Fundamentals of Data Management

Pass Tasks 1.1.1: 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 1.1.2

Pass Tasks 1.1.1 — Submission Details and Assessment Criteria

Document your solutions to the tasks in a report using a word processor and upload to Doubfire as pdf.

Subtask 1.1.1

1.

a. For VMWare users:

Install the VMWare Player or open it if it exists. Unzip the virtual machine files on your local file system. Use the VMWare Player to navigate to the .vmx file to open the virtual machine provided. Watch the recorded instructions (vm-intro.mp4) to learn how to open a command line window and use commands on a Linux system. In the VM terminal, navigate to Documents/week1. You'll find the files access.log, IPaddresses.txt, People.txt and CanPostcodes.txt in this directory.

b. For Cygwin / MinGW users:

Download the files access.log, IPaddresses.txt, People.txt and CanPostcodes.txt from Doubtfire resources into your working directory.

Open a command line window (Terminal).

2. In the Terminal window navigate to your working directory using the cd command.
3. Run the 'grep' program as follows:

   grep Mozilla access.log
4. Take a screenshot and add it to your lab submission document.