Data Librarian in the middle, creating instructional content for the digital humanities

Matthew Gertler
Data, Statistics and GIS Librarian
University of Toronto Scarborough

The Context

- The University of Toronto Scarborough (UTSC) is a campus of the University of Toronto
- 13,430 students enrolled in 16 different academic departments or centres

Departments

- Anthropology
- Arts, Culture and Media
- Biological Sciences
- Computer & Mathematical Sciences
- Critical Development Studies
- English
- French and Linguistics
- **Historical and Cultural Studies**
- Human Geography
- Management
- Philosophy
- Physical & Environmental Sciences
- Political Science
- Psychology
- Sociology

Presentation Topic

My journey as a data librarian began from a Social Science point of view. This presentation will cover:

- My experience making digital humanities content
- What I learned throughout this experience
- An overview of presentation concepts and themes

DH Tips for Data Librarians

Leverage Similarities

- Generally data structures are universal
- Many tools and visualizations are used across disciplines

1	FACILITY NAME	Full Address	Street #	Street Name	Suite	City
3	Albion Pool & Health Club	1485 Albion Road, Toronto, ON, M9V 1B2	1485	Albion Road		Toronto
4	Albion Branch (TPL)	1515 Albion Road, Toronto, ON, M9V 1B2	1515	Albion Road		Toronto
5	Theatre Francais de Toronto - Centre for Creation	21 College Street, Office 610, Toronto, ON, M5G 2B3	21	College Street	Office 610	Toronto
6	Humber Arboretum Gardens	203 Humber College Boulevard, Toronto, ON, M9W 5L7	203	Humber College Boulevard		Toronto
7	Thistletown Baptist Church	2534 Kipling Avenue, Toronto, ON, M9V 3A9	2534	Kipling Avenue		Toronto
8	North Albion Collegiate Institute	2580 Kipling Avenue, Toronto, ON, M9V 3B2	2580	Kipling Avenue		Toronto
9	Smithfield Community Services	175 Mount Olive Drive, Toronto, ON, M9V 2E3	175	Mount Olive Drive		Toronto
10	Rexdale Community Hub	21 Panorama Court, Toronto, ON, M9V 4E3	21	Panorama Court		Toronto
11	Elmbank Community Centre	10 Rampart Road, Toronto, ON, M9V 4L9	10	Rampart Road		Toronto
12	Franklin Carmichael Arts Centre	34 Riverdale Drive, Toronto, ON, M9V 1C8	34	Riverdale Drive		Toronto
13	North Kipling Community Centre	2 Rowntree Road, Toronto, ON, M9V 5G6	2	Rowntree Road		Toronto
14	Crowne Plaza Hotel	33 Carlson Court, Toronto, ON, M9W 6H5	33	Carlson Court		Toronto
15	Toronto Congress Centre	650 Dixon Road, Toronto, ON, M9W 1J1	650	Dixon Road		Toronto
16	Doubletree International Plaza Hotel	650 Dixon Road, Toronto, ON, M9W 1J3	650	Dixon Road		Toronto
17	Sheraton Toronto Airport Hotel	801 Dixon Road, Toronto, ON, M9W 1J5	801	Dixon Road		Toronto
18	Humberwood Library & Community Centre	850 Humberwood Boulevard, Toronto, ON, M9W 7A6	850	Humberwood Boulevard		Toronto
19	Msg. Percy Johnson Catholic Secondary School	2170 Kipling Avenue, Toronto, ON, M9W 4K9	2170	Kipling Avenue		Toronto

source: City of Toronto Open Data Catalogue

Understanding Differences

- Consult with faculty to understand their needs
- Review literature from the discipline
- Look at exemplar projects from the discipline

Example Project: Mapping the Republic of Letters

Understand Student Skill Level

Students from different disciplines will have very different exposures to:

- Digital tools
- Analyzing data or visualizing data
- Research Data Management

Selecting Appropriate Tools

ArcGIS

Google Fusion Tables

Digital Humanities Presentations

Project Goals:

- Create presentations that cover topics in a theoretical and practical way
- Design the presentations so that other librarians or faculty can potentially give them
- Give the students the tools they need to complete digital assignments

Five DH Presentations

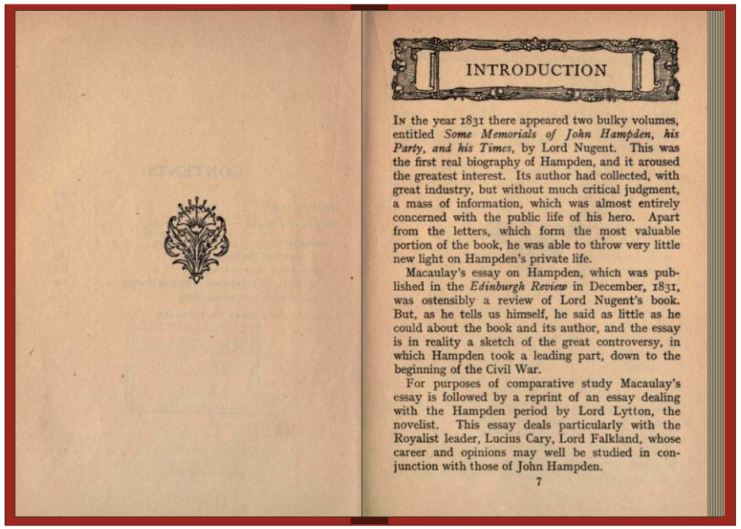
Five Introductory Presentations at the request of History Faculty:

- Structured Data
- Image Annotation
- Network Analysis
- GIS and Webmapping
- Creating Timelines

Structured Data

What is Humanities Data?

