**Python Programming Course**

"I wish I did more Pythong programming and Statistical Learniing courses."

— Every Data Science regular program students

\_\_Goal\_\_: Improve your programming skills to intermediate level with Python and get you ready to jump into our Data Science regular program curriculum on day 1 

* Regular python programming course for beginner and intermediate
* Focusing on data processing and math problem solving
* Capstone project

Please read this entire document before getting started on the individual sections.

Before classes begin, you are expected read the contents in each chapter, complete the code assignments for every chapter, and turn them in.

Please complete and hand in each chapter sequentially.

**Notes**

In the notes folder, you'll find these things:

* [Syllabus](https://github.com/alexseong/dsy_python_programming/blob/master/syllabus.md): Full syllabus including the daily schedule, instructor contact info, and course requirements
* [setup.sh](https://github.com/alexseong/dsy_python_programming/blob/master/notes/setup.sh): Script for setting up your environment
* [Workflow](https://github.com/alexseong/dsy_python_programming/blob/master/notes/workflow.md): Notes on programming workflow
* [Git issues](https://github.com/alexseong/dsy_python_programming/blob/master/notes/git_issues.md): Commonly encountered git issues

**How to submit the assignments.**

You will submit your solutions with a git [pull request](https://help.github.com/articles/using-pull-requests). Here are step-by-step instructions of how to do this:

1. Fork this repository (go to <https://github.com/alexseong/dsy_python_programming> and click "Fork")
2. [Install git](https://help.github.com/articles/set-up-git) on your computer if you haven't already
3. Clone your forked repository onto your computer: git clone https://github.com/<your username>/dsy\_python\_programming
4. Edit the assignment files with your solutions
5. Add your changes to the repository: git add <file you edited>
6. Commit your changes: git commit -m "My solutions to python programming"
7. Push your changes to your fork: git push origin master
8. Make a pull request by going to https://github.com/<your username>/dsy\_python\_programming and clicking "Pull Requests" and then "New Pull Request". Then click "Create Pull Request". Put your name in the pull request title and then click "Create Pull Request" again.

It's good idea to commit and push to your personal fork often, even when you're not done. Then you have older versions of your work in case you screw something up or lose something. We will get a notification of your submission once you submit a pull request.

**Common git issues**

1. If we've made changes to the repository after you forked it and you want to update your repository to reflect them, you can run this command: git pull https://github.com/alexseong/dsy\_python\_programming master
2. If you try to push and get "Repository does not exist" this probably means that you cloned from the Institute repo rather than your fork. Make sure you did step 1 and created a fork. Then run this command to push: git push https://github.com/<your username>/dsy\_python\_programming master

**Getting help**

Feel free to contact us directly if you have any questions!

Consult the [list of additional references and resources](https://github.com/alexseong/dsy_python_programming/blob/master/Extra_Credit/references.md) if you'd like to further cover anything. The Extra\_Credit directory also contains some additional references on operating systems and text editors as well as HTML/CSS/JavaScript.

**Course Outline**

This is the outline for the course. Each row represents a day, readings are to be completed before class on the given day.

* **Week:** Week of the Month
* **Readings:** Assigned readings for the day. To be completed before coming into class.
* **Repo:** The repo contains the day's exercise(s). You should be able to complete this in the time allotted (you will not have access until the day of).
* **Instructor:** The instructor who is the point person for the day.
* **Slides:** The day's lecture notes and slides

This document covers much of what we'll teach you, but you are each other's greatest resource and will ultimately learn the most from each other.

**Schedule**

| **Month** | **Week** | **Topic** |
| --- | --- | --- |
| 1 |  | [Python Programming Intro](https://github.com/alexseong/dsy_python_programming#month-1-python-programming-intro) |
| 2 |  | [More Python](https://github.com/alexseong/dsy_python_programming#month-2-more-python) |
| 3 |  | [Math Programming with Numpy](https://github.com/alexseong/dsy_python_programming#month-3-math-programming-with-numpy) |
| 4 |  | [Data Analysis with Pandas](https://github.com/alexseong/dsy_python_programming#month-4-data-analysis-with-pandas) |
| 5 |  | [Visualizaton](https://github.com/alexseong/dsy_python_programming#month-5-visualization) |
| 6 |  | [Capstone Project](https://github.com/alexseong/dsy_python_programming#week-6-capstone-project) |

