

Data Visualization

Dave Stinchcomb

October 12, 2016



What is Data Visualization?

- Data visualization is about **COMMUNICATION**
- Using visual design principles to facilitate communication of quantitative data
- Key questions:
 - Who is your audience?
 - What is the main message?

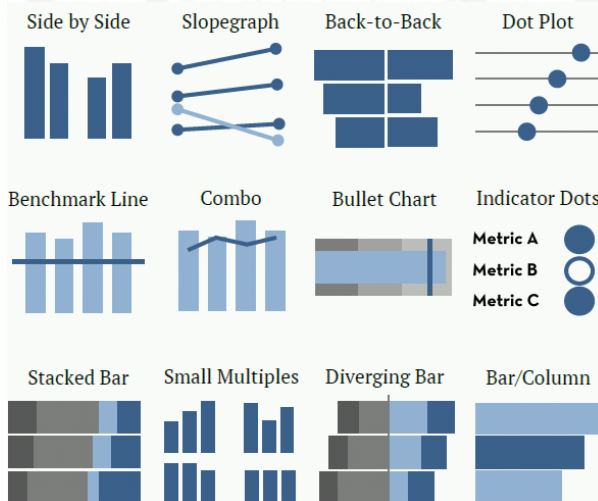


The Galaxy of Data Visualization

- Static graphs, charts, and maps
- Interactive graphics
- Dashboards
- Infographics
- Report design / web page design

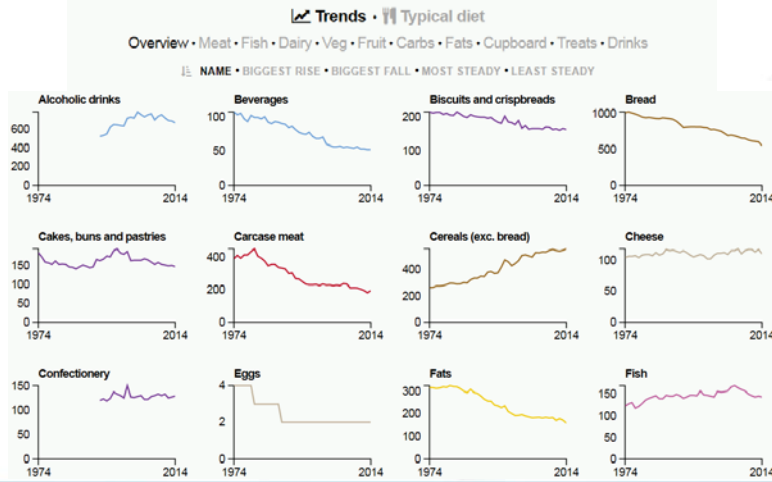


Views of the Galaxy - Static Graphs

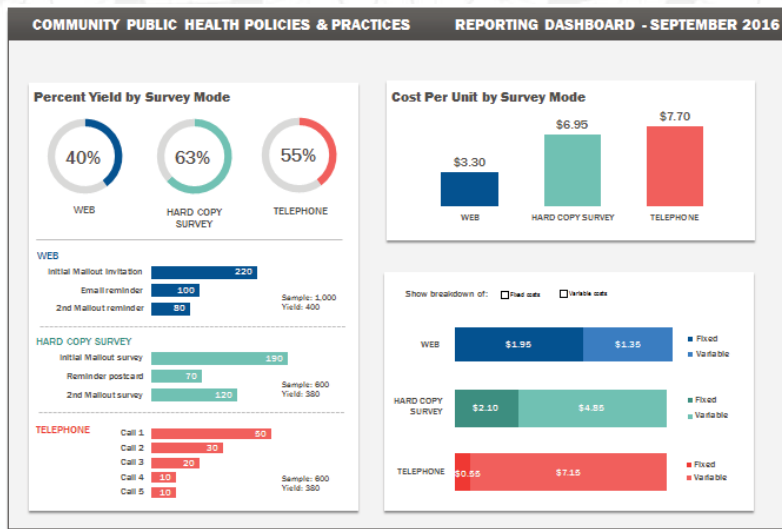


Views of the Galaxy - Interactive Graphics

Changes in the British Diet – <http://britains-diet.labs.theodi.org/>



Views of the Galaxy - Dashboards



Views of the Galaxy - Infographics

Buckle Up: Restraint Use in ARKANSAS

MOTOR VEHICLE OCCUPANT DEATHS

Number of Deaths, 2001-2012
4,933 motor vehicle occupants were killed in Arkansas

Rate of Deaths by Age (per 100,000 population), 2012

Age Group	National	Arkansas
0-20	4.8	5.8
21-34	13.34	18.9
35-44	9.3	13.8
45+	12.1	16.4
65+	7.9	11.8

Rate of Deaths by Gender (per 100,000 population), 2012

Gender	National	Arkansas
Male	8.6	12.8
Female	7.7	10.8

RESTRAINT USE

Percentage of Drivers and Front Seat Passengers Wearing Seat Belts

Category	National	Arkansas
Drivers	84%	72%
Front Seat Passengers	84%	72%

Fast Facts

- Motor vehicle crashes are a leading cause of death during the first three decades of Americans' lives.
- By wearing seat belts and properly buckling children into age- and size-appropriate car seats and booster seats, parents can reduce the risk of serious injury and even in a crash by half.
- Although most drivers in the United States follow these safety measures on every trip, there are still millions who don't.
- These data show what's happening in your state.

