The Impact of Tree Cover on the Urban Heat Islands Effect

Christopher Prendergast,
12/14/2023

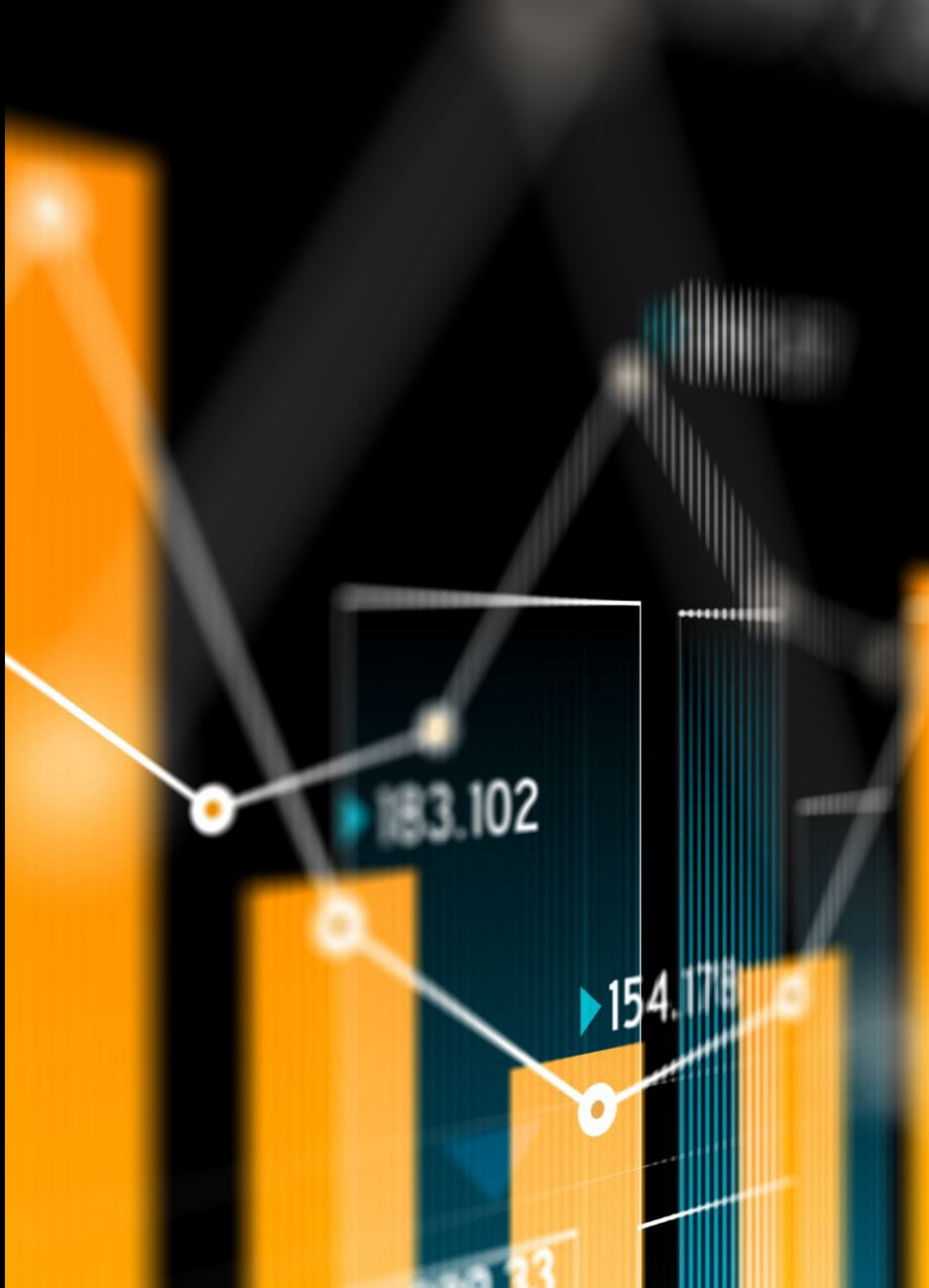
Project Description

- Explore urban heat island effect (UHI)
- Analyze spatial relationship between percent tree cover and temperature
- Study area is the City of San Jose, California

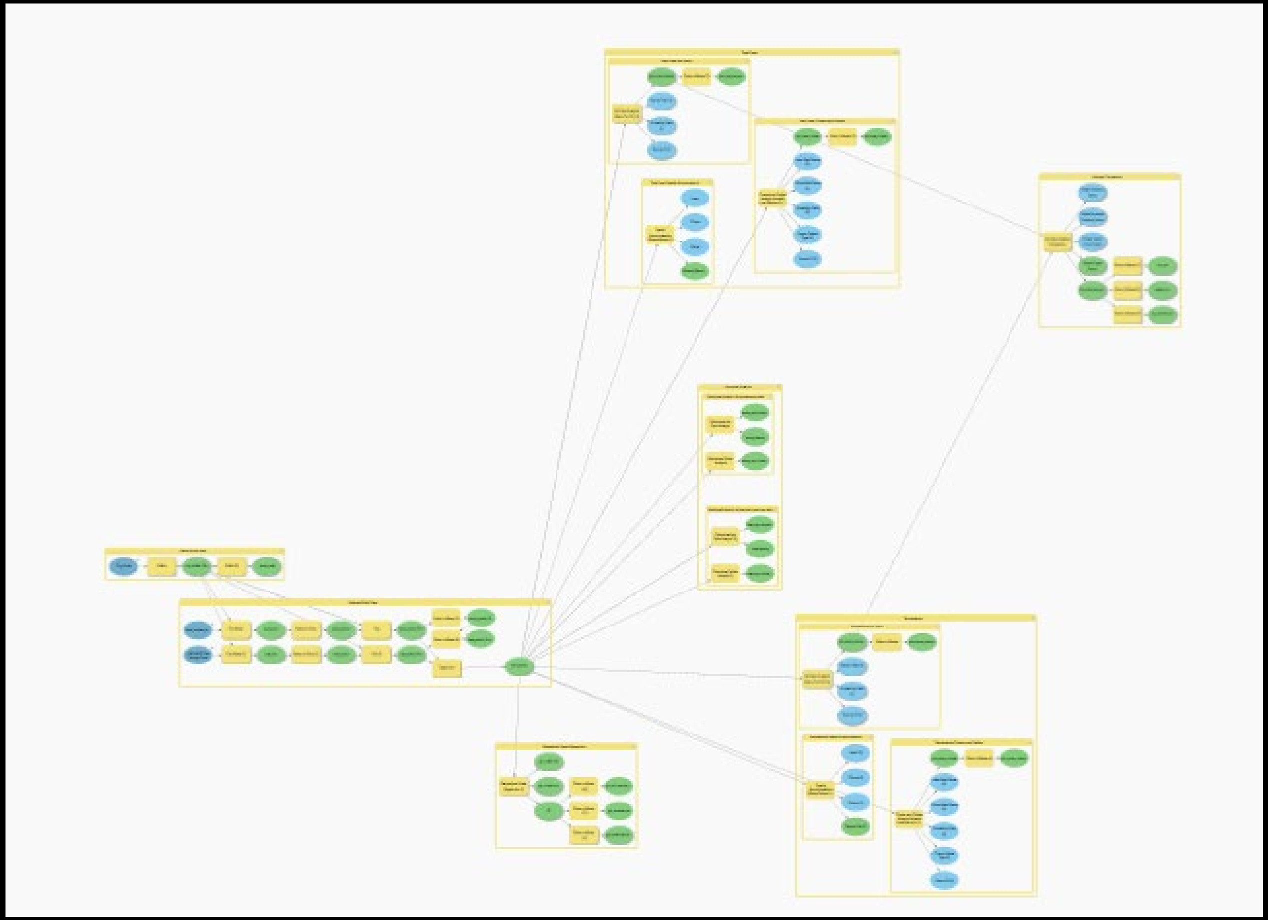
- Null hypothesis: The observed pattern of temperatures could have occurred through random chance and is unrelated to the distribution of tree cover across the study area

Analysis Steps

1. Prepare raster datasets
2. Prepare the study area polygon
3. Wrangle the point data
4. Look for spatial autocorrelation
5. Optimized hotspot, and outlier analyses
6. Hotspot (Getis-Ord Gi*) analyses
7. Cluster and outlier (Anselin Local Moran’s I) analyses
8. Compare hotspots
9. Generalized Linear Regression
10. Geographically Weighted Regression (Didn’t run)

