Data 605: Actionable Visualization and Analytics

Data Story Review

Mapping New York's Noisiest Neighborhoods

Reviewed by Success Attoni

01
02
03
04
05
06
07

Outline

1. Intro / Background
2. The Dataset, Analysis approach & Possible Tools
3. Overall Message
4. Key Takeaways
5. Analysis Strengths
6. Areas of Improvement
7. Conclusion

Intro/Background

MAPPING NEW YORK'S NOISIEST NEIGHBORHOODS

Blog Post Published on the NEW YORKER
Published: 17 January 2015

About.

An analysis of over 140,000 noise complaints (calls) logged by New York City's 311* between the winter of 2013 and the fall of 2014 to uncover insights on:

- The major noise sources
- Complaint profile across weekdays and time
- The noisiest and quietest neighborhoods in the NYC.

And…

- Producing fit-for-purpose visuals to best convey the analysis including maps of the city showing analysis

* 311 is NYC's call number for non-emergency city service and information

"Keep it simple and impactful"

Ben Wellington
LinkedIn

- Data Scientist (calls himself a "Data Storyteller")
- PhD in Computer Science (Natural Language Processing) – NYU
- B.S in Math and Computer Science – Bucknell University

- Visiting Assistant Professor, teaching statistics at the City & Regional Planning Program, Pratt Institute, Broklyn, USA.

- Deputy Head of Feature Forecasting at Two Sigma.
- Teaches Public speaking and job training skills at Cherub Improve
- Founder, I Quant NY

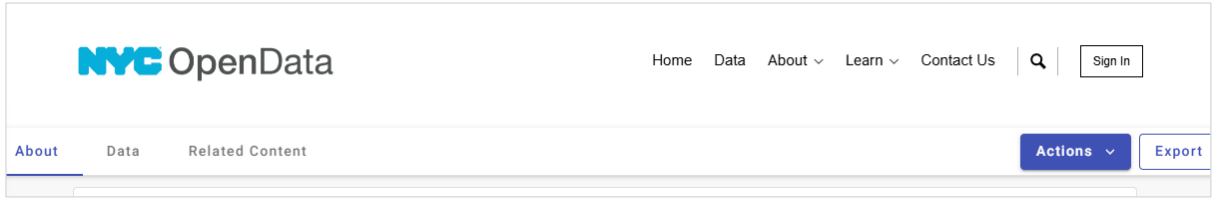
- Ted talk: Making data mean more through storytelling.

Inspiration to work on NY City Data:
Signing of NYC Open Data into law by Mayor Blomberg - 2011
Make a positive impact, cause change, improve NYC.
Married to an urban planner.

The Dataset

NYC 311 Service Requests From 2010 to Present

https://nycopendata.socrata.com/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9/about_data



Update Frequency: Daily
Date made public: October 18, 2011

Date Created: October 9, 2011
Data Format: CSV download available, etc

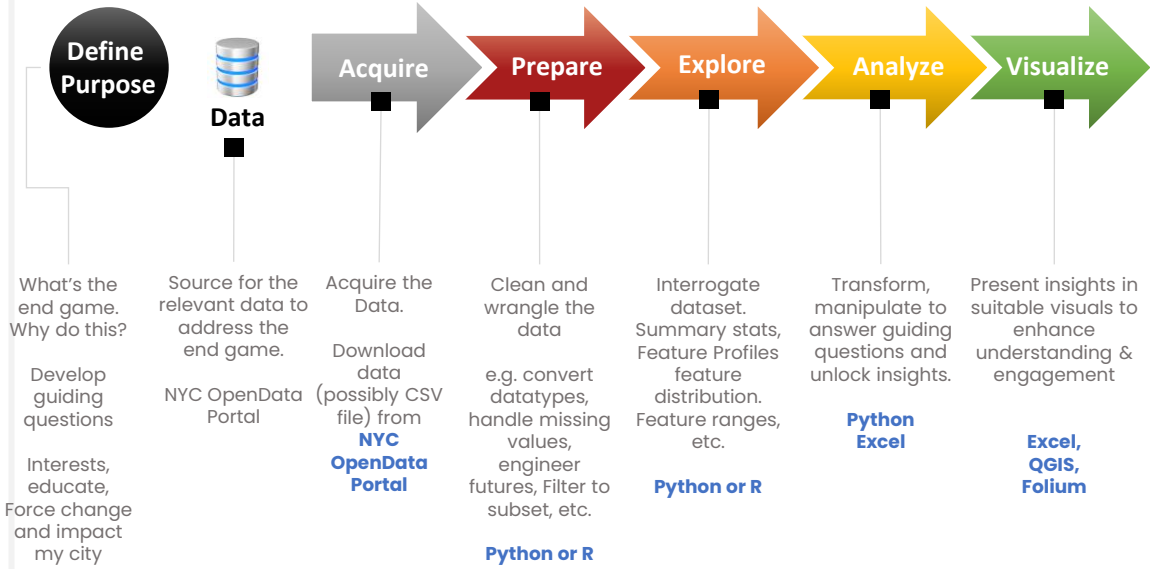
What's in the Dataset?

