Easy Interactive Charts and Maps with Tableau

Eric Monson
Data Visualization Analyst
Data and Visualization Services

Fall 2017

Data, handouts & slides: http://bit.ly/IntroTableauFall2017

Today

Using Tableau Public to create visualizations to explore a dataset:

- Overview of the program structure
- Terminology used
- Walk-through of a data visualization project
- Exercises on your own with new data

Conceptual introduction

We add things up and split them
to get an overview and see patterns

Order Date	Sales	Product Category
10/13/10	$261.54	Office Supplies
10/1/12	$10,123.02	Office Supplies
10/1/12	$244.57	Office Supplies
7/10/11	$4,965.76	Technology
8/28/10	$394.27	Office Supplies
8/28/10	$146.69	Furniture
6/17/11	$93.54	Office Supplies
6/17/11	$905.08	Office Supplies
3/24/11	$2,781.82	Office Supplies
2/26/10	$228.41	Office Supplies
11/23/10	$196.85	Office Supplies
11/23/10	$124.56	Office Supplies
6/8/12	$716.84	Office Supplies
6/8/12	$1,474.33	Technology
8/4/12	$80.61	Office Supplies
5/30/11	$1,815.49	Furniture

Conceptual introduction

We add things up and split them
to get an overview or see patterns

Sum of Sales
$14,915,600.82

Conceptual introduction

We add things up and split them
to get an overview or see patterns

Product Categories	Sum of Sales
Technology	$5,984,248.18
Furniture	$5,178,590.54
Office Supplies	$3,752,762.10
Grand Total	**$14,915,600.82**

Conceptual introduction

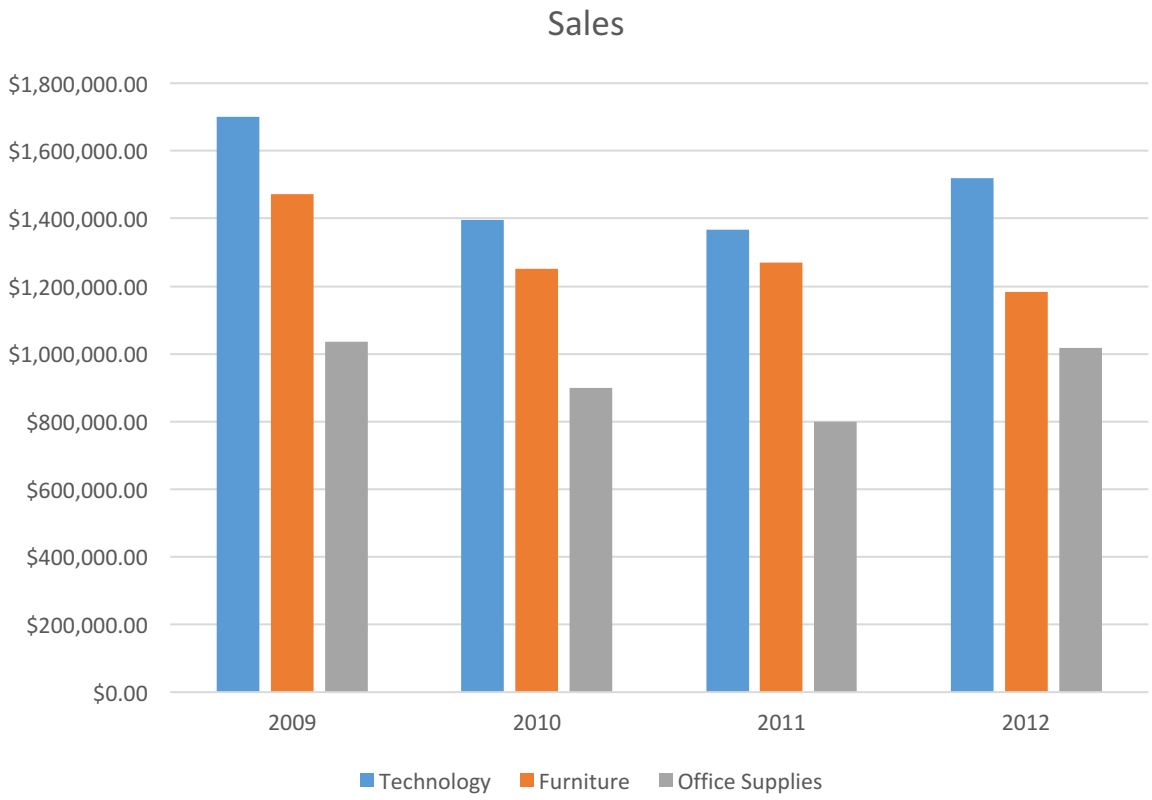
We add things up and split them
to get an overview or see patterns

