COURSE PREP: DATA SCIENCE

GETTING STARTED WITH DATA SCIENCE

Congratulations on joining us at General Assembly for Data Science! We are excited to have you join our community and are looking forward to working with you soon!

In order to best prepare and ensure success in the course, we've gathered a list of resources that we'd like you to complete. We want to ensure that when you arrive on your first day, you will be up to speed and familiar with the necessary skills needed to succeed in this course.

PRE-COURSE CHECKLIST

- ❏ Complete Enrollment Agreement: This must be completed before the first day of class.
- ❏ Join Slack (you'll receive an invite to your inbox)
- ❏ Review our Terms and Conditions
- ❏ Install Anaconda by Continuum. We will be using Python 2.7 for this program.
- ❏ Data Science Basics Tutorials

SOFTWARE INSTALLATIONS

- Install Anaconda by Continuum. We will be using Python 2.7 for this program. Questions pertaining to Anaconda installation will be handled by your instructional team in class

PRE-COURSE TUTORIALS

- **Command Line Interface (CLI):** If you're a complete beginner, start here. If you want to flesh out your skills, go here.
- **Python:** If you're new to Python, start with this free 13-hour short course on Codecademy. If you want to flesh out your skills, go through the free online book *Learn Python The Hard Way*.
- **Python + Data:** You will use the *pandas* library a lot during this course, therefore strongly recommended you learn the basics before starting!
    - To understand what *pandas* is used for, watch this 10-minute video.
    - If you're entirely new to Python and tools like the IPython Notebook, I'd recommend *Python for Data Analysis* by Wes McKinney (he wrote the *pandas* library).
    - If you're reasonably comfortable with Python, skip the book and jump straight into the tutorial lessons here. This page also contains additional resources for advanced use cases.
- **Stats:** Learn what statistical significance, confidence intervals, chi-square tests, Bayes' Theorem and much more are in Think Stats, a free book that uses real-world examples written in Python.
- **Podcasts:** Learn about k-means while you clean, or about random walks on a random walk! *Data Skeptic* and *Linear Digressions* are two of our favourites.