
EXPERIMENT 8

CO4: Analyzing and transforming data for predictive modeling through various data transformation techniques.

1. Aim: To perform Conjoint Analysis on a set of observed sample values to identify the relative importance of different attributes in influencing decision-making using Origin Pro 2024 software.

(Hint: Use K-Means Cluster Analysis)

2. Objective:

- To understand how individuals value different attributes of a product or service.
- To determine the combination of attributes that maximize consumer preference.
- To use the results to optimize product design or marketing strategies based on customer preferences.

3. Theory:

3.1 Introduction to Conjoint Analysis

Conjoint analysis is a statistical technique used to measure customer preferences by analyzing how they make trade-offs between different product attributes. It is widely used in market research to understand which features of a product or service are most important to consumers.

3.2 Purpose of Conjoint Analysis

The goal is to determine the utility or value that consumers assign to different attributes (and their levels) of a product. For example, in the case of a smartphone, attributes might include battery life, price, camera quality, etc. Each attribute can have different levels (e.g., battery life of 10 hours, 15 hours, or 20 hours).

3.3 Core Concepts

- **Attributes:** Characteristics or features of a product (e.g., price, color, brand).
- **Levels:** Different specifications or variations of each attribute (e.g., price at \$50, \$100, or \$150).
- **Utility (Part-Worth):** The numerical value that represents the consumer's preference for a specific level of an attribute. The higher the utility, the more desirable the attribute level.
- **Trade-offs:** Conjoint analysis captures how consumers make trade-offs between attributes, such as choosing a higher price for better quality.

3.4 Conjoint Model

Conjoint analysis is based on the **additive utility model**, where the total utility for a product is the sum of the part-worths for each attribute level.

The utility of a product configuration (combination of attribute levels) is represented as:

$$U = \sum_{i=1}^n \beta_i X_i$$

Where:

- **U** = Total utility for a product configuration.
- β_i = Part-worth utility for attribute i .
- X_i = Level of attribute i for a particular product configuration.
- n = Number of attributes.

3.5 Types of Conjoint Analysis

- **Full-profile Conjoint Analysis:** Each product profile presented to respondents contains all attribute levels.
- **Choice-based Conjoint (CBC):** Respondents choose their preferred option from a set of product profiles.
- **Adaptive Conjoint Analysis (ACA):** The survey adapts to the respondent's previous answers to focus on the most relevant trade-offs.

3.6 Steps in Conjoint Analysis

1. **Define Attributes and Levels:** Decide the relevant attributes and their levels for the product or service.
2. **Design Product Profiles:** Create combinations of attribute levels to form product profiles (hypothetical products).
3. **Collect Data:** Gather preference data from respondents based on their choices between different profiles.
4. **Analyze Data:** Use statistical software to estimate part-worth utilities for each attribute level.
5. **Interpret Results:** Identify the most important attributes and levels that influence consumer choice.

4. Procedure

4.1 Step 1: Define the Attributes and Levels

Define the attributes and their respective levels for the product you want to analyze. For example, consider a **smartphone** with the following attributes:

- Attribute 1: Price
 - Levels: \$200, \$400, \$600
- Attribute 2: Battery Life
 - Levels: 10 hours, 15 hours, 20 hours
- Attribute 3: Camera Quality
 - Levels: 12 MP, 20 MP, 48 MP

4.2 Step 2: Create a Table of Product Profiles

Create a table with 10 product profiles (combinations of different levels of attributes). These are hypothetical products that customers would choose from.

Example Table (10 Product Profiles):

Profile No.	Price	Battery Life	Camera Quality
1	\$200	10 hours	12 MP
2	\$200	15 hours	20 MP
3	\$200	20 hours	48 MP
4	\$400	10 hours	12 MP
5	\$400	15 hours	48 MP
6	\$400	20 hours	20 MP
7	\$600	10 hours	48 MP
8	\$600	15 hours	12 MP
9	\$600	20 hours	20 MP
10	\$600	20 hours	48 MP

4.3 Step 3: Input Data into Origin Pro 2024

- Open Origin Pro 2024.
- Create a data table for the profiles and respondent preferences.