- Definition of data
- Sources of Humanities Data



INTRODUCTION

In the year 1831 there appeared two bulky volumes, entitled *Some Memorials of John Hampden, his Party, and his Times*, by Lord Nugent. This was the first real biography of Hampden, and it aroused the greatest interest. Its author had collected, with great industry, but without much critical judgment, a mass of information, which was almost entirely concerned with the public life of his hero. Apart from the letters, which form the most valuable portion of the book, he was able to throw very little new light on Hampden's private life.

Macaulay's essay on Hampden, which was published in the *Edinburgh Review* in December, 1831, was ostensibly a review of Lord Nugent's book. But, as he tells us himself, he said as little as he could about the book and its author, and the essay is in reality a sketch of the great controversy, in which Hampden took a leading part, down to the beginning of the Civil War.

For purposes of comparative study Macaulay's essay is followed by a reprint of an essay dealing with the Hampden period by Lord Lytton, the novelist. This essay deals particularly with the Royalist leader, Lucius Cary, Lord Falkland, whose career and opinions may well be studied in conjunction with those of John Hampden.

7

UTSC Library

Why Used “Structured” Data

Introducing students to the concept of machine readable data for the first time

Tabular

	A	B	C	D	E	F
1	Timestamp	1. What age range do you fall in?	Sex	How would you describe your level of physical health?	What level of education have you reached?	What is your household income?
2	13-04-10 12:10	35 to 49	Male	Good	Post graduate	$100K+
3	13-04-10 12:30	18 to 34	Male	Excellent	University degree	$40K to $59K
4	13-04-10 12:33	50 to 64	Male	Good	University degree	$40K to $59K
5	13-04-10 12:52	50 to 64	Male	Good	4 years university, no degree	$80K to $99K
6	13-04-10 13:24	18 to 34	Male	Very good	College or trade school diploma	$40K to $59K
7	13-04-10 13:26	35 to 49	Male	Very good	College or trade school diploma	$40K to $59K
8	13-04-10 13:27	18 to 34	Male	Good	College or trade school diploma	$100K+
9	13-04-10 13:28	35 to 49	Male	Very good	Post graduate	$100K+
10	13-04-10 13:31	35 to 49	Female	Fairly good	Post graduate	$80K to $99K
11	13-04-10 13:33	18 to 34	Male	Good	College or trade school diploma	$60K to $79K
12	13-04-10 13:43	18 to 34	Male	Very good	College or trade school diploma	Under $20K
13	13-04-10 13:44	35 to 49	Male	Very good	High school diploma	$80K to $99K
14	13-04-10 13:45	18 to 34	Male	Very good	University degree	$80K to $99K
15	13-04-10 13:45	18 to 34	Male	Excellent	University degree	$20K to $39K
16	13-04-10 13:45	18 to 34	Female	Very good	University degree	$40K to $59K
17	13-04-10 13:55	50 to 64	Female	Fairly good	High school diploma	$80K to $99K
18	13-04-10 12:54	35 to 49	Male	Good	College or trade school diploma	$60K to $79K
19	13-04-10 13:03	18 to 34	Female	Good	High school diploma	$20K to $39K
20	13-04-10 13:04	50 to 64	Female	Good	University degree	$80K to $99K
21	13-04-10 13:06	18 to 34	Female	Poor	University degree	$60K to $79K
22	13-04-10 13:11	35 to 49	Female	Good	College or trade school diploma	$40K to $59K
23	13-04-10 13:17	35 to 49	Male	Good	Post graduate	$100K+
24	13-04-10 13:55	35 to 49	Male	Very good	Post graduate	$100K+
25	13-04-10 13:57	50 to 64	Male	Very good	Post graduate	
26	13-04-10 14:00	18 to 34	Male	Fairly good	College or trade school diploma	$20K to $39K
27	13-04-10 14:02	35 to 49	Male	Very good	University degree	$80K to $99K
28	13-04-10 14:05	35 to 49	Female	Fairly good	College or trade school diploma	$40K to $59K
29	13-04-10 14:13	35 to 49	Male	Good	University degree	$100K+
30	13-04-10 14:15	18 to 34	Male	Very good	College or trade school diploma	$40K to $59K

