

Predicting Fire Risk in Atlanta

Data Science for Social Good- Atlanta Fire Rescue Department



Team Members: [Xiang Cheng](#), [Oliver Haimson](#), [Michael Madaio](#), [Wenwen Zhang](#) | Advisors: [Polo Chau](#), [Bistra Dilkina](#) | Partner: [Matt Hinds-Aldrich](#)

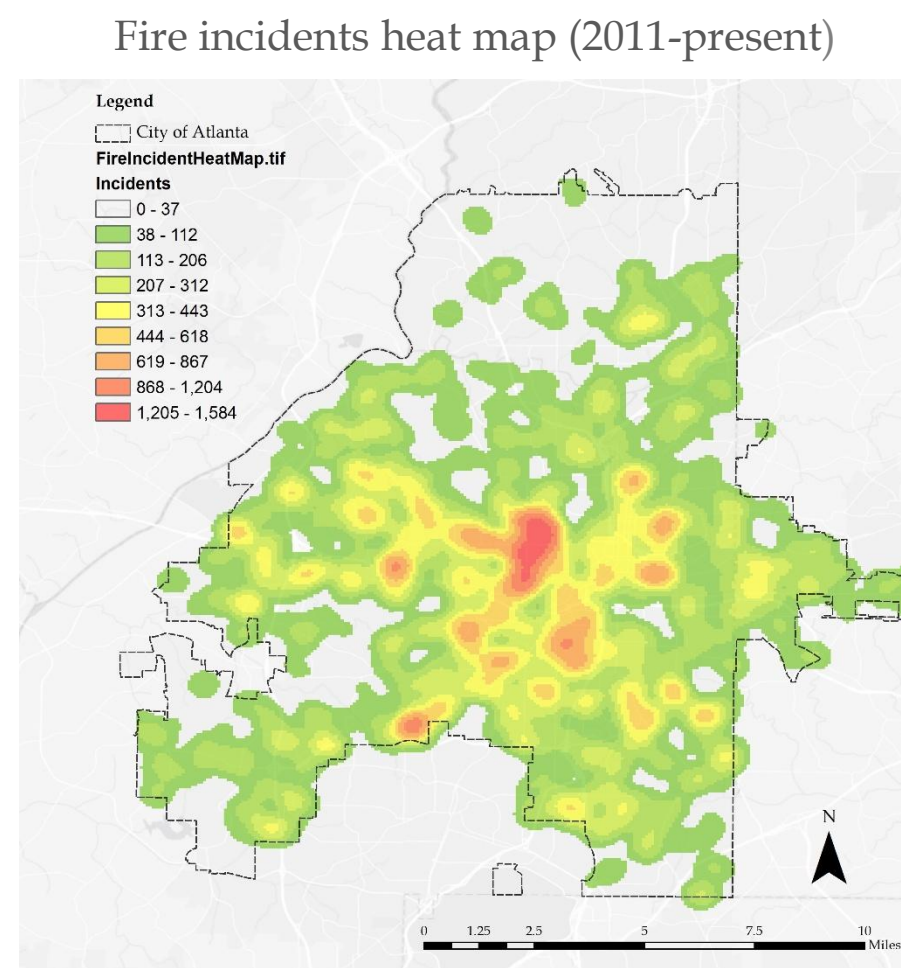
xcheng7@emory.edu ohaimson@uci.edu mmadaio@gatech.edu wzhang300@gatech.edu

polo@gatech.edu bdilkina@cc.gatech.edu

mhinds-aldrich@atlantaga.gov

Summary

Hundreds of buildings have fires in Atlanta every year. The Atlanta Fire Rescue Department (AFRD) attempts to reduce fire risk by inspecting buildings for potential hazards and fire code violations, but they currently only inspect a subset of the total buildings needing inspection. Our project aims to reduce fire risk in Atlanta by identifying and prioritizing buildings that should be inspected by AFRD.