Sum of Sales	2009	2010	2011	2012	Grand Total
Technology	$1,701,825.48	$1,397,142.14	$1,366,806.58	$1,518,473.98	$5,984,248.18
Furniture	$1,472,671.72	$1,252,665.11	$1,269,661.09	$1,183,592.62	$5,178,590.54
Office Supplies	$1,034,642.25	$899,873.56	$800,349.03	$1,017,897.26	$3,752,762.10
Grand Total	**$4,209,139.46**	**$3,549,680.80**	**$3,436,816.70**	**$3,719,963.86**	**$14,915,600.82**

Conceptual introduction

Sum of Sales					
	2009	2010	2011	2012	Grand Total
Technology	$1,701,825.48	$1,397,142.14	$1,366,806.58	$1,518,473.98	$5,984,248.18
Furniture	$1,472,671.72	$1,252,665.11	$1,269,661.09	$1,183,592.62	$5,178,590.54
Office Supplies	$1,034,642.25	$899,873.56	$800,349.03	$1,017,897.26	$3,752,762.10
Grand Total	**$4,209,139.46**	**$3,549,680.80**	**$3,436,816.70**	**$3,719,963.86**	**$14,915,600.82**

And then visualize to make those patterns and comparisons more clear



Conceptual introduction

But how do we start from this individual event data and explore and communicate patterns?

Order Date	Sales	Product Category
10/13/10	$261.54	Office Supplies
10/1/12	$10,123.02	Office Supplies
10/1/12	$244.57	Office Supplies
7/10/11	$4,965.76	Technology
8/28/10	$394.27	Office Supplies
8/28/10	$146.69	Furniture
6/17/11	$93.54	Office Supplies
6/17/11	$905.08	Office Supplies
3/24/11	$2,781.82	Office Supplies
2/26/10	$228.41	Office Supplies
11/23/10	$196.85	Office Supplies
11/23/10	$124.56	Office Supplies
6/8/12	$716.84	Office Supplies
6/8/12	$1,474.33	Technology
8/4/12	$80.61	Office Supplies
5/30/11	$1,815.49	Furniture
11/25/09	$248.26	Office Supplies
2/14/12	$4,462.23	Furniture
2/14/12	$663.78	Furniture
4/15/12	$834.90	Technology
4/15/12	$2,480.92	Technology

Tableau will visualize by…

- ***Categorical data***
  (text [string], date or boolean [True/False] values)
  Split up the numbers – set the granularity, or the level of detail
- ***Numerical values***
  Aggregated (sum, mean, etc)
  broken up according to the categories

What can Tableau make?

- Text tables
- Heat maps
  a grid representing variables by size and color
- Highlight tables
  a grid representing variables by text and color
- Maps (symbol, filled)
- Pie charts
- Horizontal bars
- Stacked bars
- Side-by-side bars
- Treemaps
  a grid representing variables by size
- Circle views
- Side-by-side circles
- Lines/Area charts
- Lines/Area charts (discrete)
- Dual lines
- Dual combination
- Scatter plots
- Histograms
- Box-and-whisker plots
- Gantt charts
- Bullet graphs
- Packed bubbles/ Word cloud

Tableau Public

- Free program
- Data limits
  - 15,000,000 rows of data
  - Excel, Google Sheets (live every 24 hrs), CSV, MS Access
- Publishing limits
  - Can only save work to the **web** (10GB storage space per account)
  - Can now disable data download, though

Tableau Desktop

- Allows you to save files privately
- Additional data connection options (databases, etc.)
- No data size limits

Free for:

