CSC148 Worksheet 7 Solution

Hyungmo Gu

April 19, 2020

Question 1

- Noticed that there are 11 students in total.
- Students should be grouped by year as closest as possible.

Notes:

- 형모 해낼 뚜 있쬬!
- 형모 화이팅!

Question 2

Name	Year	College	
Priya	3	Victoria	
Alain	2	New	Group 1
Zoe	3	Woodsworth	
Francesco	3	Victoria	
Mohammed	4	Woodsworth	
Xiaoyuan	5	New	Group 2
Rohit	2	New	
Yimin	3	Trinity	
Grace	5	Woodsworth	
Claire	1	Woodsworth	Group 3
Kai	1	Woodsworth	

Question 3

- First, we need to find the group as homogenous as possible in terms of year students are in.

The definition tells us group needs to be in 4, and the following table tell us there are 4 3^{rd} year students.

Student Year	Number of Students
1	2
2	2
3	4
4	1
5	2

It follows from these facts that the group of 3^{rd} year students best satisfy this criterion.

Next, we need to find the group as not homogenous as possible in terms of year students are in.

The same table tells us with 2 5^{th} year students, 1 4^{th} year students and 1 3^{rd} year student, a group spanning 3 years can be created.

Since we know there can't be a group spanning 4 years, we can conclude the group of 3 years (2 5^{th} year students, 1 4^{th} year students and 1 3^{rd} year student) best satisfy this criterion.

Correct Solution:

Group 1 is the most homogeneous.

Group 3 is the least homogeneous.

Question 4

- We will calculate the group score based on the code below. The code is also included in *worksheet_7_q4_solution.py*

```
    def get_group_score(group):
        """Evaluates the group score

            Precondition: len(group) == 4
        """
        n = 4

```

```
        max_year = get_max_year(group)
        min_year = get_min_year(group)
        similarity_list = []

        i = 0

        while i < 4:
            j = 0
            while j < 4:
                # find the scaled distance
                scaled_distance = 0
                if max_year != min_year:
                    scaled_distance = abs(group[i]['year'] - group[j]['year']) / float(max_year - min_year)

                # find the similarity
                similarity = 1 - scaled_distance

                # add to list
                similarity_list.append(similarity)

                j += 1
            i += 1

        # find the average
        average = float(sum(similarity_list))/len(similarity_list)
        return average.as_integer_ratio()

    def get_max_year(group):
        """returns max value of year in group"""

        max_value = -1

        for student in group:
            max_value = max(student['year'], max_value)

        return max_value

    def get_min_year(group):
        """returns min value of year in group"""

        min_value = 100 # this is impossible value

        for student in group:
            min_value = min(student['year'], min_value)

        return min_value


    if __name__ == '__main__':
        group_1 = [{'name': 'Primya', 'year': 3},
        {'name': 'Zoe', 'year': 3},
        {'name': 'Francesco', 'year': 3},
        {'name': 'Yimin', 'year': 3}]
```

```
61
62          group_2 = [{'name': 'Primya', 'year': 5},
63          {'name': 'Zoe', 'year': 5},
64          {'name': 'Francesco', 'year': 4},
65          {'name': 'Yimin', 'year': 3}]
66
67
68          score_1 =  get_group_score(group_1)
69          print(score_1)
70
71          score_2 =  get_group_score(group_2)
72          print(score_2)
73
74
75
```

For the most homogeneous group, the group score is 1.

For the least homogeneous group, the group score is $\frac{9}{16}$.

Correct Solution:

We will calculate the group score using the code below. The code is also included in *worksheet_7_q4_solution.py*

```
1      def get_similarity_list(group):
2          """Evaluates the list of similarities""" # <- Correct solution
3          similarity_list = []
4
5          i = 0
6
7          while i < len(group): # <- correct solution
8              j = i + 1 # <- correct solution
9              while j < len(group): # <- correct solution
10                 # find the scaled distance
11                 scaled_distance = abs(group[i]['year'] - group[j]['year
   ']) / 5.0 # <- correct solution
12
13                 # find the similarity
14                 similarity = 1 - scaled_distance
15
16                 # add to list
17                 similarity_list.append(similarity)
18
19                 j += 1
20             i += 1
21
22         return similarity_list
23
24
```

```
if __name__ == '__main__':
    # Correct Solution
    group_1 = [{'name': 'Primya', 'year': 3},
    {'name': 'Alain', 'year': 2},
    {'name': 'Zoe', 'year': 3},
    {'name': 'Francesco', 'year': 3}]

    # Correct Solution
    group_2 = [{'name': 'Mohammed', 'year': 4},
    {'name': 'Xiaoyuan', 'year': 5},
    {'name': 'Rohit', 'year': 2},
    {'name': 'Yimin', 'year': 3}]

    # Correct Solution
    group_3 = [{'name': 'Grace', 'year': 5},
    {'name': 'Claire', 'year': 1},
    {'name': 'Kai', 'year': 1}]

    # Correct Solution
    list_1 =  get_similarity_list(group_1)
    print(list_1)

    # Correct Solution
    list_2 =  get_similarity_list(group_2)
    print(list_2)

    # Correct Solution
    list_3 =  get_similarity_list(group_3)
    print(list_3)


```

We will calculate the score of each group in parts.

