Data Analysis & Findings

Clustering Performance

To explore segmentation based on scalp severity scores, **K-Means clustering** was applied with **4 clusters**. The quality of clustering was evaluated using the **Silhouette Score**:

- **Silhouette Score**: **0.1179**

   Interpretation:
    The low silhouette score (close to 0) indicates poor separation between clusters. This suggests that the defined clusters exhibit considerable overlap and lack clear, distinguishable boundaries. It implies potential limitations in the clustering performance for this particular segmentation.

---

2 Demographic Associations via Chi-Square Test

A Chi-Square test was conducted to examine the association between **NCCS (socioeconomic classification)** and **Age Group**.

- **Chi-Square Statistic**: 2.1

- **Degrees of Freedom**: 4

- **P-value**: 0.7174

   Conclusion:
    ❌ No statistically significant association was found between **NCCS and Age Group**. The high p-value indicates that any apparent relationship may be due to random variation rather than a meaningful pattern.

---

3 Descriptive Statistics: Consumer Responses

Purchase Intent & Packaging Impact

Question	N	Mean	Median	Std. Dev.	Min	25%	50%	75%	Max

Q1. Likelihood to buy razor in future	30	2.85	3	1.24	1	2	3	4	5
Q2. Influence of pack design on buying	0	-	-	-	-	-	-	-	-

Interpretation: Q1 – Likelihood to Buy Razor in Future

- **Central Tendency:**

  - The mean is 2.85 and the median is 3, indicating a neutral to slightly positive inclination overall.

  - Responses are fairly centered around the midpoint, with a slight lean toward "less likely".

- **Distribution Spread:**

  - A standard deviation of 1.24 suggests moderate variability, meaning that respondents vary quite a bit in their future buying intent.

  - The range spans the full scale from 1 (Very Unlikely) to 5 (Very Likely).

- **Quartile Summary:**

  - 25% of respondents gave scores of 2 or lower.

  - 50% of respondents gave scores of 3 or lower.

  - 75% of respondents gave scores of 4 or lower.

  - Only a quarter of respondents gave the highest likelihood rating (5).

- **Insight:**

  - Overall likelihood to buy is moderate. There is potential to increase purchase intent through interventions such as improved product features, pricing strategies, promotional offers, or better brand communication.

Interpretation: Q2 – Influence of Pack Design on Buying

- No data has been collected (N = 0), which means statistical analysis is not possible for this question.

- This could be due to:

  - Survey logic error or routing issue.

  - Data loss during collection or extraction.

  - A question that was unintentionally skipped or not shown to respondents.

Recommendation

- Review the survey programming and data pipeline to identify why this question has zero responses.

- If this variable is important for decision-making (e.g., pack design influence on purchase), consider re-fielding the question or conducting a follow-up study.

Packaging Perceptions

Question	N	Mean	Median	Std. Dev.	Min	25 %	50 %	75 %	Max
Q7. Pack Uniqueness	300	3.32	3	1.03	1	3	3	4	5
Q8. Pack Believability	300	3.64	4	1.03	1	3	4	4	5
Q9. Pack reflects personal grooming style	0	-	-	-	-	-	-	-	-

Attention-Grabbing & Price Perception

Question	N	Mean	Median	Std. Dev.	Min	25%	50%	75%	Max
Q15. Likelihood of picking product A (1–10)	300	6.48	6	1.53	4	5	6	8	9
Q21. Likelihood of pack catching attention	0	-	-	-	-	-	-	-	-
Q22. Perceived price of razor based on appearance	0	-	-	-	-	-	-	-	-

Interpretation:

- **Central Tendency:**

    - The mean (3.32) and median (3) show that respondents generally perceive the pack as moderately unique.

    - A large portion of responses cluster around 3 and 4, suggesting average to above-average perceived uniqueness.

- **Spread:**

    - With a standard deviation of 1.03 and a range from 1 to 5, the perception of uniqueness varies but stays within a moderate spread.

- **Quartiles:**

    - 25% of respondents gave a score of 3 or less, and 75% gave 4 or less, showing that while a majority rate the pack positively, very high uniqueness (score of 5) is less frequent.

Insight:

- The product's packaging is seen as reasonably unique, but not exceptionally so. Further differentiation in visual design, structure, or branding could enhance uniqueness perceptions.

Interpretation:

- **Central Tendency:**

    - The mean of 3.64 and median of 4 indicate a strong perception of believability in the packaging.

    - Most participants rated it 4, suggesting high credibility and trust in the packaging claims and design.

