Intro to Mapping

Alex Leslie

10/2/2018

This workshop

Welcome! If you'd like to follow along on your own computer, you can download these slides and the sample data used later in this workshop from github.com/azleslie/MappingIntro.

Preparations

Should We Map?

The first question we ask before making a visualization should always be: will this visualization add something that description alone can't say (or can't say efficiently)?

The first question we ask before making a map should always be: is the spatial distribution of this information vital to understanding it?

Narrative vs. Thematic

There are two main types of map:

- Narrative maps, which visualize a story by the way it unfolds in space.
- Thematic maps, which plot one or several variables onto a representation of geographic space.

While we're most accustomed to encountering thematic maps, narrative maps can also be useful for a number of purposes and situations.

Considerations

- Is this a map of a narrative, or of data?
- How many observations does your data include?
  - Just a handful? Ten to twenty? Over a hundred?
- What will the map be used for?
  - A presentation? Teaching? An online platform or blog? A publication?
- Are there other meaningful relationships between observations in your data?
  - Is there a temporal dimension?
  - Are the observations also agents in a network?

Narrative Mapping: StoryMapJS

StoryMapJS

StoryMapJS is an online tool built by the Knightlab at Northwestern University with journalism in mind, but it's gained a following among humanists and educators.

I encourage you to follow along with the tutorial that follows either by starting your own test map on the StoryMapJS website or by checking out my own humble example..

StoryMapJS

Why You Might Use It	Why You Might Not
Great for telling a narrative, especially about an individual person / thing moving through space	Not amenable to quantitative data, comparison, or representing multiple points at once
Easy to learn	Relatively limited customization
Well adapted for presentations	Not appropriate for publication
Works well as a classroom assignment	

StoryMapJS

- Getting started
  - Google account required
  - You can get started right away on the StoryMapJS website by simply clicking "Make a StoryMap"
  - StoryMap is basically a slideshow/powerpoint. The first screen is a title slide, so this is where
- Making your map
  - After this, click "Add Slide" on the left-side navigation bar to add a location slide.
  - Either type in a location in the search bar at the bottom of the map or drag the icon on the map to the desired location.
  - Enter a title and narrative description in the lower right pane, and add any media you'd like in the lower left pane.
  - StoryMapJS automatically adjusts the zoom for each location in the journey based on how far apart they are and how specific your location entry is.

StoryMapJS: Options

- Additional options
  - General options - font, map style, language, size - can be edited by clicking “Options” in the upper left.
  - For any given location slide, you can change the map icon with “Marker Options” and change the background to another color or an image with “Background Options,” both in the lower right corner.

StoryMapJS: Sharing

- Clicking "My Maps" in the upper left will allow you to view your different creations.
- StoryMapsJS automatically saves when you move to edit a new slide.
- Click "Share" in the upper right corner to generate a shareable link for your narrative map. (You can also select a thumbnail and blurb here too.)
- Don't forget to click "Publish Changes" in the upper left if you make changes after sharing!

StoryMapJS: Teaching

Ease of use and ease of access make StoryMapJS amenable to classroom assignments, either as research or review. For example:

- Students could map a character's journey in a novel, such as Huck's odyssey down the Mississippi in *The Adventures of Huckleberry Finn*.
- Students could trace the travels of a historical actor, such as a regiment in the Napoleonic Wars.
- Students could track a socio-historical process, such as a spice's typical route on the Silk Road.

Thematic Mapping

Data Structure

The first step to making a thematic map is structuring your data appropriately.

Most software will require tidy data structured as a spreadsheet (like Microsoft Excel or Google Sheets), comma-separated values (a .csv file), or some other tabular format. This means that:

1. Each observation - each point on the map - should have its own row.
2. Each variable - each category of information you have for each point, geographic or otherwise - should have its own column.
3. Each value should have its own cell.

For Example

```
subs <- read.csv("subscribers.csv")
subs[240:245,] %>%
  select(town, state, country)
```

```
##             town state       country
## 240 Locust Grove    VA United States
## 241    Watertown    NY United States
## 242    Watertown    NY United States
## 243       Marion    NC United States
## 244       Auburn    AL United States
## 245        Clark    NJ United States
```

Take Note!

