

## Geography N80 – Lab 05

### A more complicated map (of a more complicated story)

|                   |  |
|-------------------|--|
| <b>Materials:</b> | Bcourses: Files → Lab Materials → Lab05 → San_Francisco_Evictions.zip              |
| <b>Tools:</b>     | Carto  |
| <b>Outputs:</b>   | An html page including your web map and the answers to the comprehension questions |
| <b>Due date:</b>  | Monday, November 19 <sup>th</sup> , 9pm  |

### Introduction

In this lab you will look particularly at options for presenting point data in a different way, focusing on the *intensity* or *density in space* of a set of points. This method is very often used in the form of ‘heat maps’ with which you may be familiar. While such methods can produce attractive visual effects, it is not entirely clear how useful they are and that is something you will need to think about as you work on this assignment..

### Materials

The context for this lab is still in the Bay Area, but with an angle more relevant to our conversations about redlining and gentrification. We are looking at Ellis Act Evictions data from 1997-2013. We obtained these data from the Anti-Eviction mapping project, which you should take a look at here: <http://www.antievictionmap.com/>

Note also that the data may not exactly be absolutely correct in their spatial locations. The data we obtained included street addresses, and had to be *geocoded* to latitude-longitude coordinates. This was accomplished using the service at Texas A&M University at <http://geoservices.tamu.edu/>. One side-effect is that while some points were geocoded exactly to the land parcel, some were only geocoded to the street segment, and some only to the ZIP code area (there is a column in the data table that tells you which). The data on the anti-evictions website are more accurate, since they have been manually located by locals who know the area.

A number of spatial and attribute data tables are available collated into a single zip file in the bCourses files folder for this lab:

|   |  |
|---|--|
| evictions-geocoded.csv  | 2490 geocoded Ellis Act eviction records from 1997 to 2013. These include associated dates, the petition number, and the street address, and the geocoding accuracy.   |
| SF_tracts.zip   | Shapefile of 2010 census tracts for San Francisco county   |
| SF_Tracts_2010_ageSex.xls<br>SF_Tracts_2010_race.xls<br>SF_Tracts_2010_tenure.xls | Tables of 2010 census data relating to various factors (age and sex, race, housing tenure) for San Francisco tracts  |
| tech_bus_stops.json   | GeoJSON data on the locations and some details of ‘tech’ shuttle stops. These were collected by Stamen. See <a href="https://hi.stamen.com/the-city-from-the-valley-57e835ee3dc6">https://hi.stamen.com/the-city-from-the-valley-57e835ee3dc6</a> for details. |

There are at least two maps of a potential relationship between tech shuttle locations and changes in San Francisco real estate. One is this one at the evictions mapping website:

<http://www.antievictionmappingproject.net/techbusevictions.html>

Another is this one by a 'data journalist' highlighted by CityLab:

<https://www.citylab.com/transportation/2014/01/visual-study-just-how-bad-tech-shuttles-are-san-francisco/8221/>

The general aim of this assignment is to make your own map of the evictions, to think about the process and decisions you make in doing so, and also to consider the limitations of the tools you are working with for accomplishing the effects you would like. See the end of this document for a more extended description of what we are looking for.

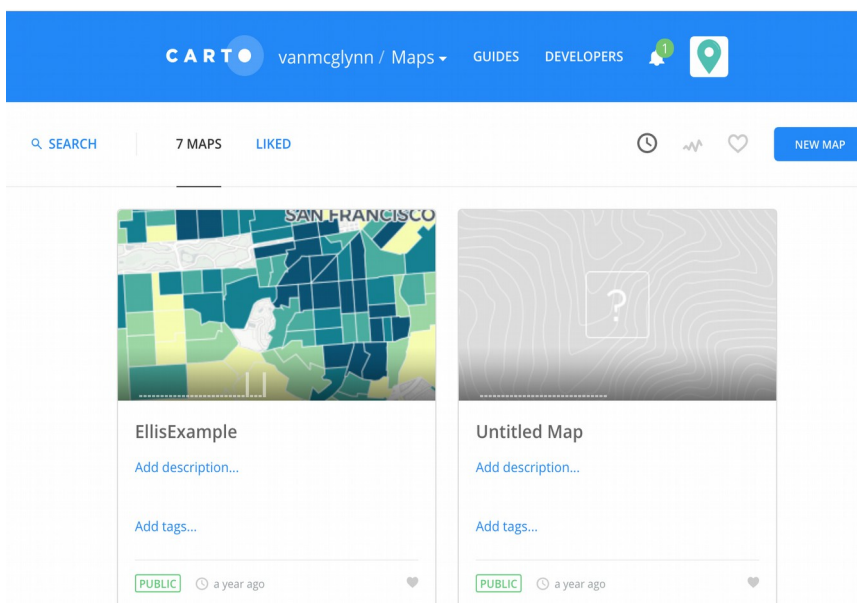
## ***An intro to Carto + a nostalgic rant***

