

# Linux Commands

## Data Science Tools 1

Fall 2021



UNIVERSITY *of*  
DENVER

# Agenda

- Review Wk 2 Async
- Review Wk2 Homework
- Questions
- Additional Topics



- Additional Git commands
- Common Linux Commands
- Grep
- Awk

- git merge
  - What causes merge conflict?
- git diff
- git status
- git tag -a "v1.1" -m "Version 1.1 released"
- git reset - soft, mixed and hard
- git cherry-pick
- <http://hiddenlayer.net/understanding-git-as-version-control-system/>

- echo, cat, cut, head, tail, more, sort, uniq, curl, wget, tee
- ls, cd, sudo, rm (-r), mkdir, chmod, apt-get (yum), mv (rename a file or directory), source, cwd, dir, curl -o (wget)
- Three default file streams
  - Stdin - 0 (change to file using <)
  - Stdout - 1 (change to file using >)
  - StdError - 2 (change to file using 2>)
- Piping commands in Linux using |

- Basic Regex Engine (default)
  - ^
  - \$
  - .
  - \* (zero or more)
  - How to negate a pattern match? (-v, ^( pattern)
- Extended Regex Engine -> -E
  - ?, (optional)
  - +, (one or more)
  - \ (escape) Example - \(

- Transform each line of a file
- Format
  - `awk 'BEGIN {FS=OFS=","} NR>1 {print $2}' movies.csv`
- What does the above line do?
- \$1 - first column
- \$2 - second column
- \$3 - third column



- Stream Editor
- Used mainly for search and replace
- `s/colour/color/g`
- Handle on matched pattern is in `&`
  - `s/colour/&-british/g => colour-british`

Soft, mixed and hard reset

- soft - resets commit

- mixed - resets commit and index

- hard - resets commit, index and working tree

What option should you use if you want to print NUM lines of trailing context after matching lines?

- Look for answer in man grep

- *Python for Data Analysis, 2e*, by Wes McKinney, O'Reilly, 2018, 978-1-491-95766-0.
- *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow, 2e*, by Aurelien Geron, O'Reilly, 2019, 978-1-492-03264-9.
- *Pro Git, 2e*, by S. Chacon & B. Straub, Apress, 2014, 978-1-4842-0077-3.
- *Ubuntu Unleashed, 2019e*, by Matthew Helmke, Pearson Education, 978-0-13-498546-6.
- *Python Machine Learning*, by Sebastian Raschka & Vahid Mirjalili 2nd edition





- Three modes
  - insert mode (i), command mode (default or esc), and last-line mode (:)
- Common commands - command mode
  - Beginning and end of file (gg and G)
  - / - searching a pattern
- Basic editing - insert mode
  - dd - delete a line
  - yy - copy a line
  - p - paste
  - u will undo the last operation

- Last line mode
  - `:w` - save file
  - `:wq` - save and quit
  - `:set number` - show line numbers
  - `:set paste` - preserved paste formatting
  - `:10` - moves to line 10
- search and replace
  - `:%s/color/colour/g`

