Contents

1. Introduction:

This report analyses customer preferences for a new circular product using conjoint analysis and principal component analysis (PCA). It focuses on six key factors influencing value creation: environmental friendliness, delivery time, service level, price, material quality, and marketing proficiency. Through 18 product profiles and customer ratings, it computes attribute part-worths, willingness to pay (WTP), and market segmentation potential. PCA examines 32 product models and 11 attributes, calculating singular values, loading factors, and variance explained. It culminates in a perceptual map illustrating the relationship between product attributes and customer perceptions.

2. AIM:

1. Product Profile Selection Rationale: Understand why 18 product profiles were chosen and why they represent an optimal selection.
2. Part-Worths Computation and WTP Calculation: Calculate part-worth utilities for each attribute and respondent and determine the Willingness to Pay (WTP).
3. Market Segmentation through Conjoint Analysis: Explore how conjoint analysis can be utilized to segment the market based on customer preferences.
4. PCA for Attribute Importance Determination: Utilize PCA to identify the most important attributes considered by customers when purchasing the product.

3. Methodology:

Task	Description	Reason
Data Preparation and Attribute Selection	Initially identified 50 attributes were distilled into six key factors: Environmental Friendliness, Delivery Time, Service Level, Price, Quality of Material, and Marketing Proficiency. Attribute levels reflect real market differences.	Ensures relevance and applicability of the study by focusing on significant factors and choosing attribute levels that accurately represent market variations.
Product Profile Selection	18 diverse product profiles chosen to represent various attribute combinations, ensuring coverage across different levels.	Vital for capturing a wide range of consumer preferences and enhancing study's robustness and reliability.
Conjoint Analysis Implementation	Utilization of conjoint package facilitated creation of fractional factorial designs and analysis of preference data, ensuring efficient generation of experimental designs, and streamlined analysis process.	Saves time and resources by providing tools for efficient analysis, enabling a comprehensive examination of consumer preferences.
Part-Worths Calculation and Visualization	Computation of part-worth utilities for each attribute level and respondent using standard conjoint analysis techniques. Visualisations like bar charts or scatter plots employed for effective communication of results.	Provides insights into relative importance of attributes and allows derivation of Willingness to Pay (WTP) crucial for pricing strategies. Visualisations enhance accessibility and understanding of results.

Market Segmentation	Hierarchical clustering is used to segment consumers based on preferences derived from part-worth utilities. Elbow plot analysis and dendrogram visualization are employed to determine an optimal number of clusters.	Enables targeted marketing strategies tailored to specific consumer segments, maximizing the effectiveness of marketing efforts. Elbow plot analysis ensures meaningful segmentation for actionable insights.
PCA Implementation	PCA conducted using 'promp' function to identify the most important attributes influencing customer perceptions. Singular values, loading factors, and PVE are computed to quantify attribute contributions.	Provides insights into attributes driving customer perceptions, aiding in prioritization of improvements. Quantification of attribute contributions facilitates informed decision-making.
Perceptual Mapping	Perceptual mapping attributes are created from factor loadings for first two principal components. The visualization technique aids in understanding customer perceptions relative to different attributes.	Facilitates strategic decision-making in product positioning and branding by providing a visual representation of customer perceptions.
Feature Importance	Assessment of attribute importance based on the absolute sum of loadings for the first two principal components. Insights guide prioritization of product features and enhancements.	Helps in understanding which attributes have a significant influence on overall customer perceptions, guiding strategic decisions regarding product development and improvement.
Scree-Plot Analysis	Calculation of proportion of variance explained by each principal component. Screen plot generated to visualize cumulative proportion of variance explained.	Facilitates interpretation of PCA results by providing an overview of variability captured by principal components, aiding in informed decision-making processes.

4. Result and Discussion:

4.1 Why were only 18 product profiles chosen?

Conducted a conjoint analysis using fractional factorial design to generate different sets of product profiles. The number of cards (product profiles) varied across three experiments: 25, 20, and 18.

Experiment	Number of Cards	Description
1	**25**	a. The initial experiment generated 25 product profiles using a fractional factorial design. b. Each profile represents a unique combination of attribute levels, such as environmental friendliness, delivery time, service level, price, quality of material, and marketing proficiency. **c.** The correlation matrix shows the correlation between the attributes within the generated profiles.
2	**20**	a. In the second experiment, the number of cards was reduced to 20 while maintaining the same attribute combinations. b. The correlation matrix indicates the correlation between attribute levels within the 20 product profiles.

3	**18**	a. Further reduction in the number of cards to 18 was made in the third experiment. b. The correlation matrix displays the correlation between attribute levels within these 18 product profiles.
4	**13**	a. Conducted with only 13 cards, resulting in a smaller set of product profiles. b. Correlation matrix illustrates attribute correlations within these 13 product profiles.

Only 18 product profiles were chosen likely due to a need for efficient resource utilization and statistical efficiency. This reduced number of profiles allows for meaningful analysis while conserving resources. Additionally, fractional factorial designs enable capturing essential information about attribute interactions with fewer experimental runs. Therefore, the 18 chosen profiles strike a balance between gathering sufficient data and minimizing resource expenditure.

