A Data Visualization Course at an Art School

Peter M. Border
Minneapolis College of Art and Design and
University of Minnesota
School of Physics and Astronomy
border@mail.physics.umn.edu

Abstract

We describe an online class at the Minneapolis College of Art and Design on data visualization. The class aims to give a visual, intuitive, hands-on understanding of statistical data presentation suitable for art students. Formal mathematics is kept to an absolute minimum. The curriculum is divided into three sections: a section on fundamentals of what graphs mean, a section on three graphical designers, and a section on state-of-the-art visualization. We use a system developed at MCAD to encourage community-building and mutual critique in an online enviroment. At the time of this writing, the class is beginning its third year and is doing well.

1 Introduction

This paper descirbes an online class in data visualization given in the BS:Visualization program at the Minneapolis College of Art and Design. The BS:Visualization program has some overlap with the BFA program, but is more technically oriented and emphasizes communication and process. The reasoning behind this course is summed up by a provocative quote from Edward Tufte: "Lurking behind the inept graphic is a lack of judgment about quantitative evidence. Nearly all those who produce graphics for mass publication are trained exclusively in the fine arts and have had little experience with the analysis of data." [Tufte, 1992] This course is an attempt to remedy the situation, and to give some experience with data analysis to the fine arts students so disdained by the irascible Prof. Tufte. Our students learn about data analysis and visualization from fundamentals to state-of-the-art techniques, and learn to produce graphs that are meaningful as well as beautiful.

The course is an 11-week, 4-credit online course at MCAD, currently in its third year. Normally it has from 5-10 students, usually a mixture of BS:Visualization and BFA students at MCAD, with a few professionals from outside the College. The text this year is Tufte's *Visual Explanations* [Tufte 1997], supplemented by instructor-written notes on Fundamentals of Data Analysis, and works by Nigel Holmes [Holmes, 1984] and Richard Saul Wurman [Wurman, 1999]. We also rely heavily on web-based materials.

2 Minimal Formal Mathematics

This course is unusual in that the use of mathematics is kept to an absolute minimum, and every attempt is made to explain concepts in a visual, hands-on, concrete, intuitive fashion. Mathematics is obfuscatory at best for many of our students, and we find that a visual approach is perfectly adequate for the very practical level of this course. We do no derivations, algebra or calculus and rely instead on descriptive, visual explanations. We make heavy use of spreadsheets and their built-in functions, which are very helpful for minimizing formal mathematics, since we can show things instead of proving them. Being able to use a spreadsheet is a highly employable skill in its own right which adds to the take-away value of the course.

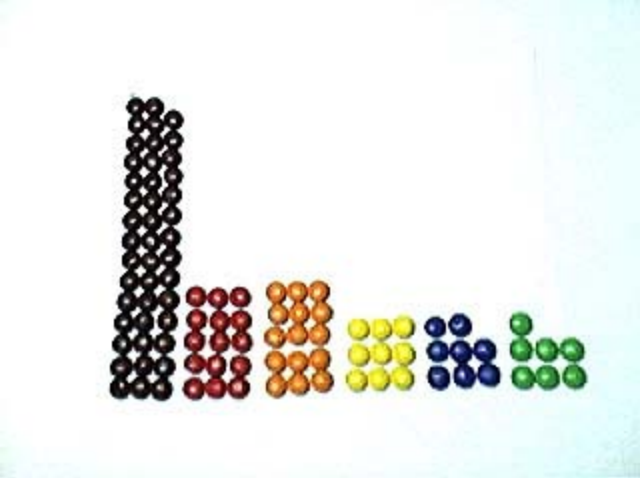

Figure 1. A hands-on explanation of a histogram.

3 Curriculum

The curriculum divides into three sections, described below.

Fundamentals and spreadsheets A week each on data (with units and metadata), histograms (introducing spreadsheets), common distributions, time-dependent plots (with an outline of curve fitting and error analysis) and multi-variable plots (with an outline of covariance). Although we discuss some advanced topics, explicit mathematics is kept to a minimum by a generous use of illustrations, spreadsheet files, examples and hands-on work. For example, histograms are explained by pushing around a random pile of candies into a bar graph sorted by color (see Fig. 1), and distributions are simply classified visually into flat, bell, or skewed bell categories.

