

2012
01/10

CATEGORY

Uncategorized

[Write comment](#)

About the Course

CS4460 Information Visualization

Fall 2012

Instructor: Prof. James Foley

GTA: Megha Sandesh


Information visualization goes beyond presenting information as static charts, graphs and maps by leveraging the power of computer interaction to help people analyze, understand and make decisions from data. Dozens of companies – including Google, Microsoft, IBM, Oracle and SAP – offer Information Visualization tools. Thousands of companies and governments use the tools for daily operations and for longer-term strategic planning.

Information visualization methods are applied to data from many different application domains, including:

- Political reporting and forecasting – as seen on TV and in the papers in election season.
- News reporting – look at the interactive visualizations used by the New York Times, Wall Street Journal, Slate, etc.
- Social science and economics data, such as census and other surveys, and micro and

INFORMATION

Change this sentence and title from admin Theme option page.

 [RSS FEED](#)



PAGES

[About](#)

[Homework Assignments](#)

[Paper Presentation](#)

[Project](#)

[Resources](#)

[Schedule](#)

META

[Register](#)

[Log in](#)

macro economic trends.

- Social networking and web traffic, to understand patterns of communication
- Business intelligence and business dashboards – to forecast sales trends, understand competitive marketplace positions, allocate resources, manage production and logistics.
- Text analysis – to determine trends and relationships for literary analysis and for information retrieval.
- Criminal investigations – to portray the relationships between event, people, places and things.
- Performance analysis of computer networks and systems.
- Software engineering – developing, debugging and maintaining software.
- Bioinformatics, to understand DNA, gene expressions, systems biology.

Course objectives are:

- Learn the principles involved in designing effective information visualizations.
- Understand the wide variety of information visualizations and know what visualizations are appropriate for various types of data and for different goals.
- Understand how to design and implement information visualizations.
- Know information visualizations use dynamic interaction methods to help users understand data.
- Learn to apply an understanding of human perceptual and cognitive capabilities to the design of information visualizations.
- Develop skills in critiquing different visualization techniques in the context of user goals and objectives.
- Learn how to use and critique existing systems for creating information visualizations.

The course will follow a lecture/seminar style with discussions, demonstrations of InfoVis software, viewing of videos and hands-on experience with research and commercial information visualization software.

[Entries RSS](#)

[Comments RSS](#)

[WordPress.com](#)

Text: *Now You See It*, Stephen Few, Analytics Press, 2009. I suggest you team up with another student to buy and share. It is also on reserve in Library.

Optional texts:

For those interested in design: Any of Edward Tufte's three books: *The Visual Display of Quantitative Information*; *Envisioning Information*; and *Visual Explanations*.

For those interested in business intelligence and business dashboards: Wayne Eckerson, *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*, Wiley, 2005, ISBN 978-0471724179

For those interested in Network Visualization, particularly Social Networks: Hansen, Shneiderman and Smith, *Analyzing Social Media Networks with NodeXL*, Morgan Kaufman, 2011, ISBN 978-0-12-382229-1.

For those interested in the psychological/perceptual factors affecting information visualization: Colin Ware, *Information Visualization: Perception for Design*, 2nd Edition, Morgan Kaufman Elsevier 2004, ISBN 978-1558608191.

Students from a variety of disciplines are invited to take the class. Programming experience will be useful but is not required – while each of you will work on a project that will produce an information visualization program, there are many tasks involved in the process, not all of which require programming.

Grading:

Grading will be based on class participation, a class presentation, homework, use and analysis of some information visualization tools, and a project. There will be two tests during the semester, and a final project. Grading weights are:

Test 1 18 points

Test 2 18 points

Homeworks 29 points

Project 35 points

See the project and homework tabs for details.

Collaboration and Academic Honesty

Unless explicitly stated otherwise, you are expected to do your home works on your own. Your project work may borrow libraries and code fragments from sources on the web that you integrate into an overall working system. Your source code should indicate what code is imported and used as is, what code is imported and modified, and what code is original. It is appropriate to discuss your project with others to gain ideas and feedback and help with sticky problems. It is not appropriate to find an infoviz system, modify it and submit it as your own work. If in doubt, confer with your instructor.

Software

One of the assignments is to analyze data using Tableau, Spotfire, or NodeXL. Tableau's [data visualization software](#) is provided through the Tableau for Teaching program.

