

Question #3: Sentiment score

Objective: String and file

Write a Python program to calculate the sentiment score based on two lists of sentiment type words: positive and negative. The program receives text input from the user and matches each word to its sentiment type. The lists of negative and positive words are stored in two separate files named "neg.txt" and "pos.txt", as shown in the figure below respectively. Words that do not occur in either list are counted as neutral.

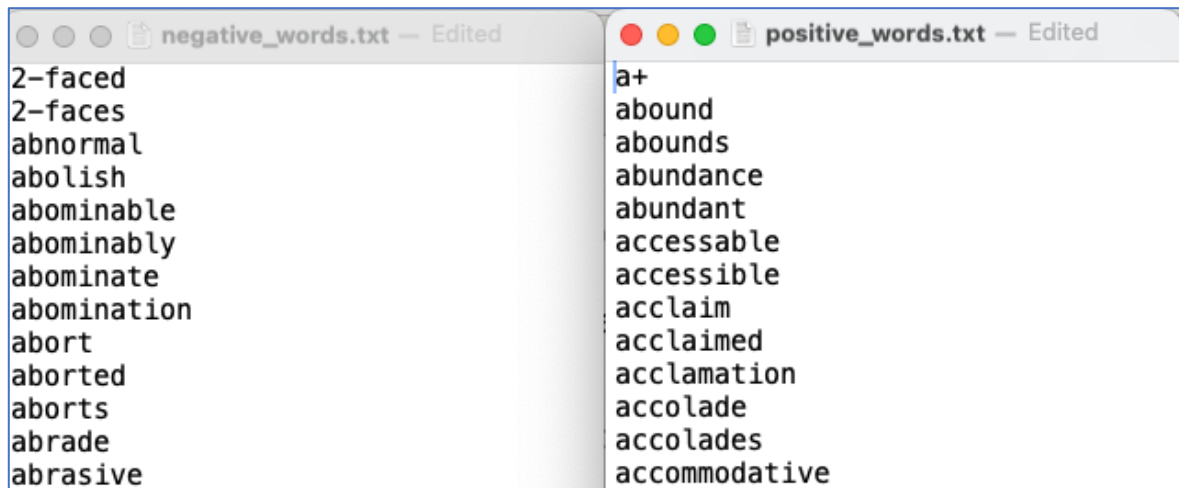


Figure 1 (Left) neg.txt and (Right) pos.txt

Sentiment scores are calculated by deducting the number of negative words (each negative word is assigned -1) from the number of positive words (each positive word is assigned +1). Words that are not found in either the positive or negative word lists are considered neutral and do not affect the score. The sentiment score is categorized as follows:

- **Negative:** Score is lower than 0
- **Neutral:** Score is equal to 0
- **Positive:** Score is higher than 0

Example:

- **Input:** "Peter is HANDSOME!, but arrogance and greedy."
- **Cleaned input:** "peter is handsome but arrogance and greedy"
- **Analysis:**
 - Total words: 7 words
 - Positive words: 1 word (handsome) – found in the positive word list
 - Negative words: 2 words (arrogance, greedy) – found in the negative word list
 - Neutral words: 5 words – not found in either list

- **Sentiment score** = 1 (positive) – 2 (negatives) = -1
- **Sentiment:** Negative (based on the score)

The program should return:

- The total count of words in the text.
- The number of words in each sentiment type (positive, negative, neutral).
- The calculated sentiment score.
- The text's overall sentiment type (positive, negative, or neutral).

We have provided the resource zip file (as shown in the figure below), which contains three files: pos.txt, neg.txt, and example_input.txt. Please place pos.txt and neg.txt in the same directory as your Python code. To submit your answer, you just need to submit only your Python code.



INPUT

There are two input lines:

- 1st Line: “the positive file name” “the negative file name”, e.g., “pos.txt neg.txt”.
- 2nd Line: A single line of text for sentiment score calculation. The input may contain non-alphabetic characters. Make sure to remove those characters before processing the text.

Hint: Use `isalpha()` function to check if a character is an alphabet.

OUTPUT

The program should return two lines of text processing results. For example, given the input: “Peter is HANDSOME!, but arrogance and greedy.”

Line 1: Displays the total number of words and the count of words in each sentiment category (positive, negative, neutral). For example: “7 1 2 5” represents Total:7 POS:1 NEG:2 NEU:5.

Line 2: Displays the overall sentiment score along with its label. For example: “-1 Negative” represents a sentiment score of -1, which is labeled as Negative.

EXAMPLES

Positive words are highlighted in green, while negative words are highlighted in yellow. Since the input text is quite lengthy, we've provided 'example_input.txt' for you to copy all the inputs.

#	Input (from keyboard)	Output (on-screen)
1	pos.txt neg.txt Python is very super easy but the exam is difficult	10 2 1 7 1 Positive
2	pos.txt neg.txt She is cuTe and loVely ^^...,	5 2 0 3 2 Positive
3	pos.txt neg.txt I hate python so badly	5 0 2 3 -2 Negative
4	pos.txt neg.txt Peter is HANDSOME !, but arrogance and greedy .	7 1 2 4 -1 Negative
5	pos.txt neg.txt I love this movie so bad !	6 1 1 4 0 Neutral
6	pos.txt neg.txt Today is exam day!?!	4 0 0 4 0 Neutral
7	pos.txt neg.txt	0 0 0 0 0 Neutral

#	Description for each case
1	Total:10 POS: 2 NEG: 1 NEU:7 Score: 1 (Positive)
2	Total:5 POS: 2 NEG:0 NEU:3 Score: 2 (Positive)
3	Total:5 POS:0 NEG: 2 NEU:3 Score: -2 (Negative)
4	Total:7 POS: 1 NEG: 2 NEU:4 Score: -1 (Negative)
5	Total:6 POS: 1 NEG: 1 NEU:4 Score: 0 (Neutral)
6	Total:4 POS:0 NEG:0 NEU:4 Score: 0 (Neutral)
7	Total:0 POS:0 NEG:0 NEU:0 Score: 0 (Neutral)

Test Cases in Grader

Testcases will be grouped. Each group has the following criteria:

Testcases quantity	Testcase characteristics
30%	Positive sentiment (without & with cleaned text)
30%	Negative sentiment (without & with cleaned text)
40%	Neutral sentiment (without & with cleaned text) + empty string