CST2555
Operating Systems & Computer Networks
Week 3

Exercises

Download the Shakespeare's Twelfth Night from:
`http://shakespeare.mit.edu/twelfth_night/full.html`

Convert this to text, using "html2text". If not installed then on Ubuntu type:
`sudo apt-get install html2text.`

Once installed then convert the html to text by typing:
`html2text TwelftNight.html > TwelfthNight.txt`

The grep searches are to be completed on TwelfthNight.txt.

1 grep basics

Answer the following questions:

- What lines does the word 'surfeit' occur on?

- What lines does the word 'heaven' occur on?

- How many lines does Curio have?

- How many lines does Sebastian have?

- How many empty lines of text are there?

- How many times does the word 'surfeit' occur?

- How many times does the word 'heaven' occur?

- How many times does the word 'soul' occur?

2   `grep` special characters and regular expressions

- How many lines?

- How many lines begin with the word 'but'?

- How many 9 letter words occur?

- How many 10 letter or more words occur?

- How many Acts are in the play?

- How many Scenes are in the play?

- How many 3 or more letter words that don't contain the letter 'e'.

- How many 3 or more letter words don't contain a vowel

- How many numbers ?

- How many non-alphabetic sequences?

- How many words?

- Identify words that have a double 'bb' in them?

- Identify words that have a double 'bb' in them and are preceded and succeeded by the same letter?

References

[1] Richard Blum and Christine Bresnahan. *Linux Command Line and Shell Scripting Bible*. Wiley, 3 edition, 2015.