Data Visualization Training Webinar

Michael López
Nola du Toit
Ned English
Tweet with us!

And follow us @NRCHispanic for discussion on this webinar and other topics important to low-income Hispanic children and families.
Presenters

Michael López, Ph.D.
Vice President, Education and Child Development, NORC at the University of Chicago
Co-Principal Investigator, NRCHCF
López-Michael@NORC.org

Nola du Toit
NORC at the University of Chicago
Dutoit-Nola@NORC.org

Ned English
NORC at the University of Chicago
English-Ned@NORC.org

A special thanks to Rebecca Berger, Ana Leon-Santos, Carina Hoyer & Katarina Yang for their behind the scenes support and contributions.
Conduct research and provide research-based information to inform ACF programs and policies supporting low-income Hispanic children and families around:

- Poverty and self-sufficiency
- Healthy marriage and responsible fatherhood
- Early care and education

We do this through:

- Building research capacity
- Dissemination and outreach
Research Capacity Building

Strengthen the capacity of the research field and expand the pipeline of scholars focused on Hispanic children and families.

- Provide resources to the field
- Support emerging scholars
- Engage key stakeholders
Disclaimer

The views expressed in this presentation do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.
Goals of Today’s Data Visualization Training

• Components of good data visualization
• Storytelling with data
• Maps and Mappable Data
• Q&A
Components of Good Data Visualization
What is the DATA story?

- Comparison of groups
- Distribution
- Composition or proportion
- Over time
- Geographic

What does the data look like?

Purpose

Audience

Best Practices
What is the INFORMATION story?

- Progress or process
- Flowchart
- Timeline
- List

Information = Data
Purpose

Policy makers, potential funders, community members, advocacy groups

Audience

General public, teenagers, elderly
Journalists

Best Practices

What do I want them to…?
Think… Know… Do…

Why is this important for them?
What matters to them?
How do they regard the issue?

How much do they need to know?

• Raise awareness of an issue
• Garner support
• Attract new funders or increase funding
• Prompt a specific action
Best Practices

Data formatting

Purpose

Audience

Tools

Free online library of JavaScript code, cool graphics, interactive, programming language

Free, cool graphics, interactive, programming language

Mapping geographic data

Easy, drag-and-drop, good for common graphics, good defaults; expensive licensing fees, easy to make mistakes

Some graphics are easy, everyone has access to it, easy to share, defaults are bad

Easy, drag-and-drop, good for common graphics, good defaults; expensive licensing fees, easy to make mistakes

Some graphics are easy, everyone has access to it, easy to share, defaults are bad

Free online library of JavaScript code, cool graphics, interactive, programming language

Free, cool graphics, interactive, programming language

Mapping geographic data

MORE
Checklist

No decimals points

Remove clutter and redundancy

How much information?
Is the data clear?
Storytelling with Data
Purpose: Children in Chicago Compared by Race/Ethnicity

Audience: General public

Data: ACS data 2012-2017

Time Periods: 2012: Chicago presents itself as an immigrant friendly city
               2016: National immigration context

Tool: Excel
One Third of Children in Metro Chicago are Hispanic…

Percent of Children in Chicago, by Race/Ethnicity

- Hispanic: 31%
- Black: 18%
- Asian: 5%
- White: 43%

Compare across race groups

Excel (bar chart)
And the numbers have not changed much over the past few years.

Percent of Children in Chicago, by Race/Ethnicity

- Hispanic: 30% (2012), 31% (2017)
- Black: 19% (2012), 18% (2017)
- Asian: 5% (2012), 5% (2017)
- White: 43% (2012), 43% (2017)

Excel (bar chart)
And the numbers have not changed much over the past few years

Percent of Children in Chicago, by Race/Ethnicity
For 2012 and 2017

<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>2012</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hispanic</td>
<td>30%</td>
<td>31%</td>
</tr>
<tr>
<td>Black</td>
<td>19%</td>
<td>18%</td>
</tr>
<tr>
<td>Asian</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>White</td>
<td>43%</td>
<td>43%</td>
</tr>
</tbody>
</table>
Maps and Mappable Data

• Want to take advantage of spatial data and so consider geography

  • All things are related but near things more related than far things (Tobler’s first law of geography)
  • Lots of data (fractal detail)

• GIS tools, techniques for analysis/management of spatial data
  • Preoccupied with representation

• Cartography deals with the presentation of spatial data (make nice maps), visualization
  • Preoccupied with presentation
What is Special about GIS?

