5-21-2018

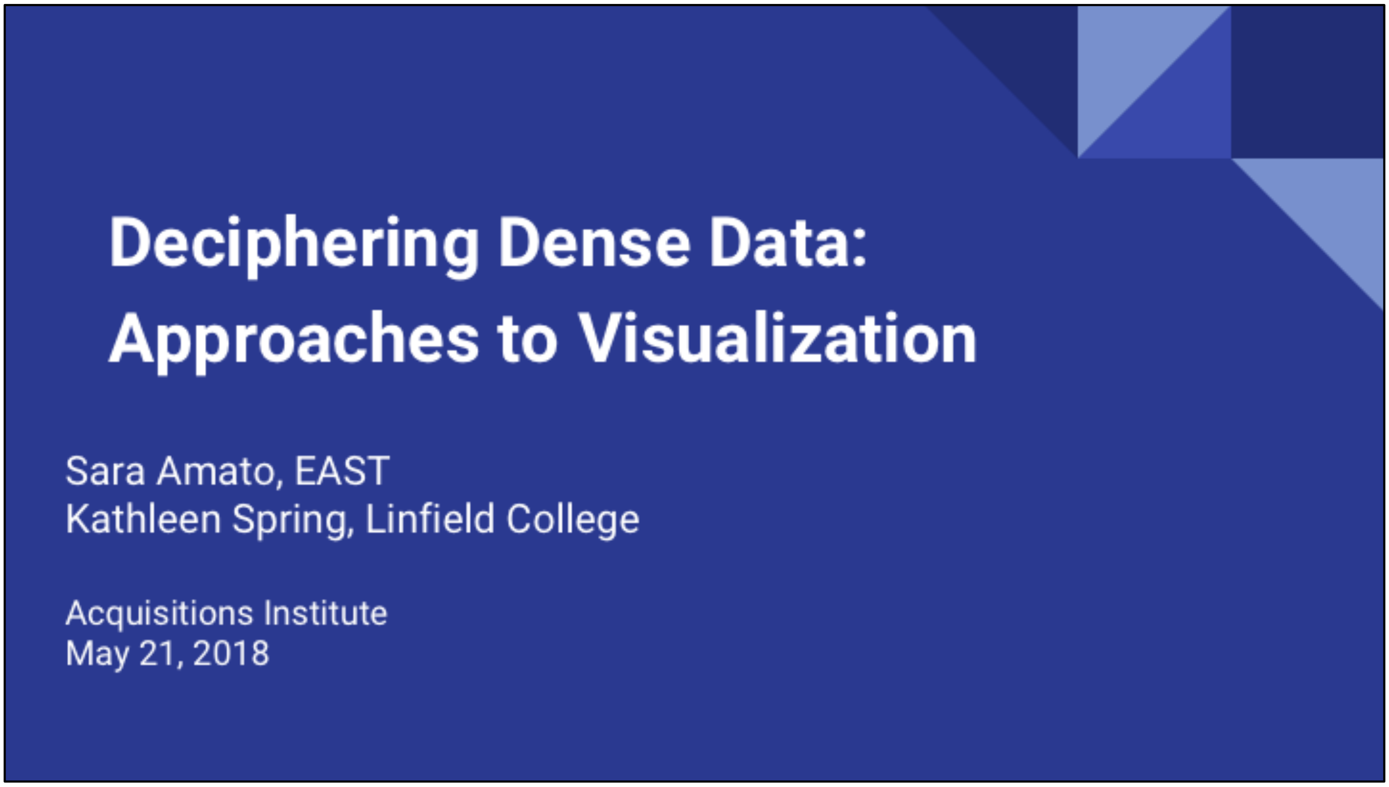
# Deciphering Dense Data: Approaches to Visualization

Sara Amato
*Eastern Academic Scholars' Trust*, samato@blc.org

Kathleen Spring
*Linfield College*, kspring@linfield.edu

## Recommended Citation

Amato, Sara and Spring, Kathleen, "Deciphering Dense Data: Approaches to Visualization" (2018). *Faculty & Staff Presentations.* Presentation. Submission 12.
https://digitalcommons.linfield.edu/librariesfac_pres/12



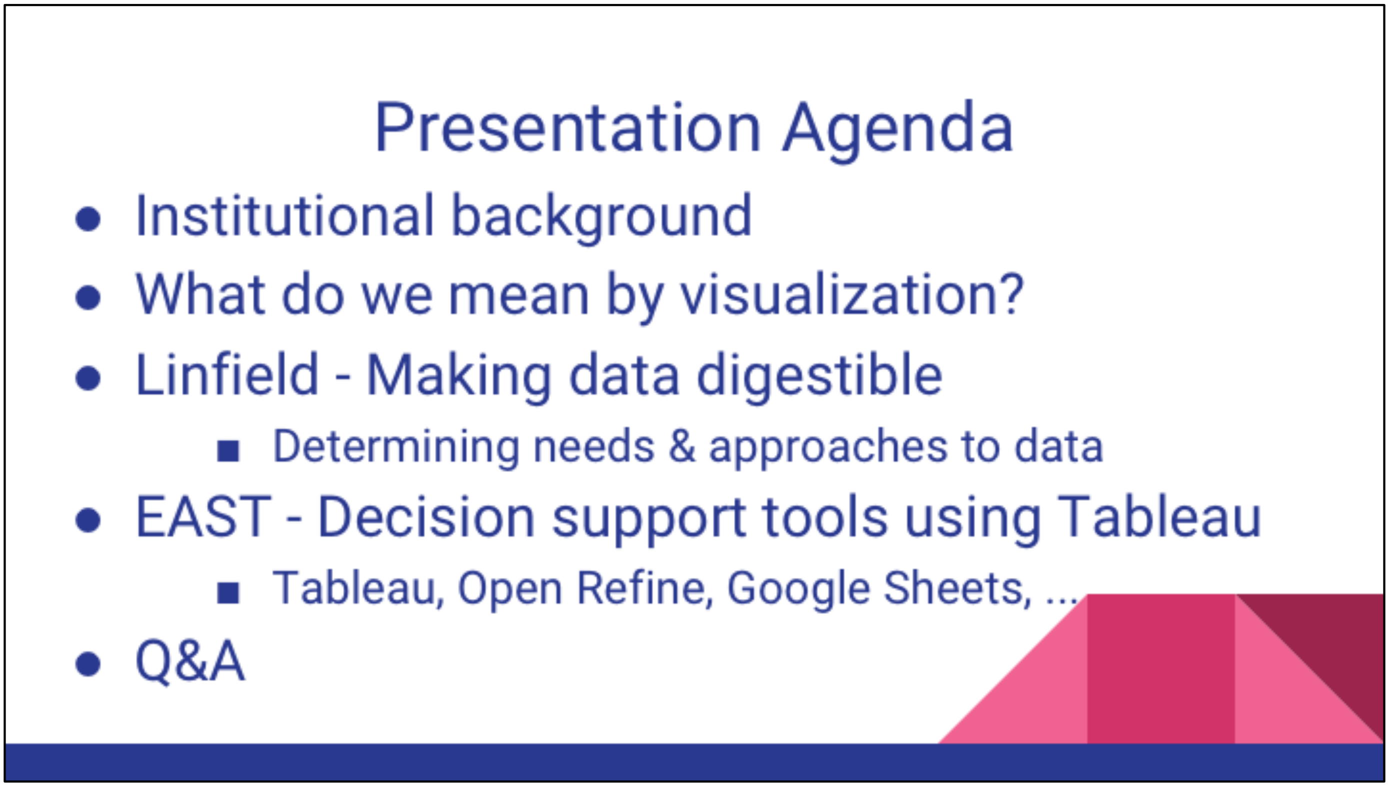
# Deciphering Dense Data: Approaches to Visualization

Sara Amato, EAST
Kathleen Spring, Linfield College

Acquisitions Institute
May 21, 2018

The goal of our session is to show how data visualization can take many forms, and can meet various needs of your organization, from analysis and decision-making support to marketing and storytelling - and that there are lots of easy free or inexpensive tools to do it! Data viz doesn't have to be scary, and you're probably already doing it even if you're not thinking of it as data visualization.

In our handout you'll find links to the tools we mention, along with some others, and a link to these slides.



# Presentation Agenda

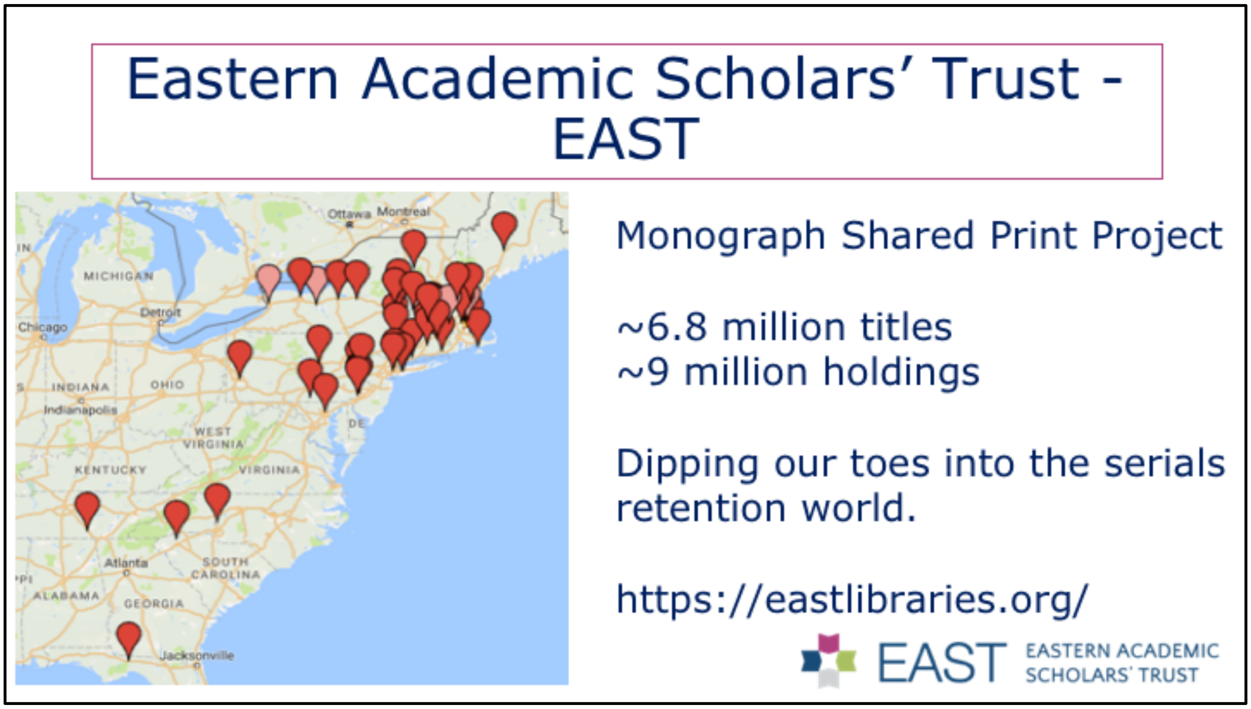
- Institutional background
- What do we mean by visualization?
- Linfield - Making data digestible
  - Determining needs & approaches to data
- EAST - Decision support tools using Tableau
  - Tableau, Open Refine, Google Sheets, ...
- Q&A

Quick overview of what we'll be covering - we'll be using stories from our organizations, focusing on ways to make data digestible and the tools we've used to visualize data, in order to address specific problems.

As we began putting together this presentation, we realized that we might not actually think of data visualization in the same way. So we want to begin by hearing a few brief responses from all of you about how *you* define visualization.

What comes to mind for Kathleen when she hears "data viz" is more in line with what Sara will show you, and Sara's first instincts are more in line with what Kathleen will talk about. So, it's a good thing we're paired together.

The important thing to remember is that visualization is meant to help answer questions or to make something easier to see for those stakeholders who don't live with the data on a regular basis, and it can take form with a variety of levels of complexity - from standard pie charts to interactive dashboards.



# Eastern Academic Scholars' Trust - EAST

Monograph Shared Print Project

~6.8 million titles
~9 million holdings

Dipping our toes into the serials retention world.

https://eastlibraries.org/

EAST EASTERN ACADEMIC SCHOLARS' TRUST

Starting with our Institutional Backgrounds, EAST – the Eastern Academic Scholars' Trust – is a shared print collaboration that currently includes 60 academic and research libraries from Maine to Florida. EAST has been funded by a grant from the Mellon Foundation, with some additional support for collection analysis from the Davis Educational foundation. Grant funding ends in June, at which point EAST will be solely member funded. Staffing includes 4 part-time staff: an executive director, a project manager, a shared print consultant, and myself, as data librarian. The Boston Library Consortium is our fiscal host.

We began work on shared print monographs in July of 2015 and completed our initial retention commitments from our first cohort with just over 6 million holdings in mid-2016. We wrapped up a second cohort of monograph retention partners in April 2018, and we now retain approximately 6.8 million unique titles representing ~9 million holdings.

As we began to dip our toes into the serials world in mid-2017, the question of data analysis loomed large, and that will be the focus of what I talk about today. Let's just say for now that we were very envious of WEST and Agua and their decision analysis tool.

