

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 뚝 있쥬!
- 형모 화이팅!

Question 2

Name	Year	College	
Priya	3	Victoria	Group 1
Alain	2	New	
Zoe	3	Woodsworth	
Francesco	3	Victoria	
Mohammed	4	Woodsworth	Group 2
Xiaoyuan	5	New	
Rohit	2	New	
Yimin	3	Trinity	
Grace	5	Woodsworth	Group 3
Claire	1	Woodsworth	
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 3rd 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 3rd 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 5th year students, 1 4th year students and 1 3rd 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 5th year students, 1 4th year students and 1 3rd 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*

```

1  def get_group_score(group):
2      """Evaluates the group score
3
4          Precondition: len(group) == 4
5      """
6      n = 4
7

```

```

8     max_year = get_max_year(group)
9     min_year = get_min_year(group)
10    similarity_list = []
11
12    i = 0
13
14    while i < 4:
15        j = 0
16        while j < 4:
17            # find the scaled distance
18            scaled_distance = 0
19            if max_year != min_year:
20                scaled_distance = abs(group[i]['year'] - group[j][
'year']) / float(max_year - min_year)
21
22            # find the similarity
23            similarity = 1 - scaled_distance
24
25            # add to list
26            similarity_list.append(similarity)
27
28            j += 1
29            i += 1
30
31        # find the average
32        average = float(sum(similarity_list))/len(similarity_list)
33        return average.as_integer_ratio()
34
35    def get_max_year(group):
36        """returns max value of year in group"""
37
38        max_value = -1
39
40        for student in group:
41            max_value = max(student['year'], max_value)
42
43        return max_value
44
45    def get_min_year(group):
46        """returns min value of year in group"""
47
48        min_value = 100 # this is impossible value
49
50        for student in group:
51            min_value = min(student['year'], min_value)
52
53        return min_value
54
55
56    if __name__ == '__main__':
57        group_1 = [{'name': 'Primya', 'year': 3},
58                   {'name': 'Zoe', 'year': 3},
59                   {'name': 'Francesco', 'year': 3},
60                   {'name': 'Yimin', 'year': 3}]

```

```

61     group_2 = [{'name': 'Primya', 'year': 5},
62                {'name': 'Zoe', 'year': 5},
63                {'name': 'Francesco', 'year': 4},
64                {'name': 'Yimin', 'year': 3}]
65
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
12             ']) / 5.0 # <- correct solution
13
14             # find the similarity
15             similarity = 1 - scaled_distance
16
17             # add to list
18             similarity_list.append(similarity)
19
20             j += 1
21             i += 1
22
23     return similarity_list
24

```

```

25 if __name__ == '__main__':
26     # Correct Solution
27     group_1 = [{'name': 'Primya', 'year': 3},
28               {'name': 'Alain', 'year': 2},
29               {'name': 'Zoe', 'year': 3},
30               {'name': 'Francesco', 'year': 3}]
31
32     # Correct Solution
33     group_2 = [{'name': 'Mohammed', 'year': 4},
34               {'name': 'Xiaoyuan', 'year': 5},
35               {'name': 'Rohit', 'year': 2},
36               {'name': 'Yimin', 'year': 3}]
37
38     # Correct Solution
39     group_3 = [{'name': 'Grace', 'year': 5},
40               {'name': 'Claire', 'year': 1},
41               {'name': 'Kai', 'year': 1}]
42
43     # Correct Solution
44     list_1 = get_similarity_list(group_1)
45     print(list_1)
46
47     # Correct Solution
48     list_2 = get_similarity_list(group_2)
49     print(list_2)
50
51     # Correct Solution
52     list_3 = get_similarity_list(group_3)
53     print(list_3)
54
55

```

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

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

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

$$[0.8, 1.0, 1.0, 0.8, 0.8, 1.0] \quad (2)$$

Then, using these facts, we can conclude that

$$\text{Group Score} = \frac{0.8 + 1.0 + 1.0 + 0.8 + 0.8 + 1.0}{6} \quad (3)$$

$$= \frac{\frac{4}{5} + 1 + 1 + \frac{4}{5} + \frac{4}{5} + 1}{6} \quad (4)$$

$$= \frac{\frac{4}{5} + \frac{5}{5} + \frac{5}{5} + \frac{4}{5} + \frac{4}{5} + \frac{5}{5}}{6} \quad (5)$$

$$= \frac{\frac{27}{5}}{6} \quad (6)$$

$$= \frac{27}{30} \quad (7)$$

$$= \frac{9}{10} \quad (8)$$

Part 2 (Score of Group 2):

We need to calculate the score of group 2.

The problem tells us the group score is

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

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

$$[0.8, 0.6, 0.8, 0.4, 0.6, 0.8] \quad (10)$$

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

$$\frac{0.8 + 0.6 + 0.8 + 0.4 + 0.6 + 0.8}{6} = \frac{\frac{4}{5} + \frac{3}{5} + \frac{4}{5} + \frac{2}{5} + \frac{3}{5} + \frac{4}{5}}{6} \quad (11)$$

$$= \frac{\frac{20}{5}}{6} \quad (12)$$

$$= \frac{20}{30} \quad (13)$$

$$= \frac{2}{3} \quad (14)$$

Part 3 (Score of Group 3):

We need to calculate the score of group 3.

The problem tells us the group score is

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

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

$$[0.2, 0.2, 1.0] \quad (16)$$

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

$$\frac{0.2 + 0.2 + 1.0}{3} = \frac{\frac{1}{5} + \frac{1}{5} + \frac{5}{5}}{3} \quad (17)$$

$$= \frac{\frac{7}{5}}{3} \quad (18)$$

$$= \frac{7}{15} \quad (19)$$

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

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

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

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

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

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

$$= \frac{\# \text{ of elements in list of similarities}}{\# \text{ of elements in list of similarities}} \quad (4)$$

$$= 1 \quad (5)$$

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

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

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

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

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

$$= \frac{0}{\# \text{ of elements in list of similarities}} \quad (9)$$

$$= 0 \quad (10)$$

Notes:

- 형모야. 차분히. 한걸음 더.
- 문제 안풀린다고 주저 앞지마.
- 사랑하는 내 여보 향해 가는거야.
- 보고싶은 내 여보 향해.
- 형모야 화이팅 :)

Question 6

Question 7

Question 8

Question 9

Question 10