



Visualizing and Sharing Multi-Platform Library Analytics with Tableau

Joan Kolarik
IGeLU Developer's Day
Trondheim, Norway
8 Sep 2016

Agenda

Who?

Tableau: Why Do We Care?

Tour Tableau

Collaboration

More Information

Who?

▶ Weizmann Institute of Science

- A leading basic-research graduate institution with faculties covering Biology, Biological Chemistry, Chemistry, Physics, Mathematics, and Computer Science
- [Fast Facts about Weizmann](#)

▶ MELI

חשתמשי אקס ליבריס - ישראל
Ex Libris Users Group - Israel

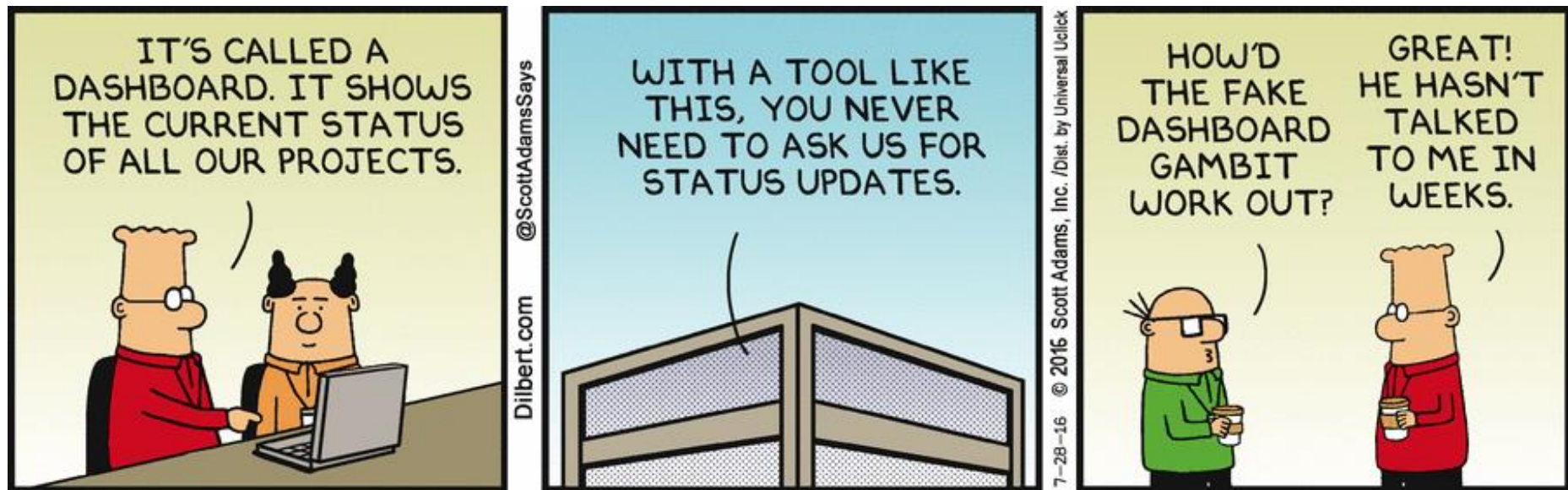


Tableau: Why Do We Care?

»» Agenda

What is Tableau?

- ▶ Tableau helps people see and understand data
- ▶ Ground-breaking data visualization software
- ▶ Tableau allows for instantaneous insight by transforming data into visually appealing, interactive visualizations
- ▶ [Video](#)



Why Use Tableau?

- ▶ Query - blend - explore - understand - discover - analyze
- ▶ Handle
 - small and large data sets
 - Data blended from multiple sources
- ▶ Create your own “viz” (graphs, dashboards, stories)
- ▶ Empower local users
- ▶ Make data more accessible
- ▶ Improve communications

How Libraries Use Tableau

- ▶ Collection Analysis – Use, Coverage, Growth, Quality
- ▶ Expenditure Analysis
- ▶ E-Resource Cost and Use (Ezproxy, USTAT, SFX, Alma)
- ▶ Library Assessment, Communication, Support of Scholarship
- ▶ ARL Statistics, School Comparisons
- ▶ Library Service Use and Contribution
 - Circulation Desk Activity – Staffing decision support
 - Reference Activity, LibGuides and LibAnswers
 - Gate Count
 - Library Sites, including Discovery
 - ILL/Circ Use Correlation to GPA (Laura Murray, non-public)