- **Spread:**

    - With a standard deviation of 1.03, there's moderate variability in opinion, but overall sentiment is positive.

- **Quartiles:**

    - Half the respondents gave a score of 4 or higher, and 75% gave 4 or less, suggesting a strong clustering around high ratings.

Insight:

- The pack is perceived as trustworthy and authentic. This is a strength that can be leveraged in communication strategies emphasizing product efficacy and transparency.

Interpretation:

- No responses were recorded for this question.

- This could be due to a programming error, routing issue, or data omission during export.

Recommendation:

- Review the survey logic and data collection process to identify why this question had zero responses.

- If this is an important metric for evaluating emotional or aspirational alignment with consumers, consider re-administering it or collecting feedback through qualitative methods.

Likeability & Brand Recommendation

Question	N	Mean	Median	Std. Dev.	Min	25%	50%	75%	Max
Q23. Likelihood of using razor if gifted	0	-	-	-	-	-	-	-	-
Q24. Overall Likeability	300	3.62	4	1	1	3	4	4	5
Q36. Likelihood to recommend Gillette (0–10)	300	7.6	8	1.71	4	6	8	9	10

Note: Several questions (Q2, Q9, Q21, Q22, Q23) appear to have missing or unrecorded data (N = 0), which limits insight generation from these items. Data cleaning or follow-up may be required for a complete analysis.

---

4 Key Takeaways

- **Segmentation** using clustering shows weak separation—may not be effective in its current form.

- **No significant demographic dependency** (e.g., NCCS vs Age), suggesting broad-based perceptions.

- **Design elements like believability and uniqueness** are moderately rated (~3.3–3.6), while **overall likeability** and **brand recommendation** scored higher (~3.6–7.6).

- **Purchase intent** is moderate (Mean ~2.85 on a 5-point scale), indicating room for improvement in packaging influence and value perception.

Q1. How likely are you to buy this razor in the future?

Future Purchase Intent – Bar Chart
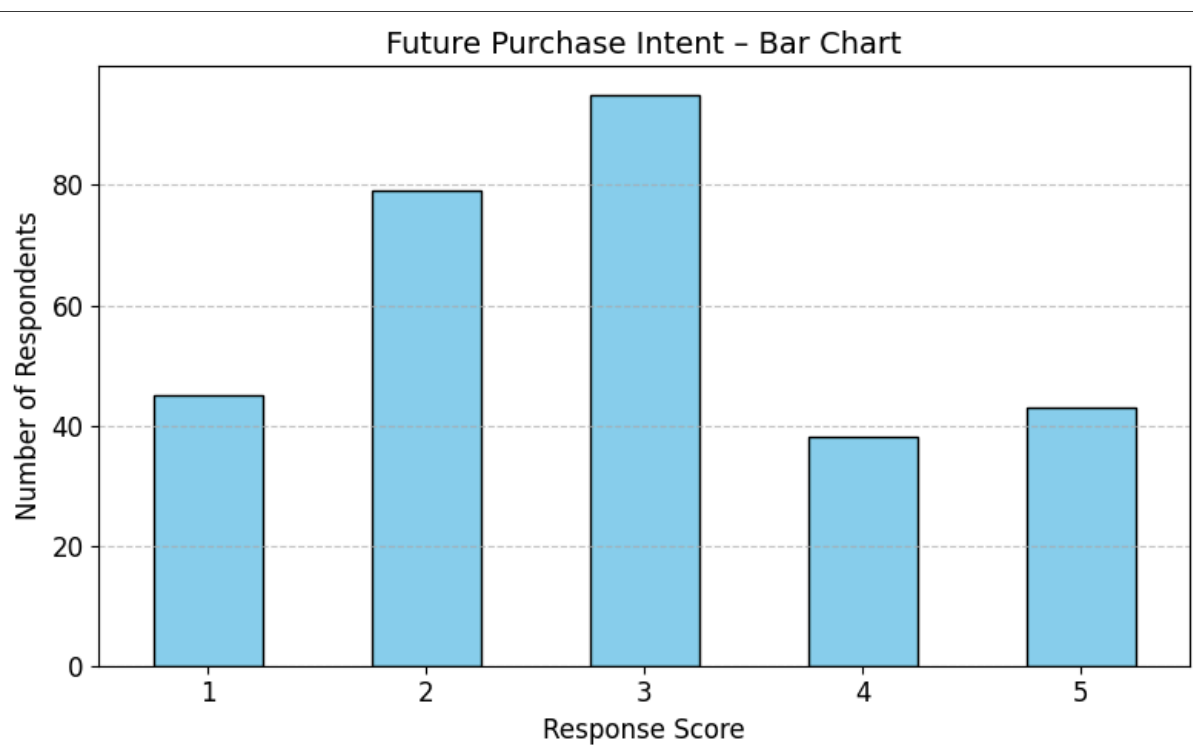
Interpretation:
 The majority of respondents chose a neutral stance (Rating 3), indicating moderate purchase intent. Lower scores (1–2) cumulatively outweigh higher ones (4–5), which suggests that while a portion is positively inclined, there's notable hesitancy or lack of enthusiasm toward immediate purchase.

