Mosaicplot forms

- Which mosaicplot form is good for what?
- Examples from
  - Titanic
  - Bowling Alone 1998
  - Rochdale
  - Divorce
  - ...
- What information is there and what can be seen with which plot?

Multivariate categorical infos

- Association between variables
- Dependence between variables
- Very frequent combinations
- Empty or rare combinations
- Form of conditional distributions
- Similarity of conditional distributions
- …

Intepreting plots

- Does everyone see the same things?
- How can the information be described verbally?
  - "a picture is worth a 1000 words"
  - twitter is limited to 140 characters (how many are in a plot?)
- How can the information be assessed statistically?
- How important are
  - background knowledge
  - scales and labelling
  - title, caption, legend, guides, annotations, accompanying text?

Mosaicplot options

- Choice of variables (in effect aggregation)
- Type of plot
- Order of variables
- Whether variables are plotted horizontally or vertically
- Order of categories within variables
- Arregation of categories within variables
- Plotting only data subsets
- Formatting: plotsize, aspect ratio, spacing between levels
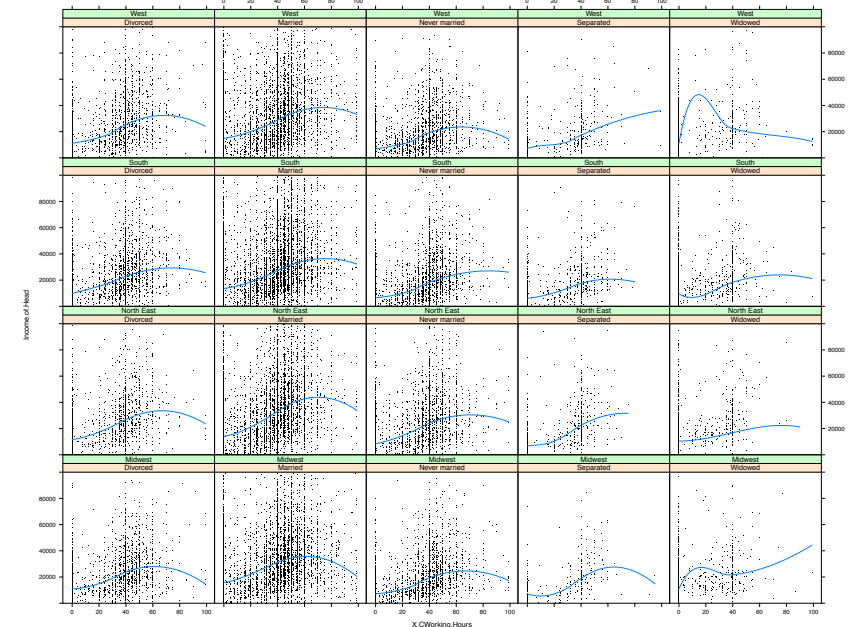
Structure of Mosaicplots

- Variable category combinations are represented by rectangles
- Rectangle area is (usually) proportional to frequency
- Rectangles may have equal width (height), so that height (width) is proportional to frequency
- Rectangles may be aligned in various ways
- Rectangles may be rotated
- Rectangles may be coloured

Small multiples

- Small multiples display a collection of subsets of the data, each subset in its own plot. All plots are of the same kind.
- Grammar of Graphics (Lee Wilkinson) and ggplot2 (Hadley Wickham) refer to facets.
- With Trellis (Rick Becker & Bill Cleveland) or lattice (Deepayan Sarkar) the subsets are based on conditioning on the values of other variables and subsets may overlap.
- With micromaps (Dan Carr) the plots are maps and the subsets are map areas chosen by conditioning on variables.
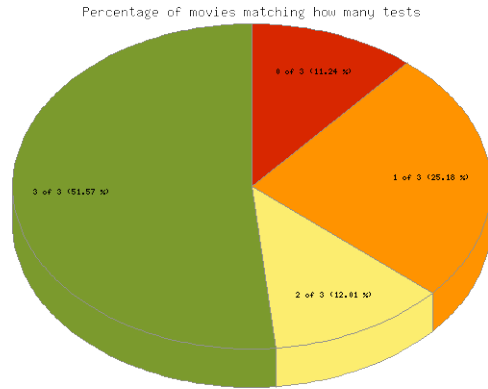- Multiple barchart mosaics are small multiples.

Trellis plots

- A panel plot is chosen (classically a scatterplot).
- Conditioning on categories of categorical variables and discretised continuous variables a panel plot is drawn for each of the subsets.
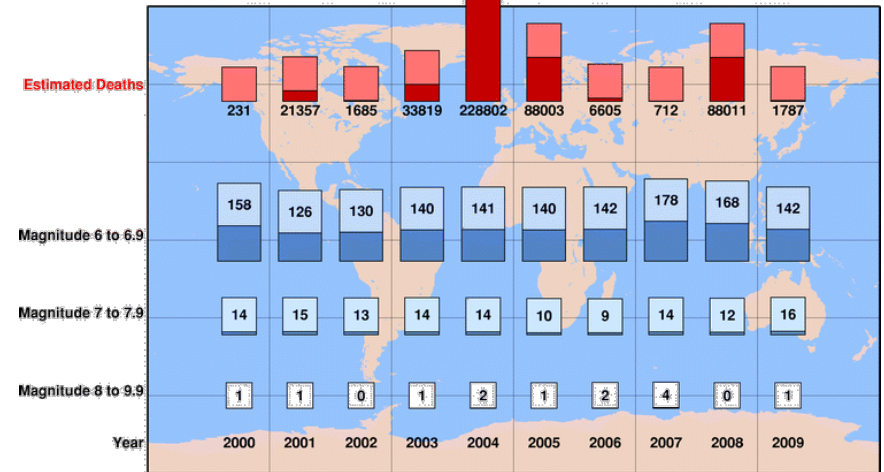- Each panel plot is scaled the same.

Bechdel Test for Films

1. It has to have at least two women in it

2. Who talk to each other

3. About something besides a man

Percentage of movies matching how many tests

0 of 3 (11.24 %)
1 of 3 (25.18 %)
2 of 3 (12.01 %)
3 of 3 (51.57 %)

bechdeltest.com/statistics/

Worldwide Earthquakes: 2000 - 2009 *

Estimated Deaths	231	21357	1685	33819	228802	88003	6605	712	88011	1787
Magnitude 6 to 6.9	158	126	130	140	141	140	142	178	168	142
Magnitude 7 to 7.9	14	15	13	14	14	10	9	14	12	16
Magnitude 8 to 9.9	1	1	0	1	2	1	2	4	0	1
Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009

* Located by the USGS National Earthquake Information Center