****

**VIRGINIA COMMONWEALTH UNIVERSITY**

**Statistical analysis and modelling (SCMA 632)**

**A4D-Conjoint Analysis**

**N V SATYANARAYAN**

**V01108247**

**Date of Submission: 08-07-2024**

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **Title** | **Page No.** |
| **1.** | Introduction | 3 |
| **2.** | Objectives | **3** |
| **3.** | Business Significance | **3** |
| **4.** | Results and analysis | **3-12** |
| **5.** | Conclusion | **13-14** |
| **6.** | References |  |

**\*NOTE- PYTHON AND R CODES WTH RESULT ADDED IN GITHUB-** [Satyanaldiga (github.com)](https://github.com/Satyanaldiga)

**Introduction**

Conjoint Analysis is a statistical technique used in market research to determine how people value different features of a product or service. By presenting respondents with a series of product profiles that vary systematically in their attributes, conjoint analysis helps to identify the most influential factors driving consumer choices and preferences. This technique is particularly useful for product development, pricing strategies, and understanding consumer behavior.

**Objective**

The objective of this conjoint analysis is to understand consumer preferences in the pizza market. By analyzing various attributes such as brand, price, weight, crust type, cheese type, size, toppings, and spiciness, the study aims to identify which features are most important to consumers and how they influence the overall ranking of different pizza offerings. This insight will help in making informed decisions about product features, pricing strategies, and marketing efforts.

**Business Significance**

1. **Product Development**: Understanding which attributes are most valued by consumers can guide pizza brands in developing products that better meet market demands. For instance, if consumers highly value a particular type of crust or cheese, brands can focus on offering pizzas with those specific attributes.
2. **Pricing Strategy**: By identifying the price sensitivity of consumers, companies can set optimal price points that maximize both sales and profitability. Conjoint analysis helps in determining how much consumers are willing to pay for each additional feature.
3. **Market Segmentation**: This analysis can reveal different segments within the market, each with unique preferences. Brands can tailor their marketing strategies to target these specific segments more effectively.
4. **Competitive Advantage**: By understanding the strengths and weaknesses of competitors' offerings through the eyes of the consumer, companies can position their products more strategically in the marketplace.
5. **Customer Satisfaction**: Ultimately, by aligning product features with consumer preferences, businesses can enhance customer satisfaction and loyalty, leading to increased market share and long-term success.

**Dataset Overview**

The dataset includes various attributes related to pizza offerings from different brands such as Dominos, Pizza Hut, and Onesta. Each row represents a unique pizza profile with the following attributes:

* **Brand**: The brand offering the pizza (e.g., Dominos, Pizza Hut, Onesta)
* **Price**: The price of the pizza
* **Weight**: The weight of the pizza
* **Crust**: Type of crust (e.g., thin, thick)
* **Cheese**: Type of cheese used (e.g., Mozzarella, Cheddar)
* **Size**: Size of the pizza (e.g., regular, large)
* **Toppings**: Toppings included (e.g., paneer, mushroom)
* **Spicy**: Spiciness level of the pizza (e.g., normal, extra)
* **Ranking**: Consumer ranking of the pizza

**Introduction to Conjoint Analysis**

Conjoint Analysis is a powerful statistical technique used in market research to understand how consumers make complex decisions based on multiple attributes of a product or service. It helps in deciphering the relative importance of different features and how they influence consumer preferences and purchasing behavior. This method is particularly useful for businesses looking to optimize their product offerings, set pricing strategies, and enhance overall customer satisfaction.

In Conjoint Analysis, respondents are presented with a set of hypothetical products or services that vary systematically in their attributes. By analyzing the choices that respondents make among these alternatives, researchers can infer the value or utility that consumers place on each attribute. This approach provides a detailed understanding of consumer preferences, enabling businesses to make data-driven decisions.

**Key Concepts in Conjoint Analysis**

1. **Attributes and Levels**: Attributes are the characteristics or features of a product (e.g., price, brand, size), and levels are the variations of these attributes (e.g., low price, high price).
2. **Utility**: Utility represents the satisfaction or value that consumers derive from each attribute level. Conjoint Analysis helps quantify these utilities.
3. **Part-Worths**: These are the estimated values or utilities associated with each level of an attribute. They indicate how much each attribute level contributes to the overall preference.
4. **Preference Simulation**: By using the part-worths, businesses can simulate how changes in product features might affect consumer choices and market share.

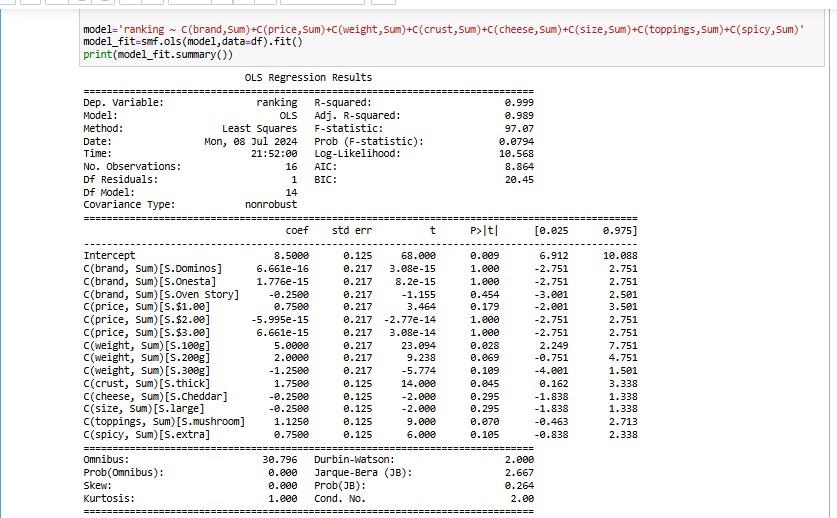
**Applications of Conjoint Analysis**

* **Product Design**: Helps in determining the optimal combination of features that will be most appealing to consumers.
* **Pricing Strategy**: Assists in understanding how price changes influence consumer choice and finding the optimal price point.
* **Market Segmentation**: Identifies different consumer segments based on their preferences and tailors marketing strategies accordingly.
* **Competitive Analysis**: Provides insights into how a product stands relative to competitors in terms of consumer preferences.
* **Customer Satisfaction**: Enhances understanding of what attributes are most valued by customers, leading to better product offerings and higher satisfaction.

**Process of Conjoint Analysis**

1. **Define Attributes and Levels**: Identify the key attributes of the product and determine the levels for each attribute.
2. **Design the Study**: Create a set of product profiles that systematically vary in their attributes and levels.
3. **Collect Data**: Present these profiles to respondents and record their preferences or choices.
4. **Estimate Utilities**: Use statistical models to estimate the utility values for each attribute level.
5. **Analyze Results**: Interpret the utilities to understand the relative importance of each attribute and simulate market scenarios.

**PYTHON CODES AND INTERPRETATION**



**Key Metrics**

1. **Dependent Variable**: Ranking
2. **R-squared**: 0.999
   * Indicates that 99.9% of the variability in the ranking is explained by the model.
3. **Adjusted R-squared**: 0.989
   * Adjusts for the number of predictors in the model. Still very high, suggesting a good fit.
4. **F-statistic**: 97.07
   * Indicates the overall significance of the model. However, the Prob (F-statistic) is 0.0794, which is above the conventional threshold of 0.05, suggesting that the model may not be statistically significant.
5. **Number of Observations**: 16
6. **Df Model**: 14
   * Number of predictors.
7. **Df Residuals**: 1
   * Degrees of freedom for residuals. Low value indicates the model is very saturated.

