Session V: What would Edward Tufte say about the Charts & Graphs coming out of Washington

Helping audiences understand data with graphics, maps, visualizations, & interactive web products

Vicki Lancaster, Ph.D.
Biocomplexity Institute & Initiative
Social Decision & Analytics Division
val7zv@virginia.edu

UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE

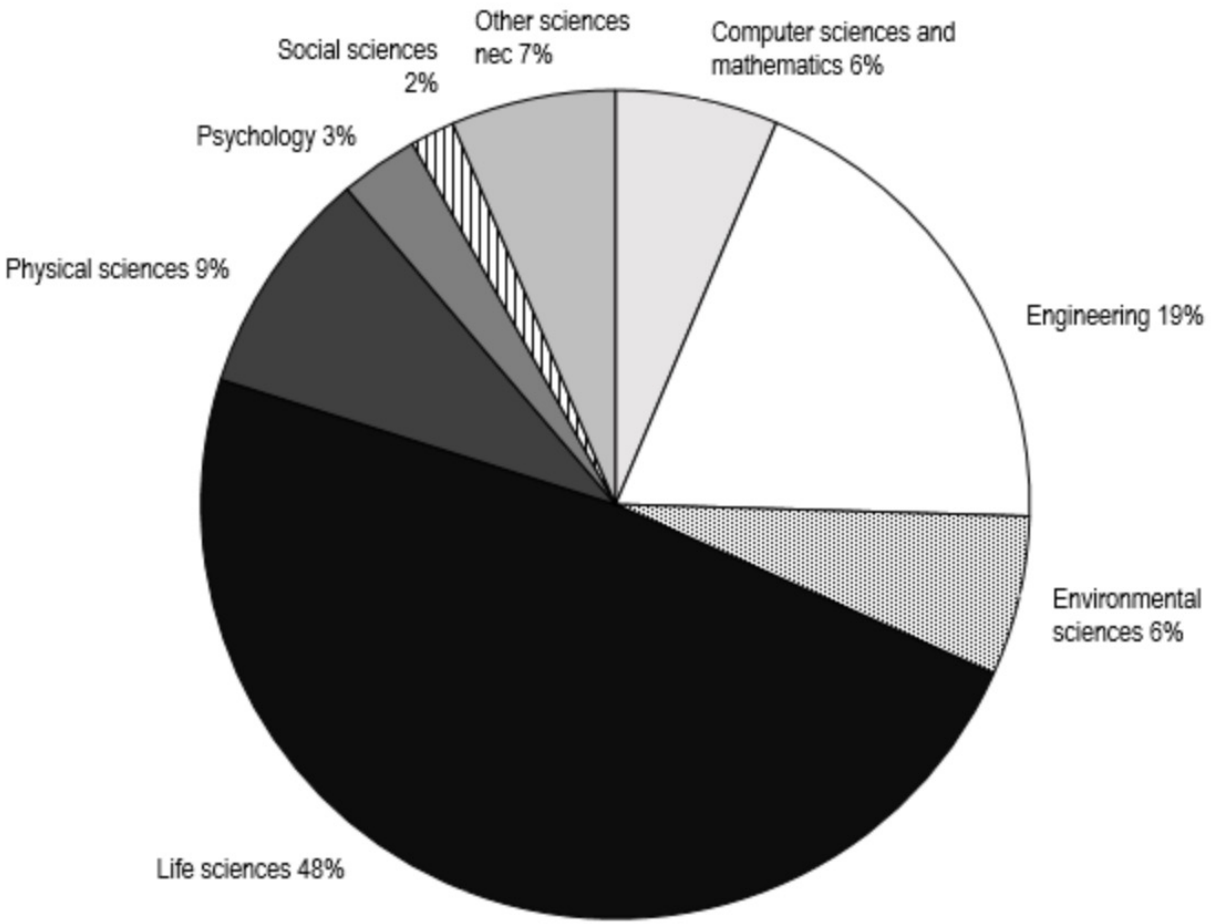
Introduction

This is a project is being done with May Ayden at the National Center for Science and Engineering Statistics that explores how to "better" represent current NCSES visualizations.

Visualization Criteria:

- contain statistical content that highlights NCSES or linked data,
- can provide additional statistical insights,
- can be reproducible in an automated production environment,
- require minimal cognitive effort to interpret,
- can appeal to a wide audience,
- can be provided to authors as alternative options for consideration in future publications.

1 Categorial Variable: Pie Chart



One Variable

> Nominal: Federal obligations to S&E fields (8 levels)

Comments:

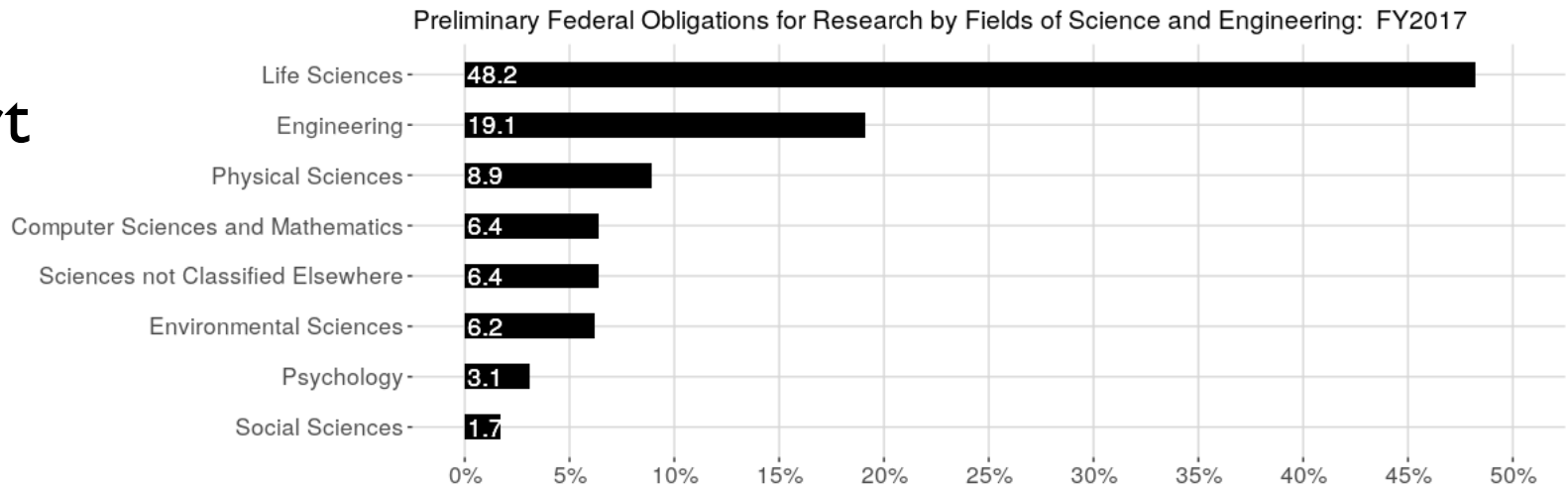
> lack of color & small slices make distinguishing between them difficult