4.2 Explain why this combination of product profiles is the optimal selection ?

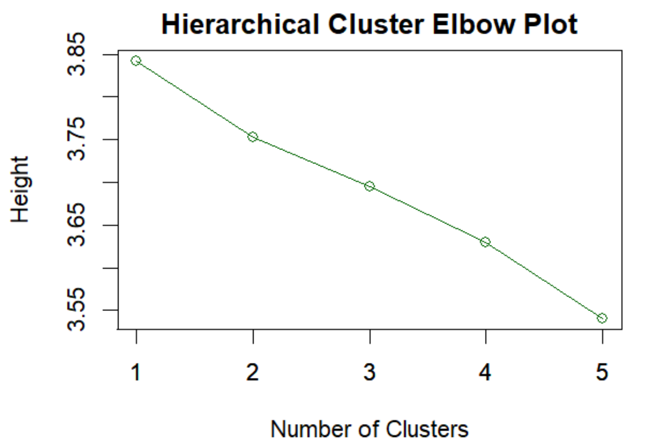
This combination of 18 product profiles is likely the optimal selection for several reasons:

Criteria	Explanation
Efficient Resource Utilization	With limited resources such as time, budget, or materials, it's essential to design experiments that provide meaningful results without unnecessary expenditure. By selecting 18 product profiles, the experiment achieves a balance between gathering sufficient data and conserving resources.
Statistical Efficiency	Fractional factorial designs are known for their efficiency in estimating main effects and interactions with fewer experimental runs compared to full factorial designs. The 18 chosen profiles were likely selected to maximize information while minimizing the number of experiments needed.
Coverage of Attribute Space	The chosen profiles likely cover a broad range of attribute combinations, ensuring that important variations and interactions are captured within the experimental space. This comprehensive coverage enables researchers to draw meaningful conclusions about how different attributes influence the outcome.
Correlation Analysis	The correlation analysis conducted on the selected profiles helps ensure that they provide diverse and independent information. By examining correlations between attributes, researchers can confirm that the chosen profiles effectively represent the variability in the experimental factors.
Balanced Representation	The selection process likely aimed to balance representation across different attribute levels and combinations. This balanced approach prevents bias and ensures that no subset of profiles dominates the experiment's outcomes.

4.3 Analysing Part-Worths and Willingness to Pay (WTP) in Conjoint Analysis.

Conjoint analysis is conducted to compute part-worth utilities for each attribute and each respondent and then calculated willingness to pay (WTP). Here's a breakdown of the steps:

Step	Description
Data Reading and Preparation	a. Start by reading two datasets: one containing preference data and another containing product profiles. b. Attribute levels in the product profiles dataset are transformed to more understandable labels using the mutate function from the dplyr package.
Computing Part-Worth Utilities	a. Compute part-worth utilities for each respondent using the caPartUtilities function. b. Establish baseline cases for each attribute and adjust the intercept accordingly. c. Adjust part-worths for each attribute level based on the baseline cases. d. The computed part-worth utilities are stored in a dataframe.
Customer Segmentation using Hierarchical Clustering	a. Hierarchical clustering is performed on the part-worth utilities using the hclust function. b. An elbow plot is created to determine the optimal number of clusters.

<table>
<tr>
<td></td>
<td>


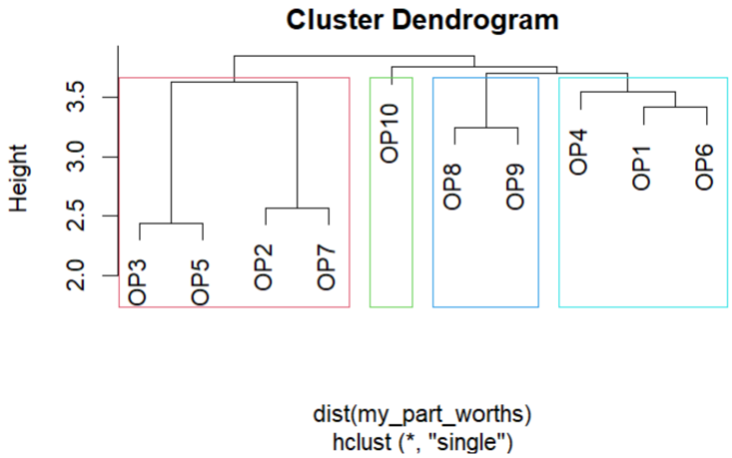
c. Between 2 and 3 Clusters:

Transition from 2 to 3 clusters, the plot’s slope may still exhibit a noticeable change, albeit less dramatic than the initial elbow point.
Adding a third cluster allows for finer granularity in grouping data points. It captures additional nuances that were previously merged into larger clusters.
Practical Implications: If the business context demands more detailed segmentation (e.g., customer personas or product categories), moving from 2 to 3 clusters might be beneficial (Li et al., 2022).

d. Beyond 3 Clusters:

When continue adding clusters beyond 3, the diminishing returns become evident.
Height values decrease further, but the improvement in model performance becomes marginal.
The risk is overfitting—creating too many clusters that might not generalize well to unseen data.
Practical Implications: Avoid excessive clustering; it can lead to complexity without substantial gains. Strive for a balance between granularity and simplicity (Li et al., 2022).

e. The dendrogram of hierarchical clusters is plotted, and clusters are delineated based on the elbow plot.



In this dendrogram, the clusters are divided into four main branches. Each branch represents a cluster generated at different heights. These clusters represent different groupings of items based on their Euclidean distance from the centroid of each cluster.
These clusters provide insights into the structure of the items being clustered. They can be used to analyze the underlying patterns in the data and to identify any specific groups or patterns that may be of interest. For example, Cluster 10 might represent a group of items with high-value Euclidean distance from the centroid, indicating that they may have a significant impact on the overall analysis.

f. Cluster labels are added to the part-worths data frame.
</td>
</tr>
<tr>
<td>Writing Results to a CSV File</td>
<td>a. The final data frame containing part-worth utilities along with cluster labels is written to a CSV file for further analysis.</td>
</tr>
</table>

