Fundamentals of Data Management

Pass Tasks 1.1.5: Text Processing and Regular Expressions

Overview

As the first step, install Cygwin or VMWare Player or MinGW and open the Ubuntu/Linux environment. Then work through the text processing tasks and document your answers to the questions.

Purpose

Learn how to use command-line tools to manipulate text without the need for an editor. Use regular expressions to search for different variations of text.

Task

Use command line tools and regular expressions to find strings in files.

Reference: Regular Expression Cheat Sheet - https://www.rexegg.com/regexquickstart.html

Time

This task should be completed in your first lab class and submitted for feedback at the end of lab 1.

Resources

- Online modules (from Canvas)

- Regular Expressions tutorial: https://www.rexegg.com/regex-quickstart.html

Feedback
- Discuss your solutions with the tutorial instructor.

Next

Get started on Task 1.2.1. If your target is a Credit or higher then complete 1.2.1 and 1.2.2

Pass Tasks 1.1.5 — Submission Details and Assessment Criteria

Document your solutions to the tasks in a report using a word processor and upload to Doubfire as pdf. Your tutor will discuss the tasks with you have uploaded them.

Subtask 1.1.5

Consult the grep manual (command: man grep) and find out how you can add line numbers to your output.

Also, find out from the manual how to write a command that returns the number of matches rather than printing the matches.

Try out both commands using the command for subtask 1.1.3.