**Drill: Unsupervised or supervised?**

*For each of these scenarios decide if you could use supervised or unsupervised techniques, or both!*

1. Define the likelihood that an individual will contract a specific disease

I think since we know what specific disease we have an outcome variable and we can predict or measure the likelihood using a Supervised model.

1. Translate a set of images into variables for modeling

Unsupervised learning since it can group or form clusters or identify similarities by itself

1. An ecommerce company wants to identify power users

We know the outcome variable and it is supervised learning based on shopping through internet etc.

But we don’t have the test data since we never classified before so it may be unsupervised learning.

1. That same company wants to see shopping patterns in users

Here we are not sure of what is the outcome and no test data so this is unsupervised learning and one of methods this can be done is by clustering

1. You want to reduce the number of variables inputting into your random forest model

This is unsupervised learning done through PCA which is a form of unsupervised learning.

If any of these scenarios are unclear feel free to follow up with your mentor.

­­­­