UCLA Sandbox

(https://sandbox.idre.ucla.edu/sandbox/)

≡ Menu

Intro to GIS: Got data? Let's map it!

Pre-class poll

Welcome! Before we begin our adventure on "spatial thinking," please take a minute to take this poll (https://onlinepoll.ucla.edu/polls/2665/responses/enter_password).

Pre-workshop poll (https://onlinepoll.ucla.edu/polls/2665/responses/enter_password)

(password: "ucla")

After the workshop, please take this survey:

Post-workshop survey (https://bit.ly/2RLCsBY)

Part 1: Introduction to GIS and ESRI

1. Background: Geographical information in the U.S.
2. Hello Map: Thematic Mapping and the basics
    - Connecting a folder
    - Adding Layers
    - Setting the projection
    - Attributes
    - Symbolization
    - Labeling
    - Choropleth Maps

Part 2: Working with spatial data

1. Acquiring data
2. The LA Data Portal
3. Project the data
4. Hot spot analysis

Part I: Introduction to GIS and ESRI

Gettin' down and dirty with ESRI

The first step in this tutorial is to understand that we are covering the basics of desktop GIS analysis using ESRI's ArcGIS software suite. This is by no means an all encompassing "entirety of GIS" tutorial, but rather a view on how GIS can be used to build maps from ESRI's perspective, limited by the functionalities of the software being covered. There are many other tools you may want to consider to do your spatial analysis, including R (http://spatial.ly/r/), Python, Carto (https://carto.com/), Mapbox (https://www.mapbox.com/), D3 (https://d3js.org/), and qGIS (https://www.qgis.org/en/site/).

The core function of the ESRI ArcGIS suite lies within two programs:

1. ArcCatalog – for managing GIS datasets
2. ArcMap – for mapping GIS datasets

What about OpenSource alternatives?

QGIS (http://www.qgis.org/en/site/) is as an alternative to ArcGIS that is free and openly available to the public on all computing platforms. Despite the accessibility of QGIS, there is a steeper learning curve for those learning GIS for the first time. However, those seeking a free low-cost alternative to ArcGIS can apply the concepts learned in this workshop with that program.

A little geo-background: Geographical information in the U.S.A.

Demographic information in the USA is typically arranged in a hierarchical geography. Starting from States, information gets broken down into Counties or Metropolitan Statistical Areas (MSAs). Each of those are comprised of Census Places which are similar to cities in their size and composition. The neighborhoods of each city are broken down into a Census Tract. Census Tracts are then subdivided further into Census Block Groups. Finally, Census Block Groups compose of Census Blocks, but data is not usually published at this level for privacy concerns.

In short US geography is organized like this:

States → Counties / Metropolitan Statistical Areas → Census Places → Census Tract → Census Block Group → Census Block

Hello Map

With geographical ideas in mind, it is finally time to map something! For this exercise, you are provided with a Workshop geodatabase, which is a collection of GIS datasets. A GIS dataset can be any of the following:

1. a vector layer – points, lines or polygons
2. a raster layer – an image, Satellite imagery, elevation data
3. tabular data – excel spreadsheet, csv, etc.

Vector vs. Rasters

Geographic data is stored either as vector (http://support.esri.com/sitecore/content/support/Home/other-resources/gis-dictionary/term/vector) data (as points, lines, or polygons) or raster (http://support.esri.com/sitecore/content/support/Home/other-resources/gis-dictionary/term/raster) data (as pixel grids).

Because of these differences in data storage, vector data is best suited for a human geography context (ex. urban planning, transportation forecasting, asset mapping), while raster data are best used for storing data on physical geography (ex. satellite imagery, elevation, watersheds, vegetation).

In ArcGIS, vector data is stored as individual .shp files or feature classes within a geodatabase. Raster data is stored as .tiffs, .jpgs, or other image formats.

Our geodatabase contains multiple GIS feature classes.

- Workshop Data Download Link (https://sandbox.idre.ucla.edu/Workshops/workshopData2018.zi

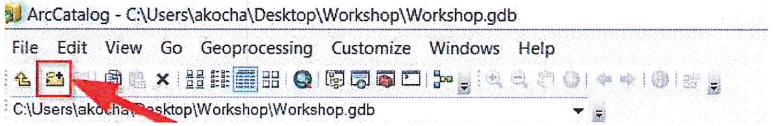
Download and extract Workshop.zip. Locate Workshop.gdb, and put it in a project folder for this workshop. You will learn how to inspect the geodatabase data in ArcCatalog, then use ArcMap to create some maps.

Here is a look at our Workshop geodatabase:

```
Workshop.gdb
|--us_cities
|--us_counties
|--us_states
```

Connecting a folder in ArcCatalog

Open up ArcCatalog and click the second button to the left, which is the "Connect Folder" button.