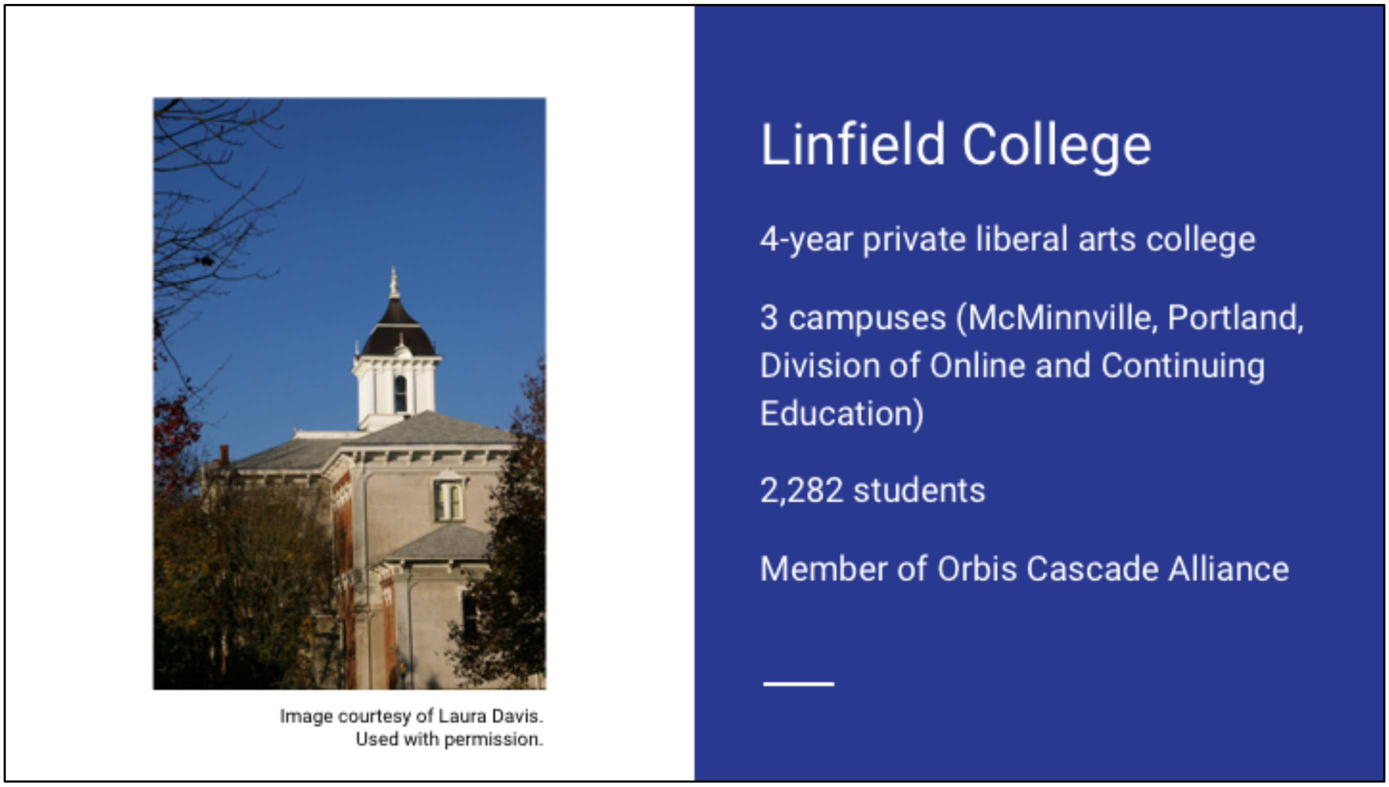


Image courtesy of Laura Davis.
Used with permission.

## Linfield College

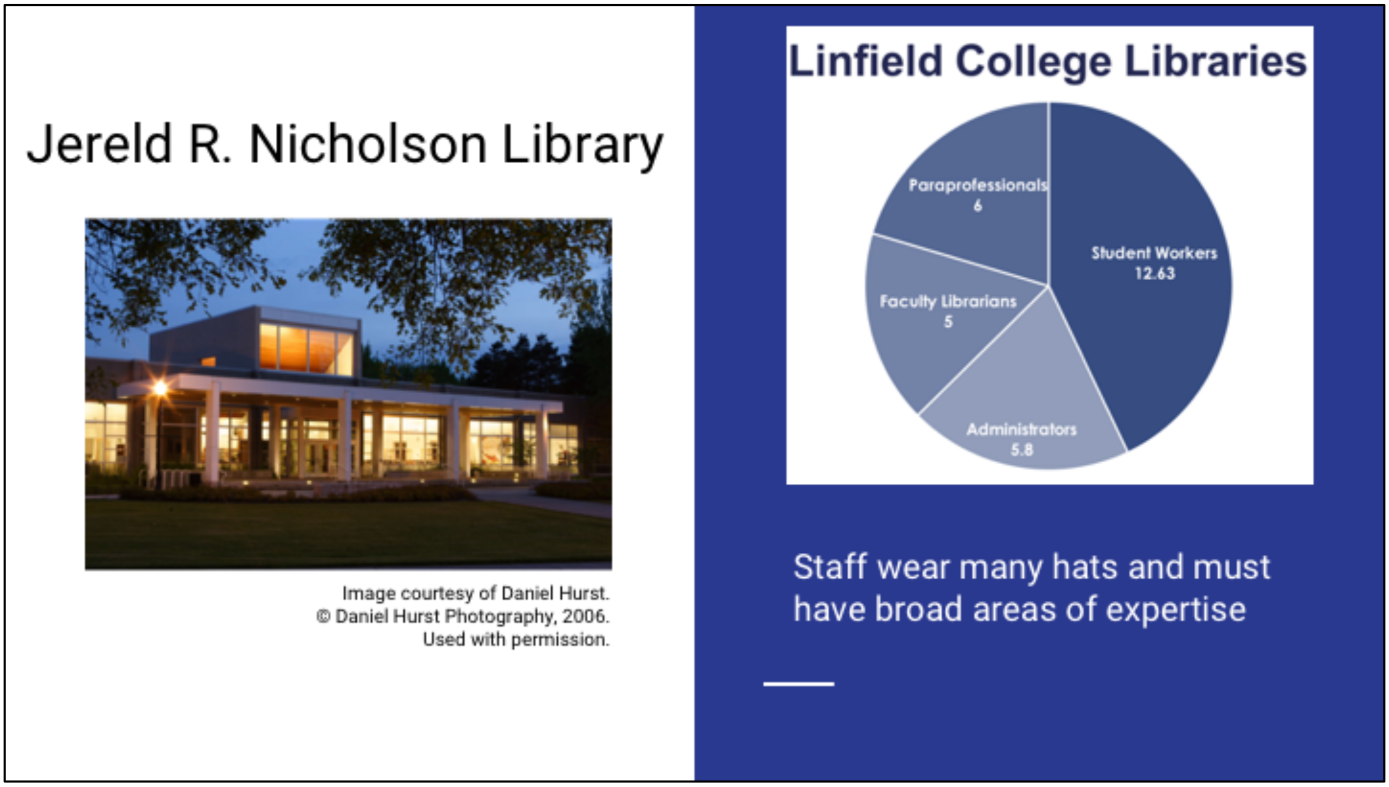
4-year private liberal arts college

3 campuses (McMinnville, Portland, Division of Online and Continuing Education)

2,282 students

Member of Orbis Cascade Alliance

Rather than diving into detail right away, Sara will tell you more about EAST and the tools she's used later in the presentation. But for now, let's just get our feet wet, and I'll do that by letting you know how we approach data visualization at Linfield College. Linfield is a small, private liberal arts college, founded in 1858, with its main campus in McMinnville, Oregon, right in the heart of wine country. We have a campus in Portland which houses our nursing program, and we also have a Division of Online & Continuing Education to deliver our distance ed programs. Together, all three campuses have a total enrollment of just under 2300, with a little more than 1600 students on the Mac campus. The college has a historical affiliation with the American Baptists.



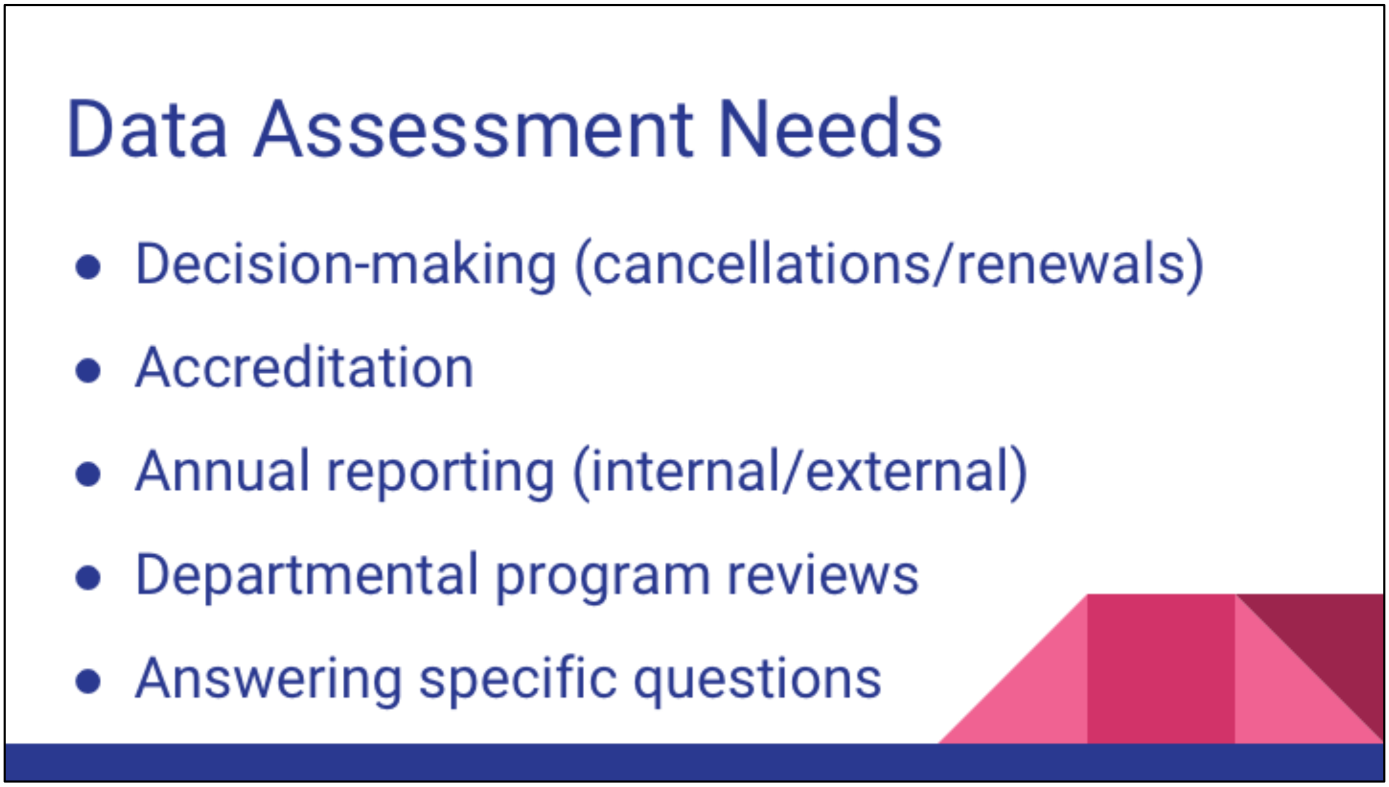
# Jereld R. Nicholson Library

Image courtesy of Daniel Hurst.
© Daniel Hurst Photography, 2006.
Used with permission.

## Linfield College Libraries

Staff wear many hats and must have broad areas of expertise

Collections management responsible for acquisitions (which means budgeting), cataloging, e-resource management, institutional repository. No true systems librarian on staff; no programmer on staff; no single person focused on assessment or statistics gathering.

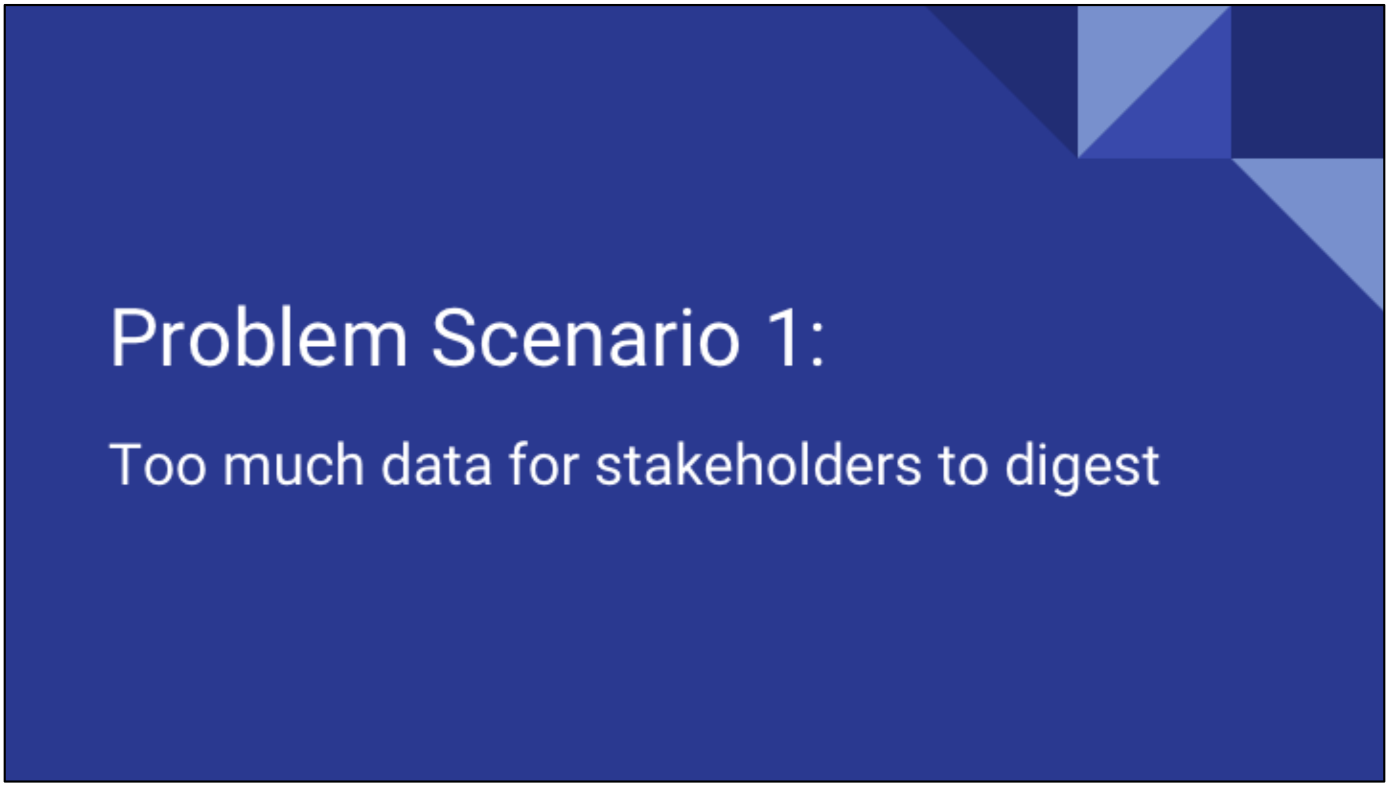


# Data Assessment Needs

- Decision-making (cancellations/renewals)
- Accreditation
- Annual reporting (internal/external)
- Departmental program reviews
- Answering specific questions

Assessment needs are local, consortial, and institutional; we utilize staff and student workers across the library to address these needs.

