Alicja Wilk
Assignment 2
BANA 271
Winter 2019

SEGMENTATION AT STICKS KEBOB
CASE ANALYSIS

Abstract

The purpose of this case study was to identify the customer profile for Sticks Kebob based on different factors presented in the business case analysis " Segmentation at Sticks Kebob Shop"[1] .Sticks Kebob is looking to open another restaurant in Virginia. In this report six main topic were covered: 1) How do people choose the fast-food restaurant to visit? What is important? 2) Who do you think are Stick's customers and what are their motivations for visiting Sticks? 3) What does the survey data tell us about differences between customers and non-customers? 4) What survey questions would you use to identify the customer segments? 5) How many customer segments can you estimate from the survey data? 6) What is the best location for the next Sticks Kebab Shop?

Data

In order to analyze the customer data and segment the population of Sticks customers, first the data was cleaned using Alteryx and R software. Data manipulation was performed and 3 data sets were used to come up with the main segmentation data set that we will use to perform segmentation and clustering analysis. Missing values were replaced with median, since the data frame is integer. For the purpose of this case, data for customers was used to identify Sticks customers. Current customers are the most reliable source to use when creating a customers profile. Once, metrics on the data were performed, 190 non-customers were surveyed and 204 current customers were surveyed.

Analysis

Sticks Kebob has a hypothesis that their customers stop by for weekday lunch more often than for weekend dinner. They also categorizing their customers as people that are active, professional that are looking for a quick meal or active parents that come to grab a snack with their children after school or practice. Convenience matters as well as the fact that the meals are healthy. Sticks has variety of food that is nutritious and not too expensive. Sticks also did not want to position as too healthy or connect to any ethic foods. Their marketing strategy is to stay neutral. Hypothetical goal is to follow their previous strategies and attract students, but mainly local residents and families. Students are important but not as much as the locals.

How do people choose the fast food restaurants to visit?

Different factors like location, price, assortment or cuisine were taken into account when analyzing customer skewness towards some factors vs others. Answers to question 7 from the survey data provided in the case study was used to identify the most important factor that customers look at when deciding where to eat.

Statistical analysis

Table1

```
##          vars   n   mean     sd median trimmed    mad min max range  skew
```

```
## Q7a      12 190 1.49 0.65      1    1.41 0.00    1    5     4  1.54     4.12
## Q7b      13 188 1.69 0.70      2    1.59 1.48    1    5     4  0.98     1.76
## Q7c      14 189 1.33 0.58      1    1.24 0.00    1    4     3  1.86     3.88
## Q7d      15 189 1.48 0.69      1    1.35 0.00    1    4     3  1.40     1.63
## Q7e      16 188 1.07 0.25      1    1.00 0.00    1    2     1  3.37     9.40
## Q7f      17 189 1.60 0.61      2    1.57 0.00    1    5     4  1.18     4.45
## Q7g      18 189 1.75 0.65      2    1.69 0.00    1    5     4  0.87     2.64
## Q7h      19 189 1.28 0.51      1    1.20 0.00    1    4     3  2.14     6.37
## Q7i      20 187 2.12 0.94      2    2.04 1.48    1    6     5  0.99     1.64
```

Table 1 above shows descriptive statistics per response for question "Please indicate how important the following factors are when you visit a restaurant"(Question 7), plus importance based on computed central tendencies.

a. Convenient place to eat (5^{th} most important)
b. Variety of menu options (7^{th} most important)
c. Good value for money (3^{rd} most important)
d. Healthy menu options (4^{th} most important)
e. Food taste and satisfaction (1^{st} most important)
f. Friendly staff (6^{th} most important)
g. Pleasant ambiance (8^{th} most important)
h. Consistency/reliability (2^{nd} most important)
i. Part of community (9 – least important)

The answers from current customers ranked from 1(most important) to 6(least important). The glimpse into the mean analysis gives us a first insight that on average customers consider taste and satisfaction being most important and being a part of sticks community is least important.

Further statistical T-test was performed per each question to examine statistical significance between customer data and each feature for customers and non-customers(1,0 respectively)

T-test:

H_0 = True difference in means is equal to 0
H_1 = True difference in means is not equal to 0

Table2

```
## $Q7a
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 1.8986, df = 392, p-value = 0.05835
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.005120258  0.293355552
## sample estimates:
## mean in group 0 mean in group 1
##        1.600000        1.455882
##
##
## $Q7b
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = -0.16468, df = 392, p-value = 0.8693
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1442088  0.1219178
## sample estimates:
## mean in group 0 mean in group 1
##        1.694737        1.705882
##
## $Q7e
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 4.3199, df = 392, p-value = 1.981e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1144879 0.3057392
## sample estimates:
## mean in group 0 mean in group 1
##        1.347368        1.137255
##
##
## $Q7f
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 0.45174, df = 392, p-value = 0.6517
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1008401  0.1610053
## sample estimates:
## mean in group 0 mean in group 1
##        1.652632        1.622549
##
```

```
## $Q7c
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = -0.86427, df = 392, p-value = 0.388
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.16678154  0.06492395
## sample estimates:
## mean in group 0 mean in group 1
##        1.257895        1.308824
##
##
## $Q7d
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 4.6329, df = 392, p-value = 4.915e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2368484 0.5860618
## sample estimates:
## mean in group 0 mean in group 1
##        1.852632        1.441176
##
```

```
## $Q7g
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 0.35075, df = 392, p-value = 0.726
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1140601  0.1635957
## sample estimates:
## mean in group 0 mean in group 1
##        1.789474        1.764706
##
##
## $Q7h
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 2.0482, df = 392, p-value = 0.0412
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.005189651 0.253427480
## sample estimates:
## mean in group 0 mean in group 1
##        1.384211        1.254902
##
```

```
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 2.109, df = 392, p-value = 0.03558
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.01337986 0.38146018
## sample estimates:
## mean in group 0 mean in group 1
##        2.305263        2.107843
```

Significant variables:

Given the statistics above we can conclude that responses : *Healthy menu options, Food taste and satisfaction, Consistency/reliability* and *Part of community* are statistically significant and 5% level of significance. Healty and taste were also most important factors among consumers.

Business Insights

Therefore we can conclude that it is crucial for business in the QRS industry to consider health when designing their menu options as well as putting more emphasis on advertising this fact. Sticks should drop their old belief of not advertising the healthy part of their menu and instead should focus on giving this part of their product line more attention. Overall customer satisfaction and food quality are more importance and significant than other factors.

