

## MARKET SEGMENTATION ANALYSIS

Brandi Rodriguez

VLJ466

### 1. BUSINESS CASE

Clinique's Product Development team is developing a new line of skin care products (i.e. moisturizers, creams, lotions) made of natural ingredients to compete in a market with a growing share of consumers who prefer natural or organic products made with non-toxic ingredients and less chemicals.

### 2. TARGET POPULATION

The target population for the new skin care line is women in the U.S. age 18 years or older, which represents a large portion of Clinique's customer base. There were 11,184 female respondents in the Simmons National Consumer Survey whose responses were incorporated into this segmentation analysis.

### 3. VARIABLES

See **Appendix A** for Frequency Tables

#### Statistical Driver / PCA Factor 1

- naturalbeauty: When I shop for health and beauty care products, I look for organic/natural items
- localgrown: I make an effort to buy locally grown food
- artificial: I prefer foods without artificial additives
- nutritional: Nutritional value is the most important factor in the foods I eat

#### Statistical Driver / PCA Factor 2

- lookyoung: It is important to keep looking young
- spendwhatever: I will spend whatever I have to, to make myself look younger
- attractive: It is important to look attractive to others
- fashion: I keep up with changes in styles and fashions

#### Other Statistical Drivers

- social media: I like to follow my favorite brands or companies on social media sites
- ad recall: I remember advertised products when shopping

#### Descriptor Variables:

- clinique: The brand I use most often for moisturizers/creams/lotions is Clinique
- sephora: I've shopped at Sephora in the last 30 days
- ulta: I've shopped at Ulta in the last 30 days
- apparel: I've spent \$1000 or more on women's apparel in the last three months
- bravo: I've viewed Bravo in the last 7 days
- onlineinfo: I have gathered information online for shopping in the last 30 days
- orderonline: I've ordered cosmetics online in the last 3 months
- cigarettes: I smoke cigarettes
- reason antiage: I use facial cleansing/medicated products/toners for anti-aging

#### 4. FACTOR ANALYSIS

The variables selected for the factor analysis were intended to measure two latent, unobserved constructs:

F1: consideration for the ingredients and chemicals contained in products consumed

F2: desire to have a youthful and attractive appearance

These factors represent the two main motivations for using our product. The first factor is the variable **allnatural**, which will help gauge how strongly a customer feels about consuming products with natural non-toxic ingredients that complement their sense of health or well-being. The second is the variable **attractiveness**, which will provide insight on the importance of looking young and attractive.

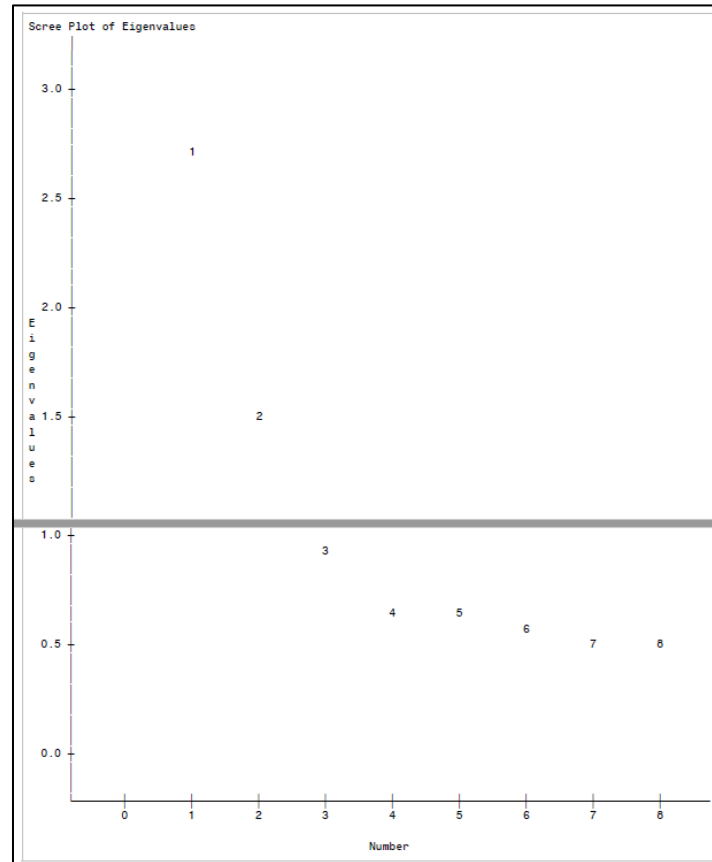
The factors will help distinguish differences in the main psychological drivers of interest between clusters, where there will likely be polarized segments with some in between that are more neutral. For example, there are customers who shop for cosmetics with the sole intention of buying the most promising products to boost their confidence and make them feel attractive, with little thought or care about the chemicals involved in achieving their ideal appearance. Others are very meticulous about vetting what they expose themselves to, and if a product is perceived to be harmful to their health, they will not make a purchase just for the sake of beauty. Then, there will be neutral consumers who may care about both ideas to varying degrees, and can still be swayed to try our product depending on the promises of the product and the marketing tactic used.

