Getting Started with Data Science

Welcome to Data Science at General Assembly!

We are all really excited that you are joining the General Assembly community. In order to ensure your success in the course, we have created an on-boarding checklist that will take most students between 15 to 20 hours to complete.

- [ ] Review the **Technical Guide**
- [ ] Work through the **On-boarding Modules**
- [ ] Complete the **On-boarding Assessment**

Please complete these pre-work items before your first day so that both you and your computer will be prepared for class. In the meantime, feel free to reach out with any questions (or just to say hi)!

Pre-Course Checklist

Please complete the following by the first day of class. Feel free to shoot us an email with any questions:

1. **Accounts**
2. **Install**
3. **Practice**
4. **Resources**

Accounts

1. [ ] Create a **GitHub account**, if you don not already have one.

Installation

Please follow the installation instructions below. If your installation is not complete or if you have any problems, plan to arrive at least 30 minutes early to the first class so that our TAs' can help you get set up!

1. [ ] Install **Python 2.7**

- **Note**: Make sure to install version 2.7, **NOT Python 3**, which has significant differences from 2.7 and is not "industry standard." *We will be using Python 2.7.*

2. [ ] Install **Anaconda**

- Test for successful installation. For example, on a Mac you can by open a Terminal window (using the Anaconda Launcher app, if needed) and type `jupyter notebook`. In a few moments, your browser should open to a window titled "Jupyter".

3. [ ] Install **Pip**, the recommended package manager for Python.

- **Note**: Most versions of Python come with Pip pre-installed. Check by opening your Terminal and running `pip install -U pip` to see if you have the latest version.

Start Practicing

Review the following resources as an introduction to some commonly used concepts and tools.

1. [ ] Complete Codecademy's free **"Learn Python"** course to practice your Python syntax. Pay close attention to lists, dictionaries, and functions.

- Bonus Option: Check out **Learn Python the Hard Way** and work through exercises 1-10.

2. [ ] Review our **Command Line** tutorial to familiarise yourself with basic Terminal commands. Pay particular attention to folder navigation and file creation.

- Bonus Option: Run through these additional exercises from **Learn Command Line the Hard Way**.

3. [ ] Read through this **10 minute guide to Pandas** to learn about a popular library used for data analysis that we will be using in this course.

- Bonus Option: Practice your Python by working through **Project Euler's computational problems**.

Additional Resources (Optional)

Consider the following reading recommendations. These are not required but may help supplement your learning experience:

- **The Quick Python Book**
- **Data Analysis With Open Source Tools**
- **Python for Data Analysis**

Links

- **On-boarding Modules**, https://docs.google.com/document/d/1xoC18JHk880C9n_cAPCsGnsNqexPgTIIkmUd7cTJQw/edit?usp=sharing

- **On-boarding Assessment**, https://mobilega.typeform.com/to/DqgDfO
- **Create a Github account**, https://github.com/join
- **Python 2.7**, https://www.python.org/downloads
- **Anaconda**, https://www.continuum.io/downloads
- **Learn Python**, https://www.codecademy.com/learn/python
- **Learn Python the Hard Way**, http://learnpythonthehardway.org/book
- **Command Line**, http://generalassembly.github.io/prework/cl/#/
- **Learn Command Line the Hard Way**, http://cli.learncodethehardway.org/book
- **10 minute guide to Pandas**, http://pandas.pydata.org/pandas-docs/stable/10min.html
- **Project Euler's computational problems**, https://projecteuler.net
- **The Quick Python Book**, http://www.amazon.com/Quick-Python-Book-Second-Edition/dp/193518220X
- **Data Analysis With Open Source Tools**, http://www.amazon.com/Data-Analysis-Open-Source-Tools/dp/0596802358
- **Python for Data Analysis**, http://www.amazon.com/Python-Data-Analysis-Wrangling-IPython/dp/1449319793