Source: E-Bike Survey Response Results, City of Toronto Data Catalog

Tagged

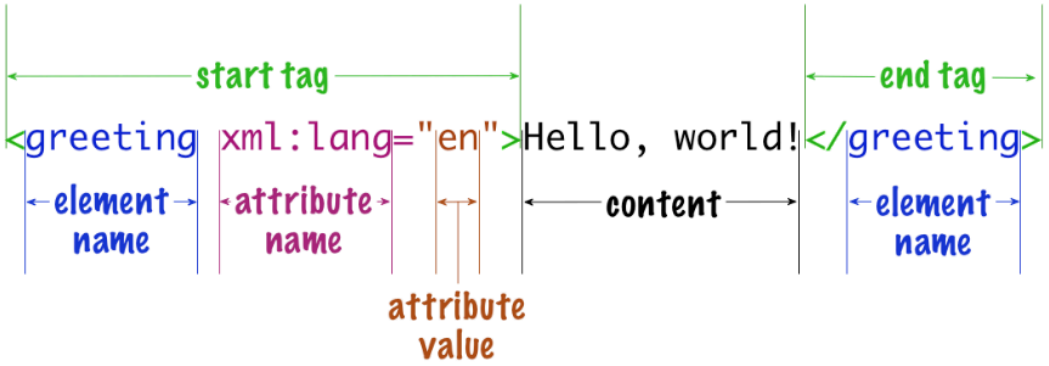


Image credit: Syd Bowman, Julia Flanders. Brown University.

Data Organization and Management

- The layout of a table
- How to properly document data
- Good data management practices

	A	B	C	D	E	F
1	Timestamp	1. What age range do you fall in?	Sex	How would you describe your level of physical health?	What level of education have you reached?	What is your household income?
2	13-04-10 12:10	35 to 49	Male	Good	Post graduate	$100K+
3	13-04-10 12:30	18 to 34	Male	Excellent	University degree	$40K to $59K
4	13-04-10 12:33	50 to 64	Male	Good	University degree	$40K to $59K
5	13-04-10 12:52	50 to 64	Male	Good	4 years university, no degree	$80K to $99K
6	13-04-10 13:24	18 to 34	Male	Very good	College or trade school diploma	$40K to $59K
7	13-04-10 13:26	35 to 49	Male	Very good	College or trade school diploma	$40K to $59K
8	13-04-10 13:27	18 to 34	Male	Good	College or trade school diploma	$100K+
9	13-04-10 13:28	35 to 49	Male	Very good	Post graduate	$100K+
10	13-04-10 13:31	35 to 49	Female	Fairly good	Post graduate	$80K to $99K
11	13-04-10 13:33	18 to 34	Male	Good	College or trade school diploma	$60K to $79K
12	13-04-10 13:43	18 to 34	Male	Very good	College or trade school diploma	Under $20K
13	13-04-10 13:44	35 to 49	Male	Very good	High school diploma	$80K to $99K
14	13-04-10 13:45	18 to 34	Male	Very good	University degree	$80K to $99K
15	13-04-10 13:45	18 to 34	Male	Excellent	University degree	$20K to $39K
16	13-04-10 13:45	18 to 34	Female	Very good	University degree	$40K to $59K
17	13-04-10 13:55	50 to 64	Female	Fairly good	High school diploma	$80K to $99K
18	13-04-10 12:54	35 to 49	Male	Good	College or trade school diploma	$60K to $79K
19	13-04-10 13:03	18 to 34	Female	Good	High school diploma	$20K to $39K
20	13-04-10 13:04	50 to 64	Female	Good	University degree	$80K to $99K
21	13-04-10 13:06	18 to 34	Female	Poor	University degree	$60K to $79K
22	13-04-10 13:11	35 to 49	Female	Good	College or trade school diploma	$40K to $59K
23	13-04-10 13:17	35 to 49	Male	Good	Post graduate	$100K+
24	13-04-10 13:55	35 to 49	Male	Very good	Post graduate	$100K+
25	13-04-10 13:57	50 to 64	Male	Very good	Post graduate	
26	13-04-10 14:00	18 to 34	Male	Fairly good	College or trade school diploma	$20K to $39K
27	13-04-10 14:02	35 to 49	Male	Very good	University degree	$80K to $99K
28	13-04-10 14:05	35 to 49	Female	Fairly good	College or trade school diploma	$40K to $59K
29	13-04-10 14:13	35 to 49	Male	Good	University degree	$100K+
30	13-04-10 14:15	18 to 34	Male	Very good	College or trade school diploma	$40K to $59K

