

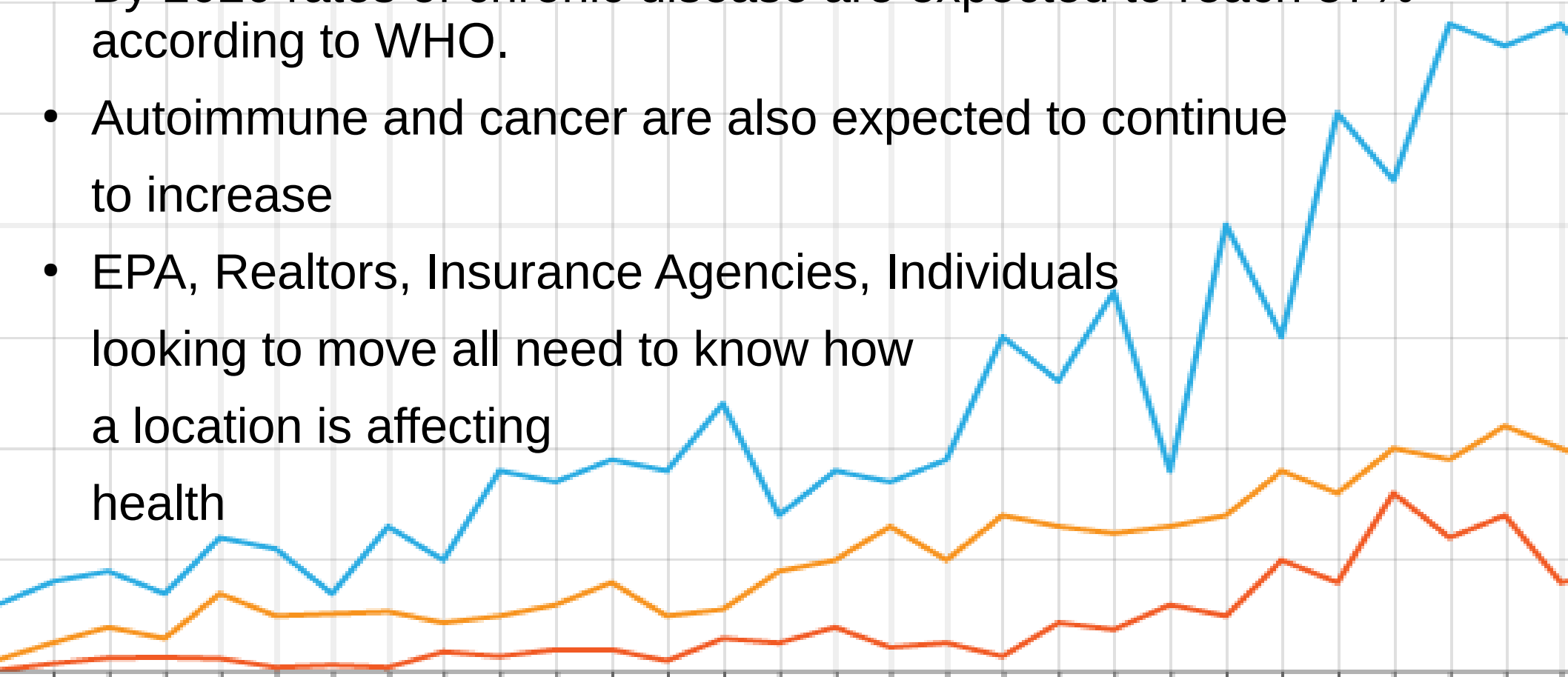
Is Your Neighborhood Contaminated?

