CATEGORICAL DISPLAYS OF A MILLION CASES

by

Heike Hofmann
Iowa State University

January 23, 2004

ABSTRACT

Fortunately, categorical displays are relatively independent from the number of cases they display. Unfortunately, the number of categories of a variable tends to grow with an increase in the number of cases. The challenge for categorical displays therefore is to deal with large number of categories and variables with skew marginal distributions. The goal in mind must - as always with visual displays - be the representation of both a global trend and individuals or small groups not fitting the overall trend. In an interactive framework techniques like re-ordering and grouping of categories help to produce clearer images of ‘what is in the data’. Querying mechanisms and variations of the display are used for extracting summarized information.