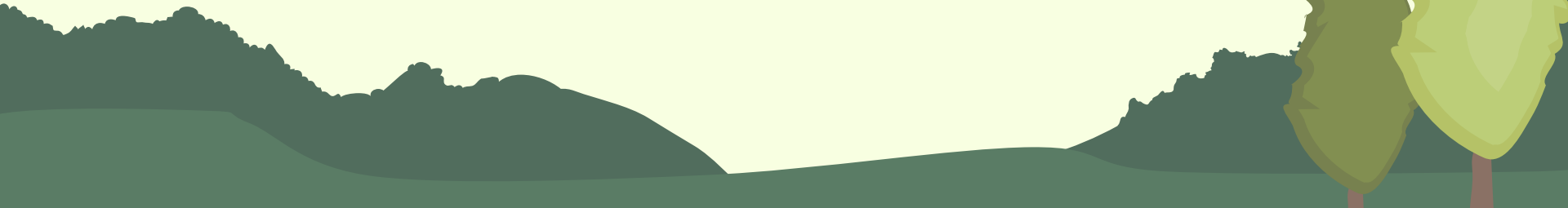




NYC Environmental Health: Analyzing the Relationship between Urban Trees and Air Quality

Team Members: Aaron, William,
Simon, and Bryan

Project Goal: To investigate patterns between
NYC's street tree distribution, air quality
metrics, and neighborhood characteristics



DATA SOURCES AND COLLECTION



NYC Open Data 2015 Street Tree Census

- 666,134 trees
- Species, health, and location data
- Geographic distribution

NYC Open Data Air Quality Surveillance Data

- Multiple pollutant measurements
- Historical trends (2008-2022)
- Geographic distribution by neighborhood

Supporting Geographic Data:

ZIP Code Land Area Data

- Web-scraped from USA.com with BeautifulSoup
- Automated collection for all NYC ZIP codes

Geographic Crosswalk Tables

- UHF34 (United Hospital Fund) to ZIP code mapping
- Integration with borough data
- Enables density calculations

DATA PROCESSING HIGHLIGHTS

Tree Data Features:

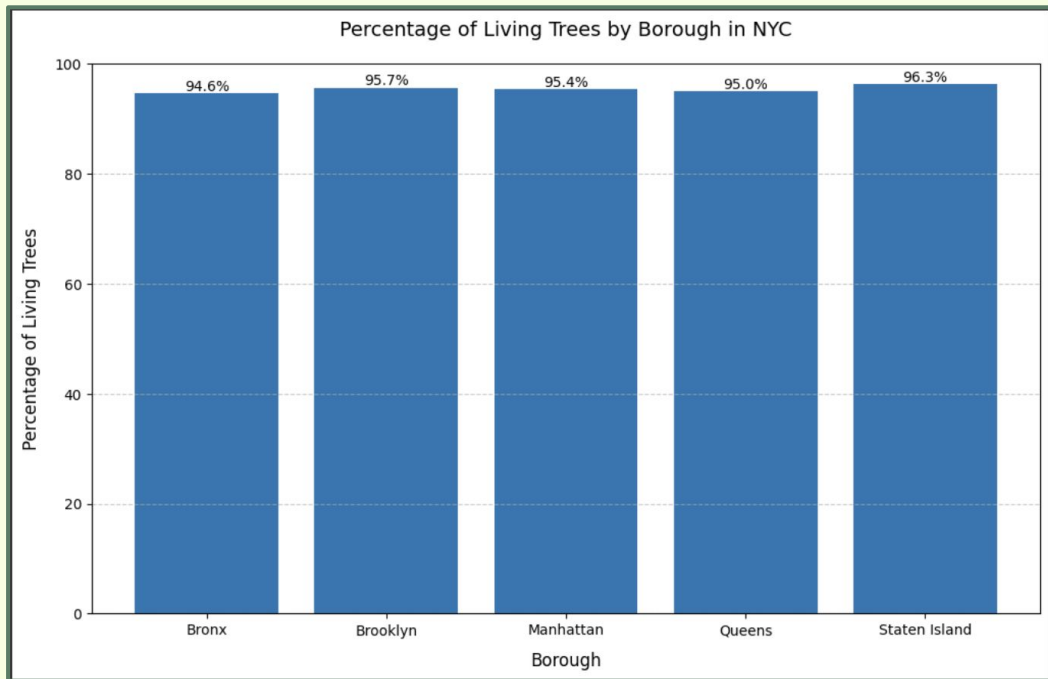
1. Status categorization
(Alive/Stump/Dead)
2. Location type
(OnCurb/OffsetFromCurb)
3. Health indicators
4. Survey metrics
5. Infrastructure impact
(sidewalk damage)

