Interactive Data Visualization: Foundations, Techniques, and Applications

COMPSCI 590DV
Professor Georges Grinstein Office – LGRC 353A
Classroom – GSMN 64 Days/Time – Tuesday and Thursday 5:30-6:45
TAs: John Fallon, Steve Giguere , Arvand Neelakantan

Course Description: In this course, you will learn the fundamental algorithmic and design principles of visualizing and exploring complex data. The course will cover multiple aspects of data presentation including human perception and design theory; algorithms for exploring patterns in data such as topic modeling, clustering, and dimensionality reduction; as well as a wide range of statistical graphics and information visualization techniques. We will explore numerical data, relational data, temporal data, spatial data, graphs and text. Hands-on projects will be based on Python or JavaScript with D3. 3 credits.

Required Textbooks

- *Interactive data visualization: foundations, techniques, and applications.* Matt Ward, Georges Grinstein, Daniel Keim. 2nd Edition, CRC Press. 2014.

Recommended Textbooks

- *Visualization Analysis and Design.* Tamara Munzner, CTC Press. 2014.

Required Software

The assignments will be split between Python (including iPython Notebook, NumPy, Pandas, SciKitLearn, MatplotLib and Seaborn) for desktop data visualization and exploration applications, and JavaScript with D3 for web-based interactive data visualization and exploration. All required software is freely available.

Course Website

The course website will be hosted on Moodle at https://moodle.umass.edu/. The course website will host lecture notes, assignments, and pointers to readings and videos. We will use Piazza (https://piazza.com/) for a course discussion forum.

Accommodation Statement

The University of Massachusetts Amherst is committed to providing an equal educational opportunity for all students. If you have a documented physical, psychological, or learning disability on file with Disability Services (DS), you may be eligible for reasonable academic accommodations to help you succeed in this course. If you have a documented disability that requires an accommodation, please notify me within the first two weeks of the semester so that we may make appropriate arrangements.

Academic Honesty Statement

Since the integrity of the academic enterprise of any institution of higher education requires honesty in scholarship and research, academic honesty is required of all students at the University of Massachusetts Amherst. Academic dishonesty is prohibited in all programs of the University. Academic dishonesty includes but is not limited to: cheating, fabrication, plagiarism, and facilitating dishonesty. Appropriate sanctions may be imposed on any student who has committed an act of academic dishonesty. Instructors should take reasonable steps to address academic misconduct. Any person who has reason to believe that a student has committed academic dishonesty should bring such information to the attention of the appropriate course instructor as soon as possible. Instances of academic dishonesty not related to a specific course should be brought to the attention of the appropriate department Head or Chair. Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent (http://www.umass.edu/dean_students/codeofconduct/acadhonesty/)

Homework and Schedule

[15 points] Homework 1 – Simple Data Analytics

[25 points] Homework 2/3 – Interactive Web Page of Coordinated Visualizations

[15 points] Homework 4 – Design for Final Project

Grading Plan

Graded Homework Assignments	45%
Midterm Exam	20%
Interactive Visualization Web Page	10%
Final Project	25%

[25 points] Final Project (due week 14, presentation on Final Day)
Select your own data set. Finish the coding of your VA coordinated visualization system by supporting brushing, linking, and any other operators for enhancement, as well as any additional analytics you want. Prepare (1) an interactive final presentation demo, (2) a debrief report, and (3) github source code.

Week	Date	TOPIC	Readings		Goal	Suggested progress work	HW due date
			WGK	TM			
1	1/24	Introduction and examples; definitions of visualizations	Ch 1	TBD	Engage students		
	1/26	Visualizing data with simple visualizations	Ch 1		Know definitions of large number of classic visualizations and their "parameters"	Implement an interactive visualization in D3. Compare 2 different normalizations with your visualization.	
2	1/31	Data preprocessing	Ch 2		Deal with data (access to, missing or erroneous data)	Get an IDE up and running (can be a simple text editor). You will need Python and D3 as base tools. Lots of different ways to set this up. See the TAs for support if you need it. Run some statistics on HW 1 data sets	
	2/2	Data and Databases for visualization	Ch 2		Other data issues	What happens if data does not fit into memory?	HW 1
3	2/7	Perception & cognition	Ch 3		How we see/hear – the early systems	Get several other visualizations up in your environment. Explore color maps.	
	2/9	Perception & cognition	Ch 4		How we remember/understand – the advanced systems	Those interested can explore sound output.	
4	2/14	Interactions (not covered due to snow day)	Ch 11		Basic interactions and queries	Get simple interactions and start thinking about users wanting to change visualization parameters via menus or other mechanisms	
	2/16	Spatial Data (not covered due to questions on D3 and others)	Ch 5		3D physical/scientific data – HPC – first encounter with big (and massive) data	What happens if data is so large that transfer is inhibitive?	
5	2/21	Interactions with classic visualizations	Ch 6		The world – data sizes	Use a graphics rendering library 3D physical data? Test if it works with your visualization environment	
	2/23	Geospatial Data, Time Oriented Data	Ch 7		Time is special. Sensor and LIDAR examples – again large data	Get a map up and provide simple interactions (and ideally labeling)	
6	2/28	Interactions, Multivariate techniques 1	Ch 8		N-dimensional – classic data but just wide	Generate a plot of some time series data. Use some statistical approaches (windows, averaging, ….), regressions, …	

	3/2	Multivariate Techniques 2	Ch 8		N-dimensional – modern visualizations	Explore parallel coordinates, circular scatter plots (RadViz, Star Coordinates), heatmaps	HW2
7	3/7	Review/uncertainty visualizations			Review using displaying uncertainty as examples	Understand most visualizations and know how to read them and their limitations. Have tried Tableau public.	
	3/9	**In class midterm**	Ch 1-8				
	3/13 -17	Break					
8	3/21	Graphs and networks	Ch 9		Relational data – Example applications and visualizations. Issues with data structures. Big data as in social network, tweets.		
	3/23	Text and visualization (single document and corpora)	Ch 10		Unstructured data – example applications and issues. Value of metadata and ontologies. Bigger data.	Generate 2 graph visualizations, one a class node-link and the other a circular one (Circos-like)	HW3
9	3/28	Text and visualizations	Ch 10		More unstructured data examples	Use the graph visualization to generate text visualizations.	
	3/30	Dimensional reduction techniques			N-D to a reasonable number with pitfalls	Apply the graphvis to some social network data that is large. What needs to happen?	
10	4/4	Clustering and other analytics			Analytics and ML	Read tweets and visualize something	
	4/6	Animation and visualization and interaction	Ch11		Improving user experience	Analyze HW3 datasets.	
11	4/11	More interaction	Ch 12		Touch, AR, VR, complex gestures	Visualize the results of a sentiment computation.	HW4
	4/13	Coordinated visualization systems			Exploratory systems. Dashboards.	Understand design questions	
12	4/18	**Monday schedule**					
	4/20	Designing Visualization Systems	Ch 13, 15		Tasks, Goals. Examples of novel and new interaction techniques.	Understand advanced vs simple visualizations in terms of users	
13	4/25	Evaluation	Ch 14		Utility vs Usability debate. Need for measures for evaluation and theories.	Select a visualization based on the task.	
	4/27	Open topic/Research topics	Ch 16		Thesis topic examples	Give examples of weak visualizations and good user interfaces	
14	5/2	Last class – Research topics					
		Final Project					Final Project