Working together, we can help keep people safe on the road every day.

Perceptions of Patient Safety Culture

The highlights presented here from the 2014 AHRQ Hospital Survey on Patient Safety Culture Comparative Database Report provide a glimpse of how those working in hospitals perceive 12 different areas of patient safety culture.

Top 3 Areas

- 81%** Employees within units (staff and non-staff) together on teams
- 76%** Superiors respect expectations & autonomy
- 73%** Organizational learning

Bottom 3 Areas

- 44%** Disciplinary response to error
- 47%** Harassment & mistreatment
- 55%** Staffing

64%
Average percent positive across all 12 areas of patient safety culture

419,281 staff from 653 hospitals responded

68% **81%**

Smaller hospitals reported higher safety culture perceptions than larger hospitals.

75% **63%**

Management reported higher safety culture perceptions than clinical staff.

62% **49%**

More staff in ICU units reported at least 1 error in the last 12 months than those in Emergency departments.

Full details, tables, charts, and data: www.ahrq.gov/patient-safety/culture

Views of the Galaxy – Report Design

National Autism Indicators Report Transition into Young Adulthood

Table of Contents

- Executive Summary: Where we stand today
 - Key Findings 2015
- Youth on the Autism Spectrum: Characteristics at age 17
- What happens to youth with autism as they enter adulthood?
 - Transition Planning: Building a bridge to the future
 - The Services Cliff: Services decreased significantly in young adulthood
- What do we know about key outcomes?
 - Young Adult Outcomes and Disconnection: The big picture
 - Health, Mental Health and Health Care: Complex needs
 - Postsecondary Education: Few continued beyond high school

Transition Planning: Special Education and Services

Building a bridge to the future

The Services Cliff

Services decreased significantly in young adulthood

Young Adult Outcomes and Disconnection

Health, mental health and health care

Postsecondary Education

Few continued beyond high school

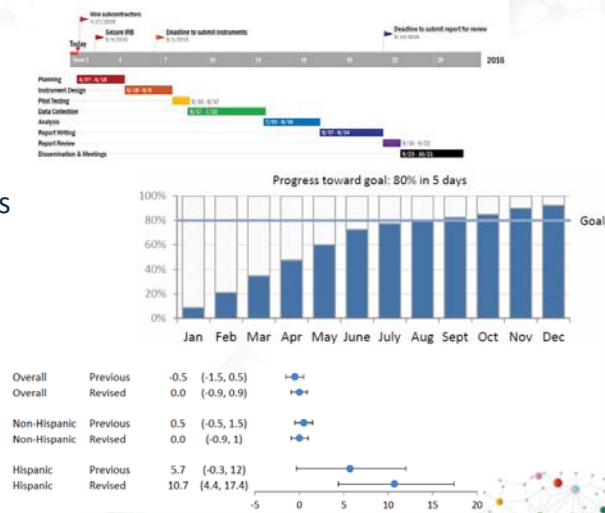
Health, Mental Health and Health Care

Complex needs

<http://stephanieevergreen.com/design-of-an-award-winning-report/>

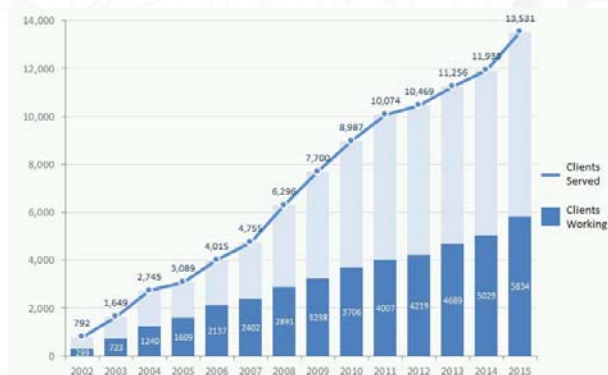
DataViz in Cancer Registry Projects

- Internal planning and tracking
 - Internal schedule tracking
 - Reporting to leadership
- Data collection
 - Tracking progress toward goals
- Analysis
 - Exploratory data visualization
- Surveillance reports
- Grant proposals
- Academic papers
- Conference presentations



Principles of Visual Design – Visual Hierarchy

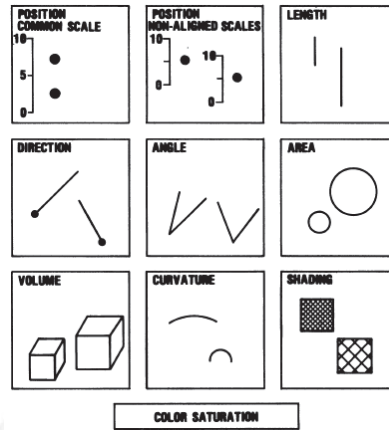
- Visual hierarchy
 - What stands out?
 - What draws your eye?
 - Elements:
 - Size
 - Color and contrast
 - Typography
 - Spacing
 - Composition
 - Do the “squint” test – what stands out?



