| 36-721 | Fall 2015 |
| Statistical Graphics and Visualization | 6.0 units |

Course Syllabus

Instructor: Jerzy Wieczorek
jerzy@cmu.edu
http://www.stat.cmu.edu/~jwieczor/
Office hrs: Weds 11:30am-1:30pm, BH 132M

TA: Jining Qin
jiningq@andrew.cmu.edu

Office hrs: Fri 1-2pm, BH 132M

Class meetings: Tues & Thurs, 3:00-4:20pm, PH A18C

Website: https://blackboard.andrew.cmu.edu/

Exam dates: None, but last assignments must be submitted by 5:00pm on Sat Oct 24

Prerequisites: None

Course objectives

An effective graphic is a powerful tool for analyzing data and communicating insights. By tapping into the human brain's efficient visual processing centers, a good statistical graphic can quickly provide a rich understanding of the data. Upon completing this course, you should be able to:

1. Explore raw data visually and assess statistical models' fit using graphical diagnostics

2. Critique and redesign statistical graphics based on the principles below

3. Produce legible, self-contained, informative graphics using statistical software

4. Plan effective statistical graphics using the principles of human visual perception

5. Model statistical graphics according to the Grammar of Graphics principles

6. Design multi-chart static works (conference posters, infographics) using the principles of graphic design

7. Generate interactive data visualizations following the principles of interaction design

8. Synthesize the data visualization research literature to justify recommendations for graphical practice

Assessment philosophy

This course will use a competency-based grading system that is designed to give you **control** over your final grade and **transparency** about your progress. You will know exactly what to do to earn a given grade, which will directly reflect your level of mastery of the learning objectives. We will provide **feedback** on each submission, and you will be able to **revise and resubmit** if you are not happy with your initial performance.

Official course listing

"Graphical displays of quantitative information take on many forms to help us understand both data and models. This course will serve to introduce the student to the most common forms of graphical displays and their uses and misuses. Students will learn both how to create these displays and how to understand them. The class will also cover some principles of visual perception and estimation. We will start with univariate and bivariate data, looking at some commonly used graphs and, after discussing their advantages/disadvantages, then turning to more sophisticated tools. We will then explore some three-dimensional tools, group structure/clustering, and projections of higher dimensional data. As time permits, the course will consider some more advanced graphical models such as statistical maps, networks, and the usage of icons."

Textbooks and software

Required textbooks:

- *The Functional Art* by Alberto Cairo, available at campus bookstore

- *Fundamental Statistical Concepts in Presenting Data* by Rafe Donahue, `http://biostat.mc.vanderbilt.edu/wiki/Main/RafeDonahue`

Recommended textbooks: Not required, but free via CMU library website:

- *The Grammar of Graphics*, 2nd edition, by Leland Wilkinson, `http://link.springer.com/`

- *ggplot2* by Hadley Wickham, `http://link.springer.com/`

- *Interactive and Dynamic Graphics for Data Analysis with R and GGobi* by Dianne Cook and Deborah Swayne, `http://link.springer.com/`

- *Interactive Data Visualization for the Web* by Scott Murray, `http://chimera.labs.oreilly.com/books/1230000000345/`

Software: MSP students in the Statistics department are required to use R for most assignments, in order to prepare you for other courses in the program.

Other students may turn in assignments using any statistical-graphics or scientific-computing software, such as Python, MATLAB, etc. However, the instructor and TA are most familiar and best able to help you with R.

We will demonstrate in-class examples using some of the following software:

- R and RStudio, free at
  `http://cran.r-project.org/` and `http://www.rstudio.com/`

- D3.js, free at `http://d3js.org/`

- Inkscape, free at `https://inkscape.org/`

- Tableau, free 1-year student license at `http://www.tableau.com/academic/students`

but again, you may use other tools as appropriate.

Assignments

You will complete up to 7 assignments: three short, focused **homeworks**; three self-directed **projects**; and one **critique** of a classmate's project draft. Your grade will depend on which assignments you complete and how well, according to the table below.

