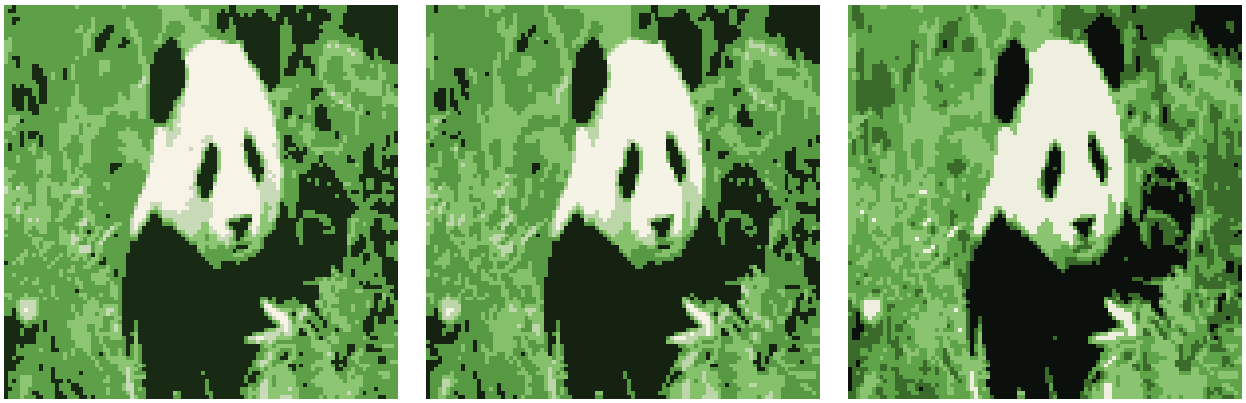
Jacob Kerstetter
4406239

Problem Set #1

Below are my observations of the parameters and how they affect the output on a variety of images.

For an image with a lot of colors, more segments are necessary to get a representative image of the original picture. For example, for image 3, the progression from 3 to 5 to 7 segments is shown below. The ability to tell that the picture is a landscape of Pittsburgh gets easier and easier as more segments are added because more colors are available to make objects distinct.

I also found as testing image 1, with more clusters, such as 5 or 7, having more random restarts helps to lower the SSD and make the color segmentation more accurate. In the first output below, for example, some of the shrubbery is matched in the same cluster as the dark color of the panda. This becomes less and less as the number of restarts increases from 5 to 15 to 20 in the last output.



For image 2, a standout observation was what happened when the iterations increased. With K=5, increasing the iterations past 15 caused a degradation in the segmentation. The increase from 5 to 15 took out a few outliers, but 20 iterations made the distinctions between segments less noticeable. Additionally, these images make it clear that SSD is not a perfect metric.

Not pictured, but also notable, was the fact that images 2 and 3 were difficult to take observations from at 3 clusters because of the intricacy of the images, but image 1 was segmented fairly well at 3 clusters.