Data Visualization for Architecture, Urbanism and the Humanities

- Columbia University | GSAPP and A&S | ARCHA4892 | Spring 2017
- Fridays 9am - 11am | Studio @ Butler
- Office hours: Mondays 10am - 12pm (previous email required)
- Professor: Juan Francisco Saldarriaga (jfs2118)
- Teaching Assistant: Buck Fivel (brw2103)

Course Overview

This course provides an introduction to data visualization theory and methods for students entirely new to the fields of computation and information design. Through a series of in-class exercises and take-home assignments, students will learn how to critically engage and produce interactive data visualization pieces that can serve as exploratory and analytical tools. The course is part of a larger initiative, hosted by the Center for Spatial Research (http://c4sr.columbia.edu/) to teach courses in the critical use of digital tools across fields in architecture, urbanism, and the humanities.

The course will be centered around a semester long data visualization group project, through which the students will learn the basics of data visualization, data analysis, data collection, programming and versioning control. However, even though the course will teach specific visualization tools, the main conceptual thread will be centered around how to work with data, both in the humanities and in architecture and urbanism. Students will define their final projects around their own interests, and will bring their own datasets into their final projects.

General Topics

- Data visualization history and concepts
- Working with data: collection and analysis
- Basic programming skills and web languages (HTML, CSS, Javascript)
- Interactivity and online data visualization
- Collaborative work and versioning control (Github)
- Working with text data
- APIs and Web-scraping

Evaluation and Grading

- 10% Class participation and discussion
- 20% Individual assignments and tutorials
- 15% Midterm presentation
- 20% Data visualization critique
- 35% Final project and final presentation and report

Resources and Materials

Course files, tutorials and presentations will be located on Courseworks, on the Center for Spatial Research (http://c4sr.columbia.edu/) website and on this repository.

The readings for the class will be duly uploaded to Courseworks. Similarly, students will be required to submit their assignments by uploading them to Courseworks. Finally, the class will also rely heavily on submissions to the blog. Students will be required to upload some of their own work as well as inspirational material, encouraging and developing a critical stance and visual skills.

Link to the blog (http://mapping2016fall.tumblr.com/) . **----Needs new Link----**

Link to number of posts (https://docs.google.com/spreadsheets/d/106JrEOQ64bOK1pqhBBwC-_z5eOU_4yphc7Fz0otWysI/edit?usp=sharing) . **----Needs new Link----**

Schedule

Week 1: Introduction to course and setup

January 20

- Course administration and syllabus
- Overview of the course
- Assignments and final project
- Resources
- Why data visualization (discussion)
- Basic infrastructure (stack)
- Introduction to HTML and CSS
- Setup of local server

Week 2: Basic web concepts and Github

January 27

- Versioning control basics
- Setup a github repository
- Setup a github projects page

Week 3: Data visualization matrix and basic programming concepts

February 3

- Introduction to p5.js
- Introduction to JavaScript
- Console
- Basic programming exercises
- **Assignment 1 Due**: Matrix of data visualization and data types

Week 4: Introduction to data visualization

February 10

- Data visualization history, examples and theory
- Minard, Snow, Nightingale, Bertin, Tufte
- Programming exercises:
    - Variables
    - Loops
    - Conditionals

Week 5: Presentations

February 17

- In class presentations
- **Assignment 2 Due**: Detailed presentation of data visualization project

Week 6: Contemporary data visualization

February 24

- Contemporary concepts and examples of data visualization
- Basic data analysis concepts and techniques
- Programming exercises:
    - Functions
    - Objects

Week 7: Working with text data

March 3

- Working with text data: presentation by Michelle McSweeney
- **Assignment 3 Due**: Visualization of one dataset

Week 8: Advanced programming exercises

March 10

- Advanced programming exercises:
    - Functions
    - Objects

Week 9: Spring Break (no class)

