

EXERCISE 1

Spatial Statistics: Measuring Geographic Distributions



Introduction

In this exercise, you will become familiar with the spatial statistics toolset by exploring the spatial distribution of your data. Measuring the geographic distribution of your data will enable you to identify any spatial patterns or clusters that exist.

Objectives

- Learn to use the Spatial Statistics Toolset for measuring geographic distributions
- Examine and interpret the locations and distribution of spatial data
- Exploring spatial relationships within your spatial data

Prerequisites

- ArcGIS Pro installed on your machine



USDA Non-Discrimination Statement

In accordance with Federal civil rights law and U.S. Department of Agriculture (USDA) civil rights regulations and policies, the USDA, its Agencies, offices, and employees, and institutions participating in or administering USDA programs are prohibited from discriminating based on race, color, national origin, religion, sex, gender identity (including gender expression), sexual orientation, disability, age, marital status, family/parental status, income derived from a public assistance program, political beliefs, or reprisal or retaliation for prior civil rights activity, in any program or activity conducted or funded by USDA (not all bases apply to all programs). Remedies and complaint filing deadlines vary by program or incident.

Persons with disabilities who require alternative means of communication for program information (e.g., Braille, large print, audiotope, American Sign Language, etc.) should contact the responsible Agency or USDA's TARGET Center at (202) 720-2600 (voice and TTY) or contact USDA through the Federal Relay Service at (800) 877-8339. Additionally, program information may be made available in languages other than English.

To file a program discrimination complaint, complete the USDA Program Discrimination Complaint Form, AD-3027, found online at [How to File a Program Discrimination Complaint](#) and at any USDA office or write a letter addressed to USDA and provide in the letter all of the information requested in the form. To request a copy of the complaint form, call (866) 632-9992. Submit your completed form or letter to USDA by: (1) mail: U.S. Department of Agriculture, Office of the Assistant Secretary for Civil Rights, 1400 Independence Avenue, SW, Washington, D.C. 20250-9410; (2) fax: (202) 690-7442; or (3) email: program.intake@usda.gov.

USDA is an equal opportunity provider, employer, and lender.

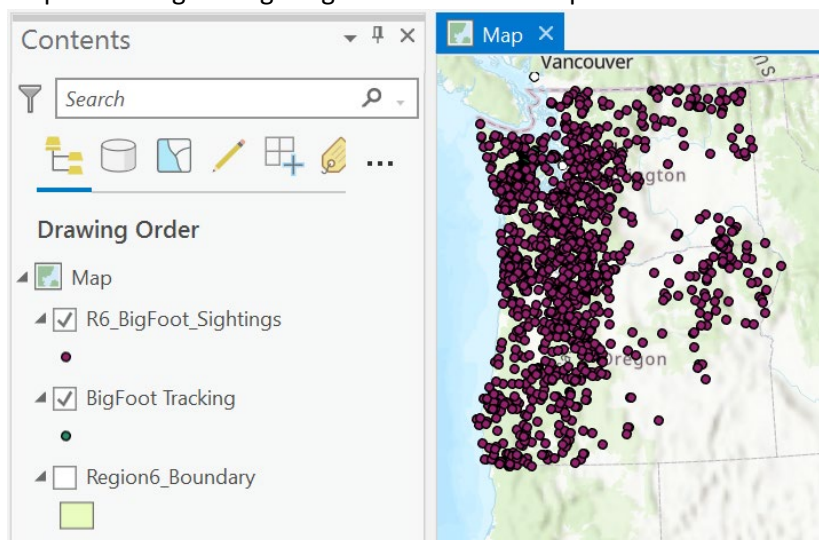
Table of Contents

Part 1: Explore the Project Data	4
A. Open the Exercise 1 project file.....	4
B. Calculate the total yearly count and averages for Bigfoot sightings	4
Part 2: Explore Spatial Statistics Tools.....	6
A. Open the Geoprocessing Pane and explore the Spatial Statistics Toolset.	6
Part 3: Explore the Spatial Distribution of Data	8
A. Locate the average lat/long of all BigFoot sightings.....	8
B. Locate the average location of BigFoot sightings based on the Class of the sightings.....	9
C. Locate the average location of BigFoot sightings based on the Year of the sightings.	10
D. Find the most centrally located BigFoot sighting.	13
E. Determine how BigFoot sightings are dispersed around the mean.	14
F. Determine how BigFoot sightings are dispersed around the mean over time.....	16
G. What is the distribution and orientation of recent BigFoot sightings?.....	19
H. Monitor the seasonal movement of a radio collared BigFoot.	21

Part 1: Explore the Project Data

A. Open the Exercise 1 project file

1. Open Windows Explorer and navigate to your course folder
....\ArcGISProSpatialStatistics\Data\Exercise_1
 - i. Open the **Exercise_1** ArcGIS Project File.
2. Inspect the Bigfoot sighting locations in the map.



B. Calculate the total yearly count and averages for Bigfoot sightings

1. In the Contents pane, right-click the **R6_BigFoot_Sightings** layer then select **Attribute Table**.

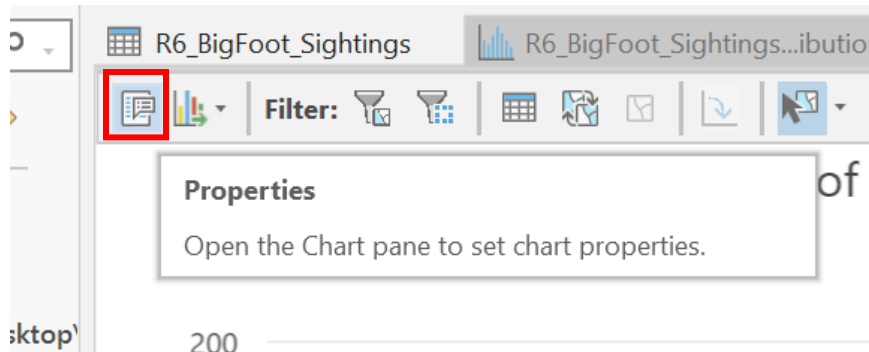
R6_BigFoot_Sightings									
Field:	Selection:								
OBJECTID *	Shape *	Class	Year	Number	Date	Latitude	Longitude	Weight	
1	Point ZM	B	2008	1	1/1/2008	-121.6885	45.23679	2	
2	Point ZM	B	2008	1	1/1/2008	-118.2287	48.7375	2	
3	Point ZM	B	2008	1	1/1/2008	-123.3687	46.99385	2	

The attribute table contains information on the type or class of the sighting, the year and how many were seen and/or heard at each location. Their classes and weights include:

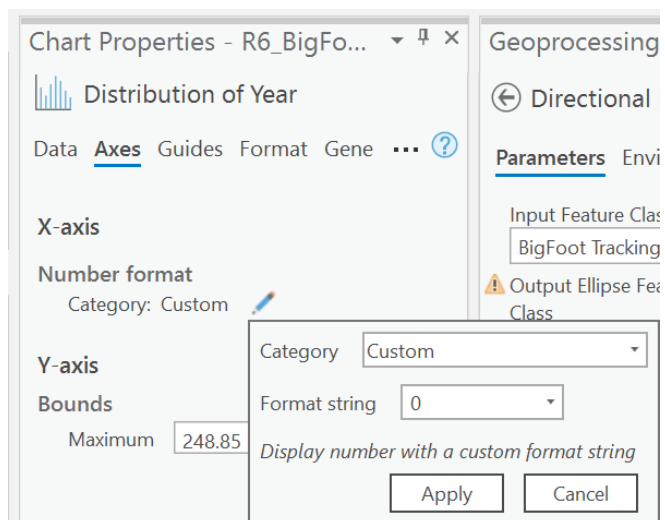
Type	Class	Weight
Physical	A	3
Audio	B	2
Footprint	C	1

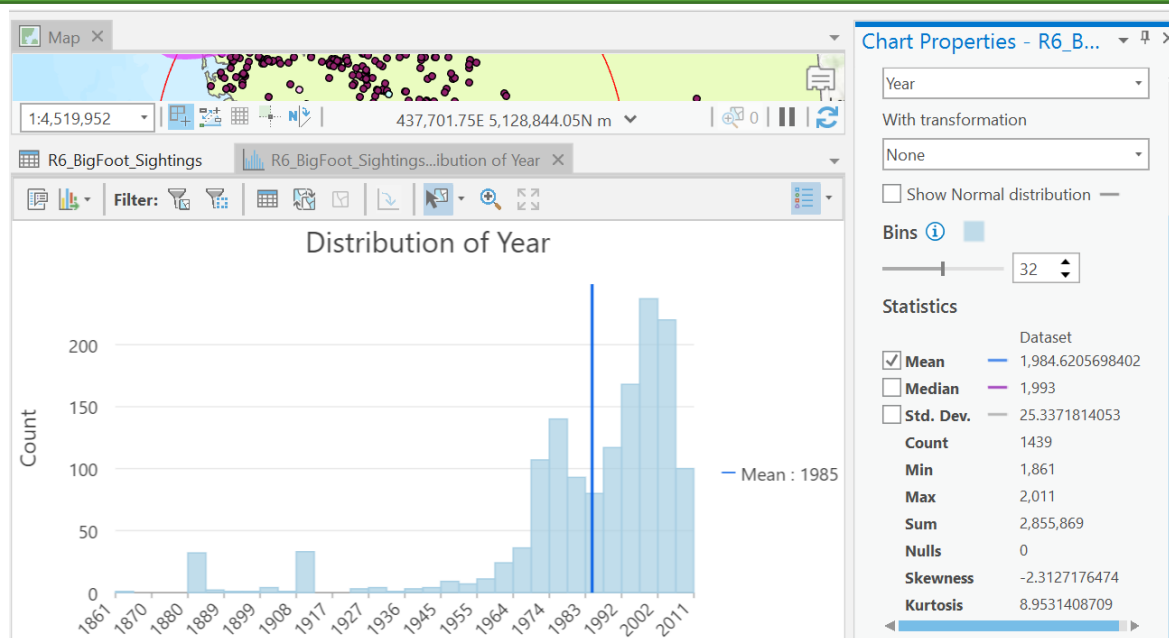
Descriptive Statistics in ArcGIS are information that can be mathematically calculated based on data values. These tools quantify or identify characteristics of features and answer questions like, where is the center or how are the features distributed around the center? They help us summarize the spatial characteristics for a set of features. You can easily get descriptive statistics about your data by opening the attribute table, right clicking a column heading... and choosing Statistics

2. In the attribute table, right-click the field heading **Year** then select **Statistics**. If the Chart Properties window doesn't appear, click the **Properties** button in the upper left corner of the attributes window.



3. To have the years in the chart displayed as whole numbers, select **Axes** in the Chart Properties.
 - i. Set the **X-axis Number format** to **Custom** and the **Format string** to **0**.





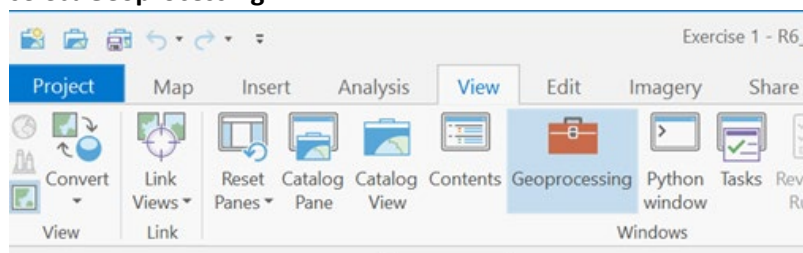
- ii. QUESTION - What is the total amount of sightings for all years?
 - iii. QUESTION - Is there a trend such as an increase or decrease of sightings over time?
4. **Close the Statistics and Attribute windows.**

Part 2: Explore Spatial Statistics Tools

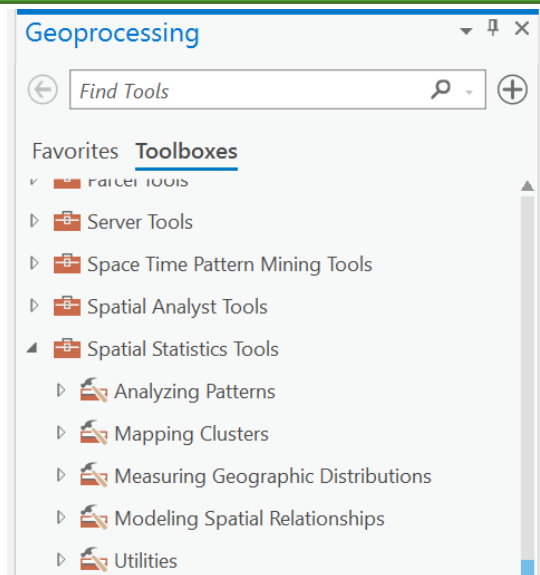
ArcGIS Pro has statistical tools for analyzing spatial distributions, patterns, processes, and relationships. These tools incorporate space (proximity, area, connectivity, and/or other spatial relationships) directly into their mathematics.

A. Open the Geoprocessing Pane and explore the Spatial Statistics Toolset.

1. If the Geoprocessing pane isn't already open, go to the **View** ribbon (Windows section) and select **Geoprocessing**.



2. In the Geoprocessing pane, click the **Toolboxes** tab.
3. Scroll down to the **Spatial Statistics Tools** and expand the toolbox.

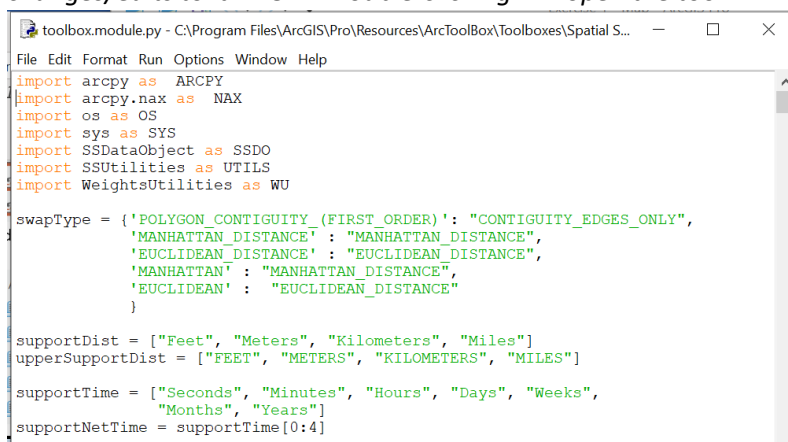


4. Expand the **Analyzing Patterns** toolset. Notice the Python Script Symbol located next to each tool.



NOTE: Most of the tools in the Spatial Statistics toolbox are actually delivered as Python Scripts.

