



Prof. Antony Unwin
Dept of Computer-Oriented Statistics and Data Analysis
Institute for Mathematics
University of Augsburg
<http://stats.math.uni-augsburg.de/>

Stochastik IV – Graphical Data Analysis

Exercise Sheet 5: Grammar of Graphics and *ggplot2*

Tutorial: Tuesday, 22nd November, 2011, 10.00 - 11.30 Uhr, Room 3029

Background

Many plots can be drawn in either the base package *graphics* or in *ggplot2*. Carry out a graphical analysis of the variables in the dataset *Salaries* from the package *car* and draw each plot you use with both packages. Discuss what the plots show, which plot of each pair you think is better and which is easier to prepare. Start with plots of the individual variables and then try bivariate displays. If you have time (and are feeling adventurous) consider drawing multivariate displays.

Are there plots that you would like to draw, which can only be drawn with one of the packages or cannot be drawn with either of them?

Extra: Graphics for what politicians say

The website

politilines.periscope.com

offers an interactive graphic to allow you to explore what the Republican candidates for the party's nomination for US president said in the televised debates. The graphic is also discussed on Nathan Yau's blog *Flowing Data*:

flowingdata.com/2011/11/11/politilines-shows-what-candidates-talk-about-during-debates/

Is this an effective graphic? What information can you draw from it? Can you suggest improvements or alternatives?

A main feature of this graphic is the interactivity it offers. What static graphics might be used instead?