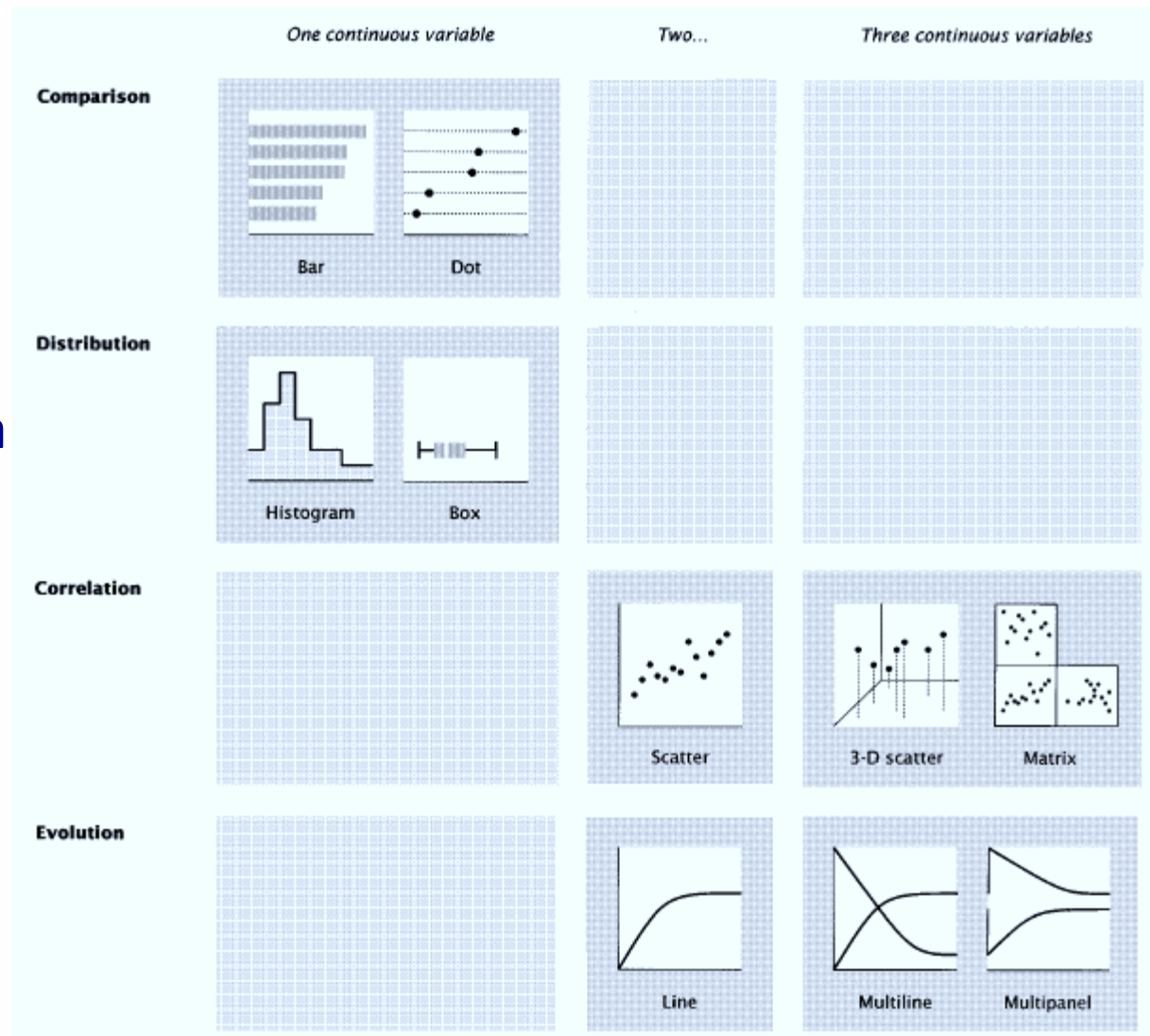


Choosing the right graph

ME447 Visualizing Data
Fall 2017–18

Richard Layton



To choose a graph, answer these basic questions about your data and your story

What story are you telling?

- distribution
- comparison
- correlation between variables
- evolution over time

How many variables are relevant to the story?

Each variable is of what type?

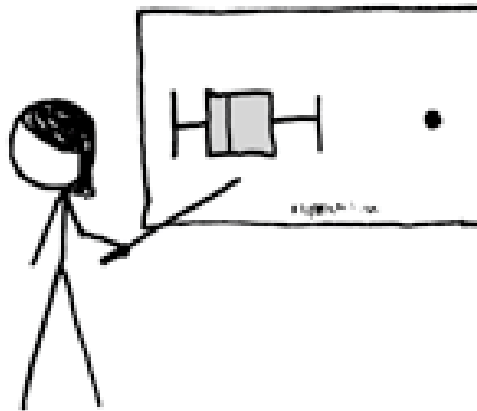
- quantitative (continuous)
- categorical (discrete)

story: distribution of a quantitative variable

I'M NOT YOUR
BOYFRIEND!

/ YOU TOTALLY ARE.
I'M CASUALLY
DATING A NUMBER
OF PEOPLE.

BUT YOU SPEND TWICE AS MUCH
TIME WITH ME AS WITH ANYONE
ELSE. I'M A CLEAR OUTLIER.



YOUR MATH IS
IRREFUTABLE.

FACE IT—I'M
YOUR STATISTICALLY
SIGNIFICANT OTHER.

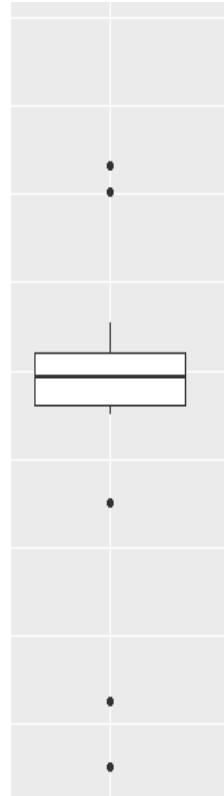
Story: **distribution**
Data: **1 variable (quantitative)**

point plot or box plot

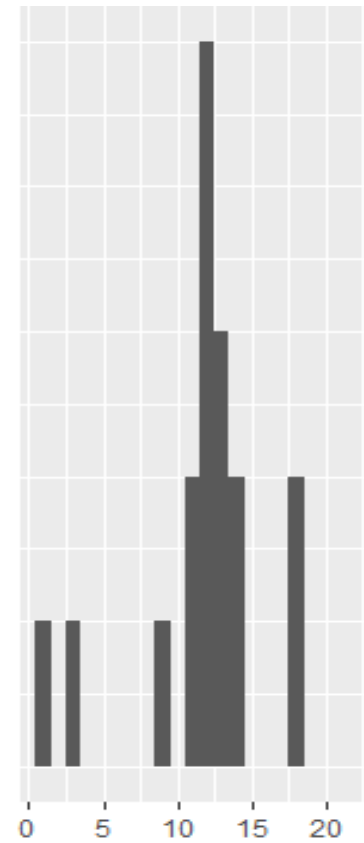


point plot
preferred

↑ quantitative variable ↓

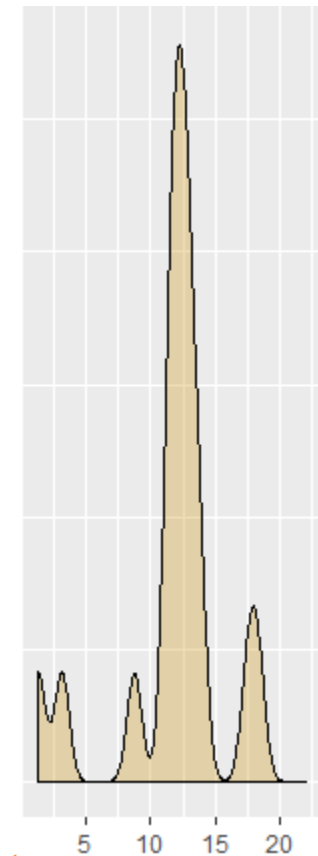


box & whisker
preferred



histogram

(prone to rhetorical malpractice)

