INTRODUCTION TO DATA & INFORMATION VISUALIZATION

iSkills
University of Toronto Libraries

February 10, 2017
Jesse Carliner
COMMUNICATIONS & REFERENCE LIBRARIAN

Nicholas Worby
GOVERNMENT INFORMATION & STATISTICS LIBRARIAN

University of Toronto Libraries
WHY DATA VISUALIZATION?

- COMMUNICATION
- GAIN INSIGHTS
VISUALIZATION WORKFLOW

WORKFLOW DESIGN

Audience & Purpose → Select and Prepare Data → Share → Select Visualization Form → Select Visualization Elements → Visualize Data
STAGE 1: AUDIENCE & PURPOSE

Audience & Purpose
STAGE 2: SELECT & PREPARE DATA

WORKFLOW DESIGN

Audience & Purpose → Share → Select Visualization Elements

Select and Prepare Data → Select Visualization Form

Visualize Data
STAGE 2: TYPES OF DATA

- QUANTITATIVE
- QUALITATIVE
STAGE 2: MEASUREMENT SCALE TYPE

- CATEGORICAL (NOMINAL)
- ORDINAL
- INTERVAL
- RATIO
IDENTIFY THIS DATA SCALE TYPE?

<table>
<thead>
<tr>
<th>FAMILIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ridolfi</td>
</tr>
<tr>
<td>Salviatti</td>
</tr>
<tr>
<td>Albizzi</td>
</tr>
<tr>
<td>Ginori</td>
</tr>
</tbody>
</table>
IDENTIFY THIS DATA SCALE TYPE?

<table>
<thead>
<tr>
<th>POPULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
</tr>
<tr>
<td>131</td>
</tr>
<tr>
<td>47</td>
</tr>
<tr>
<td>39</td>
</tr>
</tbody>
</table>
**IDENTIFY THIS DATA SCALE TYPE?**

<table>
<thead>
<tr>
<th>Quality of Sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very good</td>
</tr>
<tr>
<td>Good</td>
</tr>
<tr>
<td>Fair</td>
</tr>
<tr>
<td>Poor</td>
</tr>
</tbody>
</table>
STAGE 2: TYPES OF DATA

- GEOGRAPHIC DATA
- NETWORK DATA
- TEMPORAL DATA
- TOPICAL DATA
- NUMERIC DATA
STAGE 2: UNIT OF OBSERVATION

Microdata: Unaggregated, responses at the individual level

Statistics: Aggregated from microdata, responses grouped into categories; shows counts, percentages, averages, etc.
STAGE 2: LEVEL OF ANALYSIS

- SMALL ↔ LARGE
STAGE 2: OTHER FACTORS

- PRIVACY?
- HOW MANY VARIABLES?
STAGE 2: PREPARE YOUR DATA

WORKFLOW DESIGN

Audience & Purpose

Select and Prepare Data

Share

Select Visualization Elements

Select Visualization Form

Visualize Data
STAGE 2: PREPARE YOUR DATA

- Aggregating
- Dealing with missing values & outliers
- Cleaning, sorting, labelling variables
- Summary statistics, algorithms, correlation coefficients, etc.
- Merging with other datasets

It’s iterative!
STAGE 2: PREPARE YOUR DATA

<table>
<thead>
<tr>
<th>Country</th>
<th>OECD Status</th>
<th>Identifier</th>
<th>Date</th>
<th>Spending on services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>OECD</td>
<td>124</td>
<td>2012-01-01</td>
<td>75000000</td>
</tr>
<tr>
<td>Canada</td>
<td>OECD</td>
<td>124</td>
<td>2013-01-01</td>
<td>77000000</td>
</tr>
<tr>
<td>Germany</td>
<td>OECD</td>
<td>276</td>
<td>2012-01-01</td>
<td>42000000</td>
</tr>
<tr>
<td>Germany</td>
<td>OECD</td>
<td>276</td>
<td>2013-01-01</td>
<td>41000000</td>
</tr>
<tr>
<td>Germany</td>
<td>OECD</td>
<td>276</td>
<td>2014-01-01</td>
<td>44000000</td>
</tr>
<tr>
<td>Germany</td>
<td>OECD</td>
<td>276</td>
<td>2015-01-01</td>
<td>47000000</td>
</tr>
<tr>
<td>China</td>
<td>Non-OECD</td>
<td>156</td>
<td>2012-01-01</td>
<td>79000000</td>
</tr>
<tr>
<td>China</td>
<td>Non-OECD</td>
<td>156</td>
<td>2013-01-01</td>
<td>75000000</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Non-OECD</td>
<td>188</td>
<td>2012-01-01</td>
<td>95000000</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Non-OECD</td>
<td>188</td>
<td>2013-01-01</td>
<td>90000000</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Non-OECD</td>
<td>188</td>
<td>2014-01-01</td>
<td>92000000</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Non-OECD</td>
<td>188</td>
<td>2015-01-01</td>
<td>80000000</td>
</tr>
</tbody>
</table>

