

**World Health Organization (WHO),  
Latin American branch**

# **Estimation obesity levels based on eating habits and physical conditions**

By HIEN Victor, LUTTENBACHER Léa and MENU Victor





# Table of contents

**Part 1**      The ins and outs

---

**Part 2**      Our database

---

**Part 3**      Our answers

---

**Part 4**      Our model

---

**Part 5**      Conclusion

# ***What is obesity?***



**According to the World Health Organization (WHO), obesity and overweight are defined as "an abnormal or excessive accumulation of body fat that may impair health". This accumulation is manifested by an increase in the size and quantity of fat cells (adipocytes) in the body.**

# ➤ The ins and out



**In South America:  
6 out of 10 inhabitants affected by overweight.**

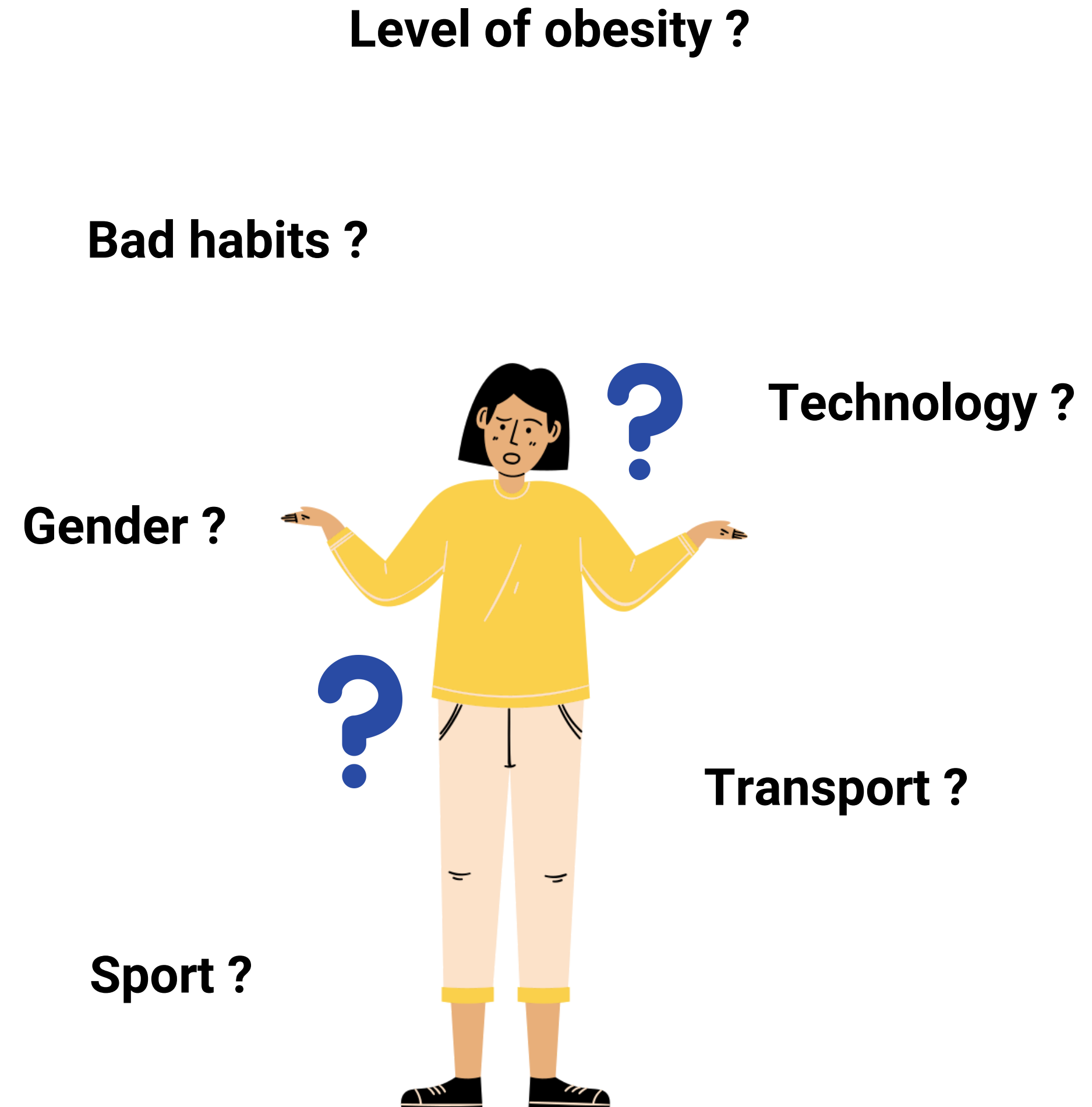
Every year there are 3.6 million more obese people in Latin America



**Provide an appropriate tool in the scope of giving a viable solution to limit and reduce the level of obesity.**

For this, we will use our dataset to derive a reliable model that will allow us to advise and guide professionals and individuals thereafter.

**What are the factors that influence a person's obesity rate?**



# ➤ Our database

- To answer our problematic , we use a dataset which includes **data for the estimation of obesity levels in individuals from the countries of Mexico, Peru and Colombia**, based on **their eating habits and physical condition**.
- The **dataset is the result of a 16 questions-survey** on a web platform. But, it should be noted that **77% of the dataset was generated synthetically** using the Weka tool and the SMOTE filter.

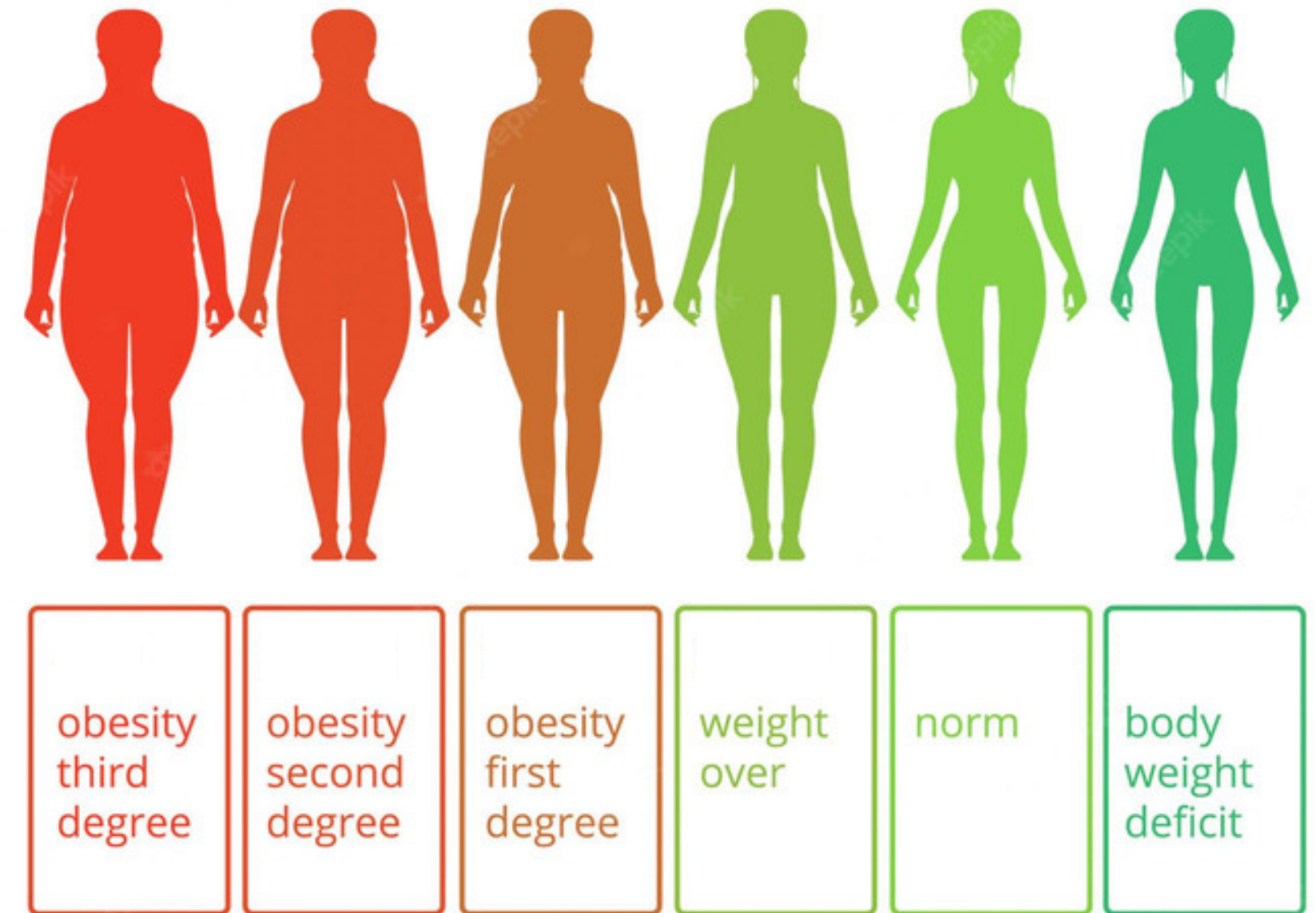




# Our variable of interest is the Level of Obesity

It exists **7 different categories** from under weight  
to obesity third degree.