Row
The unit of observation

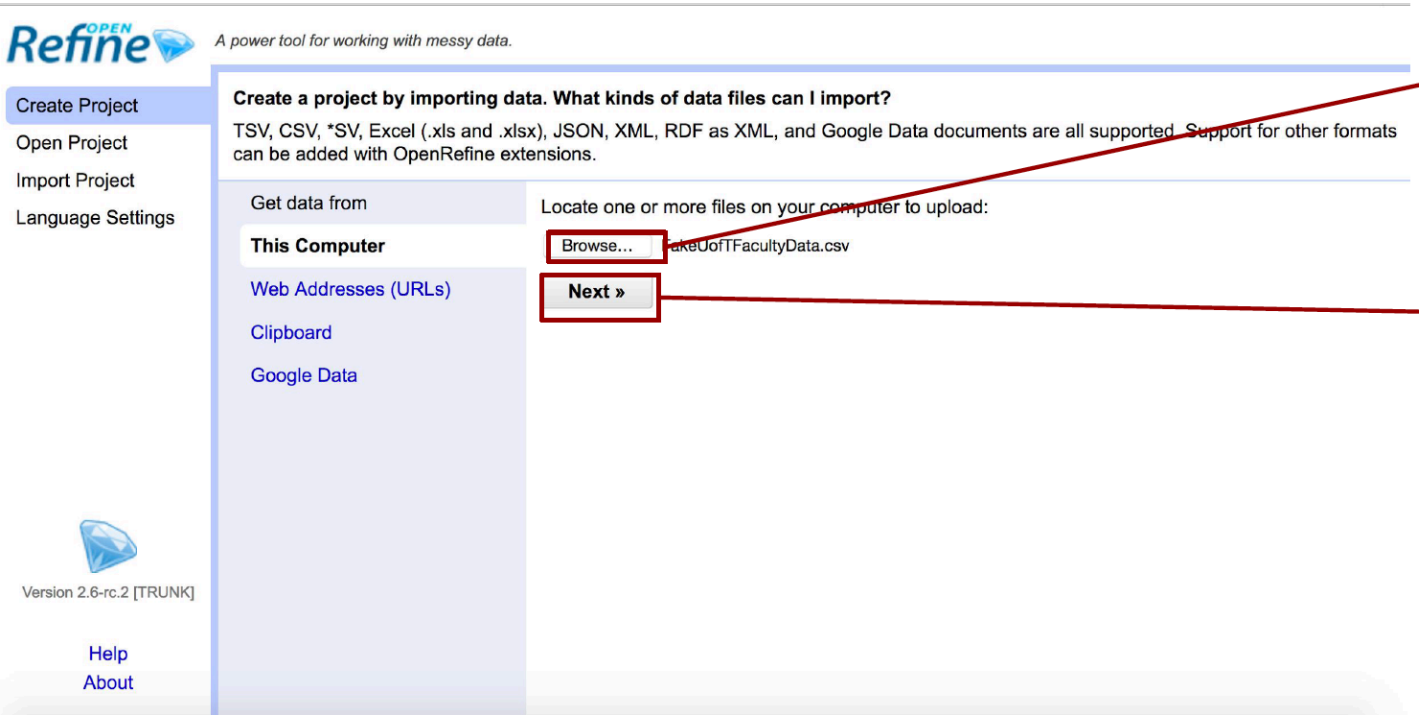
Column
One variable or question

Cell
One Observation

Source: E-Bike Survey Response Results, City of Toronto Data Catalog

Data Cleaning Workshop

Tool used: Excel and Open Refine



Click the browse button find your FakeUofTFacultyData in the chosen directory and select it.

Click Next

UNIVERSITY OF TORONTO SCARBOROUGH

Image Annotation Presentation

Why Annotate in History Research

Annotations of images

Research looking at realism vs. mythmaking around Benjamin West's painting *The Death of General Wolfe*

Some argue the depiction of indigenous warrior represents the perceived relationship of Britain with the First Nations

Painting was controversial at the time of creation as artist depicts subjects in contemporary clothing as opposed to Roman dress as was traditional in the style of neoclassicism.

Humanities Annotation Projects and Models

The OAC Model

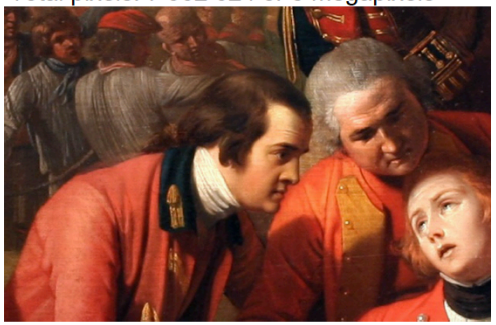
Article: Rereading the Indian in Benjamin West's" Death of General Wolfe"

Working with Images

Image Resolution Implications

The Death of General Wolfe by Benjamin West

Starting Image Resolution: 3456 x 2304 px
Total pixels: 7 962 624 or 8 megapixels

