NYC Analytics

2016

What Is MODA?

The Mayor's Office of Data Analytics (MODA) is led by NYC's Chief Analytics Officer, Dr. Amen Ra Mashariki. As part of the Mayor's Office of Operations, MODA partners with agencies to translate data-driven insight into action.

Mission: Actionable Insight for NYC Government

1. Support more effective delivery of services to New Yorkers for greater equity, safety, and quality of life

2. Grow and advance analytics throughout the City

Tenant Harassment project summary

Mayor Office of Data Analytics (MODA)
in partnership with the
Tenant Harassment Prevention Task Force (THPTF)

DRAFT version
NOT for distribution

Outline

- Tenant Harassment context
- Task Force
- MODA's role and process
- Results and success metrics
- Lessons Learned and Best Practices around analytics projects

Tenant Harassment: Context

NYC Current Situation:

high market rate rents **+** rent regulation laws and deregulation loopholes **→** tenant harassment
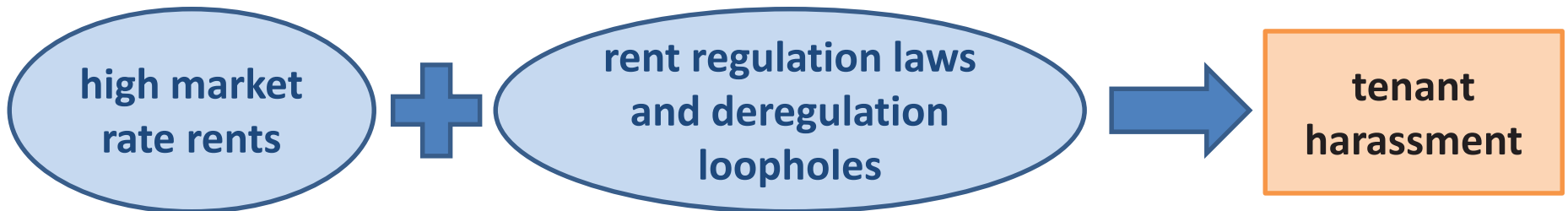
- Clear financial incentive to force out tenants
- Harassment can include anything from offering buyouts, threats, eviction notices, court dates, to deliberately making the apartment unlivable with the intent of forcing the tenant out.

Government Intervention:
Tenant Harassment Prevention Task Force (THPTF)

- Collaboration between City and State entities including HPD, DOHMH, DOB and the State Attorney General
- Goal: Coordinate joint inspections, enforcement actions and litigation strategies to prevent tenants from being forced out of their home by landlords who create unsafe living conditions.
  - Inspection teams with the capacity to issue violations, fire watch, vacate orders, etc., assess the situation.
  - Followed by litigation in cases where actions by owner rise to criminality.

http://www1.nyc.gov/site/hpd/renters/thpt.page

High profile examples

Brooklyn Slumlord Daniel Melamed Arrested for Harassing Tenants

BY NICK LUCCHESI   WEDNESDAY, JUNE 17, 2015 AT 2:55 P.M.
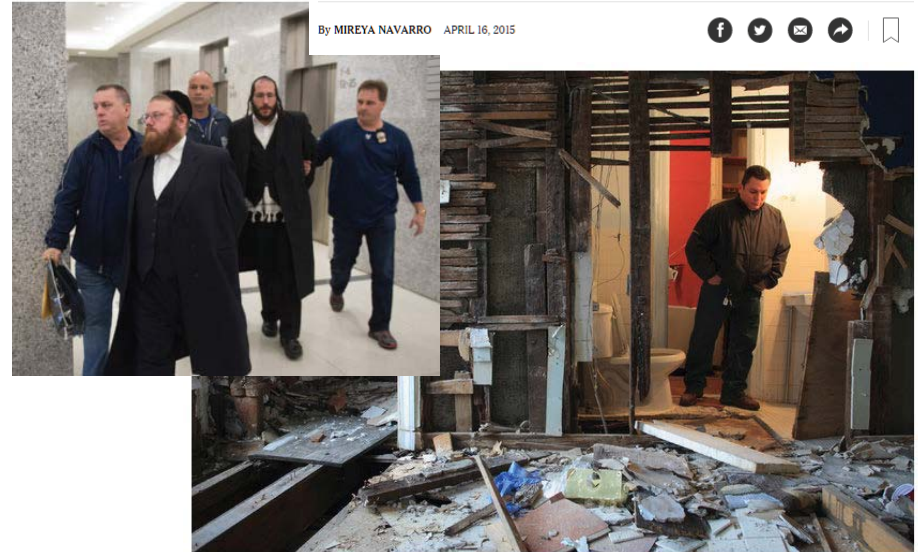
Owner: Daniel Melamed
Address: 1578 Union St.
City Data:
- BBL: 3014000042, BIN: 3037671
- Sold in 12/2012
- DOB Job in 02/2014
- 14 Months Sale to Job
- In 2014 there were 48 complaints (HPD, HPD no services, Construction, DOB illegal work)