It's also a good idea to keep some form of a notes variable field. We need to impose structure on our data in order to produce analysis, but it's always important to make record of what changes were necessary to get it that way. Since I copied my data from a subscription book from the 1850s, I made a note when I wasn't sure about the handwriting or when I made any changes in order to make an observation mappable.

```
subs[240:245,] %>%
  select(notes)
```

```
##                                                                    notes
## 240 Technically Wilderness VA, but this is the nearest postal town.
## 241
## 242
## 243
## 244
## 245                          Picton is evidently subsumed into Clark
```

Messy Data?

- If you're starting a new project, it's best to set up your process so that your data collection/transcription is tidy from the start.
- When in doubt, always err on the side of recording more and using more variables! It's always easier to combine variables later than to separate them.
- If your data isn't up to this level of tidiness, there are a lot of possible strategies to help you get it there without spending tons of time changing everything manually. To learn more, attend one of the "Data 101" workshops this semester!

Geocoding

Geo-What?

Most mapping programs will require data that is already "geocoded": that is, that already includes the latitude and longitude coordinates for each point you wish to map. It should look something like this:

```
subs[1:6,] %>%
  select(town, state, country, latitude, longitude)
```

```
##              town state       country latitude longitude
## 1 Mackinac Island    MI United States 45.84918 -84.61893
## 2 Mackinac Island    MI United States 45.84918 -84.61893
## 3         Madison    GA United States 33.59568 -83.46794
## 4       Lewisburg    PA United States 40.96453 -76.88441
## 5          Monroe    LA United States 32.50931 -92.11930
## 6     River David                Canada 41.46275 -81.85089
```

Of course, humanists almost never encounter data in this form; more often, we have the names of towns and states that we need to transform into it.

Geocoder Options

There are a number available options for generating geographic coordinates; many are pay-to-use, but a few are free. Here are two I like:

- The Google Docs Add-On Geocode by Awesome Table is convenient because it can add values directly to your existing sheet. Its limitation, however, is that the Google API it relies on limits the number of calls you can make to one thousand per day.
- The geocoder hosted by UCLA Geography is conveniently fast. It works well even when you only have information for town and state, but you'll need to make sure your state values are stored in the two-character postal format (NE instead of Nebraska): otherwise it can often return incorrect coordinates. You'll need to copy-paste the results and save them in a .csv format.

Geocoding Prep

If you're working in a spreadsheet editor like Google Sheets or Microsoft Excel, you can do this by using the `=textjoin` command and selecting all the observations and variables you wish to merge into one. (For example, `=textjoin(", ", 1, K1, J1)`).

The same effect can be achieved easily in R - especially if you've attended one of our "Data 101" workshops! All you need to do is `paste` together the desired variables `separate` them with just a comma, and `mutate` the existing data frame to add the resulting new variable, "full_loc".

```
subs[1:6,] %>%
  select(town, state, country) %>%
  mutate(full_loc = paste(town, state, country, sep=","))
```

```
##               town state       country                          full_loc
## 1 Mackinac Island    MI United States Mackinac Island,MI,United States
## 2 Mackinac Island    MI United States Mackinac Island,MI,United States
## 3          Madison    GA United States          Madison,GA,United States
## 4        Lewisburg    PA United States        Lewisburg,PA,United States
## 5           Monroe    LA United States           Monroe,LA,United States
## 6      River David              Canada               River David,,Canada
```

Geocoding Prep

If you need to turn a state name into a state abbreviation, there are several options available. In R, we can take advantage of two built-in vectors: "state.abb" for abbreviations and "state.name" for names. Since the states are listed in the same order in both of these vectors, all we have to do is find the position of each of our state names in the "state.name" vector and then find the value in the same position in "state.abb" by indexing.

```
some_states <- c("Wisconsin", "Minnesota", "Wisconsin", "New Jersey")
state.abb[match(some_states, state.name)]
```

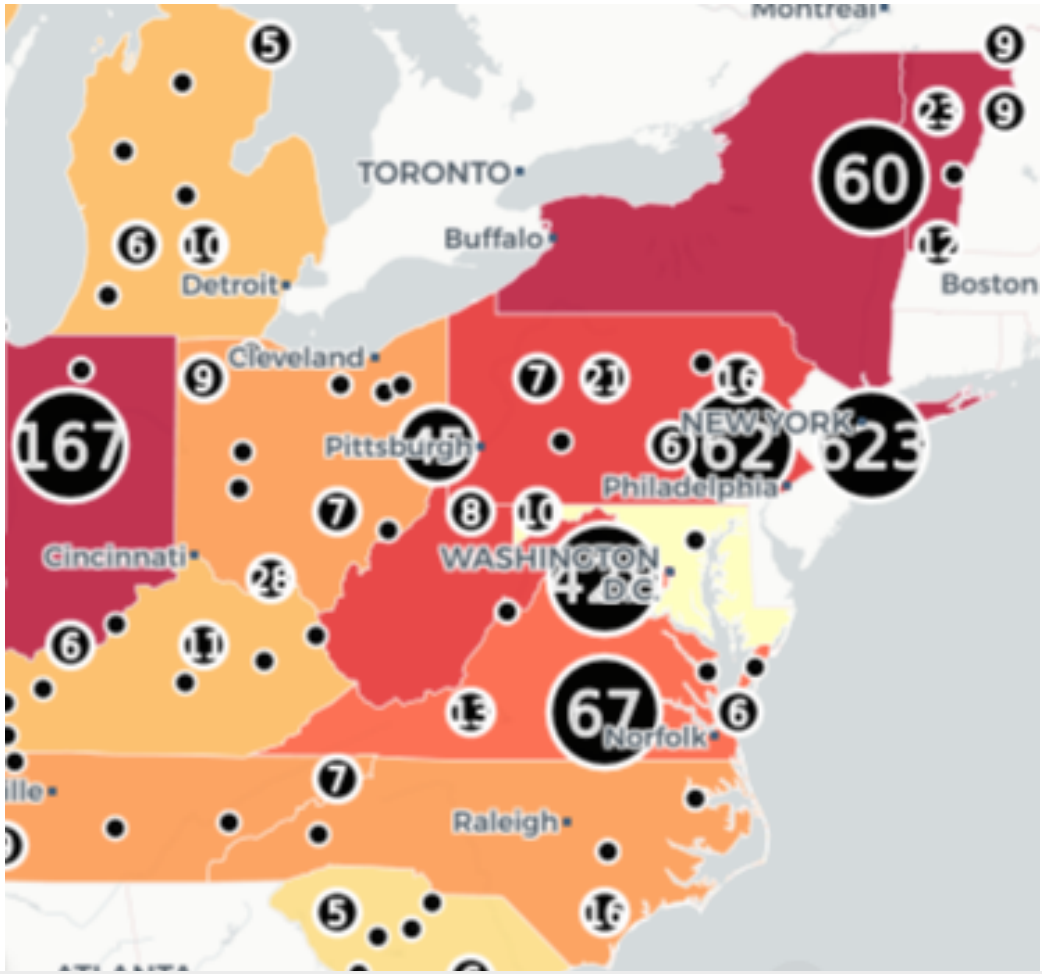
```
## [1] "WI" "MN" "WI" "NJ"
```

Once you've got tidy data with latitude and longitude coordinates, you're ready!

Choosing Your Tool

Carto

Carto is a browser-based resource primarily marketed towards business analytics, but for a time it also saw considerable use by digital humanists.



