

# DS-DC-13 LECTURE NOTES

*Jim Simpson*  
*Data Scientist, Sotera*

LESSON 05

# LESSON 5

## LEARNING OBJECTIVES

- ☐ IDENTIFY THE DATA SCIENCE TOOLKIT
- ☐ NAVIGATE GIT & THE COMMAND LINE
- ☐ FAMILIARIZE WITH DATA VISUALIZATION

## STACK

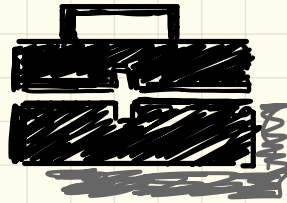
PYTHON

ANACONDA

NUMPY

PANDAS

Auto  
PEP8  
↔  
Style  
Guide



## REMOTE SERVER

AWS EC2

SPOT INSTANCE

CLONING AMI

CONNECT OVR SSH

Git  
↔  
Medium  
sized  
data

## EDITORS

JUPYTER NOTEBOOK

SUBLIME TEXT

ATOM

## COMMAND LINE

Data Science at  
the Command Line

Data Science Toolbox

Unix: head, grep, sort, ...  
csvkit: csvcut, csvstat, ...

# LESSON 5

## LEARNING OBJECTIVES

- ☒ IDENTIFY THE DATA SCIENCE TOOLKIT
- ☒ NAVIGATE GIT & THE COMMAND LINE
- ☒ FAMILIARIZE WITH DATA VISUALIZATION