Letting the collections tell their story

Using Tableau for collection evaluation

LITA Pre-Conference Full-Day Workshop, 2016

Karen R. Harker, MLS, MPH
Janette Klein, MLIS, Doctoral student
Priya Parwani, MLIS candidate
Logistics

- Seating arrangements
- Wi-Fi
  - Username: lita
  - Password: omni2016
- Restrooms are located straight out the doors, down the hall and to the right. Go past the elevator.
- Breaks!
  - Thursday: 3:15-3:35 - snacks
  - Friday: 9:15-9:35 - breakfast
Program Overview

Thursday
- Visualization
- Introduction to Tableau
- Data management
- Developing dashboards
- Using dashboards

Friday
- Lecture
- Q&A
- Small group discussions
- Exercises
- Etherpad for –
  - Sharing notes or ideas
  - Asking questions
  - Displaying results
  - [http://goo.gl/AwvaHz](http://goo.gl/AwvaHz)
Who are you and why are you here?

Your Story in Tableau:  
http://goo.gl/JY9vaO
Getting to know your group

- Share your stories with your group:
  - Your institution & role
  - Your experience with data
  - What you hope to do with data visualization
  - Ways your ideas overlap and ways to collaborate
Why visualizations? To make it easier to see...

Patterns  Relationships  Trends
Why visualizations?
To make decisions on where to focus.
Why visualizations?

- To demonstrate value of the library to stakeholders
  - Efficient use of resources
  - Effectiveness
  - Value to the institution
Great Visualizers
William Playfair
Jacques Charles Minard
Napoleon’s disastrous Russian campaign (1869)
The greatest value of a picture is when it forces us to notice what we never expected to see.

— John Tukey —
Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space.

— Edward Tufte —
Why is visualization so hot now?
Capabilities

115 Years of Moore’s Law

open

DATA

useful

usable

desirable
Visualizations in Libraries

- The Tableau Trio: Murphy-Buhler-Lewellen
- ARL Webinar Series
  - Early 2015

Rachel Lewellen, UMass Amherst
Jeremy Buhler, U British Columbia
Anne Murphy, Ohio State U
More data & visualizations in libraries

- Dando, Priscilla. 2010. Say it with data: A concise guide to making your case and getting results. Chicago, IL: ALA.
- Data Science and Visualization Institute for Librarians: NCSU
  https://www.lib.ncsu.edu/datavizinstitute - Spring 2017 Sessions!
- Library Data Visualization blog:
  http://librarydatavisual.blogspot.com/
Examples of library visualizations

- Collections
  - Dentographs
  - Bubble chart
  - Map
- Reference desk/questions
  - Ohio State
- Circulation
- Gate counts
  - Ohio State
- Duke University
  - Ithaka Survey
  - Student Survey
- National statistics
  - Public Libraries
  - ARL Expenditures
Categorizing Visualizations by Dimensions

- 1D: Tables
- 2D-Planar: Maps
- 2D-Temporal: Line charts
- nD: Multidimensional
- Tree or Hierarchical
- Network
Kinds of Visualizations

Exercise 1: Draw a chart

1. Use data in table
2. Think of one point to make about the data
3. Draw a chart to best illustrate that one point
Systems & Tools Available

- Many, many tools
- Kinds of systems
  - Chart tools in software (e.g. Excel, R)
  - Javascript chart libraries (e.g. Google Charts, d3.js)
  - Visualization software (e.g. Tableau, Plotly, Qlik)
- Who uses what (from survey)
  - Why?
  - What’s good & not-so-good?
Graph tools embedded in software

**Good**
- Direct connection to data
- Uses same infrastructure

**Caveats**
- Embedded with software
- Sharing & embedding may be limited
- Wide variety of quality and types of charts
- Limited to data in the software
Examples of tools embedded in software
Script chart libraries

Good
- Plethora of libraries
- Most are free or very cheap (<$100)
- Can be embedded
- Many are OpenSource

Caveats
- Require programming
- Variation in quality and chart selection
Examples of Script Chart Libraries

D3

Google Developers CHARTS

plotly

Chart.js
Visualization software

Good
- Point-and-click interface
- Shallower learning curve
- Able to link data from different sources

Caveats
- Limited types of charts
- Variety of interactivity
- May not be easily embedded
Examples of Visualization Software

- Power BI
- Qlik
- SAS Visual Analytics
- Tableau
Good
- 25GB space on free version
- Web interface
- Easy to build charts
- Easy sharing (see caveats)

Caveats
- No central library of visualizations
- Steeper learning curve
- More IT intensive
Good
- Free
- Library of visualizations
- No programming req’d
- Shallow learning curve

Caveats
- Public means Public
- Pricing model
- Not quite Excel
- Advanced programming helps
Questions? Comments?
Diving in: Getting started with Tableau

Priya Parwani
How excited are we?