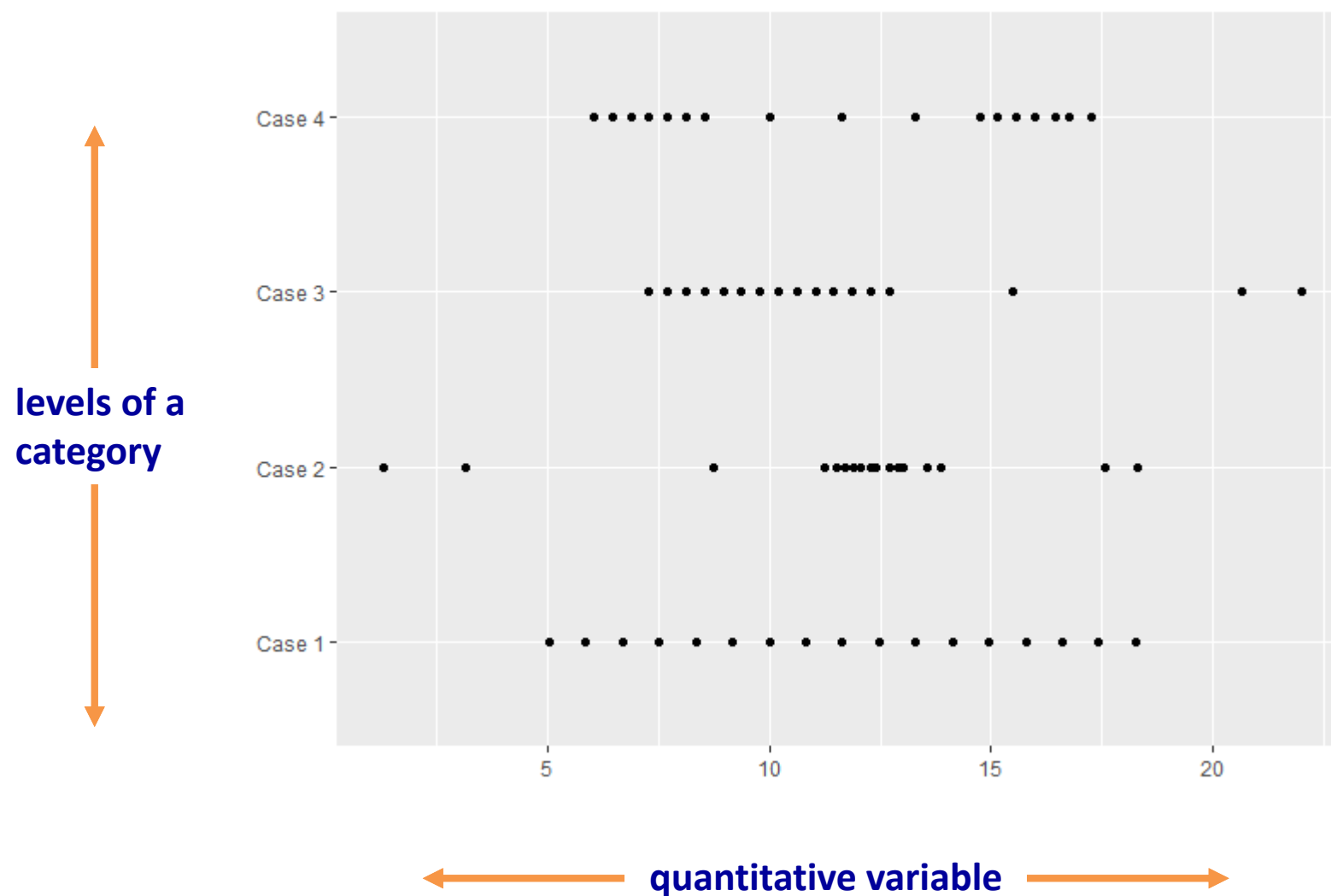


density

Story: **comparing distributions**

Data: **2 variables (1 quantitative, 1 categorical)**

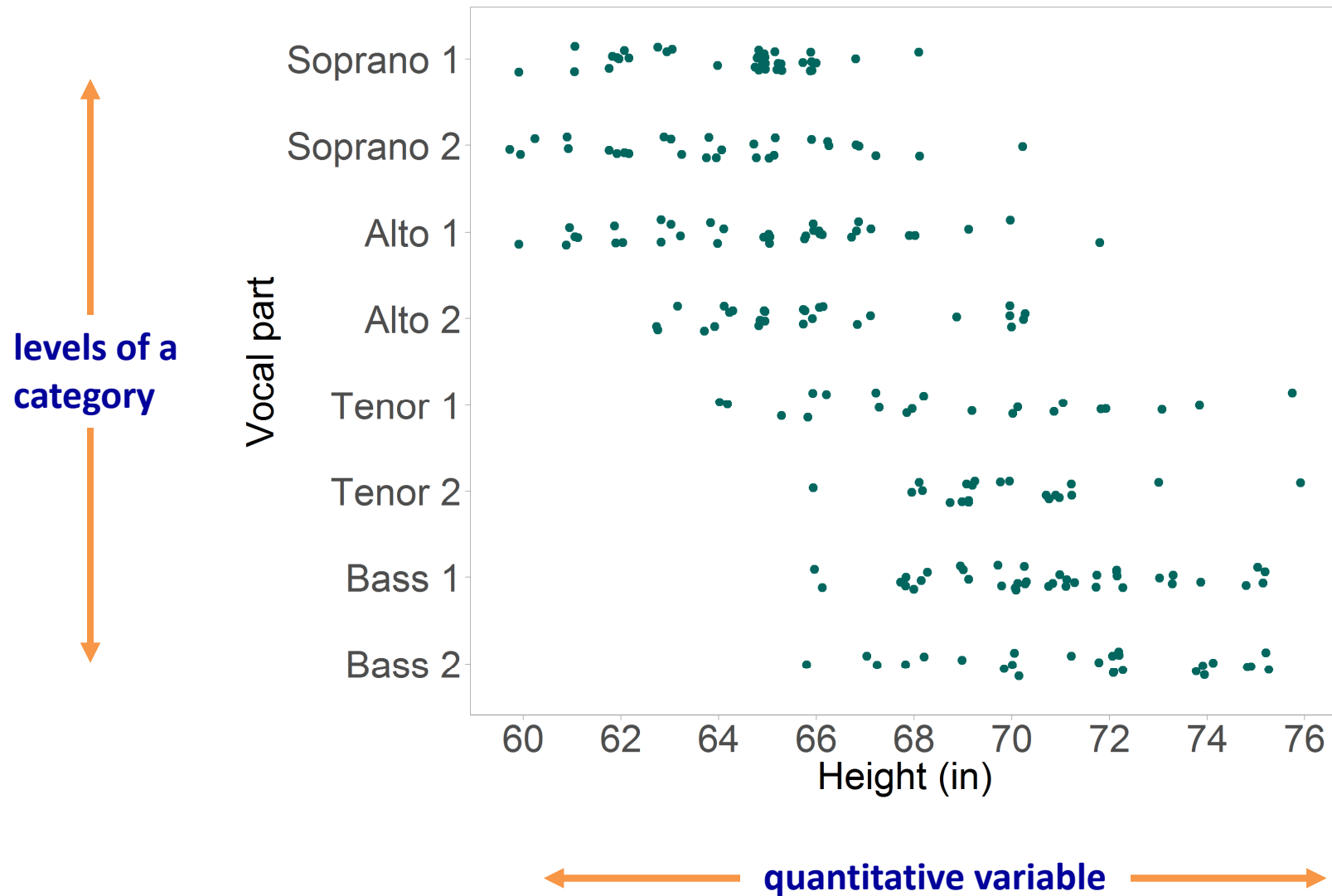
point plot



Story: comparing distributions

Data: 2 variables (1 quantitative, 1 categorical)

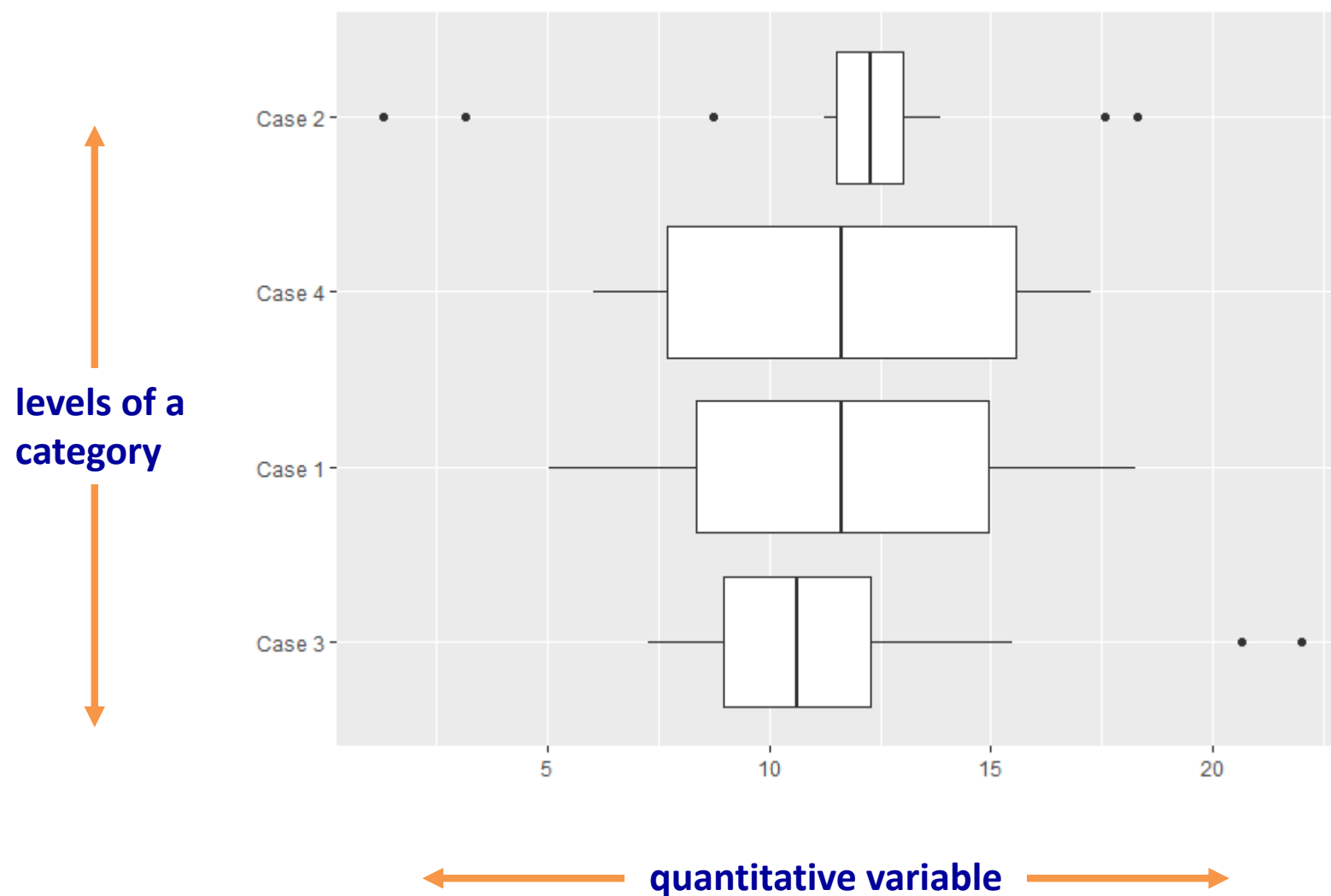
point plot



Story: comparing distributions

Data: 2 variables (1 quantitative, 1 categorical)

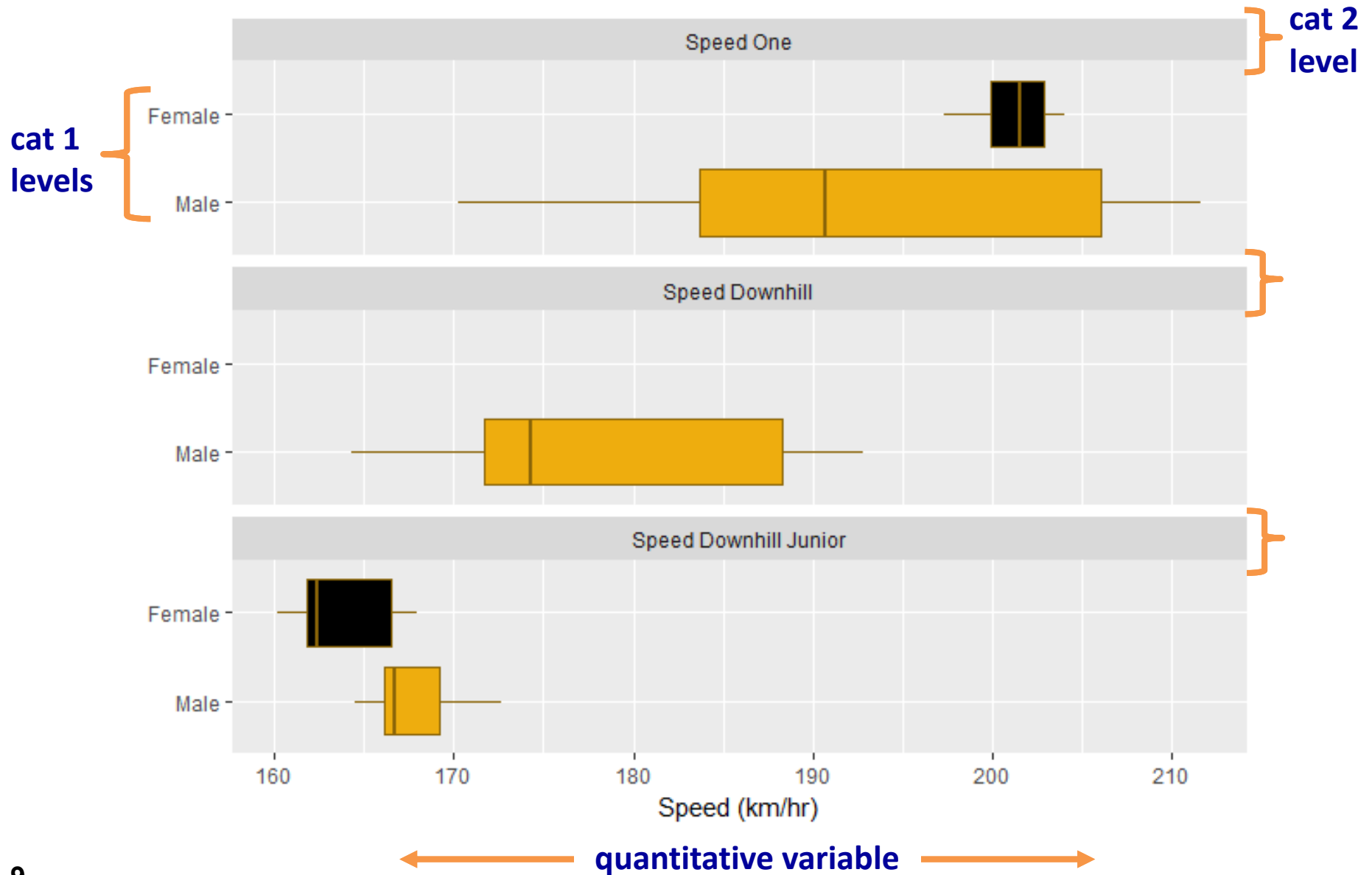
box plot



Story: **comparing distributions**

box plot

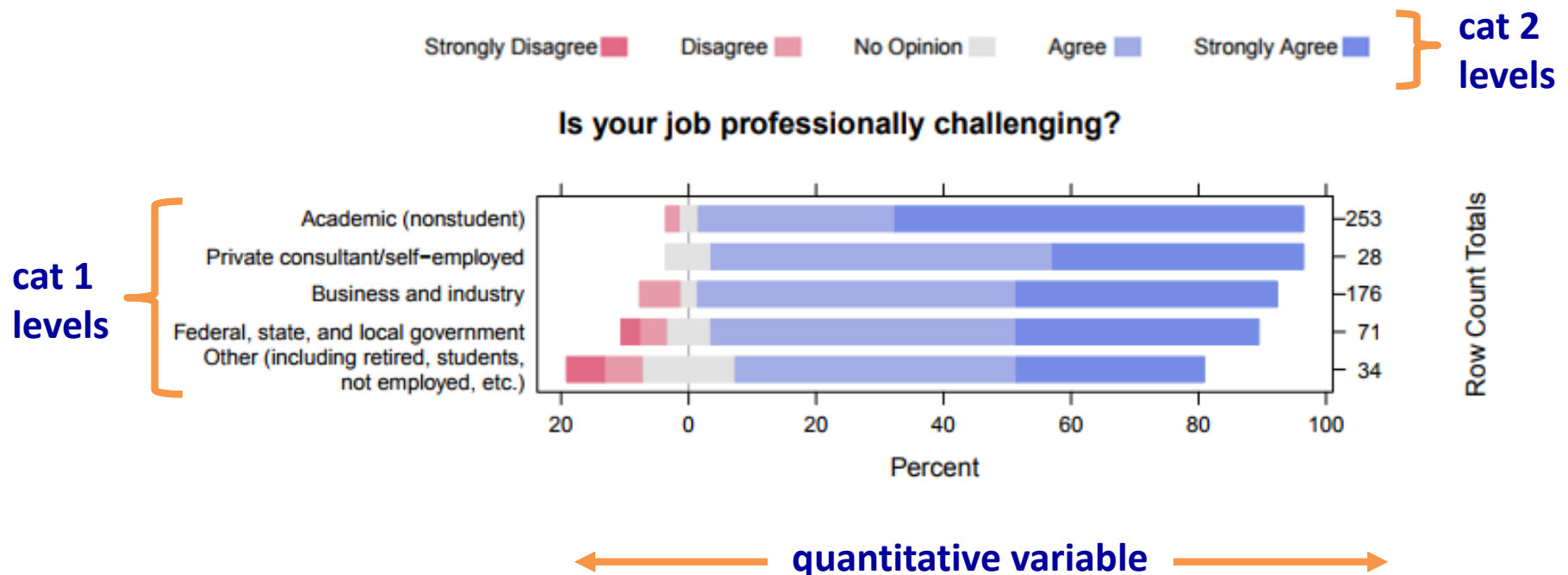
Data: **3 variables (1 quantitative, 2 categorical)**



