----Problem 1a----

['colorless', 'green', 'ideas', 'sleep', 'furiously']

----Problem 1b----

ordlu

----Problem 1c----

['colorless', 'green', 'ideas']

----Problem 1d----

colorless green ideas

----Problem 1e----

colorless

furiously

green

ideas

sleep

----Problem 2----

Enter a string

hello hi there hello

Word: hello Frequency: 2

Word: hi Frequency: 1

Word: there Frequency: 1

Words in sentence: ['hello', 'hi', 'there', 'hello'] Total no. of words: 4

----Problem 3----

Regex for locating determiners is: \b\s(a|an|the)\s

Regex for capturing arithmetic expression is \d\*[\\*|\+]\d\*[\\*|\+]\d+

----Problem 4----

Following is the text

austen-emma.txt:hart@vmd.cso.uiuc.edu (internet) hart@uiucvmd (bitnet)... austen-emma.txt:Internet (72600.2026@compuserve.com); TEL: (212-254-5093) ... austen-persuasion.txt:Editing by Martin Ward (Martin.Ward@uk.ac.durham)...blake-songs.txt:Prepared by David Price, email ccx074@coventry.ac.uk ...

Below is the list of email ids

hart@vmd.cso.uiuc.edu

hart@uiucvmd

72600.2026@compuserve.com

Martin.Ward@uk.ac.durham

ccx074@coventry.ac.uk

----Problem 5----

Enter filename:file.txt

['99!', '99.', '88.', '77.']

----Problem 6----

The provided code reads a file. If a blank line or a '#' is encountered at the start of the line, the program terminates. The re.sub command replaces the opening braces () with '(.' and ').'. The final for loop returns text present within the parenthesis.

So, the input file should contain text present within the parenthesis, so that the text between them could be retrieved

----Problem 7a----

Please wait...

Statements assigned to speakers

----Problem 7b & 7c----

Please wait...

Text preprocessed

---\* Top words used by Lehrer and its frequencies \*---

Porter Stemmer

[('governor', 25), ('two', 24), ('go', 23), ('minut', 22), ('presid', 21), ('right', 20), ('let', 20), ('govern', 15), ('mr', 15), ('first', 14)]

Snowball Stemmer

[('governor', 25), ('two', 24), ('go', 23), ('minut', 22), ('presid', 21), ('right', 20), ('let', 20), ('govern', 15), ('mr', 15), ('first', 14)]

Lancaster Stemmer

[('govern', 40), ('presid', 26), ('two', 24), ('minut', 22), ('right', 20), ('let', 20), ('mr', 15), ('first', 14), ('go', 13), ('seg', 12)]

---\* Top words used by Romney and its frequencies \*---

Porter Stemmer

[('peopl', 72), ('tax', 69), ('get', 69), ('go', 67), ('that', 46), ('presid', 45), ('plan', 42), ('cut', 40), ('one', 34), ('said', 33)]

Snowball Stemmer

[('peopl', 72), ('tax', 69), ('get', 69), ('go', 67), ('that', 46), ('presid', 45), ('plan', 42), ('cut', 40), ('one', 34), ('said', 33)]

Lancaster Stemmer

[('peopl', 72), ('tax', 70), ('get', 69), ('going', 52), ('presid', 48), ('that', 46), ('plan', 42), ('cut', 40), ('govern', 36), ('stat', 35)]

---\* Top words used by Obama and its frequencies \*---

Porter Stemmer

[('make', 55), ('governor', 49), ('go', 47), ('romney', 43), ('weve', 35), ('tax', 34), ('that', 34), ('got', 32), ('know', 29), ('cut', 29)]

Snowball Stemmer

[('make', 55), ('governor', 49), ('go', 47), ('romney', 43), ('weve', 35), ('tax', 34), ('that', 34), ('got', 32), ('know', 29), ('cut', 29)]

Lancaster Stemmer

[('govern', 58), ('mak', 55), ('romney', 43), ('going', 36), ('wev', 35), ('tax', 34), ('that', 34), ('got', 33), ('know', 29), ('cut', 29)]

----Problem 7d----

Positive words loaded from file...

Porter stemmed the positive words...

----Problem 7e----

Top positive words of Lehrer

[('right', 20), ('care', 8), ('well', 8), ('health', 7), ('close', 5), ('educ', 4), ('excus', 4), ('support', 4), ('afford', 3), ('candid', 3)]

Top positive words of Romney

[('state', 31), ('well', 27), ('right', 25), ('care', 22), ('like', 17), ('health', 16), ('kind', 14), ('time', 13), ('come', 12), ('sure', 11)]

Top positive words of Obama

[('well', 24), ('sure', 21), ('help', 20), ('care', 17), ('come', 15), ('health', 14), ('like', 14), ('educ', 12), ('right', 12), ('state', 12)]

----Problem 7f----

Lehrer Statistics

Number of positive words: 52

Romney Statistics

Number of positive words: 155

Obama Statistics

Number of positive words: 165