Rows **38.9M**

Columns **41**

Each row is a **311 Service Request.**

Possible Data Processing Workflow and tools used:



Columns (41)

Column Name	Description	API Field Name	Data Type
Unique Key	Unique identifier of a Service Request (SR) in the open data set	unique_key	Text
Created Date	Date SR was created	created_date	Floating Timestamp
Closed Date	Date SR was closed by responding agency	closed_date	Floating Timestamp
Agency	Acronym of responding City Government Agency	agency	Text
Agency Name	Full Agency name of responding City Government Agency	agency_name	Text
Complaint Type	This is the first level of a hierarchy identifying the topic of the incident. Complaint Type may have a corresponding	complaint_type	Text

Overall Message

- Noice pollution is a critical issue in NYC
- Noise sources within the city is diverse, influenced by urban planning, social behaviour and even cultural differences. It also varies with respect to time of the day and weekdays.
- A data-informed approach is crucial for policy makers to design tailored noise control strategies.

Over

140K

Noise Complaints logged between winter of 2013 and the fall of 2014

Key Take Aways

1. Noise is a Major persistent issue
Average of one complaints every four minutes

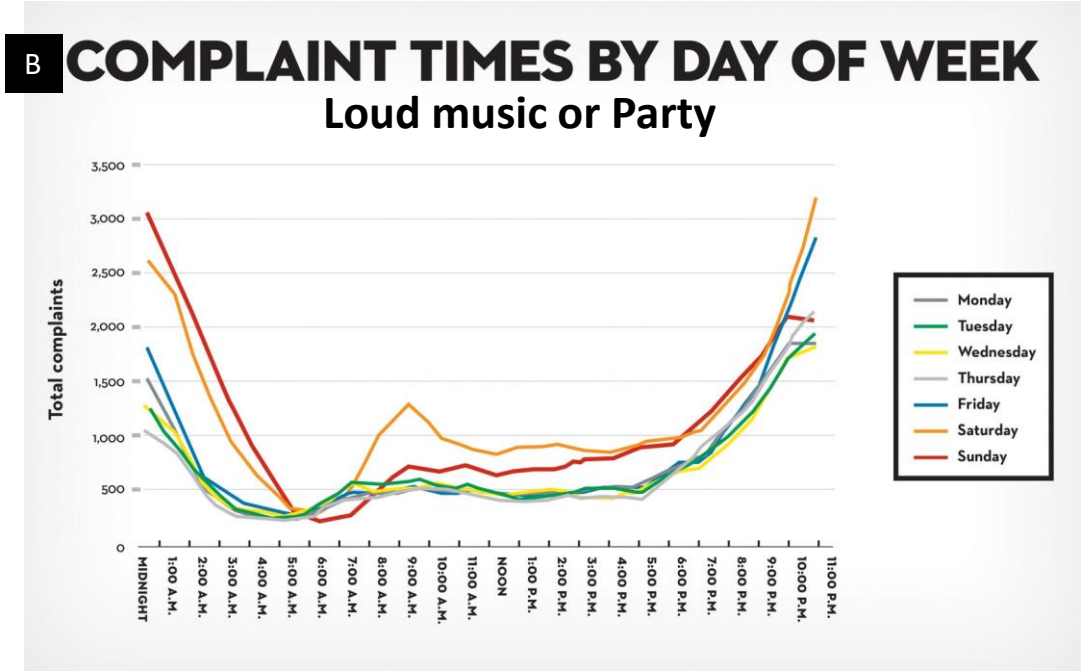
2. Social behaviour drives many noise complaints (fig A)
Most common noise complaints: loud music and parties followed by after hours construction and loud talking.

3. Nose complaints follow distinct patterns (fig B and C)
Complaints peak at 11 pm and midnight. Especially on weekends when parties and loud social gatherings are prevalent.

