Cassandra Morgan

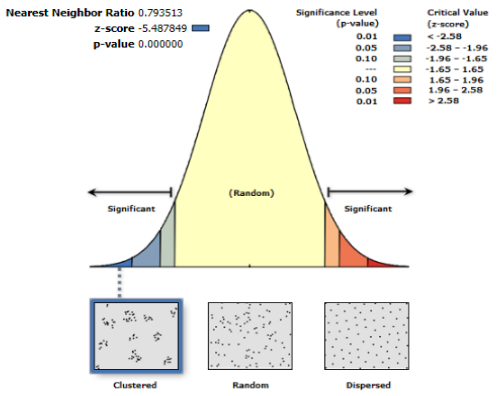
GIS 494

# California Ozone Network Dispersion Analysis

Clustering, a lack of dispersion, can cause over or under representation of subgroups within a sample. This can lead to biases in the dataset. To analyze the dispersion of the California Ozone measurement network, we will use 3 types of analysis: nearest neighbor, quadrat, and analysis of Thiessen polygons.

First, we will use nearest neighbor analysis. According to the nearest neighbor evaluation results below, we have a nearest neighbor ratio of .79, which is below 1 (the z-score associated with this value is -5.49). This implies that the network is relatively clustered. A more ideal distribution pattern would be in the dispersed range, which has a nearest neighbor value of 2.15 or above. This means there is less than a 1% likelihood that this clustering pattern could be the result of random chance.

**Average Nearest Neighbor Summary**




Given the z-score of -5.487849, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

**Average Nearest Neighbor Summary**

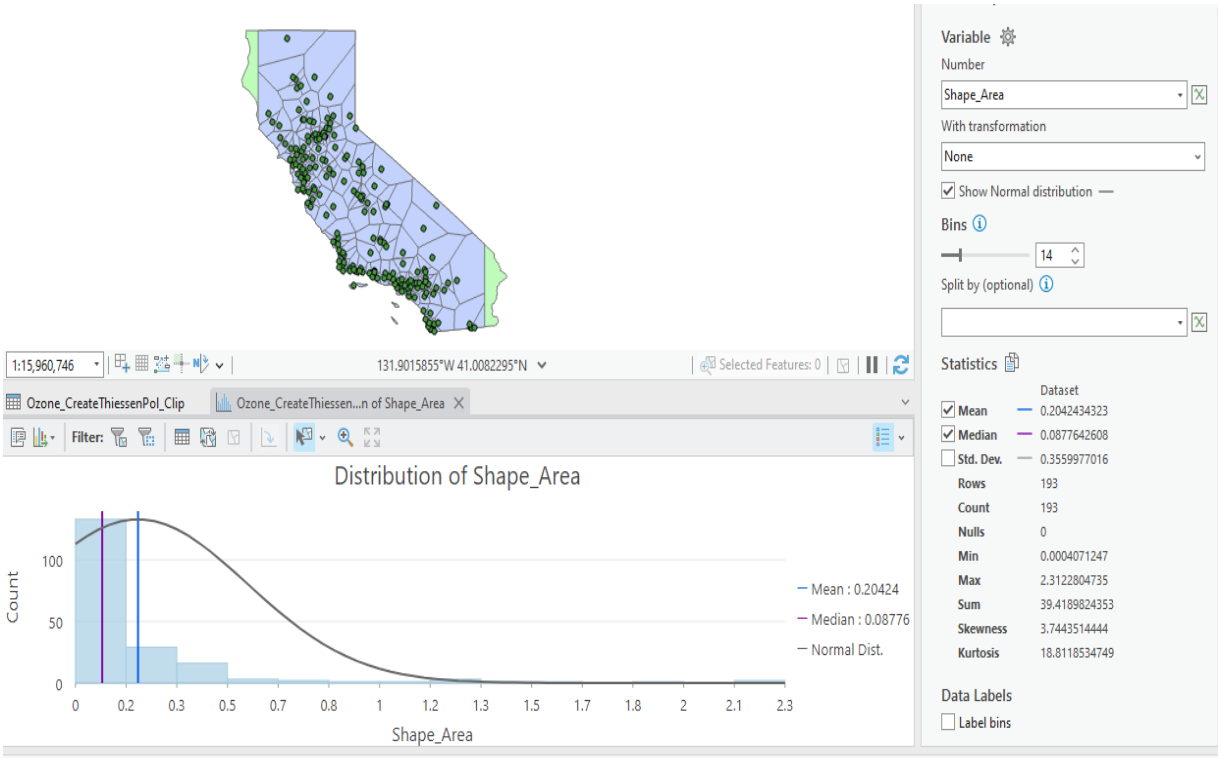
| | |
|---|---|
| **Observed Mean Distance** | 18256.4029 meters |
| **Expected Mean Distance** | 23007.0604 meters |
| **Nearest Neighbor Ratio** | 0.793513 |
| **z-score** | -5.487849 |
| **p-value** | 0.000000 |

**Dataset Information**

| | |
|---|---|
| **Input Feature Class:** | Ozone |
| **Distance Method:** | EUCLIDEAN |
| **Study Area:** | 408638766993.000000 |
| **Selection Set:** | False |

Next, we will use Quadrat analysis to find the critical z-statistic from the network distribution data. The formula used to calculate this z value is $z = \sqrt{\frac{(m-1)}{2}} * (VMR - 1)$, where m = the number of cells, and VMR is the variance of the joint count field divided by the mean of that field. In this case, the z-statistic is 11.30. A z value above +1.96 confirms a clustered pattern at a confidence level of 95%. This implies that our network is very clustered and agrees with the nearest neighbor analysis. The image below shows the map of the quadrats used to find the critical z-statistic.

Finally, we will analyze a map of the Thiessen polygons generated from our network distribution data. The kurtosis of the shape area of the polygons is 18.8 which indicates that the areas measured are very similar in size to each other.




The Thiessen polygon analysis implied that the areas measured are similar in size, but the nearest neighbor and quadrat analysis both imply that the network is not very dispersed in terms of distance. This clustering probably occurred because cities produce more ozone pollution. Additionally, measuring ozone near cities might be deemed more important due to potential health effects on a larger population.