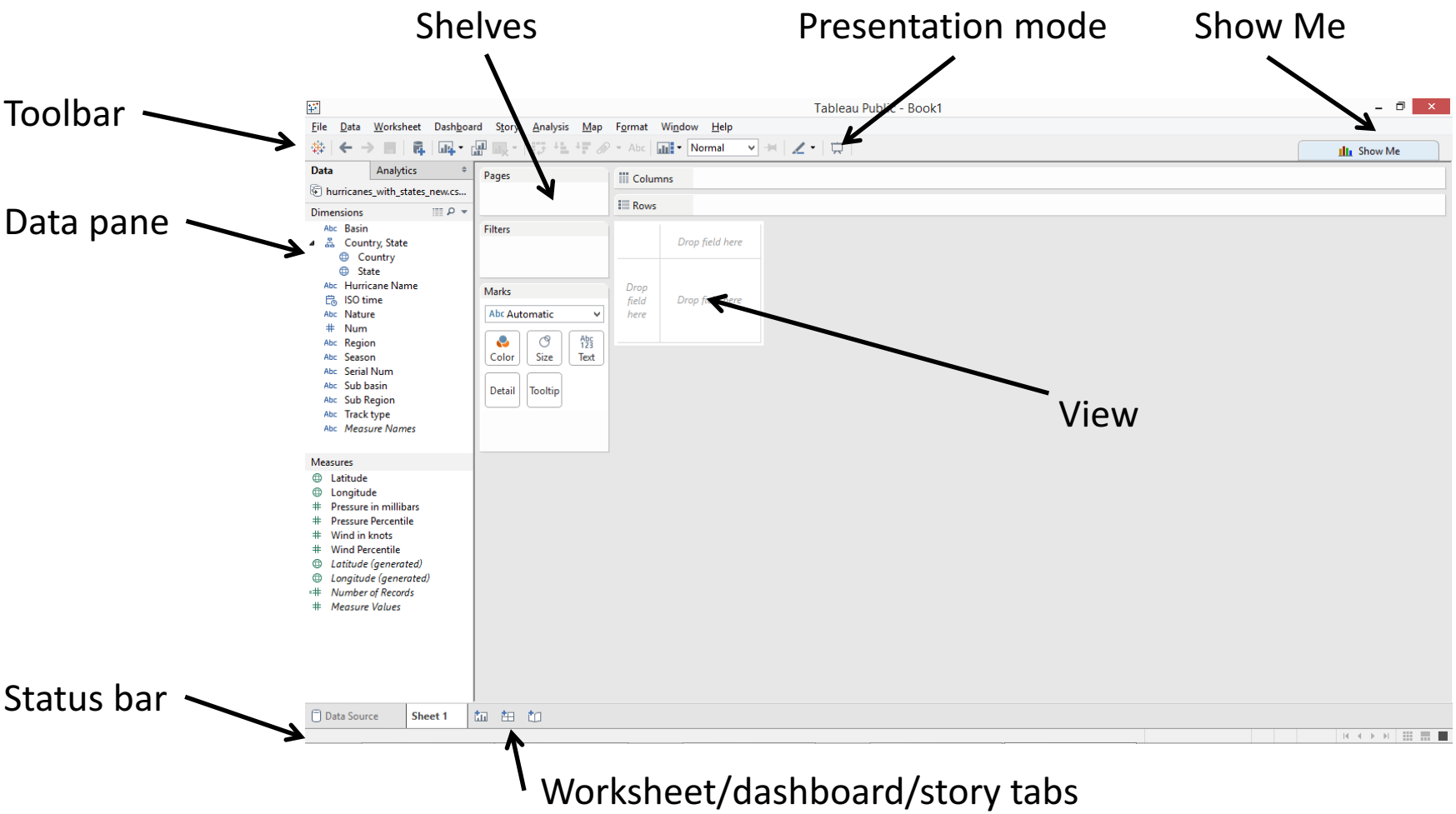
- 2 week trial
  http://www.tableau.com/products/trial
- students (year license – renewable)
  http://www.tableau.com/academic/students
- teachers using it in a class, semester license
  http://www.tableau.com/academic/teaching

Workbooks (Tableau .twbx files)

Contain one or more:

- Worksheets
  - Like a spreadsheet (a working space where data are organized/analyzed), but for creating visualizations (or “views”)
  - one (possibly complex) view per worksheet
- Dashboards
  - A presentation space where the views created in worksheets can be arranged and linked to produce a more complete visualization environment

