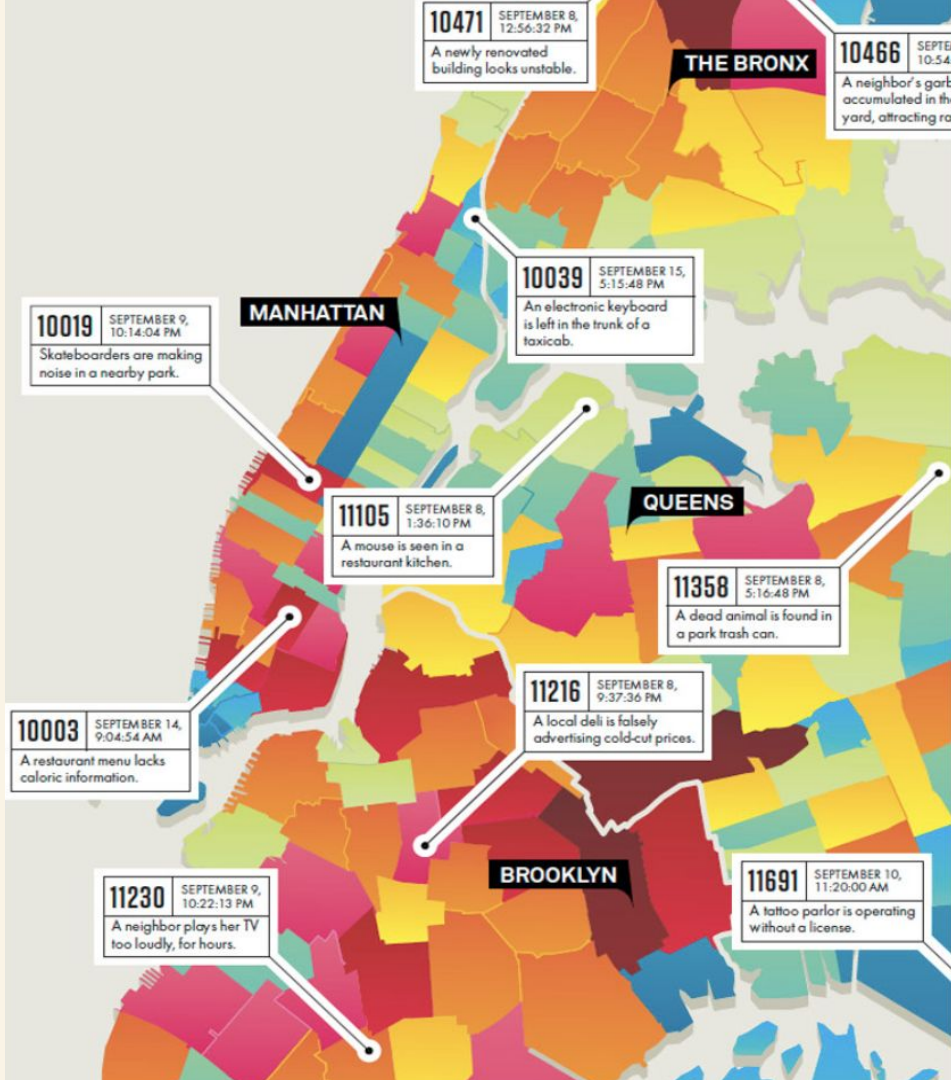
NYC 311: Fast-Tracking Complaint Resolutions

Section B1, Team 1:
Aastha Surana | Atishay Jain | Kendall Sims | Soham Sarvade



Problem Statement

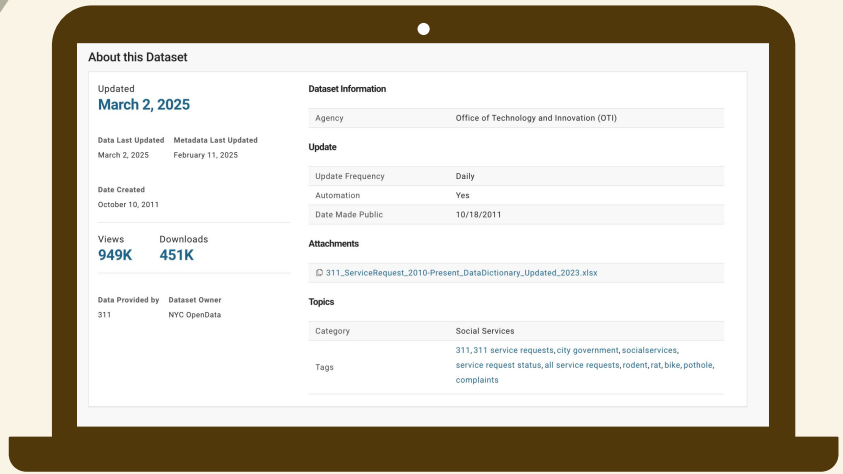
NYC 311 Concerns

- Hard to identify complaint patterns and response times
- Inefficient resource allocation impacting service quality

Our Approach

- Analyze complaint trends & response times
- Map geographic clusters for better staffing
- Extract key themes from resolution data

311 Service Requests Dataset



Source: NYC Open Data – 311 Service Requests (2010–Present)

Size: 28 million+ records, 41 features

Sampling: Selected 1% of the total dataset using stratified sampling, ensuring proportional representation from each year

Data Cleaning & Preprocessing

01

Handling Missing Data

02

Feature Engineering

03

Scaling & Encoding

04

Dimensionality Reduction

EDA

Complaint Clusters

Found patterns in geography, time, and type

Temporal Trends

Weekday peaks (Tuesdays highest), midnight noise spikes

Complaint Channels

Phone (49.1%) > Online (21.4%)

Agency Workload

NYPD, HPD, and DOT handle most complaints

Resolution Time

Some delays indicate inefficiencies

Association Rules Analysis

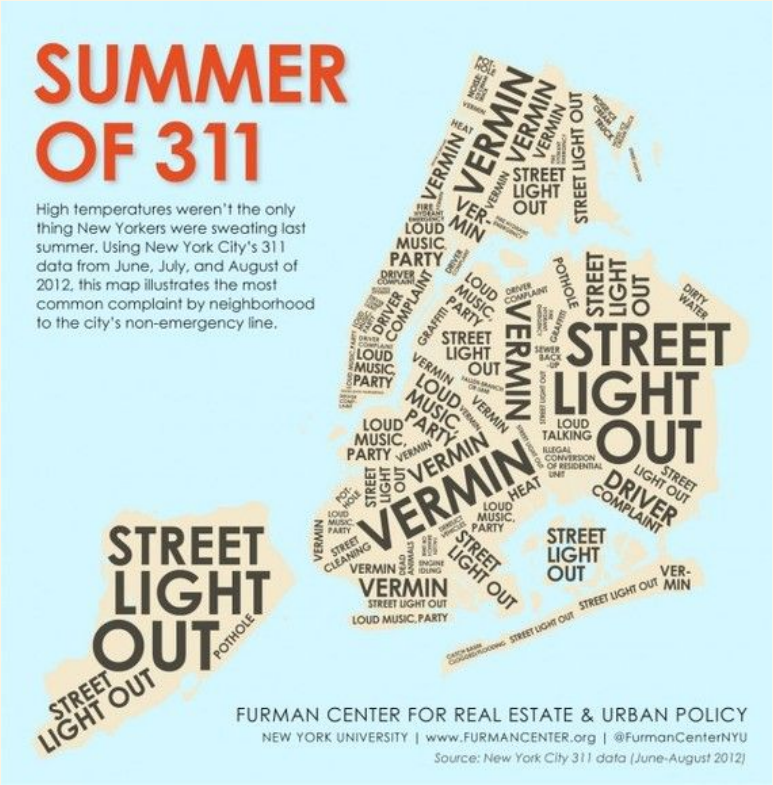
Goal: identify complaint patterns and improve city response strategies.

Key Findings

- Noise: Manhattan & Brooklyn, weekends
- Parking: Queens & Brooklyn (100% handled by NYPD)
- Heating: Bronx & Brooklyn, winter spikes

Impact

- Optimized Resource Allocation – Targeted interventions
- Better Response Strategies – Borough-specific prioritization
- Public Awareness – Reduce misdirected complaints



Increase staffing in high-complaint areas for faster response

Goal: Identify service request hotspots to optimize resource allocation

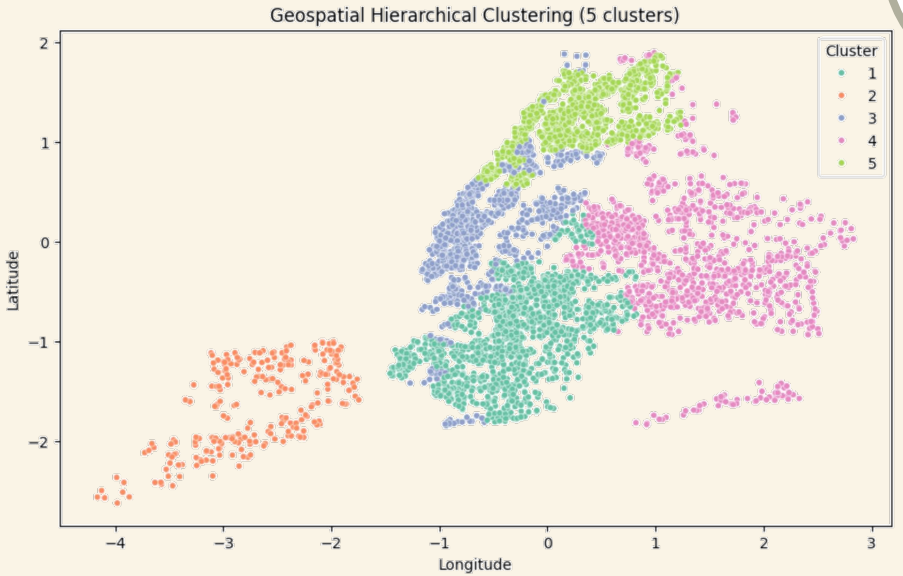
Model: Hierarchical Clustering (Ward's method), K-means clustering (with PCA)

Key Findings

- Identified **5 clusters** that aligned with NYC's boroughs
- **High-density clusters**: Bronx, Manhattan, Brooklyn
- **Low-density clusters:** Staten Island, Queens

Implications

- Targeted resource allocation for faster resolution
- Early detection of emerging service issues



Enhancing complaint resolution through categorization & sentiment insights

Goal: Optimize 311 response by identifying key themes, prioritizing urgent complaints, and analyzing sentiment based on resolution time and type
Models: Bag of Words, LDA Topic Modeling, Clustering, Sentiment Analysis