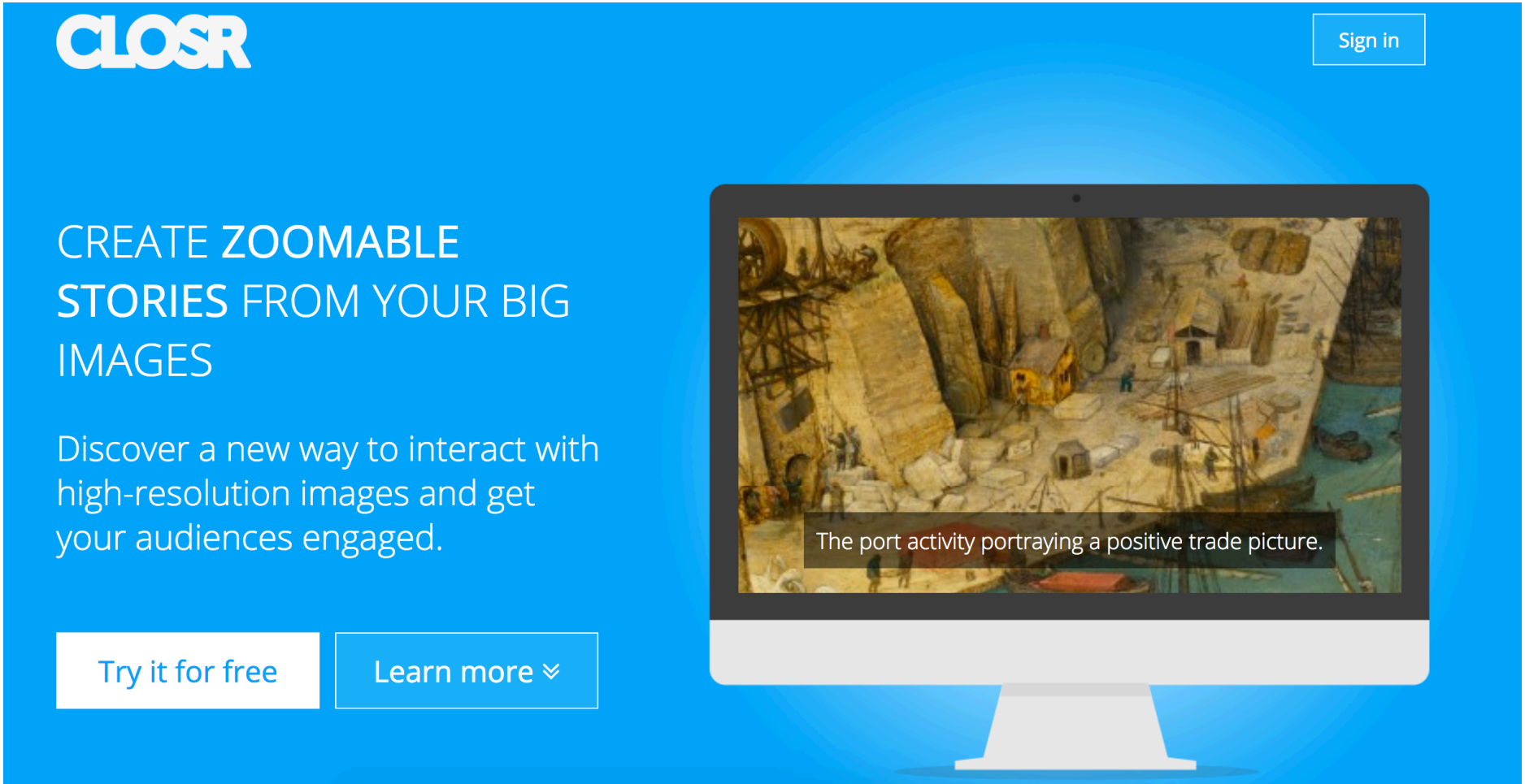


Starting Image Resolution: 640 x 427 px
Total pixels: 273 280

UNIVERSITY OF TORONTO SCARBOROUGH

Workshop Component

- Annotate images with closr.it



Network Analysis Presentation

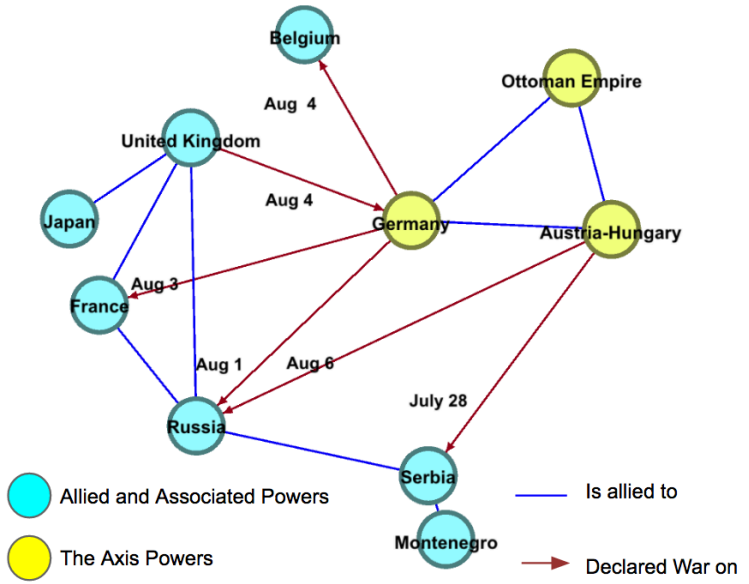
About Network Analysis

- Why Conduct Network Analysis
- How does Network Analysis Work

More than one type of edge

