

Market Intelligence - Assignment 5

Section A, Team # 18

Arwen Wang, Kewei Jiang, Kieu Anh Nguyen, Taru Dharra and Vedant Sahay

Question 1:

Jae's currently has two alternative options: open a new location near the South Point Mall or open a new location near the UNC campus. Considering that these two locations are likely to have very different customer base, Jae's should open a new restaurant in a location that is the most likely to reach their current target customer base. In order to do this analysis, we are clustering Jae's customer base in its current location and cluster the respondents in the prospective locations to see which location's customer base matches the biggest segment in the current location.

Question 2:

First, we calculated the variance for each attribute, and we calculated the mean of all variances as our threshold, which is 1.66. Then we selected attributes with variance above the mean. Those variables are *Num eat out*(2.016436), *Imp healthy options*(2.980975), *Imp taste*(2.8675), *Num jae's last month*(4.732748), *Compare convenience*(2.154396), *Compare healthy options*(3.014747), *Compare taste*(3.269391), *Compare friendly staff*(2.619779).

Then we implemented factor analysis. We used the elbow method to get the 3 components. Based on the factor loading, we chose the top absolute value factors that as following: *Imp Value*, *Imp taste*, *Compare value*, *Compare taste*, *Imp convenience*, *Imp Menu Variety*, *Compare menu variety*, *Num jaes last*, *Compare convenience*, *Imp friendly staff*, *Compare friendly staff*, *Compare healthy options*, *Imp healthy options*.

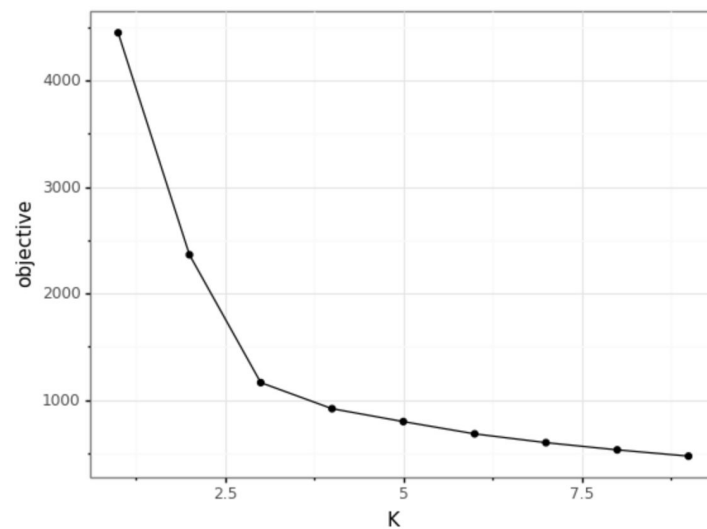
Question 3:

We input variables selected from post variance analysis into K-means, and used the 3 centers that was computed through elbow method. We run the analysis in the biggest cluster, but we didn't get strong results. So, we did not choose to go ahead with this method.

Then we run the K-means with the variables that we selected from factor analysis and used 3 centers that computed by the elbow method. We got the result of three segmentations of Jae's current customers. We defined those segmentations as "Young Foodies", "Healthy Nuts", and "Middle of pack". The "Young Foodies" is our biggest segment.

Question 4:

To cluster our current customers into different segments, we used the K-means clustering algorithm. According to the result of the elbow method, we identified the optimal number of clusters is 3 in this case.



Based on the K-mean clustering algorithm, we performed both the Variance analysis and the Factor Analysis methods to identify the differences of the customers in different clusters. For the variance analysis method, we used the mean of the variance of each variable as the threshold to select out those columns with a variance larger than the mean.

However, between those two methods, we still chose to use Factor Analysis. The result of variance analysis gave us the biggest segment with no strong preference of any characteristics of the restaurant, and customers in this segment have relatively low eating out frequency among the three clusters. On the contrary, the largest cluster based on the Factor Analysis indicates a strong preference toward taste, convenience, and value. Meanwhile, customers in this cluster have a higher eating frequency in the current Jae's restaurant compared to the other two clusters. Thus, we believe the result of the Factor Analysis is more representative.

Question 5:

To analyze the Non Customer Data, we performed two analyses:

- **K- Means on Post Variance analysis:** We found the variance of each attribute and in order to decide which columns would go into the K-Means algorithm, we found the average variance of all attributes, set the mean as our threshold and included those columns in K-Means which are above this threshold.
- **K-Means on Post Factor Analysis:** We performed the Factor Analysis on the dataset and on the basis of the output of this analysis, we decided on which columns would be a part of the K-Means algorithm.