Each assignment will directly assess one of the learning objectives and will be graded on a rubric. We will assess whether, in our best professional judgment, your work meets the specifications laid out in the rubric. The homework and critique rubrics will have two possible grades (Competent / Not yet competent), and the projects will have three (Sophisticated / Competent / Not yet competent).

Course Objective 1 will be part of each assignment. Each homework and the critique will assess basic competence on one of Course Objectives 2-5 and should not take long.

Each project will assess one of Course Objectives 6-8. Ideally, the projects will also be polished enough for your portfolio or CV, and hence may require more effort. I will provide suggested project topics. You are welcome to come up with your own topic instead, although in that case I encourage you to discuss your project ideas with me before you start.

Revisions

To encourage you to learn from our feedback, you can revise and resubmit each assignment multiple times (up to once a week per assignment, to give us time for grading and turnaround).

However, don't worry that you'll be expected to revise many times! This multiple-revisions policy is there for flexibility, not to set unreasonable expectations. Most likely, you will only need one or two turnins for each assignment. One should be enough for homeworks and the critique. Projects may need a draft and a final submission. (Still, we reserve the right to cap the total number of revisions if necessary.)

Also, once you earn a Competent score on a given project, it is "safe": you can revise and resubmit (aiming for Sophisticated) without worrying that the score will drop.

Deadlines

All assignment rubrics will be made available early in the semester, in case you want an early start on homeworks or projects. Critiques will be assigned after your classmates' project drafts have been turned in.

Assignments will be due on Saturdays at 5:00pm. Each first submission must show sincere effort; you cannot just turn in a blank HW for later revision. If you miss enough deadlines, your grade will be affected.

Revisions (if needed) can be submitted once a week, until 5:00pm on Saturday October 24. No revisions will be accepted after that point. We will need this time to prepare your final grades.

Deadline (5pm)	Assignment
9/5	HW 1: Legible Graphics
9/12	HW 2: Visual Perception
9/19	HW 3: Grammar of Graphics
9/26	Project 1: Graphic Design
10/1 (Thu, 3pm)	Critique
10/10	Project 2: Interaction Design
10/17	Project 3: Research
10/24	Last opportunity to submit revisions

Attendance and participation

I expect you to attend class and be actively engaged. I am working hard to make class time worthwhile: we will discuss, critique, and redesign graphics, work through practical exercises, demonstrate useful software, etc. These are not things you can recreate by reading the slides afterwards. If you miss enough classes, your grade will be affected.

I encourage you to think of this course as participating in a lively, temporary learning community for 8 weeks—not as getting a packet of course materials to read on your own time later.

Course grade requirements

Note: any grade below B- is a failing grade for CMU graduate students. Therefore, to pass, you **must** earn a Competent score on all three HWs, the Critique, and the first two Projects.

Grade	Requirements
R	Fail to meet requirements for D
D	Earn Competent grade on the HW1 Legible Graphics
C	Earn D **and** Competent grades on HW2 Visual Perception, HW3 Grammar of Graphics, and the Critique
B	Earn C **and** Competent grades on P1 Graphic Design and P2 Interaction Design
B+	Earn B **and** Competent grade on P3 Research
A-	Earn B **and** Sophisticated grade on one of P1 or P2
A	Earn A- **and** Competent grade on P3 Research; **or** Earn A- **and** Sophisticated grade on both P1 and P2, but nothing on P3
A+	Earn B **and** Sophisticated grade on both P1 and P2 **and** Competent grade on P3

If you miss class or assignment deadlines, your course grade may drop below the level earned in this table. For every 4 days of missed classes or late submissions, your course grade will drop by one +/- increment.
(Example: you earn an A, but you miss one class, submit a homework 2 days late, and submit a project 1 day late. That's a total of 4 missed / late days, so your grade drops to A-.)

Summary of grades B through A+, depending on whether Project 3 is attempted (all assuming that requirements for C are also met):

	P1	
P2	**Comp.**	**Soph.**
Comp.	B	A-
Soph.	A-	A

(a) Without P3

	P1	
P2	**Comp.**	**Soph.**
Comp.	B+	A
Soph.	A	A+

(b) With Competent P3

Finally, CMU undergraduate students are graded without +/- scores.