Alicja Wilk
Assignment 2
BANA 271
Winter 2019

Sticks loyal customers keep on coming back due to consistency of the menu, they value time they have for lunch and appreciate quick service and tasty, healthy food. The less menu options the better as multiple research have shown that consumers deal better with fewer options, especially if Stick wants this to stay QRS business. Restaurants that keep their menu to minimum are InnOut , Chipotle, Five Guys or Chick-filA. All successful chains that build their business on this type of marketing strategy.

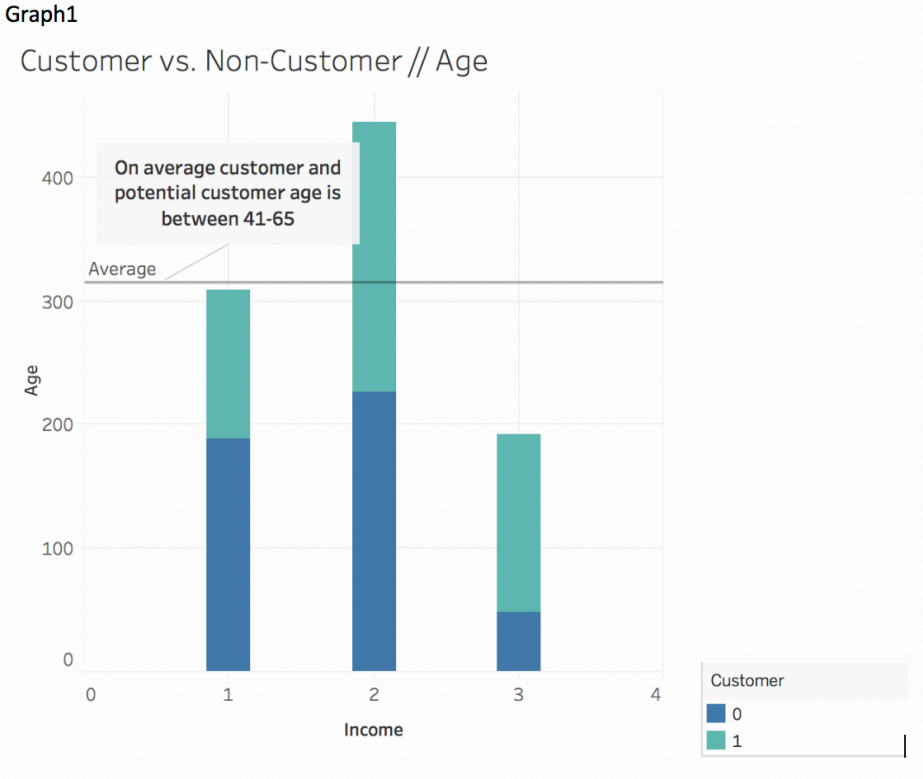
Who are Sticks customers and what is their motivation to visit?

Customers vs.Non - Customers

Note in the case report, a typical Sitcks customer would be a soccer mom that decideds on a place to eat for her and her family, a working professional that grabs a quick healthy bite during his lunch break or like recenty examined, active older customer base. What makes customers keep coming back?

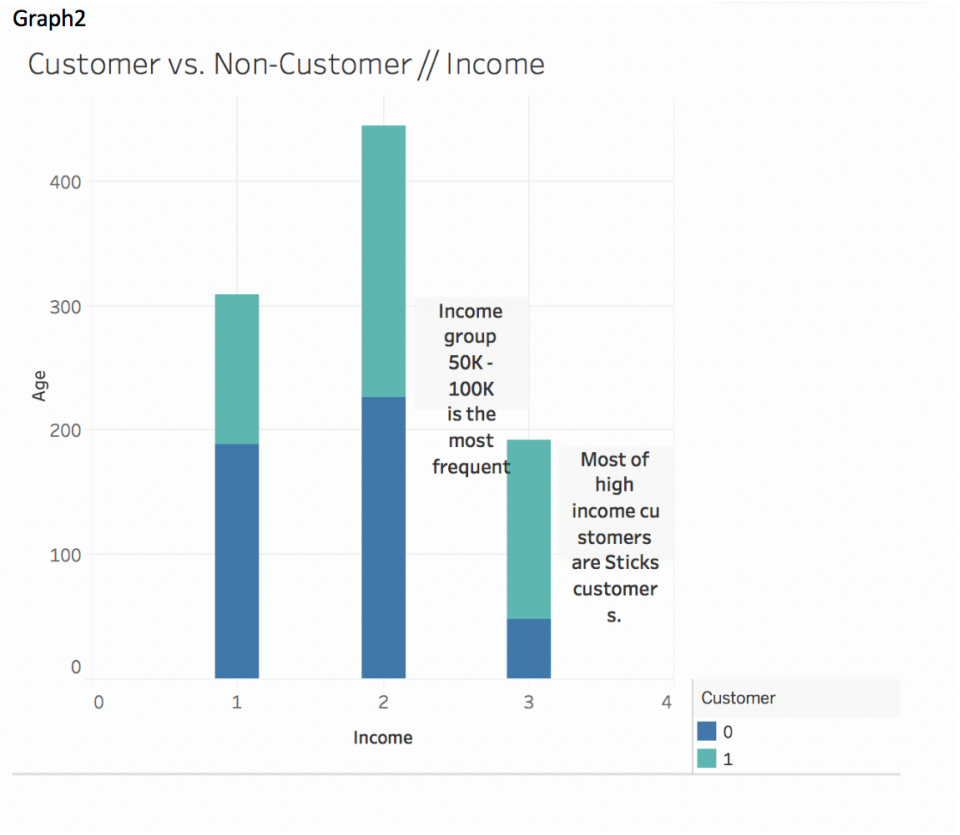
To test a typical Sticks customer profile, customers survey data was examined to come up with results. Demographics data was examined and results were visualized in Tableau. (Graph1,2,3)

