

# Jessica Freaner

DATA SCIENTIST

# Who Am I ?

Psychology BA - NYU

Self taught programmer

4+ years experience in Trade Finance as  
programmer

World traveler - Multilingual

Data Scientist



# Gentrification

Identifying Early Signals of Impending Change

“New York, concrete jungle where dreams are made  
There's nothing you can't do  
Now you're in New York”

—Jay Z



# Gentrification

# Gentrification

**DEFINITION:** rapid change of demographics in city

# Gentrification

**DEFINITION:** rapid change of demographics in city



# Gentrification - Signals

- **Property Values:**

( % ) Change from  
Previous Year



# Gentrification - WHY does it happen Where it does???

Historical NYC Examples:

SoHo  
East Village  
Williamsburg

# Gentrification

## FIRST - Artists / Young People



# Gentrification - Signals



# Gentrification - Signals THEN Development



# Gentrification - Signals AND Tourists



# How to Capture Change???

# How to Capture Change???

**HYPOTHESIS:** people's voluntary movements (i.e., not for work) signal movements of demographics

# How to Capture Change???

**HYPOTHESIS:** people's voluntary movements (i.e., not for work) signal movements of demographics

Where do people go for FUN?

# How to Capture Change???

**HYPOTHESIS:** people's voluntary movements (i.e., not for work) signal movements of demographics

Where do people go for FUN?

**INSIGHT:** look at MTA turnstile data during  
Fri. / Sat. Night  
Sat. / Sun. Brunch



# NOT Morning Commute



# Nite Life on Subway



# Brunch!!!



# Gentrification - Product

- **Interactive Visualization:** historic Property Values / MTA Activity  
( 2010 - 2014 )

