**Lab 6:**

**Mapping Point Data with Carto**

**Adding Layers**

**Embedding Carto Maps in WordPress**

**Introduction**

Lots of datasets include information about point locations stored as coordinates, such as longitude and latitude. Cities often report vehicle crashes, 311 calls, or crimes in this format. These are “geographic information” in the sense that they have a geographic component, but they can’t simply be dropped on a map like a shapefile. To see coordinates stored in a table as points we need to tell our GIS which columns represent each component of the coordinate, and what projection these coordinates are in, so they can be transformed into points. You may want to add another layer on your map to see how crime is distributed, let’s say, by supervisorial district and then publish your work! In this lab, we will learn:

* how to map a table with longitude and latitude columns using Carto
* how to add a layer
* how to embed the map in WordPress

**Preprocessing (20 mins)**

We are going to upload our point data to Carto as a CSV file. While Carto is usually pretty smart about identifying file types and formats, it can be a little fussy about how longitude and latitude are formatted (let’s start calling them “lon” and “lat” for brevity). Lon and lat need to be stored as numbers in separate columns.

Let’s look at the San Francisco crime data file provided on bCourses. It is a list of reported assaults between November 1 and December 31, 2016. The geographic information is in the “Location” column as sets of lat and lon coordinates.

The geographic information is stored as text in a single column. How can we fix this? We are going to start by separating the data into two columns:

* Open the file “SF\_CrimeStat” in Excel. Create two new columns after the Location column. We are going to fill these columns with the lat and lon values.
* The easiest way to “parse” the combined lat and lon into separate columns is to use the “Text to Columns” wizard, accessed with a button in the data tab. Highlight the Location column and then click “Text to Columns.”
* It will ask you how to identify where to split the data into columns. We are going to use the “Delimited” option, which allows us to specify certain characters the “delimit” the location of each column within each row. Click Next.
* Select “Comma” as the character that separates the longitude and latitude values. Now, if you were to click “Finish,” you would have separated the two coordinates at the locations of the commas, but you will still have parentheses left at the beginning or end of each coordinate. To get rid of the opening parentheses, you can add “(“ under “Other.” This will treat the parentheses as an additional delimiting character and remove it. Click “Finish.”
* Clean up any blank or unused columns (the original Location column may have been overwritten) and add headers, “Lat” and “Lon,” to your new columns.
* You will still have lingering closing parentheses on the lon coordinates. One way to get rid of them is through “Find and Replace.” Select the whole “Lon” column to specify that you only want to find and replace values within it. Under the Home tab, click the down arrow next to “Find & Select,” then click on “Replace…”
* You want to find all the “)” characters and replace them with nothing. Add “)” to the “Find what:” box; leave the “Replace with:” box blank; then click “Replace All.” Excel will tell you that it made a couple of thousand replacements almost instantaneously. A lot faster than manually editing!
* Now that there are no text characters adjacent to either the lat or lon values, Excel will automatically identify and format them as numbers.
* Save the changes to your CSV file. We’re ready to bring it into Carto!

**Bringing Data into Carto (15 mins)**

Once you are logged in your Carto account, let’s upload the data we are using today. Start by uploading the **‘Supervisor Name’ shapefile** following the steps you learned in lab 5. We will use this data set to overlay supervisorial districts on top of our crime data

Now let’s add the **crime data**. The process for uploading the CSV table is the same as we followed for the shapefiles. When you upload the data, notice that it automatically recognizes your columns named “Lat” and “Lon” and uses them to construct points. It will also recognize columns that are named “Latitude” and “Longitude,” but not if you name them other things, such as “X” and “Y.”

Note: If you don’t have recognizable column names, the “the\_geom” column will have null values and when you add the dataset to a map it will not draw any points. You will need to manually tell Carto which columns to use for lon and lat coordinates by clicking on the “Geocode” button.

Geocoding is the process of taking numeric or text information, such as longitudes and latitudes or city names, and converting it into geographic points. Some software and websites even allow you to geocode postal addresses. You can do this in Carto, but they charge you for it. Google “geocoding” to explore other free and paid options if you need to map addresses.