Geographic Integration:

1. ZIP code area
normalization
2. UHF34 neighborhood
matching
3. Borough-level aggregation

VISUALIZATION OF TREE DATA

KEY VISUALIZATIONS - TREES



Distribution Patterns

Species diversity, Health metrics

Summary Statistics:

Total number of trees: 426224

Number of trees by borough:

```
borough
Queens      155708
Brooklyn    116072
Staten Island 57008
Bronx       49857
Manhattan   47579
Name: count, dtype: int64
```

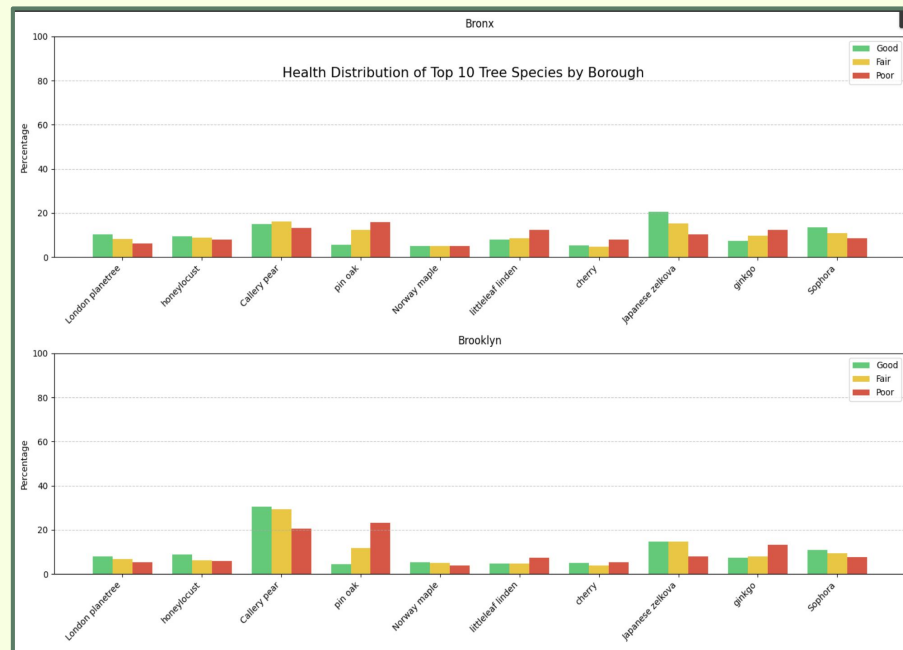
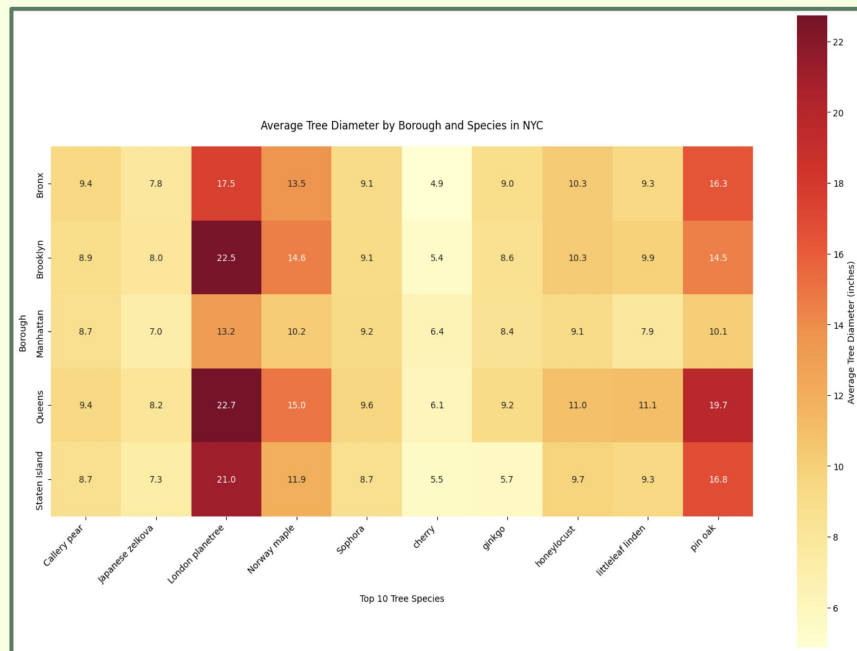
Number of trees by species:

```
spc_common
London planetree 87014
honeylocust      64264
Callery pear     58931
pin oak          53185
Norway maple     34189
littleleaf linden 29742
cherry           29279
Japanese zelkova 29258
ginkgo           21024
Sophora          19338
Name: count, dtype: int64
```