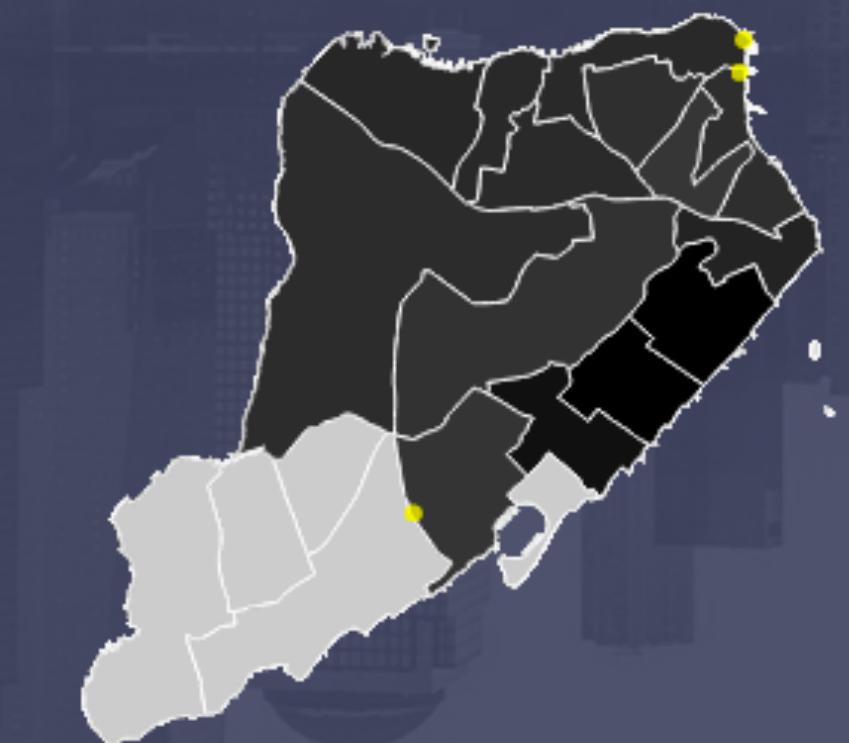
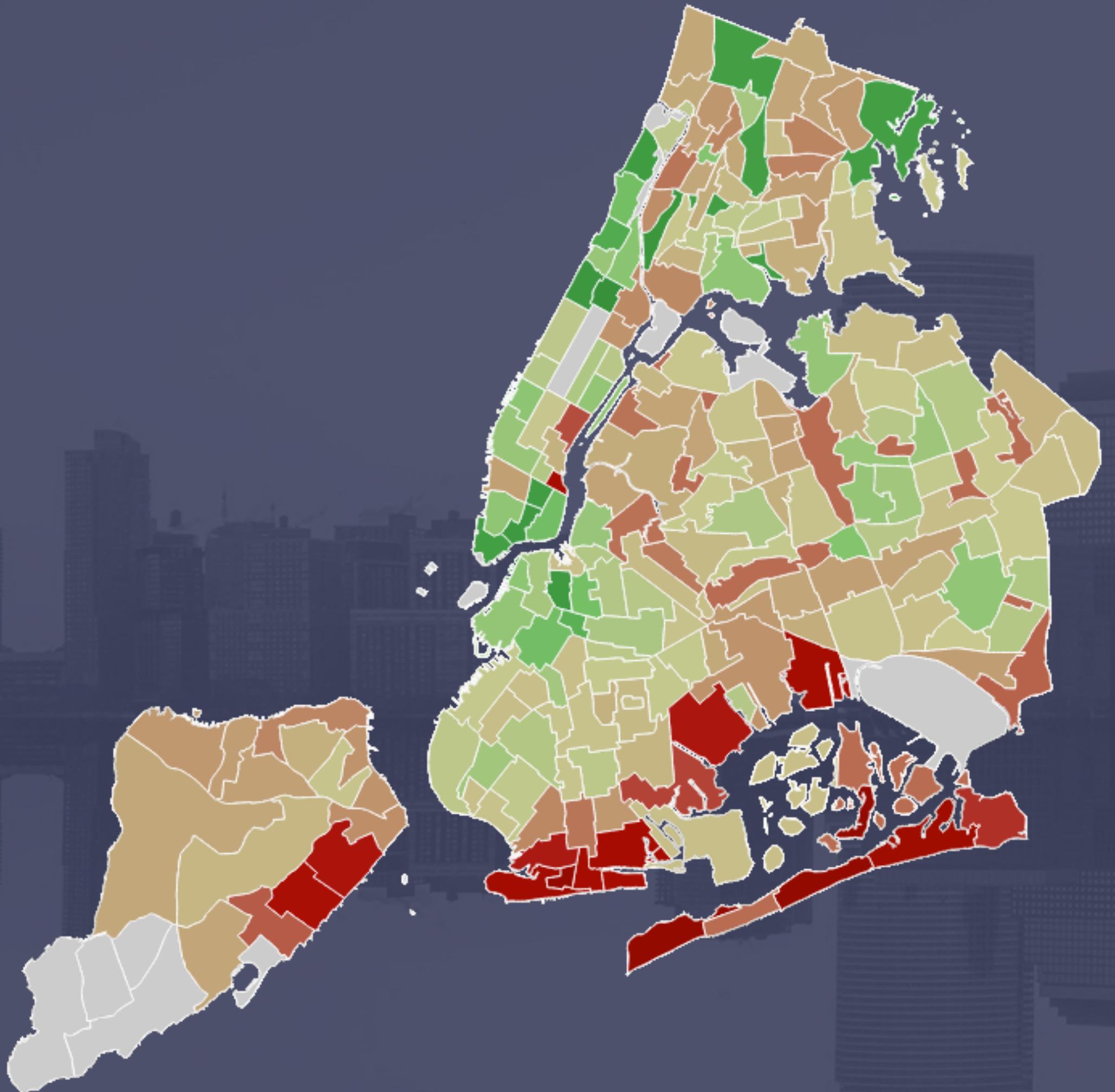
# Visualizing Residential Property Values & MTA Traffic in NYC