Goal 1: Find new properties to inspect

- List of new properties: from external business and property databases
- Prioritized list: using risk scores from the model
- Interactive map to view inspected properties, fire incidents, and potential inspections in Atlanta

Goal 2: Prioritize inspections

- Integrated database of buildings with the most complete property information
- Make a predictive model to generate risk score for properties

Approaches

Geographic Information System (GIS)

We used GIS to join and merge various datasets. The merged datasets can help us to identify the type of businesses that require permits and understand the building features that may explain the occurrence of fires.



Machine Learning Algorithms

We are using several machine learning algorithms to predict fire incidents. Linear Regression, Logistic regression, random forest tree, and Support Vector Machine (SVM) models are all tested and SVM is selected as the most appropriate model. More than 250 variables are tested and 58 variables are included in the final results.



D3.js Data Visualization

To make it easier to interpret the model results, we are going to use d3.js library to visualize the model and analysis results.



Data

SIZE: >2 GB, ~160,000 Records

Data	Source
Fire Incidents (AFRD)	Atlanta Fire Department
Fire Permits Data (FSAF)	Atlanta Fire Department
CoStar Data	Atlanta Fire Department
Neighborhood Planning Unit	ARC*
Parcel Data	City of Atlanta
Atlanta Business License	City of Atlanta
SCI Data	City of Atlanta
Business Location Data*	Google Places API
Address Based Coordinates	Google Geocode API
Demographic Data	U.S. Census Bureau
Socio-economic Data	U.S. Census Bureau

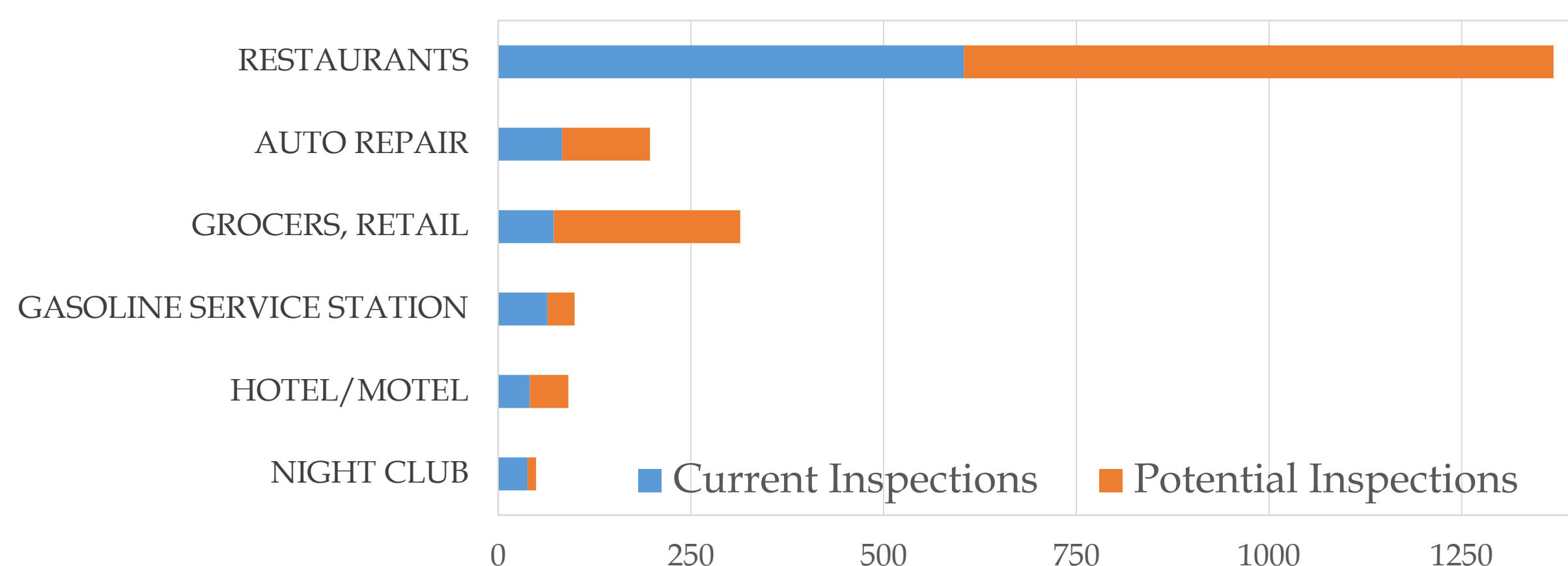
*ARC: Atlanta Regional Commission
*The type of business (and number) obtained from Google Place API includes: Bars (629), car repairs (686), laundry (280), liquor stores (117), night clubs (163), painter (94), plumber (167), restaurant (1661), roofing contractors (139), schools (617), and hospitals (441).