The New York Times

2 Brooklyn Landlords, Accused of Making Units Unlivable, Are Charged With Fraud

By MIREYA NAVARRO   APRIL 16, 2015

Owner: Joel and Aaron Israel
Address: 98 Linden St.
City Data:
- BBL: 3033320008, BIN: 3076261
- Sold in 02/2013
- DOB Job in 10/2013
- 8 Months Sale to Job
- In 2014 there were 35 complaints (DOB, DOB illegal work, Construction, HPD no services)

MODA's Role

- Task Force was already in place and conducting joint inspections before MODA's involvement.
- An HPD analyst was in charge of choosing places for the TF to go, but only had access to HPD data, not data from other agencies.
- Places were routed based on recommendations in an ad hoc way.
- MODA helped put together a procedure of finding likely cases of tenant harassment based on city data.

MODA project pyramid
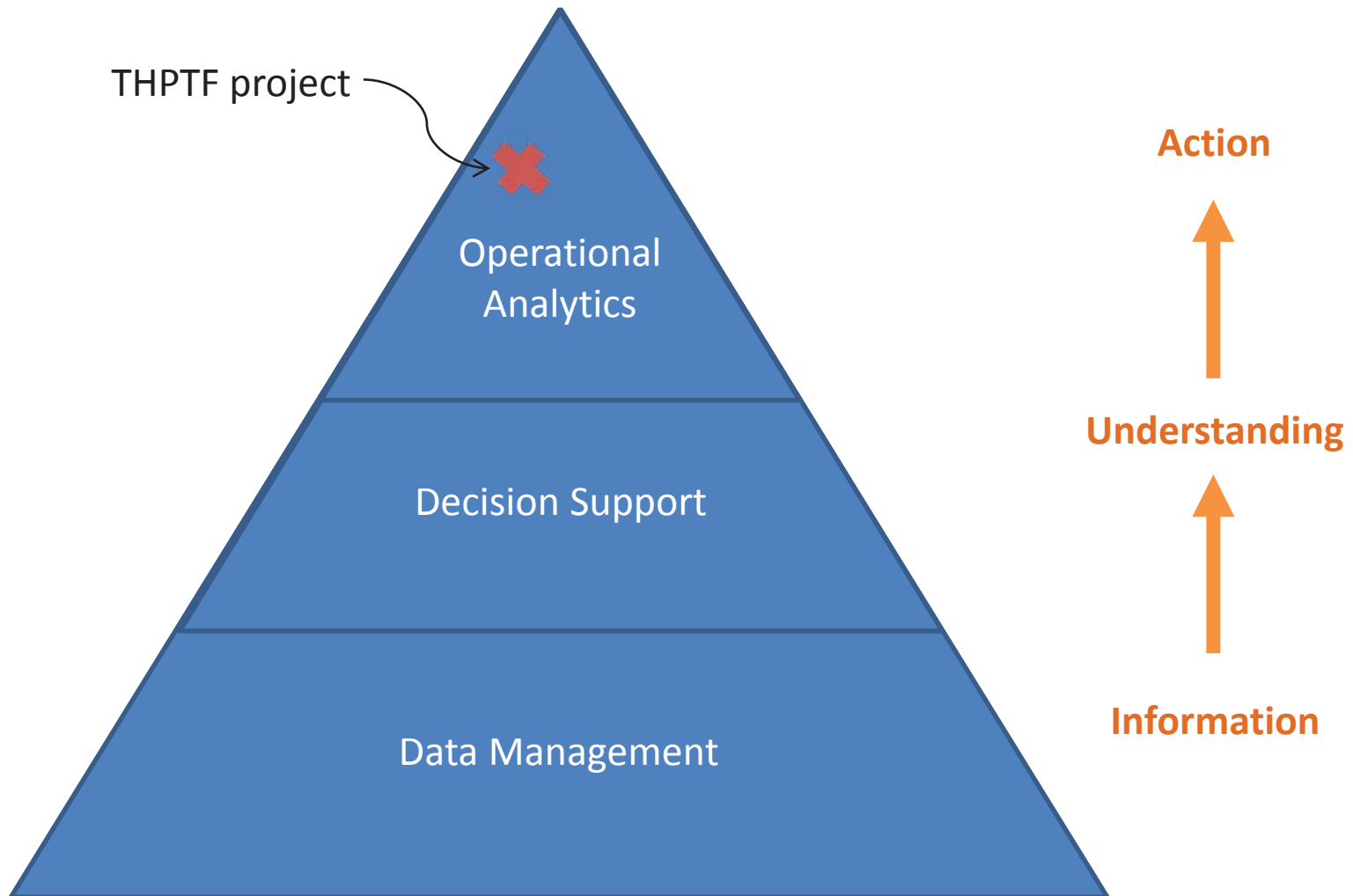
MODA process

- **Scoping** – context around the problem and operational goals.
- **Data** – what data is available and how it can be used. This includes understanding how and why the data was collected.
- **Analysis** – using operational goals as a guide, formulate and answer analysis questions using the available data.
- **Field testing** – deliverable for the agency to implement into their operations. Continue to evaluate and refine.
- **Implementation and hand off** – giving the agency the tools necessary to continue the project using their own resources.

scoping → data → analysis → field testing → implementation and handoff

Scoping

- Established Sponsor and Point of Contact: HPD
- Met with HPD Analyst who was routing inspectors. Also State AG Analysts who were deciding which cases to prosecute.
  - Came up with a working definition of "tenant harassment"
  - Formulated "indicators" to test
  - Discussed deliverable format
- Researched tenant harassment in NYC

Defining Tenant Harassment
for the Task Force

- What the TF was looking for:
  - Unlivable situations
  - Illegal construction
  - Hazardous to health of tenants
  - Unsafe conditions
  - "worst of the worst"
- How to find in city data?

Possible indicators
of tenant harassment in city data:

- Recently sold buildings
- Recently sold buildings followed by intense construction
- Illegal construction
- Complaints from residents
- Large difference between market rate and rent regulated apartments
- Landlords taking tenants to court

Analysis question: which of these indicators is more likely to result in the loss of rent regulated units?

