Assignment 2: Segmentation

Contents

Part (A) Product affinity-based segmentation.

Considering the primary attributes and demographic details, the below screenshot shows how the product segmentation lined up using the affinity-based segmentation approach.

	A	B	C	D	E	F	G	H	I	J	K
1		P1	P2	P3	overallMean	segmentP1Mean	segmentP2Mean	segmentP3Mean	segmentP1MeanLog	segmentP2MeanLog	segmentP3MeanLog
2	IPr	1937.852272	5494.033521	1626.114207	29.12540193	17.88986478	38.64108391	26.8789784	-0.211664869	0.122777373	-0.034859169
3	Iin	1502.624048	1303.750726	579.6252259	10.88745981	13.87192483	9.169645764	9.580959231	0.105210162	-0.074574006	-0.055517573
4	ICp	1744.016696	1720.859407	725.1238968	13.47266881	16.10041349	12.10328851	11.98599057	0.077383396	-0.046550248	-0.050779702
5	ICl	2020.968895	2095.87272	1318.158386	17.47588424	18.65718082	14.74086267	21.78859923	0.028406861	-0.073916259	0.095790151
6	ICn	2210.445774	1766.834563	1016.719663	16.05787781	20.40639349	12.42664472	16.80594496	0.104078108	-0.111334267	0.019774788
7	IBr	1418.134138	1832.528845	786.337017	12.9807074	13.09193086	12.88869109	12.9978175	0.003705343	-0.003089546	0.000572074
26	income	6525.885729	7084.240275	3493.873996	54.99678457	60.24567246	49.82545555	57.75225579	0.039588558	-0.042886021	0.021231654
27	age	5063.635195	6103.744253	2822.620552	44.98392283	46.74646785	42.92935108	46.65672097	0.016691476	-0.020303002	0.015856888
28	sports	53.09780762	25.38785965	25.51433272	0.334405145	0.490188345	0.178559962	0.421741102	0.166090031	-0.272488866	0.100772978
29	gradschl	40.23444276	36.2495959	27.51596134	0.334405145	0.371436332	0.25495361	0.454827175	0.045611432	-0.117811785	0.133573454
30											

Below is the heat map view of primary attributes and interaction attributes that are developed with the combination of primary attributes.
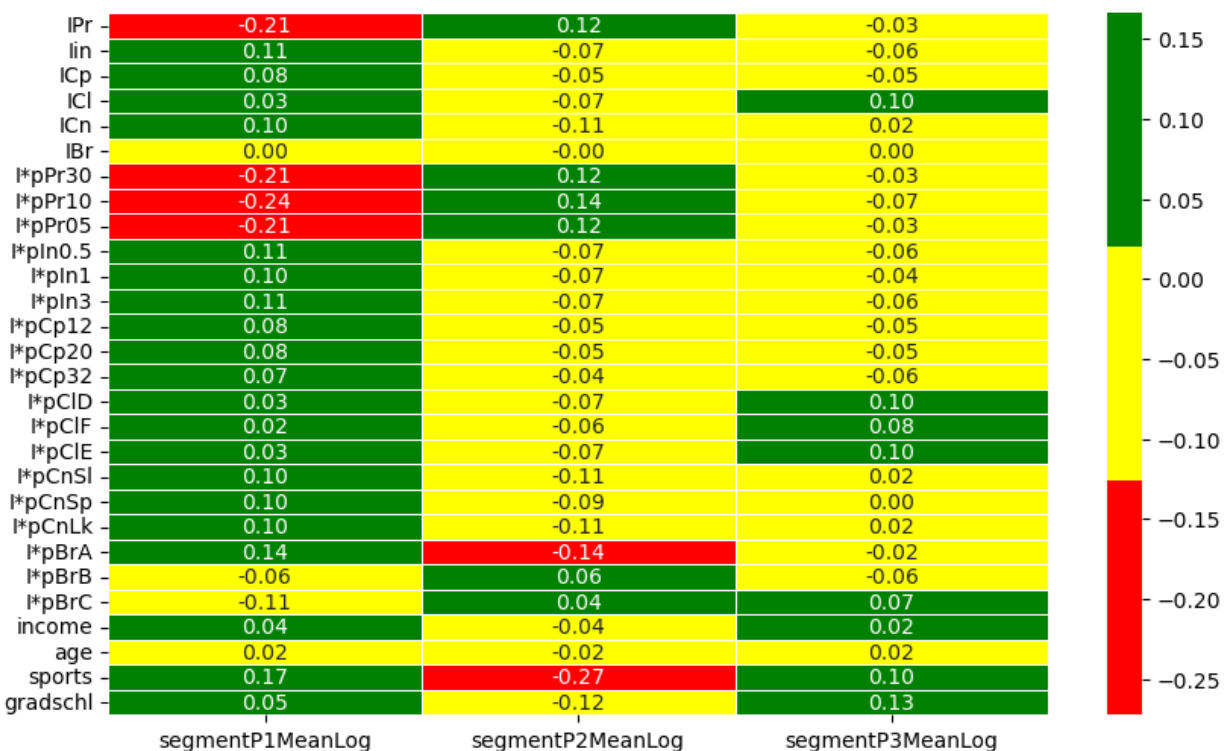
From the above screenshots product P2 segment is highly sensitive towards price of the product, whereas product P1 segment is less sensitive towards the price. Similarly, if we look at the other attributes of each product. We found that product P1 and P2 targets the customer segment which are quite opposite to each other like, insulation is very much required for customer segment who prefer product P1 whereas customers of segment of product P2 are not very much concern about the same.

