

Stat645

nd plots

Hadley Wickham

1. Mosaic plot – happiness data

2. Parallel coordinates plot – olives data

# Mosaic plots

# Happiness

Extracted from the general social survey

~51,000 surveys 1972–2006

8 variables related to happiness:

age, sex, marital status, relative financial status, health.

Plus happiness, year of survey, id variable, sampling weight

# New tools

Colour brush: cmd/ctrl + b

Mosaic plot!

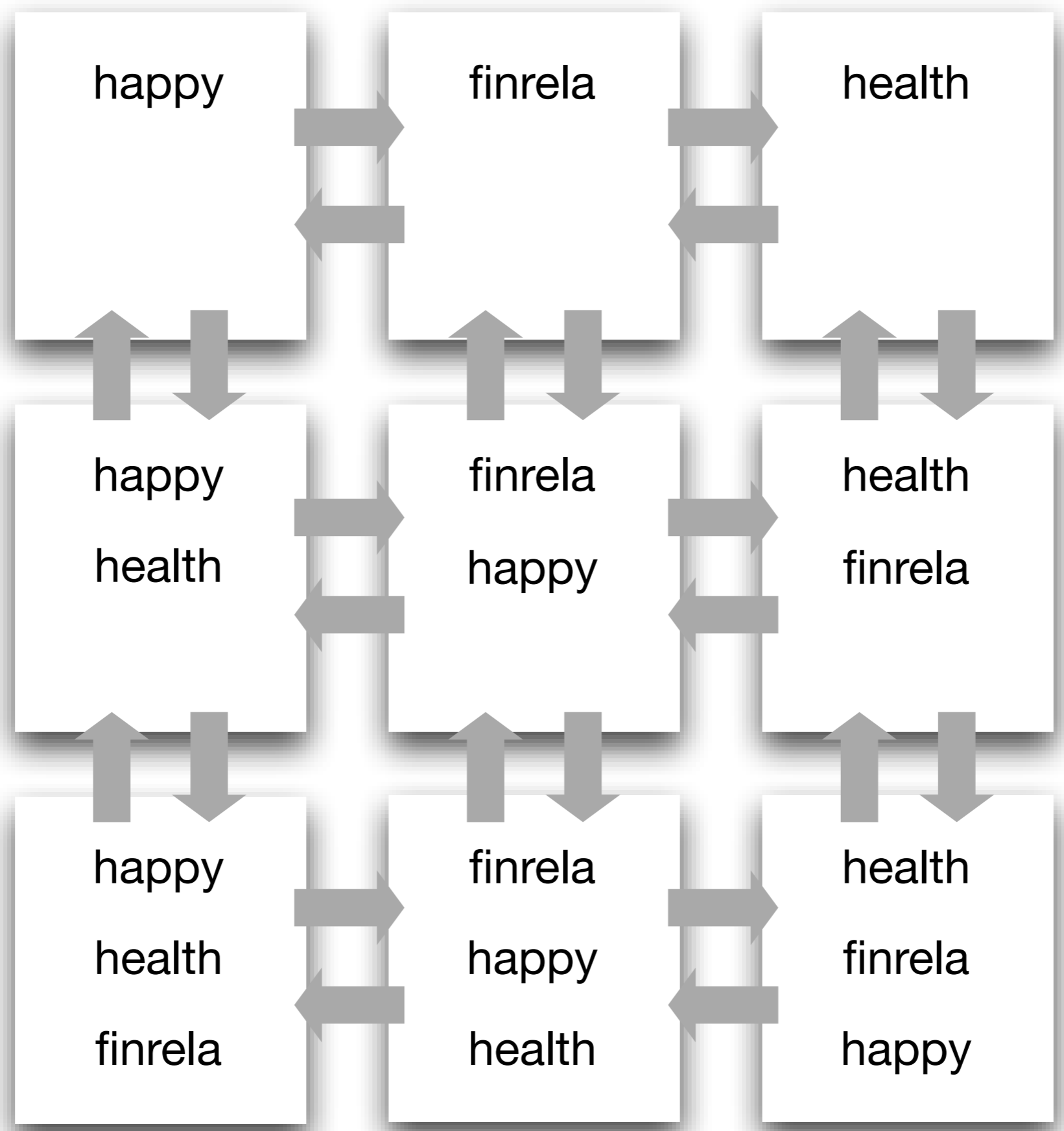
Should have basic idea how it works from the readings, just need a few practical tips.

Double click to change variable type

# Your turn

Use a mosaic plot to explore how marital status changes with age?

What are the advantages/disadvantages of this method vs a bar chart + histogram?



# Display as...

Fluctuation diagram: shows joint 2d distribution

Same bin size: useful for detecting patterns of missingness



# Your turn

How is happiness affected by health and wealth?

**PCP**

# Olives data

572 olive oils from 9 regions in 3 areas in Italy.

Composition of 8 different types of fatty acid.

Can we use the different variables to distinguish different regions?

# PCP

Alt-drag to rearrange axes

Click on name, then press ctrl/cmd + I to  
invert axis