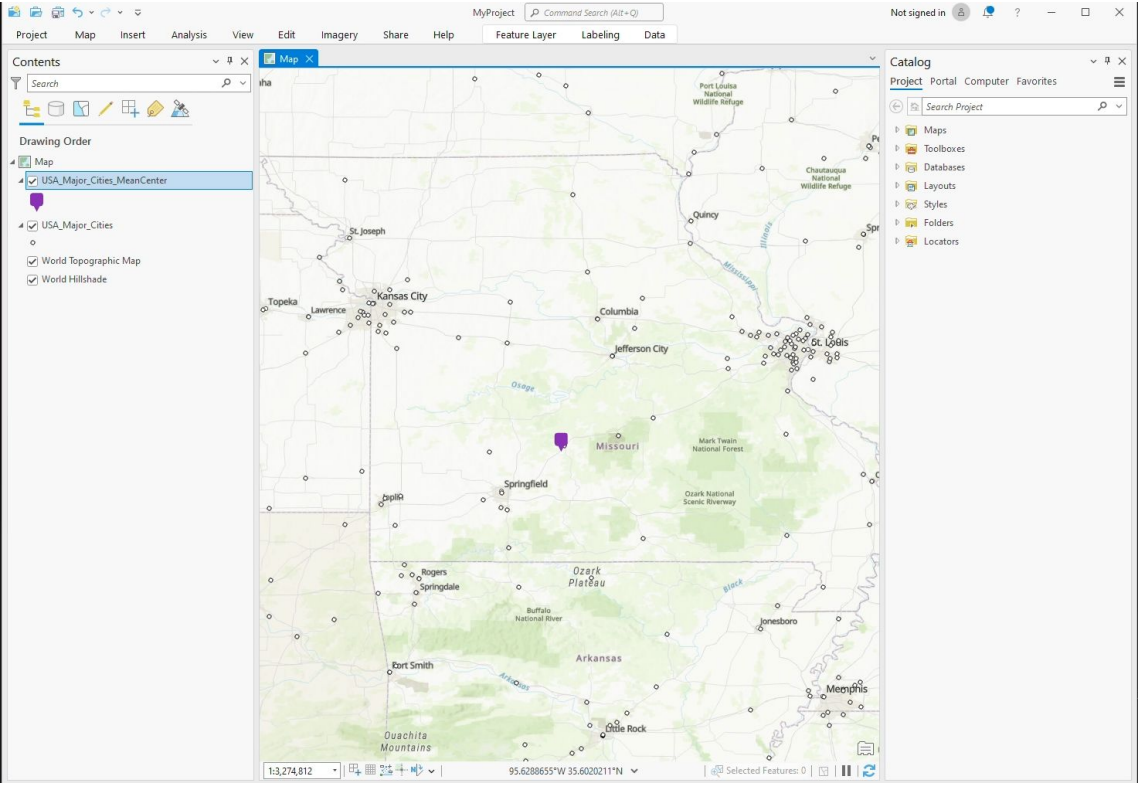
# Point Data Analysis of US Major Cities

By: Devin Webb

# Unweighted Central Tendency(UCT)

- The Unweighted Central Tendency(UCT) is the mean center based on the geographical coordinates of all the cities
  - Or the average location of all the points in the dataset, in this case the average location of all the cities
  - The UCT is represented by the Purple sign