5. Right-Click the **Average Nearest Neighbor** tool and choose **Edit** from the available menu; this will open up the underlying code that was written in Python allowing you to make changes/edits to it. **NOTE:** Double-clicking will open the tool.



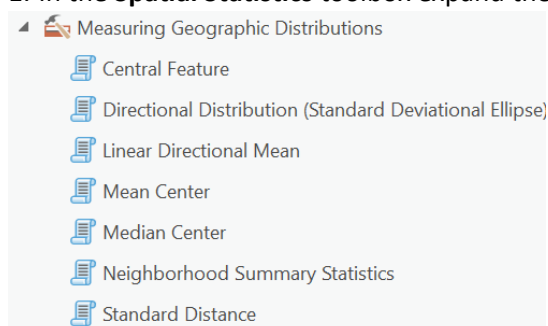
6. Close the Python script. We will not be editing the scripts, but if you did you would then have to Save As a new filename.

Part 3: Explore the Spatial Distribution of Data

*The analysis environment has already been setup for each exercise in this course but usually you must do it before using any geoprocessing tools. To open the analysis environment window, go to the **Analysis** tab (**Geoprocessing** group) and click the **Environments** button. In the analysis environment you can specify a working directory, establish the processing extent, and select an output coordinate system among other things.*

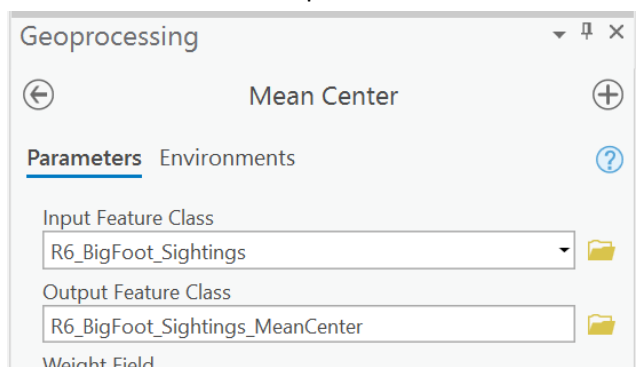
A. Locate the average lat/long of all BigFoot sightings.

1. In the **Spatial Statistics** toolbox expand the **Measuring Geographic Distributions** toolset.

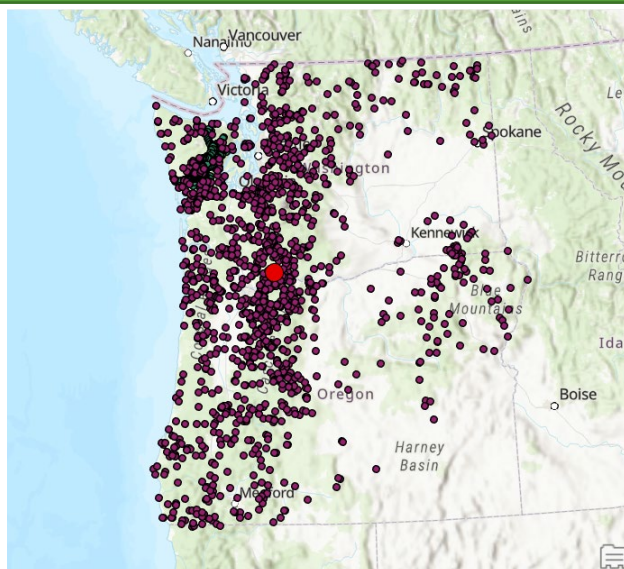


2. Single-click **Mean Center**. The Mean Center tool identifies the geographic center of the features in a dataset.

- i. For **Input Feature Class**: select **R6_BigFoot_Sightings** from the dropdown. This layer includes BigFoot sightings from 1861-2011.
- ii. Confirm that the Output Feature Class name reads: **R6_BigFoot_Sightings_MeanCenter**.



3. Accept the other defaults and click **Run**.
4. In the Contents pane, *change the symbology* of **R6_BigFoot_Sightings_MeanCenter** to a **large red dot**.



The average center of the distribution of all sightings is located where you see the red point. This indicates that the mean center of the sightings is west of the center of the study area and that the sightings are clustered or closer together, to the west of the study area.

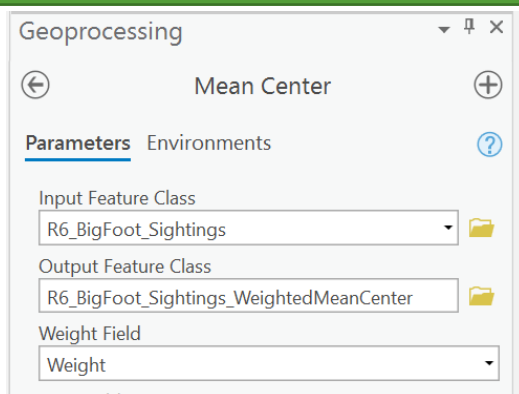
You can use the mean center tool to find the center of features based on an attribute value by choosing an attribute to use as a weight field. Let's run the tool again using the Class of the sightings as a weight.

Type	Class	Weight
Physical	A	3
Audio	B	2
Footprint	C	1

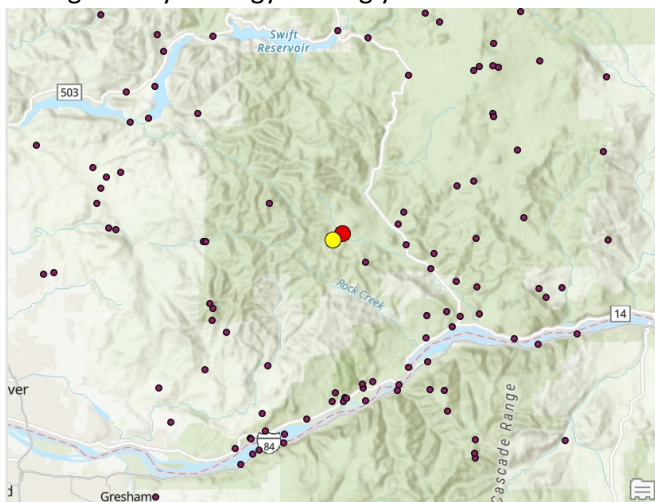
B. Locate the average location of BigFoot sightings based on the Class of the sightings.

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

- For *Input Feature Class*: select **R6_BigFoot_Sightings** from the dropdown.
- For *Output Feature Class*: type **R6_BigFoot_Sightings_WeightedMeanCenter**.
- For the *optional Weight Field*, choose **Weight**. The Weight field contains values that correspond with the Class of the Sightings.
- Accept the remaining defaults and click **Run**.