2 Packaging Influence on Purchase

Q2. How much does the overall pack design influence your willingness to buy this razor?

Pack Design Influence – Bar Chart
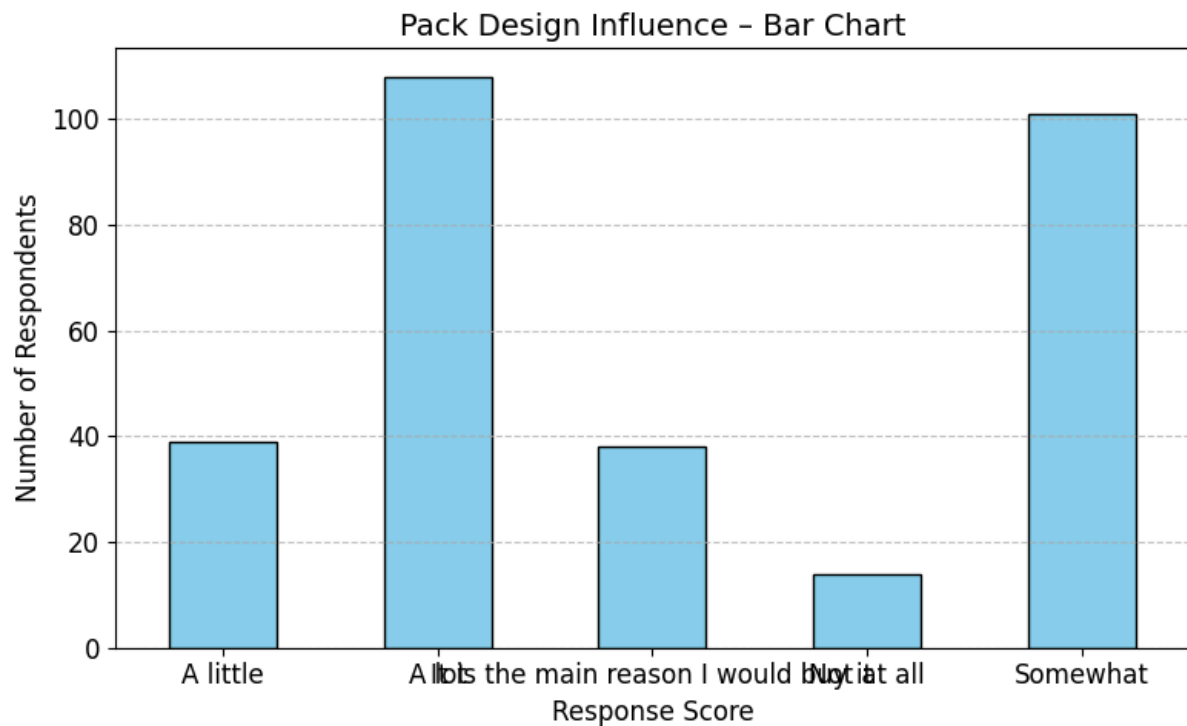
Interpretation:
 Packaging has a **significant influence** on purchase decisions. Nearly **83%** of respondents indicated that the design influences them **to some degree** ("A lot", "Somewhat", or more). However, only **12.7%** stated it as the *main reason*, signaling that while packaging is impactful, other factors also weigh heavily in the decision-making process.

