

CSI 703 / IT 875/ STAT 875
Scientific and Statistical Visualization
Syllabus: Spring 2008

Instructor: Professor Daniel Carr
Email: dcarr@gmu.edu
Room: ST2 R149
Phone: 703-993-1671
Office Hours: Tuesday 3:00-4:00 PM, briefly after class, and by appointment

Instructor: Assistant Professor Fernando Camelli
Email: fcamelli@gmu.edu
Room: Research 1 R343
Phone: 703-993-4073
Office Hours: Tuesday 3:00-4:00 PM and by appointment

Web Site: www.mason.gmu.edu/~dcarr/vis - with links to additional resources
www.mason.gmu.edu/~dcarr/vis/sched.htm - class schedule with topics, assigned reading, links to notes, and links to assignments.

Prerequisites: CSI 703: CSI 700, 701 or permission of Instructor
IT 875 / Stat 875: One of CS652, Stat 554, 662, 663, 751, 871

Required Texts: Ware, C. 2004. *Information Visualization, Perception for Design*, Kaufman Morgan Inc.

Suggested Wegman, E. and D. Carr. 1992. "Statistical Graphics and Visualization"
This technical report minus graphics is available as a .pdf file. (see web site)

Murrel, Paul. R Graphics 2006, Chapman and Hall/CRC

Chen, Chaomei. 2006. *Information Visualization, Beyond the Horizon*. Springer

Venables, V. N. and B. D. Ripley. 2002. *Modern Applied Statistics with S*,
Strongly Recommend: Fourth Edition, Springer

Cleveland W. S. 1994 *The Elements of Graphing Data*. Hobart Press

Software: **R**: Google CRAN, the comprehensive R Archive Network, to access binaries.
Binaries for R are available for the Windows, Linux and MacOS operating systems.

The basic syntax for this shareware is the same as the commercial package called

SPlus that was used for this class over many years. With recent efforts to increase compatibility, those working for federal agencies and businesses that require use of commercial software and can make an easy transition to SPlus.

Researchers around the world contribute packages to extend R. Several packages will be use in this class. The rapid advance of software tools is exciting.

Students are expected to learn enough R syntax and commands to modify scripts, to produce quality graphics, and to answer mid-term questions.

Other software: Class will have some demonstrations and assignments that will involve Windows oriented shareware. Links to the software will appear in the class schedule. Examples:

CCmaps: Java-based

CrystalVision, GLISTEN: C++ and OpenGL based

Class assignments will likely involve some C++ and OpenGL. This is also based on what is taught in class.

Grades: 5% Participation: very short in-class quizzes
 20% Homework: due 2 weeks from date given.
 20 % Midterm exam
 20 % Redesign: paper and presentation of selected papers
 35% Final project: paper and presentation

Instructors reserve the right to drop or give fewer quizzes and increase the final project percent correspondingly.

Instructors reserve the right to wave the late homework penalty for individuals with special circumstances or extend the due data of a long assignment for the whole class. Homework less than week late is penalized 20%.

Honor Code Students are encouraged to discuss homework if they have questions or ideas. However, it is an honor code violation to submit homework produced by someone else. Students are expect to modify scripts, run the scripts, write answers and produce the homework themselves.

In general it is an honors code violation to claim explicitly or implicitly claim credit for the work of others. Provide citations for material coming from the web and other sources.

Students work independently on the midterm and must follow the directions given concerning notes and access to external information.

Documents are available on the class web site concerning expectations for

redesign and final projects. The redesign should be primarily individual work but student can seek help in areas such as software syntax and proofreading. Student are also free to obtain guidance from the instructors.

The final project is often a team project. Some differences in contribution levels are to be expected. However, when there is good evidence of substantially different levels of contribution, the instructors reserve the right to grade team members individually and to avoid the problem of one or more members taking undo credit.

Submitting the work of others outside the team without appropriate citation is an honors code violation and puts the whole team at risk. If this happens, efforts will be made to restrict the violation to the appropriate individuals. However, it is wise for the team members to check each other's claimed contributions and to avoid the problem.

Teams are encouraged to obtain guidance from the instructors and others with expertise. Acknowledgement is appropriate if the contribution is nontrivial.

Course Topic Discussion

This course brings together material from many disciplines to provide an overview of scientific and statistical visualization. The scope is interpreted broadly to include contributions from the fields of information visualization and the emerging field of visual analytics.

The goals of visualization include data description and analysis, discovery, hypothesis generation, analysis, understanding, presentation and education. In this class the active agent is the analyst not passive user. A graphics design challenge is to engage the analyst. Toward this end the class begins by stressing human cognition and perception as a foundation for graphics that are useful for analysts. This is followed by topics from several disciplines. Topics from computer graphics cover some of the basics of image transformations and rendering. Topics from statistics include graphical methods for describing and comparing distributions, for modeling relationships and for assessing quantitative models. Topics from geography, cartography and earth systems address visualization in a geospatial and temporal contexts. The course will touch on a data mining, a domain with contributions coming from several disciplines.

The class stresses the importance of data (observational data or/and simulation data) and data models in driving the graphics. Data comes from many domains. The class will feature a wide range of the example produced by the instructors on such topics such gene expression, peptide binding on human immune system molecules, cancer surveillance, and views of atmospheric data from JPL. A whole course could be devoted visualization in one individual domain such as satellite imagery, medical imaging, genomics, proteomics, computation fluid dynamics, social networks, simulation, and so on. However there is great merit is seeing what is done across

disciplines. Student redesigns and final project presentations and sometimes guest lectures help to provide coverage of different domains. The topics evolve over time depending class composition and advances in visualization and visual analytics. The schedule on the web site indicates topics for a particular semester that generally include

Topics Typically Addressed

- Psychology: human perception and cognition
 - Psychophysics Laws
 - The eye and visual channels
 - Three forms of blindness
 - Rapid discrimination
 - Perceptual grouping
 - Color
 - Perceptual accuracy of extraction
 - Memory constraints
 - Motion
- Computer Graphics
 - 2-D and 3-D viewing transformations
 - Rotation and stereo
 - Rendering
 - Alphablending
 - Graphics card capabilities
- Statistical Graphics
 - Types of data
 - Distribution representation and comparison
 - Functional relationships
 - Transformations
 - Encodings
 - Layouts and sorting
 - Dynamic Methods
 - Brushing and linking
 - Filtering
 - Grand tour and projection pursuit
- Large data set methods
 - Case reduction
 - Dimension reduction
 - Complexity reduction
 - Overviews
- Regression and diagnostics
- Clustering and classification
- Supervised classification
- Representing uncertainty
- Conditioned views
- Linked views
- Visualization for
 - Fluid flow
 - Medical Imagery
 - Satellite Imagery
 - Proteomics and genomics
 - Graphs and networks
- Cartography
 - Map projection
 - Symbolization
- Computer Human Interface
 - Usability Assessment
 - Accessibility
 - Navigation
- Software and web resources
- Web graphics
- Education – Multimedia Learning
 - Diagrams and spoken text

The class strives to encourage creativity. Student redesigns improve poor graphics and most projects study new data sets. Some student projects develop new tools. Hopeful all students leave the class seeing quantitative graphics in a new light.