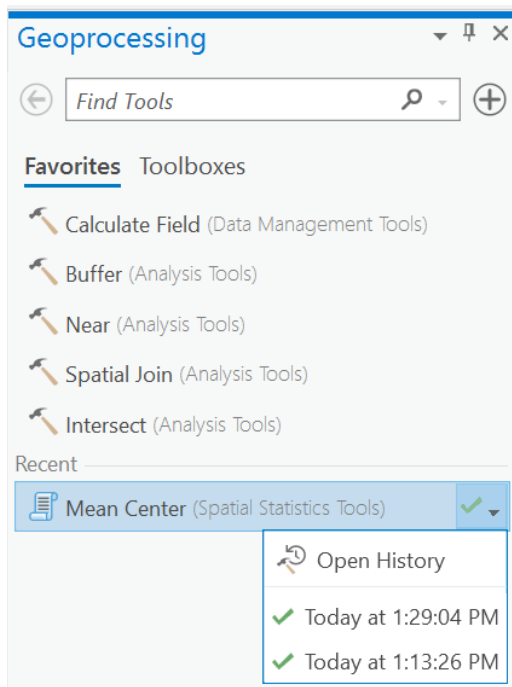
2. Change the symbology to a big yellow dot.



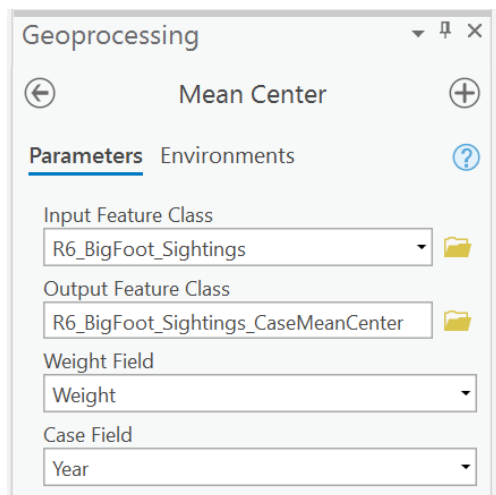
Notice the location of the weighted mean center is very close to the location of the mean center. This tells us that there are no attribute values that vary too much from the average attribute value; there are no significant outliers in this case.

C. Locate the average location of BigFoot sightings based on the Year of the sightings.

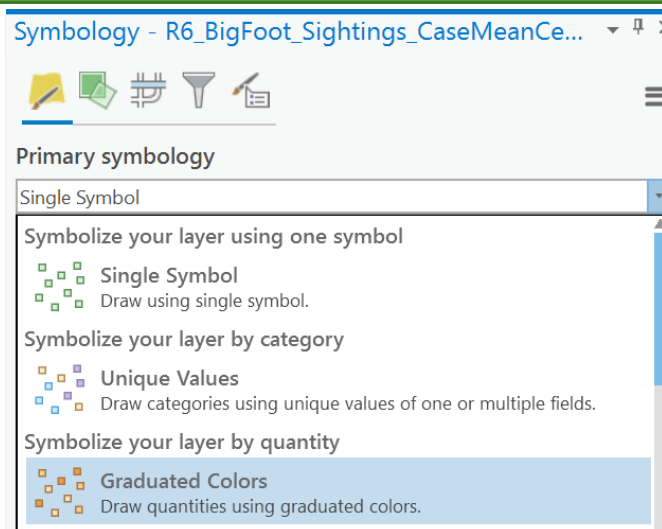
1. Now let's run the Mean Center tool again using a case field. You can reopen your previous executions of a tool by going to the **Favorites** tab of the Geoprocessing pane.
 - i. In the Recent section, click the dropdown arrow next to Mean Center then select the most recent time you ran the tool.



2. The tool reopens with the parameters set up exactly as before. We only need to make the following changes:
 - i. Update *Output Feature Class* to: **R6_BigFootSightings_CaseMeanCenter**.
 - ii. For *Case Field*: select **Year**. The Input & Weight Fields will remain the same as before.



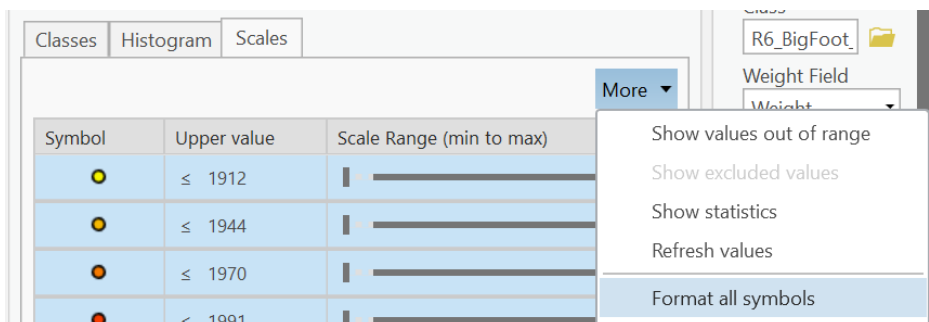
3. Click **Run**. Your output should have a separate point representing the average location for each year.
4. Change the symbology to display each year differently.
 - i. In the *Contents pane*, right-click **R6_BigFoot_Sightings_CaseMeanCenter** then select **Symbology**.
 - ii. In the Symbology pane, go to the Primary symbology menu and select **Graduated Colors**.



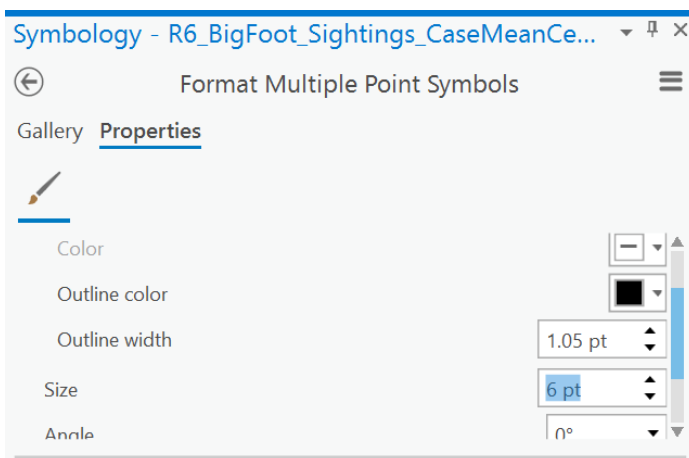
iii. Set the *Field* to **Year**.



iv. Click the **More** menu next to the symbols then select **Format All Symbols**.

