Report for Nutritional Dietary data

Bangcale, Kurt Andrei A.[1]
Era, Chloe Josephine F.[2]

1Bangcale, Kurt Andrei A.,2Era, Chloe Josephine F., Data Science for the Health Sciences/CON29, School of Health Sciences, Mapúa University

INTRODUCTION

The dataset contains nutritional dietary data from 1000 patients. This dataset contains the nutrients taken by the individuals, where:

-Body fat percentage: The percentage of a person's body weight that comes from fat. It's a key metric in determining body composition.

-Muscle maass (kg): The total mass of muscle in the body and the indicator of physical fitness and strength.

-BMI: The measure of body weight in relation to height.

-Physical activity hours per week: The number of hours an individual spends engaging in physical activity per week..

-Daily caloric intake (kcal):Total calories consumed by patients daily.

-Water intake (ml): The amount of water consumed by the individual.

The micronutrients(vitamin C and iron) and macronutrients(protein, carbohydrate, and fat) show the sources of energy taken in the body.

The dataset shows patterns and trends of each variable to the patients, as well as their varying effects on each other too.

The objective is to utilize the key variables of the dataset such as physical activity hours per week and its connection to other variables such as body fat and muscle mass. This would show that the dataset is helpful in determining a patients' physical state and capabilities.

METHODS

The researchers used a personal laptop with RStudio and uploaded the *3_Nutritional_Dietary_data_Group_017.csv* file.

It was cleaned first and removed the outliers. After that, the researchers plotted the data from physical activity hours, muscle mass, and body fat percentage first before comparing those variables.

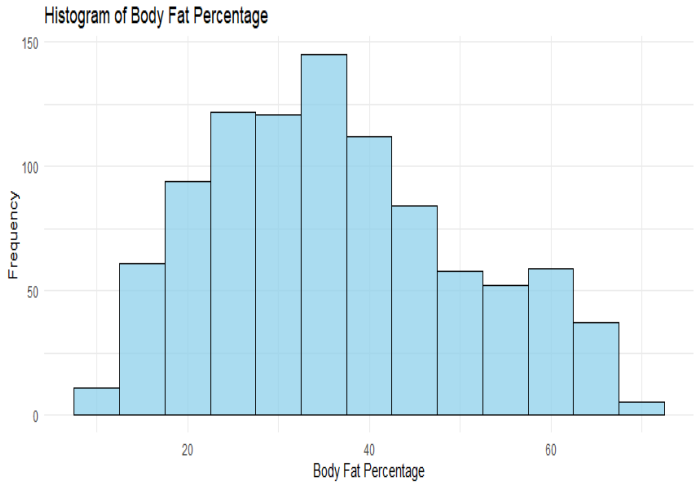

Figure 1: Body Fat % Histogram

The histogram shows the body fat percentage of all the patients in the dataset. The majority of the individuals in the dataset have a body fat percentage between 20%-40%.

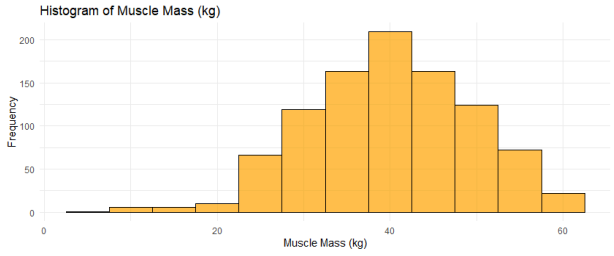

Figure 2: Muscle Mass Kg Histogram

The histogram for Muscle Mass Kg. This shows that almost all of the patients have more than 20 kilograms of muscle mass in their body.

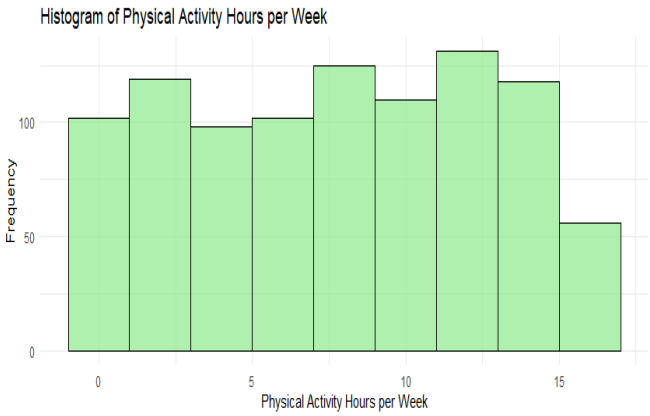

Figure 3: Physical Activity hours per week histogram

This shows the plot for the physical activity hours per week of all patients. The physical activity hours per week varies on many individuals.

