

Locality Lookout

Identifying and predicting complaints in your locale

Dahlia Nadkarni





Complaint dataset



Complaint dataset

- Non-emergency complaints: Noise, Sewer, Rodent, Mold, Heating,...



Complaint dataset

- Non-emergency complaints: Noise, Sewer, Rodent, Mold, Heating,...
- Around 8 million complaints registered since 2010



Complaint dataset

- Non-emergency complaints: Noise, Sewer, Rodent, Mold, Heating,...
- Around 8 million complaints registered since 2010
- Location: 5 borough, ~ 60 districts, ~200 zip code, lat-long



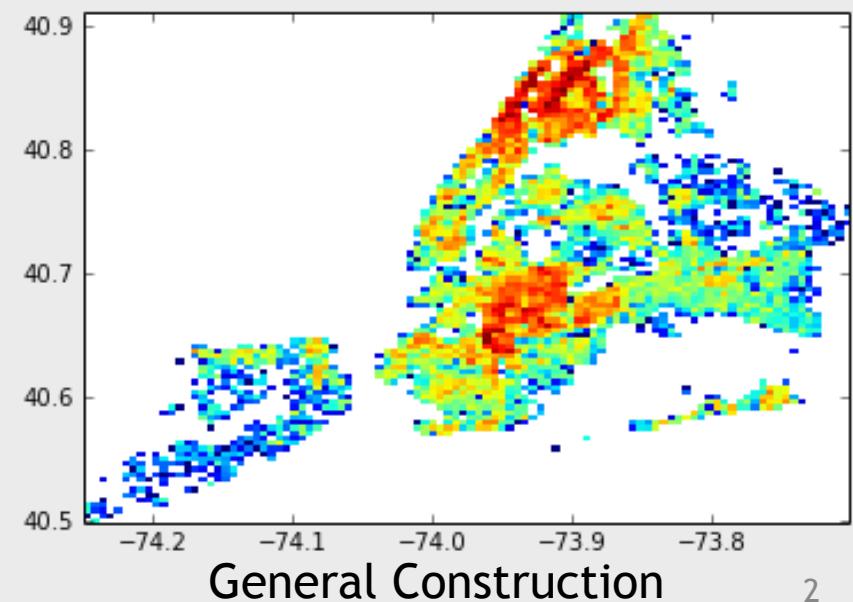
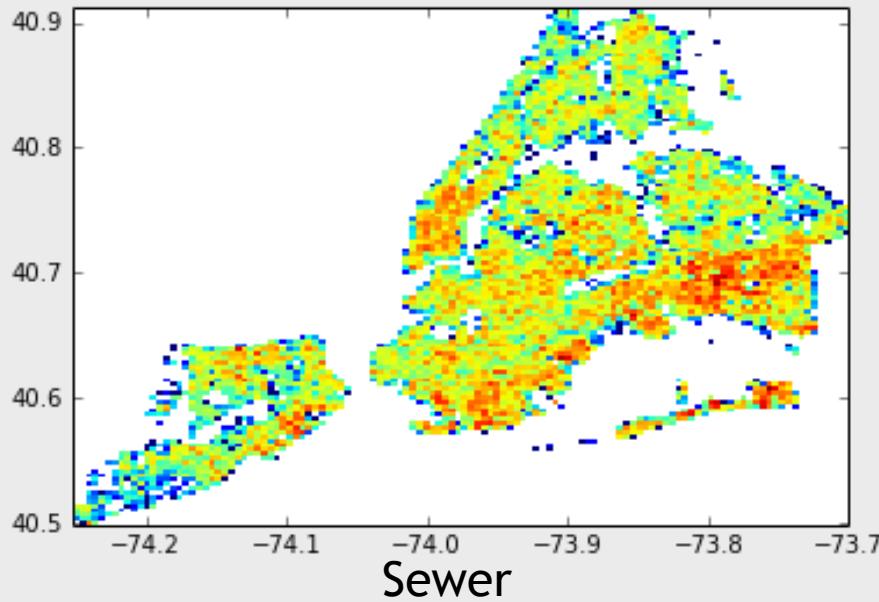
Complaint dataset

- Non-emergency complaints: Noise, Sewer, Rodent, Mold, Heating,...
- Around 8 million complaints registered since 2010
- Location: 5 borough, ~ 60 districts, ~200 zip code, lat-long
- Timestamps: created, resolved, due date