4.4 Visualizing Part-Worth Utilities and Willingness to Pay (WTP):

4.4.1 Average Attribute Part Worths:

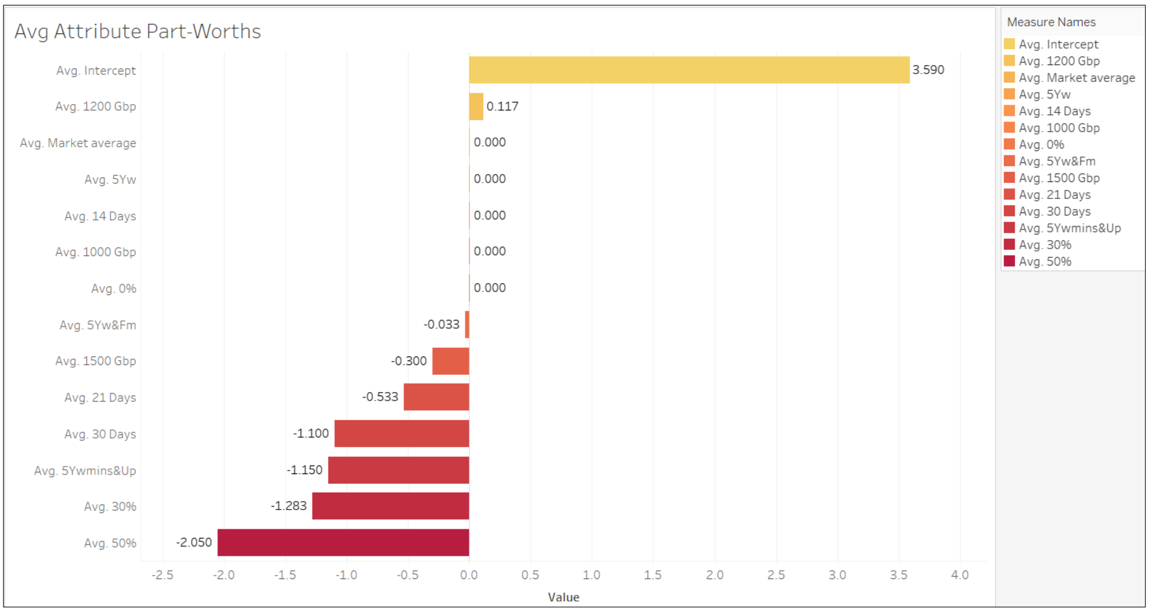


Fig.1. Attribute levels influence consumer preferences for product profiles.

a. Baseline Preference (Intercept):

The average intercept represents the baseline preference for all product profiles. Higher utility scores indicate greater preference, while lower scores indicate less preference.

b. Price Levels:

Price levels have a significant impact on preference. Higher prices, such as 1500 GBP, negatively affect preference compared to lower prices like 1000 GBP or 1200 GBP.

c. Delivery Time:

Longer delivery times are associated with reduced preference. The largest negative impact is observed for the longest delivery time of 30 days.

d. Environmental Friendliness:

Increasing levels of CO2 reduction have a diminishing positive impact on preference. Customers may not highly value environmental friendliness in this context.

e. Service Level:

While basic service levels (5YW and 5YW&FM) don't significantly affect preference, higher service levels with maintenance, installation, and upgradeability (5YWMinS&UP) negatively impact preference.

f. Quality:

Products with quality levels higher than the market average are less preferred compared to those with market-average quality.

g. Marketing Proficiency:

Higher proficiency in marketing and communication skills is preferred by customers, as indicated by the notably negative average part worth for profiles with poor communication skills.

4.4.2 Average Attribute Importance:

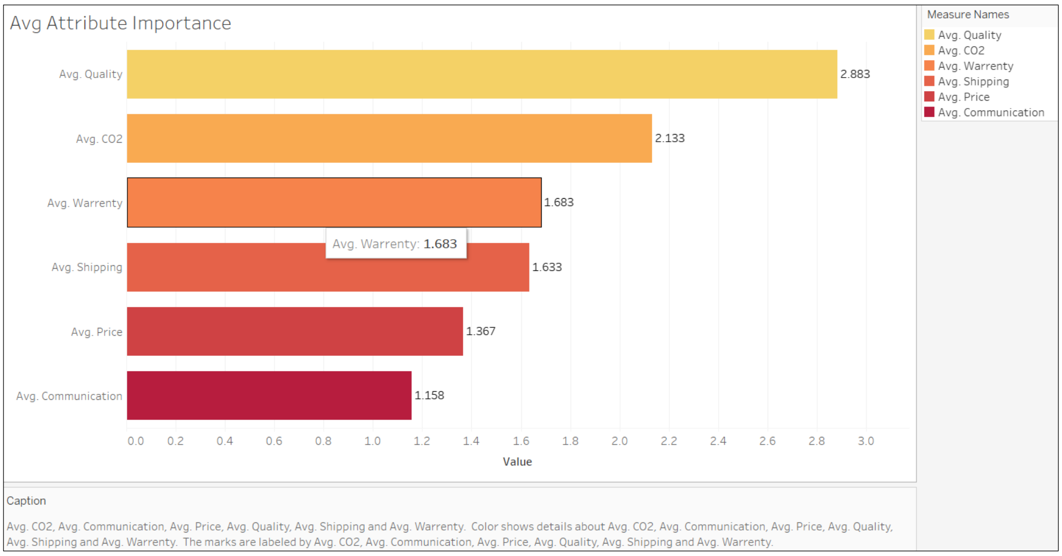


Fig.2. Average attribute importance values.