I’m so excited.

SO EXCITED

I CAN'T WAIT!!

I can’t wait

I’m so excited
Introduction to Tableau

- Opening Tableau app
- Creating a Tableau Public account (if not already)
- Browsing, searching Tableau Public for cool examples & for open data sets
- Training opportunities
  - Tableau Public
  - Lynda.com
- Communities
Tableau Introduction

- **Publish Data Sources and Workbooks** Share your work with others using Tableau Online, Tableau Server, the Tableau mobile app, or on Tableau Public.

- **The Tableau Environment** Learn what elements of the Tableau UI are called.

- **Tableau Concepts** Dig into the ideas behind the terms and behaviors you see in Tableau.

[http://goo.gl/jzLND0](http://goo.gl/jzLND0)
Exercise 2: Re-create hand-drawn chart

Open Tableau
Recreate the hand-drawn chart

Connect to a data source
Place dimensions and measures on the canvas
Thinking Visually
Janette Klein
12 Chart Types in Brief – A Snapshot

Dissemination of Results

Audience
- Library administration
- Collection Development Manager
- Subject specialist librarian
- Library staff
- College Administration
- Requesting Entity
- Faculty governance committee
- General public
- Other constituent

Enrollment by Program Level

College
- Arts and Sciences
- Level
  - Doctoral
  - Masters
  - Post-Bac
  - Undergraduate

Year of Semester
- 2011
- 2012
- 2013
- 2014
- 2015
- 2016

LEED Buildings by state per million people

Where are Data Tools Used or Interested in Using

Annual Enrollment Category

Enrollment by Category & Level

Mon 18 Jan 2010
- Category
  - Arts & Nu Manifes
  - Business

Mon 25 Jan 2010
- Category
  - Arts & Nu Manifes
  - Business

Mon 01 Feb 2010
- Category
  - Arts & Nu Manifes
  - Business

Mon 08 Feb 2010
- Category
  - Arts & Nu Manifes
  - Business

Implementation Phase 1
- Preparation work
- Build prototype

Assign resources
- John
- Mike, John

Gather documents
- Mike, John, Kent

Report to management
- Mike, John, Kent
- Kent

LibPAS
- Access
- Visualization

SPSS
- Used
- 0%
- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%

LibPAS
- Access
- Visualization

Excel
- Used
- 0%
- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%
Break
What Chart When?

Comparison

What would you like to show?

Relationship

Distribution

Composition

Changing Over Time

Static

Few Periods

Many Periods

Over Time

Variable Width Column Chart

Table or Table with Embedded Charts

Bar Chart

Column Chart

Circular Area Chart

Line Chart

Column Chart

Line Chart

Two Variables per Item

Many Categories

Many Items

Few Items

Cyclical Data

Non-Cyclical Data

Single or Few Categories

Many Categories

One Variable per Item

Few Categories

Many Periods

Few Periods

Among Items

Scatter Chart

Bubble Chart

Relative and Absolute Differences Matter

Relative and Absolute Differences Matter

Only Relative Differences Matter

Only Relative Differences Matter

Simple Share of Total

Accumulation or Subtraction to Total

Components of Composers

Stacked 100% Column Chart

Stacked Column Chart

Stacked 100% Area Chart

Stacked Area Chart

Pie Chart

Waterfall Chart

Stacked 100% Column Chart with Subcomponents

Single Data Points

Many Data Points

Two Variables

Three Variables

3D Area Chart
Data source management

Priya
Tableau: Data Management Issues

- Source of data - direct or indirect connections?
- Frequency of updating
  - Static
  - Monthly
  - Weekly
  - Daily
  - Hourly
  - Real Time
- Common fields - standardizing attributes
Exercise 3: Tableau Worksheet Creation

- **Exercise Goal:** Create a worksheet using an area chart showing circulation starting from FY 2011 by collection

- Excel file needed for this exercise: *Checkouts by Collection Month and FY*
Exercise 4: Tableau - Provide an Answer

- Exercise Goal: Create a worksheet showing the trend of circulation starting from FY 2011 by category with color for category and detail for collection as percentage of total checkouts.

- Exercise Question: What is the % of circulation for FY 2015, Category Business, Collection Merchandising?

- Excel files needed for this exercise: Checkouts by Collection Month and FY
Day 1 Wrap-up

- Concluding Thoughts
- Questions and Answers
Looking Ahead to Day 2

- Review of Data Wrangling Part I
- Data Wrangling Part II
- Developing Dashboards
- Using Dashboards for Decisions
- Wrap-up
Exercise 5 Part 1: Tableau – Using the Tableau Pivot Function

Part 1:

Exercise Goal: Create a worksheet showing holdings data by collection using Tableau pivot functions

Exercise Question: What can interpretations could be derived from this data set? What impact could this information have upon your audience?

