Crawler Coding Challenge

The goal of this challenge is to implement a console application that displays the most common words used in a portion of a webpage.

# Requirements

The code should be written in C#, Python, or Java.

The code should return the most common words used and the number of times they are used. The following should be configurable:

* The number of words to return (default: 10)
* Words to exclude from the search

Your code (only the source code, no binaries) should be returned as a zip directly to my email. The code should build into an executable console application.

# Page to crawl

<https://en.wikipedia.org/wiki/Microsoft>

Only words from the section “history” should be accounted for.

# Example of the expected result

|  |  |
| --- | --- |
|  | # of occurrences |
| The | 205 |
| Microsoft | 113 |
| in | 110 |
| of | 88 |
| and | 88 |
| a | 81 |
| to | 79 |
| on | 59 |
| Windows | 55 |
| for | 50 |