Navigate to the **Folder** where you extracted the "Workshop.zip" file and then select "OK".

Do not try to connect a file!

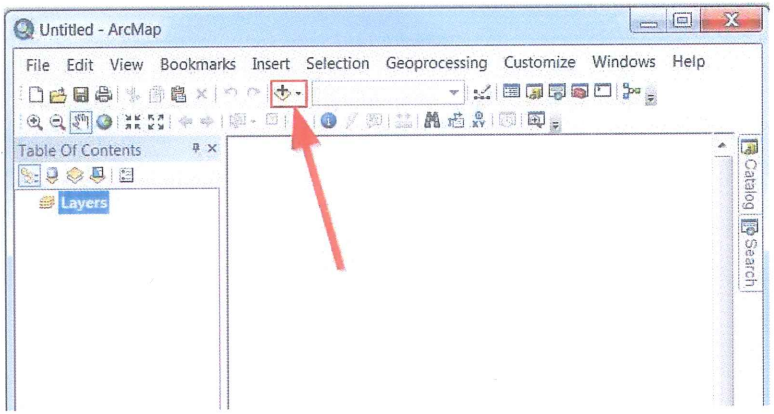
If you try to connect files, you will notice that the "OK" button is grayed out, connecting folders allows you only to choose folders.

View and Preview the data

After you've connected the folder, you can check **Folder Connections** and open the Folder which you've connected. Locate "**Workshop.gdb**" and double click it to view its contents. Browse for **us_states** and click the "**Preview**" tab.

Adding Layers

Now the time has come to fire up ArcMap and become a digital cartographer! The first step for any GIS project is to have data (more on this later). In order to add data to your project click on the "**Add data**" button:

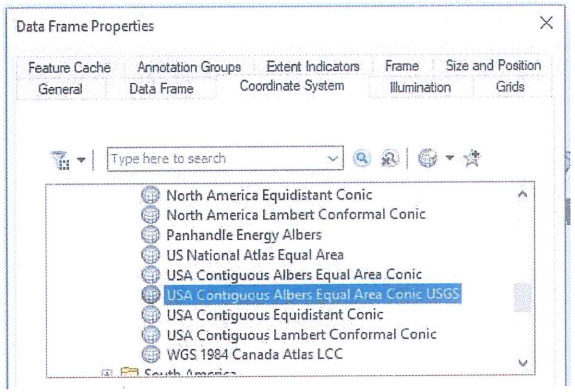


Notice how the connected folder can be selected and datasets be added now? Also, if your map is feeling a bit empty, you can add base maps by clicking the upside down triangle next to the **Add Data** button. Adding a basemap only provides reference information and nothing else.

Setting the projection

The datasets provided in this workshop are in a geographic coordinate system (**GCS_WGS_1984**). Geographic coordinate systems are measured in decimal degrees, and are useful when your data is global and/or comes with latitude and longitude coordinates. However, because of its angular units, it is not recommended for spatial analysis. Instead, consider projecting your data to a projected coordinate system that is suited to the region of analysis. Given that we are only working with US based data, we can choose to visualize our maps with a more "US-centric" perspective. Let's set our projection with this in mind:

1. Right click on **"Layers"** and go to **"properties"**
2. Select the **"coordinate systems"** tab
3. Go to **"Projected Coordinate Systems"**, **"Continental"**, **"North America"**, and choose **"USA Contiguous Albers Equal Area Conic USGS (http://desktop.arcgis.com/en/arcmap/latest/map/projections/albers-equal-area-conic.htm)"**



It's on the fly!

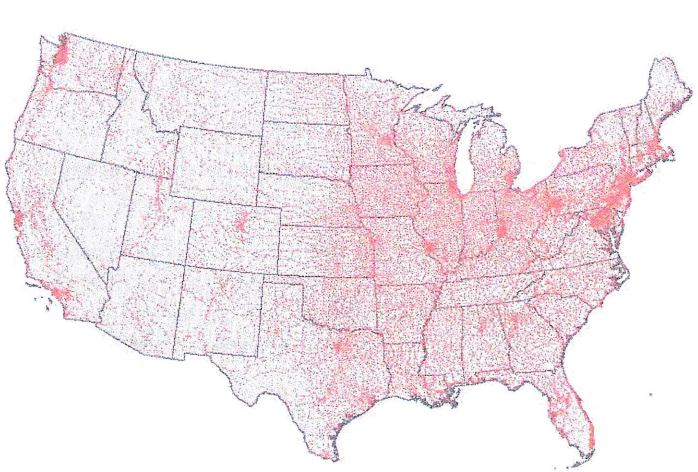
The software will warn you that you are projecting your datasets on the fly (note that it is not *reprojecting* the actual data, it is doing so only within the scope of this project space). If you want to perform spatial analysis, it is recommended that all layers in your project be reprojected to an appropriate coordinate system. More information on how to do this can be found here (https://support.esri.com/en/technical-article/000004886).