Under “Select data type,” select “Latitude and Longitude” from the “Data Type” dropdown it if isn’t already selected. Then, under “Define your parameters,” select the columns where you respective lon and lat data are stored. Finally, click “Apply.” Carto will add point geometries to your dataset.

**Visualizing our Point Data (15 min)**

Let’s play with options for data visualization!

Click on the “**Style**” tab in the Layer Option for your crime points. The tab will look different than it did for the other data types. In particular, there are several aggregating options will allow you to summarize the point data.

By default, “Points” is selected and displays one point for each data record. This is a straightforward representation, but the map can get cluttered if there are many overlapping points. The next two options, “**Squares**” and “**Hexbins**,” aggregate the results that fall within square or hexagonal grid cells, similar to a raster representation of density. This can be helpful if there are lots of overlapping or closely spaced points.

You can also show data over time using animation. Scroll to the right under “**Aggregation**” to find the “**Animated**” option. Then, look under “**Style**” to find the column with which your data will be animated. Pick the “**date**” column. If date is unavailable to select, look at the attribute table for the dataset (go back out to the dashboard and click on the dataset). What data type is the “date” column? Carto will need to recognize that your “date” data are stored in the date data type in order to animate it. Change the data type of the date column by clicking the vertical ellipses next to the column header, selecting “Change data type,” and then “Date.” Confirm the change.

Now that your dates are recognized as dates, return to your map. Select the “date” column under the animation options back in your Style menu. Now Carto will play through the dataset as if it were a movie, with crimes for each date flashing up as the play head moves through time. It looks really cool, although it can be challenging to draw analytical inferences through an animated map. Explore other ways you could symbolize the dataset based on dates or other columns.

**Data Widgets (15 min)**

The “Data” tab provides similar functionality as it did in the previous lab. It allows us to create widgets that sort the data into different bins. For example, add a widget for the “**resolution**” column. This sorts the reported crimes into resolved and unresolved crimes. We could use this to determine if there is a geographic pattern for where crimes have and have not been resolved.

Adding a widget for day of week would allow us to see if crimes were more common in specific locations on certain days. A widget has likely already been created for date (this happened when we created an animated map). This widget will show a range from November 1 to the end of the year. This allows us to select for specific dates. With a data set that covers the whole year, you could see if crime patterns are different during different seasons. Play around with different widgets and styles to tell different stories with your map.

**Adding a Layer (10 min)**

* Go to the “Layers” dialog. From here, you can add another dataset as a map layer;
* click the “+ Add New Layer” button at the top.
* Select the “supervisor\_name” dataset from “Your Datasets.”
* Then click the “Add Layer” button in the lower left.
* Overlaying this layer with the crime data will allow us to see which districts have more crime.
* Drag the new layer below your sf\_crimestat layer so you can see the points
* Play around with the layer style. Any idea how to add labels (hint: there is a label box under Style)

**Analysis Tab (15 min)**

* You can add multiple analysis functions to a single layer.
* In the “Analysis” tab, select the **“+ Add New Analysis**” button.
* Select “**Intersect and Aggregate**” function. This function allows you to find overlapping geometries from a second layer and aggregate its values in the current layer. Click on the tile and select add analysis. This adds a new tile in the Analysis panel.
* Your base layer should be ‘supervisor\_name’. Under “Intersect and Aggregate,” click on the “**Target Layer**” box andselect **“sf\_crime…”**
* Under “Measured by”, click on the Operation box, and select ‘Count’
* Click “Apply” at the bottom when you are finished.
* Did your map change?
* Go to Style and change Polygon Color by selecting ‘By value’ and then choosing the new column ‘count\_vals’ and a color scheme. Did your map change?
* Select the ‘Pop-Up’ menu, select an ‘Style’ and click on some of the variables under the ‘Show items’ list (e.g. supdist and count\_vals). Now hover over your map and click on one of the districts.

**Embedding a Carto Map in WordPress (10 mins)**

* Go back to the Layers menu
* Click Publish in the Layers menu
* After Publishing your Carto map, you will see two boxes. Copy the text that is in the Get the link box. This is the link to your map.
* Now go to your previous post in WordPress and add a new block with the plus (+) icon.
* Then you will see a pop-up window with multiple alternatives (page break, paragraph, image, embed, etc.). Click on Embed.
* You will see a new pop-up window where you need to paste your map link and click on
* Embed again
* Now you should see your map on your post. Go and click on Preview and then Update. You should see your interactive Carto map right there! Amazing!
* Use your post URL link to see how others will see your post. Play around with the map (zooming in and out, use filters if the map has an active widget, etc.).
* Go ahead and add another map or image.
* Think about the story you want to tell in assignment two and how WordPress is useful for it.

Note: Some years ago, Wordpress wanted you to use something called an 'iframe' (instead of the map URL) to embed a map. If you are curious, google the term 'iframe' and explore in Carto how to obtain the iframe of your map (hint: look at steps 11 and 12). Other content management systems may require an iframe.

**DIY: Figuring Out What Options Are Available**

This week we provided a basic introduction to Carto, but we have only scratched the surface. You will need to play around with Carto to see what it can and can’t do. Sometimes it will please you and sometimes it will be frustrating. When you want to try out a new tool, we recommend starting with a small dataset or a subset of your records. This helps to clarify whether any issues are related to the underlying structure of your data or is simply a limitation posed by a large number of records. Unfortunately, Carto can sometimes be overwhelmed by very large datasets.

Happy mapping!

**APPENDIX**

**Longitude and Latitude**

The most common coordinate system for geographic information divides the globe into “degrees” of longitude and latitude. Lines of latitude demarcate distance north or south from the equator (remember, latitude = ladder). Latitude ranges from -90° (or 90° south) at the south pole, to 0° at the equator, to 90° (or 90° north) at the north pole. Lines of latitude are always parallel— they are the same distance apart from one another, no matter the latitude. For this reason, they are sometimes called “parallels.” Lines of longitude, which are sometimes called “meridians,” describe extents to the east or west of the “prime meridian,” which runs north-south through Greenwich, England. (Can you guess which country formalized the modern system of longitude and latitude?) Longitude ranges from -180° (or 180 west), running approximately along the International Date Line through the middle of the Pacific Ocean, to 0° at Greenwich, to 180° (or 180° east) back at the International Date Line. Lines of longitude converge at the north and south poles—they are not parallel—so the length of a degree of longitude varies at different latitudes. The lengths of a degree of longitude and latitude are equal at the equator, and a degree of longitude approaches zero length near the poles.

Using the longitude and latitude coordinate system, we can describe the location of any point on the (approximately) spherical globe. Berkeley is at about -122° longitude and 38° latitude. We can get more precise by adding decimal places (e.g., -122.2730°, 37.8715°; these are called “decimal degrees”) or through a more old-fashioned system that divides degrees into 60 “minutes,” and minutes into 60 “seconds,” often with decimal seconds (e.g., 122° 16' 22.0116'' W, 37° 52' 17.9976'' N; these are called degree-minute-second, or DMS, coordinates). Most GIS software prefers decimal degrees, so you’ll need to do some math to convert DMS coordinates if you run across them.

Because it is conventional to make maps where north is oriented upward, we usually use longitude to represent the horizontal (“X”) coordinates, and latitude for the vertical (“Y”) coordinates.

**Projections in Carto**

Carto assumes that you’re working with longitude and latitude; the basic online interface won’t let you work with other projections (though you can write some custom Javascript if you’re feeling ambitious and want to rebel against Web Mercator). Conversely, in a full-fledged desktop GIS software, like ArcGIS, you would need to specify which coordinate system your coordinates are stored in. If you have a dataset with coordinate values that aren’t within the typical lon and lat ranges, you may be dealing with a different coordinate system.

UTM coordinate systems, for example, use meters as their primary unit and often have values in the 10,000s, 100,000s, or millions. If this is the case, you may need to use a more advanced GIS software or find an online tool to convert between coordinate systems.