Table 1: Spending on services
Prepared March 14, 2016

<table>
<thead>
<tr>
<th>OECD Countries</th>
<th></th>
<th>Spending ($ Millions)</th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Identifier</td>
<td>2012-01</td>
<td>2013-01</td>
<td>2014-01</td>
<td>2015-01</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>124</td>
<td>7.5</td>
<td>7.7</td>
<td>9</td>
<td>8.4</td>
<td>32.6</td>
</tr>
<tr>
<td>Germany</td>
<td>276</td>
<td>4.2</td>
<td>4.1</td>
<td>4.4</td>
<td>4.7</td>
<td>17.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-OECD Countries</th>
<th>Identifier</th>
<th>Spending ($ Millions)</th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Identifier</td>
<td>2012-01</td>
<td>2013-01</td>
<td>2014-01</td>
<td>2015-01</td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>156</td>
<td>7.9</td>
<td>7.5</td>
<td>7.7</td>
<td>7.6</td>
<td>30.7</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>188</td>
<td>9.5</td>
<td>9</td>
<td>9.2</td>
<td>8</td>
<td>35.7</td>
</tr>
</tbody>
</table>
WORKFLOW DESIGN

- Audience & Purpose
  - Select and Prepare Data
- Share
- Select Visualization Elements
  - Select Visualization Form
- Visualize Data
STAGE 3: CHOOSE VISUALIZATION TYPE

- Based on purpose
  - Data exploration or conveying thesis?
  - Overall trends or accurate judgments?
- Based on audience
  - Academic or hospital stakeholders?
- Based on data
  - Quantitative, qualitative, categorical, scale (ordinal, interval), scope, temporal & geospatial considerations
Graphical perceptual accuracy, Cleveland & McGill 1984

Source: Cairo, Alberto. *The Functional Art*, p.120
Jock Mackinlay, 1986

Quantitative
- Position
- Length
- Angle
- Slope
- Area
- Volume
- Density
- Saturation
- Hue
- Texture
- Connection
- Containment
- Shape

Ordinal
- Position
- Density
- Saturation
- Hue
- Texture
- Connection
- Containment
- Length
- Angle
- Slope
- Area
- Volume
- Shape

Nominal
- Position
- Hue
- Texture
- Connection
- Containment
- Density
- Saturation
- Shape
- Length
- Angle
- Slope
- Area
- Volume

[Mackinlay, Automating the Design of Graphical Presentations of Relational Information, 1986]
RESOURCES FOR CHOOSING VISUALIZATION TYPES

http://datavizcatalogue.com/

Handout
PIE CHARTS

- Lack an inherent reference system (no x, y, z axis or grid lines)
- Use angle to convey difference
BAR GRAPHS

• Use with when comparing quantitative variables across discrete categorical variables
• Counts or sums rather than characteristics of a distribution
• Length or height more accurate than angles
BAR GRAPHS

• Use consistent axes when comparing two graphs
BAR GRAPHS

Use a zero baseline!
SCATTER PLOTS

• Effective for comparing how one variable is affected by another, e.g. correlation, cause & effect relationships
• Pros: Easy to spot outliers & trends
• Cons: Hard to use with REALLY large datasets, occlusion issues with multivariate scatterplots
Best practices:
• Prevent occlusion & ambiguity issues with easily distinguishable symbols
• Consider colour, size, & symbols to capture more dimensions of your data
SCATTER PLOTS

Separate multivariate scatter plots with occlusion issues into multiple scatter plots.
Consider scatterplot matrices (aka Trellis plots) for large, multivariate sets.