Order your layers

Vector layers are also referred to as "feature classes" in ESRI-Land. All GIS datasets can be added in this same way. Now drag each layer and re-order them. If you are familiar with Adobe Photoshop or Illustrator, you will recognize conceptual similarities with layering. What happens when layers are re-ordered? How does this dictate your strategy on building a single flattened map with multiple layers?

Challenge Exercise

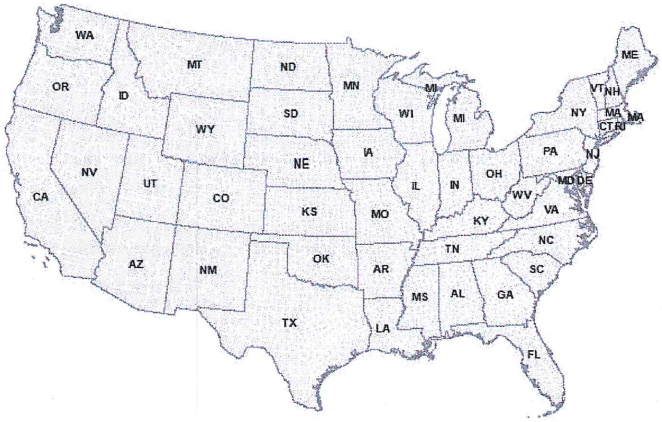
Modify your map by changing fill colors, outline colors, symbol sizes, symbol colors to make it look like this:

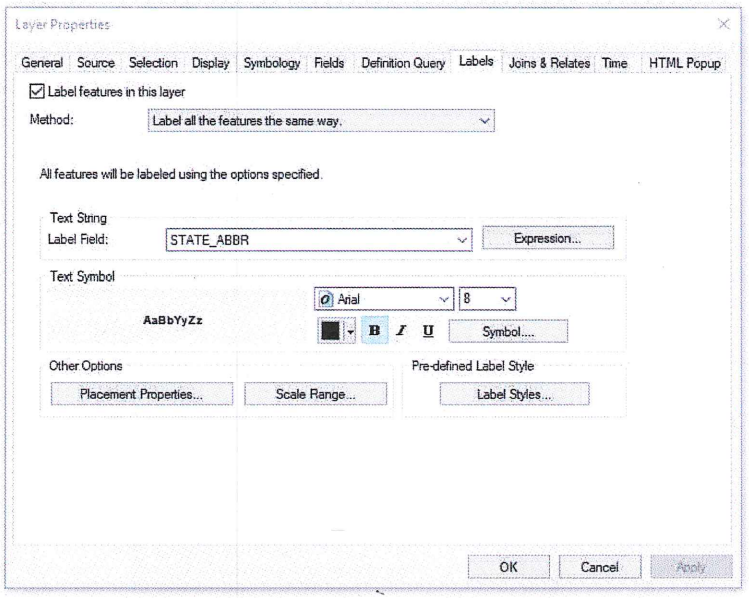


Symbolization

Outlines, fills, colors, weight, action! Here is where the artist in you comes out and the design phase of creating a map begins. Consider color choices: grayscale? Color schemes? Color hierarchy? Inevitably, you will find yourselves in the throes of ESRI's symbolization quagmire...

Labeling



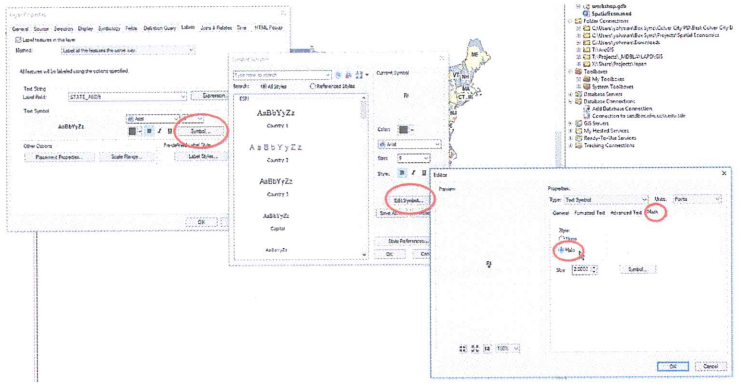


Map elements need labels at times. Consider what needs to be labeled, and what does not. Label sizes, fonts, weights, placement, colors are all things to consider for your map. Understand the relationship between labels, attributes, and layers.