Figure 2 reveals the relative significance of various product attributes in influencing consumer preferences.

a. Quality: Ranked highest in importance, indicating that consumers prioritize the quality of materials used in the product.
b. CO2 Reduction: Indicates a significant preference for environmentally friendly products, as measured by carbon emission reduction.
c. Warranty: Consumers value longer warranty periods, seeking assurance and protection against potential issues.
d. Shipping Time : Faster shipping times are preferred, reflecting a preference for convenience and prompt delivery.
e. Price : While important, price holds less sway compared to attributes like quality and environmental friendliness.
f. Communication : Least influential among the attributes considered, suggesting that effective marketing communication is not as critical as product quality and other tangible features.

4.4.3 Percentage Average Importance Attribute:

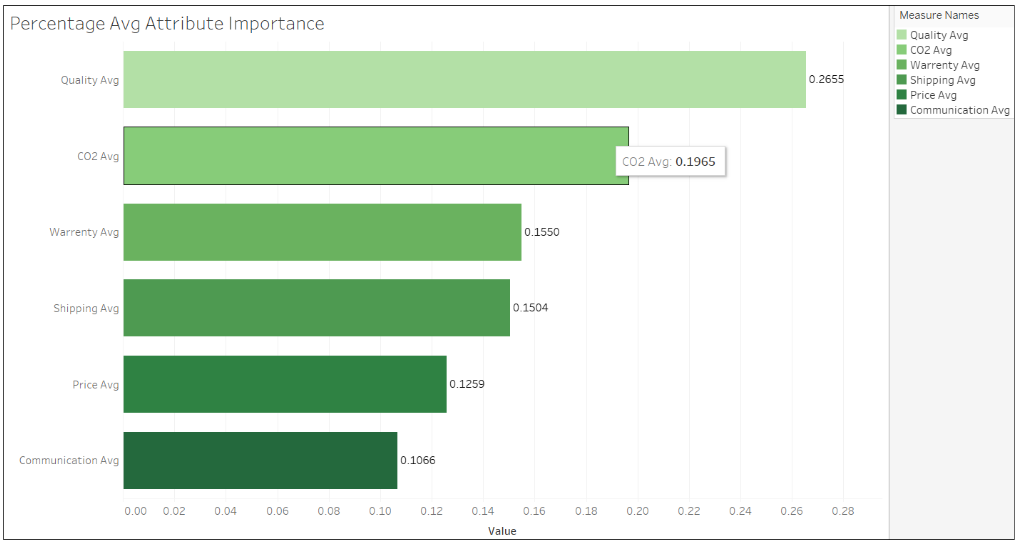


Fig.3. The chart of percentage average importance attributes highlights key factors influencing consumer preferences.

a. Quality Avg (26.55%) : Indicates high priority on product quality.
b. CO2 Avg (19.65%) : Reflects growing consumer demand for environmentally friendly products.

c. Warranty Avg (15.50%) : Longer warranties are preferred for product assurance.

d. Shipping Avg (15.04%) : Faster delivery times are favored for convenience.

e. Price Avg (12.59%) : Price is a consideration but ranks lower in importance.

f. Communication Avg (10.66%) : Effective marketing communication matters, but less pivotal than product features.

4.4.4 Average Willingness for a Attribute:

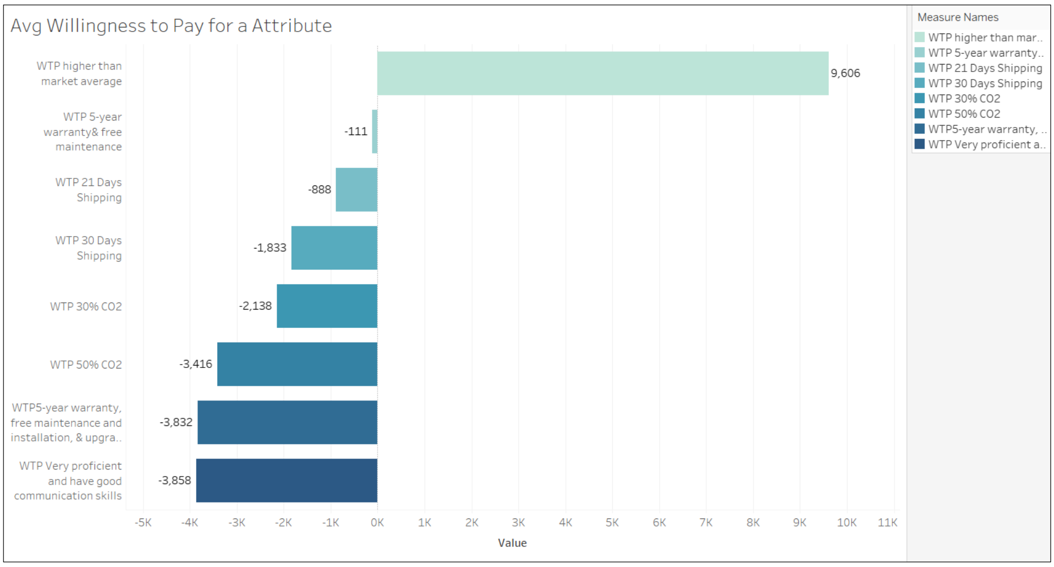


Fig.4. The Average Willingness to Pay (WTP).

Figure 4 values reveal consumer preferences for various product attributes:

a. Consumers show a considerable willingness to pay more for products of higher quality, as indicated by a WTP of 9,606.13 GBP for higher-than-market-average quality.

b. However, there's little inclination to pay extra for a 5-year warranty with free maintenance, with a negative WTP of -111.30 GBP.

c. Faster shipping times are valued, with consumers willing to pay less for longer shipping durations, as shown by negative WTP values of -888.37 GBP for 21 days shipping and -1,832.56 GBP for 30 days shipping.

d. Environmental friendliness, represented by CO2 reduction levels, has a diminishing impact on willingness to pay, with WTP decreasing as CO2 reduction levels increase. For instance, the WTP is -2,138.12 GBP for 30% CO2 reduction and -3,415.53 GBP for 50% CO2 reduction.

e. Comprehensive service packages, such as a 5-year warranty with free maintenance and installation, and upgradeability, don't significantly influence willingness to pay, as indicated by a negative WTP of -3,832.39 GBP.

f. Proficient marketing and communication skills also don't drive consumers to pay more, with a negative WTP of -3,858.05 GBP for products with very proficient communication.