After plotting all the necessary variables, the researchers performed the codes for a compatible correlational test, which is the pearson's correlation. It measures the strength and direction of the linear relationship between two continuous variables, and since our variables have linear trends, it is compatible for a pearson's correlational test.

Also, both relationships are measured in simple linear regression, which is due to the fact that all three variables are continuous. This makes the linear regression model ideal.

RESULTS

After inputting and running the codes, the researchers found the results.

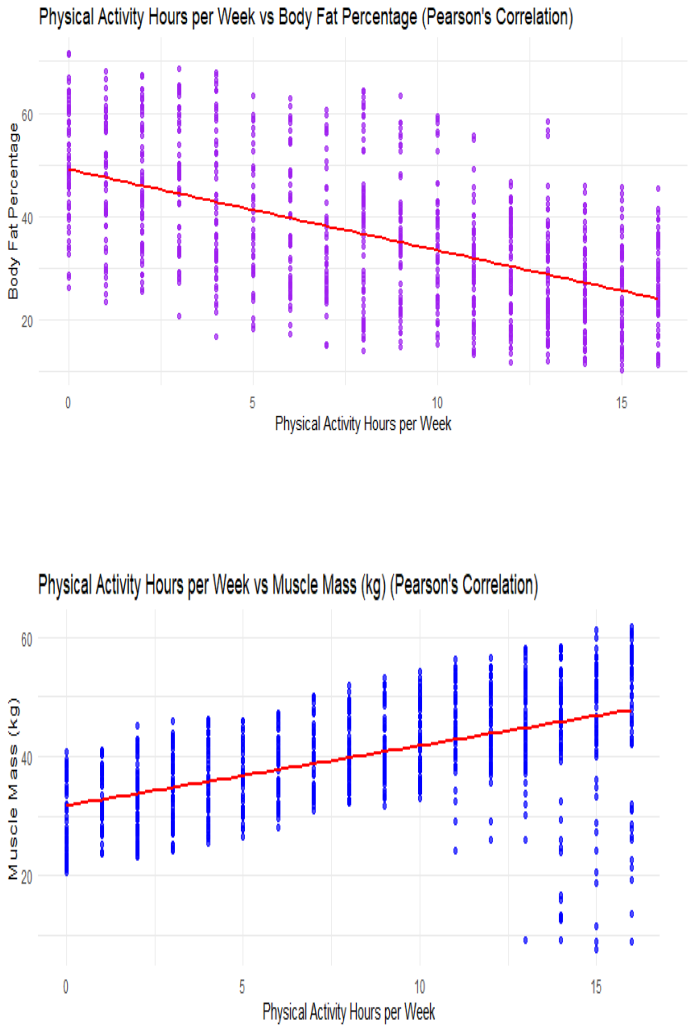

Figure 4. Physical Activity Hours per Week vs Body Fat %/Muscle Mass Pearson's Correlation plots

About the results from using the code to perform simple linear regressions on both relationships it shows that the relationship between physical activity and body fat percentage/ muscle mass per kilogram is linear, as both are linear when plotting those relationships.

-The p-value for the physical activity-body fat relationship (figure 4, above) is lower than < 2e-16, indicating a strong relationship, and the slope going downwards also

highlight that the two variables have a significant but negative correlation, meaning that physical activity hours per week increases, the more body fat percentage decreases in a person. In other words, an individual would be a bit healthier and weigh less if they do exercises for hours per week.

-The p-value for the physical activity-muscle mass relationship(figure 4, below) is also ower than 2e-16, meaning that it has a strong relationship. The slope is in an upward trend, where it shows that the two variables have a strong positive relationship. From the graph's interpretation, the more an individual exercises, the more their muscle masses increase by kg. They become physically stronger.

In terms of pearson's correlational coefficient results, the t-value of -20.51 for physical activity-body fat % and 19.08 for physical activity-muscle mass shows that both relationships are highly significant.

CONCLUSION

Both regression models and correlational tests reveal significant and meaningful relationships between physical activity and body composition.

For the body fat percentage, the model shows that physical activity hours are inversely related to it, with an estimated reduction of 1.56% for every additional hour of physical activity per week.

For Muscle Mass, the model indicates that physical activity is highly correlated with muscle mass, with an increase of 1.01 kg for every additional hour of physical activity per week.

Not only the findings did highlight the important role of physical activity plays in both reducing body fat and increasing muscle mass, but also showcased that the dataset really shows the details of the variables to each patients very clearly.

It also means that other factors may also contribute to the variability in these outcomes. Overall, the findings and the dataset shows how important detailed nutritional dietary data is, as every individual has a different set of macronutrient and micronutrient composition