by Jessica Freaner

2014

Adjusted for Inflation

Fri Nite



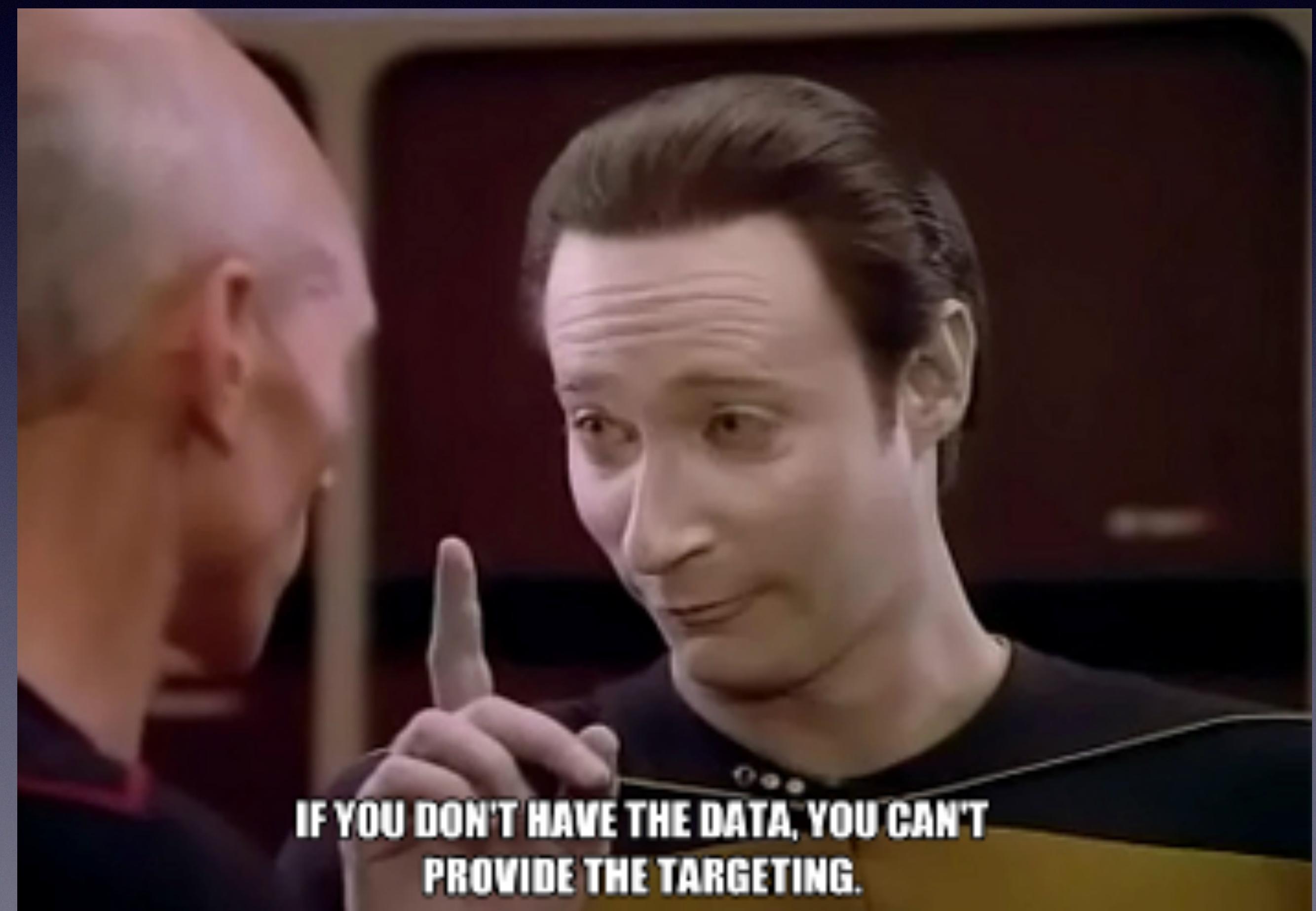
# Utility

- **Developers:** profit
- **Communities:** empower community boards to influence incoming development to serve their interests
- **Renters / Buyers:** guide for neighborhoods to avoid
- **D3 Users:** explore utility of D3 (will be published on GitHub)

# Data Sources

- **Turnstile Counts** - web scraped (BeautifulSoup) urls & wget for MTA
- **MTA Station Lat./Long.** - GitHub repository
- **NTA GeoJSON & Property Values** - Extracted from d3 visualization by MIT Labs
- **Inflation Adjuster** - Wolfram Alpha

The image features a large, semi-transparent watermark in the center that reads "DIRTY DATA". The letters are bold and have a slight drop shadow. The background consists of a grid of binary digits (0s and 1s) arranged in rows, creating a digital or data-oriented aesthetic.