Visual Comparisons

- Cleveland and McGill's comparison accuracy scale:

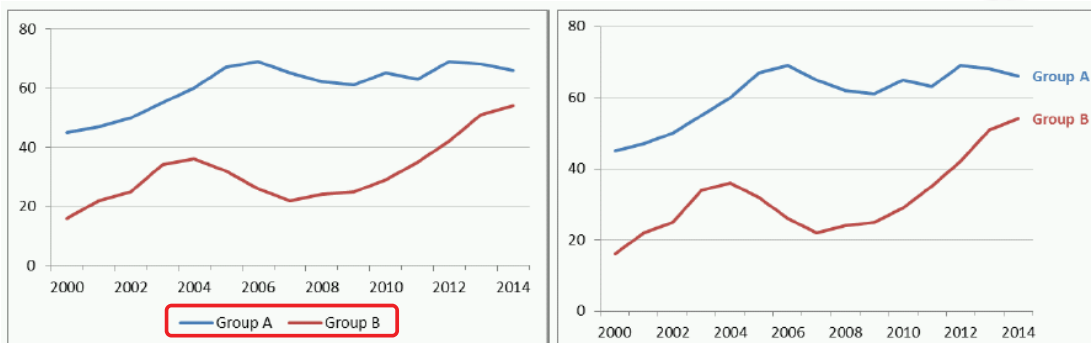
1. Position along a common scale
2. Position on identical but nonaligned scales
3. Length, angle & direction
4. Area
5. Volume, curvature
6. Shading, color saturation



Cleveland & McGill, JASA, 1984

Make It Easy on the Viewer

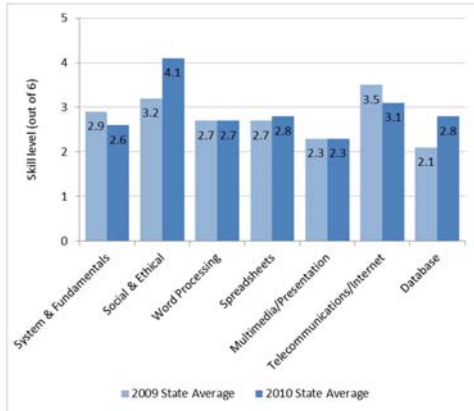
- Avoid legends and encoded meaning



- If a legend is needed, use placement, order, and color to make interpretation easier

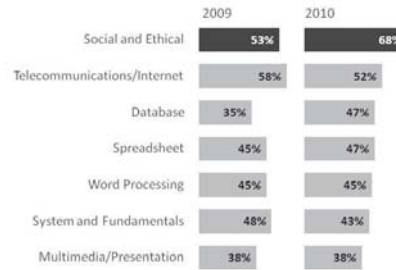
Make It Easy on the Viewer

- Avoid diagonal or vertical text



Better redesign:

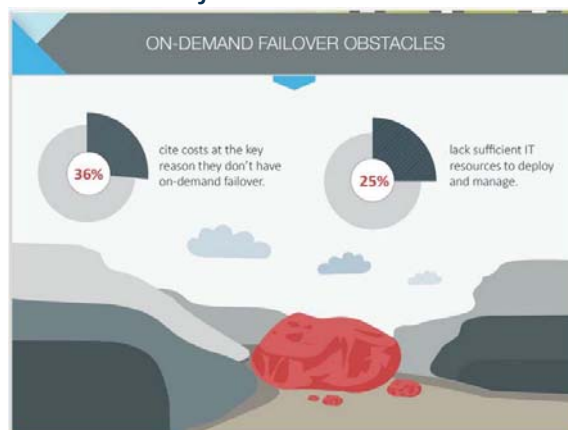
In 2010, students scored highest on Social and Ethical Skills
They also scored well on Telecommunications/Internet and Database skills.



Adapted from <http://annkemery.com/avoiding-diagonal-text/>

Make It Easy on the Viewer

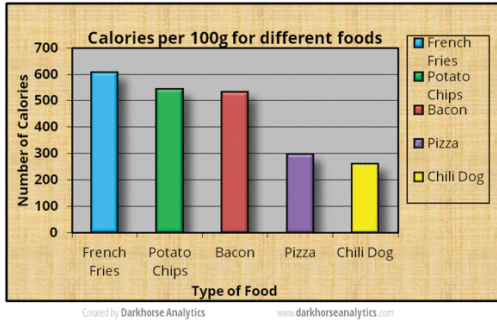
- Avoid clutter and “chart junk”



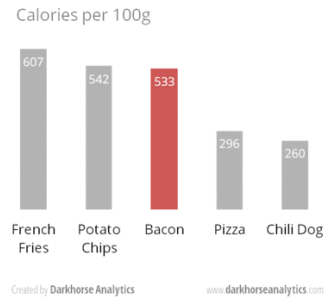
<http://visual.ly/top-4-disaster-recovery-findings>