Our aim is to understand which factors can  
explain a certain level of obesity .



**In our original data, we had various explanatory variables**



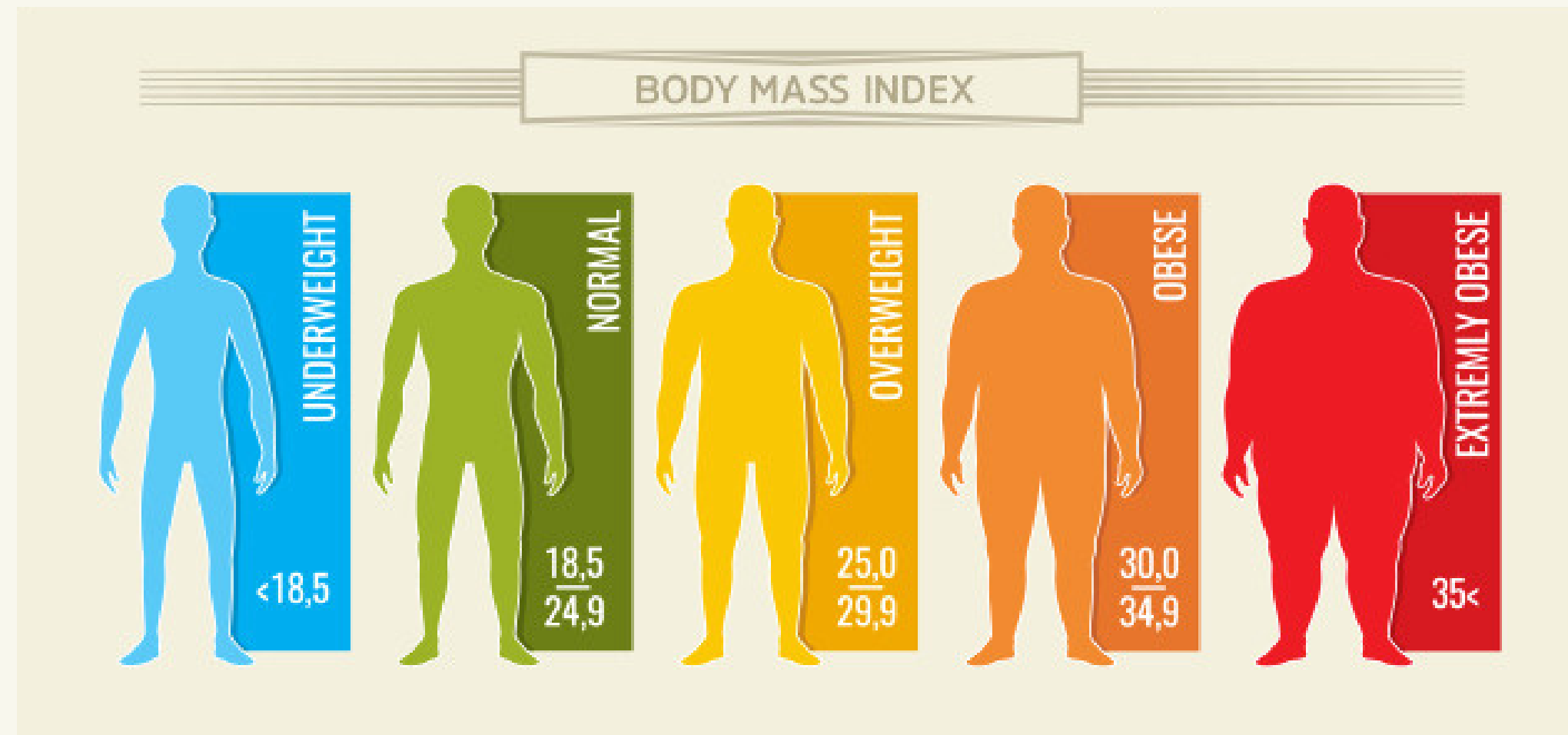
- **Gender**
- **Age**
- **Height**
- **Weight**
- **Family History with Overweight**
- **Information about the person's food habits**
- **Frequent Consumption of High Caloric Food**
- **Frequency of Consumption of Vegetables**
- **Number of Main Meals**
- **Consumption of Food between Meals**
- **Smoke**
- **Consumption of Water Daily**
- **Calories Consumption Monitoring**
- **Physical Activity Frequency**
- **Time using Technology Devices**
- **Consumption of Alcohol**
- **Transportation Used**



# We added a new value the Body Mass Index (BMI) :

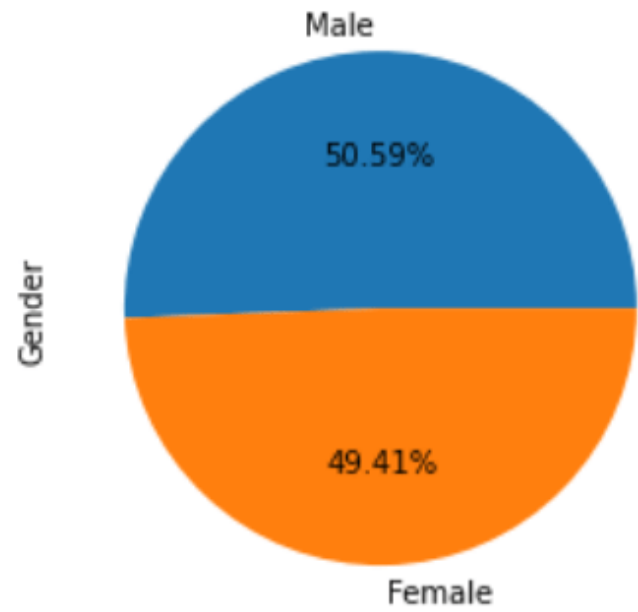
- A universally used variable which makes it possible to categorize the degree of obesity