Attribute	Segment P1	Segment P2	Segment P3
Overall	Prefer products that has attributes like best insulation, high capacity,	Prefer products that has attributes like best ROI, low insulation, medium	Prefer products that has attributes like medium ROI, medium insulation,

Attribute	Segment P1	Segment P2	Segment P3
	high containment, good brand and doesn't care much about the product price.	capacity, least cleanability, least containment, no brand preference and care too much about the product price.	low capacity, medium containment, not very much brand preference but care too much about the product cleanability.
Price (overall mean is 29.1)	~17.9 which is much lower than the overall mean of the product price	~38.6 which is 9 points higher than the overall mean of the product price	~26.9 which very close to the overall mean of the product price compared to product P1 & P2
Insulation (overall mean is 10.9)	~13.9 which is highest among the competitors and better than overall mean	~9.1 which is lower than the overall mean and lowest among the competitors	~9.6 closer to the overall mean and in between the other two products
Capacity (overall mean 13.5)	~16.1 which is highest among the competitors and much better than the overall mean	~12.1 which is closer to overall mean and in between the other two products	~11.9 less than overall mean and least among the competitors
Cleanability (overall mean 17.5)	~18.6 better then overall mean however one of the competitors is providing better option	~14.7 much less then overall mean and least among the competitors	~21.8 best in the category and much higher than overall mean and competitors
Containment (overall mean 16.0	~20.4 best in the category and much higher than the overall mean.	~12.4 very less then the overall mean and least in among the competitors	~16.8 very close to the overall mean and targeting the right customer expectations
Brand (overall mean 12.9)	~13.1 best in the category and very conscious about the product brand.	~12.9 nearly as close as the overall brand and least among the competitors	~13.0 in between the two competitor and catering both types of customers who are brand conscious and not brand conscious.
Persona Description	**PrudentCleanse Seekers**: The identified group places a high emphasis on affordability, seeking products at a lower price point, and prioritizes easy cleaning features. They exhibit a moderate inclination towards well-established brands. This persona is characterized as budget-conscious consumers with a keen focus on cleanliness and a preference for recognizable brands.	**Discerning Enthusiast**: The discerning enthusiasts demonstrate a willingness to invest a moderate premium in exchange for a reputable and well-known brand. Their strong inclination towards brand recognition is evident, as they actively seek a harmonious balance between price and the esteemed reputation of the brand. This persona exhibits a moderate concern for cleanability, while showing lower levels of emphasis on insulation and containment.	**Practical Purity Enthusiast**: Customer's willingness to invest in premium products that stand out for their exceptional cleanability. For them, brand recognition takes a back seat, as their focus lies more on prioritizing product features and performance. Seeking top-notch cleanliness, this persona values quality and functionality over strong brand loyalty.
Persona Story	Meet Raha, the "PrudentCleanse Seekers" savvy consumers in search	Introducing Micheal, the "Discerning Enthusiast," a consumer with a refined	Meet Sam, the "Practical Purity Enthusiast" consumer who values

