

# Categorical Displays of a Million Cases

Heike Hofmann  
Department of Statistics  
Iowa State University  
hofmann@iastate.edu

**Keywords:** Categorical Data, Exploratory Data Analysis, Massive Data Sets, Mosaic Plots.

## **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. 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.