Visualizing Data for Decision-Making at Your Library

Catherine Stollar Peters, Ph.D.
Assistant Director of Operations
Bethlehem Public Library

New York Library Association
Annual Conference
November 15, 2019
Case study: Geographic location of library card holders

Case study: Dashboard

Case study: Mapping applications

Application at your institution
Bethlehem Public Library

- Delmar, New York
- Service area 28,000+
- School district public library
- Upper Hudson Library System
- 650,000 Annual Circulation
Bethlehem Public Library

- Delmar, New York
- **Service area 28,000+**
- School district public library
- Upper Hudson Library System
- **650,000 Annual Circulation**
Data Visualization Tools

- Mapping tools
  - Tableau
  - ArcGIS
  - Policy Map
  - Google Earth

- Working with data
  - Tableau
  - R
  - OpenRefine
  - Python

- Many many more
  - Google Analytics
  - Power BI
  - DIVE (MIT)
Data Visualization Tools

Mapping tools
- Tableau
- ArcGIS
- Policy Map

Working with data
- Tableau
- R
- OpenRefine
- Python

Many many more
- Software libraries
- Google Analytics
- Power BI
Tableau Public: dashboard creation
Policy Map: geographic data resource
ArcGIS Online: sophisticated mapping tool
Case Study: Geographic Location of Library Card Holders
Start with a question

First Case study:
• Are we meeting the needs of our users with our current location?
• Does the geography of our district and the location of the library shape how residents use the physical library?
Fry’s Seven Stages of Visualizing Data

Acquire

Obtain the data

- Patron address data
- Patron aggregated circulation data
- Map of library district
Parse
Provide some structure for the data’s meaning, and order it into categories

- Standardized export of address data
- Clean any data (wrong zip code, typos)
- Create spreadsheet categories
Filter

Remove all but the data of interest

- Allowed duplicate address entries
- Removed non-district outliers (staff personal, staff agency cards)
- Remove patron name
- Kept number of total checkouts
Mine

Apply methods from statistics or data mining as a way to discern patterns or place the data in mathematical context

- Consider super users
- Look for frequency of zip codes
- Looked for frequency of duplicate address
- Looked at household density (single family house vs. multiple family house)
- Statistical applications
Represent
Estimated population density between 2009-2013.

Estimated population density between 2009-2013. The population density was determined by taking the count of people and dividing it by the square mileage of the area. Areas for which data is not reported are represented as having "Insufficient Data" in the map.
Refine

Improve the basic representation to make it clearer and more visually engaging

- Find new internal and external data
- Use different tools
- Ask more questions
District Super Users
Finding the Data

External Data
• Needed more granular data
  – District-wide data not specific
  – Zip code borders non-congruous
  – Census tract data too large and not exact
  – Find list of addresses in the district
• Use Census block population
  – Decennial census (2010)
  – American Fact Finder
  – Data.census.gov

Internal Data
• Library card data included zip codes
• Smallest unit of data without using geocoding
• Allowed patron privacy while still examining use statistics
American Fact Finder (2018)
Interact

- Present the map and data in an interactive tool
Chart in Tableau
Public
Results

• Data analysis informed Outreach
• Confirmed anecdotal survey data 2018
• Allows targeted patron registration activities
POP-UP LIBRARY!

Jennifer Weiner
Mrs. Everything

Susan Mallery
The Summer of Sunshine & Margot

Angie Thomas
On the Come Up

James Patterson
Kat vs. Dogg

Tara Westover
Educated

BETHELHEM PUBLIC LIBRARY
Case Study: Dashboard
OverDrive Dashboard

Trends

Year over year / User activity

Year over year

Starting month: January

Checkouts: Year over year

Unique users: Year over year
OverDrive Dashboard
What you need to get started

Tableau Public
• Desktop software for creating viz
• Public account for sharing viz
• Data to create dashboard
• Wireframe design
Case Study: Mapping Applications

Catherine Stollar Peters, Ph.D. NYLA 2019
Intro to GIS Data Model

Data Types

• Layers are comprised of two data types
  – Spatial data (where, location)
  – Spatial data is stored in a shape file, geodatabase or geographic file
  – Attribute data (how much, when, what)
  – Attribute data is stored in a table

All geographic data has 4 properties: projection, scale, accuracy and resolution

Source: USGS
Intro to GIS Data Model

Types of data representation

• Vector formats with features
  – Points
  – Lines
  – Polygons

• Raster formats
  – Grid representation

Source: Arcgis.com
ArcGIS Online

- https://arcg.is/19OjzH
- https://arcg.is/0afRXq
- https://arcg.is/18HDWf
- https://arcg.is/PWj4W
5. Don't overload the chart.

Adding too much information to a single chart eliminates the advantages of processing data visually; we have to read every element one by one!

Try changing chart types, removing or splitting up data points, simplifying colors or positions, etc.

(Kaiser Fung, Ruining the Cake with Too Much Icing; Cole Nussbaumer, Death to Pie Charts)
Visualizing Data for Decision-Making at Your Library

Catherine Stollar Peters, Ph.D.
Assistant Director of Operations
Bethlehem Public Library

New York Library Association
Annual Conference
November 15, 2019