Each week, students find some data of their own choice and produce a graph of the assigned type. Choosing their own data forces them to think about the subject, and leads to some excitement and competition when they view each other's work. Notable examples from last year included a histogram of the lengths of songs on a students mp3 player and a time chart of the number of cards in the discard pile of a card game. There is also a weekly reading assignment from the text.

Designers: We spend a week each reading works by Edward Tufte [Tufte, 1997][Tufte, 2001], Nigel Holmes [Holmes, 1984] and Richard Saul Wurman [Wurman, 1999], and discuss their similarities and differences. Students are generally quite surprised to find that there are schools of thought about data visualization, and this section has some very lively discussion. The project is to do three graphs of the same data in the style of

the three different designers.

State of the Art: The first week of the last section is about using a demonstration version of commercial graphing product (Kaleidagraph, Origin, etc.). Students work through the tutorials and produce a graph with it. Unfortunately, the online nature of and the budget for this course precludes actually purchasing these products, but demonstration versions are enough for our purposes. The remainder of this section is spent perusing and critiquing the posters from the InfoViz conference. Students also work on their final projects during this time.

4 Projects

On most weeks, the students do a project involving visualizing some data of their own choosing, but there are also two large-scale projects:

Spring Break Survey: Over Spring break students usually interact with a much wider variety of people than they do when school is in session. We make use of this time by having our students give a questionnaire on Learning Styles to at least 10 people, keeping track of respondents age, sex, handedness and occupation. This gives them some experience with actually doing a survey and collecting raw data.

Final Project: For the last three weeks of the class, students are divided into groups of two or three and work together on a thorough visualization of a multi-variable data set. Their data-set is their own choice, but must be approved by the instructor. At least one group usually analyzes the data from the Spring Break Survey. Last year's standout project was an interactive graph showing how gasoline prices have changed over the last decade, breaking the costs down into raw oil, refining, transport and markup pieces. The students involved used Processing (from processing.org, a product of Ben Fry and Casey Reas) to handle the interactivity. Another group produced a sample brochure on AIDS statistics.

5 Building Community Online

Working online can be a lonely experience unless measures are taken to build community in the class. We use a system worked out at MCAD by Prof. Briggs for a class on art history, and subsequently used in online Studio Art classes.

Posting projects: Every week, students post their projects in the website's discussion boards at least 24 hours before the weekly chat session. After posting their work, they are required to look at each others work and write comments on at least 3 of the other students postings. The comments prove that they have studied the piece and not just glanced at it. Discussing each other's work is interesting and very helpful for community-building. It can also inspire some healthy competition.

Weekly chat: We have a weekly chat session where we check in and discuss the last weeks project, and get ready for the next weeks. We also critique a "Graph of the week" in this session, which is a selected commercially-produced piece with some interesting points, sometimes good and sometimes not.

6 Results and Future Plans

The course is currently in its third year, and is doing well. Students have been receptive to this course and the evaluations have been quite positive. Actually measuring the success of a course like this is rather difficult, and we invite suggestions as to a methodology.

There are several directions this effort could grow into in the future. An advanced version of the course, focussing on computer-aided and internet-based data visualization would be very interesting. A research program is possible, though it will have to be on a low level in the immediate future. Developing data visualization tools that can rapidly communicate large amounts of information has immediate practical consequences, and would make some wonderful student projects.

On an immediate level, MCAD is starting a games program in Fall 2006, and game interfaces present a superb example of rapidly tracking large amounts of data with complicated relationships. Simulation games usually have a plethora of data analysis tools built in, but the visualization behind them tends to be rather primitive. Another possibility is online science courses, which are always in need of some first-rate data visualization and information design, and students in an art-based visualization program are in a very good position to work on such problems.

7 References

HOLMES, NIGEL. 1984. *Designer's Guide to Creating Charts and Diagrams.* Watson-Guptill Publications.

TUFTE, EDWARD R. 1992. *The Visual Display of Quantitative Information.* Graphics Press.

TUFTE, EDWARD R. 1997. *Visual Explanations.* Graphics Press.

WURMAN, RICHARD SAUL. 1999. *Understanding USA*. Ted Conferences.

WURMAN, RICHARD SAUL. 2001. *Information Anxiety 2*. Que Press.