Which buildings have rent regulated units?

- Landlords file annual registration with NY State Division of Housing and Community Renewal (DHCR)
  - PDFs of building list available on their website, but does not include unit counts
- Weird fact: City charges a $10 tax for each apartment that is rent stabilized.
  - Not Open Data, but it has been FOILed to get unit counts per building 2007-2014

https://github.com/talos/nyc-stabilization-unit-counts

Rent Regulated Units

- Over 40,000 buildings with rent stabilized units

- In any year over 10% of these buildings will loose 1 or more rent stabilized units

- Overall there are 10,000 to 20,000 units lost every year

Available Data

- Data we had access to:
  - HPD Complaints and Violations
  - DOB Complaints and Violations
  - 311 Complaints (for all agencies)
  - DOB Construction Permits
  - DOF Sales
  - DOF Rent Regulation Tax
- Data we did not have access to:
  - DHCR State data on rent regulated units
  - OCA Housing Court data

What is DataBridge?

DataBridge is a citywide platform that facilitates data sharing, storage, and use for operations, analytics, and reporting.

Data Management

Data Repository ①

Master Data
(Buildings, Property,
Zoning, Etc) ④

311 ③

Agency 1 ②

Agency 4

MMR

Agency 2

Agency 5

Collisions

Agency 3

Agency X

Data Integration ⑤

① **Data Repository**
Contains all the data available in DataBridge. Scalable, secure storage for almost any type of data

② **Agency Data**
Physically separated and secured. Loaded on an independent schedule.

③ **Shared data**
Citywide data sets are available to complement agency analysis

④ **Master data**
Data that can be used by any agency to enhance analysis; in particular, geospatial analysis

⑤ **Data Integration**
Tools to acquire, cleanse, harmonize, and load data

DataBridge leverages existing data systems to give analysts, inspectors, program managers, . . . , first responders information and tools to do their jobs most effectively.

17

Why DataBridge?

Achieving the City's goals requires coordination and sharing information across agencies.

- Complex problems are not single agency issues.

- DataBridge integrates agency collected information about people, places and businesses, and enables citywide analysis and coordinated action.

- Greater and deeper use expands City capabilities while increasing efficiencies.

Housing Court Data – could be another indicator?

