CP 101 Introduction to Urban Data Analytics (4 credits)
Lecture: Tues/Thurs 11:00 am–12:30 pm; Wurster 112
Labs: (1) Thus 2:00–3:30 pm, (2) Fri 12:30–2:00 pm; Wurster 214C

Course Description
This course (1) provides a basic intro to census and economic data collection, processing, and analysis; (2) explores visualization and story mapping techniques in planning; (3) teaches methods of urban analytics; and (4) provides a socio-economic-political context for the urban analytics movement, focusing on data ethics and governance.

Synopsis
CP 101 introduces students to the systematic analysis of urban data in its institutional context. Recognizing that defining this context relies on critical thinking with regard to economic, social, and environmental outcomes, this course explores how stakeholders conceptualize “smart” and inclusive urbanity. Accordingly, this course teaches students systematic approaches to collecting, analyzing, modeling, and interpreting quantitative and qualitative data used to inform robust research, and, ultimately, urban planning practice and policymaking. Beyond instruction in urban data science and analytics, students will be introduced to theory and critical discourses on topics such as big data, open data and e-governance. Instructors will expect students to engage with technical and theoretical - with particular focus placed on ethical - considerations associated with these subjects in lecture and laboratory sections. The course will introduce students to programming in Excel and Python, using open source software, accessing open and scraped data, and other tools and techniques for urban analysis.

The course will be structured following 3 modules:

Module 1: Introduction to Urban Data
During this module students will be introduced to fundamental data applications and ethical dilemmas in urban planning. They will be instructed on sourcing data, analyzing data via statistical testing, and presenting data through written reports and visualizations. In Module 1, students will gain skills in working with Census and economic data, statistical testing, and static data visualization. The deliverable for this module will be a descriptive profile of a Bay Area neighborhood.

Module 2: Mapping the City
In the course’s second module, students will learn different tools to make maps. We will gain an understanding of the basic elements of maps, how to map with online programs and geographic information systems software (Carto), and how to construct story map websites. Students will produce a story map as the product for this module.
Module 3: Data Science for Planners: Big Data and Analytics

In the course’s final module, students will use knowledge acquired in earlier modules to explore urban data science techniques. Classes will cover topics such as big data, open data, volunteered geographic information, smart cities, and civic hacking; and students will gain skills in real-time and crowd-sourced data collection and use. As the final project for the class, students will use novel sources of data to answer a research question of their choice.

Prerequisites
CP 101 reserves seats for CED majors, as well as others from around campus (e.g., Data Science majors); others can enroll with the permission of the instructor. No prior statistics coursework is assumed. This class provides a foundation to pursue further undergraduate data science courses at UC-Berkeley. For Urban Studies CP 101 satisfies one of the four additional City Planning courses for Upper Division Urban Studies Core. For the City Planning minor, CP 101 satisfies one of the four additional City Planning courses for Upper Division courses under List 1. For SED, CP 101 can count as an Upper Division outside SED major for Fall 2016 admits and later.

Course Requirements
All computer assignments will involve the use of data and software available online, through campus IT, or through the DCRP Computer Lab. We will be holding labs in a CED computer lab - lab computers use Windows' operating system. You may use your personal computer in lab, if you’d like. We'll be teaching labs for both PC and Mac users. If you are unfamiliar with either Windows or DCRP’s computer lab and want to use lab computers, you are encouraged to take advantage of the CED’s library services (Environmental Design Library Website) or consult with course instructors during office hours.

Students will be expected to have a working knowledge of the Microsoft Office Suite - specifically Excel, PowerPoint, and Word - for this course. Students are welcome to explore the use of free, open software interfaces, such as the R language for statistical computing, the RStudio software package; and the QGIS project; however, please check with the graduate student instructor before using any alternative software. Labs will be offered to introduce students to Excel, American FactFinder & Social Explorer, Carto, WordPress, and Python programming.

