**Net Id: ngr140030**

**3. K Means Report**

* The output images for different values of K refer to folder Koala and Penguins.
* Took 3 reading for each k value
* Compression ratio = original size of image / compressed size(kmeans op image).
* Variance 1/5 \* summation of (xi-mean)^2 .

**K means Compression for Koala.jpg (Original size= 780 KB)**

| K | Compression | Ratio |  |  |  | Mean | Variance |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 2 | 5.9672 | 5.9672 | 5.9988 | 5.9994 | 5.9988 | 5.9863 | 0.000303 |
| 5 | 4.2154 | 4.4229 | 4.2154 | 4.4462 | 4.4229 | 4.3446 | 0.013992 |
| 10 | 4.7916 | 4.7784 | 4.7943 | 4.7547 | 4.7761 | 4.7791 | 0.000248 |
| 15 | 4.9661 | 4.9839 | 4.9475 | 4.9694 | 4.9327 | 4.9599 | 0.0004 |
| 20 | 4.9835 | 5.0235 | 5.0435 | 5.0472 | 4.9656 | 5.0126 | 0.001332 |

**K means Compression for Penguine.jpg (Original size= 777.8 KB)**

| K | Compression | Ratio |  |  |  | Mean | Variance |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 2 | 9.1496 | 9.1496 | 9.1516 | 9.1516 | 9.1496 | 9.1504 | 0.01318 |
| 5 | 7.3499 | 7.1457 | 7.3496 | 7.1461 | 7.3496 | 7.2682 | 0.01246 |
| 10 | 6.6511 | 6.5885 | 6.8137 | 6.8470 | 6.6486 | 6.7098 | 0.012881 |
| 15 | 6.8402 | 6.6984 | 6.6741 | 6.6551 | 6.7863 | 6.7308 | 0.006261 |
| 20 | 6.7534 | 6.8615 | 6.8113 | 6.7727 | 6.8294 | 6.8057 | 0.001883 |

**Conclusions from K values:**

From above; it is clear that increase in k value decreases the compression but increases quality.

As the K value increases the size of the image increases (sometimes it decreases) on running it multiple times

The value k=10 looks good tradeoff for quality and compression. (Ideally) it is best when you know the exact number of distinct colors and with proper initialization.