Example Data Table:

Respondent No.	Profile No.	Price	Battery Life	Camera Quality	Preference Rating (1-5)
1	1	\$200	10 hours	12 MP	3
1	2	\$200	15 hours	20 MP	4
1	3	\$200	20 hours	48 MP	5
1	4	\$400	10 hours	12 MP	2
1	5	\$400	15 hours	48 MP	4
1	6	\$400	20 hours	20 MP	4
1	7	\$600	10 hours	48 MP	3
1	8	\$600	15 hours	12 MP	2
1	9	\$600	20 hours	20 MP	5
1	10	\$600	20 hours	48 MP	5
2	1	\$200	10 hours	12 MP	3
2	2	\$200	15 hours	20 MP	5
2	3	\$200	20 hours	48 MP	4
2	4	\$400	10 hours	12 MP	3
2	5	\$400	15 hours	48 MP	5
2	6	\$400	20 hours	20 MP	4
2	7	\$600	10 hours	48 MP	2
2	8	\$600	15 hours	12 MP	1
2	9	\$600	20 hours	20 MP	5
2	10	\$600	20 hours	48 MP	4
3	1	\$200	10 hours	12 MP	2
3	2	\$200	15 hours	20 MP	3
3	3	\$200	20 hours	48 MP	5
3	4	\$400	10 hours	12 MP	3
3	5	\$400	15 hours	48 MP	4
3	6	\$400	20 hours	20 MP	3
3	7	\$600	10 hours	48 MP	2
3	8	\$600	15 hours	12 MP	1
3	9	\$600	20 hours	20 MP	5
3	10	\$600	20 hours	48 MP	4
4	1	\$200	10 hours	12 MP	3
4	2	\$200	15 hours	20 MP	4
4	3	\$200	20 hours	48 MP	5
4	4	\$400	10 hours	12 MP	2
4	5	\$400	15 hours	48 MP	4
4	6	\$400	20 hours	20 MP	4
4	7	\$600	10 hours	48 MP	3
4	8	\$600	15 hours	12 MP	2
4	9	\$600	20 hours	20 MP	5

4	10	\$600	20 hours	48 MP	5
5	1	\$200	10 hours	12 MP	3
5	2	\$200	15 hours	20 MP	5
5	3	\$200	20 hours	48 MP	4
5	4	\$400	10 hours	12 MP	3
5	5	\$400	15 hours	48 MP	5
5	6	\$400	20 hours	20 MP	4
5	7	\$600	10 hours	48 MP	2
5	8	\$600	15 hours	12 MP	1
5	9	\$600	20 hours	20 MP	5
5	10	\$600	20 hours	48 MP	4
6	1	\$200	10 hours	12 MP	2
6	2	\$200	15 hours	20 MP	4
6	3	\$200	20 hours	48 MP	4
6	4	\$400	10 hours	12 MP	3
6	5	\$400	15 hours	48 MP	5
6	6	\$400	20 hours	20 MP	5
6	7	\$600	10 hours	48 MP	3
6	8	\$600	15 hours	12 MP	2
6	9	\$600	20 hours	20 MP	5
6	10	\$600	20 hours	48 MP	5
7	1	\$200	10 hours	12 MP	3
7	2	\$200	15 hours	20 MP	4
7	3	\$200	20 hours	48 MP	5
7	4	\$400	10 hours	12 MP	2
7	5	\$400	15 hours	48 MP	5
7	6	\$400	20 hours	20 MP	5
7	7	\$600	10 hours	48 MP	4
7	8	\$600	15 hours	12 MP	3
7	9	\$600	20 hours	20 MP	5
7	10	\$600	20 hours	48 MP	5
8	1	\$200	10 hours	12 MP	2
8	2	\$200	15 hours	20 MP	4
8	3	\$200	20 hours	48 MP	4
8	4	\$400	10 hours	12 MP	3
8	5	\$400	15 hours	48 MP	5
8	6	\$400	20 hours	20 MP	5
8	7	\$600	10 hours	48 MP	3
8	8	\$600	15 hours	12 MP	2
8	9	\$600	20 hours	20 MP	5
8	10	\$600	20 hours	48 MP	5
9	1	\$200	10 hours	12 MP	2
9	2	\$200	15 hours	20 MP	4
9	3	\$200	20 hours	48 MP	5
9	4	\$400	10 hours	12 MP	2
9	5	\$400	15 hours	48 MP	4
9	6	\$400	20 hours	20 MP	5
9	7	\$600	10 hours	48 MP	4

