CE C265 & PBHLTH 285

Spring 2025

Assignment 4: Crash Clustering

In this assignment, you will create a crash density plot and identify clusters of crashes to pinpoint locations with unusually high crash counts. Your goal is to locate areas where distracted driving or driving under the influence has led to crashes with serious outcomes.

You will use data from the Transportation Injury Mapping System (TIMS): https://tims.berkeley.edu. Detailed information about the dataset is available here: https://tims.berkeley.edu/help/SWITRS.php.

1. Select a Study Area:
   Choose a large city along with a neighboring smaller city (e.g., Oakland and Berkeley).
2. Review the Data:
   Examine the dataset and variable definitions to understand the available fields.
3. Filter and Visualize:
   a. Use the provided code as a starting point to.
   b. Filter for high-severity crashes (Fatal, Serious, and Visible Injuries) involving either distracted driving or driving under the influence.
   c. Generate a density plot (e.g., via Kernel Density Estimation) to visualize crash concentrations.
   d. Identify and highlight clusters with unusually high crash counts.
4. Report:
   Submit a concise report that includes:
   a. A narrative summary of your findings.
   b. The Python code used for data processing and visualization.

You may complete this assignment individually or in pairs.

Note:

The spatial data from TIMS is provided in latitude and longitude. Ideally, these coordinates should be projected to a geographic coordinate system for accurate spatial analysis. However, for simplicity, you may use the Point_X and Point_Y fields without projection.