Share this:



Twitter



Facebook

Like this:

Like

Be the first to like this.



Write comment



Comments RSS

Trackback (0)

Comments (1)



[Mr WordPress](#)

January 10th, 2012

► [REPLY](#) | [QUOTE](#)

Hi, this is a comment.

To delete a comment, just log in, and view the posts' comments, there you will have the option to edit or delete them.

Leave a Reply



Email **(required)**

(Address never made public)

Name **(required)**

Website

☐ Notify me of follow-up comments via email.

Post Comment

Project

You will form four-person project teams (three people in a few cases). I will facilitate some in-class discussions about project groupings, but you should explore ideas amongst yourselves as well. I want the teams to be balanced in terms of background and experience – some from computational media and the people thread, some from the information and other threads, and from outside CoC.

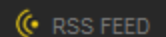
The idea of the project is to take the InfoVis knowledge that you are acquiring this semester and use in a new, creative effort. A real key to a successful project is to select a topic/data that people want to know more about, and that is of interest to your team.

There are many sources of data on the web – financial, social/demographic, economic. Census Bureau, CDC, CIA, WHO, UN, NBA, FAA, Many Eyes. This semester I am encouraging focus on health-related data, and CDC has a LOT of data. **It is important that you choose data that is already in a usable form**, such as CSV, that is easily imported into your program. Else you will spend time on data formatting etc. rather than on information visualization!!!

No matter what topic/data you choose, I am expecting a high-quality project. In particular, I'm seeking creative projects showcasing interesting ideas. A stellar project consists of an implementation of a (potentially new) novel visualization design with multiple views of the data and good interaction, designed to allow users to answer interesting questions and gain

INFORMATION

Change this sentence and title from admin Theme option page.



PAGES

[About](#)[Homework Assignments](#)[Paper Presentation](#)[Project](#)[Resources](#)[Schedule](#)

META

[Register](#)[Log in](#)

insights. But note that I am explicitly NOT expecting user testing and evaluation.

You are free to choose any software development environment and graphics/visualization support library that you want. There are many. That said, if you have no previous experience with an InfoVis toolkit, then I recommend that you use d3 (<http://mbostock.github.com/d3/>). Your GTA will be knowledgeable in its use, and with a lot of you using it, helping one another will be easier. Even if you don't use d3, you should use a tool that is web-deployable so that your system can be shown to everyone in the world!

If you slack off in sharing the workload of the project, you will receive a lower grade than your partners. Because the project counts 35%, slacking off is not a good idea.

There are four project milestones. First, you must form your team and settle on a topic. Second, you will submit a detailed design summary about halfway through the term. Third, you will make a final presentation about your project during the last week of class. Finally, the software system you build will be delivered, along with a short write-up describing the problem and solution.

Milestones (see schedule for due dates):

- Milestone 1 – Initial project description. One page document listing project members and topic/data to be addressed. Also who is filling the following team roles: 1) Team Leader; 2) Keeper of “to do” list; 3) Convener of meetings; 4) Mid-way progress report integrator (everyone contributes, one person integrates); 5) PPT presentation integrator (everyone contributes, one person integrates); 6) Final report integrator (everyone contributes, one person integrates). Hard copy submission in class.
- Milestone 2 – an elaboration on your initial project description, to include 3 or 4 questions that users of your system should be able to investigate/answer; your data source and format of the data (spread sheet, SQL data base, ...). Hard copy submission in class.
- Milestone 3 – Paper sketches of the information visualizations you expect to provide and listing of the interaction methods you expect to use. What software you will be using, discussion of how the data will be organized and accessed, and a description of how your team is dividing up responsibilities for carrying out the project. This should not be more than 5 pages, but MUST include at least two screen sketches. Hard copy submission in

Entries RSS

Comments RSS

WordPress.com

class.