--

**Month 1: Python Programming Intro**

| **Week** | **Readings** | **Repo** | **Instructor** | **Slides** |
| --- | --- | --- | --- | --- |
| Week 1 | [Development Workflow](https://github.com/alexseong/dsy_python_programming/blob/master/notes/workflow.md) | [Assessment 1](https://github.com/alexseong/dsy_python_assessment-1)  [Python Intro 1](https://github.com/alexseong/dsy_python_intro_1) | Alex Seong | [git-intro](https://github.com/alexseong/dsy_lectures/tree/master/intro-git)  [slides](https://github.com/alexseong/dsy_lectures/tree/master/python-intro/python_intro_1) |
| Week 2 | [Python Syntax 1](https://github.com/alexseong/dsy_python_programming/blob/master/notes/python_syntax) | [Python Intro 2](https://github.com/alexseong/dsy_python_intro_2) | Alex Seong | [slides](https://github.com/alexseong/dsy_lectures/tree/master/python-intro/python_intro_2) |
| Week 3 | * [Learn Python the Hard Way (ex 40-42)](http://learnpythonthehardway.org/book/ex40.html) * Extra: Learn Python the Hard Way ex 43 | [OOP](https://github.com/alexseong/dsy_oop) [Python Intro 3](https://github.com/alexseong/dsy_python_intro_3) | Alex Seong | [slides](https://github.com/alexseong/dsy_lectures/tree/master/OOP) |
| Week 4 | * [Intro to IPython Notebook](http://nbviewer.ipython.org/github/jvns/pandas-cookbook/blob/master/cookbook/A%20quick%20tour%20of%20IPython%20Notebook.ipynb) * [Python Syntax 2](https://github.com/alexseong/dsy_python_programming/blob/master/notes/python_syntax) | [Python Intro 4](https://github.com/alexseong/dsy_python_intro_4) | Alex Seong | [slides](https://github.com/alexseong/dsy_lectures/tree/master/python-intro/python_intro_4) |
| -- |  |  |  |  |

**Month 2: More Python**

| **Week** | **Readings** | **Repo** | **Instructor** | **Slides** |
| --- | --- | --- | --- | --- |
| Week 1 | Prealgebra with Python | Assessment 2 [Prealgebra](https://github.com/alexseong/dsy_python_prealgebra) | Alex Seong | slides |
| Week 2 | * Prealgebra with Python * Algebra1 with Python | [Prealgebra](https://github.com/alexseong/dsy_python_prealgebra) [Algebra1](https://github.com/alexseong/dsy_python_algebra1) | Alex Seong | slides  slides |
| Week 3 | * Prealgebra with Python * Algebra1 with Python * Algebra2 with Python | [Prealgebra](https://github.com/alexseong/dsy_python_prealgebra) [Algebra1](https://github.com/alexseong/dsy_python_algebra1) [Algebra2](https://github.com/alexseong/dsy_python_algebra2) | Alex Seong | slides slides slides |
| Week 4 | * Prealgebra with Python * Algebra1 with Python * Algebra2 with Python * Geometry with Python | [Prealgebra](https://github.com/alexseong/dsy_python_prealgebra) [Algebra1](https://github.com/alexseong/dsy_python_algebra1) [Algebra2](https://github.com/alexseong/dsy_python_algebra2) [Geometry](https://github.com/alexseong/dsy_python_geometry) | Alex Seong | slides slides slides |

--

**Month 3: Math and Scientific Programming with NumPy, SciPy, SymPy**

| **Week** | **Readings** | **Repo** | **Instructor** | **Slides** |
| --- | --- | --- | --- | --- |
| Week 1 | NumPy | Assessment 3 Introduction to NumPy, SciPy, SymPy | Alex Seong | slides |
| Week 2 | Agebra with Numpy | Algebra with NumPy | Alex Seong | slides |
| Week 3 | Probability 1 | Probability 1 | Alex Seong | slides |
| Week 4 | Probability 2 | Probability 2 | Alex Seong | slides |