Part 1 (Score of Group 1):

We need to calculate the score of group 1.

The problem tells us the group score is

$$\text{Group Score} = \frac{\sum \text{similarities}}{\#\text{ of elements in list of similarities}} \tag{1}$$

and the above program tells us that the list of similarities are

$$[0.8, 1.0, 1.0, 0.8, 0.8, 1.0] \tag{2}$$

Then, using these facts, we can conclude that

$$\text{Group Score} = \frac{0.8 + 1.0 + 1.0 + 0.8 + 0.8 + 1.0}{6} \tag{3}$$
$$= \frac{\frac{4}{5} + 1 + 1 + \frac{4}{5} + \frac{4}{5} + 1}{6} \tag{4}$$
$$= \frac{\frac{4}{5} + \frac{5}{5} + \frac{5}{5} + \frac{4}{5} + \frac{4}{5} + \frac{5}{5}}{6} \tag{5}$$
$$= \frac{\frac{27}{4}}{6} \tag{6}$$
$$= \frac{27}{30} \tag{7}$$
$$= \frac{9}{10} \tag{8}$$

Part 2 (Score of Group 2):

We need to calculate the score of group 2.

The problem tells us the group score is

$$\text{Group Score} = \frac{\sum \text{similarities}}{\#\text{ of elements in list of similarities}} \tag{9}$$

and the above program tells us that the list of similarities are

$$[0.8, 0.6, 0.8, 0.4, 0.6, 0.8] \tag{10}$$

Then, we can conclude that the value of group score is

$$\frac{0.8 + 0.6 + 0.8 + 0.4 + 0.6 + 0.8}{6} = \frac{\frac{4}{5} + \frac{3}{5} + \frac{4}{5} + \frac{2}{5} + \frac{3}{5} + \frac{4}{5}}{6} \tag{11}$$
$$= \frac{\frac{20}{5}}{6} \tag{12}$$
$$= \frac{20}{30} \tag{13}$$
$$= \frac{2}{3} \tag{14}$$

Part 3 (Score of Group 3):

We need to calculate the score of group 3.

The problem tells us the group score is

$$\text{Group Score} = \frac{\sum \text{similarities}}{\#\text{ of elements in list of similarities}} \tag{15}$$

and the above program tells us that the list of similarities are

$$[0.2, 0.2, 1.0] \tag{16}$$

Then, we can conclude that the value of group score is

$$\frac{0.2 + 0.2 + 1.0}{3} = \frac{\frac{1}{5} + \frac{1}{5} + \frac{5}{5}}{3} \tag{17}$$

$$= \frac{\frac{7}{5}}{3} \tag{18}$$

$$= \frac{7}{15} \tag{19}$$

Notes:

- Realized that 'group' in question means 'group 1', 'group 2' and 'group 3' in question 2 :(.

Question 5

- We need to find the highest and the lowest possible score on this criterion.

  We will do so in parts.

  Part 1 (Calculating highest possible group score):

  We need to find the highest possible group score on this criterion.

  The definition of group score tells us that

$$\text{Group Score} = \frac{\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1}\left(1 - \frac{|student\_year_i - student\_year_j|}{5}\right)}{\#\text{ of elements in list of similarities}} \tag{1}$$

By observation, we can conclude that the highest possible group score happens when $|student\_year_i - student\_year_j| = 0$.

Then, using this fact, we can calculate that

$$\text{Group Score} = \frac{\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1}\left(1 - \frac{|student\_year_i - student\_year_j|}{5}\right)}{\#\text{ of elements in list of similarities}} \tag{2}$$

$$= \frac{\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1} 1}{\#\text{ of elements in list of similarities}} \tag{3}$$

Then, because we know $\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1} 1 = \#$ of elements in list of similarities, we can conclude

$$= \frac{\#\text{ of elements in list of similarities}}{\#\text{ of elements in list of similarities}} \tag{4}$$

$$= 1 \tag{5}$$

Part 2 (Calculating lowest possible group score):

We need to find the lowest possible group score achievable on this criterion.

The definition of group score tells us that

$$\text{Group Score} = \frac{\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1}\left(1 - \frac{|student\_year_i - student\_year_j|}{5}\right)}{\#\text{ of elements in list of similarities}} \tag{6}$$

By observation, we can conclude that the lowest possible group score happens when $|student\_year_i - student\_year_j| = 5$.

Then, using this fact, we can calculate that

$$\text{Group Score} = \frac{\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1}\left(1 - \frac{|student\_year_i - student\_year_j|}{5}\right)}{\#\text{ of elements in list of similarities}} \tag{7}$$

$$= \frac{\sum_{i=0}^{n-1}\sum_{j=i+1}^{n-1} 0}{\#\text{ of elements in list of similarities}} \tag{8}$$

$$= \frac{0}{\#\text{ of elements in list of similarities}} \tag{9}$$

$$= 0 \tag{10}$$

Question 6

Question 7

Question 8

Question 9

Question 10