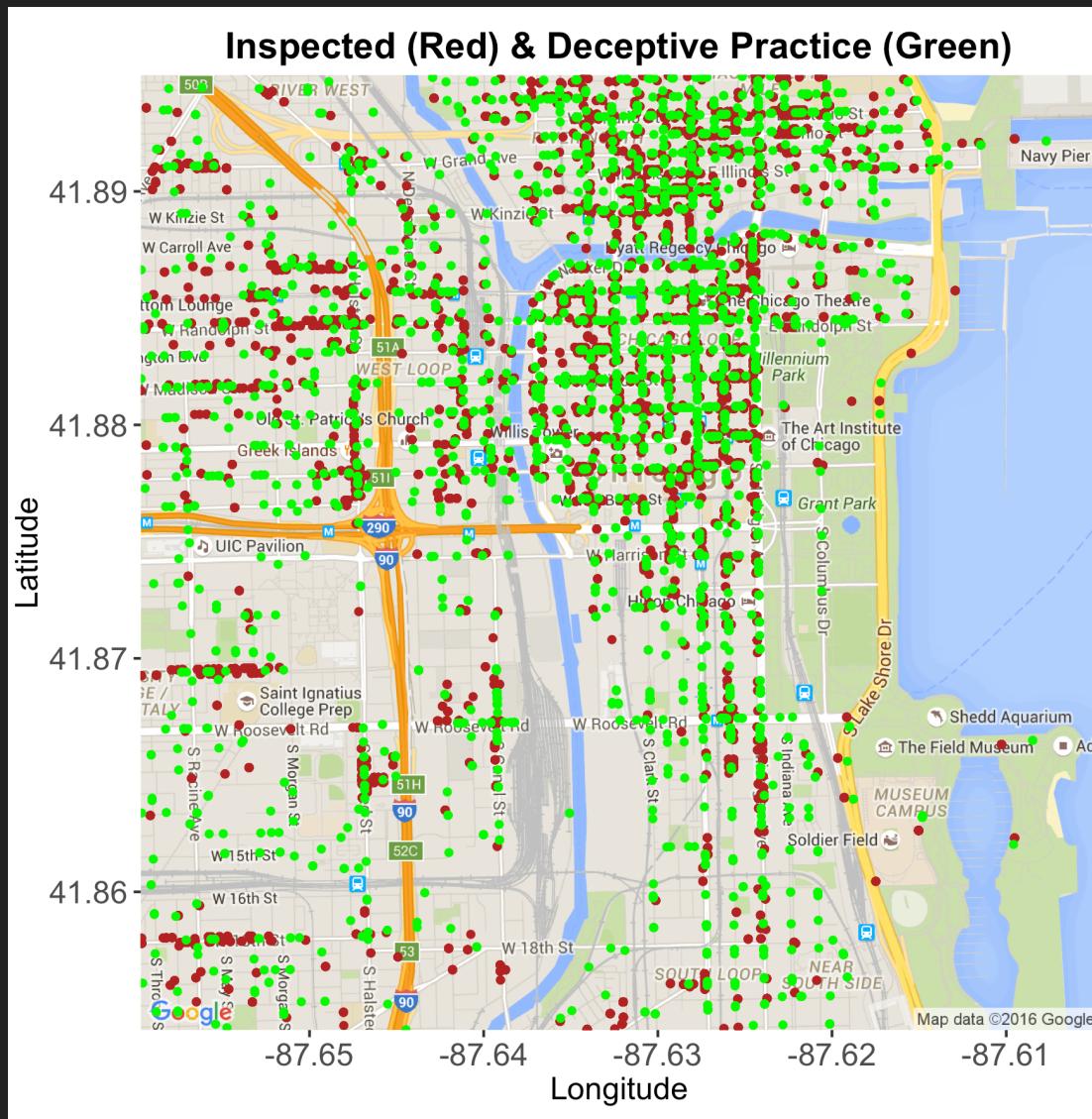


Jeongku Lim, Ph.D.

