edX

GTx: CSE6040x

FA18: Computing for Data Analysis

# Module 0: Fundamentals (bootcamps)

## Topic 1: Python Essentials Notebook 1

[MSA Python Bootcamp – August 2018](http://datamastery.gitlab.io/msabc/august2018.html)

ITCS means [Interactive Think CS Textbook](http://interactivepython.org/runestone/static/thinkcspy/index.html)

IP means [Introducting Python](http://shop.oreilly.com/product/0636920028659.do)

TP means Think Python, 2nd Edition, by Allen B. Downey, O'Reilly Media, December 2015. Available free at <http://greenteapress.com/wp/think-python-2e/> and from O'Reilly at <http://shop.oreilly.com/product/0636920045267.do>

* Course Intro and Tool Set-up
  + ITCS – Introduction
    - Algorithm: solution created through the problem solving process. Step by step list of instructions that if followed exactly will solve the problem under considerations.
    - Goal in computer science is to take a problem and develop an algorithm that can serve as a general solution. Once we have such a solution, we can use our computer to automate the execution.
    - Programming is a skill that allows a computer scientist to take an algorithm and represent it in a notation (a program) that can be followed by a computer.
    - Python is a high-level language.
    - Machine language (low-level language) is the encoding of instructions in binary so that they can be directly executed by the computer.
    - High level languages:
      * Much easier to program
      * Take less time to write
      * Portable – can run on different kinds of computers with few or no modifications
    - Two kinds of programs process high-level languages into low-level languages:
      * Interpreters
      * compilers
  + TP 1
  + Video: Getting Started (24.03)
    - Cs2316.gatech.edu > resources
    - “Computers are like onions, they are organized in layers and when you dig in they make you cry”
    - Shell: application that gives you direct access to the operating systems services
    - Folder = directory
    - CMD – primary do two things: navigate to files and execute programs
  + Video: Intro to Python (29:43)
    - Python is general purpose programming language, a glue language
    - Different than e.g. SQL whose primary purpose is manipulating data
    - [www.python.org](http://www.python.org) -> <https://docs.python.org>
    - <https://github.com/csimpkins/course-tools>
      * Python script to upload course schedule to a website every semester
      * Finding best time to dive based on input of time range and dates from a list of ~1400
    - Python is an interpretive programming language 🡪
      * You create a file e.g. hello.py
      * Feed that file into a python interpreter
      * Get a running program
    - Contract to a compiled language e.g. C
      * Create a file e.g. hello.c
      * Feed that file into a compiler (e.g. gcc)
      * Get another file: hello.exe
      * Get a running program
    - Java is a hybrid of these two approaches
      * Create a file e.g. hello.java
      * Feed that file into javac
      * Get another file hello.class
      * Feed that into java program
      * Get a running program
    - 19:30 – directories? Sub directories
    - Python interpreter has a REPL -> read, evaluate, print, loop
    - Print() is a function that doesn’t return a value – if you don’t specify a value, it returns a special value – none.
    - Get out of the REPL: exit() or Ctrl +Z
    - Bash – born again shell -> install on Windows?

To-do:

Upgrade PIP: CMD > python -m pip install --upgrade pip

Install python packages: CMD>

python -m pip install requests 🡪 DONE

python -m pip install beautifulsoup4 🡪 DONE

python -m pip install sqlite 🡪 DID NOT INSTALL

python -m pip install pyqt 🡪 DID NOT INSTALL

python -m pip install ipython 🡪 DONE

python -m pip install numpy 🡪 DONE

python -m pip install pandas 🡪 DONE

python -m pip install matplotlib 🡪 DONE

python -m pip install jupyter 🡪 DONE

python –m pip install statsmodels scipy scikit-learn 🡪 DONE

Questions:

* Multiple tabs in command prompt?
* Default command prompt to a file location?

## Assignment: Notebook 1 (Due Sept 3, 2018 at 11:59 UTC)

<>

# Module 1

## Topic 1

<>

## Assignment:

<>