[See detailed slides at end](#)

Tableau Products



▶ *Public*

- Basic data sources
- Storage free, public, and hosted

▶ *Personal*

- Basic data sources
- Local private file storage
- Free for students and instructors at accredited academic institutions

▶ *Professional*

- Full data sources
- *Server* or *Online* (private local or hosted data storage)
- Enterprise-wide deployment

Tableau Tools

▶ Data Sources

- basic = Excel, Text (e.g., csv), Access, Statistical Files (SAS, SPSS, R), Odata, [Web Data Connectors](#)
- full = basic + over 35 others, including Google Analytics, SQLServer, MySQL, Oracle databases, Salesforce.com, ODBC 3.0 compliant sources

▶ Workbook files (.twbx)

- Can be saved from all Desktops except Public
- Can be downloaded from Tableau Public Server

▶ Reader

- View and interact with visualizations via .twbx files
- Cannot create or change visualizations

Tableau Free / Public Limitations

- ▶ Data sets up to 10 million rows
- ▶ Storage limit = 10 GB

- ▶ NOT a Limitation:
 - Can prevent file downloads
 - Authors can keep even free workbooks proprietary

Tableau SDK

- ▶ Supports Windows, Linux, and Mac
- ▶ Uses C, C++, Java, and Python
- ▶ Extract API
 - populate Tableau Data Extract (.tde) files
 - connect to unsupported data sources and write to .tde
 - previously called TDE API
- ▶ Server API
 - publish extracts as data sources to a server
- ▶ JavaScript API
 - integrate vizes from any Tableau web server in web pages
 - export vizes to image or PDF file

Tour Tableau

»» Agenda

Looking Around Tableau

- ▶ Connect to Data Sources
- ▶ Customize and Save
- ▶ Export and Share
- ▶ Univariate Charts (1 measure)
- ▶ Bivariate Charts (2 measures)
- ▶ Create Multivariate Charts
- ▶ Calculate User-defined Fields
- ▶ Explore Advanced Features

Tableau Lessons Learned

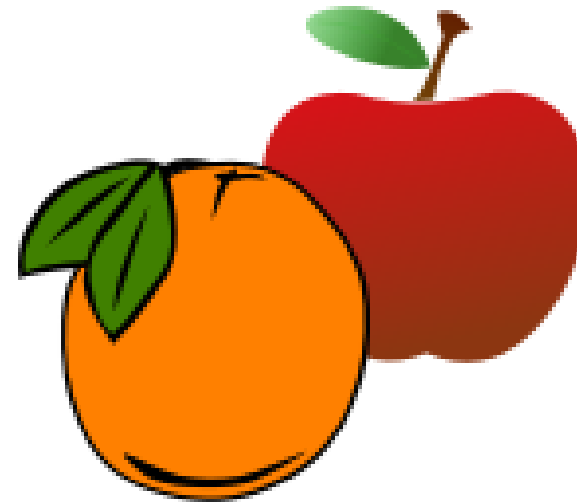
- ▶ Best structure: each row is 1 data point
- ▶ Unions can combine data for separate years located in different data sources
- ▶ **IMPORTANT:** If you're using only the free, public version and saving on the web, be sure to download a backup frequently
 - It's really easy to delete a visualization!