Complaint dataset

- Non-emergency complaints: Noise, Sewer, Rodent, Mold, Heating,...
- Around 8 million complaints registered since 2010
- Location: 5 borough, ~ 60 districts, ~200 zip code, lat-long
- Timestamps: created, resolved, due date



Snapshot of the data

| Unique Key | Created Date | Closed Date | Agency | Complaint Type | Incident Zip | City | Status | Due Date | Resolution Action Updated Date | Latitude | Longitude |
|------------|---------------------|------------------------|--------|-------------------------|--------------|----------|----------|------------------------|--------------------------------|-----------|------------|
| 25408482 | 2014-07-31 09:22:47 | NaN | DSNY | Graffiti | 11208 | BROOKLYN | Open | NaN | 07/31/2014 09:22:47 AM | 40.681150 | -73.880894 |
| 25411862 | 2014-07-31 09:06:10 | 06/18/2013 12:00:00 AM | DSNY | Graffiti | 10469 | BRONX | Open | NaN | 07/31/2014 09:06:10 AM | 40.861803 | -73.848244 |
| 25416040 | 2014-07-31 08:31:20 | 07/29/2013 12:00:00 AM | DSNY | Graffiti | 11213 | BROOKLYN | Open | 08/30/2014 08:31:20 AM | 07/31/2014 08:31:20 AM | 40.671239 | -73.928093 |
| 25409622 | 2014-07-31 08:27:35 | 07/29/2013 12:00:00 AM | DSNY | Graffiti | 10458 | BRONX | Open | 08/30/2014 08:27:35 AM | 07/31/2014 08:27:35 AM | 40.856835 | -73.888855 |
| 28582487 | 2014-07-31 02:20:50 | NaN | DOT | Street Condition | NaN | NaN | Open | NaN | NaN | NaN | NaN |
| 28581269 | 2014-07-31 02:12:05 | NaN | NYPD | Blocked Driveway | 11377 | WOODSIDE | Open | 07/31/2014 10:12:05 AM | NaN | 40.736517 | -73.919085 |
| 28582144 | 2014-07-31 02:10:22 | 07/31/2014 02:48:27 AM | NYPD | Illegal Parking | 10454 | BRONX | Closed | 07/31/2014 10:10:22 AM | 07/31/2014 02:47:07 AM | 40.809466 | -73.917764 |
| 28583360 | 2014-07-31 02:09:07 | NaN | NYPD | Noise - Street/Sidewalk | 10032 | NEW YORK | Assigned | 07/31/2014 10:09:07 AM | 07/31/2014 02:43:27 AM | 40.837369 | -73.946984 |
| 28587135 | 2014-07-31 02:06:52 | NaN | NYPD | Noise - Vehicle | 11216 | BROOKLYN | Open | 07/31/2014 10:06:52 AM | NaN | 40.675203 | -73.950001 |

Who uses this data?

Who uses this data?

Citizens

Who uses this data?

Citizens

City Council
Members

Who uses this data?

City Council
Members

Citizens

City Agencies
(NYPD, etc.)

Who uses this data?