Average diameter by species:

```
spc_common
London planetree 21.560657
pin oak          16.867707
Norway maple     14.330516
honeylocust      10.210958
littleleaf linden 10.045827
Sophora          9.254628
Callery pear     8.958307
ginkgo           8.625476
Japanese zelkova 7.863559
cherry           5.691041
Name: tree_dbh, dtype: float64
```

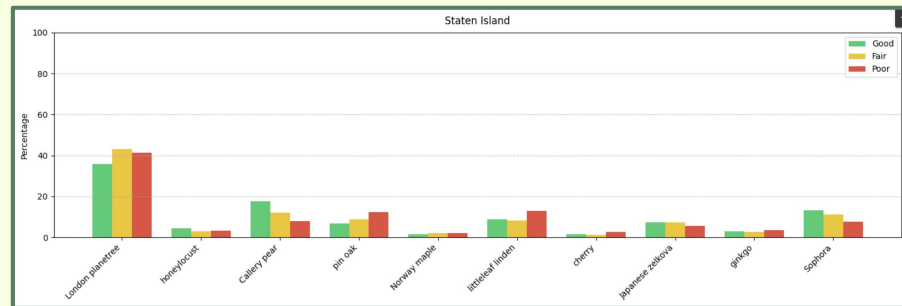
KEY VISUALIZATIONS - TREES



Top 10 Tree Species Analysis:

- Interactive visualization showing diameter variations
- Color intensity represents average tree diameter in inches
- Borough-by-borough comparison reveals geographic patterns
- Highlights species adaptation to different urban environments

KEY VISUALIZATIONS - TREES



Summary Statistics:

Overall health distribution for top 10 species:

```

health
Fair    14.3
Good    81.9
Poor     3.8
dtype: float64
    
```

Health distribution by species (percentages):

health	Fair	Good	Poor
spc_common			
health			
Callery pear	14.9	81.6	3.5
Japanese zelkova	10.9	86.4	2.7
London planetree	13.2	84.3	2.5
Norway maple	26.8	62.1	11.1
Sophora	14.9	81.9	3.2
cherry	11.8	83.8	4.5
ginkgo	13.2	81.5	5.3
honeylocust	13.3	84.8	1.9
littleleaf linden	14.9	79.3	5.8
pin oak	12.0	85.7	2.3

Health distribution by borough (percentages):

health	Fair	Good	Poor
borough			
Bronx	12.5	84.2	3.3
Brooklyn	13.9	82.8	3.2
Manhattan	18.3	76.3	5.5
Queens	14.3	81.9	3.8
Staten Island	13.7	82.6	3.8

Top 10 Tree Species Analysis:

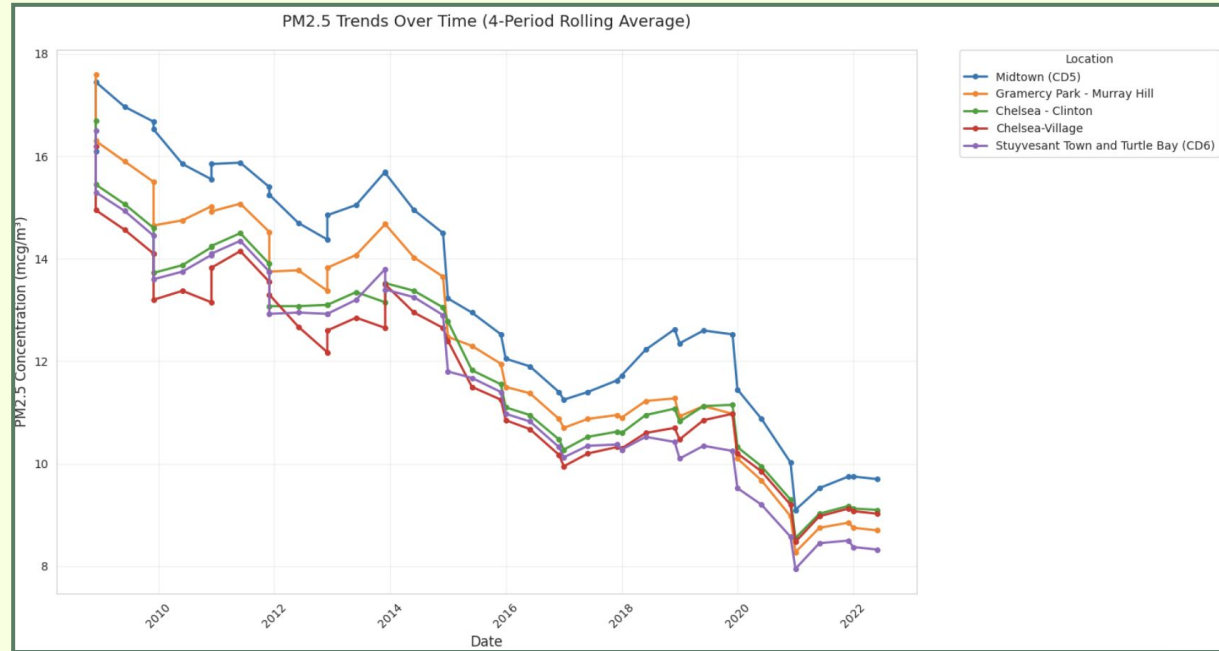
- Focus on living trees only (filtered dead/stump data)
- Top 10 most common species across NYC
- Health categories distribution
- Species resilience patterns
- Correlation with location and environment

Visualization of Air Data

KEY VISUALIZATIONS - AIR

PM2.5 Trends (2008-2022):

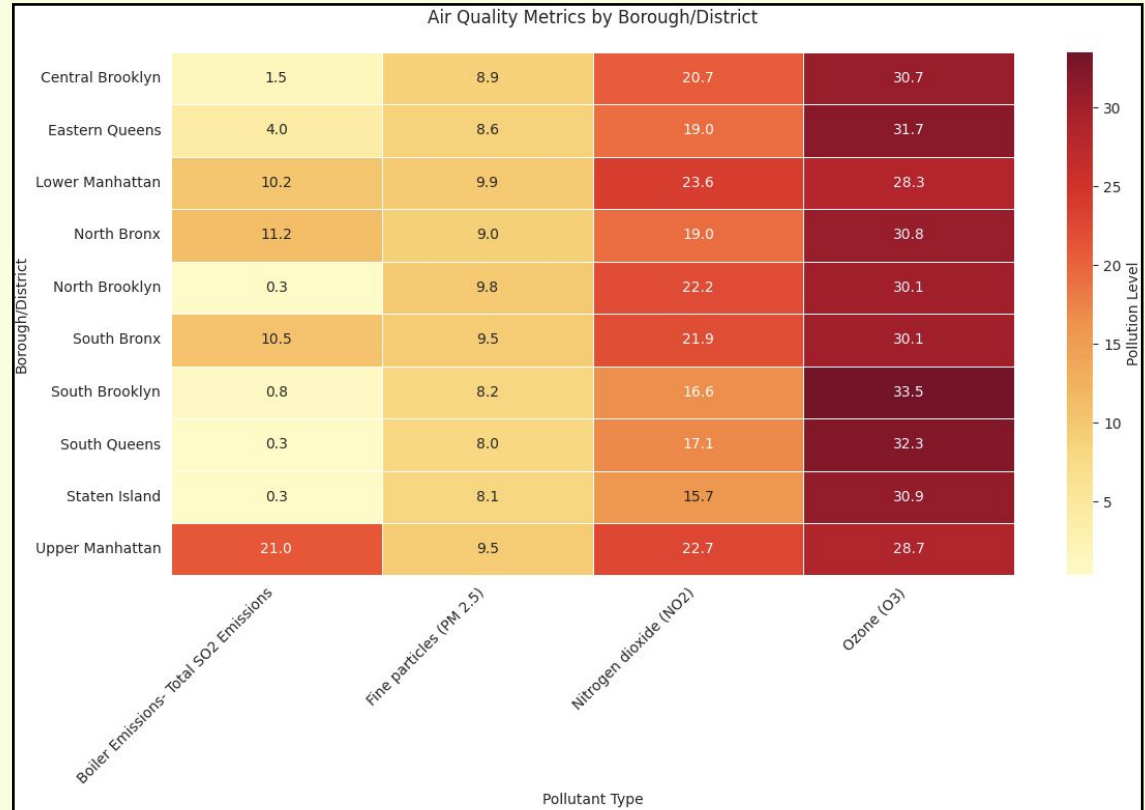
- Line graph showing dramatic improvement
- Top 5 worst-performing locations tracked
- Nearly 50% reduction across all monitored areas
- Year-over-year improvement patterns
- Seasonal variation considerations