3 Overall Likeability

Q24. Overall, how much do you like the razor packaging?
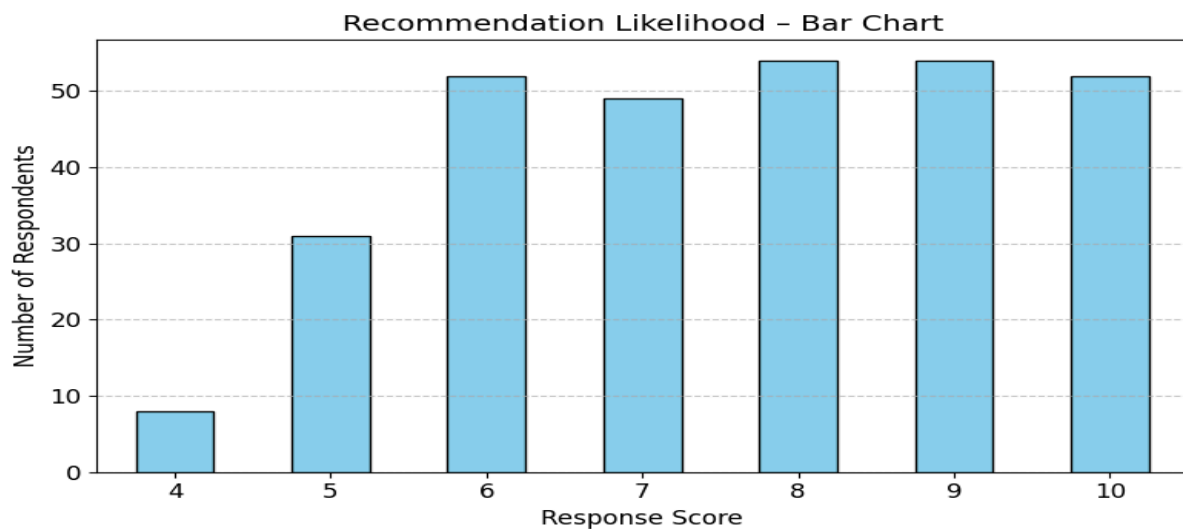

Overall Likeability – Bar Chart
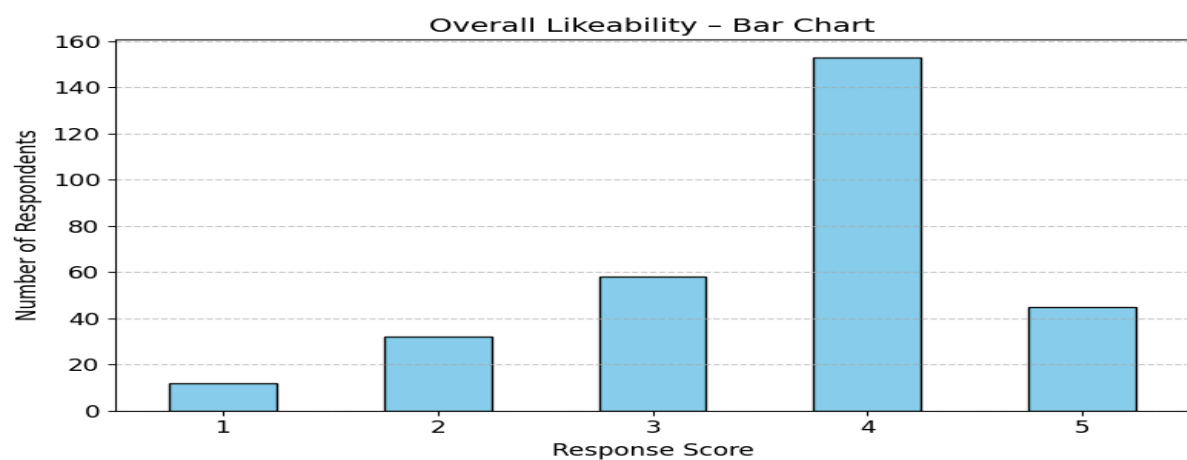
Interpretation:

The sentiment is **positively skewed**. Over **65% of participants** rated the likeability as 4 or 5. This suggests strong resonance with the packaging design, reinforcing positive brand perception.

---

4 Brand Advocacy

Q36. How likely are you to recommend Gillette to a friend? (Scale: 0–10)


Recommendation Likelihood – Bar Chart

Interpretation:

Recommendation scores are relatively high, with **27% giving 9 or 10**, and another **18% giving 8**. This indicates strong brand advocacy potential. However, a considerable proportion gave mid-level ratings (5–7), which may point to areas for improvement to convert neutrals into promoters.

