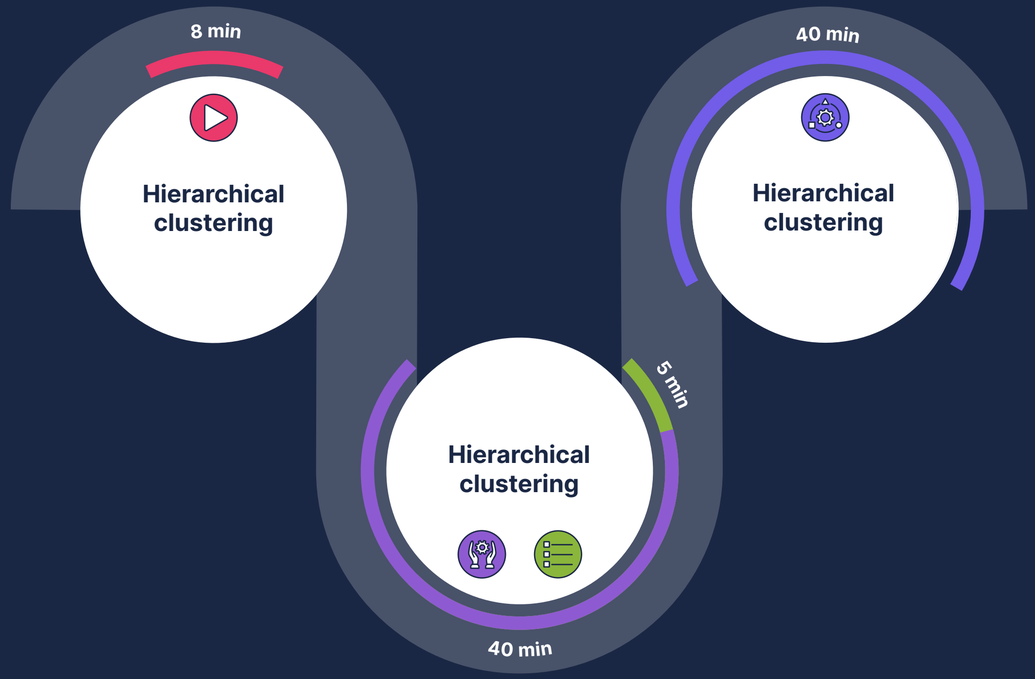
Hierarchical clustering – Lesson overview

Video
Knowledge questions
Examples
Exercise

Hierarchical clustering is a method of cluster analysis used in data science. It facilitates the exploration of complex datasets by organising them into a **hierarchical structure**. Through this method, **data points are grouped together based on their similarities**, forming clusters that can be further subdivided into smaller clusters.

In this module, we will explore hierarchical clustering, covering its **algorithm** and the **dendrogram interpretation**. We will **implement** this method using sklearn and SciPy, and learn how to determine the optimal **number of clusters**.



Learning objectives

- Understand hierarchical clustering.
- Interpret dendrograms.
- Implement agglomerative clustering in sklearn and SciPy.
- Determine the optimal number of clusters from a dendrogram.