Assignment 1

Q1. Use the Boats data and perform market segmentation analysis using appropriate software and method of choice. Use train/test approach to validate the segments.

A. Code for factor analysis and market segmentation could be found in the ipynb file attached.

The initial step involves conducting Exploratory Data Analysis (EDA) to ensure data integrity and consistency across all responses. Of particular importance is verifying that all 2813 respondents have provided answers for each question in the survey. During this process, one outlier observation was identified and subsequently removed from the Boat dataset. This outlier pertained to a customer preference for a boat length of 325 feet, notably exceeding the lengths indicated by other respondents. Its substantial deviation from the dataset's distribution, even surpassing the second-highest preference by over 100 feet, deemed it an anomaly warranting exclusion.

Factor Analysis:

As an initial step, we have decided to conduct factor analysis on the 29 part questions of Q1 for several reasons:

1. **Interpretability**: By simplifying the data and grouping it into a few meaningful sets or factors, we can provide a more interpretable representation of the data. This approach allows us to focus on understanding the underlying constructs or themes reflected in the factors, enabling us to design more actionable marketing plans.
2. **Improved Model Performance**: Clustering algorithms often struggle with high-dimensional data due to the curse of dimensionality. By reducing dimensionality through factor analysis, we aim to improve the performance of clustering algorithms.
3. **Efficient Use of Information**: Factor analysis enables us to extract the most important information from the Q1 parts of the Boats dataset while discarding noise and redundant information. By focusing on the underlying factors capturing the variability in the data, we can gain a more nuanced understanding of the relationships among the questions and the structure of the data.

The analysis commenced with factor analysis applied to the responses corresponding to Question 1 of the survey, aiming to condense the dataset by consolidating the 29 components of this question into a smaller set of meaningful constructs. Initial exploration involved extracting eigenvalues, serving as indicators of the variance explained by each extracted factor. Employing the criterion of retaining factors with eigenvalues exceeding 1, a total of 5 factors were identified as optimal.

Utilizing the principal factor method for factor extraction, latent factors were extracted from the 29 components of Question 1. Subsequently, the dataset was transformed by computing factor scores via the dot product operation with the factor loadings matrix. This transformation facilitated the elucidation of relationships between the original variables and the identified latent factors.

Ultimately, the analysis yielded 5 distinct factors, each representing specific underlying constructs within the dataset. These factors are interpreted as follows:

Factor 0
Considers boating as a significant part of lifestyle, reflecting a commitment to and passion for the activity.
Key drivers :

Q id	Influence Direction (Loading sign)	Description
28	Positive	Boating is my true passion in life
14	Positive	When buying accessories for my boat, I tend to buy the latest and greatest
20	Positive	When not on my boat, I often find myself doing boating-related activities
16	Positive	People tend to come to me for advice about boating
27	Positive	Boating is the number one thing I do with my spare time

Factor 1
Emphasizes the psychological and emotional benefits of boating, including relaxation, connection with nature, adventure, and social bonding.
Key drivers :

Q id	Influence Direction (Loading sign)	Description
21	Positive	Boating helps me escape from everyday life and relax
19	Positive	Boating allows me to experience nature
18	Positive	Boating gives me a feeling of adventure
24	Positive	Boating gives me an outlet to socialize with family and/or friends
17	Negative	I consider myself more knowledgeable about boating than the rest of my boating peers

Factor 2
Highlights a pragmatic approach to boating expenditures, balancing cost, functionality, and quality.
Key drivers :

Q id	Influence Direction (Loading sign)	Description
12	Positive	When it comes to boating, I tend to prefer a basic boat with little to no frills
2	Positive	When buying a boat, getting the lowest price is more important than the boat brand
8	Positive	When buying a boat, functionality is more important than style
5	Negative	I am willing to pay a premium for a brand with a reputation for high quality
11	Positive	I tend to perform minor boat repairs and maintenance on my own

Factor 3
Showcases the importance of knowledge and expertise in boating, valuing informed decisions and practical involvement.
Key drivers :

Q id	Influence Direction (Loading sign)	Description
10	Positive	When buying a boat, I rely more on expert opinion than my own
9	Positive	I see my boat as a status symbol
2	Positive	When buying a boat, getting the lowest price is more important than the boat brand
16	Negative	People tend to come to me for advice about boating
11	Negative	I tend to perform minor boat repairs and maintenance on my own

Factor 4
Focuses on a discerning selection process for boats, emphasizing functionality, technological innovation, and quality above all.
Key drivers :

