CUSTOMER SEGMENTATION USING DATA SCIENCE

Customer segmentation models are often used for dividing a company’s clients into different user groups. Customers in each group display shared characteristics that distinguish them from other users.

Here is a simple example of how companies use data segmentation to drive sales:

Every time I visit an e-commerce site, I look for items that are on sale to add to my cart. If I want to buy an item of clothing and it isn’t currently on sale, I wait until I see a special offer before making a purchase.

Data scientists at e-commerce companies often build customer segmentation models to identify shared traits amongst their customers. After building such a model, they notice that there are a handful of customers like me who always wait for a special offer before making purchases.

They classify us into a segment called “thrifty shoppers.”

Every time a new promotion is released, the company’s marketing team sends me and every other “thrifty shopper” a curated advertisement highlighting product affordability.

Whenever I get notified of a special discount, I rush to purchase all the items I require before the promotion ends, which increases the company’s sales.

Similarly, all the platform’s customers are grouped into different segments and sent targeted promotions based on their purchase behavior.

The example above demonstrates how customer segmentation models add value to organizations.

Data scientists usually build customer segmentation models using unsupervised machine learning algorithms such as [K-Means clustering](https://365datascience.com/tutorials/python-tutorials/k-means-clustering/) or hierarchical clustering. These models can pick up on similarities between user groups that often go unnoticed by the human eye.

In this article, I will show you how to build a data segmentation model in Python. You will learn to prepare data for customer segmentation and to build a K-Means algorithm from scratch. We will also look at how RFM is used in marketing to analyze customer value and explore other metrics for evaluating the performance of a clustering algorithm. Finally, we’ll answer the question of how to visualize and interpret clusters for customer segmentation.