

Course Syllabus

[Jump to Today](#)

CP255 Urban Informatics and Visualization

Instructor: Paul Waddell (waddell at berkeley.edu)

GSI: Geoff Boeing (gboeing at berkeley.edu)

Office Hours: T Th, 10-11 in the DCRP office

Scheduling: Fall, 2014. Tuesdays and Thursdays, 11:00-12:30

Office Hours: T Th, 10-11 in 214 lab

Location: Wurster 214 Computer Lab

COURSE SYNOPSIS AND LEARNING OBJECTIVES

This is a hands-on course that trains students to analyze urban data, develop indicators, and create visualizations and maps using the Python programming language, open source tools, and public data. The course will first introduce the fundamentals of programming in Python before moving on to a survey of data analysis/visualization tools and technologies. Classroom sessions will include lectures and workshops. A series of assignments will reinforce the skills and topics being presented.

This course is designed to provide future city planners with a toolkit of technical skills for quantitative problem solving. It requires some tolerance for experimentation, self-directed trial and error, and an interest in learning to write code. If you are willing to roll up your sleeves and embrace some uncertainty, you will learn the fundamentals of urban spatial analysis and visualization, and might discover an entirely new lens through which to study, plan, and design neighborhoods, cities, and regions.

Topics to be covered include:

- Fundamentals of programming with Python and IPython notebooks
- Cleaning, manipulating, and analyzing urban data with Python's pandas library
- Visualizing data in Python with charts, graphs, and tables
- Accessing public data from the web and APIs (including Twitter, Google, and census data)
- Analyzing location accessibility and building simple regression models
- Developing spatial indicators and mapping urban data with open source GIS tools, Mapbox, and TileMill

COURSE PREREQUISITES

Some prior coursework (such as CP 204c) or basic experience using a GIS is required. Students are not required to have prior programming experience, although it will be beneficial. Python is an accessible language and the course will emphasize learning by doing. Prior or concurrent course work in statistics and data analysis (e.g. CP 204A or 204D) is encouraged as this course will not provide the theoretical foundations of statistical analysis.

This course is open to students from across campus, but priority enrollment will be given to students in the Master of City Planning program.

COURSE ASSIGNMENTS AND ASSESSMENT

Students will develop skills gradually through in-class exercises and tutorials, and graded assignments. There will be numerous assignments, scheduled roughly every other week -- but no exams or final projects. Assignments will typically involve writing some code and documenting it, using IPython Notebooks that can be shared and interactively run inside a web browser, and a writeup discussing the assignment and its results.

Exercises and Assignments will account for 90% of the course grade, and class participation for 10%. Exercises will be smaller tasks that demonstrate a degree of mastery of skills on a smaller scale, and will be used mainly as a means of ensuring that students are keeping up with the material and not falling behind. Assignments will be somewhat larger and more involved tasks, that require a higher degree of facility and independent work, putting pieces together to complete larger tasks. Assignments will be given generally on a Thursday, and be due by end of the day on the following Monday. Thursdays will generally involve more of a workshop format, with students working individually on exercises or assignments in class to have opportunities to interact with the instructor and GSI and classmates while gaining skills and confidence.

If a student wishes to undertake a larger project using the skills developed in this course, they may explore doing an independent study with the instructor.

READINGS AND RESOURCES











This course does not use a book or assigned text, since so much of the material is based on emerging and rapidly evolving technology. For the most part, the lecture notes, tutorials and examples provided on this website will provide a good foundation for the skills students will gain in the course. Some of the key packages we will be using have free documentation available, and you should make extensive use of these to help you climb the learning curve.





















1. Downey, Allen. 2013. Think Python: How to Think Like a Computer Scientist. Green Tea Press. Available as a [free PDF download](http://www.greenteapress.com/thinkpython/thinkpython.pdf). <http://www.greenteapress.com/thinkpython/thinkpython.pdf> This text provides a fairly comprehensive overview of Python as a programming language.
2. [IPython Documentation](http://ipython.org/ipython-doc/stable/index.html). 2013. This online documentation explains what IPython is and how to use it.
3. McKinney, Wes. 2012. pandas: powerful Python data analysis toolkit. [Python for Data Analysis](http://shop.oreilly.com/product/0636920023784.do). O'Reilly. Available as an e-book from the library <http://oskicat.berkeley.edu/record=b20726801~S1>, and as a PDF <http://it-ebooks.info/read/1041/>. You may also want to access the online [Documentation available in pdf](http://pandas.pydata.org/pandas-docs/stable/pandas.pdf) <http://pandas.pydata.org/pandas-docs/stable/pandas.pdf>.
4. [Statsmodels Documentation](http://pyvideo.org/video/970/pydata-data-analysis-in-python-with-pandas). 2013. This is online documentation for the statsmodels library we will use for developing statistical models. <http://pyvideo.org/video/970/pydata-data-analysis-in-python-with-pandas>













These e-books are available for free from the Berkeley library system and can be considered additional resources:

- Rossant, Cyrille. 2013. [Learning IPython for Interactive Computing and Data Visualization](http://www.packtpub.com/learning-ipython-for-interactive-computing-and-data-visualization/book). PACKT Publishing.
- Westra, Eric. 2013. [Python Geospatial Development - Second Edition](http://www.packtpub.com/python-geospatial-development-second-edition/book). PACKT Publishing. Available as an e-book from the library <http://site.ebrary.com/lib/berkeley/docDetail.action?docID=10439369>.