Factor analysis was used as a data reduction technique to reduce the eight selected questions into two factors. Principal Component Analysis (PCA) was chosen as the method of extraction because of its computational ease to perform. The Varimax (Orthogonal) Rotation Method was executed to create factors that are as uncorrelated to each other as possible, making it easier for the clustering algorithm to place observations in different clusters. Rotating the axes helps discriminate the variables and because the factors are orthogonal, it is more likely the clusters will be well-separated.

The Principal Component Analysis resulted in two factors using the Kaiser Criterion, which states factors with an eigenvalue greater than 1 should be used. The first two factors explain 52.34% of the variance, with Factor 1 explaining 33.71% and Factor 2 explaining 18.63%. Including a third factor would not substantially increase the cumulative variance explained, so any additional factors were excluded.

The FACTOR Procedure			
Initial Factor Method: Principal Components			
Prior Communality Estimates: ONE			
Eigenvalues of the Correlation Matrix: Total = 8 Average = 1			
	Eigenvalue	Difference	Proportion
1	2.69651352	1.20585229	0.3371
2	1.49066124	0.55846272	0.1863
3	0.93219852	0.26125622	0.1165
4	0.67094230	0.04115889	0.0839
5	0.62978340	0.03173635	0.0787
6	0.59804706	0.09648264	0.0748
7	0.50156442	0.02127488	0.0627
8	0.48028954		0.0600
			Cumulative
			0.3371
			0.5234
			0.6399
			0.7238
			0.8025
			0.8773
			0.9400
			1.0000

The scree plot supports the use of two factors as well, where the elbow – a sharp decline – can be seen at the second factor:



The PCA technique was successful in creating two separate factors. The two factors are clearly distinguished by the Rotated Factor Matrix loadings, with Factor 1 more heavily loaded on the first four variables and is renamed **allnatural**. Factor 2, renamed **attractiveness**, is more heavily loaded on the remaining four variables.

Rotated Factor Pattern				
			Factor1	Factor2
naturalbeauty	When I shop for health and beauty care products, I look for organic/natural items		0.56025	0.35695
localgrown	I make an effort to buy locally grown food		0.80095	0.03017
artificial	I prefer foods without artificial additives		0.81260	-0.00458
nutritional	Nutritional value is the most important factor in the foods I eat		0.66774	0.19281
lookyoung	It is important to keep looking young		0.07876	0.72688
spendwhatever	I will spend whatever I have to, to make myself look younger		0.08737	0.75377
attractive	It is important to look attractive to others		0.16062	0.52152
fashion	I keep up with changes in styles and fashions		0.06318	0.74022

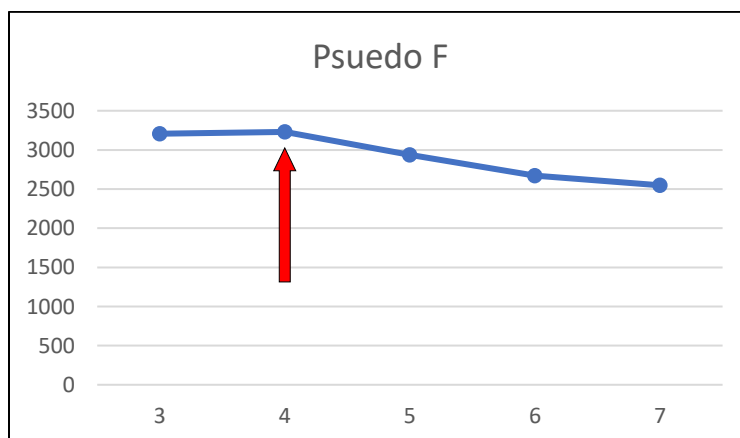
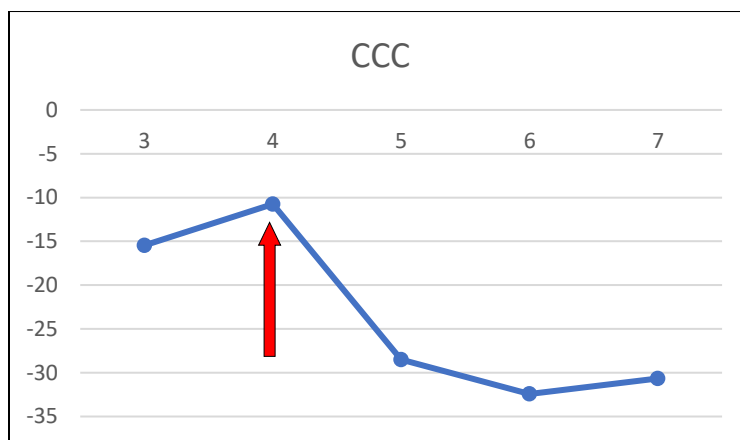
## 5. K MEANS