Grading & Assignments and Reading
Students will be expected to read a number of articles/chapters/etc. prior to lecture and lab sessions. For ten sessions, students are asked to submit response questions online on our class bCourse site. We will randomly call on students in each class. Students will also be expected to attend and actively participate in class and lab. They will take two in-class midterm quizzes and one lab midterm; complete three group projects; and have the option of a final presentation. Grades will be assigned as follows:
Assignment 1 (Neighborhood Profile): 15%
Assignment 2 (Story Map): 15%
Assignment 3 (Final Project): 25%
Midterm Quiz #1: 10%
Lab Midterm: 10%
Midterm Quiz #2: 10%
Reading Response Questions: 10%
Class Participation: 5%

Extra credit: Final presentation (5 points)

**Reading Responses**

Students are expected to respond to ten different sets of class readings by submitting at least three related questions online on the course bCourse site. Only the class sessions marked with a * are available for commentaries; these are the sessions with more theoretical or critical readings, rather than technical texts. The questions should react critically to the concepts and analyses presented. During each class, the instructor will randomly select one or more of the questions to read and discuss during the lecture. Questions should be posted by midnight the day before class to the bCourses discussion thread for the lecture (e.g., February 24 for February 25 class).

**Assignments**

All assignments will be conducted by groups of two students (though students may petition the instructor to expand the group to three). Groups will be assigned randomly by the instructors, who will match students from technical or STEM majors with environmental, social science, or humanities majors (to the extent possible). In the first two assignments (Neighborhood Profile and Story Map), students will explore the phenomenon of neighborhood change through in-depth analysis of census and economic data, as well as story-mapping. The Neighborhood Profile will be a Word document of 8-10 pages (1.5 spaces), plus appendices. The Story Map will consist of an interactive online map, with significant explanatory text.

For the third assignment, students will explore a research question of their own choosing, using big data and/or open data portals, as well as analysis and visualization techniques learned in class. Students may select their own partners for this project. This deliverable should include either an online project and a narrative of 15-20 pages (1.5 spaces), including references in proper format. Students will submit a one-page description of their research question and approach in the beginning of April and work closely with the instructors to develop a methodological approach. The semester will culminate with short (8-10 minute) presentations of Assignment 3 (during reading week); this presentation is optional and will be for extra credit.

**Exam**

There will be a short midterm quiz held during the second module to assess student learning of course topics; a midterm lab held after the second module focusing on analytic techniques
taught in the first two months of class; and a second midterm quiz held during the third module. The quizzes will consist primarily of definitions and short answers based on the class readings and lectures. The in-lab midterm will be based on the lab exercises and assignments.

**Attendance and Class Participation**
Students are expected to attend *every* class lecture. In addition to attending, students are expected to actively contribute to class discussions and ask questions. Participation will be evaluated by instructors based on a combination of attendance, and observed engagement and participation.

There are ten computer labs (plus two optional labs and one midterm lab). Students are expected to attend six of these labs (of their choice). Instructors will take attendance at the lab and lab attendance will be part of the final grade.

**Course Materials**
CP 101 has one **required** book and one **optional** book. There is no course reader. All readings that are not part of the required book will be posted to the CP 101 bCourses. Also, please see the course website, [http://www.cp101.org](http://www.cp101.org), for a variety of resources related to the course.

**Required:**

**Optional:**

We expect you to purchase the book at the student store, or via an online bookseller.

Lecture slides and other course materials will be posted on the CP 101 bCourses site.

**Office Hours**
Prof. Karen Chapple ([chapple@berkeley.edu](mailto:chapple@berkeley.edu)): Wed and Thurs 4-5 PM, 228 Wurster (sign up via website, [www.karenchapple.com](http://www.karenchapple.com))

Manuel Santana Palacios ([manuel.santana@berkeley.edu](mailto:manuel.santana@berkeley.edu)): Walk-in group office hours on Thursdays from 1 to 2 pm and one-on-one 15-minute slots on Fridays from 11 am to 12 pm (sign up via [http://calendly.com/mansanp](http://calendly.com/mansanp))

Chester Harvey ([chesterharvey@berkeley.edu](mailto:chesterharvey@berkeley.edu)): office hours TBD.

**Statement on Academic Integrity**
Any test, paper or report submitted by you and that bears your name is presumed to be your own original work that has not previously been submitted for credit in another course. You may use words or ideas written by other individuals in publications, websites, or other sources, but only with proper attribution. If you are not clear about the expectations for completing an assignment or taking an exam, be sure to ask a course instructor.

You should also keep in mind that as a member of the campus community you are expected to demonstrate integrity in all of your academic work and be evaluated on your own merits. The consequences of cheating and academic misconduct — including a formal discipline record and possible loss of future opportunities — are not worth the risks.

**Statement on Accommodations for Students with Disabilities**
If you have been issued a letter of accommodation from the [Disabled Students Program (DSP)](https://disabledstudentsprogram.com), please see the course lead instructor as soon as possible to work out the necessary arrangements. If you need an accommodation and have not yet seen a Disability Specialist at the DSP, please do so as soon as possible.

If you would need any assistance in the event of an emergency evacuation of the building, the DSP recommends that you make a plan for this in advance. (Contact the DSP access specialist at 643-6456.)

**Statement on Scheduling Conflicts**
Please notify course instructors by the second week of the term about any known or potential extracurricular conflicts (such as religious observances, graduate or medical school interviews, or team activities). We will try our best to help you with making accommodations, but cannot promise them in all cases.

**Course Schedule - Readings & Assignments**
The course schedule/readings are outlined below.

---

**Module 1: Introduction to Urban Data**

**Tuesday, January 21: Introduction to Urban Analytics**
Singleton, Spielman, and Folch (2018) Chapter 1, “Questioning the city through urban analytics”


Optional:
Thursday, January 22: Data Fundamentals for Planners*


Wheelan (2013) Chapter 7, "The Importance of Data"

Lab 1 (optional): Introduction to Excel

Tuesday, January 28: Metadata: Understanding the US Census*


B. Strasser and P. Edwards, “Big Data is the Answer... But What is the Question?” Osiris 32, 2017: pp. 328-345


Thursday, January 30: Using Census Data


Lab 2: Using data.census.gov

Tuesday, February 4: Stats and the American Community Survey
Singleton, Spielman, and Folch (2018) Chapter 6, “Explaining the city”

**Thursday, February 6: Static Data Visualization**


Optional:
Check out Piktochart for infographics, [www.piktochart.com](http://www.piktochart.com)

And the whole Tufte book is great – especially check out Chapter 1, “Graphical Excellence.”

**Lab 3: Excel Basics and Generating Charts**

**Tuesday, February 11: Neighborhood Data and Indicators: The Urban Displacement Project***


Urban Displacement Project, [www.urbandisplacement.org](http://www.urbandisplacement.org) [SKIM]

**Thursday, February 13: Introduction to Economic Data and the Longitudinal Household-Employment Data**

Lab 4: Calculating Margins of Error; Accessing Census and Economic Data via Social Explorer Accessing Local Employment-Household Dynamics Data

Assignment #1 due Tuesday, February 18!

Module 2: Mapping the City

Tuesday, February 18: Spatial Data & GIS Fundamentals

Singleton, Spielman, and Folch (2018) Chapter 4, “Visualizing the city”
Monmonier, Mark. 1996 Chapters 1, 2, 3, 4, and 10 How to Lie with Maps. University of Chicago Press.

