Instructor: Suresh K. Lodha; lodha@soe.ucsc.edu; E2-361 or E2-262
Class Location: TBD
Meeting Times: TuTh 6:00-7:45pm
Office Hours: TuTh 5:00-5:45pm, after TuTh class, and by appointment
Class Web Page: classes.soe.ucsc.edu/cmps080l/Win16
TA/Tutor: Daniel Helkey (dhelkey@soe.ucsc.edu) or TBD
Tutors: Venkat Thota(vthota@ucsc.edu) or TBD

Goals: Welcome to CMPS80L! The main goal of this course is to understand how visual representations can help in the analysis of and elucidation of complex controversial data from social sciences.

First, we ask data science questions. Can you trust the data? Is it true? What is the source? Is it consistent? What are the sources? Are they reliable? What is missing? Can it be interpreted in various ways? How large is the margin of error? How often does it change?

Second, what is really relevant in this data? What subset do I choose? Why do I choose this? What point do I want to communicate? To whom? To what purpose? Which statistical summaries do I choose – mean, mode, variance? Is there bias? Are variables correlated? What ranking or scales should I choose? How do I fit a straight line and should I? How does it transpire that same data can be used to create opposite messages? Is data extrapolated? What are the underlying assumptions?

For the more sophisticated, which variables matter most (factor analysis)? Are there clusters (cluster analysis)?

Third, ah, now that I trust and understand the data and know what I want to communicate and to whom, how do I communicate the data visually? Bar chart, line chart, pie chart, scatter plot, but why? Do they make sense? What colors do I choose and why? What legends, labels, axes, annotations, should I use and why?

Finally, how can I make an impact using visualization?!!!!??

There will be programming exercises in R that allow you to create simple visualizations. Examples will be drawn from energy, environment, economics, education, and empowerment.

Textbooks:

Readings will be assigned from the textbook in the very first class. It is important that you have ready access to this book.

Recommended (potentially helpful) Books on R (although several internet resources and readings) may suffice:
1. R in a Nutshell by Joseph Adler, O’Reilly Books

Additional Readings: There will be several additional weekly state-of-the-art readings, videos, websites, blogs, and visualizations to read/view and critique.
Overview (catalog description):

Prerequisites: Although there are no prerequisites for the class, you will be a good candidate to take this class if (i) you love beautiful visualizations, (ii) you truly care about some social issues, and (iii) sometime you truly wonder which data to believe? Although statistical computation and programming will be covered in the class, I believe, motivation will be the key to success in the class.

Evaluations: Unlike previous offerings, this class will use traditional evaluation techniques consisting of written homework, programming assignments, programming examination, quizzes and examinations. There will be opportunities to earn bonus points for class presence, meaningful engagement, effort, and outstanding contributions. There will be four components to the evaluation:

I. Homework (concepts and practice): 25 points
II. Programming Assignments: 25 points
III. Programming Examination (on Tu of the last week of classes): 25 points
IV. Conceptual understanding evaluated via a combination of a couple of quizzes-in-class, one written examination (during the 9th week of class), and a verbal examination on concepts and practice (on Th of last week of classes): 25 points
Bonus Point Opportunities: 20
Extra Credit: 10+

A minimum of 50% in each of the four components above and an overall total of 60 points is required to pass the class. Tentative grading scale is as follows (this may be tweaked only in favor of students):
<= 50 F; <= 60 D
C: 61-70; C+ 71-80
B-: 80-86; B: 87-93; B+:94-99
A-: 100-105; A: 106-110; A+:111-120

Outside Help: Students are encouraged to help and learn from each other and any other resources that they can find including peers, internet, teachings assistants, tutors, readers, and experts. All such help must be acknowledged while submitting the homework. You must list all the references that you have used at the end of the homework. If it is a website, full url weblink must be provided along with a meaningful title of the weblink. If it is a person (other than the instructor or the teaching staff), list the name and relationship (classmate, mom, friend, found from quora or facebook).

The final submitted work must be in student’s own words (no cut and paste). The final submitted programming assignments must be executed by the student himself/herself. The focus is that students can use every possible existing resource to meet the goal of the course: to create compelling accurate impactful socio-economic visualizations.

Posted Date: 09/01/2015