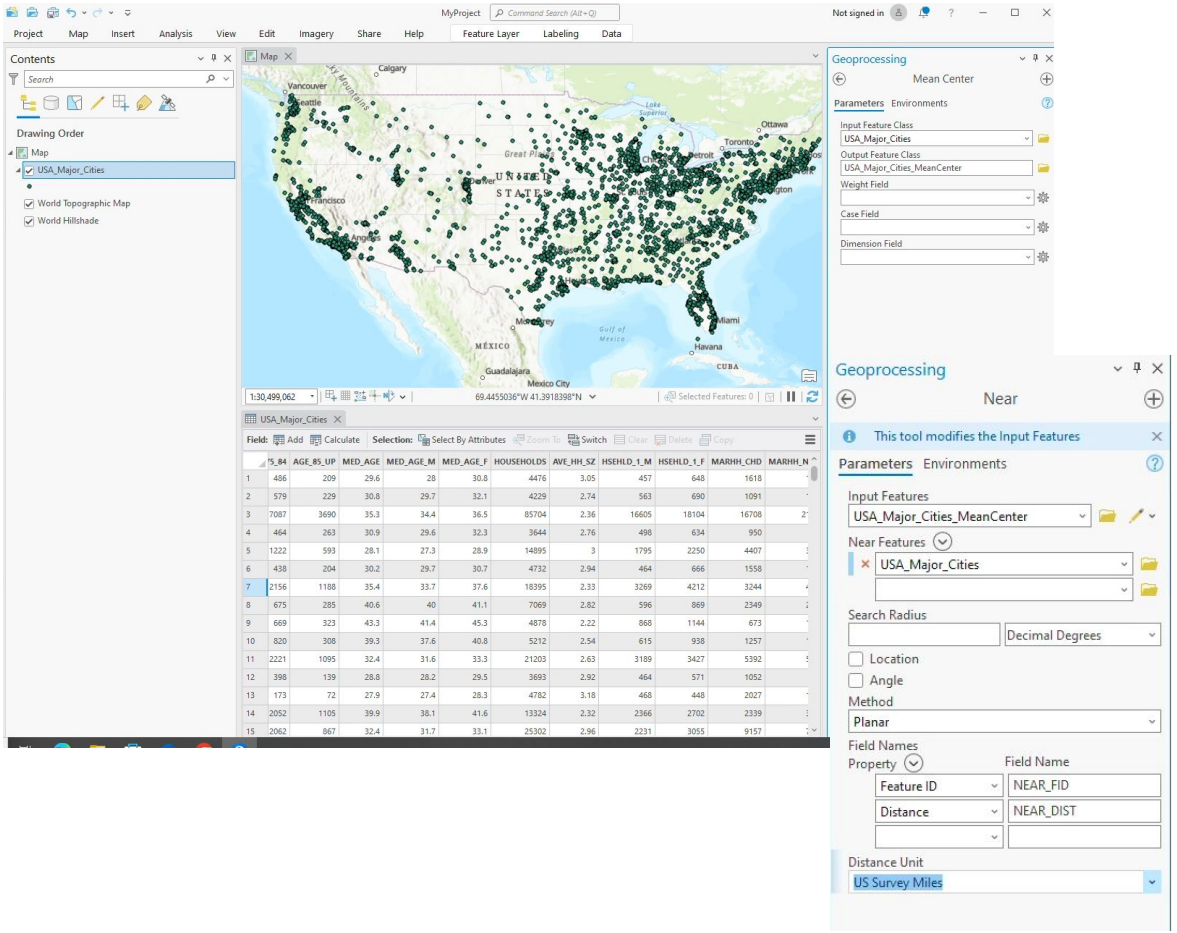



- The UCT happens to be right next to Lebanon, Missouri
- The UCT is approximately 0.071683 miles away from this city

# Tools Used

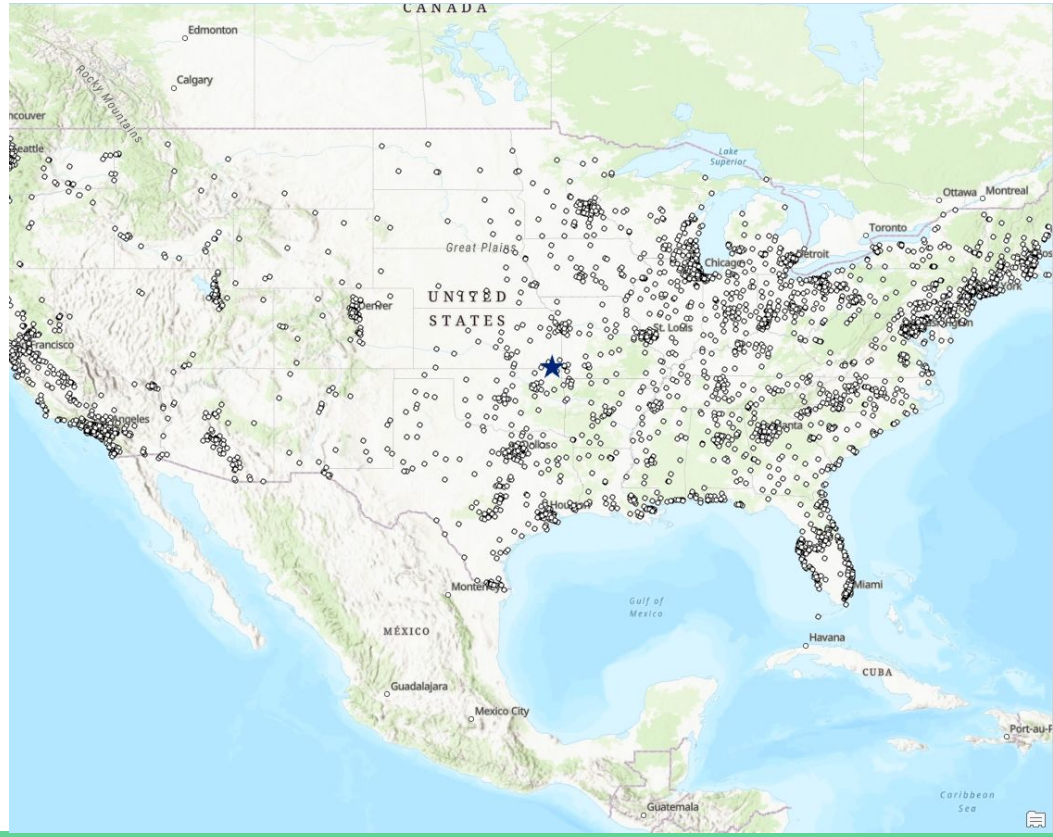
I used:

- ❏ the Mean Center tool to calculate the UCT
- ❏ the input feature as US_Major_Cities and created an output feature named USA_Major_Cities_MeanCenter
- ❏ the near tool to help calculate how far away the UCT was from the nearest Major City

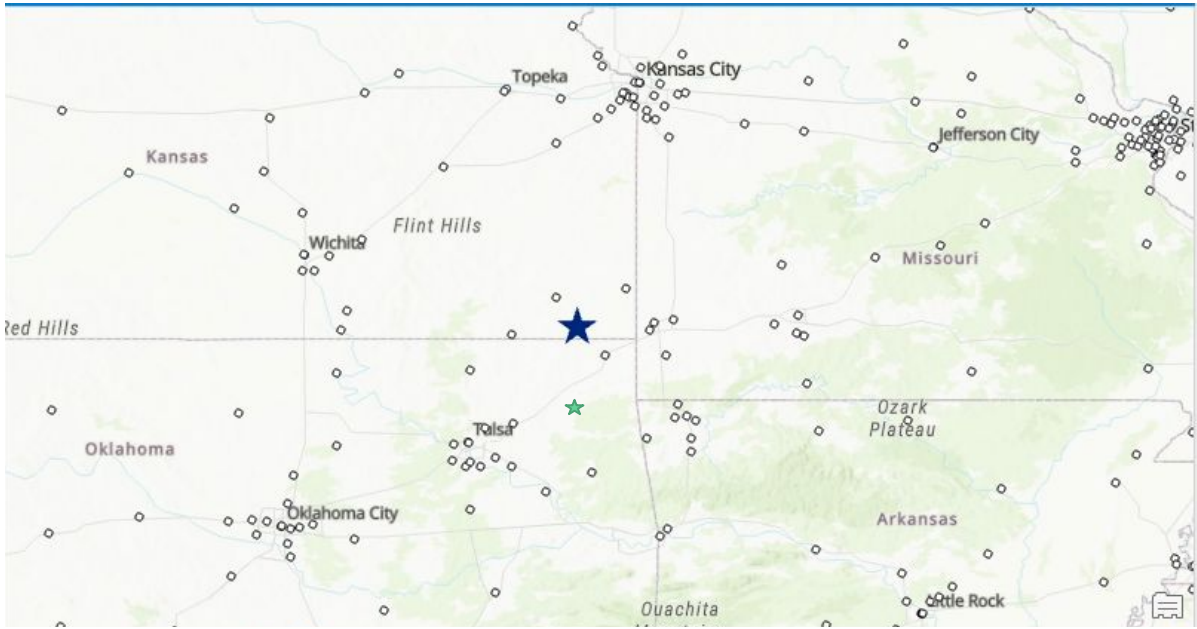


# Weighted Central Tendency

- The Weighted Central Tendency (WCT) is the mean center based on an attribute
  - In this case the attribute is ‘POPULATION’
- Each point becomes weighted by this attribute giving cities with larger population higher weight
- The WCT is represented by the blue star