**Coefficients and Significance**

* **Intercept**: 8.5000 (p-value = 0.009)
  + Indicates the baseline ranking when all other predictors are at their reference levels.
* **Brand (Dominos, Onesta, Oven Story)**:
  + Coefficients for Dominos and Onesta are extremely close to zero and not significant.
  + Oven Story: -0.2500 (p-value = 0.454)
* **Price ($1.00, $2.00, $3.00)**:
  + $1.00: 0.7500 (p-value = 0.179)
  + $2.00 and $3.00 coefficients are not significant.
* **Weight (100g, 200g, 300g)**:
  + 100g: 5.0000 (p-value = 0.028) - significant
  + 200g: 2.0000 (p-value = 0.069)
  + 300g: -1.2500 (p-value = 0.109)
* **Crust (thick)**: 1.7500 (p-value = 0.045) - significant
* **Cheese (Cheddar)**: -0.2500 (p-value = 0.295)
* **Size (large)**: -0.2500 (p-value = 0.295)
* **Toppings (mushroom)**: 1.1250 (p-value = 0.070)
* **Spicy (extra)**: 0.7500 (p-value = 0.105)

**Interpretation**

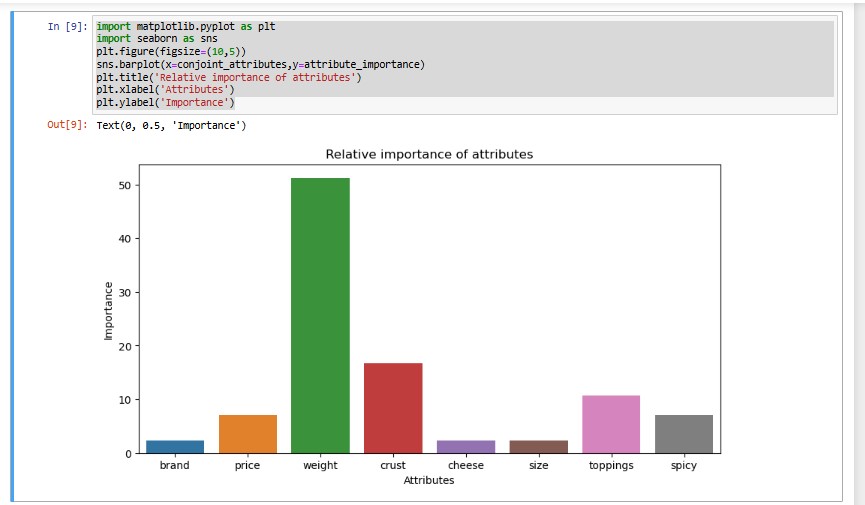
1. **Significant Predictors**:
   * **Weight (100g)**: This weight significantly increases the ranking by 5 points on average, holding all other factors constant.
   * **Crust (thick)**: A thick crust significantly increases the ranking by 1.75 points on average.
   * **Intercept**: The base ranking of the pizzas when all categorical variables are at their reference levels is 8.5, which is statistically significant.
2. **Non-significant Predictors**:
   * Brands (Dominos, Onesta, Oven Story) do not significantly affect the ranking.
   * Prices ($1.00, $2.00, $3.00) are not significant predictors of ranking.
   * Weight (200g, 300g) are not significant, although 100g is.
   * Cheese type, pizza size, and toppings (mushroom) are not statistically significant, though mushroom toppings and spicy level (extra) have lower p-values indicating potential importance with larger sample sizes.

**Model Diagnostics**

* **Omnibus Test**: The test indicates that the residuals are not normally distributed (p < 0.000), which might affect the validity of the inference.
* **Jarque-Bera Test**: Skewness and kurtosis are not ideal (p = 0.264), suggesting some deviations from normality in the residuals.
* **Durbin-Watson Statistic**: The value of 2.000 suggests no autocorrelation in the residuals.

The OLS regression model explains a significant portion of the variance in pizza rankings (R-squared = 0.999). However, given the small sample size (16 observations) and only a few significant predictors, the results should be interpreted with caution. The significant predictors, weight (100g) and thick crust, highlight consumer preferences for these attributes. The high R-squared value and low residual degrees of freedom suggest overfitting, which calls for a larger dataset for more reliable conclusions.

