

Submission Details

Grade: 2 / 2

 [Show Rubric](#)

Identifying Components of Clustering Algorithms [Videos 4.8 – 4.9 and Try-It Activity 4.2]

Roy Nunez submitted Apr 19 at 11:38pm

[Re-submit Assignment](#)

[Live Support](#)

**What spending trends do you notice in the clusters formed from your data?**

The clusters are formed from Age and Spending score for clustering where spending behavior will be observed.

We see two clusters of lower ages with centroids in the upper 20s and lower 30s. The lower ages cluster has comparable spending scores to an older age group, while the second older age group has the highest spending scores of all the clusters.

We then have a middle aged group with the lowest spending scores.

If I try to interpret these observations, the youngest cluster has less money to spend on discretionary spending than the older young cluster. The older young cluster as a young professional may be spending money on dating or other discretionary and leisure expenses, and have varying expenses with active consumption.

The middle aged group can be mostly settled down and have stabilized their expenses over time.

The older group may be living off savings and retirement funds.

Is there another variable, apart from age, that you could have used in this scenario?

Yes, one would assume annual Income influences spending habits and gender can also be considered as a variable.

However, when I generated clusters that included annual income, I observed no difference between income brackets. (Image 1) Similarly when I incorporate gender and plotted all 4 attributes I did not observe a discernible difference (Image 2) even when I tried to separate the genders (Image 3).

I then combined the ratio of spending score to annual income and labeling it as spending efficiency (Image 4).

What I observed in this plot was the age group with the highest spending scores on average were younger. I interpret this as a group of young professionals who prioritize experiences over savings or just more risk-averse. I did try to measure my performance with this cluster and got a Silhouette Score (0.372), which leads me to believe there is room for improvement on the segmentation if I wanted to go through this route.

Image [Live Support](#)