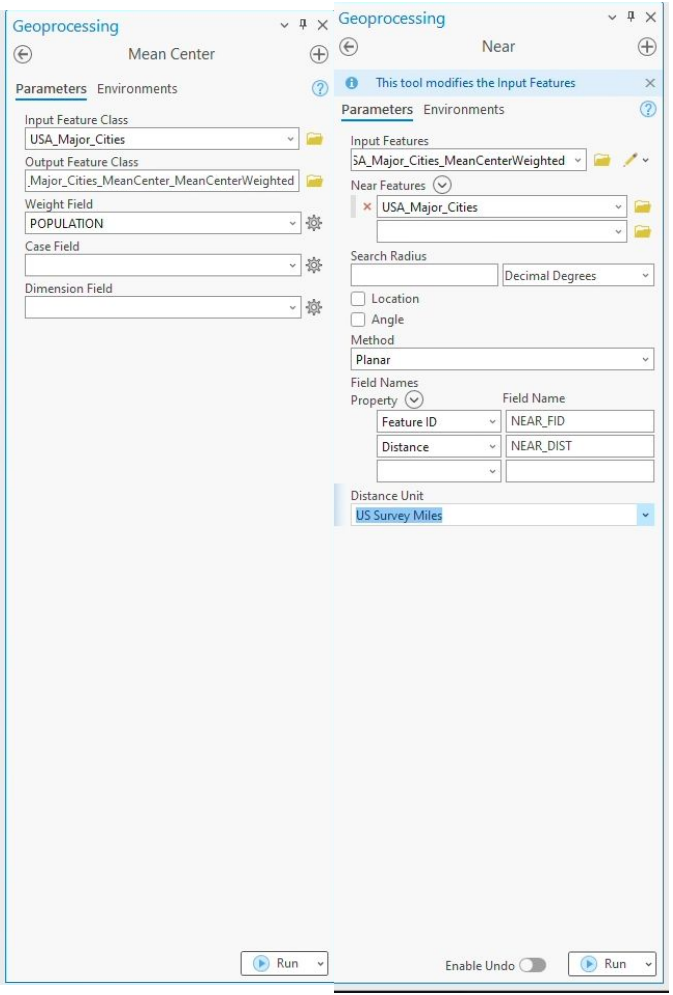


- The Nearest city to the WCT is Parsons, Kansas
- The WCT is approximately 0.284397 Miles away from Parsons
- Parsons is represented by the smaller star



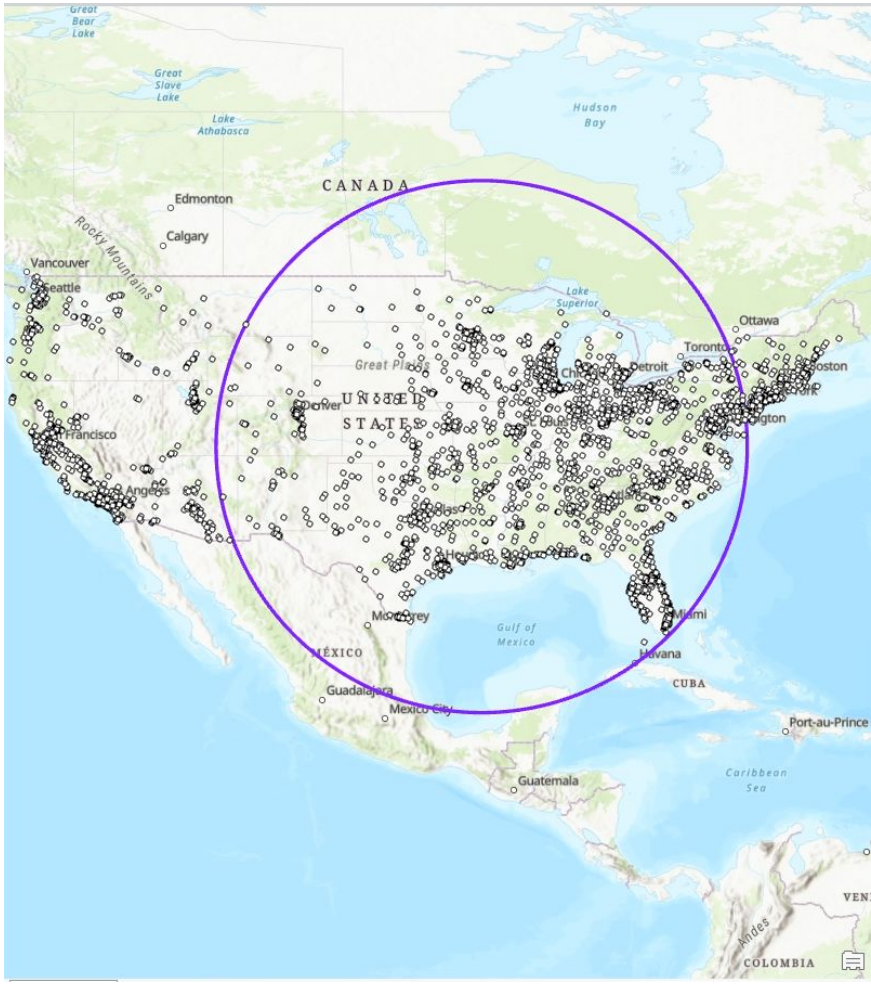
# Tools Used

- I used the Mean Center tool to calculate the WCT
- I used the input feature as US_Major_Cities and created an output feature named USA_Major_Cities_MeanCenterWeighted
- I set the Weight Field to ‘POPULATION’
- I also used the near tool to help calculate how far away the WCT was from the nearest Major City