Labels hard to read? Halo it!

Sometimes your labels may be hard to read, depending on what resides in the background. In this situation, you can add a white "halo" to your labels to make them "pop" some more. This feature is very, very hidden in ArcMap, but here is how to get to it:

1. Go the **Label** tab
2. Click **"Symbol"**
3. Click **"Edit symbol"**
4. Click **"Mask"**
5. Choose **"Halo"**



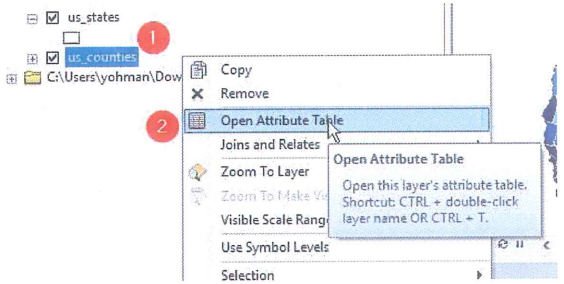
Attributes

Table

us_counties

OBJECTID *	Shape *	NAME	STATE_NAME	STATE_FIPS	CNTY_FIPS	FIPS	POP2010	POP1
1	Polygon	Autauga	Alabama	01	001	01001	54571	
2	Polygon	Baldwin	Alabama	01	003	01003	182265	
3	Polygon	Barbour	Alabama	01	005	01005	27457	
4	Polygon	Bibb	Alabama	01	007	01007	22915	
5	Polygon	Blount	Alabama	01	009	01009	57322	
6	Polygon	Bullock	Alabama	01	011	01011	10914	
7	Polygon	Butler	Alabama	01	013	01013	20947	
8	Polygon	Calhoun	Alabama	01	015	01015	118572	
9	Polygon	Chambers	Alabama	01	017	01017	34215	
10	Polygon	Cherokee	Alabama	01	019	01019	25989	
11	Polygon	Chilton	Alabama	01	021	01021	43643	
12	Polygon	Choctaw	Alabama	01	023	01023	13859	
13	Polygon	Clarke	Alabama	01	025	01025	25833	
14	Polygon	Clay	Alabama	01	027	01027	13932	
15	Polygon	Cleburne	Alabama	01	029	01029	14972	

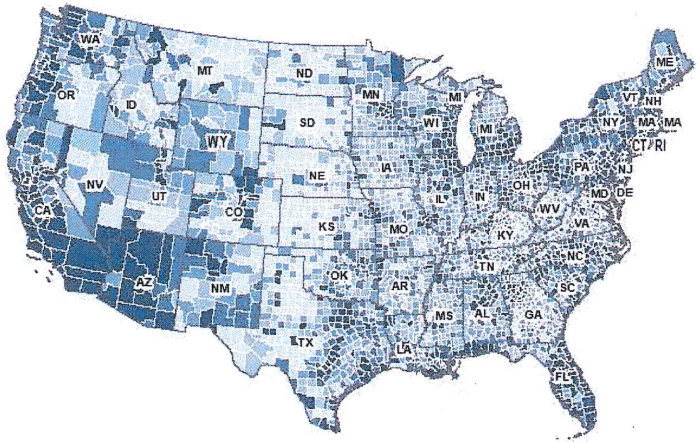
Every layer (feature class) comes with attributes. This is the all-important "information" part of geographic "information" systems mapping. Data in the attribute tables dictates what *can* get mapped. Open the attribute table of each layer (right click on the layer from the table of contents, **Open Attribute Table**):



Study how each row and column is tied to the mapped element. Questions we will answer include:

- What is the unique identifier for each row?
- What other attributes exist?
- What happens when you select a row on the attribute table?
- How do you sort elements?
- Can you build custom queries?
- Can you build graphs?

