17.2.1

Overview of Machine Learning

You're excited to get started on your machine learning journey. However, before using machine learning to solve problems, let's look at the most common methods employed.

Machine learning is the use of statistical algorithms to perform tasks such as learning from data patterns and making predictions. There are many different models—a model is a mathematical representation of something that happens in the real world—and you'll learn about several this week.

Broadly speaking, machine learning can be divided into three learning categories: supervised, unsupervised, and deep. For our purposes, we'll only discuss supervised and unsupervised learning.

Supervised Learning

Supervised learning deals with labeled data. An example of supervised learning might be to predict, based on data from previous patients, whether a new patient has diabetes.

Patient ID	Age	Body Mass Index	Has Diabetes
1	31	24	No
2	68	39	Yes
3	57	35	?

In the simplified example above, we know whether or not the first two patients have diabetes. That is, the dataset is labeled. Each row represents a patient, and the "Has Diabetes" column is the label that informs whether or not the

patient has diabetes. The status of the third patient is unknown, and based on this patient's age and BMI, the goal is to predict whether or not this person is diabetic.

Unsupervised Learning

In **unsupervised learning**, by contrast, machine learning algorithms work with datasets without labeled outcomes. In supervised learning, the labels provide the correct answers. In unsupervised learning, such correct answers, or labels, aren't provided. An example of unsupervised learning might be to task a machine learning algorithm with grouping a bag of objects as it sees fit. The algorithm isn't given labels, so it's on its own to find patterns. In this case, it might group them based on shapes, colors, or perhaps both:
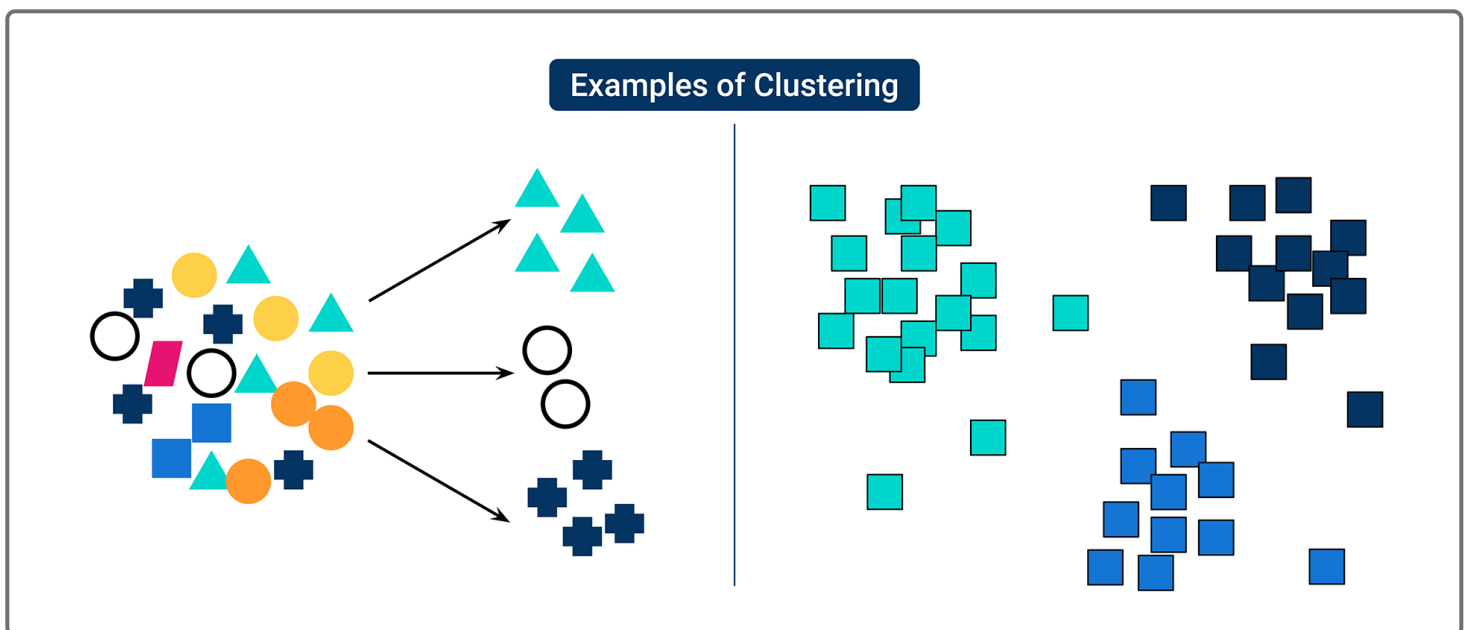
A common application of unsupervised learning is to group customers by purchasing patterns.

You'll learn more about unsupervised learning in the next module.

C Retake