CSci 343 Fundamentals of Data Science
Challenge 3

Submission Window Opens:
October 21, 2019

Points Available:
200 XP for a working demonstration
50 XP for readable & understandable code

Objectives:

- Perform basic image analysis
- Use NumPy for data conversion

Assignment:

Picture it, Sicily, 1920…errr, I mean 2019. It's two years in the future, and you've graduated! After graduating you were hired by the Italian video security and surveillance company Surveilco. They are looking to improve their security products by using basic image analysis. Being a recent UM graduate, they know you're one of the best data scientists for the job!

Their video systems provide a stream of JPEG images over a network connection. They have given you a sample dataset to get started. The dataset consists of 200 JPEG formatted images captured from real world sources. Unfortunately, the network programmers who wrote the code for the cameras are not as skillful as you. As a result, about 5% of the images are corrupted and unusable… You will need to remove these corrupted images from your dataset ***before*** you even begin writing your code. (This will require you to actually look at your data before getting started.) You can remove them by hand or write a program to do it. Regardless, you'll need to remove these bad images.

In class, we made the observation that a digital image is, in fact, a type of numerical matrix. Using this observation, we can infer that we can perform mathematical operations on an image. Indeed, we can! In class, we worked through examples of calculating *average images* and *difference images*. For this assignment, you'll be calculating an *average image* and a *standard deviation image*. Keep in mind, that your images can be treated like a regular ol' mathematical variable, so you should be able to use them

directly in your mean and standard deviation formulas (no wizardry required).

Your task is to write a program to analyze a set of images, calculate an average image, calculate a standard deviation image, and then display the average image with areas of change highlighted in red (example image included below). Your program should prompt the user to enter a threshold value that will range between 0 and 255 (be sure that your program checks to make sure that the input is reasonable). This value represents the minimum standard deviation needed to highlight a pixel. To do the highlighting, you will need to loop through the pixels of your average image and change their color to the highlight color, but only if the pixel's standard deviation is higher than the threshold value. For example, setting a pixel in the image to red would look something like this:

```
avg_image[row][column]=[255.0, 0.0, 0.0]
```

You will need to submit your *source code* and your *filtered image data*. You will also need to save an image of your output plot (there is a button on your MatPlotLib window that will do this for you) and submit it with your code.

Pro-Tip: Be sure to have thoroughly tested your program to make sure it works at least SEVERAL DAYS before you demo it. These assignments are not "night-before programs" (even for very skilled students). You will probably have questions that will need to be asked a couple days before the submission window opens. The only way to know what to ask is to have played with the assignment before hand.

Deliverables

Upload ***all your code, image dataset, and a saved image of your working output*** to Blackboard as a single ZIP file. Name your ZIP file s*piritAnimal*.zip, where s*piritAnimal* is your class user ID (not your webID or ID number). Be sure the name matches your assigned name EXACTLY. Be sure to name your main source file *SpiritAnimal*.py (where SpiritAnimal is your Spirit Animal ID). In a comment at the top of your source file, be sure to include the following information.

- Spirit Animal User ID
- Date the file was last edited
- Challenge Number
- Cite any sources that you used as a reference for code, data, and content (including title and URL)

Failure to comply with submission instructions will result in loss of points up to 25XP.

Sample Execution:

```
[jones@Computer CSCI343]$ python SpiritAnimal.py
Enter your change threshold: 45
```