- Landlord harassment strategy: take tenants to court

- Background: The NY State Office of Court Administration (OCA) sells housing court data to tenant screening companies. These companies create a tenant blacklist based on whether they have ever been to court (irrespective of the case outcome).Cost $20,000

- MODA and the State AG could not get access to this data, unable to test

What data we had to test

✓ Recently sold buildings

✓ Recently sold buildings followed by intense construction

✓ Illegal construction

✓ Complaints from residents

Indicator	Data Used
Property Sales	DOF Ownership data – when an owner name is updated
Construction	DOB Jobs filing, type alt 1-3, on multiple floors
Illegal Construction	311 Service Requests, DOB Complaints for illegal work
Complaints from residents	311 Service Requests

Breaking down 311 SR data:

There are over a thousand different complaint types.
MODA grouped into relevant categories around THPTF

Category	Definition
Air Quality - NonConstruction	DEP air quality complaints not directly related to construction
Asbestos	DEP/DOHMH asbestos complaints
Dust – Construction DEP	Dust from construction (outdoor) routed to DEP
Dust – Construction DOHMH	Dust from construction (indoor) routed to DOHMH
Construction	Complaint area containing 'construction'
DOB Illegal Work	DOB complaint containing: illegal, unsafe, safety, or permit.
DOB	Any other DOB complaint
Dirty Conditions	Includes DSNY/DOHMH rodent, mold, unsanitary, standing water, vector, sanitation condition, dirty conditions
HPD - No Services	HPD Heat/Hot water, electric, gas, refrigerator
HPD	Any other HPD complaint
Hazardous Materials	DEP hazardous materials.
Other	Everything not included in other categories. These are complaints expected to be tangentially related to tenant harassment.
NA	Agencies: DOE, TLC, EDC, DCA, DOITT ['Food Establishment','Beach/Pool/Sauna Complaint', 'Mobile Food Vendor','Street Light Condition','New Tree Request', 'Broken Muni Meter','Broken Parking Meter', 'City Vehicle Placard Complaint','Literature Request','CFC Recovery', 'Collection Truck Noise','DSNY Spillage','Employee Behavior','Snow', 'Snow Removal','Storm','Missed Collection (All Materials)', 'Bike Rack Condition','Bus Stop Shelter Complaint','Street Sign - Damaged','Street Sign - Dangling','Street Sign - Missing','Traffic']

311 Service Requests: numbers per year

Number of BBLs with 311 SR by Category
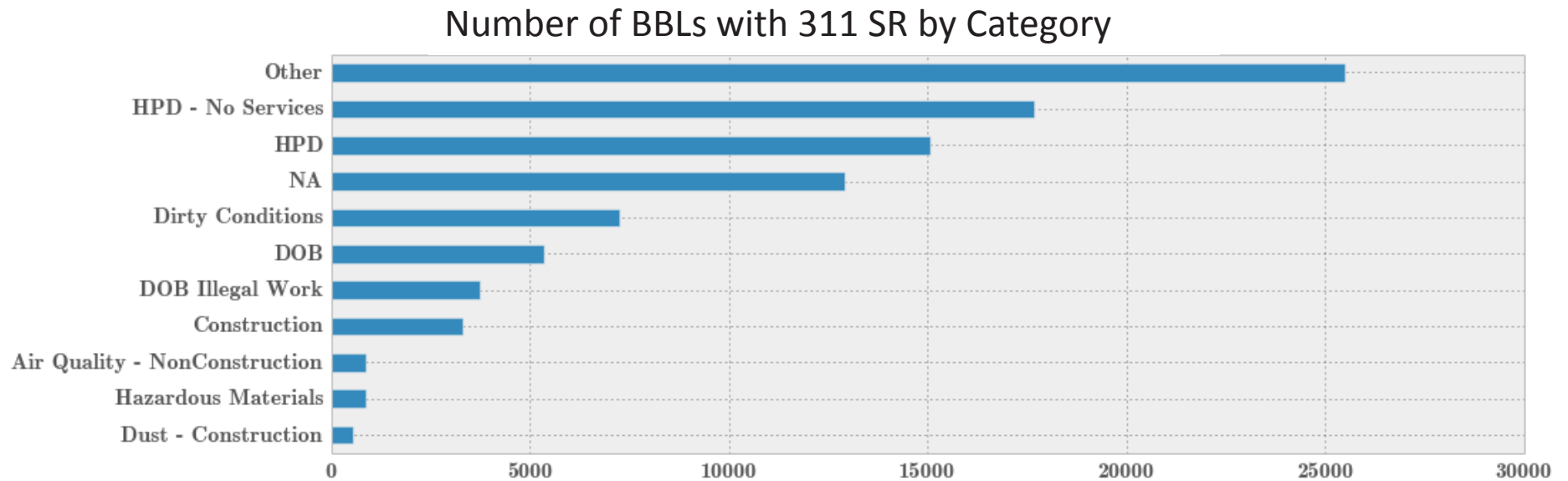
Analysis question: What indicators lead to the loss of rent regulated apartments?

