AN ANALYSIS OF THE FACTORS AFFECTING AND CONTRIBUTING TO WEIGHT IN PEOPLE

R. HOWARD MBINDA MUCHAKI

18/10/2024

# Introduction

The goal of this analysis is to identify the factors that contribute to weight gain.

# Background

Obesity is a serious and chronic disease with genetic and environmental interactions. It is defined as an excessive amount of fat tissue in the body that is harmful to health. The main risk factors for obesity include social, psychological, and eating habits. Obesity is a significant health problem for all age groups in the world. Currently, more than 2 billion people worldwide are obese or overweight. Research has shown that obesity can be prevented. In this study, artificial intelligence methods were used to identify individuals at risk of obesity. An online survey was conducted on 1610 individuals to create the obesity dataset. To analyze the survey data, four commonly used artificial intelligence methods in the literature, namely Artificial Neural Network, K Nearest Neighbors, Random Forest and Support Vector Machine, were employed after pre-processing. As a result of this analysis, obesity classes were predicted correctly with success rates of 74.96%, 74.03%, 74.03% and 87.82%, respectively. Random Forest was the most successful artificial intelligence method for this dataset and accurately classified obesity with a success rate of 87.82%.

# Data source

The source of data is from the Kaggle website((*Learn Python, Data Viz, Pandas & More | Tutorials | Kaggle*, n.d.)).With the provided CSV file and its documentation of the different columns and what each variable number means.

# Data preparation

## Data collection

The data was collected from the Kaggle dataset website. Its usability of 8.24, and lack of need for updation were the primary reasons for its selection. The data set is a survey in nature, with numbers used to represent qualitative data.

## Data cleaning

The cleaning process was done on Jupyter's notebook. With Python being the programming language used.

First identification of null values was done. But since none appeared, no cleaning was required in that aspect. Following this was the identification of duplicate values, again none were found thus the data was clean. The data formats were looked at to ensure consistency, especially with the column names and its text size. No columns were dropped as all were going to be used in the analysis of the factors. The quantitative values of the responses were converted from their numeric forms to their string form for easier understanding when the visuals were made.

# 

# Exploratory data analysis

## Descriptive statistics

Below is an overview of the numeric values in the dataset.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Age | Height | Number of meals |
| Count | 1610.0000 | 1610.00000 | 1610.000000 |
| Mean | 33.115528 | 167.741615 | 1.872050 |
| std | 9.835076 | 7.979873 | 0.683845 |
| min | 18 | 150 | 1 |
| 25% | 25 | 161 | 1 |
| 50% | 32 | 168 | 2 |
| 75% | 41 | 174 | 2 |
| max | 54 | 193 | 3 |

# 

# Analysis and results

Sentiment analysis was used to go over the data.

This shall be provided in the Power Bi file.

# 

# 

# 

# 

# 

# 

# Discussion

Going over the data. Males had a higher tendency to be obese(52.96%) and overweight(60.47%) as compared to their female counterparts. While females had the highest percentage of being underweight(80.82%) Smoking did not promote weight gain as the majority(658) of those who smoke came from those in the normal class division.

Automobiles were the most common mode of transport amongst all classes, with motorbike transport being the least used. Those who are obese and overweight took automobiles and public transport as their most common means. With a small minority opting for walking. Those in the normal class had the highest appearance with walking as a means of transport.

There was no correlation between height to weight gain, as males tended to be taller than women. But this did not show any connection to weight.

As for snacking between meals, the survey showed that most, sometimes take a snack in between meals. With the overweight class dominating the sometimes and always category.Meaning that their weight gain is related to their in-between meals. Those in the normal class dominated the rarely and usually category, meaning that the fewer meals one has in between can lead to weight loss.

The average age of those who are obese is 39 years old while that of those who are underweight is 26 years, the following normal weight is 28 years. This shows that there is a gradual gain of weight over the years. Providing evidence that age is also a factor in weight gain.

When looking at physical activity to see if this can be used to curb weight gain. It was discovered that those who do not perform any physical activity are of age 27 years, Which is the category of those of normal weight. While those who partake in activity more the 6 times a week are those of age 37. This is our obesity class, meaning that the obese do become more active, but the activity is not enough to help with weight loss.

While looking at the number of meals in a day, the obesity class had the highest median of main meals a day of 3, while the underweight class had one. Then the normal and overweight had only 2. This suggests that the main meals a day should be 2 to promote a healthy level of weight.

## 

## Limitations

The aspect that the survey data values came in numeric form, limited the speed and efficiency of performing data cleaning. As the values had to be converted into their character format for easier graphing.

Lack of defined definitions of the class. That is to say, what is the weight range for underweight, normal, overweight and obese?

# 

# 

# 

# 

# 

# 

# 

# Conclusion

In summary, the primary factor affecting weight gain is ageing as over time you will come to gain weight. Having an excess of snacks in between meals contributes to added calories. Having 2 main meals a day is needed to maintain a normal weight. Physical activity should be prioritized in earlier years of life to maintain your normal weight. Smoking does not in any way contribute to weight gain.

# References

<https://doi.org/10.33484/sinopfbd.1445215>

*Learn Python, Data Viz, Pandas & more | Tutorials | Kaggle*. (n.d.). https://www.kaggle.com/datasets/suleymansulak/obesity-dataset/data

Stevens, E. (2023, May 10). The 7 most useful data analysis Techniques [2024 Guide]. *CareerFoundry*. https://careerfoundry.com/en/blog/data-analytics/data-analysis-techniques/#sentiment