Transformation 1 – Data Look Better Naked

Before



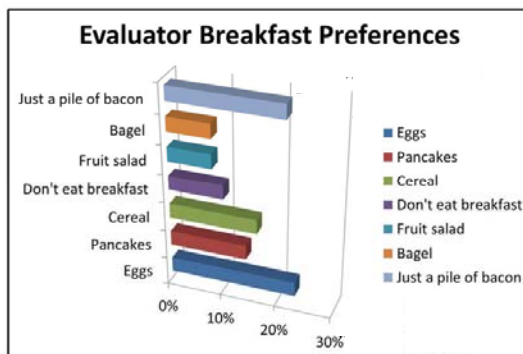
After



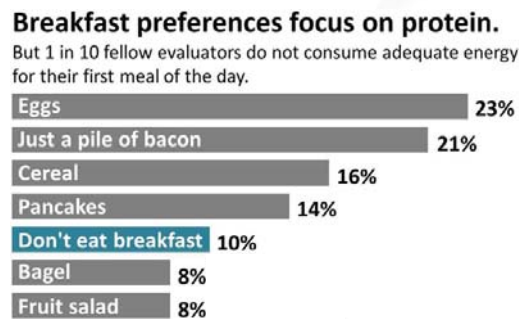
<http://www.darkhorseanalytics.com/blog/data-looks-better-naked/>

Transformation 2 – Evaluator Breakfasts

Original:



Redesign:



Stephanie Evergreen - <http://stephanieevergreen.com/dataviz-checklist/>

Makeover – Adding bars to data tables

Original:

Program/Cohort	Number of Grantees With Data	Percentage of Grantees That Improved
Cohort 1	19	73.7
Cohort 2	5	60.0
Cohort 3	10	70.0
Cohort 4	13	61.5
Cohort 5	3	66.7
All cohorts	50	68.0
All States & D.C.	51	68.6

Redesign:

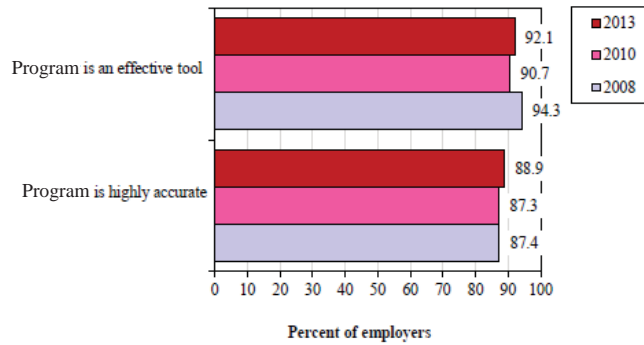
Program/Cohort	# of Grantees With Data	% of Grantees That Improved
Cohort 1	19	73.7
Cohort 2	5	60
Cohort 3	10	70
Cohort 4	13	61.5
Cohort 5	3	66.7
All cohorts	50	68
All States & D.C.	51	68.6



Makeover – Bar chart with time series

Original:

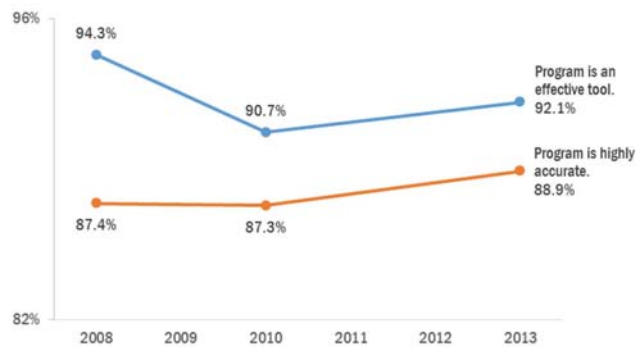
Percent of users reporting that they agree with positive statements about the Program's effectiveness: 2013, 2010, and 2008



Makeover – Bar chart with time series

Redesign: - time series line chart

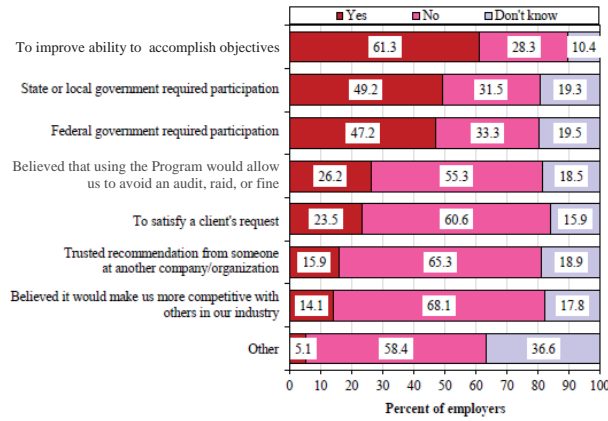
The percent of participants agreeing that Program is **effective** dropped in 2010 but both **effectiveness** and **accuracy** increased in 2013.