- Milestone 4 – Show us some code and screen shots. By now you should have some part of your project coded. Submit hard-copy of your code, and two or three screen shots. Your code will probably include some that you have picked up from other sources along with some that you have written. Highlight in some way the code you have written. No need to include libraries etc. We just want to see that your team is underway. Hard copy submission in class.
- Milestone 5 – demonstrate (partially) working system to us. We'll provide feedback to help you improve your system prior to your project presentation and final report.
- Milestone 6 – Project presentations. Each team will have approximately 20 minutes to present. The presentation is partially PPT, partially live demo. The PPT should cover: Goals of project; Examples of questions your users should be able to answer by using your project; Source of data; IDE you used (if any), and your evaluation of how useful it is; InfoVis toolkit you used, and your evaluation of how useful it is; Who did what, and what you learned as a group about project team work; Software architecture diagram. This will generally be one slide per topic. For the demo, quickly show the overall set of capabilities – types of interaction, multiple views, various types of data. Then go through 4 or 5 scenarios to answer questions of the type described in point 2 of the PPT. Hard copy submission of PPT in class.
- Milestone 7 – Final executable software is due along with a final report describing the system. Submission mechanism TBA.

The Milestone 7 final report should include:

1. Goals of project, as in the midterm progress report but possibly revised (if so, indicate revisions).
2. Examples of questions your users should be able to answer by using your project.
3. Source of data.
4. Four to six screen shots with captions explaining what is going on.
5. IDE you used, and your evaluation of how useful it is.
6. InfoVis toolkit you used, and your evaluation of how useful it is.

7. Who did what, and what you learned as a group about project teamwork.
8. Software architecture diagram.
9. For each code module in the architecture diagram, indicate how much (%%) you wrote vs. how much you were able to import.

Parts 1-3 should not be more than 3 pages total. For part 4, do one screen shot per page.

Parts 5-9 are about a page each. Part 10 is an appendix to the report, submitted as a s

For this Milestone, turn in a single zipped file that contains:

- Milestone 6 PPT project presentation (may be revised if you presented prior to this milestone).
- Milestone 7 final pdf report
- Commented source code, including comments or highlighting to indicate what code was imported and used as is, what code was imported and modified, what code was written mostly or completely from scratch. This should be in a pdf file so it can be opened without an IDE.
- All necessary source code, including libraries, in whatever file format your IDE uses such that the code can be loaded into that IDE.
- A program executable.

Grading: We will evaluate the overall quality of your project, including all milestones and components. A grading sheet is at the end of this syllabus.

Great projects will typically have 1000+ lines of code. Some of the code will be new, some heavily-modified, some from libraries on the web. I would expect that several hundreds of lines of code are new. Functionality of great projects typically includes two or more simultaneous views of data with linking between the views, details on demand, and some means of selecting which aspects of the data set are displayed (maybe dynamic queries). Whatever interactions you provide should support users in answering questions about the data!

Team Project Grading Criteria
CS4460 Information Visualization

Project is 35% of your course grade
Grading on a 35-point scale

Project Name: Grade ____/35

Project Team:

____/1 Was the mid-way progress report submitted and provide the requested information?

____/3 Does the system work, i.e, does it read in the data and present a usable interactive visualization of the data?

____/3 Is the visualization an effective representation of the data such that it can support answering your analytical questions about the data?

____/9 Overall scope/ambition of the working project

- ____ Multiple views of data
- ____ Quality and number of interactions (such as details on demand, linking/brushing, zoom/pan, animation, semantic zoom, fish-eye views, dynamic queries, etc.). I don't expect you to do all of these, but would expect at least one or two others.

____/9 Did your team do more than copy and paste imported code to complete the project?

You are encouraged to copy and paste some code and to use libraries, but I expect that your team to be "adding value" by writing and substantially modifying a non-trivial amount of code on your own, to gain experience. Projects that have an appropriate scope/ambition typically involve well over a thousand lines of code.

____/5 Was your presentation an effective discussion and portrayal of the project?

____/1 Include mention/critique of toolkit and IDE used?

____/1 Include list of analytical questions you expect users to answer

____/3 Live demonstration shows how some or all of the questions can be answered

____/5 Does your report follow the guidelines and help the reader understand your overall goals and how your system achieves those goals:

- ____/0.1 Goals of project, as in the midterm progress report but possibly revised (if so, indicate revisions in italics).
- ____/0.7 Examples of questions your users should be able to answer by using your project, as in the midterm progress report but possibly revised (if so, indicate revisions with italics).
- ____/0.1 Source of the data.
- ____/0.5 Four to six screen shots with captions explaining what is going on.
- ____/0.1 IDE you used, and your evaluation of how useful it is.
- ____/0.3 InfoVis toolkit you used, and your evaluation of how useful it is.
- ____/0.2 Who did what, and what you learned as a group about project team work.
- ____/0.5 Software architecture diagram.
- ____/0.5 For each code module in the architecture diagram, indicate how much (%%) you wrote vs. how much you were able to import.
- ____/2.0 Commented source code, including comments that indicate what code was imported and used as is, what code was imported and modified, what code was written mostly or completely from scratch.