Attribute	Segment P1	Segment P2	Segment P3
	of affordability and convenience. Raha, with a keen eye on easy cleaning features, she likes to strike a balance by leaning towards well-established brands. In their journey, she likes to navigate the market wisely, prioritizing cleanliness without compromising their budget. Raha like to embody the fusion of practicality and cleanliness in their pursuit of smart and cost-effective choices.	taste for quality. Micheal gracefully invests a moderate premium, valuing the prestige of renowned brands. Seeking the perfect blend of price and reputation, he would like to make choices that reflect their discerning nature. With a measured focus on cleanability and a lesser emphasis on insulation and containment, he curates a lifestyle where every purchase is a testament to their commitment to excellence and brand distinction.	practicality and top-tier cleanliness. His willingness to invest in premium products stems from a deep appreciation for exceptional cleanability. Unlike brand-centric buyers, Sam prioritizes product features and performance, navigating the market with a keen eye for quality. In their pursuit of top-notch cleanliness, Sam likes to stand out for their commitment to functionality and superior product attributes, setting the standard for those who appreciate substance over brand loyalty.

Part (B) Classical segmentation

Classical segmentation method is expected to run through K-means clustering technique. As per the instruction mentioned in question, K-means clustering is executed on the all the 311 customer data points across product's primary and interacted/derived attributes. From the elbow method and to find the optimal K based on inertia change, optimal K is either 6 or 5. However, looking at the inertia change, 5 is the optimal K value.
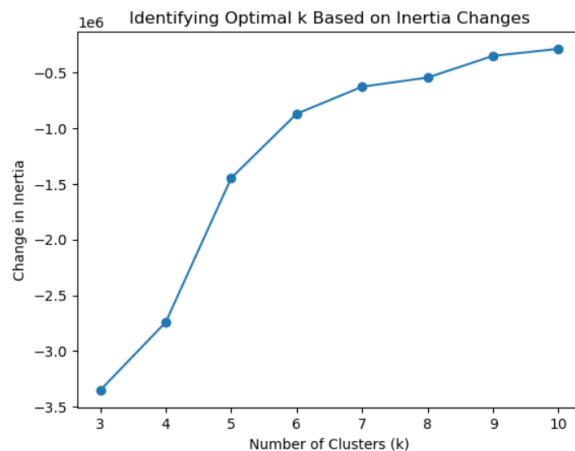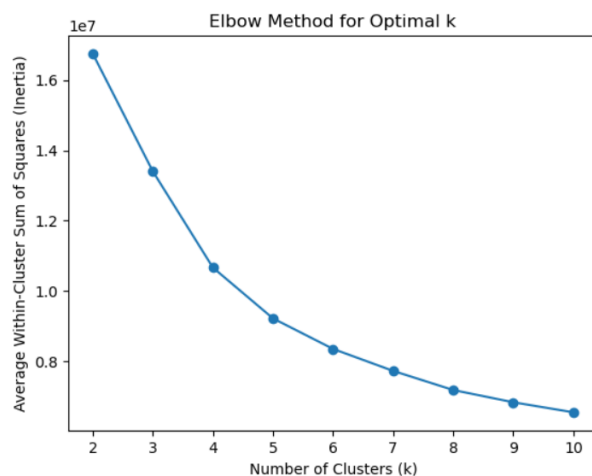
Machine Learning Clustering Approach:

From the above screenshots, the optimal value of K is 5 and primary cluster details are shown as below:

	Cluster--> 0	1	2	3	4
20 IPr	0.230189	-0.29097	0.076025	-0.41112	-0.31974
21 Iin	-0.21135	-0.04621	0.049808	0.242258	-0.10685
22 ICp	-0.24776	-0.202	-0.02439	0.38532	-0.17029
23 ICl	-0.12158	0.368303	-0.08147	-0.19882	-0.20614
24 ICn	-0.2	-0.09952	-0.08304	-0.09364	0.437258
25 IBr	-0.00415	0.011432	-0.0132	-0.00252	0.022364
44 income	-0.04791	0.081604	-0.07659	0.047671	0.051335
45 age	-0.01418	0.062785	-0.04546	0.007915	0.015472
46 sports	-0.40413	0.422482	-0.62694	0.198609	-0.03616
47 gradschl	0.096469	0.393192	-1.40509	-0.16937	-0.11534
48					