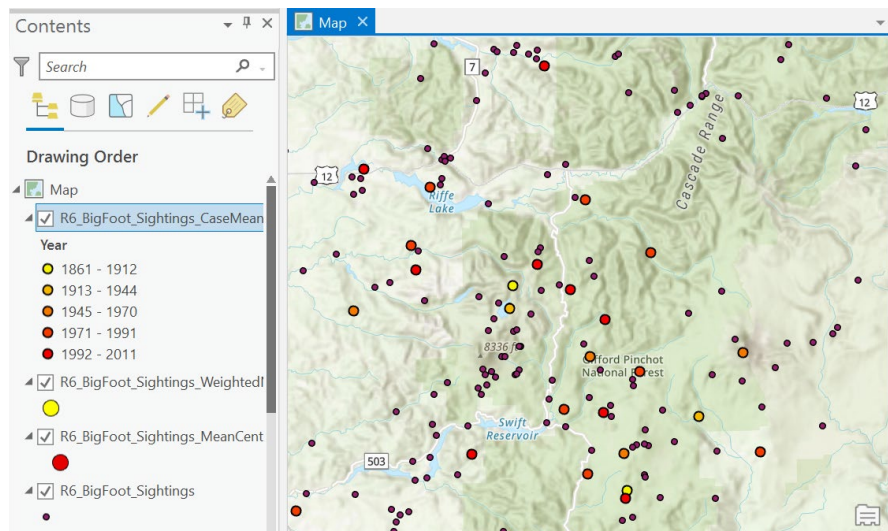


v. Set the symbol **Size** to **6**.



vi. Click **Apply**.

5. Examine the points in the map. The pattern of these points suggests that over time, the sightings became more concentrated.

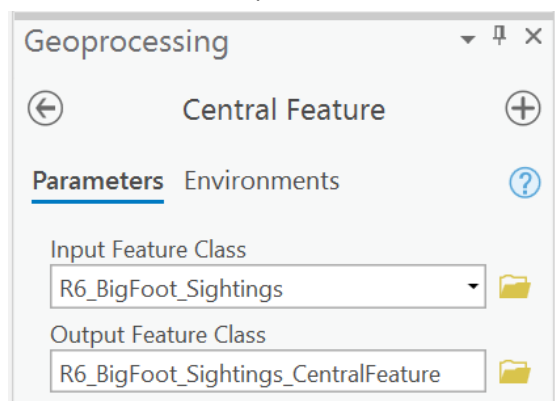


D. Find the most centrally located BigFoot sighting.

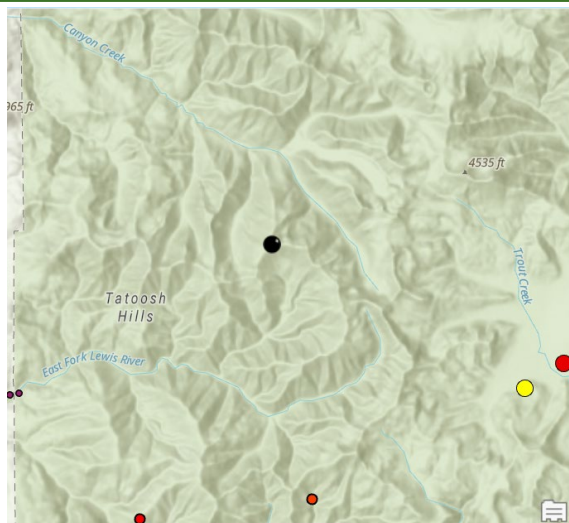
1. Go to the **Measuring Geographic Distributions** toolset, open the **Central Feature** tool.

The central feature tool identifies the feature that has the lowest total distance to all the other features. This tool answers the questions: What feature is closest to all other features? Where are the sightings most prevalent?

2. In the Central Feature dialog, enter the following parameters:
 - i. For **Input Feature Class**: select **R6_BigFoot_Sightings** from the dropdown.
 - ii. Confirm that the output feature class reads: **R6_BigFootSightings_CentralFeature**.



- iii. Accept all other defaults and click **Run**.
3. Change the point symbology to a black dot.



The location of the Central Feature compared to that of the Mean Center tells us that we do not have any extreme outliers in the dataset.

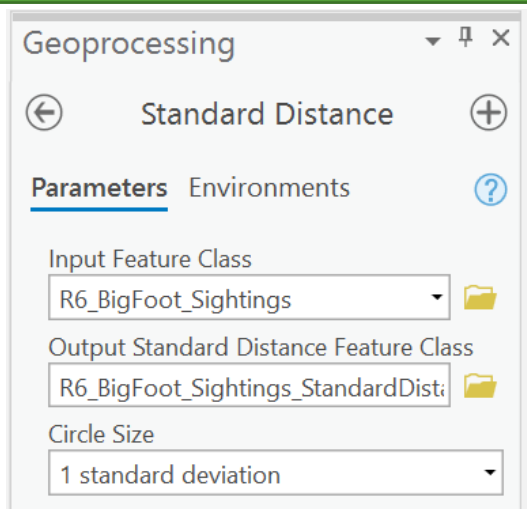
To give you a little better understanding, let's imagine a single point in the BigFoot Sighting dataset has an inaccurate XY location somewhere near the north pole. Think about how this would affect the output location of the Mean and Central Feature tools.

While both outputs would understandably be inaccurate, this would skew the output location of the Mean Center to be much farther North than that of the Central feature.

E. Determine how BigFoot sightings are dispersed around the mean.

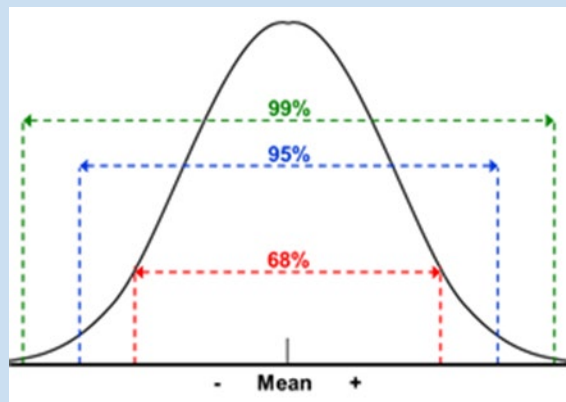
While the Mean Center and Central Feature tools tell us about the center of a distribution, they don't tell us how the features are dispersed around the center. What if we need to know information about the overall distribution of the data and how the data is dispersed around the center? The Standard Distance tool will provide this evidence.

1. Go to the **Measuring Geographic Distributions** toolset and open the **Standard Distance** tool.
2. In the Standard Distance enter the following parameters:
 - i. Input Feature Class: **R6_BigFoot_Sightings**.
 - ii. Output Feature Class: **R6_BigFootSightings_StandardDistance**.
 - iii. Circle Size: **1 Standard Deviation**.

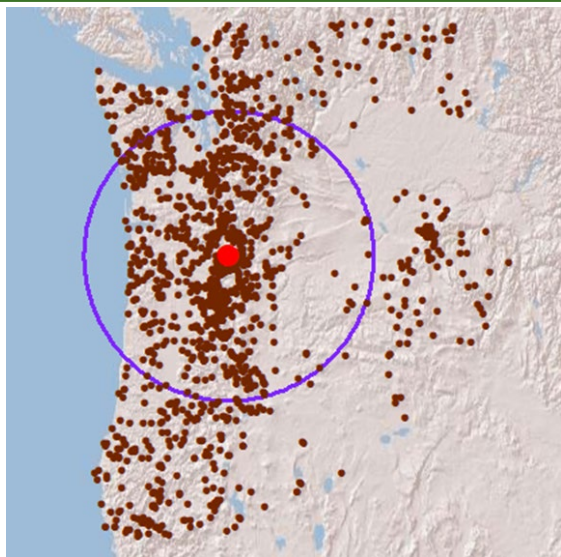


iv. Accept all other defaults and click **Run**

We chose 1 standard deviation for the sample size, 1 standard deviation includes 68% of the data that is closest to the mean in the analysis. Run the tool again with 2 or 3 standard deviations. The ellipses will include a larger percentage of points.