KEY VISUALIZATIONS - AIR

Pollution Source Breakdown

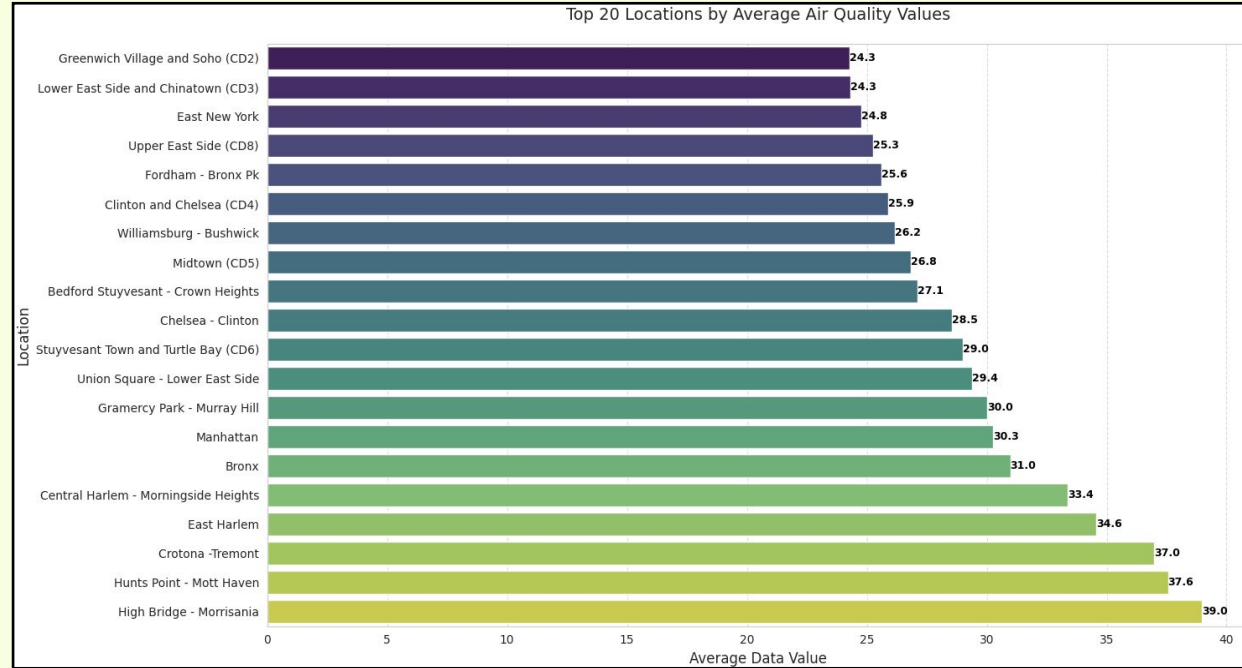
- Annual vehicle miles traveled correlation
- Building boiler emissions impact
- Combined effect analysis
- Asthma ED visits correlation
- Neighborhood-specific patterns