Starting Out With Tableau

- ▶ Data cleanup and structure, see also
 - <https://public.tableau.com/s/blog/2016/05/clean-your-excel-data-data-interpretation>
 - <https://public.tableau.com/s/blog/2016/05/data-prepping-and-data-cleaning-tableau-explained>
- ▶ An early effort:
https://public.tableau.com/profile/cunyols#!/vizhome/CUNYOLSDashboardv2/OLSDashboard2_0
 - Running Totals – whoops
 - The opposite is Difference

Aleph & Primo Source Data

- ▶ Aleph / Primo vs Google Analytics (GA)
- ▶ GA data activated from April 2015
- ▶ Required other source pre-April
- ▶ Aleph = Aleph OPAC Activity, aka Z69
- ▶ Primo = Primo Searches (pre-OBI)
- ▶ Yes, they are different sessions,
 - but you take what you can get



Aleph Source Data from Z69

- ▶ Aleph OPAC Activity, aka Z69
- ▶ Loooong history! At CUNY available from 2006
- ▶ Includes session information (as defined by Aleph)
- ▶ The following SQL extracts 1 month of data
 - Tab-delimited
 - Don't include Z69_CLIENT_ADDRESS for privacy

```
select Z69_TIME_STAMP || chr(9) || Z69_SESSION_ID || chr(9) ||  
Z69_EVENT_TYPE || chr(9) || Z69_BOR_STATUS || chr(9) || Z69_SEARCH_BASE  
|| chr(9) || Z69_SEARCH_QUERY || chr(9) || Z69_NO_HITS || chr(9) ||  
Z69_SCAN_BASE || chr(9) || Z69_SCAN_CODE || chr(9) || Z69_SCAN_TEXT  
from z69  
where Z69_TIME_STAMP > 2015120000000000000000  
and Z69_TIME_STAMP < 2016010000000000000000;
```

Source Data from Primo

- ▶ Primo Searches (pre-OBI)
- ▶ Primo Reports > [NEW] Annual search statistics
- ▶ 1 by 1 for 20 views
 - Not a viable ongoing model
 - No way to automate
- ▶ OBI
 - Can create similar Primo-defined sessions report
 - Similar to GA, not exact
 - Report scheduling
 - Output all Views in 1 report (central admin)
 - Full retrospective data from July 2014

Primo + Aleph from GA

- ▶ For most efficiency, need data reported by school
- ▶ Use Content Grouping
 - Extracts school via regex
 - Only applied from date of group creation
 - Older data must be filtered manually
- ▶ Choose site > Admin > View > Content Grouping > Content Grouping Settings > Group Using Extraction > Page > regex
 - Aleph requires 2 regex to harvest library (xxx01) + admin (+ xxx50) URLs

GA Content Group Settings

Name

Configure the grouping

Content Grouping lets you create logical collections of site or app content, and use those as primary dimensions in your reports. Use one or more of the methods below to group your content. [Learn More](#)

GROUP BY TRACKING CODE

+ Enable Tracking Code

GROUP USING EXTRACTION

+ Add extraction

Primo: vid=(..)

1. local_base=([^\s]*) local_base=([^\s]*)

2. sub_library=([^\s].) sub_library=([^\s].)

GROUP USING RULE DEFINITIONS

+ Create a rule set

Aleph:

SFX Data Source

- ▶ Based on SFX Query 2 : Number of requests (and clickthroughs) per source
 - Total by Month
 - Show All
 - Date Range = 1 month
 - Filters = HIGHLIGHT ALL Institutes
 - Output plain-text, email
- ~~▶ Configure 20 scheduled queries to collect all schools~~
 - Wrote script to harvest all queries then combine into 1 file automatically
- ▶ Remember SFX Offline Stats
 - may need to reload if historical data required

SFX Query to Excel

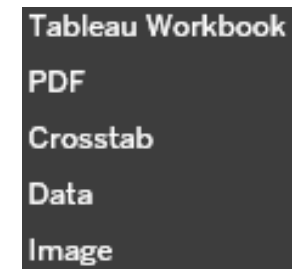
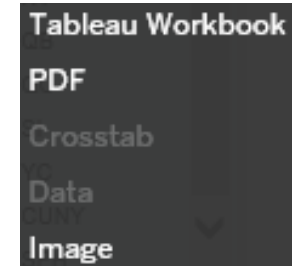
- ▶ Use Excel analytics-sfx-rq-ct.xlsx
- ▶ Receive monthly query email
- ▶ Open Excel and text file
- ▶ Copy from text file
- ▶ Paste to Excel sheet
- ▶ Automatically summarizes data by vendor using text pattern
 - =SUMIF('2016-02'!\$A:\$A,"*bx-*",'2016-02'!\$B:\$B)
- ▶ Copy to consolidate sheet

