**Lab 9: Descriptive Spatial Statistics and Point Pattern Analysis**

**Learning Objective**

In this lab, we will use a tornado touchdown points database which spans from 1950 to 2018 to introduce you to some techniques of measuring geographic distributions. Through the use of yearly and monthly mean centers and standard deviation ellipses, well explore how tornado touchdowns are distributed and then use the tracking analyst to visualize how these change over time. Next, we’ll quantify the distribution of ad hoc subsets of the data using point pattern analysis.

**What you need to submit**

**Lab 9: Answer Sheet**  
Name:

**Question 1**:  
Repeat steps 4 – 7 using the monthly\_SDellipse layer to create the temporal layer. Export the animation as yournamemonthly\_ttmov and be sure it runs properly. Upload the .avi file (just the monthly ellipses) on Blackboard.

What does this animation tell you about seasonal changes in tornado touchdown locations? Is there any obvious trend?

**Question 2**:  
Fill in the details on the following table (based on point extent): In the C/R/D column, indicate whether there is significant Clustering (z-value less than -1.96), significant Dispersion (z-value greater than 1.96) or neither (R for random) using an alpha of .05.

| **Place & Time** | **NNR** | **Z-Score** | **C/R/D** |
| --- | --- | --- | --- |
| Kansas 2005 |  |  |  |
| Washington 2005 |  |  |  |
| Alabama 2005 |  |  |  |
| USA May 2007 |  |  |  |
| USA Nov 2007 |  |  |  |

**Question 3**:  
Describe the differences in the point patterns in Kansas, Washington, and Alabama in 2005. Which (if any) of the patterns was found to be significantly different from the random distribution? Assume alpha level = .05 (the z-value associated with 95% confidence interval is +/- 1.96).

**Question 4**:  
What is the total area of the conterminous US in square miles?: \_\_\_\_\_ SQ MI.

### **Gathering and formatting data**

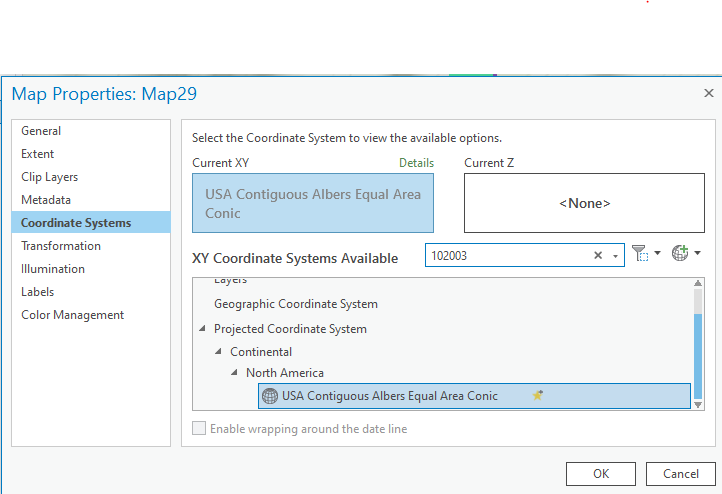
We’ll start by downloading data from the [NOAA Storm Prediction Center Severe GIS page](https://www.spc.noaa.gov/gis/svrgis/). There are a number of datasets here, but what we’re after here is the original csv data, [here](https://www.spc.noaa.gov/wcm/#data), under the Severe Weather Database Files (1950-2017) heading.

**1)** Download (at minimum) the 2005-2007\_torn.csv file, and place them in their own lab 9 folder.

* Files are listed at <https://www.spc.noaa.gov/wcm/#data>
* Be sure to [read the user guide](https://www.spc.noaa.gov/wcm/data/SPC_severe_database_description.pdf) so that you understand what it is you’ve actually downloaded.

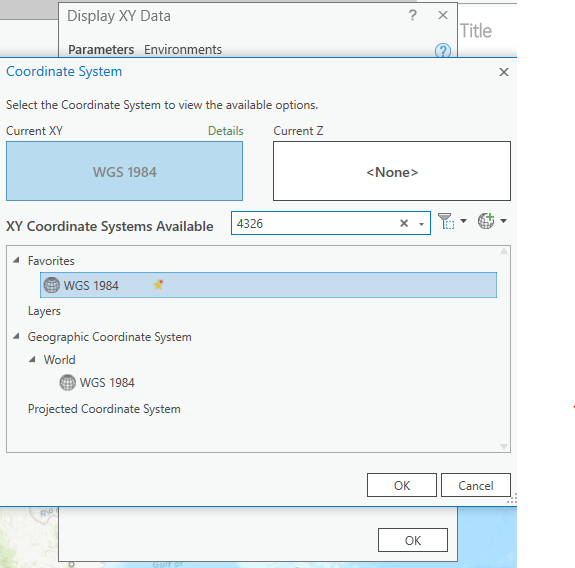
**2)** Set up and import the data

* Before we import anything, let`s set the projection of the dataframe to an equal area projection so that we can accurately visualize these within the context of dispersion.
* Insert a blank dataframe and rename it as Lab9.
* R-click the dataframe title and go to properties > Coordinate Systems and search for WKID: 102003