$$\text{BMI} = \frac{\text{weight (kg)}}{\text{height (m}^2\text{)}}$$

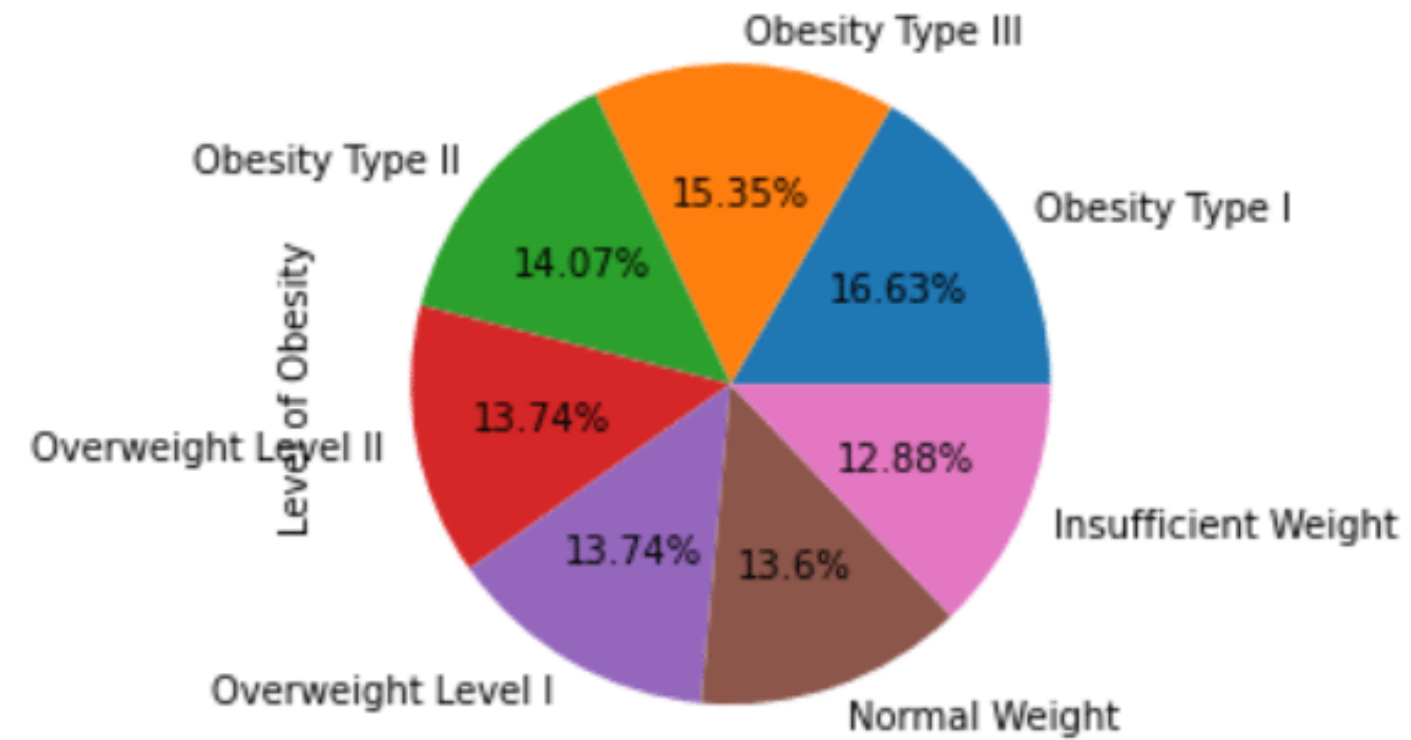


# Overview of the data set

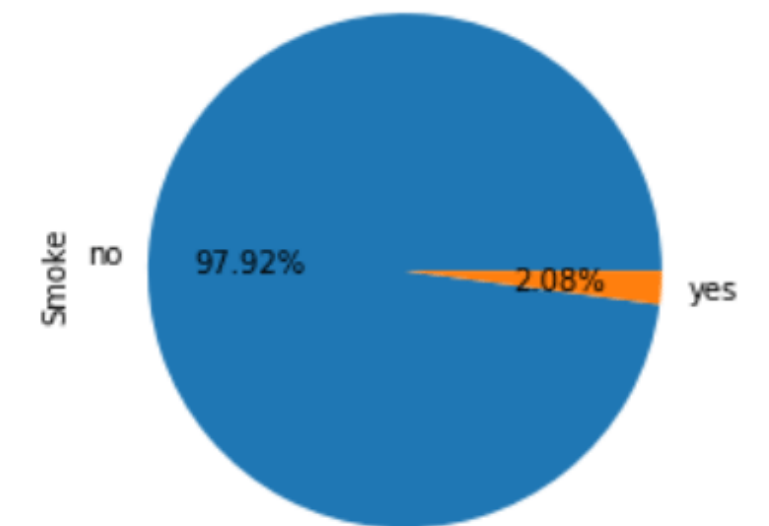
Distribution of Gender



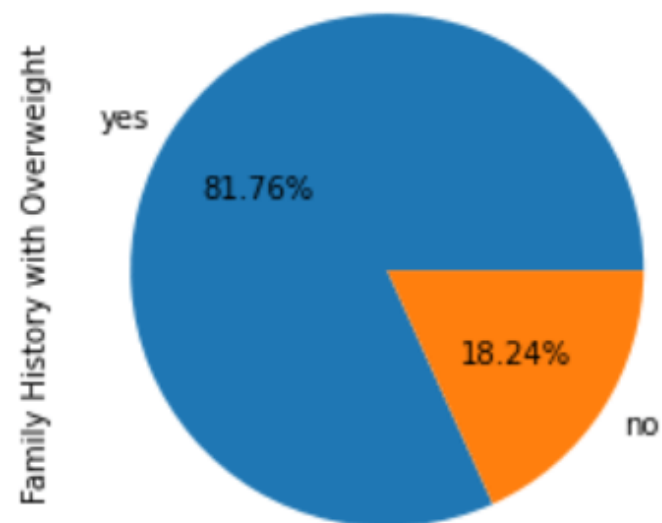
Distribution of Level of Obesity



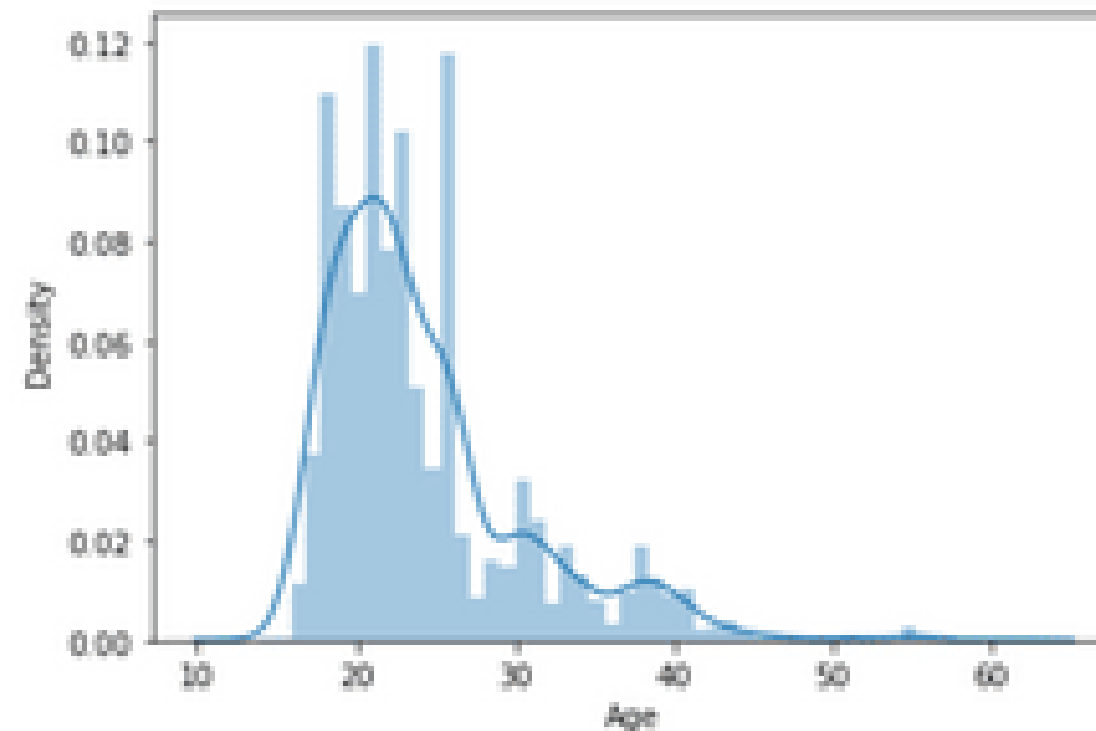
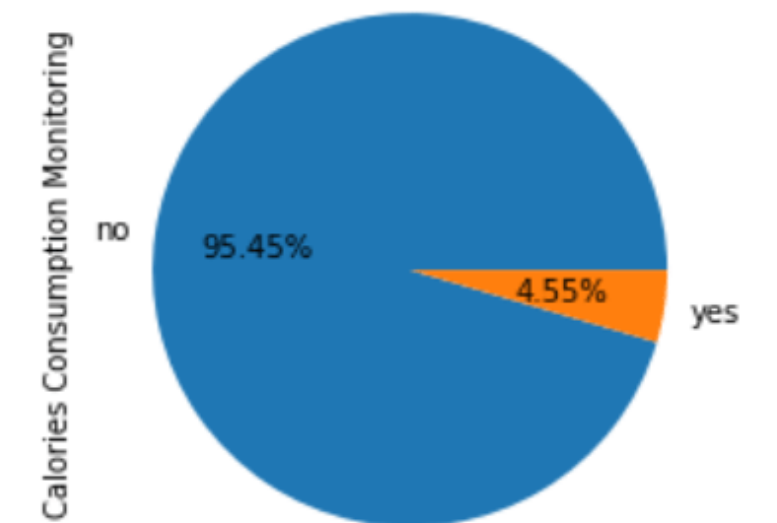
Distribution of Smoke



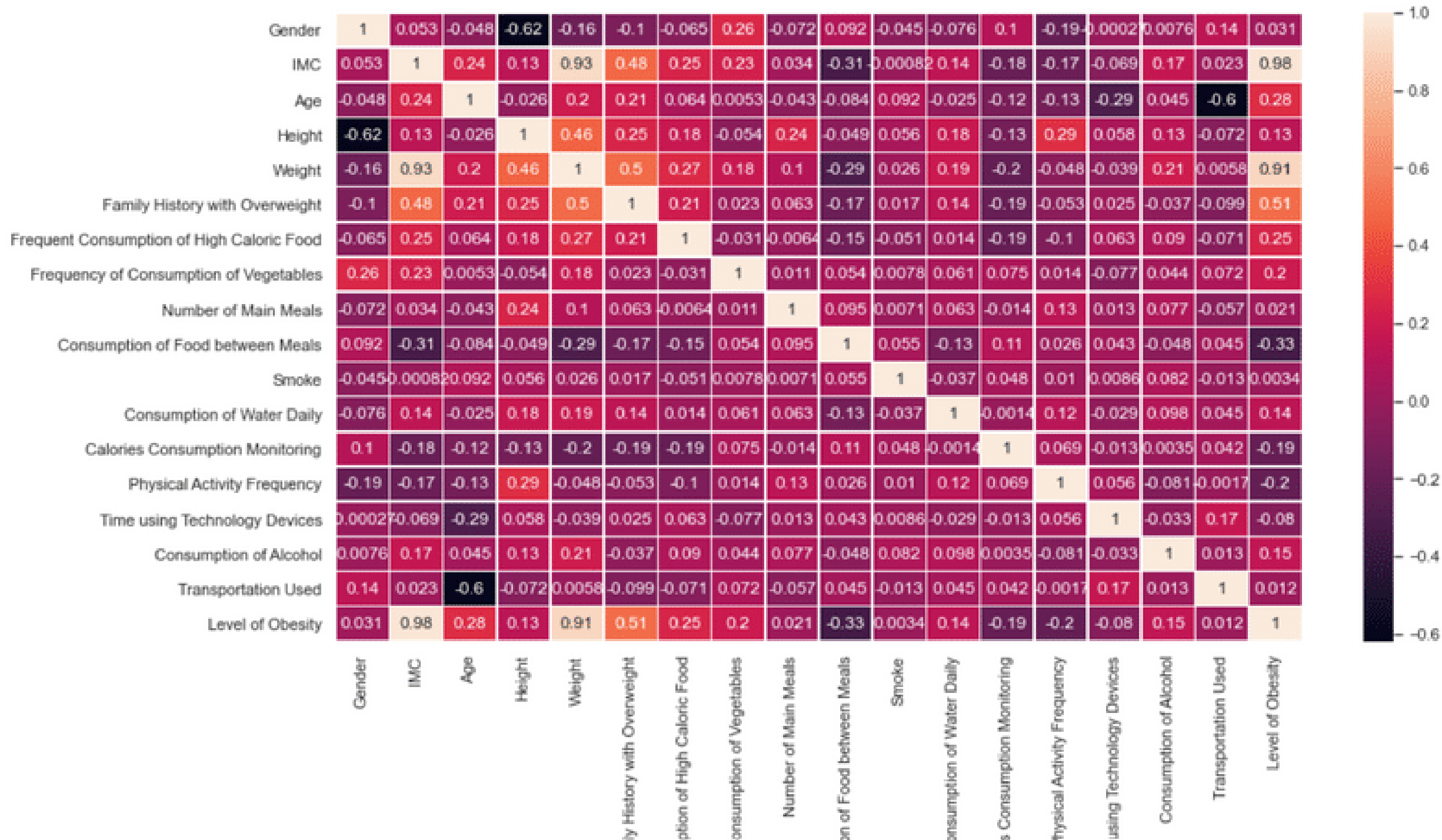
Distribution of Family History with Overweight



Distribution of Calories Consumption Monitoring



# Our answers

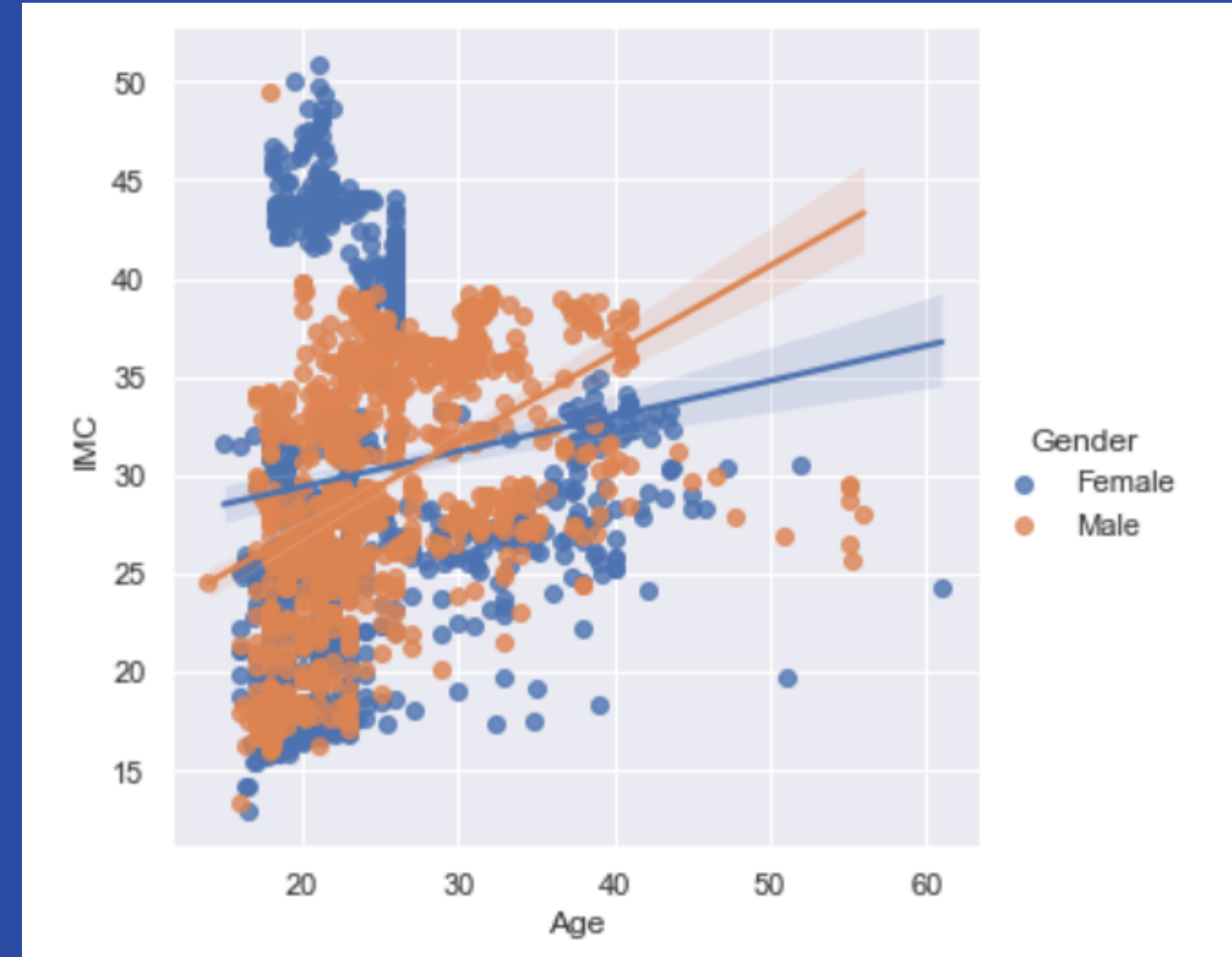
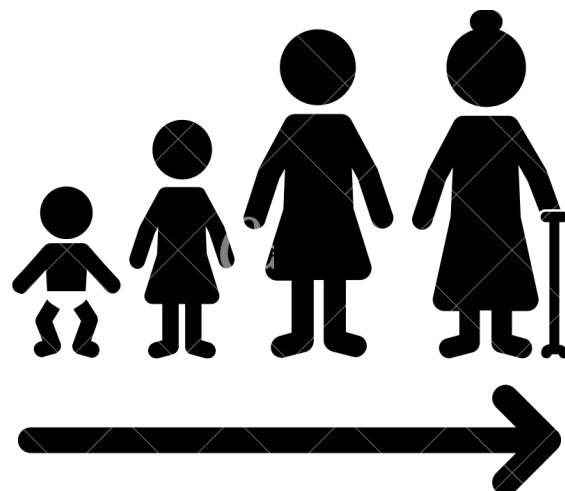


