Exercises for CHASE's Arts and Humanities in the Digital Age (ADHA) 2017 Introduction to Information Visualisation

Exercise 1: comparing n-gram tools

Time: 5 minutes plus discussion
Goal: explore the impact of source data and algorithms on input text

1. Think of two words or phrases you'd like to compare over time (e.g. Burma, Burmah).
2. Open two browser windows
3. In one, go to http://books.google.com/ngrams
4. In the other, go to http://benschmidt.org/OL/
5. Enter your words or phrases in each and compare the results
6. Discuss with your neighbour: what differences did you find, and why?

Google Ngram tips: http://books.google.com/ngrams/info. You can change the corpus and explore the difference it makes.

Bookworm tips: click the 'cog' icon in the top right-hand corner to change the time period. Click the underlined words next to the search term to change which corpus is searched (e.g. subject, language, country, gender of author). You can also compare the same word or phrase in different corpus.

Tip: if you're more interested in newspapers, try the Library of Congress's Chronicling America collection at http://arxiv.culturomics.org/ChronAm/ or Australia and New Zealand newspapers at http://dhistory.org/querypic/create/

Exercise 2: Try entity extraction

Time: c. 5 minutes. Goal: explore methods for extracting information from text or an image

You can choose between text and images.

For text:

1. Find a short paragraph of text (e.g. from a news site or digitised text) to copy and paste into the box
2. In your browser, go to http://nlp.stanford.edu:8080/ner/
3. Paste your text into the window and run the tool

For images:

1. Find an image (e.g. from a news site or digitised text) you can download and drag into the window
2. In your browser, go to https://www.clarifai.com/demo
3. Load an image and run the tool

For both:

How many of possible entities (concepts, people, places, events, references to time or dates, etc) did it pick up? Is any of the other information presented useful? Did it label anything incorrectly? What if you change classifiers or other parameters? Does it do better with different source material?

Exercise 3: exploring scholarly data visualisations

Time: 10-15 minutes plus discussion
Goal: explore visual methods for presenting narratives or dealing with dynamic data while practicing critical analysis.

Instructions

1. In your browser, go to one of the sites below
2. Take a few minutes to explore the visualisation
3. Discuss with your neighbour. Questions might include:
    - What did you learn from the visualisation?
    - What have you learned from visualisation that you might not have learned from looking at the data or reading a description of it?
    - Can you easily see where to start and how to use it?
    - Does it work better at one scale over another? Do you find it more effective at aggregate or detail level?
    - Does it present an argument (viewpoints or statements about the data) or provide a space for you to explore and develop one? What arguments does the site present? How are they expressed - framing, design, data choices?
    - Can you tell who created it? Who the intended users are?
    - Where did the data come from? Is any data left out?
    - What building blocks does the visualisation use - charts, timelines, etc?
    - Does the design suit the source material and intended users? Why, why not?
    - Are there any novel elements? How effective are they?
    - Do you trust the visualisation? Why, why not?

Be prepared to report back to the group

For example, be prepared to summarise the site's purpose, visualisation formats and data types, or to share unresolved questions or the most interesting parts of your discussion

America's Public Bible

http://americaspublicbible.org/

The State of the Union in Context

http://benschmidt.org/poli/2015-SOTU

On Broadway

http://on-broadway.nyc/

Six Degrees of Francis Bacon

http://www.sixdegreesoffrancisbacon.com/

Bristol Know Your Place
http://maps.bristol.gov.uk/knowyourplace/

HistoryPin
https://www.historypin.org/

University of Richmond, 'Visualizing Emancipation'
http://www.americanpast.org/emancipation/

City Readers
http://cityreaders.nysoclib.org/About/visualizations

'Mapping the Republic of Letters'
http://www.stanford.edu/group/toolingup/rplviz/rplviz.swf - needs Flash

Locating London's Past
http://www.locatinglondon.org/

Digital Harlem :: Everyday Life 1915-1930
http://digitalharlem.org/

Digital Public Library of America's timeline, map
http://dp.la/
Further information: http://dp.la/info/ and http://dp.la/info/news/blog/

Orbis
http://orbis.stanford.edu/

Lost Change
http://tracemedia.co.uk/lostchange/

Viral Texts
http://viraltexts.org/ (you can go straight to http://networks.viraltexts.org/1836to1899/index.html or http://viraltexts.northeastern.edu/)

Viewshare
Viewshare is a free web application supported by the Library of Congress. You can quickly create maps, timelines, charts and embed them on webpages, so it's a good way to start thinking about what's needed for your own visualisations.

As with most visualisation tools, once you've set up the structure of your spreadsheet (or database, JSON file, etc), you can tweak the values in each row or column to suit the visualisation you want to create.

Using data in Viewshare
The tutorial provides a link to a sample spreadsheet, which you can download then re-upload. You don't need to open the spreadsheet to do so.

If you're confident, you can try the instructions with your own dataset. See https://viewshare.uservoice.com/knowledgebase/articles/243875-importing-data-into-viewshare for information on importing data.

If you want to try a different dataset but don't yet have your own, you can download Tate artist data from https://drive.google.com/open?id=16AbHNcVl4YGQPTnvVqzlAInStPMs81L1gY2PpQmc (aggregated values for simple charts) and https://docs.google.com/spreadsheets/d/14f8bsvlUTznDs9NB5aC2ktaTK21zLWhwCP2VpNbcmRM/edit#gid=1790891819 (individual artist records with city names, birth and death dates)

Viewshare Exercise 1: Ten minute tutorial - getting started with Viewshare

This will teach you the basics of using Viewshare https://viewshare.uservoice.com/knowledgebase/articles/77925-10-minute-viewshare-tutorial

Viewshare Exercise 2: Create new views and widgets

Follow the instructions on the link below to create views (lists, maps, pie charts, bar charts, scatter plots, tables, timelines or galleries) or widgets (search boxes, lists, tag clouds, sliders, ranges, logos or text). The Viewshare interface should provide tips on what kinds of information are needed, and you may need to change the type of a field to enable certain views and widgets.

https://viewshare.uservoice.com/knowledgebase/articles/244379--generating-views-of-collection-data

References and links

References:
http://www.miaridge.com/resources-for-data-visualisation-for-analysis-in-scholarly-research

Explore and analyse more visualisations: http://scholarlyvision.tumblr.com/

Getting from idea to visualisation - work out what data you need and the best way to prepare and present it.http://www.dear-data.com has some lovely examples of creative sketches.

Finding tools

You can look for software in the DiRT Directory, a registry of digital research tools for scholarly use: http://dirtdirectory.org

These applications can be used with your own or public data. In order of simplicity:

- Microsoft Excel can produce many types of charts (e.g. pivot tables). You'll need something like Excel for data prep so you might as well learn all it can do.
- Viewshare http://viewshare.org - simple interactive maps, timelines, facets, tag clouds
- Google Fusion Tables http://tables.googlelabs.com
- Nodegoat can show change over time http://nodegoat.net/
- Tableau Public, download from http://tableausoftware.com

Mia Ridge for CHASE, February 2017