Cluster Details explanations

Cluster 1 (column B of above screenshot): ***Price-Conscious Capacity Prioritizers (PCCP)***

This cluster is strong negatively influence by the capacity (ICp), indicating that capacity is not the critical factor. These individuals are likely to prioritize capacity in their purchasing decision. The negative value of insulation and containment suggests that consumer may not be willing to compromise on the price. They are willing to pay more price irrespective of the insulation, capacity, and containment. There is no strong negative on the demographic variables. Infect the cluster 1 has positive impact on the graduation.

Cluster 2 (column C of above screenshot): ***Active Lifestyle Enthusiasts (ALE)***

Cluster 2 is positively influence by the multiple facts like cleanability, brand, income, age, sports and gradschl. The top two attributes that positively influence are sports and gradschl. This shows that most of the customers in this segment are educated and preferred active lifestyle and earning well. Their willingness to pay does not depends upon product's price, insulation, or containment but very much influence by cleanability of product.

Cluster 3 (column D of above screenshot): ***Value-Conscious Pragmatists (VCP)***

The customers in cluster 3 has medium impact on product attributes like containment, cleanability, capacity, insulation and price. But observed that they are very much negatively influence by brand and demographic attributes like income, age, active lifestyle, and gradschl. Their willingness to pay mostly depend upon product attributes and customer's demographics.

Cluster 4 (column E of above screenshot): ***Active Utility Enthusiasts (AUE)***

The customers in the cluster 4 are primarily sports lover and like the good insulation and high capacity. Other attributes like brand, containment is good to have but should not be costly and easy to clean. Other demographic attributes are of medium range category. Customers of this cluster doesn't like to pay extra unless cleanability factor is high.

Cluster 5 (column F of above screenshot): ***Value-Driven Longevity Seekers (VDLS)***

The customers in the cluster 5 are primarily negatively influence by cleanability of the product and highly influence by containment i.e., product lasting and its brand and it is because the customers in this group are high age group and don't want to pay high price for the product. Their willingness to pay is towards good product brand and has good product life.

Persona Stories Explanation

- PCCP Cluster
  Meet Rajesh Kumar, a 32-year-old software engineer from Bangalore, representing Cluster Price-Conscious Capacity Prioritizers (PCCP). Rajesh prioritizes product capacity in his purchasing decisions, valuing functionality over insulation and containment features. A savvy shopper, he seeks affordability without compromising on essential features. The cluster's negative influence on insulation and containment indicates a strong price-conscious mindset; Rajesh is willing to pay more for superior capacity. The positive impact on graduation in Cluster 1 suggests an educated consumer base with informed decision-making. Rajesh is likely to compare prices, participating in sales for products with higher capacity. In essence, individuals like Rajesh in Cluster PCCP are pragmatic, price-conscious consumers, making them a valuable segment for businesses to understand and cater to.

- ALE Cluster
  Meet Mia Johnson, a 28-year-old marketing professional embodying the Active Lifestyle Enthusiasts (ALE) in Cluster 2. Mia, with a graduate school degree and a well-paying job, is passionate about sports and fitness, defining her active lifestyle. In this cluster, positive influences include education, income, age, sports, and grad school, with the top attributes being sports and grad school. Mia's product preferences revolve around easy cleanability, a key factor in her purchasing decisions. Notably, the willingness to pay in ALE isn't driven by price, insulation, or containment but centers on the product's cleanliness. Mia is an informed shopper who values brands aligning with her active lifestyle, making her a distinct market segment. Understanding consumers like Mia provides businesses with insights to tailor products that seamlessly integrate into the dynamic lives of Active Lifestyle Enthusiasts.

