

CSSS 569 · Visualizing Data

GALLERY 1: SCALES AND STORYTELLING

Christopher Adolph

Department of Political Science

and

Center for Statistics and the Social Sciences

University of Washington, Seattle



Gallery

Sociology: Comparative Assault Death Statistics

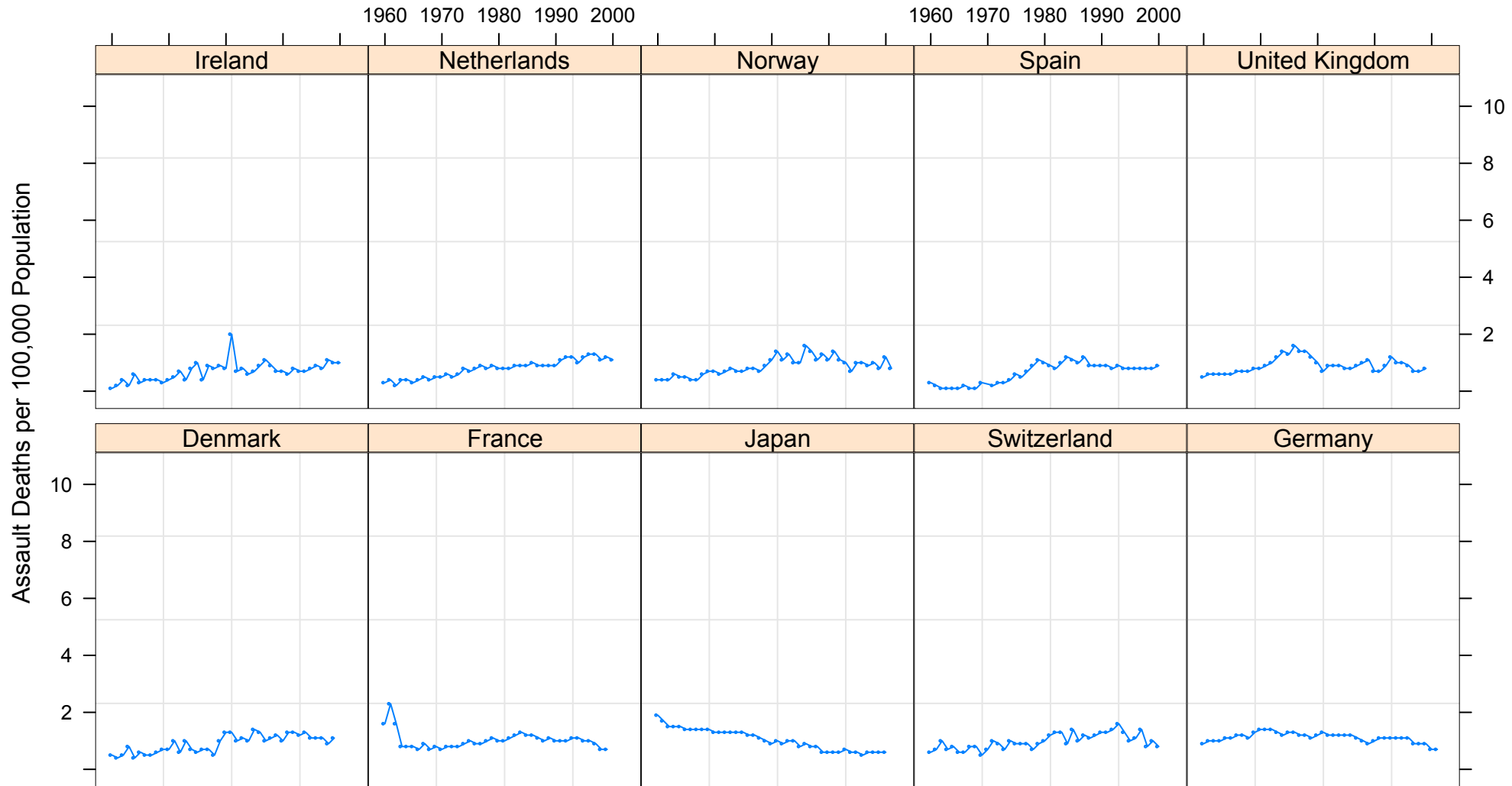
Astronomy: Detecting distant planets

Healy: Comparative Assault Deaths