Final Results

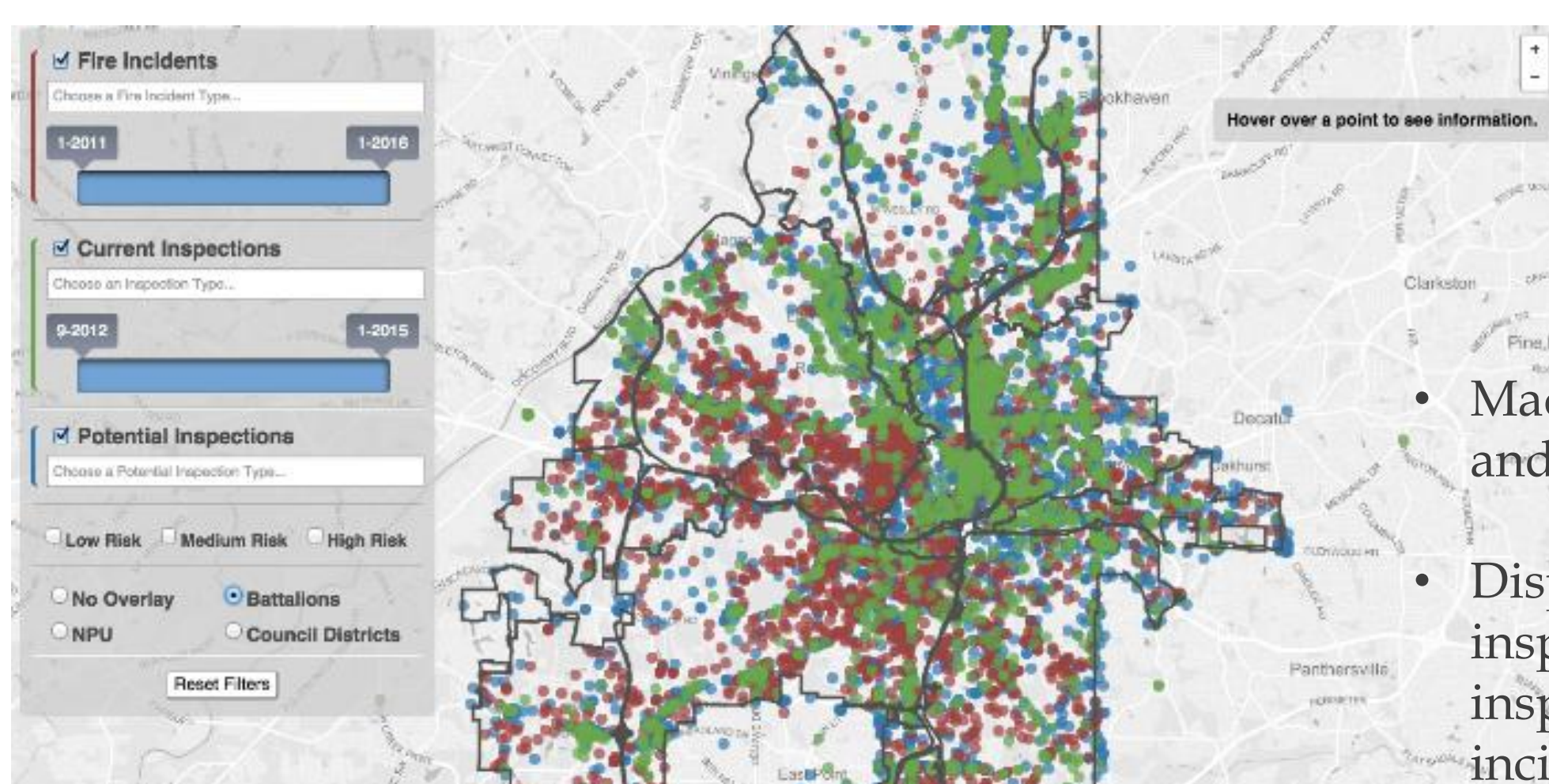
Finding Potential Inspections (Goal 1)



