

# Obesity Risk

Under supervision :

***Eng:Abd Elrahman Eid***

***Made by:***

*Hajar El Behairy*

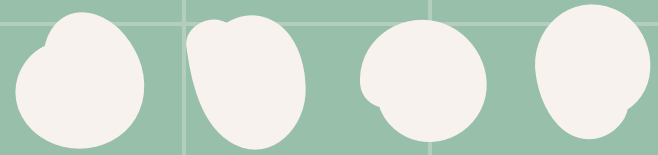
*Nagwa Mohamed*

*Mohamed Salama*



# Agenda

- Introduction
  - Problem Statement
  - Key Findings
  - Approach
  - Results
  - Visualization
  - Conclusion & Recommendations
- 
- A decorative wavy line in a light orange color, starting from the bottom left and curving upwards and to the right, ending near the bottom right corner of the slide.



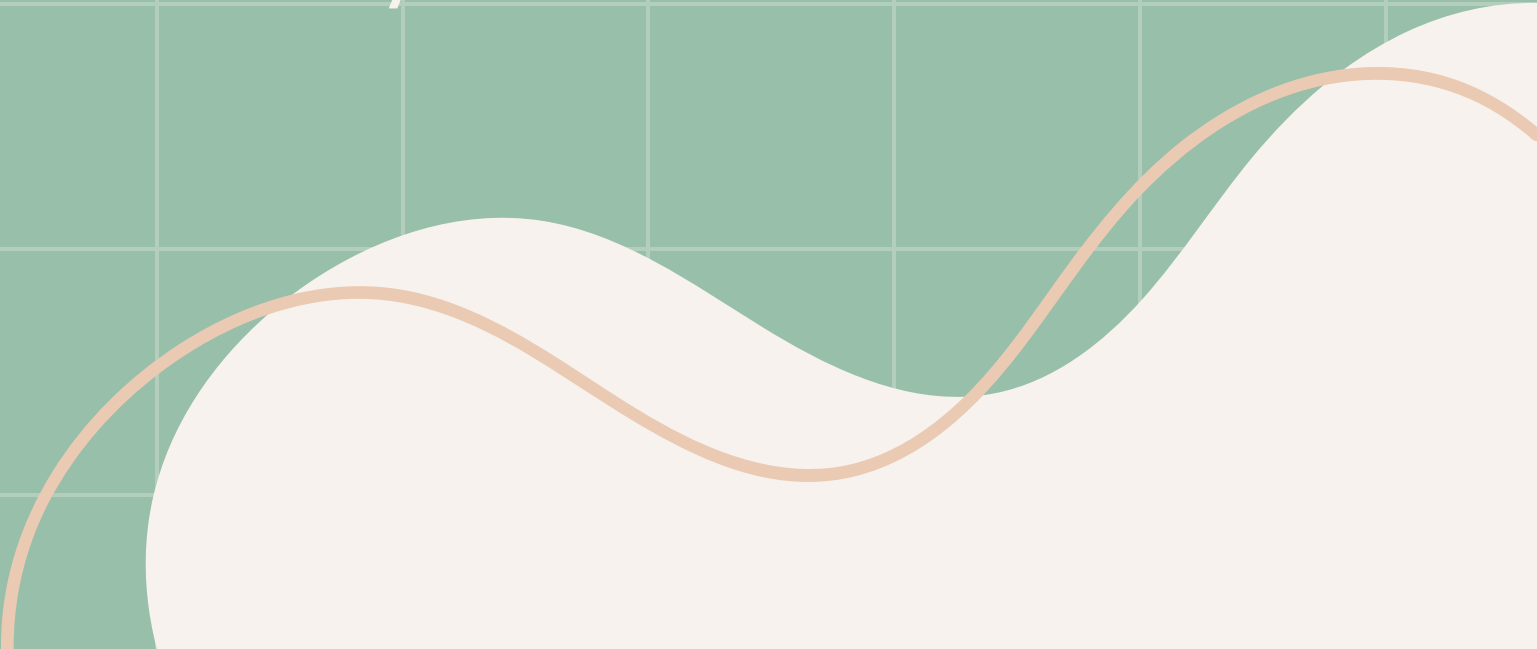
# Problem

Predicting obesity risk based on lifestyle and demographic factors.