Makeover – Stacked bar charts

Original:

Percent of employers reporting all reasons for agreeing to participate in the Program



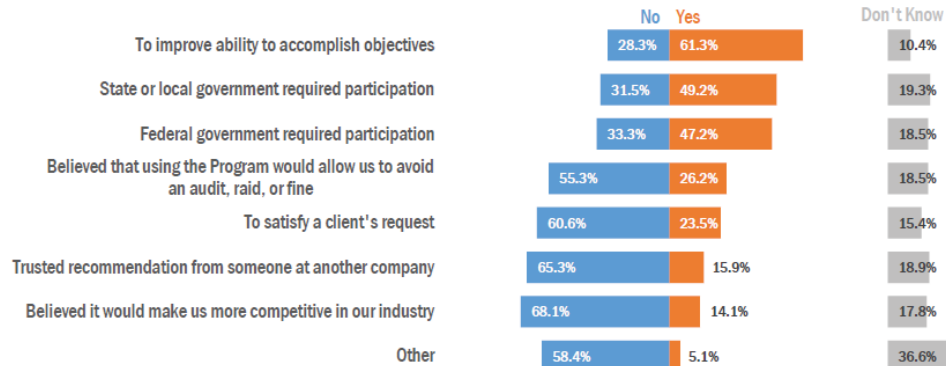
Makeover – Stacked bar charts

Redesign – diverging stacked bar chart:

Poll results:

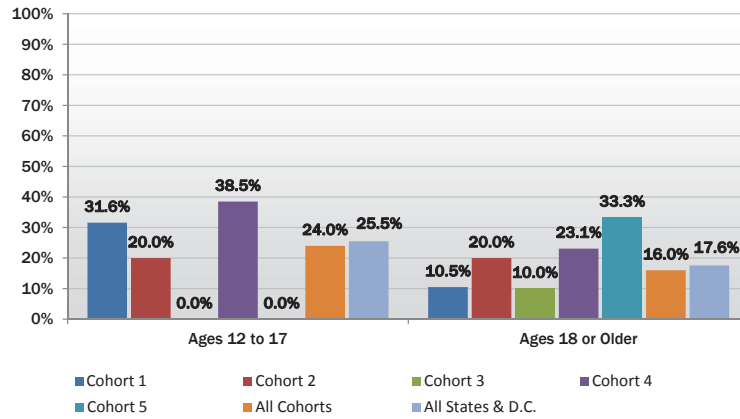
Answers	Results	Percentages
"Yes" to the left, "No" to the right	29/128	23%
Keep it as it is	55/128	43%
Doesn't really make a difference	13/128	10%
Can you repeat the question?	1/128	1%
No Answer	31/128	24%

Employers reported three main reasons for agreeing to participate in the Program



Makeover – Clustered column chart

Original:

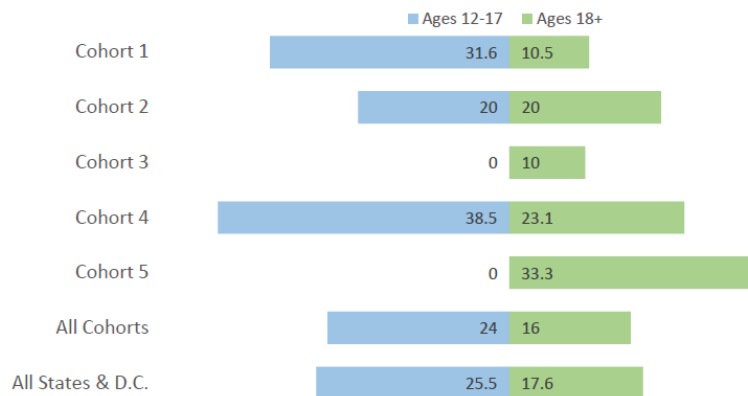


22

Data Visualization | D. Stinchcomb

Makeover – Clustered column chart

Redesign – back-to-back bar chart (with legend):

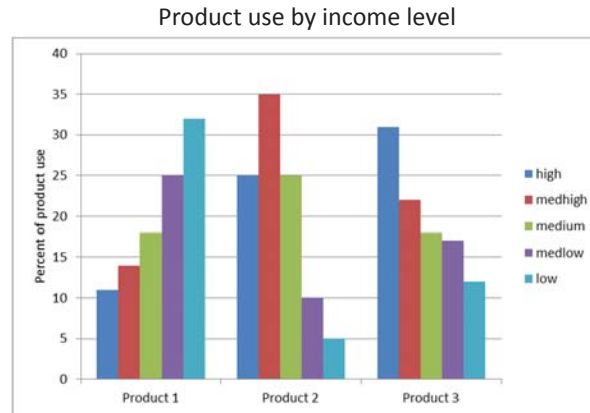


23

Data Visualization | D. Stinchcomb

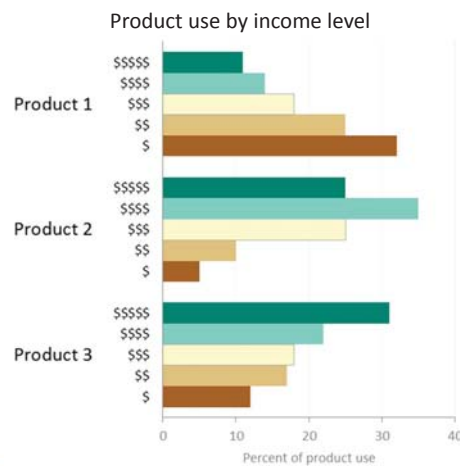
Makeover – Make it easy on the viewer

Original:



Makeover – Make it easy on the viewer

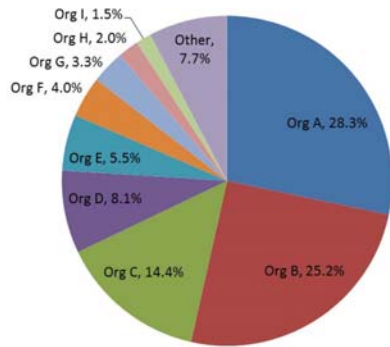
Redesign – horizontal bar chart, diverging color scheme:



Makeover – Pie charts

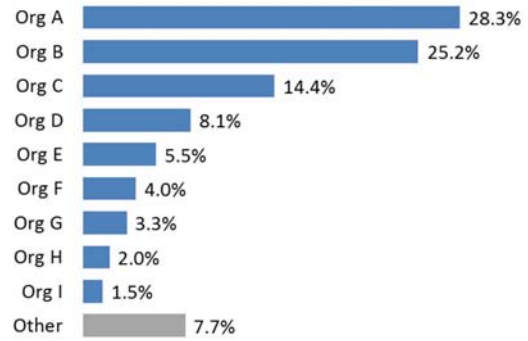
Original:

Percent Participation by Organization



Redesign:

Percent Participation by Organization



Makeover – change from baseline

Original:

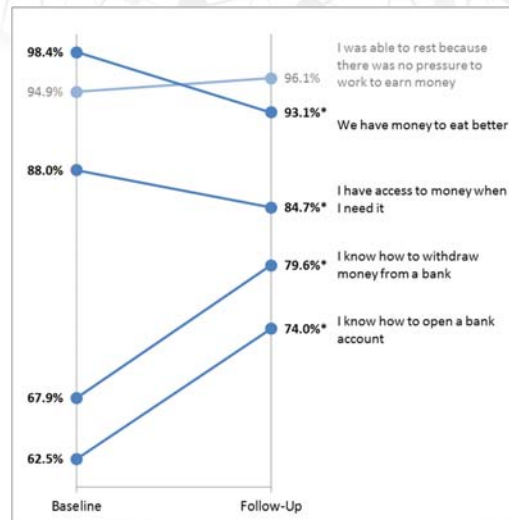
Table 6-5. Women Empowerment Indicators

Statement about Positive Impacts	Percentage of Mothers Agreeing with Each Statement	
	Baseline	Follow-Up
I was able to rest because there was no pressure to work to earn money	94.9	96.1
I know how to withdraw money from a bank	67.9	79.6*
I know how to open a bank account	62.5	74.0*
I have access to money when I need it	88.0	84.7*
We have money to eat better	98.4	93.1*
Weighted n	2,453	2,059
Unweighted n	98,853	91,321

* Change from Wave 1 to Wave 2 is statistically significant at the one percent level.

Makeover – change from baseline

Redesign – slope graph,
highlight significant results:



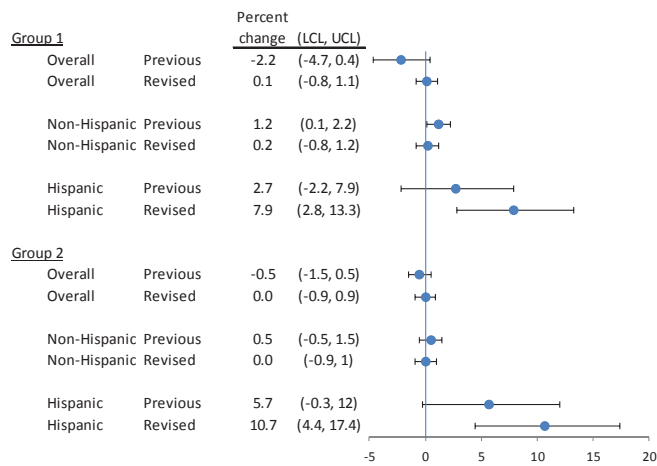
Makeover – Confidence intervals in tables

Redesign – add graph to the table:

<u>Group 1</u>		Percent	
		<u>change</u>	<u>(LCL, UCL)</u>
Overall	Previous	-2.2	(-4.7, 0.4)
Overall	Revised	0.1	(-0.8, 1.1)
Non-Hispanic	Previous	1.2	(0.1, 2.2)
Non-Hispanic	Revised	0.2	(-0.8, 1.2)
Hispanic	Previous	2.7	(-2.2, 7.9)
Hispanic	Revised	7.9	(2.8, 13.3)
<u>Group 2</u>		Percent	
		<u>change</u>	<u>(LCL, UCL)</u>
Overall	Previous	-0.5	(-1.5, 0.5)
Overall	Revised	0.0	(-0.9, 0.9)
Non-Hispanic	Previous	0.5	(-0.5, 1.5)
Non-Hispanic	Revised	0.0	(-0.9, 1)
Hispanic	Previous	5.7	(-0.3, 12)
Hispanic	Revised	10.7	(4.4, 17.4)