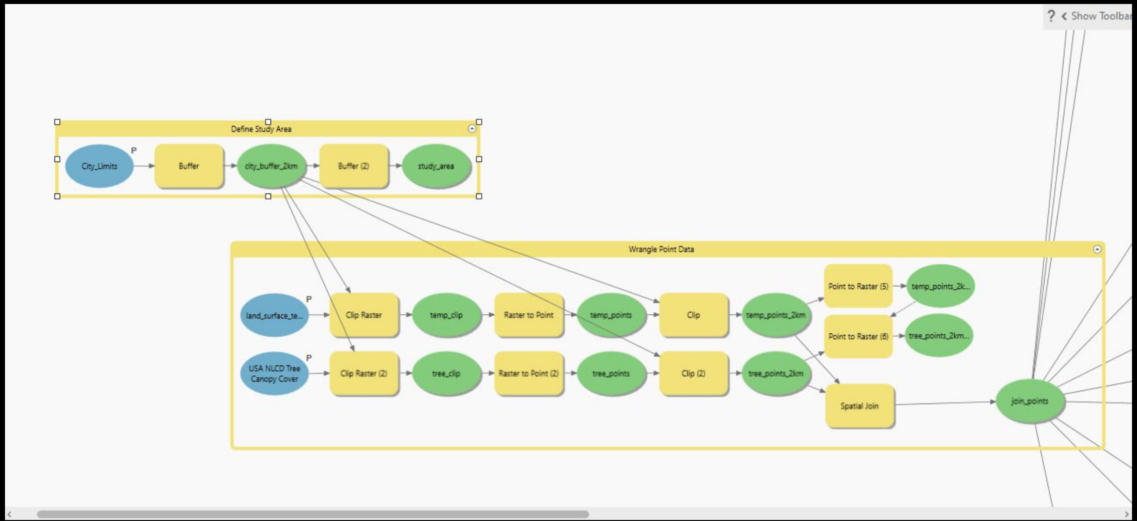


Model Builder -- Overview



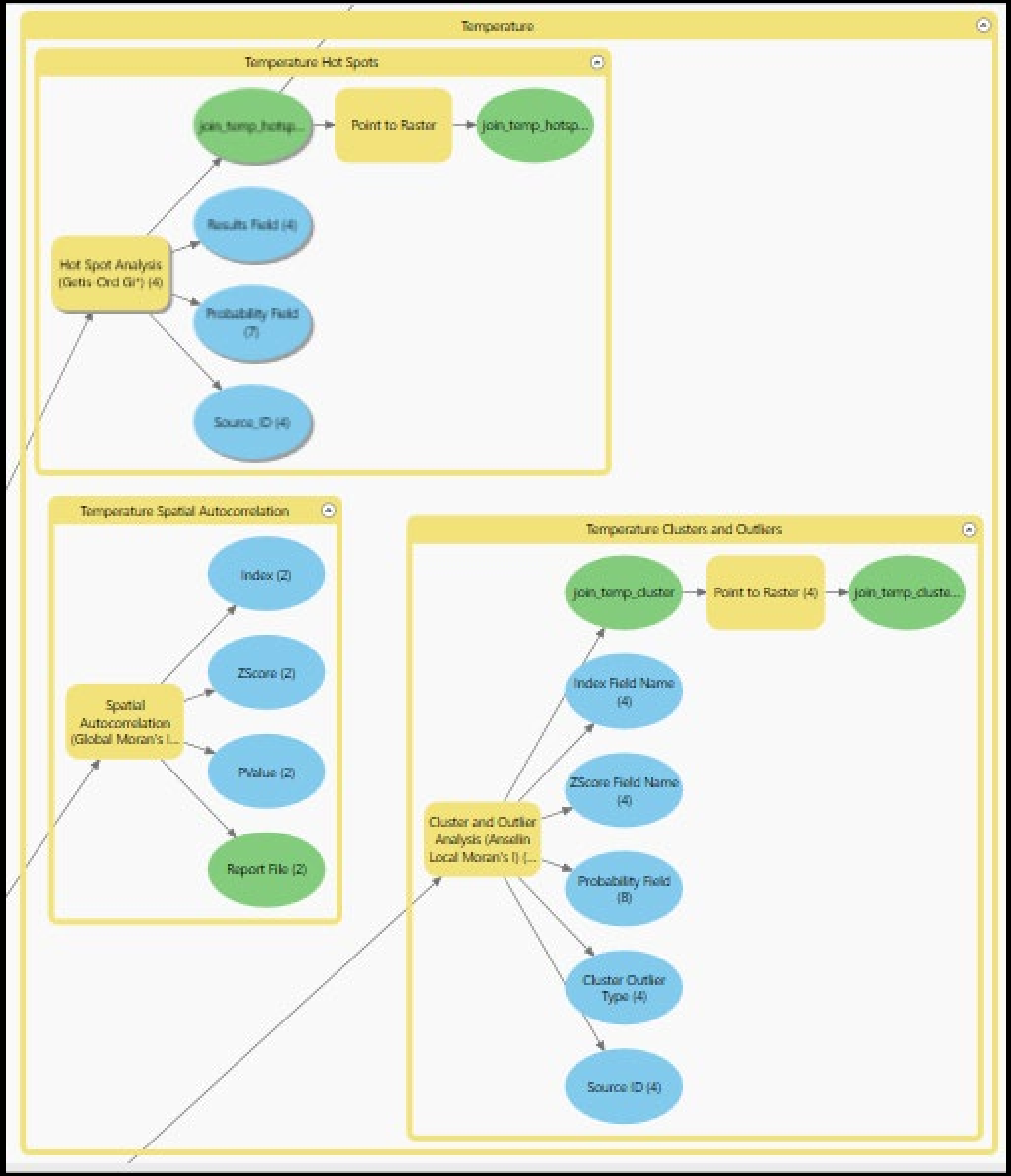
Data Wrangling

- Transform the data into the appropriate format for analysis
- Convert raster data to points
- Create a study are polygon
- Clip datasets to the study area
- Spatially join the datasets to create a single point feature class for analysis



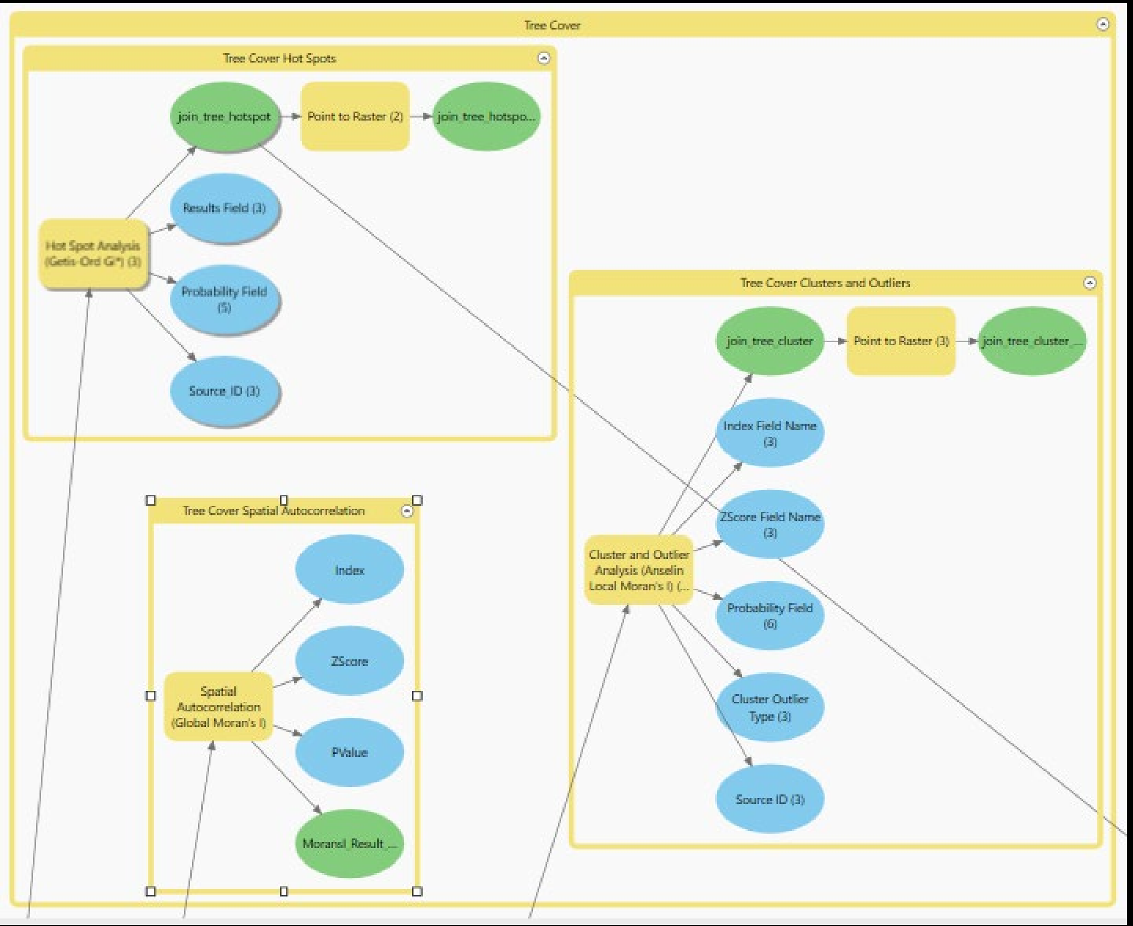
Temperature Analysis

- Is the distribution of temperature data clustered, random, or dispersed?
- Where are the statistically significant temperature hot spots and cold spots?
- Are there any outliers, where are they located?
- Convert vector (point) results to raster for display



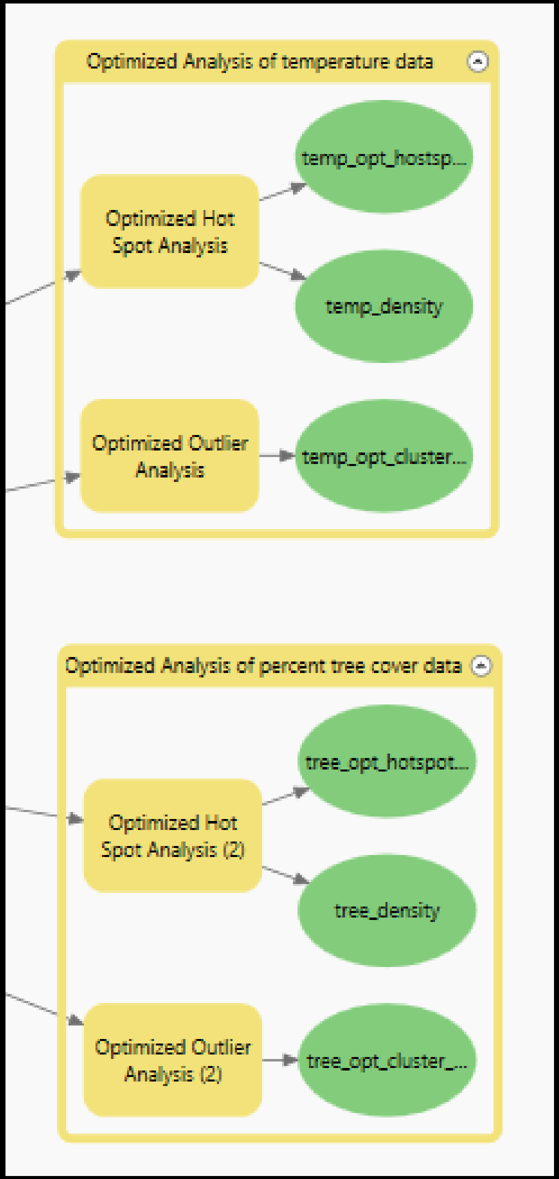
Tree Cover Analysis

- Is the distribution of tree cover data clustered, random, or dispersed?
- Where are the statistically significant temperature hot spots and cold spots?
- Are there any outliers, where are they located?
- Convert vector (point) results to raster for display



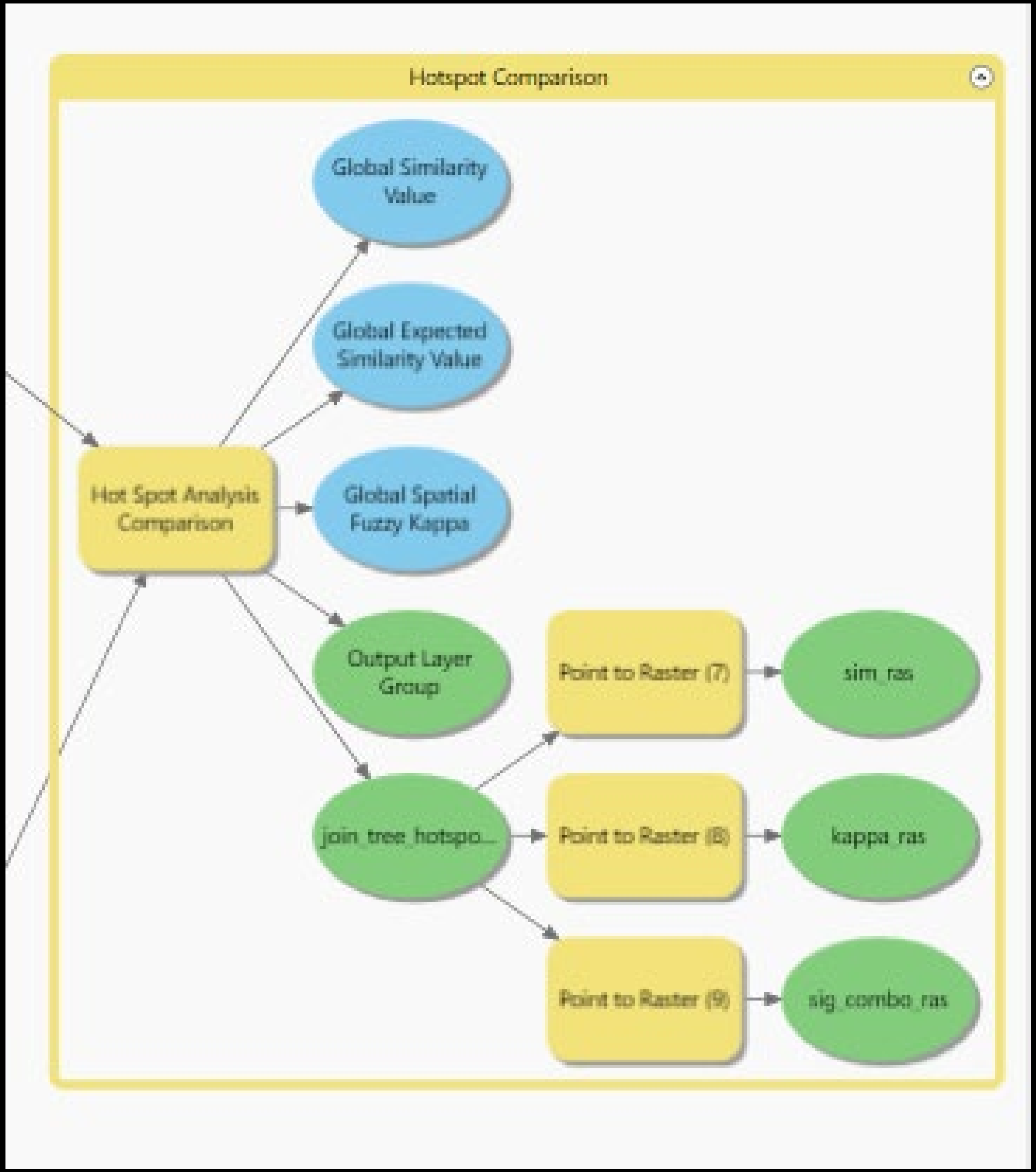
Optimized Hot Spot Analysis

- Perform optimized hot spot analysis on both temperature and tree cover datasets
- What is the optimum distance to define spatial neighborhoods for analysis?



