School Safety Sip Analysis Plan

Step 1: Why?

Distribution of school safety's work by different demographic grouping

- Purpose: To identify which communities are underserved

Step 2: How?

1. Identify all data sources necessary to complete analysis:
   * SIP Project data (why? To identify school safety projects)
   * SIP geometry data (why? To join to lion)
   * Lion (why? To join to NTA and grabbing project mileage)
   * Neighborhood Tabulation Area (NTA) data  (why? To obtain demographic data for sip project groupings)
   * NTA descriptions
2. Where are all the data sources:
   * SIP project data: public.sip\_proj in ris database
   * SIP project geom data: public.sip\_projects\_geo in ris database
   * Lion: lion in ris database
   * NTA map data: public.districts\_neighborhood\_tabulation\_areas in his database
   * NTA desc: working.v\_nta\_summary\_map in crashdata database
3. Execution:
4. Move NTA data to ris db so that all data is in same db
5. Limit sip project selection to only school safety projects; where unit\_desc = ‘School Safety’
6. Join sip projects to sip projects geo
7. Join sip projects geo to lion on geom
8. Join lion to NTA on nta codes
9. Obtain Demographic information:
   * Questions:
     1. What percent of sip mileage is in high poverty, low poverty
     2. What are race/ethnicity quintiles by sip mileages
   * Execution:
     1. Group sip data with working.v\_nta\_summary\_map with poverty column data
     2. Group sip data with working.v\_nta\_summary\_map with race column data
10. Summary:
    * From compiled data, get sum of different categories in NTA descriptions and group by sip project
    * Include sip project length by taking sum of st\_length of sip geometry grouped by sip project
11. Future: Normalize by school age population