Example scatter plots:

- (a) shows a positive correlation between Variable 1 and Variable 2.
- (b) shows a negative correlation between Variable 1 and Variable 2.
- (c) shows a less clear relationship between Variable 1 and Variable 2.
- (d) shows a strong clustering of points in Variable 2 at lower values of Variable 1.
- (e) shows a point cloud with a distinct outlier.
- (f) shows a dense cluster of points in Variable 2 for a specific range of Variable 1.

The scatterplot matrix on the right consists of individual scatter plots for each pair of variables:

- age
- high_school
- bach_degree
- adv_degree
- unemploy
- hh_income
- rent

The plots display the relationship between different variables, with age, high_school, bach_degree, adv_degree, unemploy, hh_income, and rent as the variables of interest.
“SMALL MULTIPLES” & “SPARKLINES”
HEAT MAPS

NBA per game performance of top 50 scorers

2008-2009 season

[Heat Map Image]
VISUALIZING PARTS OF A WHOLE: TREEMAPS

Accuracy
- position
- length
- direction
- angle
- area
- volume
- curvature
- shading
- colour
VISUALIZING PARTS OF A WHOLE: STREAMGRAPHS
NETWORK GRAPHS

• Track activity & relationships across temporal, geo., & categorical variables
• Can analyze hubs & clusters of activity
• May need to apply algorithms to fix layout issues
• Labels can often get obscured
TABLES

- Use for a small set of quantitative data that can’t be quickly summarized in text
- Edward Tufte’s rule: Fewer than 20 numbers, just use a table*
- Takes up less space than graphs
- Author submission guidelines often spell out when to just use a table

1. I want to compare obesity rates across countries across a 5 year span.

- Which variables will you need?
- What are you trying to observe? (e.g. distribution, relationships, data over time, etc.)
- Which visualization type(s) would you use?

<table>
<thead>
<tr>
<th>Country</th>
<th>Obese % of population 2005</th>
<th>Obese % of population 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>23.7</td>
<td>25.4</td>
</tr>
<tr>
<td>New Zealand</td>
<td>26</td>
<td>27.8</td>
</tr>
<tr>
<td>United States</td>
<td>33</td>
<td>36.1</td>
</tr>
<tr>
<td>Spain</td>
<td>etc.</td>
<td>etc.</td>
</tr>
<tr>
<td>Russia</td>
<td>etc.</td>
<td>etc.</td>
</tr>
</tbody>
</table>

Use [http://datavizcatalogue.com](http://datavizcatalogue.com) or the handout
2. I want to explore the relationship between obesity rates across countries, across a 5 year span in relation to GDP per capita.

- Which variables will you need?
- What are you trying to observe? (e.g. distribution, relationships, data over time, etc.)
- Which visualization type(s) would you use?

<table>
<thead>
<tr>
<th>Country</th>
<th>Obese % of population 2005</th>
<th>Obese % of population 2010</th>
<th>GDP per capita 2005</th>
<th>GDP per capita 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>23.7</td>
<td>25.4</td>
<td>36000</td>
<td>47465</td>
</tr>
<tr>
<td>New Zealand</td>
<td>26</td>
<td>27.8</td>
<td>27525</td>
<td>32845</td>
</tr>
<tr>
<td>United States</td>
<td>33</td>
<td>36.1</td>
<td>44307</td>
<td>48377</td>
</tr>
<tr>
<td>Russia</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
</tbody>
</table>

Use [http://datavizcatalogue.com](http://datavizcatalogue.com) or the handout
STAGE 4: SELECT VISUAL ELEMENTS

WORKFLOW DESIGN

Select and Prepare Data → Share → Select Visualization Form → Select Visualization Elements → Visualize Data → Audience & Purpose
STAGE 4: SELECT VISUAL ELEMENTS

- POSITION
- FORM
- COLOUR
- OPTICS
- TEXTURE
STAGE 4: SELECT VISUAL ELEMENTS

POSITION
STAGE 4: SELECT VISUAL ELEMENTS

POSITION
STAGE 4: SELECT VISUAL ELEMENTS

POSITION
STAGE 4: SELECT VISUAL ELEMENTS

FORM
STAGE 4: SELECT VISUAL ELEMENTS

Visualization from Information is beautiful: http://www.informationisbeautiful.net/visualizations/best-in-show-whats-the-top-data-dog/
STAGE 4: SELECT VISUAL ELEMENTS

