

Conclusion and References

WE'LL COVER THE FOLLOWING ^

- References *
- Bash
- REGEX
- AWK
- SED
- GREP
- Big data
- Companion books

Thank you for joining us!

Through each lesson, I have covered a some commands, analyzed some data, but most importantly, asked a few interesting questions and used command line tools to answer them. I've also introduced with a concise beginner friendly guide to the big data landscape including an overview of the critical Big Data tools such as HDFS, MapReduce, YARN, Flume, Hive and more.

With all the topics we discussed, you are now be well-equipped to do some data analysis of your own!

References

Bash

- [Linuxconfig.org](http://linuxconfig.org) Bash Scripting Tutorial*
- Bash guide on Greg's wiki
- Steve Parker's shell scripting guide
- Advanced Bash Scripting Guide (ABS)

- IBM developerWorks “Bash by example”
- Bash Programming Introduction HowTo (TLDP)*
- LinuxCommand.org: Writing shell scripts.
- Beginner Linux Tutorial
- Beginner Bash Scripting Tutorial
- Wikipedia: BASH

(*some examples are used in this course)

REGEX

- Regular Expressions Cookbook by Jan Goyvaerts and Steven Levithan
- Teach Yourself Regular Expressions in 10 Minutes by Ben Forta
- Mastering Regular Expressions by Jeffrey Friedl
- Java Regular Expressions by Mehran Habibi
- Oracle Regular Expressions Pocket Reference by Jonathan Gennick & Peter Linsley
- Regular Expression Pocket Reference by Tony Stubblebine
- Regular Expression Recipes by Nathan Good
- Regular Expression Recipes for Windows Developers by Nathan Good
- Wikipedia: REGEX

AWK

- The GNU Awk User’s Guide at gnu.org
- awk(1) OS X Manual Page at developer.apple.com
- Unix awk(1) manual page at man.cat-v.org
- awk POSIX specification at opengroup.org
- Getting started with awk at cs.hmc.edu
- Wikipedia:AWK

SED

- GNU sed user’s manual as one page at gnu.org
- sed(1) OS X Manual Page at developer.apple.com
- Unix sed(1) manual page at man.cat-v.org
- Sed FAQ at sed.sourceforge.net

- Sed by example, Part 1 at ibm.com

- Wikipedia:sed

GREP

- grep Pocket Reference by John Bambenek
- GNU grep user's manual as one page at gnu.org
- grep(1) OS X Manual Page at developer.apple.com
- Unix grep(1) manual page at man.cat-v.org
- Wikipedia:grep

Big data

- Hadoop For Dummies Dirk Deroos
- Hadoop Operations by Eric Sammer
- Hadoop: The Definitive Guide, 4th Edition
- Agile Data Science: Building Data Analytics Applications with Hadoop by Russell Journey
- Learning Spark
- Programming Hive
- Professional Hadoop Solutions by Boris Lublinsky, Kevin T Smith, Alexey Yakubovich
- MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop Donald Miner
- Learning Spark: Lightning -Fast Big Data Analysis by by Holden Karau, Andy Konwinski, Patrick Wendell and Matei Zaharia
- Advanced Analytics with Spark: Patterns for Learning from Data at Scale by Sandy Ryza, Uri Laserson, Sean Owen and Josh Wills

Companion books

- Hello Big Data @ Bash: Explore real-world data at the command line, [Leanpub.com](http://leanpub.com)
- Data Science at the Command Line by Jeroen Janssens
- Adventures in Data Science, by Robert Aboukhalil