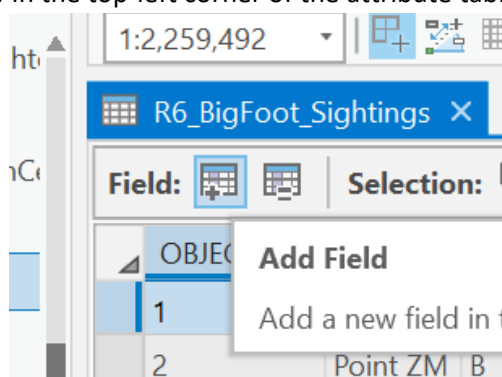
The standard distance ellipse shows that a majority of BigFoot sightings occur within a specific region and indicates possible clustering.



F. Determine how BigFoot sightings are dispersed around the mean over time.

Let's include time in our analysis. We have 150 years of BigFoot sighting data. By creating a new "Case" field for R6_BigFoot_Sightings that includes 25 year increments, we can monitor the change of sightings over time.

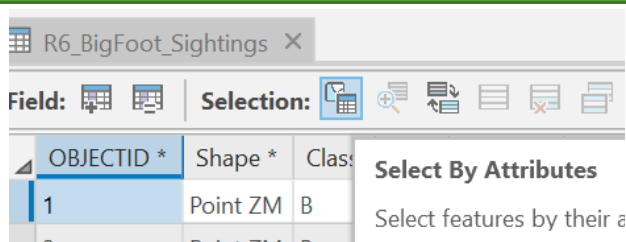
1. In the Contents pane, right-click **R6_BigFoot_Sightings** the select **Attribute Table**.
2. In the top-left corner of the attribute table click the **Add Field** button.



- i. Set the new Field Name to **Case**.
 - ii. **Data Type: Short Integer**.
 - iii. Close the *Fields* table then click **Yes** when asked 'Save all changes?'.
3. Verify the new field was added on the right end of the attribute table
4. Now, one at a time, we will select each span of years as outlined in the chart below and assign the appropriate Case value.

Time	1861—1886	1887—1911	1912—1936	1937—1961	1962—1986	1987—2011
Case	1	2	3	4	5	6

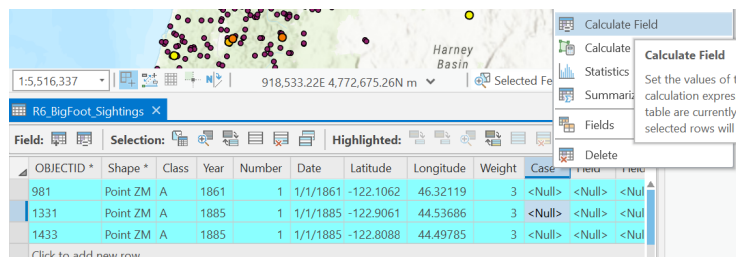
5. In the Selection section of the attribute table click the **Select by Attribute** button.



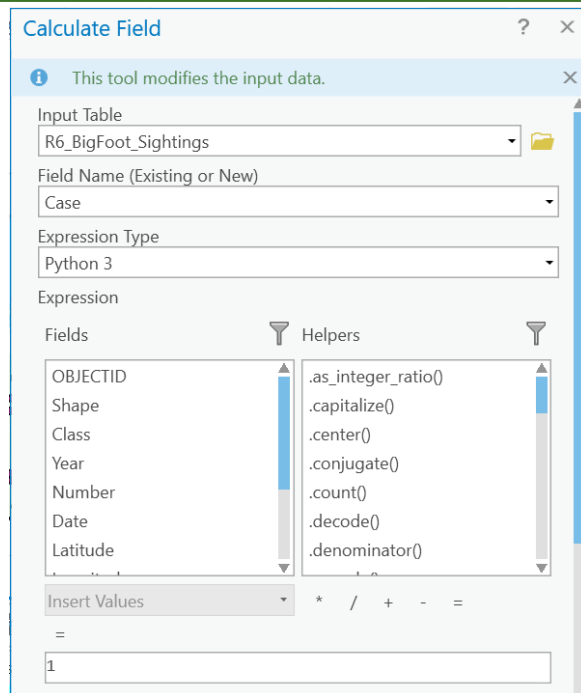
6. In the Select by Attribute dialog, select **'New expression'** then use the dropdown menus to select **Where Year is greater than or equal to 1861 And Year is less than or equal to 1886.**



7. Click **Apply**. There should now be 35 features selected.
8. We want this selection to have a Case value of 1. *Right click Case field header then click Calculate Field.*



9. Enter the value **1** in the field below the = sign.

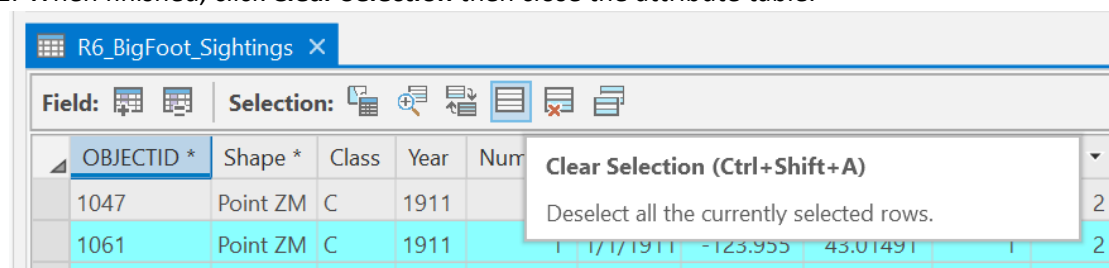


10. Click **Apply**. The Case field for the selected records is updated with the number 1.

11. Using the chart below, repeat those steps for the remaining time periods.

Definition Query	Case
"Year" >= 1861 AND "Year" <= 1886	1
"Year" >= 1887 AND "Year" <= 1911	2
"Year" >= 1912 AND "Year" <= 1936	3
"Year" >= 1937 AND "Year" <= 1961	4
"Year" >= 1962 AND "Year" <= 1986	5
"Year" >= 1987 AND "Year" <= 2011	6

12. When finished, click **Clear Selection** then close the attribute table.



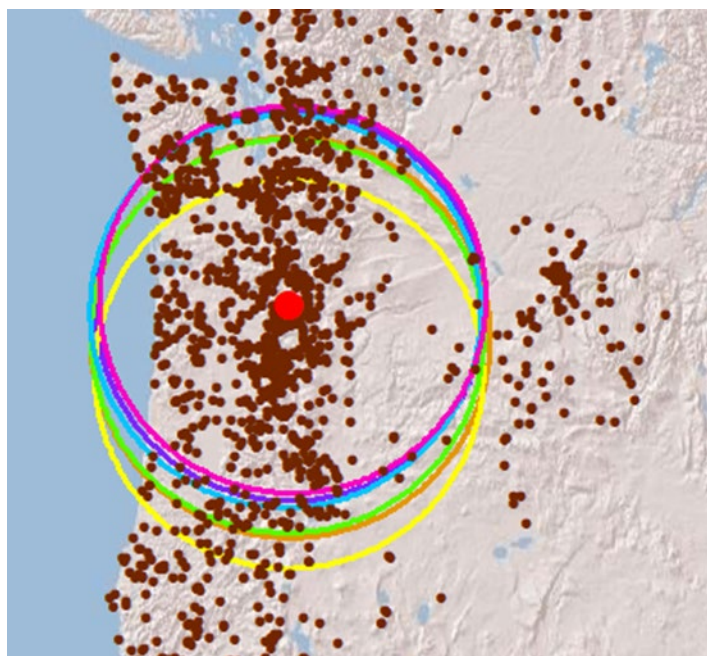
13. Now that we have a case field, we can create our standard distance ellipses. Open the **Standard Distance** tool again.