Top 6 Current Inspected Business Types and Potential Inspections

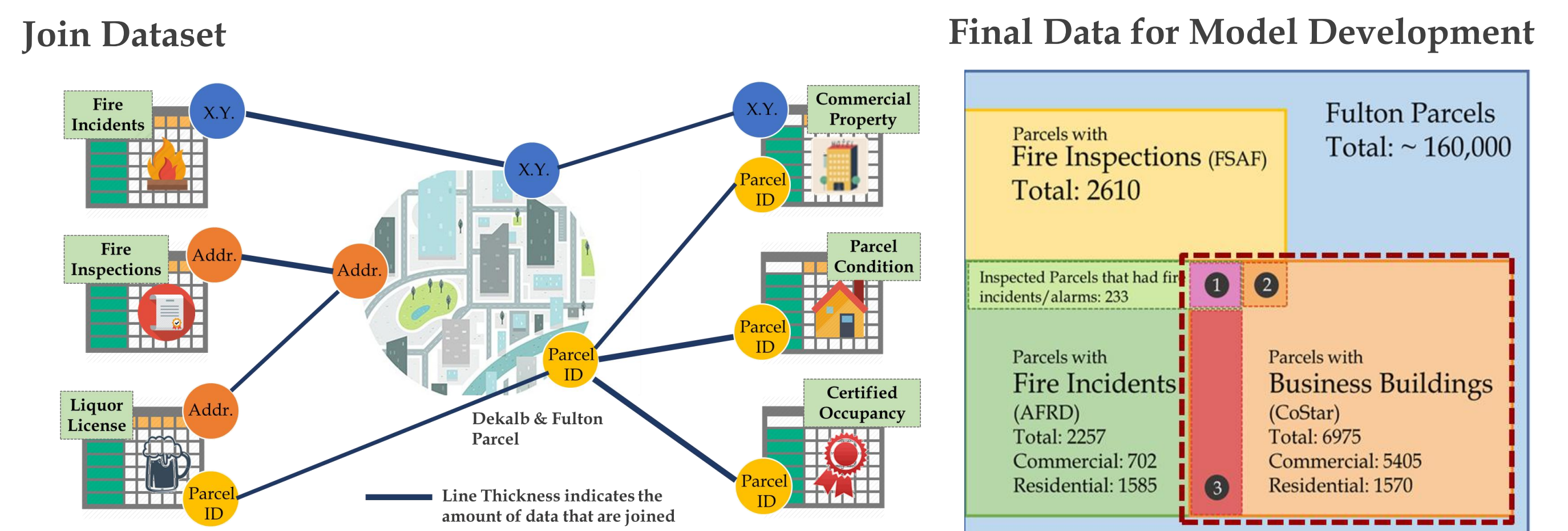


Interactive Inspection Map



- Made with D3, Leaflet, and Mapbox
- Displays the current inspections, potential inspections, and fire incidents