Excel files needed for this exercise: Holdings by Collection and Publication Year for Dashboard

Small Group Discussion Point: What type of charts perform well with this data set? What ones did not work well? Why?
DAY 2
Day 1 Review

- Visualizations
- Dashboard design
- Collections data for decision making
- Introduction to Tableau
- Data Management Issues
- Uploading data to Tableau
Data Wrangling with Tableau

Janette
Collections Data for Decision Making – Needs

- Enrollment Data
- Institution Fact Book
- Institutional Fact Sheets

<table>
<thead>
<tr>
<th>Enrollment</th>
<th>US Citizen</th>
<th>Permanent Resident</th>
<th>N/R Alien</th>
<th>Residence Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texas</td>
<td>39,670</td>
<td>897</td>
<td>56</td>
<td>33,822 (90.2%)</td>
</tr>
<tr>
<td>Out-of-State</td>
<td>1,432</td>
<td>39</td>
<td>0</td>
<td>1,471 (3.9%)</td>
</tr>
<tr>
<td>N/R Alien</td>
<td>0</td>
<td>0</td>
<td>2,182</td>
<td>2,182 (5.9%)</td>
</tr>
<tr>
<td>University Total</td>
<td>34,002</td>
<td>936</td>
<td>2,237</td>
<td>37,175 (100%)</td>
</tr>
</tbody>
</table>

(Per IPEDS Definition)

<table>
<thead>
<tr>
<th>Student Classification:</th>
<th>Breakdowns of Interest:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduate Total: 6,672</td>
<td>First Time in College* 4,661</td>
</tr>
<tr>
<td>Undergraduate Total: 30,003</td>
<td>New Undergraduate Transfers 4,037</td>
</tr>
<tr>
<td>Total Enrollment: 37,175</td>
<td>Post-Bac – 915 Master – 3,910 Doctoral - 1,803</td>
</tr>
<tr>
<td>Special Professional- 44</td>
<td></td>
</tr>
</tbody>
</table>

(*Includes TAMS and Dual Admits)

<table>
<thead>
<tr>
<th>Full/Part Time and FTE:</th>
<th>Full-time</th>
<th>Part-time</th>
<th>Total</th>
<th>FTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergraduate</td>
<td>24,957</td>
<td>5,546</td>
<td>30,503</td>
<td>26,010</td>
</tr>
<tr>
<td>Graduate</td>
<td>3,101</td>
<td>3,571</td>
<td>6,672</td>
<td>3,872</td>
</tr>
<tr>
<td>Total</td>
<td>28,058</td>
<td>9,117</td>
<td>37,175</td>
<td>29,882</td>
</tr>
</tbody>
</table>

Average age:
- Undergraduate – 22.1
- Graduate – 31.3
- All students – 23.7
Collections Data for Decision Making – Needs