A study of potential
contamination via reported
hospital diagnosis



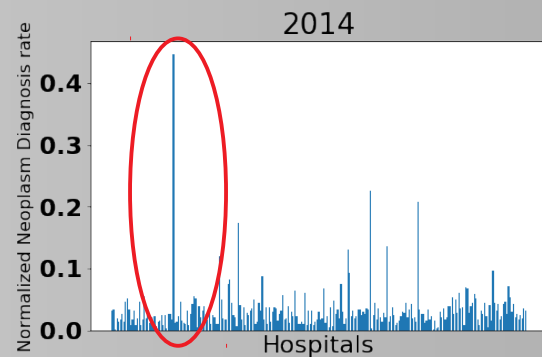
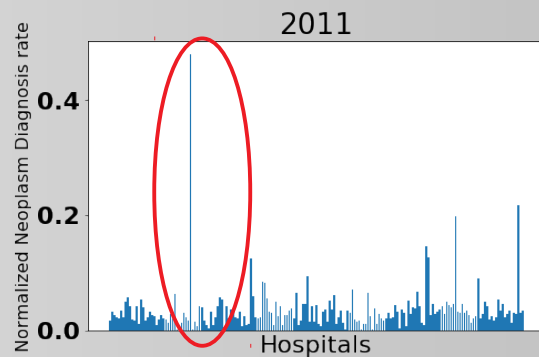
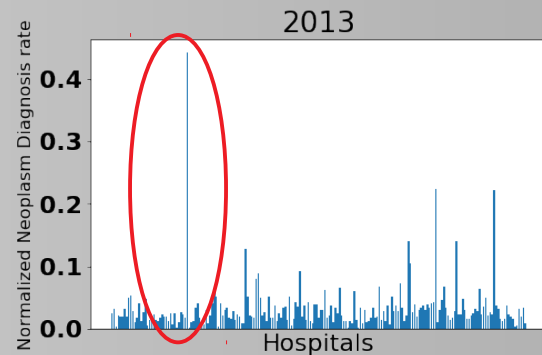
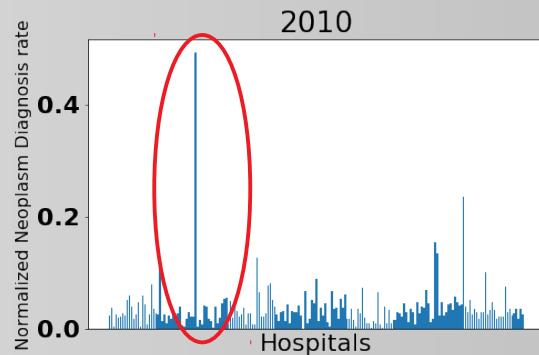
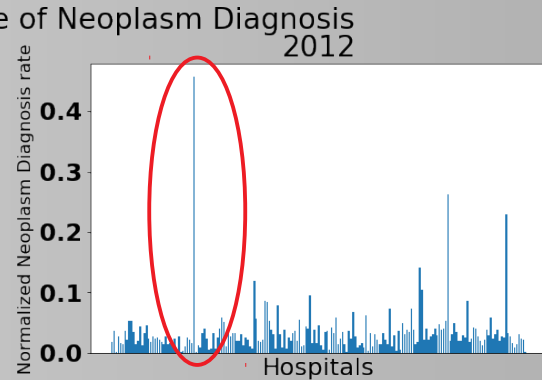
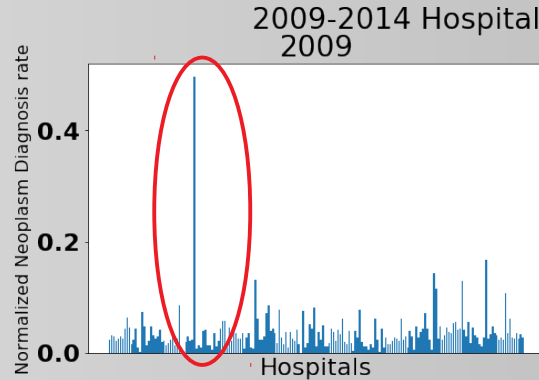
Overview / Description

- By 2020 rates of chronic disease are expected to reach 57% according to WHO.
- Autoimmune and cancer are also expected to continue to increase
- EPA, Realtors, Insurance Agencies, Individuals looking to move all need to know how a location is affecting health



