Marketing Analytics

GetFit Running Shoes (Segmentation)

GetFit is an American company that operates in specialty athletics equipment, with particular focus on running shoes. In the last decade or so, America has seen a boom in the long distance running activity as something to do for fun or to get fit. While marathons and half-marathons were once restricted to "serious" competitive runners, they are now seeing more registrations than ever. These additional registrations come from non-competitive runners who use these events as means to motivate themselves to get fit or stay fit. Needless to say, this has led to an explosion in demand for specialty training equipment, especially running shoes.

GetFit has been quite successful at growing its customer base with its small but well-regarded product line. So far, GetFit has been selling its products only in its own branded stores located in most major cities in America. In these stores, customers get a free "diagnostic" of their foot shape, running posture, and so on using a series of comprehensive tests using sophisticated biomechanical devices. A small crew of trained sales staff personally assists customers and recommends the right shoe for them.

GetFit's focus has also been primarily on the male demographic. Now GetFit is looking to expand its offerings for the female demographic. Their existing running shoes for women vary only slightly from their running shoes for men; different shoe sizes, a few more colors, and a few more size and shape variations. The GetFit product design team has come up with prototypes for two new brands of shoes.

GazelleFit is a running shoe built after extensive research on the running posture and biomechanics of professional female long distance runners. It is made from a new material that is light but sturdy. Its heels, insoles, and body last longer than the average GetFit shoe. It is also costly to make, so the pricing people at the company expect to sell it for about $150-250 depending on colors and extra accessories such as insoles and specialty socks. **SleekFit** is a running shoe that is only a minor improvement on GetFit's existing shoes for women. Its design enhancements include variations in color, styles, and additional set of shoelaces, and an ergonomic design that makes them appear more like "general" sneakers. Using this brand, GetFit hopes to appeal to a broader base of female customers who might not be as serious about long distance running, but want to have a shoe that is of a better athletic design than the generic sneaker. This brand will be priced at tentatively $75-125.

GetFit has also been considering widening the distribution of their products which so far has been restricted exclusively to company owned stores. There are some in the company who believe that at least a limited number of brands should also be stocked in retail outlets such as Target, Dick's Sporting Goods, and so on. This idea has been met with resistance from "traditionalists" in

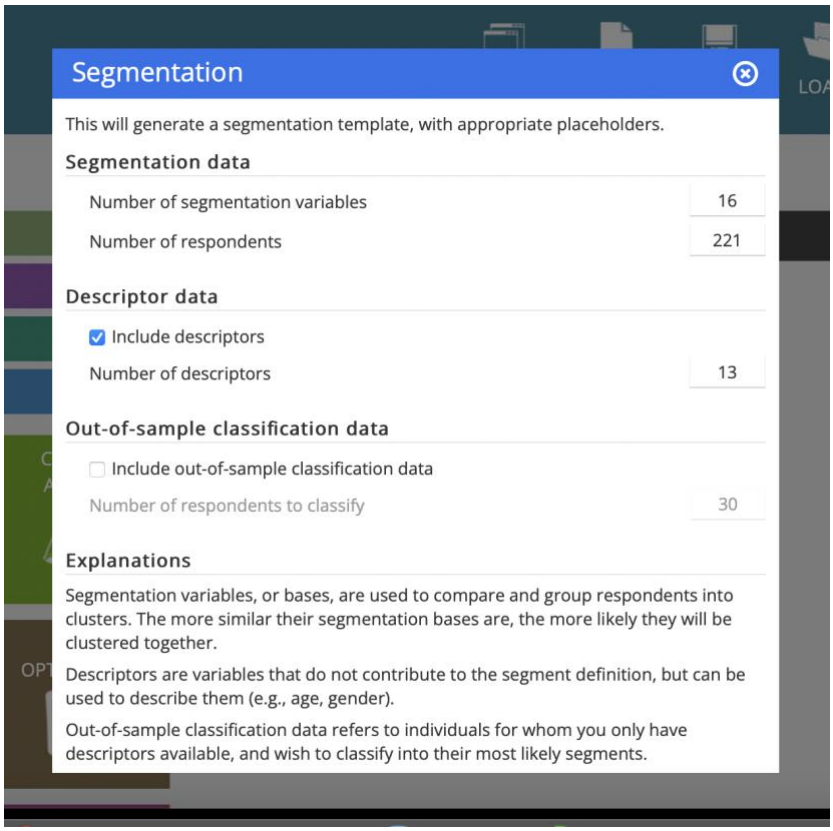
the company who believe that the personalized attention that GetFit offers at its store, especially to amateur runners, is part of their brand identity.

