

Set Operations

Problem ID: a15p02setoperations

Write a program that, given two files, collects the words from each file removing all punctuations, then computes and prints the following metrics using set operations:

- Common words in both files.
- Words that only occur in one file but not both.
- Words that only appear in the first file.

For each of these, you should both print how many words there are, and the words themselves in alphabetical order. Note: You should remove any punctuation from the start and end of each word (e.g. “Transylvania,” should become “Transylvania”).

Input

Input consists of two lines. Both lines each contain one filename, the files that should be compared.

Output

Output consists of three sections. Each section starts with a specific string, they are as follows:

1. First section starts with, “The two files have $\{n_1\}$ words in common.”, where n_1 is the number of words the two files have in common.
2. Second section starts with, “There are $\{n_2\}$ words that are only in either file but not both.” where n_2 is the number of words present in one but not both files.
3. Third section starts with, “There are $\{n_3\}$ words that only appear in the first file.”, where n_3 is the number of words exclusive to only the first file.

Each of these strings is then followed by the string “These words are as follows:”, unless the number of words in question is $n_i = 0$, in which case no additional informations should be given for that particular section. The above line should be followed by an enumeration of the relevant words; the words that the two files have in common, that are present in only one file, or only present in the first file, depending on the section. In the enumeration, each word should be separated by a comma and a space, except the final word, which should be separated from the previous one with an “ and ”.

Sample Input 1

```
pangram_1.txt
pangram_2.txt
```

Sample Output 1

```
The two files have 1 words in common.
These words are as follows:
The

There are 13 words that are only in either file but not both.
These words are as follows:
boxing, brown, dog, five, fox, jump, jumps, lazy, over, quick, quickly, the and wizards.

There are 5 words that only appear in the first file.
These words are as follows:
boxing, five, jump, quickly and wizards.
```

Sample Input 2

```
dracula_1.txt  
dracula_2.txt
```

Sample Output 2

The two files have 34 words in common.

These words are as follows:

I, a, able, and, are, as, at, but, by, could, had, have, here, in, is, it, known, my, no,

There are 217 words that are only in either file but not both.

These words are as follows:

Before, Bistritz, Borgo, British, Bukovina, Carpathian, Carpathians, Castle, Count, Dracula,

There are 76 words that only appear in the first file.

These words are as follows:

Bistritz, British, Bukovina, Carpathian, Castle, Count, Dracula, Europe, Having, London, M