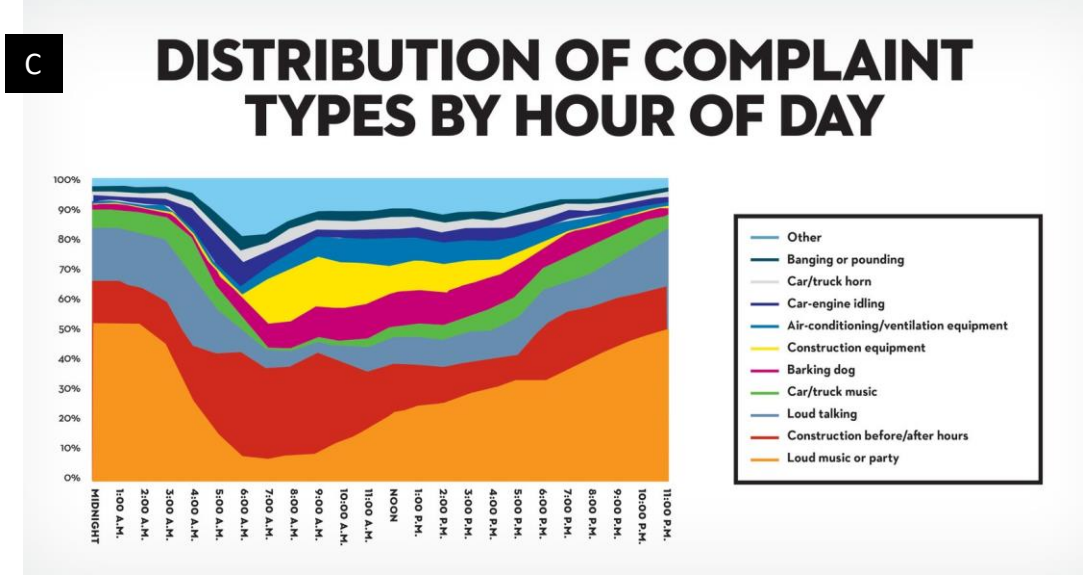
A Distribution of Complaint Type

COMPLAINT TYPE	COUNT	PERCENTAGE
Loud music or party	52,368	37%
Construction before/after hours	23,180	16%
Loud talking	18,210	13%
Car/truck music	8,962	6%
Barking dog	7,480	5%
Construction equipment	5,819	4%
Air-conditioning/ventilation equipment	4,200	3%
Car-engine idling	3,886	3%
Car/truck horn	3,374	2%
Banging or pounding	3,087	2%
Other	10,098	7%

B



C



Key Take Aways

4. Noise distribution varies by neighborhood. (fig A)

- Midtown Manhattan has the highest complaint
- Quieter Neighbourhoods, such as Co-op City (Bronx) planned as a residential development has low level of complaints.

5. Different Neighborhoods have different noise issues (See maps)

- Loud parties: most common in Inwood/Marble Hill.
- Loud talking: mostly reported in Norwood (Bronx)
- Construction noise: concentrated in Midtown, Lower Manhattan, and North Brooklyn

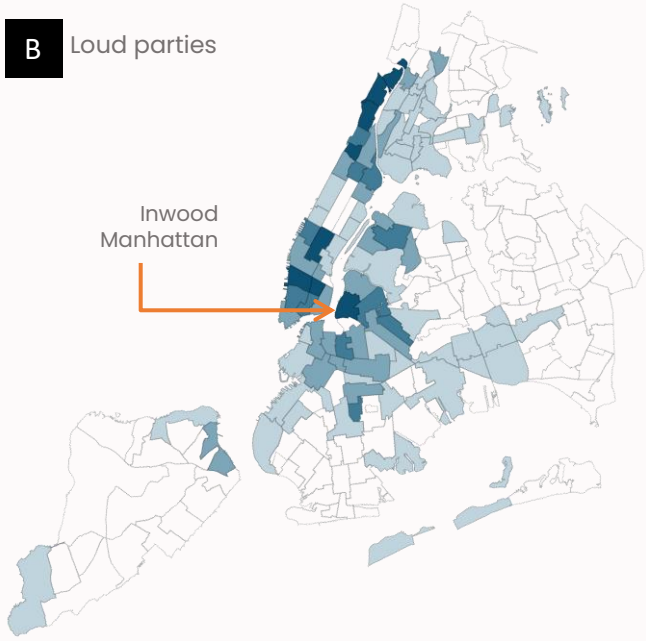
Etc…

A Distribution of Complaint Type

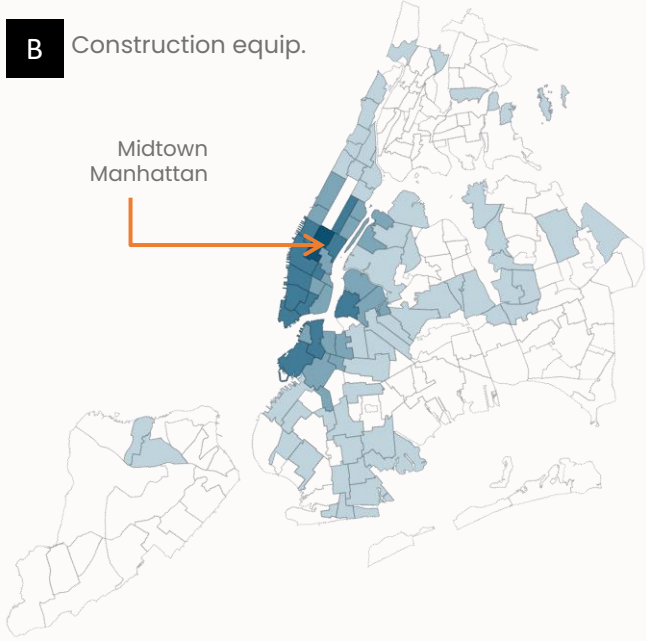
RANK	NEIGHBORHOOD	COMPLAINTS PER THOUSAND RESIDENTS
1	Midtown/Midtown South	104.51
2	North Side/South Side	75.59
3	SoHo/Tribeca/Civic Center/Little Italy	75.57
4	Battery Park City/Lower Manhattan	59.98
5	West Village	59.96
...	...	...
184	Todt Hill/Emerson Hill/Heartland Village/Lighthouse Hill	3.16
185	Oakland Gardens	3.05
186	Rossville/Woodrow	3.03
187	Starrett City	2.25
188	Co-op City	1.10

Mapping of Complaint type

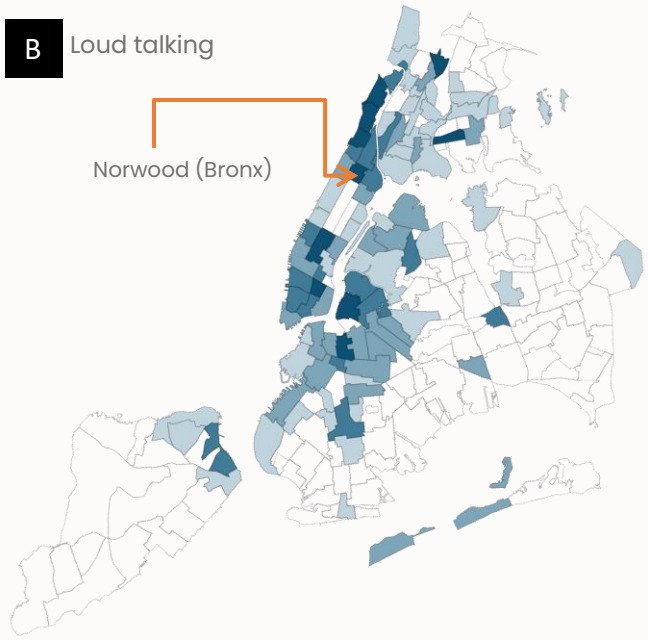
B Loud parties



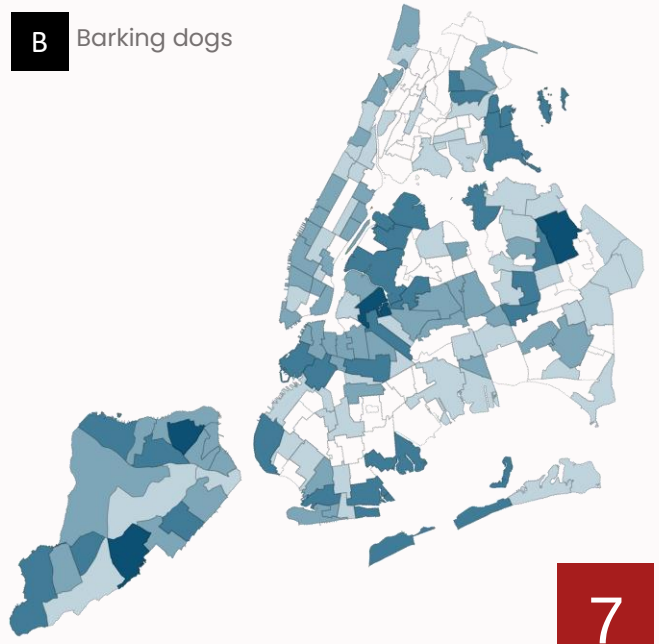
B Construction equip.