4.4.5 Product Profile Average Utility:

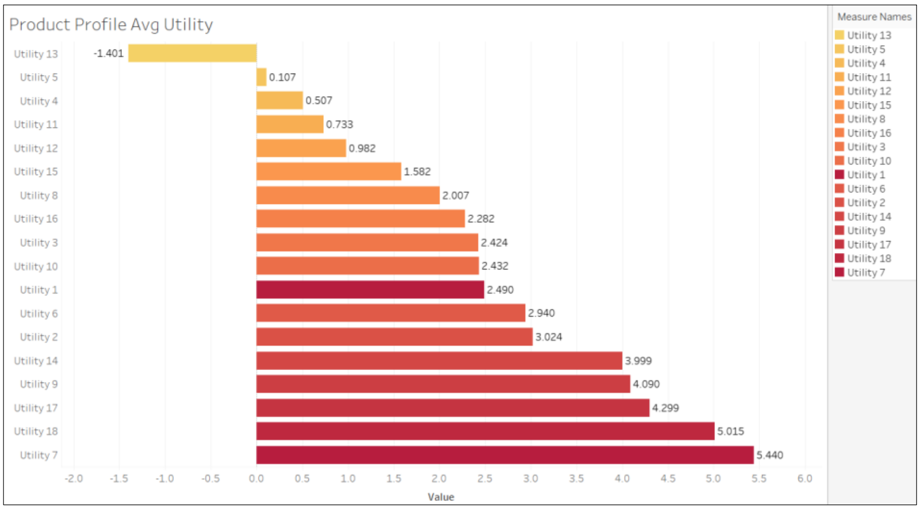


Fig.5. Product Profile Average Utility.

The average utility scores provided in the chart are calculated using calculated fields that incorporate attribute part worths and product profile data. These scores represent the aggregated preference for each product profile, considering the importance of different attributes as reflected in the part worths. They serve as a quantitative measure of consumer preference, enabling businesses to assess the relative attractiveness of different product offerings and optimize their marketing strategies accordingly.

4.5 Explain how conjoint analysis can be used to segment the market?

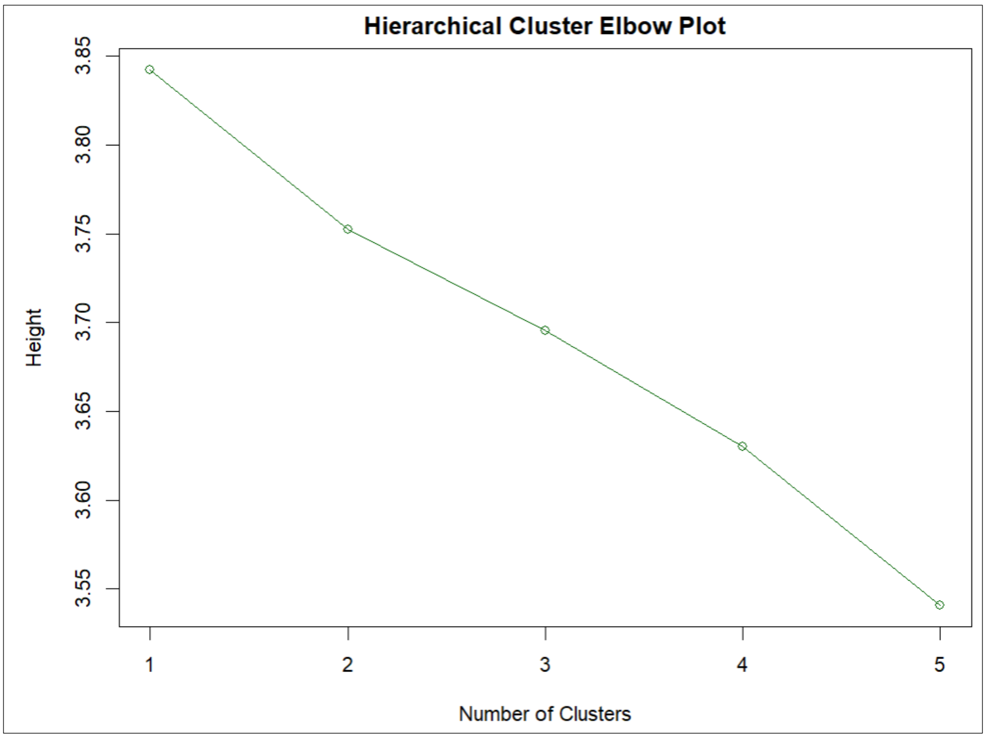


Fig. 6. Hierarchy Cluster Elbow Plot.

Conjoint analysis, coupled with clustering methods like k-means clustering, is a potent approach for market segmentation. By pinpointing crucial attributes and employing clustering techniques, businesses can group customers with similar preferences. When a kink appears after the 3rd point in the elbow plot, adding a third cluster captures subtle differences in preferences. These segments allow tailored marketing strategies and optimized product development, ensuring products resonate with specific customer preferences and needs.

4.6 Determine the most important attributes that customers consider when buying this product?

To determine the most important attributes that customers consider when buying the product based on the PCA results, we can examine the factor loadings of the principal components. In PCA, attributes with higher absolute values of factor loadings contribute more to the variation explained by the principal components. Therefore, attributes with higher absolute factor loadings on Factor 1 and Factor 2 can be considered more important.