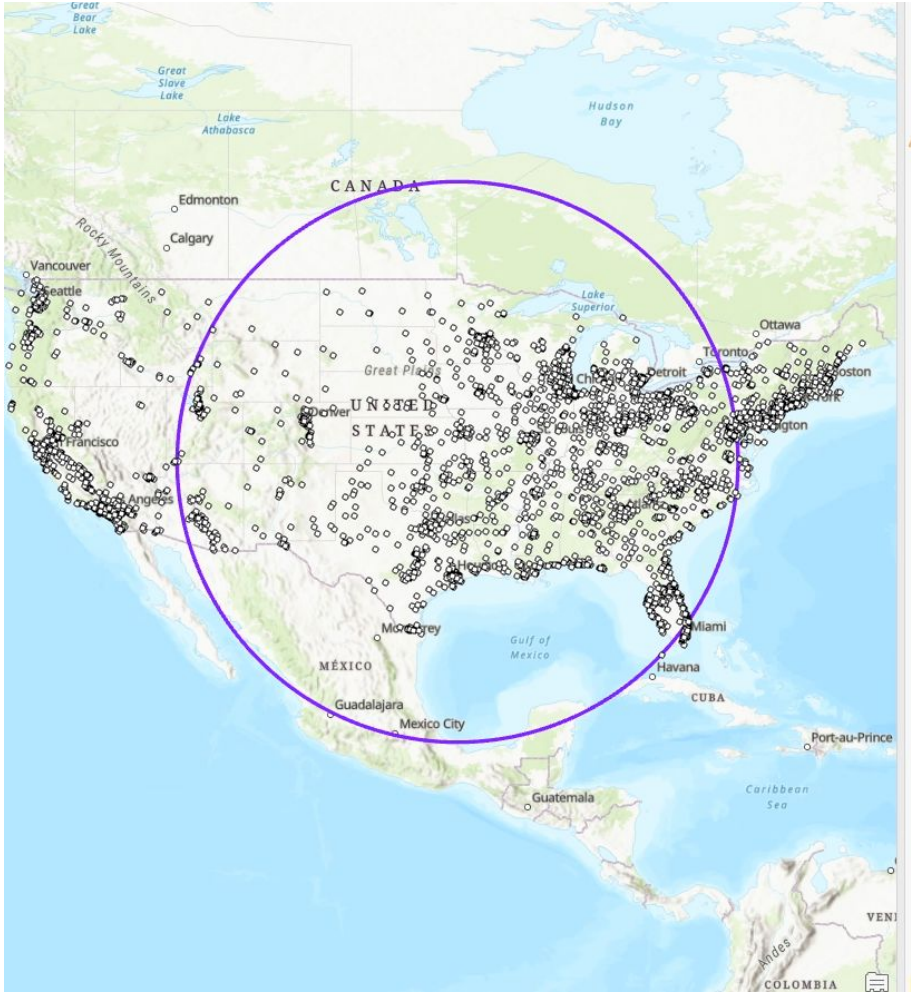
# Dispersion for UCT

- This is the standard distance 1 standard deviation away from the mean center or the UCT