March 17

Week 10: Midterm review

March 24

Week 11: APIs

March 31

- APIs and web-scrapping techniques

Week 12: Interactivity

April 7

- Interactivity concepts and techniques
- **Assignment 4 Due**: Data visualization critique

Week 13: Work in class

April 14

- Work in class

Week 14: 3/4 Review

April 21

Week 15: Work in class

April 28

- Work in class

Week 16: Final review

May 5

Assignment Schedule (Due Dates)

- **January 27:** Basic website hosted on Github
- **February 3:** Data and visualization matrix
- **February 17:** Data visualization project presentation
- **March 3:** One dataset visualization
- **March 24:** Midterm review
- **April 7:** Data visualization project critique
- **April 21:** 3/4 review
- **May 5:** Final review

References

Books

- Data Visualization:
  - *Data Flow: Visualizing Information in Graphic Design*
  - *Data Flow 2: Visualizing Information in Graphic Design*
  - *Data Points*, Nathan Yau
  - *Atlas of Shrinking Cities*, Beyer Elke
  - *Visualizing Information for Advocacy*, Tactical Technology Creative
  - *Design for Information*, Isabel Meirelles
  - *Semiology of Graphics*, Jacques Bertin
  - *The Visual Display of Quantitative Information (2nd Edition)*, Edward R. Tufte
  - *Envisioning Information*, E. R. Tufte
  - *Visualization Analysis and Design*, Tamara Munzer
  - *Dear Data*, Giorgia Lupi, Stefanie Posavec
  - *Show Me the Numbers: Designing Tables and Graphs to Enlighten*, Stephen Few
  - *Now You See It: Simple Visualization Techniques for Quantitative Analysis*, Stephen Few
  - *Information Visualization: Perception for Design*, Colin Ware
  - *The Functional Art: An Introduction ton Information Graphics and Visualization*, Alberto Cairo
- Programming:
  - *Generative Design*, Hartmut Bohnacker, Benedikt Gross, Julia Laub, Claudius Lazzeroni
  - *Processing: A Programming Handbook for Visual Designers (Second Edition)*, Casey Reas and Ben Fry
  - *The Nature of Code: Simulating Natural Systems with Processing*, Daniel Shiffman

    - *Eloquent JavaScript* *(http://eloquentjavascript.net/)* , Marijn Haverbeke
    - *Beginning JavaScript (4th Edition)*, Paul Wilton
    - *JavaScript: The Definitive Guide (4th Edition)*, David Flanagan
    - *HTML & CSS: Design and build websites*, Jon Duckett
    - *Program or Be Programmed: Ten Commands for a Digital Age*, Douglas Rushkoff
- Typography:
    - *Thinking With Type*, Ellen Lupton

Blogs & Websites

- Visualizing Data (http://www.visualisingdata.com/)
- Flowingdata (http://flowingdata.com)
- Periscopic (http://periscopic.com)
- Visualizing.org (http://visualizing.org)
- Accurat (http://accurat.it)
- Moritz Stefaner (http://truth-and-beauty.net/)
- Nocholas Felton (http://feltron.com)
- Infosthetics (http://infosthetics.com)
- Visualcomplexity (http://visualcomplexity.com)
- The Economist - Graphic Detail (http://www.economist.com/blogs/graphicdetail)
- New York Times - The Upshot (http://www.nytimes.com/upshot/)
- Visualoop (http://visualoop.com/)
- FiveThirtyEight (https://fivethirtyeight.com/datalab/our-33-weirdest-charts-from-2014/)
- Huffington Post (http://www.huffingtonpost.com/2014/12/22/huffpost-infographics-201_n_6351828.html)
- LA Times (http://graphics.latimes.com/2014-in-graphics/)
- Wall Street Journal (http://graphics.wsj.com/wsj-interactives-2014/)
- Washington Post (https://www.washingtonpost.com/graphics/national/2014-in-graphics/)
- Quartz (http://qz.com/318339/all-of-the-charts-we-made-in-2014/)
- New York Times - Interactive Storytelling (http://www.nytimes.com/interactive/2014/12/29/us/year-in-interactive-storytelling.html?_r=0#data-visualization)
- Fathom (http://fathom.info/)
- Data Canvas (http://map.datacanvas.org/#)
- Waze Global Driver Satisfaction Index (http://blog.waze.com/2015/09/global-driver-satisfaction-index.html)
- Lapham's Quaterly Maps (http://www.laphamsquarterly.org/archive/maps)
- Lapham's Quaterly Charts and Graphs (http://www.laphamsquarterly.org/archive/charts-graphs)
- Territory (http://themapisnot.com/)