4.6.1 From the PCA results:

- Factor 1 (my_factor1) indicates the primary dimension of variability among the attributes.
- Factor 2 (my_factor2) indicates the secondary dimension of variability among the attributes.

Attributes with higher absolute values of factor loadings on these components are more influential in shaping customer perceptions.

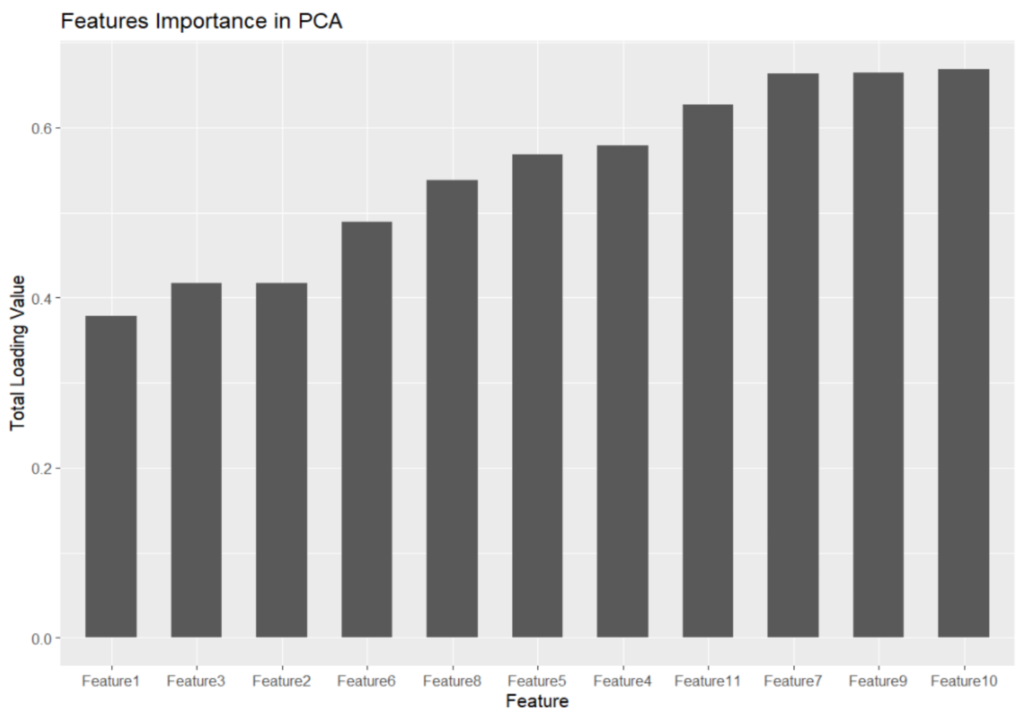
	Attribute	my_factor1	my_factor2	my_path
1:	Feature1	-0.9319502	-0.02625094	1
2:	Feature2	0.9612188	-0.07121589	1
3:	Feature3	0.9464866	0.08030095	1
4:	Feature4	0.8484710	-0.40502680	1
5:	Feature5	-0.7561693	-0.44720905	1
6:	Feature6	0.8897212	0.23286996	1
7:	Feature7	-0.5153093	0.75438614	1
8:	Feature8	-0.7879428	0.37712727	1
9:	Feature9	-0.6039632	-0.69910300	1
10:	Feature10	-0.5319156	-0.75271549	1
11:	Feature11	0.5501711	-0.67330434	1
12:	Feature1	0.0000000	0.00000000	0
13:	Feature2	0.0000000	0.00000000	0
14:	Feature3	0.0000000	0.00000000	0
15:	Feature4	0.0000000	0.00000000	0
16:	Feature5	0.0000000	0.00000000	0
17:	Feature6	0.0000000	0.00000000	0
18:	Feature7	0.0000000	0.00000000	0
19:	Feature8	0.0000000	0.00000000	0
20:	Feature9	0.0000000	0.00000000	0
21:	Feature10	0.0000000	0.00000000	0
22:	Feature11	0.0000000	0.00000000	0
	Attribute	my_factor1	my_factor2	my_path

4.6.2 Based on the factor loadings provided:

Key findings:

- Feature 2 is highly influential, with the highest factor loadings on both Factor 1 and Factor 2.
- Feature 3 also plays a significant role, showing high factor loadings on both factors.
- Feature 6 and Feature 4 are important attributes, with notable loadings on both factors.
- Feature 7, Feature 8, and Feature 11 have moderate to high loadings on both factors, indicating their importance.
- Feature 1, Feature 5, Feature 9, and Feature 10 contribute less to the overall variability captured by the principal components.

4.6.3 Features importance in PCA:



This chart helps to identify which features contribute most to the variance in the data and are therefore the most significant in the PCA model.

4.7 Compute singular values, loading factors and PVEs. Which attributes are the most important?

4.7.1 Singular Values:

```
> #Singular Values
> my_pca$sdev
 [1] 2.5706809 1.6280258 0.7919579 0.5192277 0.4727061 0.4599958 0.3677798 0.3505730 0.2775728
[10] 0.2281128 0.1484736
```

- The singular values represent the importance of each principal component in explaining the variance of the original data.
- The first two singular values are approximately 2.5707 and 1.6280, respectively, indicating that the first two principal components explain a significant portion of the variance in the data.