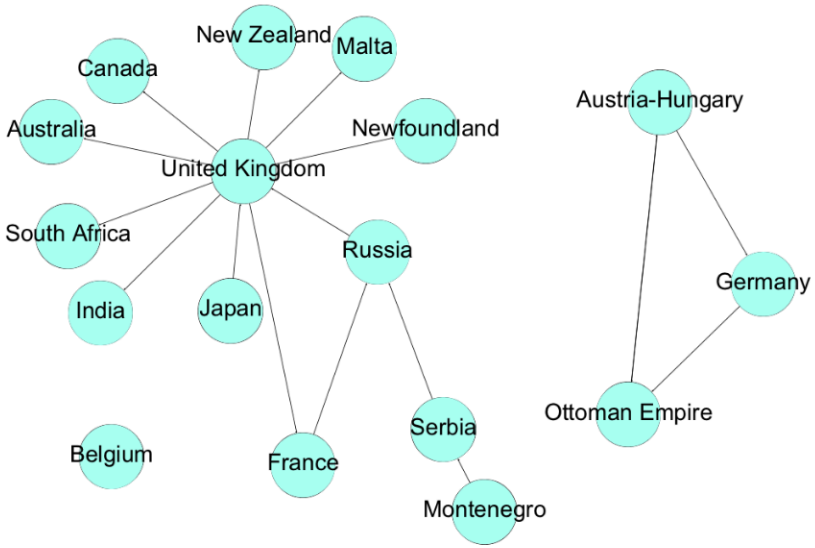
Edges can be directed, undirected and mixed. The Edge type gives directionality to a relationship.

National Declarations of War in the first 10 days of WWI



Examples from Scholarship

Researchers decisions affect the story



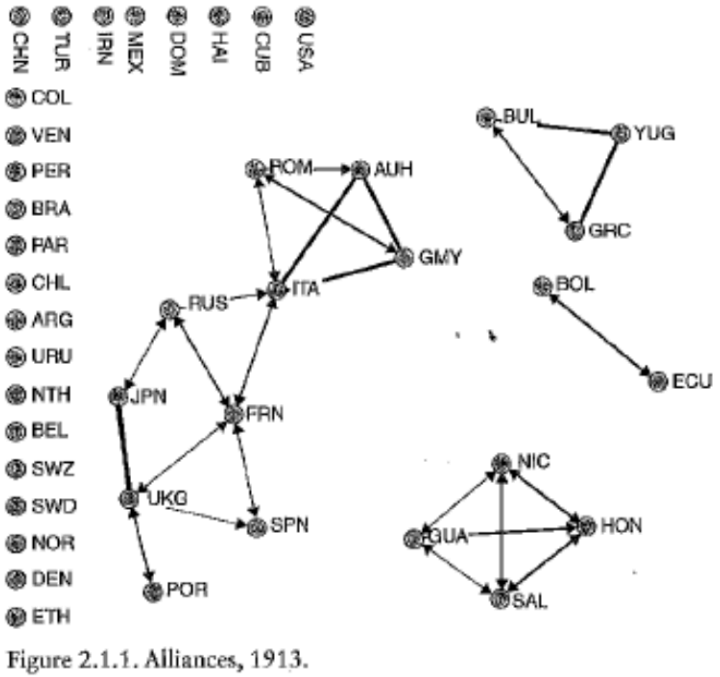
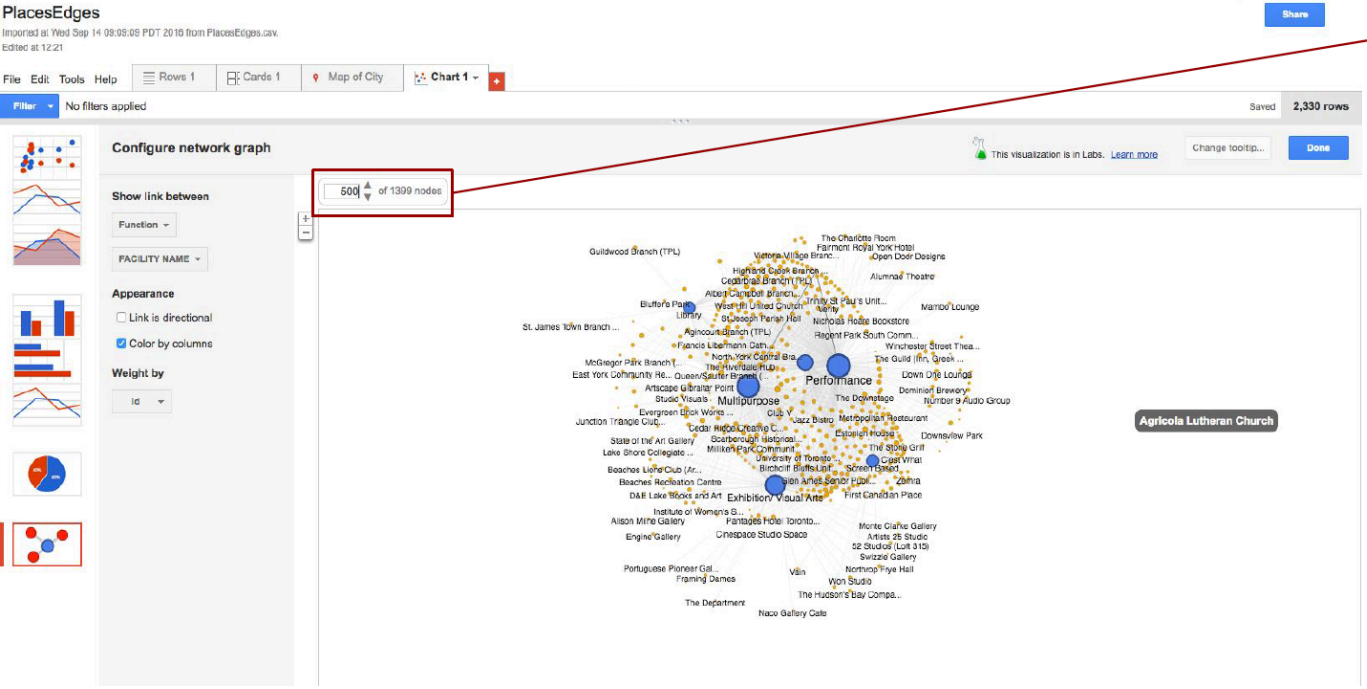


Figure 2.1.1. Alliances, 1913.