> insufficient number of shades of grey to differentiate between the 8 slices

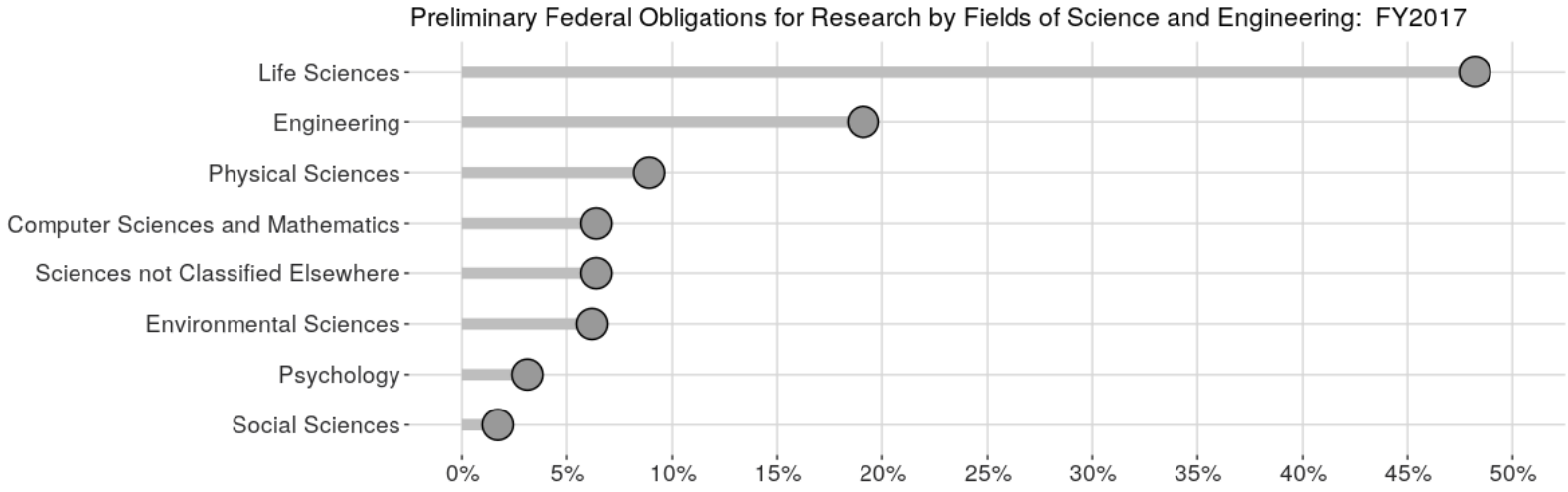
> normally employing a different aesthetic would signal a different variable