There is debate within the Marketing department at GetFit about the need to understand the various segments and their motivations better. The CEO, believing that market research could provide guidance about how to move forward, commissioned a survey of customers who visited GetFit stores.

Based on the data provided, conduct a segmentation analysis and also discriminant analysis. **IMPORTANT – Make sure you are choosing the Discriminant sheet in the drop down menu, NOT the segmentation sheet again.**
Identify the different segments and describe them, both psychographically and demographically. Explain which (if any) segments should be targeted more than the others. Use your findings to also make recommendations about the two new brands to be introduced, the distribution strategy, and how the new brands should be promoted.

Solution:

For GetFit, it was unable to directly load the file on Enginius. It kept showing the error. So I had to choose the segmentation template and then copy and paste the data from excel file.



After running the segmentation analysis on 3 & 4 segments, I found 4 segments to be more understandable and clearer.

For 3 number of segments, I got:

	Population	Segment 1	Segment 2	Segment 3
Size	221	67	62	92
Relative size	100%	30%	28%	42%

Segment size.

And for 4 number of segments, I got:

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Size	221	67	62	63	29
Relative size	100%	30%	28%	29%	13%

Segment size.

Segment 1 has highest people of 30% and segment 4 has lowest customer size of 13%.

After selecting 4 segments for analysis, a segment descriptor table for analysis is found in the generated report.

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Comfort	3.20	2.33	3.74	3.11	4.28
Price	4.67	4.91	5.39	4.67	2.59
Distance	3.76	3.91	5.34	2.46	2.90
ShpTime	4.08	4.13	5.29	3.11	3.48
CasualOcc	5.72	5.37	5.68	6.05	5.93
Percept	4.27	4.16	4.55	3.97	4.59
Range	4.17	4.58	3.60	3.90	5.00
ProcSat	5.27	5.21	5.18	4.98	6.24
StaffSat	4.67	4.52	4.77	4.00	6.21
Premium	3.66	2.52	3.68	4.16	5.14
Display	3.48	2.87	3.71	3.75	3.86
Advice	4.41	3.78	4.50	4.60	5.28
Weight	5.39	4.55	5.55	5.76	6.21
Sleek	5.86	5.37	5.82	6.11	6.55
Color	5.22	4.12	5.63	5.65	5.97
Match	3.99	2.24	4.50	4.65	5.48

Each segment has different attributes and that says about their preferences and by this, segments can be clustered differently. These are named based on their attributes.

Segment 1 – Irresolute class

This segment is kind of tricky. People in this segment do not want any of the basic shoe things like premium status, color, weight, and sleekness of shoe and more importantly, comfort level of the shoe. These customers also score low on Casual occasion. They score high in Range, which is a term where there are wide range of choices to make, these people find it difficult to make a choice. So, this segment is loaded with indecisive people

Segment 2 – Working professionals

People in this segment look for comfortable shoes with a good price range. Not only those counts, but it is also important to them that shoe shopping should not take much of their time and the shop where they buy shoes from should be in a near distance. They look for good colors in shoes and that color should also match their other wearables. At the end it is important to them that shoes they wear reflect their personality to others. And these population is very good at making judgement on their choices.

Segment 3 – Fitness freaks

This population does not care if they wear a running shoe to casual occasion. They are willing to pay a high price from premium manufacturers. They do not care about the shoe shop distance and how the staff in shop behaves or how much they have knowledge about the shoe. But it is important analysis that overall, they care about weight, sleekness, and color of shoe. They also want their shoe to match their accessories like segment 1 people.

Segment 4 – Trend-setters

People in this segment want the most comfort. They are not price conscious and prefer to buy shoes from nearest shop. Not only they care about weight, color, richness of the shoes but they also want staff to treat and advise them well on their purchase as they get confused while making decisions. They want to have a good overall purchase experience. They are willing to spend good money on shoes every year.

Now after this I analyzed descriptor data,

Descriptor analysis:

	Population	Segment 1	Segment 2	Segment 3	Segment 4
FitFreak	2.67	2.04	2.71	3.16	3.00
FitOthers	3.34	2.78	3.39	3.75	3.66
OthersThnk	3.63	3.06	3.76	3.90	4.10
OtherAdv	3.35	2.70	3.32	3.81	3.90
FitMag	3.93	2.91	3.95	4.67	4.62
OtherPpl	3.72	2.78	4.31	3.98	4.10
MagSource	3.90	3.15	4.53	4.25	3.55
Tvcomm	4.48	3.57	5.16	4.84	4.31
Internet	2.25	1.73	2.52	2.59	2.14
Age	2.45	2.79	2.31	2.33	2.21
Expend	2.03	1.84	1.77	2.25	2.55
Income	2.70	2.43	2.65	3.00	2.79
Education	1.85	1.82	1.98	1.86	1.62

Based on descriptor data here are my findings,

Segment 1:

People in this segment scored high only in age factor. Other than that, they were lowest on almost all criteria. This indicates that this segment contains people with relatively older age. This justifies their indecisiveness in making a choice. At the same time one other factor that conflicts this finding is the comfort factor, which is low, so this cannot be the eldest group.