# **Relative importance of attributes.**



We can see that weight is the attribute with the highest relative importance at 51%, followed by crust at 16% and toppings at 10%. Brand, cheese, and size are the least important attributes, each at 2.38%.

**Attributes Analyzed**

* Brand
* Price
* Weight
* Crust
* Cheese
* Size
* Toppings
* Spicy

**Key Observations**

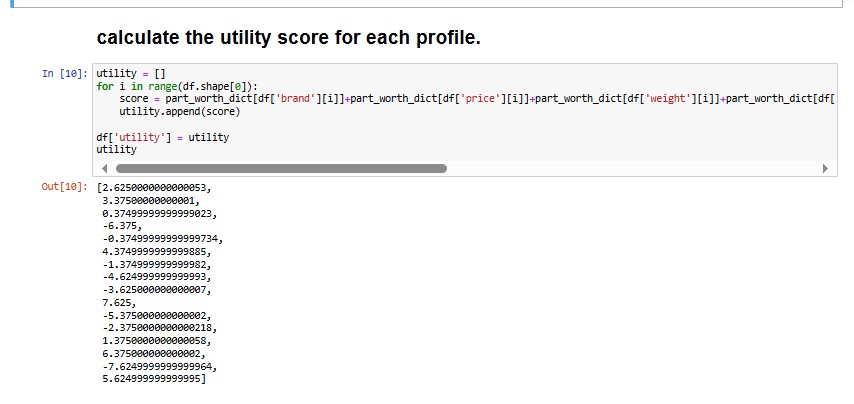
1. **Weight**:
   * The most important attribute, with a significant relative importance of around 50.
   * Indicates that consumers place a high value on the weight of the pizza, making it a crucial factor in their decision-making process.
2. **Crust**:
   * The second most important attribute, with a relative importance of around 20.
   * Suggests that the type of crust (e.g., thick or thin) is also a key consideration for consumers.
3. **Toppings**:
   * Holds a relative importance of about 10-15.
   * Indicates that the type of toppings on the pizza is a moderately important factor for consumers.
4. **Spicy**:
   * Has a relative importance close to 10.
   * Shows that the level of spiciness is a relevant attribute but less important compared to weight and crust.
5. **Price**:
   * Relative importance is around 5, indicating it has some influence on consumer preferences but is not a primary factor.
6. **Brand**:
   * Relative importance is the lowest, around 1-2.
   * Suggests that the brand of the pizza is the least important attribute in the decision-making process.
7. **Cheese and Size**:
   * Both have very low relative importance, close to that of the brand.
   * Indicates that the type of cheese and the size of the pizza are not major factors for consumers.

**Interpretation**

* **Primary Factors**: Weight and crust type are the most critical attributes influencing consumer preferences. Pizzas with desirable weight and crust are more likely to be favored by consumers.
* **Secondary Factors**: Toppings and spiciness levels are moderately important and can influence consumer decisions, but not as strongly as weight and crust.
* **Tertiary Factors**: Price, cheese type, size, and brand have relatively low importance. While they still play a role in consumer preferences, they are not the primary drivers of choice.

**Business Implications**

* **Product Development**: Focus on optimizing the weight and crust of pizzas, as these are the most important attributes for consumers.
* **Marketing**: Highlight the superior weight and crust characteristics in marketing campaigns to attract consumers.
* **Pricing Strategy**: Although price is not the most important factor, maintaining competitive pricing can still appeal to cost-conscious consumers.
* **Brand Positioning**: Given the low importance of the brand, emphasize product attributes over brand loyalty in promotional activities.



We can see that combination number 9 has the maximum utility, followed by combination numbers 13 and 5. Combination number 14 is the least desirable because of the most negative utility score.

