MAT 345 - PROJECT #4
due Monday, November 23, 2020 at 10:00PM.

OBJECTIVE: In this project, you will perform k-means clustering.

GRADING: The assignment is worth 8% of your course grade.

INSTRUCTIONS: Students will work individually on this project, but they may ask questions and clarification from classmates and the instructor. Students must submit their projects on Moodle.

SUBMIT THE FOLLOWING: A copy of your code and a Project Report. Make sure your name is on all files submitted.

PROJECT: You will run the k-means clustering algorithm to re-color an image with k colors only. Run the clustering algorithm for 3 different images, with different features.

1. Choose a picture and read each pixel's RGB info. This will be your data set $\mathcal{D}$.
2. For $3 \leq k \leq 10$
   - cluster the data set $\mathcal{D}$ into k clusters.
   - use the centroids from each cluster to re-color the image.
   - run the algorithm a few times and observe the resulting image.
3. Print the resulting image for each k and include in the written report.
4. Decide which k would be better to use for your image. Explain your choice in the written report.

Project Report: in your report, you must include, but are not limited to:

- Your Name
- The programming language you used for the project
- For each of the 3 images:
  - printout of a resulting clustering for each k;
  - best choice of k and explanation why;
  - any relevant discussion of changes in output for specific choices of k, and when k varies.