

## 18.6.1

## Understanding Hierarchical Clustering

**You** both covered a lot of ground today with unsupervised learning, and Martha now feels ready to complete her project for her client. She does have one last question: Is K-means the only sort of clustering algorithm? There is another type, hierarchical clustering.

Similar to K-means clustering, hierarchical clustering, also known as agglomerative clustering, works with groups (clusters) of data points. The algorithm starts by declaring each point with its own cluster, then merges the two most similar clusters until a declared stopping point has been reached. This stopping point is implemented within your code.

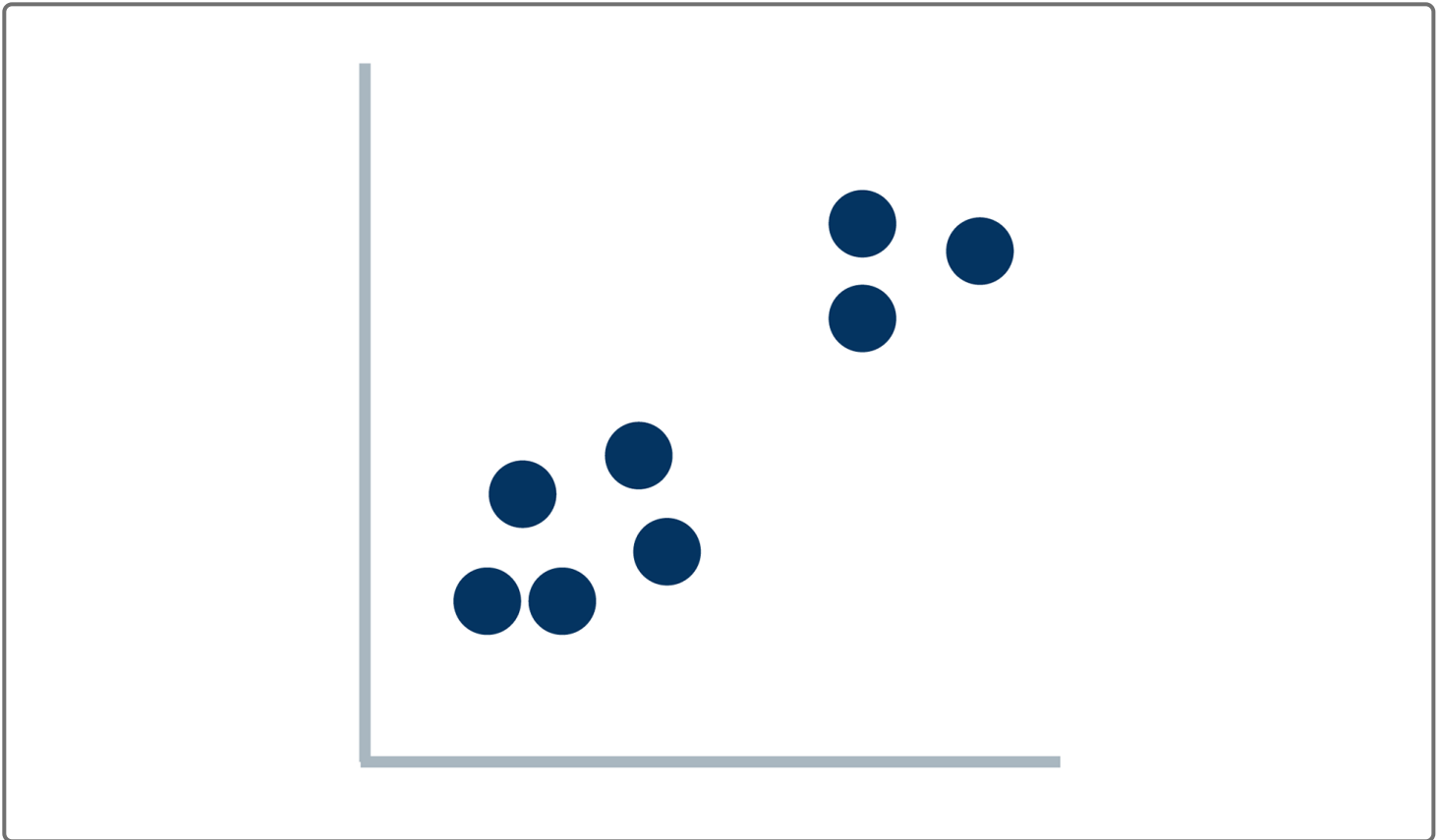
Hierarchical clustering has three methods for determining how points are linked: ward, average, and complete.

**Ward** is the algorithm's default setting. Simply put, this function selects the two clusters that, when merged, will mean the least amount of variance between all remaining clusters. This often leads to clusters that are relatively equal in size.