The FASTCLUS procedure was performed in SAS using the two aforementioned PCA factors in addition to two other statistical drivers selected:

- socialmedia: I like to follow my favorite brands or companies on social media sites
- ad\_recall: I remember advertised products when shopping

The procedure was executed for 3 to 7 clusters, resulting in the diagnostics displayed below. The CCCs were negative but there was a promising first local maximum at 4 clusters for both the CCCs and Pseudo F measures. Although negative CCCs are not ideal, this was the optimal and most viable solution out of the trials that were executed using various combinations of statistical drivers (see **Appendix B** for trial results).

K	Pseudo F	R <sup>2</sup>	CCC
3	3206.31	0.40333	-15.449
4	3229.51	0.49169	-10.75
5	2936.63	0.56591	-28.51
6	2670.32	0.60381	-32.43
7	2547.91	0.63327	-30.653



## 6. GAP ANALYSIS

The HPCLUS procedure was executed in SAS using the same statistical drivers, and it suggested the use of five clusters. This was where the first peak occurred, indicating the point of the most separation between clusters.

The HPCLUS Procedure

ABC Parameters			
Minimum Cluster	Maximum Cluster	Reference Distribution Count	Alignment Method
2	6	20	PCA

ABC Statistics					
Number of Clusters	Logarithm of Within-Cluster SSE		Gap	Simulation Adjusted Standard Deviation	One Standard Error Adjusted Gap
	Input	Reference			
2	9.6807	10.9357	1.2549	0.0139	1.2411
3	9.5235	10.5524	1.0289	0.0165	1.0124
4	9.4030	10.3658	0.9628	0.00892	0.9539
5	9.2997	10.2914	0.9917	0.0223	0.9694
6	9.2214	10.1075	0.8861	0.0132	0.8729

There's a .9694 difference (the one standard error adjusted gap) between the .9917 gap for cluster = 5 and the simulation adjusted standard deviation of .0223. This difference is larger than the .9628 gap number for cluster = 4, indicating that cluster = 5 is farther than 1 standard deviation from the previous gap number. SAS confirms this is the first peak, suggesting the optimal number of clusters is 5:

Estimated Number of Clusters	
Criterion	Number of Clusters
FIRSTPEAK	5

## 7. SEGMENTS

After reviewing the results and cluster means of both solutions, the market segmentation was completed using the K Means (FASTCLUS) Procedure. The four clusters that resulted from this procedure were more evenly distributed, with no "clumps" in the distribution of the observations within each cluster. There were 2,319 observations in Cluster 1, 3,274 in Cluster 2, 2,099 in Cluster 3, and 3,107 in Cluster 4. The means in the selected solution had a good amount of distance, indicating the solution did a good job of separating the clusters into four distinct groups (see **Appendix C** for means tables).

cluster	allnatural	attractiveness	socialmedia	ad_recall
1	-0.8145	0.3004	1.5461	3.6332
2	0.6065	-0.7294	1.3825	2.4583
3	-0.7654	-0.3801	3.5315	1.8350
4	0.4482	0.7214	3.7424	3.5416