5 Correlation Analysis

A correlation matrix was generated to understand relationships among key consumer response metrics:

Correlation Matrix: Loyalty Drivers
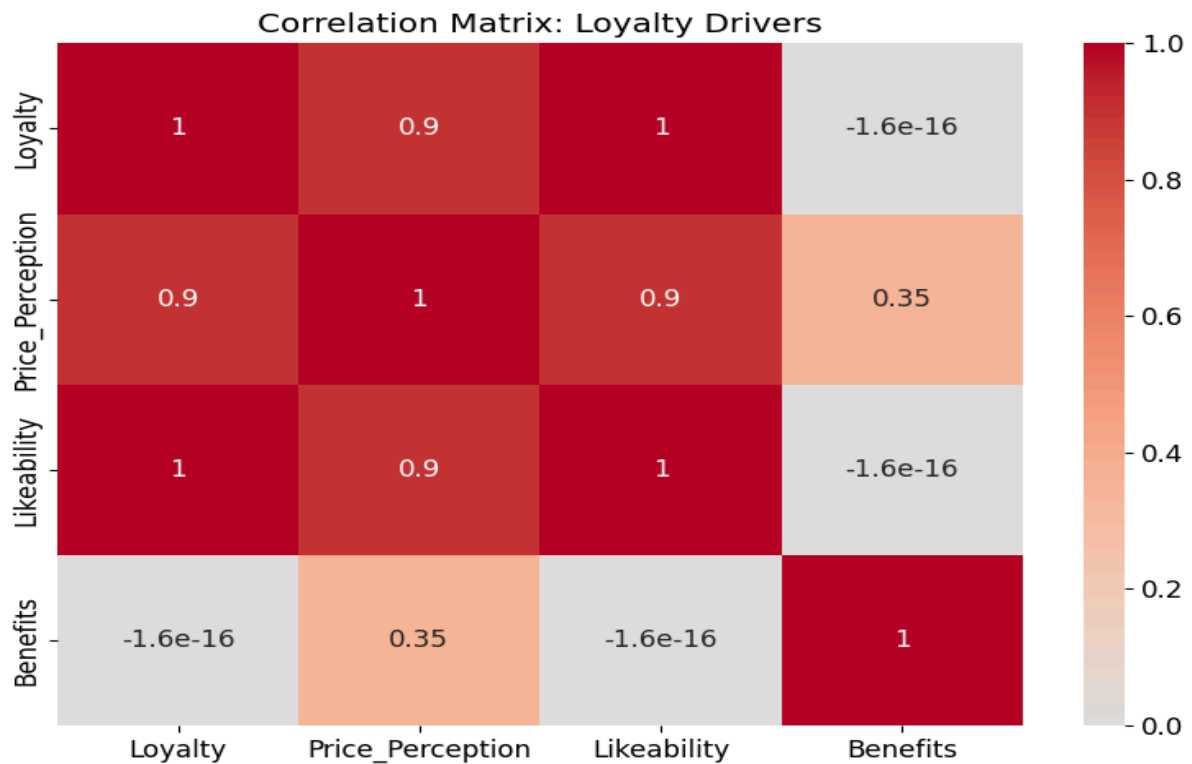
Interpretation:

- **Loyalty and Likeability** are **perfectly correlated** (1.0), suggesting that improving emotional engagement with the product (design, aesthetics, experience) directly boosts loyalty.

- **Price Perception** is also **strongly correlated** (0.9) with both loyalty and likeability, showing that perceived value is closely tied to consumer satisfaction.

- **Benefits** perception, however, is **not strongly correlated** with loyalty or likeability, which may indicate that functional claims are either unclear or undervalued by consumers.

Model Diagnostics: Residual Analysis

To assess the quality and validity of the regression model used in this study, we conducted a residual diagnostics check using key statistical tests. The results are as follows:

Metric	Value	Interpretation
Omnibus	1.467	Tests overall normality of residuals. Low value suggests no strong departure from normality.
Prob(Omnibus)	0.480	P-value > 0.05 indicates residuals are normally distributed; no significant deviation.
Jarque-Bera (JB)	1.217	Another normality test; low JB value supports assumption of normal distribution.
Prob(JB)	0.544	P-value > 0.05 confirms that residuals are normally distributed.
Skew	0.064	Indicates symmetry of residuals. A value close to 0 implies balanced distribution around the mean.
Kurtosis	3.285	Measures the "tailedness" of the distribution. A value close to 3 suggests normal kurtosis.
Durbin-Watson	2.272	Measures autocorrelation in residuals. A value near 2 suggests no autocorrelation.
Condition Number	536	Assesses multicollinearity. Values below 1000 generally indicate stable models without severe multicollinearity.

Conclusion:
The residual diagnostics confirm that the assumptions of linear regression have been reasonably met:

- Residuals are **normally distributed**.

- There is **no significant autocorrelation**.

- **Skew and kurtosis** values are close to normal thresholds.

- The **condition number** is well within acceptable limits, indicating no serious multicollinearity concerns.