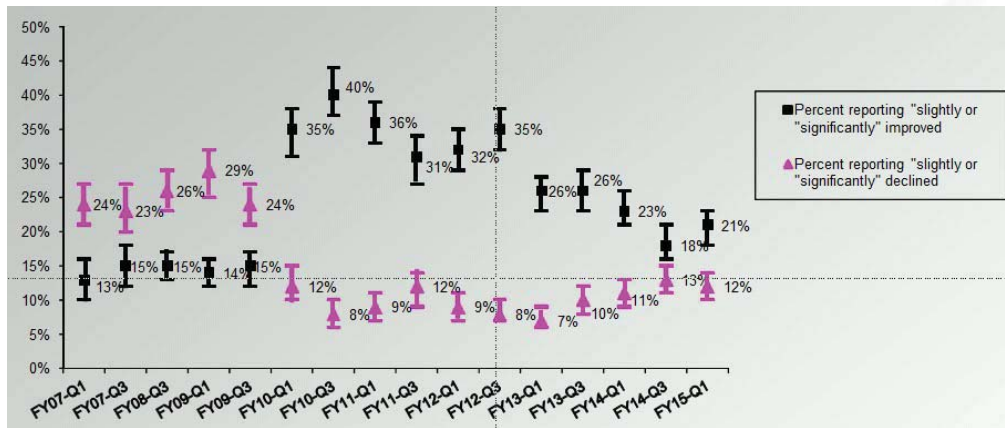
Makeover – Confidence intervals in tables

Redesign – add graph to the table:



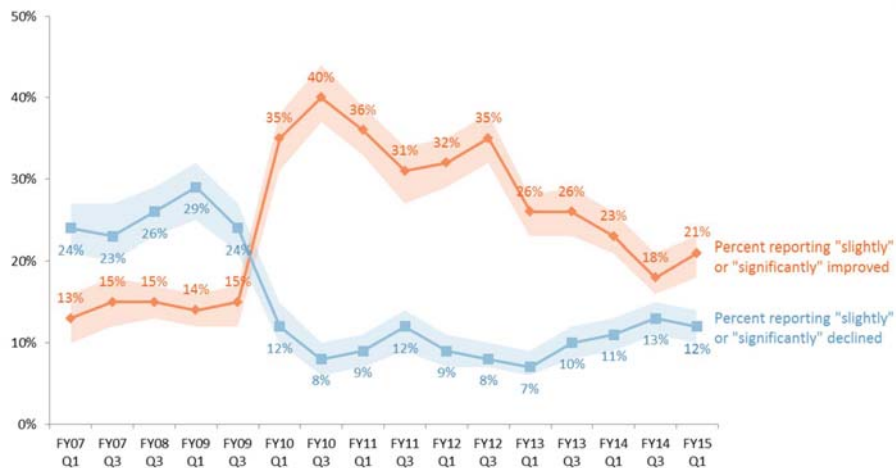
Makeover – Confidence intervals in graphs

Original:



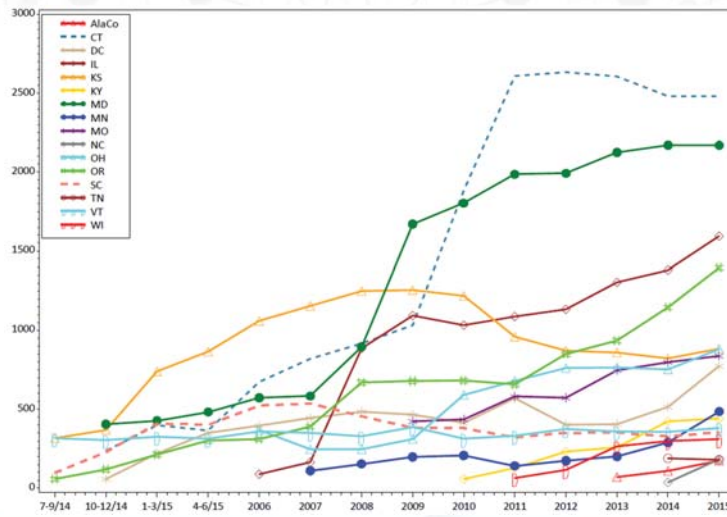
Makeover – Confidence intervals in graphs

Redesign – error bands:



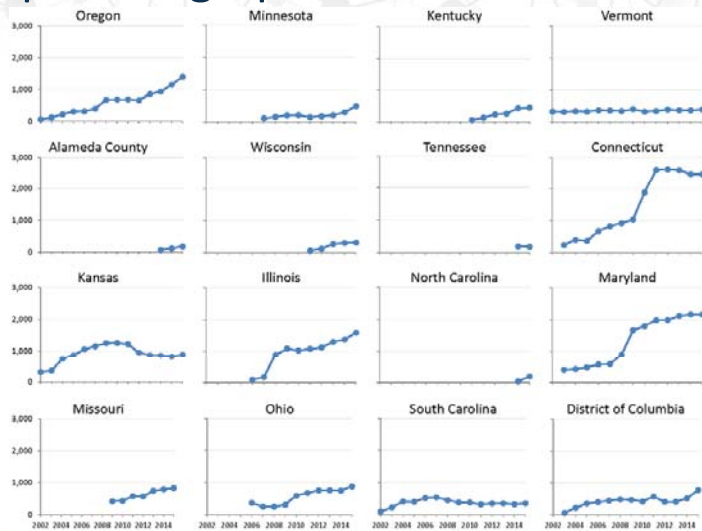
Makeover – Multiple line graph

Original:



