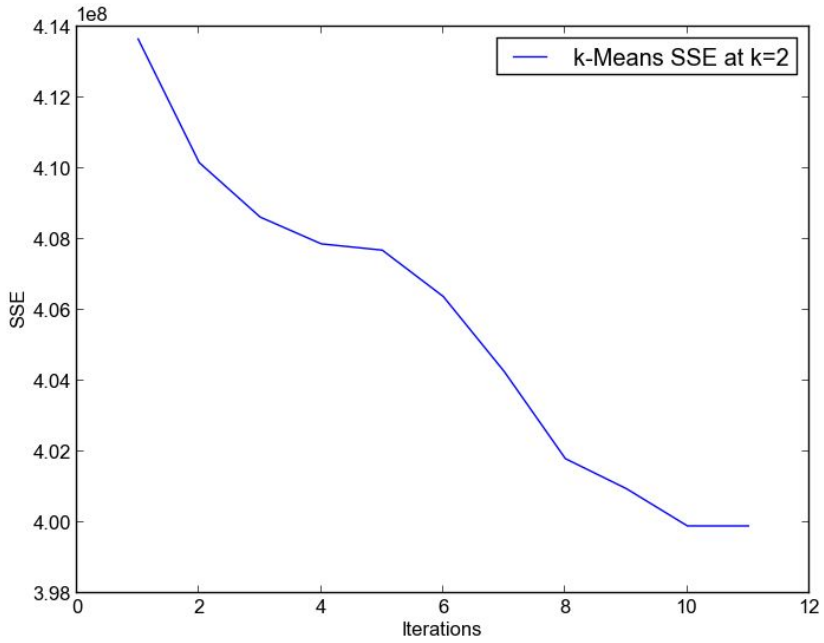
Trevor Swope
William Selbie
CS 434
May 27, 2018

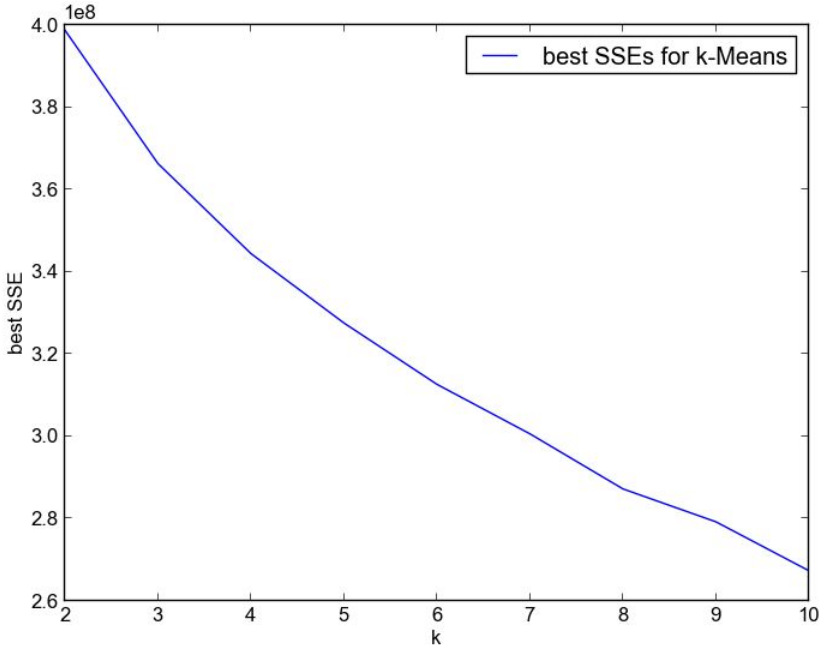
Implementation Assignment #4

Part 2

1. The following is the result of a typical k-means run on k=2



2. Based on this and other graphs from different runs, I would suggest that k=6 would be a good choice for this dataset; the most obvious knee appears somewhere around there



Part 3

1.

EigenValue 0: 352868.691256
EigenValue 1: 267895.866871
EigenValue 2: 227632.699244
EigenValue 3: 174703.490258
EigenValue 4: 130486.76236
EigenValue 5: 115542.502682
EigenValue 6: 99726.4367264
EigenValue 7: 90576.0578785
EigenValue 8: 85326.5368082
EigenValue 9: 71547.9660125

2.

The images look like numbers, although the resemblance is not perfect. There's a lot of noise / distortion that makes it difficult to distinguish which number each image actually is. The eigenvector based images seem to capture the general shape of the number though. For instance, the eigenvector associated with the highest eigenvalue shows what appears to be a 3 or a 7, though it is definitely a 7.

3. The images look quite similar to the eigenvector images, although the eigenvector images do still require a little bit of deciphering to understand. The similarities are most clear with images 0, 2, and 6. There's a general trend that 3s and 8s are easily confused, as are 4s and 7s.