The Tableau Workspace

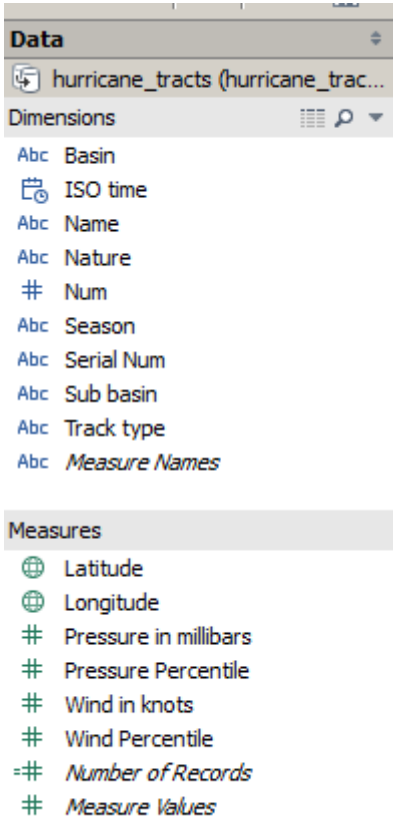


Online help:

http://onlinehelp.tableau.com/current/pro/online/windows/en-us/help.htm

Data Types

- **Dimensions** (categorical – text, date or boolean values)
  Set the level of detail (LOD)
- **Measures** (numerical values)
  Aggregated (sum, mean, etc) to the LOD of the Dimensions in the view



http://www.tableausoftware.com/public/knowledgebase/understanding-data-types-and-roles

EXAMPLE DATA SET: HURRICANE TRACTS

Source:
http://www.ncdc.noaa.gov/ibtracs/index.php?name=wmo-data
(downloaded July 2014)

Discrete versus continuous dates

Arrow pop-up menu on the date pill

- Options on the top are **Discrete date parts**
- Options on the bottom are **Continuous date truncations**

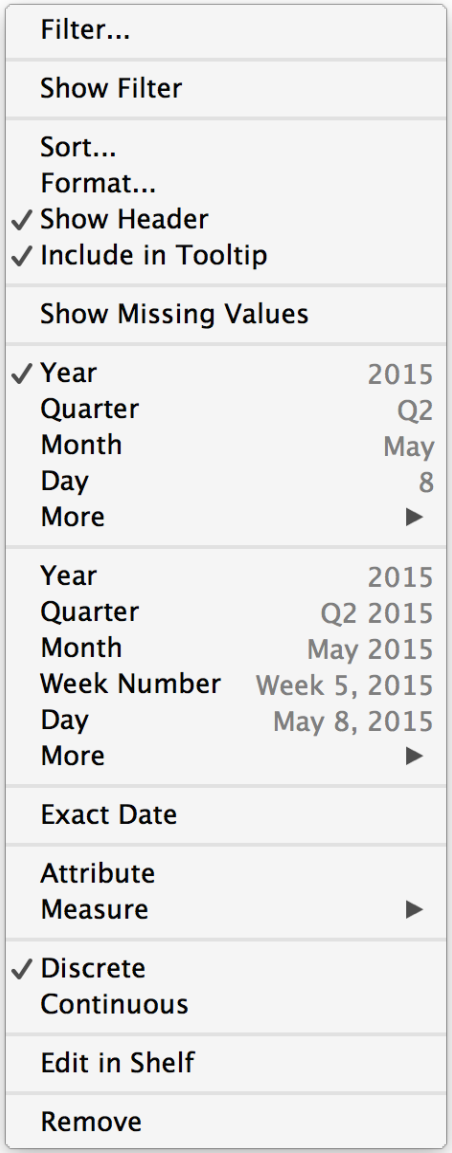


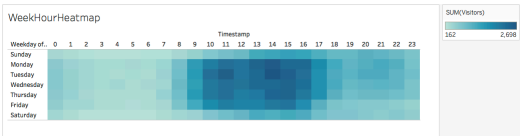
Tableau has a really good tutorial
http://www.tableau.com/learn/tutorials/on-demand/understanding-pill-types

