

Word Frequency

Write a program to fetch a Wikipedia page and report the top n words on that page.

The parameters to the program should be 'page_id' and 'n'.

The URL to fetch a page from Wikipedia is

`https://en.wikipedia.org/w/api.php?action=query&prop=extracts&pageids=[PAGE ID]&explaintext&format=json` (replacing [PAGE ID] with the requested page id). This returns a JSON object where the page's text is available under `query.pages.[PAGE ID].extract` and the title under `query.pages.[PAGE ID].title`.

The output should look like this (where `n = 5` and `page_id = 21721040`):

URL:

`https://en.wikipedia.org/w/api.php?action=query&prop=extracts&pageids=21721040&explaintext&format=json`

Title: Stack Overflow

Top 5 words:

- 21 stack, questions
- 16 overflow
- 11 that, users
- 10 site
- 7 answers, question

When two or more words have the same frequency, include them all on the same line separated by a comma.

A word is defined as a sequence of at least four alphabetic characters.

Include unit tests and a README describing how to compile (if necessary) and run your solution.

NOTE: Wikipedia pages change all the time so the results of this example may differ for you

Confidentiality

This assessment is intended for the recipient only so we request that it is treated in a confidential manner and not share or released on any 3rd party sites.