- Define Significant Unit Loss for a building (or BBL)
  - Loosing 5 or more rent stabilized units per year
  - Loosing 40% or more rent stabilized units per year
  - Combination

- $P(loss) = \dfrac{BBLs \ with \ loss}{All \ rent \ stab \ BBls}$

- $P(loss|indicator) = \dfrac{BBLs \ with \ loss \ and \ indicator}{All \ rent \ stab \ BBls \ with \ indicator}$

- $Risk \ Ratio = \dfrac{P(loss|indicator)}{P(loss)}$

  Comparing the two to understand if the presence of an indicator is more likely to determine unit loss vs a random BBL

Analysis Results:

Risk Ratio = $\dfrac{P(loss|indicator)}{P(loss)}$

Probability building has significant unit loss given indicator was present. Over the probability building has significant unit loss.

Indicators:

DOB Illegal Work
Sale+Construction
Sale
311SR Other
311SR NA
311SR Hazardous Materials
311SR HPD - No Services
311SR HPD
311SR Dust - Construction DOHMH
311SR Dust - Construction DEP
311SR Dirty Conditions
311SR DOB Illegal Work
311SR DOB
311SR Construction
311SR Asbestos
311SR Air Quality - NonConstruction
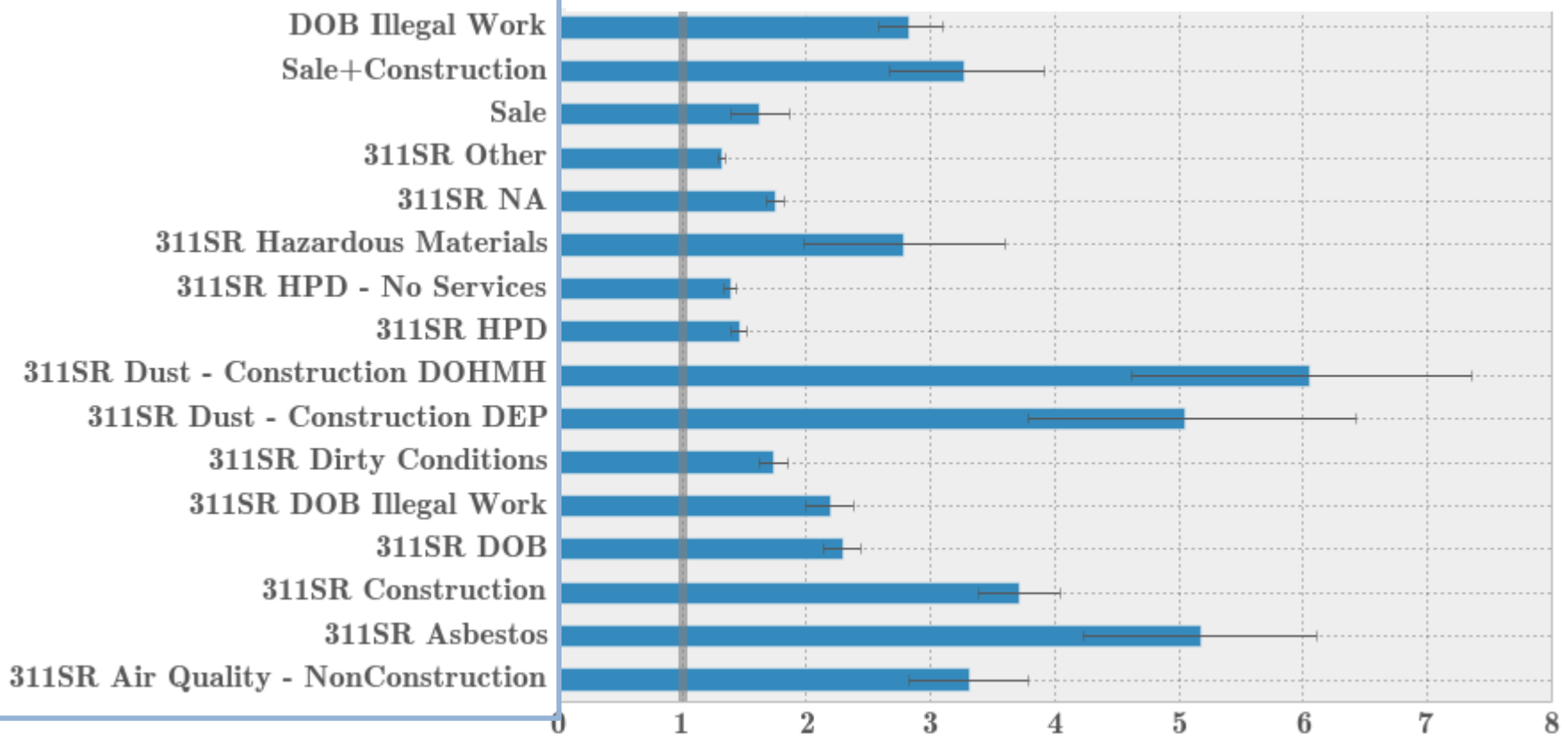
*error bars indicate 90% confidence interval

What we found

- Buildings with **dust** or **asbestos complaints** were 4-7 times more likely to experience unit loss in the following year compared to a random rent stabilized building.