Carto

Why You Might Use It	Why You Might Not
Automatically geocodes data	Can't export coordinates after geocoding
Easy to use, including for dynamic maps	Carto logo and aesthetics seem to "brand" exported results
Easy to share	Data and maps become public
SQL API allows for additional customization	Limited to geospatial analysis
But...	Now only free with a (free) GitHub student developer pass

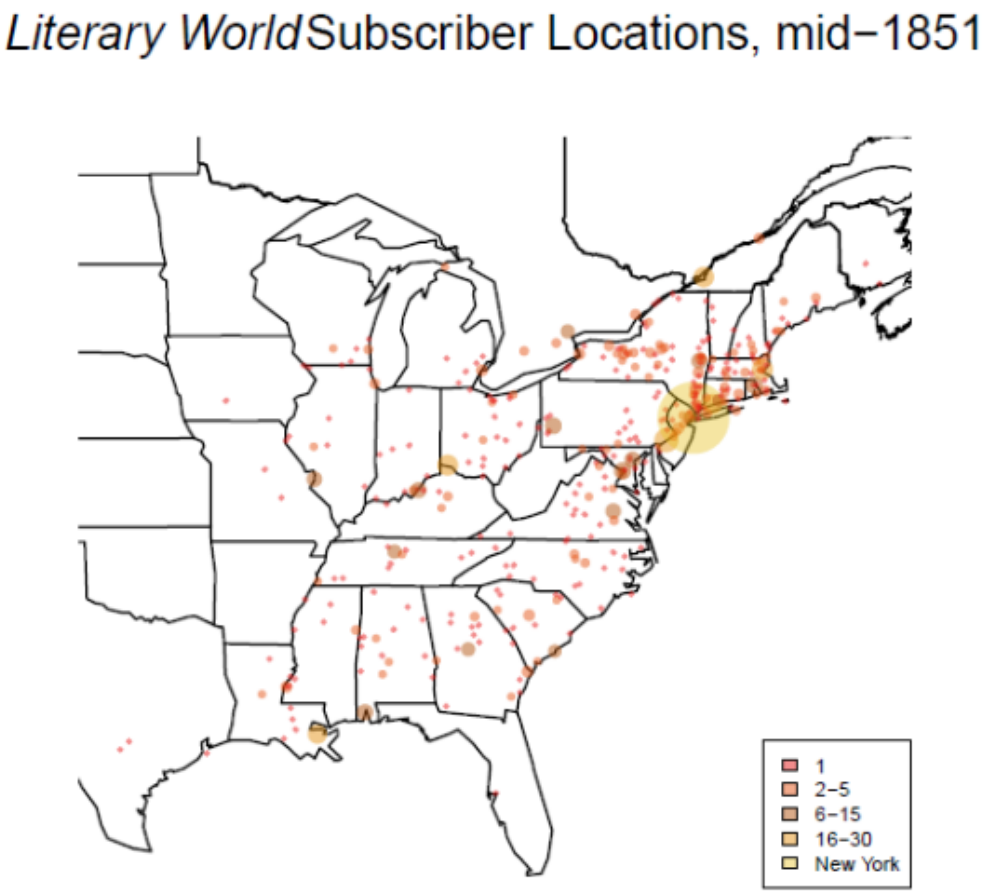
Tableau

Tableau is a more expansive software built for business analytics but also used by some humanist scholars.

Why You Might Use It	Why You Might Not
Can move between maps and myriad graph options	Downloaded software, no online version
Really striking visual capacities	Relatively complex interface
But...	Not free

R

It is also possible to map in R using the `ggplot2` package, part of the same group of packages as `dplyr`.



R

Why You Might Use It	Why You Might Not
Blandly professional and no software logos: great for publishing	"Blandly professional"
Highly customizable	Customization requires more trial and error than more accessible software
If you're already working in R, it's easy to move directly from data analysis to visualization	Requires familiarity with R

Thematic Mapping: Palladio

Palladio

Palladio is an online tool built by the Stanford Humanities + Design Lab. It's built specifically with Digital Humanists in mind to function as an accessible multi-function tool for analyzing and visualizing humanistic data.

For my examples I'll be working with metadata of the American Antiquarian Society periodicals holdings. If you'd like to follow along on your own computer, you can download this data here.