CS 4460 Project Team Self and Group Evaluation

Each project member completes and emails to foley@cc.gatech.edu

Your name:

Project name:

The goal is to understand your role in the group project, and the role of others. For each question, your self-evaluation is first, then your evaluation of each of your partners. Enter your partners' names in the places provided.

For each question, after each name circle whichever number applies, using this scale:

1 – way below average

2 – below average

3- average

4 – above average

5- way above average

Effort: How would you rate the relative effort (that is, time commitment) of each team member:

My effort 1 2 3 4 5

<<Name 1>> 1 2 3 4 5

<<Name 2>> 1 2 3 4 5

<<Name 3>> 1 2 3 4 5

Communications: How would you rate the communications of each team member
(responding/initiating email, coming to meetings, contributing to meetings)?

My effort 1 2 3 4 5

<<Name 1>> 1 2 3 4 5

<<Name 2>> 1 2 3 4 5

<<Name 3>> 1 2 3 4 5

Flexibility/teamwork: How would you rate each team member on their willingness to pitch in
and do whatever needed to be done (as opposed to "I can only do this")?

My effort 1 2 3 4 5

<<Name 1>> 1 2 3 4 5

<<Name 2>> 1 2 3 4 5

<<Name 3>> 1 2 3 4 5

Specific tasks you carried out on the project:

Other comments:

Share this:



Twitter



Facebook

Like this:

Like

Be the first to like this.

 [Write comment](#)

 [Comments RSS](#)

[Trackback \(0 \)](#)

[Comments \(0 \)](#)

Leave a Reply



Email **(required)**

(Address never made public)

Name **(required)**

Website

☐ Notify me of follow-up comments via email.

Post Comment

Homework Assignments

[Homework 0](#) Wiki page – August 28 (-1 if not completed)

[Homework 1](#) Survey - August 23 (-1 if not completed)

[Homework 2](#) Viz Examples - August 23 (2 pts)

[Homework 3](#) Data Source – August 30 (1 pts)

[Homework 4](#) Processing – Sept 6 (5 pts)

[Homework 5](#) d3 – Sept 18 (5 pt)

[Homework 6](#) Test 1 Question – October 2 (2 pts)

[Homework 7](#) Touchgraph – October 9 (5 pts)

[Homework 8](#) Tableau – October 25 (7 pts)


[Homework 9](#) Test 2 Question – November 13 (2 pts)

Homework 0 – Wiki Page

Put up on the class T-Square Wiki

INFORMATION

Change this sentence and title from admin Theme option page.

 RSS FEED



PAGES

[About](#)

[Homework Assignments](#)

[Paper Presentation](#)

[Project](#)

[Resources](#)

[Schedule](#)

META

[Register](#)

[Log in](#)

1. A paragraph about yourself, containing at least:

- a) Home town / state / country
- b) Undergrad major / school
- c) GT major
- d) Work experience – both general and UI-related
- e) Programming languages / tools
- f) Design skills / tools
- g) Something most people don't know about you

2. A picture of yourself

Also bring a hard copy to class on the due date.

[Top](#)

Homework 1 – Survey

Complete the programming background survey by accessing the URL below:

[Infoviz Preliminary Student Survey](#)

Homework 2 – Two interesting visualizations

Find two visualizations of data from a domain that is of interest to you (health, politics, finance, economics, world demographics, sports, etc etc etc). They can be information presentations (static) or information visualizations (interactive). For each visualization, do a one page write-up. Each one-pager should have the visualization (about half the page) and then:

- 1. What message(s) the visualization is intend to convey.
- 2. A critique of the visualization – both pros and cons – ways in which the visualization does a

[Entries RSS](#)

[Comments RSS](#)

[WordPress.com](#)

good job, ways in which it could be improved.

Turn in at the start of class. I will ask some of you to show your visualization and share your pros and cons with the class.