Pie Chart Alternatives

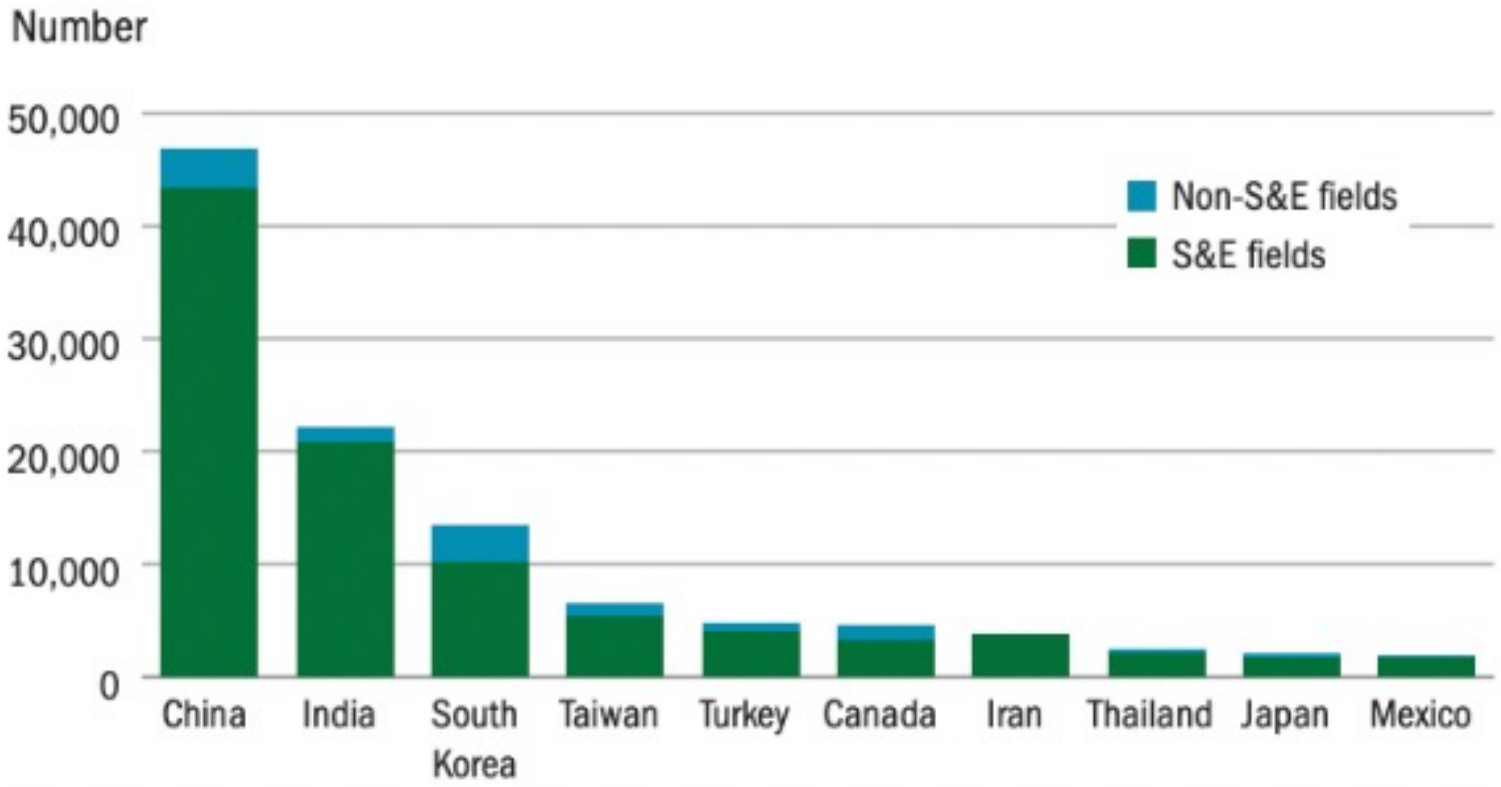
Horizontal Bar Chart



Lollipop Chart



2 Categorial Variables: Simple Stacked Bar Chart



Two Variables:

> Nominal: S&E field (2 levels)

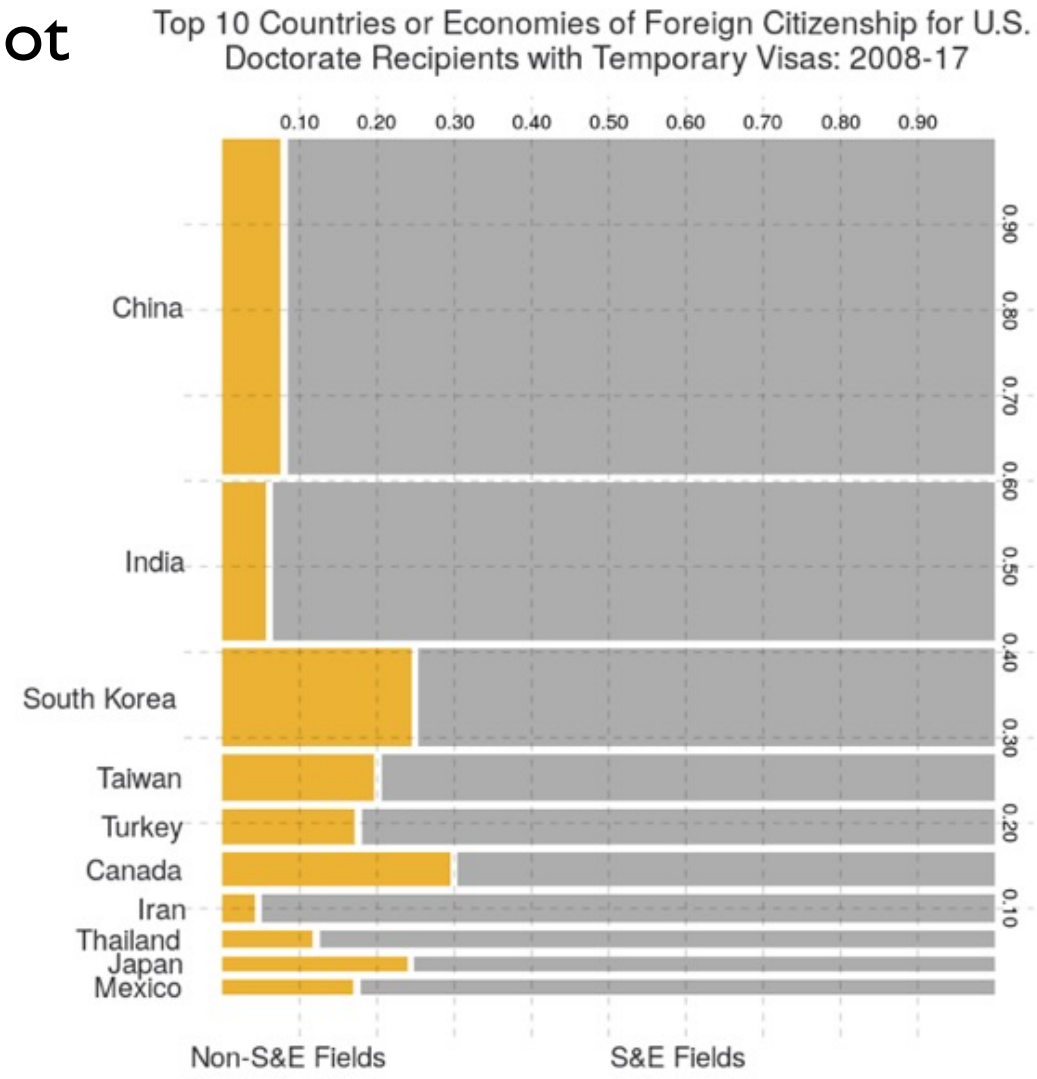
> Nominal: Country (10 levels)

Comments:

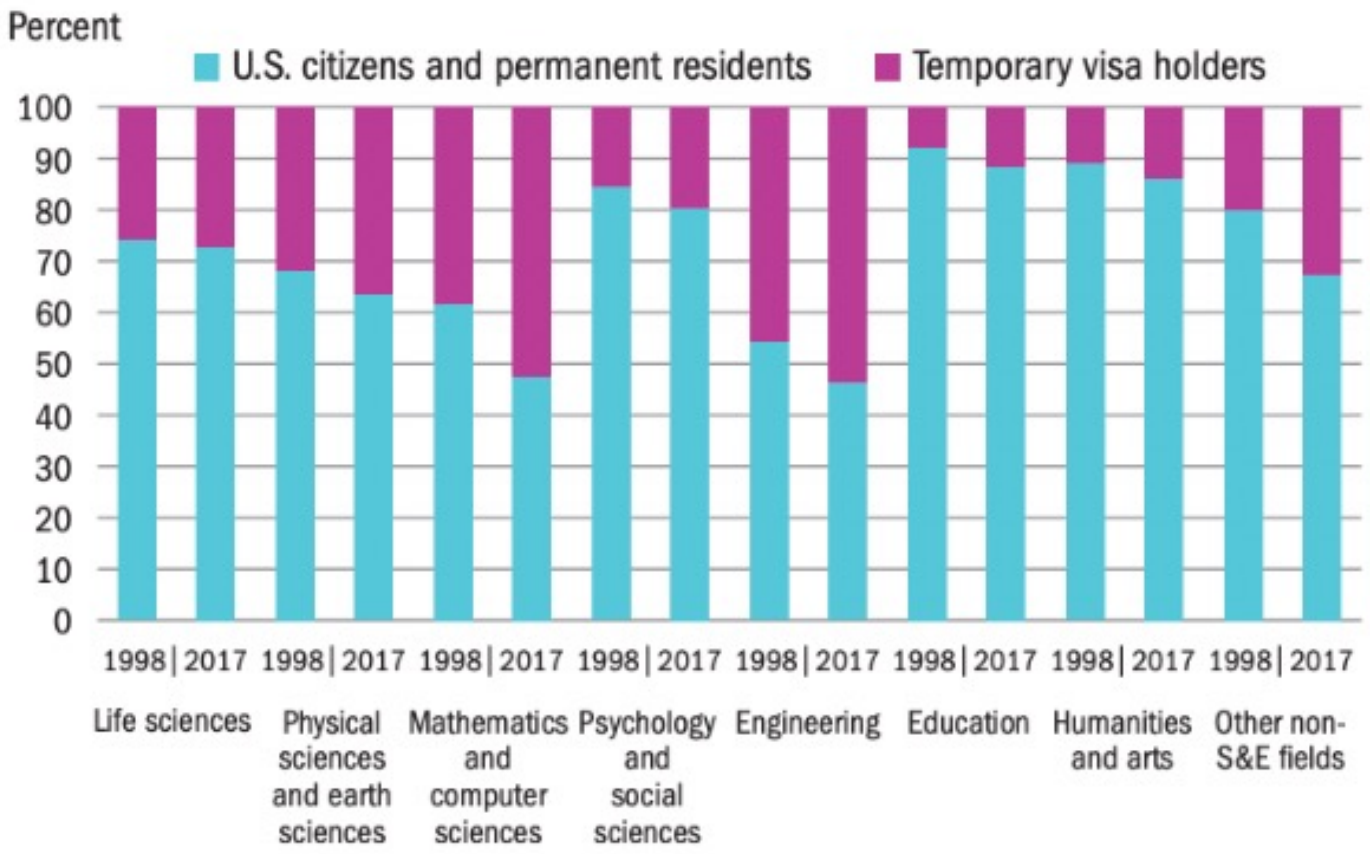
> large differences between heights of the bars make the stacks hard to interpret

Simple Stacked Bar Chart Alternative Visualization

Mosaic or Marimekko Plot



>2 Categorial Variables: Grouped Stacked Bar Chart



Two Variables:

> Nominal: Broad field of study (8 levels)

> Nominal: Race/Ethnicity (5 levels)

> Interval: Time (2 years)

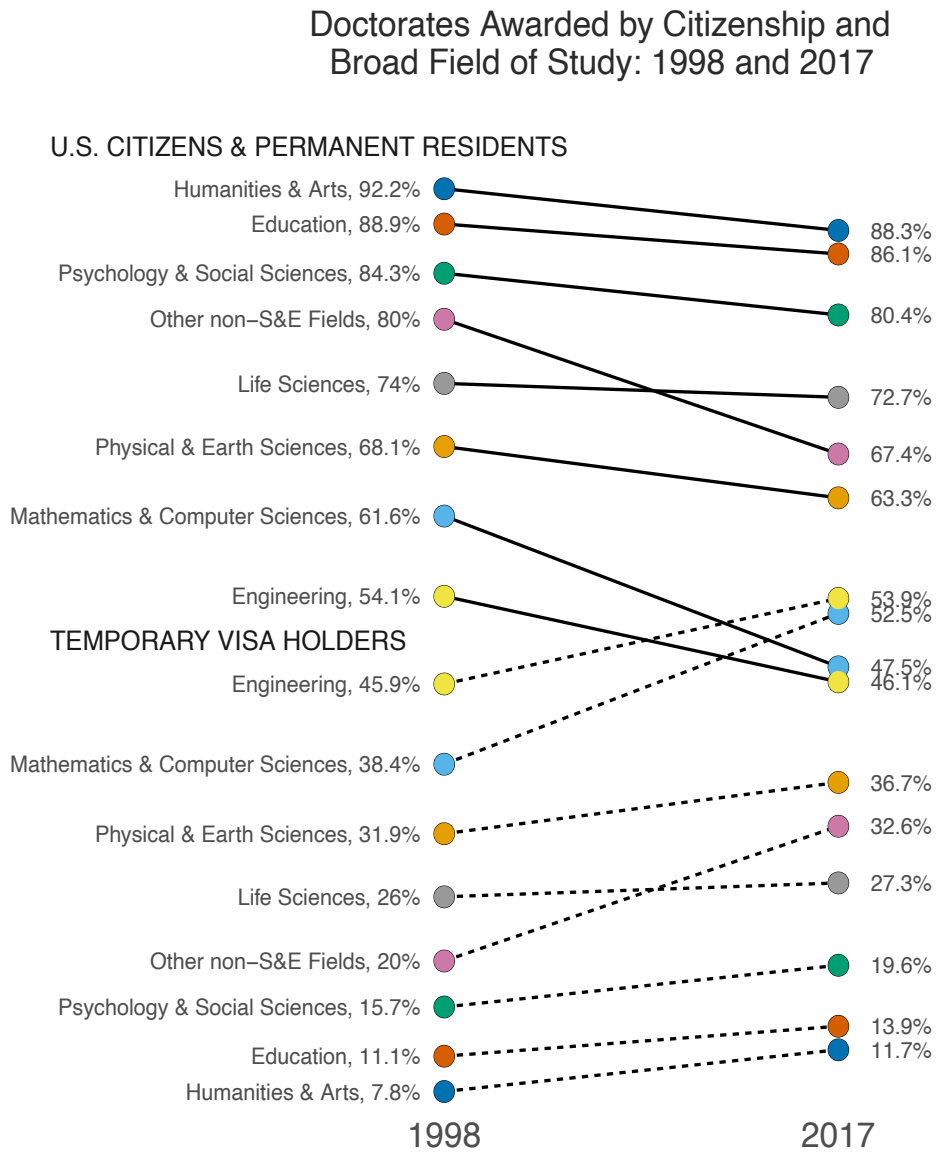
Comments:

> color is used to identify the levels of only one of the three variables

> the figure does not tell the story of the data without exerting a lot of mental energy

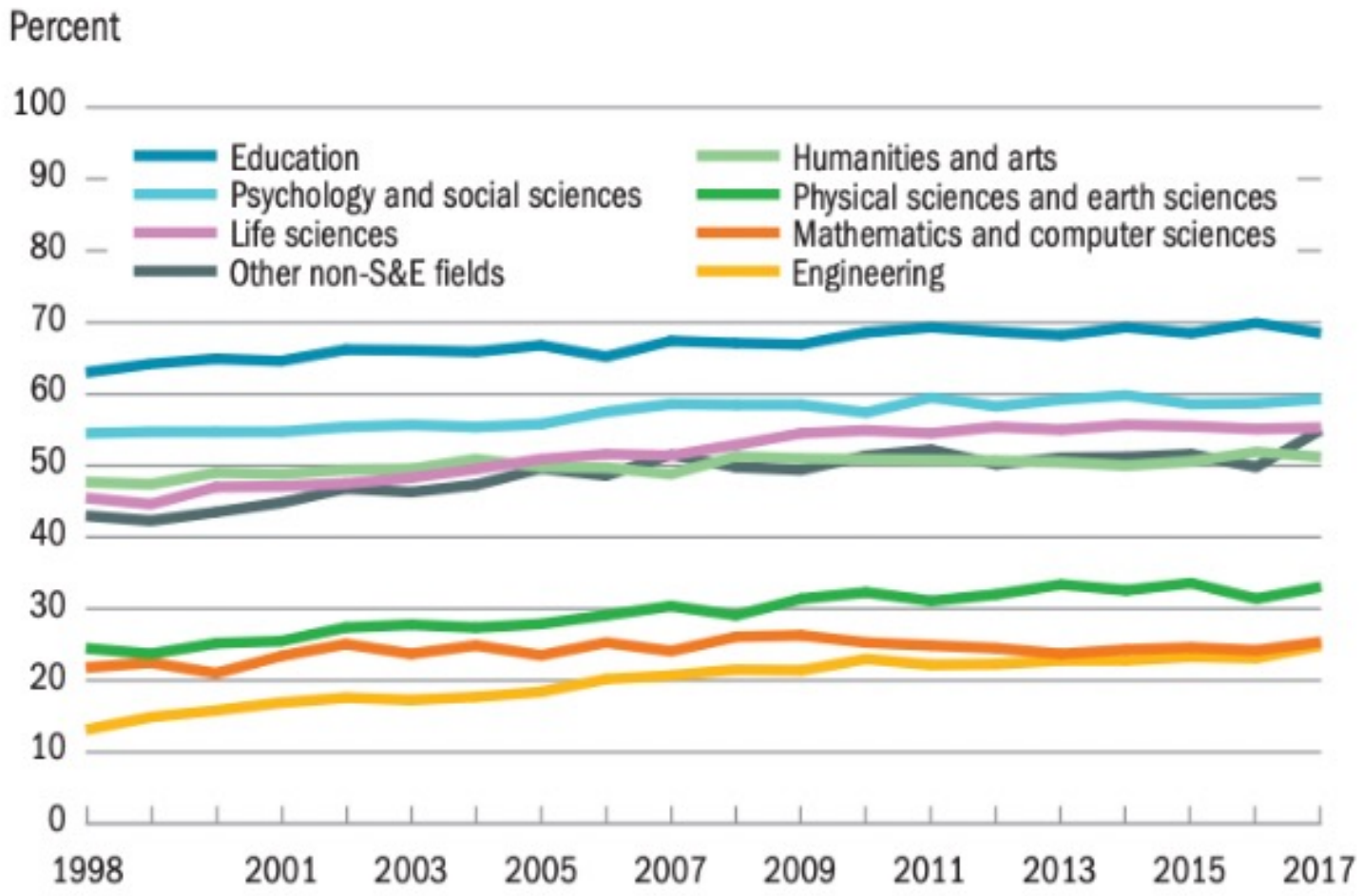
Grouped Stacked Bar Chart Alternative Visualization

Slope Plot



UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE

Longitudinal Variable



Two Variables:

> Nominal: Broad field of study (8 levels)

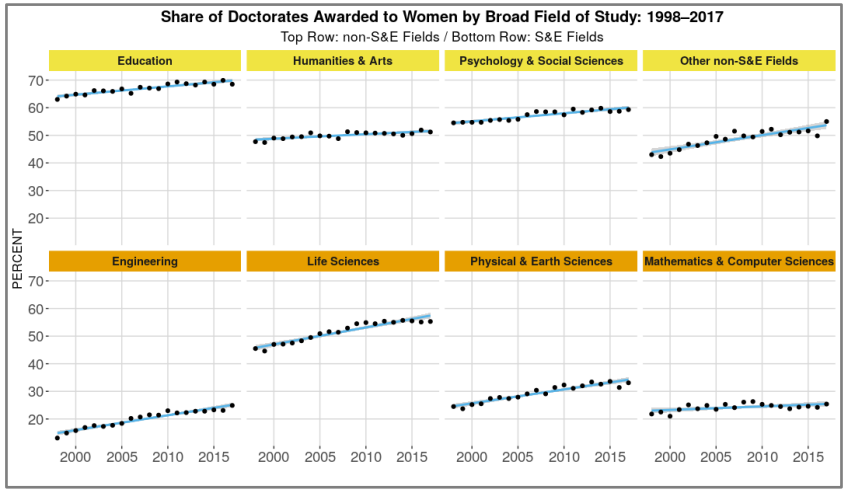
> Interval: Time (19 years)

COMMENTS:

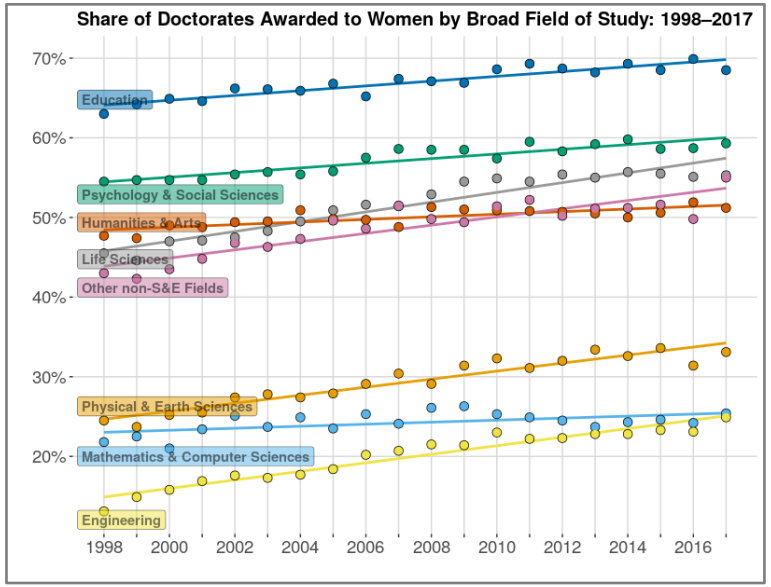
> change the vertical axis from [0-100] to [10-70] to provide greater separation between the lines

Longitudinal Variable Alternative Visualizations

Lattice Scatter Plot with Simple Linear Regression Line

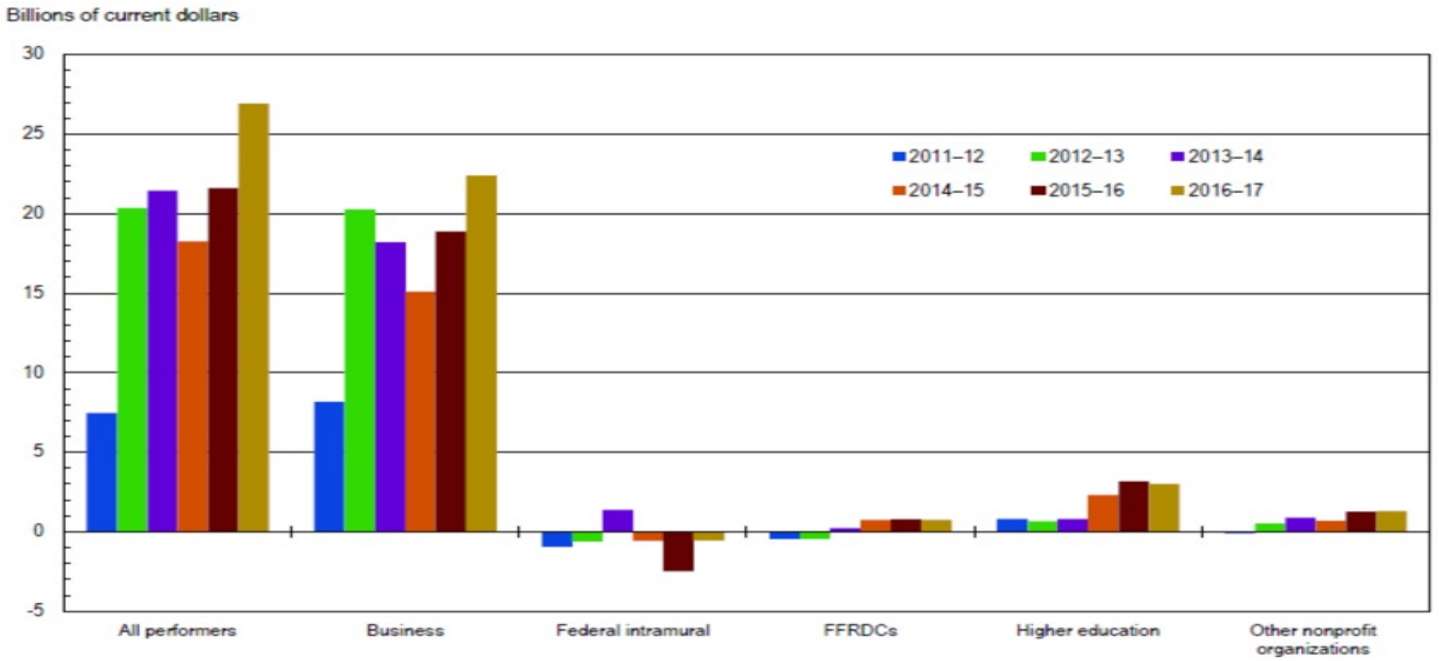


Scatter Plot with Simple Linear Regression Line



UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE

Difference Between Variables: Multi-set Bar Chart



Two Variables:

> Nominal: Performer (5 levels)

> Interval: Time (6 1-year intervals)

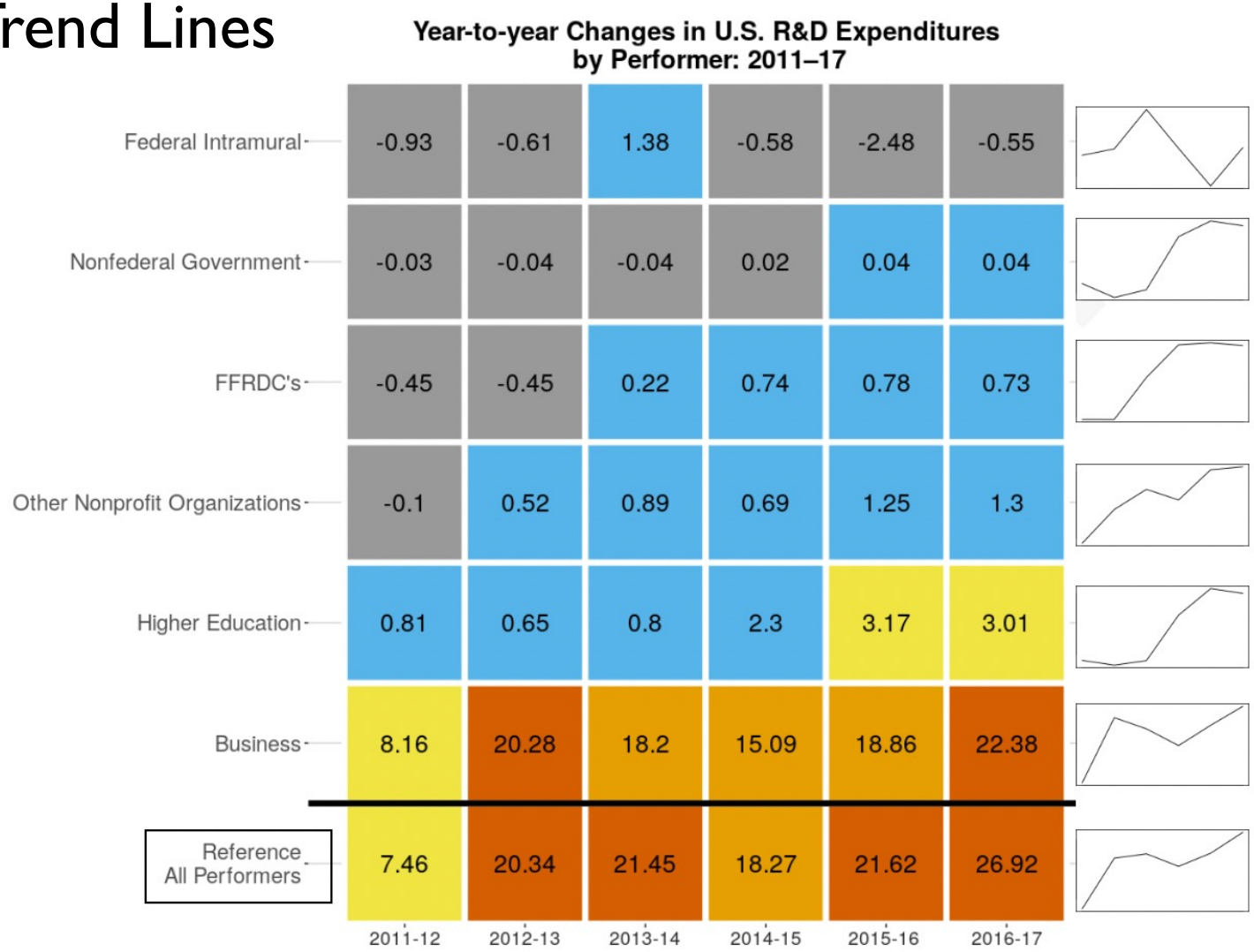
> One reference value: All performers

COMMENTS

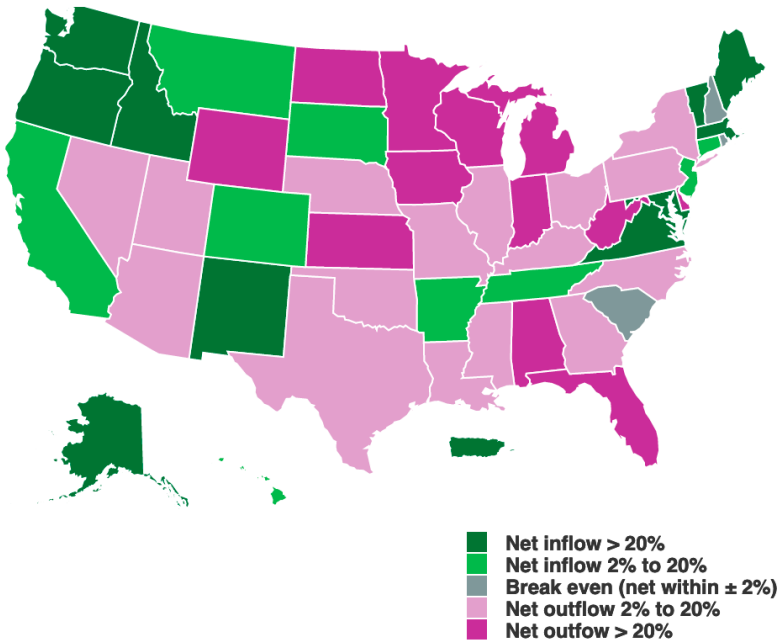
> multi-set bar or grouped bar charts are difficult to interpret especially in cases where the values differ by an order of magnitude

