GENERAL ASSEMBLY

INTRO TO DATA SCIENCE
LESSON 3:
PYTHON FOR DATA ANALYSIS

AGENDA

- ▸ **INTRODUCTION**
- ▸ **IPYTHON ENVIRONMENT**
- ▸ **DATA BASICS**
- ▸ **NUMPY**
- ▸ **PANDAS**
- ▸ **MATPLOTLIB**
- ▸ **LAB**
- ▸ **DISCUSSION**

INSPIRATION

But first, you must complete your training

GETTING COMFY

Navigating Your State of the Art Python Environment

GETTING COMFY

‣ IPython Notebooks - You can live here

‣ Comments

‣ Help

‣ Tricks

PYTHON FOR DATA ANALYSIS

PYTHON DATA BASICS

SLITHERING TOWARDS AWESOME

‣ Control Flow

‣ Functions

‣ Data Structures

  ‣ Lists, Sets, Dictionaries, Tuples,

‣ Modules

‣ Classes

PYTHON FOR DATA ANALYSIS

NUMPY

NUMPY OVERVIEW

‣ ndarray

‣ Indexing and Slicing

‣ Boolean Indexing

PYTHON FOR DATA ANALYSIS

PANDAS
For DataFrames

OBJECT CREATION

VIEWING DATA

SELECTION

MISSING DATA

GROUPING

RESHAPING

TIME SERIES

PLOTTING

I/O

MATPLOTLIB

Because "Math" takes too long to say.

MATPLOTLIB OVERVIEW

‣ History

‣ Plot()

‣ Figures and Text

‣ Bar, Histograms, ScatterPlots

LAB. NEW MUNGE TIMES

DISCUSSION