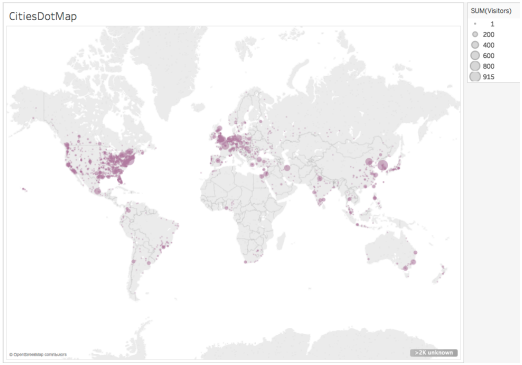
Library web site page views exercises

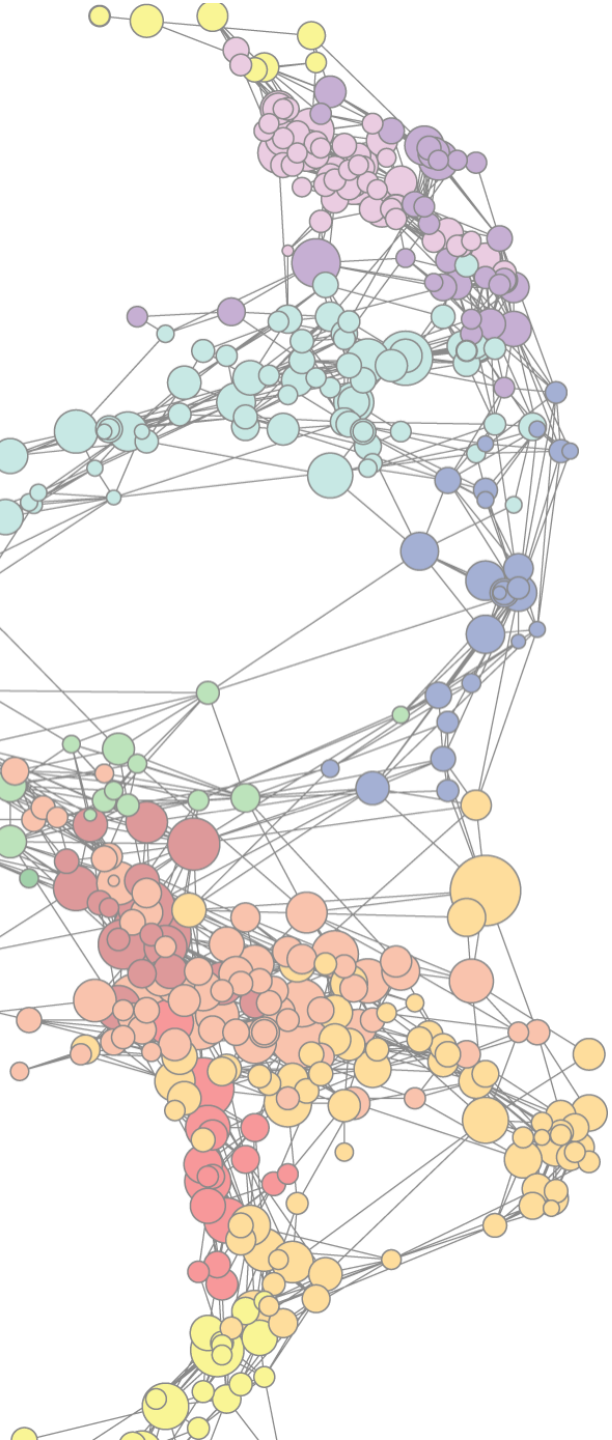
pageviews_2012_subset.xlsx

- Identify winter, summer and spring break from a continuous line plot of the sum of visitors per day.
- Summed over all the data, what hour of the day has peak site viewing?
- Create a heatmap of sum of visitors with weekday vs hour of day as the axes.
- Show the hour of day time shift in the sum of visitors peak between Canada and the UK.
- Make a proportional symbol map of sum of visitors for each city, but exclude all North Carolina cities.