FORM

Visualization from Information is beautiful: http://www.informationisbeautiful.net/visualizations/snake-oil-supplements/
STAGE 4: SELECT VISUAL ELEMENTS

COLOUR
STAGE 4: SELECT VISUAL ELEMENTS

COLOUR ACCESSIBILITY

COLOUR BLINDNESS SIMULATOR: http://www.color-blindness.com/coblis-color-blindness-simulator/
STAGE 4: SELECT VISUAL ELEMENTS

COLOUR

COLOUR BREWER:
http://www.colorbrewer2.org/#
STAGE 4: SELECT VISUAL ELEMENTS

QUALITATIVE COLOUR PALETTE

Visualization from Information is beautiful: http://www.informationisbeautiful.net/visualizations/oil-well-every-cooking-oil-compared/
STAGE 4: SELECT VISUAL ELEMENTS

SEQUENTIAL COLOUR PALETTE

Visualization from Guardian: http://www.theguardian.com/money/ng-interactive/2014/may/23/-sp-see-how-house-prices-have-risen
STAGE 4: SELECT VISUAL ELEMENTS

DIVERGING COLOUR PALETTE

Visualization from Guardian: http://www.theguardian.com/society/ng-interactive/2015/sep/02/unaffordable-country-where-can-you-afford-to-buy-a-house
STAGE 4: SELECT VISUAL ELEMENTS

WHAT COLOUR PALETTE (S) WOULD YOU CHOOSE FOR THIS DATA SET?

COLOUR BREWER: http://www.colorbrewer2.org/

<table>
<thead>
<tr>
<th>FAMILIES</th>
<th>ACRES OF LAND</th>
<th>RANK</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIDOLFI</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>SALVATTI</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>ALBIZZI</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>GINORI</td>
<td>10</td>
<td>2</td>
</tr>
</tbody>
</table>
STAGE 4: CREATE TEXT

- INCLUDE ALL NECESSARY CONTENT
- CLARITY
- LESS IS MORE
- TYPOGRAPHIC BEST PRACTICES
STAGE 4: LAYOUT AND CREATION

- DATA-INK RATIO
- GESTALT
- GRIDS
STAGE 4: DATA-INK RATIO

STAGE 4: LAYOUT AND CREATION

GESTALT
STAGE 4: LAYOUT AND CREATION

- SIMILARITY
- PROXIMITY
STAGE 4: LAYOUT AND CREATION

- SIMILARITY
- PROXIMITY
STAGE 4: LAYOUT AND CREATION

GROUPING
STAGE 4: LAYOUT AND CREATION

VISUAL COMPLETION

OR

LESS IS MORE
STAGE 4: LAYOUT AND CREATION

AVOID DISTRACTIONS

THIS OR THIS?
STAGE 5: SHARE YOUR VISUALIZATION

WORKFLOW DESIGN

Audience & Purpose → Share → Select Visualization Elements → Visualize Data

Select and Prepare Data → Share → Select Visualization Form
STAGE 5: SHARING YOUR VISUALIZATION

- SHARE
- GET FEEDBACK
- WAS IT EFFECTIVE?
- REVISE
EXERCISE: IMPROVING VISUALIZATIONS

- In groups of 2 or 3, examine the visualization on your handout.
- These are problematic visualizations. What are the issues and how could they be improved?
Chart 2 - Total Expenditures on Health as a Percentage Share of GDP, by OECD Country, 2004

Note: For the United States the 2004 data reported here do not match the 2004 data point for the United States in Chart 1 since the OECD uses a slightly different definition of "total expenditures on health" than that used in the National Hi-

Figure 10: Public- and private-sector jobs (000s) in Ontario, 1993–2013

Source: Statistics Canada, CANSIM Table 282-0089; employment by class of worker and sex, seasonally adjusted and unadjusted. Ontario, Public sector and private sector employees; Both sexes; Seasonally adjusted (x 1,000).
What’s up with the x axis?