Recommendations:

Suggestions would be to have a knowledgeable staff that helps and advice to make good decisions on their purchases.

Segment 2:

For shoe purchase this segment cares about views from other people and what shoes are trending in current fitness magazines as well as TV ads. People here do not want to spend much every year on shoe purchases

Recommendations:

These individuals are working class and probably purchase shoes on other's recommendations while it comes to purchase. It would be a good strategy to spend on TV & magazine advertisement as it is the best way to reach them.

Segment 3:

From descriptor analysis it came to light that most people in this segment are female and are fitness freaks that considers themselves fit than other female. They like having shoes information from reading Fitness magazines, TV ads and Internet. People takes advice from them as they are fitness freaks. They also love to spend more on shoes every year.

Recommendations:

To target this audience, it would be wise to focus on a personalize experience for running shoes that are lightweight, have sleek design and they can look appealing so that it could be worn casual occasions too.

Segment 4:

Analysis is that this segment has lowest age group people. They take information for shoe purchase from other fitness magazines, other peoples. It is important to them that other people think they are fit. This can justify their trend-setter attributes. They are willing to spend highest on shoe every year.

Recommendations:

To target this audience, it would be wise to have a celebrity brand ambassador as people here are attracted towards brand and will pay a higher price for premium shoes.

AutoData (Segmentation)

In the past few decades, the variety of automobiles available for customers have splintered along many different features such as luxury, make, model, and fuel type, among others. Previous discriminant variables related to customers such as age, ethnicity, gender, and income are also not as clean cut as they used to be in terms of identifying key segments.

This dataset is a sample from actual customer level data about their choices in car buying AND their demographic characteristics. Be careful loading it and running the analysis, because there are two different sheets. The variable names are self-explanatory.

First run just the segmentation analysis, first hierarchical, then the appropriate k-means, and suchlike. Describe the segments, first based purely on the segmentation variables. Then run the analysis with the discriminants (make sure you choose the right option in the drop down menu). And then present an overall analysis about the segments here, how you would go about naming and describing them, and what demographics stand out.

The dataset here is limited in terms of the number of variables, so please make sure to thoroughly do different scenario analysis for the most sensible segments.

Solution:

For AutoData question the data given in this file consists of two different files one is segment data that has factors such as Customerid, Segments, Make, Model and Fuel type. Another file has discriminant data that has factors such as Customerid, Age, Enhanced Ethnic Groups, Gender and Income.

For AutoData, first I ran hierarchical segmentation analysis without forcing number of segments. As we can see in below screenshot, it automatically selected 3 segments. Also, to point out segment 3 had lower population compared to segment 1.

	Population	Segment 1	Segment 2	Segment 3
Size	2 000	1 549	263	188
Relative size	100%	77%	13%	9%

Segment size.

Now, just to check if segment 4 makes more sense, I forced segmentation analysis with 4 number of segments.

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Size	2 000	1 078	263	471	188
Relative size	100%	54%	13%	24%	9%

Segment size.

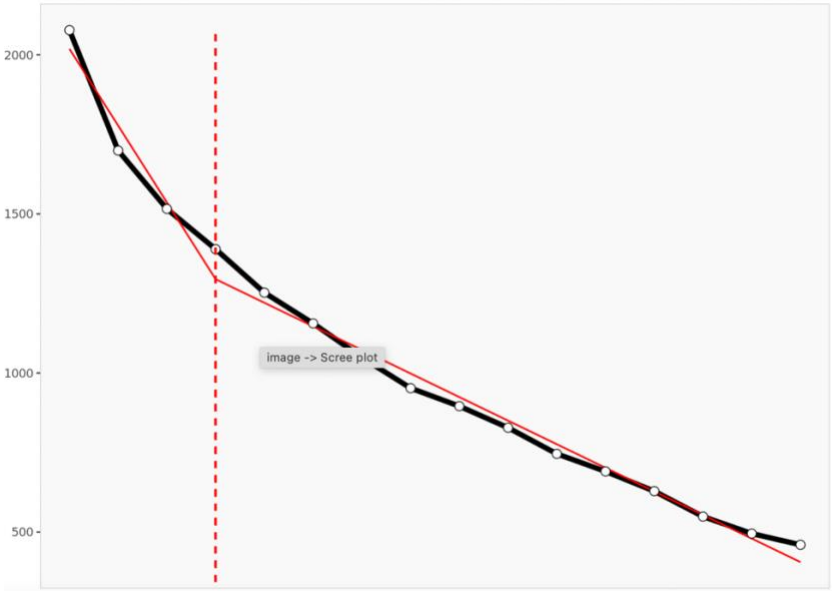
As it can be seen, relative size distribution of segments made more sense. Moreover, scree plot after taking 4 number of segments was better than taking 4 number of segments.

