Data Science for Smart Cities

T 2:00P-4:00P | Wurster 106
Units: 2

Instructor: Marta C. González

Design and operation of smart, efficient, and resilient cities nowadays require data science skills. This course provides an introduction to spatial analysis using data from transportation systems, power grids, and communication networks, putting into context their population access via census data. Methods include: network science, introduction to machine learning and complex systems thinking to study urban systems. Programming language is Python.

Prerequisites: An undergraduate-level understanding of statistics, and programming skills is assumed

List of lectures

Lec 1: January 23rd	Introduction to Network Science	
Lec 2: January 30th	Network Metrics	
Lec 3: February 6th	Network Models (SW and RG)	Assignment 1
Lec 4: February 13th	Network Models (Barabasi Model)	Assignment 2
Lec 5: February 20th	Clustering Census Data (cenpy + sandiego)	
Lec 6: February 27th	Clustering Census Data (Georgia example)	Assignment 3_I
March 5th:	No class	Assignment 3_II
Lec 7: March 12th	Discussion of Research Ideas	
Lec 8: March 19th	Centralities and Network Resilience	
March 26th	Spring Recess	
Lec 9: April 2nd	Clustering Networks (Community Detection ,Airports Paper, Guest Lecture)	
Lec 10: April 9th	OSMNx (Google API)	Assignment 4
Lec 11: April 16th	Weighted Networks + BART (GTFS)	
Lec 12: April 23rd	Class Summary Tools of Spatial Analysis	Paper Draft
Final Presentations	Project Presentations (April 30th, May 1st and May 2nd)	

- Problem sets

	Covers	Assigned	Due
Assignment 1	L1-L2	01/30	02/08
Assignment 2	L3-L4	02/10	02/18
Assignment 3 part 1	L5-L6	02/20	02/28
Assignment 3 part 2	L5-L6	03/01	03/10
Assignment 4	L7-L10	03/12	04/02
Checkup (paper draft)	L1-L12	04/04	04/18
Final Project	L1-14		04/30, 05/01 and 05/02

- Assignments
  - 4 Homework assignments (problem sets)
  - 1 Assignment for Project Plan (a Paper Draft)

- 1 Final project: Paper report and oral presentation

- Unless stated otherwise, problem sets should be solved individually. Electronic versions must be uploaded to the bcourses site.
- 20% for late assignments reduction for late assignments more than a 48 hours grace period
- 15% for each problem set (4x15 =60%)
- 10% Class Participation
- 15% Project check-up (paper draft)
- 10% Final project paper
- 5% Project presentation

- Lectures: Tuesdays 2-4pm

- Contact information:

  Prof. Marta C. Gonzalez
  Office hours: Tuesdays 4:00-5:30 PM @Room 406c; martag@berkeley.edu
  Thursdays 9:00-10:00 by appointment via email (20 min slots)
  Via ZOOM https://berkeley.zoom.us/j/4124372482

Refer to bourses for the most up-to-date information.
The piazza site is at: https://piazza.com/berkeley/spring2023/ce88cp88

Textbook:

1) *FIRST COURSE IN NETWORK SCIENCE,* ***Author:*** *MENCZER,* ***ISBN:*** *9781108471138*

2) Geographic Data Science with Python:

https://geographicdata.science/book/intro.html (*Clustering & Regionalization)*

Desired Student Outcomes:

1. An ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics
2. An ability to apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors
3. An ability to communicate effectively with a range of audiences
4. An ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions