6.S097 Project

Understanding Crime & Traffic for Safer Urban Planning

Subby Olubeko
Marwa Abdulhai
Mark Wang
Kevin Lyons
Problem

- Find clusters of crime within Boston and see how these areas are correlated to traffic patterns.
- By finding areas where crime is prevalent, will be able to identify places which need more security/enforcement/etc.
Algorithm and Implementation

• Used the Boston crime data database provided on the 6.S097 website
• Divided crime data into several categories, then ran K-Means clustering with various K
• Shaved the clusters to find areas where crime is prevalent
• Combined with public Boston traffic data to find patterns and inform rider safety
Disruptive
Violent
Performance

• Original goal was to use hierarchical clustering, but runtime up to $O(n^2)$
• Impractical for large sets $n > 10,000$
• Went with focused K-means on crime set
Conclusions

• High Traffic Areas
  – North End
  – Boston University/Kenmore
  – South Boston (near UMass, Boston)
  – South End
  – Roxbury
Combining Data
Results/Future

• Seems to be a positive correlation between traffic volume and the crime rate in a given area.

• Could be due to number of factors – population density, lots of human interaction, economic activity.

• Could use additional datasets/clustering techniques.
6.S097 Project

Understanding Crime & Traffic for Safer Urban Planning

Subby Olubeko
Marwa Abdulhai
Mark Wang
Kevin Lyons