4.7.2 Loading Factors (Importance):

```
> #Loading Factor
> my_pca$rotation
                 PC1         PC2         PC3          PC4         PC5         PC6          PC7
Feature1  -0.3625305 -0.01612440  0.22574419 -0.022540255 -0.10284468  0.10879743 -0.367723810
Feature2   0.3739160 -0.04374371  0.17531118 -0.002591838 -0.05848381 -0.16855369 -0.057277736
Feature3   0.3681852  0.04932413  0.06148414  0.256607885 -0.39399530  0.33616451 -0.214303077
Feature4   0.3300569 -0.24878402 -0.14001476 -0.067676157 -0.54004744 -0.07143563  0.001495989
Feature5  -0.2941514 -0.27469408 -0.16118879  0.854828743 -0.07732727 -0.24449705 -0.021119857
Feature6   0.3461033  0.14303825 -0.34181851  0.245899314  0.07502912  0.46493964  0.020668302
Feature7  -0.2004563  0.46337482 -0.40316904  0.068076532  0.16466591  0.33048032 -0.050010522
Feature8  -0.3065113  0.23164699 -0.42881517 -0.214848616 -0.59953955 -0.19401702  0.265780836
Feature9  -0.2349429 -0.42941765  0.20576657 -0.030462908 -0.08978128  0.57081745  0.587305101
Feature10 -0.2069162 -0.46234863 -0.28977993 -0.264690521 -0.04832960  0.24356284 -0.605097617
Feature11  0.2140177 -0.41357106 -0.52854459 -0.126789179  0.36131875 -0.18352168  0.174603192
                   PC8          PC9        PC10         PC11
Feature1   0.754091423  0.235701617 -0.13928524  0.124895628
Feature2   0.230824925  0.054035270  0.84641949  0.140695441
Feature3  -0.001142134  0.198427848 -0.04937979 -0.660606481
Feature4   0.222358441 -0.575830072 -0.24782351  0.256492062
Feature5  -0.032193501 -0.046901228  0.10149369  0.039530246
Feature6   0.008571929  0.359498251 -0.09439426  0.567448697
Feature7   0.231840021 -0.528377185  0.27067295 -0.181361780
Feature8  -0.025935128  0.358582624  0.15903909 -0.008414634
Feature9   0.059746952 -0.047403982  0.17778541 -0.029823537
Feature10 -0.336150240 -0.001735039  0.21382515  0.053507085
Feature11  0.395629107  0.170640677 -0.07225950 -0.319594676
```

- The loading factors indicate the correlation between the original variables and the principal components.
- Based on the absolute sum of loadings for the first two principal components, the most important attributes are determined.

4.7.3 Proportion of Variance Explained (PVE):

```
> my_cumulative_PVE
 [1] 0.6007637 0.8417153 0.8987332 0.9232421 0.9435558 0.9627918 0.9750884 0.9862612 0.9932655
[10] 0.9979960 1.0000000
```

The Proportion of Variance Explained (PVE) measures the amount of variability captured by each principal component. The cumulative PVE shows the proportion of total variance explained by the first k principal components. Here are the cumulative PVE values:

4.7.4 Most Important Attributes:

```
> my_attribute_importance
          Attribute Importance
Feature10 Feature10  0.6692649
Feature9   Feature9  0.6643605
Feature7   Feature7  0.6638312
Feature11 Feature11  0.6275887
Feature4   Feature4  0.5788409
Feature5   Feature5  0.5688455
Feature8   Feature8  0.5381583
Feature6   Feature6  0.4891416
Feature2   Feature2  0.4176597
Feature3   Feature3  0.4175093
Feature1   Feature1  0.3786549
```

Based on the absolute sum of loadings for the first two principal components, the most important attributes are Feature10, Feature9, and Feature7. These attributes contribute the most to the variability captured by the principal components.

4.8 Screen Plot:

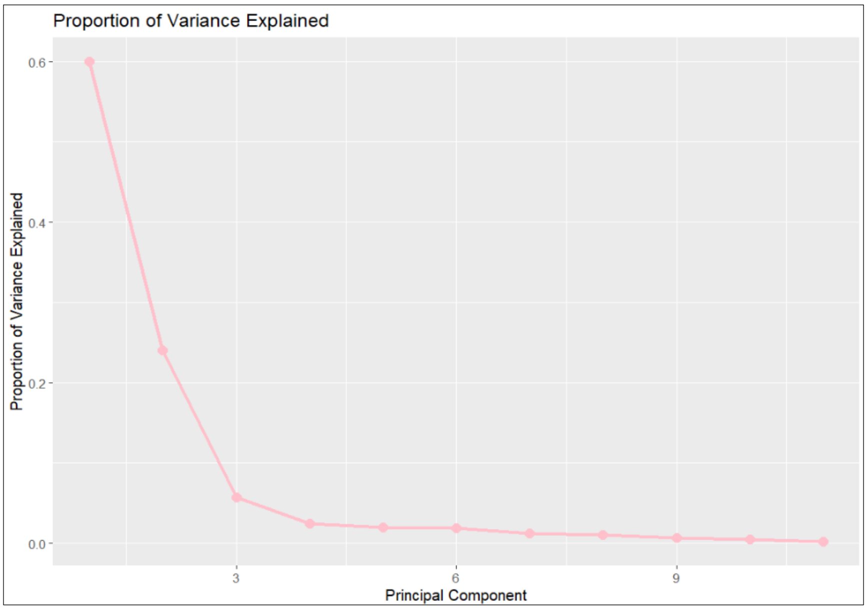
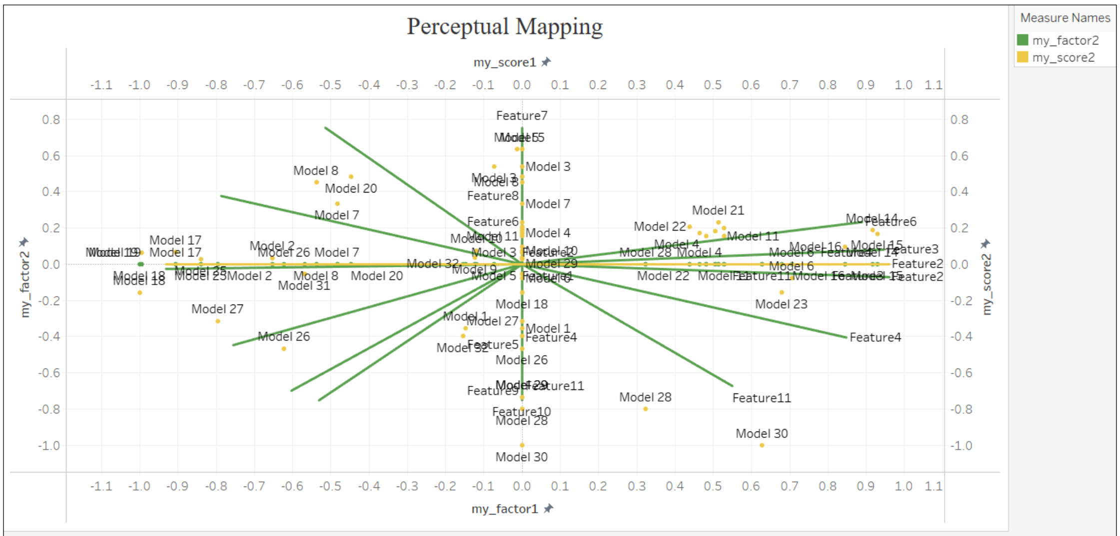