I'll walk through four problem scenarios briefly to give you a sense for a problem we've faced and the different ways we're trying to visualize our data in order to solve that problem. The hope is that you'll see something here or in Sara's portion later on that will click for you and that you'll be able to take back to your institution for further consideration.

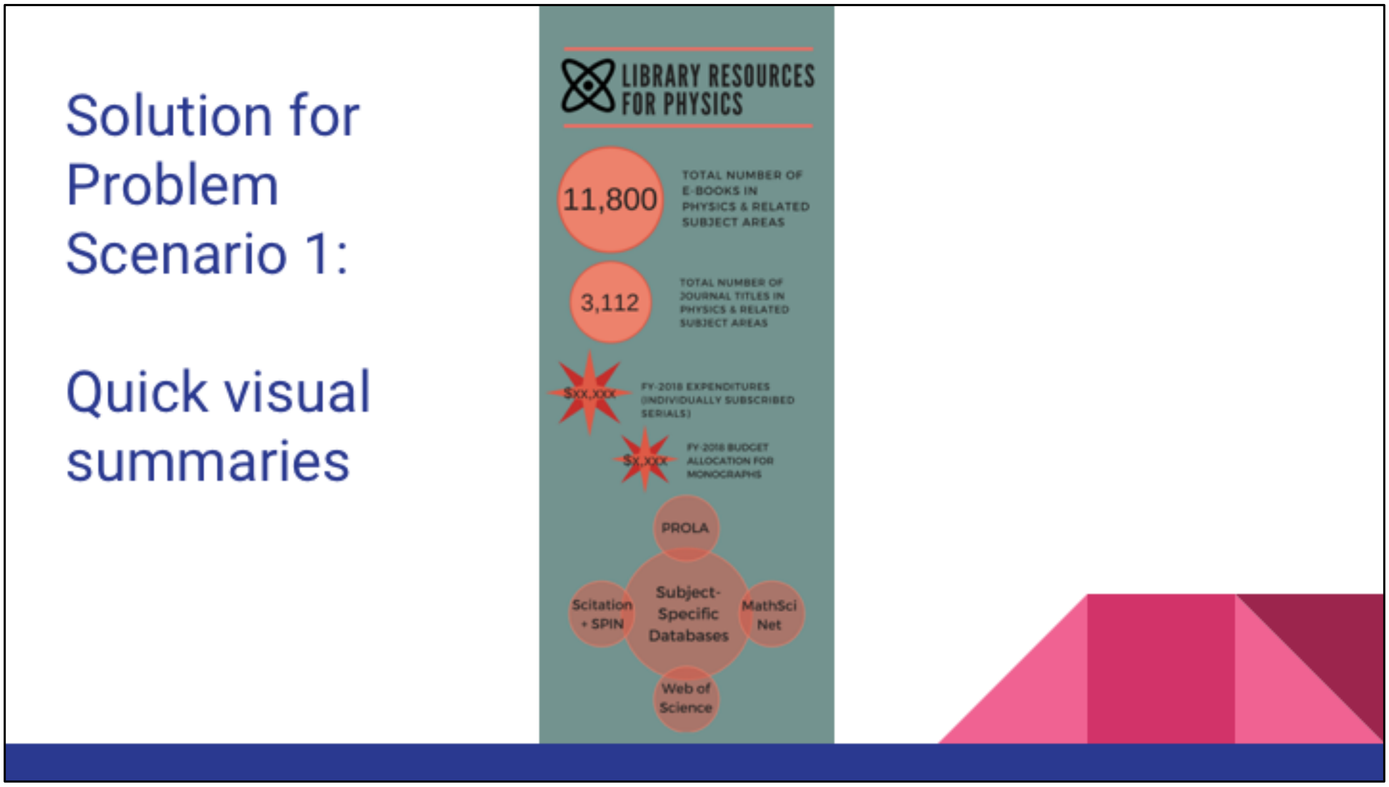


# Problem Scenario 1:

## Too much data for stakeholders to digest

Library is involved in departmental program reviews that happen every 7 to 10 years. An internal reviewer from another department at Linfield is appointed to lead the review, along with an external reviewer from a comparator school. The department produces a self-study that goes to the reviewers in advance, and the library has historically provided information on access to serials as part of that report. When the external reviewer comes to campus, they meet with a number of relevant stakeholders; the library usually gets about 45 minutes to discuss issues related to instruction partnerships, services, and collections.

The data we've given departments in the past has been unwieldy - a giant spreadsheet of titles listing access via different packages, along with any print subscriptions. It soundly ignored other collection components (such as monographs or subject-specific databases). I didn't even want to look at it, so I'd be surprised if the reviewers did. There was just too much data to digest.



# Solution for Problem Scenario 1:

## Quick visual summaries

LIBRARY RESOURCES FOR PHYSICS

11,800 TOTAL NUMBER OF E-BOOKS IN PHYSICS & RELATED SUBJECT AREAS

3,112 TOTAL NUMBER OF JOURNAL TITLES IN PHYSICS & RELATED SUBJECT AREAS

$xx,xxx FY-2018 EXPENDITURES (INDIVIDUALLY SUBSCRIBED SERIALS)

$x,xxx FY-2018 BUDGET ALLOCATION FOR MONOGRAPHS

Subject-Specific Databases

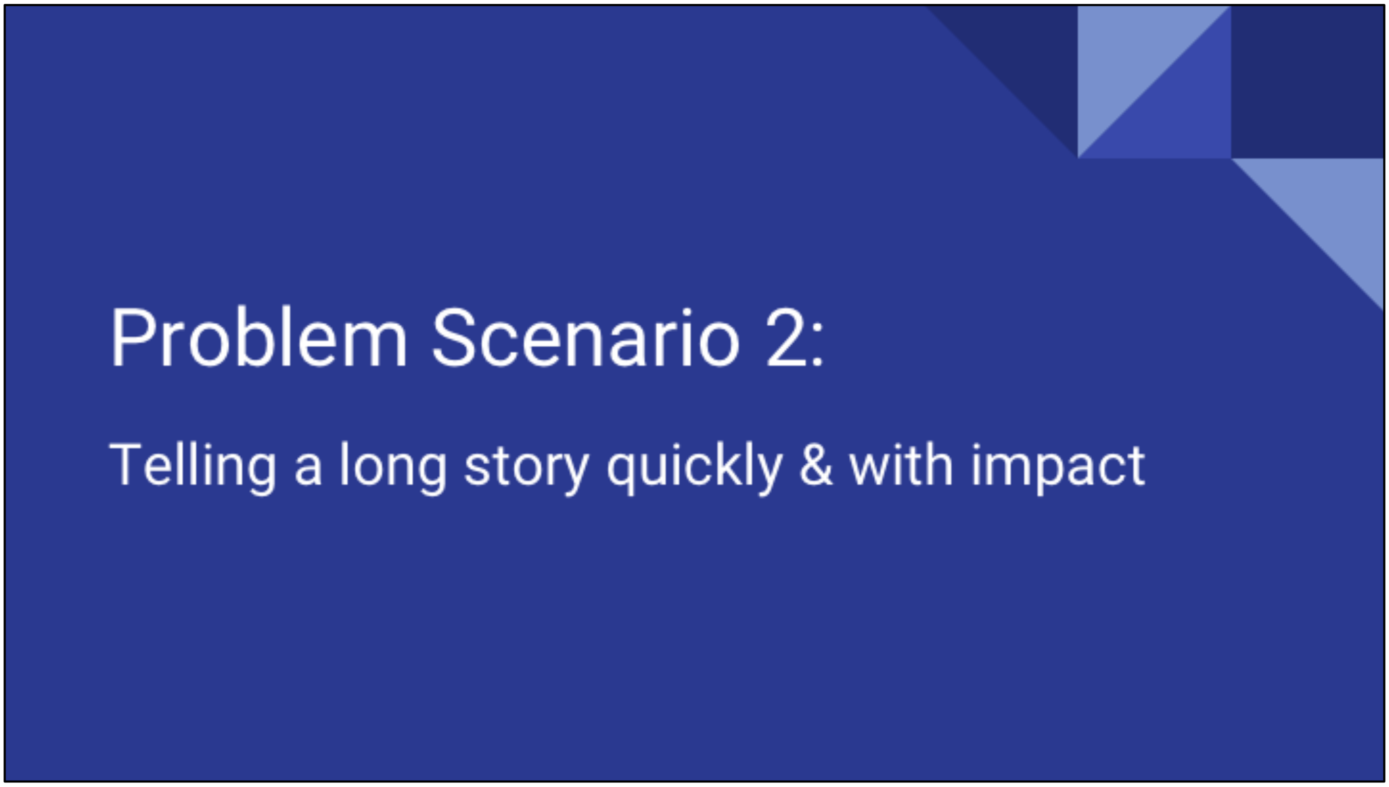
PROLA

Scitation + SPIN

MathSci Net

Web of Science

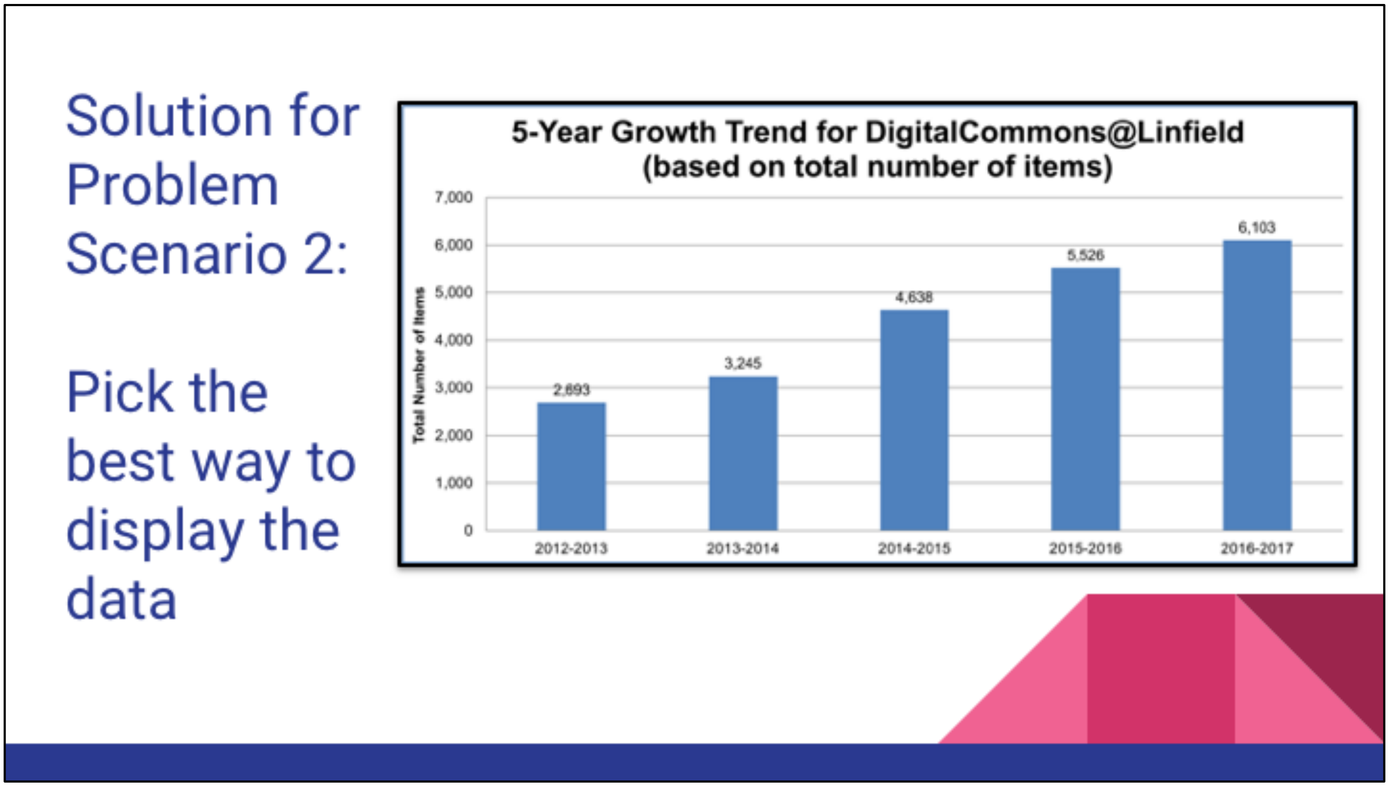
Idea is to provide a quick visual summary in addition to providing the full data for departments and external program reviewers - this is a less overwhelming way to present key data about resources and may spur questions that wouldn't otherwise have arisen because stakeholders open the spreadsheet, see the mass of data, and don't end up engaging with it at all. This type of summary can be iterated without too much customization for each new department, which means our return on investment is pretty good.