KEY VISUALIZATIONS - AIR

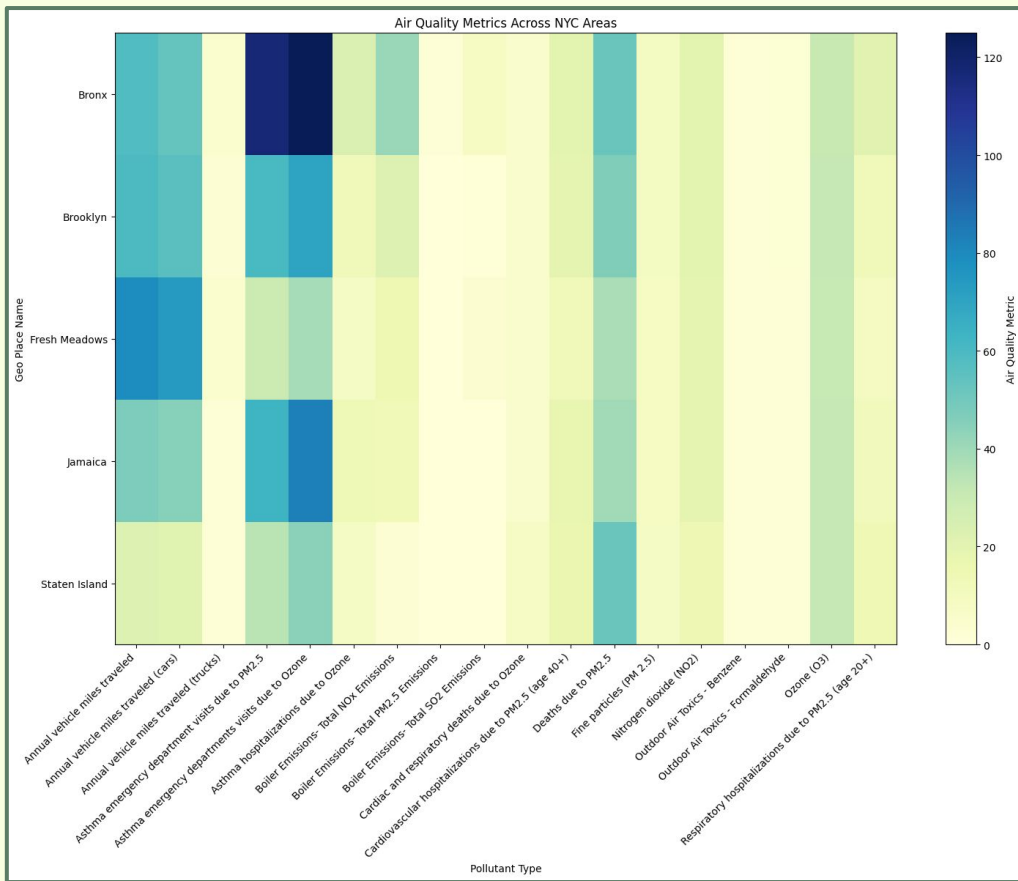
Pollution Source Breakdown

- Heat map of pollution concentration
- Upper Manhattan/Harlem as notable concern area
- Bronx pollution patterns and clusters
- Ozone levels showing unique distribution
- Cross-neighborhood comparison