--

**Month 4: Data Analysis with Pandas**

| **Week** | **Readings** | **Repo** | **Instructor** | **Slides** |
| --- | --- | --- | --- | --- |
| Week 1 | * 10 Minutes to Pandas * Pandas Top 10 | Assessment 4 Introduction Pandas | Alex Seong | slides |
| Week 2 | EDA with Pandas | EDA | Alex Seong | slides |
| Week 3 | Data Wrangling with pandas | Data Wrangling | Alex Seong | slides |
| Week 4 | Plotting | pandas-matplitlib | Alex Seong | slides |

--

**Month 5: Visualization and Basic Statistics**

| **Week** | **Readings** | **Repo** | **Instructor** | **Slides** |
| --- | --- | --- | --- | --- |
| Week 1 | More plotting | Seaborn and Plotly | Alex Seong | slides |
| Week 2 | Basic Linear Algebra and Numpy | Linear Algebra and EDA | Alex Seong | slides |
| Week 3 | Sampling and Estimation | Sampling and Estimation | Alex Seong | slides |
| Week 4 | Hypothesis Testing | A/B Testing(or Linear Regression Intro) | Alex Seong | slides |

--

**Month 6: Capstone Project**

| **Week** | **Readings** | **Repo** | **Instructor** | **Slides** |
| --- | --- | --- | --- | --- |
| Week 1 | Project Proposal | Web Scraping | Alex Seong | slides |
| Week 2 | Project Proposal | Web Products(Flask) | Alex Seong | slides |
| Week 3 | Project Review | Project Review | Alex Seong | slides |
| Week 4 | Project Presentation | Web or PPT | Alex Seong | slides |

--

--

**Textbooks**

We will focus on a few canonical texts for the class and readings will be assigned from them. If they are not in a physical form in our library, they are located in digital form on the Time Capsule or the Internet.

* [Think Python](http://www.greenteapress.com/thinkpython/html/index.html)[(pdf)](http://www.greenteapress.com/thinkpython/thinkpython.pdf)
* [Dive into Python 3](http://histo.ucsf.edu/BMS270/diveintopython3-r802.pdf)

**References**

**Data Analysis with Python**

* [Python for Data Analysis](http://shop.oreilly.com/product/0636920023784.do): Some of the pandas methods have changed (always reference pandas [online documentation](http://pandas.pydata.org/)) but a solid book on data analysis in Python.

**Numpy**

* [Official Numpy Tutorial](http://wiki.scipy.org/Tentative_NumPy_Tutorial)
* [scipy Lectures](https://scipy-lectures.github.io/intro/numpy/index.html)
* [Crash Course in Python for Scientist](http://nbviewer.ipython.org/gist/rpmuller/5920182)
* [Scientific Python Lectures](http://nbviewer.ipython.org/github/jrjohansson/scientific-python-lectures/blob/master/Lecture-2-Numpy.ipynb)
* [Numpy Broadcasting](http://wiki.scipy.org/EricsBroadcastingDoc)
* [Pyton Bootcamp Lectures](http://nbviewer.ipython.org/github/profjsb/python-bootcamp/blob/master/Lectures/05_NumpyMatplotlib/IntroNumPy.ipynb)

**Scipy**

* [scipy Lectures](https://scipy-lectures.github.io/)

**scikit-learn**

* [Introduction to Machine Learning with sklearn](http://researchcomputing.github.io/meetup_spring_2014/python/sklearn.html)
* [scikit-learn workshop](https://github.com/jakevdp/sklearn_pycon2014)
* [Machine Learning Tutorial](https://github.com/amueller/tutorial_ml_gkbionics)
* [Introduction to scikit-learn](http://nbviewer.ipython.org/github/tdhopper/Research-Triangle-Analysts--Intro-to-scikit-learn/blob/master/Intro%20to%20Scikit-Learn.ipynb)
* [Data analysis with scikit-learn](http://sebastianraschka.com/Articles/2014_scikit_dataprocessing.html)
* [Advanced Machine Learning with scikit-learn](https://us.pycon.org/2013/community/tutorials/23/)

**Extra**

* [University of Colorado Computational Science workshops](http://researchcomputing.github.io/meetup_spring_2014/)