- Quartz Atlas Charts (https://www.theatlas.com/)
- Sensory Maps (http://sensorymaps.com/)
- Library of Congress - Maps (https://www.loc.gov/maps/collections/)
- The National Geologic Map Database (http://ngmdb.usgs.gov/ngmdb/ngmdb_home.html)

Podcasts:

- Data Stories (http://datastori.es/)
- PolicyViz (http://policyviz.com/the-policyviz-podcast/)

Tools

- Text editors:
    - SublimeText (https://www.sublimetext.com/)
    - TextWrangler (http://www.barebones.com/products/TextWrangler/)
- Python IDEs:
    - SublimeText (https://www.sublimetext.com/)
    - Anaconda (https://www.continuum.io/downloads)
    - Canopy (https://www.enthought.com/products/canopy/)
    - Idle (https://en.wikipedia.org/wiki/IDLE_(Python)) - Not recommended but it's installed with ArcGIS on GSAPP computers.
- Raw (http://raw.densitydesign.org/)
- Color:
    - Colorgorical (http://vrl.cs.brown.edu/color)
    - ColorHexa (http://www.colorhexa.com/)
    - ColorBrewer (http://colorbrewer2.org/#type=sequential&scheme=BuGn&n=3) - Mostly for maps but it's a good resource.
    - Adobe Color CC (https://color.adobe.com)
    - i want hue (http://tools.medialab.sciences-po.fr/iwanthue/)
    - Color Picker for Data (http://tristen.ca/hcl-picker/#/hlc/6/1/15534C/E2E062)
- Visualization toolkits:
    - D3 (https://d3js.org/) - a JavaScript library for manipulating documents based on data.
    - Vega (https://vega.github.io/vega/) - a declarative format for creating, saving, and sharing interactive visualization designs.
    - Vega-lite (https://vega.github.io/vega-lite/) - is a high-level visualization grammar. It provides a concise JSON syntax for supporting rapid generation of visualizations to support analysis.
    - Processing (https://processing.org/) - a flexible software sketchbook and a language for learning how to code within the context of the visual arts.
    - p5.js (http://p5js.org/) - a JavaScript library that starts with the same goal as

Processing, to make coding accessible for artists, designers, educators, and beginners, and reinterprets it for today's web.
    - Protovis (https://mbostock.github.io/protovis/) - JavaScript visualization language, predecessor of d3.
    - Leaflet (http://leafletjs.com/) - an open-source JavaScript library for mobile-friendly interactive maps.
    - Visdown (http://visdown.amitkaps.com/) - visualization with markdown
    - g9.js (http://omrelli.ug/g9/) - automatic interactive graphs
    - Bamboo DiRT (http://dirtdirectory.org/) *Nearly comprehensive list of tools to use for DH projects*
- Text Analysis Tools
    - AntConc (http://www.laurenceanthony.net/software/antconc/) *Does basic text analysis NLTK-style. Great for non-Pythonic approach to distant reading a text. Best used with The Programming Historian's Tutorial (http://programminghistorian.org/lessons/corpus-analysis-with-antconc) .*
    - MALLET (http://mallet.cs.umass.edu/) *Topic modelling tool. Best used with The Programming Historian's Tutorial (http://programminghistorian.org/lessons/topic-modeling-and-mallet) .*
- Data Cleaning Tools
    - Google Open Refine (http://openrefine.org/) . The Programming Historian's Tutorial (http://programminghistorian.org/lessons/cleaning-data-with-openrefine) is very helpful here.