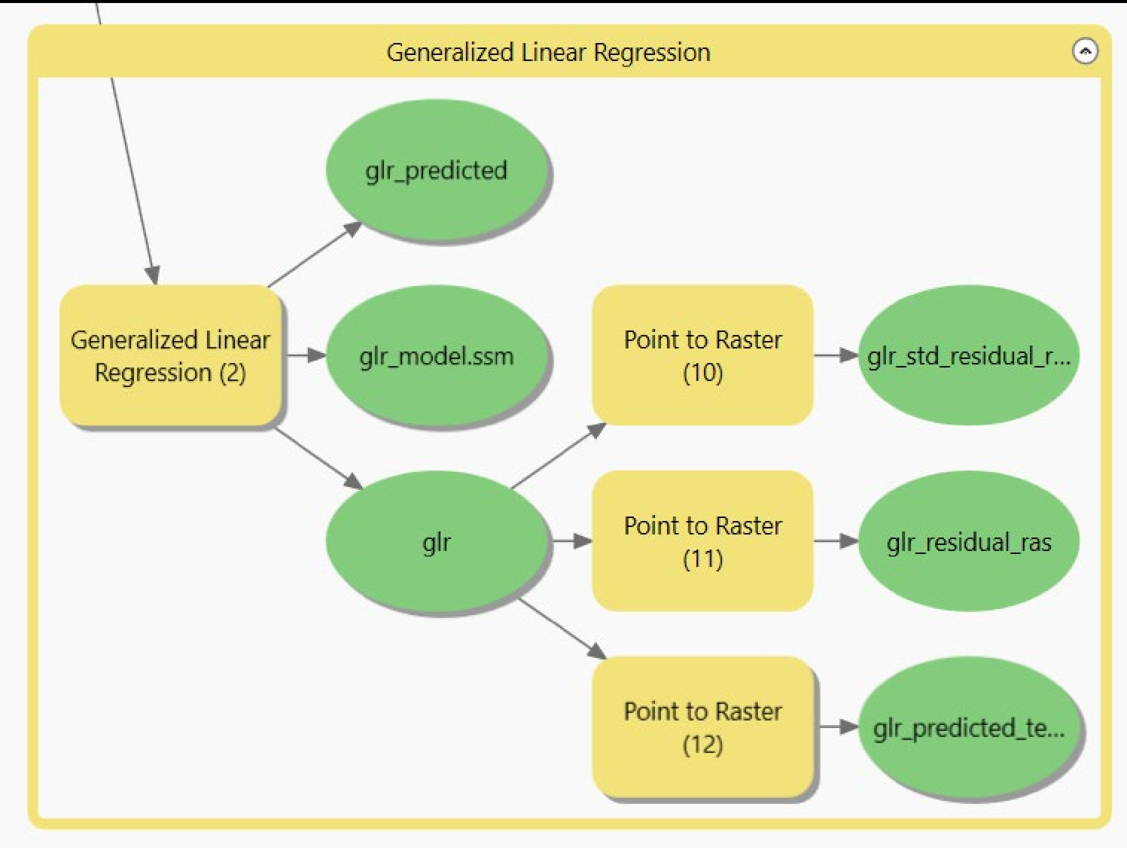
Hot Spot Comparison

- How are the statistically significant hot spots in the temperature data related to statistically significant hot spots in the tree cover data?
- Where are the areas of agreement and disagreement?
- Convert vector (point) results to raster for display



Generalized Linear Regression

- What is the relationship between the tree cover and the temperature?
- What is the best fit equation to describe this relationship?
- How good is this equation at making predictions?
- Where does this model work well? Where does it perform poorly?
- Convert vector(point) results to raster for display



Interpretation and Evaluation of Results

Percent tree cover and the temperature datasets display:

- Statistically significant positive spatial autocorrelation
- Statistically significant hot spots and cold spots
- Temperature dataset has almost no outliers
- Tree cover has relatively few outliers

- Hot spot between datasets are correlated
- Linear regression shows negative correlation between the percent tree cover and temperature
- Hot spot comparison shows strong association between temperature hot spots and tree cover cold spots

Challenges

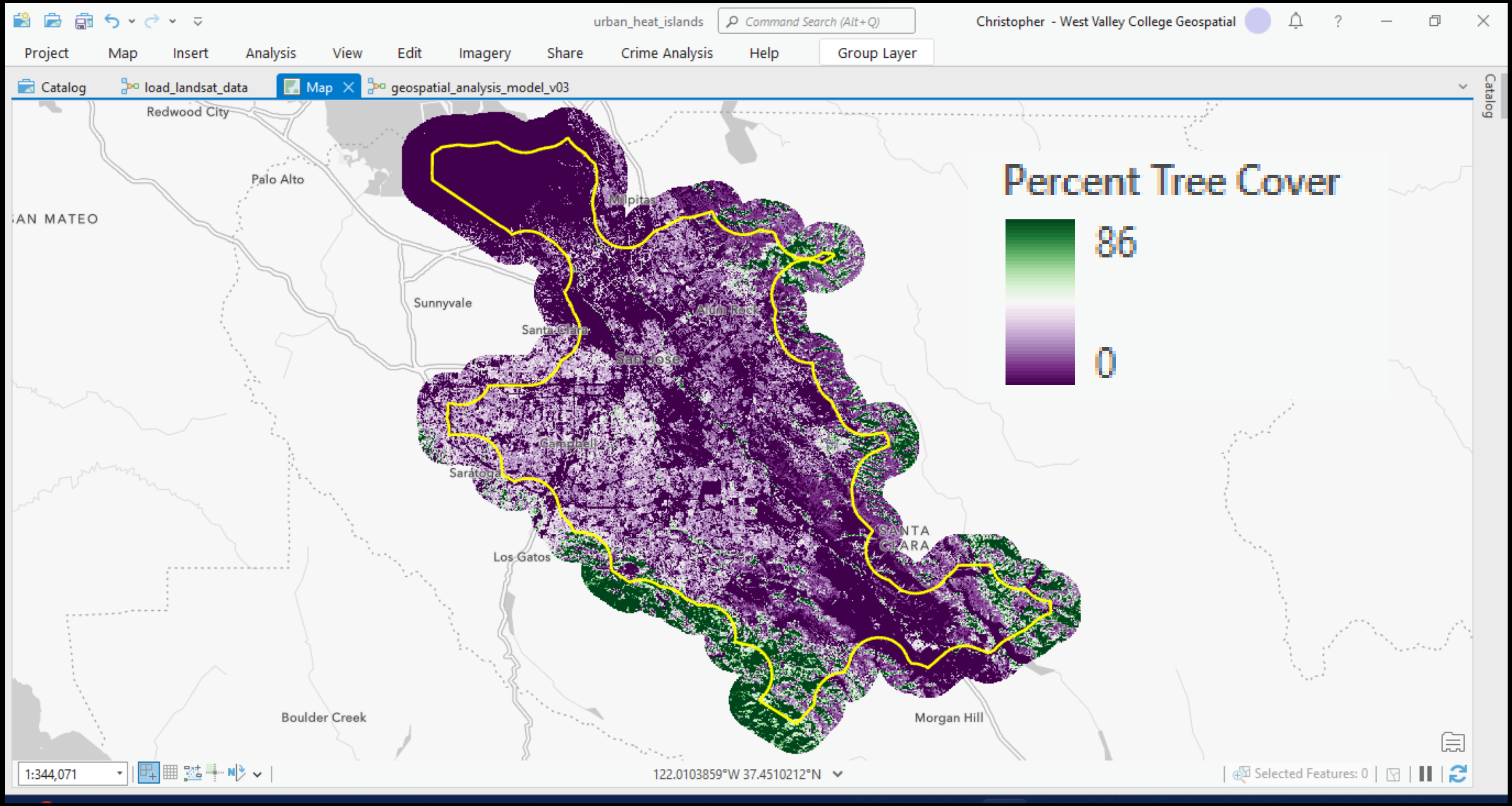
- Rater data does not align so a spatial join is required to prepare the data
- Large number of points (>1 million) created from 30 m rasters take a lot of resources to process
- Working with large point datasets is problematic for display (solution is to convert back to raster)
- Optimized analysis could not find an optimal distance – so, used a distance of 120 m
- GLR does not make good predictions for areas of water
- It was not possible to run GWR on this dataset (multicollinearity error)

Future Directions

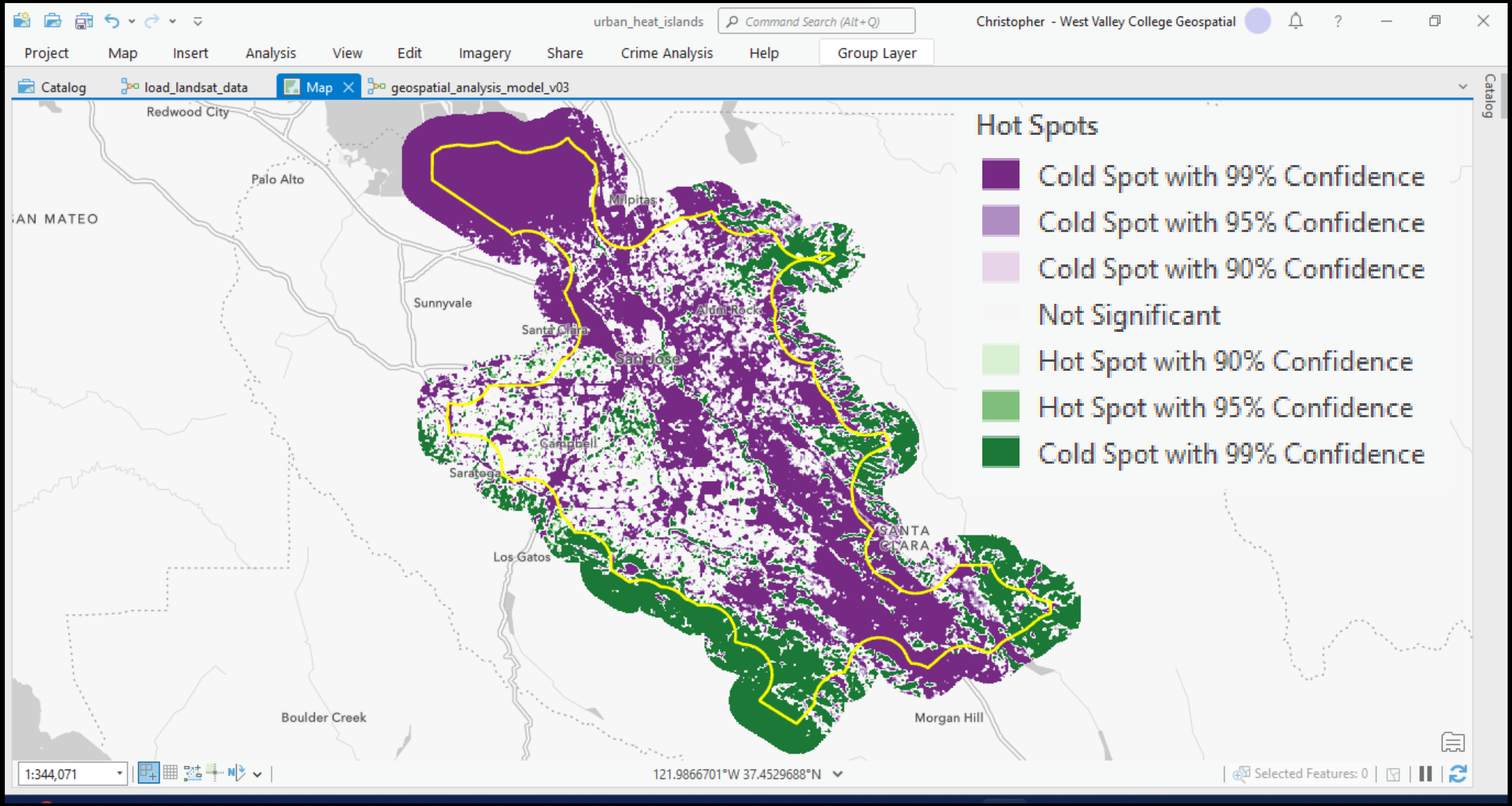
- Include additional factors in multivariate linear regression – candidate factors to explore might include elevation, aspect, and land cover type
- Rerun analysis with data for additional dates and different seasons
- Aggregate data into larger areal units, for example at census block level and rerun the analysis