Multi-class classification problem to categorize individuals into different obesity risk levels.



# Key Findings

- Physical activity levels, caloric intake, and age are significant factors influencing obesity risk.
  - Various models were tested, with the highest accuracy achieved using **XGBoost (92% on train, 91% on test)**.
  - Feature importance analysis revealed that lifestyle factors are strong predictors of obesity risk.
- 
- A decorative wavy line in a light orange color, starting from the bottom right and curving upwards and to the left, ending near the bottom center of the slide.

# Approach

## ***Data Preprocessing:***

This step involves managing encoding categorical variables and scaling numerical features to ensure consistency and enhance model performance.

## ***Feature Engineering:***

Identifying and selecting the most relevant features that contribute significantly to obesity risk prediction.

# Approach

- ***Model Training:***

Implementing and evaluating various algorithms, including:

- Logistic Regression
- Decision Tree
- Random Forest
- XGBoost

***Hyperparameter Tuning:***

Applying GridSearchCV to optimize model performance by selecting the best combination of hyperparameters

# Results

**Best model:**

**XGBoost** with approximately **92% accuracy** on the *training* set and **91% accuracy** on the *test* set.

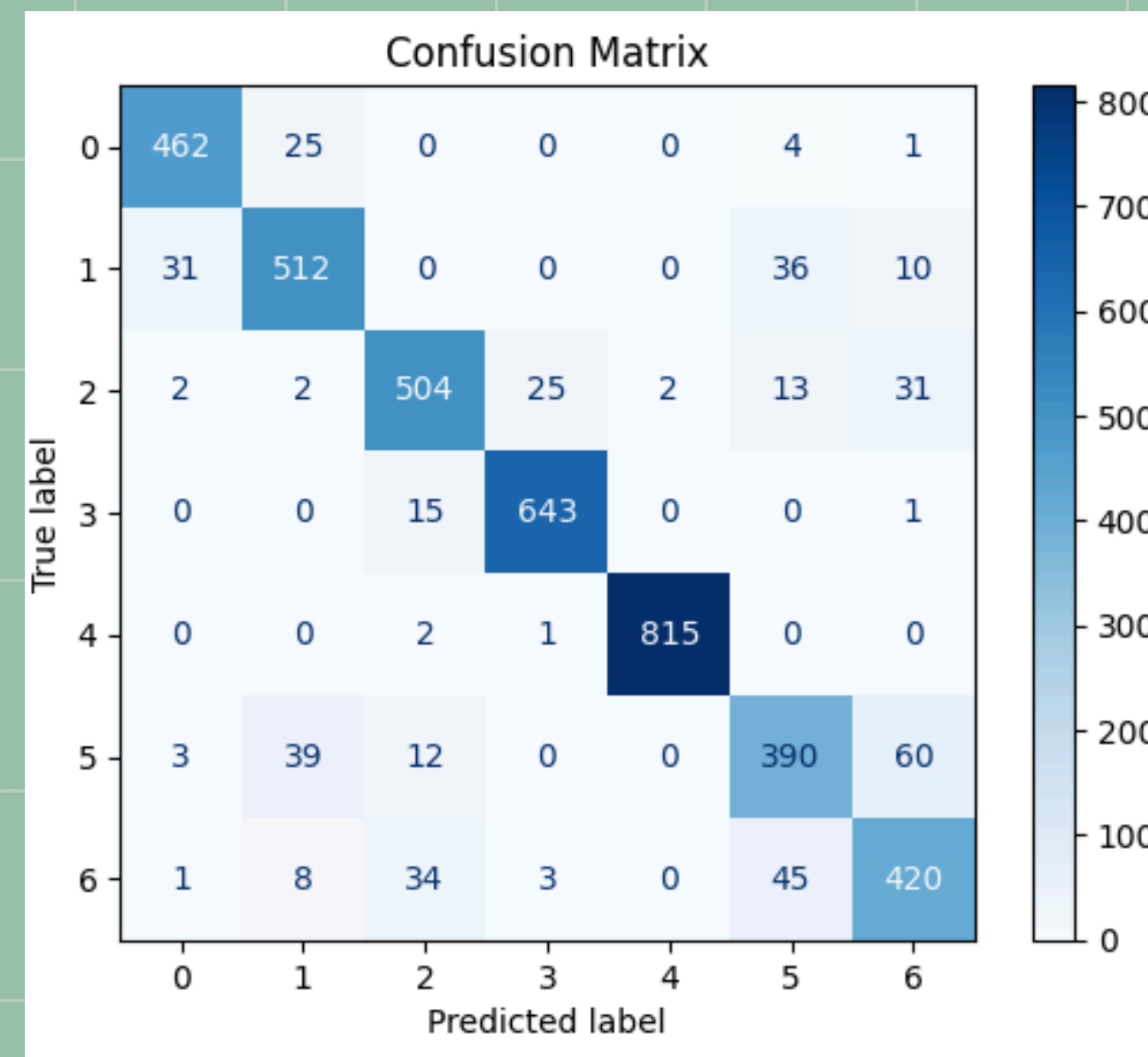
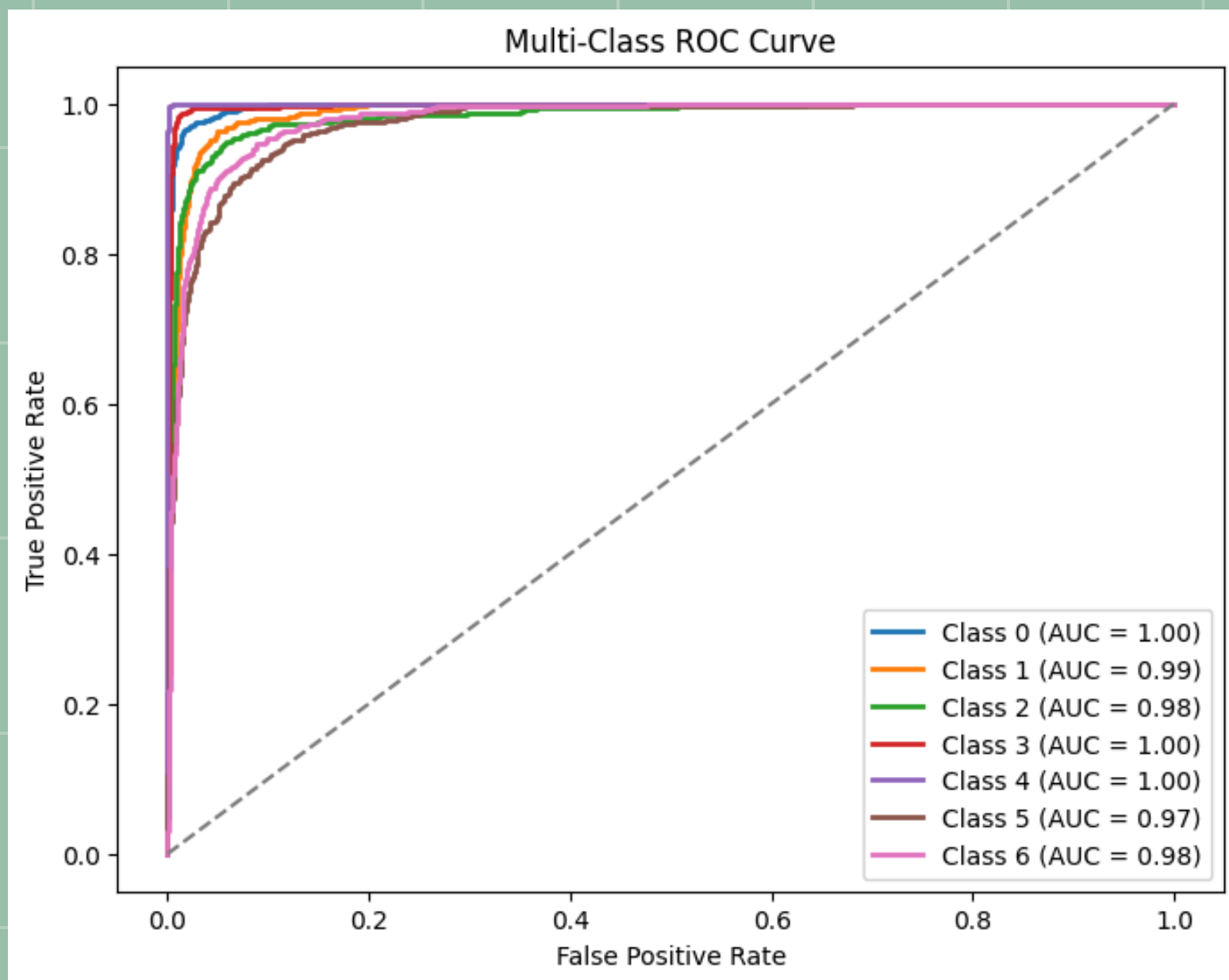
**XGBoost** outperformed other models by effectively capturing complex patterns in the data.

