

CS101: HW 4

Due: April 30, 2018

Submission Instructions: Type your answers to the following questions in a .doc or .txt file and submit on Canvas. You will also need to submit your code **in a separate file**. You will also need to submit your images.

Image Filter Exercises

Complete the exercises at <http://web.stanford.edu/class/cs101/image-10-exercises.html>

1. Describe the edge detection algorithm we used in class. How does it work? Describe in English what the code is doing. What is the purpose of the if statement?
2. How would you modify the faded images code to merge two images completely (i.e. without filtering based on color - we just want to merge two pictures)? What about to have a weighted average, favoring the original image twice as much as the back image? For this question, a written description is fine, though you may include code if you'd like.

Thursday's Exercises

Spreadsheet Exercises

For the following exercises, you can find the datasets here: https://drive.google.com/drive/folders/1v4LenzHEnLXBNbCz-FGjLeFyAZuXjst_?usp=sharing.

You should submit a copy of the chocolate dataset with the formulas for answering the questions below. Except for the last question, all the questions below pertain to the chocolate dataset.

3. How many distinct Broad Bean Origins are included, excluding empty Broad Bean Origins?
4. What is the average rating for chocolate bars sourced from Hawaii?
5. Which origin has the most number of bars ranked?
6. Which origin has the highest average rating?
7. Which origin has the greatest minimum rating (i.e. the minimum rating for a region is the smallest rating - which region has the largest minimum rating)?

8. How does chocolate percentage affect the rating? Specifically, what are the average ratings for chocolate of at least 75%, between 60 and 74.9%, and below 59.9%.
9. Which company location is most likely to source their chocolate from Madagascar?
10. What percentage of chocolate made by companies in U.S.A. scored a rating of at least 4?
11. Find your own dataset (generally, a .csv file is best) and open it in Google Sheets or Excel. Describe three findings you made about your dataset, and include the formulas you used to calculate them. These findings should be non-trivial (i.e. more than just that the dataset has this number of rows or an average value of that number). Please also include a link to your dataset.

Note: if you're having trouble finding a dataset, you can try FiveThirtyEight (<https://github.com/fivethirtyeight/data>) or Kaggle (<https://www.kaggle.com/datasets>).