ML HW 3

Part 2

K-means clustering on images
Penguins Data

K=2

```
AVG Compression Ratio= 0.10950908611723567
VAR Compression Ratio= 1.925929944387236e-34
```

K=5

```
AVG Compression Ratio= 0.13564419189159654
VAR Compression Ratio=1.3000434469784282e-06
```

K=10

AVG Compression Ratio=0.15052678267241767
VAR Compression Ratio=7.90350138079267e-07

K=15

AVG Compression Ratio=0.14944017690127084
VAR Compression Ratio=1.914849767047207e-06

K=20

AVG Compression Ratio=0.149766700362575363
VAR Compression Ratio=2.009065021278021e-06

Koala Data-
K=2

AVG Compression Ratio=0.16926838201864425

VAR Compression Ratio=0.0

K=5

AVG Compression Ratio=0.22570697116277402
VAR Compression Ratio=0.0

K=10

AVG Compression Ratio=0.20955238713626892
VAR Compression Ratio=3.158447294813235e-07

K=15

AVG Compression Ratio=0.20246532220160315

```
VAR Compression Ratio=1.4523074356112616e-06
```

K=20

```
AVG Compression Ratio=0.20011628636670417
VAR Compression Ratio=7.350916321271544e-07
```

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?

Yes, the less the value of K, the lesser the image quality. K=15 would be a good value where there's not much loss of image quality but at the same time, a good amount of compression has been achieved.