Tyler Gates  
GEOG:3050  
Final Project

Prevalence of Crime Near Public Transit Stations

*Goal*

The goal of this project is to investigate the occurrence of crimes within the area of public transit stop locations. The main algorithm used for this analysis is the Colocation Analysis tool included in the latest release of ArcPro (V 2.5). This tool measures local patterns of spatial association between two point categories or shapefiles. This script is being developed in hopes to determine which cries occur most often in which area, if stops need to add security, or to find out which crimes are most common in certain areas of the city. I believe that there will be collocated points around the most-used transit stops around a city of interest.

*Methodology*

A screenshot of a cell phone

Description automatically generated

Figure 1: Workflow for colocation analysis

This study begins with a crime statistic csv file, including fields with x and y positions, specific crime types, and community codes. It also begins with a point file containing locations of public transit stations, and a shapefile defining regions of a city which can be determined by the user. All point files are clipped to the size of the city shapefile in case any routes extend out of the boundaries. From the xy coordinates, a temporary event layer is created, and transformed into a feature class to be permanent. The crime type of this xy point data is describing is determined by the user during the function input. These categories can be found by filtering the field in the excel sheet prior to executing the function. This is also clipped to the size of the city shapefile.

The colocation analysis is ran on both the crime data points and the public transit points. The category of interest in this analysis are the crime data points, while the neighboring categories are the transit locations. A distance band of 1000 feet, or around 2 blocks, is used as a spatial relationship between points. This ensures that only points within a defined area around the stops will be considered.

After the colocation analysis had been run, they were filtered to only show the significantly collocated points. These points have a p-value of <0.05, lower than any other category. These are filtered by using the field LCLQTYPE, whose value is 0 for this particular category.

A spatial join is performed on the city shapefile and the new, filtered point file in order to see how many significant points are located within each neighborhood. These results are placed into a new field, which is designated as a short integer field – there will only be whole numbers involved, so there is no need to create a float or double type field. Using this data, the city shapefile can be colored on a gradient scale in order to display which neighborhoods are most likely to have the chosen type of crime to occur near a public transit stop.

*Data and Descriptive Statistics*

The example used in this study is performed on data taken from the Chicago Data Portal. The neighborhood shapefile includes all the historic neighborhoods of the City of Chicago. The transit locations being used are the elevated train line stops, and the crime statistics are from 2019 , with a focus on a portion of assault cases (csv lines 0-1000).

There were a total of .34% of points which had been determined to be significantly collocated using the specific filters and inputs. In this particular portion of the dataset, there did not seem to be much colocation between assault crime occurrences and train stops in Chicago.