### **Cluster 1: Anti-aging Beauty Consumers**

This segment is motivated by their desire to stay looking young and attractive. They care the absolute least about the all-natural aspect. This lack of interest is bolstered by the fact that this group had the second highest response for smoking cigarettes (15.4% compared to 15.5% seen in Cluster 3, which also cared very little about the all-natural aspect). They are not as inclined to order cosmetics online and have a lower tendency to gather information online for making shopping decisions. They're not interactive with brands on social media, but do have the highest ability to recall advertised products while shopping. If they purchase our product, it is more than likely because they saw it recently advertised, not because the all-natural benefits stood out to them. Online tactics will not be successful at engaging or reaching this segment. It's best to reach them through more traditional means (magazine ads, mailers, commercials) or through interactions at the stores they shop at like Sephora and Ulta. Just because they are not interested in the all-natural aspect, it is still a viable segment for our product, as they are still very interested in the anti-aging aspect and are more affluent shoppers, implied by their higher response to spending more than \$1000 or more on women's apparel in the last three months.

### **Cluster 2: Health Minded Consumers**

This segment cares the absolute most about all-natural ingredients and the absolute least about attractiveness. They are more serious shoppers and do their research before making a purchase, having the highest response for gathering information online for shopping. With the lowest response for using skin care products because of the anti-aging aspect of it, this segment more likely uses products for more practical reasons, like facial cleansing. Practical is a good way to describe this segment. They spend less money and time on frivolous things like watching Bravo or spending thousands of dollars on clothes. They are the least interactive with social media and have a lower propensity to recall advertised products, relying more on their own research. The best way to reach this segment is through the use of facts that highlight the natural aspect of our skin care line, not promises of youthful looking skin.

### **Cluster 3: Neutral Consumers**

This cluster is more neutral and seem to be unmotivated by either construct. They are interactive with brands on social media, but are least likely to recall advertised products. The most effective and cost-efficient way to reach our neutral consumers is through social media, perhaps organically by word-of-mouth from their social media connections or by contracting with beauty bloggers to highlight our new line of skin care. This group had the lowest response rate for shopping at Sephora, Ulta, and purchasing Clinique products. They also had the lowest response for using skin care products for anti-aging purposes. If they purchase our product, it is likely by chance or convenience and more for practical reasons (a need for a facial cleanser suitable for their skin type), as they are not concerned about being attractive or the ingredients in products consumed (they had the highest response for smoking cigarettes). Although this segment seems uninterested, there may still be a low-cost opportunity to reach them through social media, where they tend to be more engaged with companies and brands than other segments.

### **Cluster 4: Natural Beauty Consumers**

This segment cares highly about both concepts of all-natural ingredients and attractiveness, with the highest mean out of all clusters for attractiveness. They are the most likely to follow their favorite brands on social media and have a high ability to remember advertised products when shopping. Additionally, they had the highest response rate for shopping at Sephora and Ulta, and they do purchase Clinique products already. They like to watch Bravo, so a TV commercial or magazine ad would be effective and worth the marketing spend, as this segment has the highest inclination to recall product advertisements and are also our biggest spenders. Out of all segments, this one had the highest response for using skin

care products for anti-aging reasons. This is our ideal segment for gaining traction in the all-natural skin care space.

## **8. EXECUTIVE SUMMARY**

The segmentation system successfully identified two polarized segments, the Health Minded Consumers and the Anti-aging Beauty Consumers, with two more neutral groups in between. While one was generally apathetic and thus not a great target, the other is exceptionally ideal for gaining traction in the all-natural skin care space – the Natural Beauty Consumers. They are highly interested in both the all-natural and anti-aging concepts, are receptive to advertisements, and are engaged online for following their favorite brands online or using it to scope out product information. They already shop at stores where our product is traditionally sold such as Sephora, Ulta, and are noticeably bigger spenders. By engaging with this promising segment, there could be trickle down effects to other segments who may gain exposure to our product from online influencers or other marketing tactics initially geared towards the Natural Beauty Consumers. This strategy would ensure we reach our most receptive segments first while indirectly gaining the interest of other segments where there are overlapping interests and commonalities.

## APPENDIX A: FREQUENCY TABLES

Gender				
gender	Frequency	Percent	Cumulative Frequency	Cumulative Percent
female	11184	100.00	11184	100.00

The brand I use most often for moisturizers/creams/lotions is Clinique				
clinique	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	11125	99.47	11125	99.47
yes	59	0.53	11184	100.00

When I shop for health and beauty care products, I look for organic/natural items				
naturalbeauty	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	3407	32.12	3407	32.12
disagree a little	1934	18.23	5341	50.36
neither agree nor disagree	3494	32.94	8835	83.30
agree a little	1142	10.77	9977	94.07
agree alot	629	5.93	10606	100.00
Frequency Missing = 578				

I make an effort to buy locally grown food				
localgrown	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	1312	12.37	1312	12.37
disagree a little	1289	12.16	2601	24.53
neither agree nor disagree	3617	34.11	6218	58.64
agree a little	2496	23.54	8714	82.18
agree alot	1890	17.82	10604	100.00
Frequency Missing = 580				

I prefer foods without artificial additives				
artificial	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	1171	11.08	1171	11.08
disagree a little	1069	10.11	2240	21.19
neither agree nor disagree	3339	31.59	5579	52.78
agree a little	2439	23.07	8018	75.86
agree alot	2552	24.14	10570	100.00
Frequency Missing = 614				



Nutritional value is the most important factor in the foods I eat				
nutritional	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	1249	11.81	1249	11.81
disagree a little	1450	13.71	2699	25.51
neither agree nor disagree	3286	31.06	5985	56.57
agree a little	2614	24.71	8599	81.28
agree alot	1980	18.72	10579	100.00
Frequency Missing = 605				

It is important to keep looking young				
lookyoung	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	1008	9.46	1008	9.46
disagree a little	1624	15.23	2632	24.69
neither agree nor disagree	4093	38.40	6725	63.09
agree a little	2717	25.49	9442	88.57
agree alot	1218	11.43	10660	100.00
Frequency Missing = 524				

I will spend whatever I have to, to make myself look younger				
spendwhatever	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	4027	37.94	4027	37.94
disagree a little	2372	22.35	6399	60.29
neither agree nor disagree	3020	28.46	9419	88.75
agree a little	784	7.39	10203	96.14
agree alot	410	3.86	10613	100.00
Frequency Missing = 571				

It is important to look attractive to others				
attractive	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	632	5.99	632	5.99
disagree a little	687	6.51	1319	12.50
neither agree nor disagree	3187	30.19	4506	42.69
agree a little	3666	34.73	8172	77.42
agree alot	2383	22.58	10555	100.00
Frequency Missing = 629				

I keep up with changes in styles and fashions				
fashion	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	2756	26.03	2756	26.03
disagree a little	2283	21.56	5039	47.59
neither agree nor disagree	3590	33.90	8629	81.49
agree a little	1466	13.84	10095	95.33
agree alot	494	4.67	10589	100.00
Frequency Missing = 595				

I like to follow my favorite brands or companies on social media/networking sites				
socialmedia	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	1754	32.46	1754	32.46
disagree a little	909	16.82	2663	49.28
neither agree nor disagree	1302	24.09	3965	73.37
agree a little	972	17.99	4937	91.36
agree alot	467	8.64	5404	100.00
Frequency Missing = 5780				

I remember advertised products when shopping				
ad_recall	Frequency	Percent	Cumulative Frequency	Cumulative Percent
disagree alot	1447	13.72	1447	13.72
disagree a little	1968	18.67	3415	32.39
neither agree nor disagree	3968	37.64	7383	70.03
agree a little	2492	23.64	9875	93.66
agree alot	668	6.34	10543	100.00
Frequency Missing = 641				

I've shopped at Sephora in the last 30 days				
sephora	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	10953	97.93	10953	97.93
yes	231	2.07	11184	100.00

I've shopped at Ulta in the last 30 days				
ulta	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	10957	97.97	10957	97.97
yes	227	2.03	11184	100.00

I've spent \$1000 or more on women's apparel in the last three months				
apparel	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	11102	99.27	11102	99.27
yes	82	0.73	11184	100.00

I've viewed Bravo in the last 7 days				
bravo	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	10637	95.11	10637	95.11
yes	547	4.89	11184	100.00

I have gathered information online for shopping in the last 30 days				
onlineinfo	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	9452	84.51	9452	84.51
yes	1732	15.49	11184	100.00

I've ordered cosmetics online in the last 3 months				
orderonline	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	11094	99.20	11094	99.20
yes	90	0.80	11184	100.00

I smoke cigarettes				
cigarettes	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	9645	86.24	9645	86.24
yes	1539	13.76	11184	100.00

