Homework Assignment 1 (1/12/2009)

25 (+ 5 EC + $n \times 3$ EC) Points — Due 1/23/2009 (by 1pm)

(i) (10 Points) Read Chapters 1, 2, and 3 of Tufte (1983) "*The Visual Display of Quantitative Information*". Then take a closer look at the figures on top of p. 55 ("Comparative Annual Cost ...") and on top of p. 69 ("Accroissement ...").

(∗) For each of these figures, explain which rule(s) (how to construct a bad graphic) from our lecture notes the graphics designer has followed, i.e., list the rule(s) and explain why it has been followed. Demonstrate how these poor graphics might be improved. Using the data from the graphic (or your best approximation if necessary), construct a superior representation of the same information, using R. Include a short write–up (about half a page to a page) as to how you believe your version improves on the poor original.

(ii) (10 Points) Repeat (∗) for the two recent graphics included below.

(a)

(b)
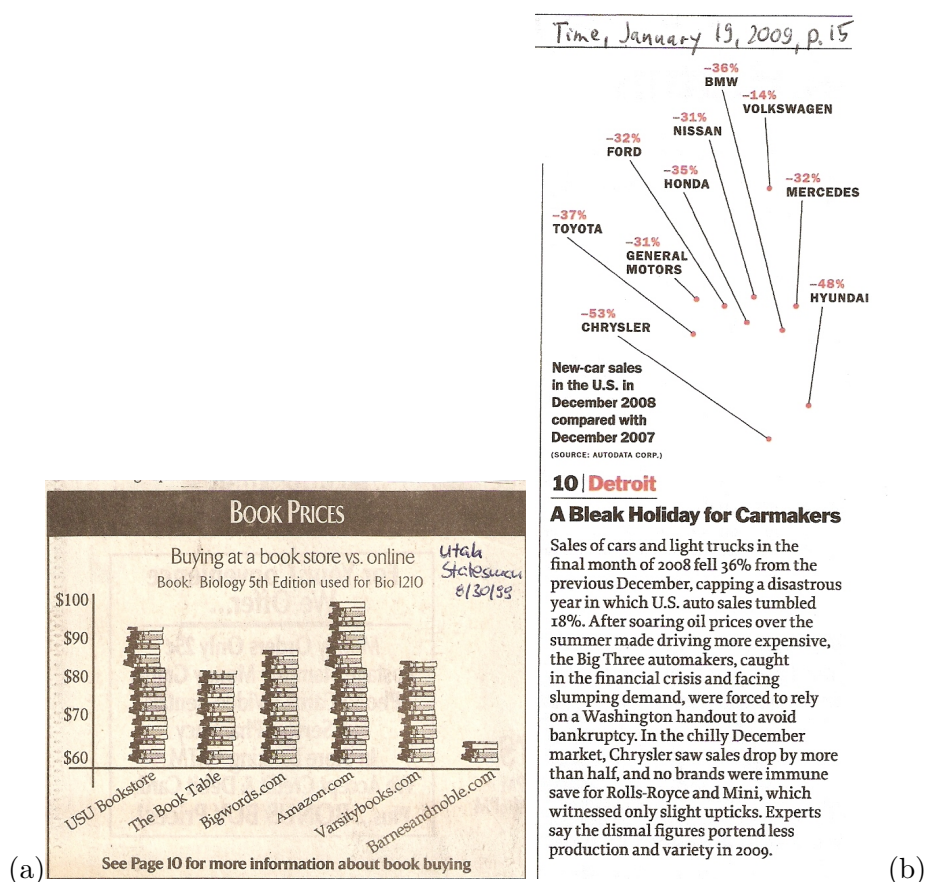
(iii) (5 Points) Find a fresh example of a poor statistical graphic. Do not choose your example from one of the books for this class, or another book specifically on graphics and charts, but from an original (preferably recent) source. Journal articles, newspapers, magazines, and scholarly books are all appropriate sources. Repeat (∗) for your graphic. Turn in a scan of the original together with precise documentation (source and page, URL, etc.) where you found that figure.