Correlation between the level of obesity  
and other variables

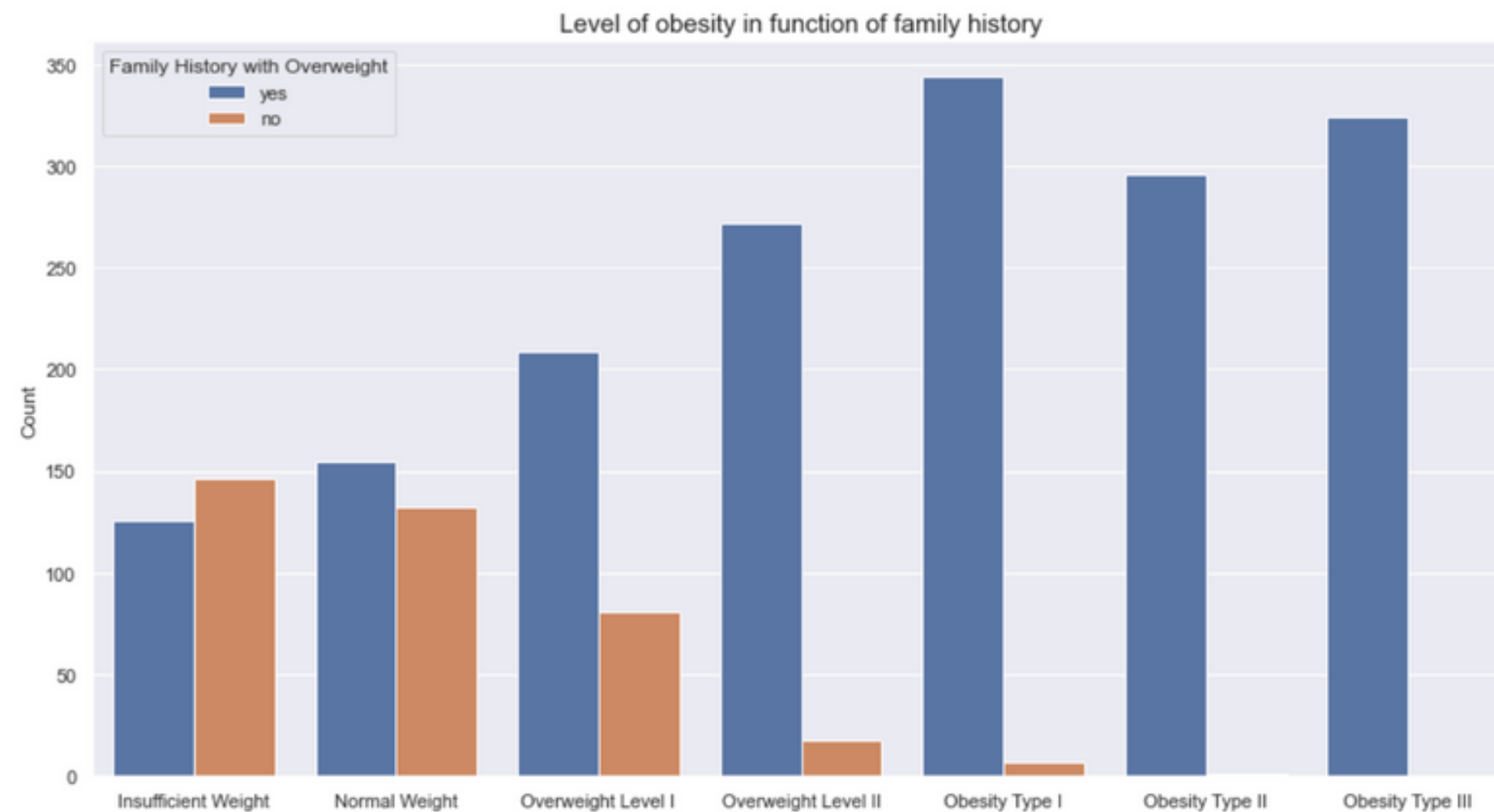
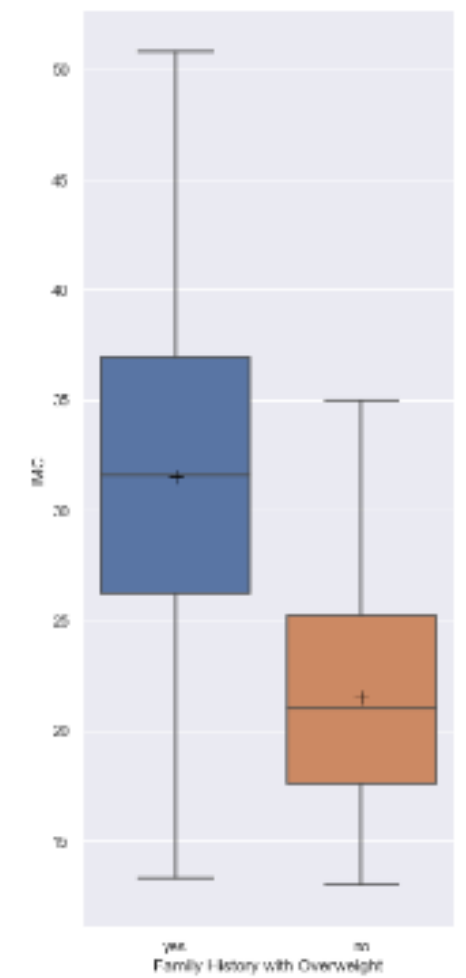
# How does the age affect the level of obesity?

With age, the BMI (and thus the level of obesity) tends to increase, and this more strongly in men.

However, age doesn't play a direct role, it is more about the lifestyle habits which change when we get older (less sport....).