**Exhibit A-1**

Enrollment by Classification and Level

Fall Semesters 2006-2015

<table>
<thead>
<tr>
<th>Year</th>
<th>First-Time Freshmen</th>
<th>Total Freshmen</th>
<th>Sophomore</th>
<th>Junior</th>
<th>Senior</th>
<th>Undergraduate Total</th>
<th>Post-Bacc²</th>
<th>Master's</th>
<th>Doctoral³</th>
<th>Special Professional</th>
<th>Graduate Total</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>06</td>
<td>3,522</td>
<td>5,643</td>
<td>5,529</td>
<td>6,623</td>
<td>8,603</td>
<td>28,698</td>
<td>1,256</td>
<td>4,080</td>
<td>1,500</td>
<td>0</td>
<td>6,845</td>
<td>33,443</td>
</tr>
<tr>
<td>07</td>
<td>3,388</td>
<td>5,437</td>
<td>5,856</td>
<td>6,689</td>
<td>9,250</td>
<td>27,242</td>
<td>1,191</td>
<td>4,207</td>
<td>1,513</td>
<td>0</td>
<td>6,911</td>
<td>34,153</td>
</tr>
<tr>
<td>08</td>
<td>3,300</td>
<td>5,113</td>
<td>5,861</td>
<td>7,068</td>
<td>9,737</td>
<td>27,779</td>
<td>1,135</td>
<td>4,261</td>
<td>1,498</td>
<td>25</td>
<td>6,919</td>
<td>34,698</td>
</tr>
<tr>
<td>09</td>
<td>3,175</td>
<td>5,024</td>
<td>5,846</td>
<td>7,497</td>
<td>10,089</td>
<td>28,474</td>
<td>1,349</td>
<td>4,683</td>
<td>1,577</td>
<td>40</td>
<td>7,649</td>
<td>38,123</td>
</tr>
<tr>
<td>10</td>
<td>3,367</td>
<td>5,199</td>
<td>5,973</td>
<td>7,183</td>
<td>9,928</td>
<td>28,283</td>
<td>1,432</td>
<td>4,660</td>
<td>1,656</td>
<td>36</td>
<td>7,784</td>
<td>36,067</td>
</tr>
<tr>
<td>11</td>
<td>3,607</td>
<td>5,237</td>
<td>5,878</td>
<td>7,325</td>
<td>9,642</td>
<td>28,282</td>
<td>1,258</td>
<td>4,374</td>
<td>1,743</td>
<td>37</td>
<td>7,412</td>
<td>36,694</td>
</tr>
<tr>
<td>12</td>
<td>3,937</td>
<td>5,558</td>
<td>5,908</td>
<td>7,278</td>
<td>10,095</td>
<td>28,911</td>
<td>1,157</td>
<td>3,887</td>
<td>1,785</td>
<td>38</td>
<td>6,867</td>
<td>35,778</td>
</tr>
<tr>
<td>13</td>
<td>3,889</td>
<td>5,632</td>
<td>6,233</td>
<td>7,525</td>
<td>10,091</td>
<td>29,481</td>
<td>1,142</td>
<td>3,678</td>
<td>1,829</td>
<td>38</td>
<td>6,867</td>
<td>36,168</td>
</tr>
<tr>
<td>14</td>
<td>3,810</td>
<td>5,487</td>
<td>6,397</td>
<td>7,713</td>
<td>10,126</td>
<td>29,723</td>
<td>967</td>
<td>3,646</td>
<td>1,785</td>
<td>43</td>
<td>6,441</td>
<td>36,164</td>
</tr>
<tr>
<td>15</td>
<td>4,015</td>
<td>5,680</td>
<td>6,456</td>
<td>8,143</td>
<td>10,224</td>
<td>30,503</td>
<td>915</td>
<td>3,910</td>
<td>1,603</td>
<td>44</td>
<td>6,672</td>
<td>37,175</td>
</tr>
</tbody>
</table>

¹First-Time Freshmen counts exclude dual admit and TAMS students per new definition as of Fall 2001 from THECB.
²Students are designated post-baccalaureate if they have completed their bachelor's degree and are continuing their education but have not been admitted into a graduate program. Post-baccalaureate category was first used by UNT in Fall 1993.
³Doctoral includes those enrolled in PhD programs.
Collections Data for Decision Making – Needs Based Data

- ILL borrowing
  - Interlibrary loans and distance learning requests
  - Data extraction from ILLiad & WorldShare
- Key data:
  - transaction number, request type, document type, patron status, department, LC Call Number, publication date, request date.
Collections Data for Decision Making – Holdings Data

- **Source:** OCLC - World Share Conspectus Evaluation System
  - By Format
  - By Year of Publication
  - By Collection
Collections Data for Decision Making – Use Data

- Circulation (via III Sierra)
- Monthly and annual summaries and Cross-tabs
  - Patron Type
  - Patron’s Major/Department
  - Item Type
  - Call Number Range
  - Location
Collections Data for Decision Making – Collections Data

- Collection map
  - Collection
  - Subject

Manage the Collection Map
- View Full Conspectus
- Modify Program-Based Collections
- View Collections by Full Conspectus
- Manage Collections by Conspectus Subject
- Manage Collections by Call Number Range
- Assign an LCCN Range to a Collection
- Assign Subjects to a Division to a Collection
- View YBP Approval Plan
What is the UNT Collection Map? – A Brief Background

A Traditional Past

Individual fund
- Funding
- Assessment data

Collection 1

Collection 2

An Assessment-Based Future

Topic 1
Topic 2
Topic 3

Assessment data
UTN Collection Map Development

WorldShare® CES
- 11,251 Subjects
- Call number ranges
- aka Conspectus

UNT Libraries'
- 59 Subject-based Collections

Collection Map
Mapping **Items** to **Subjects** to Collections

- **Collections**
- **Subjects** with call # ranges
- **Items**, each with a call #

**Map Collections**

- **Subjects** assigned to 1 or more **Collections**
- Some **Subjects** assigned to “Out of Scope” Collection

**Map Items**

- List of **Items**, etc. with call number
- Assigned a **Subject** based on call #
- **Subject** assigned to **Collection(s)**

Database
Map Subjects to Collections

Subject:
Sociology:
Social Sciences, General:
Theory, Methodology
H61

Collection:
Anthropology
H61
HQ 12-472

Subject:
Sociology:
Family, Marriage, Sexual Life:
Sexual life, Erotica
HQ 12-472