Graph1

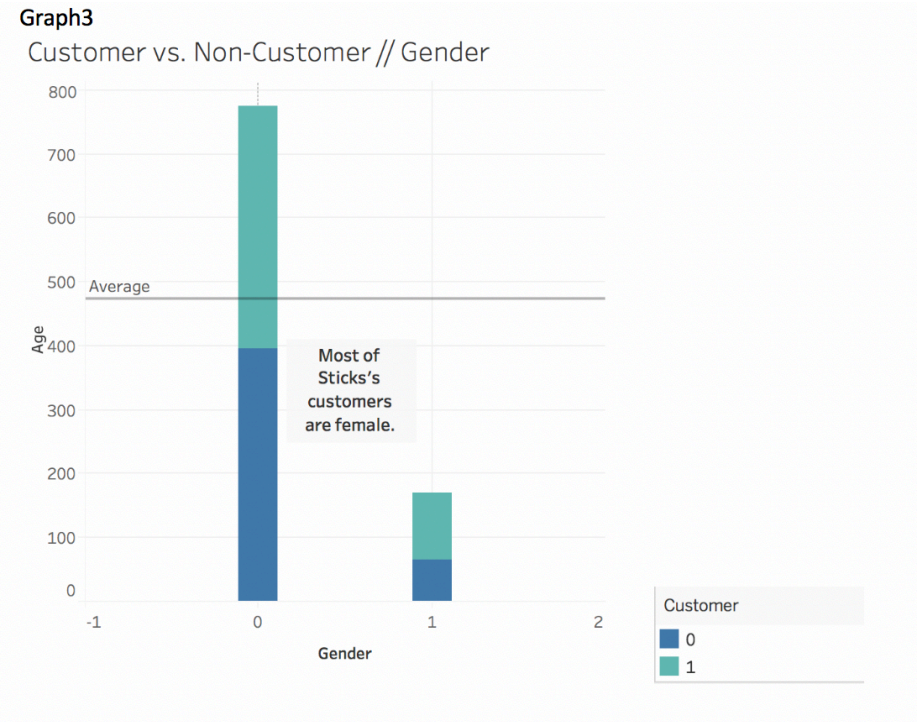


Alicja Wilk
Assignment 2
BANA 271
Winter 2019

Graph2



Graph3



Visual representation of the graphs states that Sticks customers are people in the age of 41-65 that have income 50K plus and most of Sticks customers are women.

Following statistical analysis(T-test) was performed on demographic variables of Sticks customers and non-customers(Table3).

H_0 = True difference in means is equal to 0
H_1 = True difference in means is not equal to 0

The finding were that Sticks customers are younger than non customers on average of similar restaurants and that females consist of most of the data. Gender and Age are not statistically significant at 5% level of significance.

Statically significant variable is Income with p= .0001109 at 5% level of significance. When comparing means of two groups, Sticks customers are higher income than non customers.

Motivation

Further analysis of question 16 and question 1 helped to discover what the factors are that might driver customers and non-customers in making the food decisions. Statistical computations are shown in Table4 for question 16 and Table5 for question 1(only statistically significant variables).

Sticks customers tend to either make food at home(on average 3 times a week compared to 4 times a week for non customers) or make a run and buy food out(also about 3 times average). At the same time they are willing to skip lunch and get a small snack instead(about 1.6 times on average a week).

Sticks customers also find it relatively important to purchase products that are made locally and they once again, do consider their health when making a choice on what to eat. (Both ranked highest on the scale of importance)(Graph4). As the figure shows customers gave more importance on questions d) and c) making these two a heavy skew to the left.

Graph4

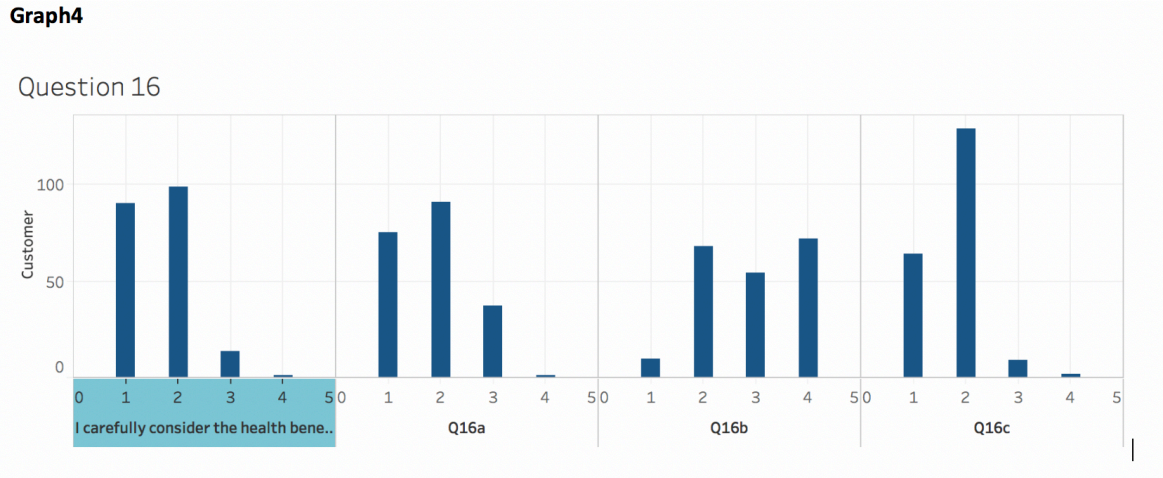


Table3

```
$Age

 Two Sample t-test

data:  v by as.factor(df$customer)
t = 1.0018, df = 392, p-value = 0.3171
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.07639049  0.23511082
sample estimates:
mean in group 0 mean in group 1
       2.442105        2.362745


$Income

 Two Sample t-test

data:  v by as.factor(df$customer)
t = -3.9051, df = 392, p-value = 0.0001109
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.413643 -0.136615
sample estimates:
mean in group 0 mean in group 1
       1.705263        1.980392


$Gender

 Two Sample t-test

data:  v by as.factor(df$customer)
t = -1.4522, df = 392, p-value = 0.1473
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.12789504  0.01922631
sample estimates:
mean in group 0 mean in group 1
      0.1368421       0.1911765
```

Alicja Wilk
Assignment 2
BANA 271
Winter 2019

Table4

```
## $Q16c
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 4.8525, df = 392, p-value = 1.763e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1768859 0.4178509
## sample estimates:
## mean in group 0 mean in group 1
##        2.047368        1.750000
##
##
## $Q16d
##
##  Two Sample t-test
##
## data:  v by as.factor(df$customer)
## t = 5.4166, df = 392, p-value = 1.062e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2377865 0.5087563
## sample estimates:
## mean in group 0 mean in group 1
##        2.010526        1.637255
```

Table5

```
$Q1a

        Two Sample t-test

data:  v by as.factor(df$customer)
t = 5.7583, df = 392, p-value = 1.718e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.710260 1.446706
sample estimates:
mean in group 0 mean in group 1
       4.357895        3.279412


$Q1d

        Two Sample t-test

data:  v by as.factor(df$customer)
t = -3.5988, df = 392, p-value = 0.0003606
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.7683633 -0.2254448
sample estimates:
mean in group 0 mean in group 1
       2.473684        2.970588


$Q1e

        Two Sample t-test

data:  v by as.factor(df$customer)
t = 3.2571, df = 392, p-value = 0.001223
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.1397602 0.5653998
sample estimates:
mean in group 0 mean in group 1
       1.994737        1.642157
```

Business Insights

Knowing that our customers have higher household income and do care about health benefits could make us consider creating a special sub-menu that could be priced a little premium and also budget money for advertising products that are made locally. Sticks customers expect great customer service with fresh and healthy product and menu options.