Now when I ran k-means analysis without forcing number of segments, it automatically selected 4 segments. K-means helped in breaking down segments.

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Size	2 000	991	380	412	217
Relative size	100%	50%	19%	21%	11%

Segment size.

Scree plot:



The scree plot also shows promising elbow shape.

So, I decided to do analysis using k-means method of segmentation for AutoData.

After proper segment distribution it was time to analyze segments.

Segmentation data analysis:

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Segments = NON LUXURY TRADITIONAL MID SIZE	0.061	0.095	0.071	0.000	0.000
Segments = NON LUXURY COMPACT CUV	0.121	0.000	0.066	0.000	1.000
Segments = NON LUXURY TRADITIONAL COMPACT	0.078	0.107	0.132	0.000	0.000
Segments = LUXURY MID SIZE CUV	0.085	0.000	0.105	0.316	0.000
Segments = NON LUXURY MID SIZE CUV	0.081	0.144	0.050	0.000	0.000
Segments = LUXURY COMPACT CUV	0.086	0.000	0.037	0.381	0.000
Make = KIA	0.050	0.064	0.024	0.000	0.129
Make = FORD	0.064	0.092	0.055	0.000	0.074
Make = TOYOTA	0.075	0.081	0.155	0.000	0.051
Make = BMW	0.063	0.000	0.024	0.284	0.000
Make = CHEVROLET	0.057	0.083	0.050	0.000	0.060
Make = MERCEDES-BENZ	0.053	0.000	0.050	0.211	0.000
Fuel Type = GASOLINE	0.798	1.000	0.000	0.998	0.894
Fuel Type = HYBRID	0.079	0.000	0.413	0.000	0.000

So, while doing analysis, I have found different attributes of all segments. This made various clusters for segments.

Segment 1- Middle class working people

In this segment I have noticed that all the people in this segment prefer non-Luxury cars. They have a serious dislike for luxury cars. Another thing that I have noticed from segment data is that people in this segment only prefers gasoline as fuel type, they do not want a hybrid car.

Segment 2 – JDM lovers

Analysis on this segment was a bit tricky because people in this segment does not score 0 in any car segment, hence in this segment people like to experiment different car types. They do not want gasoline as their car fuel they just want hybrid cars. Toyota is the most proffered make for them and out of all car segments they have highest liking for non-luxury traditional compact. It may be possible that they have love for Japanese domestic motorcars (JDMs). Toyota is a JDM so, they prefer it more

Segment 3 – Deluxe class

People in this segment are only interested in luxury CUVs of all types. They want gasoline powered car and not a hybrid one.

Segment 4 – Casual drivers

This segment is very simple and clear. People only want a non-luxury compact CUV and out of all makes, preferably a KIA. They do not want hybrids, only gasoline as fuel type.

Descriptor Analysis:

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Age = 55-64 YEARS OLD	0.161	0.170	0.147	0.155	0.152
Age = 35-44 YEARS OLD	0.180	0.184	0.205	0.177	0.124
Age = 45-54 YEARS OLD	0.182	0.171	0.168	0.199	0.221
Age = 65-74 YEARS OLD	0.152	0.168	0.137	0.146	0.120
Age = 25-34 YEARS OLD	0.157	0.145	0.208	0.141	0.152
Age = 75+ YEARS OLD	0.118	0.113	0.100	0.146	0.120
Enhanced Ethnic Groups = ASIAN	0.132	0.128	0.124	0.148	0.129
Enhanced Ethnic Groups = HISPANIC	0.169	0.155	0.197	0.182	0.152
Enhanced Ethnic Groups = CAUCASIAN	0.311	0.306	0.324	0.308	0.318
Enhanced Ethnic Groups = OTHER	0.166	0.178	0.129	0.180	0.152
Enhanced Ethnic Groups = AFRICAN AMERICAN	0.144	0.151	0.158	0.109	0.147
Gender = MALE	0.545	0.561	0.592	0.517	0.438
Income = 0-49999	0.304	0.306	0.313	0.248	0.387
Income = 50000-99999	0.189	0.184	0.145	0.223	0.221
Income = 150000-199999	0.150	0.147	0.168	0.148	0.129
Income = 100000-149999	0.183	0.191	0.179	0.187	0.143
Income = 250000+	0.097	0.095	0.095	0.119	0.065

After running descriptor data, I did analysis and concluded for each segment.