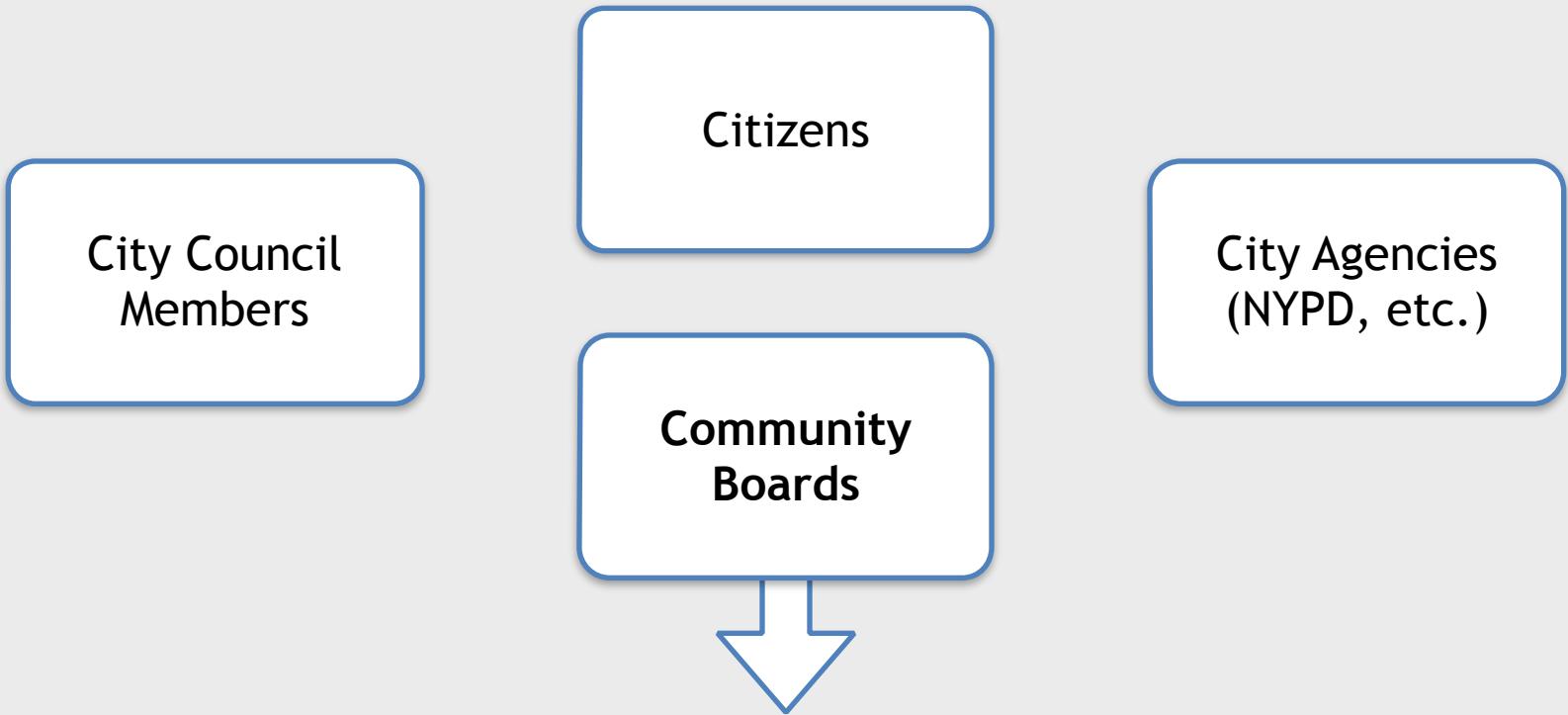
City Council
Members

Citizens

City Agencies
(NYPD, etc.)

Community
Boards

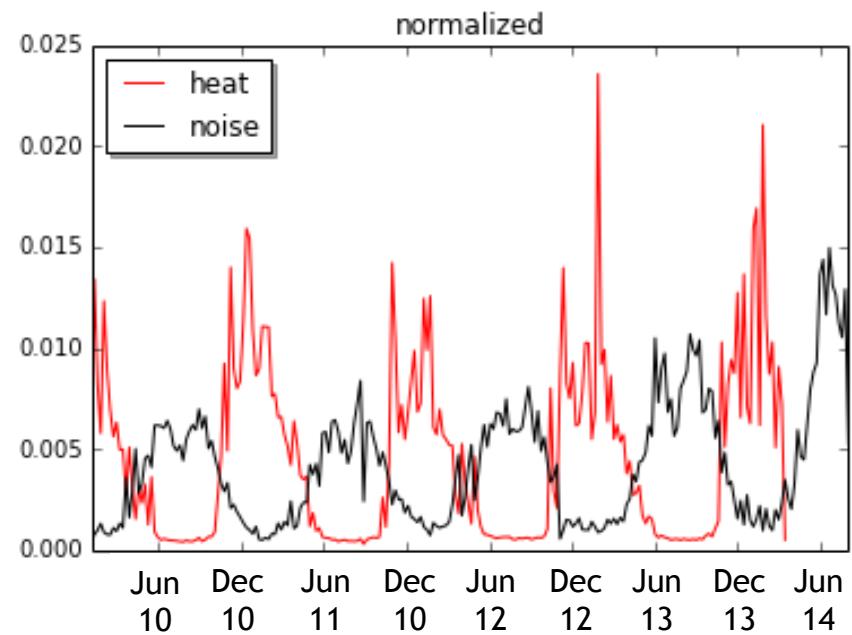
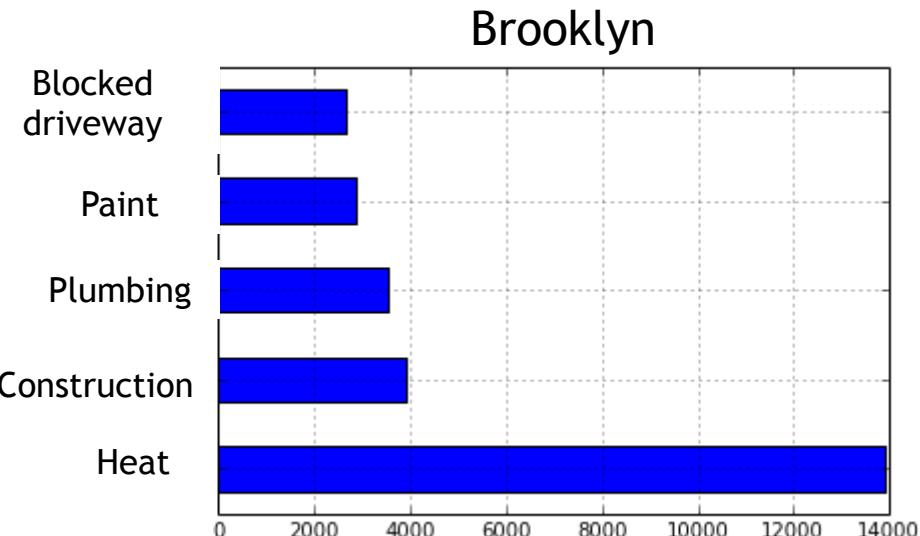
Who uses this data?



- **Advocacy:** “I can do better advocacy, if I had data that is easier to look at.”
- **Efficiency:** “If we had easier to understand 311 data, we could have more productive CB meetings.”
- **Effectiveness:** "I could set better agenda topics"

Identifying high-priority problems

How do you quantify ‘high-priority’?



LocalityLookout.com

Home Visualize Slides More About Me

Identify high-priority complaints in your NYC locale

Enter your borough:

Enter your Community Board Number:

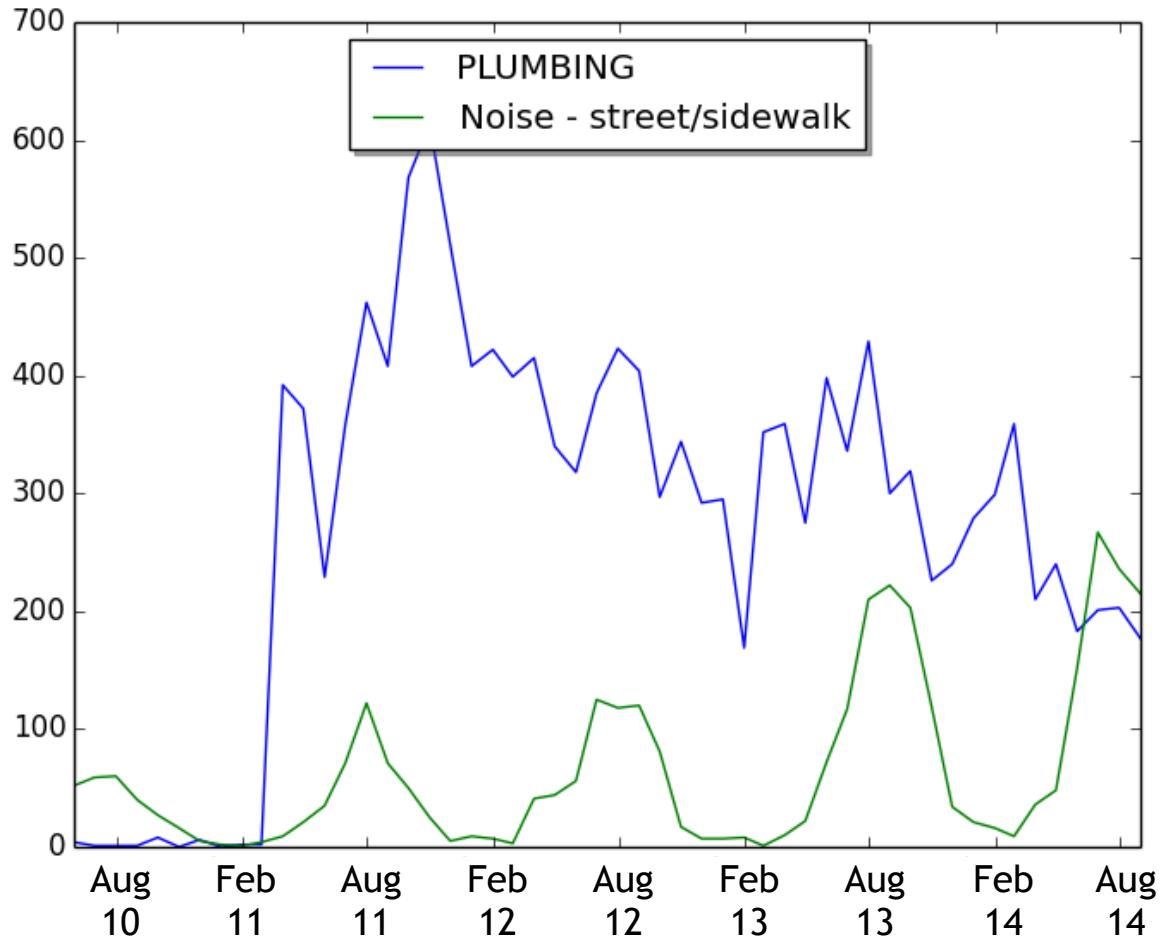
Pick a time period to focus on:

- Last 1 month
- Last 3 months
- Last 6 months

[View top complaints by priority](#)

CB 04 Bronx

(last 6 months)



| Complaint Type | Total Count | Priority | Priority score* |
|-------------------------|-------------|-------------------------------------|-----------------|
| Noise - street/sidewalk | 954 | High | 2.35 |
| Electric | 571 | Complaint volume less than expected | -1.09 |
| Water system | 814 | Complaint volume less than expected | -1.69 |
| Plumbing | 1214 | Complaint volume less than expected | -2.42 |

Analysis

Analysis

- Model the time series: Linear regression & Gaussian processes

Analysis

- Model the time series: Linear regression & Gaussian processes
 - Categorical features: for periodicity (quarter, day of the week)

Analysis

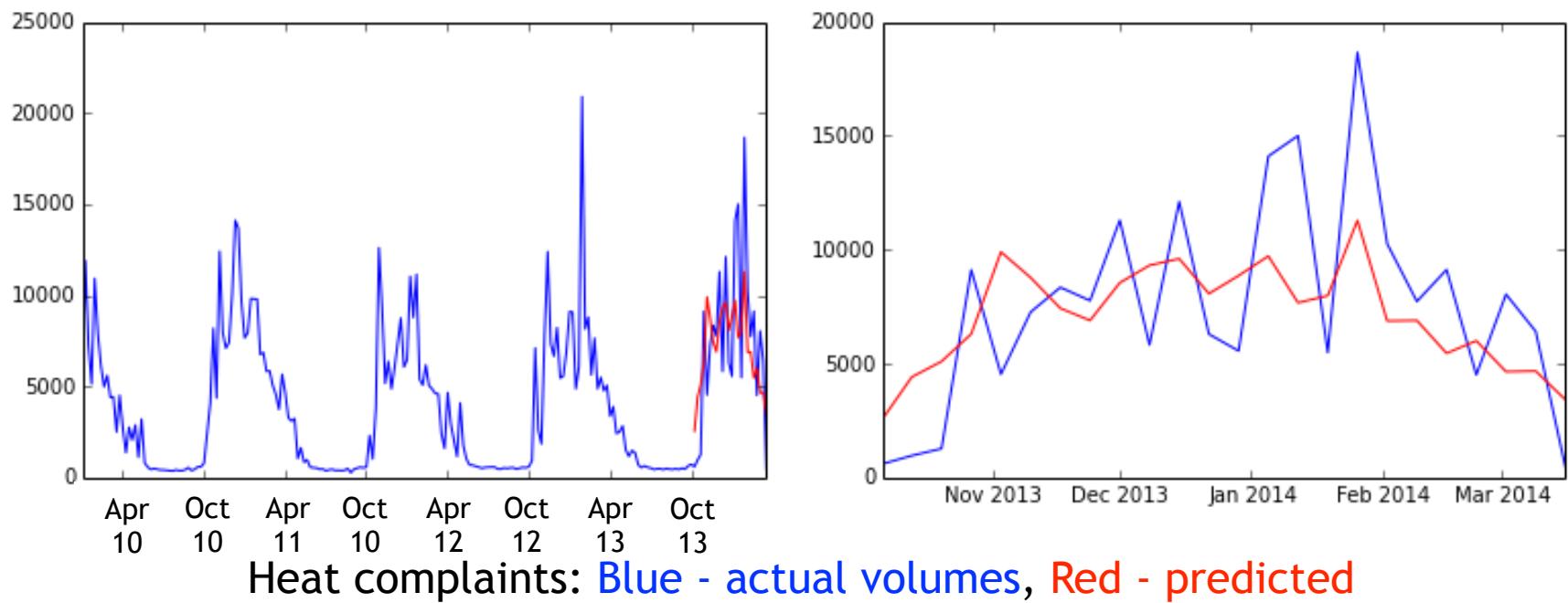
- Model the time series: Linear regression & Gaussian processes
 - Categorical features: for periodicity (quarter, day of the week)
 - Linear features: for linear trends

