Marketing Proposal

M. He   C. Suarez   S. Zhu   M. Santos

Menu

The Company

""We decided long ago that we didn't want Chipotle's success to be tied to the exploitation of animals, farmers, or the environment, but the engagement of our customers." -Steve Ells

Mission statement:

"To provide food with Integrity."

Vision statement:

"We believe that food has the power to change the world"

Values:

Improving health

Reasonable prices

Food with integrity

Where is Chipolte Mexican Grill (CMG) headed?

"A goal without a plan is just a wish." — Antoine de Saint-Exupéry, writer and pioneering aviator
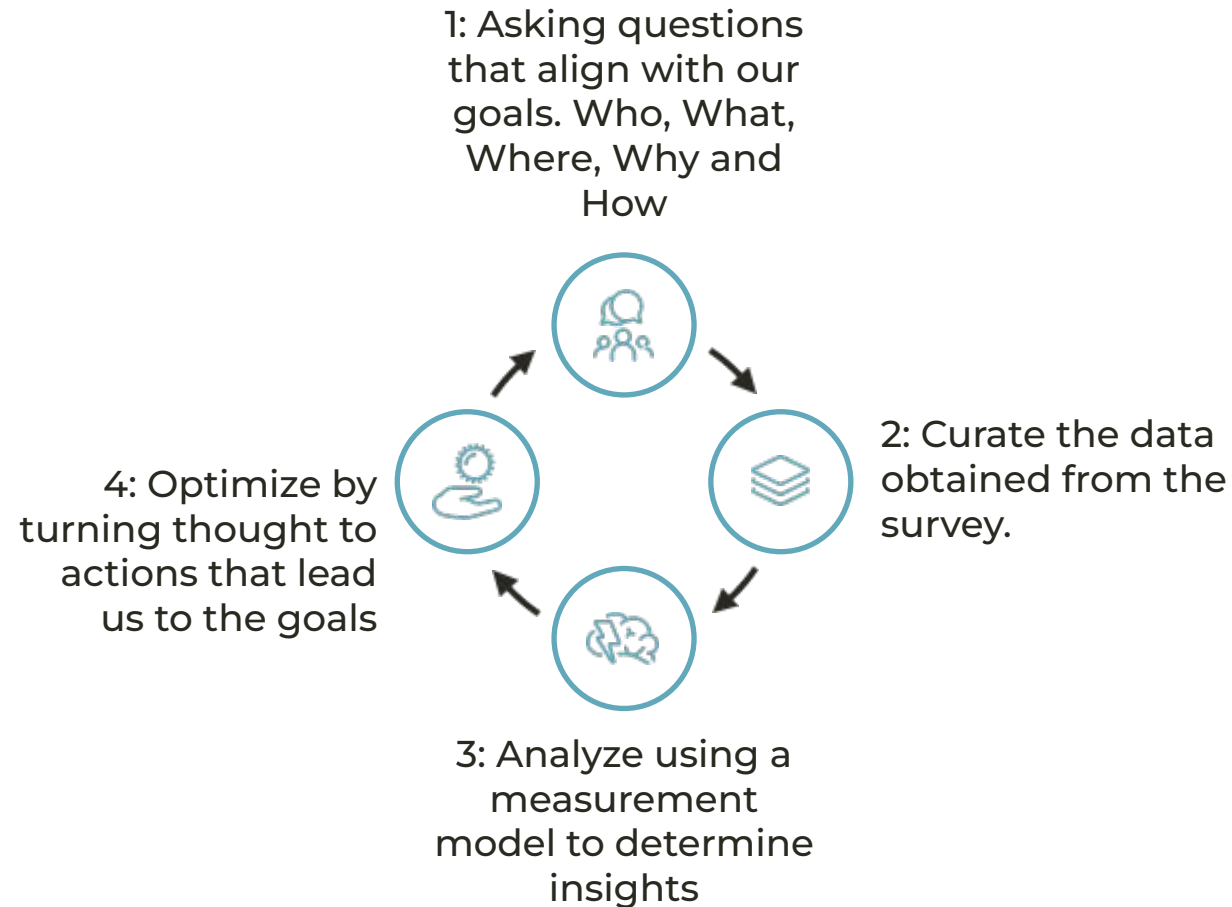
Goal 1: Retention

Goal 2: Acquisition

Increase customer retention by 10 percent over the previous quarter by the end of this quarter.

Increase customer acquisition by 7 percent over the previous year by the end of this year.

Decision-making Framework

A proven process to lead us to the goals

1: Asking questions that align with our goals. Who, What, Where, Why and How

2: Curate the data obtained from the survey.

3: Analyze using a measurement model to determine insights

4: Optimize by turning thought to actions that lead us to the goals

- **Specific**
  Our objective is clear

- **Measurable**
  We have a clear target and KPI(s) all actions can be tracked and validated

- **Achievable**
  Chipolte has a proven business model and resources to achieve our goal

- **Time-bound**
  We have 90 days to get there

Feature Engineering and Data Preparation

The original data set has 362 entries with 25 features (24 numerical Dtype and 1 String)

Labeling and typos

The categorical variable top1 were inconsistently labeled and with typos.

We Converted all strings to lowercase and removing any non alphabetic characters and spaces. Then, Mapped through rows and correct all misspelled words using wildcards and regex.

Missing variables

There were 504 missing data points from the original survey.

We removed all the rows that are missing more than 75% survey (9 rows removed).

Then filled all null data with the medium data result of that column.

Range

Some of the data entries from the survey exceeded the set range of 1 thru 5.

We looped through the columns to replace those entries with a 5 response.

