

Fundamental Exam Data Science

Score (45)

1. You are now a data scientist in a company, your first task is to detect the outlier data by calculate the IQR (Interquartile Range).
For example I have data = [71, 70, 73, 70, 70, 69, 70, 72, 71, 300, 71, 69].
To identify the outlier data , read the instruction below.
 - a. Sort the data , data = [69, 69, 70, 70, 70, 70, 71, 71, 71, 72, 73, 300]
 - b. Find the median = $70 + 71 / 2 = 70.5$
 - c. After finding the median, you have to find the Q1 and Q3 data. Because the example data length is 12, so the Q1 is equal to the median of first six data. And Q3 is equal to the median of the last six data.
 - d. $Q1 = 70 + 70 / 2 = 70$
 - e. $Q3 = 71 + 72 / 2 = 71.5$
 - f. $IQR = Q3 - Q1 = 1.5$
 - g. Find the lower limit and the upper limit, lower limit = $Q1 - 1.5 \times IQR$, upper limit = $Q3 + 1.5 \times IQR$
 - h. if a data point is below the lower limit or the data point is above the upper limit, then the data point is outlier

`remove_outlier([71, 70, 73, 70, 70, 69, 70, 72, 71, 300, 71, 69]) => [71, 70, 73, 70, 70, 69, 70, 72, 71, 71, 69]`

the answer must be the picture below,

```
data asli = [71, 70, 73, 70, 70, 69, 70, 72, 71, 300, 71, 69]
data setelah di sort = [69, 69, 70, 70, 70, 70, 71, 71, 71, 72, 73, 300]
setengah data pertama = [69, 69, 70, 70, 70, 70]
setengah data terakhir = [71, 71, 71, 72, 73, 300]
q1 adalah = 70.0
q3 adalah = 71.5
lower limit adalah = 67.75
upper limit adalah = 73.75
data yang tidak outlier = [71, 70, 73, 70, 70, 69, 70, 72, 71, 71, 69]
```

Another example input and output

```
data asli = [60, 63, 64, 62, 69, 80, 1, 60, 63, 64, 60]
data setelah di sort = [1, 60, 60, 60, 62, 63, 63, 64, 64, 69, 80]
setengah data pertama = [1, 60, 60, 60, 62]
setengah data terakhir = [63, 64, 64, 69, 80]
q1 adalah = 60
q3 adalah = 64
lower limit adalah = 54.0
upper limit adalah = 70.0
data yang tidak outlier = [60, 63, 64, 62, 69, 60, 63, 64, 60]
```

Score(20)

2. Return the number (count) of vowels in the given string. We will consider a, e, i, o, and u as vowels for this case. The input string will only consist of lower case letters and/or spaces.

countVowel('budi pergi ke pasar') → 7

countVowel('purwadhika') → 4

Score(30)

3. Given a two-dimensional list of integers, return the flattened version of the list with all the integers in the sorted (ascending) order.

given([[3, 2, 1], [4, 6, 5], [], [9, 7, 8]]) → [1, 2, 3, 4, 5, 6, 7, 8, 9]

given([[3,4,2,1], [1,2,3], [5,4,3,1]]) → [1,1,1,2,2,3,3,3,4,4,5]

Score(30)

4. Given an list of digital numbers, return a new list of length number containing the last even numbers from the original list (in the same order).

The function contain two parameter, list and integer.

findEven([1, 2, 3, 4, 5, 6, 7, 8, 9], 3) → [4, 6, 8]

findEven([-22, 5, 3, 11, 26, -6, -7, -8, -9, -8, 26], 2) → [-8, 26]

findEven([6, -25, 3, 7, 5, 5, 7, -3, 23], 1) => [6]

Score(30)

5. Given string of words, you have to calculate the count of every word in the sentences

countWords('jangan jangan kamu adalah aku')
should return

```
Jumlah kata 'Jangan' ada sebanyak 2
Jumlah kata 'Kamu' ada sebanyak 1
Jumlah kata 'Adalah' ada sebanyak 1
Jumlah kata 'Aku' ada sebanyak 1
```