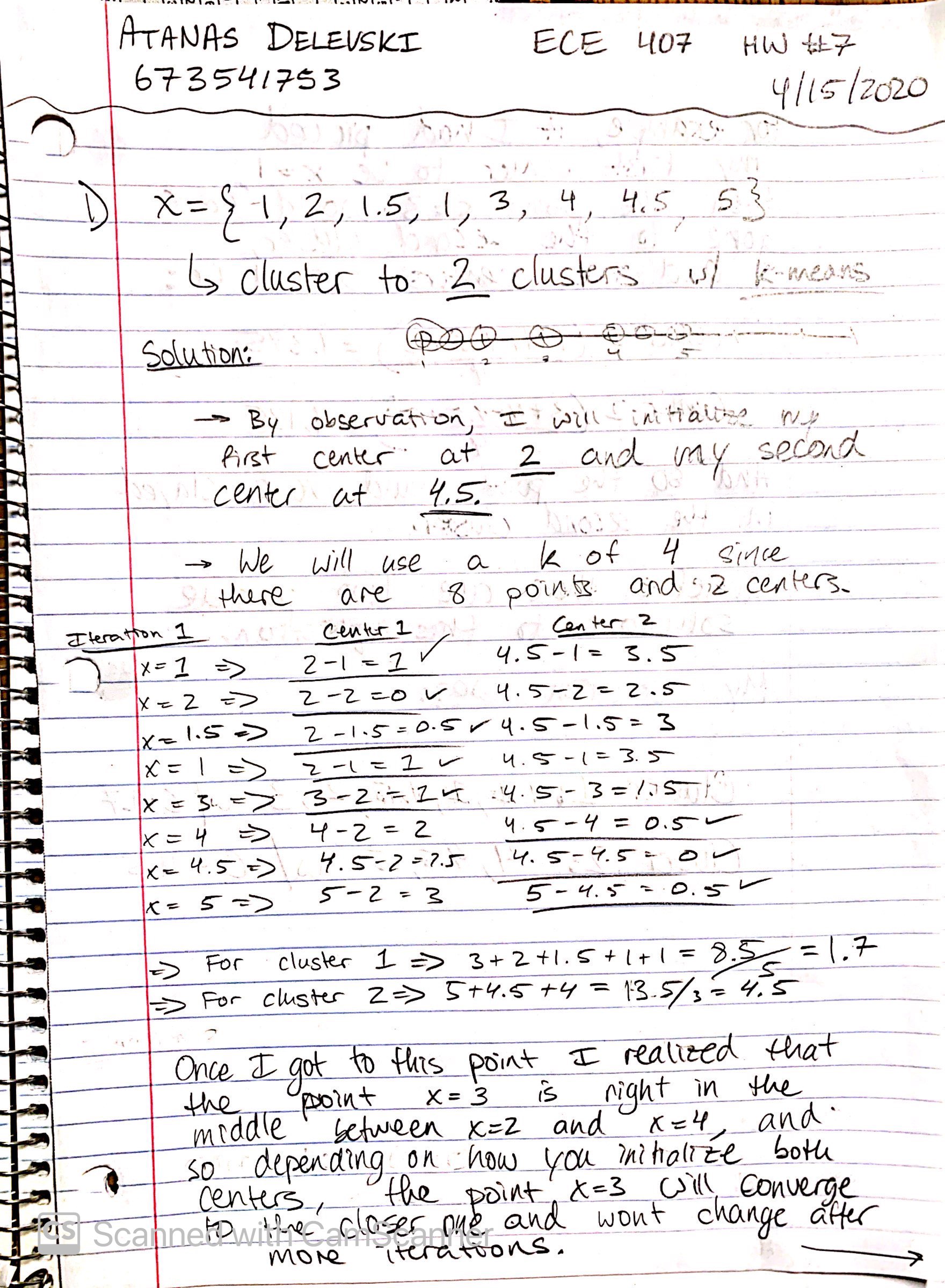
Atanas Delevski

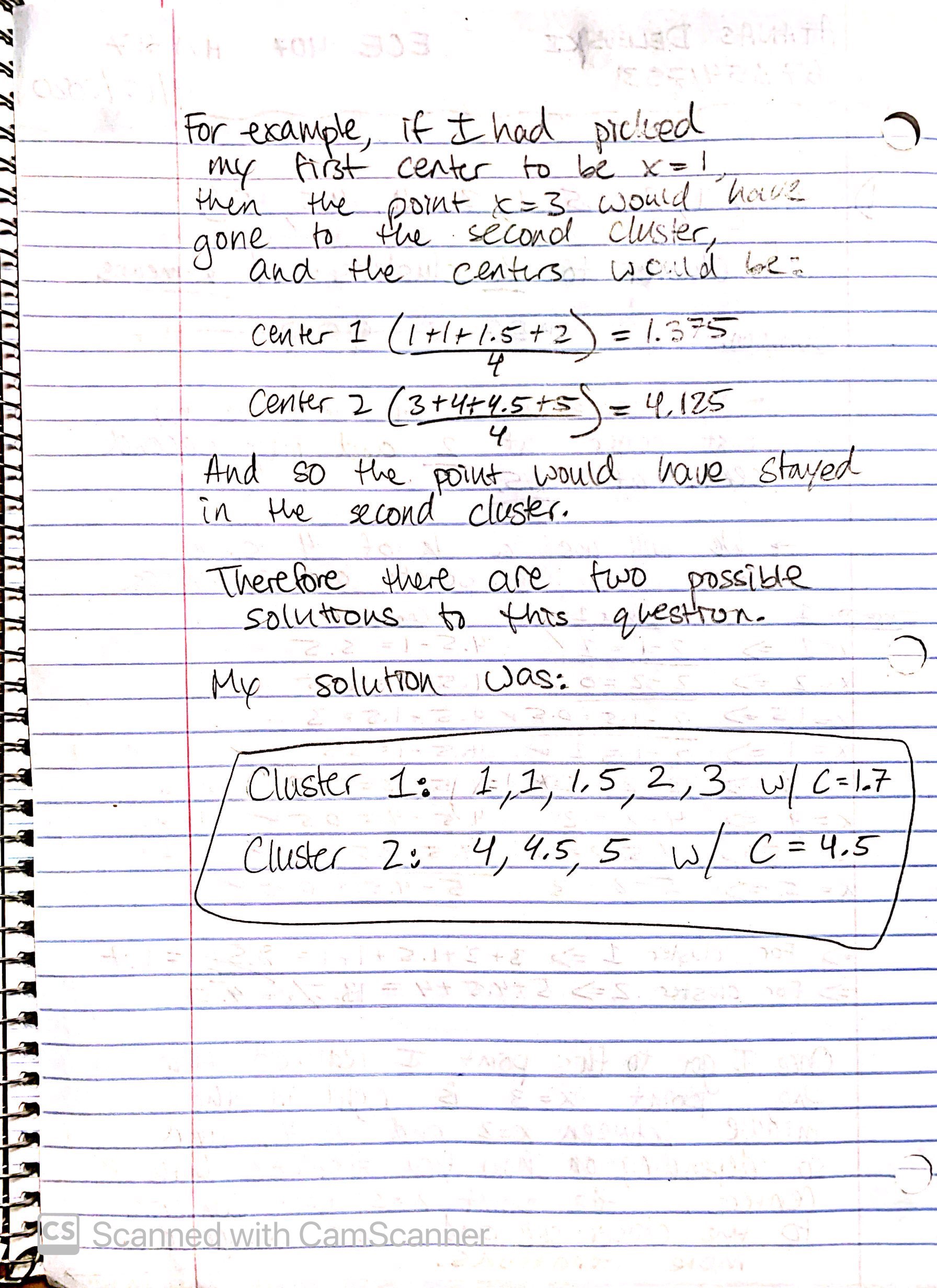
ECE 407 HW #7

4/15/2020

Problem 1: (I did on paper [2 pages]) (black box is covering a mistake)







Problem 2:

As in the previous homework, I decided to once again use Python for this problem. I used the same libraries as last time; **pandas** for data manipulation and **matplotlib** for data visualization.

At first I tried to program the K-means algorithm by myself but even after looking at multiple resources online, I found it way too complicated to turn into code even though I understood the algorithm.   
 Therefore, after hours of trying, I decided to just use the **scikit-learn** library in order to use it’s k-means function. I wish I could have gone without using it but it turned out to be too hard for this homework assignment.

After using the k-means clustering function from scikit-learn and setting the number of clusters to 4, I got these centers: (It is worth mentioning that K-Means algorithm has a downfall of falling into local minima in terms of optimization. This means that every time I ran the program, I got different centers. This is just one out of many different solutions depending on how the program first initializes the centers.)

|  |  |  |
| --- | --- | --- |
|  | X | Y |
| Center 1 (red) | 4.75 | 3.15 |
| Center 2 (green) | 6.352 | 2.92 |
| Center 3 (blue) | 5.2833 | 3.70833 |
| Center 4 (yellow) | 5.52 | 2.62 |

**Part B:** Given the test vector, (4.9, 6.2), this instance of the algorithm would classify that vector as Class 3 (blue).