- VCP Cluster
  Meet Rohit Sharma, a 35-year-old IT professional representing Cluster Value-Conscious Pragmatists (VCP). Rohit exhibits a pragmatic approach, with moderate influences on product attributes like containment, cleanability, capacity, insulation, and price. Notably, he is negatively influenced by brand and demographic factors such as income, age, active lifestyle, and grad school. Rohit's willingness to pay hinges on a balanced consideration of product attributes and personal demographics, showcasing the discerning nature of VCP. Practicality and functionality take precedence in his purchasing decisions, and he tends to favor affordability over brand recognition. Engaging in informed product comparisons and seeking value for money, Rohit and his counterparts in VCP present a market segment with preferences that businesses can cater to by providing practical solutions with a focus on functionality and reasonable pricing.

- AUE Cluster
  Meet Arjun Singh, a 30-year-old fitness trainer embodying Cluster Active Utility Enthusiasts (AUE). Arjun's lifestyle revolves around sports, making insulation and high capacity key factors in his product choices. Within Cluster 4, other attributes like brand and containment are desirable but should not be costly, and cleanability is crucial. Arjun falls into the medium range for demographic attributes, showcasing a balanced consideration of factors like age and income. His willingness to pay is tied to the product's cleanability, emphasizing the importance of easy maintenance in his purchasing decisions. Arjun's shopping habits involve exploring brands that cater to active individuals, reflecting a market segment where businesses can thrive by offering products that prioritize utility, affordability, and easy maintenance for sports enthusiasts like Arjun.

- VDLS Cluster

Meet Mrs. Gupta, a 60-year-old retiree representing Cluster Value-Driven Longevity Seekers (VDLS). This consumer segment, characterized by a high age group, prioritizes containment, brand reputation, and product longevity. Mrs. Gupta, like others in VDLS, values durability and reliability over easy maintenance, as indicated by the negative influence of cleanability. Despite being cost-conscious, Mrs. Gupta is willing to invest in a good product brand that ensures a lasting product life. As part of her shopping habits, she explores brands known for their reliability and longevity, relying on word-of-mouth and reviews. Mrs. Gupta exemplifies the discerning nature of VDLS, showcasing a consumer segment that emphasizes quality, longevity, and brand reputation in their product choices. Understanding the preferences of VDLS can guide businesses in tailoring products that align with these values.

Marketing Way Clustering

The clustering process has chosen a value of K=4. This decision stems from the recognition that each demographic segment naturally exhibits three levels. Moreover, the ML clusters display some overlap, indicating that opting for either 3 or 4 clusters would be more pragmatic. Targeting numerous clusters with less distinct products could result in increased costs. Hence, to enhance resource optimization and streamline marketing endeavors, the choice of K=4 is considered appropriate for this analysis.

	A	B	C	D	E
1	Cluster -->	0	1	2	3
20	IPr	-0.26149	0.190217	-0.31974	-0.38149
21	Iin	-0.01948	-0.10363	-0.10685	0.235688
22	ICp	-0.20045	-0.15249	-0.17029	0.367706
23	ICl	0.348808	-0.13361	-0.20614	-0.18111
24	ICn	-0.09065	-0.16013	0.437258	-0.07643
25	IBr	0.006087	-0.00537	0.022364	-0.00779
44	income	0.069498	-0.06592	0.051335	0.03998
45	age	0.054094	-0.0296	0.015472	0.003201
46	sports	0.380593	-0.58931	-0.03616	0.174697
47	gradschl	0.33385	-0.12347	-0.11534	-0.22324
48					

Marketing Cluster Details explanations

Cluster 1 (column B of above screenshot): ***Clean Lifestyle Connoisseurs (CLC)***
This cluster is strongly influenced by cleanability expectation and active lifestyle of consumer. Consumers in this cluster are not only prefer active lifestyle but also earning good income and well educated. They don't care much about the capacity of the product. Instead, good product bran with ease of cleanability is more preferred over higher capacity, long containment, high insulation and even high price.

Cluster 2 (column C of above screenshot): ***Budget-Conscious Essentials Seekers (BCES)***
This cluster is strongly positively influenced by product price and negatively influenced by products insulation, containment, and brand. Demographically, consumers are medium level of education, or some are still studying and don't have good source of income and not much active in their lifestyle. Consumers in this cluster prefer to have good value for money product that has medium capacity and cleanability.

Cluster 3 (column D of above screenshot): ***Brand-Conscious Durability Seekers (BCDS)***
The cluster is very conscious about the containment & brand attributes i.e. prefer the products of good brand and last long. Not much interested in cleanability and insulation attribute of the product and little conscious

about product price. Consumer of this cluster preferred a well-known product brand which last long to support their financial condition.