# Problem Scenario 2:

## Telling a long story quickly & with impact

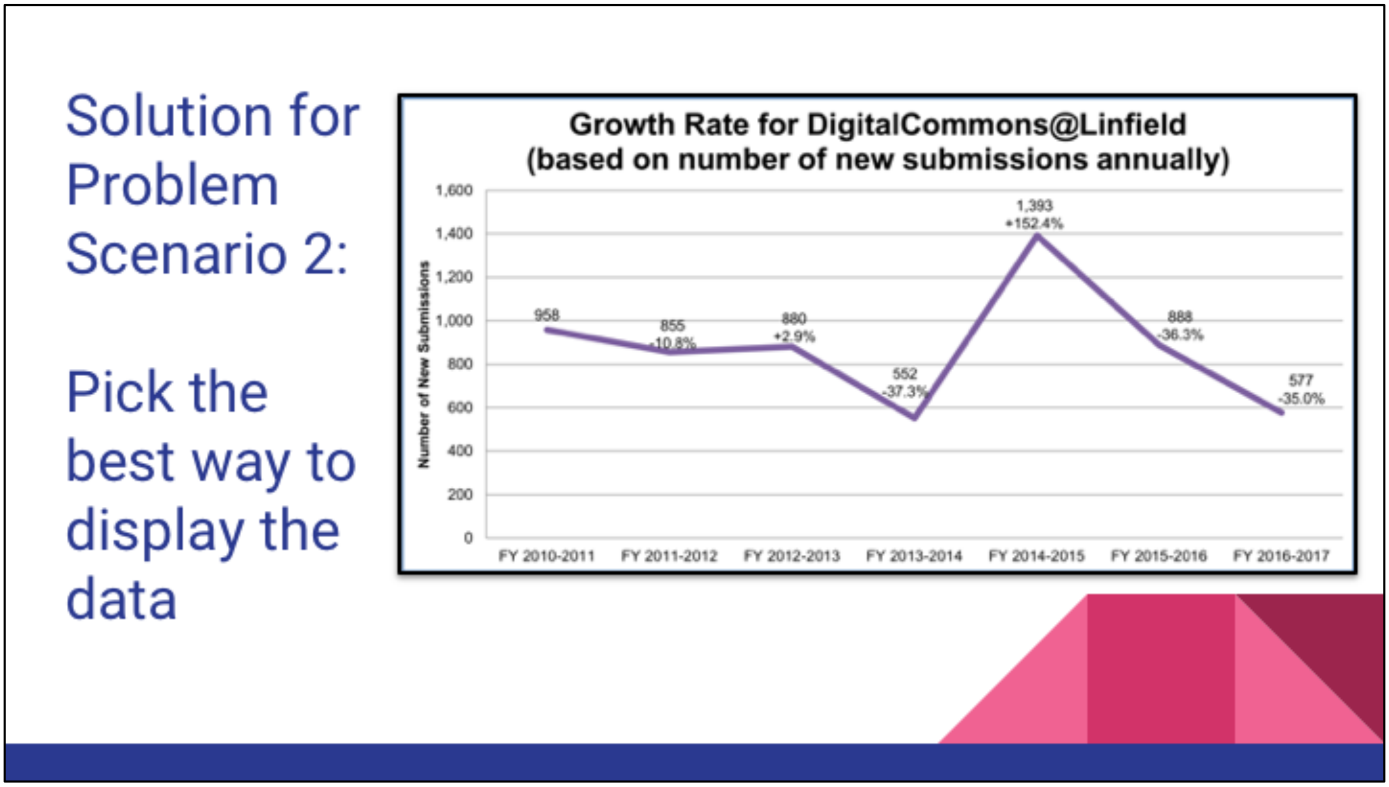
Sometimes you find yourself in a situation where you’re only going to have 2 or 3 minutes to explain an issue to someone and make them understand the importance of what they’re seeing. In this scenario, you need to be accurate, but you also need to be at your most persuasive. How can you do that quickly and with impact?



## Solution for Problem Scenario 2:

## Pick the best way to display the data

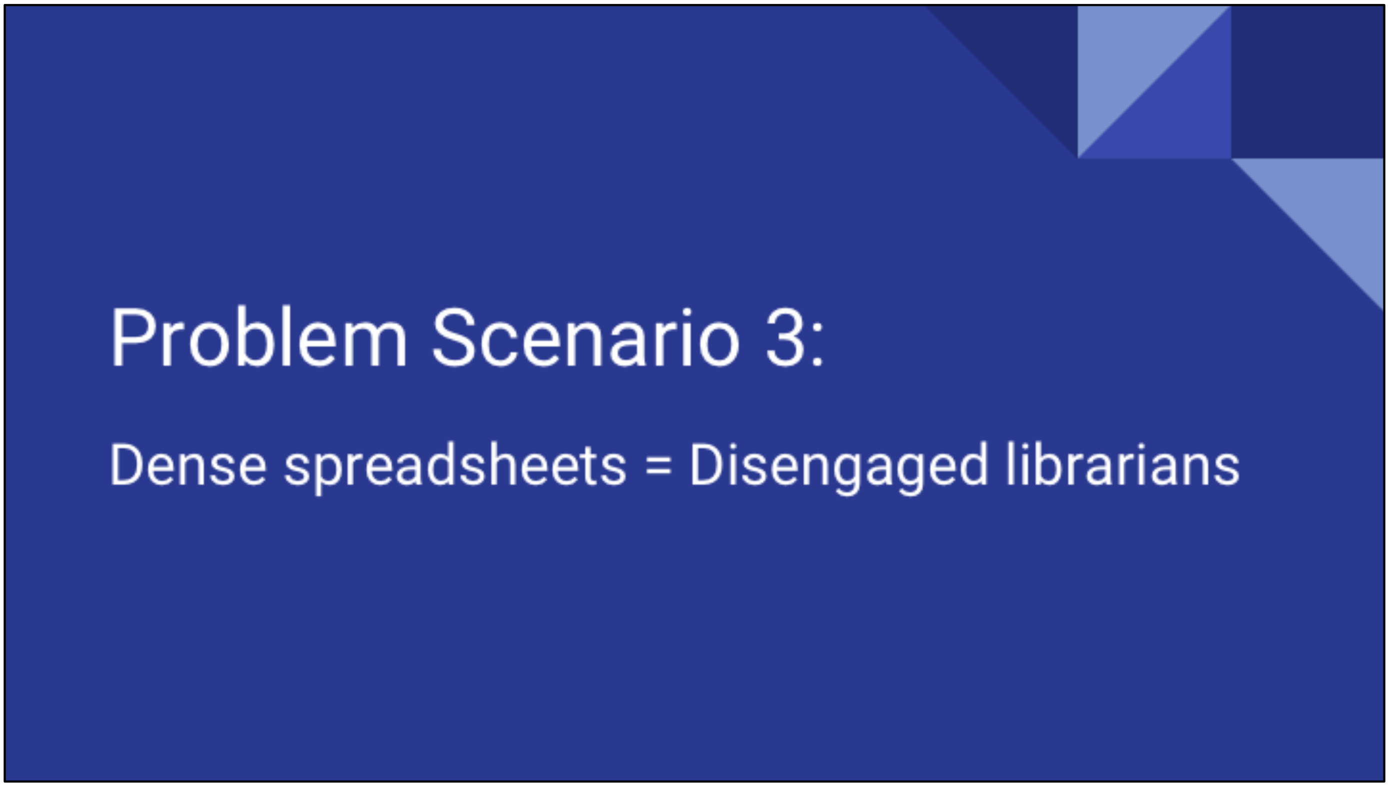
Consider this example. I wanted to make the case that there was a definite sweet spot for the level of growth for our institutional repository. A couple of years ago, I felt like the number of submissions we had added was way out of whack with previous years and was unsustainable given our existing staffing. So, I turned to the data to see if my theory was supported. Here the data show a fairly steady climb in terms of growth across fiscal years - the gap between FY 13-14 and FY 14-15 is bigger than the others but doesn't look arrestingly large. This graphic is true, but it wasn't persuasive in the way I wanted it to be.



## Solution for Problem Scenario 2:

## Pick the best way to display the data

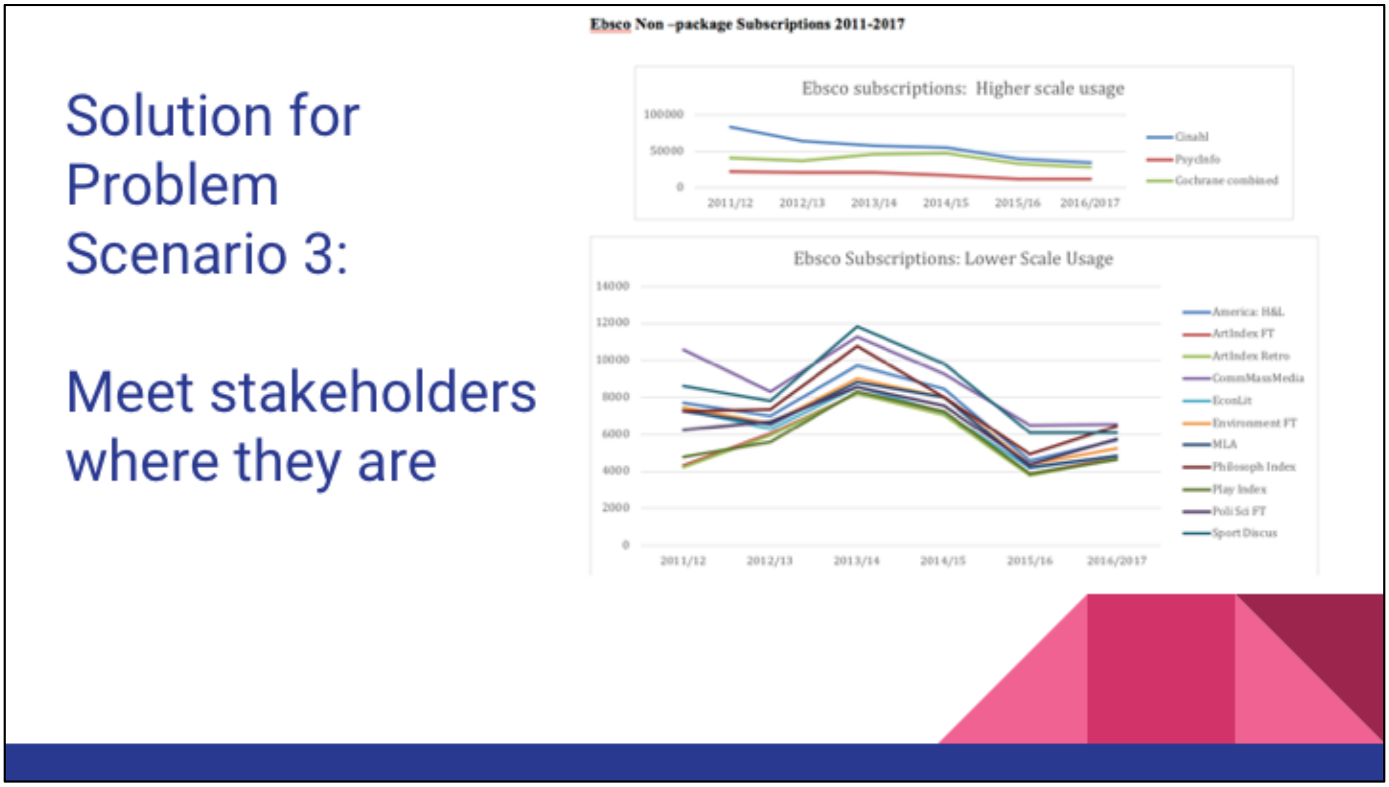
So I created an equally true alternate representation from a slightly different set of data that get at the same general concept. With this graph, it becomes much easier to see large swings in terms of the number of items contributed (which translates to large swings in workload). It's also easier to see where that "sweet spot" is across years. And, this version makes it easier to recognize where other factors contributed to decreased growth (in FY13-14, it was our ILS migration; in FY16-17, it was my promotion and tenure application).



# Problem Scenario 3:

## Dense spreadsheets = Disengaged librarians

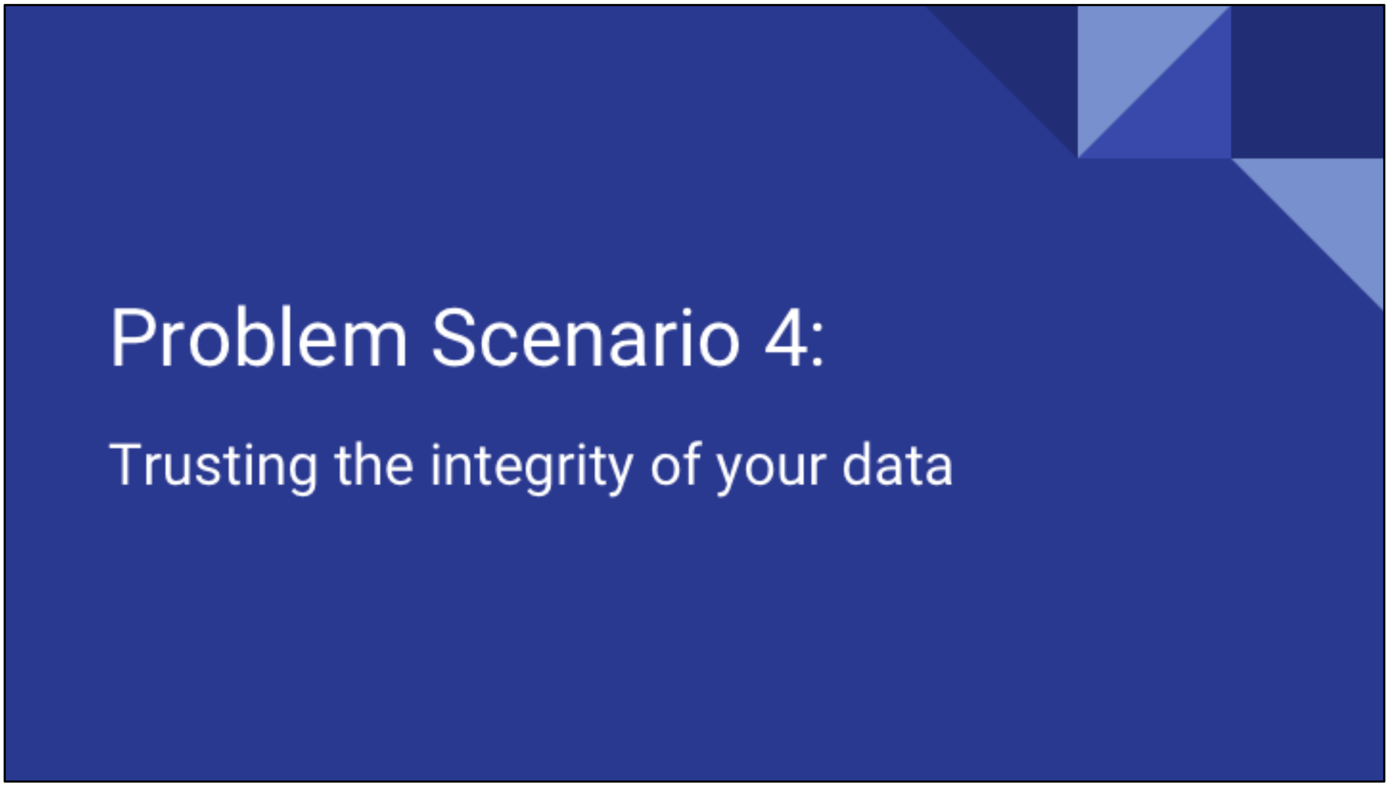
External stakeholders don’t usually relish dense data, and internal stakeholders aren’t really all that different. What do you do to avoid the disengaged librarians trap when, for example, you’re reviewing continuing resources?



