Day 7: Temperature Predictions

Objective
In this challenge, we practice predicting values. Check out the Resources tab for some tips on approaching this problem.

Task
Given a record containing the maximum and minimum monthly temperatures at a particular station. The record shows the temperature information for each month in a data range from $\textbf{January 1908}$ to $\textbf{March 2012}$; however, some of the temperature values have been blanked out! Estimate and print the missing values.

Input Format

The first line contains an integer, N, denoting the number of rows of data in the input file.
The second line contains the header for the tab-separated file; this line can be ignored, and is simply there to make the test case easier to read.
The N subsequent lines each describe the respective $year$, $month$, $maximum\ temperature$, and $minimum\ temperature$ data as a row of tab-separated values. In some of the rows, The $minimum$ or $maximum$ temperature field has been blanked out and replaced by: $Missing\_1$, $Missing\_2$, etc.

Constraints

- $1 \leq N \leq 1500$
- $1908 \leq Year \leq 2013$
- $-75 \leq Maximum/Minimum\ Temperature \leq 75$
- $\textbf{The } month \textbf{ field will always contain a fully spelled out month (January - December).}$
- $\textbf{Data is provided for each month in chronological order.}$
- $\textbf{Data may be missing for certain months or years.}$

Scoring

The score seen upon hitting $\textbf{Run Code}$ is the score against the sample test case (of 444 rows) only. It is normalized and will always lie between 0 and 1.
Upon hitting $\textbf{Submit Code}$, the score seen is determined *solely on the basis of the hidden test case*.

Details on the Scoring Formula

We compute the average of the magnitude of difference between your predicted value and the actual recorded value for each of the missing terms. If this average exceeds 5.0 degrees, you will be assigned a score of zero.

For each of the values predicted by you (p), we will compute an $error$. The $error$ is the difference of the predicted value (p) and the actual temperature at that location. Hence,
$error = |p - actualTemperatureRecording|$
$M = \textbf{total missing values in the test case}$

We will compute the *average* of all these error terms over all rows of data in the input file, and record it as

$avg = \frac{\sum errors}{M}$

Your score for this challenge will be $(1 - \frac{avg}{5.0}) \times (MaxScore)$
Here, $MaxScore = \text{Total number of points assigned to this challenge}$.

Output Format

Print each missing value on a new line.

Sample Input

```
20
yyyy   month   tmax   tmin
1908   January 5.0 -1.4
1908   February   7.3 1.9
1908   March   6.2 0.3
1908   April   Missing_1   2.1
1908   May Missing_2   7.7
1908   June   17.7   8.7
1908   July   Missing_3   11.0
1908   August  17.5   9.7
1908   September   16.3   8.4
1908   October 14.6   8.0
1908   November   9.6 3.4
1908   December   5.8 Missing_4
1909   January 5.0 0.1
1909   February   5.5 -0.3
1909   March   5.6 -0.3
1909   April   12.2   3.3
1909   May 14.7   4.8
1909   June   15.0   7.5
1909   July   17.3   10.8
1909   August  18.8   10.7
```

The above test case is for explanatory purposes only, which is why we included only 20 lines.
The sample test case, which is run upon hitting Run Code, has 444 rows of data.
The hidden test case, which is used at the time of submission, has over $1.2k$ rows of data. The sample test case rows are a subset of it.

Sample Output

The four missing values ($Missing\_1$, $Missing\_2$, $Missing\_3$, and $Missing\_4$) are:

```
8.6
15.8
18.9
0.0
```

Your task is to predict values as close as possible to these.