Assignments Summary:

Date	Details	
Thu Aug 28, 2014	 Introduction and Course Overview	11am to 12:30pm
Tue Sep 2, 2014	 Data visualization using GeoCanvas and QGIS	11am to 12:30pm
Thu Sep 4, 2014	 Data visualization using CartoDB	11am to 12:30pm
Mon Sep 8, 2014	 Assignment 1: Data Visualization	due by 11:59pm
Tue Sep 9, 2014	 Working with strings, lists, and dictionaries	11am to 12:30pm
Thu Sep 11, 2014	 Working with strings, lists, and dictionaries	11am to 12:30pm
	 Exercise 1: Python Script	due by 11:59pm
Tue Sep 16, 2014	 Conditions, loops, functions	11am to 12:30pm
Thu Sep 18, 2014	 Conditions, loops, functions	11am to 12:30pm
	 Exercise 2: Python Script	due by 11:59pm

Date	Details	
Tue Sep 23, 2014	 From Rows, to Lists, Arrays, and DataFrames	11am to 12:30pm
Thu Sep 25, 2014	 Data Wrangling and Reverse Geocoding	11am to 12:30pm
Mon Sep 29, 2014	 Assignment 2: Python + Map Visualization (https://bcourses.berkeley.edu/courses/1247349/assignments/5337444)	due by 11:59pm
Tue Sep 30, 2014	 Pandas 1: Reviewing Basic Operations	11am to 12:30pm
Thu Oct 2, 2014	 Pandas 1: Basic Operations Continued	11am to 12:30pm
	 Exercise 2.5: Review of Basic Pandas Operations (https://bcourses.berkeley.edu/courses/1247349/assignments/5579473)	due by 11:59pm
Fri Oct 3, 2014	 Exercise 3: Pandas (https://bcourses.berkeley.edu/courses/1247349/assignments/5337445)	due by 11:59pm
Tue Oct 7, 2014	 Pandas Data Wrangling	11am to 12:30pm
Thu Oct 9, 2014	 Pandas Data Wrangling and Visualization	11am to 12:30pm
Fri Oct 10, 2014	 Exercise 4: Pandas Visualization (https://bcourses.berkeley.edu/courses/1247349/assignments/5337448)	due by 11:59pm
Tue Oct 14, 2014	 Data science with Pandas	11am to 12:30pm
Thu Oct 16, 2014	 Data science with Pandas	11am to 12:30pm
Tue Oct 21, 2014	 Spatial analysis in Python	11am to 12:30pm
Wed Oct 22, 2014	 Assignment 3: Data Science with Pandas (https://bcourses.berkeley.edu/courses/1247349/assignments/5337449)	due by 11:59pm
Thu Oct 23, 2014	 Spatial analysis in Python	11am to 12:30pm
Tue Oct 28, 2014	 Pandana: spatial analysis on networks	11am to 12:30pm
	 Exercise 5: Spatial Analysis (https://bcourses.berkeley.edu/courses/1247349/assignments/5337450)	due by 11:59pm
Thu Oct 30, 2014	 No Class: ACSP Conference	11am to 12:30pm
Tue Nov 4, 2014	 Local Code: Nicholas de Monchaux	11am to 12:30pm
Wed Nov 5, 2014	 Exercise 6: Network Analysis (https://bcourses.berkeley.edu/courses/1247349/assignments/5337451)	due by 11:59pm

Date	Details	
Thu Nov 6, 2014	 Using APIs and Web Mapping with Leaflet: Sam Maurer	11am to 12:30pm
Fri Nov 7, 2014	 Exercise 7: Maps with Leaflet https://bcourses.berkeley.edu/courses/1247349/assignments/5337455	due by 11:59pm
Tue Nov 11, 2014	 Cal Holiday	11am to 12:30pm
Thu Nov 13, 2014	 Web data	11am to 12:30pm
Fri Nov 14, 2014	 Exercise 8: Web data https://bcourses.berkeley.edu/courses/1247349/assignments/5337458	due by 11:59pm
Tue Nov 18, 2014	 Open Data: Code for America: Michal Migurski, CTO	11am to 12:30pm
Thu Nov 20, 2014	 Berkeley Open Data: Donna LaSala	11am to 12:30pm
Tue Nov 25, 2014	 LocalData - Prashant Singh	11am to 12:30pm
Mon Dec 1, 2014	 Assignment 4: Using Open Data APIs https://bcourses.berkeley.edu/courses/1247349/assignments/5337457	due by 11:59pm
Tue Dec 2, 2014	 CivicInsight - Eddie Tejada	11am to 12:30pm
Thu Dec 4, 2014	 Wrap-up and Final Assignment	11am to 12:30pm
Sun Dec 14, 2014	 Assignment 5: LocalData https://bcourses.berkeley.edu/courses/1247349/assignments/5337459	due by 11:59pm