Various models were tested, with the highest accuracy achieved using XGBoost (**92% on train, 91% on test**).



# Results

Confusion matrices and classification reports indicate strong performance across all categories.



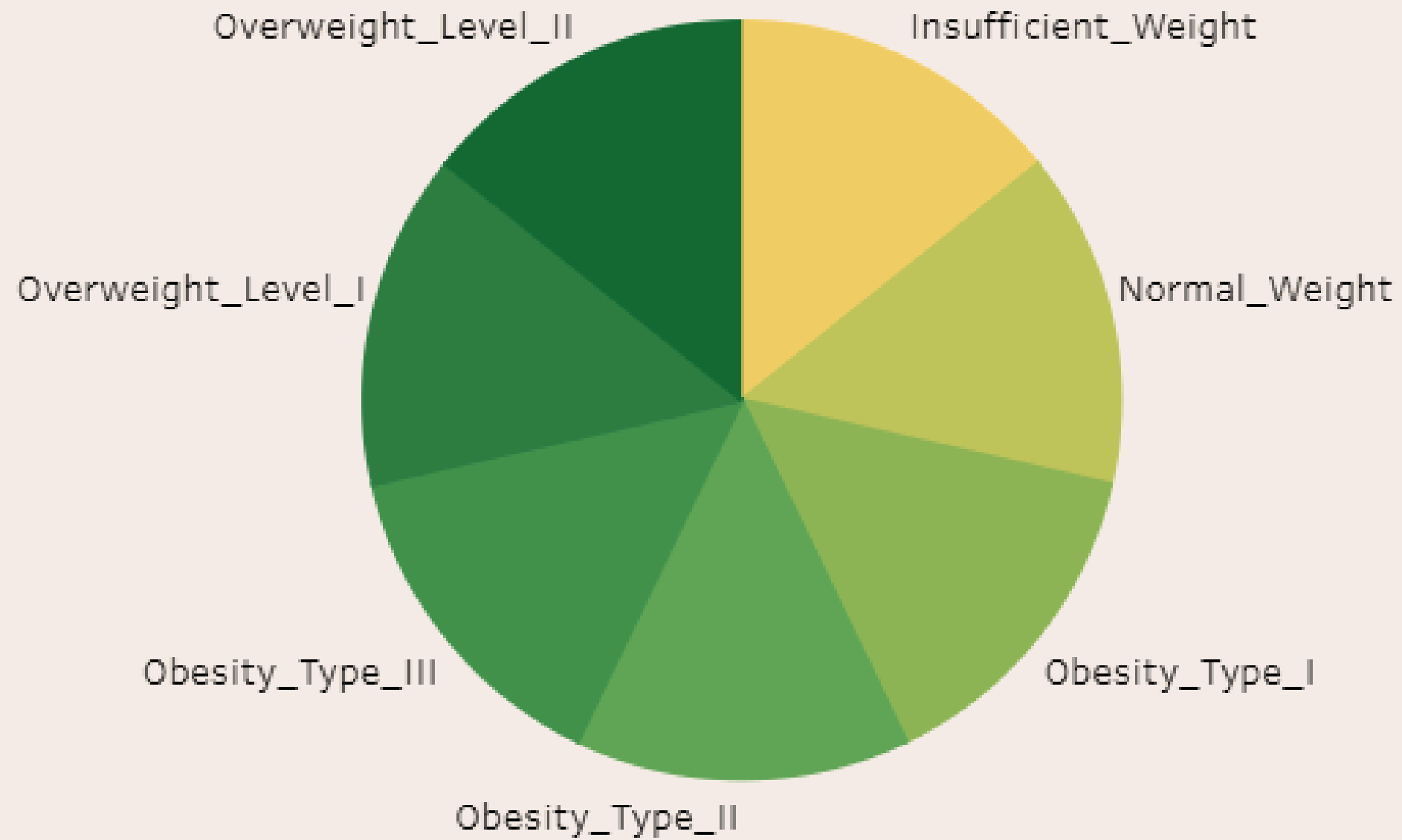


# Visualizations

	count	mean	std	min	1%	25%	50%	75%	99%	max
id	20758.000000	10378.500000	5992.462780	0.000000	207.570000	5189.250000	10378.500000	15567.750000	20549.430000	20757.000000
Age	20758.000000	23.841804	5.688072	14.000000	16.913841	20.000000	22.815416	26.000000	41.000000	61.000000
Height	20758.000000	1.700245	0.087312	1.450000	1.507106	1.631856	1.700000	1.762887	1.889104	1.975663
Weight	20758.000000	87.887768	26.379443	39.000000	42.000000	66.000000	84.064875	111.600553	137.855041	165.057269
FCVC	20758.000000	2.445908	0.533218	1.000000	1.000000	2.000000	2.393837	3.000000	3.000000	3.000000
NCP	20758.000000	2.761332	0.705375	1.000000	1.000000	3.000000	3.000000	3.000000	4.000000	4.000000
CH2O	20758.000000	2.029418	0.608467	1.000000	1.000000	1.792022	2.000000	2.549617	3.000000	3.000000
FAF	20758.000000	0.981747	0.838302	0.000000	0.000000	0.008013	1.000000	1.587406	3.000000	3.000000
TUE	20758.000000	0.616756	0.602113	0.000000	0.000000	0.000000	0.573887	1.000000	2.000000	2.000000

Descriptive Analysis

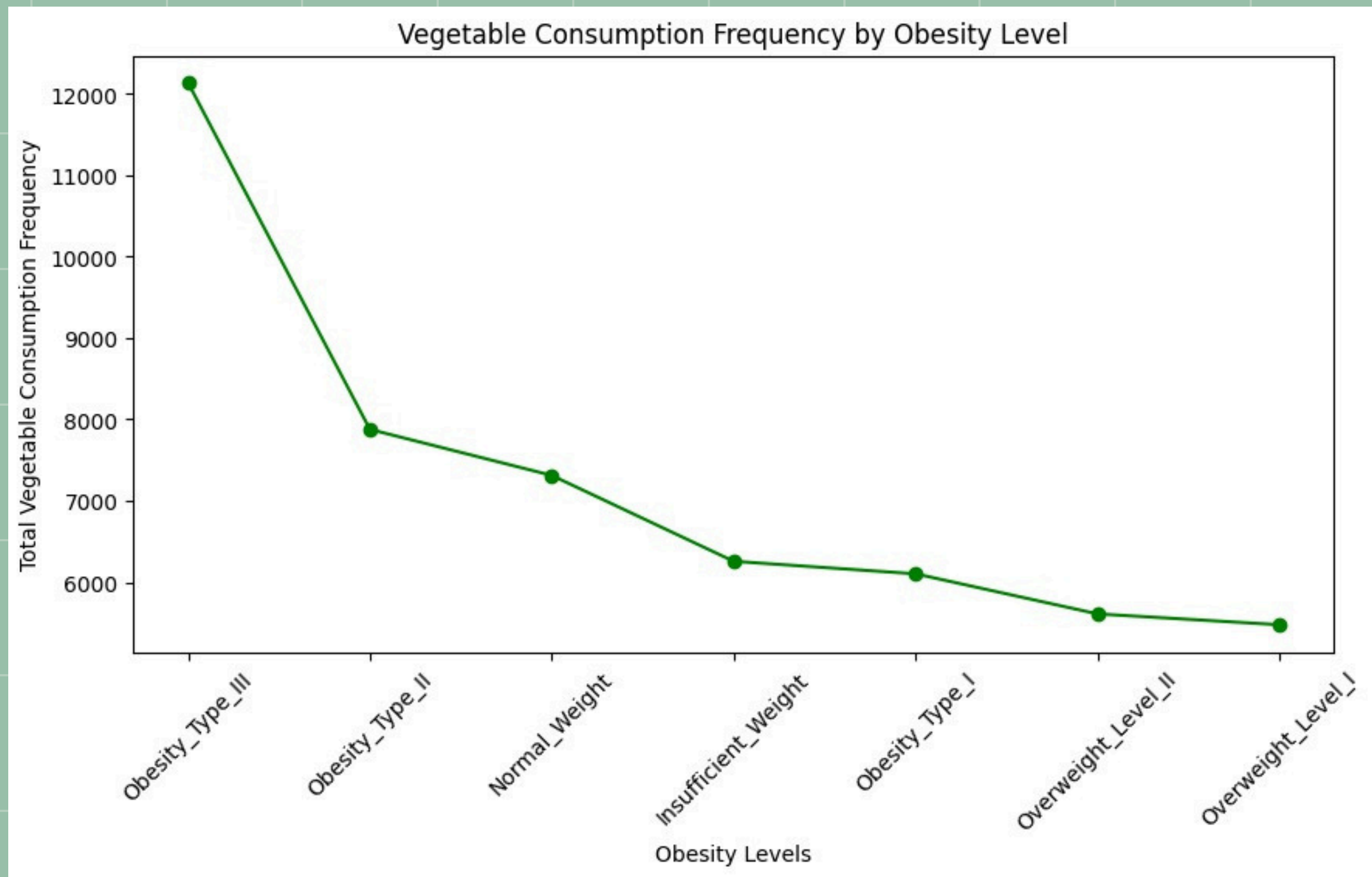
# Visualizations



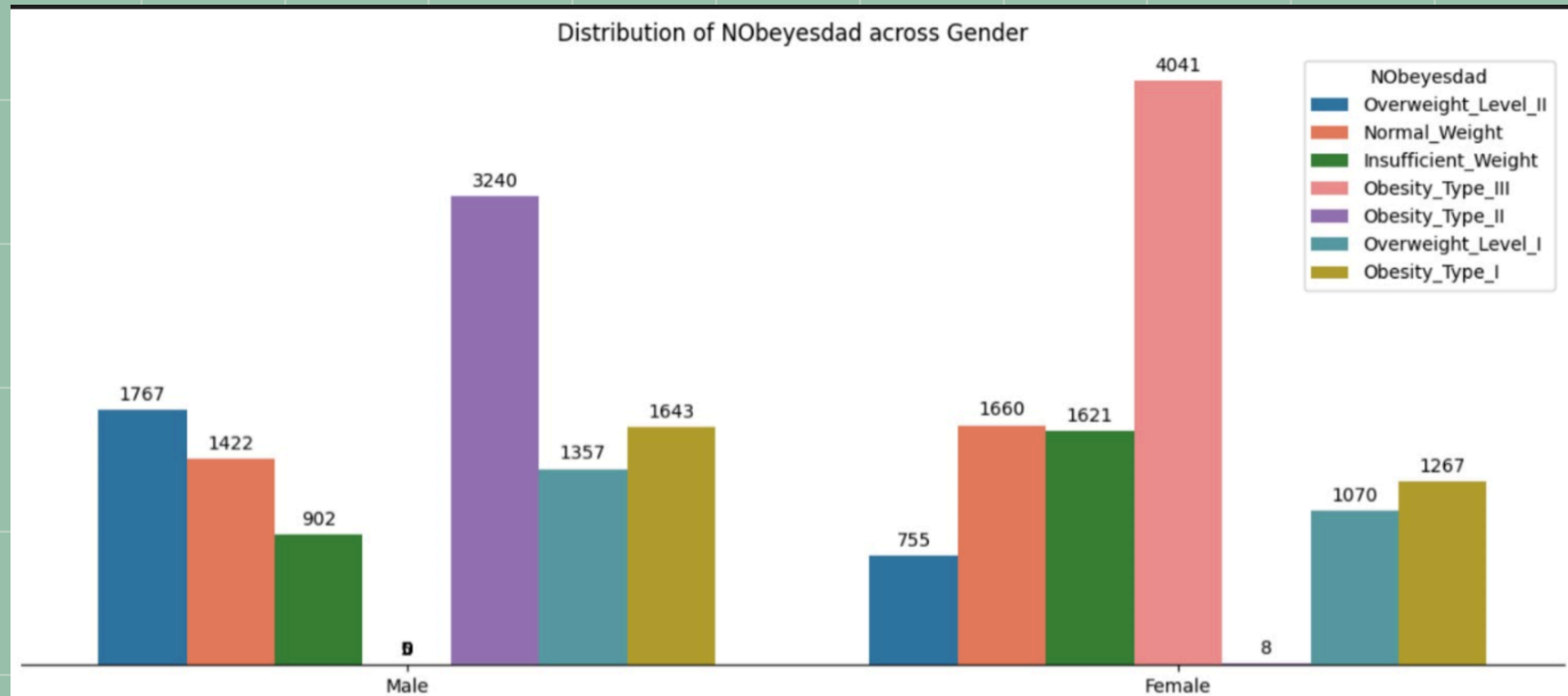
# Visualizations



# Visualizations



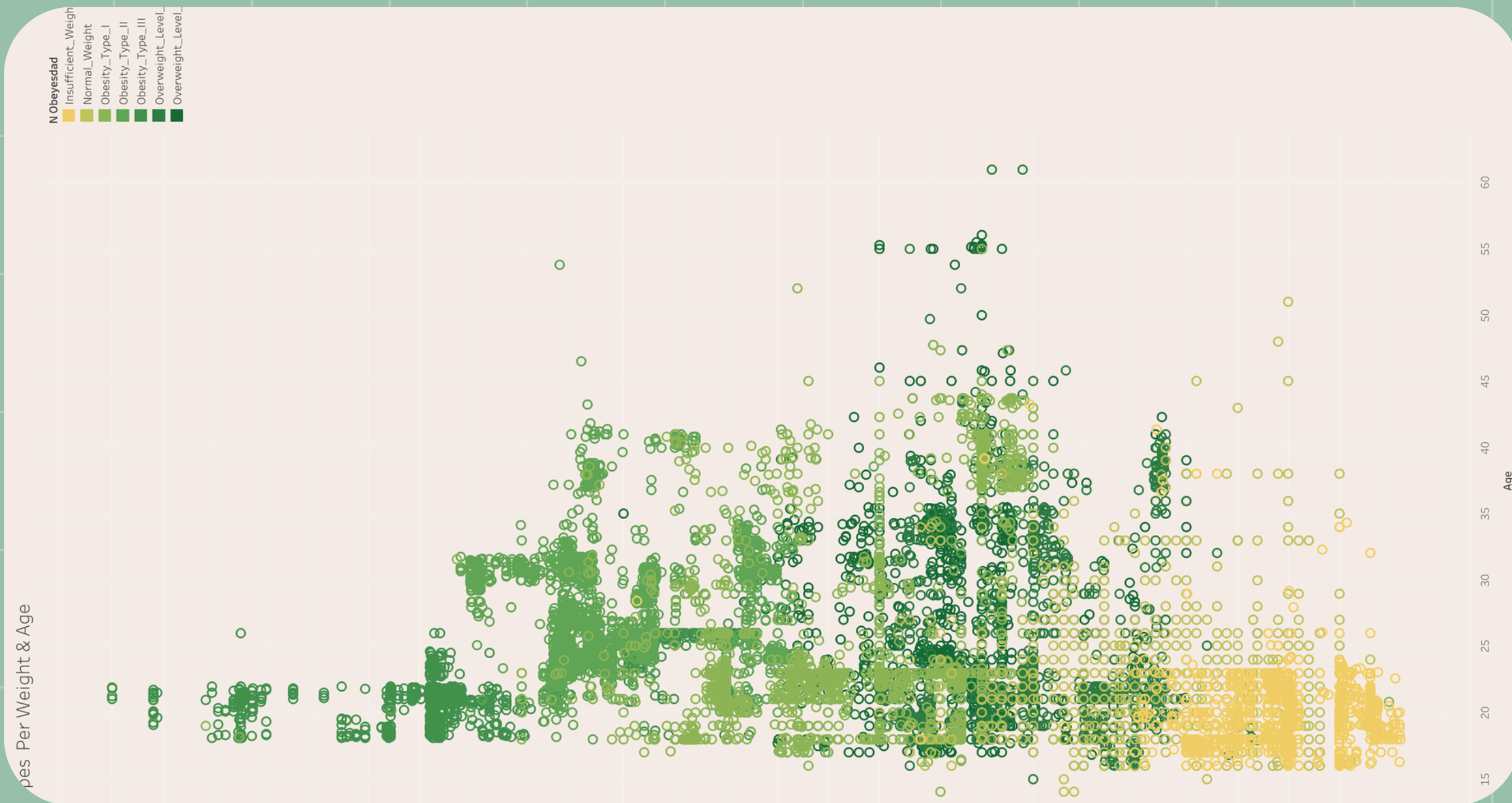