(1) K-means on Post Variance Analysis on Non Customer Data:

The breakdown table of Cluster sizes:

Cluster Name	Cluster Size
Cluster 0	44
Cluster 1	45
Cluster 2	33

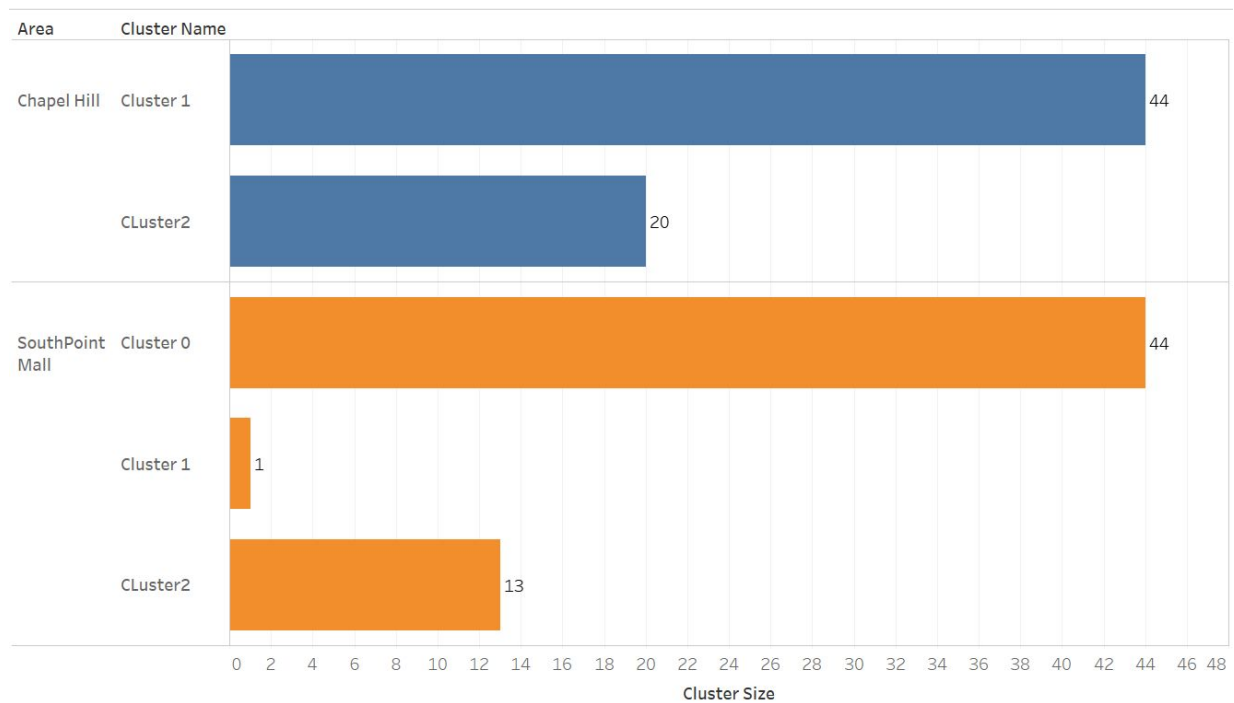
We see that Cluster 1 is the largest cluster and below are the attributes associated with it:

Attribute	Value	Compared to other clusters	Label
Imp taste	5.02	High	Young Foodies
Imp value	4.69	High	
#times/wk eat lunch/dinner out	4.64	High	
Imp convenient	4.33	High	
Imp ambience	2.02	Medium	
Imp friendly staff	1.98	Medium	
Imp healthy options	1.33	Low	

According to the table above, we can see that the population within Cluster 1 can be labelled as '**Young Foodies**' due to the high value attributes such as 'Important Taste', 'Important Value', 'Frequency of Dining', etc. And some of the characteristics of the restaurant that customers cares are good value, convenient location, decent ambience.

Also, the same attributes were compared within Cluster 0 and Cluster 1 (the elbow method indicated 3 optimal clusters) and we rated these attributes on the basis of the attribute values. For example, let's consider the attribute 'Imp taste', amongst all the clusters, it had the highest attribute value in Cluster 1 and hence labelled as 'High' when compared to other clusters, whereas if we consider the attribute 'Imp healthy options', it's attribute value was the lowest in Cluster 1 when compared to Clusters 0 and 2 and hence labelled as 'Low' when compared to other clusters.

Also, in terms of determining the location, we conclude that Chapel Hill would be a better location for Jae to open the second restaurant as reinforced by the bar graph below.