Q id	Influence Direction (Loading sign)	Description

| 10	Positive	When buying a boat, I rely more on expert opinion than my own
8	Positive	When buying a boat, functionality is more important than style
4	Positive	I only consider buying a boat from a reputable brand
15	Positive	I am serious about the technology on my boat
5	Positive	I am willing to pay a premium for a brand with a reputation for high quality

Market Segmentation:

For market segmentation purposes, our initial step involved feature selection to determine the most valuable and appropriate features or questions to include in our segmentation model. After careful consideration, we have decided to include the following features in our model, as indicated in the table below:

Features / Questions Included	Reason
Q1	Attitudes towards boating are instrumental in segmenting the boating market, influencing product development, marketing strategies, and customer relationship management efforts. By discerning consumers' attitudes, businesses can identify distinct segments characterized by factors like passion for boating, viewing it as a status symbol, prioritizing value, or valuing expert opinions. This segmentation facilitates tailored offerings and communication strategies, aligning with the unique preferences and motivations of each segment. Ultimately, this approach optimizes resource allocation, enhances customer satisfaction, and fosters long-term success in the dynamic boating industry.
Q3	The Purchase Intent Score serves as a critical feature for gauging the likelihood of a segment purchasing a new boat in the future and estimating sales within the respective market.
Q4	This feature helps us understand the target customer profile and positioning of our brand within the market. Similar to the distinction between luxury brands like Chanel and more accessible brands like H&M, this input will guide our strategic decisions as we enter the market.
Q5 Q6	While these questions may primarily pertain to product features, their inclusion in the analysis is justified by their significant impact on revenues and profits. For instance, if a segment willing to invest significantly in boats (Q4) is identified, it's crucial to consider additional product features such as horsepower (Q5) and size (Q6). These features not only influence production costs but also play a pivotal role in determining pricing strategies and profit

	margins. By examining the interplay between customer willingness to spend and product specifications, market segments likely to yield the highest profits can be pinpointed.

Features / Questions Discarded	Reason
Q2	The count reveals that 723 individuals do not own boats, while 1723 individuals own one boat. Given that the majority (87%) are not avid boat collectors, the ownership count feature may not contribute significantly to our analysis. Instead, our focus lies on understanding their attitudes and preferences regarding boats and boating, as these factors are likely to provide more valuable insights for our segmentation and strategic decisions.
Q2 (Cluster)	Lack of clarification of what this feature is.
Q7	The size of the boat (Q6) likely indicates that individuals are not boating alone, as larger boats typically accommodate multiple passengers. These are also attitude questions about boating (correlated with Q1 part 24) which poses the question whether they use boating as an outlet to socialize with family and friends.
Q8	Majority of the respondents falls in the range of Intermediate to Advance (no super beginners and no super experts). Thus, it does not provide much value to include
Q9	NO strong opinions on these set of activities, majority answers in the range of 3, "Sometimes".
Q10	Majority use their boat for a total of 1 month (between 40 and 55 days). So, we will be discarding this feature as there are no unique characteristics for any customers.
Q12	Discarding because majority clusters average around 14-15, with exception of a small group of individuals (261 counts) who make less than $5000 a year. This segment raises concerns about the reliability of the income question, as it seems implausible for individuals with low incomes to demonstrate a decent to high willingness to purchase boats and afford such purchases. Therefore, excluding this segment ensures that our marketing efforts are focused on more viable and profitable market segments.
Q13	Discarding because we feel that there are other factors more valuable to marketing than being married/living with a partner and alone, such as the attitudes to boating (Q1 Part 24), whether they like to boat alone or with family/friends.
Q14	Information extracted from this is overlapping with a lot of the other factors. We are not really interested in whether they are working or not as an employee or self-employed, but rather the money they make and willingness to spend (Q3), and the friends and coworkers they are likely to invite to go boating with (Q1 part 24)
Q15	Discarding because of data quality issues. All of the ages are between 1 and 4.
Q16, Q17, Q18	We will be discarding this question because it focuses on branding data.