Choropleth Maps



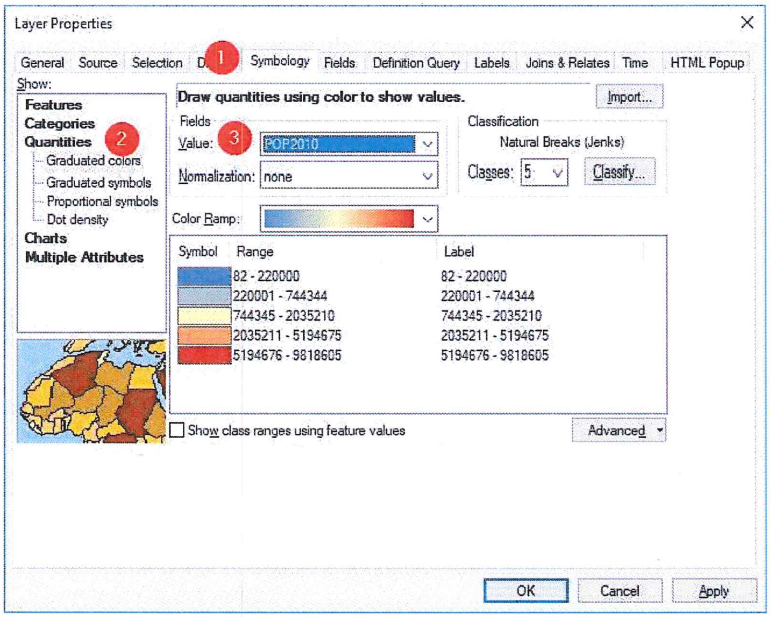
For this section, we will focus on creating a choropleth (which just means a colored map based on numerical data)!

When creating a choropleth the following needs to be considered:

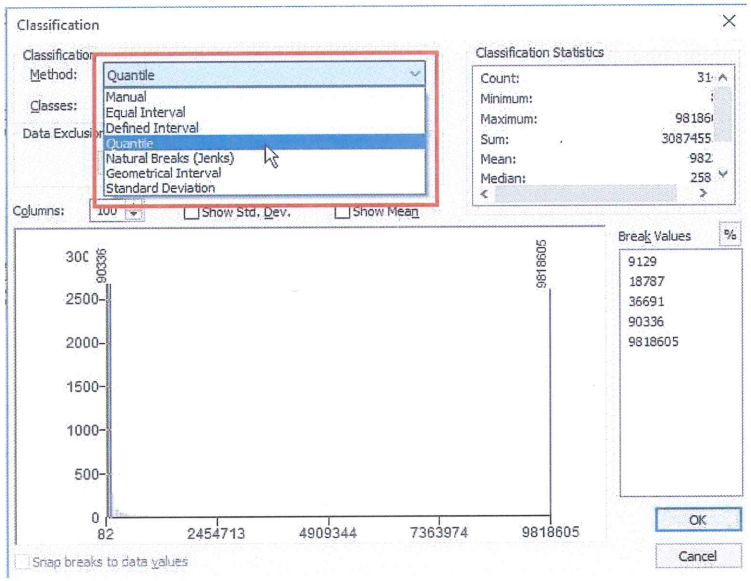
1. Is the data *choropleth-able*?
    1. Choropleths work best when representing data where boundaries are important
    2. Conversely, choropleths do not work well when attempting to show data where boundaries are NOT important/irrelevant
2. Do you have the data in the geographic scale you wish to map it at?
3. Can you connect the data to an existing layer?
4. Which coloring style best represents your data?
    1. If your information is continuous then use a single color gradient
    2. If your information has a positive or negative range, use an opposite color scheme

To create a choropleth map, follow these steps:

1. Right click on **us_counties** and go to **properties** (or just double click it!)
2. Select the **Symbology** tab, click on **Quantities**, and select **POP2010** for the Value field.



Now click on the **Classify** button. There are several methods to choose from. Look at the following documentation (http://pro.arcgis.com/en/pro-app/help/mapping/layer-properties/data-classification-methods.htm) to determine which method is best suited for your data.



Part 2: Working with spatial data

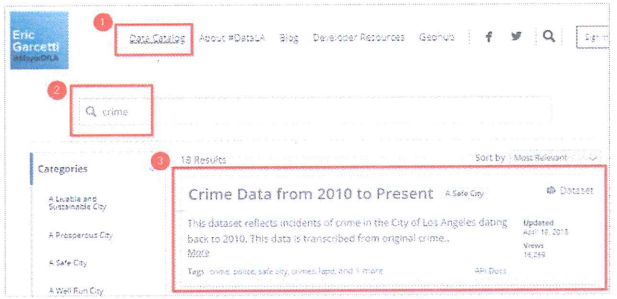
Acquiring data

The open data movement has made more and more data available for academics to download and use for their research. But how can we map this data? This workshop will take you through the process of acquiring data from the Los Angeles Open Data portal (https://data.lacity.org/) and visualizing it on ArcGIS for further analysis.