Data Collection and Cleaning

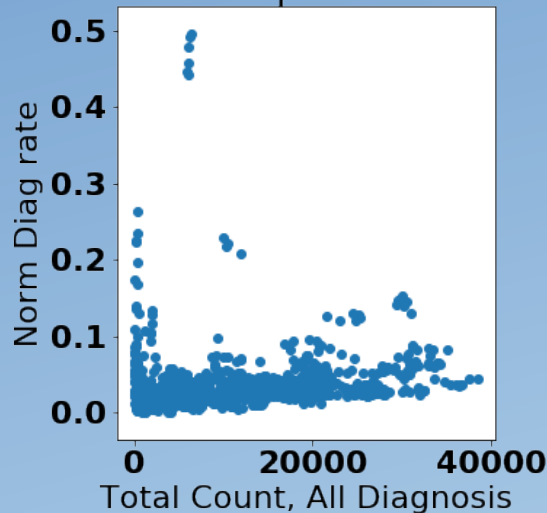
- Collect Data on Hospitals / Diagnosis
 - census.gov, epa.gov, or data.chhs.ca.gov
- Data contains
 - >40,000 rows
 - 18 different diagnosis types
- Clean and Combine relevant information
 - Attach longitude and latitude to hospital diagnosis
 - Create total diagnosis per hospital
 - Create diagnosis ratios per hospital
- Remove outliers from tests



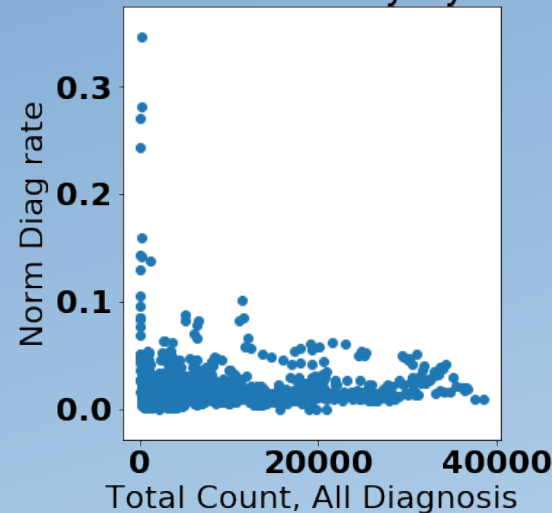
Does more
people Equal
more illness?



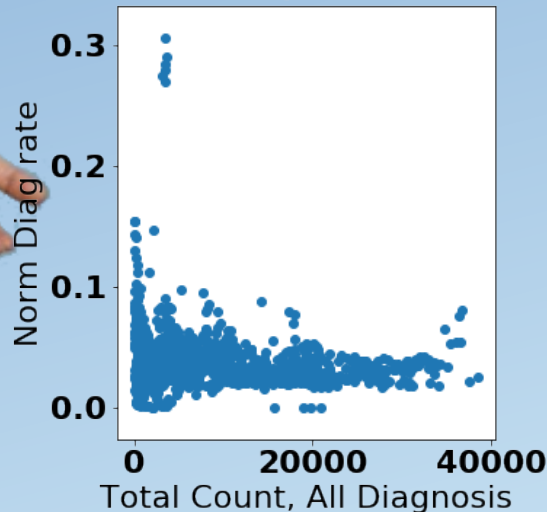
Total Patients vs Normalized Diagnosis
Neoplasms



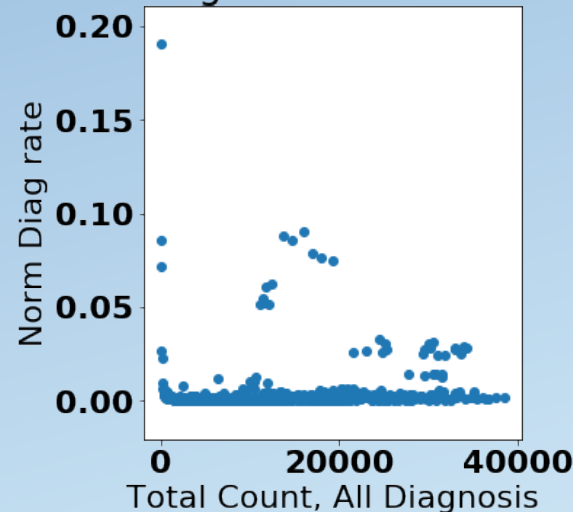
Nervous & Sensory Systems



Endocrine/Metabolism



Congenital Anomalies



Guess
Not

STATS!