Following the feature selection process, the application of the k-means clustering algorithm was initiated for market segmentation. This algorithm begins with the selection of k cluster centroids, followed by the assignment of each data point to the nearest centroid. Subsequently, centroids are recalculated based on the mean of data points within each cluster. This iterative process continues until centroids stabilize, signifying minimal changes or a specified number of iterations is reached. Determination of the optimal value for k was conducted by assessing the within-cluster sum of squares (via the elbow method), evaluating silhouette scores for various k values, and iterative experimentation. Through this process, a value of k=3 was determined to be optimal for our market segmentation.
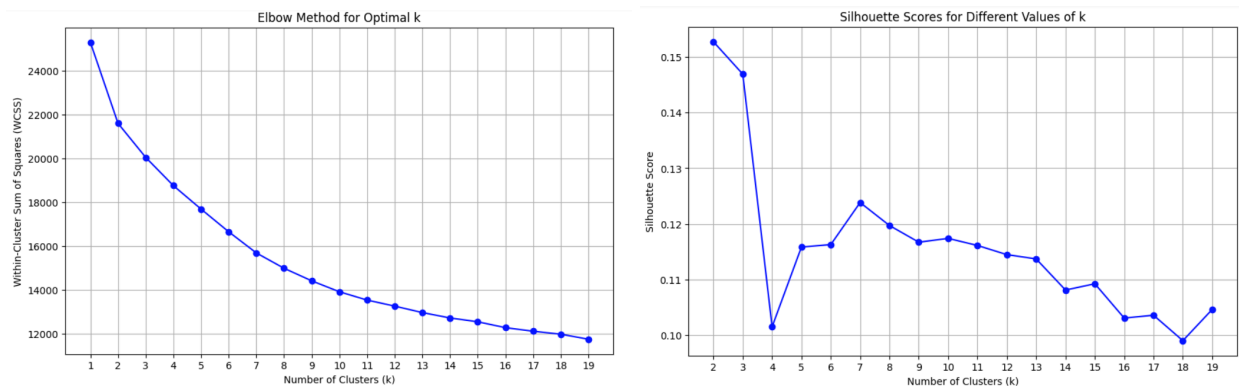
Here are the results obtained from the k-means clustering analysis with k=3:

Distribution of Clusters	
Cluster	Count
1	1416
2	137
3	1259

Mean of Clusters per Feature / Question								
Cluster	Q3	Q4	Q5	Q6	Factor1	Factor2	Factor3	Factor4	Factor5
1	3.98	3.50	2.64	20.14	42.75	4.85	5.86	1.39	3.33
2	4.31	5.70	4.34	51.12	52.03	3.78	5.64	1.61	3.30
3	4.33	4.19	3.82	22.54	54.64	4.06	5.81	1.33	3.24

Q2.**Build profile of segments. What can you do with these segments? Build comprehensive Marketing Plans for each of the segments.**

A.We can use the model developed to break the given data into 3 clusters. The features and qualities of each cluster are mentioned

Clusters

Cluster 1: Moderately likely to buy a new boat, but are least likely compared to the other 2 clusters.Cluster members are likely to spend between $20-35k on the new boat. Members are not looking for a very powerful engine in the boat, anywhere between 15-115 hp would work well for their purposes. Would ideally want the boat to be approximately 20 feet long. The cluster is not very status conscious and the boat is not meant to be a status symbol. The boat is a tool to relax and have a good time. Members are bargain hunters and are looking for good deals, they are most likely to look for discounts, promotions, etc compared to other clusters for the best bang for their buck. The cluster would prefer a no frills boat, that isn't too technical, just to have fun on. Where they buy the boat from is important and they are receptive to expert opinions when it comes to buying boats.

Cluster 2: Quite likely to purchase the boat. Members are willing to spend a significant amount of money on buying the boat, anywhere between $50-100K, the most compared to other clusters. Looking for a powerful engine, between 116 - 170 hp. The cluster is also interested in a boat that is roughly, 51 feet long. Overall the members are quite status conscious and want to buy a boat primarily as a status symbol fitted with impressive tech. Though the primary motivation is status,the cluster members do not want to over pay for the boat. They generally trust experts the most compared to other clusters when it comes to making the purchase decision on the boat.

Cluster3: Quite likely to buy a boat, and have the highest likelihood of purchasing a boat compared to the other clusters. Members want to spend between $40-50K to buy the boat. Generally interested moderately powered engine making anywhere between 61-120 hp. The cluster is looking for a boat that is roughly 22 feet long. This cluster is quite status conscious and wants to get a boat as a status symbol primarily. To a lesser degree than cluster 1, they also want to enjoy the boat as a tool for relaxing. Similarly, to a lesser degree than cluster 1, the members are bargain hunters and are on the lookout for good deals. They are not willing to overpay for the boat. Expert opinions and reputable sellers are important but are not the biggest factors in their decision to buy a boat.

