

Problem 1

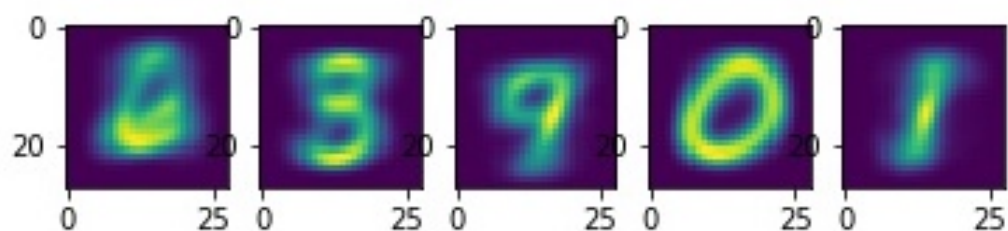
a.)

(See code)

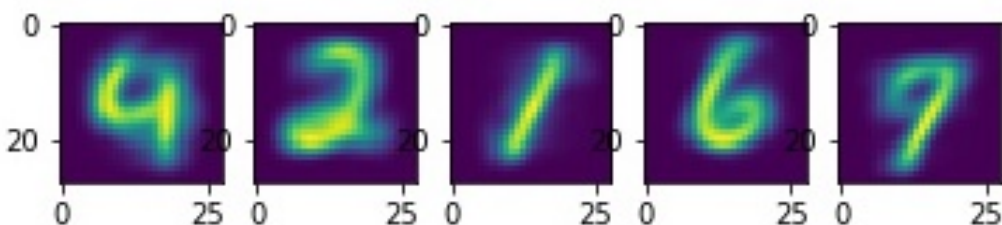
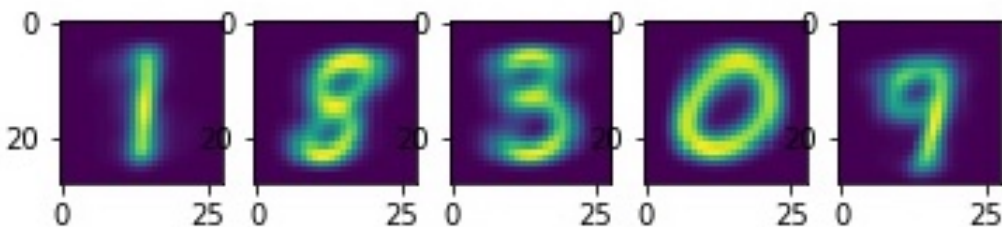
b.)

Visualization of MNIST for k -means

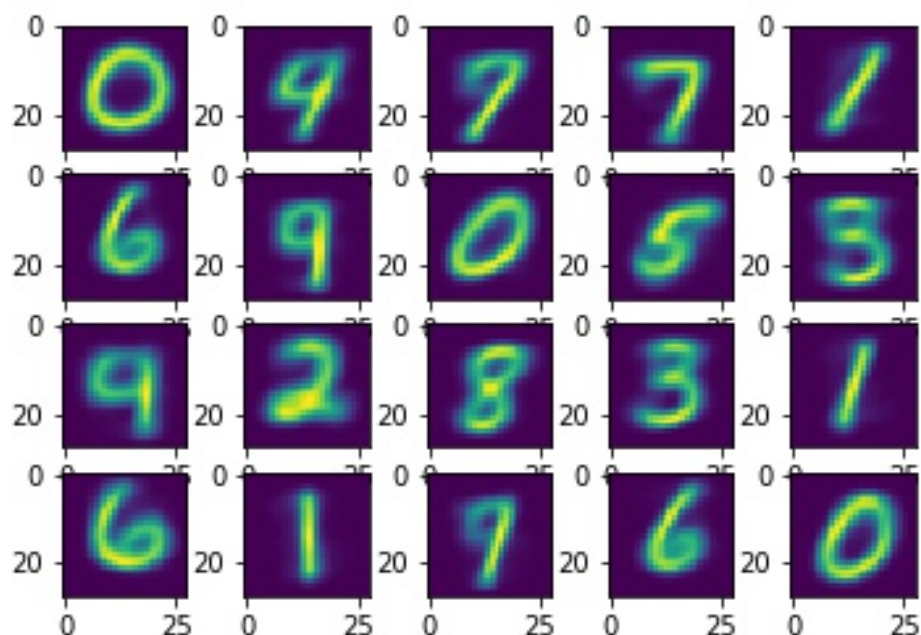
$k = 5$



$k = 10$



$k = 20$



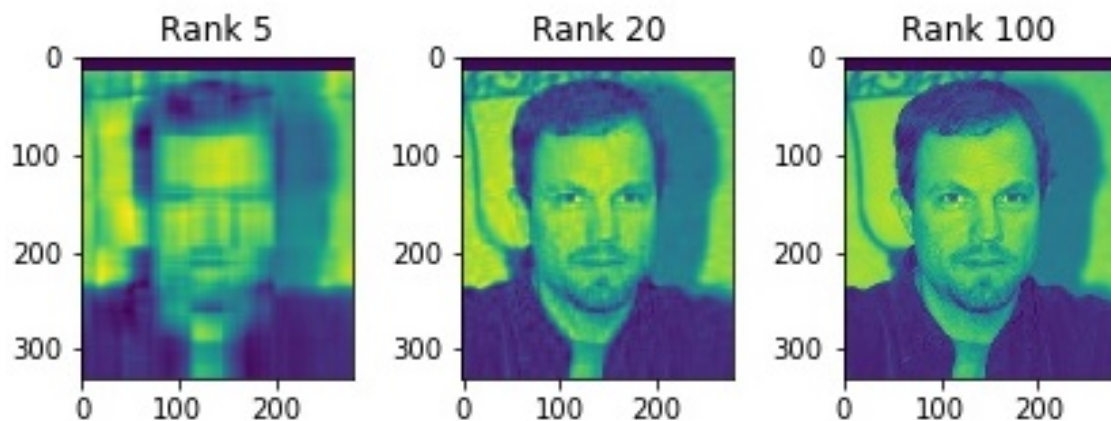
It becomes apparent that the resolution of the numbers that are picked out by the clusters improves. For example, it is difficult to identify a number corresponding to the first cluster in the $k = 5$ set. In the $k = 10$ set we are able to identify most of the numbers (to at least two possible digits) even though not all digits 1-10 are represented. Finally, in the $k = 20$ clustering, we have at least one clustering corresponding to each digit 1-10 and the resolution of each cluster is significantly improved.