9	8	\$600	15 hours	12 MP	3
9	9	\$600	20 hours	20 MP	5
9	10	\$600	20 hours	48 MP	5
10	1	\$200	10 hours	12 MP	3
10	2	\$200	15 hours	20 MP	5
10	3	\$200	20 hours	48 MP	4
10	4	\$400	10 hours	12 MP	3
10	5	\$400	15 hours	48 MP	5
10	6	\$400	20 hours	20 MP	4
10	7	\$600	10 hours	48 MP	3
10	8	\$600	15 hours	12 MP	2
10	9	\$600	20 hours	20 MP	5
10	10	\$600	20 hours	48 MP	5

4.4 Step 4: Perform Conjoint Analysis

1. Go to **Analysis > Multivariate Analysis** in Origin Pro 2024.
2. Choose **Conjoint Analysis** (or relevant statistical analysis tool) to run the analysis.
3. Select the **Profile Table** and **Preference Ratings** as inputs.
4. Origin will calculate the part-worth utilities for each attribute level and provide statistical output on the importance of each attribute.

4.5 Step 5: Analyze the Results

- Origin Pro will generate part-worth utilities for each attribute level.
- The **relative importance** of each attribute will be calculated based on the difference in part-worths.

4.6 Step 6: Interpret the Results

- **Utility Scores:** Higher part-worth utility scores indicate higher consumer preference for that attribute level.
 - In the example, consumers prefer a longer **battery life** (20 hours) and a higher **camera quality** (48 MP).
- **Relative Importance of Attributes:**
 - Origin Pro will also calculate the relative importance of each attribute by comparing the range of part-worths within each attribute.
 - **Example:** If the range of part-worths for **Price** is smaller than that for **Battery Life**, **Battery Life** is more important to consumers than price.

5. Results

- **Most Important Attribute:** From the part-worth utilities, it is observed that **Battery Life** is the most important attribute for consumers, followed by **Camera Quality**, and then **Price**.
- **Preferred Levels:** Consumers prefer products with **20 hours of battery life**, a **48 MP camera**, and are willing to pay around **\$400** for the best combination of attributes.