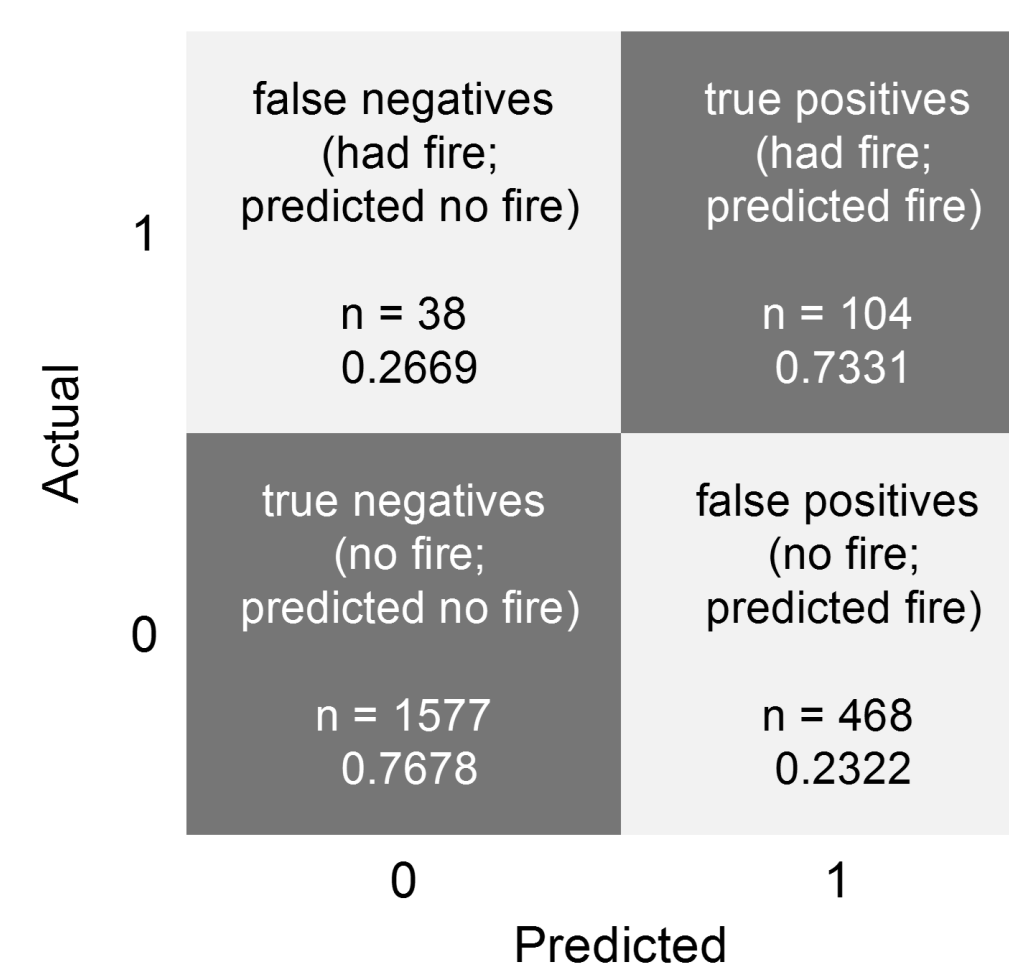
Predictive Fire Risk Model (Goal 2)