- Undergraduates who meet requirements for B through A- will earn a B.

- Undergraduates who meet requirements for A or A+ will earn an A.

5

Academic integrity

All students are expected to comply with the CMU policy on academic integrity:
`http://www.cmu.edu/academic-integrity/`
Always ask if you are unsure whether your actions comply with the assignment instructions. Always acknowledge any help received on assignments: list the names of the people you worked with, and cite any external sources you used. You are encouraged to discuss assignments with your classmates, but the work you submit must be your own.

Cheating or copying of any sort are typically grounds for course failure. At the very least, you will receive no credit for the assignment, and we reserve the right to drop you down a letter grade. We are obliged to report any incidents to the appropriate university authorities.

Laptops and mobile devices; video/audiotaping

You are encouraged to bring a laptop to class for course-related use (following along with course examples or software demos). However, please silence any laptops or mobile devices. We reserve the right to disallow their use in class if it becomes disruptive.

No student may record or tape any classroom activity without the express written consent of the instructor.

Communication and email

Assignments, updates, and other class information will be posted on Blackboard. Help with using Blackboard is available at `http://www.cmu.edu/blackboard/gettingstarted/`

We are obligated to communicate with you about the course only through your `...@cmu.edu` or `...@andrew.cmu.edu` account. Please check your CMU mail regularly, or set up email forwarding if you normally use another email service.

Disability services

If you have a disability and need special accomodations in this class, please contact the instructor and the Disability Resources office: 412-268-2013, `access@andrew.cmu.edu` `http://www.cmu.edu/hr/eos/disability/`

Tentative schedule

Date	Topic
Week 1	**Data Visualization Basics**
Tu 9/1	Course introduction; history, classic examples; installing R and `shiny`
Th 9/3	Best practices for core 1D & 2D charts (and tables); `xtable`, `knitr`;
	base R plots (bar, box, hist, scatter, line, KDE); image formats & resolution
Sa 9/5	HW 1 Legible Graphics due 5pm
Week 2	**Data Visualization Principles**
Tu 9/8	Dataviz as 'external cognition'; preattentive processing and perceptual tasks;
	weaknesses of pies, 3D bars, glyphs, etc.
Th 9/10	Grammar of Graphics; `ggplot2`, Tableau
Sa 9/12	HW 2 Visual Perception due 5pm
Week 3	**Design Principles**
Tu 9/15	Graphic design; layout; visual style; color theory; communicating the story;
	sketching; titles & annotations; Inkscape (or Illustrator)
Th 9/17	Animation & interactive graphics; interaction design; affordances;
	Shneiderman's mantra; brushing and selection; `shiny`, `animation`, `rgl`; D3.js
Sa 9/19	HW 3 Grammar of Graphics due 5pm
Week 4	**Research and Communication**
Tu 9/22	Doing dataviz research; open Qs: displaying uncertainty in maps & rankings,
	mult. comparisons, survey weighting, missing data, . . . ;
	Inkscape tutorial continued
Th 9/24	Shiny and D3 practice
Sa 9/26	Project 1 Graphic Design due 5pm
Week 5	**Statistical Analysis and Maps**
Tu 9/29	Binwidths/bandwidths for histograms and KDE; regression diagnostics;
	plotting math functions (contours, 3D densities)
Th 10/1	Maps; principles of cartography; map projections;
	Critique due 3pm (in class)
Week 6	**Special Topics**
Tu 10/1	*No class*
Th 10/8	High-dim. data: brushing & linking, projection pursuit, GGobi
Sa 10/10	Project 2 Interaction Design due 5pm
Week 7	**Special Topics**
Tu 10/13	Networks and hivemaps; trees and treemaps
Th 10/15	Wrap-up; bonus topics
Sa 10/17	Project 3 Research (optional) due 5pm
Week 8	**Finals Week**
Tu & Th	No class; office hours in usual class time/location
Sa 10/24	All final revisions due 5pm