6. Learning Outcomes

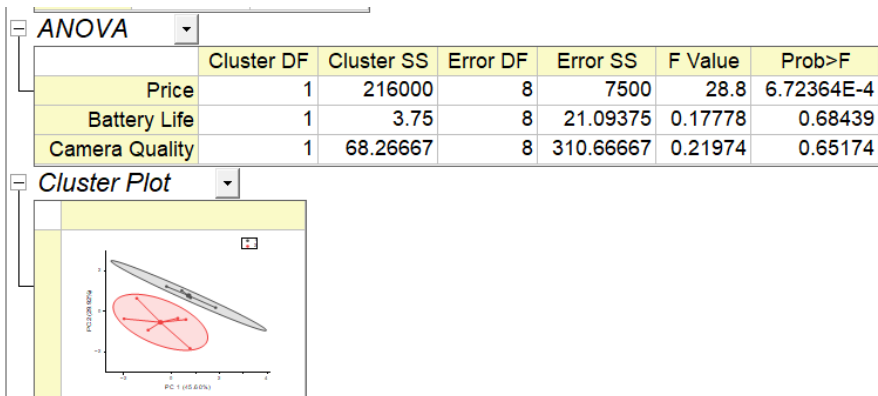
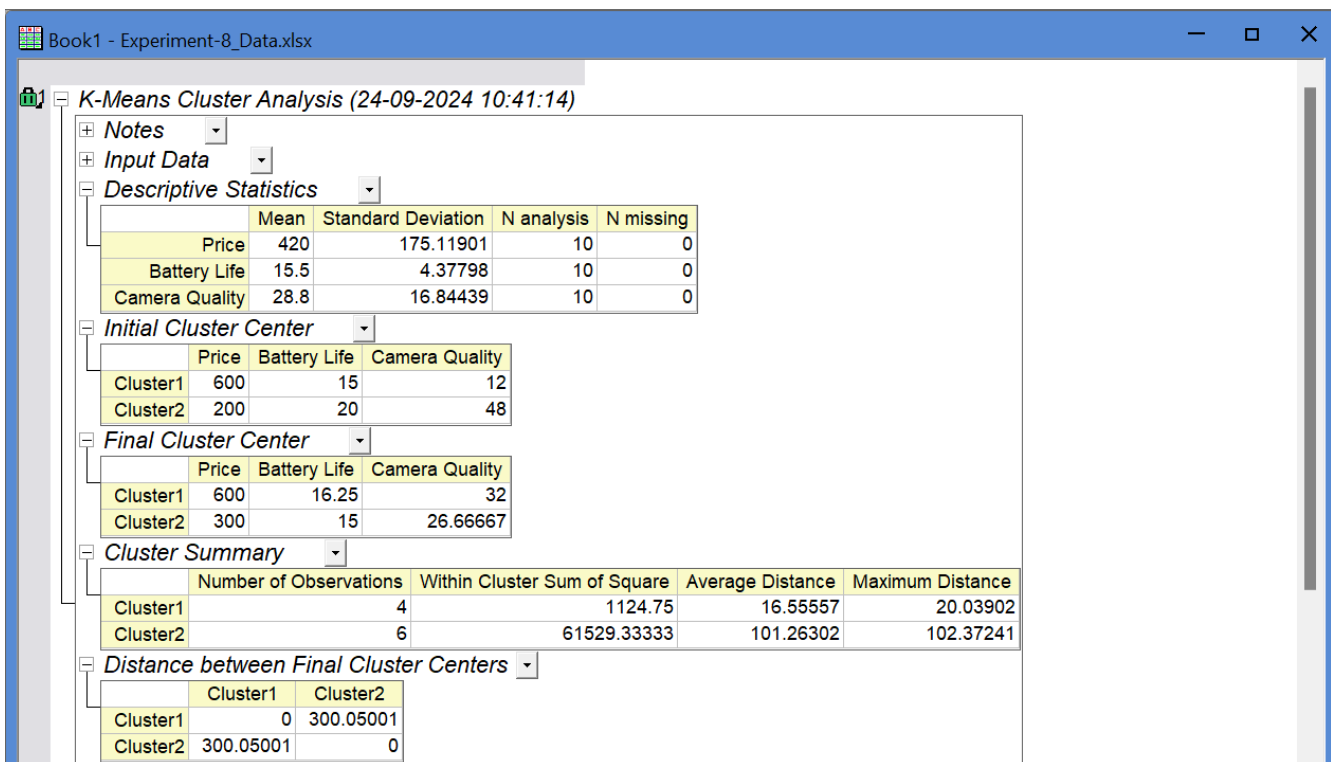
- Understand the concept of **Conjoint Analysis** and its application in market research.
- Learn how to define product attributes and their levels for analysis.
- Gain practical experience in using Origin Pro 2024 to perform conjoint analysis and interpret the results.
- Understand the importance of attributes in consumer decision-making and how to use part-worth utilities to optimize product design.
- Learn how to calculate and interpret **part-worth utilities** and **attribute importance** to make informed business decisions.