# Do family history have an effect on obesity?

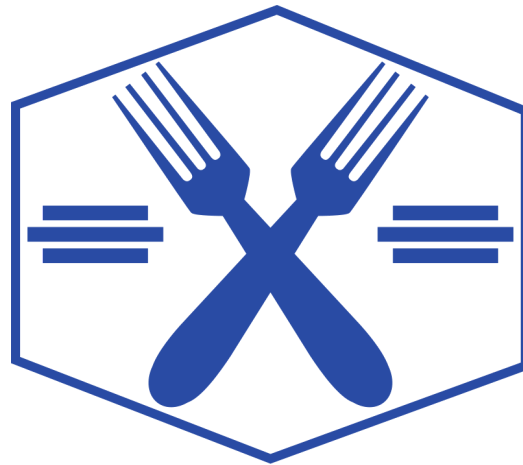


Above a certain level of obesity, overweight level I, the majority of people have a family history of being overweight. Moreover, as the level of obesity increases, so does the proportion of people with a family history of overweight. Furthermore, **all people of obesity type II and III have a family history of overweight.**

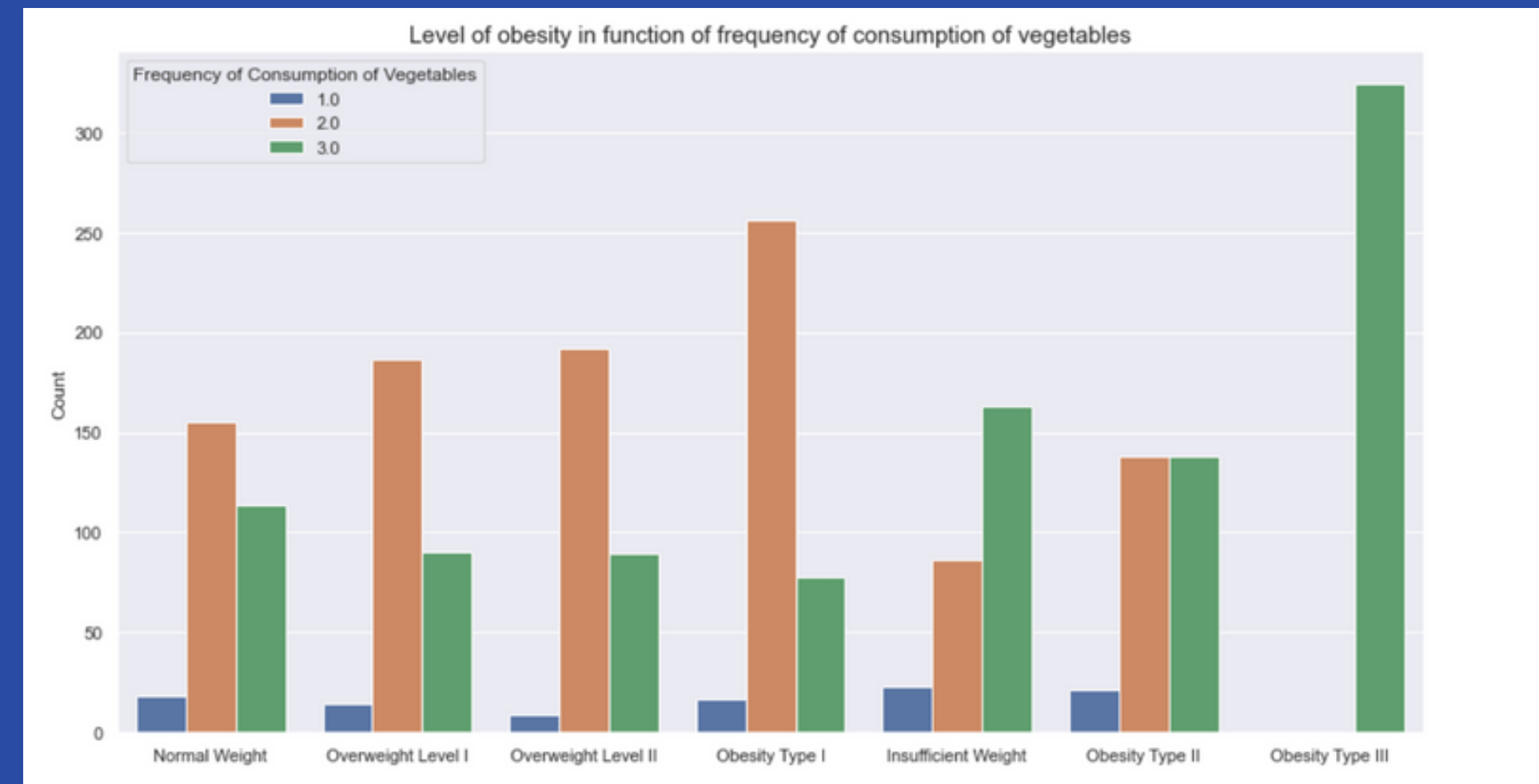
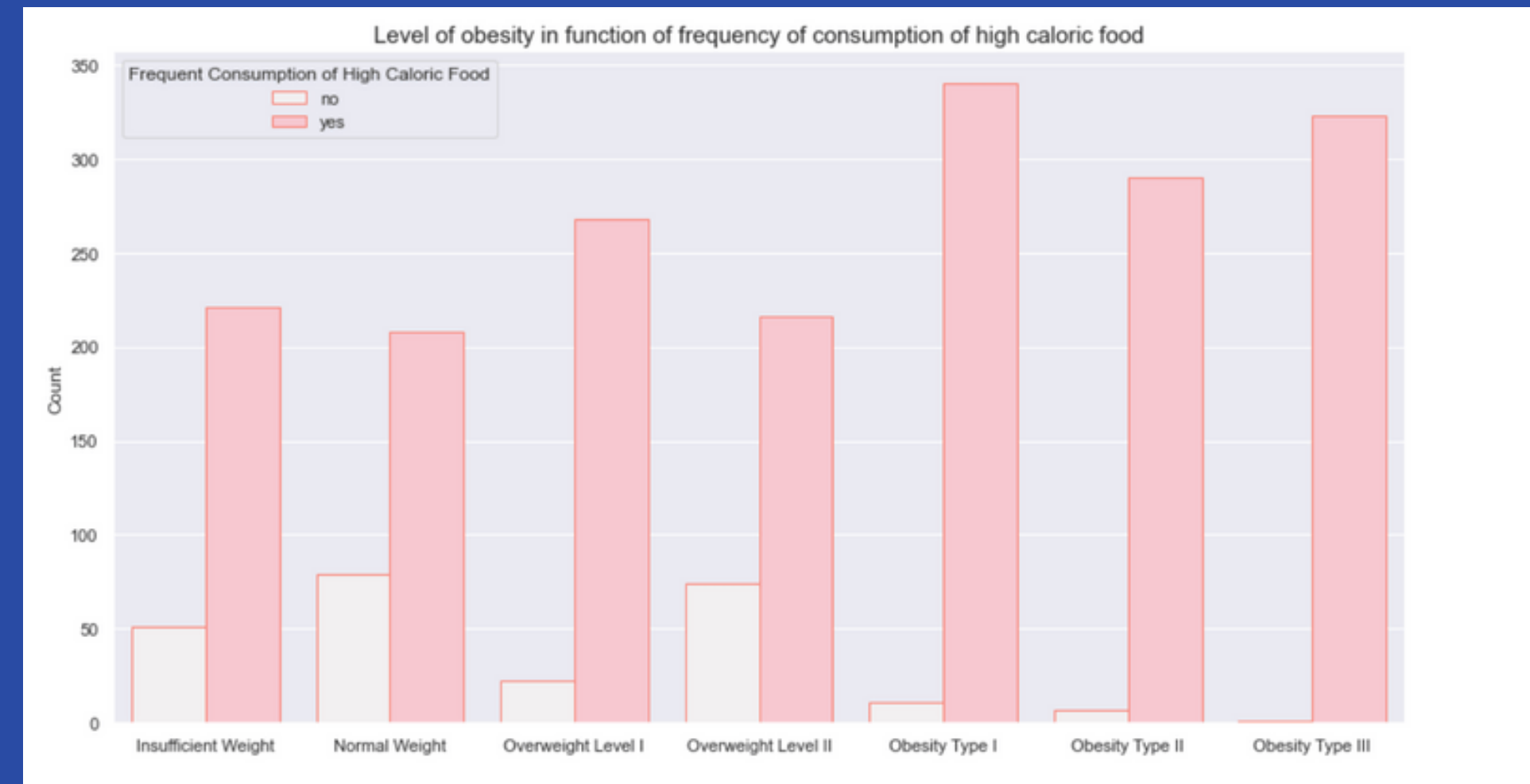
Conclusion: Having a family history of overweight largely favours obesity.

