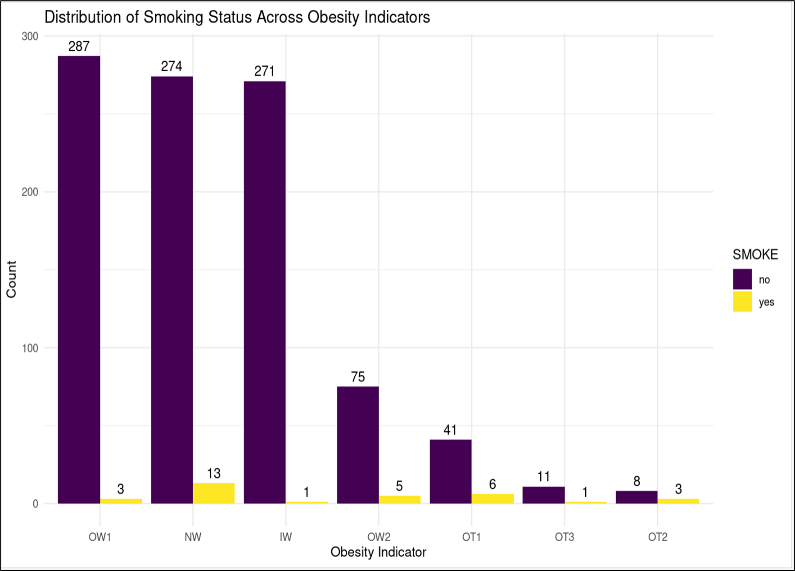
****

Figure 3: Smoke vs Obesity

Figure 4: Smoke vs Obesity

The visual representations did not provide a very clear association between smoking and obesity.

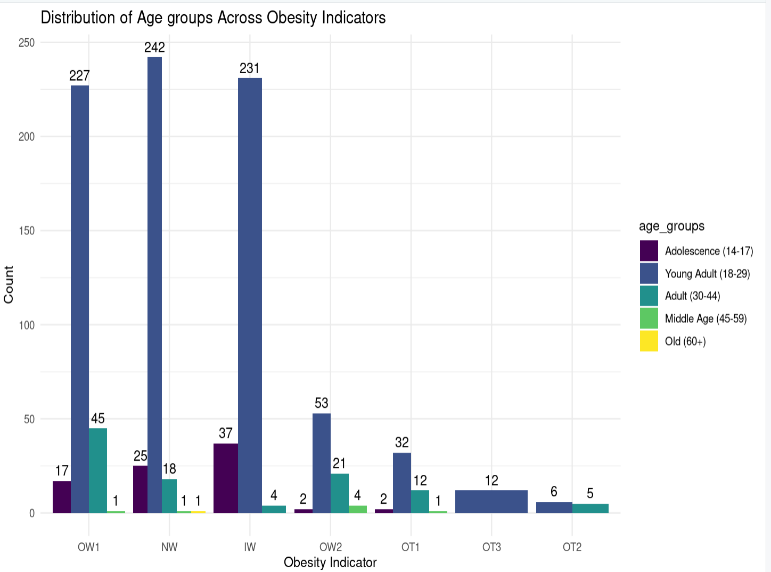
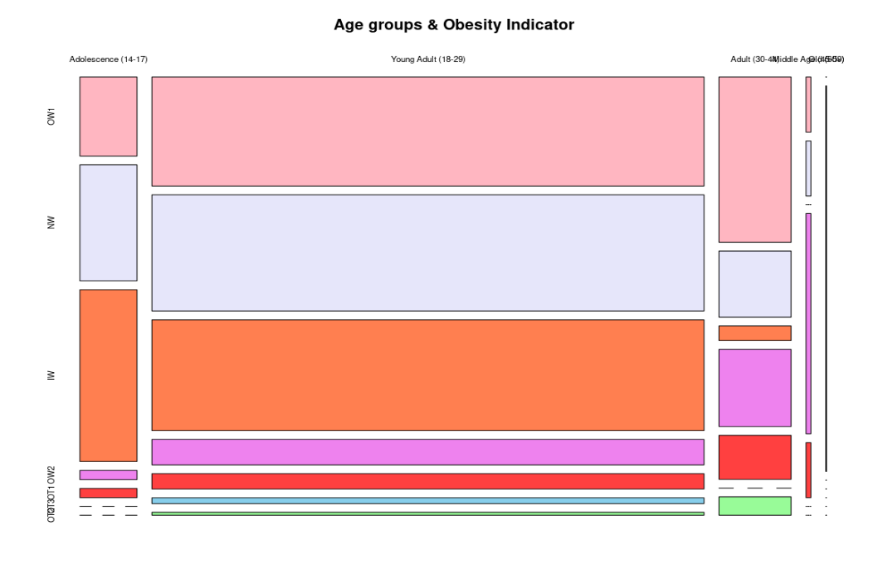