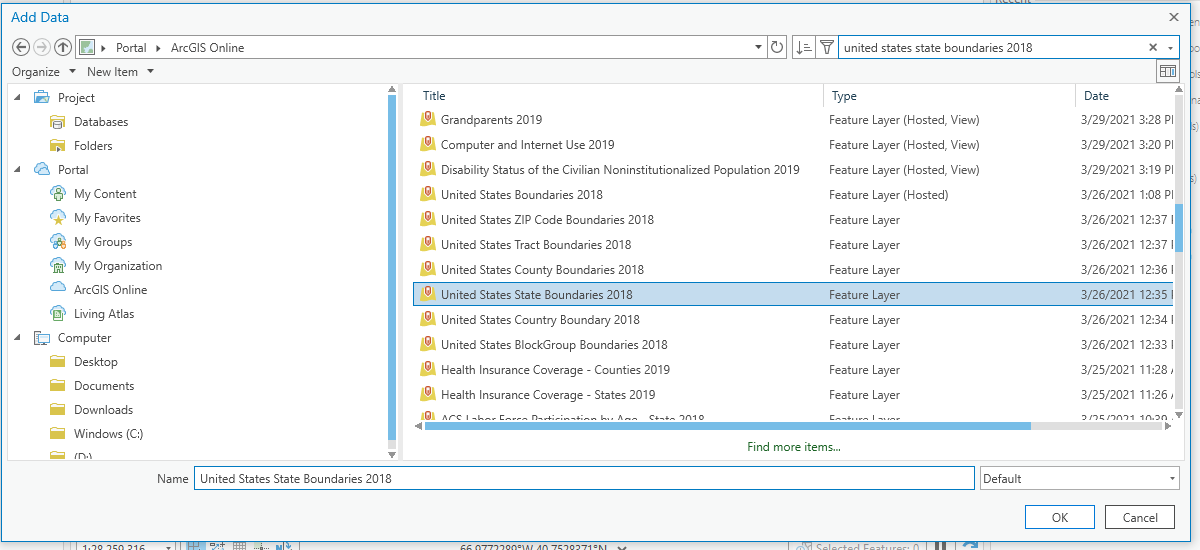
### Importing the tornado data

* Let’s next bring in the csv. Drag the csv file from the Catalog panel into the dataframe.
* Right click on the csv and “Display XY Data”.
  + Set X Field to slon
  + Set Y Field to slat, and then edit the coordinate system.
  + Set the spatial reference of the data to WKID: 4326
  + Save the output shapefile into your lab#9 folder.
  + OK your way through the tool.



### Adding ancillary data

* Next add a state boundaries shapefile from your provider of choice.
  + The quickest way to acquire this is to use the arrow dropdown next to add data and add data from ArcGIS Online (left side).
  + You will need to sign in to an ArcGIS account in the upper right hand corner.
  + Search for “United States State boundaries 2018”, and make sure the layer you add is a feature layer.
  + Use select tool (e.g. polygon) to select the lower 48 states (and DC) from the layer that loads in.
  + R-click the USA\_State> Selection > Make layer from selected features and then rename the new layer as *lower48.shp*.
* If this fails for some reason, you can grab [state boundaries from TIGER](https://www.census.gov/cgi-bin/geo/shapefiles/index.php).

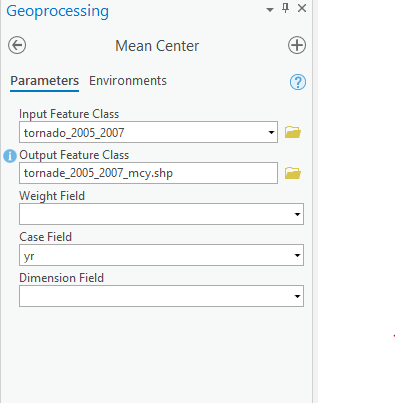


# Part One: Visualizing the distributions

### **Calculating and Visualizing Descriptive Spatial Statistics**

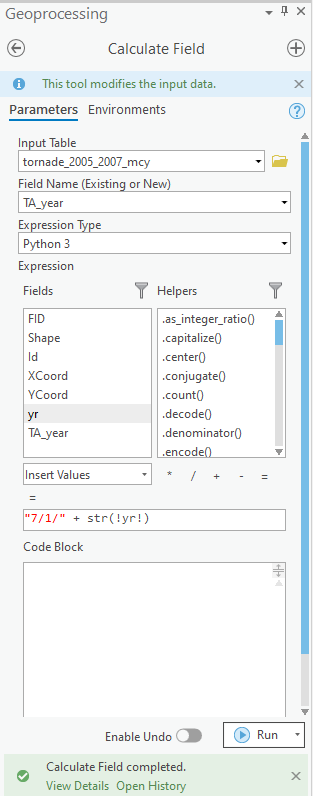
**1)** Open the **Mean Center** tool.

