

Bluegrass Data Science Group

Python Immersion Course – 2018

Joe Blankenship
Tosha Fraley

Overview

- 6 month co-learning course
 - 5 working + 1 capstone
- Modular
 - Whole course vs. module
- Learning Objectives
 - Python Language
 - Web apps & APIs
 - ETL processes (data munging) & data storage
 - Visualization & advanced methods (ML/DL)

Overview

Month 1	Month 2	Month 3	Month 4	Month 5	Month 6
February	March	April	May	June	July
Intro Overview Setup	Python Basic Project	Web Application Project	Data Analysis Project	Machine Learning Project	Capstone Final Projects

Overview

- Introductions and Why Python?
 - Newcomers to data science
 - Overview & course expectations
 - Operating Systems, Virtual Environments
 - IDEs
 - Additional Resources
 - Q&A
 - Some Setup (if time permits)

Overview

- Programming in Python
 - For newcomers, refresher for others
 - Foundation for the rest of the course
 - Use Python 3!
 - Computer Science Perspective
 - Zelle, Lutz, Slatkin
 - **Choose a Project**
 - **Find an area of interest**
 - **Find related data**
 - **Use Python!**

Overview

- Web Development in Flask
 - Understanding a web framework
 - Build an API
 - Hosting
 - Publish your data visualizations
 - HTML/CSS/JS
 - Testing and TDD
 - **Choose a Project**
 - **Build a portfolio website**