Story: **comparing frequency**
Data: **3 variables (Likert-like scales)**

diverging stacked bar

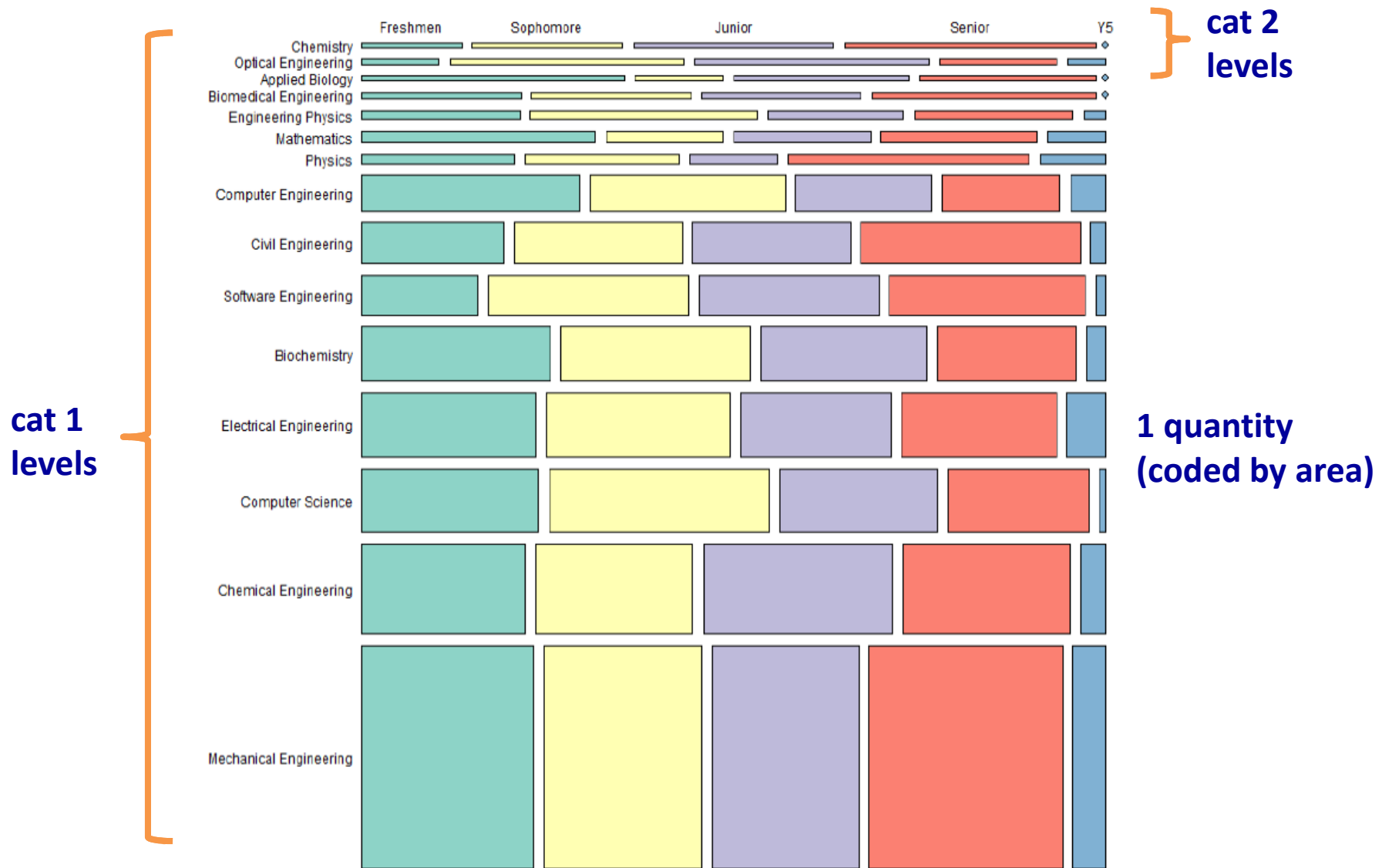
A response belongs to a mutually exclusive combination of levels.



