Nhóm 13

Tên thành viên:

Lê Trọng Nghĩa (Nhóm trưởng)

Nguyễn Thiên Phúc

Bành Đức Khánh

Lê Viết Tố Khoa

Lê H0àng Nguyên

**Project 3**

CONDUCTING HYPOTHESIS TESTING ON GENERAL ASSUMPTIONS ABOUT NUTRITION

1. **Introduction**

* Background and Motivation

The consumption and nutrition of plant-based and plant-based foods have always been a subject of widespread attention. With the development of humanity, people are increasingly focusing on healthy and more effective eating habits. This project aims to explore, depending on the case (weight gain/weight loss), which foods we should use to be the most effective depending on the purpose.

* Hypothesis and Objectives

Hypothesis 1: People who lose weight should eat plant-based foods.

Hypothesis 2: People who need to gain muscle should eat animal-based foods.

Objective: To analyze the nutrients obtained in foods as a general point of that food for human health, thus using statistical analysis and test to these hypotheses.

1. **Methodology**

* Data Collection

To collect data on the nutritional composition of dishes, we use Parsehub to perform web scraping from tasty.com website. After using Parsehub, we end up with around 2500 URLs of dishes both plant-based and animal-based. We used Beautiful Soup, a Python library used to analyze and extract information from HTML and XML documents. Beautiful Soup will help us extract the ingredients of the dish from tasty.com.

* Data Preprocessing

Since our data is gathered by accessing the URL using Beautiful Soup. The corrupted links won’t yield any data. Therefore, our raw data file is relatively clean. Minimal work is done in data preprocessing.

* Feature Extraction

To add detailed nutritional information to the dishes, we use the Calorie Ninjas. This API will return a list of nutrients from the ingredients as input. The nutrients are:

* sugar\_g
* fiber\_g
* sodium\_mg
* potassium\_mg
* fat\_saturated\_g
* fat\_total\_g
* cholesterol\_mg
* protein\_g
* carbohydrates\_total\_g

The information collected from the Calorie Ninjas will be integrated into the existing datasets to enable more detailed nutritional analysis.

* Creating Target Feature Y

After collecting sufficient nutritional information and related indicators from various sources, we standardize and weigh the characteristics. Weight is done manually and differently based on each hypothesis. This process helps evaluate the importance of each characteristic in the analytical model and gives further assessments of the health effects of nutrients and weight loss/muscle growth targets.

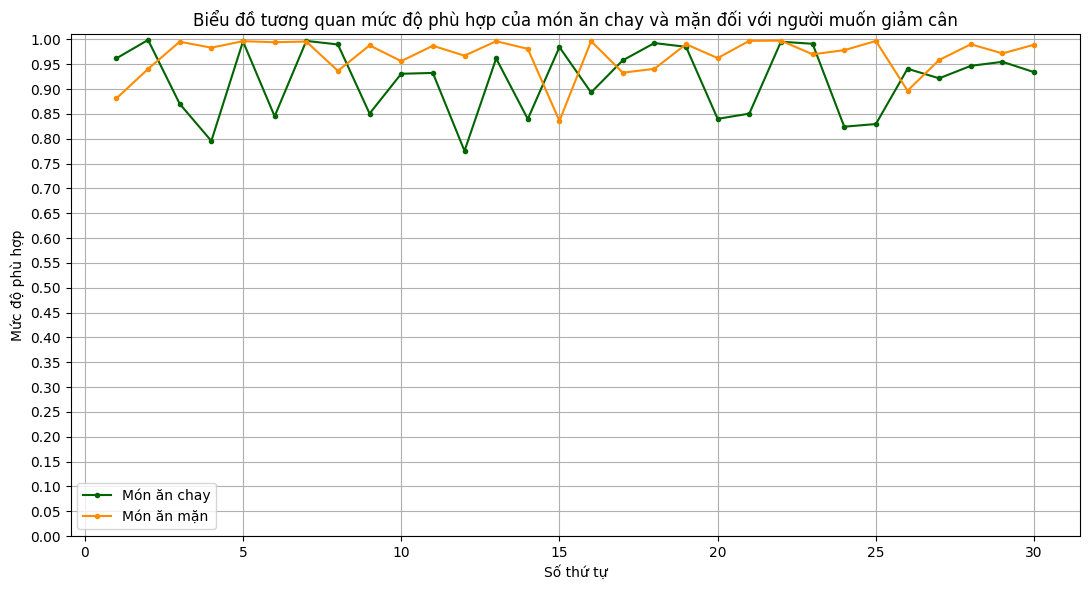
1. **Data Analysis**

* Statistic analysis

First, we'll perform statistical analysis by creating plots and boxplots to compare plant-based food and animal-based food nutrients.