Analysis

- Model the time series: Linear regression & Gaussian processes
 - Categorical features: for periodicity (quarter, day of the week)
 - Linear features: for linear trends
- Predict future complaint volumes: useful to allocate resources

Analysis

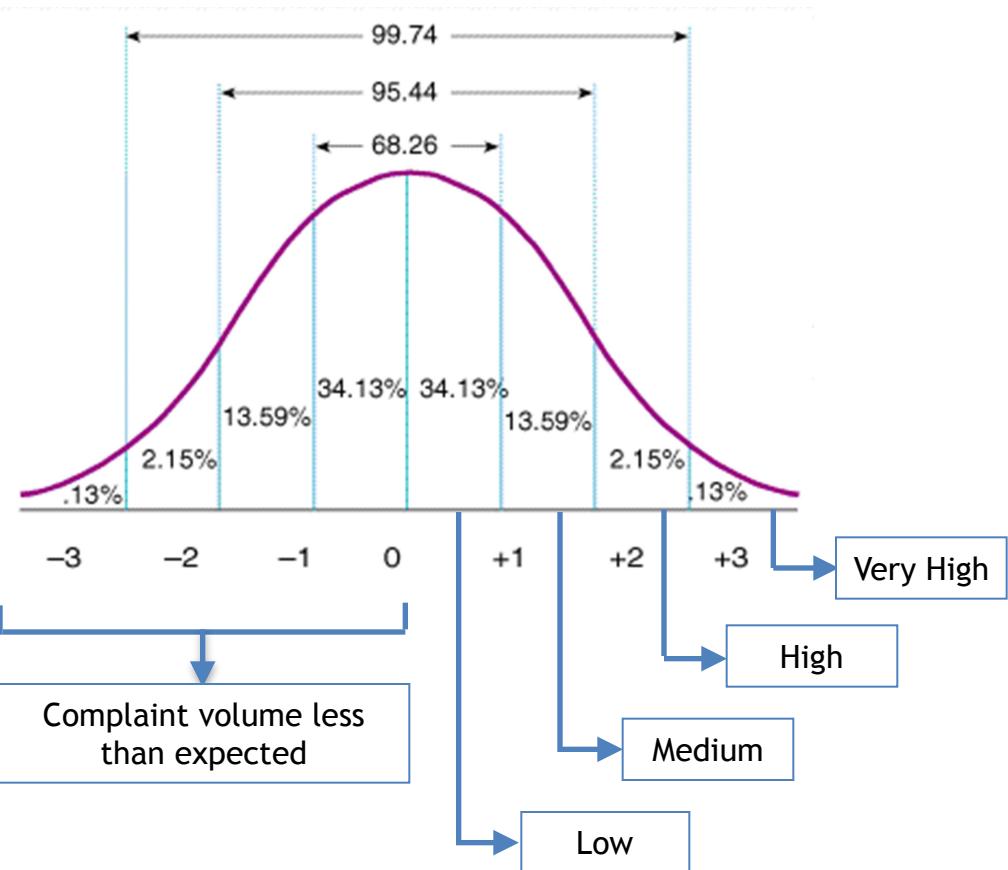
- Model the time series: Linear regression & Gaussian processes
 - Categorical features: for periodicity (quarter, day of the week)
 - Linear features: for linear trends
- Predict future complaint volumes: useful to allocate resources