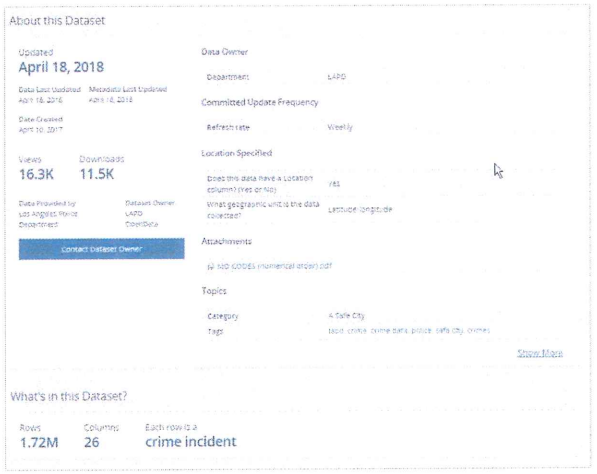
Los Angeles Open Data portal (https://data.lacity.org/)

The Los Angeles Open Data Portal

Search for crime data

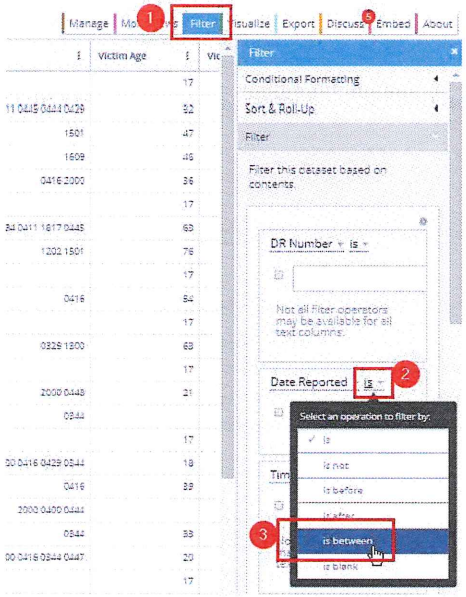


Inspect the data

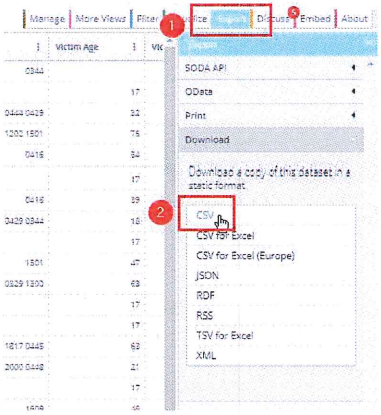


Almost 2 million records! Let's filter it down to something more manageable.

Now add the filter to narrow down the data to one month:



Export the data



Cleaning up those coordinates

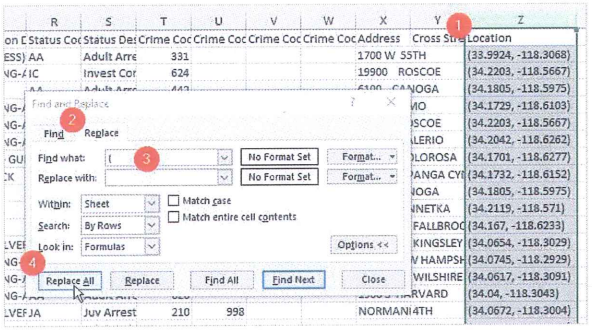
Open the downloaded data in Excel. Scroll to the right until you see the Location column.

V	W	X	Y	Z
oc Crime Coc	Crime Coc	Address	Cross Stre	Location
		1700 W	55TH	(33.9924, -118.3068)
		19900	ROSCOE	(34.2203, -118.5667)
		6100	CANOGA	(34.1805, -118.5975)
		5600	COMO	(34.1729, -118.6103)
		19900	ROSCOE	(34.2203, -118.5667)
		22900	VALERIO	(34.2042, -118.6262)
98		23000	DOLOROSA	(34.1701, -118.6277)
		6100	TOPANGA CYN	(34.1732, -118.6152)

Hmm, that's strange, the latitude and longitude columns are in the same column! ArcGIS does not like this. Let's clean this up.

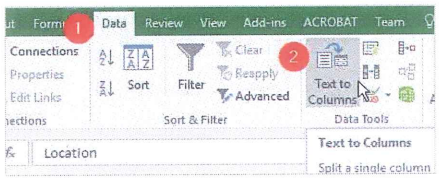
First, find and replace the brackets.

1. Select the Location column
2. Bring up the find and replace tool (ctrl-h)
3. For "Find what", enter an open bracket "("
4. Click Replace All

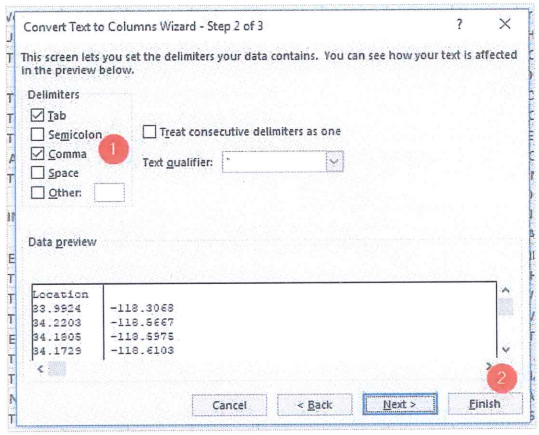


Repeat for the closing bracket.