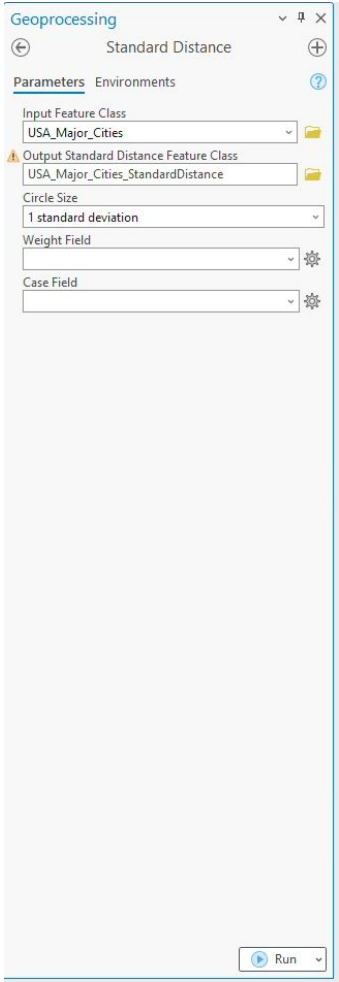
# Dispersion for WCT

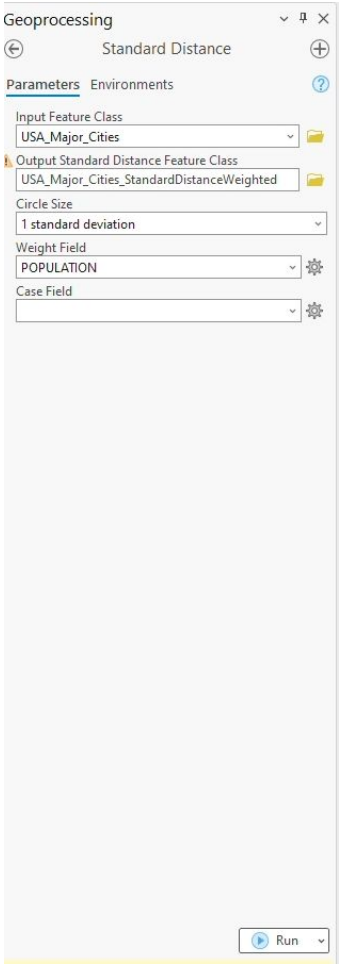
- This is the standard distance 1 standard deviation away from the mean center or the WCT weighted by the ‘POPULATION’
- This circle is includes more cities to the left indicating the populations on the left have more population but there are more cities on the right in general



# Tools Used

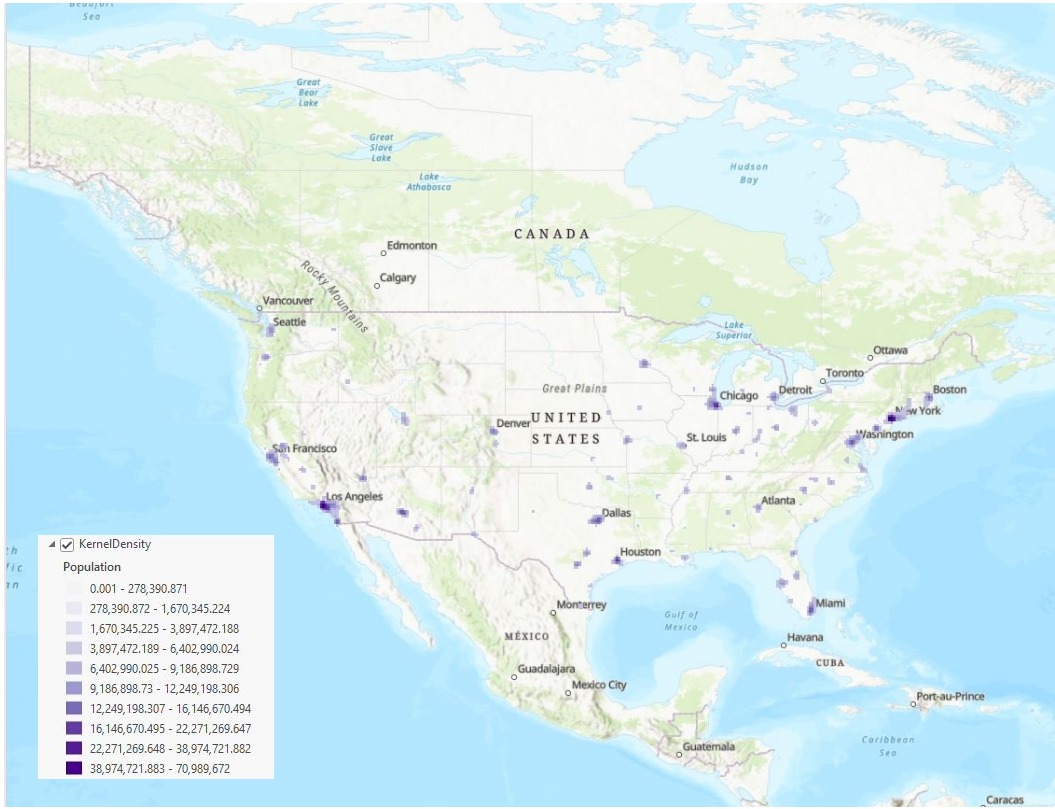
- I just used the Standard distance tool
- I set the circle size to 1 standard deviation which should include about 68% of the data.
- The image on the right is for the WCT



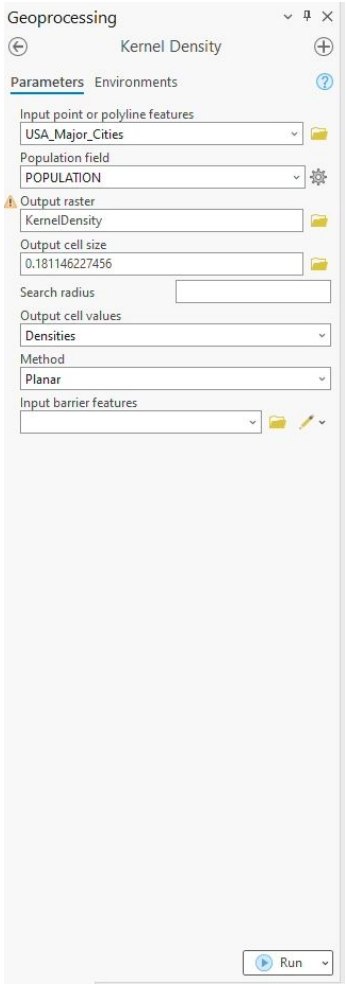
# Cluster/Density

- Cluster/Density is clustering or the density based off an attribute
  - In this case the Attribute is ‘POPULATION’
- As the color gets darker the density of these areas gets higher as seen on the key