Figure 5: Age vs Obesity Indicator

Figure 6: Age vs Obesity Indicator

The visual representations provide a very clear association between age and obesity particularly in adults

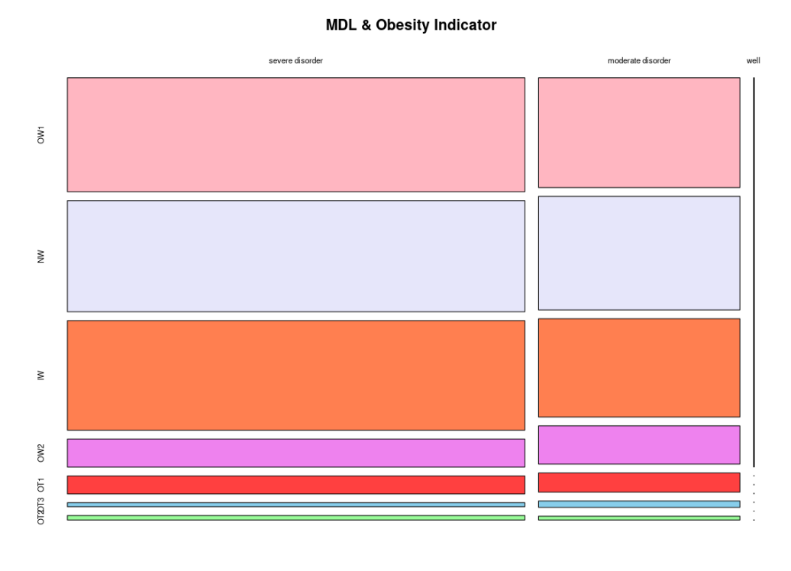
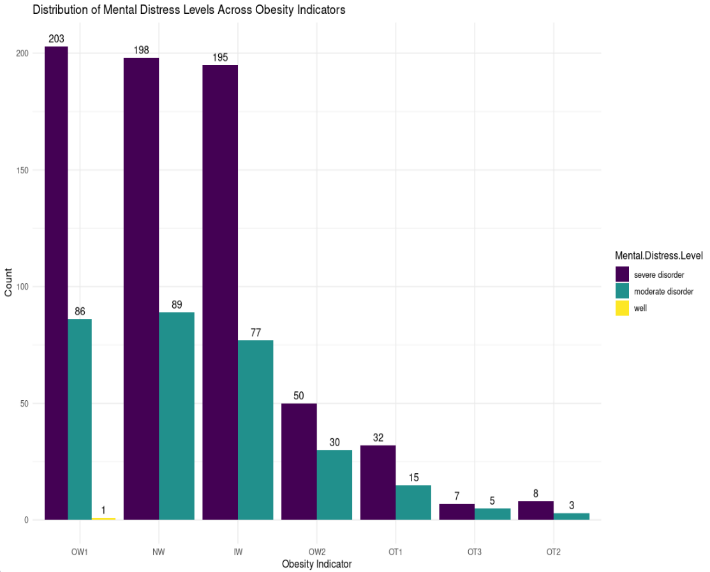
****

Figure 7: Mental Distress Level vs Obesity Indicator

Figure 8: Mental Distress Level vs Obesity Indicator

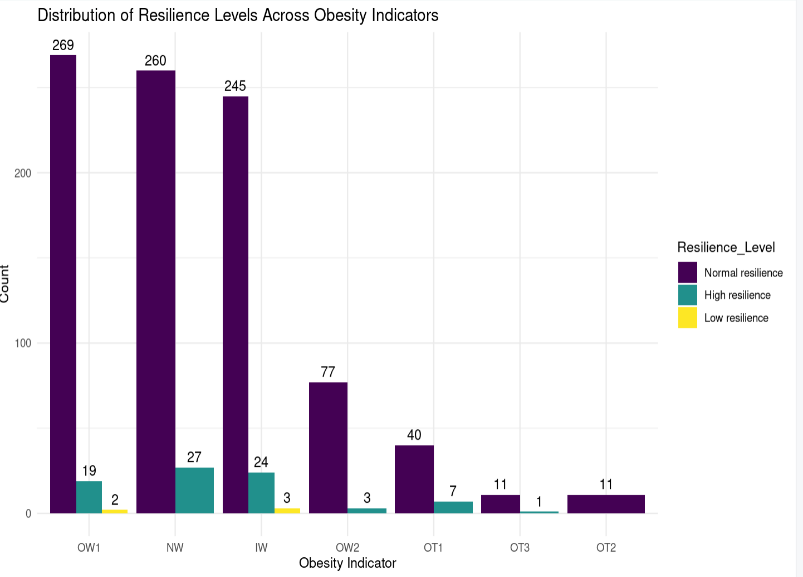
****The visual representations provided a very clear association between Mental Distress Level and obesity