Makeover – Multiple line graph

Redesign –
small multiples:



Tools – Choosing the Right Chart Type

- Data Visualization Catalog <http://www.datavizcatalogue.com/>

Tools – DataViz Checklist

<http://stephanieevergreen.com/dataviz-checklist/>

Data Visualization Checklist

by Stephanie Evergreen & Ann K. Emery
May 2014

Guideline	Rating
Text size is hierarchical and readable Titles are in a larger size than subtitles or annotations, which are larger than labels, which are larger than axis labels, which are larger than source information. The smallest text - axis labels - are at least 9 point font size on paper, at least 20 on screen.	2 1 0 n/a
Data are labeled directly Position data labels near the data rather than in a separate legend (e.g., on top of or next to bars or pie slices, and next to lines in line charts). Eliminate/embed legends when possible because eye movement back and forth between the legend and the data can interrupt the brain's attempts to interpret the graph.	2 1 0 n/a
Graph is two-dimensional Avoid three-dimensional displays, bevels, and other distortions.	2 1 0 n/a
Color is used to highlight key patterns Action colors should guide the viewer to key parts of the display. Less important or supporting data should be a muted color.	2 1 0 n/a
Color is legible for people with colorblindness Avoid red-green and yellow-blue combinations when those colors touch one another.	2 1 0 n/a

Tools – Creating static graphs

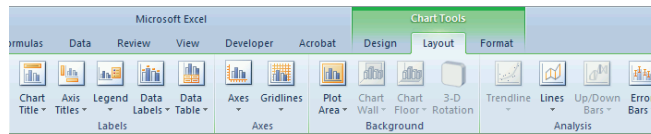
- Excel is your friend
 - Most examples have been done with Excel



- But Excel graph defaults are **NOT** your friends!
 - Get in the habit of modifying each graph

<http://thewhyaxis.info/defaults/>

The Why Axis



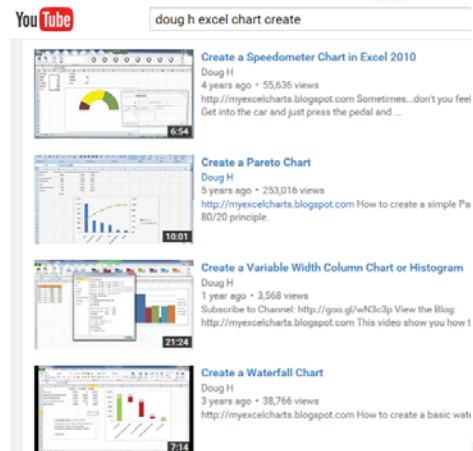
Breaking Excel Defaults – A Government Chart Remake

Guest author Jon Schwabish – an economist with the U.S. federal government and creator of policy-relevant data visualizations. You can reach him at jschwabish@gmail.com or by following him on Twitter @jschwabish.

- Programming tools for static graphs:
 - R and the “ggplot2” package
 - SAS statistical graphics: SG procedures (SGPLOT, SGSCATTER, SGPANEL)

Tools – Excel How-To Guides

- EvergreenData step-by-step
<http://stephanieevergreen.com/tag/step-by-step/>
- YouTube videos by Doug H



Interactive Data Visualization Tools

- Interactive data visualization:

- Tableau



- Commercial tool for quick interactive graphs and dashboards
- Inexpensive for public data; can be pricey for private data

- Browser-based JavaScript libraries:

- Highcharts



- D3 (popular and powerful)



- Dygraphs



General DataViz Resources

- DataViz Blogs:

- Evergreen Data – Stephanie Evergreen
 - <http://stephanieevergreen.com/blog/>
- Peltier Tech Blog – Jon Peltier
 - <http://peltiertech.com/>
- FlowingData – Nathan Yau
 - <http://flowingdata.com/>

- Books by:

- Edward Tufte
- Stephen Few
- Stephanie Evergreen
- Cole Nussbaumer Knaflic

Conclusions – Key Points

- Data visualization is about communication
 - Bring data to life
 - Who is the audience; what is the key message?
- Build data visualization into registry plans
 - Operations, data collection, analysis, reporting
- Visual design principles:
 - Visual hierarchy
 - Effective visual comparisons
 - Making it easy on the viewer

Conclusions – The Bottom Line

- Data visualization can make the work you do
 - More accessible
 - More impactful
 - More effective



Thank You

Useful Links

- DataViz Catalog
<http://www.datavizcatalogue.com/>
- DataViz Checklist
<http://stephanieevergreen.com/dataviz-checklist/>
- Breaking Excel Defaults
<http://thewhyaxis.info/defaults/>
- EvergreenData Step-by-step
<http://stephanieevergreen.com/tag/step-by-step/>
- Color selection:
<http://colorbrewer2.org/>
- Colorblind testing tools:
<http://www.color-blindness.com/coblis-color-blindness-simulator/>
<http://www.vischeck.com/vischeck/>