# Tools Used

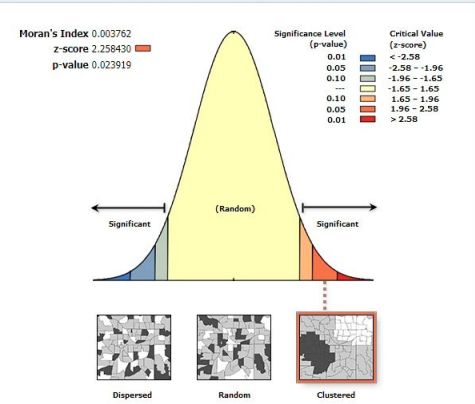
- The tool I used for this method is Kernel Density
- Kernel Density estimates the density across the major cities  and helps visualize the data.
- I left the output cell size the same and set the population field to POPULATION



# Global Spatial Autocorrelation

- Here is the report generated using the Global Moran's 1 Tool
- The Moran's index is slightly positive autocorrelation. This indicates there happens to be a slight tendency for cities with similar populations to be near each other
- The expected index  is -0.000257 and by observed index is higher than this which supports the presence of some level of clustering.

**Spatial Autocorrelation Report**




Given the z-score of 2.25843, there is a less than 5% likelihood that this clustered pattern could be the result of random chance.

**Global Moran's I Summary**

| | |
|---|---|
| **Moran's Index** | 0.003762 |
| **Expected Index** | -0.000257 |
| **Variance** | 0.000003 |
| **z-score** | 2.258430 |
| **p-value** | 0.023919 |

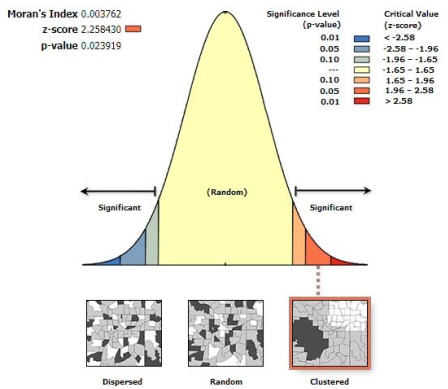
**Dataset Information**

| | |
|---|---|
| **Input Feature Class:** | USA_Major_Cities |
| **Input Field:** | POPULATION |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | True |
| **Distance Threshold:** | 920805.0047 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | False |

# Global Spatial Autocorrelation

- The very small variance suggests very little variability in the spatial autocorrelation
- The Z-Score at 2.258430 which indicates there is less than a 5% chance the observed clustering is not likely to be random
- The P-Value of 0.023919 is less than 0.05 also confirming the observed clustering is not likely to be random

**Spatial Autocorrelation Report**




Given the z-score of 2.25843, there is a less than 5% likelihood that this clustered pattern could be the result of random chance.

**Global Moran's I Summary**

| | |
|---|---|
| **Moran's Index** | 0.003762 |
| **Expected Index** | -0.000257 |
| **Variance** | 0.000003 |
| **z-score** | 2.258430 |
| **p-value** | 0.023919 |

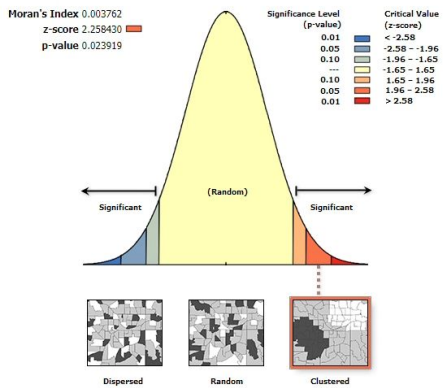
**Dataset Information**

| | |
|---|---|
| **Input Feature Class:** | USA_Major_Cities |
| **Input Field:** | POPULATION |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | True |
| **Distance Threshold:** | 920805.0047 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | False |

# Global Spatial Autocorrelation

- The population of U.S. major cities, when taken as a whole, shows a slight clustering tendency rather than being randomly distributed or evenly dispersed across the country.
- This could suggest that there are certain areas within the U.S. where cities are more likely to have similar population sizes
  - Possibly due to economic, geographical, or social factors that influence city development and population distribution

## Spatial Autocorrelation Report




Given the z-score of 2.25843, there is a less than 5% likelihood that this clustered pattern could be the result of random chance.