Segment 1

Majority of population aged between 65 and 74 are present in segment 1. They have average salary range, so they want a car that is comfortable and easy on pocket. Also, they so not want to experiment on hybrid cars. This can justify their liking for non-luxury traditional cars as they run on gasoline and are affordable.

<u>Recommendation:</u>

I would say to target this segment, we need to have a car that is easy to drive, have no complex maintenance issues as people just want to use car as a means of transportation and not fun. So, the suggestion would be affordable and comfortable car that runs on gasoline.

Segment 2

This segment has comparably young working-class adults and majority of them are male. They have a bit low income so that justifies their preference for hybrid cars that would not burst their fuel budget. This segment has mixed preference for car types. Toyota is more preferred by them.

Recommendation:

To target this segment, I would say that we need some fancy hybrid car models.

Segment 3

This segment is all about Luxury. They want luxury cars, but CUVs. They do not have that high income, but their age is more too. We can say they just want to have fun around and complete their wishes that's why they are not much concerned about having savings.

Recommendation:

Recommendation would be a fancy CUV luxury car that offers comfort and driving pleasure.

Segment 4

This segment has majority of people who are in middle of their career. They have lowest income slab and majority of them are female.

Recommendation:

They have low income so, a cheap and reliable car would be nice target.

YogaStudio (Segmentation)

This dataset is from a yoga studio's customers, each row representing a yoga session. The segmentation tab describes the characteristics of the sessions that each customer chose, the variables being
Service category – in studio, livestream, workshop
Visit type – beginner, mixed, workshop, all
Pricing option – Many different levels
Booking method – different ways the booking was made, from an app to "other"

Use these variables to first run a segmentation analysis and describe the segments you get. Next, incorporate the discriminant analysis, with variables
First visit – Whether they first visited in Jan, Feb, Mar, or Apr. This is important, given that people are most likely to join fitness services early in the year.
Number of visits since then – This indicates frequency and/or commitment
Referral type – where the customer came through.
Staff – The person they took the class from. Very important in fitness settings where people prefer certain trainers.

Given this fresh and live dataset, conduct a full segmentation and discriminant analysis, and based on your findings, come up with recommendations for the marketing people at the yoga studio.

Solution:

I ran analysis with 3 & 4 number of segments for this case analysis and found relative sizes as can be seen in figure.

	Population	Segment 1	Segment 2	Segment 3
Size	330	114	148	68
Relative size	100%	35%	45%	21%

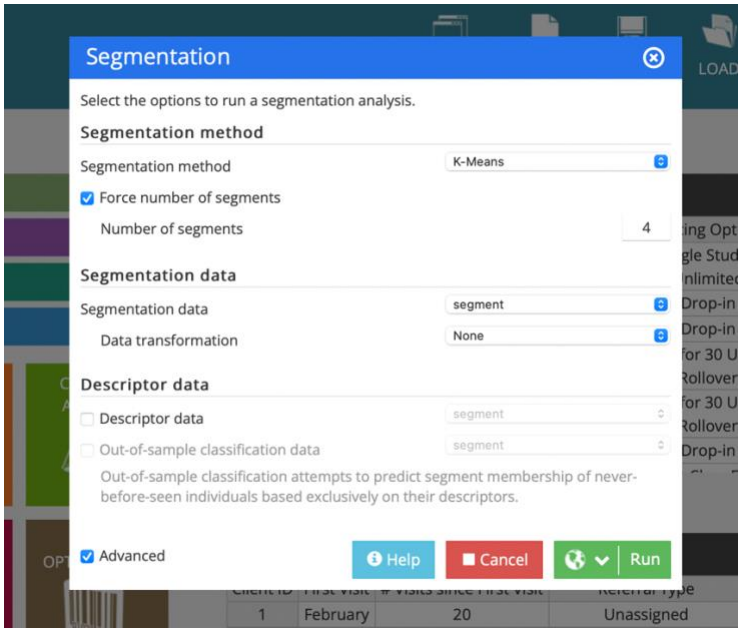
Segment size.