Analysis

- Compare expected & actual volumes

- Priority score =
$$\frac{(y - \hat{y})}{\hat{\sigma}_y}$$



Dahlia
Nadkarni



Dahlia
Nadkarni



Dahlia
Nadkarni



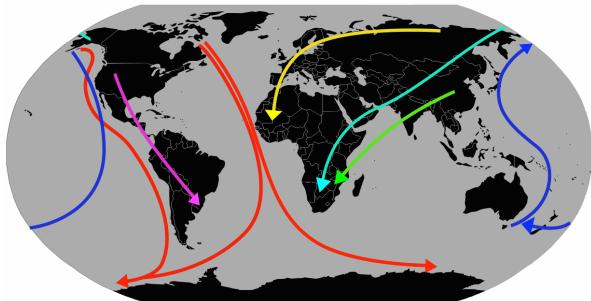
BROWN



Dahlia Nadkarni



BROWN

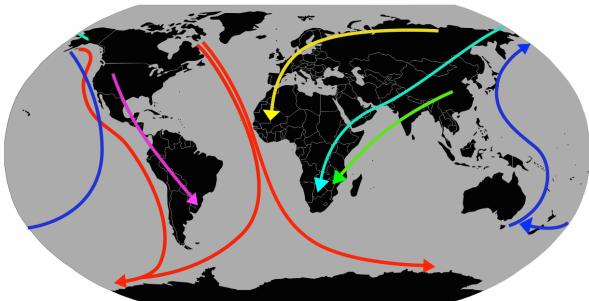
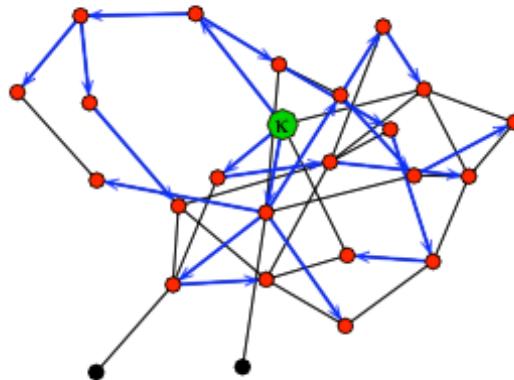




Dahlia Nadkarni



BROWN

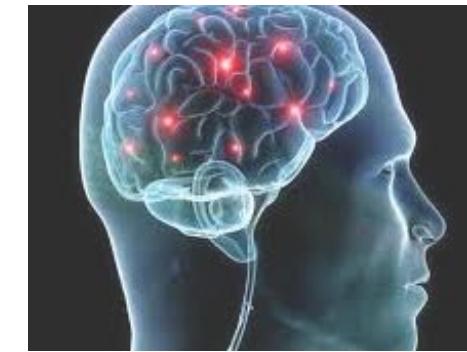
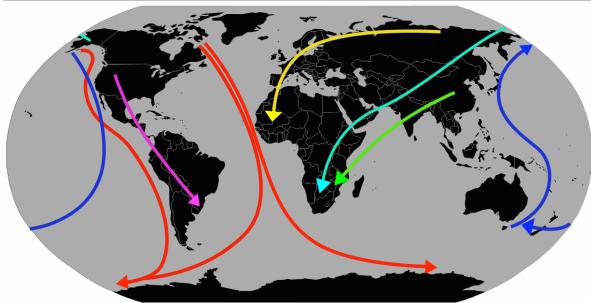
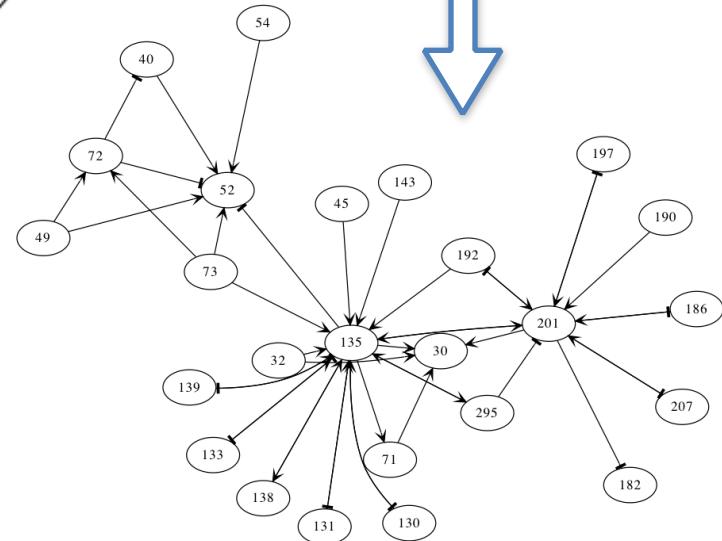
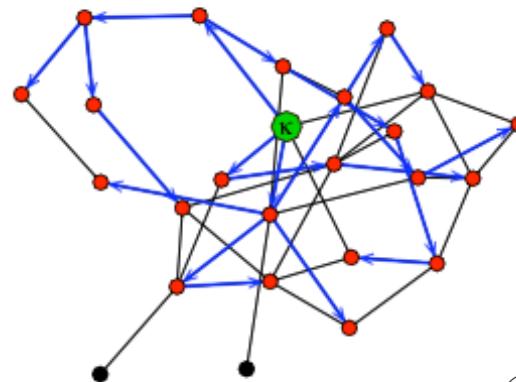




