Consolidated Missing Persons Database for Spatial Analysis

Law enforcement personnel and community activists tend to lack comprehensive, consolidated, data-driven tools to address cases involving the “critically missing.” According to District of Columbia (DC) Metropolitan Police Department’s (MPD) missing persons [website](https://missing.dc.gov/), a "critical missing person" is defined as any person under the age of 15 or over the age of 65, or anyone that, based on the specific circumstances (e.g., mentally incapacitated, patient who presents an imminent danger to him/herself or others, in a life threatening situation, real or suspected danger of foul play, etc.), is designated as such by the Patrol District’s Watch Commander.” In an effort to improve public communications regarding missing persons and hopefully expedite safe return, in 2017 the DC MPD instituted a practice of posting on Twitter whenever an individual is reported as “Critical Missing.” As with any new effort to publicize information on crime or other social problems, it can create the perception of an upward trend or surge of such issues. A lack of national consolidated data for comparative context can add to the confusion.

Similar to the [Opioid Mapping Initiative](http://opioidmappinginitiative-opioidepidemic.opendata.arcgis.com/) that utilizes spatial data to help first responders proactively address overdoses, a consolidated database of spatiotemporal data on missing persons can help law enforcement and advocacy groups identify trends in missing persons data and potentially investigate incidences of exploitation such as human trafficking or serial crime.

This project will leverage machine learning techniques, web scraping, and other tools in order to build data pipelines to populate a proof-of-concept consolidated missing persons geodatabase. This database will include spatiotemporal and demographic data on missing persons from major cities during a particular period of time (i.e. summer). Once compiled, geospatial statistical concepts such as Spatial Joins, [**Nearest Neighbor Analysis**](https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-average-nearest-neighbor-distance-spatial-st.htm)**,** [**Getis Ord Gi Analysis**](https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm), [**Point Density Analysis**](https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/point-density.htm)**,** [**Kernel Density Estimation**](https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-kernel-density-works.htm) **and** [**Regression**](http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Regression_analysis_basics)to test the hypothesis that there are a disproportionate number of critically missing women in major cities in the United States.