Results

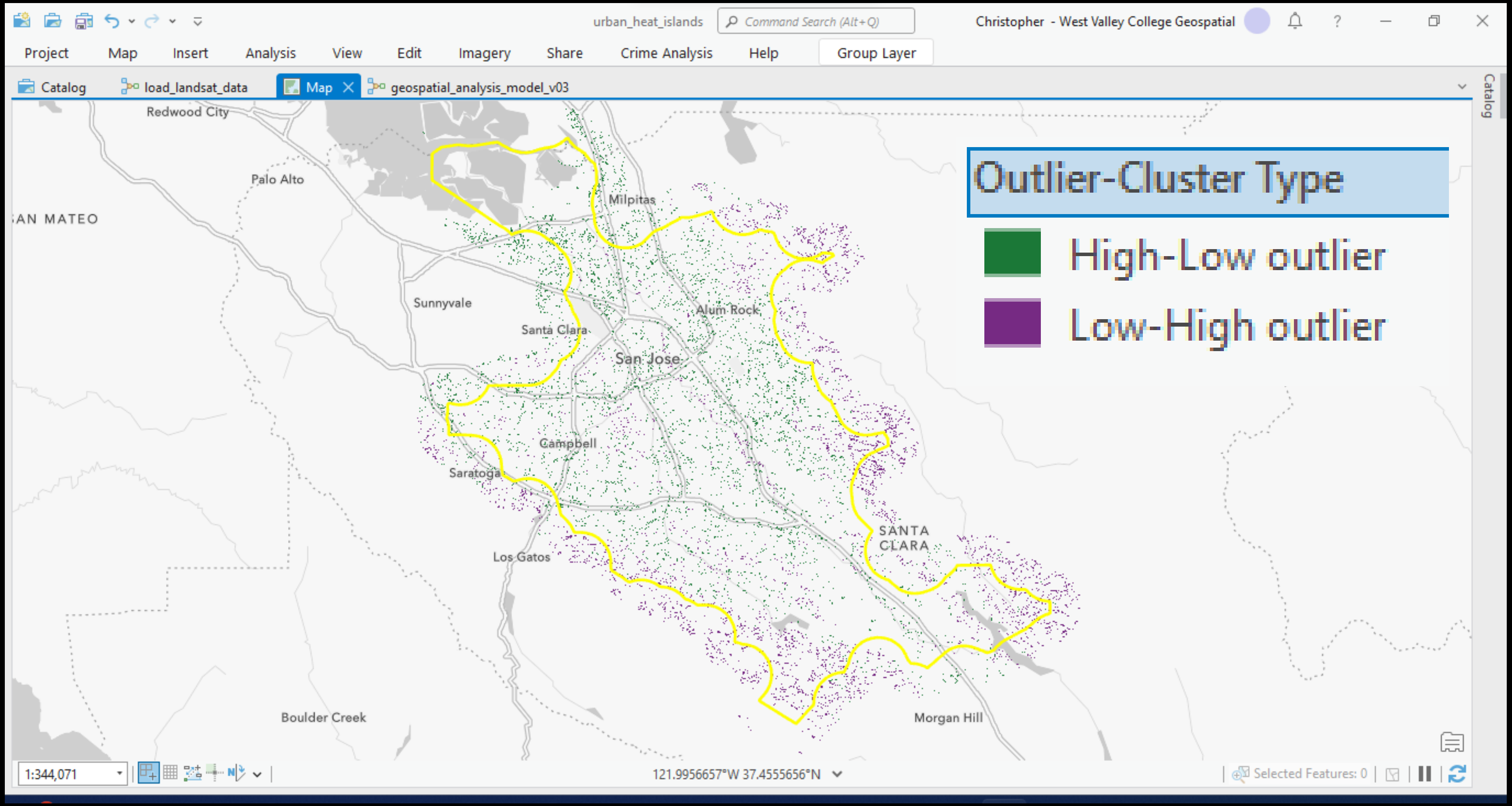
Percent Tree Cover -- Data



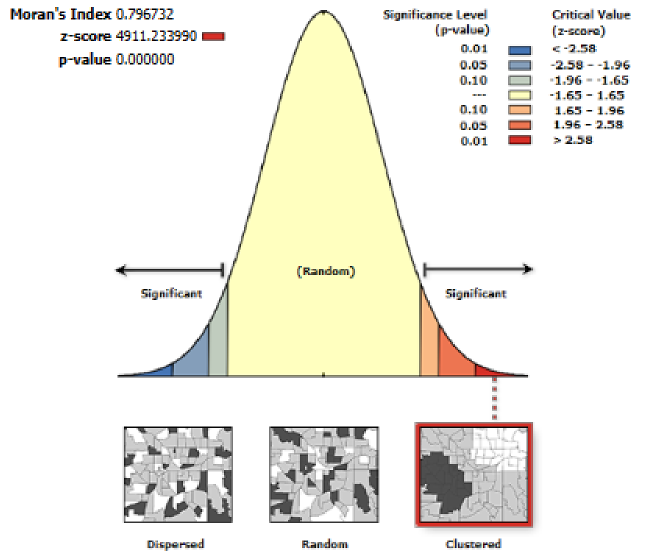
Percent Tree Cover -- Hot Spots



Percent Tree Cover -- Outliers



Percent Tree Cover -- Spatial Autocorrelation

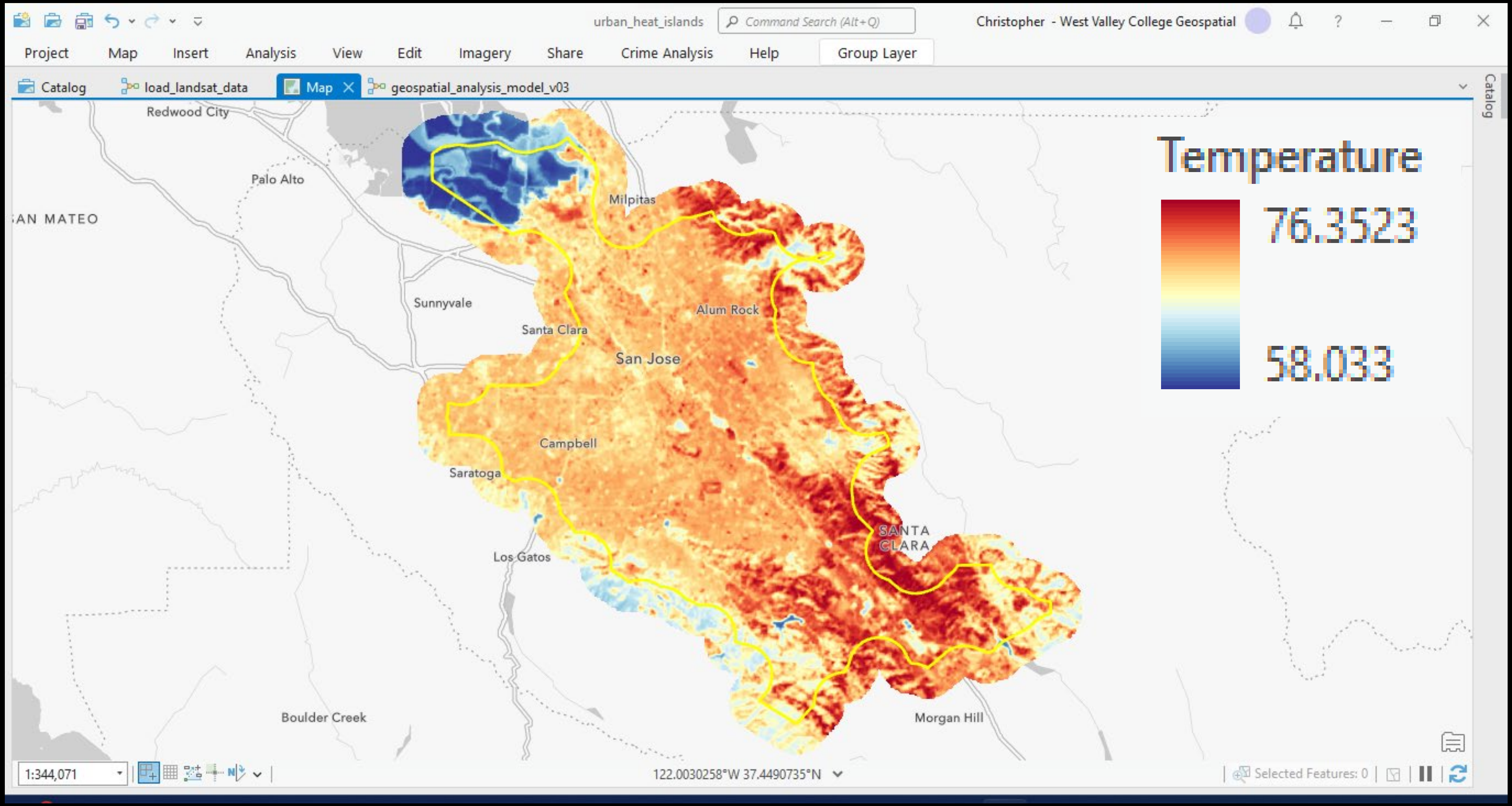


Given the z-score of 4911.23399, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