Questions:

Is a hospital's rate of diagnosis higher than the rate for the rest of California in a given year

Null Hypothesis:

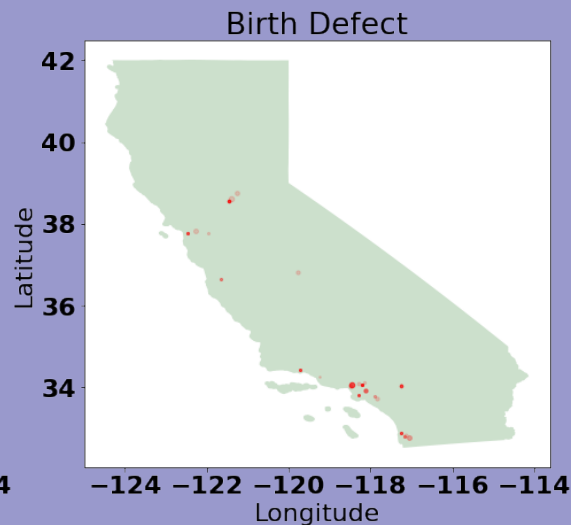
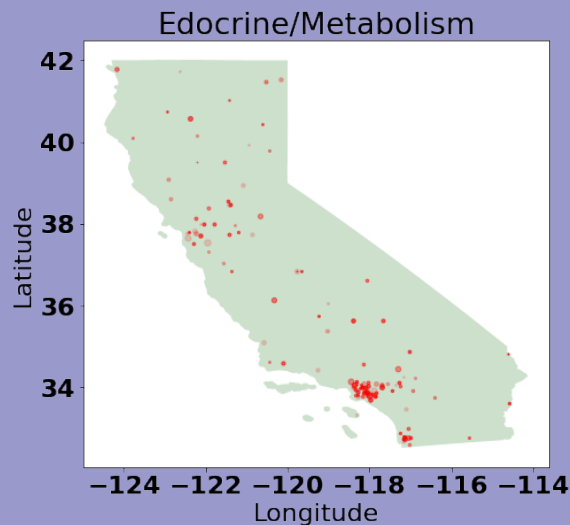
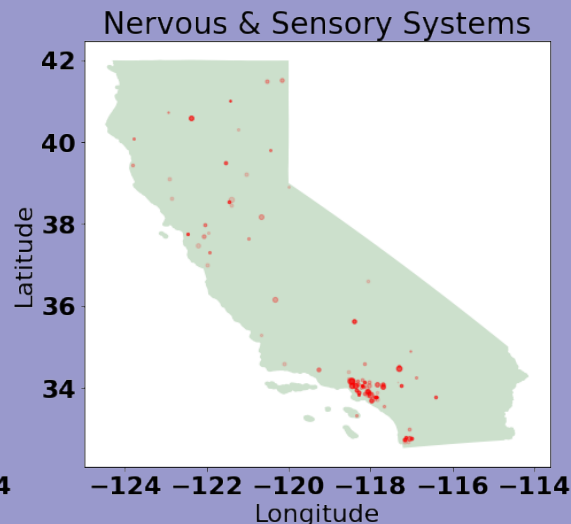
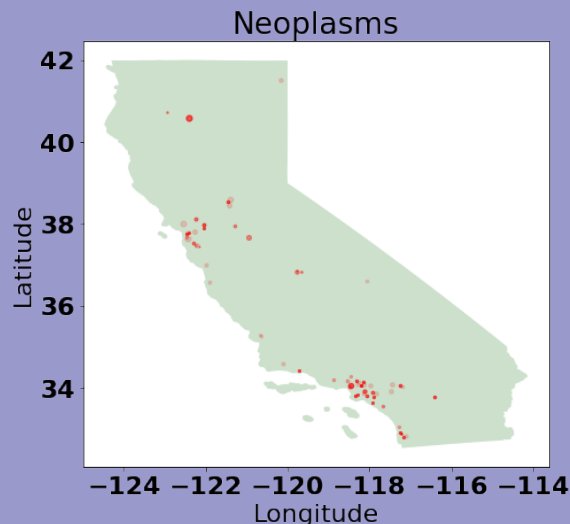
The rate of diagnosis for a hospital is the same or lower than the rate of diagnosis for the rest of California

Alternate Hypothesis:

The rate of diagnosis for a hospital is higher than the rate of diagnosis for the rest of California