[Top](#)

Homework 3 – Find a Data Source

For one of the two domains of interest from HW 2, locate a data source on the WWW from which that or similar could be obtained. Submit a URL and description of the data – how it is organized, names of data fields.

[Top](#)

Homework 4 – Programming with Processing

To be written.

You are expected to do this homework as a two-person team. You are encouraged to seek help as needed, but the actual work you submit is to be your team's. How to submit HW?

Homework 5 – Programming with d3

To be written

You are expected to do this homework as a two-person team. You are encouraged to seek help as needed, but the actual work you submit is to be your team's. How to submit HW?

[Top](#)

Homework 6

Write a short essay question for the upcoming test AND your answer to the question. No T/F or multiple choice questions! Bring to class on the due date, turn in at the beginning of class. I

will share selected questions with the entire class as part of our review.

[Top](#)

Homework 7

Use and Critique TouchGraph

Sign up for a 30-day trial of TouchGraph at <http://www.touchgraph.com/navigator/trial>. Use it with the Google Shopping and Senate files that come with the trial.

1. Write a one paragraph “elevator speech” that you would use to explain to a non-CS student what TouchGraph does.
2. Evaluating TouchGraph from an InfoVis perspective, list all the things that TouchGraph does well – that is, features that it is good to have in an InfoVis program. Example: details on demand.
3. Evaluating TouchGraph from an InfoVis perspective, list all the things that TouchGraph does poorly – that is, features that it is good to have in an InfoVis program but that are not in TouchGraph. Example: edges are drawn as straight lines, so they go thru nodes.

You want to be really exhaustive with your pro and con lists; I have a list of 18 pros and 11 cons. Work with all the commands/actions and critique as a way to build up your lists.

[Top](#)

Homework 8

Use and Critique a full-featured Information Visualization System

This assignment will familiarize you with a full-featured InfoVis system. You may choose one of three systems with which to work:

Tableau

Spotfire

NodeXL

These are all Windows programs, so we Mac users have to adjust ☹. Tableau and Spotfire are commercial systems that can be downloaded at no charge for classroom use. NodeXL is from a non-profit. Details on accessing each are at the end of this document. The three programs will be demonstrated in class. The assignment is due after the in-class demonstrations.

If you would like to earn extra credit and work a bit ahead in class by demonstrating one of the programs in class, let me know quickly.

The goals of the assignment are for you to learn the capabilities provided by one of the systems, learn the basic visualization methods that they provide, and assess their utility in analyzing data. Each program comes with some data sets that you can use.

You can write the report on this homework by yourself, or you can do it with a partner (which I encourage, it will be more fun and you will learn more). If you write with a partner, you will both receive the same grade. You may ask others for help with downloading and figuring out how to use whichever program you decide to use. The paper and its ideas should be developed by you or by your two-person team.

The assignment has four parts:

1. Gain familiarity with the systems – Familiarize yourself with the visualization techniques and the user interfaces of the three the systems by watching the on-line video tutorials, which are available at:

Tableau: [http: Find instructions in the Wiki for CS 4460 under T-Square](http://support.tableau.com/articles/4460)

Spotfire: <http://support.spotfire.com/training/jumpStarts/ProfessionalJumpStart/player.html>

NodeXL: <http://www.connectedaction.net/2009/11/11/video-using-nodexl-to-map-the-digg-mentioning-twitter-population/>

You don't need to watch all of the videos, just enough to get a flavor of what the systems do, so you can decide which system you will use for the rest of the assignment.

2. Examine the data sets – Browse several data sets to decide which one to use for the rest of this assignment. Decide on one, and then use the systems to explore it further.

3. Develop three interesting questions about the data set – Put yourself in the shoes of a data analyst, and think about all the different kinds of analysis tasks that a person might want to perform with your chosen data set. For instance, someone working with breakfast cereal data might have analysis tasks like:

- Find all the information on Cocoa Pebbles.
- Identify the cereal with the least fat that is also high in fibre.
- What is the distribution of carbohydrates in the cereals?
- Does high fat mean high calories?
- Which of the following three cereals is best for people on a diet?

Do NOT make all of your questions be about correlations or min or max values.