Survey Questions Choice for Segmentation

The purpose of segmenting customers into groups is giving firms and businesses a better understanding how their business model and business strategy are working and if what they had initially thought of for their business is actually the right target market. Target markets and customer segments are important to identify so businesses do not spend unnecessary money targeting customers that they should not be targeting. Segments also help with establishing brand recognition and customer motivation and satisfaction.

We should consider options that will help diversify our segments as well as diversify our responses. The main focus for customer segmentation analysis will be the current customers survey data. Non-customers will not be clustered on. The best reflection on Sticks customer will be done by analyzing current data and environment.

From different types of questions the choice was made to look at question 1 and 6 (behavioral) 7,8,16(psychographics) and demographics like age, gender and income.

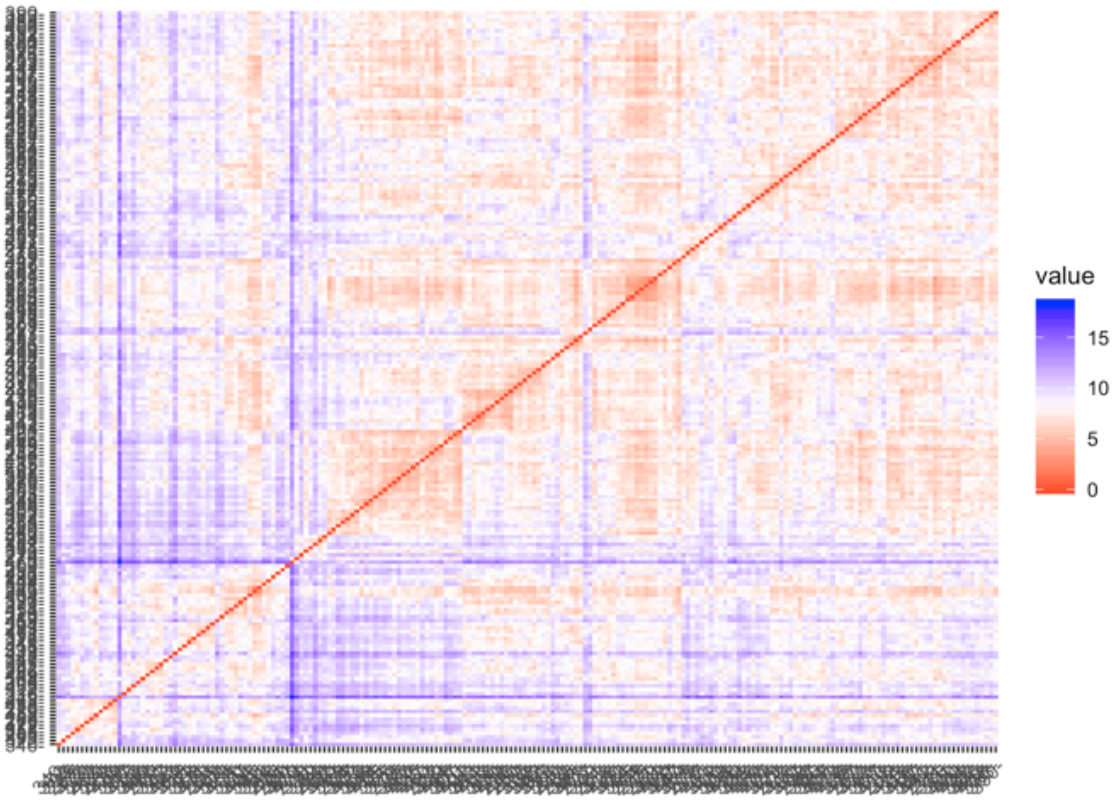
<u>Customer Segmentation</u>
Assumptions are made that we will be looking at targeted customer data which are Sticks current customers.

First, data was scaled and standardized before we combined distance measures. Data is then comparable with SD=1 and mean=0.
Snapshot of scaled data below *head()*.

```
##            Q1a        Q1b        Q1c         Q1d        Q1e        Q6a
## 268  0.4198870  0.2540848  1.4312037  0.01779553  4.8461766 -0.1343644
## 269  0.9499944  0.2540848 -0.4220216 -0.70114394 -0.5207514 -0.1343644
## 270 -0.6403277  1.4339730 -0.4220216 -0.70114394 -0.5207514 -0.1343644
## 271  0.4198870 -0.9258034 -0.4220216 -1.42008340  1.6260198 -0.1343644
## 272  0.9499944 -0.9258034 -0.4220216  0.73673500 -0.5207514 -0.6882750
## 273 -0.6403277 -0.9258034  5.1376545 -1.42008340  0.5526342 -0.6882750
      vars   n mean sd median trimmed  mad   min  max range  skew kurtosis   se
Q1a      1 202    0  1  -0.11   -0.08 1.57 -1.17 2.01  3.18  0.47    -0.94 0.07
Q1b      2 202    0  1  -0.34   -0.12 0.87 -0.93 2.02  2.95  0.71    -0.84 0.07
Q1c      3 202    0  1  -0.42   -0.27 0.00 -0.42 5.14  5.56  2.69     7.22 0.07
Q1d      4 202    0  1   0.02   -0.07 1.07 -1.42 2.89  4.31  0.61    -0.08 0.07
Q1e      5 202    0  1  -0.52   -0.24 0.00 -0.52 4.85  5.37  2.43     6.53 0.07
```

The data has n=202, “eucilidead’ metric was used as the distance measure. We compute distance to discover similarities and dissimilarities within the chosen data partition.