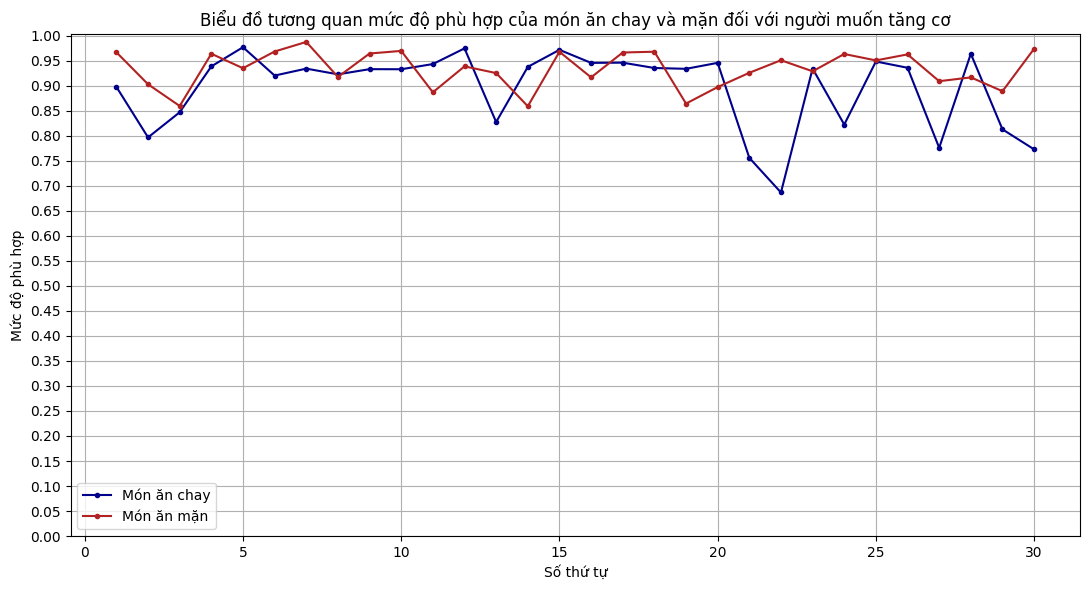
* Data Visualization

**Hypothesis 1:** People who lose weight should eat plant-based foods.



Looking at the visual chart, it's easy to see that the value of animal-based foods (yellow) tends to be higher than that of plant-based (green) in terms of how suitable they are for people who want to lose weight.

**Hypothesis 2:** People who need to gain muscle should eat animal-based foods.



Looking at the visual chart, it's easy to see how suitable animal-based foods (red) tend to be for people who want to gain weight, tending to be higher than plant-based (blue).

1. **Results**

* Statistical testing

**Hypothesis 1:** People who lose weight should eat plant-based foods.

Let x be the suitability of plant-based dishes for people on a diet, x follows a normal distribution 𝑁(𝜇1,𝜎1^2)

Let y be the suitability of plant-based dishes for people on a diet, y follows a normal distribution 𝑁(𝜇2,𝜎2^2)

Let 𝑛1=980 ,𝑛2=913

Set:

𝐻0:𝜇1=𝜇2

𝐻1:𝜇1>𝜇2

Rejection region W: (𝑧𝛼,+∞)

With a significance level of 5%, the rejection region is (1.6449,+∞)

We calculate T = (𝜇1 - 𝜇1)/ sqrt[ (𝜎1^2/n1) + (𝜎1^2/n1) ]

+ If T ∈ W then we reject the H0 hypothesis. There is enough evidence to conclude that plant-based food is more suitable for people who want to lose weight than animal-based food.

+ If T ∉ W then we cannot reject the H0 hypothesis. There is not enough evidence to conclude that plant-based food is more suitable for people who want to lose weight than animal-based food.

T1 value: -10.640064628344177, rejection domain W = (1,6449 ; + inf)  
  
Because T is not in the rejection domain W, we can not reject the H0 theory. There is not enough evidence to conclude that plant-based foods are better for people who want to lose weight than animal-based foods. So we conclude with a 5% significance, that the first hypothesis is false.

**Hypothesis 2:** People who need to gain muscle should eat animal-based food.

Let x be the suitability of plant-based dishes for people on a diet, x follows a normal distribution 𝑁(𝜇1,𝜎1^2)

Let y be the suitability of plant-based dishes for people on a diet, y follows a normal distribution 𝑁(𝜇2,𝜎2^2)

Let 𝑛1=980,𝑛2=913

Set:

𝐻0:𝜇1=𝜇2

𝐻1:𝜇1<𝜇2

Rejection region W: (-∞, - 𝑧𝛼)

With a significance level of 5%, the rejection region is (-∞, -1.6449)

We calculate T = (𝜇1 - 𝜇1)/ sqrt[ (𝜎1^2/n1) + (𝜎1^2/n1) ]

+ If T ∈ W then we reject the H0 hypothesis. There is enough evidence to conclude that animal-based foods are more suitable for people who want to gain muscle than plant-based foods.

+ If T ∉ W then we cannot reject the H0 hypothesis. There is not enough evidence to conclude that animal-based foods are more suitable for people who want to lose weight than plant-based foods.