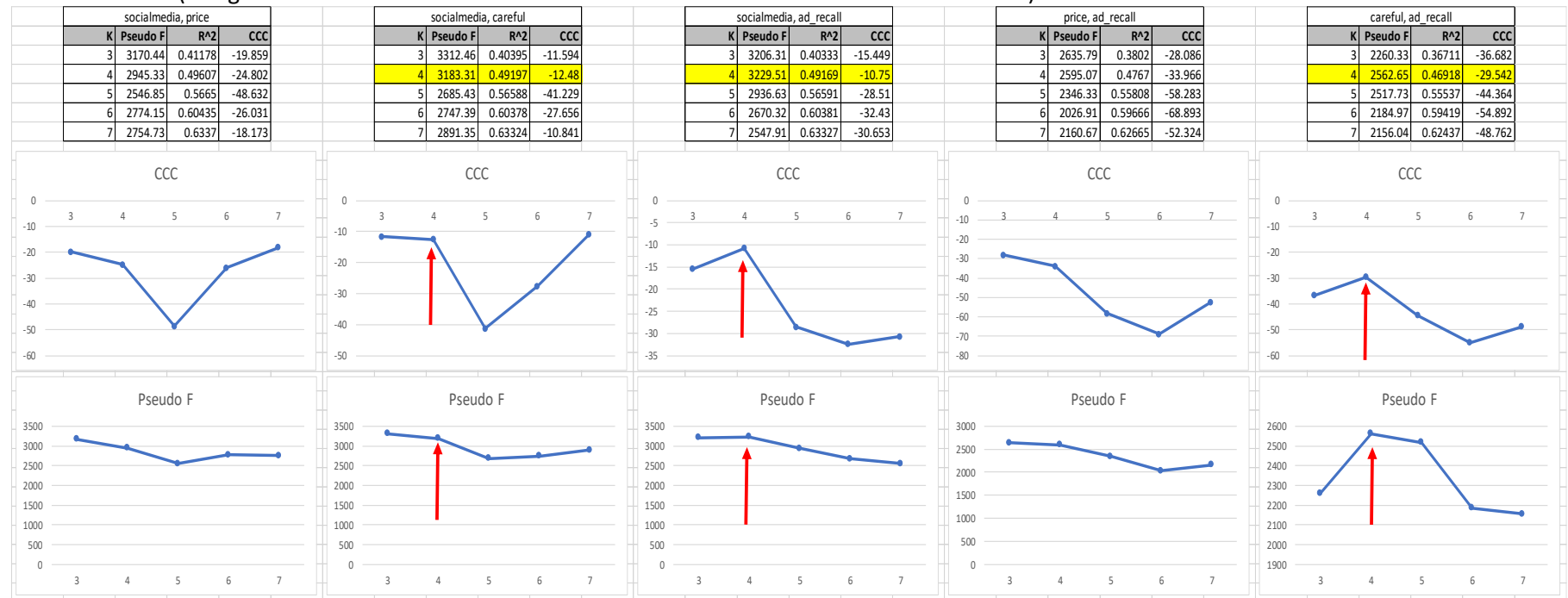
I use facial cleansing/medicated products/toners for anti-aging				
reason_antiage	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	10999	98.35	10999	98.35
yes	185	1.65	11184	100.00

## APPENDIX B: TRIALS

Statistical Drivers considered:

- price: price is not the most important factor – it is getting exactly what I want
- ad recall: I remember advertised products when shopping
- socialmedia: I like to follow my favorite brands or companies on social media / networking sites
- careful: I'm careful with money

Results of Trials (using **allnatural** and **attractive** PCA Factors as the two additional statistical drivers):



## APPENDIX C: MEANS TABLES

### STATISTICAL DRIVERS

cluster	allnatural	attractiveness	socialmedia	ad_recall
1	-0.8145	0.3004	1.5461	3.6332
2	0.6065	-0.7294	1.3825	2.4583
3	-0.7654	-0.3801	3.5315	1.8350
4	0.4482	0.7214	3.7424	3.5416

### DESCRIPTOR VARIABLES

cluster	clinique	sephora	ulta	apparel	bravo
1	0.65%	2.37%	2.07%	0.86%	5.05%
2	0.52%	1.47%	1.86%	0.49%	3.63%
3	0.24%	1.33%	1.33%	0.62%	4.19%
4	0.64%	2.83%	2.61%	0.97%	6.05%

cluster	onlineinfo	orderonline	cigarettes	reason_antiage
1	14.79%	0.60%	15.44%	2.07%
2	17.59%	0.70%	10.63%	1.01%
3	13.29%	0.62%	15.48%	0.91%
4	16.29%	1.22%	14.68%	2.45%