Visualization of Distances

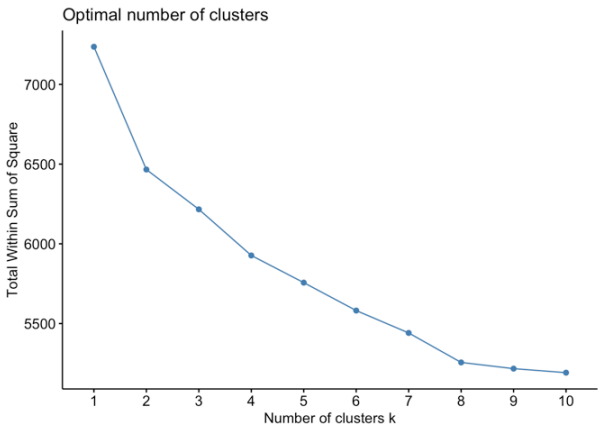


In the graph above, blue color represents low similarity values and red represents high similarity.

Further K-means data clustering is used to identify the high intra class similarity and the low inter class similarity to assign clusters to the data. The goal is to minimize the total inter cluster variation. The total sum of square measures the goodness of fit.

The elbow method was chosen to help with number of clusters.

```
set.seed(123)
fviz_nbclust(df1, kmeans, method = "wss")
```



K-Means in R was computed with 4 clusters and run 25 times to find optimal values. Output result outlined below:

```
str(k2)
```

```
## List of 9
##  $ cluster     : Named int [1:202] 1 4 4 2 3 2 4 2 3 1 ...
##   ..- attr(*, "names")= chr [1:202] "268" "269" "270" "271" ...
##  $ centers     : num [1:4, 1:36] -0.26367 0.07355 0.16465 -0.12537 -0.00984 ...
##   ..- attr(*, "dimnames")=List of 2
##   .. ..$ : chr [1:4] "1" "2" "3" "4"
##   .. ..$ : chr [1:36] "Q1a" "Q1b" "Q1c" "Q1d" ...
##  $ totss       : num 7236
##  $ withinss    : num [1:4] 975 2200 1421 1331
##  $ tot.withinss: num 5927
##  $ betweenss   : num 1309
##  $ size        : int [1:4] 38 75 54 35
##  $ iter        : int 5
##  $ ifault      : int 0
##  - attr(*, "class")= chr "kmeans"
```

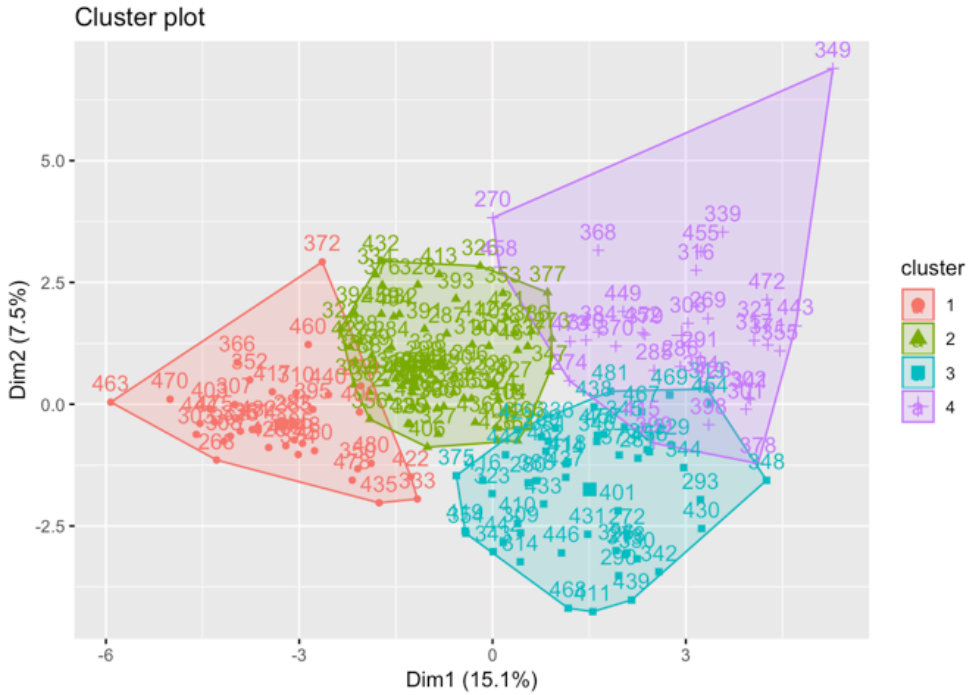
```
## Within cluster sum of squares by cluster:
## [1]  975.3131 2199.5352 1421.4974 1330.5972
##  (between_SS / total_SS =  18.1 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"
## [5] "tot.withinss" "betweenss"    "size"         "iter"
## [9] "ifault"
```

Interpreting the results:

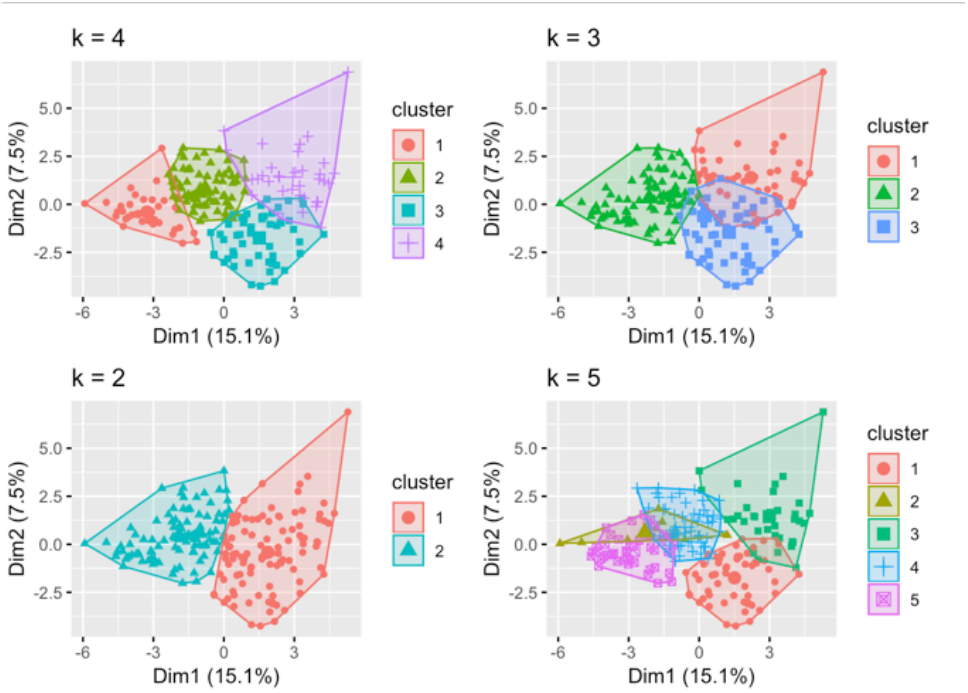
- cluster : vector of 1:202
- totss: total sum of squares = 7236
- withinss: within sum of sqares = 975,220,1421,1331
- size : 38,75,54,35 (points in each cluster)
- variation between SS = 18%