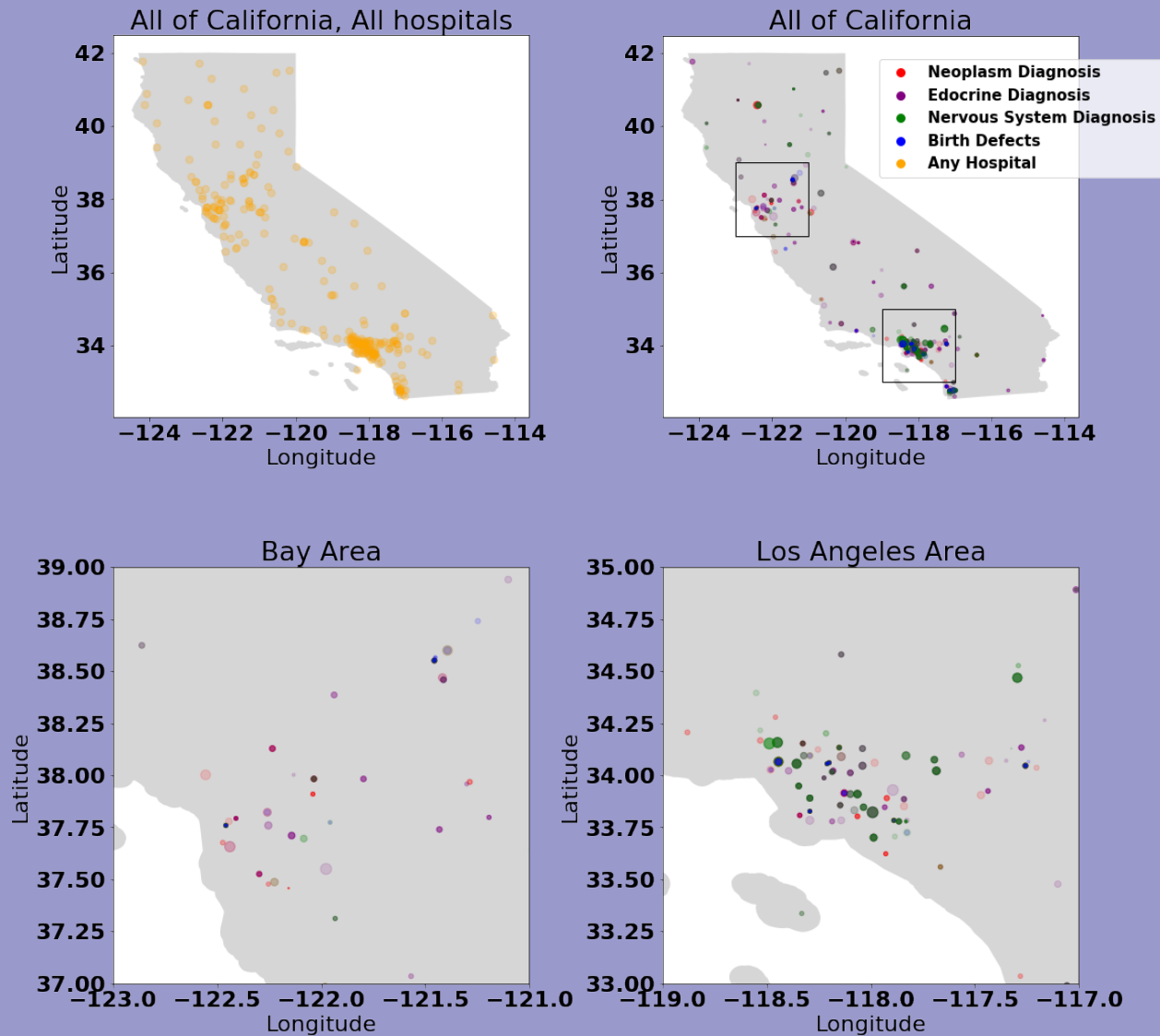
Rejection Threshold:

$\alpha = 0.05$



And GRAPHS!

Hospitals with high ratio of environmentally effected Diagnosis 2009-2014



ER!

BIGGER!

Conclusions

- Major cities have health problems but not all
- Los Angeles is pretty unhealthy
- Massive data limitation
- Going forward
 - Get more data
 - AL grouping analysis

