Day 1: Interquartile Range

5 more challenges to get your next star!

Points: 5/10

10 Days of Statistics

Problem Submissions Leaderboard Editorial

Objective

In this challenge, we practice calculating the *interquartile range*. We recommend you complete the Quartiles challenge before attempting this problem.

Task

The interquartile range of an array is the difference between its first (Q_1) and third (Q_3) quartiles (i.e., $Q_3 - Q_1$).

Given an array, X, of n integers and an array, F, representing the respective frequencies of X's elements, construct a data set, S, where each x_i occurs at frequency f_i. Then calculate and print S's interquartile range, rounded to a scale of 1 decimal place (i.e., 12.3 format).

Tip: Be careful to not use integer division when averaging the middle two elements for a data set with an even number of elements, and be sure to *not* include the median in your upper and lower data sets.

Input Format

The first line contains an integer, n, denoting the number of elements in arrays X and F.
The second line contains n space-separated integers describing the respective elements of array X.
The third line contains n space-separated integers describing the respective elements of array F.

Constraints

- $5 \leq n \leq 50$
- $0 < x_i \leq 100$, where x_i is the i^{th} element of array X.
- $0 < \sum_{i=0}^{n-1} f_i \leq 10^3$, where f_i is the i^{th} element of array F.
- The number of elements in S is equal to $\sum F$.

Output Format

Print the *interquartile range* for the expanded data set on a new line. Round your answer to a scale of 1 decimal place (i.e., 12.3 format).

Sample Input

```
6
6 12 8 10 20 16
5 4 3 2 1 5
```

Sample Output

```
9.0
```

Explanation

The given data is:

Element	Frequency
6	5
12	4
8	3
10	2
20	1
16	5

First, we create data set S containing the data from set X at the respective frequencies specified by F:

$$S = \{6, 6, 6, 6, 6, 8, 8, 8, 10, 10, 12, 12, 12, 12, 16, 16, 16, 16, 16, 20\}$$

As there are an even number of data points in the original ordered data set, we will split this data set exactly in half:

> Lower half (L): 6, 6, 6, 6, 6, 8, 8, 8, 10, 10
>
> Upper half (U): 12, 12, 12, 12, 16, 16, 16, 16, 16, 20

Next, we find Q_1. There are 10 elements in $lower$ half, so $Q1$ is the average of the middle two elements: 6 and 8. Thus, $Q_1 = \frac{6+8}{2} = 7.0$.

Next, we find Q_3.There are 10 elements in $upper$ half, so $Q3$ is the average of the middle two elements: 16 and 16. Thus, $Q_3 = \frac{16+16}{2} = 16.0$.

From this, we calculate the interquartile range as $Q_3 - Q_1 = 16.0 - 7.0 = 9.0$ and print 9.0 as our answer.

Python 3

```python
# Enter your code here. Read input from STDIN. Print output to STDOUT

from statistics import median

size = int(input())
numbers = list(map(int, input().split()))
frequency = list(map(int, input().split()))

s=[]
for i in range(size): s += ([numbers[i]] * frequency[i])
s.sort()

t=len(s)//2
if len(s)%2==0:
    L=s[:t]
    U=s[t:]
else:
    L=s[:t]
    U=s[t+1:]

print(round(float(median(U) - median(L)),1))

```

Line: 13 Col: 12

Upload Code as File

Test against custom input

Run Code

Submit Code