Data type conversion

All categorical data was converted to numerical data so out KMeans model could process using pd.get_dummies.
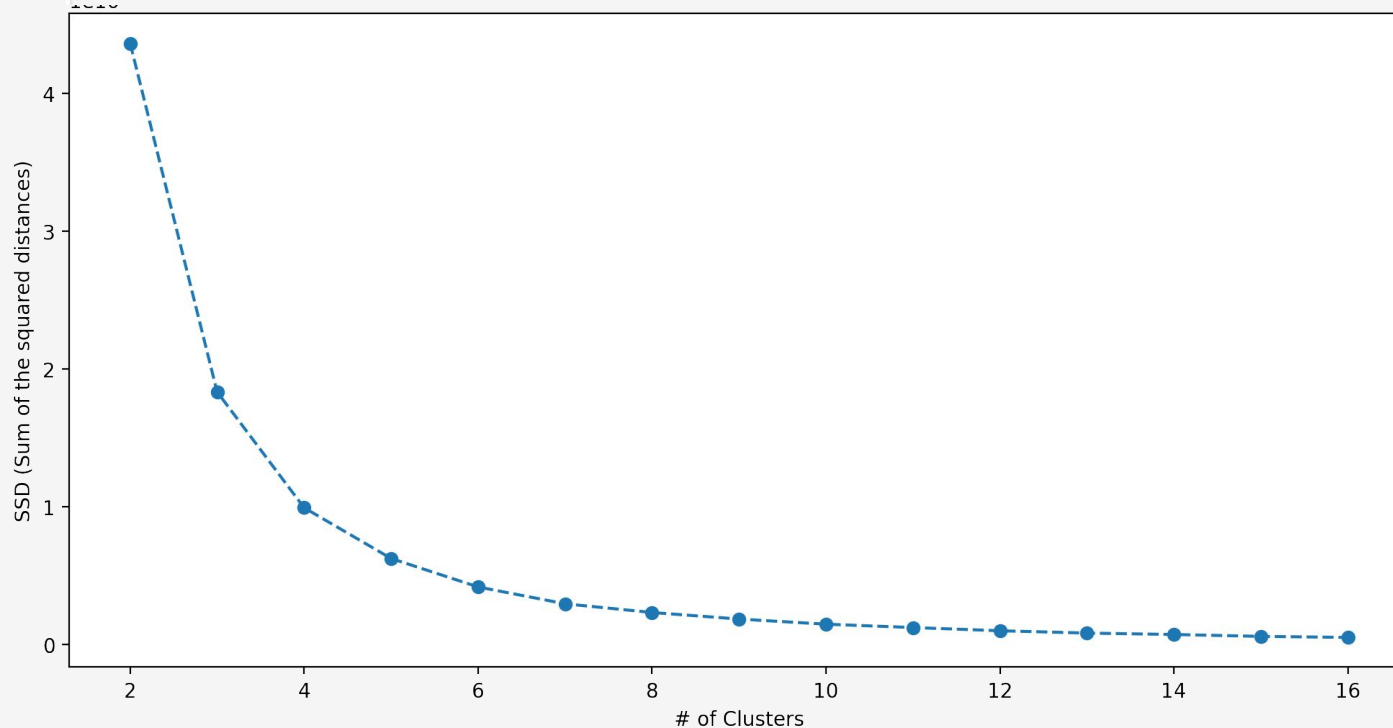
Clustering

KMeans

Elbow Test SSD

```
ssd = []
for i in range (2,17):
    model = KMeans(n_clusters= i)
    model.fit(df)

    ssd.append(model.inertia_)   ## SSD point --> cluster center
```


To choose the optimal number of clusters for the KMeans model we used the elbow/knee method. This method is used to compare and plot the differences
between the between the Sum of the Squared Distances (SSD) as # of clusters increases

we want each cluster to have at least 20 rows the maximum cluster we will loop is 353 / 20 = 17

When we increase the number of clusters.The SSD drop flattens after 6 clusters. Each cluster has 20 people and we can see how its correlated to each feature.

```
df['Cluster'].value_counts()

5    90
0    85
2    69
3    56
1    29
4    24
Name: Cluster, dtype: int64
```
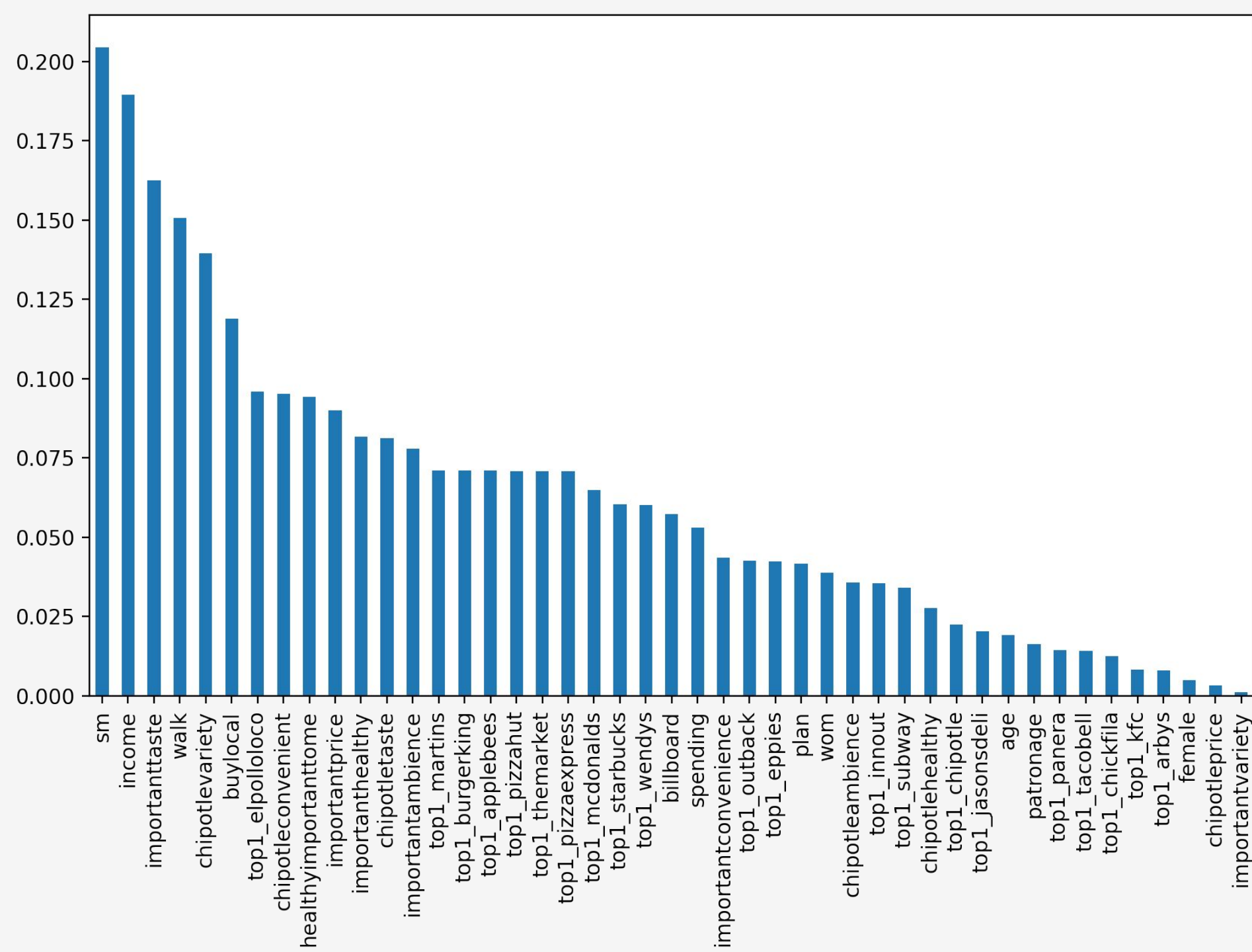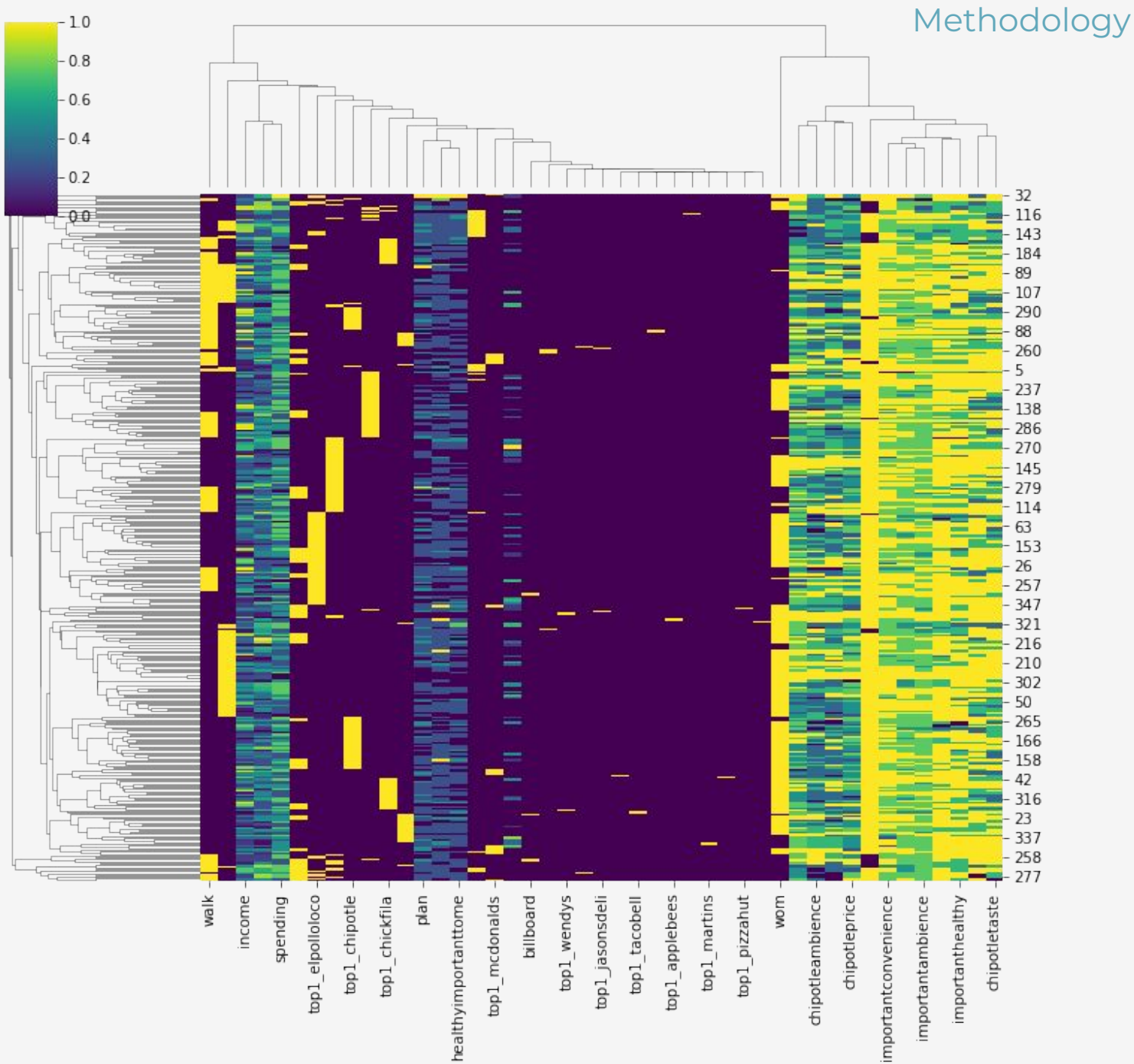
Cluster Feature Correlation

Shows how correlated each feature is to it's clusters.

How our KMeans model weight each features when clustering / doing segmentation

Methodology

Hierarchical Clustering

Hierarchical clustering using the Euclidean Distance similarity metrics.

Scale our data frame with minMaxScaler

AgglomerativeClustering using distance_threshold = 7 calculated from the hierarchy linkage matrix.

The model also suggests 6 clusters and are 72% similar with KMeans cluster results.

T-test & ANOVA

Scipy & Pingouin

```
#### Testing the assumptions
a = df[(df['Cluster'] == 0)]
b = df[(df['Cluster'] == 1)]

# homogeneity
stats.levene(a['income'], b['income'])    # The test is not significant (small p-v
# Shapiro-Wilk test for normality
stats.shapiro(a['income'])                # Some Test normality were significant,
stats.shapiro(b['income'])
```

```
LeveneResult(statistic=1.8645414140570864, pvalue=0.17483777599665506)

ShapiroResult(statistic=0.950531542301178, pvalue=0.0025769586209207773)

ShapiroResult(statistic=0.9222952127456665, pvalue=0.034873563796281815)
```

```
### Thus we can use parametric instead of non-parametric ttest

stats.ttest_ind(a['income'], b['income'])
```

```
Ttest_indResult(statistic=-60.153285029722106, pvalue=4.174981314753616e-87)
```

The null hypothesis corresponds to the hypothesis that there is no relationship or difference between our Clusters.

Levene and Shapiro-Wilk were performed to test the homogeneity and normality and determine if each cluster and dependent variables are parametric or nonparametric. (Independence of observations, Normality of data and Homogeneity of variance)

For pairs that are parametric we use stats.ttest_ind to perform T-test. And stats.wilcoxon for non parametric pairs.

T-test & ANOVA Cont.

```python
import pingouin as pg

# Run the ANOVA
aov = pg.anova(data=df, dv='income', between='Cluster', detailed=True)
aov

# P-value ≤ α: The differences between some of the means are statistically significant
```

	Source	SS	DF	MS	F	p-unc	np2
0	Cluster	1.492598e+11	5	2.985196e+10	2469.227167	9.999804e-269	0.972662
1	Within	4.195089e+09	347	1.208959e+07	NaN	NaN	NaN

```python
#  As we can see we have a p-value below the threshold, so there is a significant difference between the various Clusters!
pg.pairwise_tests(data=df, dv='income', between='Cluster')
```

	Contrast	A	B	Paired	Parametric	T	dof	alternative	p-unc	BF10	hedges
0	Cluster	0	1	False	True	-67.823840	61.583894	two-sided	1.405782e-59	3.272e+88	-12.849279
1	Cluster	0	2	False	True	42.643649	150.490950	two-sided	6.319479e-86	2.539e+82	6.797594
2	Cluster	0	3	False	True	-21.597707	103.983241	two-sided	3.165259e-40	3.73e+42	-3.826073
3	Cluster	0	4	False	True	-35.807844	36.022080	two-sided	9.279218e-30	2.535e+57	-8.382757
4	Cluster	0	5	False	True	26.715983	161.953320	two-sided	3.071858e-61	2.867e+59	4.046949
5	Cluster	1	2	False	True	104.026821	59.916574	two-sided	2.150313e-69	1.217e+96	21.600869
6	Cluster	1	3	False	True	38.011137	77.920164	two-sided	4.666719e-52	2.32e+50	7.595633
7	Cluster	1	4	False	True	14.755692	42.118642	two-sided	3.030958e-18	5.482e+16	4.118857
8	Cluster	1	5	False	True	94.827397	49.000270	two-sided	3.449330e-57	1.142e+108	19.724739
9	Cluster	2	3	False	True	-55.746346	100.576349	two-sided	1.925541e-77	1.018e+85	-10.251660
10	Cluster	2	4	False	True	-63.336467	35.714543	two-sided	2.753550e-38	7.275e+72	-15.973730
11	Cluster	2	5	False	True	-20.557348	138.649480	two-sided	6.222475e-44	5.227e+42	-3.316446
12	Cluster	3	4	False	True	-16.561585	49.401128	two-sided	8.357269e-22	6.125e+23	-3.794338
13	Cluster	3	5	False	True	43.585643	87.423814	two-sided	4.365045e-61	5.828e+80	8.028292
14	Cluster	4	5	False	True	53.490172	30.951295	two-sided	4.569564e-32	2.471e+77	14.043169

pingouin.anova
&
pingouin.pairwise_tests()
to display more than 2 cluster groups.

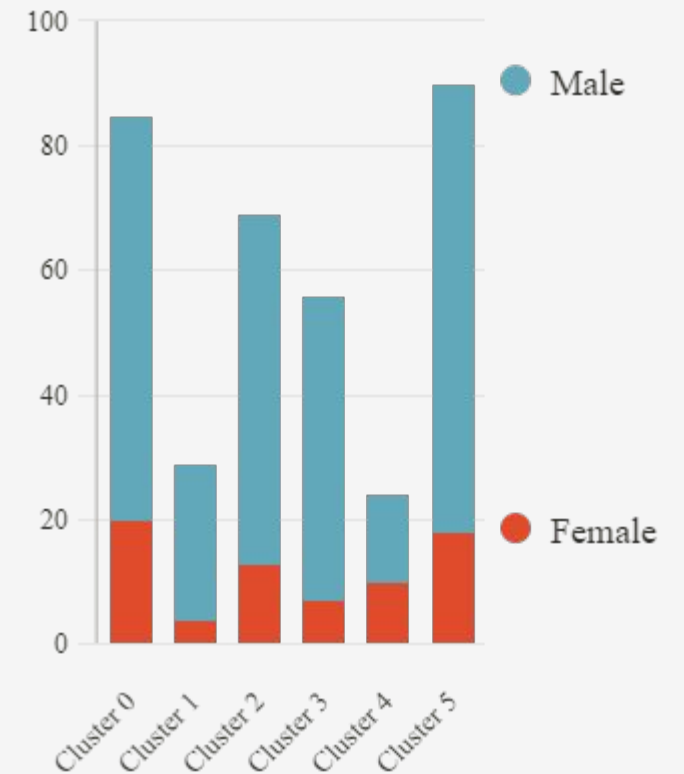Segmentation & Mean Analysis

Psychographics

Figuring out attitudes and beliefs

Planning

Buying local
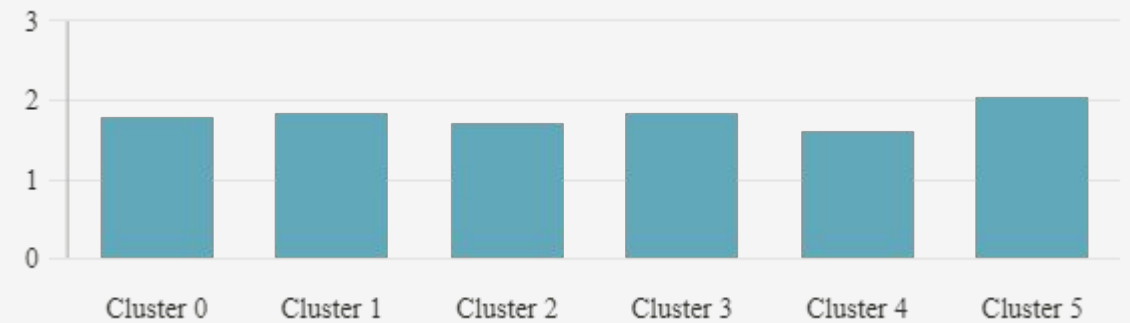
Trouble Controlling Spending