T2 value: -14.306926114721003, rejection W = (- inf ; -1,6449)  
  
Because Z is in the rejected domain W, we reject the H0 theory. There is enough evidence to conclude that animal-based food is better for muscle growth than plant food. So we conclude with a 5% significance, that the second hypothesis is true.

1. **Discussion**

Our analysis focused on comparing the nutritional content of plant-based and animal-based foods to evaluate their effectiveness in weight loss and muscle gain. The statistical tests provided insights into the nutritional differences and their implications for health.

1. **Nutritional Differences Between Plant-Based and Animal-Based Foods**:
   * Plant-based foods generally have fewer calories and more fiber compared to animal-based foods.
   * Animal-based foods tend to have higher protein content and essential amino acids, which are crucial for muscle growth.
   * These differences highlight the varied benefits of each food type in achieving specific health goals.
2. **Relationship Between Nutrition and Health Outcomes**:
   * By analyzing the nutritional factors and health scores through regression analysis, we observed that high-protein diets (typically animal-based) are more effective for muscle gain.
   * Diets with higher fiber and lower calorie content (typically plant-based) are traditionally associated with weight loss, but our data showed that satiety and appetite control from animal-based foods also contribute to effective weight management.

**Implications** The findings from our research have practical implications for dietary recommendations:

1. **Weight Loss**:
   * Although plant-based foods are often recommended for weight loss due to their lower calorie density and higher fiber content, our study suggests that animal-based foods can also be effective. Their ability to promote satiety and reduce appetite plays a significant role in weight management.
2. **Muscle Gain**:
   * For individuals aiming to gain muscle, animal-based foods are more beneficial due to their high protein content and essential amino acids. This supports the use of animal-based diets in muscle-building regimens.
3. **Customized Diet Plans**:
   * The results provide a basis for tailored dietary advice. Depending on individual health goals—whether weight loss or muscle gain—specific recommendations can be made to optimize the diet.
4. **Conclusion**

**Summary of Findings** This study aimed to test two hypotheses regarding the effectiveness of plant-based and animal-based foods for weight loss and muscle gain:

1. **Hypothesis 1: People who lose weight should eat plant-based foods**:
   * Our statistical analysis did not find sufficient evidence to support this hypothesis. While plant-based foods have benefits such as lower calorie density and higher fiber, animal-based foods also contribute to weight loss by enhancing satiety and reducing appetite.
2. **Hypothesis 2: People who need to gain muscle should eat animal-based foods**:
   * The data strongly supported this hypothesis. Animal-based foods, rich in protein and essential amino acids, were shown to be more effective for muscle growth compared to plant-based foods.
3. **Conclusions** Our research concludes that:

* Both plant-based and animal-based foods have their own unique advantages for weight loss and muscle gain.
* A balanced approach that considers individual health goals and the specific benefits of each food type is essential for effective dietary planning.

By integrating these findings, individuals can make informed dietary choices that align with their health objectives, whether it is to lose weight or build muscle.

1. **Limitations**

* **Limited sample size:** The data collected may not be large enough to represent the entire population. This may affect the generality of the results.
* **Other factors:** In addition to diet, there are many other factors that affect weight loss and muscle gain, such as physical activity, sleep, and genetics. We haven't been able to control all of these factors in our research.
* **Limited feature:** the Calorie Ninjas API doesn’t provide detailed information about other nutrients like vitamins and minerals. Which can create biases toward animal-based food.

1. **Casual Inference**

**Hypothesis 1:** While our analysis did not find strong evidence supporting the superiority of plant-based diets for weight loss, we cannot entirely rule out the possibility that they may be beneficial for certain individuals.

**Hypothesis 2:** Our findings strongly suggest that animal-based diets are more effective for muscle gain, likely due to higher protein content and the presence of essential amino acids. However, while the observational nature of the data limits causal certainty, the results are consistent with existing literature and biological plausibility.

1. **References**

[World Health Organization (WHO) - The World Health Organization often has reports and guidelines on nutrition, including diets for health and weight loss.](https://www.who.int/health-topics/nutrition?fbclid=IwAR34JUDaXMKT8VETm-CXo9vUmoW5lH0HudT3P5jp5gSVHkRgByetptofJxU#tab=tab_1)

[National Institutes of Health (NIH) - The US National Institutes of Health is also a resource that can provide research and guidance on nutrition and weight loss.](https://www.niddk.nih.gov/health-information/diet-nutrition?fbclid=IwAR1IDJaihmiHRNVuL1gDqQFEvWl3mV1dBKkpV6QSL9xsIqwf-avfY8CsZqA)

[British Journal of Sports Medicine (BJSM): This journal publishes research on sports medicine and sports nutrition, which can inform nutritional strategies to optimize muscle growth.](https://bjsm.bmj.com/?fbclid=IwAR1KVBazwuHBsLtVjgxHeosvULFOluOQd7317Ric7lFdYRYLXupTXTRMUqo)

[CalorieNinjas API](https://calorieninjas.com/?fbclid=IwAR1u0IuYswKrtnRvqquQPL72Fyg_EOjg_T86w5CZxSwDjUcuVxBYgj574RA)