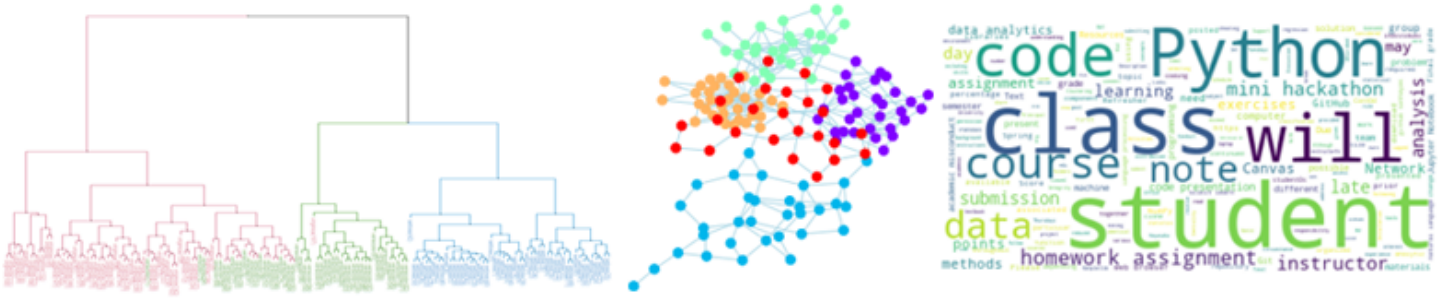
PSY 394U **Data Analytics with Python**, Spring 2019



Mid-Term Mini-Hackathon

March 12 & 14, 2019

Glasses or no glasses

The data file **`FaceData.npz`** contains 238 images (64 x 64 pixels) of people with and without glasses. In particular there are 3 arrays:

- **`Faces`**: an array containing flattened images, i.e., a series of 4096 pixel intensities.
- **`FaceImages`**: an array containing 64 x 64 images
- **`y`**: an indicator variable of whether the person is wearing glasses (1: glasses, or 0: no glasses)

Needless to say, observations are stored in the same order in these arrays. Your goal in this project is to construct a classifier to separate those with glasses and those without glasses. You can choose a classification algorithm, but you need to determine the optimal parameter(s), if necessary, so that the classification performance is optimized.

After constructing your best classifier, you need to demonstrate its performance with a 10-fold cross-validation.

Source: The Olivetti faces dataset available in Scikit-learn, https://scikit-learn.org/0.19/datasets/olivetti_faces.html