Machine Learning Results: Support Vector Machine Model

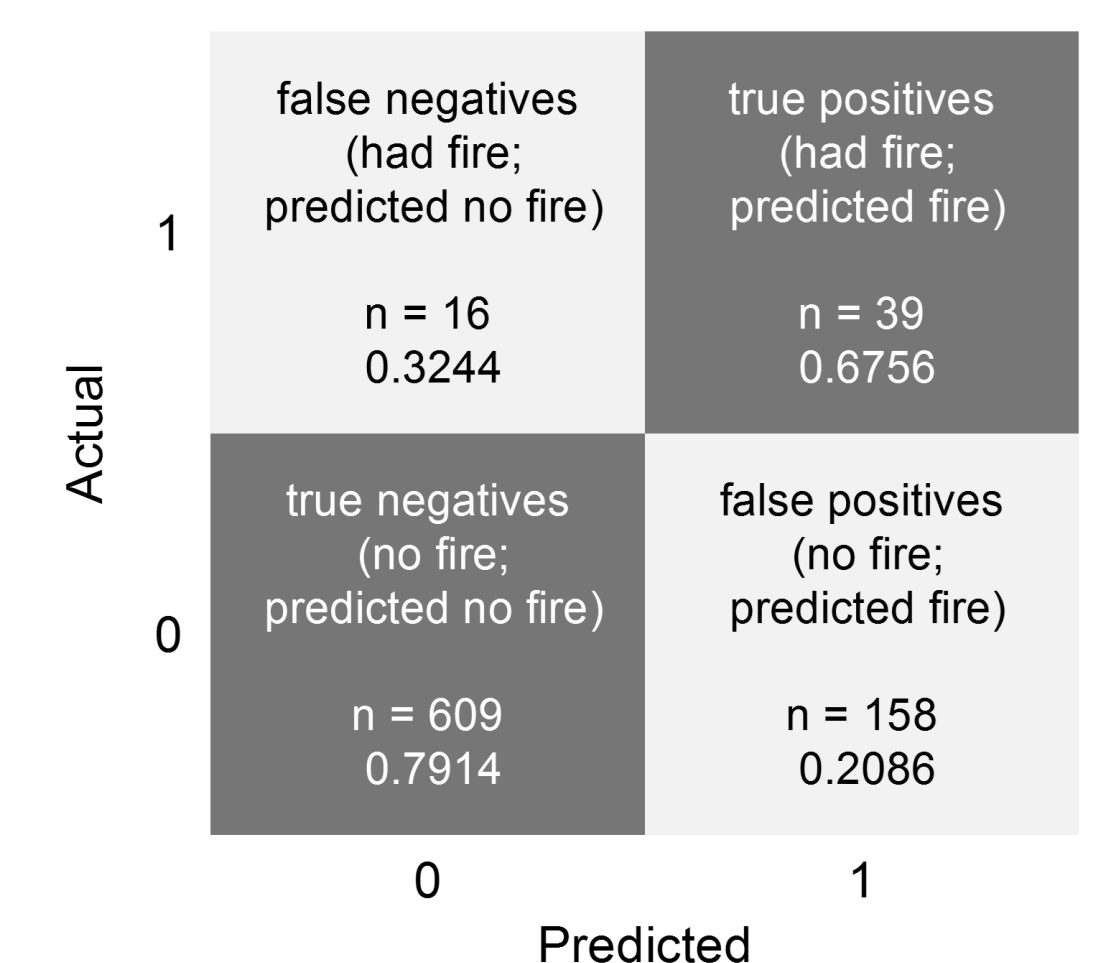
Model 1 (10 bootstrap result)

- Data from 2011-2014 and applied to 2014-2015 fire
- Average accuracy: 0.77 | Average AUC: 0.75



Model 2 (10 cross validation result)

- Data from 2011-2015
- Average accuracy: 0.78 | Average AUC: 0.73



Predictive Factors

Location	NPU (Neighborhood Planning Unit), zip code, submarket, neighborhood, tax district
Land/ property use	property/business type, land use codes, zoning
Financial	tax value, appraisal value
Time-based	year built, year renovated
Condition	lot condition, structure condition, sidewalks
Occupancy	vacancy, units available, percent leased
Size	land area, building square feet
Building	number of units, style, stories, structure, construction materials, sprinklers, last sale date
Owner	owner or property management company, owner's distance from the location of property
Demographics of location (TAZ level)	density, land use diversity, intersection features, crime density, racial makeup
Inspection	whether or not the parcel had been inspected by AFRD

Apply Model to Potential Fire Inspections

Predictive fire risk ranging from 0-1 and rescaled to 1-10