# The level of obesity in function of food habits

- On the first graph, we can clearly see that the **higher the level of obesity is, the higher the proportion of people consuming caloric food is.**

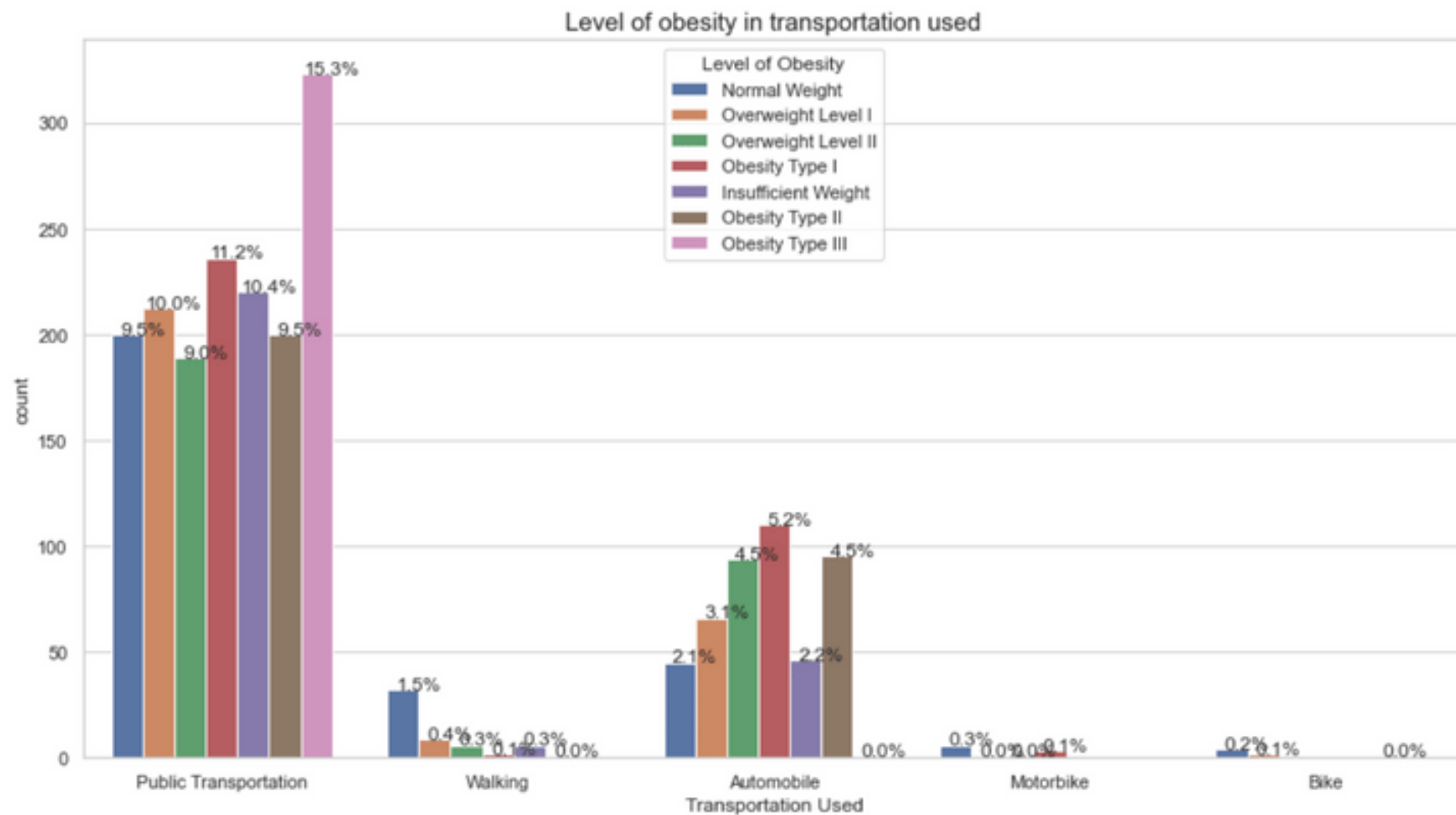


- On the second graph, we can also see very clearly that the consumption of vegetables is much more important for obese people, but also for those in insufficient weight. Moreover, we also notice that **the more obese people are, the more vegetables they eat.**





# Is there a relationship between obesity types and transportation?

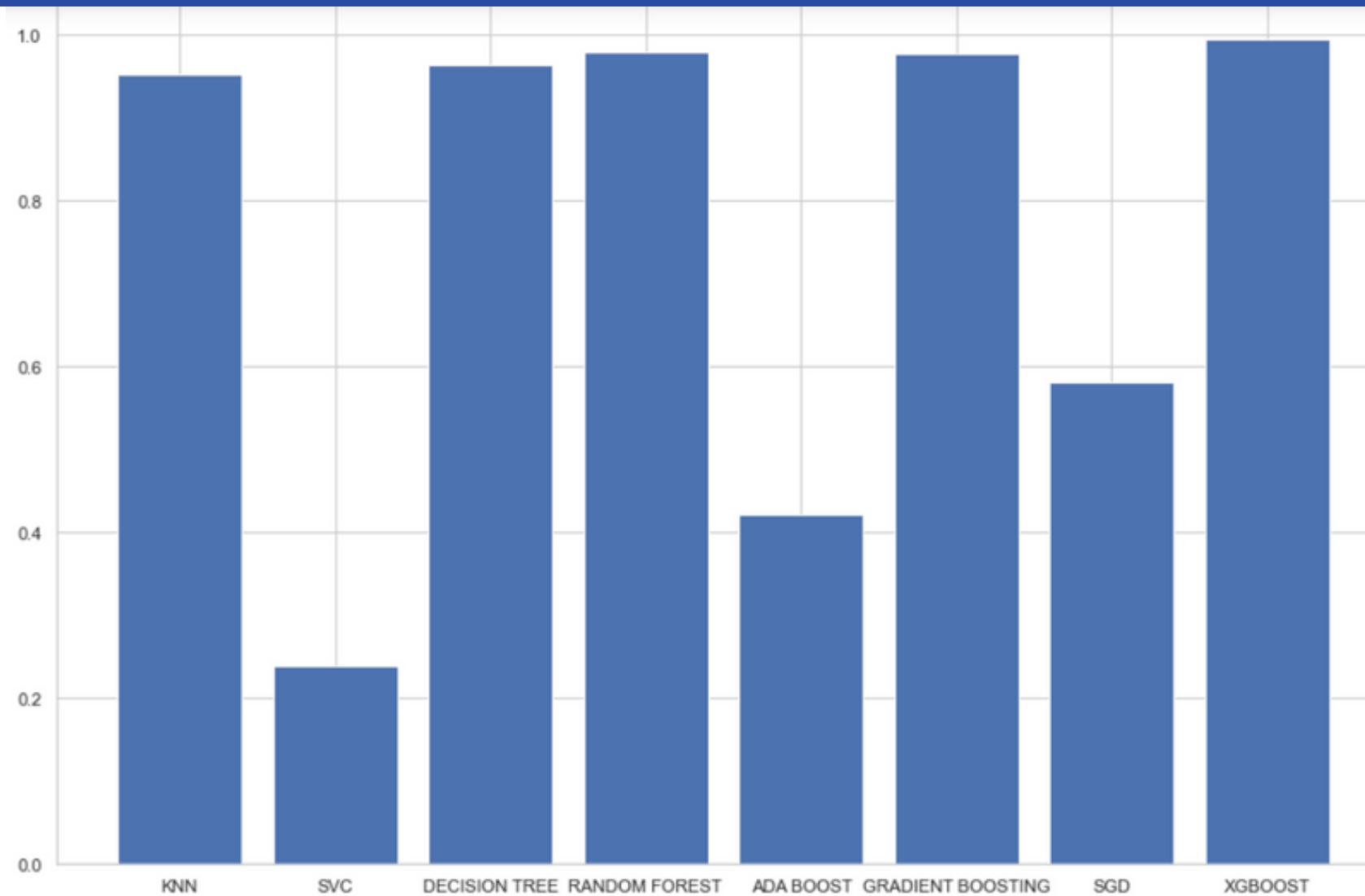


From all of types of weight, everyone tends to favor the public transport at first and secondly the automobile.

An important information of this graph is that most of the people who walk have a normal weight.

