Data-driven discovery and visualization has emerged as an important component of graduate education. Majors/minors/certificate programs have been created recently in areas such as “Big Data Social Science”, “Social Data Analytics”, and “Computational Social Science” at Stanford University https://iriss.stanford.edu/news/new-certificate-computational-social-science, UC Berkeley mdp.berkeley.edu, Penn State University bdss.psu.edu/education.html, and many other institutions. The intended audience for this course is graduate students in both computational and social sciences.

Goals: Welcome to CMPS 263! The main goal of this course is to understand how visual representations can help in the analysis of and elucidation of complex data. Data curation, including issues of data consistency, completeness, context, reliability, and uncertainty will be discussed. Key design principles and techniques for visualizing data will be explored. A rich variety of examples of infographics encompassing several applications and disciplines will be critiqued. Guiding principles of leading visualization thinkers and design artists will be presented and juxtaposed leading to assist in carving out your own unique style of visualization design.

Expected Outcomes:
1. Critically evaluate visualizations and suggest refinements or improvements
2. Identify, curate, and derive (or compute) data that supports a task or theme
3. Conceptualize effective visualizations using design, cognition, and perception principles
4. Communicate prototype visualization design for effective feedback from users
5. Design coherent and clear visualizations using storytelling

Textbooks:
1. The Functional Art: An Introduction to Information Graphics and Visualization by Alberto Cairo, 2013

Recommended Books:
2. Data Visualization: A Successful Design Process by Andy Kirk
3. Visual Thinking for Design by Colin Ware, 2008

Readings: There will be weekly state-of-the-art readings from practitioners and analysts of practical and impactful visualizations. Names of some of the practitioners include Jacques Bertin, Otto Neurath, Edward Tufte, Stephen Few, Nigel Holmes, William Cleveland, and M. E. McGill.

Websites: In addition, several active websites and blogs on visualization will be introduced. An example is http://www.nyu.edu/ipk/people/laura-noren by Laura Noren. Check out thesocietypages.org/graphicsociology.
Overview (catalog description): High-quality interdisciplinary research using socio-economic data and software available on the internet; data curation, computation, and visualization to strengthen scientific inquiry to bear on large-scale societal problems. Applications include inequality, poverty, water, energy, environment, health, education and democracy.

Prerequisites: Evidence of software training is required. Ability to navigate internet and dig down the databases such as the ones provided by US Census or the United Nations or the World Bank data is needed. Prior experience with problems of interest to society is preferred and needed to identify a class project. Skills and/or motivation to learn an-easy-to-use software such as Excel or Social Explorer to query the data will be needed for class projects. Familiarity with Powerpoint will be needed for class presentation. Additional programming skills in Python or R or D3 or Javascript or any open source web-based programming tools are additional plus but not required.

Class Requirements: Class presence is expected and required for most classes. There will be in-class participation and discussion of various visualization processes, products, and evaluation approaches. There will be regular homework consisting of reading assignments, data discovery and/or curation exercises, and visualization design and evaluation. There will be one in-class presentation from an assigned list of topics. There will be three projects – two mini-projects roughly due during the 3rd and 6th week, and one final project during the 10th week of the quarter. There will be no examination. The final class evaluation will be based as follows: 10% class participation, 20% homework, 10% class presentation, 12% each for the two mini-projects, and 36% for the final project. Extra bonus points will be provided for exceptional effort and work.

Projects: The three projects will focus on (i) data identification, curation, and computation (derivation) of publicly available data on a theme of interest to the student, (ii) designing a series of visualizations using design, cognitive, and perceptual principles, critiquing your own designs, followed by refinements, and improvements, and (iii) a finished visualization product elucidating a theme of students’ interest.

An Example of Design Critique (to be conducted as part of the class):
Consider the following visualization of droughts and deluge by New York Times:

Now address the following questions:
1. What is the main message to you from this graphic?
2. What tasks can you perform: presentation, comparisons, clustering (organizing), correlating (chapter 2 of Albert Cairo), …
3. What, in your opinion, is the main intended message of this graphic?
4. What variables are visually presented in static view? How are these variables encoded?
5. Which variables are presented through interaction? How are these variables encoded?
6. Assess various visualization elements based on the visualization wheel (Chapter 3 of Albert Cairo) and draw a radar chart.
7. Do you like or dislike this visualization? Why?
8. How would you improve this graphic to better support understanding (perhaps using the same or lesser amount of space)?
9. Would you support adding “art” elements to this graphic following Nigel Holmes? If yes, what changes would you incorporate and why?

How to enroll: Registration for this class requires permission code from the instructor. Please email the instructor lodha@soe.ucsc.edu with the subject heading: CMPS 263