**Introduction**

http://quotes.toscrape.com/

This website is a static website that lists several quotes, tags and authors

Our desire as the collaborators on this project is to scrap the website and obtain words, common tags and what authors have the most quotes so that we carry out an analysis on what words were most common, how often words were mentioned, what are most common tags, and what authors have the most quotes.

**Description of scrapper mechanism**

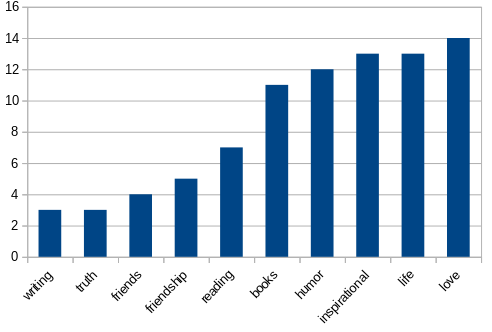
**Description of output**

The output is 7 csv files

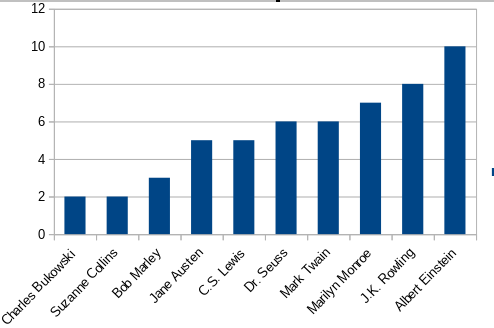
1. authors.csv shows the author names and the number of times the author appears on the website
2. quotes.csv shows the quote, the name of author and the tags associated with it
3. tags.csv shows the tags and the number of times the tag appears on the website
4. top\_authors shows the top author of quotes
5. top\_tags shows the top common tags
6. top\_words shows the top common words
7. words.csv shows the frequency of occurrence of words

**Data analysis**

Top 10 Tags

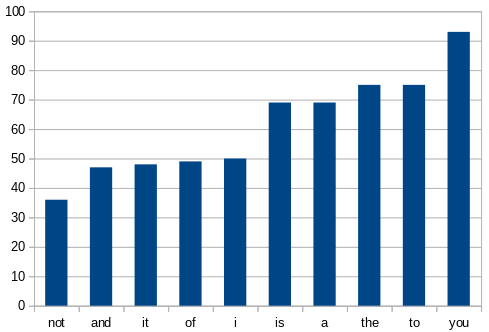


This is a simple bar chart that shows the number of times a word appears in tags

Top 10 Authors

This shows that the top 10 authors of quotes

Top 10 Words



This shows the top 10 most common words in all quotes

**Part allocation**

1. Beautiful Soup - Sandile Kholisani Nhlalo - Sibanda
2. Selenium - Sandile Kholisani Nhlalo - Sibanda
3. Scrapy - Nomthunzi Moyo
4. Analysis - Nomthunzi Moyo
5. Project file compilation – Nomthunzi Moyo