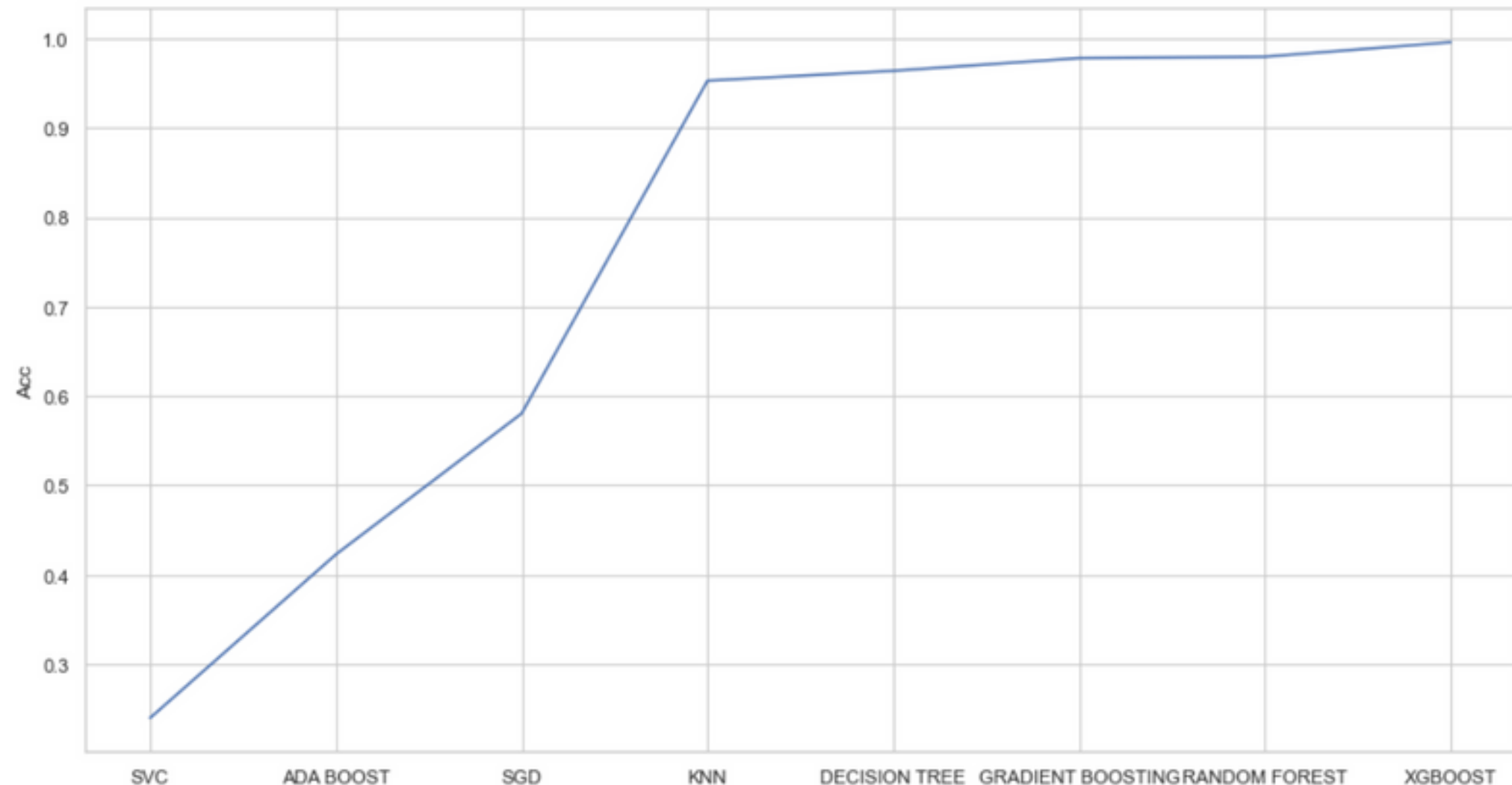
Favouring walking over other public transport is really interesting for his health.

# ➤ Our Model

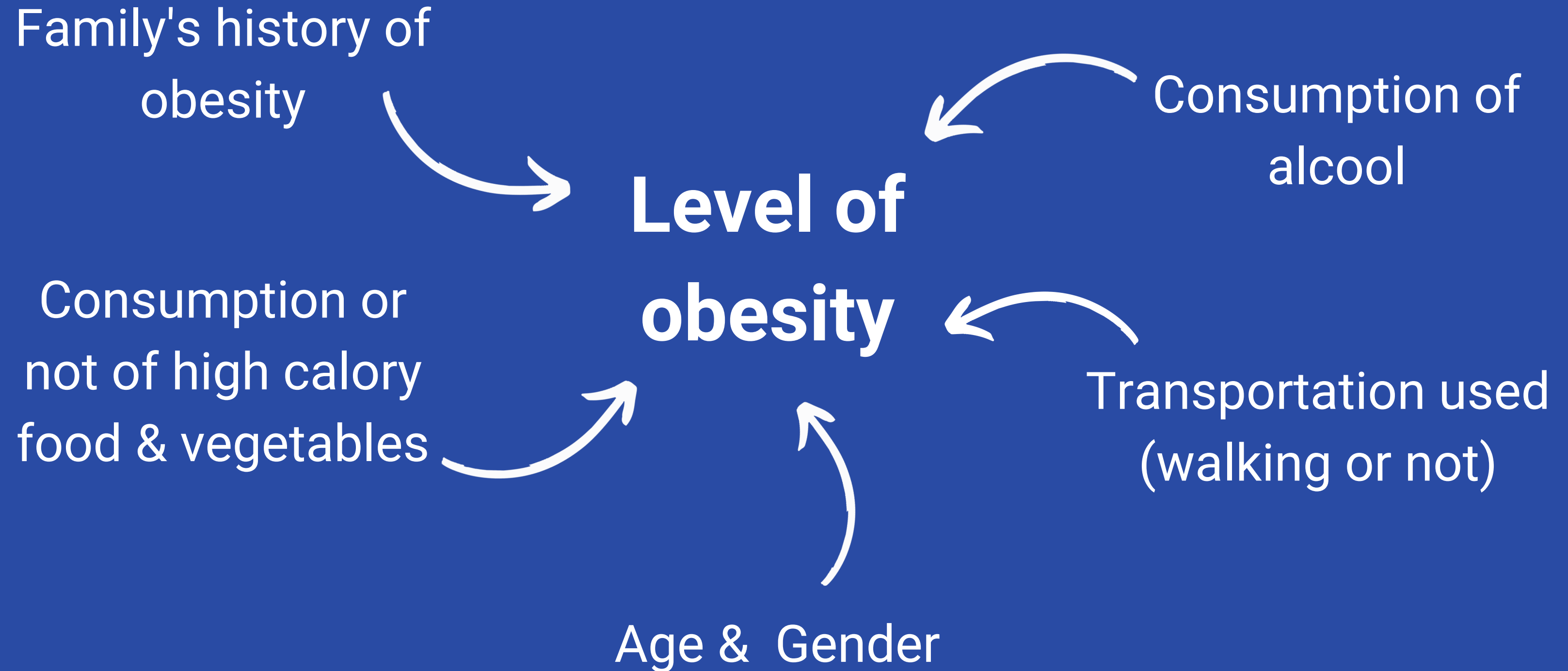


- KNN Accuracy : 0.953
- Decision Tree Accuracy : 0.959
- Gradient Boosting Accuracy : 0.977
- Random Forest Accuracy : 0.979
- XGBoost Accuracy: 0.996

- Thanks to our **modeling** part, but also especially to our encoding, our modifications of the dataset and our optimization (grid search, etc ...), we managed to obtain a very **good accuracy and precision** for our models; which allows us to affirm that **our model is reliable and will be exploitable**.



# CONCLUSION





## In Summary

- Data : Estimation obesity levels based on eating habits and physical conditions
- Crucial ins and outs for predicting and reducing obesity risk
- A data set explored and well prepared to respond effectively to our problem
- Observations and conclusions on factors influencing the level of obesity
- Creation of a reliable and usable model
- An API operation as an example

# Thank you

**Feel free to check our GitHub for more information !**

**[https://github.com/Vim17/ESILV\\_PythonforDataAnalysis\\_Project2023](https://github.com/Vim17/ESILV_PythonforDataAnalysis_Project2023)**