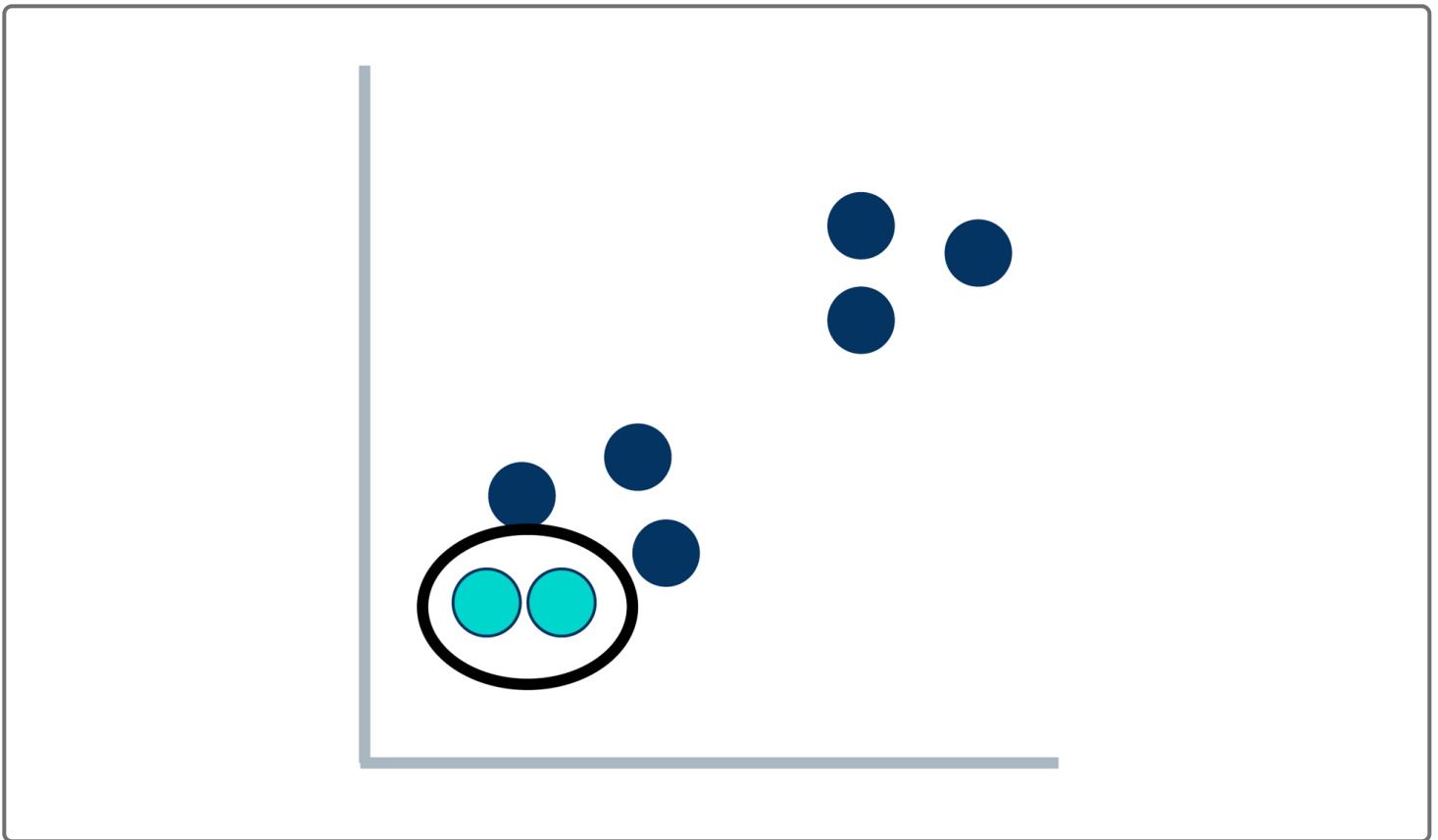
You can also link points using the **average**. This method connects clusters that have the smallest average distance between all of their points, then connects clusters on the smallest average distance between all their points.

The third method is **complete**, which links by merging clusters that have the smallest maximum distance.

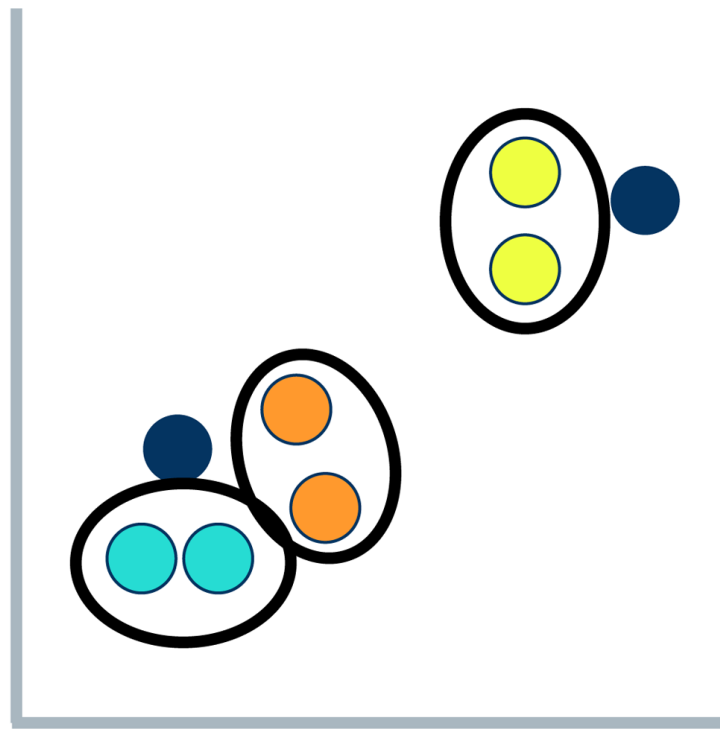
For example, say we're given the following points and our stopping point is set at two clusters:



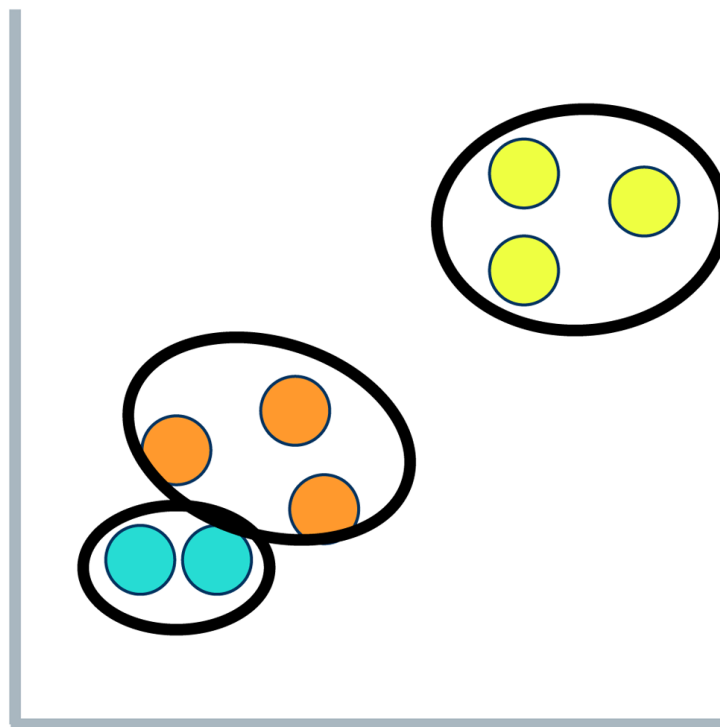
Each point, at the start, is considered a part of its own cluster, so we can say our starting point is eight clusters. We'll start by grouping together the two closest points:



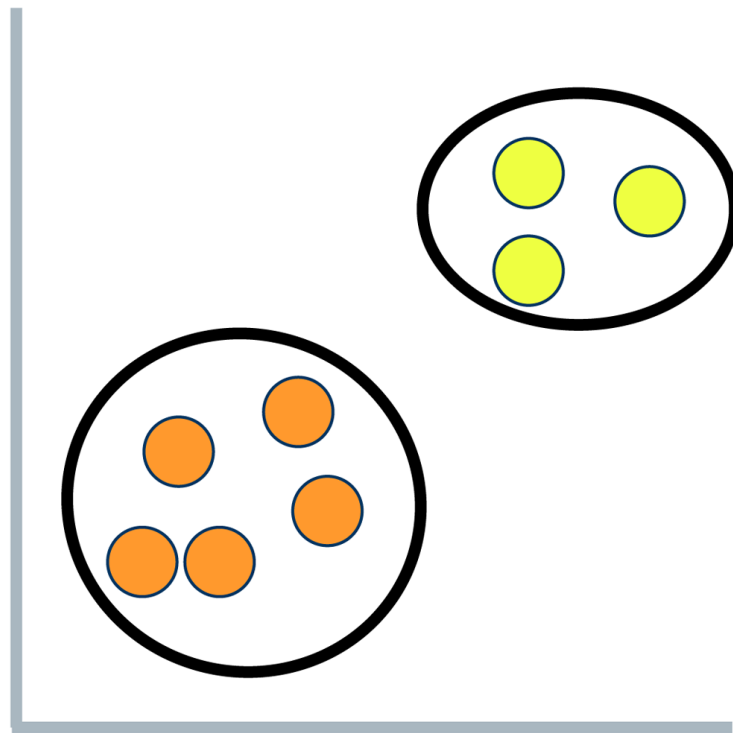
We now have seven clusters, and our stopping point of two clusters has not been met, so we keep going. The next two closest points are then grouped together, and we continue with this process:



Now, the next closest points contain a point in a cluster. When this happens, we join that point to the closest clusters. Remember, each point is considered a cluster:



The next two closest points to each other are now both contained in a cluster, since the stopping point of two clusters has still not been met, these clusters are merged:



We have now reached the stopping point of two clusters, and we just completed hierarchical clustering.

Now, you might wonder how we determine the ideal number of clusters, as we did for the elbow curve. We'll cover that in the next section with dendrograms.

© 2020 - 2022 Trilogy Education Services, a 2U, Inc. brand. All Rights Reserved.