CartoDB, not dissimilar from mapbox, was a super slick web-based tool for generating simple (and not so simple) web map analysis. While "New Carto" feels a bit like "New Coke" to the older users, there is a lot of functionality worth diving into in order to complete this lab, which we'll go over briefly below. Note that Carto keeps changing its education account standards. To date, the things you need to sign up are a .edu email address and an educational github account: <https://education.github.com/pack> (with more documentation here: <https://carto.com/help/getting-started/student-accounts/> ). Sometimes Carto will still give you a subscription error, but often these are localized to individual operations, so don't fear and make sure to read all the messages the interface throws at you. Further, the current Carto interface makes it very simple to do some really cool things but somehow obnoxiously difficult to accomplish simpler tasks. We will give you some light instruction on the interface to get started, but with the understanding that we're reviewing some key concepts in this lab so part of the job is applying concepts you already know to a new tool.

## Setup

Once you have your account set up, go to the dashboard, click new map and start adding data. The key mode in Carto is add data to a map (your own or from within Carto's repositories) and analyze within that interface, at once way more straightforward and way harder than mapbox. There's a logic to all the other details that you'll get the hang of with some experimentation, but to start off, I recommend the

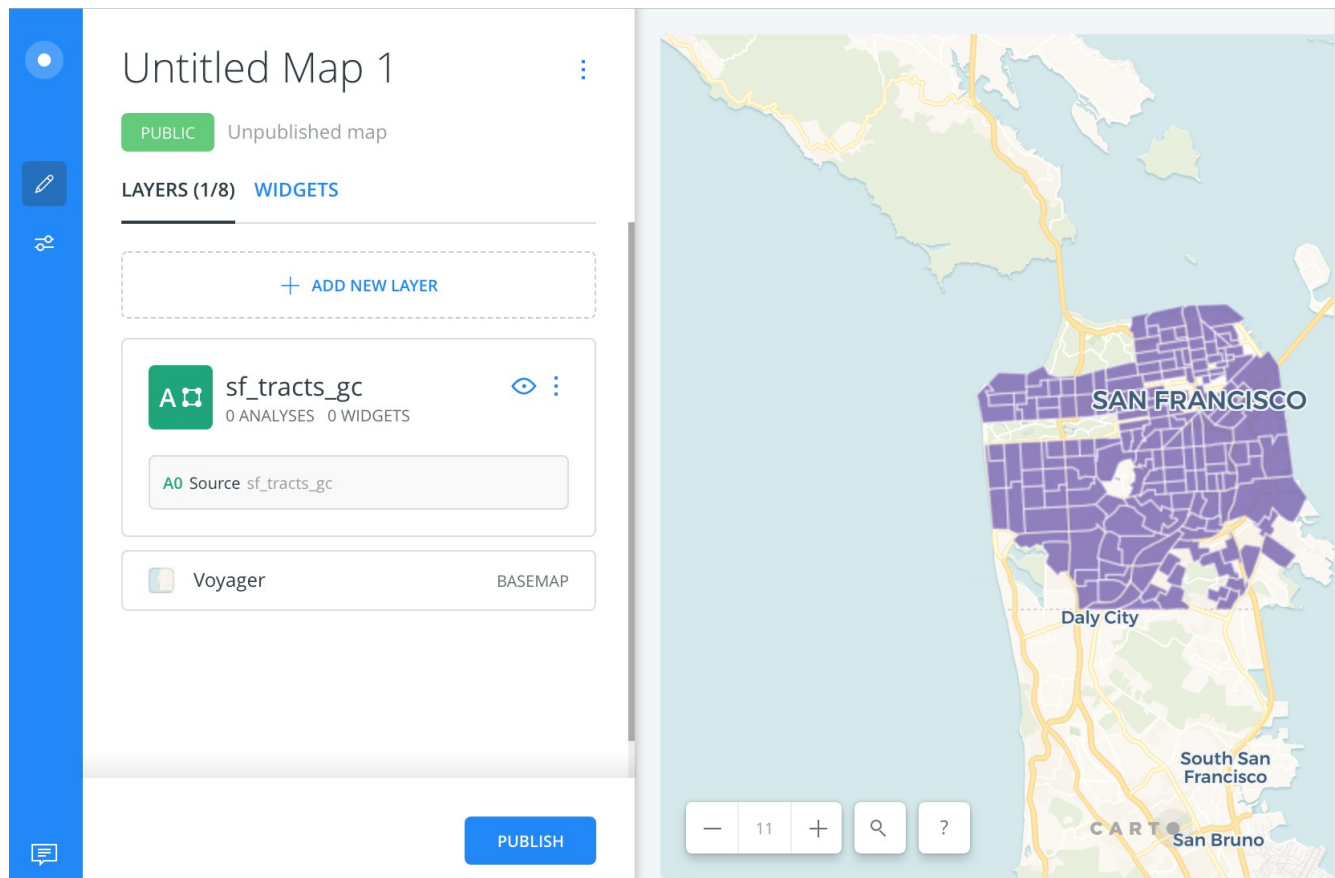
documentation from the Carto folks themselves, specifically here <https://carto.com/help/getting-started/dashboard/> and here <https://carto.com/help/tutorials/> .