Cluster 4 (column E of above screenshot): ***Insulation & Capacity Enthusiasts (ICE)***
Consumers of this cluster is very much conscious about product's insulation and capacity. They don't prefer to pay more just for the name of brand, cleanability or containment. Demographically, consumers are not highly educated but maintained active lifestyle and earning good enough to manage expenses. Consumers of this cluster are very much focused on buying product that has good insulation and higher capacity.

Persona Stories Explanation
- CLC Cluster
  Meet Aisha Malik, a 34-year-old marketing professional embodying Cluster Clean Lifestyle Connoisseurs (CLC). Aisha's vibrant persona mirrors the distinctive traits of this consumer segment. As a fitness enthusiast and successful marketing professional, Aisha places a high value on cleanliness in her dynamic lifestyle. With a strong emphasis on ease of cleanability, she seeks products that align with her commitment to a hygienic living space. Well-educated and financially stable, Aisha's preferences, typical of CLC, prioritize product brand and cleanliness over attributes like capacity, containment, insulation, or price. Her shopping habits involve exploring reputable brands with a focus on cleanability. Aisha represents a discerning consumer within CLC, reflecting a market segment where businesses can cater to those who prioritize brand quality and cleanliness in their everyday lives.

- BCES Cluster
  Meet Rahul Patel, a 26-year-old graduate student and part-time retail worker embodying Cluster Budget-Conscious Essentials Seekers (BCES). As a student with limited income and a focus on essentials, Rahul prioritizes affordability in his product choices. Influenced by product prices and characterized by medium-level education, he seeks good value for money without compromising essential features. In BCES, insulation, containment, and brand recognition have minimal impact on decisions, and Rahul prefers products with medium capacity. His shopping habits involve exploring budget-friendly options and sales, emphasizing practicality. Rahul's less active lifestyle aligns with the BCES profile, highlighting a market segment where individuals prioritize affordability and practicality in essential product choices. Understanding the preferences of BCES allows businesses to offer budget-friendly options that meet the essential needs of individuals like Rahul.

- BCDS Cluster
  Meet Priya Verma, a 40-year-old finance professional embodying Cluster Brand-Conscious Durability Seekers (BCDS). As part of this segment, Priya prioritizes containment and brand attributes in her product choices. Her focus on durability and brand reputation reflects a preference for long-lasting solutions and well-known brands. While cleanliness and insulation are considered, they don't hold as much weight in Priya's decision-making. Financially capable, Priya is willing to invest in quality products from reputable brands. Her shopping habits involve exploring established brands known for reliability, relying on reviews and brand reputation. Priya's persona represents BCDS, a market segment where consumers prioritize durability and brand recognition, presenting an opportunity for businesses to emphasize the longevity and reliability of their products to appeal to such preferences.

- ICE Cluster
  Meet Arjun Kapoor, a 38-year-old fitness instructor embodying Cluster Insulation & Capacity Enthusiasts (ICE). Arjun's lifestyle, centered around fitness, aligns with ICE's distinct traits, emphasizing insulation and capacity in his product choices. Unlike prioritizing brand recognition, cleanability, or containment, Arjun

values functionality and performance. With a moderate education level and sufficient income, he represents the demographic profile of ICE. His active lifestyle underscores the importance of insulation and higher capacity in his product preferences. Arjun's shopping habits involve seeking products known for their practical features, placing less emphasis on brand recognition. His persona signifies a market segment where individuals prioritize functionality, insulation, and capacity in their purchases, presenting an opportunity for businesses to cater to these preferences by highlighting the practical aspects of their products.

Appendix

Affinity Based Segmentation Files	
Mugs Data	mugsData.csv
Merged Mugs Data from the combination of product attributes	mugsDataMergedDF. csv
Affinity based segmentation mugs data	resultDataFrameAffini resultDataFrameAffini ty.csv ty.xlsx
Classical Segmentation Files	
Product attributes segmentation data with the K=5	logLifts5Upd.csv logLifts5Upd.xlsx
Product attributes segmentation data with the K=4	logLifts4Upd.csv logLifts4Upd.xlsx