14. In the Standard Distance tool, enter the following parameters:

- Input Feature Class: **R6_BigFoot_Sightings**
- Output Feature Class: **R6_BigFoot_Sightings_StandardDistanceCase**
- Circle Size: **1 Standard Deviation**.

- iv. **Case Field:** **Case**.
- v. Click **Run**. The output will be one standard distance ellipse for each 25 year time period.
- 15. Change the symbology of the new layer to display each Case as a separate ellipse and color.
 - i. In the *Contents pane*, right-click **R6_BigFoot_Sightings_StandardDistanceCase** then select **Symbology**.
 - ii. In the Symbology pane, go to the **Primary symbology** menu and select **Unique Values**.
 - (a) Set Field 1 to **Case**.
 - (b) Adjust the fill to hollow.

The ellipses appear to be moving slightly north but for the most part have maintained a relatively consistent location. For future sightings, a similar mean location is likely.

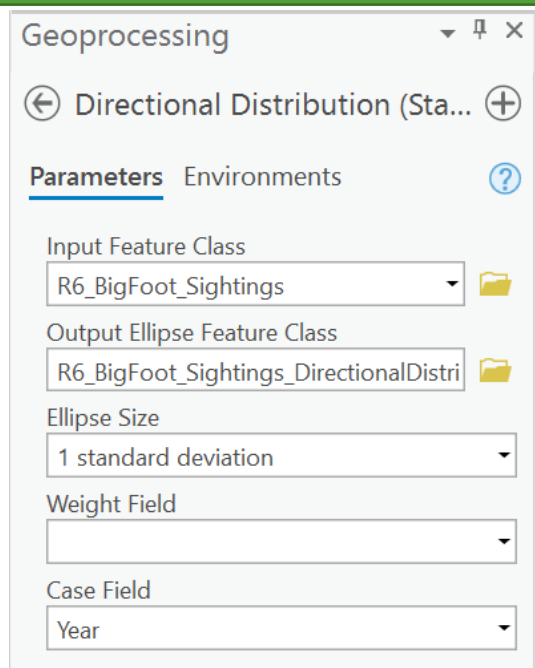


G. What is the distribution and orientation of recent BigFoot sightings?

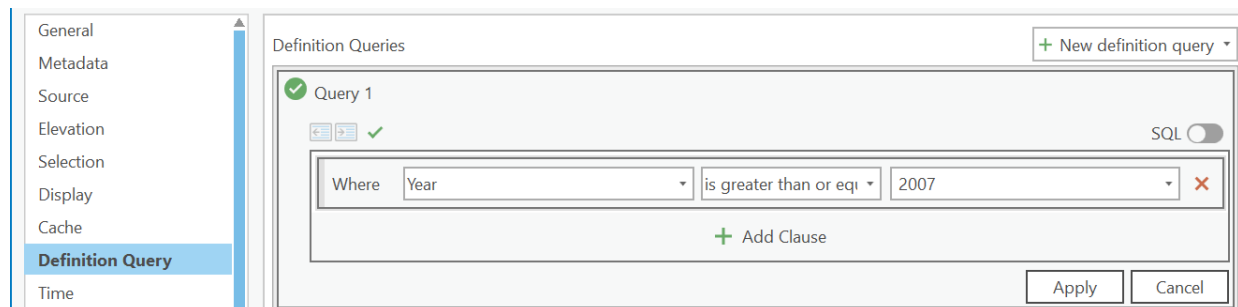
- 1. From the Measuring Geographic Distributions toolset, open the **Directional Distribution (Standard Deviation Ellipse)** tool.

The Directional Distribution (Standard Deviation Ellipse) tool will show us the distribution of the data as well as any directional trends or orientations in the distribution of the data. In this step, we will look at the variance in distribution of the sightings from 2007 - 2011.

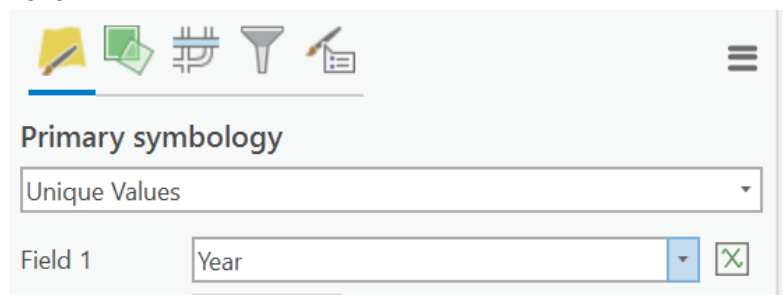
- 2. In the *Directional Distribution (Standard Deviation Ellipse)* window, enter the following parameters:
 - i. *Input Feature Class:* **R6_BigFoot_Sightings**.
 - ii. *Output Feature Class:* **R6_BigFoot_Sightings_DirectionalDistribution**.
 - iii. *Circle Size:* **1 Standard Deviation**.
 - iv. *Case field:* **Year**.
 - v. Click **Run**.



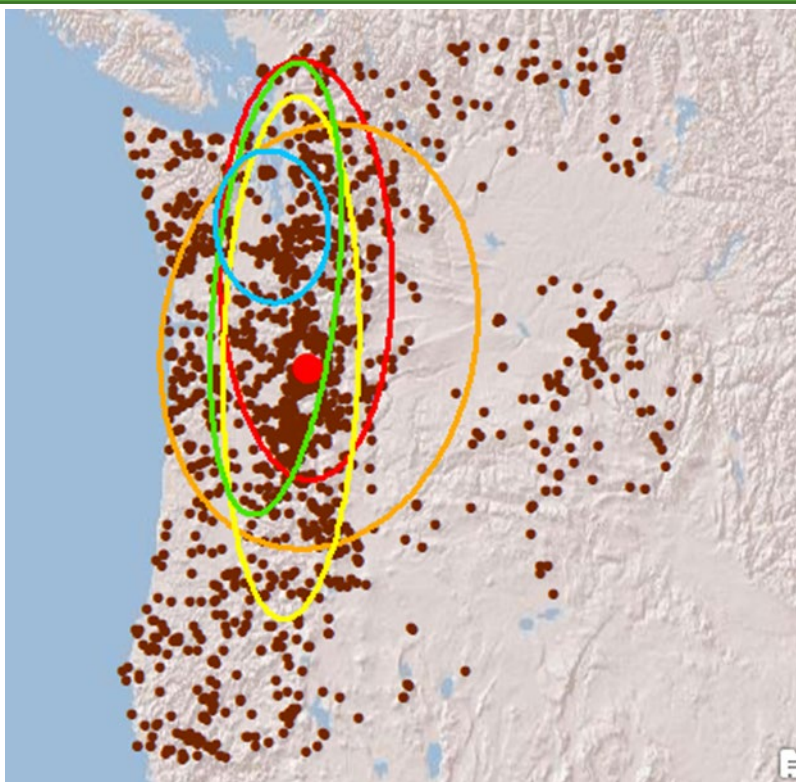
3. Remember we are only interested in the distribution of sightings from 2007 - 2011. We can create a definition query to exclude anything occurring before 2007. In the Contents pane, right-click the **R6_Big_Foot_Sightings_DirectionalDistribution** layer and select **Properties**
 - i. In the Layer Properties window click **Definition Query**
 - ii. Enter the following definition query: **Year is greater than or equal to 2007.**