Collection:
Sociology
H0-9999

Subject:
Sociology:
Family, Marriage, Sexual Life:
Women, Feminism
HQ 1101-2030

Collection:
Women’s Studies
HQ0-9999
Items connected to Collections via Subjects

- Multiple Subjects Linked to one or more Collections
- Multiple Items Linked to any one Subject

Each and All Items Linked to One or More Collections via the Subject
Results

Any one **item** may be assigned to one or more subject-based **collections**

<table>
<thead>
<tr>
<th>Collection</th>
<th>EbookID</th>
<th>LC Call Num</th>
<th>Conspicuous Description</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>529305</td>
<td>PS3057.E6 M38</td>
<td>American Literature - 19th Century</td>
<td>An insect view of its plain</td>
</tr>
<tr>
<td>English</td>
<td>10387177</td>
<td>PR3803 .B65 20</td>
<td>English Literature - Provincial &amp; Local - Ireland</td>
<td>Blighted beginnings</td>
</tr>
<tr>
<td>Psychology</td>
<td>10419977</td>
<td>RC506 .G76 200</td>
<td>Psychotherapy - Chemo-, Shock,Hypnosis,Ana</td>
<td>--but at the same time and on another level--</td>
</tr>
<tr>
<td>Counseling &amp; Higher Education</td>
<td>10419977</td>
<td>RC506 .G76 200</td>
<td>Psychotherapy - Chemo-, Shock,Hypnosis,Ana</td>
<td>--but at the same time and on another level--</td>
</tr>
<tr>
<td>History</td>
<td>1587360</td>
<td>E183.B.ISS E39</td>
<td>History - United States, General Works</td>
<td>Dual containment policy in the Persian Gulf :</td>
</tr>
<tr>
<td>History</td>
<td>966527</td>
<td>DK268.4 .J58 20</td>
<td>Soviet Regime, 1918-1991</td>
<td>Enemies of the people under the Soviets :</td>
</tr>
</tbody>
</table>

- Integrated interdisciplinary Collection Map
- **One item** – **Many collections**
- Mapping of item level data with LC call numbers
Future Data Integration with Tableau

Demand-Driven Acquisitions
- Bibliographic data from Sierra
- Transaction data from:

MINES for Libraries©
- Who uses...
- What
- From where
- When
- Why
Collection Data
Break
Data Wrangling for Tableau

- Data Collection
  - Raw Data
    - Multiple data sources
    - Messy
    - Non-standardized
  - Summary Data
    - Partially cleaned
Data Wrangling for Tableau

- Data Collection
  - “Janitorial work” – labor intensive, detail oriented,….
  - New York Times (2014) notes that data scientists spend “50 percent to 80 percent of their time mired in this more mundane labor of collecting and preparing unruly digital data, before it can be explored for useful nuggets” (para. 3).
  - Has been likened to the “modern Wild West…” (para. 5).

What is involved?

- Data Cleaning
- Data Processing
- Data Validation

Data Wrangling for Tableau

Data Collection – Raw Data

- Data Collection
- Raw Data
Raw Data - Desired Information

Source Information


Title & Author Information

Advanced techniques for counseling and psychotherapy, electronic resource] / Christian Conte.

