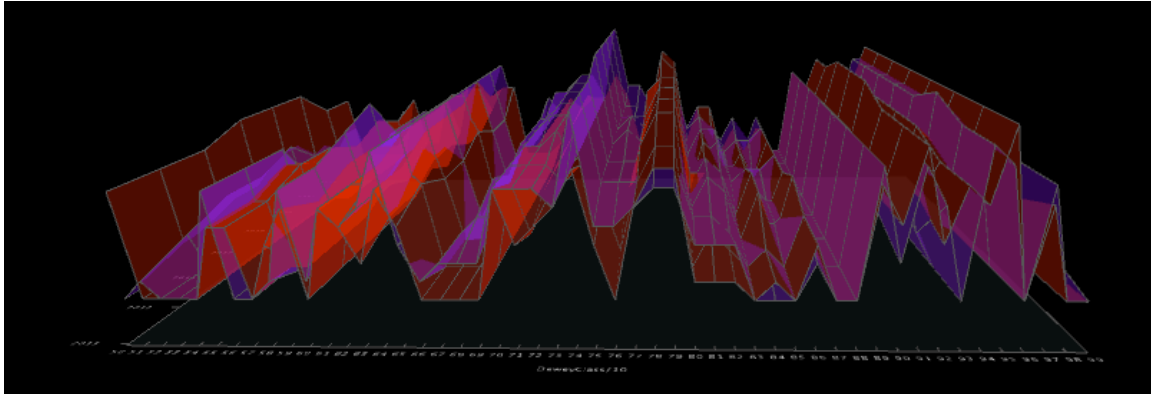


MAT259 Data Visualization

Winter 2015

Instructor: George Legrady

Elings Hall rm2611, Tues-Thurs 2:00-3:50



This ten week interdisciplinary computational course concentrates on techniques of 1) data mining, 2) data aggregation and 3) visualization in the java based Processing environment. Knowledge acquired include 1) how to identify and retrieve significant data from a dataset with MySQL 2) Develop skills in the fundamentals of visual language through programming 3) Visualize abstract data to reveal patterns and relationships 4) Normalize data to enhance legibility and coherence 5) Implement interactivity within 3D volumetric visualization.

Visual representation is a complex practice that involves knowing what data to extract from a dataset followed by the communication and creation of content through a skillful command of the use of form, space, balance, scale, dimensions, tone, color, texture, direction, motion, line, and text. The skills in knowing how to work with visual syntax is both a process which is partially rule-based, but then also perceptual, and dictated by conventions. In addition to data mining, the course introduces skills in the implementation of algorithms for data analysis; and acquisition of visual design methods to represent abstract data. Assignments include frequency mapping, 2D spatial and 3D interactive visualizations, and correlation with other data sources such as from Twitter, NY Times, Amazon, etc.

The course begins with MySQL exercises in data mining to develop skills in retrieving meaningful information from a unique multivariate resource consisting of over 80 million datasets generated from a public resource. The data has been retrieved hourly since September 2005. The dataset represents a wide breath of subject matter in books, cds, dvd's checked-out hourly by patrons of the Seattle Public Library. Each transaction contains scalar, numeric, time-based, semantic and other forms of multivariate metadata which allow for a multiplicity of subject matter. Every student in the course creates their own visualizations using the same dataset source to allow us to quickly master examples from prior student assignments and share solutions.

Resources: Data visualization is a broad discipline, and there are numerous online and published references. Books of special interest include "*The Semiology of Graphics*", Jacques Bertin, "*Graphics of Large Datasets*", by Unwin, Theus, Ben Fry's "*Visualizing Data*", O'Reilly, Katy Borner's *Atlas of Science: Visualizing What We Know*, MIT Press and others. You can review student results from previous courses at: <http://vislab.mat.ucsb.edu/courses.html>

Course Software: MySQL: <http://dev.mysql.com/doc/refman/5.6/en/index.html>, and the java-based Processing: www.processing.org. Knowledge of C or a java-based language recommended.