Dahlia Nadkarni



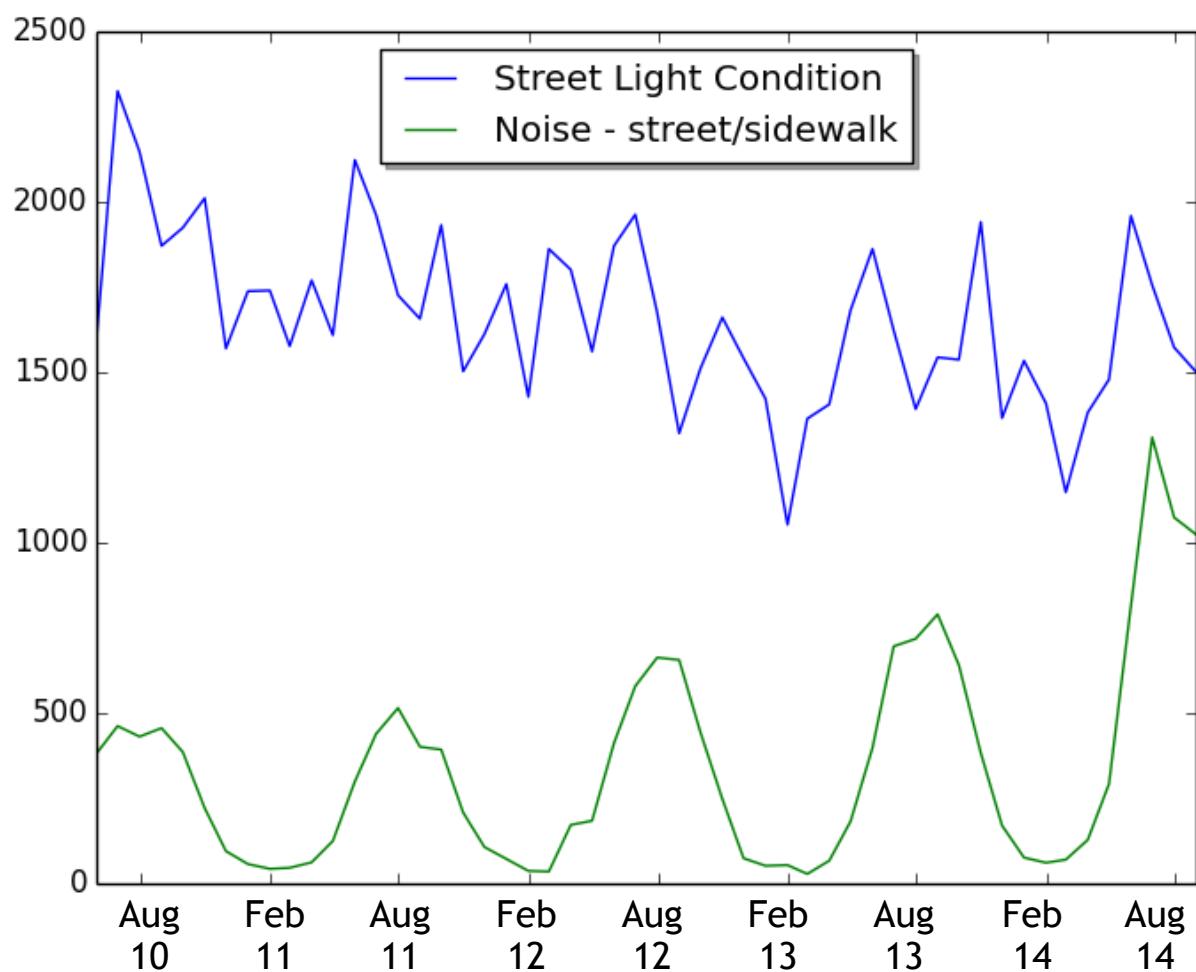
BROWN



Thank you

All of Bronx

(last 3 months)



| Complaint Type | Total Count | Priority | Priority score* |
|-------------------------|-------------|-------------------------------------|-----------------|
| Noise - street/sidewalk | 3407 | Very High | 6.05 |
| Blocked driveway | 2846 | Low | 0.78 |
| Street light condition | 4830 | Low | 0.47 |
| Plumbing | 3626 | Complaint volume less than expected | -1.86 |
| Water system | 4131 | Complaint volume less than expected | -5.40 |