- Also significant:
  - **DOB Illegal Work Complaints**,
  - **Sales followed by Construction**,
  - 311 SR related to **Construction**,
  - 311 SR for **Air Quality** (non-construction)

What was surprising

- 311 complaints about lack of services (HPD - No Services) were no more predictive of unit loss than other 311 complaints we thought had nothing to do with tenant harassment (NA and other)

- 311 complaints we thought had nothing to do with tenant harassment (NA and other), were actually slightly more likely to experience unit loss than a random rent regulated building. Perhaps complaints indicate a higher density of people or businesses – areas more at risk for gentrification?

What we missed

- People who did not call 311 (did not know or were too worried)
- Called 311 but did not report one of the significant categories.
- Landlord did not file a construction permit with the City (DOB)
- Filed construction permit, but did not match our timeline
- Aiming for: High precision, Low recall

Deliverable

- List of buildings with one or more of the significant indicators with the timeline:
  - Past 6 months for complaints
  - Sale + Construction: sale in the last 18-6 months followed by construction

  Timeline chosen to avoid sending stale places
- Information included:
  - Location: BBL, BIN, Address, Community District
  - Building information: number of units, rent stabilized
  - Type of complaint including agency it was referred to
  - Linear scoring system

Deliverable

- Because the HPD analyst was comfortable with the searching and sorting features in excel, we chose this as our deliverable format

- Format: excel spreadsheet with about 4 thousand buildings

- Updated every 2 months (or as needed)

Deliverable screenshot

Data Dictionary and Model Details

Targets Model:

Buildings that make it onto the targets list satisfied one or more of the following:

    (1) **Sale and Construction** - they had a sale in the last 18 months to the last 6 moths followed by a DOB jobs filing for alt 1-3 and multiple floors.

    (2) **Dust** - they had a 311 Service Request (SR) relating to dust from construction in the past 6 months

    (3) **Asbestos** - they had 311 SR relating to asbestos in the past six months.

    (4) **DOB Illegal Work** - they had specific DOB complaints related to illegal work in the past six months.

    In addition the BBL had to be listed in PLUTO as having 6 or more residential units.

Breakdown: Number of buildings by type of condition

	number of buildings on target list	number of rent stabilized buildings on target list
(1) Sale and Construction	829	577
(2) Dust SR	498	361
(3) Asbestos SR	418	319
(4) DOB Illegal Work complaint	2737	1921
Any of the above (full list)	3971	2772

Data was pulled on Oct 14, 2016

Targets Data Dictionary

	Column name	Column description	Source	Source building identifier
Location	BBL	Borough Block Lot		
	BIN	Building Identificaiton Number (DOB)		
	Address	address of the BBL	PLUTO 16v1	BBL
	Borough	borough	PLUTO 16v1	BBL
	CD	Community District	PLUTO 16v1	BBL
Sale and Construction	DateSale	date of the most recent sale in the last 18 months	DOF ownership table in Databridge (DOF_OWNR_NAME_UPDTD)	BBL
	DateRecentJob	date of the most recent DOB job filing in the last 18 months	DOB Jobs table in Databridge (J_PRE_FILING_DATE)	BBL,BIN
	TypeRecentJob	job type of the most recent DOB job filing in the last 18 months	DOB Jobs table in Databridge (J_JOB_TYPE_DESC)	BBL,BIN
	FloorsRecentJob	floors listed in the most recent DOB job filing in the last 18 months	DOB Jobs table in Databridge (J_FLOOR)	BBL,BIN
	Sale+Construction	yes' if the most recent sale was between 18 and 6 months ago and there has been an alt 1-3 construction job filing for multiple floors after the sale.		
	DateJobAfterSale	date of the first DOB job filing after sale (for sales between 18-6 months ago and jobs: alt 1-3 on multiple floors)	DOB Jobs table in Databridge (J_PRE_FILING_DATE)	BBL,BIN
	TimeSaleJob	time interval in days between DateJobAfterSale and DateSale		
Illegal Work Complaints	DOB IllegalWork	number of DOB complaints relating to illegal work, issued in the past 6 months (COMPLAINT_CATEGORY= 05,12,3A,5A,5G,66,71,76,83,86,90)	DOB complaints in databridge	BBL,BIN
	DustSR	number of SR for dust from construction in the past 6 months		

Intro / Targets / 311SR_detail / DOB_IllegalWork_detail / **DataDictionary**

Deliverable screenshot

How HPD and the TF use it

- Primary routing is still primarily based on recommendations
- They look for other buildings within that same neighborhood from our target list. The TF can visit multiple buildings in one trip.