# Data Complexity

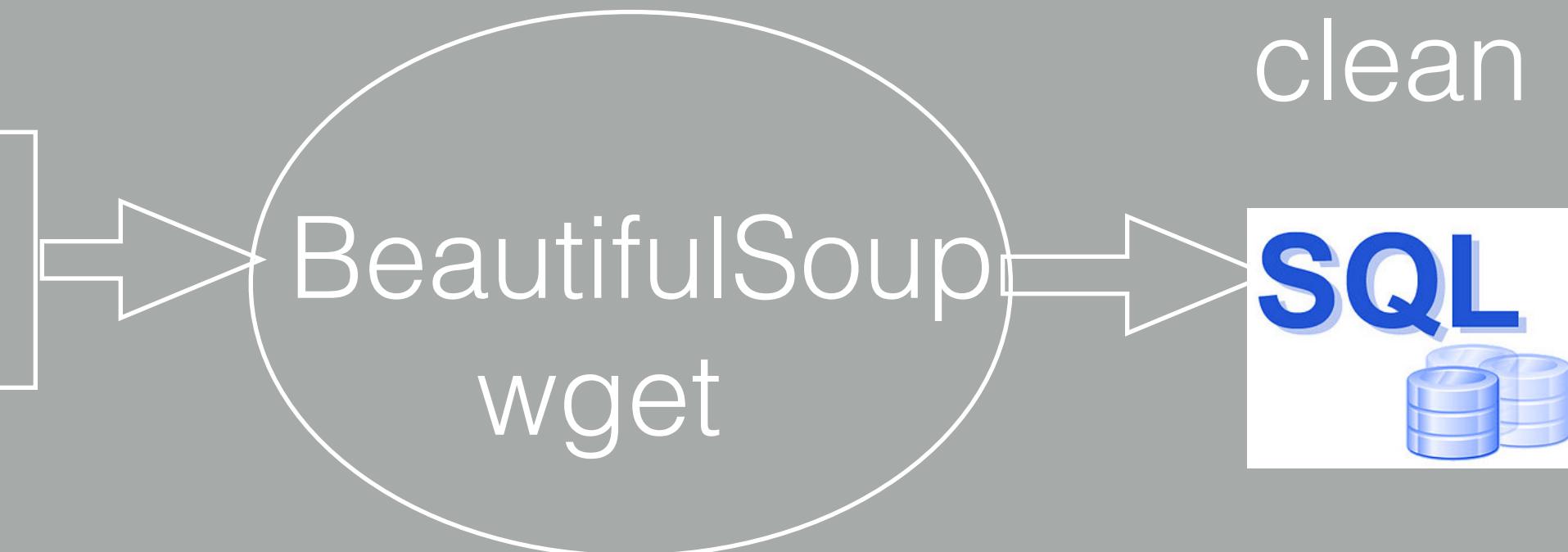
- **Turnstile Counts** -
  - combined data across **2 formats**
  - handled **cumulative counts** by turnstile; combined to station level
  - combine across **different dates** for Friday & Saturday Night (**8pm - 4am**)
  - included **PATH** stations, **MTA Card Van**
  - **errors** in counts ( e.g., 35,000,000 entries in 1 turnstile!!! )
- **MTA Station Lat/Long.** - GitHub repository
  - point-in-polygon logic: initial polygons located over **Long Island**; OOPS!
  - Some stations mistakenly identified as in **park-cemetery-etc**; OOPS!