## Solution for Problem Scenario 3:

## Meet stakeholders where they are

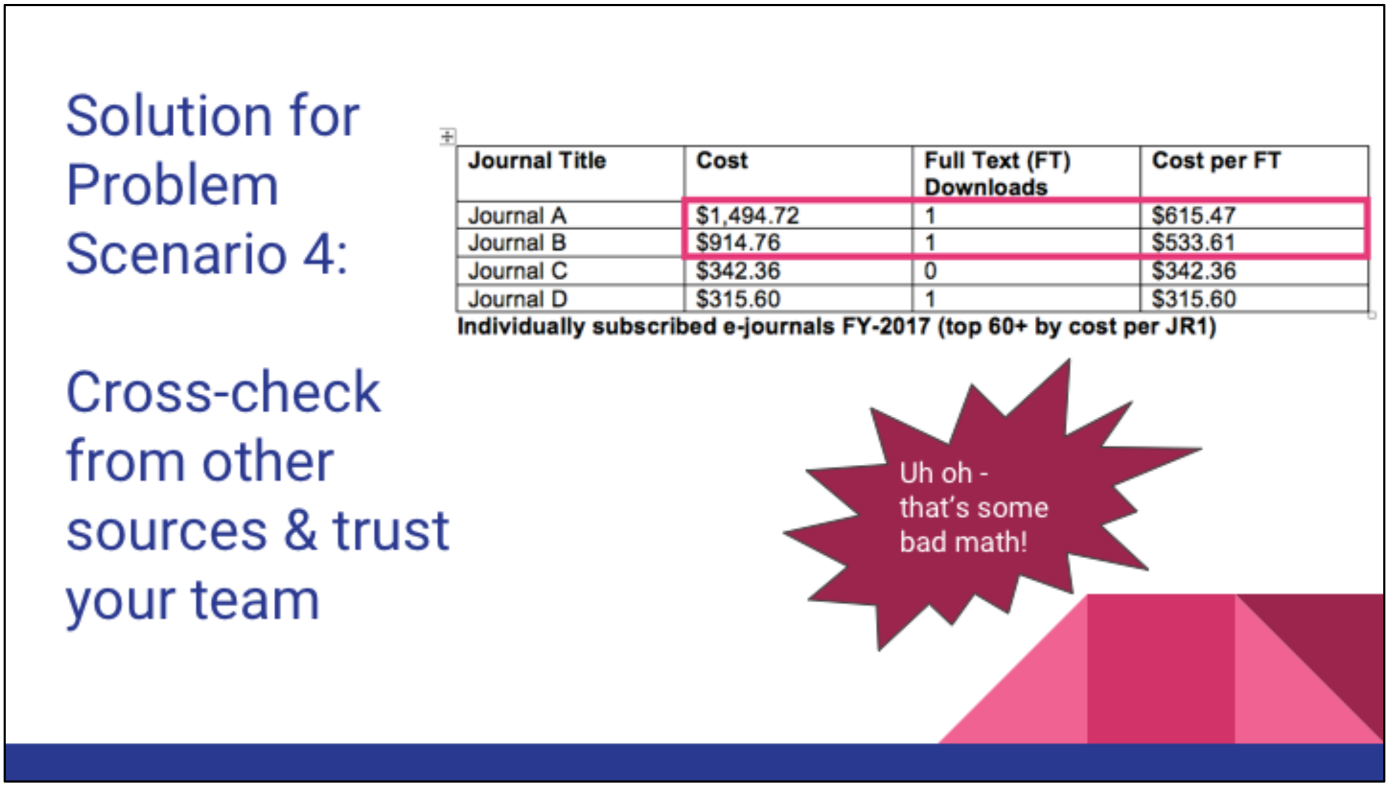
Our solution to this problem is to try to meet our stakeholders where they are. Annual e-resources report translates dense data spreadsheets into digestible chunks, visualizing COUNTER usage statistics for e-journals and databases, and providing tabular data for e-reference and other selected e-books. This report includes brief analyses to accompany the data, with bullet points highlighting resources to watch or high-performing resources. Information gleaned from this report is used throughout the year as continuing resources come up for renewal, providing supporting evidence for conversations with departments about what resources to keep/drop. The report also includes a high-level look at search and other activity in our discovery layer (Primo) which informs discussions about which (if any) local customizations we want to make to elements like facets.



# Problem Scenario 4:

## Trusting the integrity of your data

The last problem I want to address is one which plagues institutions of all sizes - how do you trust the accuracy of your data?



## Solution for Problem Scenario 4:

## Cross-check from other sources & trust your team

| Journal Title | Cost | Full Text (FT) Downloads | Cost per FT |
|---|---|---|---|
| Journal A | $1,494.72 | 1 | $615.47 |
| Journal B | $914.76 | 1 | $533.61 |
| Journal C | $342.36 | 0 | $342.36 |
| Journal D | $315.60 | 1 | $315.60 |

**Individually subscribed e-journals FY-2017 (top 60+ by cost per JR1)**

Uh oh - that's some bad math!

For us, the solution is to cross-check your data and trust your team.

Truth be told, this scenario is less about how to visualize data and more about *why* you should visualize data. Sometimes looking at individual data points out of context doesn't reveal when there are larger issues with the integrity of data. Even when you look at data points in concert with other data points, it's still possible to miss glaring problems, as was the case with this cost per full text data provided by our ILS vendor. This should be a simple calculation, pulling from data already in our ILS, but as you can see, there's some pretty bad math going on here when cost divided by the number of full-text downloads doesn't pencil out consistently. Relying on team members with vested interests in subsets of the data to review information more closely can be a lifesaver for smaller institutions. If you have access to data from different sources that will allow you to cross-check the accuracy, it's a good idea to take the time to do some spot-checks. Ultimately, data integrity isn't always a problem you can solve 100%, but using different kinds of visualization approaches to look at the data can help to reveal where there might be issues you didn't even know existed.

EAST
Serials / Journals

The search for a viable collection analysis tool

Center for Research Libraries
GLOBAL RESOURCES NETWORK

EAST EASTERN ACADEMIC SCHOLARS' TRUST

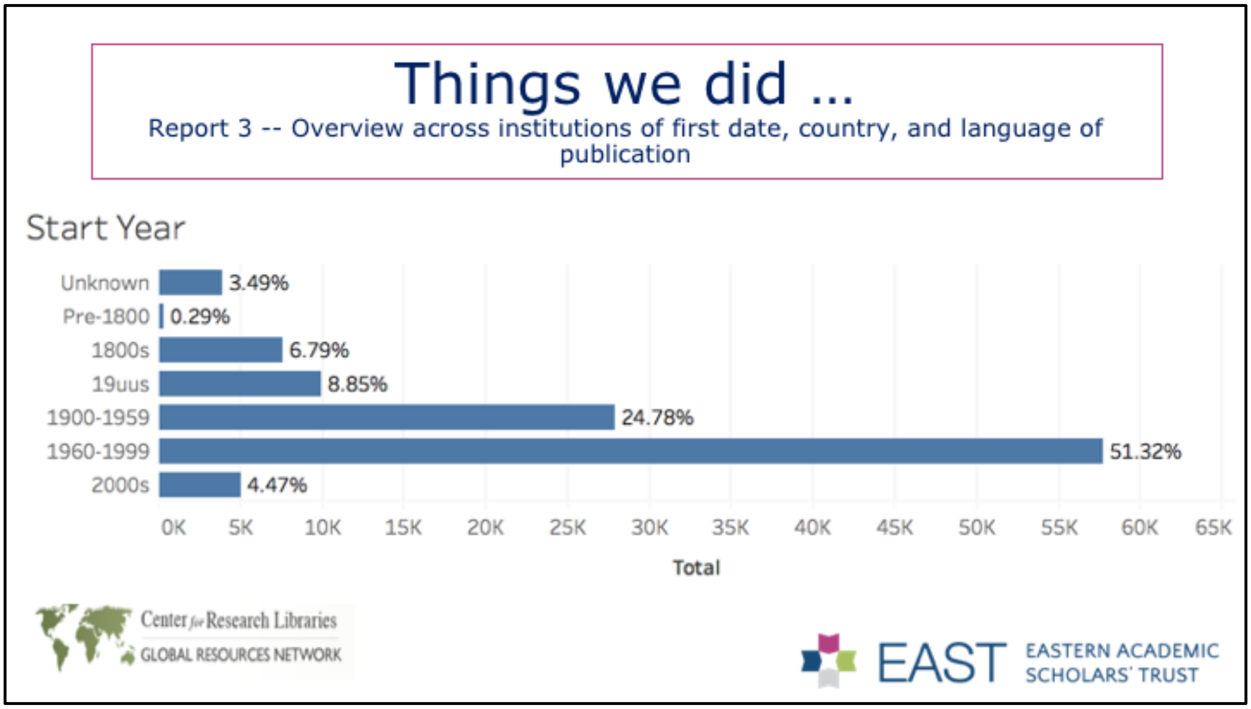
Kathleen has presented some nice visualizations using a variety of data and tools.
The story I'm going to tell is about one monster spreadsheet, with data across 21 tabs, and needing to use it for collection analysis.

After completing our first cohort's monograph retention commitments we turned our attention to Serials and Journals in an effort to accomplish some work on this before the Mellon grant funding expires. The largest question was of course how to do the collection analysis. After discussions with OCLC/SCS (who at that point were discussing developing a tool), WEST, and CRL, we determined that given our budget and time frame that using CRL for our collection analysis was our only viable option.

CRL had the advantage of being available to do the work almost immediately, had a good history of dealing with dirty serials holdings data, being the host for the PAPR database and thus making registration easy, and also highly affordable as 14 of our initial 21 serials and journals partners are also CRL members and were not charged for their participation.

The downside was that data returned involved a series of data-dense spreadsheets that were not easily digestible.

I also want to put a caveat here, that while I deal with data, I am by no means a visualization expert - this work done here was all by the seat of my pants over the last year, which means: a) there are definitely people in this room I can learn from, and b) if you are saying "no way can I do that," you might be surprised at how attenable some of this is with a little bit of focused time. I'm still learning all the time with this and find that there is a lot of community support.



# Things we did …

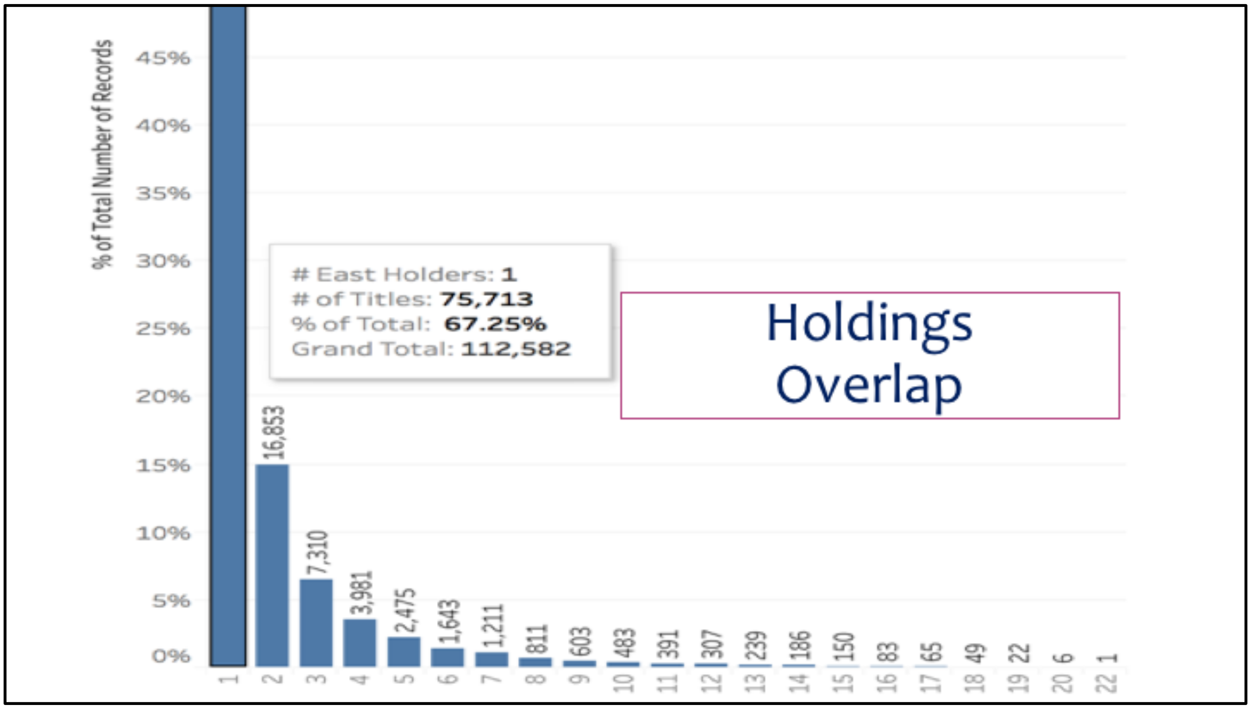
Report 3 -- Overview across institutions of first date, country, and language of publication

Center for Research Libraries
GLOBAL RESOURCES NETWORK

EAST EASTERN ACADEMIC SCHOLARS' TRUST

First I'm going to step through the end result of what we did, and then go back and talk about the tools to get there. All of these visualizations were created using Tableau and were published in the public Tableau server, and there is a link in your handout if you want to see them live.

Here is a bar chart of date for first publication per journal held. 3.49% were unknown, the rest grouped somewhat arbitrarily but the best we could do with the data, e.g. 19uu, 1900-1960, 1960-1999, 2000s.   Not surprisingly the majority were in the latter half of the 20th century.