* In the Mean Center dialog box, select *Torn\_yyyy\_yyyy.shp* as your input feature class.
* name the output feature class *Torn\_yyyy\_yyyy\_mcy.shp*.
* set the case field to yr.

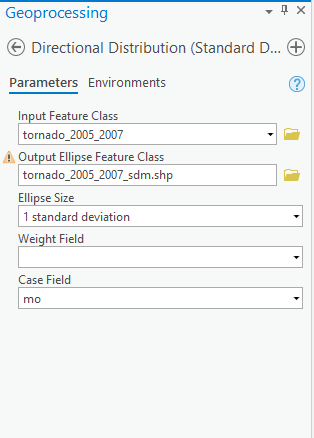


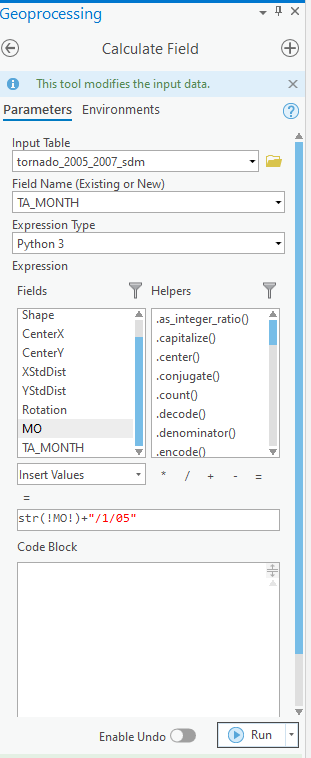
**2)** Open the attribute table for Torn\_yyyy\_yyyy\_mcy. Notice that you have n records, one for each year of data you imported. In order to use tracking analysis, the extension ArcMap uses to create videos, we need to **create a new field**, and populate it using **Field Calculator** to create a readable by the tool. We’ll assign a single day, for example, “7/1/1950” to represent 1950

* In the table options, **Add Field** to create a new field named *TA\_YEAR*
  + Set the type to Text
  + Set the length to 32 byte
* Use the **Calculate Field** to calculate new field as the following image shows:



**3)** Calculate monthly standard deviation ellipses using the **Directional Distribution (Standard Deviational Ellipse)** tool.



* Select *Torn\_yyyy\_yyyy.shp* as your input feature class.
* Name the output feature class *Torn\_yyyy\_yyyy\_sdm.shp*
* leave the ellipse size to the default
* set the Case Field to mo
* Click OK.
* Add a new field named *TA\_MONTH* (Text type, 32 byte length)
* assign the date as follows.
* 

In ArcMap, we can use tracking Analyst Tool to visualize the point track based on different years. But Pro has removed Tracking Analyst Tool. So this part will stop here. If you are interested in finding an alternative way doing that in ArcGIS Pro, Please feel free to tell me.

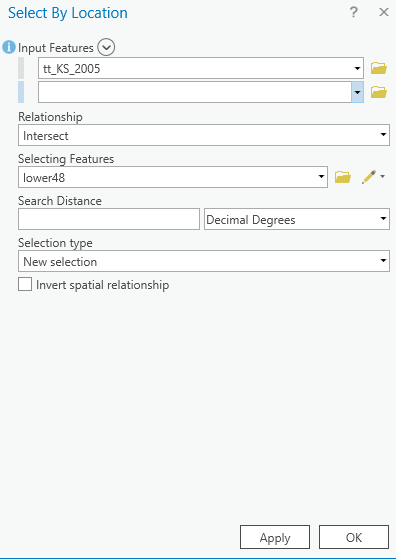
Reference of removed tools by ArcGIS Pro:

<https://pro.arcgis.com/en/pro-app/latest/tool-reference/appendices/unavailable-tools.htm>

# Part Two: Nearest Neighbor Analysis

**1)** Select tornado touchdowns within Kansas in 2005.

* Starting with *lower48* and the *Torn\_yyyy\_yyyy.shp*, **export** the tornado touchdown data for 2004
  + select by attributes, input layer *Torn\_yyyy\_yyyy.shp*, “YEAR” = ‘2005’, R-Click layer, Selection> Make layer from selected features, and call it *tt\_KS\_2005.shp*, add it to the map
* Turn off the *Torn\_yyyy\_yyyy* and *tt\_KS\_2005* layers temporarily and use the select tool to select Kansas from lower48 layer (alternatively you can use the Select by attributes tool)
* Go to Selection > Select by Location, select features from *tt\_KS\_2005* that intersect the features in *lower48*.
* Click Apply.
* Turn the *tt\_KS\_2005* layer back on to verify that only touchdown points for the year 2005 in Kansas are highlighted



**2)** Using the **Average Nearest Neighbor** tool in the Spatial Statistics Tools | Analyzing Patterns, calculate the Nearest Neighbor Index for Kansas in 2005.