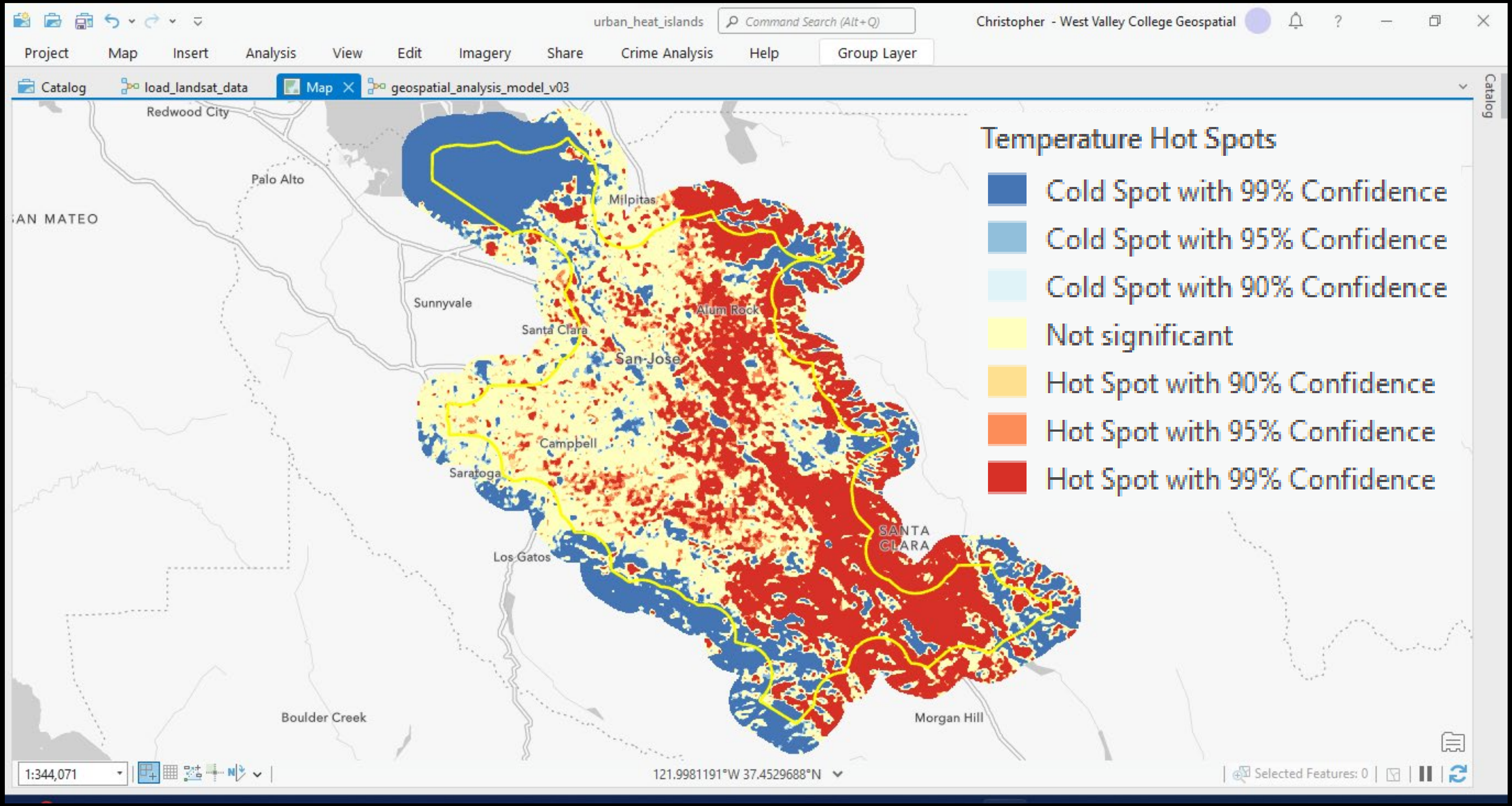
Global Moran's I Summary

Moran's Index	0.796732
Expected Index	-0.000001
Variance	0.000000
z-score	4911.233990
p-value	0.000000

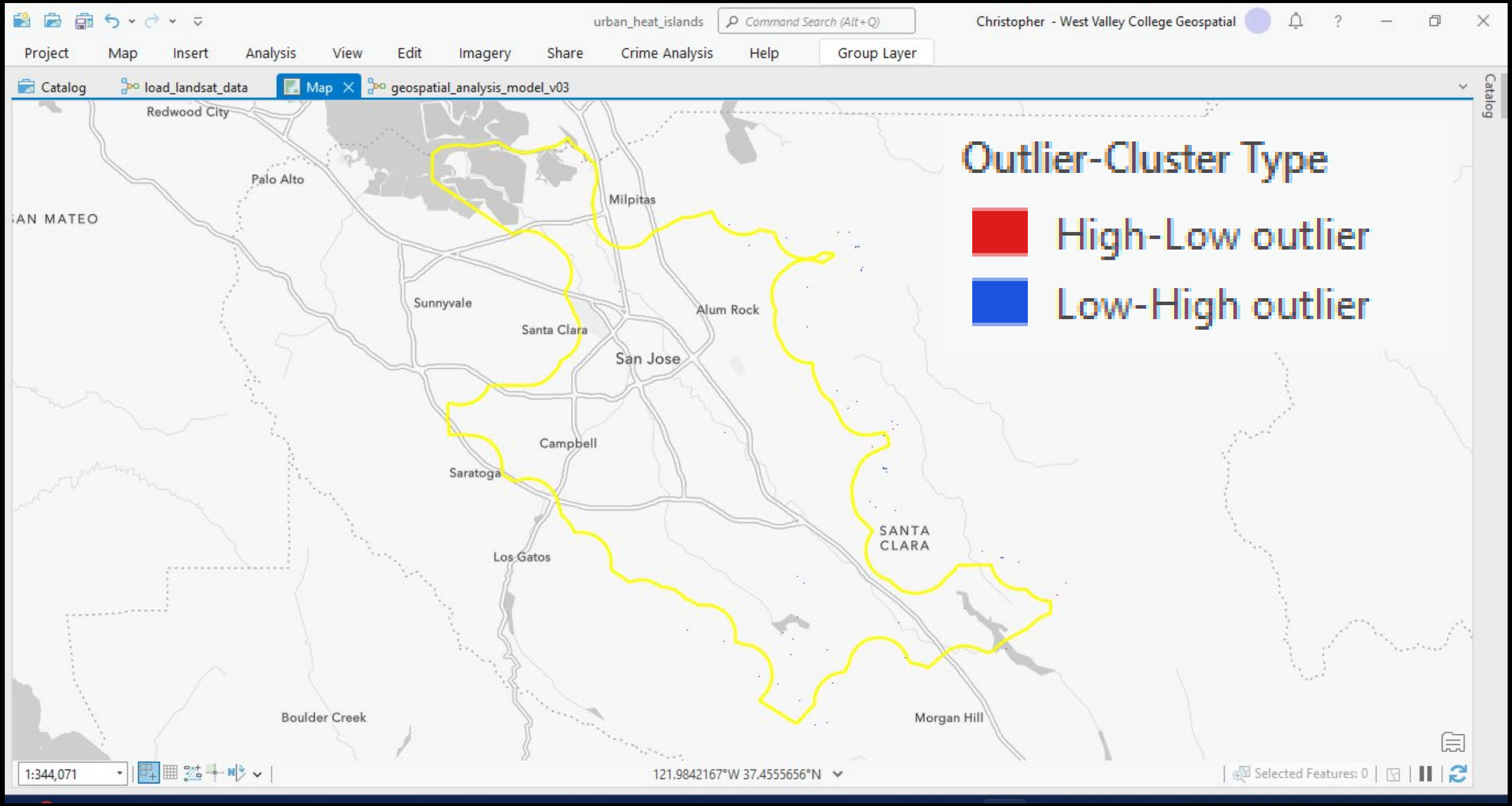
Temperature -- Data



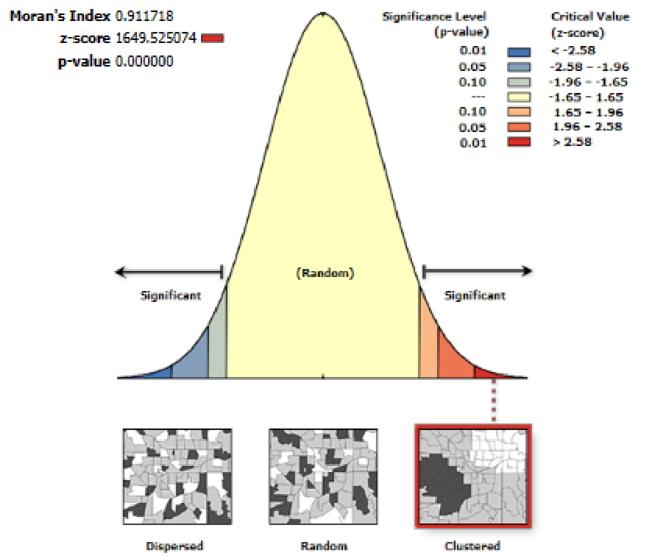
Temperature -- Hot Spots



Temperature -- Outliers



Temperature -- Spatial Autocorrelation

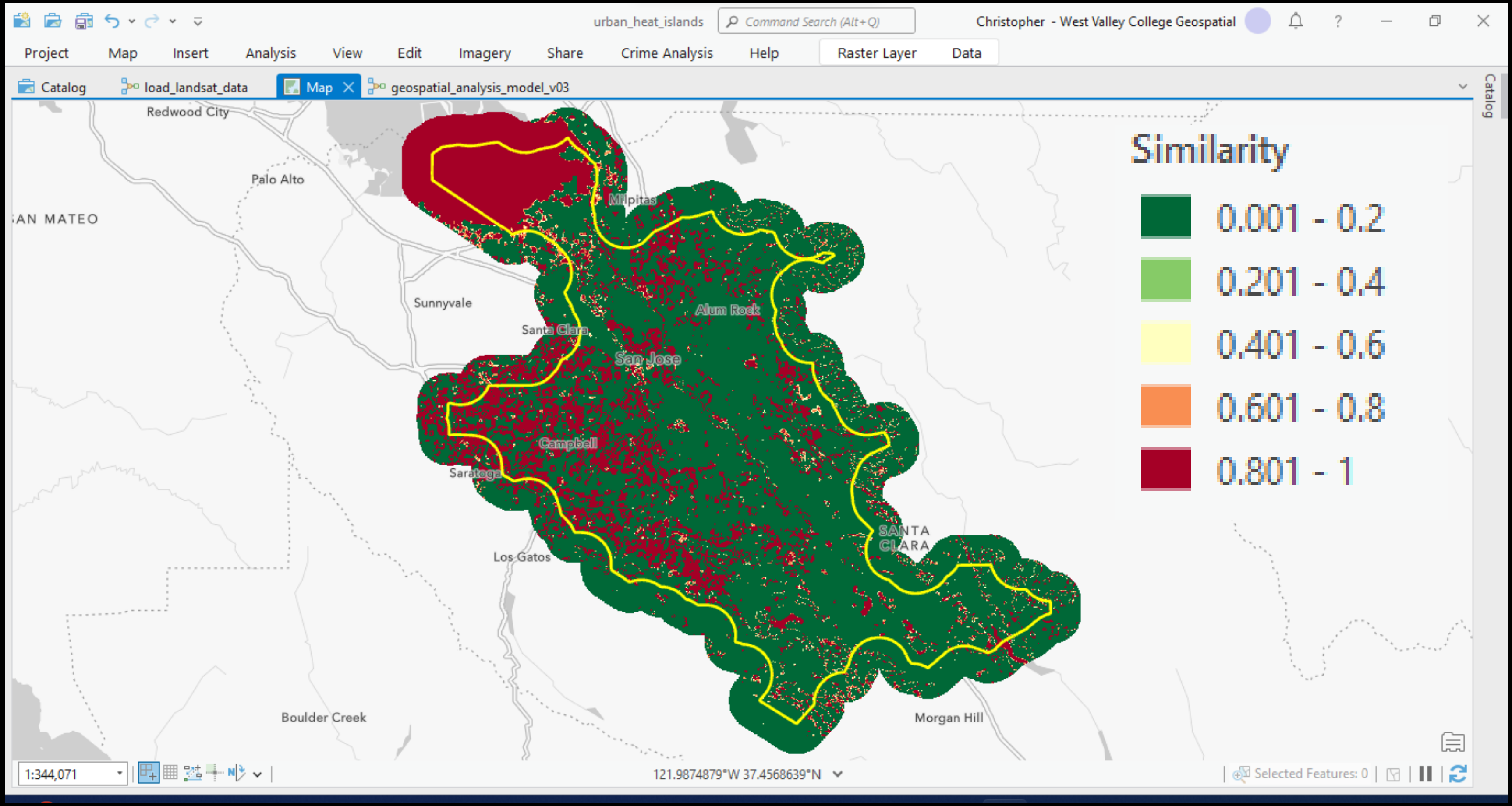


Given the z-score of 1649.525074, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