Healthy eating
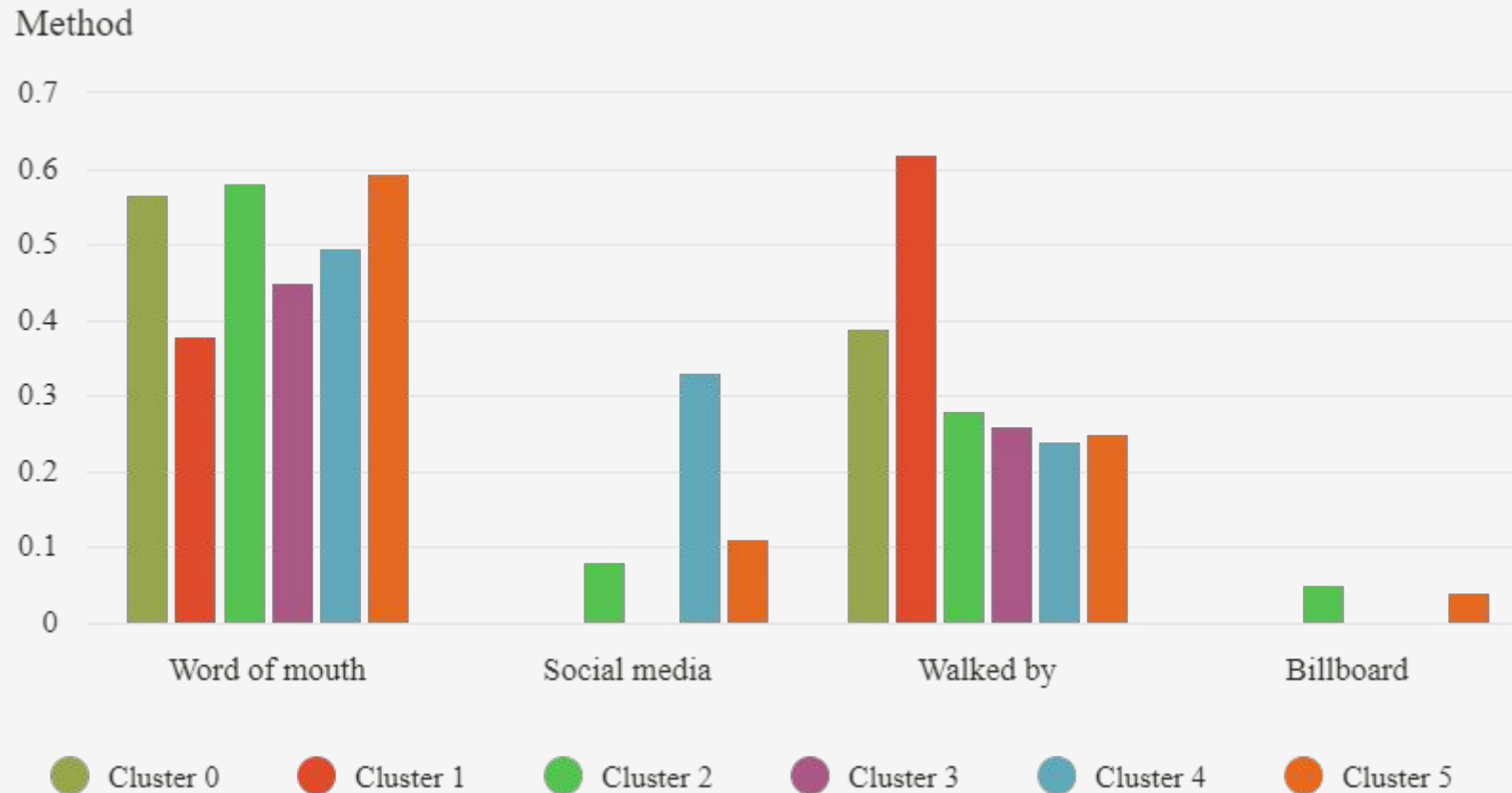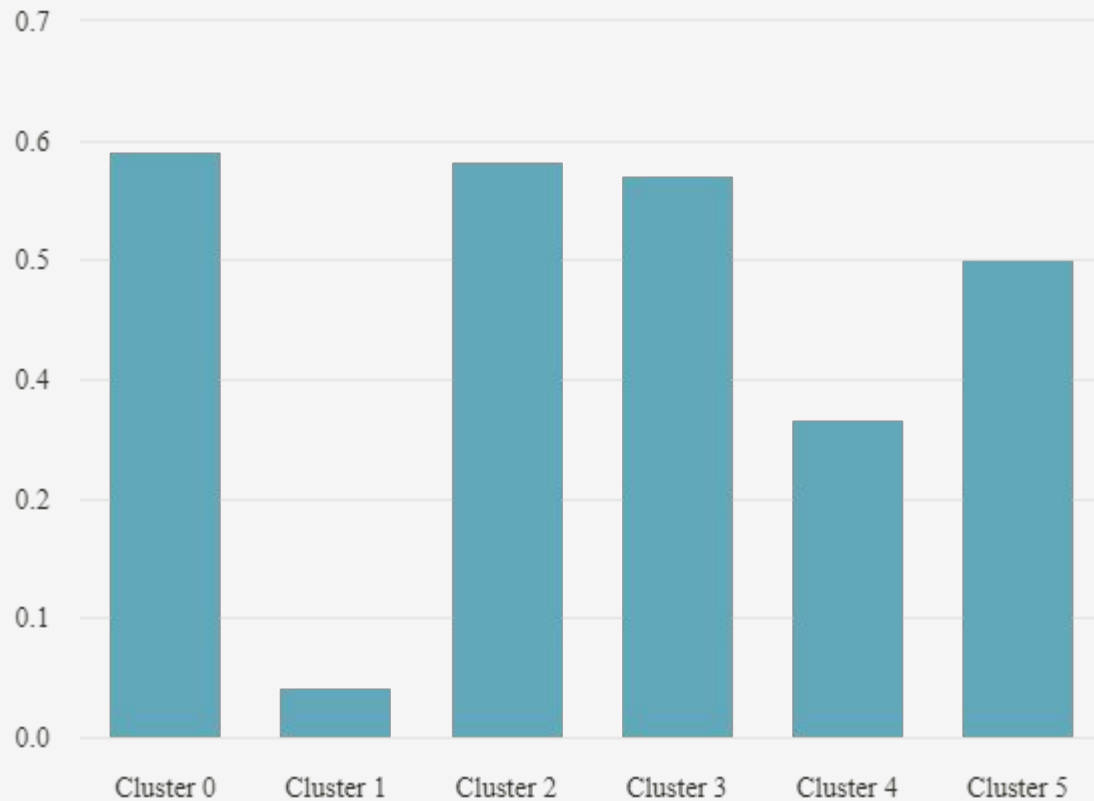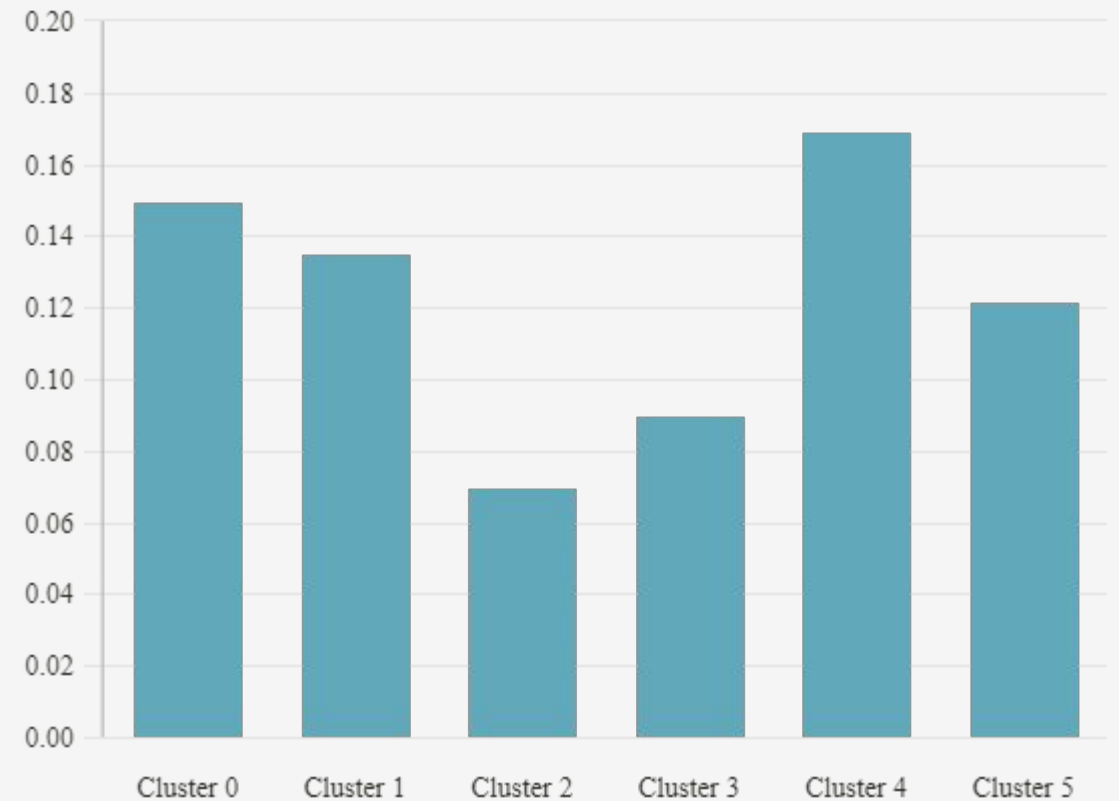
Loyalty and Brand perception (Behavioral)

How often do they dine with us & how do we rank in their eyes?