(2)K-means on Post Factor Analysis on Non Customer Data:

The breakdown of Cluster sizes is as follows -

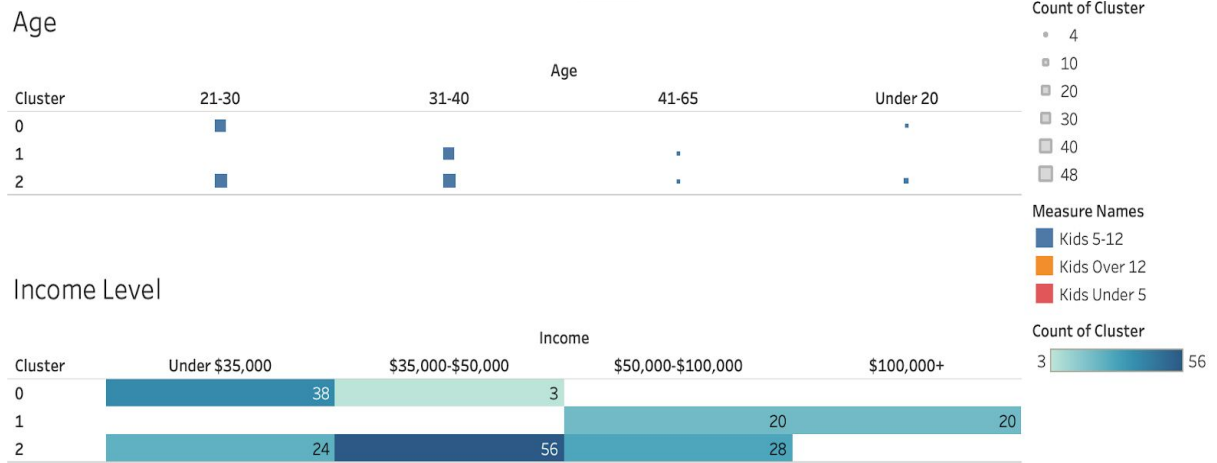
Cluster Name	Cluster Size
Cluster 0	33
Cluster 1	45
Cluster 2	44

We see that Cluster 1 again is the largest cluster and below are the attributes associated with it:

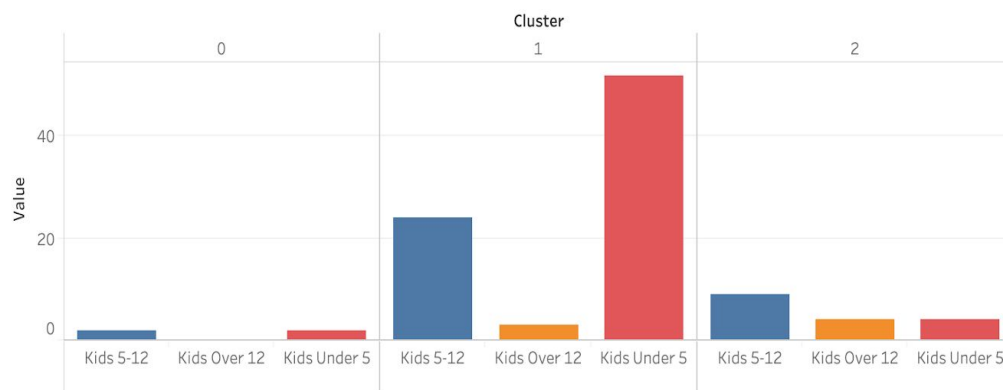
Attribute	Value	Compared to other clusters	Label
Imp taste	5.02	High	Young Foodies
Imp value	4.69	High	
#times/wk eat lunch/dinner out	4.64	High	
Imp convenient	4.33	High	
Imp menu variety	3.68	High	
Imp ambience	2.02	Medium	
Imp friendly staff	1.98	Medium	
Imp healthy options	1.33	Low	

We see that an additional attribute - '*Imp menu variety*' shows up as a 'High' attribute which was not present in the Post Variance analysis however with the inclusion of this metric only reiterates the same point that was made in the Post Variance Analysis.

Question 6:



of Kids

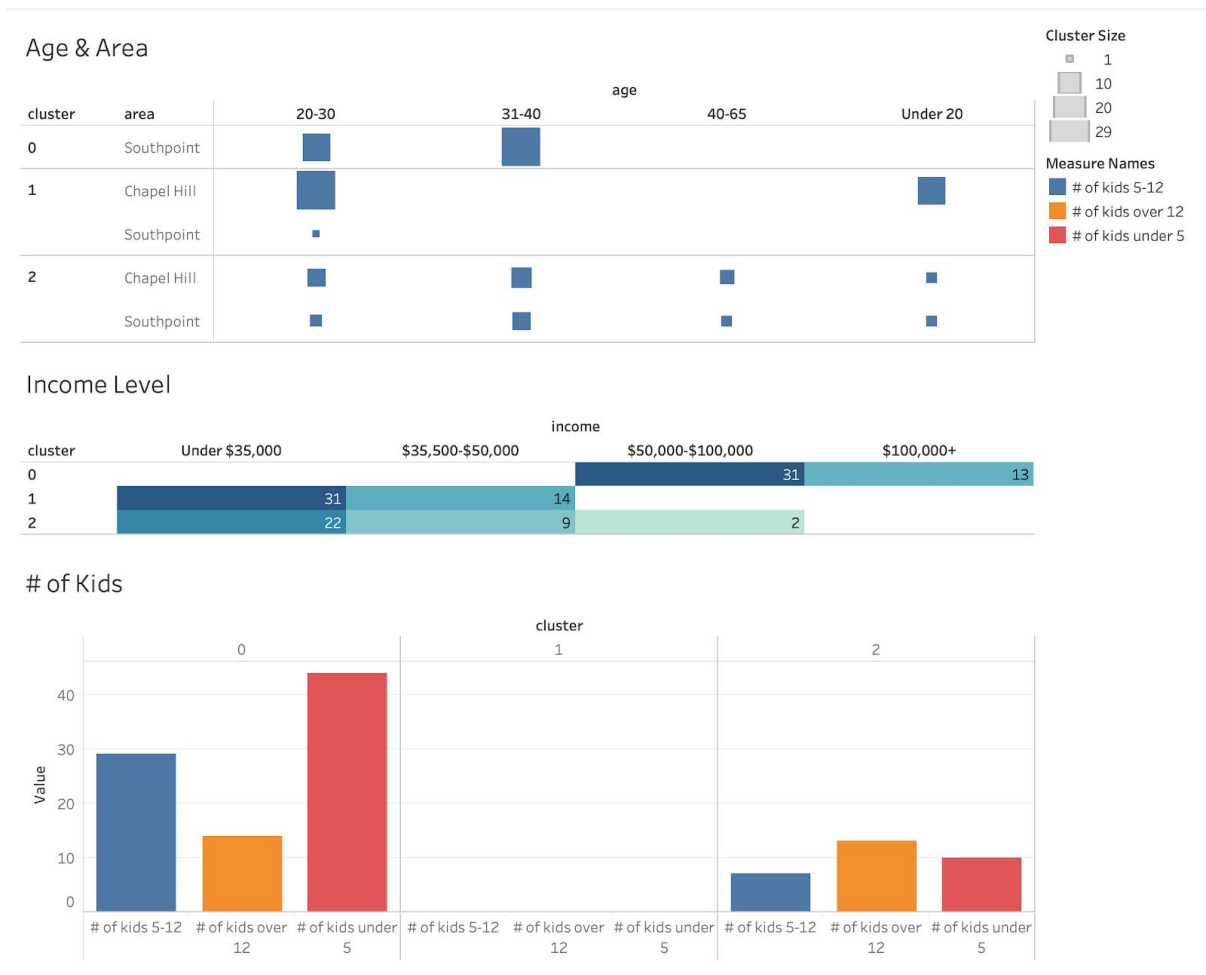


Dashboard 1: Current customers

Cluster 0: Young customers, lower income, barely have children

Cluster 1: Mid age customers, relatively high income, large number of children

Cluster 2: All ages, low to middle income, small number of children



Dashboard 2-Non customers

Cluster 0: Middle age, High income, large number of kids

Cluster 1: Young age, low income, no kids

Cluster 2: All ages, low income, small number of kids

Comparing clusters between Jae's current customers and non customers, we see that demographic characteristics of clusters within each are quite similar. They both have clusters the people are at a young age, have low income, and have no/very few kids. They both have clusters of people who are in middle age, with high income and have a large number of kids. The cluster that people are at all ages, have low income and a small number of kids is also present in both of the results.

Question 7:

- (1) As per both the analysis performed above, we would suggest Jae to pay attention to attributes in the new restaurant.
 - (a) Good tasting food.
 - (b) Value for money.
 - (c) Expect customers to dine out frequently, hence an opportunity to hook on the loyal customer.
 - (d) Convenience of the restaurant in terms of proximity, location, etc.

- (e) Variety of menu would attract customers.
- (2) As per K-means, we recommend Jae to open a new restaurant in Chapel Hill to attract a larger number of customers so as to maximize profitability.

Question 8:

When establishing a new location, it is important for Jae's to move to an area where potential customers are of the same demographic as well as psychographic as Jae's current customers. In Jae's current location, their biggest customer segment is young people under 20 years old or 20 to 30 years old who want a delicious, convenient and affordable meal. According to the survey, the customer base in Chapel Hill is a great match since it is filled with college students who are also looking for value, taste and convenience. While customers in Chapel Hill has less disposable income and are less likely to have children, this area is still a good match because they are looking for the same thing as Jae's established brand. Opening a new location in Chapel Hill will allow Jae's to focus on their strengths of providing convenient, great tasting and affordable meals, as well as maintaining a consistent brand and quality of service in the two locations.