Segment Profiles

Taking the information from each cluster we made segment profiles and highlighted some of their most prominent and defining features.

Segment 1: Value Seekers

The Value Seeker segment is looking to buy a boat because it brings them relaxation and joy. They are already interested in boats, the challenge for them is their relatively low budgets. In keeping with their low spending power however, their requirements for engine power and boat size are also reasonable. Bargains, deals and value offers are the key to getting this segment engaged. The reputation of the boat seller and the opinions of boating experts are also factors that matter a great deal to the Vale Seekers.

Key Insights

- (Low) Probably will buy a new boat in the future
- Budget: $20,000 - $35,000
- Engine Power: 15 - 115 hp
- Boat Size Preference: 20 feet
- Status Consciousness: Low
- Boat Purpose: Relaxation and Fun
- Buying Preferences: Bargain Hunter, Focus on Deals and Value, Importance of Seller Reputation, Receptive to Expert Opinions
- Consists of 1416 respondents of the survey (50.36%)

Segment 2: Clout Chasers

The Clout Chasers are ready to spend some serious money in order to get their hands on a grade A boat. For the kind of money they are spending they expect the boat to not only be large, but come with an impressive engine as well. The Clout Chasers are so named, because their primary motivation for getting a boat is to build a status symbol for themselves. This cluster is either for individuals who have already made it or are on their way to making it big. Social proof is the best way to get to this group, they are willing to rely on the experts in the field to make their buying decisions.

Key Insights

- (Medium) Probably will buy a new boat in the future
- Budget: $50,000 - $100,000
- Engine Power: 116 - 170 hp
- Boat Size Preference: 51 feet
- Status Consciousness: High
- Boat Purpose: Status Symbol, Interesting features

- Buying Preferences: Value-Conscious, Trust in Experts, Desire for Prestige, Avoid Overpaying
- Consists of 137 respondents of the survey (4.87%)

Segment 3: Value-Conscious Status Seekers

The third cluster exists in the middle ground between cluster 1 and 2. The Value-Conscious Status Seekers are willing to spend more money than the Value Seekers but not as much as the Clout Chasers. In response they want a fairly modest boat with regard to engine power and size. Their motivations are also split in the middle with evidence for both wanting a boat as a status symbol as well as wanting it as a personal relaxation tool. The best way to reach them is through enticing offers geared at barging out their inner bargain hunter.

Key Features

- (Medium) Probably will buy a new boat in the future
- Budget: $40,000 - $50,000
- Engine Power: 61 - 120 hp
- Boat Size Preference: 22 feet
- Status Consciousness: Moderate
- Boat Purpose: Status Symbol with Relaxation Option
- Buying Preferences: Bargain Hunter, Moderate Trust in Experts, Focus on Value, Seller Reputation Matters, Avoid Overpaying
- Consists of 1259 respondents of the survey (44.77%)

Segment Marketing Plans

Now we will share a marketing plan for each segment profile.

Segment 1: Value Seekers

Targeting:

All campaign targeting should focus on budget-conscious individuals who prioritise value and affordability. The focus of the campaign should be on customers who view the boat primarily as a tool for relaxation and fun rather than a status symbol.

Marketing Strategies:

Offer promotions and discounts to appeal to bargain hunters within the specified budget range. By emphasising the reliability and functionality of the boats and highlighting the no-frills design and focus on practicality the campaign may make the boat more relatable to the simple users

putting any concerns around any additional operational costs to rest. It is vital to collaborate with reputable sellers and experts to build trust and credibility for this segment. It is imperative to utilise social media platforms and online forums frequented by value-conscious consumers to build credibility for the company as well as position the brand and its partners as true experts in the field.

Messaging:

As mentioned the main thrust of this campaign is on the 'low cost fun' angle. The messaging should be

"Affordable Fun: Your Perfect Boating Companion"

"Relax and Unwind with Our Budget-Friendly Boats"

"Expert-Approved Value: Where Quality Meets Affordability"

Channels:

- Social media advertising (Facebook, Instagram) targeting budget-conscious individuals interested in boating.
- Online marketplaces with emphasis on value deals and customer reviews.
- Collaborations with boating experts for endorsements and recommendations.
- As an additional channel the company should consider producing original content for their social media pages that the customers can engage with.

Segment 2: Clout Chasers

Targeting:

The campaign will target high-income individuals who prioritise status and prestige in their purchases. It should appeal to customers who see the boat as a symbol of luxury and technological innovation.

