1. Due Monday 2/24
    a. An informal document outlining:
        i. your chosen context,
            1. Facial Recognition Technology
        ii. big question,
            1. Is it possible to identify which BOW institution a student attends based on their appearance?
        iii. Sub-question,
            1. How can we use EVD & PCA to accurately determine which BOW institution a student attends based on their LinkedIn profile photo?
        iv. the algorithm you will use to answer your sub-question
            1. Load training data into 3 [100; 100; n] matrices for each institution (where n is number of training data photos)
            2. Convert each matrix into column vectors
            3. Find covariance matrix for each institution (and experiment with how many principal components are necessary)
            4. Find eigenvectors
            5. Form eigen-bodies
            6. Put test images in a [100; 100; n] matrix → column vectors → eigen bodies
            7. Create answer matrix to run nearest neighbor classification using the eigen-bodies from each institution
        v. a plan for what analysis you will do and what kind of results you will get.
            1. Talk about ethics! While this is a really cool application specifically for our context, it can easily be used for the wrong purposes. This algorithm literally divides people into different categories and puts a label on them based purely on their looks. This program can very easily be taken one step further and no longer talk about placing students into their respective institutions. It can be placing people into more harmful categories based on looks.
            2. After seeing the accuracy of using the full profile pictures as a decimal between 0 and 1, there are a few different possible routes we can take. We will look at the overall accuracy as well as the accuracy for classifying each institution individually and determine if editing our script and process overall has to be done to achieve better results.

a. If the outcomes seem heavily skewed by gender expression (if most female students are put into Wellesley), an option is to just do Olin or Babson. This would eliminate the strong gender-biased information if that presents as a problem.

b. If using the full profiles creates a program with very low accuracy (some are headshots, some are full body, etc) we can modify our program to just look at the faces within our images. There are two ways we can go about this. We can use a program online to crop the images to just the faces, or we can research how to make our MatLab script do this. We can then repeat the process we already have just with the cropped photos and analyze the differences.