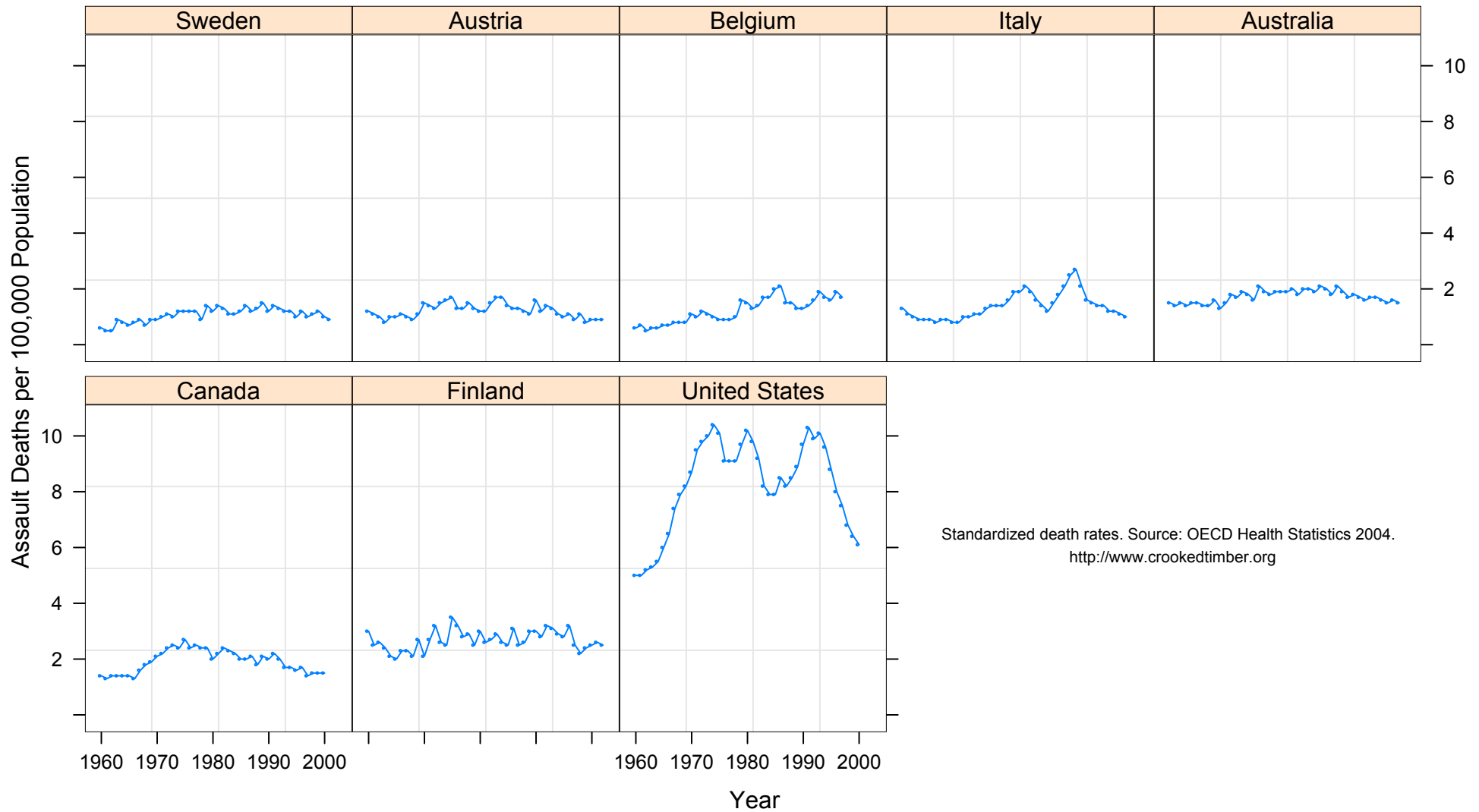
Kieran Healy, sociologist and blogger, wondered how the mean and variance of homicide rates varied across countries

To answer his questions, he made the following plots (probably in R)

Healy: Comparative Assault Deaths



Healy: Comparative Assault Deaths



Healy: Comparative Assault Deaths

Carefully designed.