(iv) (Individual EC: 5 Points, due 2/20/2009, 1pm) Construct a graphic for the *Chance* graphics contest, as described in *Chance* 21(4), p. 62, from Fall 2008. Also see `http://www.public.iastate.edu/~larsen/` and `http://www.public.iastate.edu/~larsen/graphic%20contest/graphics%20contest.pdf`. The data (in csv format) are available from these Web pages. If you want to participate at the competition, their deadline is 1/15/2009! 2/20/2009 is our internal deadline (just for this course). I could not find the original 1951 figure in a book or article published by Will Burtin, but the NY Times republished the figure as part of a book review, dedicated to his life and work.
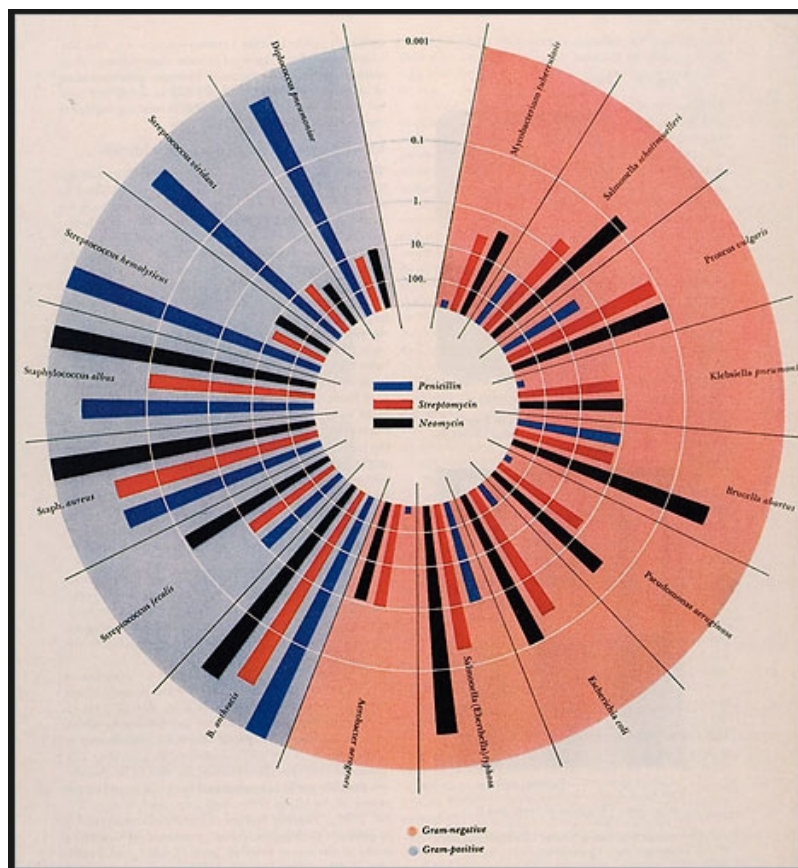
Check here `http://www.nytimes.com/2008/06/01/books/review/Heller-t.html?_r=3&oref=slogin` for the book review and the source of the figure above.

(v) (Individual EC: 3 Points each, up to 5 times) For the duration of the course (until 4/17/2009), part (iii) may be repeated up to five times for extra credit. For each example of a poor published graphic which is turned in and handled as in (∗), up to 3 points will be added to your total score.

NOTE: EC questions usually are individual questions. In fact, only the first person who notices a poor graphic can work on that particular graphic. If you notice a poor graphic, you have to send an e–mail to me and all other students in class, indicating about the following: *I found a poor graphic on page 1 of the Salt Lake Tribune on Wed 1/7/2009, showing the average tax increase and decrease per person in the case of changes to the state's cigarette tax.* From the time you send this e–mail, you have 7 days to turn in your full answer. If you send an e–mail but do not turn in your answer within 7 days, 2 points will be subtracted from your final score. This is to prevent that someone grabs all easily accessible poor graphics (from sources like CNN, Time magazine, the Utah Statesman, etc.) and prevents others to work on these graphics. Also, at most one entry per person per week.

General Submission Rules: (for homework, EC, and individual reports)

All your submissions this semester must be typeset in LaTeX. In fact, your submissions should translate via `pdflatex`. Figures (from scans or from graphical software) must accompany your LaTeX document in electronic format. R code and data sets must be directly accessible from your document. You should assume that all documents reside in the same directory. For testing, one student should finalize all documents while another student checks the intended submission for completeness on a different computer. LaTeX warnings are OK, but LaTeX error messages will result in point deductions (depending on how much effort it takes on my side to fix a problem). Submit your files via e–mail to `symanzik@math.usu.edu`. In case your submission consists of four or more individual files, you have to collect these in a zip file and just submit this single zip file. There is no need to submit the final pdf file as I will retranslate all documents on my side.

Your files should be named as follows (or in a closely related way):

```
groupI_hwJ_main.tex
groupI_hwJ_figK.jpg
groupI_hwJ_qL.R
groupI_hwJ_qM_data.xxx

lastname_firstname_EC_hwJ_qL_main.tex
lastname_firstname_EC_hwJ_qL_figK.jpg

lastname_firstname_projectN_main.tex
lastname_firstname_projectN_figK.jpg
```

where I, J, K, L, M, and N will be replaced by appropriate integers. xxx can be any acceptable extension for R data files. Include comments in your files where possible, e.g., dates, names, purpose of a file, etc. Your `groupI_hwJ_main.tex` (and your `lastname_firstname_EC_hwJ_qL_main.tex` and your `lastname_firstname_projectN_main.tex`) need to translate correctly in my `lect_main.tex` environment. All necessary templates will be provided later this week.