Figure 9: Resilience Level vs Obesity Indicator

The visual representations did not provide a very clear association betweenResilience Level association and obesity

Figure5:

Figure 10: Resilience Level vs Obesity Indicator

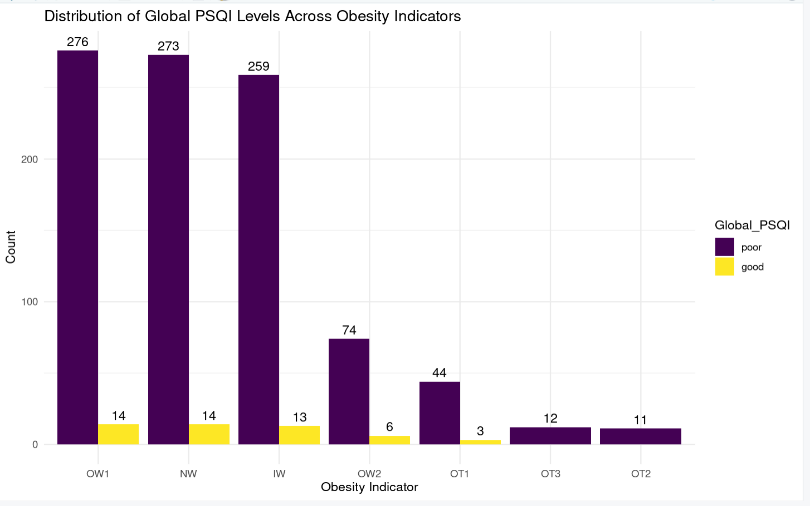


Figure 11: PSQI vs Obesity Indicator

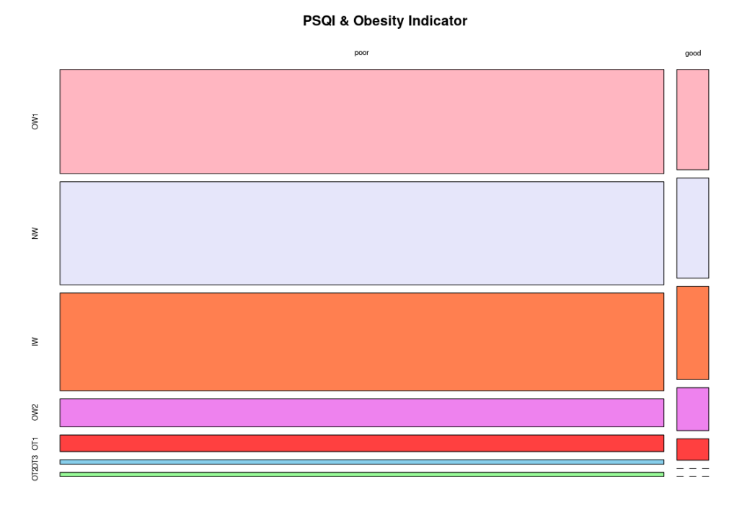


Figure 12: PSQI vs Obesity Indicator

The visual representations provided a very clear association between PSQI and obesity

Table 2:P-value between column and Obesity Indicator

|  |  |  |  |
| --- | --- | --- | --- |
| Column | Test | P-Value | Statistically significant |
| Gender | Chi test | 1.952653e-07 | Statistically significant |
| Smoke | Fishers Exact Test | 0.0004998 | Statistically significant |
| Age | Chi test | 2.2e-16 | Statistically significant |
| Metal Distress Level | Chi test | 0.9247 | Not Statistically significant |
| Resilience level | Chi test | 0.4402 | Not Statistically significant |
| PSQI | Chi test | 0.8679 | Not Statistically significant |

# 4. Recommendations and Further Work

* **For age**

Help adults address emotional eating, develop healthier life style habits, as well as improve self-discipline, through offering workshops and seminars on nutrition, exercise, and stress management to educate adults on maintaining a healthy weight.

Also, promote "support groups" or "buddy systems", where individuals can share their experiences, motivate each other, and offer practical advice.

* **For smoking**

Encourage the adoption of healthy lifestyle alternatives that can replace smoking and support weight management. For example, promote physical activities like walking, cycling, or other forms of exercise that can help individuals in managing stress and reduce the need for smoking while also combating obesity.

* **For gender**

Regulate the marketing of unhealthy foods, particularly those targeted at children and teenagers.

Furthermore, address social and cultural factors that may contribute to gender disparities in obesity. For instance, women may face societal pressures related to body image, while men might underreport weight-related issues.