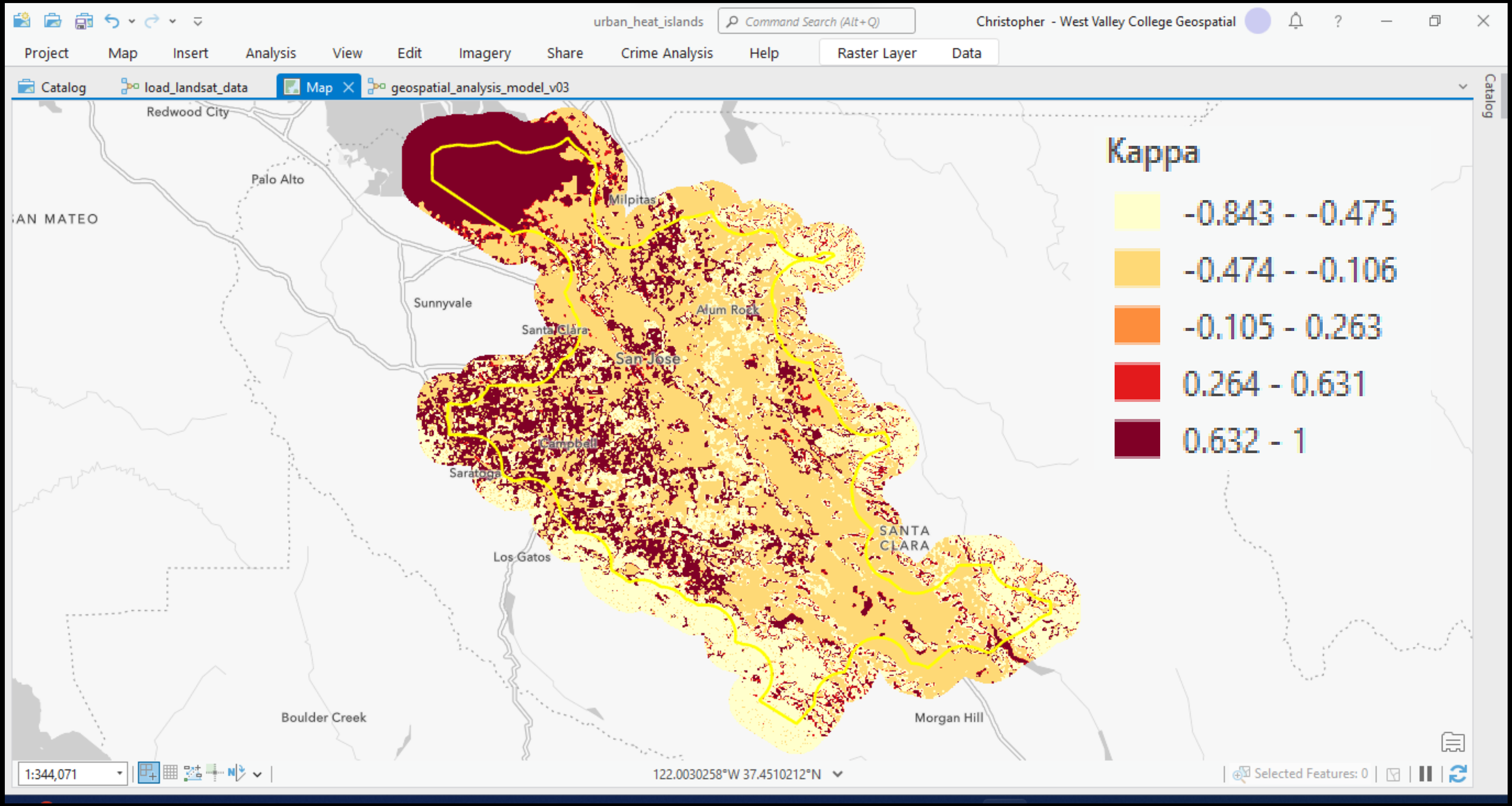
Global Moran's I Summary

Moran's Index	0.911718
Expected Index	-0.000001
Variance	0.000000
z-score	1649.525074
p-value	0.000000

