

# Categorical data 1-d examples



# Categorical data 1-d examples

Bowling Alone (DDB Lifestyle survey 1998 excerpt)

## Variables

Church

Religfun

Religion

Age

Children

Citysize

Education

Spouse education

Region

Income

20 CHURCH Attended church or other place of worship (freq last 12 months)

- 1 None
- 2 1-4 times
- 3 5-8 times
- 4 9-11 times
- 5 12-24 times
- 6 25-51 times
- 7 52+ times

208 RELIGFUN Religious fundamentalism is the greatest peril in the country today

- 1 Definitely Disagree
- 2 Generally Disagree
- 3 Moderately Disagree
- 4 Moderately Agree
- 5 Generally Agree
- 6 Definitely Agree

209 RELIGION Religion is an important part of my life

# Categorical data 1-d questions

- What information might you want to know?
  - absolute frequencies of categories
  - relative frequencies of categories
  - are categories equally frequent
  - which categories are most important (if any)
  - which categories hardly or never occur
  - what ordering of categories is there
  - is there a pattern for ordinal categories
  - ...

# Categorical data 1-d displays

- Piecharts: for emphasising shares
  - order of categories
  - piece removed
  - 3-d
- Barcharts: for absolute frequencies
  - order of categories
  - horizontal/vertical

# Categorical data 2-d displays

- Stacked bars (equal width, unequal height)
- Side by side bars
- Fourfold displays, Sieves, Rings, Jittered dotplots
- Spineplots
- Mosaicplots and variants
- Treemaps
- ...

## An unbalanced society

There is (unsurprisingly) very little data about the distribution of wealth in America. There is one source, the *Survey of Consumer Finances*, sponsored by the Federal Reserve Board, that does provide data starting from 1983.

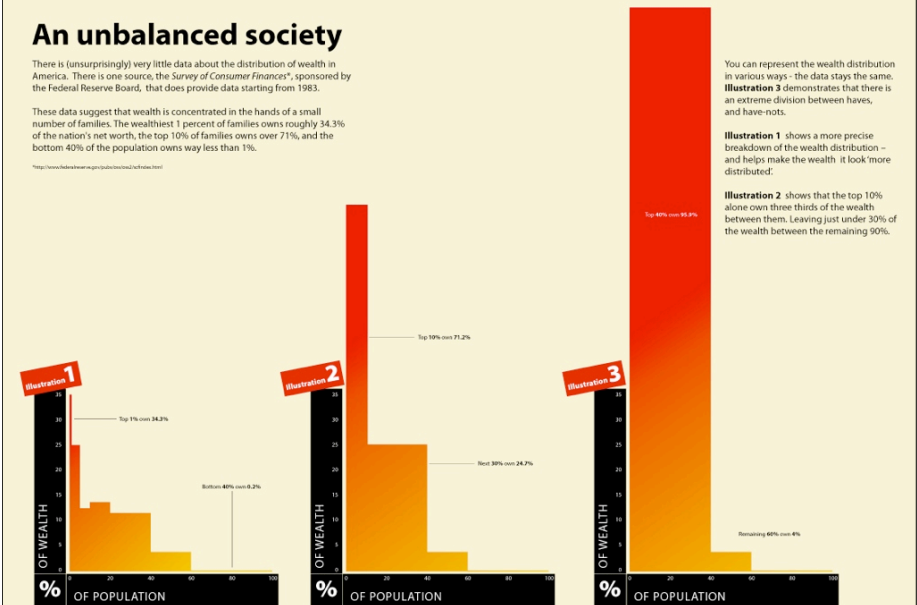
These data suggest that wealth is concentrated in the hands of a small number of families. The wealthiest 1 percent of families owns roughly 34.3% of the nation's net worth, the top 10% of families owns over 71%, and the bottom 40% of the population owns way less than 1%.

http://www.federalreserve.gov/oc/oc20020201.htm

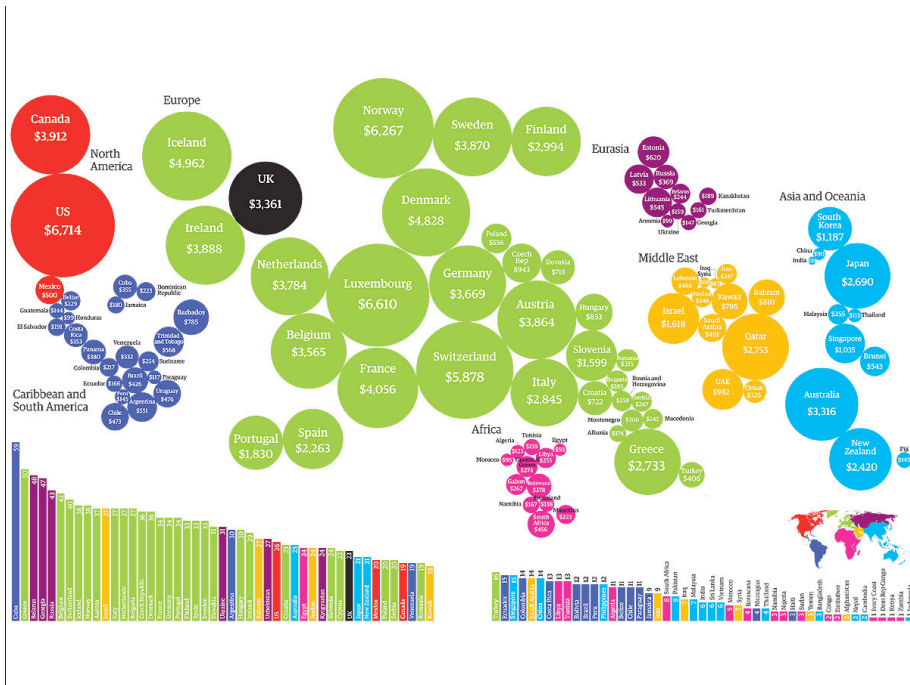
You can represent the wealth distribution in various ways - the data stays the same. **Illustration 3** demonstrates that there is an extreme division between haves, and have-nots.

**Illustration 1** shows a more precise breakdown of the wealth distribution - and helps make the wealth look more distributed.

**Illustration 2** shows that the top 10% alone own three thirds of the wealth between them. Leaving just under 30% of the wealth between the remaining 90%.



Designed by <http://www.efek.net> - 2010 • Data source 2004: <http://www.fairfield.edu/faculty/hodgson/Courses/so11/stratification/income&wealth.htm>



## Mosaicplot variants

- Classical (Observed)
- Expected (based on a model)
- Fluctuation diagram
- Same binsize
- Multiple barcharts
- Doubledecker
- rmb
- Weighted mosaics