- iii. Click **Apply**.
4. Change the layer's symbology to **Unique Values** with **Field 1** set to **Year**. Adjust the fill to hollow.



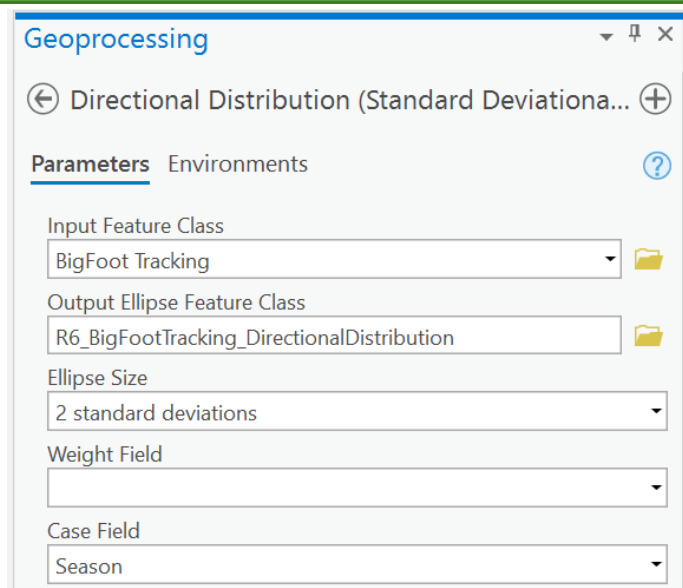
- i. **QUESTION** – Are there any trends in the data?
 - ii. **QUESTION** – What is significant about the location of sightings that occur in 2011?



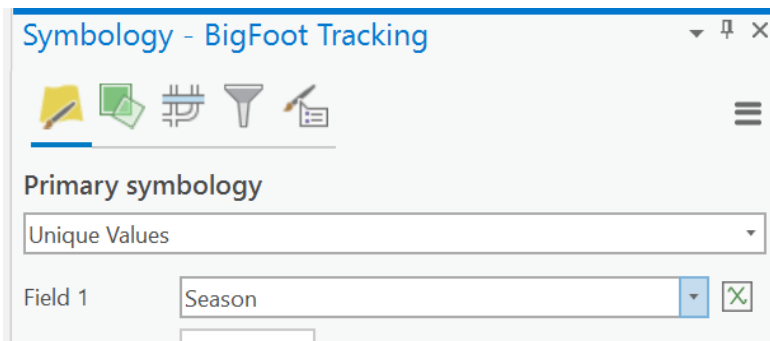
H. Monitor the seasonal movement of a radio collared BigFoot.

GOOD NEWS! We were able to capture and radio collar a BigFoot near the Olympic Peninsula. The collar will capture a GPS point every 24 hours. With this data, we hope to gain a better understanding of the species with respect to their seasonal movement in order to help maintain their basic needs of food, water, and security. Our goal is to measure and document seasonal movement patterns.

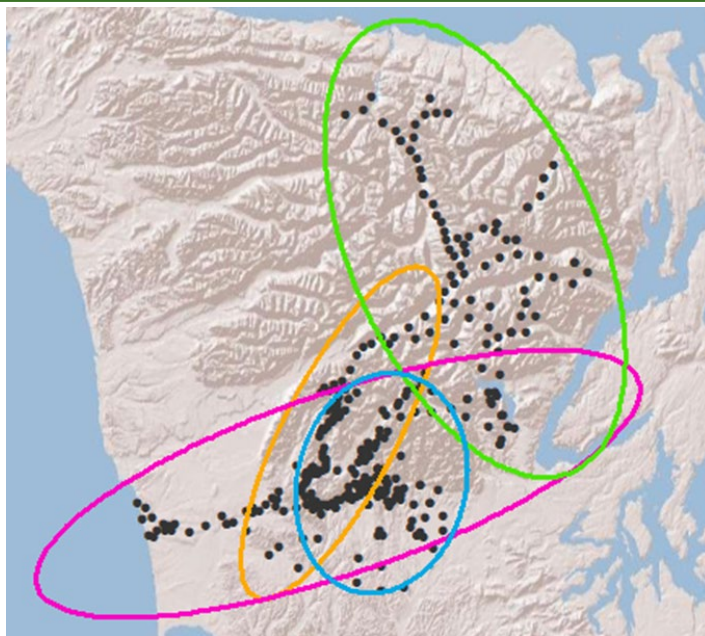
1. In the *Contents pane*, turn on the **BigFoot Tracking** layer and turn off all the other layers.
2. Right-click the BigFoot Tracking layer then click **Zoom to Layer**.
3. Right-click the BigFoot Tracking layer again then click **Attribute Table**.
 - i. The attributes for BigFoot Tracking include a date, the location and the season. We will use the *Standard Deviation Ellipse* tool to analyze the BigFoot's movement during different seasons.
4. From the Measuring Geographic Distributions toolset, open the **Directional Distribution (Standard Deviation Ellipse)** tool.
5. In the Directional Distribution (Standard Deviation Ellipse) window, enter the following parameters:
 - i. Input: BigFoot Tracking.
 - ii. Output Feature Class name: **R6_BigFootTracking_DirectionalDistribution**.
 - iii. Circle Size: **2 Standard Deviations**.
 - iv. Case Field: **Season**.



- v. Click **Run**.
6. Change the symbology of the new output layer to **Unique Values**, based on **Season**. Adjust the fill to hollow.



7. Examine the output in the map.



The ellipse allows us to see if the distribution of features is elongated and consequently has a particular orientation. In the example above, we see several ellipses with different orientations. Fall = NE/SW, Spring = E/W, Summer = NW/SE, Winter-Clustered

So... what have we learned? *The standard deviational ellipses tell us there is very little movement during the winter; the fall and spring show movement in a southwest/northeast direction but remain reasonably close to that same area. The summer months allow movement into higher elevations and further away from the average distribution and display movement in a northwest/southeast direction.*

Congratulations! You have completed Exercise 1.