Timestamp	Lcc First Letter	Visitors	City	State	Country
8/28/2011 12:00:00 ...	B	1	Manila	Manila	Philippines
8/28/2011 12:00:00 ...	B	1	Durham	North Carolina	United States
8/28/2011 12:00:00 ...	G	1	Durham	North Carolina	United States
8/28/2011 12:00:00 ...	H	1	Durham	North Carolina	United States
8/28/2011 12:00:00 ...	J	1	Durham	North Carolina	United States
8/28/2011 12:00:00 ...	J	1	Durham	North Carolina	United States
8/28/2011 12:00:00 ...	P	1	Durham	North Carolina	United States
8/28/2011 12:00:00 ...	P	1	Durham	North Carolina	United States
8/28/2011 12:00:00 ...	Z	1	Durham	North Carolina	United States
8/28/2011 1:00:00 AM	B	1	Kitchener	Ontario	Canada
8/28/2011 1:00:00 AM	J	1	Durham	North Carolina	United States
8/28/2011 1:00:00 AM	T	1	Durham	North Carolina	United States

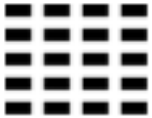






Duke Data and Visualization Services

Support Areas



Data Sources

Data Analysis

Data Management

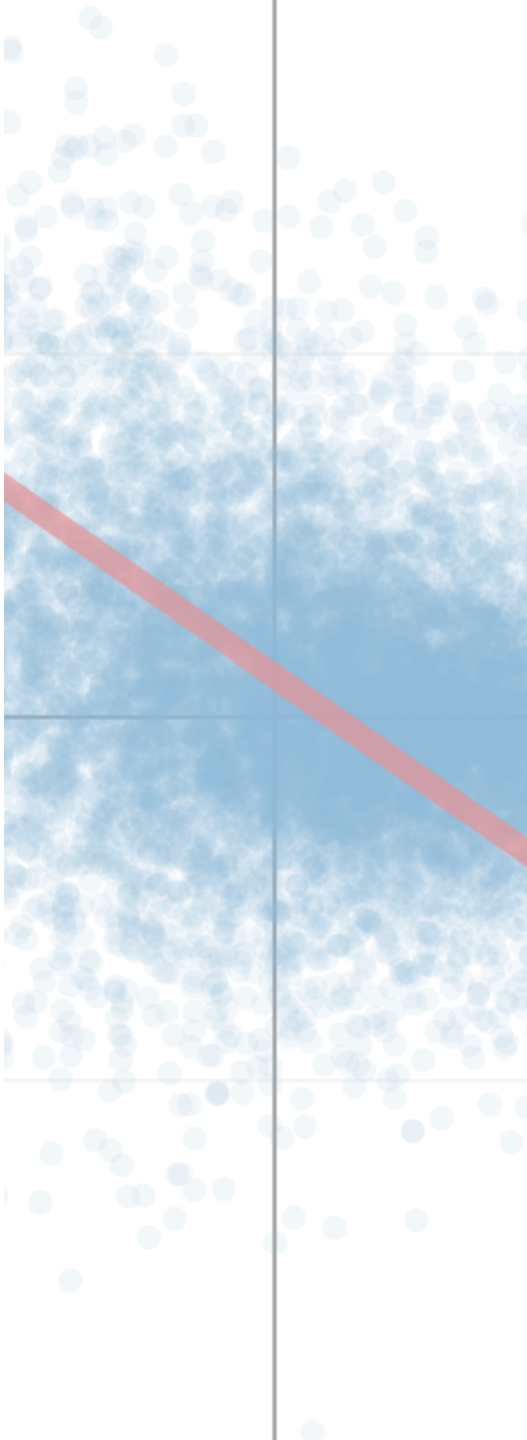