How successful is the project (MODA)

- Dust and asbestos 311 complaints, which are generally routed to DOHMH or DEP, would not have been seen by the TF since HPD did not have access and/or knowledge of that data. Because there are DOHMH inspectors on the TF, this fit in very well.
- Does the TF go on more inspections per trip than they otherwise would?
- Is the TF seeing more relevant places than it otherwise would?
- Of the places the city would not have inspected (no pending complaints) how many of them received violations or further action?

How successful is the Task Force?

- Reaching places the City would not have seen through the normal inspection routine:
  - Joint inspections allows agencies to see more places than they otherwise would. How high is the cross agency violation rate? (e.g. complaint was for DOHMH but HPD wrote violations)
- What fraction of inspections go on to further actions?
- How many inspections resulted in an intervention that would otherwise have resulted in the unit going market rate.
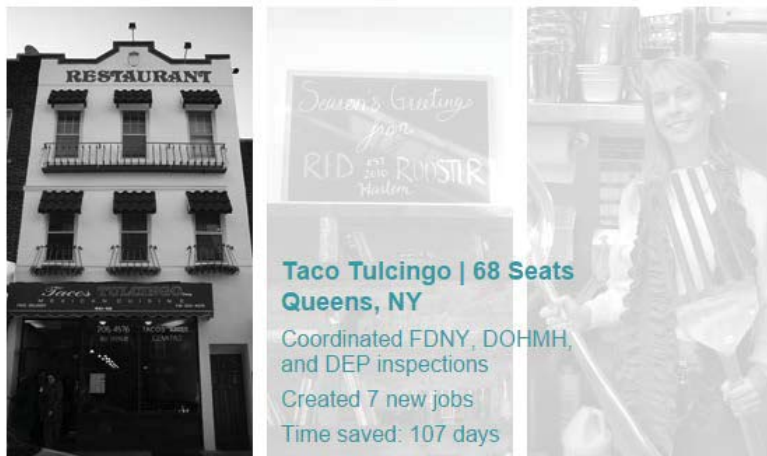
Analytics In Action

- Measuring Success

- Providing Situational Awareness

- Data Driven Enforcement

- Analytics for Equity

Measuring Success

**NYC Small Business Services
Time-to-Open Metric**

Creates end-to-end measurements to quantify effectiveness of new business processes and informs future policy decisions.

Taco Tulcingo | 68 Seats
Queens, NY

Coordinated FDNY, DOHMH, and DEP inspections

Created 7 new jobs
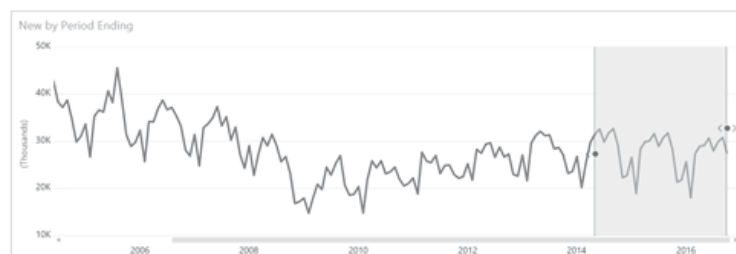
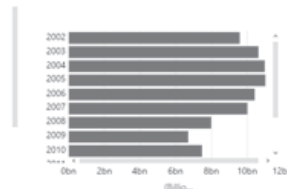Time saved: 107 days

Project Highlights

- Understanding the timeline of opening a business: as measured by NYC

- Cross-agency data from DOB, FDNY, DOHMH, and others

- Quantifying this process opens the door for leveraging advanced statistical methods to inform policy decisions.