#### Cluster=1

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
clinique	The brand I use most often for moisturizers/creams/lotions is Clinique	2319	0.0064683	0.0801825	0	1.0000000
sephora	I've shopped at Sephora in the last 30 days	2319	0.0237171	0.1521992	0	1.0000000
ulta	I've shopped at Ulta in the last 30 days	2319	0.0206986	0.1424040	0	1.0000000
apparel	I've spent \$1000 or more on women's apparel in the last three months	2319	0.0086244	0.0924863	0	1.0000000
bravo	I've viewed Bravo in the last 7 days	2319	0.0504528	0.2189246	0	1.0000000
onlineinfo	I have gathered information online for shopping in the last 30 days	2319	0.1479086	0.3550859	0	1.0000000
orderonline	I've ordered cosmetics online in the last 3 months	2319	0.0060371	0.0774805	0	1.0000000
cigarettes	I smoke cigarettes	2319	0.1543769	0.3613876	0	1.0000000
reason_antiage	I use facial cleansing/medicated products/toners for anti-aging	2319	0.0206986	0.1424040	0	1.0000000

#### Cluster=2

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
clinique	The brand I use most often for moisturizers/creams/lotions is Clinique	3274	0.0051924	0.0718821	0	1.0000000
sephora	I've shopped at Sephora in the last 30 days	3274	0.0146610	0.1202100	0	1.0000000
ulta	I've shopped at Ulta in the last 30 days	3274	0.0186316	0.1352409	0	1.0000000
apparel	I've spent \$1000 or more on women's apparel in the last three months	3274	0.0048870	0.0697466	0	1.0000000
bravo	I've viewed Bravo in the last 7 days	3274	0.0363470	0.1871806	0	1.0000000
onlineinfo	I have gathered information online for shopping in the last 30 days	3274	0.1759316	0.3808201	0	1.0000000
orderonline	I've ordered cosmetics online in the last 3 months	3274	0.0070250	0.0835334	0	1.0000000
cigarettes	I smoke cigarettes	3274	0.1062920	0.3082581	0	1.0000000
reason_antiage	I use facial cleansing/medicated products/toners for anti-aging	3274	0.0100794	0.0999043	0	1.0000000

## Cluster=3

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
clinique	The brand I use most often for moisturizers/creams/lotions is Clinique	2099	0.0023821	0.0487601	0	1.0000000
sephora	I've shopped at Sephora in the last 30 days	2099	0.0133397	0.1147520	0	1.0000000
ulta	I've shopped at Ulta in the last 30 days	2099	0.0133397	0.1147520	0	1.0000000
apparel	I've spent \$1000 or more on women's apparel in the last three months	2099	0.0061934	0.0784729	0	1.0000000
bravo	I've viewed Bravo in the last 7 days	2099	0.0419247	0.2004649	0	1.0000000
onlineinfo	I have gathered information online for shopping in the last 30 days	2099	0.1329204	0.3395696	0	1.0000000
orderonline	I've ordered cosmetics online in the last 3 months	2099	0.0061934	0.0784729	0	1.0000000
cigarettes	I smoke cigarettes	2099	0.1548356	0.3618341	0	1.0000000
reason_antiage	I use facial cleansing/medicated products/toners for anti-aging	2099	0.0090519	0.0947326	0	1.0000000

## Cluster=4

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
clinique	The brand I use most often for moisturizers/creams/lotions is Clinique	3107	0.0064371	0.0799856	0	1.0000000
sephora	I've shopped at Sephora in the last 30 days	3107	0.0283231	0.1659211	0	1.0000000
ulta	I've shopped at Ulta in the last 30 days	3107	0.0260702	0.1593697	0	1.0000000
apparel	I've spent \$1000 or more on women's apparel in the last three months	3107	0.0096556	0.0978032	0	1.0000000
bravo	I've viewed Bravo in the last 7 days	3107	0.0605085	0.2384650	0	1.0000000
onlineinfo	I have gathered information online for shopping in the last 30 days	3107	0.1628581	0.3692956	0	1.0000000
orderonline	I've ordered cosmetics online in the last 3 months	3107	0.0122304	0.1099307	0	1.0000000
cigarettes	I smoke cigarettes	3107	0.1467654	0.3539288	0	1.0000000
reason_antiage	I use facial cleansing/medicated products/toners for anti-aging	3107	0.0244609	0.1545000	0	1.0000000