Story: **comparing frequency**

Data: **3 variables (1 quantity, 2 categories)**

mosaic

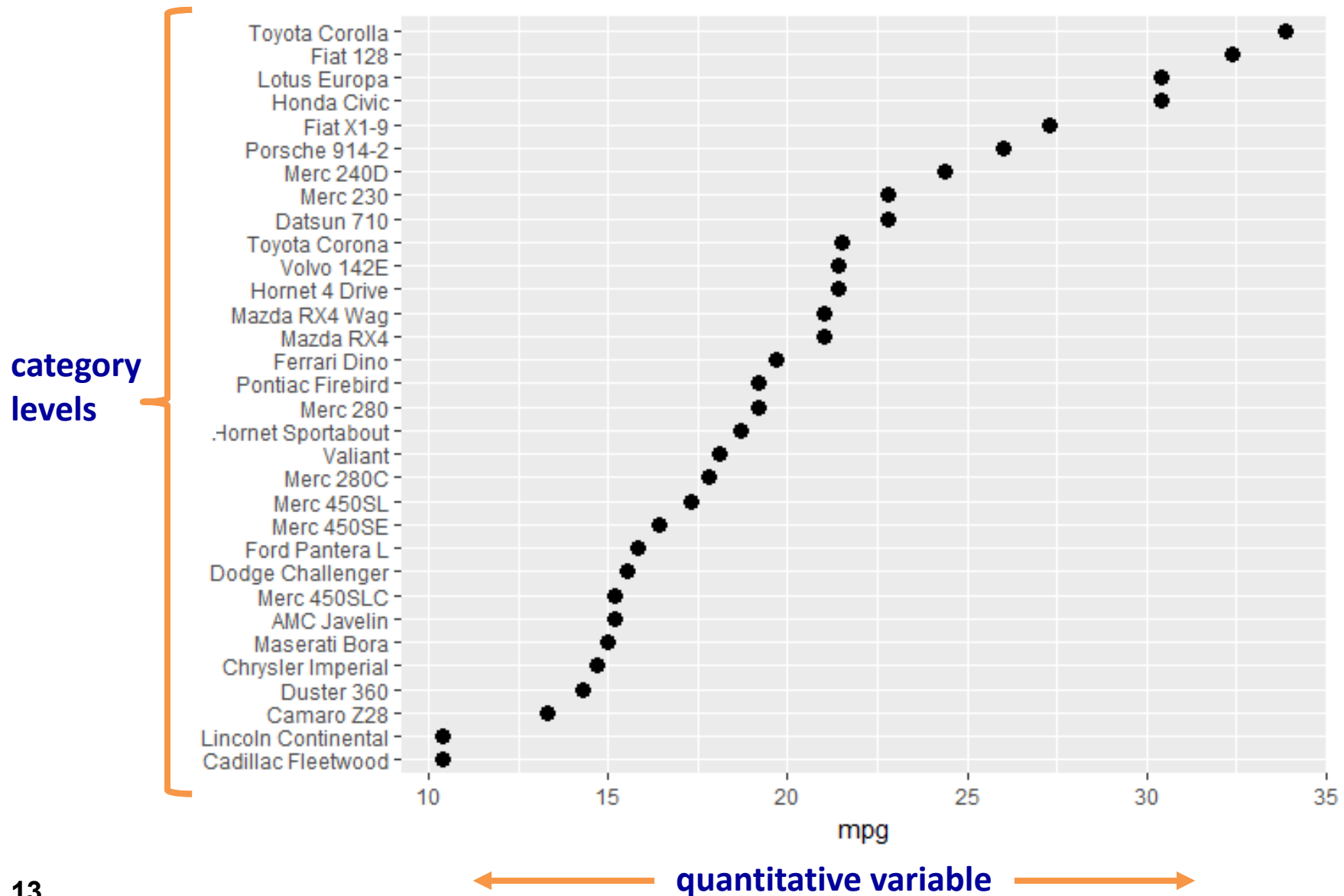


story: comparing a quantity among groups

Story: **comparing 1 quantity**

Data: **2 variables (1 quantitative, 1 categorical)**

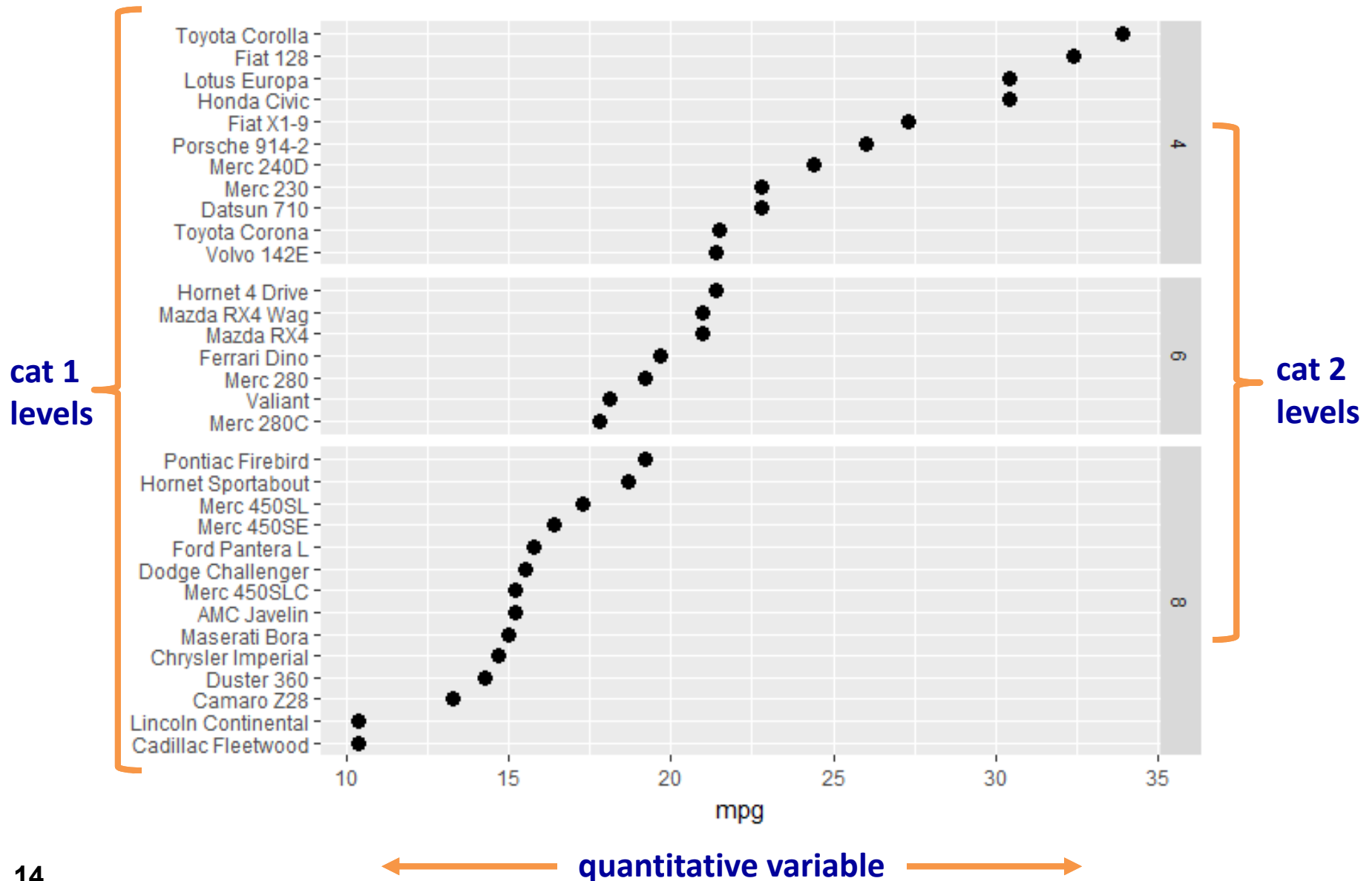
dot plot



Story: **comparing 1 quantity**

Data: **3 variables (1 quantitative, 2 categorical)**

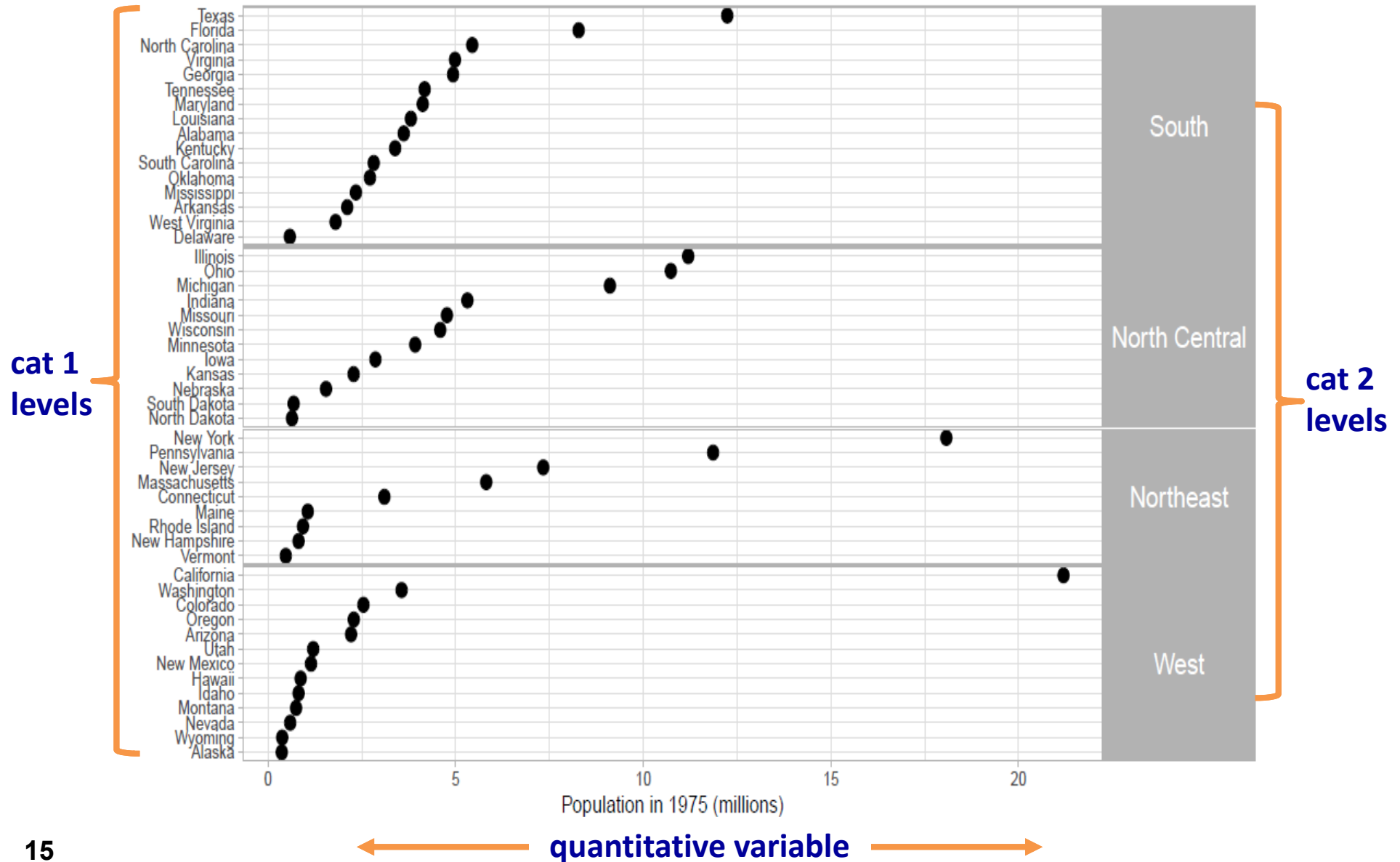
dot plot



Story: **comparing 1 quantity**

Data: **3 variables (1 quantitative, 2 categorical)**

dot plot

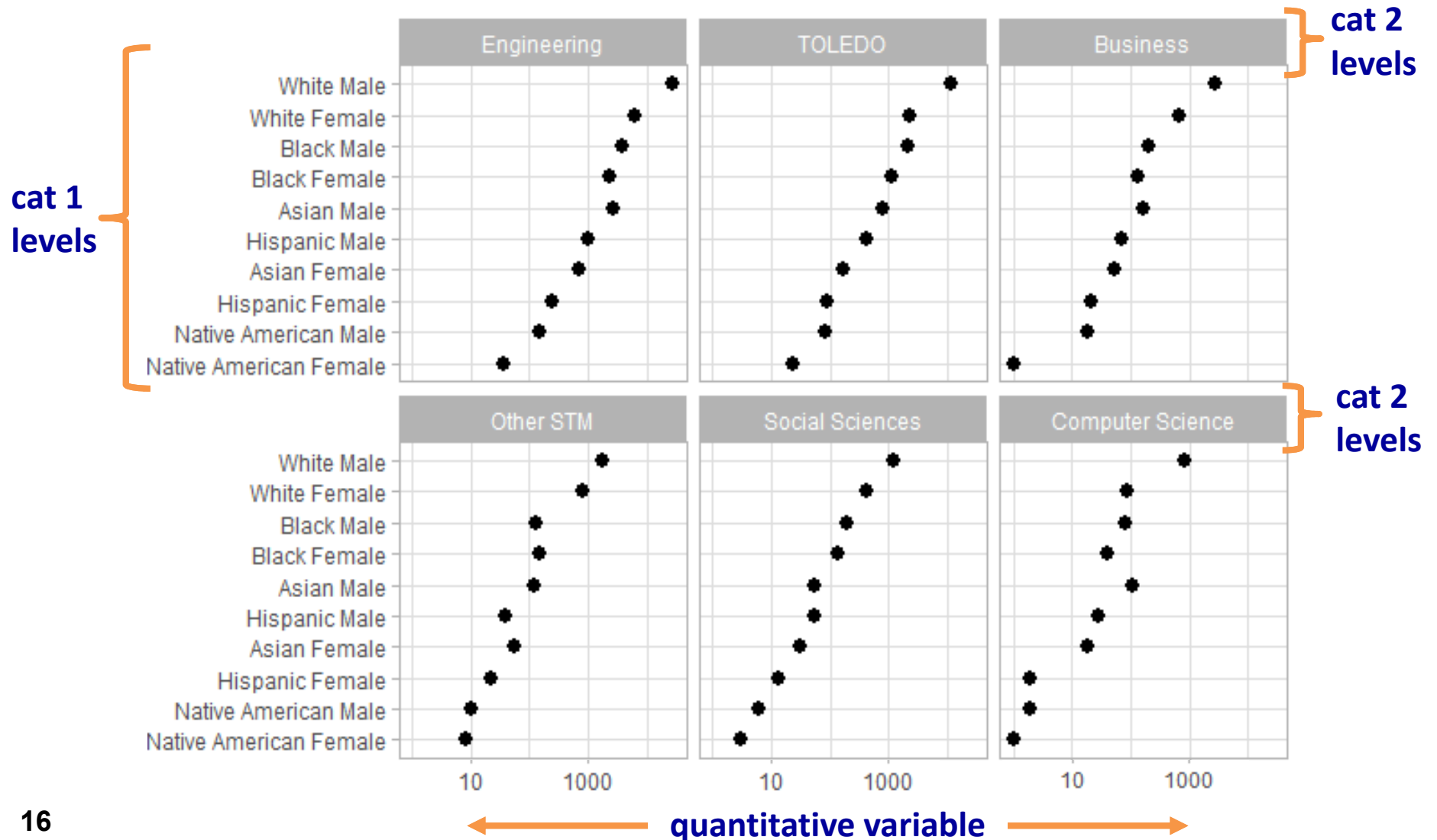


Story: **comparing 1 quantity**

Data: **3 variables (1 quantitative, 2 categorical)**

multiway

A quantitative value for *every combination* of levels of the two categorical variables.