The dashboard is where all of the maps you make reside. My user name is vanmcglynn, so my dashboard url is <http://vanmcglynn.carto.com/dashboard> which you can see to the left (complete with some maps I had already made before taking the screenshot).

You'll also see that your dashboard has links at top right to **DOCUMENTATION** (i.e. help, always useful) and to your account settings.

To get started you click the **NEW MAP** button. In the next window that appears, connect to your collected data (from the previous lab) by clicking the **CONNECT DATA** link and navigating to the relevant file on your computer (ie, the package of files we provided in bCourses). Carto can cope with GPX, KML or GeoJSON formats fine. If the data are OK, you'll get another window, where you should again click the **CONNECT DATA** button. When it finishes you will be back at your maps, page, and now you should again click the **NEW MAP** button, select the data you just added and then click **CREATE MAP**. Your data should appear in a map view.



Notice the default for the map is three layers: the labels, the base map, and the data layer you just added. Click on the layer menu to start editing map contents. Many of the available options are self-explanatory, and will repay time spent experimenting with them. It is important to notice the connection between the menu options and 'CartoCSS' (an option you can find on the bottom of the style menu) – changes made using the wizard controls will affect the CartoCSS settings, and vice-versa. The same is applicable to the 'SQL' and 'filter' controls, although you are less likely to want to experiment with these for the purposes of this assignment

As you should see from casual experimentation, there are many ways you can manipulate the map beyond what is covered here. For now, the watchword is *experiment!*

## Clusters

When mapping points, the clusters option will, at any particular zoom level, aggregate points that are close to one another, into a ‘cluster’ and label the cluster with the number of points it represents. This doesn’t need too much explanation, but may be an option to explore when showing the spatial distribution of the evictions. It has the advantage of making the map appear less cluttered. Clustering is hidden under **Analysis**, about halfway down the menu list.

## Intensity Mapping

This is a popular approach, although the default options in Carto do not really show the effect off to best effect. The default option will create a merged together mass of color, such that in areas where there are a lot of points, they will overlap and create darker regions. You can change the look of things by experimenting with the size of the points (Style → **Marker Fill** number), and their opacity (the second **Marker Fill** number, where 1 is opaque, and numbers near 0 are progressively more transparent. Setting the **Marker Stroke** will change the width (first number) and opacity (second number) of the outlines of the circles. It can dramatically change the look if you set the stroke width to 0, so that there is no outline. You might also try the **Blending** option, which gives a few options for layering data without futzing with individual controls.

To tune the intensity visualization further, you need to switch to the CartoCSS editor. There, you can modify the settings more precisely. In particular, it is worth experimenting with the ‘marker-comp-op’ setting, which determines how colors are ‘composited’ (i.e. combined) when the circles representing points overlap. There is a list of the options here

<https://carto.com/developers/styling/cartocss/#effects-of-composite-operations> (scroll up a section—the anchor on the site is broken)

along with examples of their effects, although it may be better simply to experiment. You will find that if you want to use the **lighten** option (which makes regions with more points brighter) that it works best to set the opacity to low values. An alternative is the ‘plus’ setting, which can work well.

Although you can accomplish a lot with these simple options, to achieve a more ‘classic’ heat map look, you need to overlay several sets of transparent circles of different sizes and colors and composite them. That rapidly gets complicated, and you may instead want to look at the Heatmap option (although, as I have suggested in class there are some issues with this approach also).

## Density mapping

Another option is density mapping. This imposes a set of hexagons (often called ‘hexbins’) across the region where there are points, calculates how many points occur in each hexagon, and then applies a color sequence as you would for a choropleth map. This is quite a nice effect, and you have the usual options available, to adjust the size, color ramp and opacity of the hexagons, and also the line width of the hexagons. Hexbins and a few other options can be found under **Aggregate** in the layer controls.

Unfortunately, if you want to precisely control the color sequence, you need to go into CartoCSS. There, you can precisely fix the break point values in the color ramp, which may dramatically alter the appearance and impression of where the greatest number of evictions are. The hitch is that these break points will remain fixed when the map is zoomed, and so you may find that the map only works well at one particular zoom level. Again: experiment!