Collaboration

»» Agenda

Sharing Tableau Vizes

- ▶ Grab a CUNY dataset + template @
 - <https://public.tableau.com/profile/cunyols/>
- ▶ **Template**
 - Bottom right > Download > Tableau Workbook (or top right)
 - twbx file can be opened in any Tableau Desktop
- ▶ **Data** (tab above Dashboard)
 - Download each sheet's data separately
 - Bottom right of sheet (!) > Download > Data
 - Tableau format and content, not original
 - In new window, at bottom, click "Download all rows as a text file"
 - Save and open in Excel



Standing on the Shoulders of Giants

- ▶ SUNY Buffalo

<https://public.tableau.com/profile/ub.libraries#!/vizhome/Circulation-Activity-by-Library/ActivitybyLibraryDashboard>

- ▶ CUNY <https://public.tableau.com/profile/cunyols#!/vizhome/CUNY-Circulation-Activity/CirculationDashboard>

Groups

- ▶ **IGeLU/ELUNA Analytics Special Interest Working Group (SIWG)**
 - Coordinator & Chair = [John Krug](#), University of Lancaster

- ▶ **Tableau/Ex Libris Users Group**
 - Shared platforms
 - Shared data structure
 - Let's share templates and logic!

Some Analytics SIWG Objectives

- ▶ Maintain key goals document
- ▶ Maintain open use cases and scenarios document
- ▶ Promote uses of analytics in Ex Libris products
- ▶ Promote use of analytics in ‘evidence-based librarianship’
- ▶ Ensure availability of relevant product data
 - in analytics products and via APIs and publishing / batch export, supporting uses like data warehouses, big data, visualization, and programming
- ▶ Contribute to educating the user community
- ▶ Maintain links with LOD & Interoperability SIWGs
- ▶ Coordinate analytics needs for the NERS process
- ▶ Establish mailing list for announcements and discussion

OBI / Tableau Integration ?

- ▶ [Idea Exchange](#): Tableau can connect to Oracle BI (Alma/Primo Analytics), but for security reasons Ex Libris won't open/allow such access. Currently I manually import Alma analytics xls files into Tableau to work on them (was waiting for ExL to open the connection), but with the above update I'll now start work on connecting Tableau to the Alma APIs.
- ▶ Please contact Justin if you are also interested in this development and tell him
 - how you currently use Tableau and Alma Analytics
 - if you have plans to use the APIs with Tableau
- ▶ [Justin Kelly](#), Library Web Developer, Swinburne University of Technology (Analytics SIWG member)

Tableau Web Data Connectors

- ▶ <https://www.tableau.com/web-data-connector>
- ▶ Available in both free and paid desktops
- ▶ Ex Libris has a problem providing institution-specific access

Tableau/ExLibris Users Group

- ▶ Conceived at ENUG, 30 Oct 2015
- ▶ 1st meeting, 16 Dec 2015
 - setting up Tableau Desktop (free); discussed MySQL use
- 2nd meeting, 24 Mar 2016
 - Kristina's NYU data; Jennifer's SUNY data; Kevin's CUNY data; analytics remixing
- ▶ 3rd meeting @ ELUNA, 6 May 2016
 - Introduction to the group; Mark's work on automatic harvesting of USTAT files
- 4th meeting, 16 Jun 2016
 - Methods of copying templates without data; NYU's work analyzing metadata
- 5th meeting, SIWG Show&Tell, Oct 2016

Tableau at Academic Libraries (1)

- ▶ Brandeis U
 - Collections; Use Cost
- ▶ California Institute of Technology *
- ▶ [Cornell U](#)
 - Assessment; Scholarship Support; Contribution; Collections; Services
- ▶ [CSU Northridge](#) *
 - Collections
- ▶ [CUNY](#)
 - Discovery; Circulation
- ▶ Emory University *
- ▶ George Mason U *
- ▶ [Lancaster U](#) *
 - Loans; EZproxy
- ▶ Lawrence Livermore National Lab *
- ▶ MIT *
- ▶ Mount Royal U *
- ▶ [MSU Michigan State U](#)
 - Reference; Discovery; Circulation
- ▶ The New School *
- ▶ [NorthEastern U](#)
 - Gate Count; Reference; ILL; Website