# Cleaning Data!!!





MTA Turnstile Data



clean



MTA Geocodes



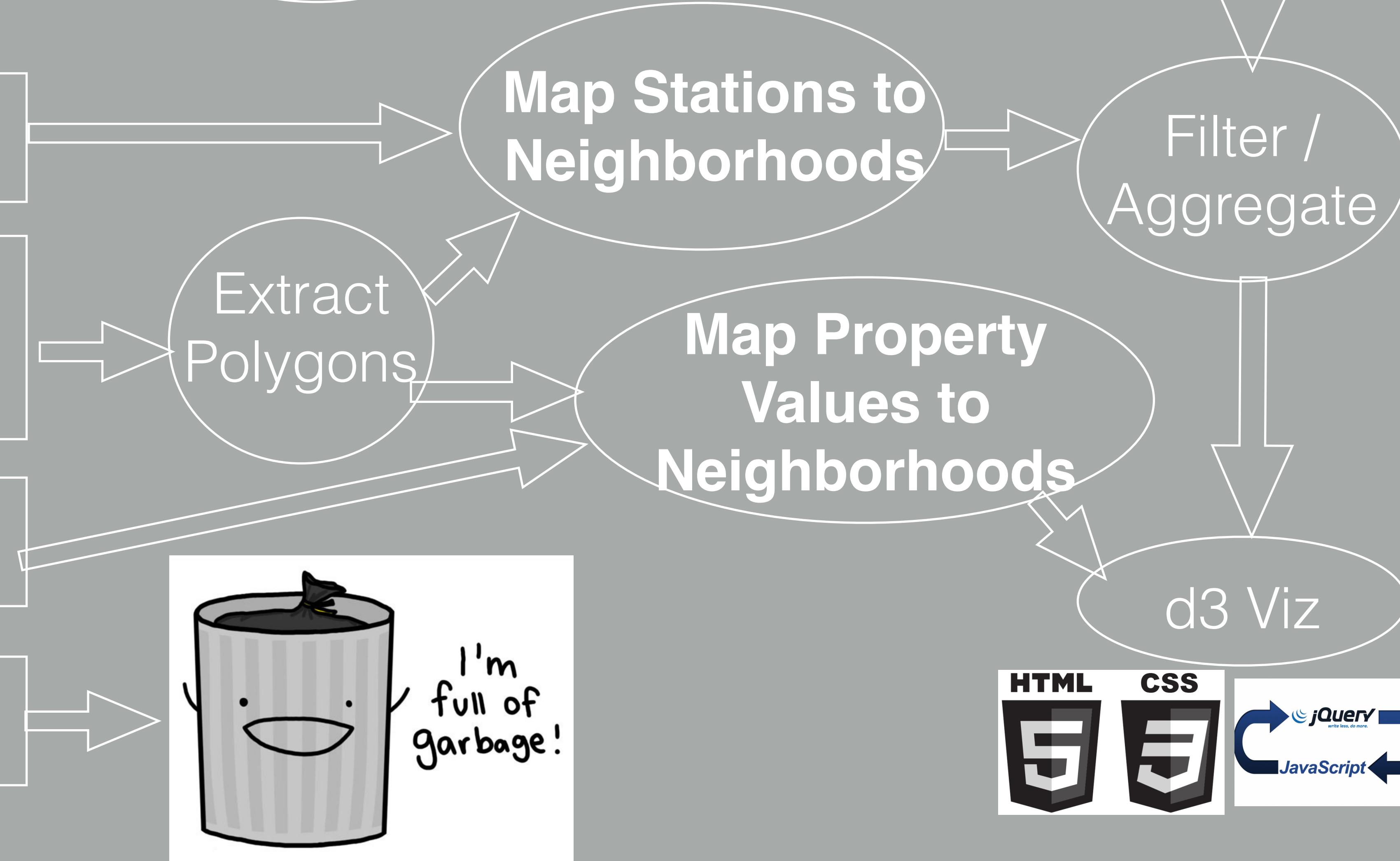
Property Values  
&  
GeoJSONs



Inflation Adjuster



OTHER DATA



HTML5  
CSS3  
jQuery  
JavaScript