Mapping and GIS

Data Cleaning

Data Visualization

Basic services

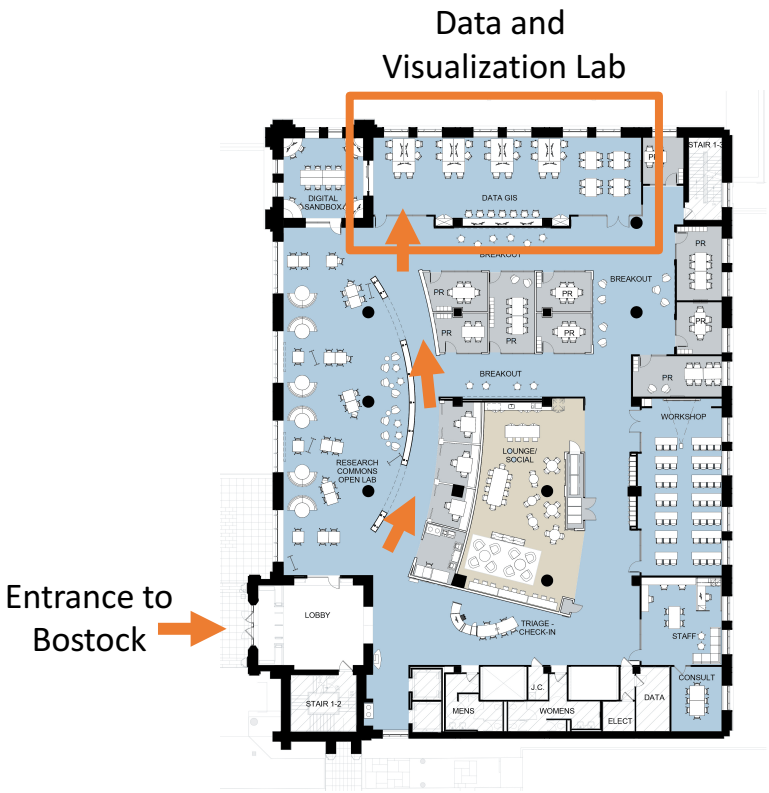
- Computer lab *(The Edge)*
- Consultations *(walk-in or scheduled)*
- Data collections
- Workshops
- Online instructional materials



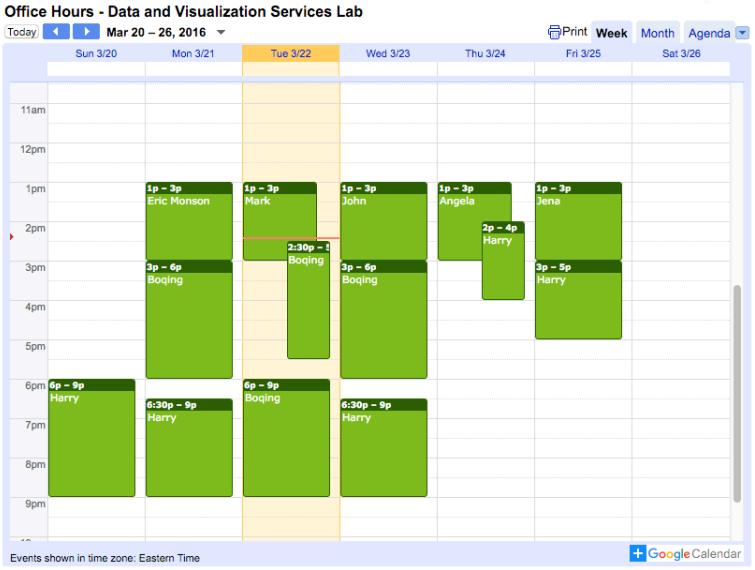
Brandaleone Family Lab for Data and Visualization Services

- **The Edge** *(1st floor of Bostock)*
- Open whenever the library is open
- 12 high-powered Dell workstations
- 3 Bloomberg financial workstations
- Software for data analysis, GIS, and visualization

http://library.duke.edu/data/about/lab



Consulting



http://library.duke.edu/data/about/schedule

…or email
askdata@duke.edu
for an appointment

Types of visualization consulting

- Look at data and brainstorm about the best visualization
- Recommend appropriate tools
- Troubleshoot software problems
- Help with cleaning and structuring data
- Offer graphic design advice for graphics and posters