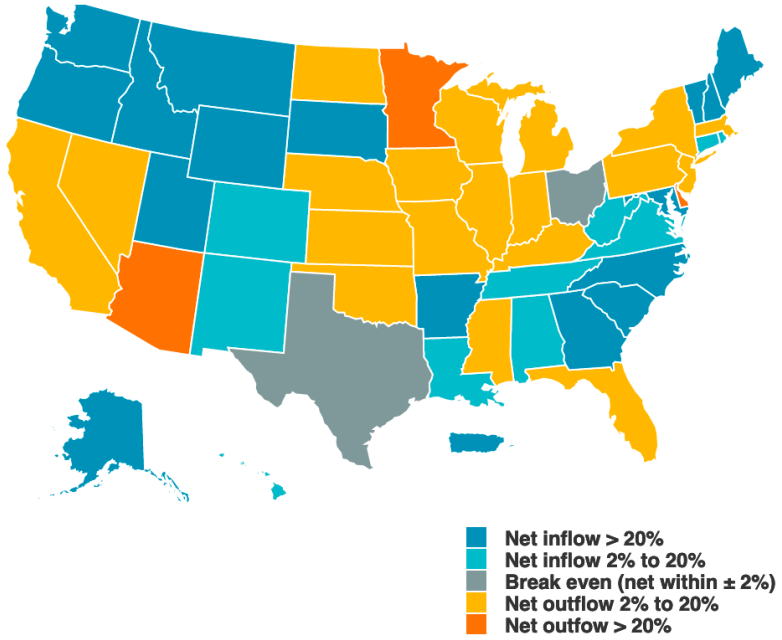
Multi-set Bar Chart Alternative Visualization

Heat Map with Trend Lines



Geospatial Variable: Chloropleth Map





Two Variables:

> Nominal: States + Puerto Rico (51 levels)

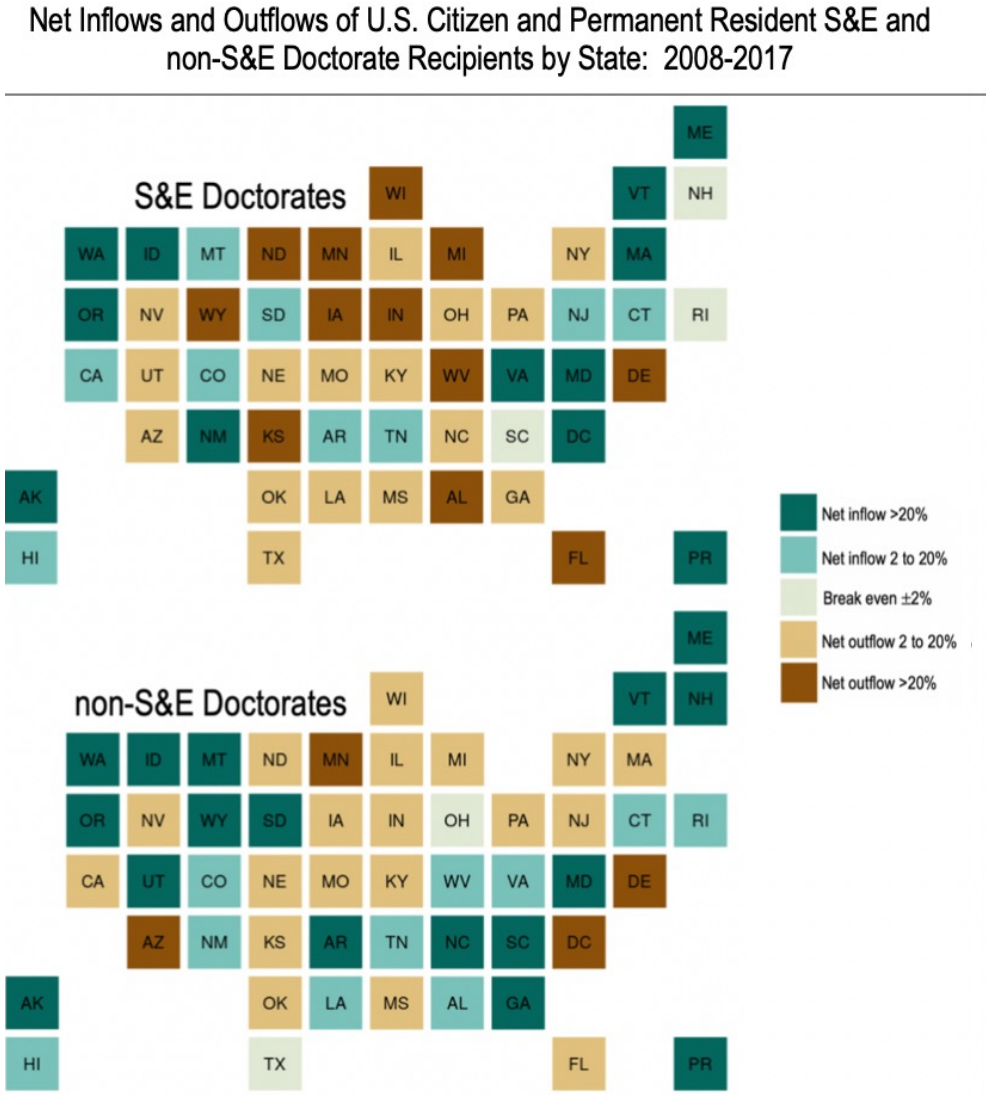
> Interval: Inflow/outflow of S&E % non-S&E doctorate recipients between 2008 & 2017

Comments:

> use the same color palette for both maps to make it easier to compare the populations in the two maps

Chloropleth Map Alternative Visualization

State Bins



Chloropleth Map Alternative Visualization

Chloropleth Map Using the same Colorblind Palette for both Maps