Providing Situational Awareness

NYC Emergency Management
Emergency Management Data

Inform on the ground operations in order to provide first responders with information at the "speed of thought"
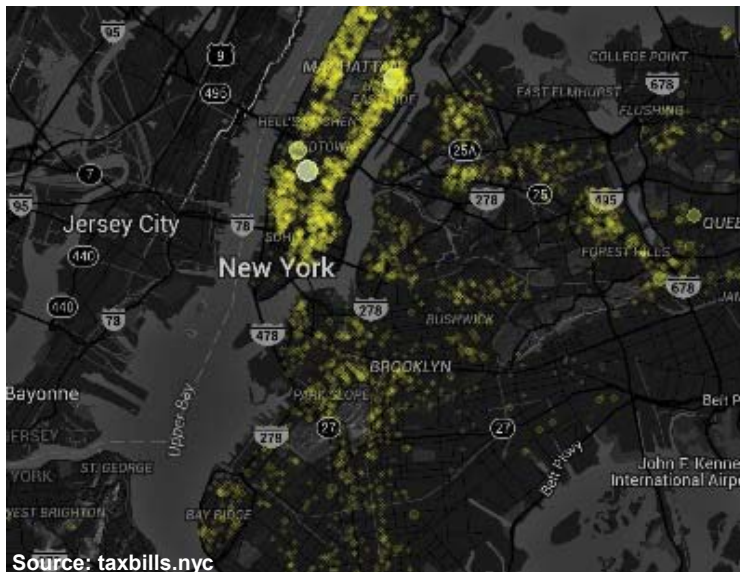

Project Highlights

- Using city data to build situational awareness

- Cross-agency data from ECB, DCA, DOHMH, and others

- Allows agency to become exponentially smarter before , during, and after emergency situations.

Data Driven Enforcement

NYC Housing Preservation and Development Tenant Harassment Task Force

Using NYC data to drive the prioritization of Task Force inspections, by identifying trends and relationships from past cases of harassment.

Rent stabilized unit loss from 2013 - 2014

Source: taxbills.nyc

Project Highlights

- Understanding the 'timeline of harassment', to paint a picture of harassment as seen by NYC data systems, so these occurrences can be proactively mitigated

- Data aggregated from multiple agencies including NYC DOF, DOB, HPD, 311 complaints as well as the NYS Attorney General

- Determining the driving factors of rent stabilized unit loss
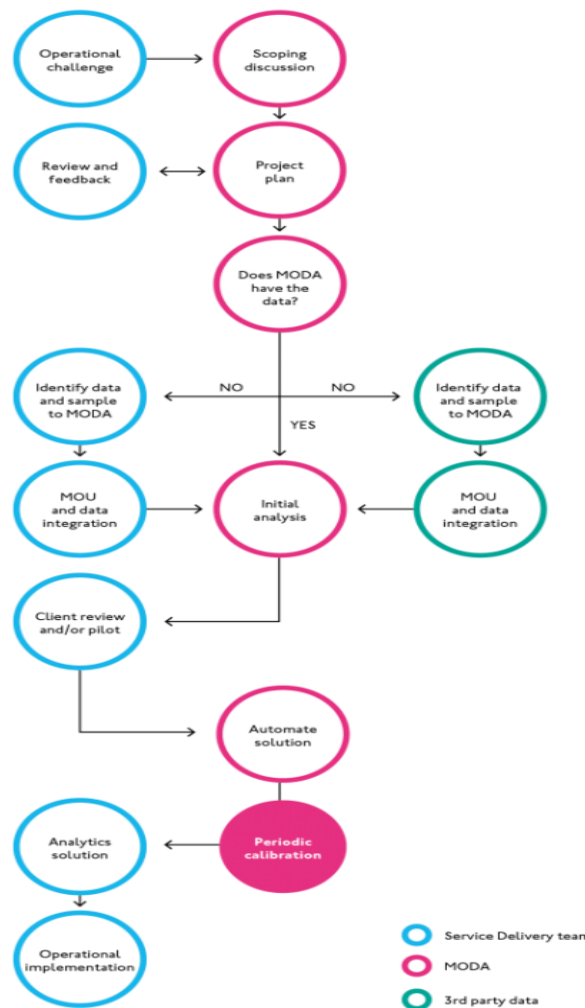
Analytics for Equity

NYC Commission on Human Rights

The goal of the project is to determine what neighborhoods are at-risk and to test whether illegal housing practices are occurring in them.

Project Highlights

- Locate where landlords may be engaging in source of income discrimination in New York City.

- Identifying the neighborhoods where few public rent vouchers are currently used, have well-performing schools, a below-average incidence of crime, and an ample stock of affordable rental housing.

- assembling portfolios of property ownership in these neighborhoods to identify the big players in those real estate markets.

Operational challenge → Scoping discussion

Review and feedback ↔ Project plan

Does MODA have the data?

Identify data and sample to MODA ← NO — NO → Identify data and sample to MODA

YES

MOU and data integration → Initial analysis ← MOU and data integration

Client review and/or pilot

Automate solution

Analytics solution ← Periodic calibration

Operational implementation

Service Delivery team
MODA
3rd party data

40

QUESTIONS ?