Thursday, February 20: Accessibility*


Optional:


Lab 5: Carto Part I

Tuesday, February 25: Introduction to Story Mapping*

Examples to review:

● Displacement in the Bay Area: http://www.antievictionmap.com/#/narratives-of-displacement/
● Mapping Segregation in DC.  
  http://jmt.maps.arcgis.com/apps/MapJournal/?appid=061d0da22587475fb969483653179091

● Creating a neighborhood change zoning plan for Spruce Hill  
  https://storymaps.arcgis.com/stories/378dcb733ffe4698b97ce8cd2884b3dd

● Gangs of Los Angeles (2015):  
  http://maps.google.com/gallery/details?id=zMC7tfcRop6s.koD9cCcyHJ_0&hl=en


● Katrina +10: A Decade of Change in New Orleans:  

You can find more examples at ESRI’s gallery: https://www.esri.com/en-us/arcgis/products/arcgis-storymaps/overview

Thursday, February 27: Participatory Mapping (Kate Beck, SafeTREC, invited)*


Check out the Street Story Project

Lab 6: Carto Part II

Tuesday, March 3: Power, Place and Mapping (Steve Spiker, invited)*


Thursday, March 5: Midterm Quiz #1

Lab 7: OCF/WordPress and Embedding Carto Maps

Assignment #2 due Tuesday, March 10!
Module 3: Data Science for Planners: Big Data and Analytics

Tuesday, March 10: Introduction to Big Data*


Thursday, March 12: Big Data – and Ethics -- for Planners*


Crawford, Kate. “The Trouble with Bias”, NIPS conference keynote, December 2017 (especially minutes 14:00 - 38:00) https://www.youtube.com/watch?v=fMym_BKWQzk


Lab 8: In-Lab Midterm

Tuesday, March 17: Complex Urban Modeling: Machine Learning (guest speaker TBD)*


Pedro Domingos, A Few Useful Things to Know About Machine Learning (2012)

Thursday, March 19: Deploying Data Science Techniques in Urban Research

NO LAB THIS WEEK!!!

Week of March 25-9: NO CLASS, SPRING BREAK!!

Tuesday, March 31: Urban Data Analytics


*Thursday, April 2: Volunteered Geographic Information (Sam Maurer, guest speaker)*


Lab 9: Python - Intro to Jupyter/Python/Pandas

*Tuesday, April 7: Open Data & Using Portals (Jason Lally, DataSF, Invited)*


*Thursday, April 9: Midterm Quiz #2*

Lab 11: Python - Web Scraping

*Tuesday, April 14: Interactive Visualizations*


Explore additional interactive visualizations here:

http://polygraph.cool/history/
http://goodcitylife.org/chattymaps/index.html
http://hubcab.org/#13.00/40.7219/-73.9484
http://218consultants.com/interactive-suitability-map/ (Look at all 3 interactive maps)
https://ourworldindata.org/a-history-of-global-living-conditions-in-5-charts/
http://www.urban.org/features/vision-equitable-dc
and of course, http://www.urbandisplacement.org!

Optional:

**Thursday, April 16: Defining Smart Cities in Theory and Practice** *(guest speaker TBD)*


**Lab 12: Python - Mapping**

**Tuesday, April 21: Smart Institutions & e-Governance** *(Raleigh McCoy, MTC, invited)*


Also look over: https://smartcitizen.me/

**Thursday, April 23: Civic Hacking and Equity** (*Emily Wasserman, Code for SF, invited)*


**Lab 12: Python - Working with Big Data**

**Tuesday, April 28: Presenting Data**


Optional:


**Thursday, April 30: The Inclusive Smart City**


**Lab 13: Open Help Session (optional)**

*Assignment #3 due Friday, May 8!*
The Data Analytics colloquium involves three central learning components. 1) regular engagement with guest presentations and community activities in data analytics, 2) group discussion featuring critical analysis and connection of themes found in the guest presentations and in related data analytics topics, and 3) preparation and refinement of professional communication skills necessary for the required internship component of the data analytics major. This is a capstone seminar for the Data Analytics major in which students work on independent research projects in a collaborative seminar setting. Problems may derive from internship experiences, courses of study at Denison, or another source subject to instructor approval.