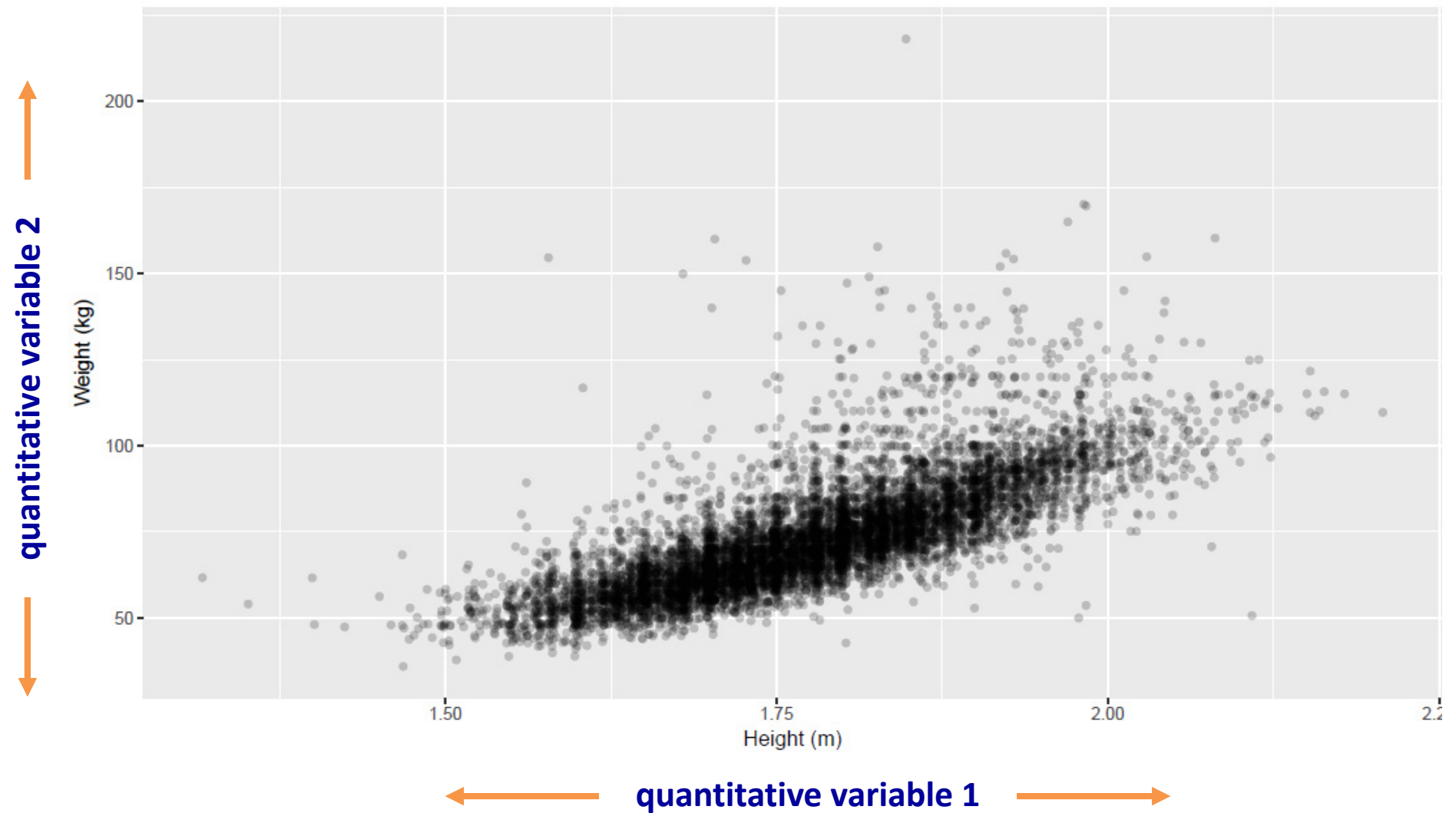


story: correlation between variables

Story: **correlation**

Data: **2 quantitative variables**

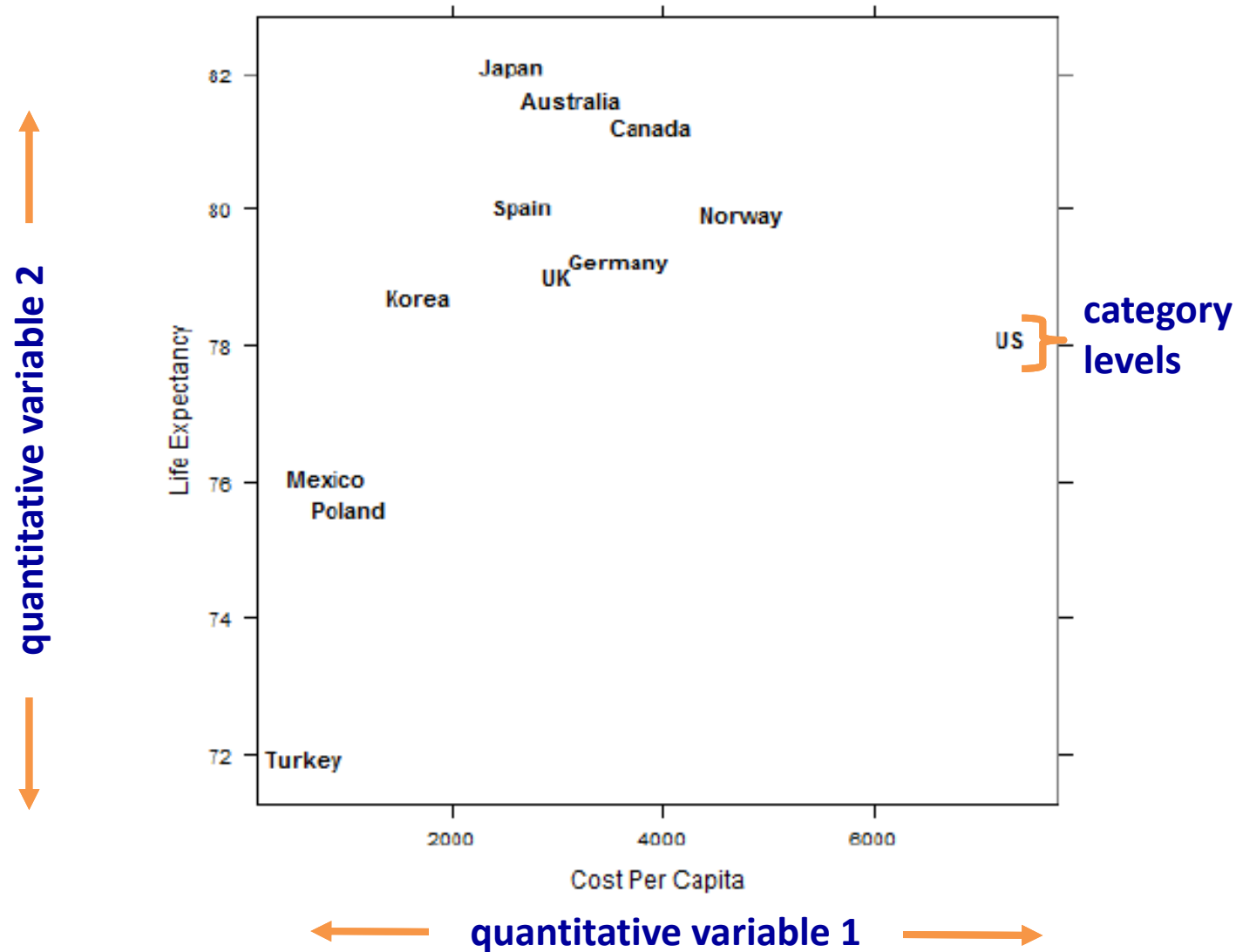
scatterplot



Story: correlation & comparison

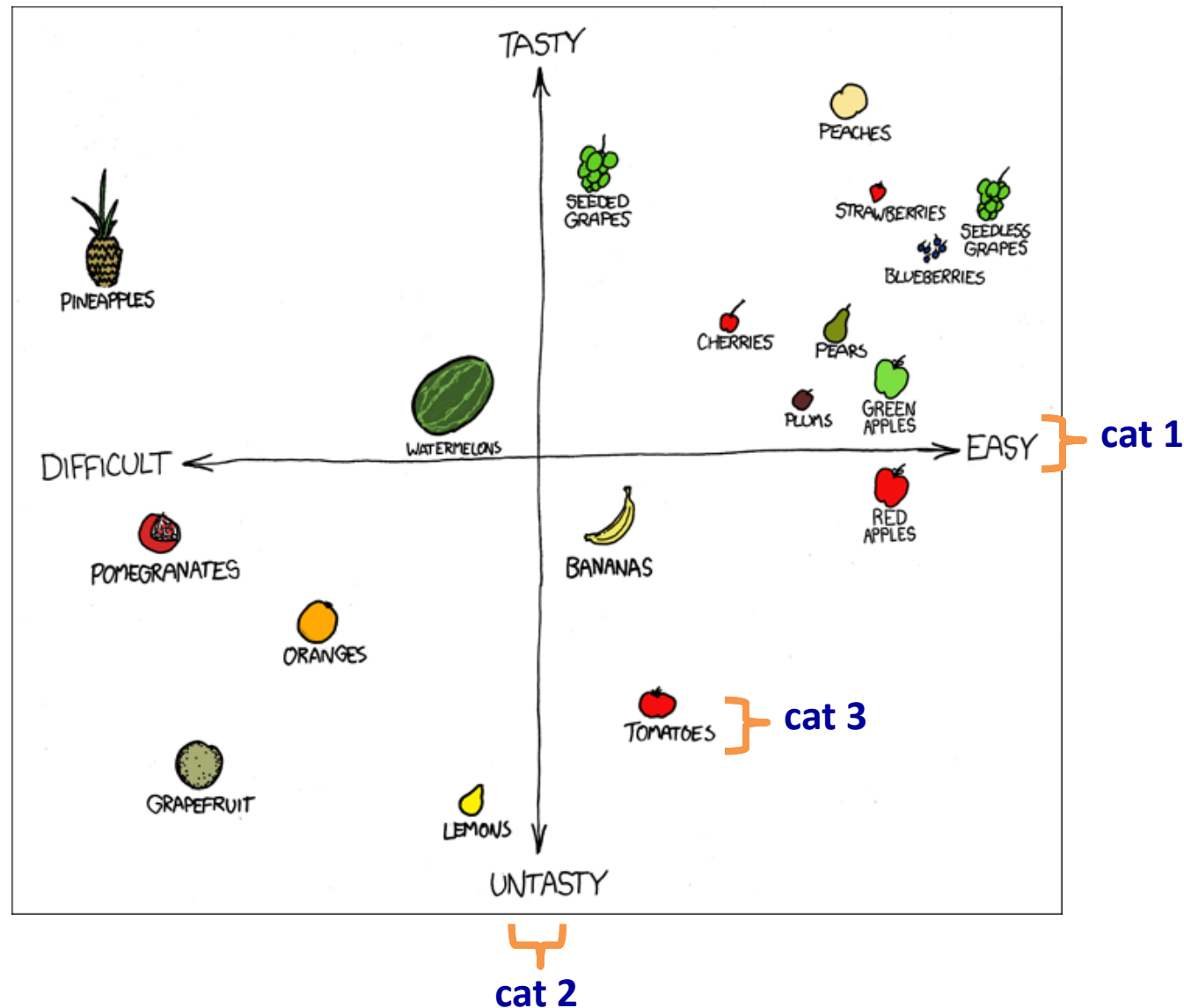
Data: 3 variables (2 quantitative, 1 categorical)

scatterplot



Story: **comparison**
Data: **3 categorical variables**

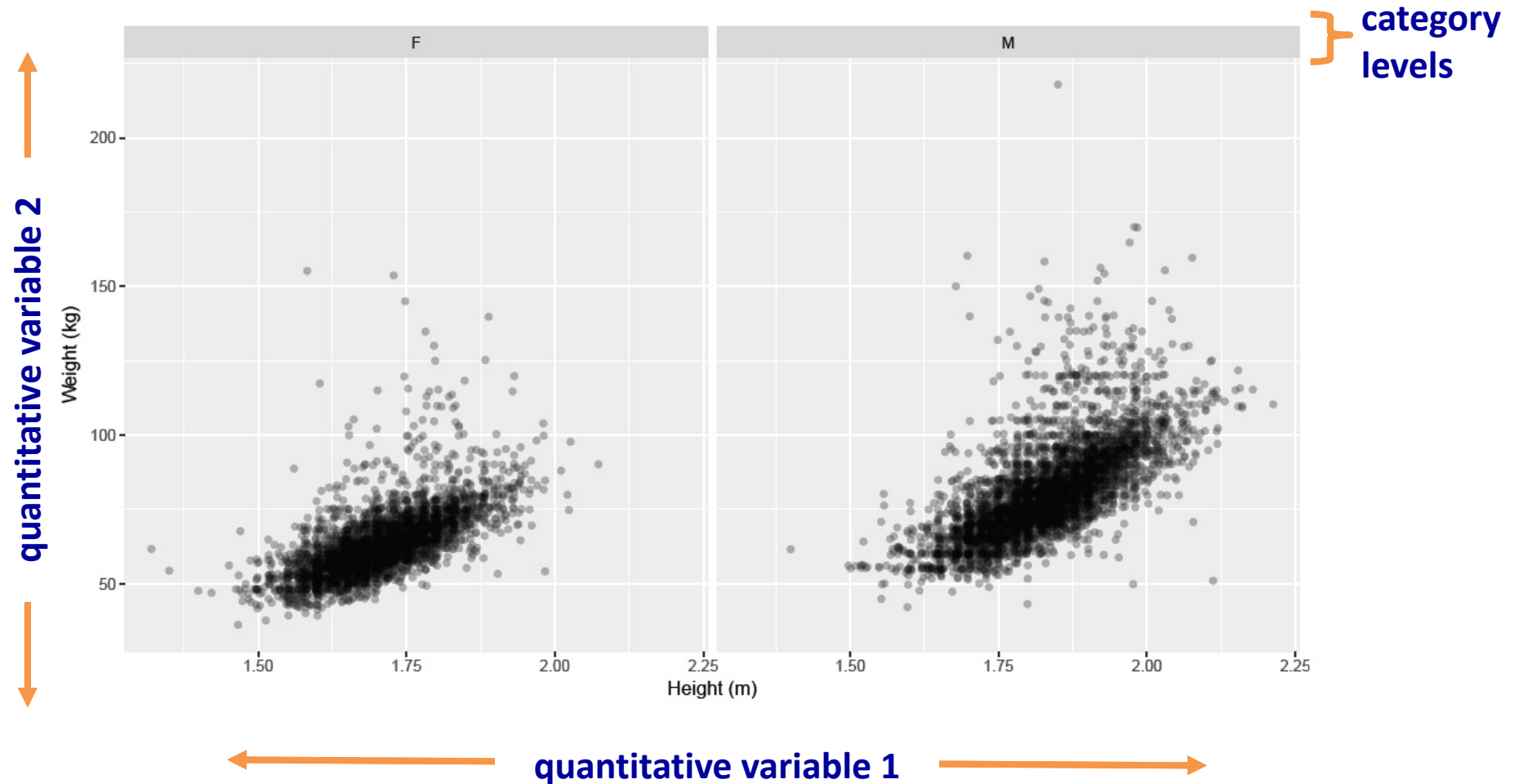
scatterplot-like



Story: **correlation & comparison**

Data: **3 variables (2 quantitative, 1 categorical)**

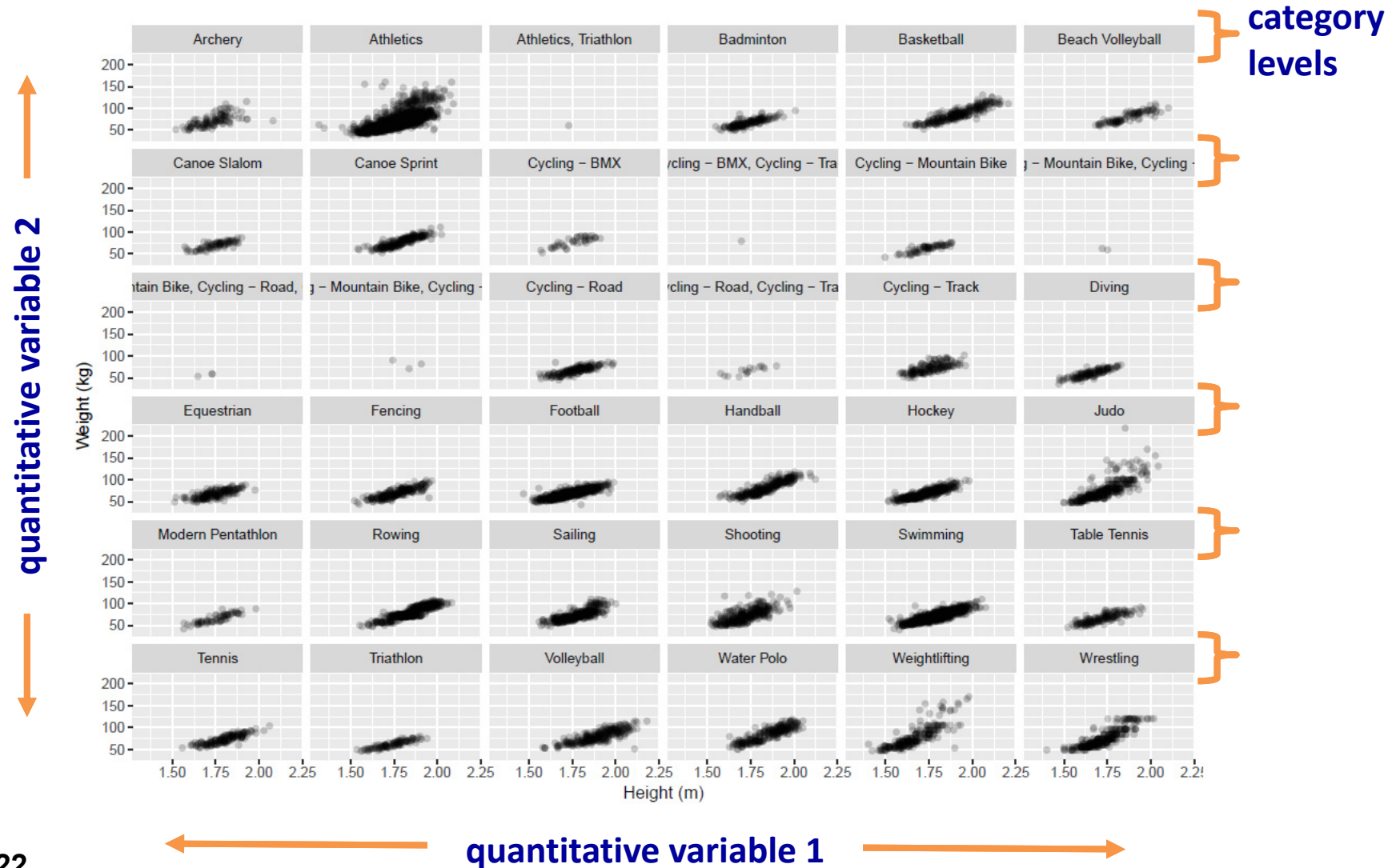
scatterplot
multi-panel



Story: correlation & comparison

Data: 3 variables (2 quantitative, 1 categorical)