Publisher & Year of Publication Information

## Data Wrangling for Tableau

### Data Collection – Summary Data

<table>
<thead>
<tr>
<th>Collection</th>
<th>Total Hold</th>
<th>Pre-1800</th>
<th>1800-1899</th>
<th>1900-1999</th>
<th>2000+</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounting &amp; Taxation</td>
<td>4</td>
<td>2785</td>
<td>1529</td>
<td>869</td>
<td>338</td>
<td></td>
</tr>
<tr>
<td>Anthropology</td>
<td>130</td>
<td>42855</td>
<td>14165</td>
<td>15998</td>
<td>12608</td>
<td></td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>198</td>
<td>43722</td>
<td>16151</td>
<td>14862</td>
<td>12626</td>
<td></td>
</tr>
<tr>
<td>Aviation Logistics</td>
<td>27</td>
<td>25408</td>
<td>8106</td>
<td>10018</td>
<td>7077</td>
<td></td>
</tr>
<tr>
<td>Behavior Analysis</td>
<td>42</td>
<td>17715</td>
<td>5092</td>
<td>6096</td>
<td>6491</td>
<td></td>
</tr>
<tr>
<td>Biological Sciences</td>
<td>799</td>
<td>39447</td>
<td>13762</td>
<td>12123</td>
<td>12891</td>
<td></td>
</tr>
<tr>
<td>Biomedical Engineering</td>
<td>7</td>
<td>497</td>
<td>74</td>
<td>141</td>
<td>280</td>
<td></td>
</tr>
<tr>
<td>Business, General</td>
<td>41</td>
<td>27480</td>
<td>11189</td>
<td>9991</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Chemistry</td>
<td>98</td>
<td>16626</td>
<td>6515</td>
<td>5639</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication Studies</td>
<td>2</td>
<td>5289</td>
<td>1183</td>
<td>2059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer Science &amp; Engineering</td>
<td>108</td>
<td>31107</td>
<td>3210</td>
<td>12715</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Consumer Experience</td>
<td>37</td>
<td>32472</td>
<td>9429</td>
<td>13632</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Counseling</td>
<td>36</td>
<td>20414</td>
<td>15015</td>
<td>12255</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Criminal Justice</td>
<td>25</td>
<td>12171</td>
<td>2773</td>
<td>3737</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Dance &amp; Theatre</td>
<td>63</td>
<td>33635</td>
<td>5962</td>
<td>10970</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>DDA Out of Scope</td>
<td>859</td>
<td>181666</td>
<td>78754</td>
<td>56935</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td>29</td>
<td>19943</td>
<td>7937</td>
<td>7433</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design Management</td>
<td>15</td>
<td>10184</td>
<td>6764</td>
<td>6042</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economics</td>
<td>47</td>
<td>29451</td>
<td>12036</td>
<td>10929</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational Psychology</td>
<td>64</td>
<td>42313</td>
<td>12924</td>
<td>14672</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Electrical Engineering</td>
<td>73</td>
<td>18395</td>
<td>3135</td>
<td>7452</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrical Engineering, General</td>
<td>145</td>
<td>18768</td>
<td>6706</td>
<td>6759</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>69</td>
<td>13383</td>
<td>5133</td>
<td>4638</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>122</td>
<td>160023</td>
<td>64479</td>
<td>48928</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Ethnic Studies</td>
<td>68</td>
<td>23191</td>
<td>7479</td>
<td>8125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forensic Sciences</td>
<td>21</td>
<td>6089</td>
<td>1414</td>
<td>1741</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalities</td>
<td>6</td>
<td>6362</td>
<td>3112</td>
<td>1720</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Functions in Excel and Tableau

- Foundational functions for working with text in Excel
  - Format as table
  - Conditional formatting & filtering for first looks
  - Find and Replace
    - Examples include (.), (/), ©, ®, [ ], etc.
- Basic functions
  - TRIM: Removes leading & trailing spaces
  - LEFT, RIGHT, and MID functions extract parts of data
    - Rely on patterns

Reality Check - data is rarely neat and tidy
Excel is not Tableau and Tableau is not Excel

- Re-structuring data from horizontal summaries (e.g. columns are for months or years) to vertical

- Tableau peculiarities:
  - Using dates instead of months and/or years
  - Vertical structure better than horizontal structure (see above Excel tip)
Excel is not Tableau and Tableau is not Excel - Example

- Example of Excel data

<table>
<thead>
<tr>
<th>Collection</th>
<th>Category</th>
<th>Pre-1980</th>
<th>1980-1999</th>
<th>2000+</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounting &amp; Taxation</td>
<td>Business</td>
<td>1529</td>
<td>869</td>
<td>338</td>
<td>49</td>
</tr>
<tr>
<td>Anthropology</td>
<td>Social Sciences</td>
<td>14164</td>
<td>15997</td>
<td>12606</td>
<td>84</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>Arts &amp; Humanities</td>
<td>16022</td>
<td>14969</td>
<td>12775</td>
<td>140</td>
</tr>
<tr>
<td>Behavioral Analysis</td>
<td>Social Services</td>
<td>5092</td>
<td>6096</td>
<td>6491</td>
<td>36</td>
</tr>
<tr>
<td>Biological Sciences</td>
<td>STEM</td>
<td>13106</td>
<td>11919</td>
<td>12252</td>
<td>181</td>
</tr>
<tr>
<td>Business, General</td>
<td>Business</td>
<td>11189</td>
<td>9991</td>
<td>6026</td>
<td>254</td>
</tr>
</tbody>
</table>