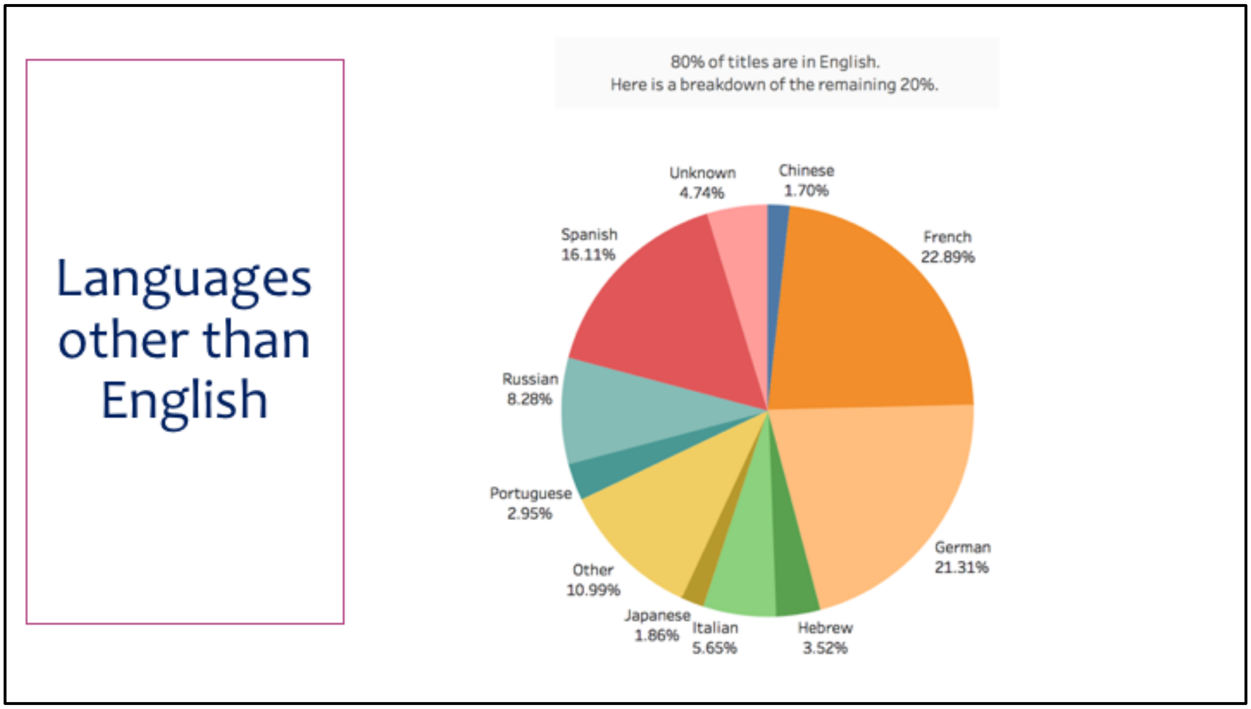


Holdings Overlap

# East Holders: **1**
# of Titles: **75,713**
% of Total: **67.25%**
Grand Total: **112,582**

And here is holdings overlap - note that you can make nice tooltips that have other calculations in them. Here we can see number of holders, number of titles in that range, percent of total, which is also seen by the height of the bar, and the total number of titles we were dealing with.

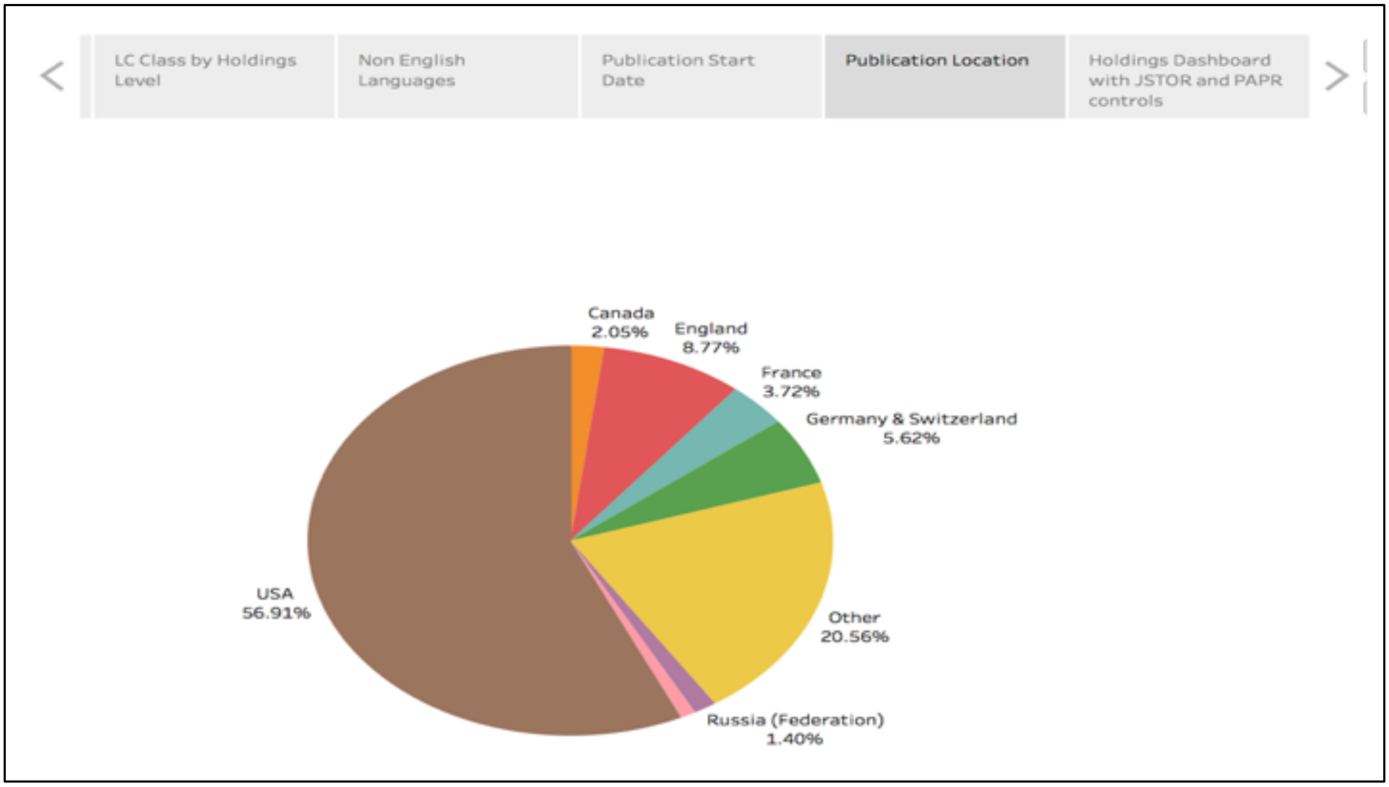
We can see from this that there was a high level of uniqueness - 67% held by only one institution. After mulling this over the SJWG (Serials & Journals Working Group) decided to initially focus on 'medium rare' titles in the 4-6 holdings range, giving some attention to preservation of 'rare' titles while still offering some weeding opportunities. (Note that a lot of the 'unique' titles were cataloging anomalies or locally important titles, e.g. Catalog of xyz university, and/or were special collections titles, which were not at risk of withdrawal.)



# Languages other than English

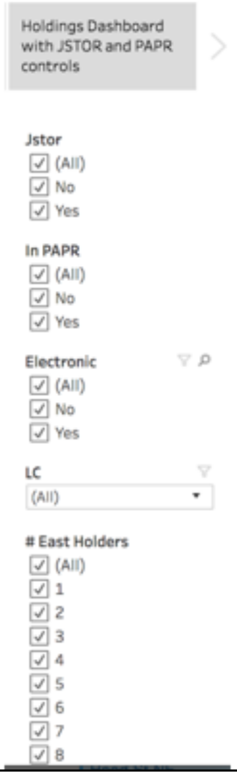
80% of titles are in English.
Here is a breakdown of the remaining 20%.

Data included language of publication - looking at titles that were non-English, we can quickly see they were mostly French and German, followed by Spanish. That 11% ‘other’ is a large list we’ll look at in a minute. This was just a quick and easy way to see your data in a way you can immediately synthesize rather than a spreadsheet. Honestly, we didn’t end up doing anything with this data, but it was nice to see just to have an overview and idea of what we had.


LC Class by Holdings Level
Non English Languages
Publication Start Date
Publication Location
Holdings Dashboard with JSTOR and PAPR controls
Canada 2.05%
England 8.77%
France 3.72%
Germany & Switzerland 5.62%
Other 20.56%
Russia (Federation) 1.40%
USA 56.91%

These visualizations can be pulled into a story dashboard in tableau, where you can click through titles at the top and tell the story of the data. This was used in our initial meetings just to look at the data returned to us and see where we wanted to dive deeper, e.g., LC class data, overlap with JSTOR and PAPR, languages, and date. It was a nice way to review the data rather than staring at spreadsheets.

# Interactive Dashboards can be added to the story board



Holdings Dashboard with JSTOR and PAPR controls

Jstor
☑ (All)
☑ No
☑ Yes

In PAPR
☑ (All)
☑ No
☑ Yes

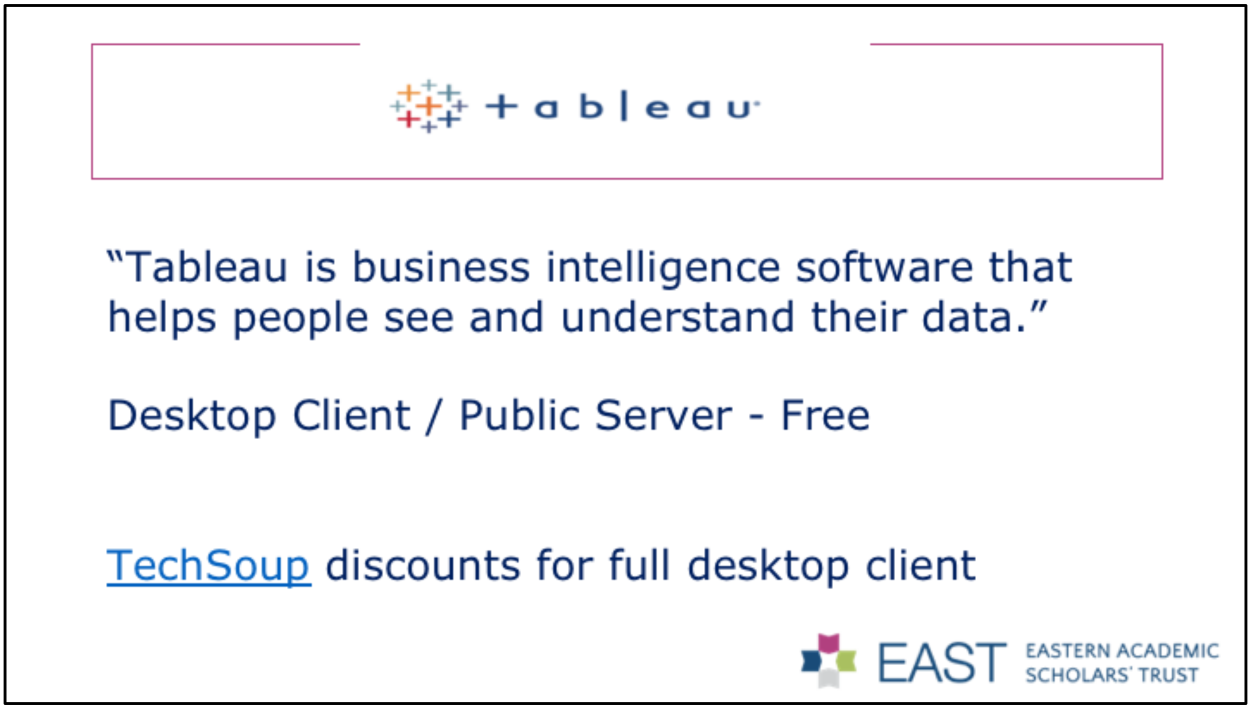
Electronic
☑ (All)
☑ No
☑ Yes

LC
(All)

# East Holders
☑ (All)
☑ 1
☑ 2
☑ 3
☑ 4
☑ 5
☑ 6
☑ 7
☑ 8

EAST EASTERN ACADEMIC SCHOLARS' TRUST

Interactive filters can give you ways to play with your data easily. We also used these filters to create title lists for each library to review when considering parameters for our retention program, e.g., a library could download the full list of titles that fit into a retention scenario they were considering and determine if the list looked like something they would be willing to retain.



tableau

"Tableau is business intelligence software that helps people see and understand their data."

Desktop Client / Public Server - Free

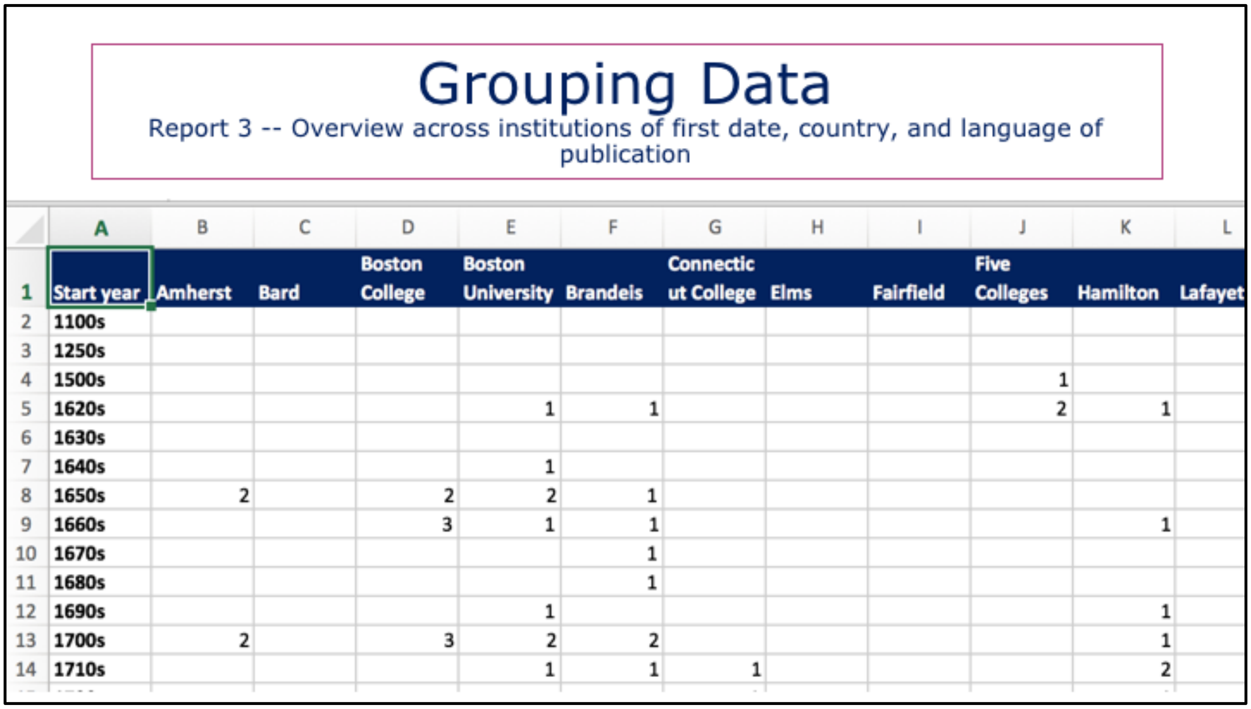
TechSoup discounts for full desktop client

EAST EASTERN ACADEMIC SCHOLARS' TRUST

As I mentioned, we used Tableau to create all of these vizzes.