4. Write a report – Part 1 – List your three questions and answers, along with a screen shot showing the visualization you used to answer each question. Try to do this with one page per question. Part 2 – Critique the system. What are the system's strengths and weaknesses? For what kinds of user tasks is the system particularly well suited? Focus more here on the visualization techniques as opposed to the particular user interface quirks, though you should feel free to comment on UI aspects when they are particularly good or bad. As with HW 4, describe characteristics of the UI using the concepts and terminology you have learned in class. This second part should be 2 to 3 pages.

Your document should be in pdf format and is limited to a maximum of 10 pages including cover sheet. Use Times Roman 12 point type with normal margins, 1.5 line spacing. Please bring a stapled hardcopy to class on the due date, AND submit the paper with T-Square.

Tableau Instructions

Go to the Wiki in T-Square. You will find a new page named Assignment 7. Open it and you will find instructions on how to download Tableau.

<https://t-square.gatech.edu/portal/site/27388.201202/page/e63834f1-2338-4701-9747-6ad9523bbf41>

Spotfire Instructions

Go to <http://registration.spotfire.com/eval/academic.asp?source=stephanie>

Fill in the form. You MUST use your GT email address for your request for a download to be accepted.

Training videos are at

<http://support.spotfire.com/training/jumpStarts/ProfessionalJumpStart/player.html>

NodeXL Instructions

NodeXL is the free and open add-in for Excel that supports network overview, discovery and exploration. The code and application can be found at <http://www.codeplex.com/nodexl>.

Technical questions can be asked on our discussion boards on the Codeplex site at <http://nodexl.codeplex.com/Thread/List.aspx>

A collection of social media network maps created with NodeXL can be found at:

http://www.flickr.com/photos/marc_smith/sets/72157622437066929/

General NodeXL news can be found at: <http://www.connectedaction.net/> and a recent video and slide deck describing the application of NodeXL to generate social media maps can be found at: <http://www.connectedaction.net/2010/06/04/june-3-and-4-2010-personal-democracy-forum-2010-nyc/>

A video tutorial for NodeXL can be found at: <http://www.connectedaction.net/2009/11/11/video->

[using-nodexl-to-map-the-digg-mentioning-twitter-population/](#)

A manuscript tutorial guide to NodeXL created at the University of Maryland, College of Information Studies can be found at

http://casci.umd.edu/images/4/46/NodeXL_tutorial_draft.pdf

A book Analyzing Social Media Networks with NodeXL: Insights from a connected world is available from Morgan-Kaufmann.

Supporting data sets can be found at http://casci.umd.edu/NodeXL_Teaching.

Recent slide decks describing NodeXL can be found at:

http://www.slideshare.net/Marc_A_Smith/2009-december-nodexl-overview and

http://www.slideshare.net/Marc_A_Smith/2010-june-personal-democracy-forum-marc-smith-mapping-political-social-media-crowds .

Video from the PDF2010 presentation is available at:

<http://www.connectedaction.net/2010/06/04/june-3-and-4-2010-personal-democracy-forum-2010-nyc/>

NodeXL allows for the import of network data in the form of edge lists, matrices, graphML, UCInet, and Pajek files along with CSV and other workbooks.

In partnership with the Uberlink corporation (<http://www.uberlink.com.au/>), the VOSON data collector component has recently been integrated into NodeXL to enable web hyperlink network extraction.

NodeXL requires Office 2007. Other versions of Excel (like 2008 on Mac, or the older 2003) do not work with NodeXL (sorry!). NodeXL works with the new Office 2010 version of Excel.

NodeXL is a project from the Social Media Research Foundation

(<http://www.smrfoundation.org/>) and receives generous support from the Microsoft Research External Projects Group.

[Top](#)

Homework 9

Write a short essay question for the upcoming test AND your answer to the question. No T/F or multiple choice questions! Bring to class on the due date, turn in at the beginning of class. I will share selected questions with the entire class as part of our review.

[Top](#)

Share this:



Twitter



Facebook

Like this:

Like

Be the first to like this.



Write comment



Comments RSS

Trackback (0)

Comments (0)

Leave a Reply



Email (required)

(Address never made public)

Name (required)

Website

☐ Notify me of follow-up comments via email.

Post Comment

[Blog at WordPress.com.](#) | Theme: [monochrome](#) by [mono-lab](#).

▲ RETURN TOP