- Example of Tableau data

<table>
<thead>
<tr>
<th>Collection</th>
<th>Category</th>
<th>Pivot Field Name</th>
<th>Pivot Field Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engineering, Gen...</td>
<td>STEM</td>
<td>1980-1999</td>
<td>4,608</td>
</tr>
<tr>
<td>Materials Sci...</td>
<td>STEM</td>
<td>1980-1999</td>
<td>2,118</td>
</tr>
<tr>
<td>Mathematics</td>
<td>STEM</td>
<td>1980-1999</td>
<td>15,816</td>
</tr>
<tr>
<td>Mechanical &amp; En...</td>
<td>STEM</td>
<td>1980-1999</td>
<td>2,727</td>
</tr>
<tr>
<td>Physics</td>
<td>STEM</td>
<td>1980-1999</td>
<td>4,598</td>
</tr>
<tr>
<td>Art Education/Ar...</td>
<td>Arts &amp; Humanities</td>
<td>2000+</td>
<td>12,775</td>
</tr>
<tr>
<td>Dance</td>
<td>Arts &amp; Humanities</td>
<td>2000+</td>
<td>474</td>
</tr>
<tr>
<td>Design</td>
<td>Arts &amp; Humanities</td>
<td>2000+</td>
<td>4,411</td>
</tr>
</tbody>
</table>

- Example of Excel data

<table>
<thead>
<tr>
<th>Collection</th>
<th>Usage Month</th>
<th>Usage Year</th>
<th>Total Checkouts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Art Education/Art History</td>
<td>1</td>
<td>2011</td>
<td>859</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>1</td>
<td>2012</td>
<td>768</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>1</td>
<td>2013</td>
<td>765</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>1</td>
<td>2014</td>
<td>613</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>1</td>
<td>2015</td>
<td>521</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>1</td>
<td>2016</td>
<td>432</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>2</td>
<td>2011</td>
<td>889</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>2</td>
<td>2012</td>
<td>862</td>
</tr>
<tr>
<td>Art Education/Art History</td>
<td>2</td>
<td>2013</td>
<td>1111</td>
</tr>
</tbody>
</table>

- Example of Tableau data

<table>
<thead>
<tr>
<th>Collection</th>
<th>Checkout Semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Art Education/Ar...</td>
<td>10/1/2010</td>
</tr>
<tr>
<td>Dance</td>
<td>10/1/2010</td>
</tr>
<tr>
<td>Dance &amp; Theater</td>
<td>10/1/2010</td>
</tr>
<tr>
<td>Design</td>
<td>10/1/2010</td>
</tr>
<tr>
<td>English</td>
<td>10/1/2010</td>
</tr>
</tbody>
</table>
What Chart When?

- **Comparison**
  - What would you like to show?
  - **Relationship**
  - **Distribution**

- **Composition**
  - Changing over time:
    - Few periods
    - Many periods
  - Static:
    - Only relative differences matter
    - Relative and absolute differences matter
    - Simple share of total
    - Components of components
  - Over time:
    - Few data points
    - Many data points
  - Stacked 100% column chart
  - Stacked column chart
  - Stacked 100% area chart
  - Stacked area chart
  - Pie chart
  - waterfall chart
  - Stacked 100% column chart with subcomponents

- **Variable Width Column Chart**
- **Table or Table with Embedded Charts**
- **Bar Chart**
- **Column Chart**
- **Circular Area Chart**
- **Line Chart**
- **Column Chart**
- **Line Chart**
- **Scatter Chart**
- **Bubble Chart**

- Two variables per item
- One variable per item
- Among items
- Over time

- Many categories
- Few categories
- Many periods
- Few periods
- Many categories
- Single or few categories
- Non-cyclical data
- Cyclical data

- Two variables
- Three variables
- Static

- Single variable
- Few data points
- Many data points

- 3D area chart
- Line histogram
- Scatter chart
Data and its Relationships

- Common fields - Collection, month/year
- Ways of linking data sources (joins)
- Connecting data in Tableau
- Overview of creating workbooks in Tableau
Combining Data: Blending vs Joining

**Join**
- Traditional
- Tables from *same* data source
  - Access database
  - Excel file
- Same as SQL:
  - Inner
  - Outer

**Blend**
- More flexible
  - From different data sources
  - Non-key fields
  - Different column labels
- At the worksheet level
  - Applies to one worksheet, but not another
  - Like a Left-Outer Join
- Simpler
Blend ~ Left-Outer Join

- Join
Exercise 5 Part 2: Tableau – Data Blending & Creating Ratios Using Calculated Fields

- **Part 2:**
  - Exercise Goal: Create a worksheet showing the trend of circulation and enrollment based upon collection and category using blended data.
  - Exercise Question: What can you interpret from change in checkouts to that of enrollment?
  - Excel files needed for this exercise: Checkouts by Collection Month and FY & Enrollment by Collection
  - Small Group Discussion: What worked? What didn’t? Why?
Dashboards

Priya & Janette
What is a dashboard?

“A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance”  
(Few, 2013, p. 26)
Why a dashboard?