For 4 number of segments:

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Size	330	114	62	86	68
Relative size	100%	35%	19%	26%	21%

Segment size.

Relative size distribution is more accurate in 4 number in segments, as a segment is divided further. But after looking at segment description table, I decided to run k-means segmentation for further analysis.



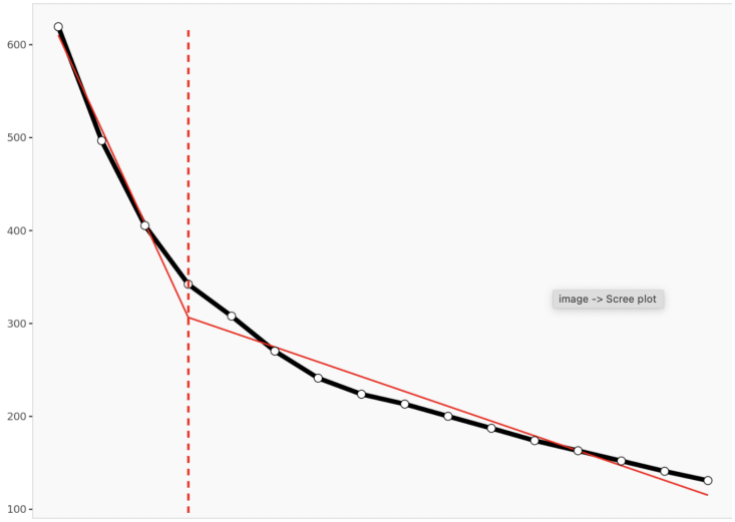
After running K-means segmentation, I found that segmentation in K-means method had some more clarity.

k-means method 4 number of segments:

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Size	330	107	57	90	76
Relative size	100%	32%	17%	27%	23%

Segment size.

Scree Plot:



The scree plot showed a better elbow, and thus it could provide a good cluster solution. Having K-means method and 4 number of segments, below is the segment description table

	Population	Segment 1	Segment 2	Segment 3	Segment 4
Service Category = In Studio Yoga class	0.955	0.925	1.000	0.922	1.000
Visit Type = mixed level	0.318	0.981	0.000	0.000	0.000
Visit Type = beginner	0.433	0.000	1.000	0.956	0.000
Visit Type = all levels	0.230	0.000	0.000	0.000	1.000
Pricing Option = Charter Single Studio Monthly Unlimited	0.100	0.112	0.035	0.167	0.053
Pricing Option = Drop-in	0.076	0.065	0.175	0.067	0.026
Pricing Option = 30 Days for 30 Unlimited Rollover	0.197	0.234	0.018	0.333	0.118
Pricing Option = First Class Free	0.152	0.150	0.246	0.089	0.158
Pricing Option = Kansas City Single Class Drop-in	0.076	0.093	0.070	0.100	0.026
Pricing Option = Free Class	0.052	0.009	0.246	0.022	0.000
Pricing Option = KIDS only Drop-In (Non-Passholder)	0.058	0.000	0.035	0.011	0.211
Booking Method = Consumer Mode	0.112	0.168	0.000	0.122	0.105
Booking Method = Business Mode	0.339	0.234	0.947	0.000	0.434
Booking Method = Other	0.506	0.523	0.000	0.844	0.461

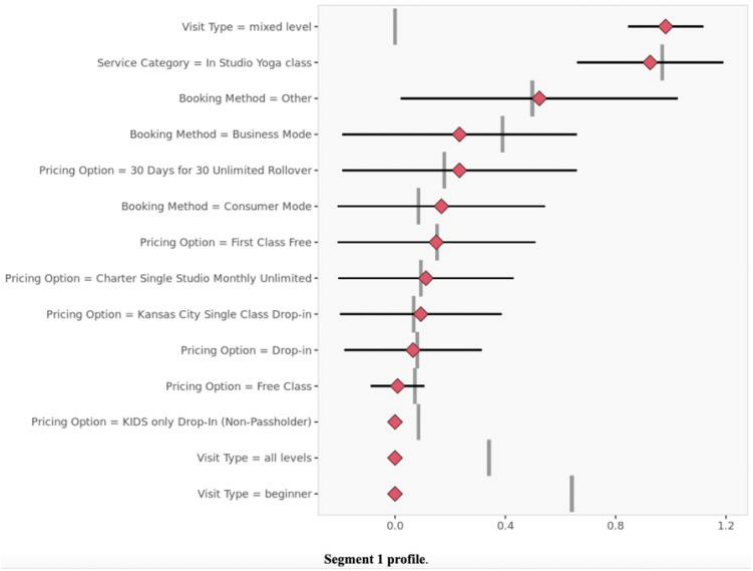
Loading the segment data, we can see that, following attributes for each segments.

I've also placed segment profile chart for each segment, where the different colored dots will be representing average of the segment and gray vertical lines will represent where other segments

are on average compared to current segment. And the horizontal lines represent standard deviation within that segment.

Segment 1 – Mixed class

People in this segment are at mixed level. None of them were beginners at all. It can be noticed that they were not interested in free class option. So, we can say that these people were not interested in any free class as they might be regulars at the Yoga studio and just want to have their sessions. They had highest population for consumer mode and lowest for business mode, hence they were regular individual customers.

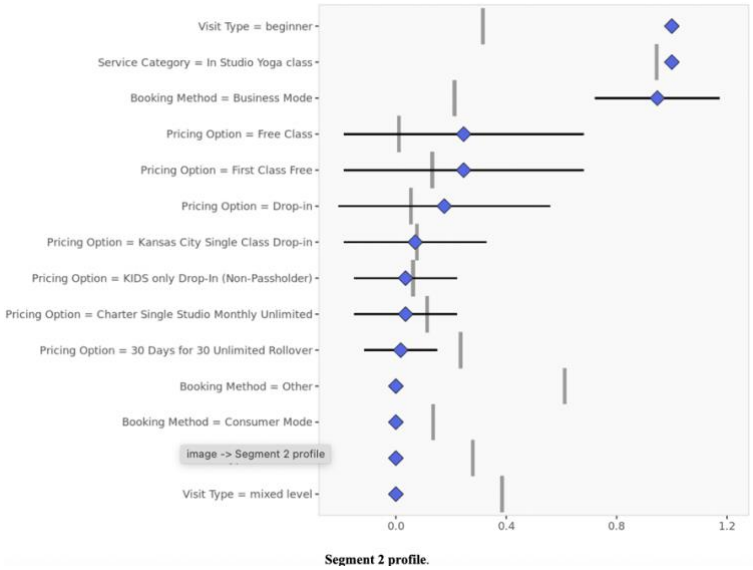


Segment 1 profile.

Here profile chart easily identifies mixed level of visit type is very far from average of other segments when comparing. From segment description chart we can back this up by noticing that all other segments are either beginners or of all levels but, not of mixed type.

Segment 2 – Corporate Beginners

For this segment it is all clear that all of them are of beginner levels. As they are beginners, without experience in Yoga classes, it is obvious that they wanted to have free class option. As a result, this segment had highest people for free class in pricing option and drop-in pricing option too. They also preferred only In Studio Yoga classes option. People in this segment had preference for business mode of booking.



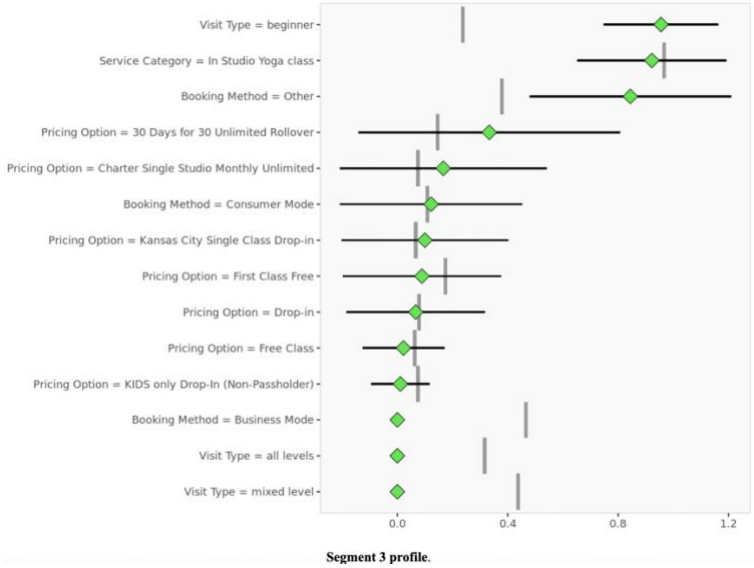
Segment 2 profile.

Here it can be noticed that vertical line of average of other segments for beginner visit type is low around 0.3, whereas segment 2 has all members having beginner status.

Based on my findings I can name this segment corporate beginners, who are working class and beginners.