## Spatial Joins (making your own choropleth)

Here's another option: count how many point events (the evictions) are inside a fixed set of polygons (like for example census tracts – which we know have *roughly* equal populations). To do this, you'll need a set of polygons, which is where the **SF\_tracts.zip** file provided comes in. Add that to your Carto account. Now go under the dataset view, go to style, and under Analysis, you'll see an option for “**intersect second layer.**”

Once that's done, you can make a choropleth map of evictions... although, really it should be *normalized by population or area* to be a really useful choropleth. Even so, it is instructive to compare how different is the impression you get of evictions from this map, compared to some of the other possibilities.

If you need to normalize by population or area... well... it's a few steps, but it's possible. First, you add one of the San Francisco census tracts data files provided—one that includes useful attributes to normalize by (such as the population or the number of households). Then merge the data in this table to the tracts file. Then add a new column to the data table to store your **eviction\_percent** result. You have to ensure that this is a ‘Number’ data type. Then, you can use SQL to do a calculation. For example if you merged the tenure data and your new column is called **eviction\_percent**, then the expression you need in the SQL editor is

```
UPDATE evictions
SET eviction_percent =
CASE
WHEN households > 0 THEN (0.0 + evictions) / households * 100
ELSE 0
END
```

to normalize by the **households** attribute (also in the tenure data table) and produce a percentage rate of evictions. It might make more sense to normalize by the number of people living in rented accommodation (the **renter** attribute). This would give you an eviction rate percentage attribute to use as the basis for a choropleth map.

Keep in mind also that you can also use the tracts data to bring in any other census data of interest. Further, for county level data, there is a baked-in spatial join function in Style → Aggregate. Scroll to “Administrative Boundaries” under **Aggregate** and you can automatically join your point layer to counties (or a few other pre-set admin layers). Now you can choose the operation (count, average, sum, etc) you want to group your point data by before honing the display further.

## Torque Mapping

Another option—one that is very popular in online maps—is animation. This is available again through dataset style, under **Animation**. When you select that option, you choose a variable for the animation to step through, which should be the ‘date’ attribute. That's more or less all there is to it, although as usual there are many tunable options to explore. This is definitely another case where experimentation is the best option.

## Data

While we have given you enough data to do the assignment, we are also always excited to see new solutions with additional data sets added in. For San Francisco, we are wealthy with options. First, you

all should already know how to download data from OSM by now. Further, Carto itself provides a lot of its own data. In your map, click **Add New Layer** and instead of the default Your Datasets tab, click Data Library to search what Carto has to offer. There are many, many, more options (web search engines are your friend), but here are three of particular interest for this project:

- <https://geodata.lib.berkeley.edu/> (Cal's library of geodata—kind of like geoOskiCat)
- <https://datasf.org/opendata/> (the city of San Francisco's open data site)
- <https://data.ca.gov/search/type/dataset> (the same sort of portal as SF but for the whole state—maybe less relevant here but keep it under your hat for lab06)

## **Assignment deliverables**

To complete this assignment, you need to do the following:

### **Make a web page (not just a map) presenting your perspective on the evictions data.**

The maps on the <http://www.anti-evictionmap.com> website give an idea of the sort of thing you could try to achieve. No particular explanation or angle on the data is expected or required: maybe you think the tech shuttle stops angle is overblown? Map something different, and don't worry about the shuttles. Maybe the time sequence is interesting to you (the rate of evictions does seem to have accelerated in more recent years).

A key challenge is how you combine different kinds of data in a single map—the evictions are points, but together they form a pattern that might be better understood in terms of its **density** or **intensity** across the city. At the same time there may be contextual factors, perhaps best understood via census data that provide a backdrop to the evictions data. Or again, you might want to use completely different data from the sources listed above. The key requirement is to use some type of density representational tool in the map, regardless of its conclusions.

To embed any maps you make (you can make more than one) in a webpage, use the Carto 'SHARE' button and copy and paste the 'Embed it' <iframe> code into the HTML page you are making. You may want to change the width of the map from the '100%' default provided, to a particular pixel width, or smaller percentage value.

The text in your webpage should explain your map, discuss design choices, and explain the perspective on the evictions that it represents. It may include links to other pages and resources, (e.g., images), and should link to any materials that you have found useful in completing this assignment.

Remember to make sure your webpage works correctly using a local server program as introduced in the early Leaflet labs.

### **Compress your HTML file, along with any additional files it needs (such as CSS files, or images) into a zip archive and submit it on bCourses in the usual way**

For guidance, any maps you make will account for 5 points of this assessment, while the commentary will carry 3 points

*David O'Sullivan, Will Payne, & Eve McGlynn July 2017-Nov 2018*