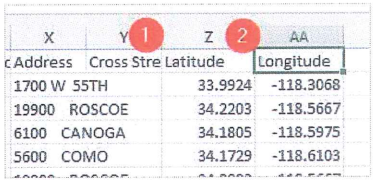
Split the column into two:



Choose, delimited, check the "Comma" box, and finish.

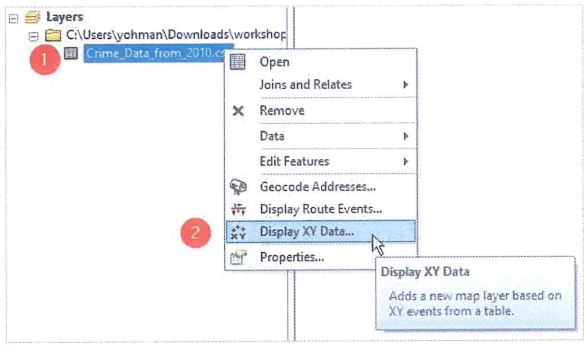


Rename the column headers to Latitude and Longitude

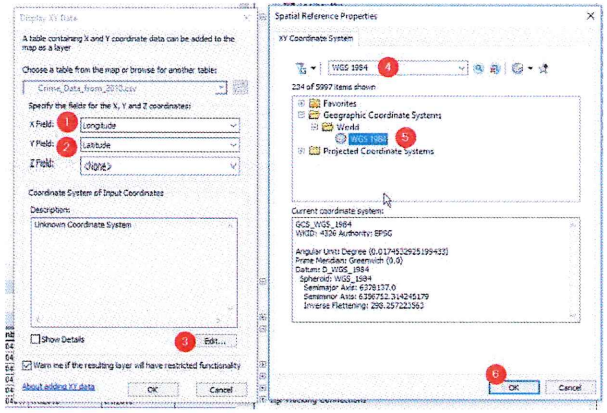


Let's map it!

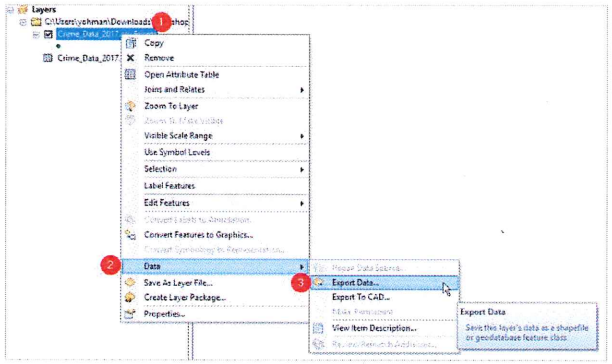
Start a brand new ArcMap project and add the csv file (remember the **Add Data** button?). Right click on the csv file and **Display XY Data**.

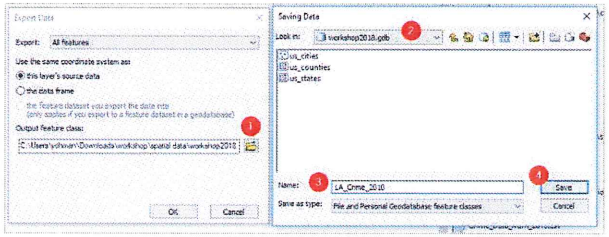


1. Set X to **Longitude**
2. Set Y to **Latitude**
3. Click **Edit** for the coordinate system
4. Enter "**WGS 1984**" in the search box
5. Choose **WGS 1984**



Now save your new layer as a shapefile, or geodatabase:



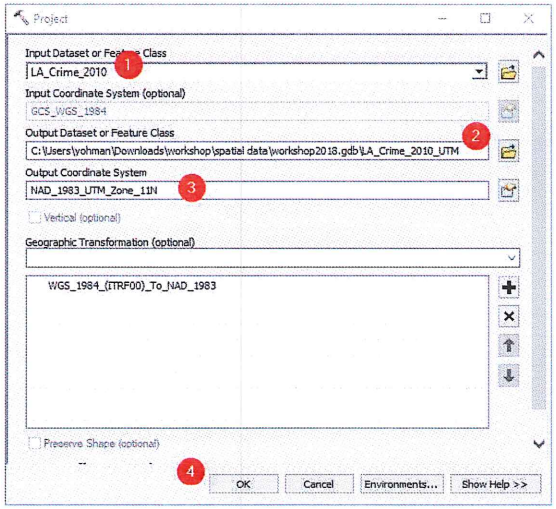


Project the data

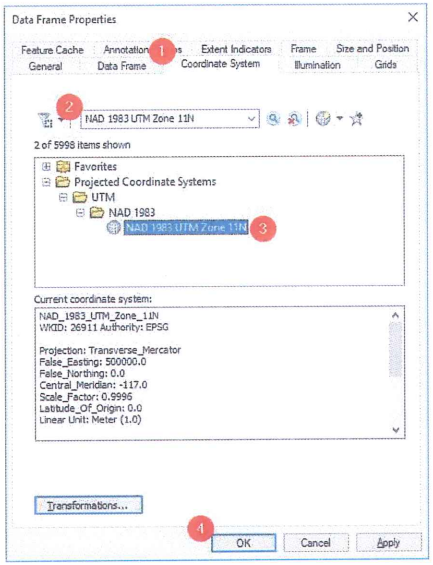
Our data is currently in a geographic coordinate system (WGS1984). Let's change this to a projected coordinate system. The UTM zone for Los Angeles is **UTM Zone 11N**.

Click on the search tool

Type "project" and click on **Project (Data Management)**



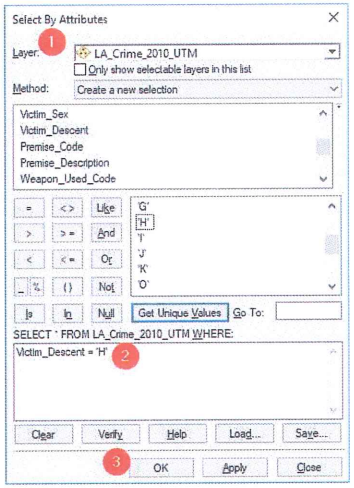
Now, set the projection of the data frame. Right click on **Layers**, and go to **properties**. Then, set the coordinate system to **NAD 1983 UTM Zone 11N**



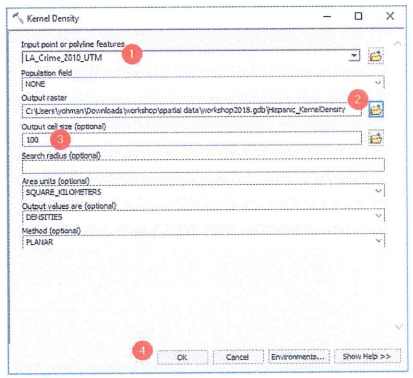
Hot spots?

Let's find crime hot spots by race. Select incidents where the person arrested was classified as Hispanic (H). In the menu bar, go to **Selection**, **Select by attribute.** Enter the following SQL statement:

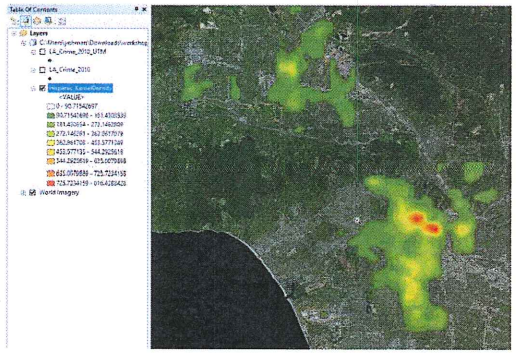
```
Victim_Decent = 'H'
```



Now perform a kernel density (http://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-kernel-density-works.htm) to visualize the density of Hispanic arrests in Los Angeles. In the search box, enter "**kernel**" and click on the **Kernel Density (Spatial Analyst)** tool. Enter the four boxes as shown below:

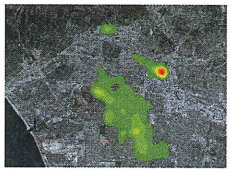


Add a **basemap**, and change the **symbology** to make the visual more powerful:



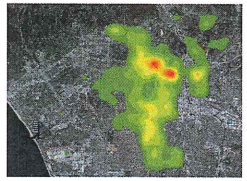
Repeat the process for other race categories:

Black



White

Hispanic



Data in this tutorial

- LAPD Crime Data 2010 (csv) (http://gis.ucla.edu/geodata/dataset/lapd-crime-data-2010)

Thanks for attending the workshop!

Please fill out the feedback form to see what we can improve upon:

Post workshop survey (https://bit.ly/2RLCsBY)

For more learning you can check out ESRI's tutorials:

http://www.esri.com/training/main/training-catalog/course-recommendations
(http://www.esri.com/training/main/training-catalog/course-recommendations)

Edit (https://sandbox.idre.ucla.edu/sandbox/wp-admin/post.php?post=3356&action=edit)