

# Assignment V



## K-Means Clustering on Images





**Ashika Prakash Acharya**  
**axa190084**

The assignment task is to display the images after data compression using K-means clustering for different values of K (2, 5, 10, 15, 20).




The algorithm runs for a total of 5 iterations.




### Koala Image Compression:

# k clusters	Image	File size	Degree of Compression
Original		781 KB	0 %
2		126 KB	84%

5		337 KB	57%
10		546 KB	31%
15		649 KB	17%
20		944 KB	120%

## Penguin image compression

#k clusters	Image	File size	degree of compression
original		778 KB	0 %
2		87 KB	89%
5		190 KB	76%

10	 A photograph of three penguins standing on a dark, flat surface. They are facing each other in a triangular formation, with their heads tilted upwards. The penguins have white bodies with blue and black markings on their heads and backs. The background is a dark, textured surface.	413 KB	47%
15	 A photograph of three penguins standing on a dark, flat surface. They are facing each other in a triangular formation, with their heads tilted upwards. The penguins have white bodies with blue and black markings on their heads and backs. The background is a dark, textured surface.	517 KB	34%
20	 A photograph of three penguins standing on a dark, flat surface. They are facing each other in a triangular formation, with their heads tilted upwards. The penguins have white bodies with blue and black markings on their heads and backs. The background is a dark, textured surface.	551 KB	30%

Is there a tradeoff between image quality and degree of compression? What would be a good value of K for each of the two images?

- From the compression images available above, it is quite evident that the image quality decreases significantly with the higher degree of compression.
- During compression, pixels with similar RGB colors will get similar common color. Thus, lower value of k signifies lesser number of colors and thereby the image quality decreases but we can achieve higher compression rates.
- Inversely, if k value gets increased, more colors can be shown in image which increases image quality, but size of the image will also increase.

**$k \text{ colors} \propto \text{Image quality} \propto \text{Image size} \propto 1/\text{degree of compression}$**

For Koala image, K=10 is a good value with 31% degree of compression and decent image quality.

For Penguin image, K=10 is a good value with 47% degree of compression and decent image quality.