Tutorials & Resources

- Codecademy (multiple courses, HTML + CSS, Python, JavaScript, D3)
- Text Visualization Browser (http://textvis.lnu.se/)
- Github:
    - Introduction to Git (https://sklise.com/2012/09/22/introduction-to-git/)
    - Git workflow for beginners (https://sklise.com/2012/10/07/git-workflow-beginner/)
    - Try Git (https://try.github.io/levels/1/challenges/1)
    - A guide to using Github pages (https://www.thinkful.com/learn/a-guide-to-using-github-pages/)
- JavaScript:
    - How to learn JavaScript properly (http://javascriptissexy.com/how-to-learn-javascript-properly/)
    - JavaScript: the right way (http://jstherightway.org/)
    - Code School: JavaScript (https://www.codeschool.com/learn/javascript)
    - JavaScript garden (https://bonsaiden.github.io/JavaScript-Garden/)
    - Mozilla Developer Network: A re-introduction to JavaScript(JS tutorila) (https://developer.mozilla.org/en-US/docs/Web/JavaScript/A_re-introduction_to_JavaScript)
    - Codecademy: JavaScript (https://www.codecademy.com/lrn/javascript)

- Debugging:
  - Chrome Debugging Tutorial (https://developer.chrome.com/extensions/tut_debugging)
  - Firebug (http://www.developerfusion.com/article/139949/debugging-javascript-with-firebug/)
- Color:
  - How to Avoid Equidistant HSV Colors (http://vis4.net/blog/posts/avoid-equidistant-hsv-colors/)
  - Your Friendly Guide to Colors in Data Visualization (https://lisacharlotterost.github.io/2016/04/22/Colors-for-DataVis/)
  - How We Created Color Scales (https://datavisualization.ch/inside/how-we-created-color-scales/)
- InfoVis Group (UBC Computer Science) - Visualization Design Resources (http://www.cs.ubc.ca/group/infovis/resources.shtml)

Datasets:

- NYC Taxis (http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml)
- Citibike Ridership Data (https://www.citibikenyc.com/system-data)
- Citibike Station Data - GBFS (General Bikeshare Feed Specification) (https://gbfs.citibikenyc.com/gbfs/gbfs.json) - And the documentation (https://github.com/NABSA/gbfs/blob/master/gbfs.md)
- Citibike Station Feed - Legacy format (https://feeds.citibikenyc.com/stations/stations.json)
- Weather data - forecast.io API (https://developer.forecast.io/)
- Census data - American Fact Finder (http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml)
- Census data - citysdk (https://uscensusbureau.github.io/citysdk/)
- World Bank data catalog (http://datacatalog.worldbank.org/)
- AWS Public Datasets (https://aws.amazon.com/datasets/)
- Campaign Finance Data (http://www.fec.gov/finance/disclosure/ftpdet.shtml#a2015_2016)
- Enigma.io (http://enigma.io/)
- Dreamtolearn - 1001 Datasets and Data repositories (Lists of lists of lists) (https://dreamtolearn.com/ryan/1001_datasets)
- Data is Plural (https://tinyletter.com/data-is-plural/archive?page=1&recs=10&sort=desc&q=)
- Visualising Data - Data Sources (http://www.visualisingdata.com/references/)
- Project Gutenberg (http://www.gutenberg.org/)
- Association of Religion Data Archives (http://www.thearda.com/)
- National Archive of Data on Arts and Culture (http://www.icpsr.umich.edu/icpsrweb/NADAC/)
- NYPL Labs Menus (http://menus.nypl.org/data)
- NYPL Digital Collections (http://digitalcollections.nypl.org/)
- The Data Visualisation Catalogue (www.datavizcatalogue.com)