B Loud talking



B Barking dogs



Strength of Analysis & Blog Post

1. Used real-world data
- Blog was about a real issue, not a fictitious one, and is relatable.

2. Distilled noise complaints by type, time and geography
- Identified top 10 sources of noise complaints, accounting for 95% of all reports.
- Effectively analyzed trends over time, mapped out noise complaints by neighbourhood area.

3. Effective use of Visualizations.
- Used time-series charts to visualize complaints over time
- Bar charts used to show category-based visuals.
- Geospatial mapping of noise distribution across the city.

4. Provided thoughtful insights and acknowledged data limitations.
- Excellent analysis and insights highlighted, including the interesting observation of spikes in Saturday morning complaint, leading to speculations about human behaviour, sleep habit, and cultural factors

5. Offered recommendation for data driven noise management solutions
- …to the NYPD and Department of Environmental Protection.

6. Accessible and Engaging Storytelling
- Blends data with storytelling – blog is engaging and draws from personal observations, humour. No technical jargon – easy to understand for non-technical readers.

Engaging, light, simple read with clear messaging.

Areas of Improvement

1. Potential biases in the data (311 complaints <> Actual Noise Levels)

- Assumes 311 complaints reflect actual noise levels. | Correlate complaint data with actual noise measurements (e.g. decibel readings from environmental sensors)
- Lacks depth on effect of cultural differences in tolerance to noise | Discussion on complaint underreporting in certain demographics could add depth.

2. Limited discussion on data accuracy and missing information.

- No discussion about data quality related issues (missing information, duplicates, misclassification, etc.) | Incorporation of a paragraph on data issues and resolution, data cleaning process would improve transparency.

3. Insufficient consideration of external factors affecting noise.

- No consideration was made to city events (e.g. parades, concerts, sports games, etc), weather consideration, etc that could result spikes in complaints.| Overlaying event schedule, weather data, etc may offer more insights.

4. Did not propose concrete solutions to address noise pollution.

- No discussion on noise reduction strategies (e.g. zoning changes, insulation requirements, etc.) | Proffer practical solutions based on the data.

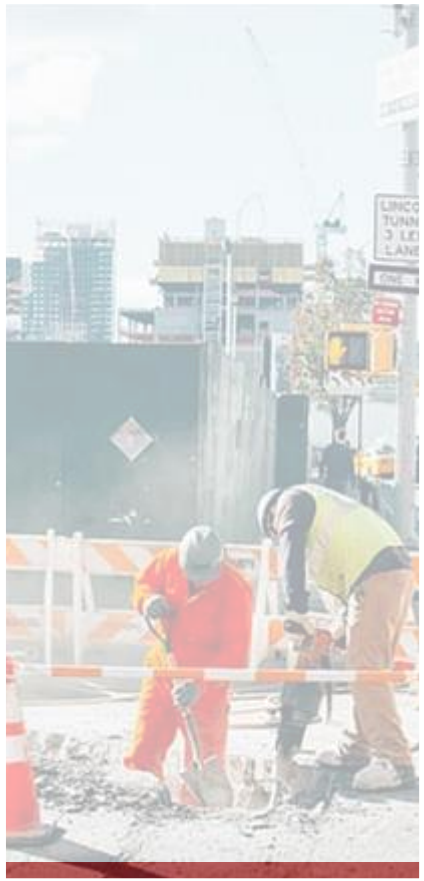
5. Visualizations could be interactive.

- Visuals (map and plots) are static, limiting the user from exploring the analysis. | Make maps and visuals interactive (possibly use plotly).
- Visuals are not numbered, and maps do not have legend. | Number all visuals.

"Complaint" is not equal to noise level.

Effect of external conditions resulting in more complaints not extensively highlighted.

Interactive visualizations (especially the maps) will offer the user the ability to explore the data more.



"A compelling data-driven analysis of NYC's noise complaints, revealing patterns in urban sound and social behavior.

Strong visuals and insights, could benefit from making visuals interactive, considering external factors in the analysis, and proposing policy solutions."

Conclusion

- Well written blog post on New York City's Noise Landscape.
- 140,000+ 311 noise complaints analyzed to reveal the city's noisiest and quietest neighbourhoods.
- Nighttime peaks for social noise, morning spikes for construction
- Noisiest neighbourhood is Midtown, planned residential areas are the most quiet.

Strengths

- Effective use of data.
- Strong visual storytelling through text, maps and charts.
- Insightful discussion on neighborhood and noise source.

Improvement Areas

- Note that noise complaints is not same as actual noise levels
- Adding external factors that may cause spikes in complaint profile (e.g. city wide parades, sporting events, etc) will make the analysis more robust
- Make visuals interactive and add legends where necessary.
- Propose strategies to tackle the menace of noise pollution.

Thank You