## The technical specification

The following is extracted from the exercise document:

1. The program accepts a **sentence**.
2. The program returns statistics of the **words** in the **sentence.**

## Data dictionary:

**Sentence:** A string ending one of the following characters: **. ? !** (ASCII decimal code 46,63, 33)

**Word:** Set of characters delimited with white space characters inside a sentence, excluding numeric representations (363.23 as example) and ones starting or ending with non-alpha numeric characters ($23 , #/ , +alpha as example). Punctuation marks are removed from the end and quotations marks from both sides.  
Lower or uppercase are not differentiated.

## Assumptions:

The data dictionary definitions above are based upon basic English language structure, no attempt to made filtering out extended or multi byte Unicode characters not compatible with this structure.

The design will consider accommodating possible future extended use case as checking a whole book, but current implementation is based on the described use case.

### The non-functional requirements

Based on the current use case:

Average length of a sentence is couple of hundred characters, max length is 1000 characters.  
The program can operate outside these limits, but won’t be tested and not guaranteed to work.