# Visualizations




correlation between targett and gender



# Visualizations



corralation between taregte and Age



# Visualizations

For better interactive  
visualization check Tableau  
DashBoard

The LINK

# ● ● ● ● Visualizations

for more exploring Descriptive  
Analysis check this notebook

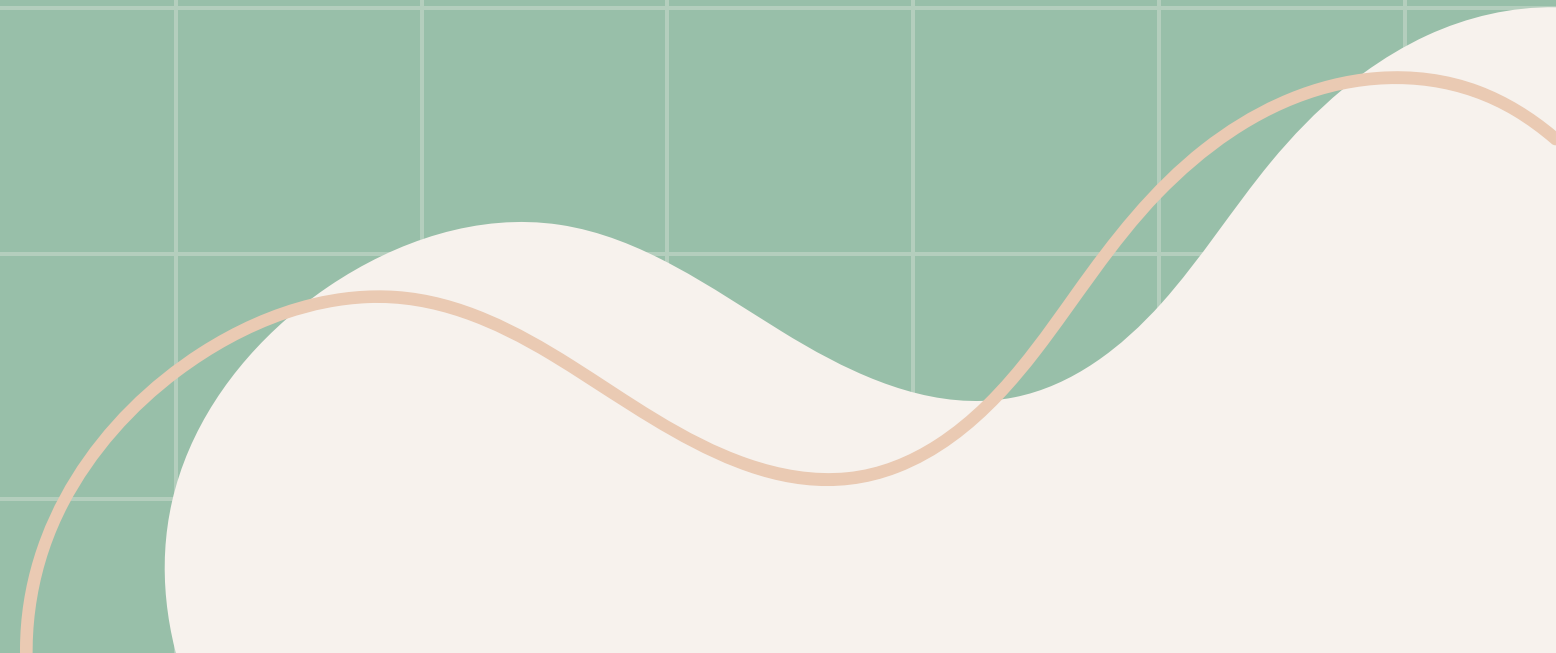
[link](#)



# Conclusion & Recommendations

**Business Insight:** Understanding obesity risk factors can guide targeted health interventions.

**Technical Insight:** Xgboost is highly effective for this problem; further improvement could be achieved through ensemble techniques or deep learning models.



Thank  
You