Which design choices affected the viewer's impression?

- Consistent axes, leading to (dramatically useful) white space
- US listed last
- Easily understood axes units, chosen to maximize comparative value (e.g., not total deaths)
- Faint gridlines for comparison of shapes *across* plots.
Faint lines don't hide the data.

Healy: Comparative Assault Deaths

Questions raised?

1. The obvious: why is the US exceptional?
2. Less obviously: does volatility vary across countries?
May need extra plots with *different* scales
3. How are the data standardized?
(how are international differences, e.g., in the definition of assault handled?)
4. What covariates would we like to see? How could they be incorporated?
 - Color code/sort countries with gun control
 - Shade periods of economic recession
 - Other variables?

Finding distant planets

In recent years, astronomers have detected planets in nearby solar systems.

Generally, these planets have been huge (i.e., Jupiter or larger)

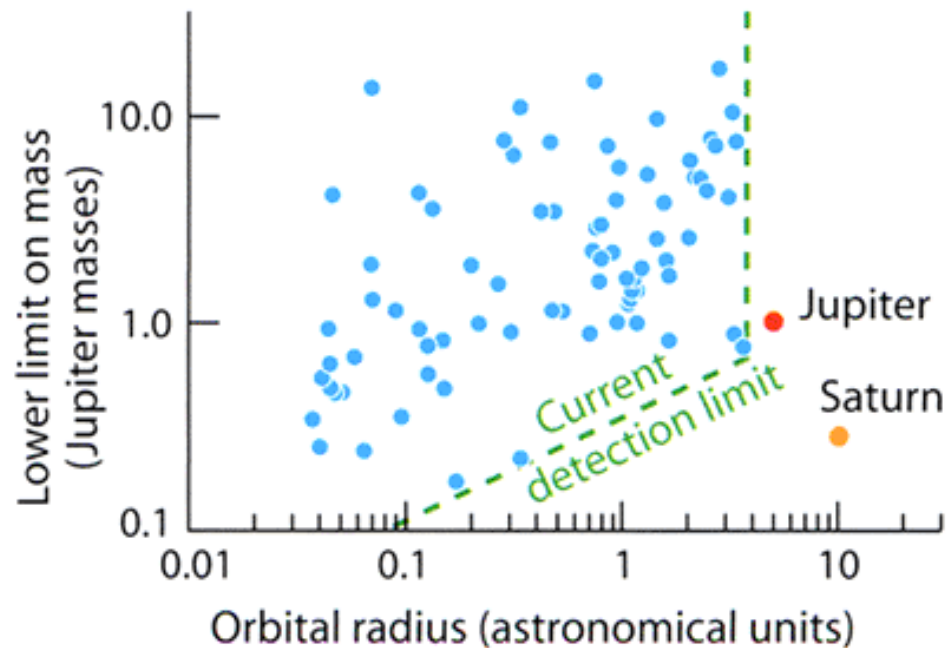
Why?

Finding distant planets

In recent years, astronomers have detected planets in nearby solar systems.

Generally, these planets have been huge (i.e., Jupiter or larger)

Why? A graphic explains:



The planets discovered so far around main-sequence stars come in a wide range of masses and orbit sizes, with little or no correlation between the two. Astronomers' ability to detect extrasolar planets is limited by two factors: the tiny changes in a star's Doppler shift (the diagonal part of the dashed line) and the size of a planet's orbit (vertical part). Stars with planets in larger orbits simply have not been observed for long enough yet, a problem that will fix itself with the passage of time.

Source: *Sky & Telescope*, March 2002, p.22

Finding distant planets

How does this look different from a social science plot?

- Long caption. Combination is self-explanatory
- Log scaling. Labels are readable in original units
Labels are $(0.01, 0.1, 1, 10)$, not $(-2, -1, 0, 1)$
- Color (alas, we seldom can afford it)

What statistical problem affects these data? Heteroskedasticity

Finding distant planets

What could be improved?

- No small planets shown as reference (e.g., Earth)
- Most exciting discovery of all would be a habitable terrestrial planet; shade region
- Terrestrial planets $\ll 0.1$ Jupiter masses. Jupiter is $317\times$ bigger than earth.
- Y-axis needs to be rescaled
- Ultimate message is depressing; can any earth-like planets be detected?