Fig. 7. Proportion of Variance Explained.

Figure 7 visually represents the proportion of variance explained by each principal component in the dataset. Key insights from the plot include:

- ***Diminishing returns:*** The plot shows diminishing returns as we move from the first principal component to subsequent ones, indicating that additional components capture less variance.
- ***Elbow point:*** An elbow or inflection point in the plot suggests the optimal number of principal components to retain. In this case, it's observed around the third or fourth component, indicating where the rate of decrease in explained variance slows down.
- ***Cumulative variance explained:*** The plot illustrates how the cumulative proportion of variance explained increases with each additional principal component. This helps in understanding how much of the total variability in the data is captured by including multiple components.
- ***Decision-making:*** The scree plot assists in deciding how many principal components to retain for further analysis. It guides in balancing the trade-off between capturing sufficient variability and avoiding unnecessary complexity.

4.9 Perceptual Map:



Caption

Sum of my_factor1 and sum of my_score1 vs. my_factor2 and my_score2. Color shows details about my_factor2 and my_score2. The marks are labeled by Attribute and Model. Details are shown for my_path.

Fig. 8 . Perceptual Map.

Values located closer together on the map are perceived to be more similar by consumers, while values farther apart are perceived as more distinct or dissimilar (eg: Model17, Model18, Model2, Model20,etc).

Values clustered together in specific regions of the map may suggest common characteristics or market segments. For example, those clustered in the upper-right quadrant may be perceived as high-quality and high-priced, while those in the lower-left quadrant may be perceived as low-quality and low-priced.

5. Recommendation:

5.1 Market Segmentation:

- Use hierarchical clustering to segment the market based on part-worth utilities, considering the optimal number of clusters determined from the elbow plot analysis.
- Transitioning from 2 to 3 clusters captures finer nuances in preferences, enabling more detailed segmentation if needed (Tynan and Drayton, 1987).
- Avoid over-segmentation beyond 3 clusters to prevent diminishing returns and complexity.

5.2 Product Development and Marketing:

- Prioritize attributes like quality, price, delivery time, and environmental friendliness based on their impact on consumer preferences.
- Incorporate insights from conjoint analysis, such as average attribute part worths and willingness to pay, into product development and marketing strategies.
- Enhance marketing proficiency to better communicate product benefits and drive consumer engagement (Ernst, Hoyer and Rübsaamen, 2010).

5.3 Visualization and Interpretation:

- Use perceptual mapping to visualize product profiles and consumer perceptions, identifying clusters of similar profiles and guiding targeted marketing efforts (Singhal, 2023).

5.4 Continued Analysis:

- Regularly monitor market dynamics and consumer preferences, adapting segmentation strategies and product offerings accordingly.
- Utilize techniques like PCA and conjoint analysis for ongoing insights to optimize marketing strategies and product development efforts (Chen *et al.*, 2023).

Implementing these recommendations can help businesses effectively segment the market, develop products aligned with consumer preferences, and optimize marketing strategies for improved competitiveness.

6. References:

Chen, J., Qi, F., Li, G., Deng, Q., Zhang, C., Li, X. and Zhang, Y. (2023) 'Identification of the Hub Genes Involved in Stem Cell Treatment for Intervertebral Disc Degeneration: A Conjoint Analysis of Single-Cell and Machine Learning', *Stem Cells International,* 2023, pp. 7055264.

Ernst, H., Hoyer, W. D. and Rübsaamen, C. (2010) 'Sales, Marketing, and Research-and-Development Cooperation across New Product Development Stages: Implications for Success', *Journal of Marketing,* 74(5), pp. 80-92.

Li, Y., Qi, J., Chu, X. and Mu, W. (2022) 'Customer Segmentation Using K-Means Clustering and the Hybrid Particle Swarm Optimization Algorithm', *The Computer Journal,* 66(4), pp. 941-962.

Singhal, V. (2023) 'Segmenting and Targeting Fashion Consumers Using Social Media: A Study of Consumer Behaviour', *Journal of Creative Communications*, pp. 09732586231203848.

Tynan, A. C. and Drayton, J. (1987) 'Market segmentation', *Journal of Marketing Management,* 2(3), pp. 301-335.