* Input feature class: *tt\_KS\_2005*, use Euclidean distance and check Generate Report
* To see the report, view details > open the ‘NearestNeighbor\_Result.html’ in the Report File window
* Fill in the appropriate information in the table (question #2 on the answer sheet)

**3)** Calculate Nearest Indices for Washington (the state) and Alabama in 2005

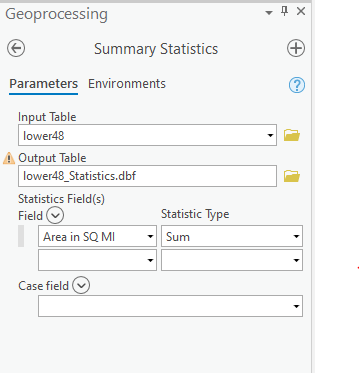
* Repeat parts of step 1 and step 2 above for both Washington and Alabama
* Fill in the table in question #2 and answer question #3 on the answer sheet

**4)** Calculating Nearest Neighbor Indices for two months (May and November) across the US in 2007

* Select the touchdown points that occurred in May of 2007 (select by attributes) and export data, naming the output *tt\_May07*
* Do the same for November and name the exported dataset appropriately
* Repeat step 2 to calculate the NNI for the two months
* Fill in the table in question #2 on the answer sheet

**5)** Calculate the total area of the 48 conterminous states and DC

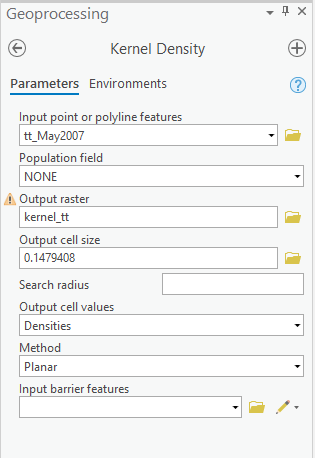
* Open the attribute table of lower 48 and observe the column names. We can see that there is one column named Area in SQ MI(Square Miles).
* Open Summary Statistics tool to calculate the total area of Area in SQ MI and answer Q#4.



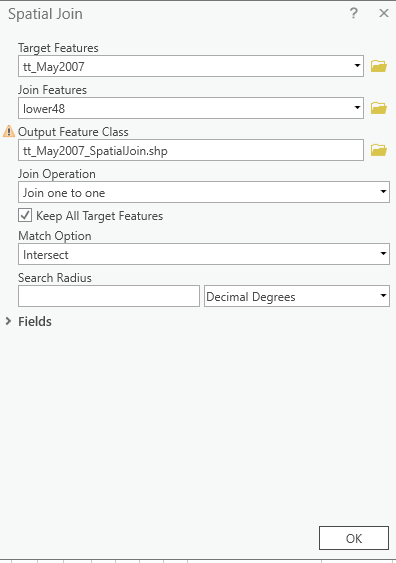
# Part Three: Kernel Estimation (KE)

Kernel estimation is to identify cluster of hot spots. Let`s first use one month of tornado point features as an example.

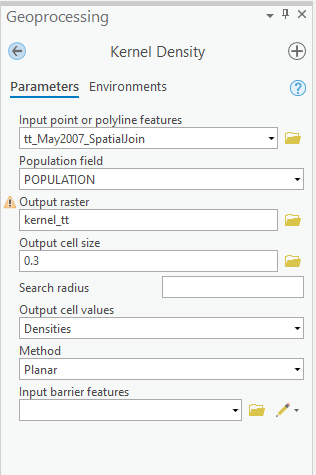
* Open Kernel Density, input point or polyline features is your tornado point features, population field is NONE, save your output into Lab#9 folder. The cell size can be default. Detailed settlings are as follows:



* Attach your shot screen of kernel density layer to the report.
* Next, let us join population data to the tornado point layer and use Kernel Estimation again.
* Select one tornado layer you want to target, R-Click > Join and Relates > Spatial Join. Join features should be lower48 layer.



* Click Ok and open the attribute table of joined layer and check if POPULATION column is there.
* Go back to Kernel Density Tool and re-run the tool with joined layer.



* Attach your result layer in the report.