CS 38003 PYTHON PROGRAMMING

Ruby Tahboub

COURSE OVERVIEW

GOALS

Provide an introduction to the Python programming language with focus on:

- *Efficient* data processing (i.e., parse, extract and process) using Python structures and libraries.
- Python libraries e.g., matplotlib, pandas, etc.

The materials are designed to provide Python programming skills needed for Data Science applications.

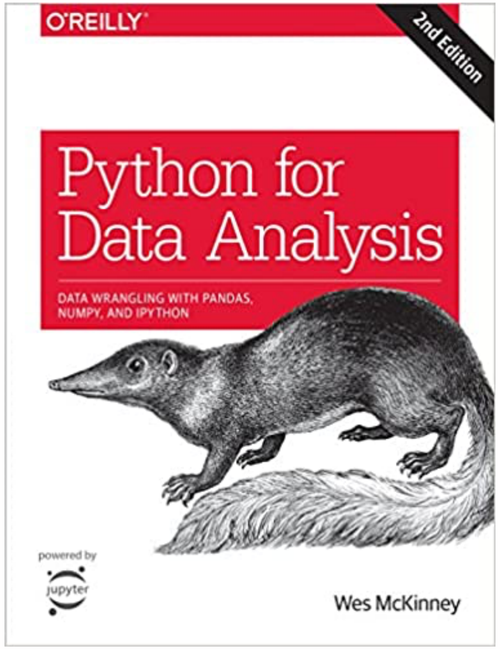
TOPICS

- Language basics: data types, decisions, loops, etc.
- Functions, libraries, and exception handling.
- Data collections: lists, tuples, sets, and dictionaries.
- Strings, regular expressions, and file processing.
- Classes and objects.
- Generating random numbers.
- Modules for Data Science: Numpy, Pandas, Matplotlib, Beautifulsoup,etc.

TEXTBOOK

- McKinney, W. (2017). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython (Second ed.). O'Reilly Media.*

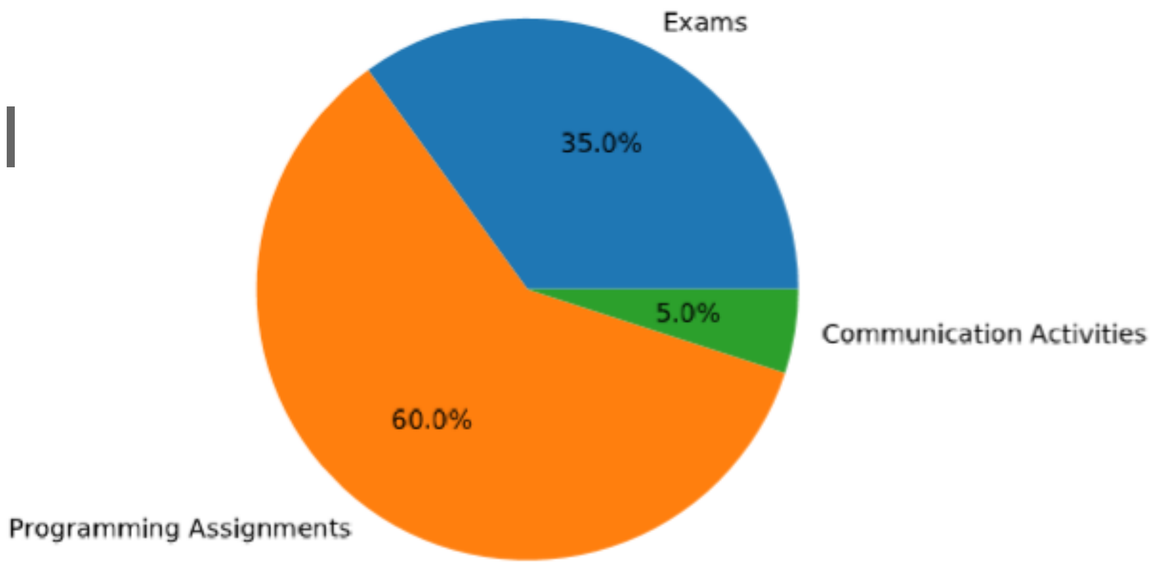
You can access this via the Purdue Library for free:



https://purdue-primo-prod.hosted.exlibrisgroup.com/permalink/f/vjfldl/PURDUE_ALMA51793440910001081

WORKLOAD and GRADING

- Five labs.
- Coding exams: midterm and final exam.
- Communication activities.



PYTHON ENVIRONMENTS

- Python 3.
- We will use Jupyter Notebook for the first part.
- Feel free to use your favorite IDE (PyCharm is a good option).
- You may use Anaconda to install Python.

CONTENT CREDITS

Parts of the materials in this class are adapted from the following sources:

- Purdue CS177, CS242, CS38003, and CS**50**023
- UPenn CIS530 python tutorial: ftp://ftp.cis.upenn.edu/pub/cis530/public_html/materials/python-tutorial.pdf

THANK YOU!