How the countries being plotted on the chart?

Appropriate use of visualization type?
Enough contextual information?

What’s up with the colour scale?

Appropriate colour choices?

Efficient way to display info?
Watch out for dual scales!

Exaggerating data?

Appropriate colour choice?
BREAK
DOWNLOAD

The prepared data sets from DROPBOX: http://bit.ly/1Gitzp4

and save to your desktop.
<table>
<thead>
<tr>
<th>Name</th>
<th># of Siblings</th>
<th>Eye Color</th>
<th>Name</th>
<th>Eye Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tata</td>
<td>2</td>
<td>Blue</td>
<td>Tata</td>
<td>Blue</td>
</tr>
<tr>
<td>Leo</td>
<td>3</td>
<td>Brown</td>
<td>Leo</td>
<td>Brown</td>
</tr>
<tr>
<td>Sam</td>
<td>1</td>
<td>Red</td>
<td>Sam</td>
<td>Red</td>
</tr>
<tr>
<td>Molly</td>
<td></td>
<td></td>
<td>Molly</td>
<td></td>
</tr>
<tr>
<td>Noa</td>
<td></td>
<td></td>
<td>Noa</td>
<td></td>
</tr>
</tbody>
</table>

**Left Join**

<table>
<thead>
<tr>
<th>Name</th>
<th># of Siblings</th>
<th>Eye Color</th>
<th>Name</th>
<th>Eye Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tata</td>
<td>2</td>
<td>Blue</td>
<td>Tata</td>
<td>Blue</td>
</tr>
<tr>
<td>Leo</td>
<td>3</td>
<td>Brown</td>
<td>Leo</td>
<td>Brown</td>
</tr>
<tr>
<td>Sam</td>
<td>1</td>
<td>Red</td>
<td>Sam</td>
<td>Red</td>
</tr>
<tr>
<td>Molly</td>
<td></td>
<td></td>
<td>Molly</td>
<td></td>
</tr>
<tr>
<td>Noa</td>
<td></td>
<td></td>
<td>Noa</td>
<td></td>
</tr>
</tbody>
</table>

**Right Join**

<table>
<thead>
<tr>
<th>Name</th>
<th># of Siblings</th>
<th>Eye Color</th>
<th>Name</th>
<th>Eye Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tata</td>
<td>2</td>
<td>Blue</td>
<td>Tata</td>
<td>Blue</td>
</tr>
<tr>
<td>Leo</td>
<td>3</td>
<td>Brown</td>
<td>Leo</td>
<td>Brown</td>
</tr>
<tr>
<td>Sam</td>
<td>1</td>
<td>Red</td>
<td>Sam</td>
<td>Red</td>
</tr>
<tr>
<td>Molly</td>
<td></td>
<td></td>
<td>Molly</td>
<td></td>
</tr>
<tr>
<td>Noa</td>
<td></td>
<td></td>
<td>Noa</td>
<td></td>
</tr>
</tbody>
</table>

**Inner Join**

<table>
<thead>
<tr>
<th>Name</th>
<th># of Siblings</th>
<th>Eye Color</th>
<th>Name</th>
<th>Eye Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tata</td>
<td>2</td>
<td>Blue</td>
<td>Tata</td>
<td>Blue</td>
</tr>
<tr>
<td>Leo</td>
<td>3</td>
<td>Brown</td>
<td>Leo</td>
<td>Brown</td>
</tr>
<tr>
<td>Sam</td>
<td>1</td>
<td>Red</td>
<td>Sam</td>
<td>Red</td>
</tr>
<tr>
<td>Molly</td>
<td></td>
<td></td>
<td>Molly</td>
<td></td>
</tr>
<tr>
<td>Noa</td>
<td></td>
<td></td>
<td>Noa</td>
<td></td>
</tr>
</tbody>
</table>