EAST already had some members using Tableau and encouraged the project team to explore its use. Tableau bills itself as "business intelligence software that helps people see and understand their data" and can be used to build everything from basic graphs and charts to dynamic data dashboards drawing from cloud-based datasets.

There is a desktop client in which the visualizations are built, and these can then be hosted on either a free public server or a for-fee private server. Our data was not sensitive and resides on the public server.  There is a free limited version of the desktop client, though as non-profit educational institutions we are eligible for a reduced fee license through TechSoup. We paid $58 for a two-year license to the full-feature desktop client, which has the great advantage of coming with tech support.



# Grouping Data

Report 3 -- Overview across institutions of first date, country, and language of publication

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Start year | Amherst | Bard | Boston College | Boston University | Brandeis | Connecticut College | Elms | Fairfield | Five Colleges | Hamilton | Lafayet |
| 2 | 1100s | | | | | | | | | | | |
| 3 | 1250s | | | | | | | | | | | |
| 4 | 1500s | | | | | | | | | 1 | | |
| 5 | 1620s | | | | 1 | 1 | | | | 2 | 1 | |
| 6 | 1630s | | | | | | | | | | | |
| 7 | 1640s | | | | 1 | | | | | | | |
| 8 | 1650s | 2 | | 2 | 2 | 1 | | | | | | |
| 9 | 1660s | | | 3 | 1 | 1 | | | | | 1 | |
| 10 | 1670s | | | | | 1 | | | | | | |
| 11 | 1680s | | | | | 1 | | | | | | |
| 12 | 1690s | | | | 1 | | | | | | 1 | |
| 13 | 1700s | 2 | | 3 | 2 | 2 | | | | | 1 | |
| 14 | 1710s | | | | 1 | 1 | 1 | | | | 2 | |

Looking behind the scenes at some of the vizzes we saw earlier, here is the data of first date of publication spreadsheet.

CRL's report 3 showed first date of publication, country, and language of publication. You can see here that data was based in columns and not aggregated, e.g., for the 1650s, Amherst, Boston College, and BU all had two titles, Brandeis had 1, etc.

# Tableau Groupings

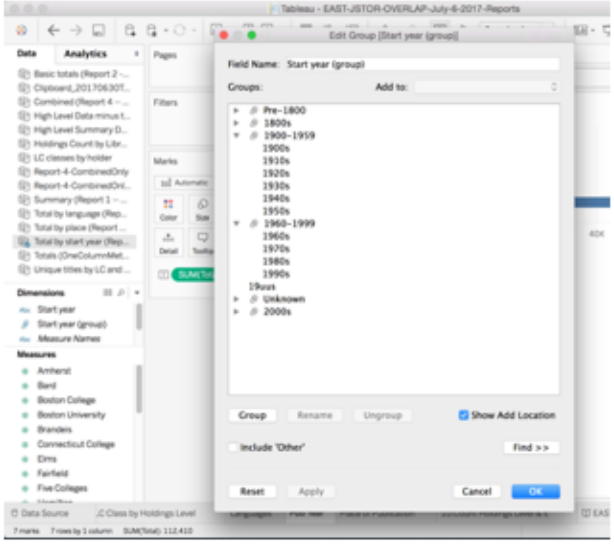


Tableau - EAST-JSTOR-OVERLAP-July-6-2017-Reports

Edit Group [Start year (group)]

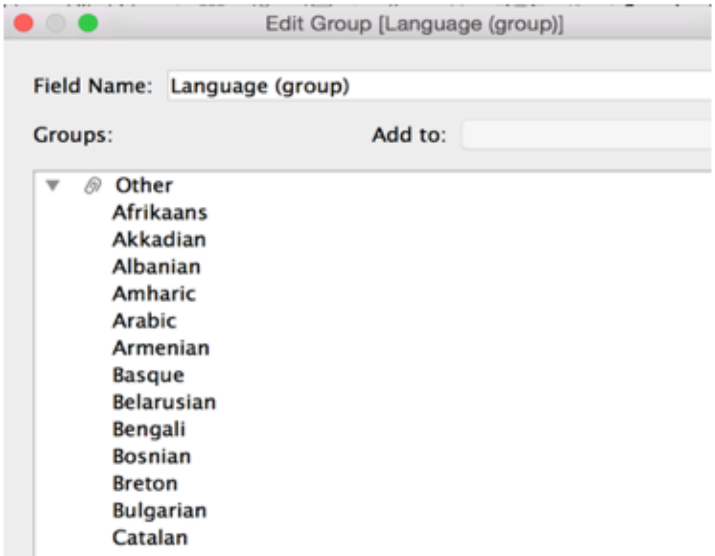
Field Name: Start year (group)

Groups: Add to:

- Pre-1800
- 1800s
- 1900-1959
  - 1900s
  - 1910s
  - 1920s
  - 1930s
  - 1940s
  - 1950s
- 1960-1999
  - 1960s
  - 1970s
  - 1980s
  - 1990s
- 19uus
- Unknown
- 2000s

☑ Show Add Location

□ Include 'Other'



Edit Group [Language (group)]

Field Name: Language (group)

Groups: Add to:

- Other
  - Afrikaans
  - Akkadian
  - Albanian
  - Amharic
  - Arabic
  - Armenian
  - Basque
  - Belarusian
  - Bengali
  - Bosnian
  - Breton
  - Bulgarian
  - Catalan

Tableau made it easy to pull in data from that spreadsheet and group it into categories. On the left here we grouped the decades into half centuries. We simply told Tableau to lump those all together, and they appear as a single category in the visualization.

Same thing with languages on the right - we took all the small languages and grouped them into an “other” category that we saw earlier in the pie chart. Other included Afrikaans, Akkadian, Albanian, etc.

# Reformatting for Database Structure

OPEN Refine

A free, open source, powerful tool for working with messy data

library carpentry

Library Carpentry

Software and data skills for library professionals

We pause this presentation to bring you a brief ad for OpenRefine. Who has used OpenRefine?

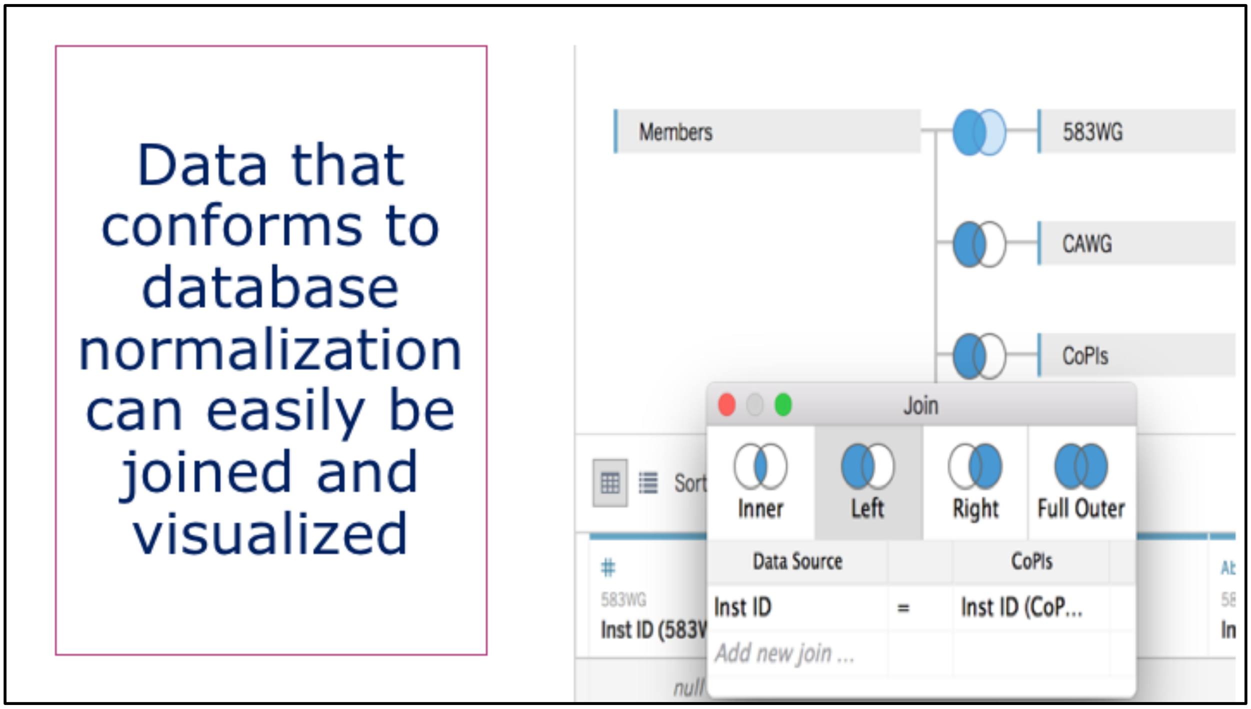
Data doesn't always come to us in a format that translates easily to visualization. We just saw that Tableau can do some nice grouping for you, but sometimes you just can't group or divide things or organize them how you want with the data in the format it was given to you.

Things you can do with OpenRefine include transforming data using regular expressions, splitting and joining cells based on delimiters, transposing rows to columns or columns to rows, faceting columns, and cleaning/merging data - useful for authority work. You can also use it to compare data to an authority file.

It is an extremely easy and user-friendly way to reformat data. You download the program and it runs locally in your browser and allows you to pull in Excel, CSV, XML, text, or JSON and parse and transform it in a variety of ways.

Also, there are an increasing number of library carpentry and data carpentry classes being offered which can get you up and running with basic data manipulation skills, e.g., regular expressions, structured query language, Unix command line tools, and, of course, OpenRefine. A few hours learning the tools can save countless hours of data manipulation.

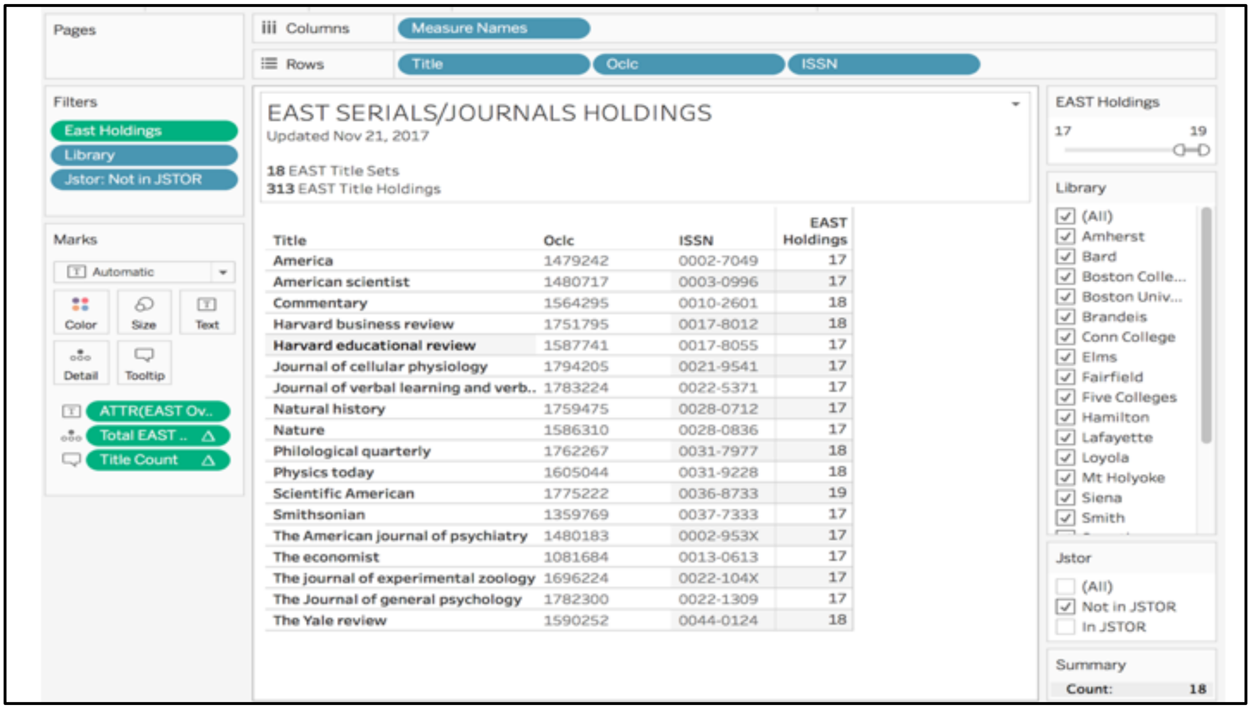
If you are dealing with large datasets with many relationships between the data, it is also worth spending an hour reading up on database normalization and making sure your data conforms. (See the Wikipedia link in the handout.) It can save lots of time down the road when trying to figure out how to make the visualization you want! Again, OpenRefine can make it easy to split/transform data into an arrangement that lets you get at the relationships you want.



Data that conforms to database normalization can easily be joined and visualized

| Data Source | | CoPIs |
|---|---|---|
| Inst ID | = | Inst ID (CoP... |
| *Add new join ...* | | |

With normalized data, Tableau can be used to create database-type structures, joining spreadsheet tabs on ID fields and making interactive visualizations.