INTEGRATED ANALYSIS: **TREES AND AIR QUALITY**

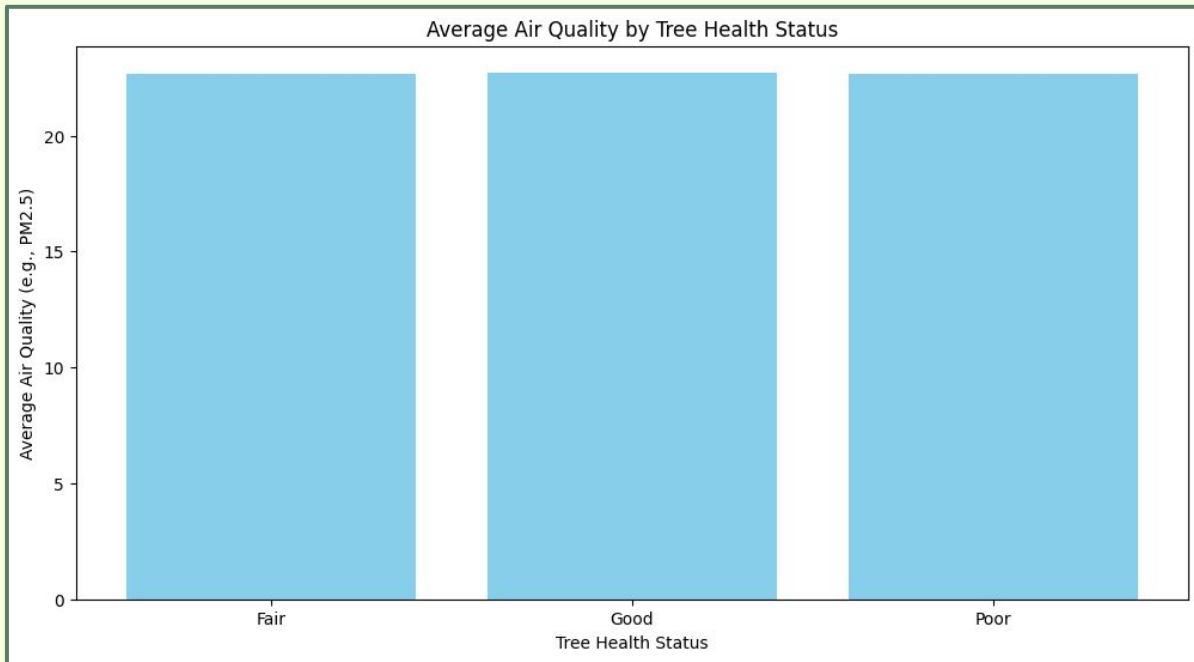
INTEGRATED ANALYSIS: TREES AND AIR QUALITY



Comprehensive Pollutant Distribution Heatmap:

- X-axis showcasing multiple pollutant types (NO2, SO2, PM2.5)
- Y-axis displaying NYC neighborhood names (Geo Place Names)
- Color gradient: darker colors indicate higher pollutant concentrations
- Side color bar scale showing precise value ranges
- Interactive tooltips showing exact values for each cell
- Clear patterns of pollution concentration by area
- Correlation overlay with tree density metrics

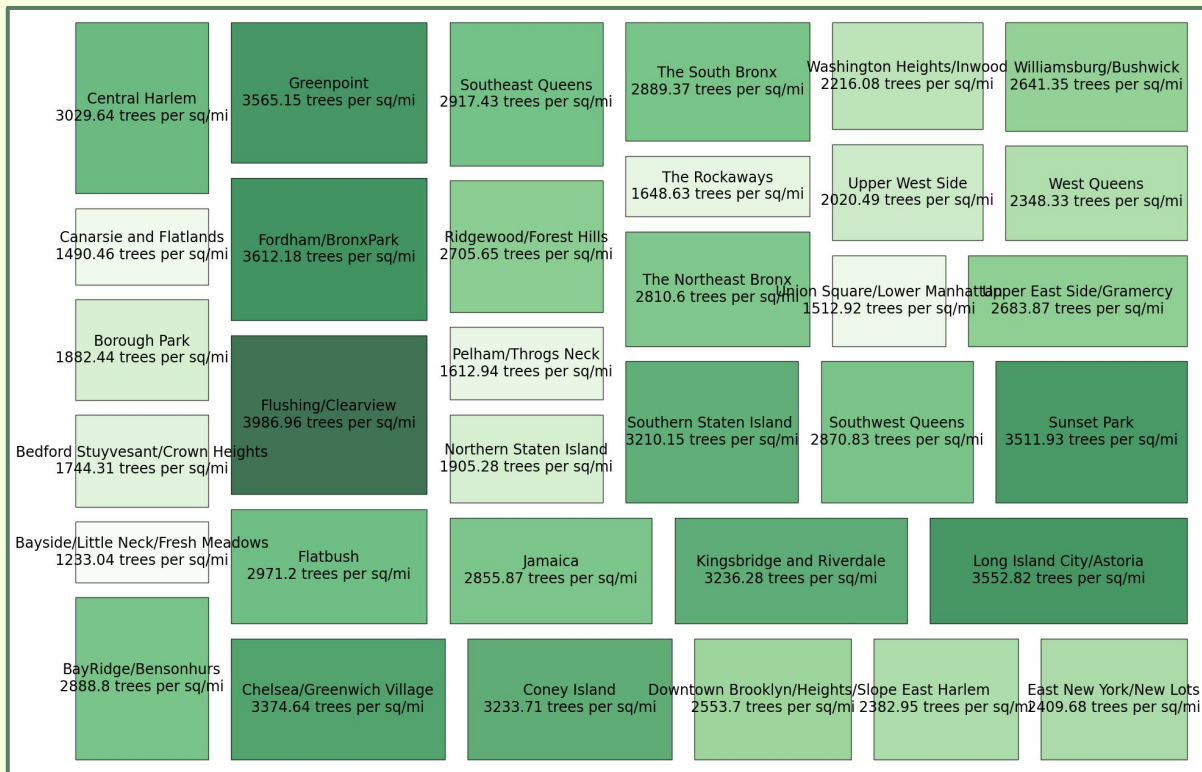
INTEGRATED ANALYSIS: TREES AND AIR QUALITY