**Outer Join**

<table>
<thead>
<tr>
<th>Name</th>
<th># of Siblings</th>
<th>Eye Color</th>
<th>Name</th>
<th>Eye Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tata</td>
<td>2</td>
<td>Blue</td>
<td>Tata</td>
<td>Blue</td>
</tr>
<tr>
<td>Leo</td>
<td>3</td>
<td>Brown</td>
<td>Leo</td>
<td>Brown</td>
</tr>
<tr>
<td>Sam</td>
<td>1</td>
<td>Red</td>
<td>Sam</td>
<td>Red</td>
</tr>
<tr>
<td>Molly</td>
<td></td>
<td></td>
<td>Molly</td>
<td></td>
</tr>
<tr>
<td>Noa</td>
<td></td>
<td></td>
<td>Noa</td>
<td></td>
</tr>
</tbody>
</table>
STATISTICAL SOFTWARE PACKAGES

SPSS:

Viz types: Standard graphs, charts, plots
Programming?: Not required
Storage: Computer hard drive
Cost: Annual license

*note: Produces low quality images, not suitable for publication, consider rebuilding visualizations in Excel or using R or Stata
STATISTICAL SOFTWARE PACKAGES

R:

Viz types: Almost any two dimensional visualization types
Programming?: Requires knowledge of R language, steep learning curve, but robust sets of tutorials & learning material online
Storage: Computer hard drive
Cost: Free, open source

*note:
- Higher quality images
- Install ggplot2 & lattice packages for excellent visualization libraries
- New packages continuously being developed
COMMERCIAL VISUALIZATION SOFTWARE

Tableau:

Viz types: Many 2-d & 3-d visualization types, including dynamic & geospatial visualizations
Programming?: None, drag & drop functionality
Storage: Computer hard drive
Cost: Expensive $$

*note:
- High quality images
- Allows users to build custom interactive dashboard
- Suggests visualization types based on your data
- Free version with limited functionality, Tableau Public
EXCEL

Viz types: Most standard 2-d visualizations
Programming?: None, easy to use drag & drop functionality
Storage: Computer hard drive
Cost: Included in Microsoft Office

*note:
- Produces serviceable visualizations
- Don’t settle for the default charts!
- Download additional Excel templates for more visualization types
- See Jorge Comoes’ blog for tutorials: http://www.excelcharts.com/blog/
OPEN SOURCE VISUALIZATION TOOLS

Google Fusion Tables

Viz types: Most standard 2-d visualizations, network graphs, & geospatial visualizations with Google Maps’ data, static & dynamic visualizations

Programming?: None, easy to use drag & drop functionality

Storage: Cloud based, attached to Google account

Cost: Free, open source
OPEN SOURCE VISUALIZATION TOOLS

Raw Density

Viz types: Dynamic visualizations, visualizations that show relationships & hierarchies, tree maps, scatterplots, network visualizations

Programming?: None, easy to use drag & drop functionality

Storage: Cloud based

Cost: Free, open source
OPEN SOURCE VISUALIZATION TOOLS

d3.js

Viz types: Interactive, geospatial, & network visualizations
Programming?: Light javascript, many tutorials
Storage: Computer hard drive, interactive visualization online
Cost: Free, open source
OPEN SOURCE VISUALIZATION TOOLS

Gephi & Sci2

Viz types: Dynamic visualizations, visualizations that show relationships & hierarchies, scatterplots, networks, geospatial visualizations

Programming?: Light Python, easy to use drag & drop functionality

Storage: Cloud based

Cost: Free, open source
**IMAGE EDITORS & COLOUR TOOLS**

**Inkscape**
- Open source image editor
- Useful for cleaning up images, editing axes or adding legends

**ColorBrewer2**
- Choose appropriate colour schemes according to brewer palates
- Creates RGB, CMYK, & HEX codes for colours
- Performs various colour blindness tests
TABLEAU TUTORIAL
EXERCISE: CREATE A DATA VISUALIZATION ON YOUR OWN

1. Choose one of the provided data sets.
2. Select 2-3 variables and create a data visualization.
3. Be sure to select a visualization form appropriate to your data selection.
4. Select appropriate visualization elements for the data.
The Data Visualization Catalogue: Detailed descriptions of data visualization types, their uses, & limitations
http://www.datavizcatalogue.com/

Tools organized by visualization type, Duke University Libraries:
http://bit.ly/1EHjs5K

Curated list of visualization tools
http://selection.datavisualization.ch/

Books

Nature Methods Points of View Columns:
QUESTIONS?

Jesse Carliner
jesse.carliner@utoronto.ca

Nicholas Worby
nicholas.worby@utoronto.ca
THANK YOU!