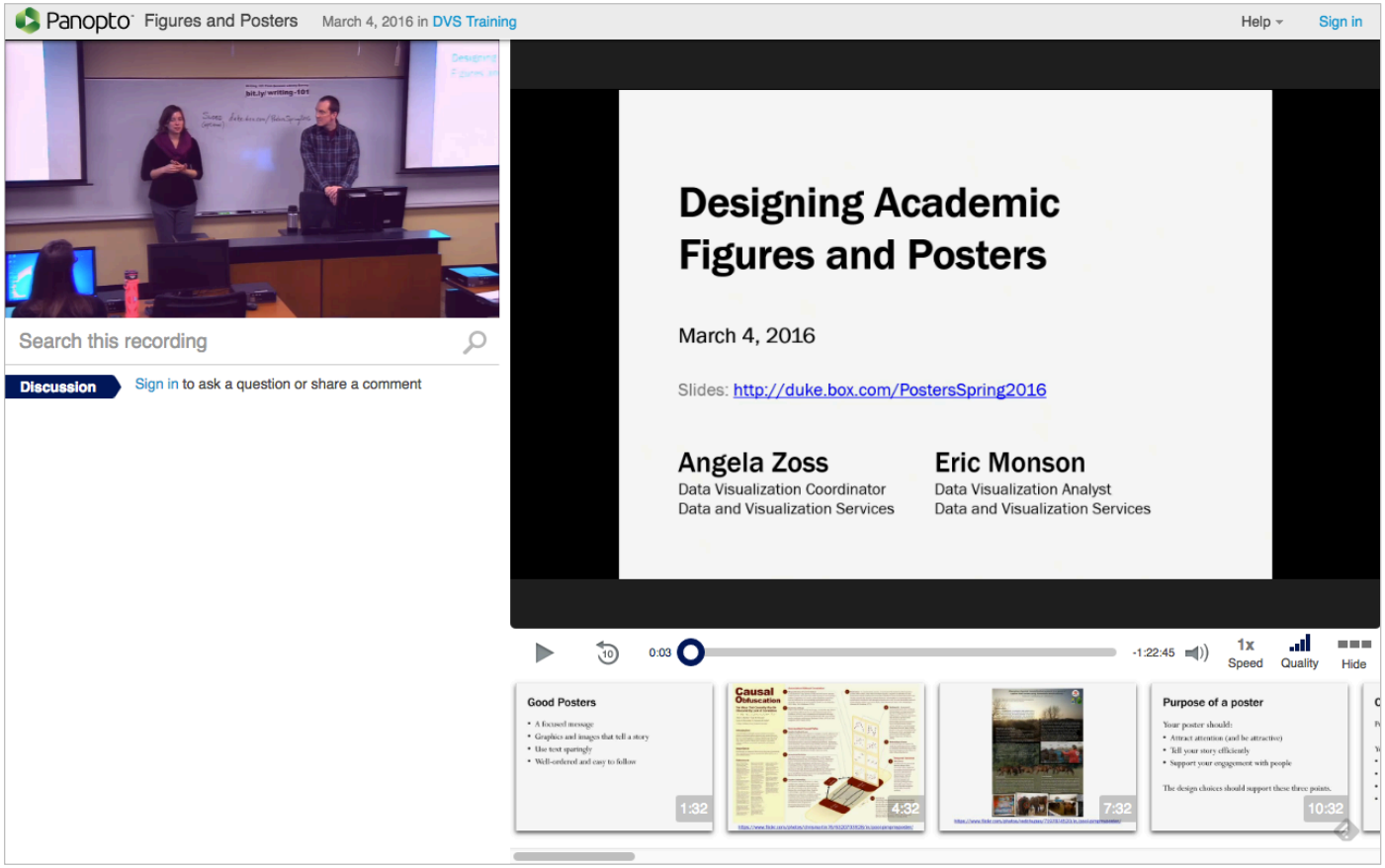
Fall 2017 visualization workshops

- Introduction to Effective Data Visualization · *Sept 12*
- Easy Interactive Charts and Maps with Tableau · *Sept 14*
- Data Visualization with Excel · *Sept 20*
- Visualization in R using ggplot2 · *Sept 25*
- Adobe Illustrator to Enhance Charts and Graphs · *Sept 29*
- Visualizing Qualitative Data · *Oct 13*
- Tidy Data Visualization with Python · *Nov 9*

Stay tuned at:
http://library.duke.edu/data/news

Videos of past workshops

http://bit.ly/DVSvideos



Visualization Friday Forum

- Every Friday during the semester
- Noon in LSRC D106
- Free lunch (sssh!)
- Live streamed and recorded
- email me to get on the mailing list

QUESTIONS

askdata@duke.edu