

## 1G Compute the Hamming Distance Between Two Strings

---

### Hamming Distance Problem

*Compute the Hamming distance between two strings.*

**Input:** Two strings of equal length.

**Output:** The Hamming distance between these strings.

TCTGAAC  
TCCGACC  
1 2

---

### Formatting

**Input:** Two DNA strings  $Text_1$  and  $Text_2$ .

**Output:** An integer representing the Hamming distance between  $Text_1$  and  $Text_2$ .

### Constraints

- The length of  $Text_1$  and  $Text_2$  will be between 1 and  $10^4$ .
- $Text_1$  and  $Text_2$  will have equal lengths.
- $Text_1$  and  $Text_2$  will be DNA strings.

## Test Cases

### Case 1

---

**Description:** The sample dataset is not actually run on your code.

**Input:**

GGGCCGTTGGT  
GGACCGTTGAC

**Output:**

3

### Case 2

---

**Description:** This dataset checks if your code isn't keeping count (i.e. returns 0 when the answer is clearly nonzero) or if your code returns a negative value, which is impossible.

**Input:**

AAAA  
TTTT

**Output:**

4

### Case 3

---

**Description:** This dataset checks if your code is finding Edit Distance (which would be 2) instead of Hamming Distance.

**Input:**

ACGTACGT  
TACGTACG

**Output:**

8

#### Case 4

---

**Description:** This dataset checks if your code is returning the number of matches (2) instead of the number of mismatches (6).

**Input:**

ACGTACGT  
CCCCCCCC

**Output:**

6

#### Case 5

---

**Description:** This dataset checks if your code works on a dataset where the two input strings have no matches.

**Input:**

ACGTACGT  
TGCATGCA

**Output:**

8

#### Case 6

---

**Description:** This dataset checks if you have an off-by-one error at the beginning (i.e. you are starting at the second character of the strings instead of the first character).

**Input:**

GATAGCAGCTTCTGAACTGGTTACCTGCCGTGAGTAAATTAAATTTATTGACTTAGGTCACTAAATACT  
AATAGCAGCTTCTCAACTGGTTACCTCGTATGAGTAAATTAGGTCATTATTGACTCAGGTCACTAACGTCT

**Output:**

15

### Case 7

---

**Description:** This dataset checks if you have an off-by-one error at the end (i.e. you are ending at the second-to-last character of the strings instead of the last character).

**Input:**

```
GATAGCAGCTTCTGAACTGGTTACCTGCCGTGAGTAAATTAAAATTTATTGACTTAGGTCACTAAATACT
AATAGCAGCTTCTCAACTGGTTACCTCGTATGAGTAAATTAGGTCATTATTGACTCAGGTCACTAACGTCT
```

**Output:**

28

### Case 8

---

**Description:** A larger dataset of the same size as that provided by the randomized autograder. Check input/output folders for this dataset.