

2H Implement DistanceBetweenPattern ndStrings

Distance Between Pattern and Strings Problem

Compute *DistanceBetweenPattern ndStrings*.

Input: A DNA string *Pattern* and a collection of DNA strings *Dna*.

Output: Distance $D(\text{Pattern}, Dna)$ between *Pattern* and *Dna*.

$$d(\text{AAA}, \begin{array}{c} \text{CTTAAAC} \\ \text{GATATC} \\ \text{ACGGCG} \\ \text{CTAAAG} \end{array}) = 4$$

Formatting

Input: A DNA string *Pattern*, followed by a space-separated collection of DNA strings *Dna*.

Output: An integer representing the output of `DISTANCEBETWEENPATTERNANDSTRINGS(Pattern, Dna)`.

Constraints

- The length of *Pattern* will be between 1 and 10^1 .
- The number of strings in *Dna* will be between 1 and 10^2 .
- The length of each string in *Dna* will be between 1 and 10^2 .
- *Pattern* and each string in *Dna* will be DNA strings.

Test Cases

Case 1

Description: This dataset checks that you are able to identify mismatches between *Pattern* and individual sequences in *Dna* when the mismatch occurs in the beginning, middle, and end of the sequence.

Input:

```
T C G
```

Output:

```
3
```

Case 2

Description: This dataset checks that you are able to identify perfect matches of *Pattern* within individual sequences of *Dna* when *Pattern* occurs in the beginning, middle, and end of the sequence.

Input:

```
G
G TT CG T TTG
```

Output:

```
0
```

Case 3

Description: This dataset checks your solution works on a standard input.

Input:

```
CG
GTG CC TTT CG C CC
```

Output:

```
2
```

Case 4

Description: This dataset checks your solution works on a standard input.

Input:

GGGG

GGGCTC C GGGTT C GG GT C T G GGT T C

Output:

4

Case 5

Description: A larger dataset of the same size as that provided by the randomized autograder.