Filters
East Holdings
Library
Jstor: Not in JSTOR

EAST SERIALS/JOURNALS HOLDINGS
Updated Nov 21, 2017

**18** EAST Title Sets
**313** EAST Title Holdings

| Title | Oclc | ISSN | EAST Holdings |
|---|---|---|---|
| America | 1479242 | 0002-7049 | 17 |
| American scientist | 1480717 | 0003-0996 | 17 |
| Commentary | 1564295 | 0010-2601 | 18 |
| Harvard business review | 1751795 | 0017-8012 | 18 |
| Harvard educational review | 1587741 | 0017-8055 | 17 |
| Journal of cellular physiology | 1794205 | 0021-9541 | 17 |
| Journal of verbal learning and verb.. | 1783224 | 0022-5371 | 17 |
| Natural history | 1759475 | 0028-0712 | 17 |
| Nature | 1586310 | 0028-0836 | 17 |
| Philological quarterly | 1762267 | 0031-7977 | 18 |
| Physics today | 1605044 | 0031-9228 | 18 |
| Scientific American | 1775222 | 0036-8733 | 19 |
| Smithsonian | 1359769 | 0037-7333 | 17 |
| The American journal of psychiatry | 1480183 | 0002-953X | 17 |
| The economist | 1081684 | 0013-0613 | 17 |
| The journal of experimental zoology | 1696224 | 0022-104X | 17 |
| The Journal of general psychology | 1782300 | 0022-1309 | 17 |
| The Yale review | 1590252 | 0044-0124 | 18 |

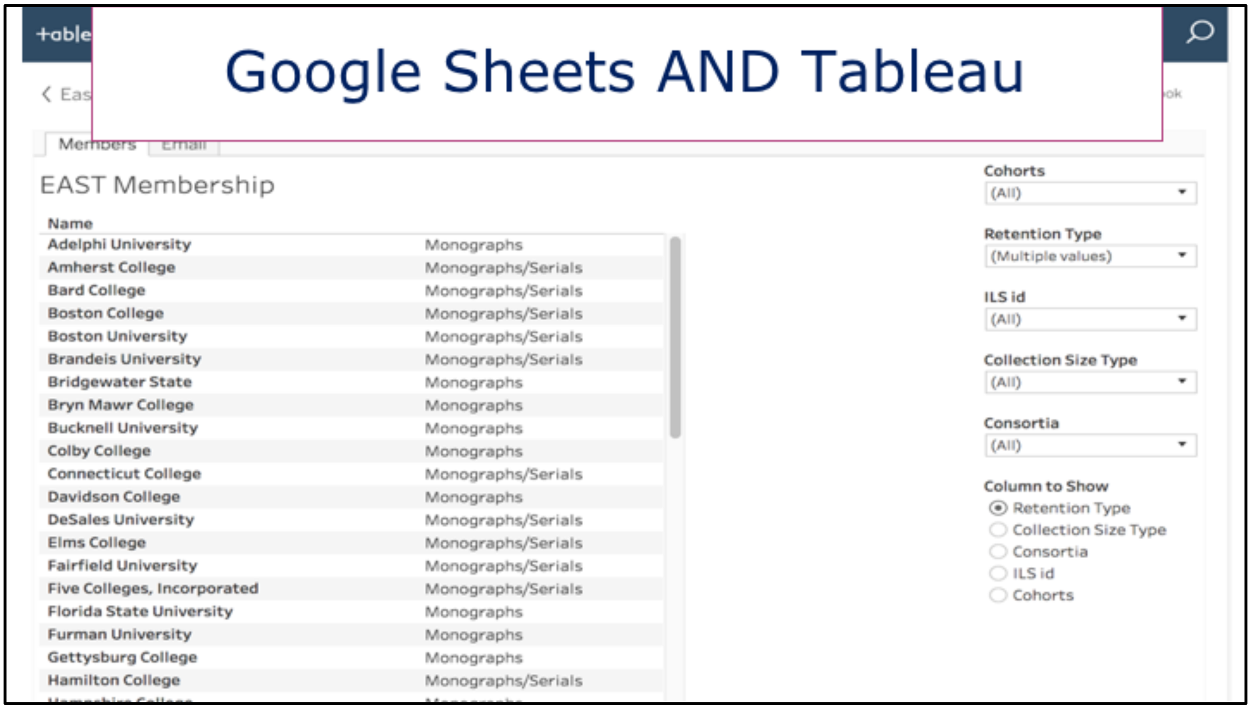
EAST Holdings
17 19

Library
☑ (All)
☑ Amherst
☑ Bard
☑ Boston Colle...
☑ Boston Univ...
☑ Brandeis
☑ Conn College
☑ Elms
☑ Fairfield
☑ Five Colleges
☑ Hamilton
☑ Lafayette
☑ Loyola
☑ Mt Holyoke
☑ Siena
☑ Smith

Jstor
□ (All)
☑ Not in JSTOR
□ In JSTOR

Summary
Count: 18

With the data transformed to one row per holding, which was not the way the data was delivered but was not too painful to achieve using OpenRefine, I could pull the holdings into Tableau and create filters/facets (basically pivot tables) to do some modeling. Here's the desktop client showing filters for holdings level, holding library, and if the title is in JSTOR.



## Google Sheets AND Tableau

EAST Membership

| Name | |
|---|---|
| Adelphi University | Monographs |
| Amherst College | Monographs/Serials |
| Bard College | Monographs/Serials |
| Boston College | Monographs/Serials |
| Boston University | Monographs/Serials |
| Brandeis University | Monographs/Serials |
| Bridgewater State | Monographs |
| Bryn Mawr College | Monographs |
| Bucknell University | Monographs |
| Colby College | Monographs |
| Connecticut College | Monographs/Serials |
| Davidson College | Monographs |
| DeSales University | Monographs/Serials |
| Elms College | Monographs/Serials |
| Fairfield University | Monographs/Serials |
| Five Colleges, Incorporated | Monographs/Serials |
| Florida State University | Monographs |
| Furman University | Monographs |
| Gettysburg College | Monographs |
| Hamilton College | Monographs/Serials |

Cohorts: (All)

Retention Type: (Multiple values)

ILS id: (All)

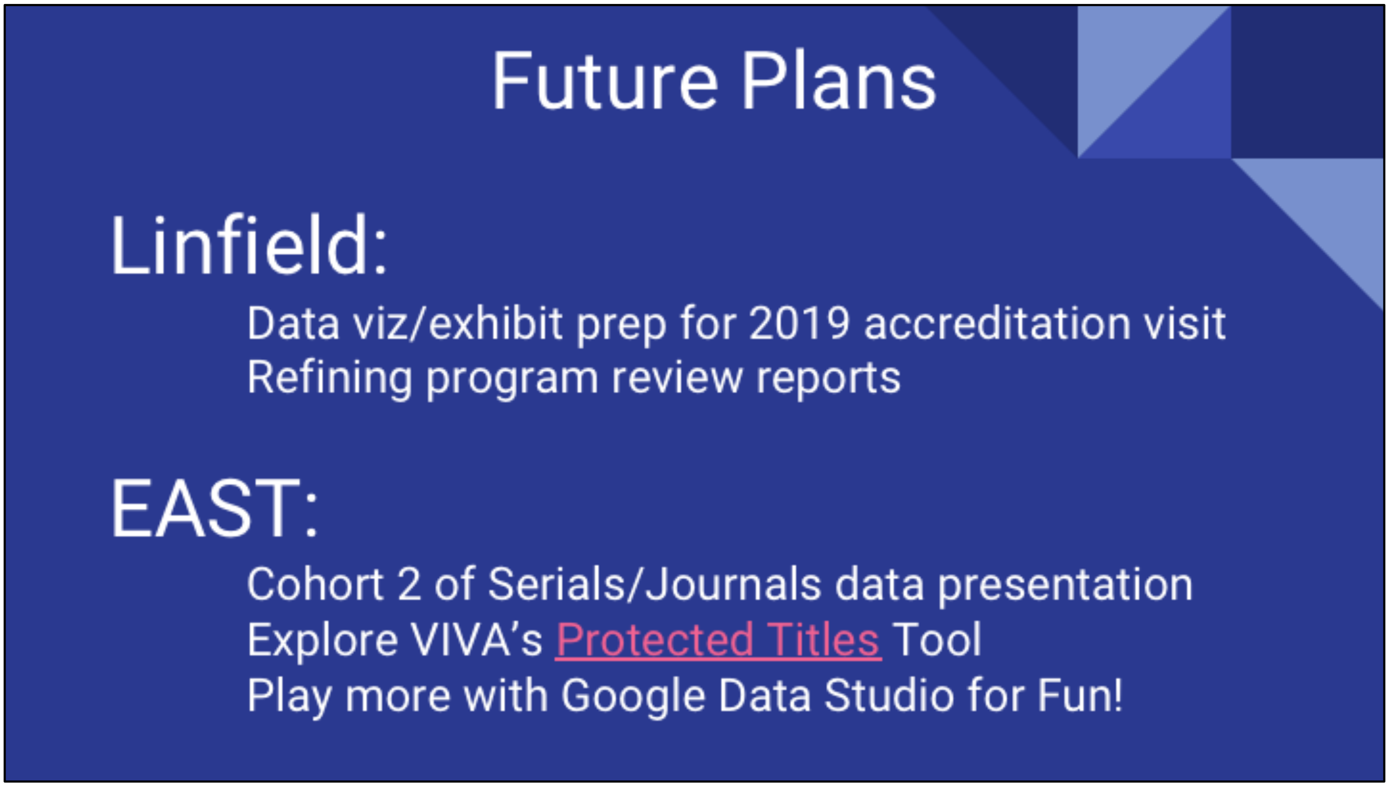
Collection Size Type: (All)

Consortia: (All)

Column to Show
- ● Retention Type
- ○ Collection Size Type
- ○ Consortia
- ○ ILS id
- ○ Cohorts

Tableau can use many data sources, including Google Sheets. We use this combination like a database to keep track of membership and email contacts, e.g., how many serials retention partners have Alma as their ILS, or how many belong to CRL. Again, this is basically like a pivot table in a spreadsheet. Data is drawn from a Google Sheet so updates happen automagically and we can use this interactive dashboard to quickly answer questions.

While I've loved using Tableau, there are lots of tools you can use to make these sorts of dashboards, including Google Sheets :) Tools are evolving all the time. While putting together this presentation, Google Sheets added a "Paste Special/Paste Transposed" option to convert columns to rows, saving me from having to pull data into OpenRefine to do that! Remember that spreadsheets are databases, and you can use SQL queries in them to pull out data you want and then make a graph based on the results. Endless fun!



# Future Plans

## Linfield:

Data viz/exhibit prep for 2019 accreditation visit
Refining program review reports

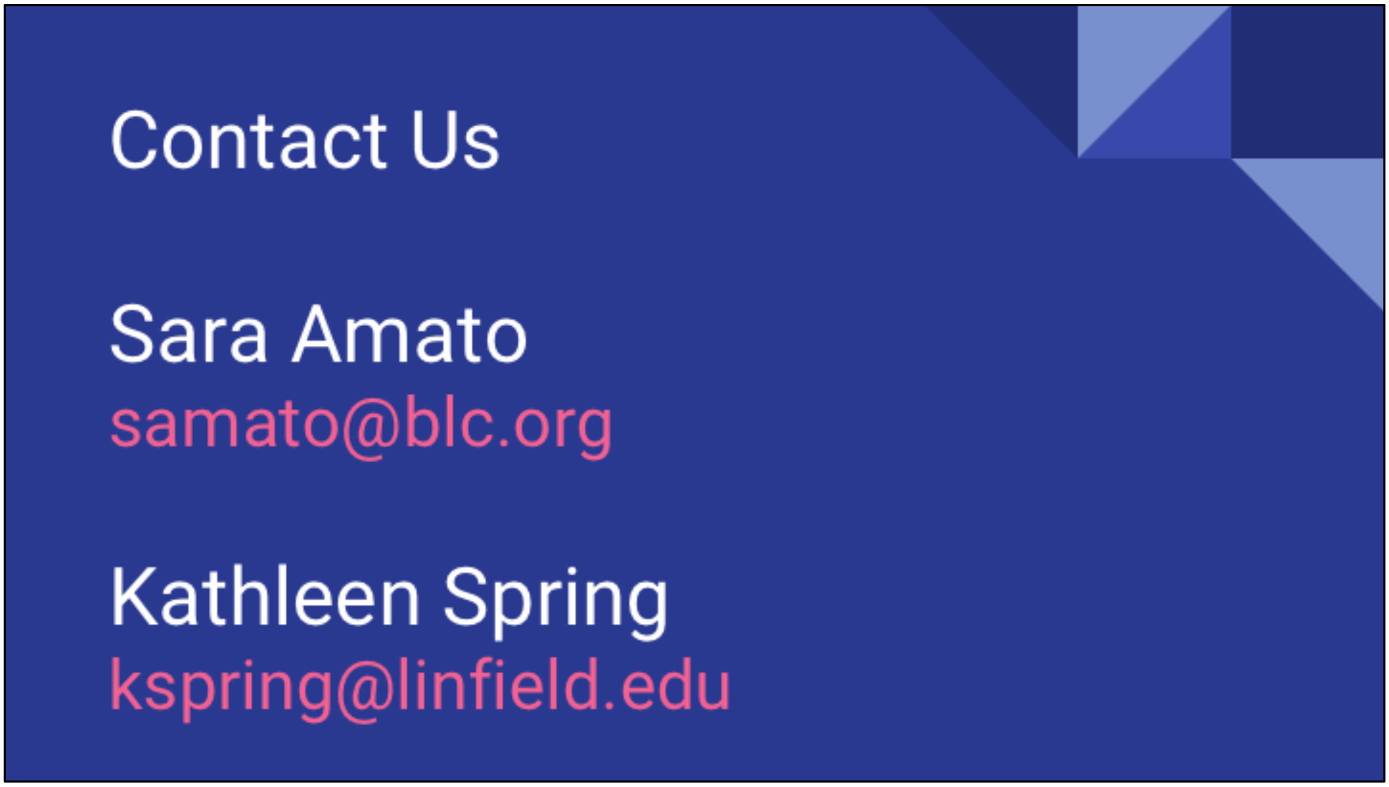
## EAST:

Cohort 2 of Serials/Journals data presentation
Explore VIVA's Protected Titles Tool
Play more with Google Data Studio for Fun!

Linfield
We're trying to take the solutions we've identified and incorporate them into ongoing data viz projects. We'll have our institutional accreditation site visit in spring 2019, and we're utilizing some of those student workers to help make our data "dance" through more visually digestible displays. We also will be focusing on reaching out to a couple of department chairs with upcoming program evaluations and meeting with them to show them our library resources mock-up to see whether it meets their specific needs or whether there are particular adjustments that might be needed.

EAST
1) 2nd cohort of serials partners coming online, will continue to use Tableau in the near-term future.
2) We currently have a very basic title/OCLC number database, but it's outgrowing its viability with the addition of 2.7 million more titles. Need to explore other options - VIVA's protected titles tool seems very cool and includes a query builder and visualizations.
3) Google Data studio keeps sending me ads :) Seems like a full-featured data viz toolset with easy connectors to cloud databases, which we are using, so it would probably be good to try.

# Questions/Discussion



## Contact Us

Sara Amato
samato@blc.org

Kathleen Spring
kspring@linfield.edu