---

# THE DATA INCUBATOR CAPSTONE PROJECT

# A SPECIFIC CRIME (DECEPTIVE PRACTICE) VS INSPECTED STORES

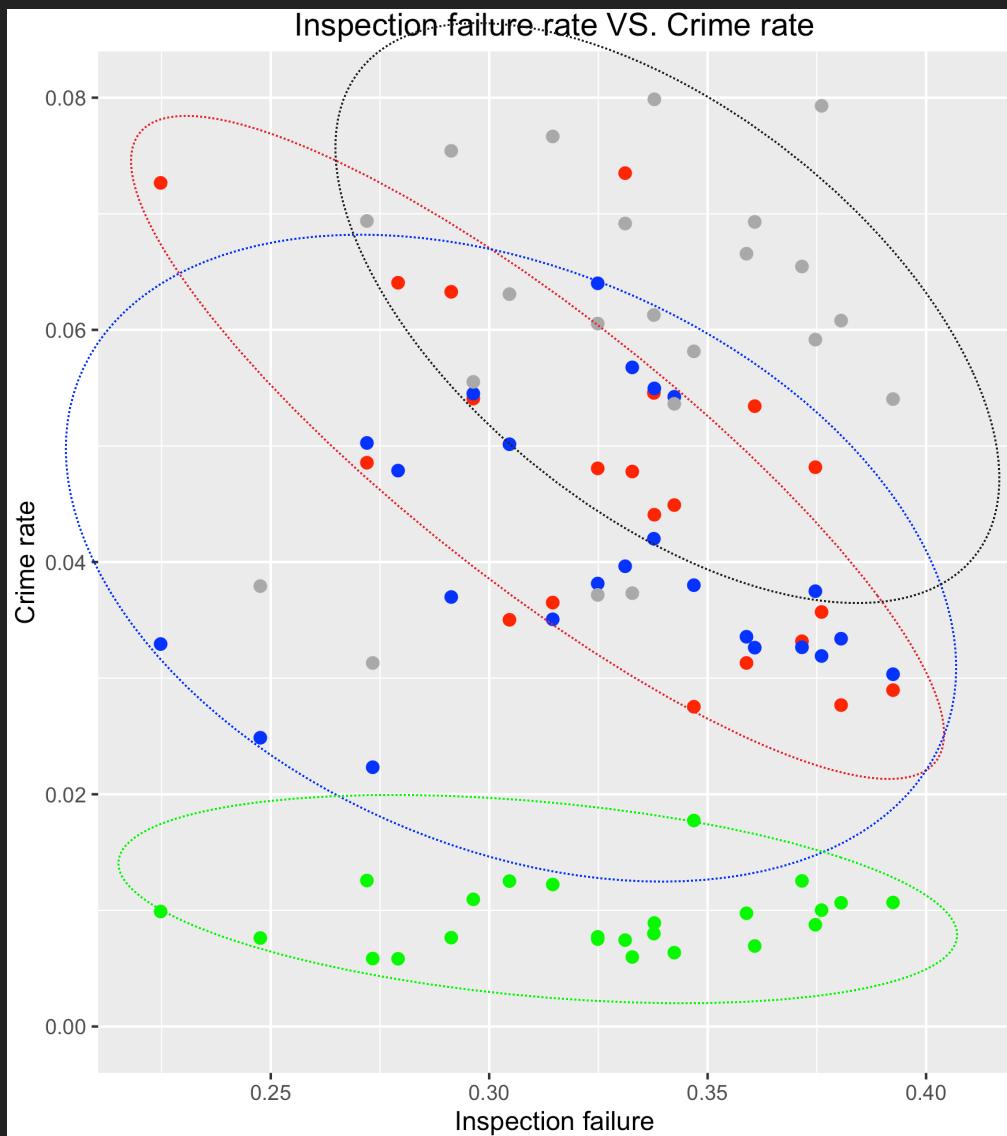


**Green** : Crime data  
**Red** : Inspection data  
**Deceptive Practice** happens near stores

'Theft' can happen anywhere



# INSPECTION FAILURE RATE VS. CRIME RATE



## Type of Crime

- Red** : Deceptive Practice
- Blue** : Motor Vehicle Theft
- Green** : Public Peace Violation
- Grey** : Other Offense

---

# LINEAR REGRESSION | INSPECTION FAILURE RATE ~ CRIMES

- ▶ All object values converted to 'fraction' in district.
- ▶ Fit summary

4 Predictors	Estimate	Std.Error	t-value	Pr(>  t )
Intercept	0.60	0.04	15.19	4.40e-12 ***
Deceptive Practice	-2.00	0.24	-8.36	8.61e-08 ***
Motor Vehicle Theft	-1.71	0.43	-3.97	8.24e-04 ***
Other Offense	-0.59	0.14	-4.17	5.20e-04 ***
Public Peace Violation	-6.48	1.92	-3.37	3.20e-03 **

Multiple R-squared: 0.8238

---

# PLANS

- ▶ Investigating crime data samples
  - Occurrence rate according to type, location, hour
- ▶ Scraping 'local business data' from 'Yelp'
  - Focus on the restaurant data for now
  - Need (Date, Location, Rating, Cost-level, Restaurant type)
- ▶ Deciding the name of this project.