Plotting the K-Means:

```
fviz_cluster(k2, data = df1)
```



To see different results, different numbers of cluster were implemented to compare graphs:

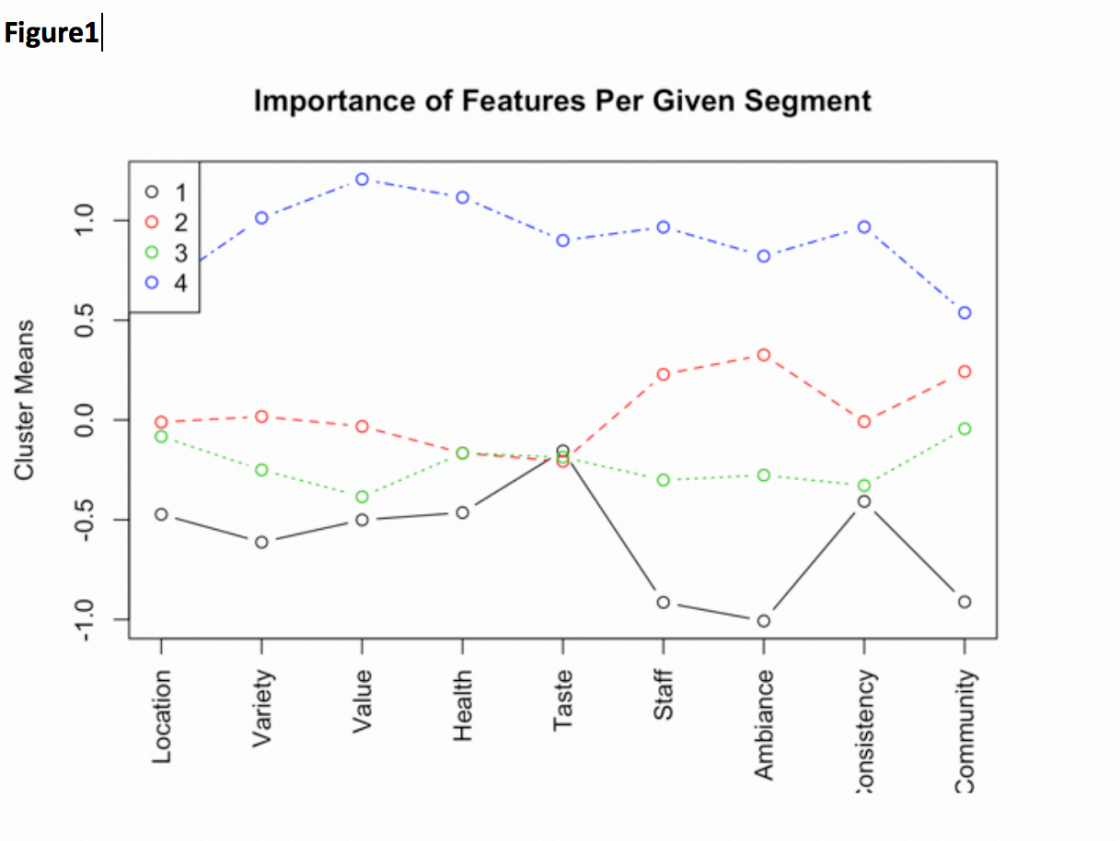


Number of clusters = 4 will stay as optimal number of clusters therefore it will become our 'customers segments'.

Following, Multivariable Plots were implemented to identify customer segments members:

Alicja Wilk
Assignment 2
BANA 271
Winter 2019

Figure1



Alicja Wilk
Assignment 2
BANA 271
Winter 2019

Figure2

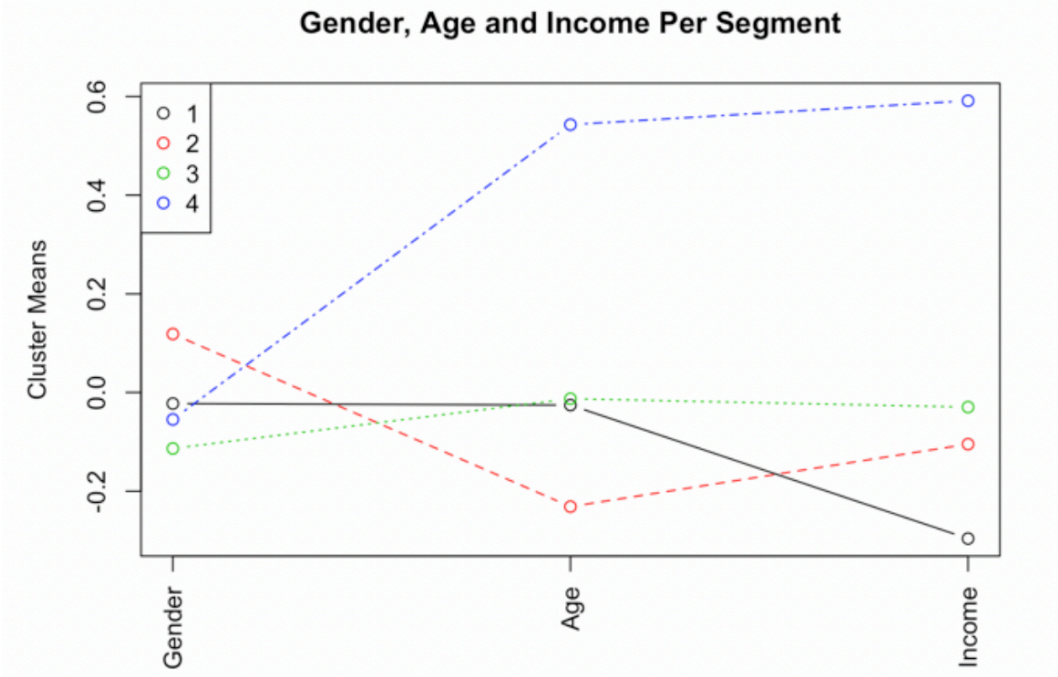


Figure3

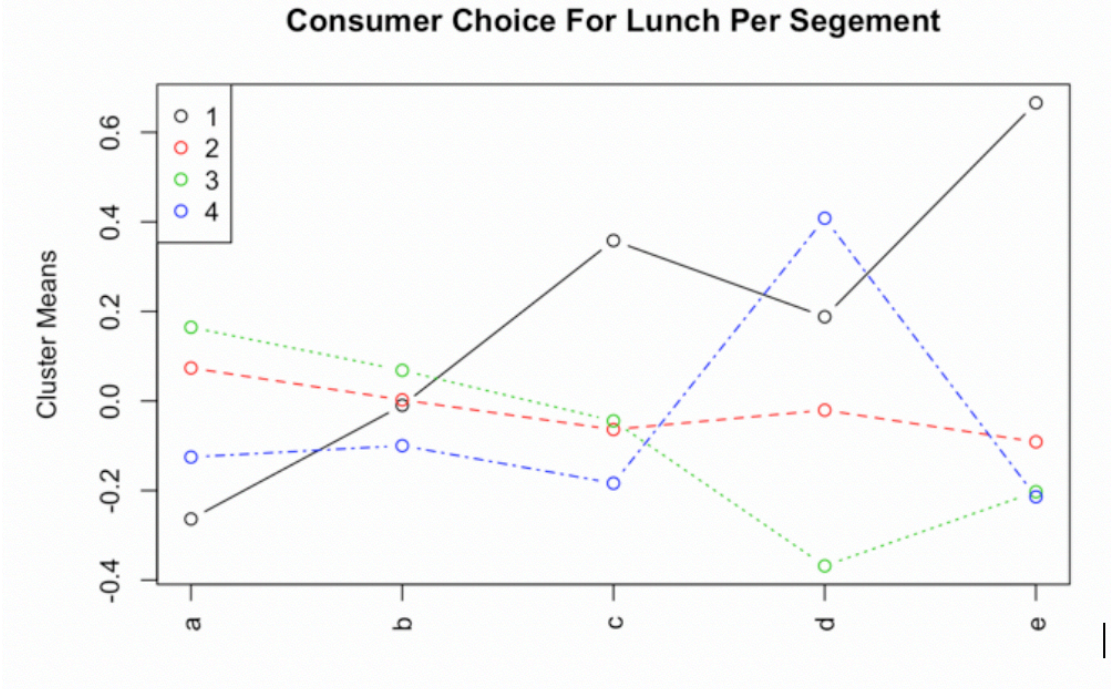


Figure4

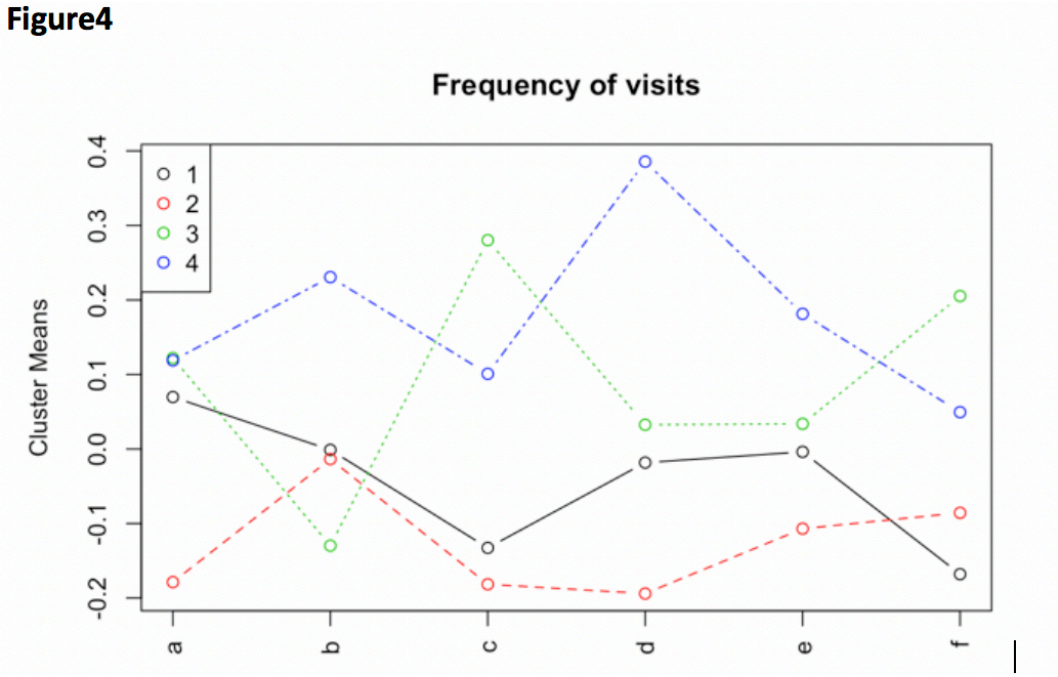


Figure5

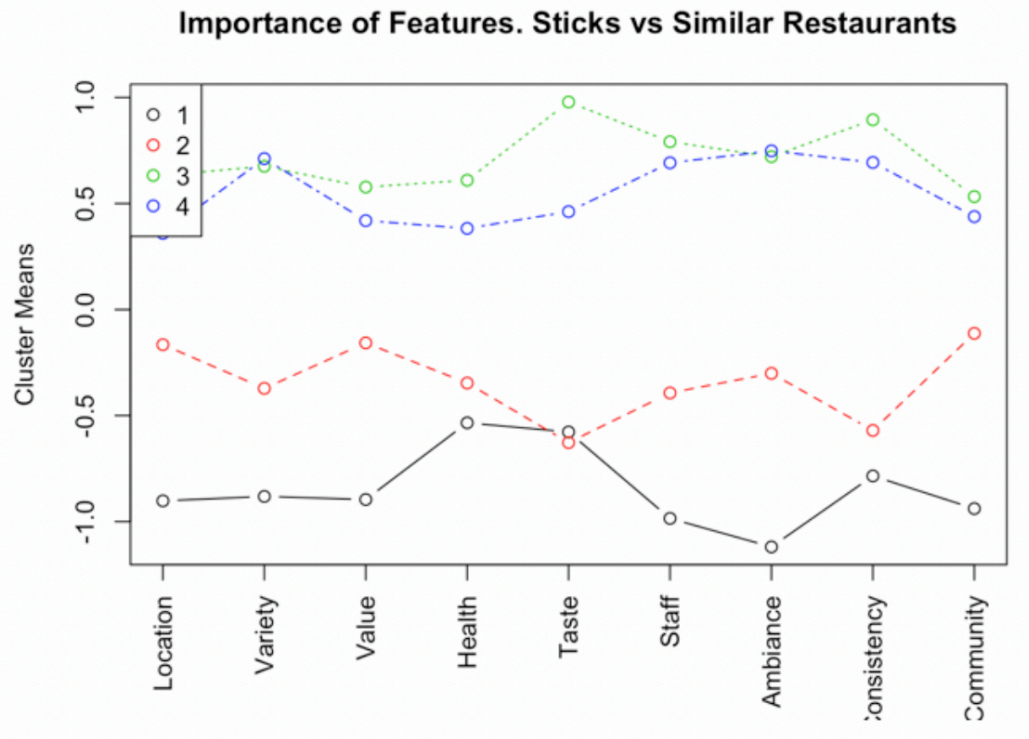
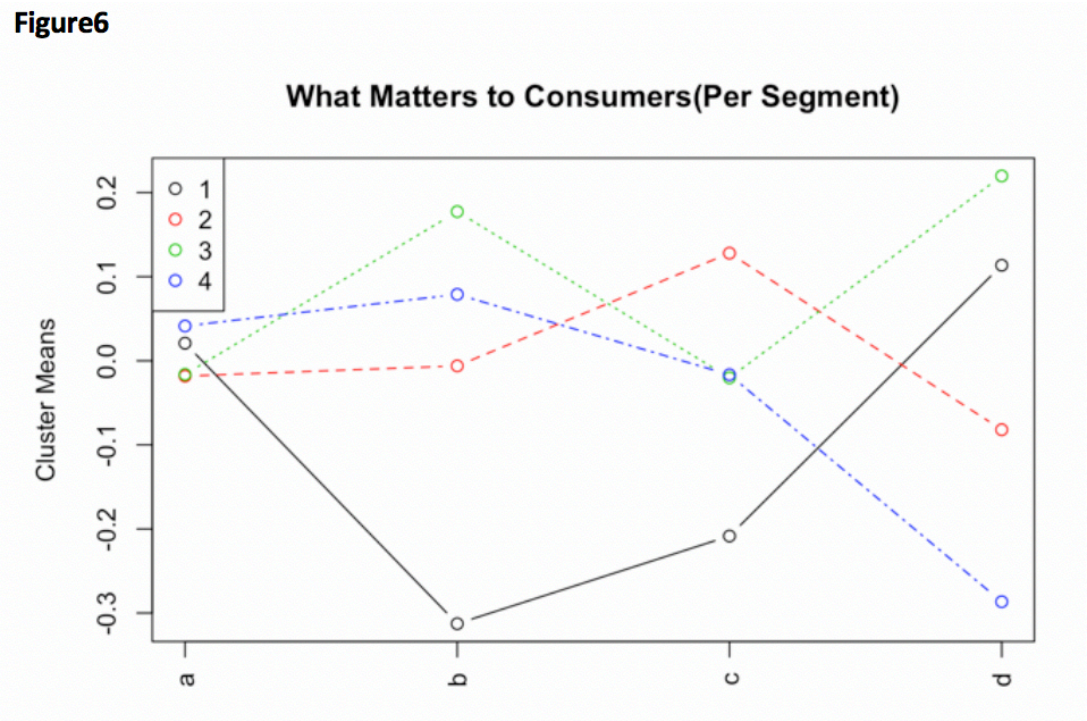


Figure6



Interpretation of Figures(plus assigned weights if applicable)

Figure 1: Importance of Features Per Given Segment
Segment 1: Staff and Ambiance(4)
Segment 2: Taste(3)
Segment 3: Value(2)
Segment 4: Location(2)

Figure 2: Gender, Age and Income Per Segment
Segment 1:younger female with small income(1)
Segment 2:younger male with lower income(2)
Segment 3: younger middle aged with steady income(4)
Segment 4: older female with high income(4)

Figure 3: Consumer Choice For Lunch Per Segment
Segment 1: buys lunch or skips it(3)
Segment 2: steady everything(1)
Segment 3: eats at home(2)
Segment 4: buys lunch out

Figure 4: Most frequent occasions
Segment 1: weekday lunch
Segment 2:weekday dinner(1)
Segment 3: weekend lunch(3)
Segment 4: weekend dinner(3)

Figure 5: Importance of Features. Sticks vs Similar Restaurants
Segment 1: Health and Taste(4)

Segment 2: Taste and Consistency(3)
Segment 3: not much importance(1)
Segment 4: not much importance(1)

Figure 6: What Matters to Consumers (Per Segment)
Segment 1:spending(3)
Segment 2:health benefits and spending(3)
Segment 3: local products purchase(1)
Segment 4:health benefits and local products(3)

Business Implications

Targeted Segment

Careful evaluation of criteria, weighted averages and possible fits in the customer segmentation pointed Segment 4 as the most attractive segment for Sticks to target.

Segment 4 does care about health benefits to its diet, visit restaurant often(even not during its prime hours) likes buying lunch out and is a female with high income and local.