Tableau at Academic Libraries (2)

- ▶ NYU *
 - Circulation Desk; Cataloging & Metadata Analysis
- ▶ OSU Ohio State
 - LibAnswers; Gate Count; ARL Comparisons; Collections
- ▶ Purdue U *
- ▶ Seneca College *
- ▶ Swinburne U of Technology
- ▶ SUNY New Paltz *
- ▶ U British Columbia
 - Library Assessment
- ▶ U Buffalo *
 - correlate ILL/Circ to GPA; Circulation; ARL Statistics; Gate Count; Discovery
- ▶ U Connecticut
 - ARL Member Expenditures
- ▶ U Houston *
- ▶ U MA Amherst
 - Collection Assessment

Tableau at Academic Libraries (3)

- ▶ [U MA Dartmouth](#) *
 - Expenditures
- ▶ [U North Texas](#)
 - Collection Assessment
- ▶ [Warwick U](#)
 - Expenditure Analysis
- ▶ William Patterson U *
- ▶ [Yale U](#) *
 - Collection; LibGuides
- ▶ [Choice.org](#)
 - Midsize MidWest Public University
Comparison of Enrollment to Library
Expenditures

More Information

»» Agenda

References (1)

- ▶ Map (<https://en.wikipedia.org/wiki/Israel>)
- ▶ Help section at right in desktop open screens →
- ▶ Blog by a Program Manager at Tableau Software (<https://boraberan.wordpress.com/>)
- ▶ Dilbert. 28 Jul 2016 (<http://dilbert.com/strip/2016-07-28>).
- ▶ How to Save and Modify Clipboard Data Source (<http://kb.tableausoftware.com/articles/knowledgebase/editing-pasted-data>)
- ▶ Preparing Excel Files for Analysis (<http://kb.tableau.com/articles/knowledgebase/preparing-excel-files-analysis>)

Discover

How-to Videos

- ▶ Overview
- ▶ Intro to the Interface
- ▶ Chart Types
- ▶ Dashboard

More how-to videos...

VIZ OF THE DAY

Clinton vs. Trump: The Battle for the Presidency →

Resources

- Blog - Viz All Things Politics for Politics Month!
- Sample Data Sets
- Live Training
- Current Status

References (2)

- ▶ How to Save and Modify Clipboard Data Source
(<http://kb.tableausoftware.com/articles/knowledgebase/editing-pasted-data>)
- ▶ Odata connector for MySQL
(<https://blogs.msdn.microsoft.com/interoperability/2012/02/09/open-source-odata-tools-for-mysql-and-php-developers/>)
- ▶ Murphy, Sarah Anne (2015). "How data visualization supports academic library assessment: Three examples from The Ohio State University Libraries using Tableau." *College & Research Libraries News*, Oct 2015 p 482-486.
- ▶ Nandeshwar, Ashutosh (2013). [Tableau Data Visualization Cookbook](#). Packt Publishing.

References (3)

- ▶ Buhler, Jeremy; Lewellen, Rachel; and Murphy, Sarah (2014). "Tableau Unleashed: Visualizing Library Data". Library Assessment Conference: Building Effective, Sustainable, Practical Assessment, 4 Aug 2014 (<http://libraryassessment.org/bm~doc/7murphypanelrev.pdf>)
- ▶ Eldridge, Alan (2013). [Best Practices for Designing Efficient Tableau Workbooks](#). Tableau white paper.
- ▶ Morton, Kristi, Bunker, Ross, Mackinlay, Jock, Morton, Robert, and Stolte, Chris (2012). "Dynamic Workload Driven Data Integration in Tableau". SIGMOD '12 Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, p 807-816.

Thank You

Joan Kolarik

Systems Librarian

Weizmann Institute of Science

joan.kolarik@weizmann.ac.il

#igelu2016 #devday16 #analytics #tableau @weizmannscience