Maoz, Z. (2011). Networks of nations: *The evolution, structure, and impact of International Networks, 1816-2001*.

Network Analysis Workshop

- Creating Networks with Google Fusion Tables
- Alternative tools: Gephi and Sci^2



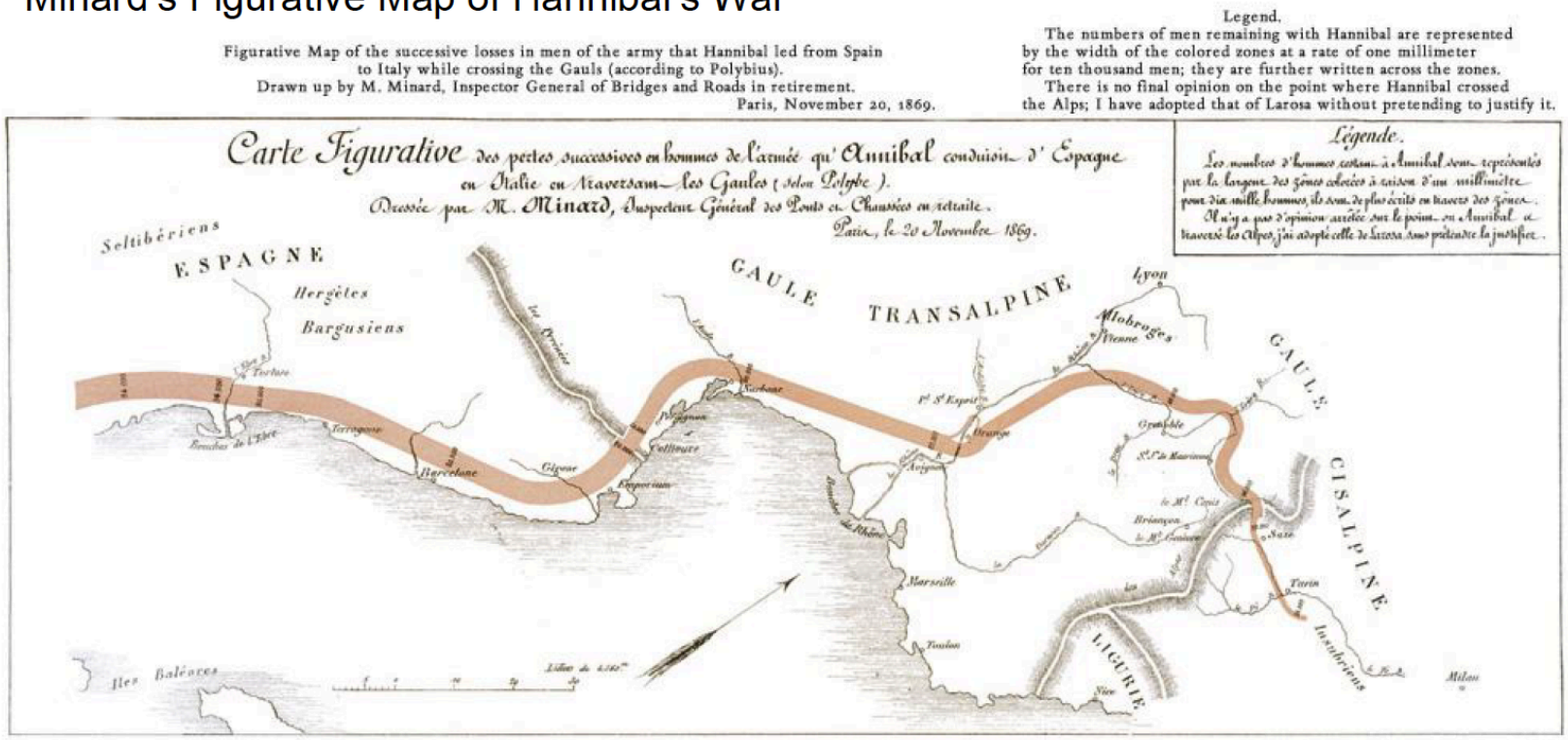
You can change the number of nodes displayed. If set to 0 you will see nothing.

