Adjusting to this assignment was a challenge, as Mat Lab is a new language, and working with it to find the results of iterative K-means. The results found show that the different maps have different levels of accuracy, and that they can vary vastly. Creating the iterative k-means required working with the given k-means method, which was designed to work with a list, instead of a matrix that had to be converted. After adjusting the images, the design for the iterative k-means required creating a matrix that had the width of k, and the length of the picture. The purpose behind this is to be able to generate an image for each k, as each column in the matrix became the equivalent of a blank image. Then, the algorithm used went through each k (column) in a pixel (row) that was analyzed in each iteration (ten were done) of the image. This led to pixels that were more likely to be in a certain cluster (represented as k-values) to have higher values. Combining a pre-given method (reshape) printed off each of the columns of the vector and created k images that had different values that displayed clustered areas together.

The results, or images, generated by this developed more noise, or irrelevant data, that can cause the clustering of the data to become more inaccurate. However, if the size of k is too small, the image isn’t completely separated, and the number of clusters are not enough to separate the data properly. Because of this, a good medium, especially considering the usefulness of creating grayscale images, is working with k between 3 to 5, as seen below.



As can be seen in the images, the 3k on the right, while not as precise as the 4k on the left, has less noise. However, this could also be attributed the fact that the image on the right has more facial region available, so the data may be slightly offset by this. However, as seen by the multiple tests ran, and the different results produced, the more facial area that was presented in an image, the more readily the algorithm was able to recognize the features, and differentiate between different areas of the image, such as the eyes, nose, and mouth verses the skin or hair.

In summary, k-means is a quick-running algorithm to perform clustering with, and using iterations of the algorithm definitely improve the accuracy of the clustering, as having multiple options to go through shows the different accuracy levels. While implementing this algorithm in a new language may be challenging, the results still show that the algorithm and methods collect the data successfully.