Patronage

Chipotle ranked as #1

Who are our Customers?

Profiles based on clustering results

Cluster 0 (The foodie):

Demographic Segmentation:

Average age: 40.3 years old (Relatively Young)

Average income: $40,417.29 (3rd Lowest)

Psychographic and Behavioral Segmentation:

Values restaurant variety, price, and taste.

Does not care about eating healthy and has trouble controlling spending.

Patronage and Loyalty:

Relatively pretty loyal when it comes to eating and dining at Chipotle.

Who are our Customers?

Profiles based on clustering results

Cluster 1 (The Chipotle Lover):

Demographic Segmentation:

Average age: 43.4 years old (2nd Oldest)

Average income: $85,210.03 (Highest)

Psychographic and Behavioral Segmentation:

Does not care about restaurant convenience and ambience but values taste.

Does not like Chipotle's performance in variety, healthiness, and ambience.

Plans things carefully, has trouble controlling spending and also does not value eating healthy.

Patronage and Loyalty:

Most loyal customers. They don't eat out a lot but when they do choose Chipotle a lot.

Who are our Customers?

Profiles based on clustering results

Cluster 2 (The Indecisive Eater):

Demographic Segmentation:

Average age: 34.4 years old (Youngest)

Average income: $16,647.03 (Lowest)

Psychographic and Behavioral Segmentation:

Values restaurant convenience and healthiness the most.

Enjoys Chipotle's ambience the highest among the clusters.

Does not plan things carefully and does not have trouble controlling spending.

Patronage and Loyalty:

Least loyal customers. They eat Chipotle a lot but they do not choose to go to Chipotle the most.

Who are our Customers?

Profiles based on clustering results

Cluster 3 (The Picky Eater):

Demographic Segmentation:

Average age: 44.6 years old (Oldest)

Average income: $55,495.70 (3rd Highest)

Psychographic and Behavioral Segmentation:
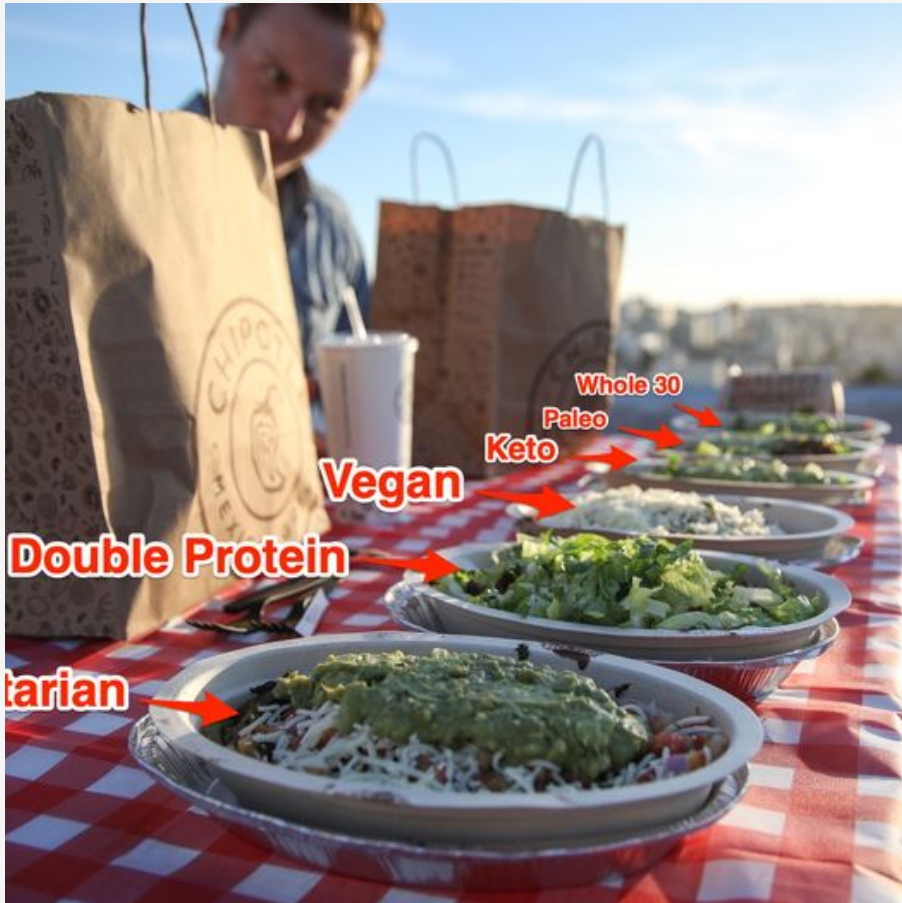
Values restaurant ambience the most.

Does not like the taste of Chipotle's food.

Patronage and Loyalty:

Fairly unloyal customers.

Who are our Customers?

Profiles based on clustering results

Cluster 4 (The Taste Tester):

Demographics:

Average age: 43.3 years old (Relatively Old)

Average income: $71,420.50 (2nd Highest)

Psychographic and Behavioral Segmentation:

Does not care for restaurant variety, price, and healthiness but cares about taste.

Is not satisfied with Chipotle's convenience and price but does enjoy its healthiness and taste.

Has trouble controlling spending and also does not like to buy from local.

Patronage and Loyalty:

Fairly loyal. They eat Chipotle and choose to dine at Chipotle.

Who are our Customers?

Profiles based on clustering results

Cluster 5 (The Penny Pincher):

Demographics:

Average age: 39.7 years old (2nd Youngest)

Average income: $26,963.40 (2nd Lowest)

Psychographic and Behavioral Segmentation:

Does not care about restaurant taste but does care about price.

Values Chipotle's performance in convenience, variety, and price.

Likes to buy from local and eating healthy.

Patronage and Loyalty:

Somewhat loyal.

Targeting

Based on the Statistical techniques of cluster analysis, we strongly recommend that the target market should focus on Cluster #1, Cluster #2 ,Cluster #4.