GIS and Webmapping

Why Map in the Humanities

An aspect of your research is explained or affected by geography

Minard's Figurative Map of Hannibal's War



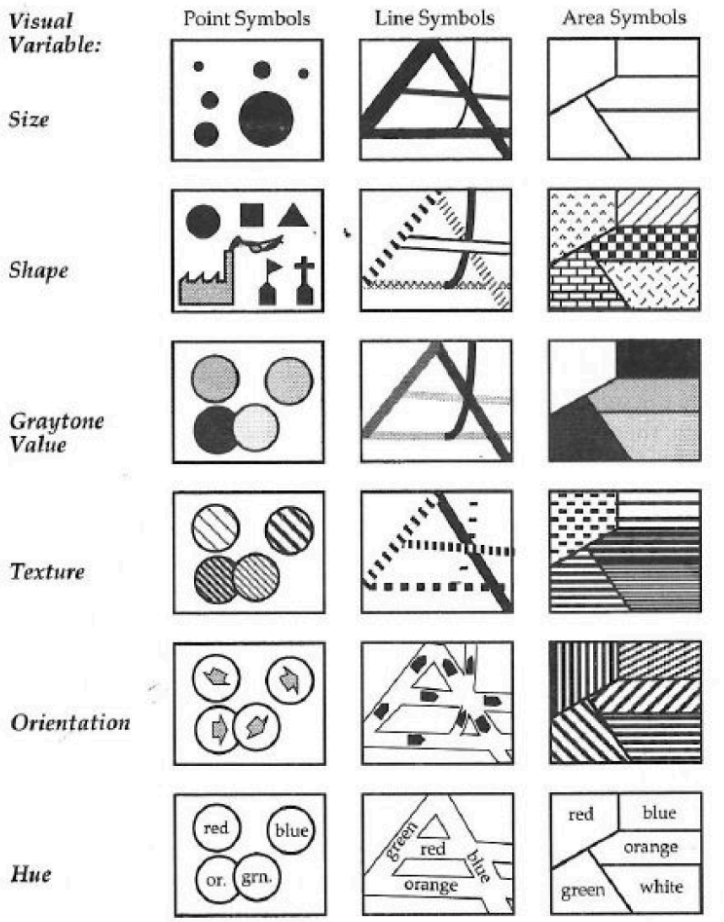
Source: Tufte.com

About GIS

Symbolization

Graphic symbols communicate location information such as places or features on a map. Examples may include:

- points for locations of places
- lines for roads,
- area symbols for political boundaries.

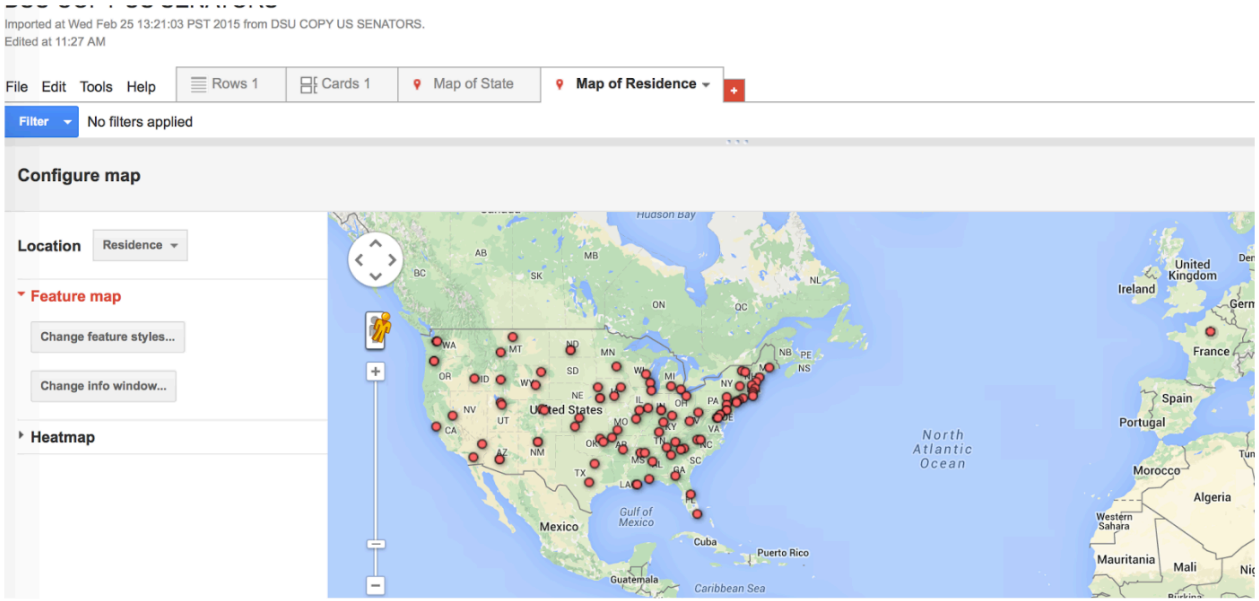


Monmonier, M. S. (1996). *How to lie with maps.*

Webmapping Workshop

- Webmapping with Fusion Tables
- Alternative tools: ArcOnline, Carto, Mapbox

Geocoding Place Names



Acknowledgements

The five DH presentations were made in collaboration with:

- Whitney Kimble, UTSC Library
- Sarah Forbes, UTSC Library
- Kirsta Stapelfeltd, UTSC Library
- Natalie Rothman, Department of Historical and Cultural Studies

Questions?