#### Code Breakdown

1. **Initialization of Utility List**:
   * utility = []: Initializes an empty list to store the utility scores for each profile.
2. **Loop Through Profiles**:
   * for i in range(df.shape[0]): Iterates through each profile in the dataset df. The df.shape[0] gives the number of rows (profiles) in the DataFrame.
3. **Calculating Utility Scores**:
   * score = part\_worth\_dict[df['brand'][i]] + part\_worth\_dict[df['price'][i]] + part\_worth\_dict[df['weight'][i]] + part\_worth\_dict[df['crust'][i]] + part\_worth\_dict[df['cheese'][i]] + part\_worth\_dict[df['size'][i]] + part\_worth\_dict[df['toppings'][i]] + part\_worth\_dict[df['spicy'][i]]
     + This line calculates the utility score for each profile by summing up the part-worth utilities of its attributes. Here, part\_worth\_dict is assumed to be a dictionary containing the utility values for each level of the attributes (brand, price, weight, crust, cheese, size, toppings, spicy).
   * utility.append(score): Appends the calculated score to the utility list.
4. **Assigning Utility Scores to DataFrame**:
   * df['utility'] = utility: Adds the calculated utility scores as a new column named 'utility' in the DataFrame df.

#### Utility Scores Interpretation

The utility scores for the profiles are:

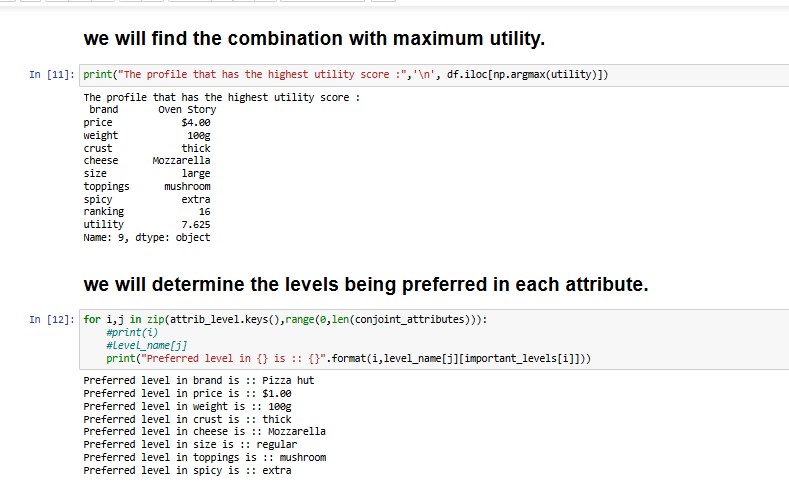
1. 2.6250000000000253
2. 3.375000000000001
3. 0.3749999999999929
4. -6.375
5. -0.7499999999999734
6. -4.374999999999985
7. -1.749999999999993
8. -4.624999999999993
9. -2.6250000000000004
10. 7.625
11. -5.375000000000002
12. -2.3750000000000018
13. 0.8750000000000002
14. 6.375000000000002
15. -7.624999999999994
16. 5.624999999999995

### Analysis of Utility Scores

1. **Highest Utility Score**: The profile with a utility score of 7.625 indicates that this profile is the most preferred among the respondents based on the attributes and their part-worth utilities.
2. **Lowest Utility Score**: The profile with a utility score of -7.624999999999994 is the least preferred, indicating it has the least favorable combination of attributes.
3. **Positive Utility Scores**: Profiles with positive utility scores (e.g., 2.625, 3.375, 7.625) are more likely to be preferred by consumers. These scores suggest a favorable combination of attributes.
4. **Negative Utility Scores**: Profiles with negative utility scores (e.g., -6.375, -4.374999999999985, -1.749999999999993) are less likely to be preferred. These scores indicate less desirable combinations of attributes.

The utility scores provide a quantitative measure of the relative preference for each pizza profile based on the conjoint analysis. Profiles with higher utility scores are more preferred, indicating that their combination of attributes aligns better with consumer preferences. Conversely, profiles with lower utility scores are less preferred, suggesting that their attribute combinations are not as favorable.