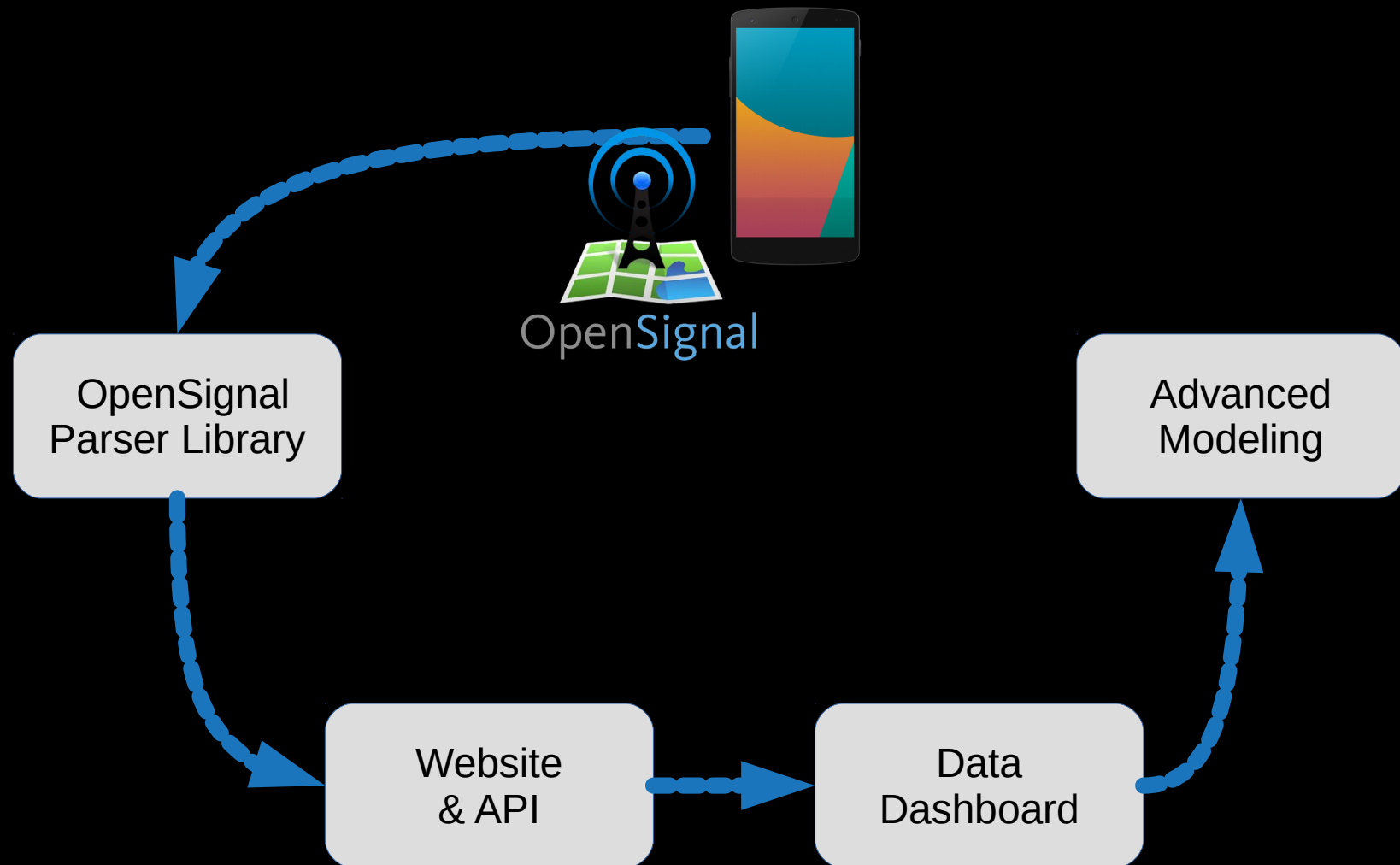
Overview

- Data Analysis with Python
 - Libraries and techniques
 - Statistics
 - ETL and “data munging”
 - Transformation/Normalization
 - Analytic methods
 - Visualizations
 - **Choose a project**
 - **Build a data dashboard for your portfolio**
 - **Explain rationale/methods/results**
 - **Source your work!**

Overview

- Machine Learning with Python
 - Concepts and applications
 - Choosing the right method
 - Deep learning briefly covered
 - **Choose a Project**
 - **Write a blog post on ML/DL experiment**
 - **Include rationale/method/results/limitations**
 - **Include visualizations when possible**
- Capstone Exercise
 - Share experiences/projects/thoughts on the course
 - **Hackathon & presentations**

Project - Example



Overview

- Additional Expectations
 - Co-learners are encouraged to find alternative learning materials
 - Frequent participation and discussion on Slack
 - Bluegrass Developers Guild Slack
 - <https://www.bluegrassdevs.org>
 - #python-immersion
 - 2 hour meeting once a month
 - Discuss lessons from previous month
 - Share experiences
 - Cover additional materials
- **Team up!**

Questions/Comments?

Operating Systems

- Go here for your respective OS download
 - <https://www.python.org/>
- Windows
 - Download executable installer or ZIP
 - Add to Path!
- Mac OS
 - Python for Mac OS
 - For latest versions, have to download
- Linux
 - Just the best ;)
 - Already have Python and/or Legacy Python

Package Management

- PyPI
 - Python Package Index
 - <https://pypi.python.org/pypi>
 - pip package utility
 - May need to install depending on OS
 - Anyone can publish their packages here
- Many other package managers
 - Do your research
 - Aptitude, Homebrew, Anaconda, Enthought

Virtual Environments

- Why?
 - Maintain project environment
 - Prevent clutter
 - System packages – Python Standard library
 - Site packages - 3rd party libraries)
 - Prevent package version issues
 - Makes sharing your research/projects easier
- Command
 - *python3 -m venv [name of virtual environment]*
 - *source [name of virtual environment]/bin/activate*

Virtual Environments

- Notes
 - Anaconda is a popular data science environment
 - Package management/virtual environments are unique to their system
 - Enthought is another data science-focused environment
 - You can go one of these routes if you choose
 - Keep in mind these are companies with partially proprietary setups

IDEs

- Integrated Development Environment
 - Jupyter Lab
 - Jupyter Notebooks
 - iPython
 - Spyder
 - PyCharm
 - Atom
 - Sublime
 - Visual Studio Code
 - Rodeo
 - Any text editor really...

Resources

- Pycoder's Weekly
 - <http://pycoders.com/>
- KDNuggets
 - <https://www.kdnuggets.com/news/index.html>
- Data Science Weekly
 - <https://www.datascienceweekly.org/newsletters/data-science-weekly-newsletter-issue-217>
- DataCamp
 - <https://www.datacamp.com/>
- Codecademy
 - <https://www.codecademy.com/>
- Public Library
 - Books, E-books, and online education resources

Questions/Comments?

Setup and Discussion

Additional Resources

- Python Packages
 - <https://packaging.python.org/tutorials/installing-packages/>
- ptpython
 - <https://github.com/jonathanslenders/ptpython>
- PyCon talks
 - <http://pyvideo.org/>
- Full Stack Python
 - <https://www.fullstackpython.com/>
- Real Python
 - <https://realpython.com/>
- Dan Bader's stuff
 - <https://dbader.org/>