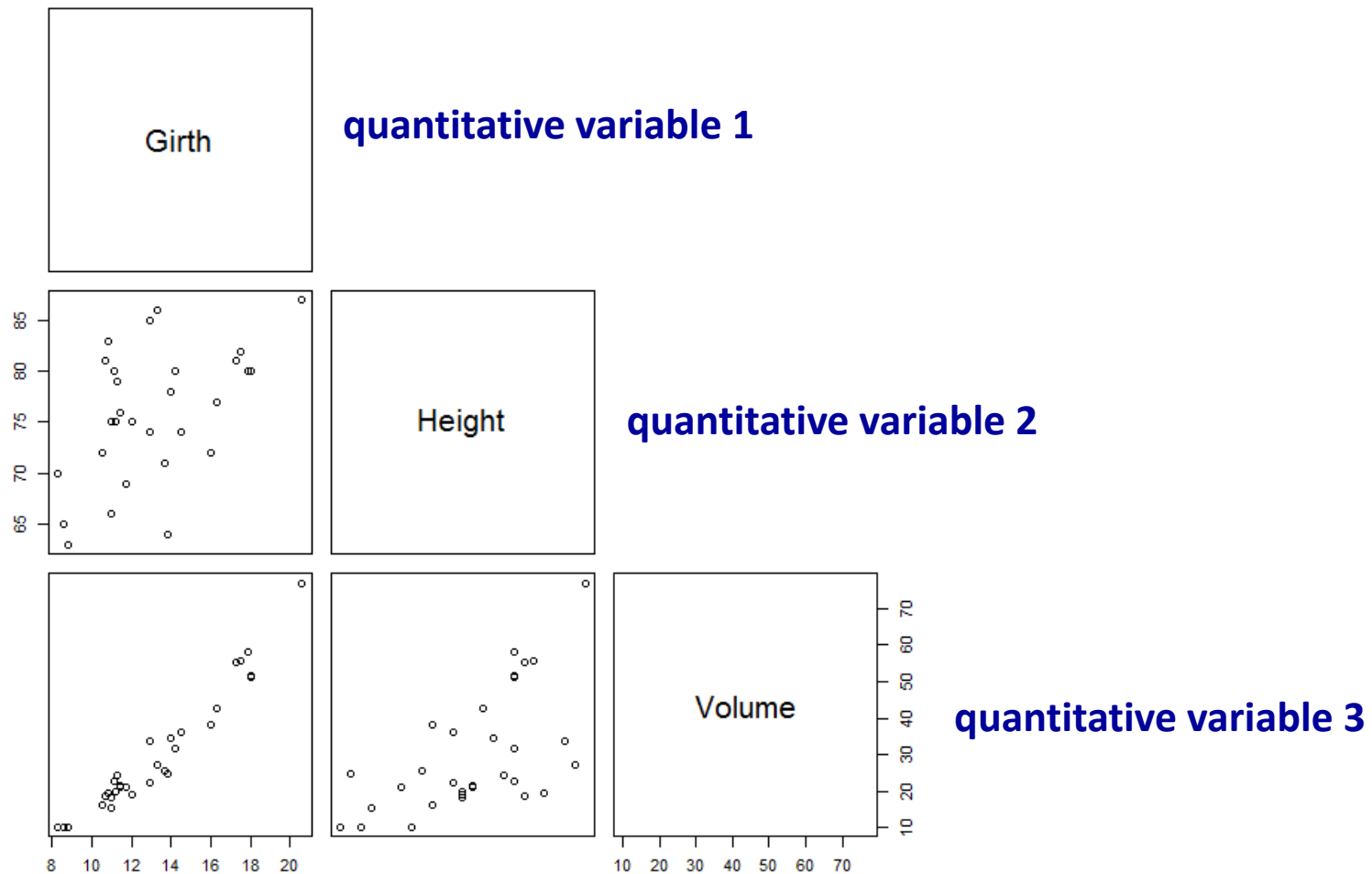
scatterplot
small multiples



Story: correlation
Data: 3 quantitative variables

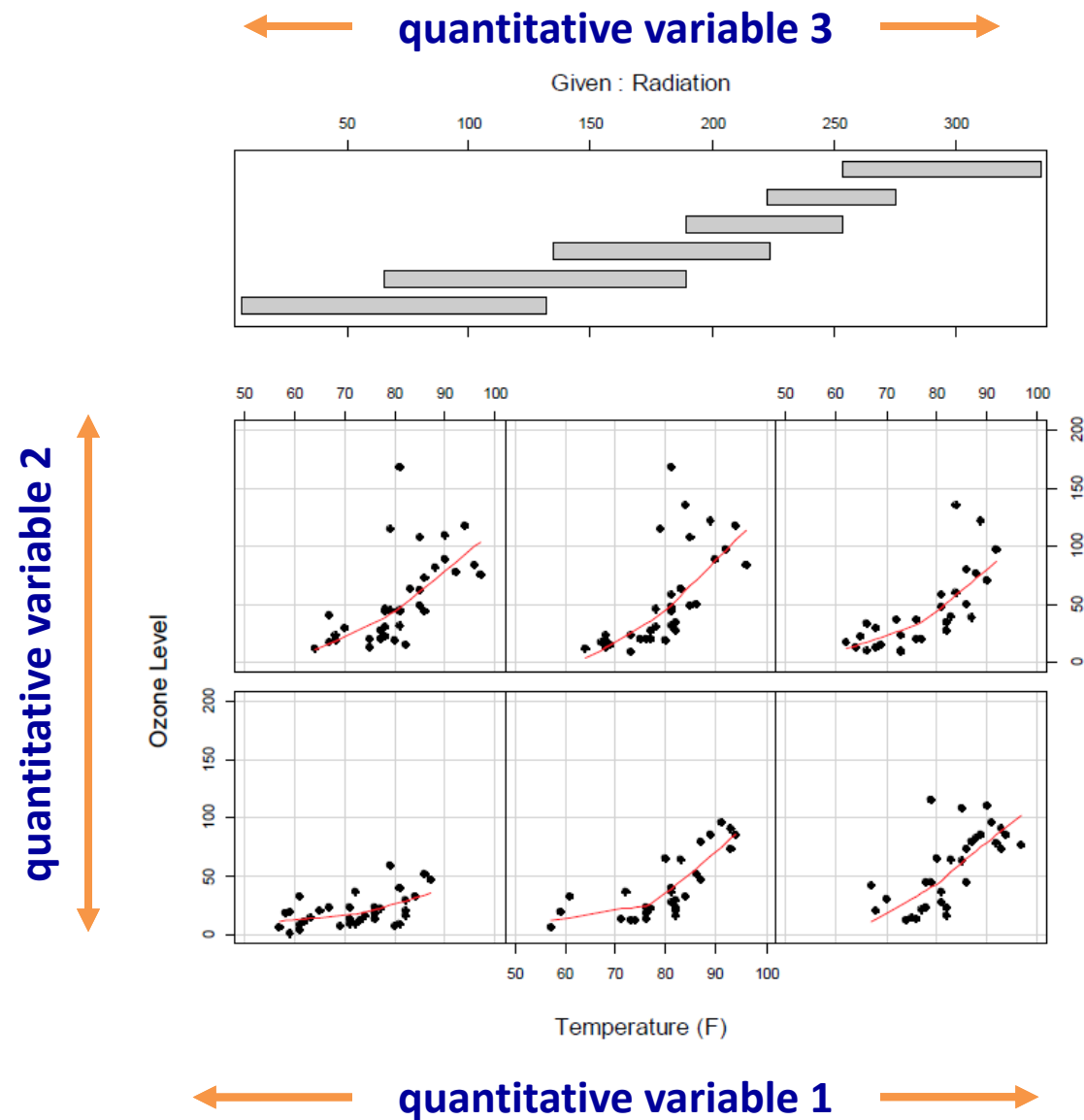
scatterplot matrix

Correlations, one pair of variables at a time. Suitable for any N quantitative variables.



Story: **correlation**
Data: **3 quantitative variables**

co-plot



story: evolution over time

Story: **evolution**
Data: **2 variables (time, 1 quantity)**

line plot



Story: **evolution & comparison**

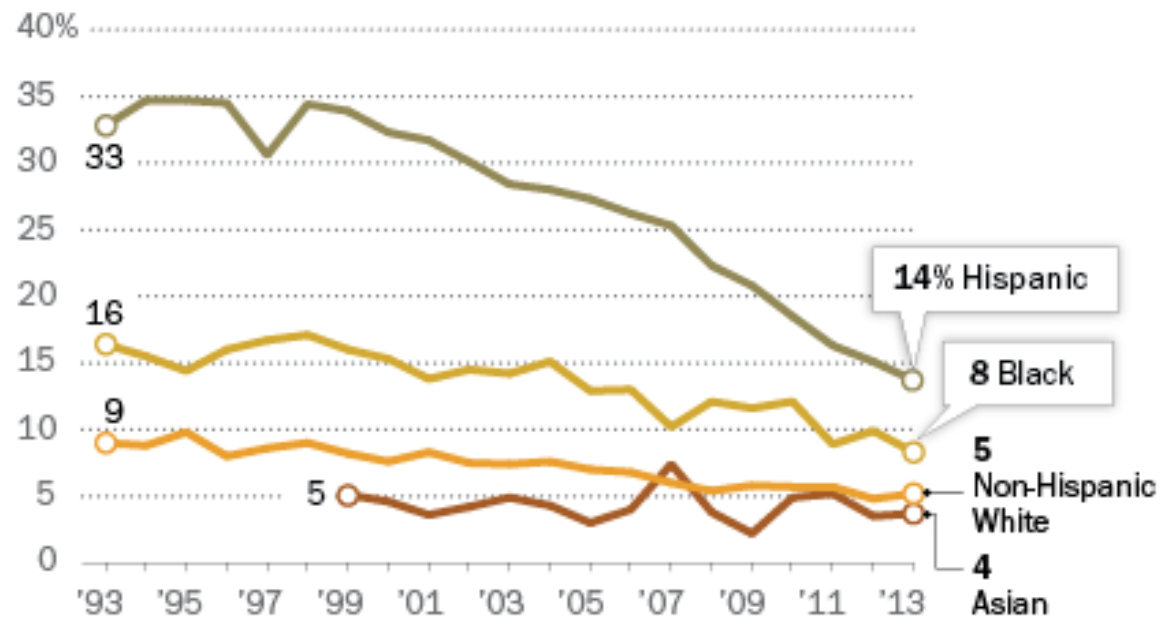
Data: **3 variables (time, 1 quantity, 1 category)**

multi-lines

↑
quantitative variable
↓

Hispanic, Black High School Dropout Rates Reach Record Lows

Percent of 18-to 24-year-olds dropping out of high school



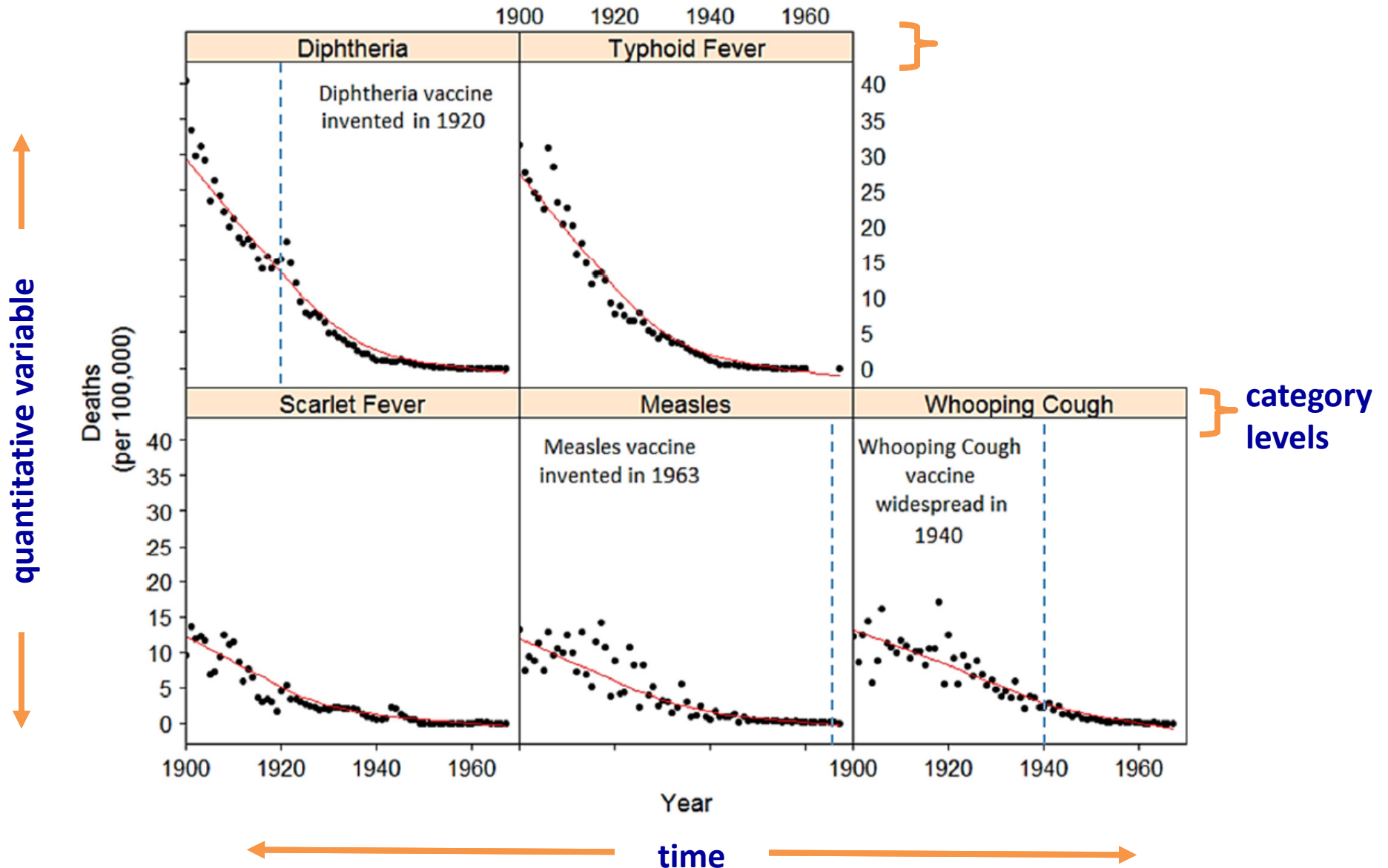
cat 1
levels

← time →

Story: evolution & comparison

Data: 3 variables (time, 1 quantity, 1 category)