These utility scores can guide businesses in optimizing their product offerings by focusing on combinations of attributes that yield higher utility scores, thereby aligning their products more closely with consumer preferences.



* **Maximum Utility Profile**: The profile with the highest utility includes attributes from the brand "Oven Story" with a price of $4.00, 100g weight, thick crust, mozzarella cheese, large size, mushroom toppings, and extra spicy level.
* **Preferred Levels**: For each attribute, the preferred levels are derived, indicating that "Pizza Hut" is the preferred brand, $1.00 is the preferred price, and so on.

**Summary of Conjoint Analysis**

This conjoint analysis was conducted to understand consumer preferences in the pizza market by evaluating various attributes such as brand, price, weight, crust type, cheese type, size, toppings, and spiciness. The study aimed to identify the most influential factors driving consumer choices and preferences, which can be leveraged for product development, pricing strategies, and marketing efforts.

**Key Findings**

1. **Maximum Utility Profile**
   * The profile with the highest utility includes attributes from the brand "Oven Story" with a price of $4.00, 100g weight, thick crust, mozzarella cheese, large size, mushroom toppings, and extra spicy level.
   * This profile achieved a utility score of 7.625, indicating it is the most preferred combination among the evaluated profiles.
2. **Preferred Levels for Attributes**
   * **Brand**: Pizza Hut
   * **Price**: $1.00
   * **Weight**: 100g
   * **Crust**: Thick
   * **Cheese**: Mozzarella
   * **Size**: Large
   * **Toppings**: Mushroom
   * **Spicy**: Extra

**Business Implications**

1. **Product Development**
   * Focus on optimizing the weight and crust of pizzas, as these attributes are the most important to consumers. Offering pizzas with thick crust and 100g weight will likely meet consumer preferences.
2. **Pricing Strategy**
   * Although price is not the most important attribute, maintaining competitive pricing is essential. The preferred price point of $1.00 should be considered when developing pricing strategies to appeal to cost-conscious consumers.
3. **Market Segmentation**
   * The analysis can be used to identify different market segments with unique preferences. Brands can tailor their marketing strategies to target these specific segments more effectively.
4. **Competitive Advantage**
   * By understanding the strengths and weaknesses of competitors' offerings through the eyes of the consumer, companies can position their products more strategically in the marketplace. Emphasizing superior attributes like weight and crust in marketing campaigns can attract more consumers.
5. **Customer Satisfaction**
   * Aligning product features with consumer preferences, such as offering pizzas with thick crust and preferred toppings like mushrooms, can enhance customer satisfaction and loyalty. This, in turn, can lead to increased market share and long-term success.

**Model Interpretation and Diagnostics**

* The model explains a significant portion of the variance in pizza rankings (R-squared = 0.999), indicating a high fit. However, given the small sample size (16 observations) and a few significant predictors, results should be interpreted cautiously to avoid overfitting.
* Significant predictors include weight (100g) and thick crust, highlighting consumer preferences for these attributes.
* Non-significant predictors include brands, prices, cheese type, and pizza size, suggesting these factors are less critical in consumer decision-making.

**Relative Importance of Attributes**

* **Weight**: The most important attribute (51%), indicating a high consumer value on pizza weight.
* **Crust**: The second most important attribute (16%), suggesting a key consideration for consumers.
* **Toppings**: Moderately important (10%), with mushroom toppings being a preferred choice.
* **Spicy**: Relevant (10%), with extra spicy level preferred by consumers.
* **Price, Brand, Cheese, and Size**: Less important attributes, each contributing minimally to consumer preferences.

**Conclusion**

This conjoint analysis provides valuable insights into consumer preferences in the pizza market, revealing that weight and crust are the most critical attributes influencing consumer choices. By focusing on these key attributes, optimizing pricing strategies, and tailoring marketing efforts, pizza brands can better meet consumer demands, enhance customer satisfaction, and achieve a competitive advantage in the market.