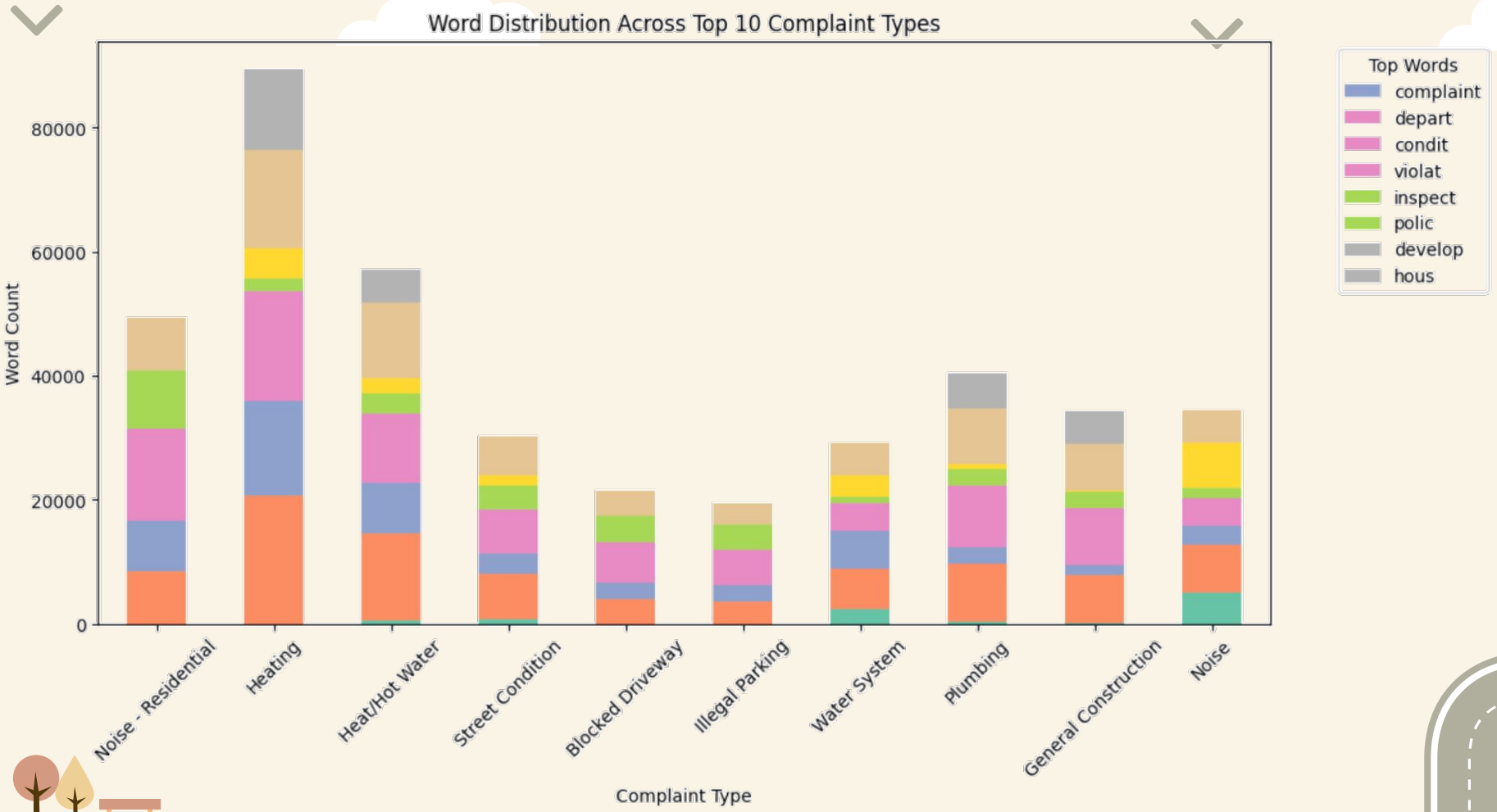
Key Findings

- **4 Key Complaint Topics**: General, Public Space, Urgent, Housing
- **Sentiment Trends**: Positive peaks in April-May, lowest in November

Impact

- Improved Complaint Categorization → Faster resolutions
- Better Resource Planning → Targeted Interventions
- Enhanced Public Satisfaction

Challenges: Switched to LDA to enhance topic modeling, did not use TF-IDF clustering results


Word Distribution Across Top 10 Complaint Types
Word Count
0
20000
40000
60000
80000
Noise - Residential
Heating
Heat/Hot Water
Street Condition
Blocked Driveway
Illegal Parking
Water System
Plumbing
General Construction
Noise
Complaint Type
Top Words
complaint
depart
condit
violat
inspect
polic
develop
hous

Key Takeaways

- High-Complaint Areas: Brooklyn, Manhattan, Bronx
- Peak Complaint Times: Weekdays, midnight spikes for noise
- Top Complaints: Noise, heating, parking dominate 50%+
- Resolution Delays: High-density areas take longer

Practical Impact

- Faster Response Times – Smarter resource allocation
- Better Urban Planning – Addressing recurring service gaps
- Proactive Emergency Response – Deploying teams ahead of seasonal spikes

Thank You!

Questions?