### Global Moran's I Summary

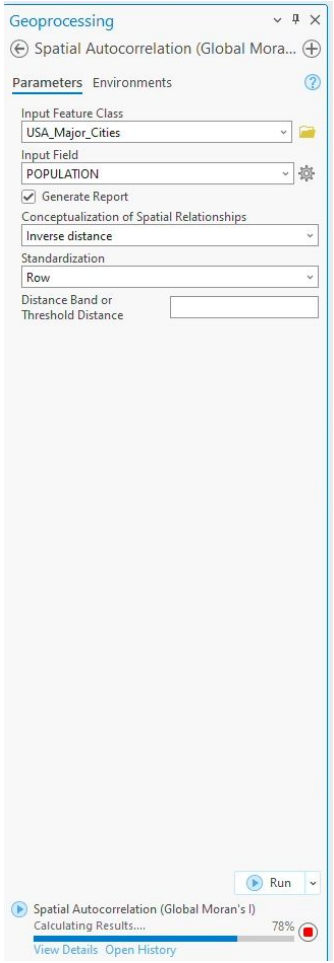
| | |
|---|---|
| **Moran's Index** | 0.003762 |
| **Expected Index** | -0.000257 |
| **Variance** | 0.000003 |
| **z-score** | 2.258430 |
| **p-value** | 0.023919 |

### Dataset Information

| | |
|---|---|
| **Input Feature Class:** | USA_Major_Cities |
| **Input Field:** | POPULATION |
| **Conceptualization:** | INVERSE_DISTANCE |
| **Distance Method:** | EUCLIDEAN |
| **Row Standardization:** | True |
| **Distance Threshold:** | 920805.0047 Meters |
| **Weights Matrix File:** | None |
| **Selection Set:** | False |

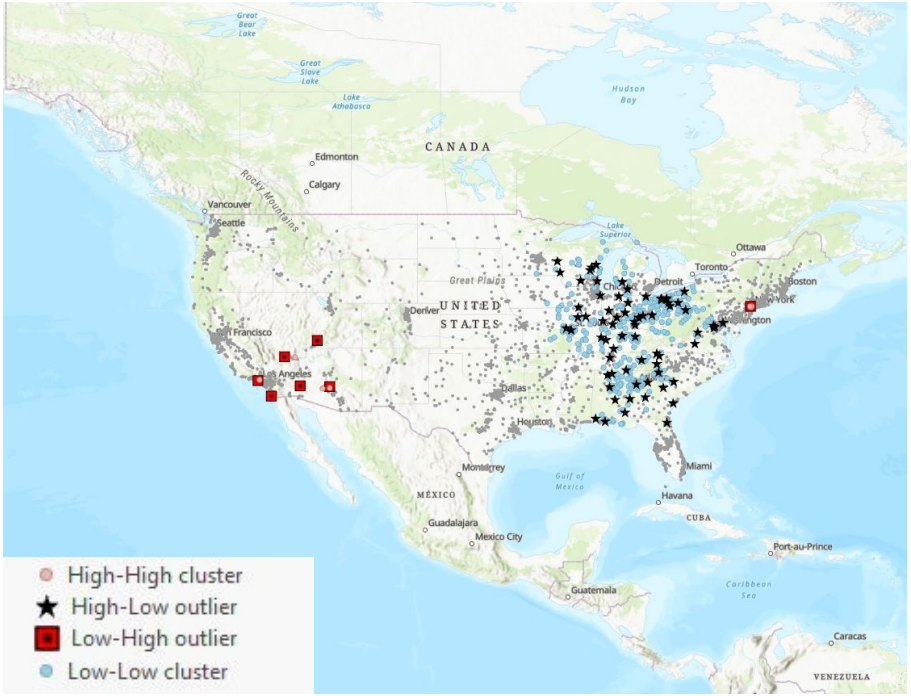
# Tools Used

- I used the Spatial Autocorrelation Global Moran's 1 tool
- I used POPULATION as the input field to see more information about the population in major cities
- I set the standardization to Row because the population amounts are very spread apart



# Local Spatial Autocorrelation Analysis

- This is the clustering of populations using the Anselin Local Moran's Law
- The High - High Cluster Represents cities with large populations near large populations cities
- The Low - Low Cluster Represents cities with low populations near low population cities
- The High-Low outlier represents high populations surrounded by low populations
- The Low-High outlier represents low populations surrounded by high populations



# Tools Used

- The tool I used is the Cluster and Outlier Analysis
- I selected USA_MAJOR_Cities and Population
- I made suser to use apply false discovery rate to limit false positives in the created map