Segment 3 – Determined Beginners

People in this segment are not only beginners, but they also prefer 30 days for 30 unlimited rollover option for pricing. Their booking method is not identified. It is important to notice that single studio monthly unlimited option was preferred by this segment



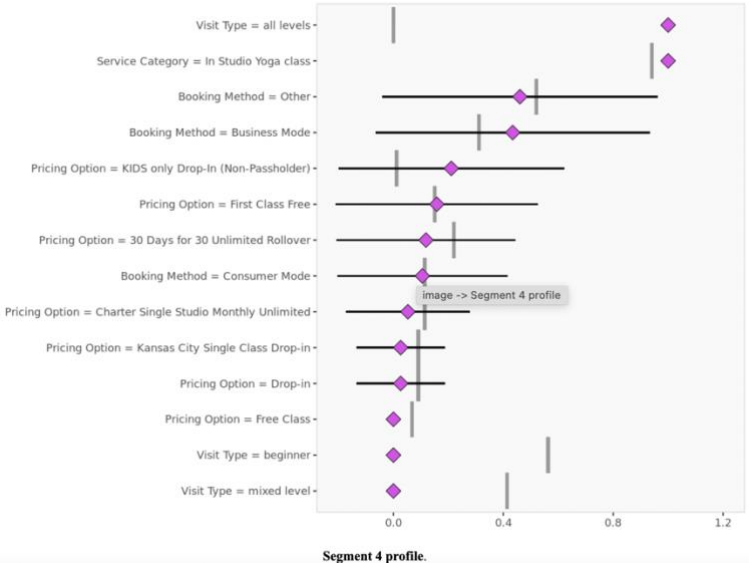
Segment 3 profile.

When analyzing segment profile, I found big difference in beginner visit type compared to other segments and booking method variance from other segments.

I could say this people in this segment are beginners but with determination. They want monthly unlimited option and they also want 30 days roll over option.

Segment 4 – Housemakers

All the people in this segment want In Studio Yoga class option. They want KIDS drop-in option indicating that they could be people that have kids and do not want to leave them alone at home. Their visit type ranks highest in all levels.



Segment 4 profile.

Looking at segment profile data, it is noticeable KIDS only drop-In option is relatively higher. In Studio service category is also higher than average.

Based on finding I want to name this segment Housemakers.

Now I ran with descriptor data.

	Population	Segment 1	Segment 2	Segment 3	Segment 4
First Visit = February	0.203	0.234	0.228	0.167	0.184
First Visit = March	0.148	0.140	0.175	0.178	0.105
First Visit = January	0.564	0.486	0.579	0.644	0.566
# Visits since First Visit	8.28	8.11	5.21	10.88	7.72
Referral Type = Unassigned	0.633	0.589	0.842	0.544	0.645
Referral Type = Another client	0.167	0.178	0.088	0.178	0.197
Staff = Gough, Kelly	0.079	0.131	0.053	0.100	0.000
Staff = Topliff, Sydney	0.197	0.178	0.281	0.233	0.118
Staff = Birdsong, Cory	0.052	0.000	0.070	0.111	0.039
Staff = Lesmeister, Casey	0.094	0.037	0.070	0.200	0.066
Staff = Belyea, Patrick	0.055	0.000	0.088	0.056	0.105
Staff = Messina, Blair	0.058	0.000	0.000	0.000	0.250
Staff = Rea, Kristin	0.133	0.019	0.158	0.144	0.263
Staff = Fiacco, Jessie	0.052	0.159	0.000	0.000	0.000
Staff = Tajchman, Shannon	0.061	0.047	0.246	0.011	0.000

Segment 1

Highest number of first visits in February was from this segment. Majority of people in this segment had Gough, Kelly and Fiacco, Jessie as their staff.

Recommendation:
They are mixed class, but they only prefer certain staff members such as Gough, Kelly and Fiacco, Jessie as their staff. So, to target them Yoga studio should make these staff member's availability more to people in this segment.

Segment 2

Majority of individuals have unassigned referral type & preferred Tajchman, Shannon as their staff as staff members could give timings to them in their busy schedule. One good analysis is that people in this segment had lowest visit count after their first visit because corporate beginners in segment are busy and come up with many things that disrupt their Yoga class schedule. This segment has staff allocation in almost every staff in Yoga studio.

Recommendation:
To target suggestion would be a free class option and In Studio Yoga class option.

Segment 3

This segment had people with highest number of visits after their first visit, as they are determined beginners. They preffered Birdsong, Cory and Lesmeister, Casey as their staff.

Recommendation:
To target recommendation would be to give unlimited single chartered monthly option and give more roll over options.

Segment 4

This segment did not prefer staff Tajchman, Shannon at all. This could be because this are homemakers and want to explore more in Yoga

Recommendation:
Suggestion is to have KIDS only drop-in option and have In Studio yoga classes option.