Market Segmentation

Bret Jaco

8/15/2021

Market Segmentation

The data we are analyzing was collected in the course of a market-research study using followers of the Twitter account of a large consumer brand called NutrientH20. Over the course of a seven-day period, the follower's tweet were categorized using 36 different categories, each representing a broad area of interest.
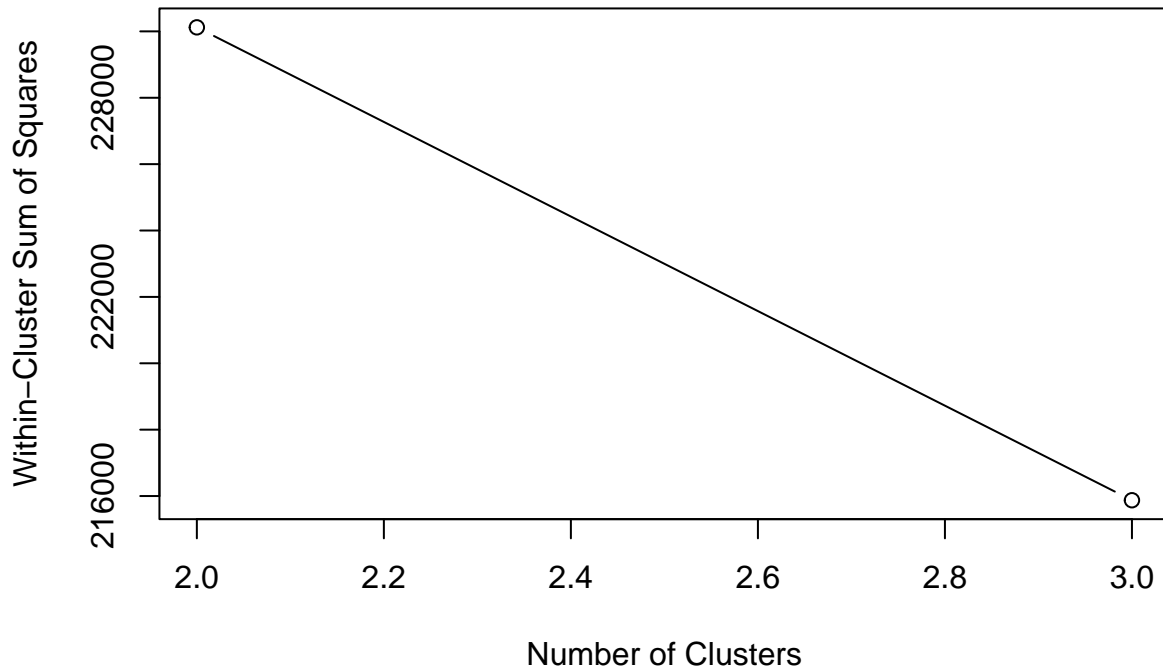
K-Means Clustering

To help NutrientH20 better understand its social-media audience, we used k-means clustering to group its twitter followers into different potential marketing segments. Prior to fitting the model, we removed the following variables, as they would not provide any beneficial insight to our problem: chatter, uncategorized, adult, and spam.

Choosing the Number of Clusters

One of the difficulties in using K-Means Clustering is deciding how many clusters to use. We decided to use the Elbow method which looks at the total Within-Cluster Sum of Squares as a function of the number of clusters; One should choose a number of clusters so that adding another cluster doesn't significantly improve the total WSS.

WSS vs Number of Clusters

Looking at the elbow plot above, we'll use 6 clusters for our analysis.

Another downside of K-means clustering is choosing the initial values, or "seeds", for the clusters. We use the K-means++ method to address this issue. The clustering will show us what tweet topics are typically posted together and help us form consumer personas for the marketing department to target.

Cluster 1

Table 1: Young Parents

	x
sports_fandom	5.90
religion	5.26
food	4.57
parenting	4.06
school	2.71
photo_sharing	2.63

Young Parents: With interests such as sports fandom, religion, food, parenting, and school, this segment represents young parents. They will want to provide their children with high quality water.

Cluster 2

Table 2: Fitness Gurus

	x
health_nutrition	12.01
personal_fitness	6.45
cooking	3.27
outdoors	2.74
photo_sharing	2.69
food	2.13

Fitness Gurus: With health nutrition and personal fitness the predominant categories, this grouping represents the healthy and fit archetype. They are likely to have a favorite water brand and stick to it.

Cluster 3

Table 3: Millennial Influencers

	x
cooking	10.91
photo_sharing	6.12
fashion	5.53
beauty	3.89
health_nutrition	2.28
shopping	2.02

Millennial Influencers: This segment is interested in cooking, photo_sharing, fashion, and beauty and likely consists of social media influencers. They like to share things with their network and would be good marketers of NutrientH20.

Cluster 4

Table 4: College Student

	x
college_uni	10.36
online_gaming	9.25
photo_sharing	2.87
sports_playing	2.57
tv_film	1.94
health_nutrition	1.73

College Student: This segment shows interest in college, online gaming, photo sharing, and sports. This sounds like a typical college student who likely isn't too picky about what kind of water they drink.

Cluster 5

Table 5: Traveling Businessman

	x
politics	8.93
travel	5.61
news	5.29
photo_sharing	2.54
computers	2.47
automotive	2.34

```
##      politics        travel          news photo_sharing      computers
##      8.933042      5.608443      5.294032      2.542940       2.474527
##     automotive
##      2.336245
```

Traveling Businessman: This segment has interests in politics, travel, and news and represents the traveling businessman. They would be a good group to target as they are frequently in airports and buying water bottles.

Cluster 6

Table 6: Average Consumer

	x
photo_sharing	2.28
current_events	1.44
shopping	1.28
travel	1.10
health_nutrition	1.10
politics	1.01

Average Consumer: This cluster seems to be very balanced among various categories and doesn't provide too much information into a specific archetype.

Market Segments

1) **Young Parents**
2) **Fitness Gurus**
3) **Millennial Influencers**
4) **College Student**
5) **Traveling Businessman**
6) **Average Consumer**

NutrientH20 can target these market segments to improve their sales and popularity!