Problem 2

a.)

Low-Rank Approximation: Face Image

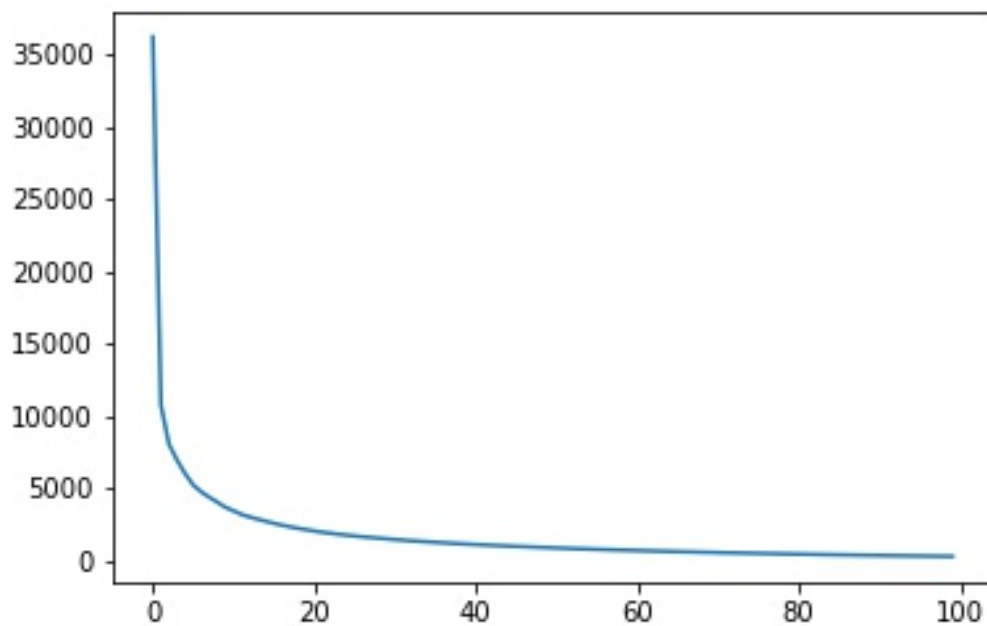
Rank-5, rank-20, and rank 100 approximations:



b.)

Mean-Squared-Error: Face Image

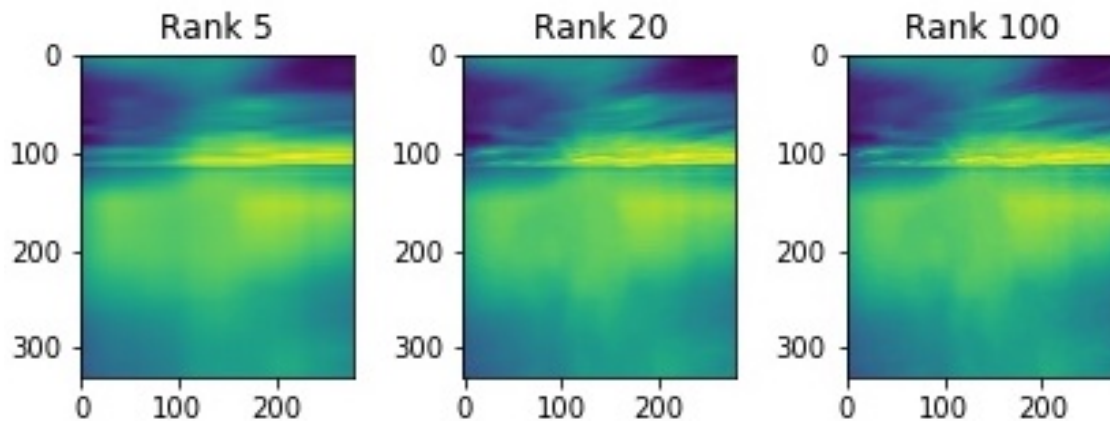
Mean-Squared Error for rank-1 through rank-100 approximations of the face image compared to the original.



c.)

Low-Rank Approximation: Sky Image

Rank-5, rank-20, and rank 100 approximations:



d.)

For the face image, a rank 40 approximation shows significantly degraded resolution (especially in the forehead area), and a rank 45 approximation still seems a bit fuzzy. I find the rank 50 approximation to be sufficiently close for the two images to be considered identical. (See Jupyter notebook for evidence)

For the sky image, the rank 20 image shows a slight degradation in quality in definition of the colors in the bottom left. This effect is significantly reduced in a rank 25 image, and by a rank 30 approximation, the original and the low-rank image are nearly indistinguishable.

The difference between the sky and face images is likely caused by the level of detail. The sky image has less distinct features, so reducing the rank will likely have less impact on the clarity of each of those features. (Additionally, from a biological perspective, the human brain is incredibly adept at identifying other human faces. Even very slight differences in the face image may trigger the brain to reject the images as identical, whereas it is more apt to overlook those slight differences in the sky image.)

[See Jupyter notebook for evidence]