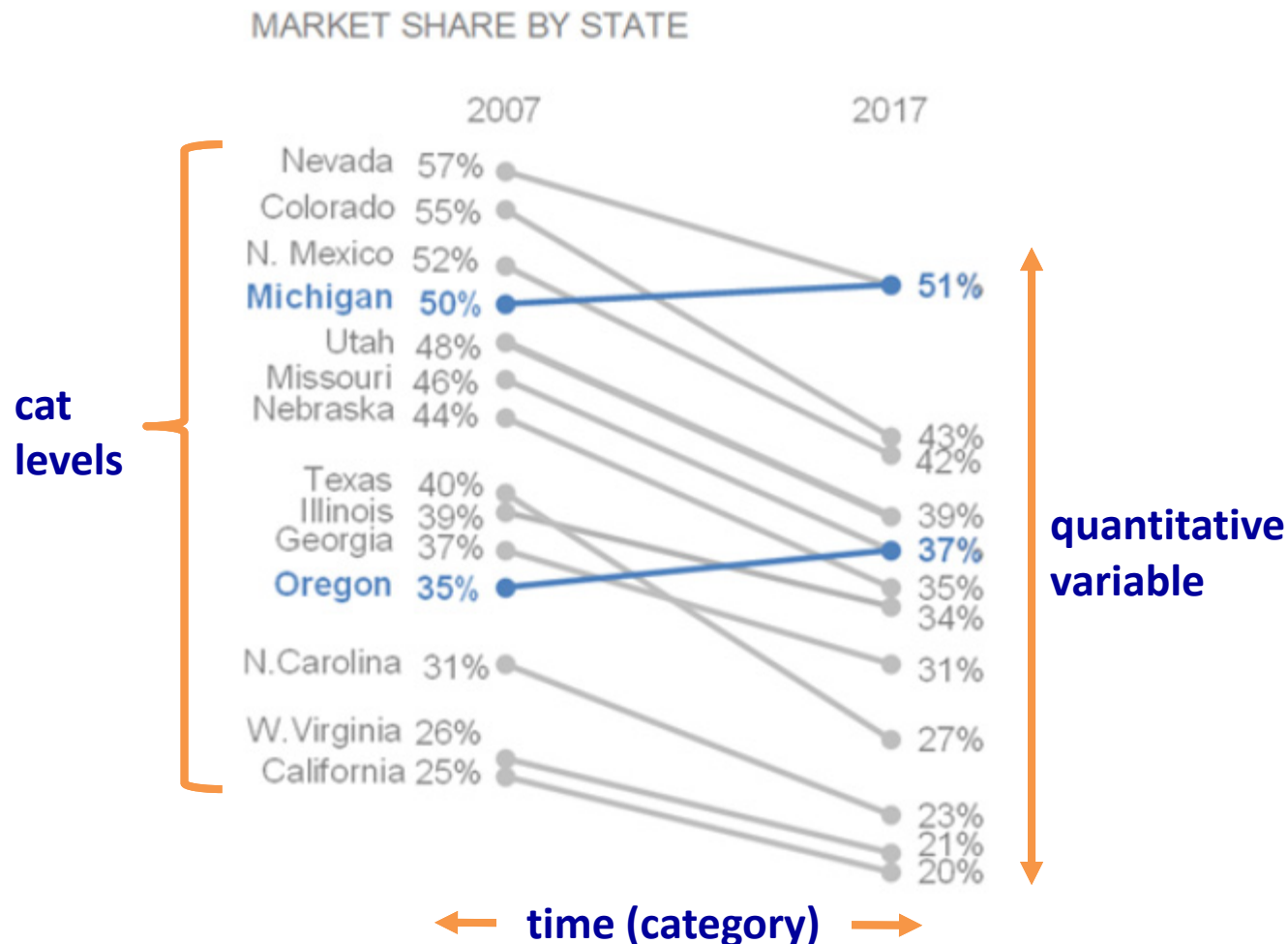
multi-panels



Story: evolution & comparison

Data: 3 variables (1 quantity, 2 categories)

slopegraph

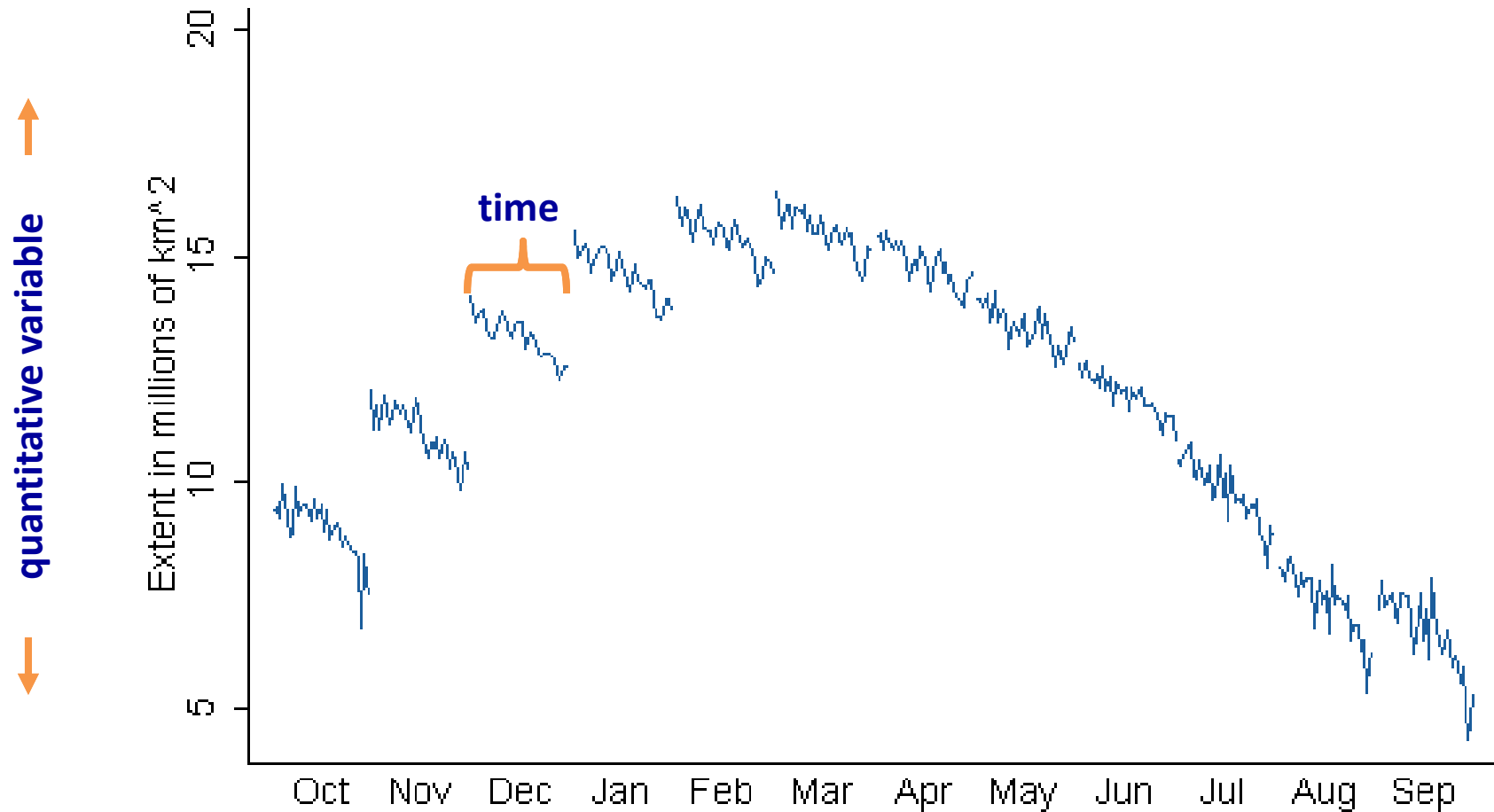


Story: **evolution & comparison**

Data: **3 variables (time, 1 quantity, 1 category)**

cycle plot

Yearly trends in arctic ice by month, 1978–2010



graph: L. Hamilton 7/20/2010

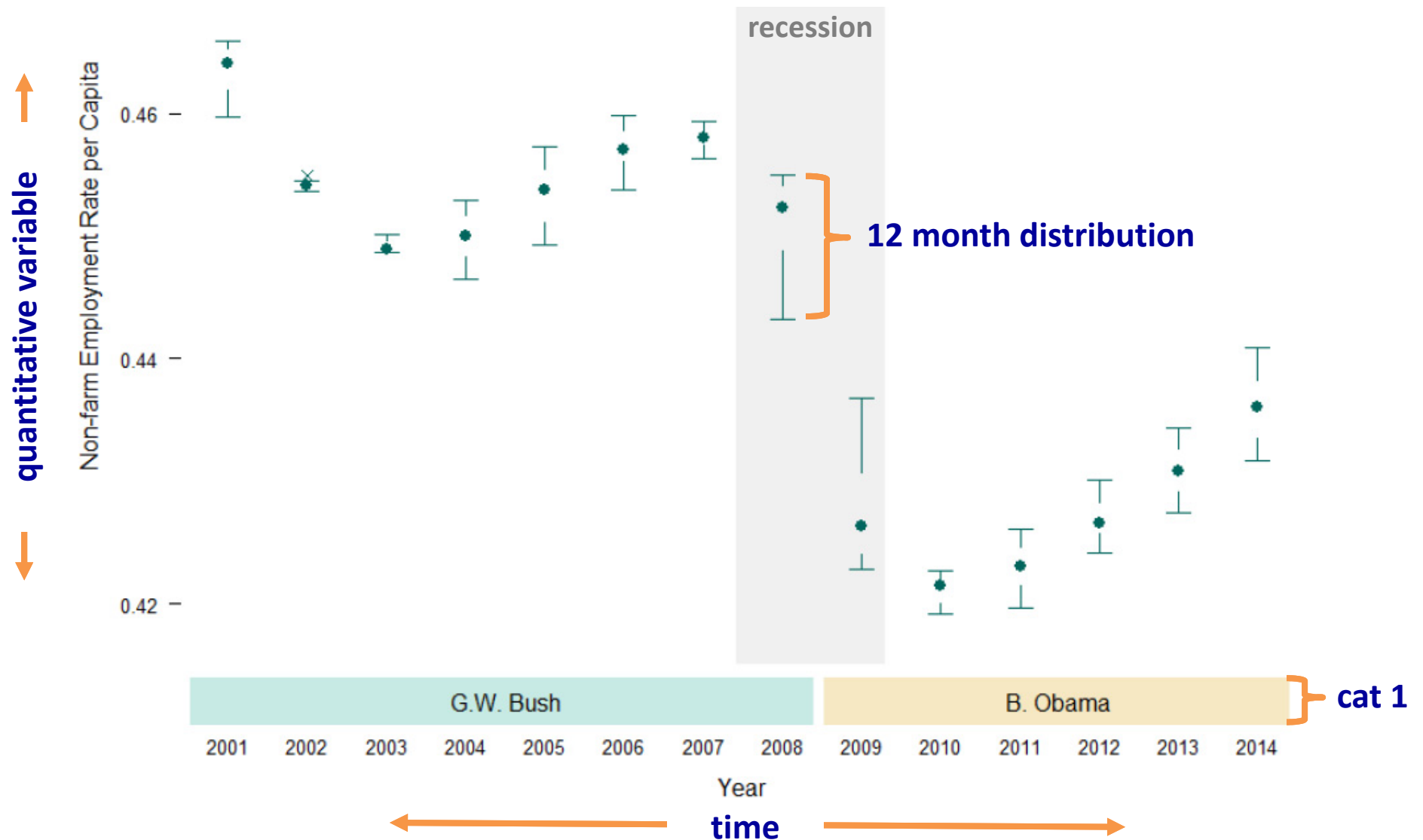
data: NSIDC

← **ordered category** →

Story: evolution of distributions

Data: 3 variables (time, 1 quantity, 1 category)

