

## Supervised vs. unsupervised machine learning exercise

**Exercise:** Sort candy into categories using two different machine learning algorithms: Procedure A and Procedure B. Half the class will follow each procedure, and we will then discuss outcomes and what we've learned.

### Procedure A

1. Assign one group member to sort candy into two categories: candy they like and candy they dislike. The person doing the sorting should NOT explain to other group members why they are sorting the way they are (i.e., don't say "I only like chocolate"). This is analogous to a training dataset with perfect classification accuracy.
2. Remaining group members should then sort the "test dataset" - another collection of candy - into the categories based on similarities to the training dataset.
3. Have the group member who originally sorted the candy score accuracy. Did the group members sorting the test dataset intuit the original sorter's preferences correctly? What rules did you identify?

### Procedure B

1. Examine traits of candy - what traits vary, and how much?
2. Establish criteria for sorting into two categories of candy. Try to maximize similarity within each cluster, and differences among clusters. Be creative - there are multiple right ways to do this, and our discussion will be most interesting if groups come to different solutions.
3. Now split the candy into three categories. Which number of clusters do you think is a better fit for the range of variation in your "dataset"? Would more clusters be advantageous?
4. Add in a new subset of candy. Do the original categories still hold, or do you have to rearrange to accommodate the new "data"?

### Follow-up discussion questions

1. Which approach is supervised learning, Procedure A or Procedure B? Which is unsupervised? Why?
2. How might sample size affect outcomes of your algorithm? What about the range of variation in your dataset?
3. What did you learn about underlying logic for assigning individual pieces of candy to one category or another? Is it easier to detect these mechanistic relationships between traits and cluster membership with supervised or unsupervised learning?