

## Project 2 Outline:

### Topic:

For project 2, I will be working alone. The topic for my presentation will be doing supervised machine learning in R. Specifically, I want to learn how to effectively implement k-means clustering in R.

### Resources:

In terms of the resources I will use, I have found a very relevant Data Camp course on supervised learning in R, specifically how to implement k-means clustering in R and also how to visualize and interpret results! <https://learn.datacamp.com/courses/unsupervised-learning-in-r>

### Project Format:

For my format, I will do a classic presentation with an accompanying slide deck, because that is how I personally learn best. The slide next will not include any interactive clicker questions, because I before a more traditional simple slide deck.

### Outline:

Below is an outline of the topics I plan to cover in my project:

Introducing the three main kinds of machine learning (putting into context what linear regression is since that's one that most people know) - 5 mins

Explaining the two main goals of supervised learning (clustering and pattern finding in data) - 2 mins

Introducing k-means clustering algorithm - 5 mins

- discuss theoretical algorithm for k-means
- identifying a good number k using 'elbow plot'

Showing how to implement k-means clustering in R - 4 mins

- show how to build model
- determine good number of k using 'elbow plot'
- visualize the k-means clustering algorithm

### Link to Materials:

The class materials can be found in the GitHub repo 'K-MeansClustering Presentation' at the link <https://github.com/Courtney-E-Miller/K-MeansClusteringPresentation>. The link to clone the repo is <https://github.com/Courtney-E-Miller/K-MeansClusteringPresentation.git>