Marketing Strategies:

The strategy is to showcase high-end boats with advanced technology features and customizable options. The aim is to position the brand as synonymous with prestige and exclusivity. It is important that the segment be offered personalised purchasing experiences, such as VIP tours and exclusive events. The brand must also highlight endorsements from industry experts and influencers to reinforce credibility.

Messaging:

The message being sent is "Owning a boat makes you better than those who do not"
The message needs to be

"Elevate Your Status: Luxury Boats for the Elite"

"Experience Luxury Like Never Before: Your Symbol of Success"

"Innovation Meets Elegance: Your Gateway to Prestige"

Channels:

- Luxury lifestyle magazines and publications targeting affluent demographics.
- Exclusive boat shows and events attended by high-net-worth individuals.
- Partnerships with luxury brands and venues to reach affluent consumers.
- Brand deals with other status markers (e.g. discounts for members of certain clubs)
- Partner up with luxury lifestyle content creators across different social media

Segment 3: Value-Conscious Status Seekers

Targeting:

The campaign will target consumers who seek a balance between value and status in their boat purchases. Aim is to appeal to individuals who prioritise affordability but still desire a certain level of status.

Marketing Strategies:

The plan is to offer mid-range boats with competitive pricing and attractive financing options. The campaign will highlight the value proposition of the boats, emphasising a combination of quality and affordability. Showcase the boats as symbols of status without the premium price tag. The company must leverage customer testimonials and reviews to build trust and credibility.

Messaging:

The messaging needs to revolve around the idea of "Luxury for the regular person".
The message will comprise of

"Affordable Luxury: Your Ticket to Status and Savings"

"Upgrade Your Lifestyle: Premium Boats, Practical Prices"

"Where Value Meets Status: Your Smart Choice for Boating"

Channels:

- Digital marketing campaigns targeting middle-income consumers interested in recreational boating.
- Local boat shows and expos with emphasis on value and affordability.
- Collaborations with local influencers and boating enthusiasts to reach target audience.
- Bundle deals with other avenues of low cost luxury (e.g. certain car dealerships can offer discounts on certain boat models to their customers)

Q3.What data elements are missing that will you add to this data set for performing the segmentation? List the variables (with the actual wording of the questions, the scale of measurement).

Data Elements that can add value to the dataset for performing segmentation include:

1. How often do you currently participate in water activities?

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5

2. How do you feel about the maintenance and upkeep required for owning a boat?

A rewarding challenge ………………………(1)

A necessary task ……………………………….(2)

No big deal ……………………………………….(3)

Not for me …………………………………….(4)

Not Opinion…………………………………………..(5)

3. What level of responsibility do you feel towards environmental conservation while boating?

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5

4. What are your expectations regarding the learning curve for operating a boat?

Challenging …………………………………………………………(1)

Somewhat Challenging……………. …………………….(2)

Enjoyable…………………………………………………..(3)

Neutral ………………………………………………………..(4)

Not Sure………………………………………………………(5)

5. What role does safety play in your decision-making process when considering purchasing a boat?

Very Important ………………………………………………………(1)

Important …………………………………………………………….(2)

Somewhat Important ………………………………………………..(3)

Not Important…………………………………………………………...(4)

Not Sure…………………………………………………………………(5)

Q4. **Would you use a variable weighting scheme? If so, how would you choose the weights?**

We believe that using a variable weighting scheme might be beneficial to improve the segmentation result, because there are some variables which seem to be more important than others.
Based on the previous steps of the segmentation, we decided to use questions 1,3,4,5, and 6 to build the segments.

Question 1 consists of 29 statements regarding customers' behaviors and points of view regarding boating. We have simplified the data by grouping those statements into 5 different factors using factor analysis method.
Question 3 asked about the possibility of purchasing a new boat in the future, and question 4 asked about the willingness to pay for the new boat. Along with question 1, question 3 and 4 refer to the customer behaviors towards boating.
On the other hand, question 5 asked about the horsepower of the boat owned/desired, and question 6 asked about the length of the boat desired. These questions are more about the boat itself rather than the customers.

As we want to do customer segmentation, we would put larger weights on customer behavior related variables rather than on product related variables. Therefore, the weights are as following :

Q1_0 (Question 1 factor 0) = 2
Q1_1 (Question 1 factor 1) = 2
Q1_2 (Question 1 factor 2) = 2
Q1_3 (Question 1 factor 3) = 2
Q1_4 (Question 1 factor 4) = 2
Q3 = 1.5
Q4 = 1.5
Q5 = 1
Q6 = 1