

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 1: Graphics for some R datasets

**Tutorial:** Monday, 24th October, 2011, 12.15 - 13.45 Uhr, Room 2001 T

#### Background

Everitt and Hothorn's book 'A Handbook of Statistical Analyses Using R' discusses many examples, emphasising analyses rather than graphics, though including some graphics as well. Consider the following datasets from their book, available in the R package *HSAUR2*. Vignettes for most of the book's chapters can be found on the package's webpage on CRAN.

#### Datasets

- water (Vignette *Simple Inference*)
- pistonrings (Vignette *Simple Inference*)
- weightgain (Vignette *Analysis of Variance*)

#### Analysis

Analyse the datasets graphically and compare your results with Everitt and Hothorn's.

1. Do your graphics confirm their results?
2. Using your graphics, comment on the assumptions required for the models and testing used by Everitt and Hothorn.
3. Are there additional features which Everitt and Hothorn do not examine which might be worth investigating? How would you test for them?
4. What do you think of the graphics used by Everitt and Hothorn?

**Extra: see next page**

## A Graphic from the Media: IBM Commuter Pain Index

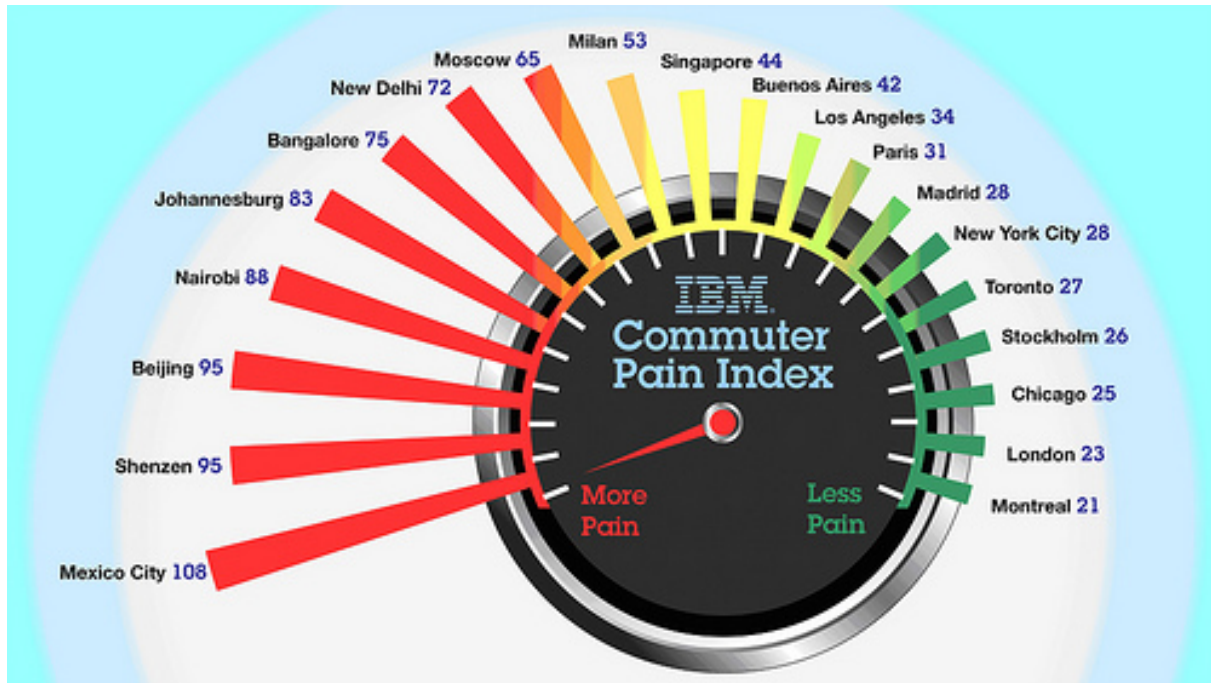


Figure 1: The IBM Commuter Pain Index  
(Source: [www-03.ibm.com/press/us/en/pressrelease/35359.wss](http://www-03.ibm.com/press/us/en/pressrelease/35359.wss))

What do you think of it? Could it be improved or should the data be displayed differently?