Palladio

Why You Might Use It	Why You Might Not
Can easily move between maps and a number of possible graphs	Less customization than available in other programs
Particularly amenable for working with networks	Doesn't yet support detailed quantitative network analysis
Online platform that takes no time to get started: no software download	Visualizations are not interactive and cannot be shared online: only by sending static exported copies
Free to use	

Palladio

- Getting started
  - First, you'll need to make sure that your geo coordinates are in the same variable, in the form "latitude, longitude," rather than separate variables for latitude and longitude.
  - It's best to also make sure that any dates are either in YYYY or YYYY-MM-DD format.
  - Visit the Palladio website and enter your data by copy-paste or drag-and-drop to begin.
  - If Palladio doesn't like some of your data formatting, it'll let you know with a red dot.
  - Give your data a title and then give your project a title.
  - The tabs on the navigation bar at the top of the screen allow shifting between the different features: Map, Graph, Table, Gallery.

Palladio: Tables

- The Table feature allows for some basic analysis of your data.
  - The "Row dimension" function allows you to either view each individual observation in your data by selecting a variable that is unique for every observation (ex., "Publication Name") or summarize your data by selecting a variable for which observations have the same value (ex., "Publisher State").
  - The "Dimensions" function allows you to select which variables are displayed.
- The lower-left corner provides options to generate timeline or timespan charts, with optional faceting.
  - Clicking on "Timeline" or "Timespan" will open a chart and an options bar where you can determine which variables of your data get used for each aspect of the chart. We could, for example, make a timeline of start dates organized by publication name or a timespan of start dates to end dates organized by publication name.
  - Once you've got a chart the way you like it, you can click the tiny icon with the arrow pointing down in order to export just that as a .svg image file.

Palladio: Maps

- Click "New Layer," leave the type tab as "Data," and give your layer a name.
  - In the "Places" drop menu, select whichever variable has your geo coordinates (ex., "coords").
  - In the "Tooltip" drop menu, select whichever variable you want to show up when you hover over a point on the map (ex., "Publication Name").
  - For most data, you'll want to check the "Size points" box so that the size of the points on the map corresponds to the number of observations for each location.
  - . . . unless you want it to correspond to something else! In which case, change the "According to" drop menu from "Number of Untitled" to that other variable. Otherwise, leave it.
  - Click "Add layer".
- We can add other layers too, either for additional data, shapes, or tiles. I for one need to add a layer for "Streets" (which includes state borders) under the "Tiles" tab.

Palladio: Filtering

- One of the handiest aspects of Palladio is the ease with which it allows users to filter and facet data.
  - Let's revisit our charts by clicking the arrow in the bottom right corner; leave the timespan chart minimized for now.
  - Click and hold somewhere on the timeline chart and slowly move the cursor to the right.
  - Wherever you let go will remain an active filter, listed on the navigation bar; you can delete a filter by clicking the red x next to it here.
  - For a map of all the periodicals active at any point during a particular time range rather than a map of all the periodicals started during a particular time range, filter with the timespan chart instead.

Palladio: Faceting

- We can also facet data: that is, select only specific variables from a particular variable to explore.
  - Click on the "Facet" tab next to the "Timeline" tab on the lower navigation bar.
  - In the "Dimensions" drop menu, pick one of the variables. For this example, pick "Subjects."
  - All the possible values will now be shown, grouped by total occurrences. Click on one or more group to make the map reflect just the locations of the corresponding observations; click a second time to remove a group.
- If you return to any previous timelines or timespan charts, you'll find they now reflect this facet.

Palladio: Other Features

For time purposes, we won't touch on Palladio's Graph (network) and Gallery features. If you'd like to learn more about these, you can read more on Palladio's tutorial page or check out Miriam Posner's more comprehensive introduction to Palladio - or come to my office hours!

Palladio: Don't Forget!

Unfortunately, while Palladio is good for exporting graphs it is not yet good at exporting maps... at all. Following the advice of the Palladio developers themselves, then, I recommend taking a screenshot of your map.

Palladio doesn't save your work for you: instead, you need to download your work using the "Download" button in the upper right corner. This will dump a .json file onto your hard drive. Next time you visit the Palladio homepage, upload this file instead to pick up where you left off!

Thanks for Coming!

We would really appreciate it if you took a minute to fill out our brief feedback survey at https://rutgers.ca1.qualtrics.com/jfe/form/SV_a3itiZN18dY3fc9.