• Take advantage of spatial relationships
• Permits complex spatial analyses
• Geographic queries added to database
• Consider spatial relationships e.g. distance, proximity, adjacency, topology
• Adds third dimension to statistical problems
• Whole realm of spatial statistics, autocorrelation, interpolation
Snow’s Cholera Map: 1864
Lots of Bad Maps...
Common Applications of GIS in Social Sciences

1. Presentation of data, results through maps
   - Often choropleth

2. Creation of new data layers
   - Overlays, spatial joins
   - Geocoding

3. Linking disparate data sets together
   - Without *a priori* key field

4. Spatial data analysis
   - Finding patterns related to geography
Choropleth Maps

- Often want to map statistical surface
  - Quantitative attribute values referenced to areas
  - A “lattice”, type of vector
- Population of Census tracts
- Median household income of counties
- Etc.

- *Choropleth* maps portray a statistical surface with area symbols
  - Data coincide with data collection regions i.e. “lattice data”
We Are Familiar with Uneven Lattices
Choropleth Maps: Design Considerations

1. Size and shape of unit areas
   • Should be about the same size and shape
   • If too big, too variable, no good

2. Number of classes

3. Class limit determination
   • Can get very different results depending
   • As described in Monmonier (1991) How to Lie with Maps
Manipulating Breaks for Different Looking Maps

From Monmonier 1991
Multi-Variate Maps

• Can be hard to map more than one variable at once
  • And be legible
  • Want to be efficient with information
• How can we accomplish this?

1. Superimpose Features
  • Use symbols that are transparent, but distinct
  • Dots, isolines, proportional-point symbols
  • Advantage: simple to understand, effective with few variables
  • Disadvantage: impractical with more variables, difficult to convey relative importance
Multivariate Maps Contd.

2. Segmented Symbols

• Could divide each symbol so more information
  • E.g. pie chart

• Can show complex relationships, but hard to understand
  • Pie charts difficult to interpret
  • Map users have trouble comparing proportions, parts of symbols
Segmented Symbols
3. Cross-variable mapping
   • nxn matrix legend
   • Can express a lot of information effectively
   • Hard to understand at >3 classes
   • Should emphasize distinction between classes
   • See a lot in popular media: NY Times, Economist
Cross-variable mapping 1

Note: Due to limitations in the Census data, foreign-born populations are not available in all areas for all years.

Sources: Social Explorer, www.socialexplorer.com; Minnesota Population Center; U.S. Census Bureau

Matthew Bloch and Robert Gebeloff/The New York Times
Map Set 1- Multivariate Choropleth

**Purpose**: Proportion of children by modal race in Chicago Census tracts

**Audience**: General public

**Data**: ACS data 2013-2017

**Time Periods**: 2012: Chicago presents itself as an immigrant friendly city, 2016: National immigration context

**Tool**: ArcGIS
• Different hue = different modal race/ethnicity
• Lighter value = more mixed tracts
Map Set 2- Multivariate Choropleth

**Purpose**
Two maps of the proportion of single parents and non-single parents living in poverty with a related child in the household, by the modal race/ethnicity in Chicago tracts

**Audience**
General Public

**Data**
Five Year ACS data 2013-2017

**Time Periods**
2012: Chicago presents itself as an immigrant friendly city
2016: National immigration context

**Tool**
ArcGIS
Single-Parent Households

Non-Single-Parent Households

"The Loop"
Map Set 3- Animated Choropleth

Purpose: Animated map of the proportion of children under 5 who are Hispanic in Chicago tracts, with home-based and center-based childcare locations.

Audience: General Public.

Data: Chicago Child Care Licensing Data and ACS data 2012-2017.

Time Periods:
- 2012: Chicago presents itself as an immigrant friendly city.
- 2016: National immigration context.

Tool: ArcGIS.
Map Set 4 - Difference Map

Purpose
Map of the change of Hispanic children in Chicago from 2012 - 2017

Audience
General Public

Data
ACS data 2012-2017

Time Periods
2012: Chicago presents itself as an immigrant friendly city
2016: National immigration context

Tool
ArcGIS
• Blue = Loss of Hispanic Children
• Red = Gain in Hispanic Children
Thanks to our Funders!
Thank you!

Michael López – Lopez-Michael@NORC.org
Nola du Toit - Dutoit-Nola@norc.org
Ned English - English-Ned@norc.org

www.HispanicResearchCenter.org

@NRCHispanic

National Research Center on Hispanic Children & Families