Cluster (#2): Our less loyal customers. They are young, frequent-diners, but not at chipotle. They are reachable through increasing our product variety. Segment Attractiveness: 20% Market Share, longer customer life cycle, social media savy. Strong potential to turn into brand promoters .
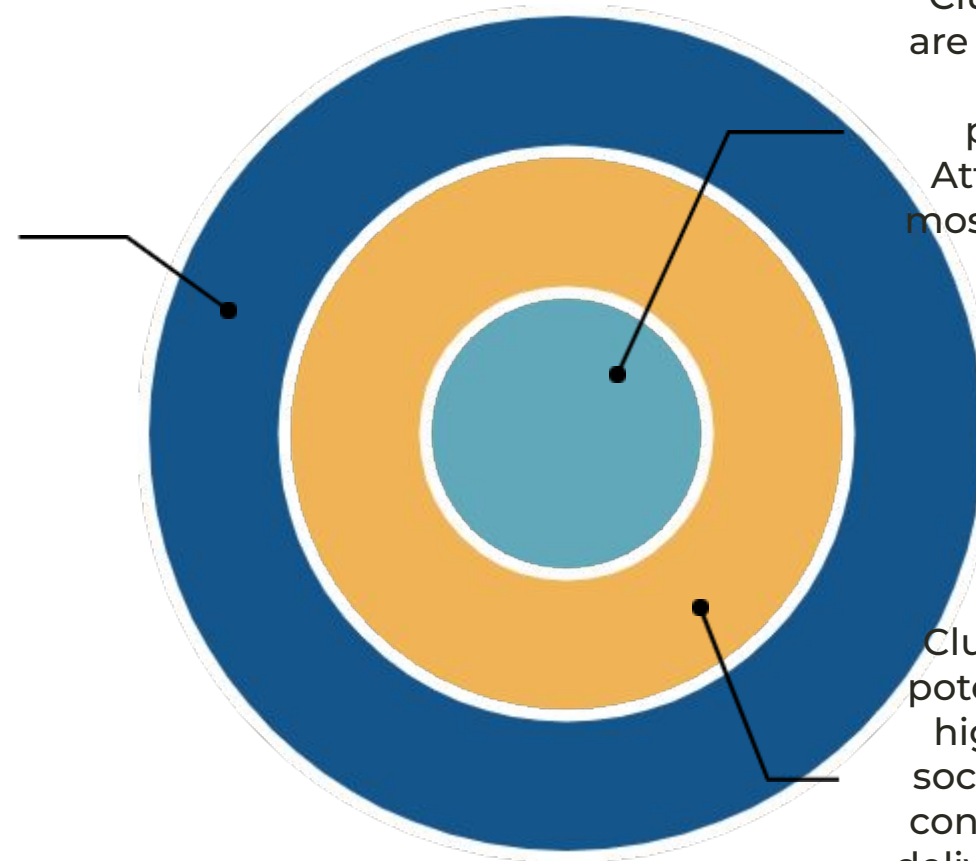
Cluster (#1) are our most loyal customers. They are likely to have more disposable income , price sensitive and reachable through in store promotion (Discount / Coupon). Segment Attractiveness: Highest income group and our most loyal segment. Goal is to increase customer retention rate.

Cluster (#4) are frequent customers with great potential to turn into loyal customers. They are a high income group with high engagement in social Media. They are reachable through better convenient dining experience (Online ordering / delivering) Segment Attractiveness: High income, High potential growth rate.
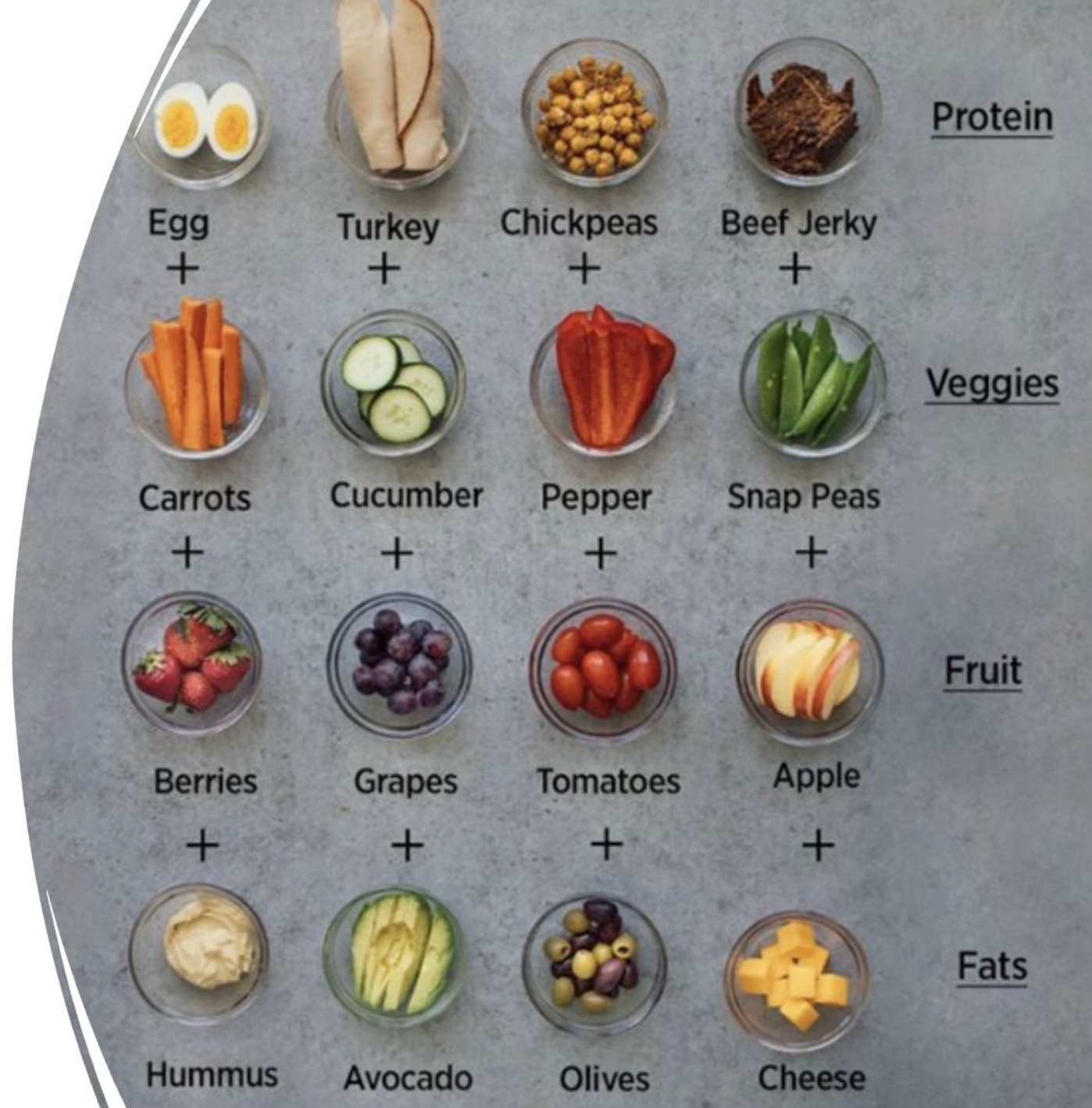
Positioning : Cluster (#2)

Product:

Increase our product variety by introducing new products/menus

1. Nutritionally balanced menu

2. Muscle-building menu

3. Vegetarian Menu

4. Fusion menu

Fasting/KETO/Clean Eating

Positioning : Cluster (#2)

Place:

They like the pleasant ambience and convenient locations of Chipotle. Possible enhancement by updating to modern furniture or design because this segment is the youngest.
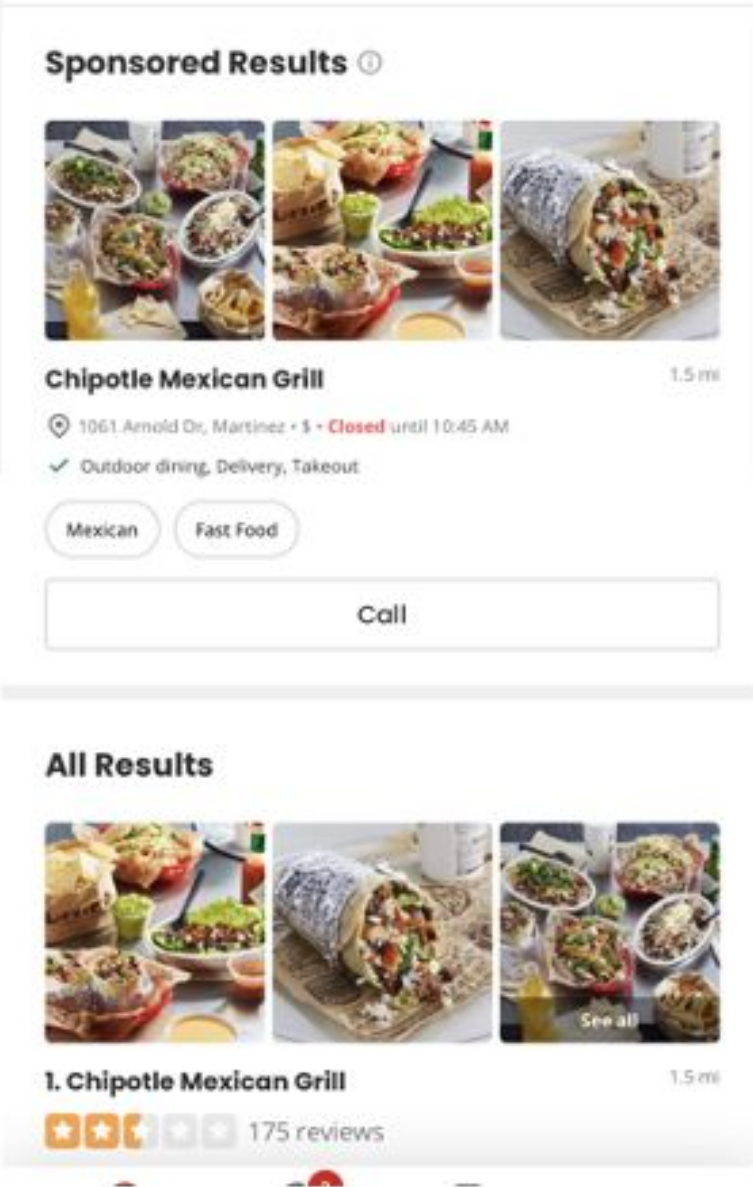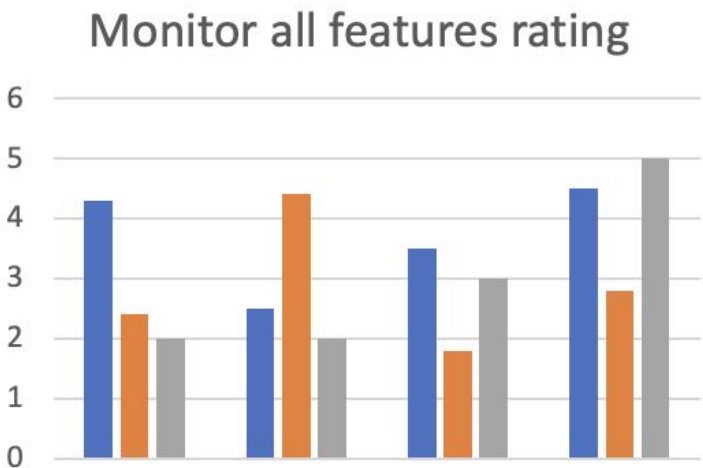
Positioning : Cluster (#2)

Price:

They are satisfied by the price of Chipotle products. We will have more room to experiment pricing strategies especially with new products that they like.

Promotion:

Ad campaigns on social media (Facebook, youtube and Instagram)

Paid ad campaigns on restaurant ranking/review sites such as Yelp. (Frequent diner)

Monitor all features rating

**Positioning :
Cluster (#1)**
Product:
Increase our product variety by introducing new products/menu
Place:
They like the pleasant ambience and convenient locations of Chipotle. Possible enhancement by increasing exterior decoration to attract walk by customers.

Positioning : Cluster (#1)

Price:

They are price sensitive. Offer coupons and value meal deals to reward frequent visits and customer loyalty

Promotion:

Decrease targeted Social media ads for this segment since no traffic was brought by social media. Instead, spend the budget on local/in-store events (Themed/holiday events/ raffle, etc.)

Positioning :

Cluster (#4)Product:

They like our products. Because they value health and taste the most. We should have more healthy choices on our menu.

Place:

They do not like the convenience of Chipotle's location. We should consider pop-up stores or invest more on online ordering and delivery
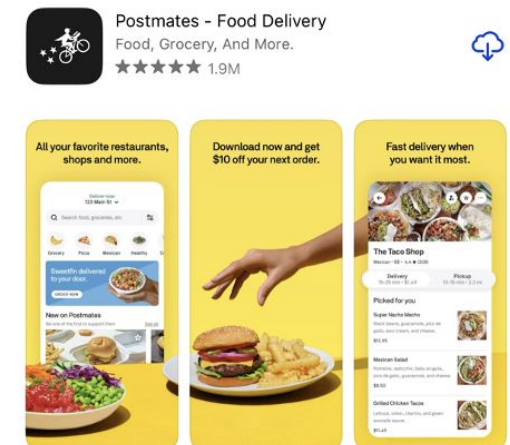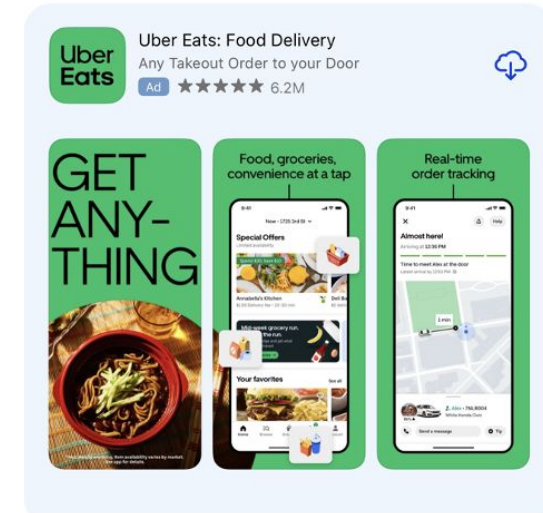
We are one click away.

Positioning :Cluster (#4)

Price:

They are price insensitive but are not happy with Chipotle's prices. Offer coupons and value meal deals to reward frequent visits and customer loyalty

Promotion:

Increase targeted Social media ads for this segment since most traffic was brought by social media. Ad campaigns on social media (Facebook, youtube and Instagram)

Thank You!
Q&A