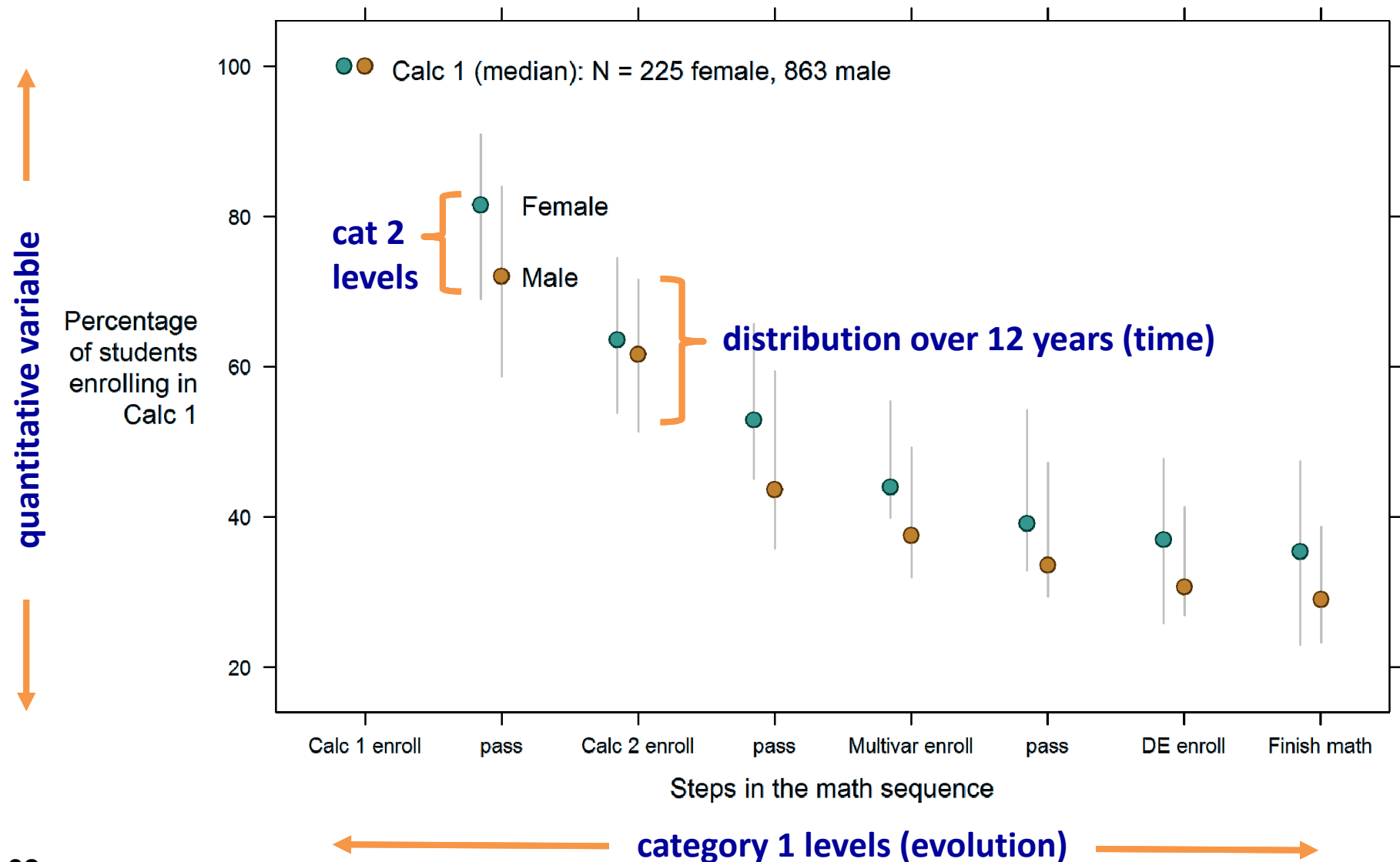
combination
evolution of boxplots



Story: evolution of distributions

Data: 3 variables (1 quantity, 2 category)

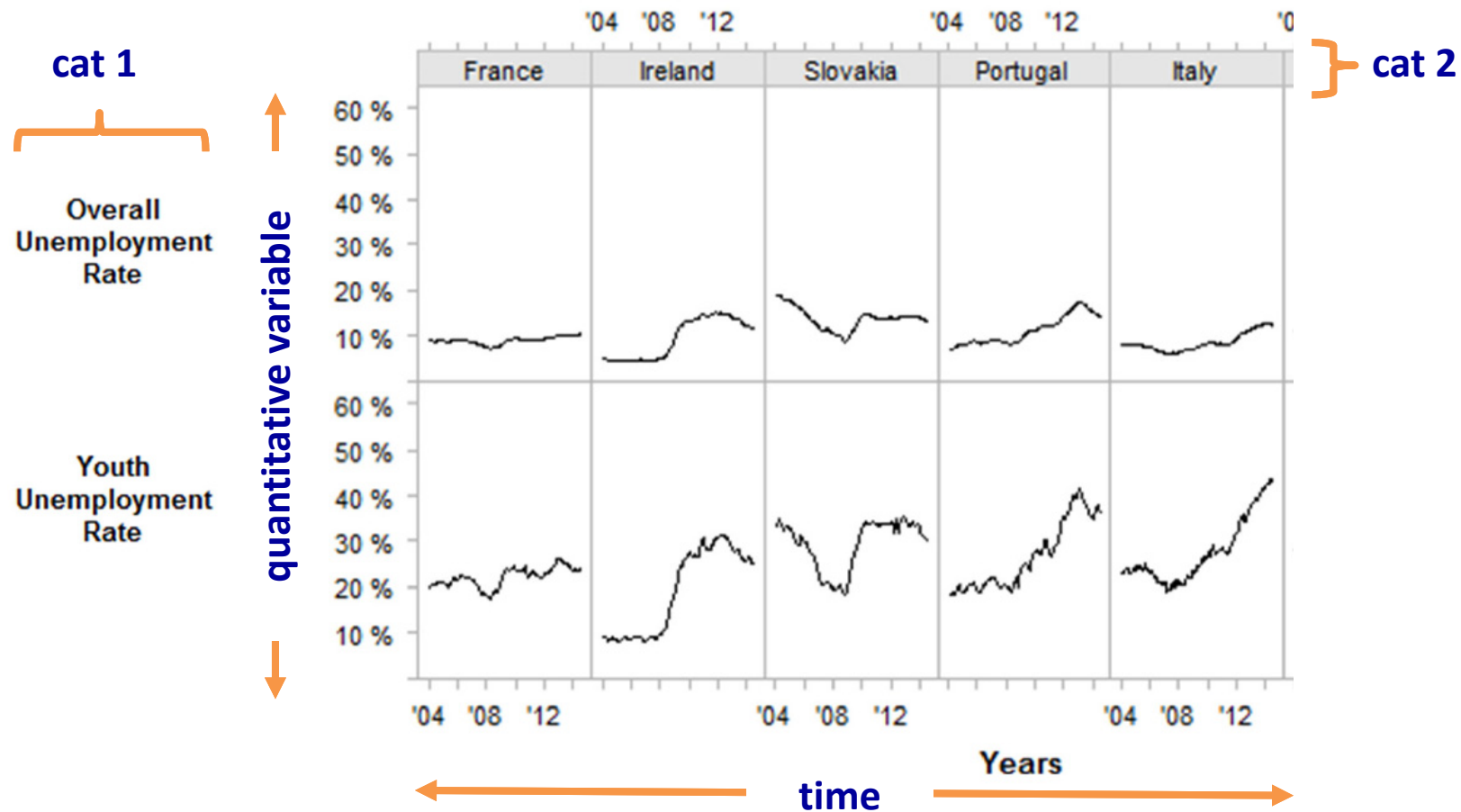
combination
evolution of boxplots



Story: evolution & comparison

Data: 4 variables (time, 1 quantity, 2 categories)

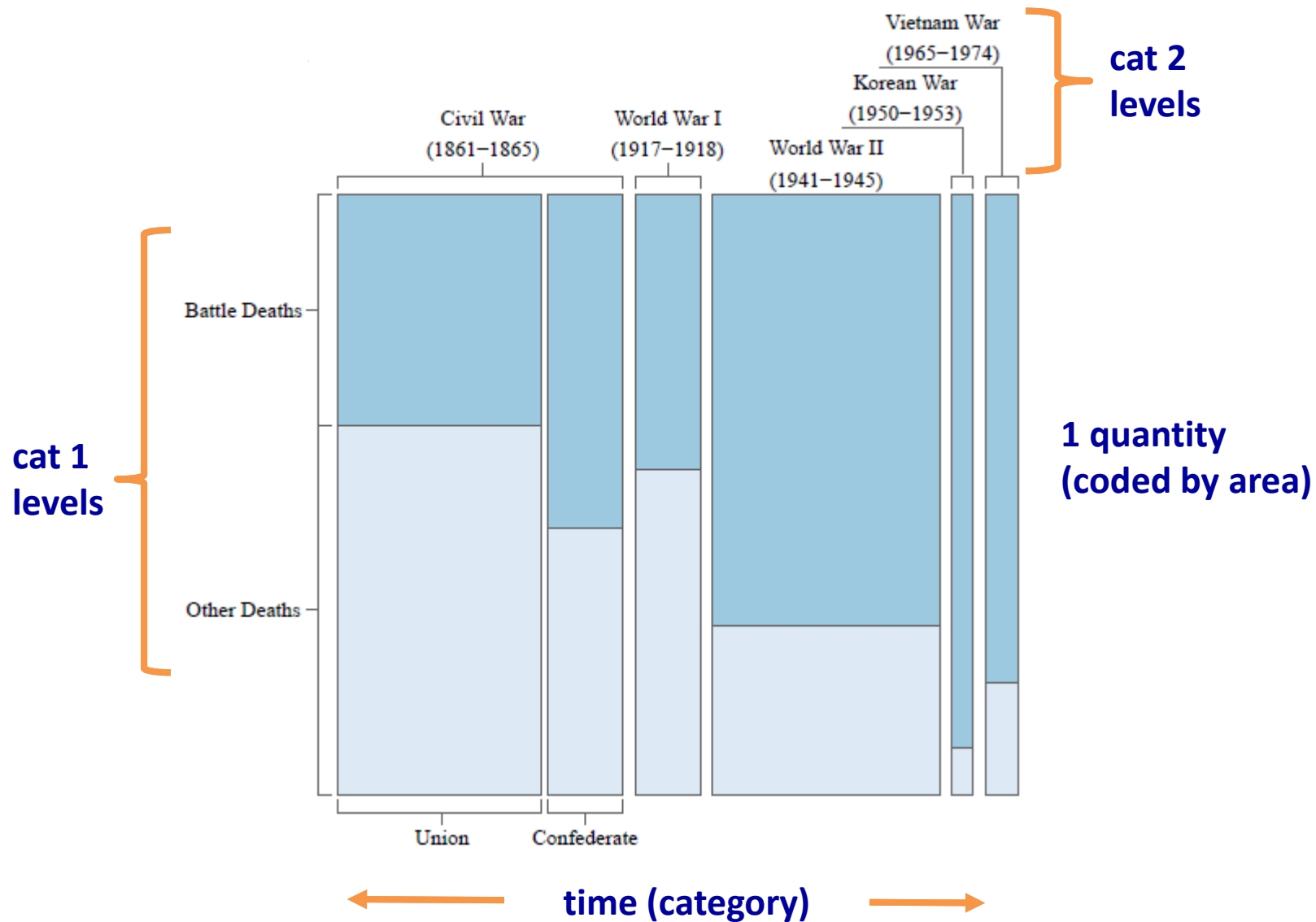
line plot
small multiples



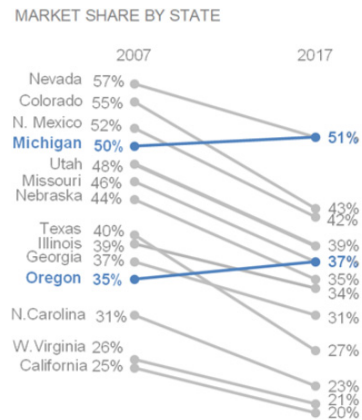
Story: evolution of frequency

Data: 4 variables (1 quantity, 3 categories)

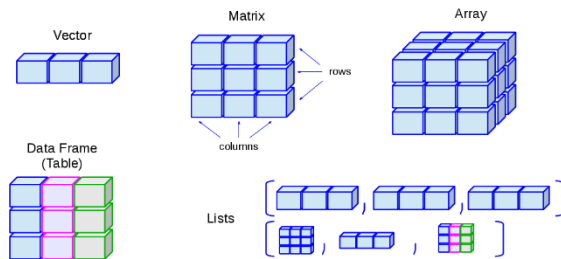
mosaic



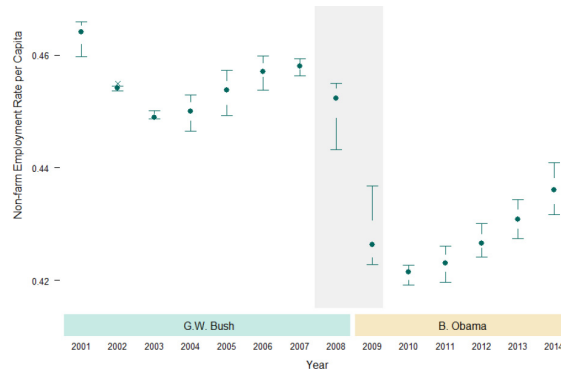
Implications for the designer



Find a story



Grasp the data structure



Choose a graph that suits both