Customer segment 4 fits for what the Stick executives are hoping for, attracting the local customer base, active sports moms and people with higher income. Especially that those variables were also statistically significant in our analysis we should consider a marketing strategy that builds on those factors and variables as foundation.

Future Location

Based on the choice of the targeted segment above, Location A is a perfect fit for Sticks Kebob. The Median Age is higher, which fits the current customer base. Median Income is highest, and like also noted above income for Sticks customers is 50K-100K with strong consumer based in the segment of 100K+ which indicates that more wealthy families look for a healthy quick supplement to their daily routine and do not mind spending extra.

Sticks managers should implement a plan for the new location that will carry the principles of the segmentation discoveries. Sticks should take advantage of their customer base income and the fact that people are aligned towards healthy meals. By promoting the menu to have options that will increase the marginal revenue of the location. For example, we could add a side vegan organic dishes that will have extra healthy ingredients add ins per customer's request(avocado, seeds, almond butter, smoothies etc). Incremental revenue would increase as well by shifting the menu even more toward the healthy side. Sticks managers should target through online and mobile advertisement since most of its customers will have a mobile device and also through coupons send to work or home(to increase local presence)

References:

1. Gibbs, Rajkumar VenkatesanShea. "Segmentation at Sticks Kebob Shop." *Harvard Business Review*, hbr.org/product/segmentation-at-sticks-kebob-shop/UV7031-PDF-ENG.