Data scientist assessment score

September 19, 2020

Problem Statement: Analytics Comapany conducted an assessment test to hire a Data Scientist. The candidates were evaluated on 5 different subjects A, B, C, D, and E. The marks of 5 different subjects out of 100 are given below. Help Analytics Vidhya to find out the answers of the following questions. Who scored the highest marks in the subject B? What is the average marks scored in the subject C? Who scored the highest percentage of marks? If considered only top-4 subjects of a candidate, then who scored the highest percentage of marks?

```
In [1]: # student marks
        student_marks = [['Name',  ['A','B','C','D','E']],
                         ['Ankit',  [41, 34, 45, 55, 63]],
                         ['Aravind',[42, 23, 34, 44, 53]],
                         ['Lakshay',[32, 23, 13, 54, 67]],
                         ['Gyan',   [23, 82, 23, 63, 34]],
                         ['Pranav', [21, 23, 25, 56, 56]]
                        ]
```

Who scored the highest marks in the subject B?

```
In [2]: # declear string variable to store person name who scored highest marks in B
        highest_B = None
        # declear string variable to store highest marks in B
        highest_B_marks = 0
        print(f"List of names with marks in subject B:")
        # loop to go through each item in list and find who scored highest in B
        for item in student_marks[1:]:
            print(item[0],item[1][1])
            if highest_B_marks < item[1][1]:
                highest_B_marks = item[1][1]
                highest_B = item[0]
            else:
                pass
        print(f"Highest scored person in subject B {highest_B, highest_B_marks}")
```

```
List of names with marks in subject B:
Ankit 34
Aravind 23
Lakshay 23
Gyan 82
```

```
Pranav 23
Highest scored person in subject B ('Gyan', 82)
```

What is the average marks scored in the subject C?

```
In [5]: # from statistics import mean function to calculate average
        from statistics import mean
        # declear empty list to store C marks
        avg_marks_C = []
        print(f"List of names with marks in subject C:")
        # loop to go through each item in list and store C marks in list avg_marks_C
        for item in student_marks[1:]:
            print(item[0],item[1][2])
            avg_marks_C.append(item[1][2])
        # print mean of marks in C
        print(f"Average marks in C: {mean(avg_marks_C)}")
```

```
List of names with marks in subject C:
Ankit 45
Aravind 34
Lakshay 13
Gyan 23
Pranav 25
Average marks in C: 28
```

```
In [6]: # function to calculate percentage
        def percentage(a, b):
            return round(a / b * 100, 2)
```

```
In [7]: # function to find top percentages
        def top_percent(total_marks, student_list, start_index, end_index, subject_number):
            # variable to store highest percentage
            highest_percent = 0
            # variable to store name who scored highest percentage
            person_scored = None
            print(f"List of names with percentages:")
            # loop to go through each item in list and find who scored highest percentage
            for item in student_list[1:]:
                total = sum(item[1][start_index:end_index])
                percent = percentage(total, total_marks)
                print(item[0], percent)
                if highest_percent < percent:
                    highest_percent = percent
                    person_scored = item[0]
                else:
                    pass
            print(f"Highest percentage of marks from {subject_number} subjects {person_scored,
```

Who scored the highest percentage of marks?

```
In [8]: # End index of marks list
        index_marks = len(student_marks[0][1])
        # number of subjects to find percentage
        subject_number = 5
        # find percentage
        top_percent(500, student_marks, 0, index_marks, subject_number)
```

```
List of names with percentages:
Ankit 47.6
Aravind 39.2
Lakshay 37.8
Gyan 45.0
Pranav 36.2
Highest percentage of marks from 5 subjects ('Ankit', 47.6)
```

If considered only top-4 subjects of a candidate, then who scored the highest percentage of marks?

```
In [9]: # End index of marks list
        index_first_4 = len(student_marks[0][1])-1
        # number of subjects to find percentage
        subject_number = 4
        # find percentage
        top_percent(400, student_marks, 0, index_first_4, subject_number)
```

```
List of names with percentages:
Ankit 43.75
Aravind 35.75
Lakshay 30.5
Gyan 47.75
Pranav 31.25
Highest percentage of marks from 4 subjects ('Gyan', 47.75)
```