Dashboards are about communication!
Kinds of Dashboard Based on Role

- **Operational**
  - Facilitate the *operation* of the business
  - Requires real-time data

- **Strategic/Executive**
  - Provides Key Performance Indicators (KPI’s)
  - Tracked periodically – real-time updating not required

- **Analytical**
  - Operational or strategic data
  - Drill-down for exploring
Purpose of a Dashboard

- Hovis (2002) states
  “The real value of dashboard products lies in their ability to replace hunt-and-peck data-gathering techniques with a tireless, adaptable, information-flow mechanism. Dashboards transform data repositories into consumable information” (p. 37).

- Improve “span of control” of your data
Know Your Audience

- Public Libraries
- Academic Libraries
- Other Institutions

- Who are the Stakeholders?
  - Internal Management
  - Advisory Board
  - Governing Board
  - Accreditation Board
  - Steering Committees
  - Working Groups
  - Patrons
  - Library support groups
  - Grant/Fundraising

Focus on a single type of user
Determining the Data Sources

- Organization-wide
- Departmental
- Individual
- Cross-segments

The data source used highlights quantitative measures of the current activity.
Exercise 6: Tableau – Dashboard Creation

**Exercise Goal:** Create a dashboard using worksheets generated throughout the preconference sessions to convey specifically targeted information.

**Exercise Question:** What does the data selected convey to your target audience? Why did you select the worksheets that you did for the final dashboard?
The Collections Dashboard

Karen
The Collections Dashboard: Beginnings

- Drowning in a sea of data.
- Data silos.
- Inputs & Outputs, but no Impact.
The Collections Dashboard: Setting the Vision

We expect the first iteration of the dashboard to help answer the following key questions:

1. What are our users’ needs?
2. In what ways do the UNT Libraries’ collections meet our users' needs?
3. How well are we meeting those needs?
The Collections Dashboard: Defining the Scope

- **What** are those needs?
  - Enrollment by subject & level

- **In what ways** do we meet our users' needs?
  - Distribution by subject and format of...
    - DDA
    - Holdings
    - ILL requests
    - Circulation

- **How well** are we meeting our users' needs?

  - **Compare** of the distribution of enrollment (need) with distributions by subject and level of...
    - Holdings
    - DDA
    - ILL requests
    - Circulation

  - **Compare** needs with holdings, usage, and ILL requests
The Collections Dashboard: Scaffolding

- Gather the data
  - Sources
  - Queries
  - Anonymization

- Standardize the data
  - College/Department (enrollment) <-> Affiliation (ILL) <-> Major (Circulation)
  - Years of coverage (greatest common factor)
  - Fiscal years (all start in September)
The Collections Dashboard: Views of Each Source

Absolute View
- Actual counts
- Shows trends over time
- Represents absolute values

By Collection Map
- Sum of its parts is greater than the whole.
- Relative distribution
- Allows comparisons
The Collections Dashboard: Overall Dashboards

- Answers the 3 questions:
  - What are our needs?
  - In what ways do we meet those needs?
  - How well do we meet those needs?
- Common factor of comparison: By Collection
- Trends over time
- Scatterplots of percentages
Assessing Effectiveness: Ratios

Ratios = efficient measures of effectiveness

\[
\frac{\text{Percentage of dependent factor}}{\text{Percentage of independent factor}}
\]

Examples:

- Percentage of Expected Use (PEU) = \% of Circs / \% of Holdings
- Percentage of Expected Holdings (PEH) = \% of Holdings / \% of Enrollment
- Ratio of Borrowings to Holdings (RBH) = \% ILL borrowing requests / \% Holdings
How well? Evaluating Ratios

Zero to One
Numerator < Denominator
E.g. % Holdings < % Enrollment

Around One
About right
The percentages are close.

Greater than One
Numerator > Denominator
E.g. % Holdings > % Enrollment
The UNT Collections Dashboards
http://goo.gl/O4RjJV

- Enrollment (Needs)
- Holdings (ways to meet needs)
- Circulation (how well needs are met)
- ILL Requests (both ways & how well)

- Overall dashboard:
  - Enrollment trends
  - Circulation trends
  - ILL requests trends
  - Ratios:
    - PEH/PEU
    - PEH/RBH
Dashboards as Decision Making Tools

3 questions of our dashboards

- Where to spend an extra $15,000 in funds? Or, conversely, Where to cut an additional $250,000?
- ACS (American Chemistry Society) accreditation report – how well does your library meet the needs of chemistry students?
- Grant to survey users about library collections. Which user-groups should you survey first?
End of Program Review

- Evaluation Survey: https://www.surveymonkey.com/r/litaforumprecon2

- Communities of Practice
  - Tableau
  - Library Assessment
    - ARL Library Assessment
    - Library Assessment Community
  - ARL-Assess Listserv: ARL-ASSESS@arl.org

- Collection Assessment
  - Coll-Assess Listserv: Coll-Assess@arl.org
References