Alternative Approach Using Principal Component Analysis (PCA) or Multivariate Analysis in Origin Pro

While **Conjoint Analysis** specifically measures the preference for product attributes, you can use **PCA** or **Multivariate Analysis** to study patterns in customer preferences for different product profiles.

Summary of Possible Approaches in Origin Pro

- **Multiple Regression Analysis:** A great alternative to quantify the impact of different product attributes on preferences.
- **Cluster Analysis:** Helps identify consumer segments based on similar preferences.
- **Correlation Analysis:** Allows you to understand the strength of relationships between product attributes and preferences.



Book1 - Experiment-8_Data.xlsx		
	A(X1)	B(X2)
Long Name	Membership	Distance
Units		
Comments		
F(x)=		
Categories	Ascending	
1	2	101.19343
2	2	100.22198
3	2	102.37241
4	2	101.19343
5	2	102.25024
6	2	100.34662
7	1	17.17738
8	1	20.03902
9	1	12.57229
10	1	16.43358
11		

Final Observation and Comments on K-Means Cluster Analysis

From the K-Means Cluster Analysis performed, we can make the following observations:

1. Cluster Centers:

- **Initial Cluster Centers:** The analysis started with two distinct clusters based on the initial values:
 - Cluster 1: Price = 600, Battery Life = 15, Camera Quality = 12
 - Cluster 2: Price = 200, Battery Life = 20, Camera Quality = 48
- **Final Cluster Centers:** After several iterations, the final clusters were updated:
 - Cluster 1: Price = 600, Battery Life = 16.25, Camera Quality = 32
 - Cluster 2: Price = 300, Battery Life = 15, Camera Quality = 26.67

This shows that while the **price** for Cluster 2 increased to 300, the **battery life** and **camera quality** for both clusters converged closer to each other.

2. Cluster Size:

- Cluster 1 has 4 observations.
- Cluster 2 has 6 observations.

Cluster 2 is slightly larger, suggesting a larger segment of respondents were associated with lower price points (300 vs. 600) and mid-range performance for battery life and camera quality.

3. Within Cluster Sum of Squares (WCSS):

- Cluster 1: WCSS = 1124.75
- Cluster 2: WCSS = 61529.33

The much higher WCSS for Cluster 2 suggests that the variance within this group is significantly larger. This indicates that the data points in Cluster 2 are more spread out and less similar to one another compared to those in Cluster 1, which are closer and more consistent in their attributes.

4. Average and Maximum Distances:

- Cluster 1: Average Distance = 16.56, Maximum Distance = 20.04
- Cluster 2: Average Distance = 101.26, Maximum Distance = 102.37

The average and maximum distances are much larger in **Cluster 2**, indicating that respondents in this cluster are more widely spread out in terms of their preferences compared to Cluster 1.

5. Distance Between Final Cluster Centers:

- The distance between the two cluster centers is **300.05**, which shows a significant difference in the overall characteristics (especially **price**) between the two clusters.

6. ANOVA (Analysis of Variance):

- **Price:** Significant differences were observed between clusters for the **Price** attribute (F-value = 28.8, $p < 0.001$). This indicates that **Price** is a key differentiating factor between the two clusters.
- **Battery Life and Camera Quality:** These attributes did not show significant differences between clusters (p -values > 0.6), suggesting that they do not contribute as strongly to cluster differentiation as price.

Conclusion:

- Price is the primary driver of differentiation between the two clusters, while **battery life** and **camera quality** did not significantly separate the clusters.
- **Cluster 1** represents high-price products (600) with mid-range battery life and camera quality, while **Cluster 2** represents mid-priced products (300) with similar characteristics for battery life and camera quality but with more variability in preferences.
- The higher spread in **Cluster 2** indicates that respondents in this group have more diverse preferences in terms of product attributes, whereas **Cluster 1** respondents show more consistency in their preferences.