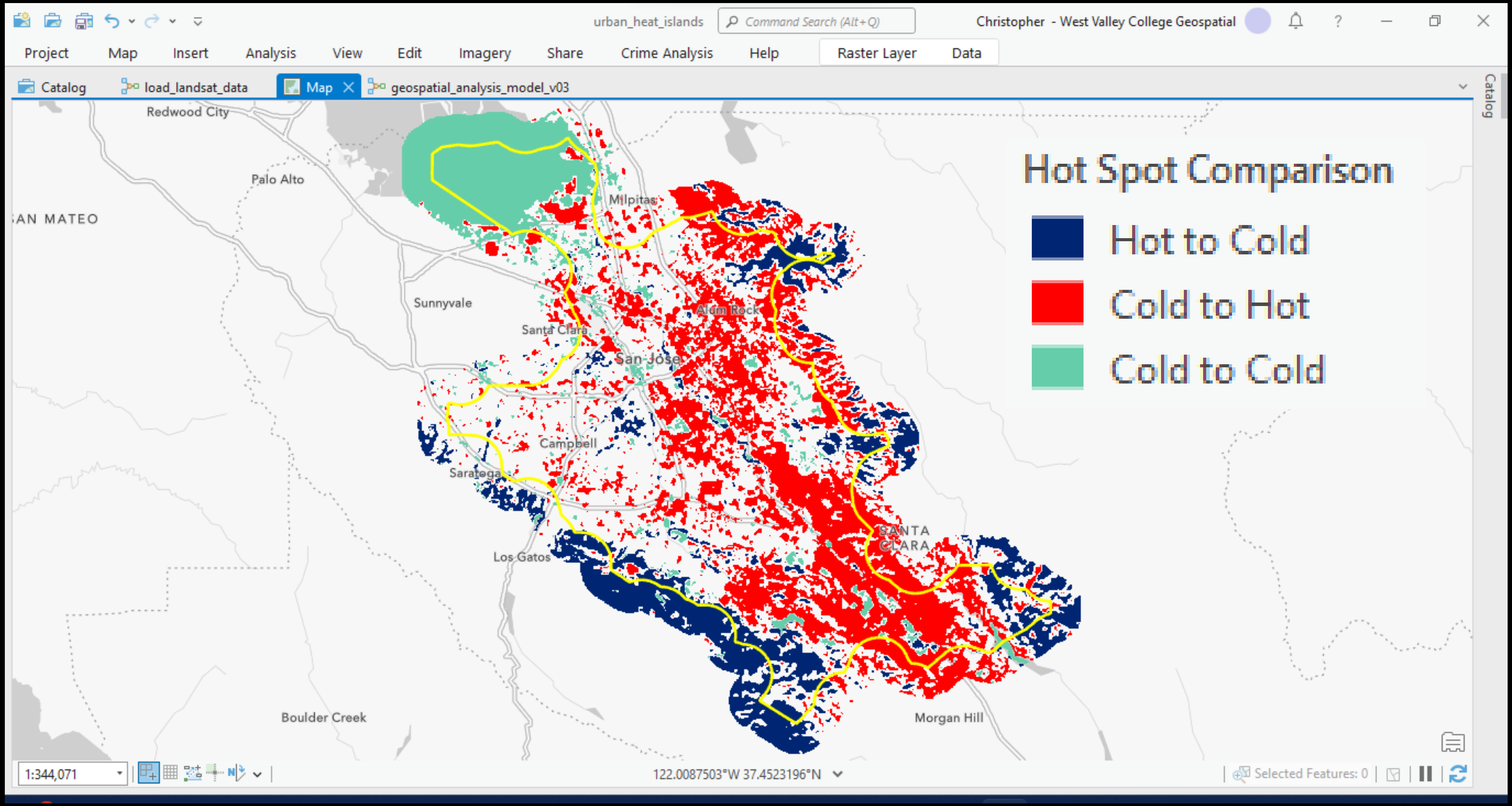
Hot Spot Comparison -- Similarity



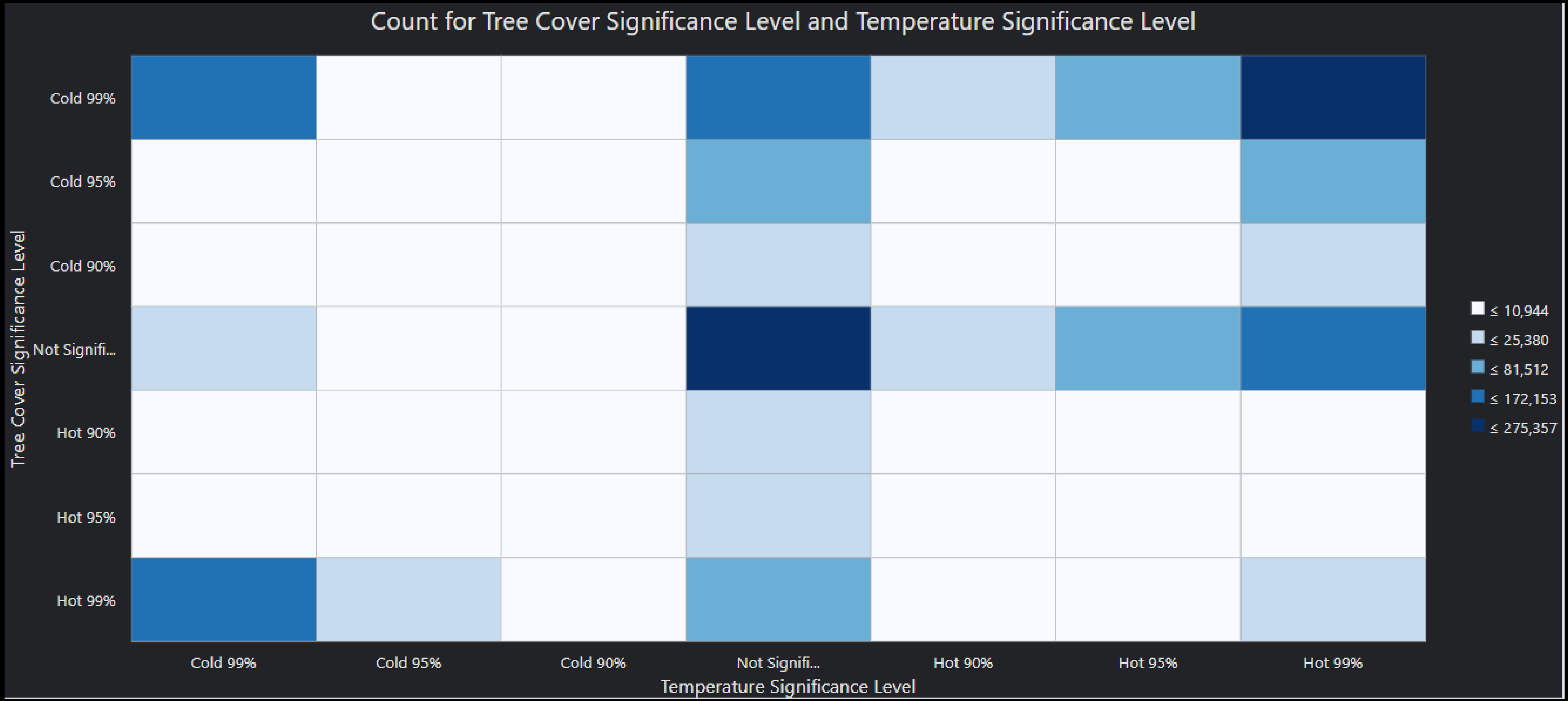
Hot Spot Comparison -- Kappa



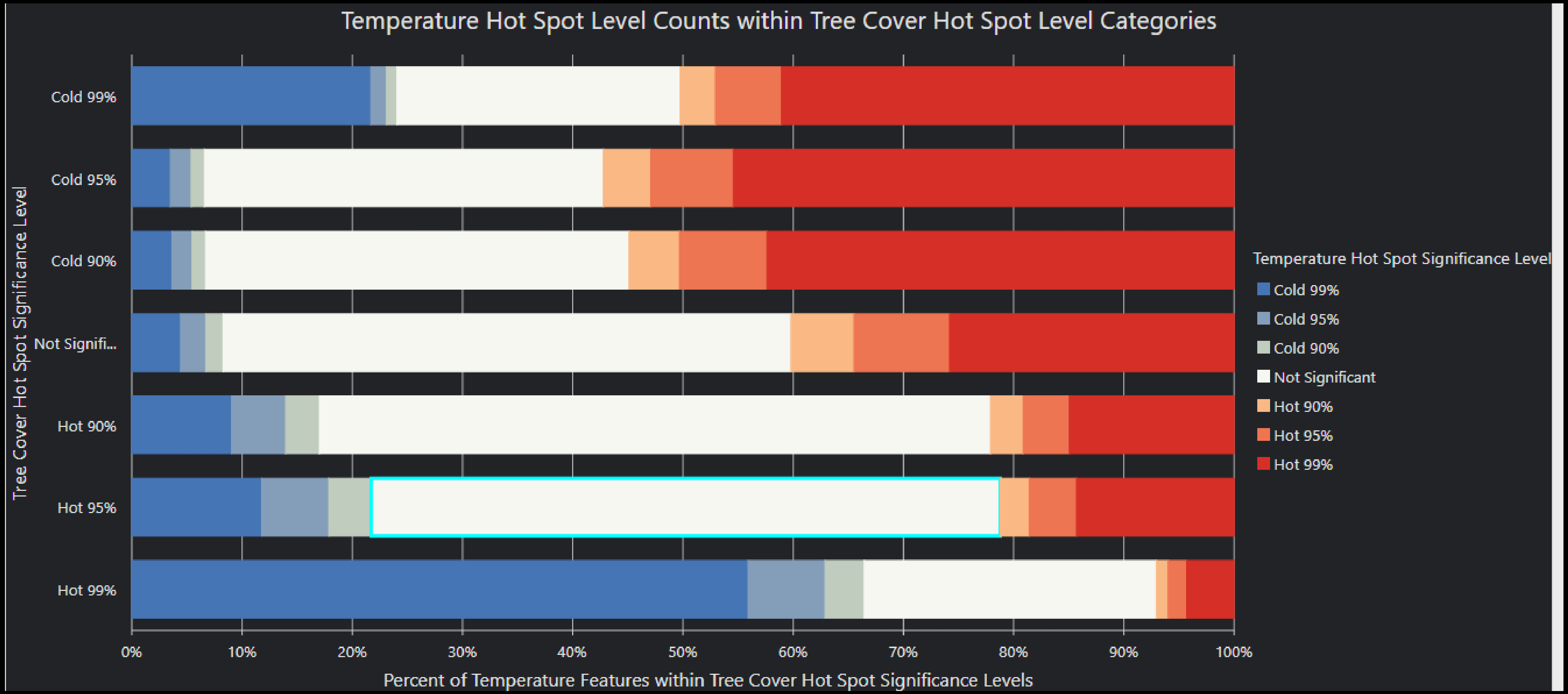
Hot Spot Comparison -- Categories



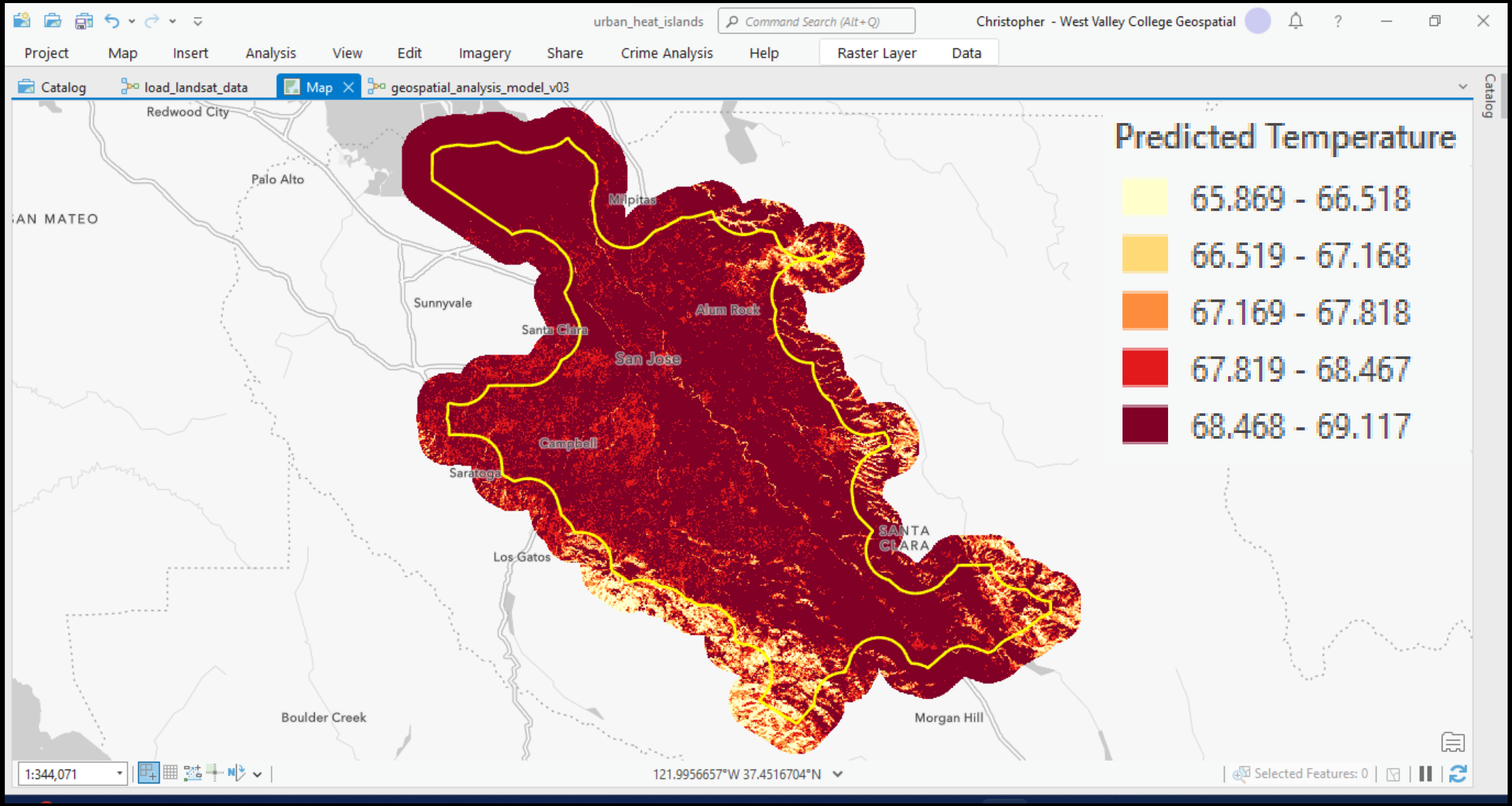
Hot Spot Comparison



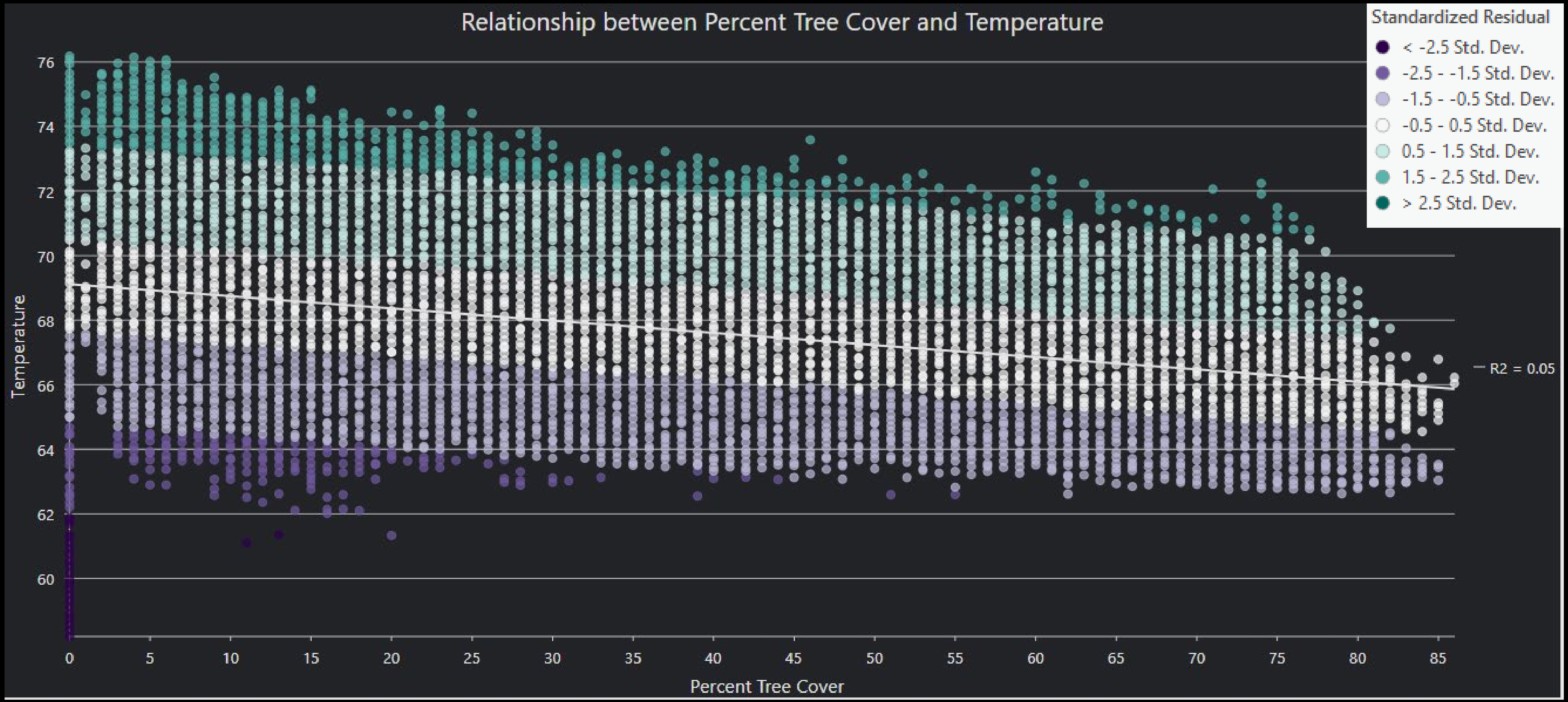
Hot Spot Comparison



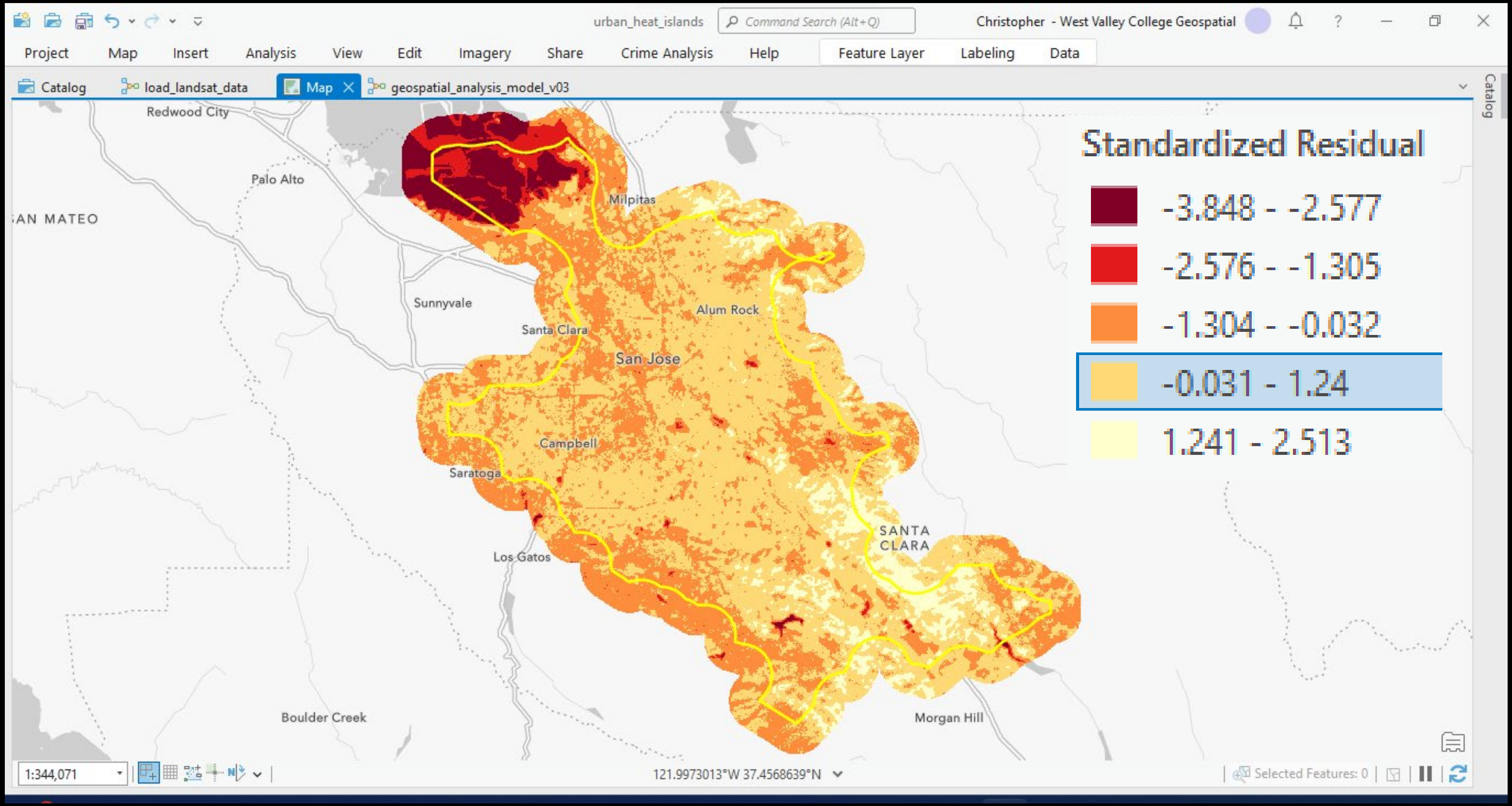
GLR -- Predicted Temperature



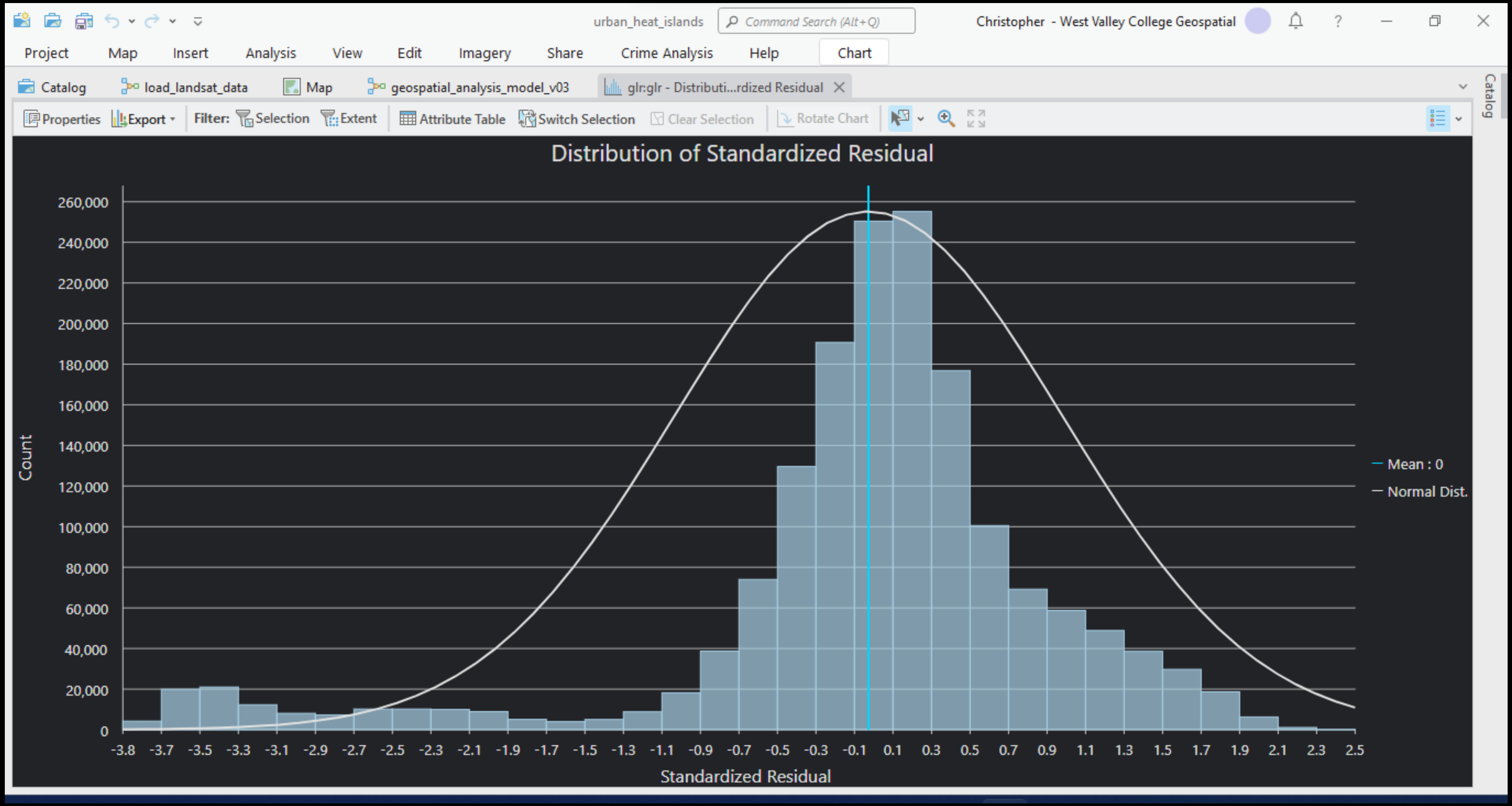
GLR -- Relationship between Variables



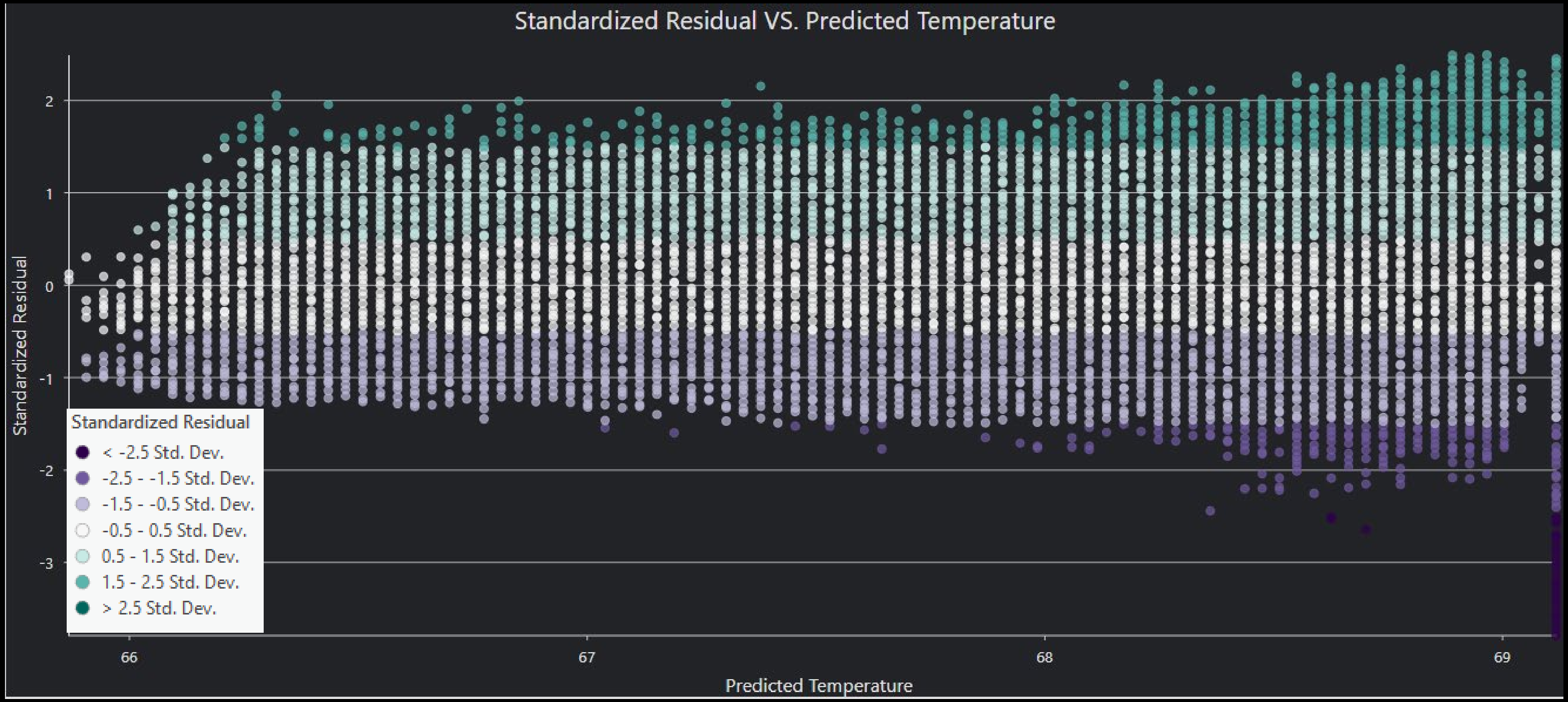
GLR -- Standardized Residuals



GLR -- Distribution of Standardized Residual



GLR -- Residual vs. Predicted Temperature



Decision

Analysis found:

- Statistically significant clustering (spatial autocorrelation) in both tree cover and temperature input datasets
- We can reject the null hypothesis that the distribution of temperature is random (p-value = 0, z-score = 1649.5)
- Hot spot comparison show there is a strong relationship between tree cover and temperature (spatial fuzzy kappa = -0.0435 shows hot spots for tree cover are strongly associated with cold spots for temperature)
- Ordinary least squared regression found a weak negative correlation between tree cover and temperature (R^2 value of -0.05 is not conclusive)