Tree Health and Air Quality Relationship Bar Chart:

- X-axis categorizing tree health status (Good/Fair/Poor)
- Y-axis showing average PM2.5 concentration levels
- Error bars indicating confidence intervals
- Color-coded bars for easy health status distinction
- Annotations highlighting significant differences
- Sample size indicators for each category
- Statistical significance markers (p-values)

INTEGRATED ANALYSIS: TREES AND AIR QUALITY

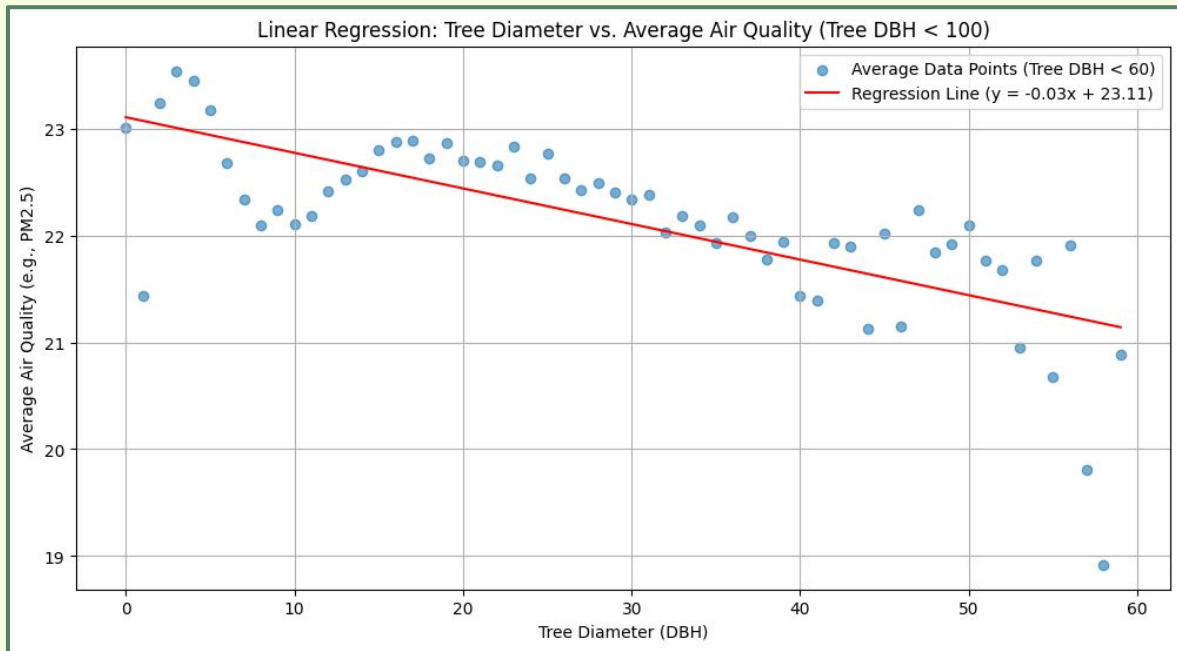


Tree Density By NYC UHF 34 Neighborhoods

- UHF34 Neighborhoods is a collection of zip codes, which stands for United Hospital Fund (UHF 34) Neighborhood Index
- Top three UHF Neighborhoods with the highest density:
 - Flushing/Clearview
 - Fordham/Bronx Park
 - Greenpoint

FITTING OF TREE AND AIR DATA

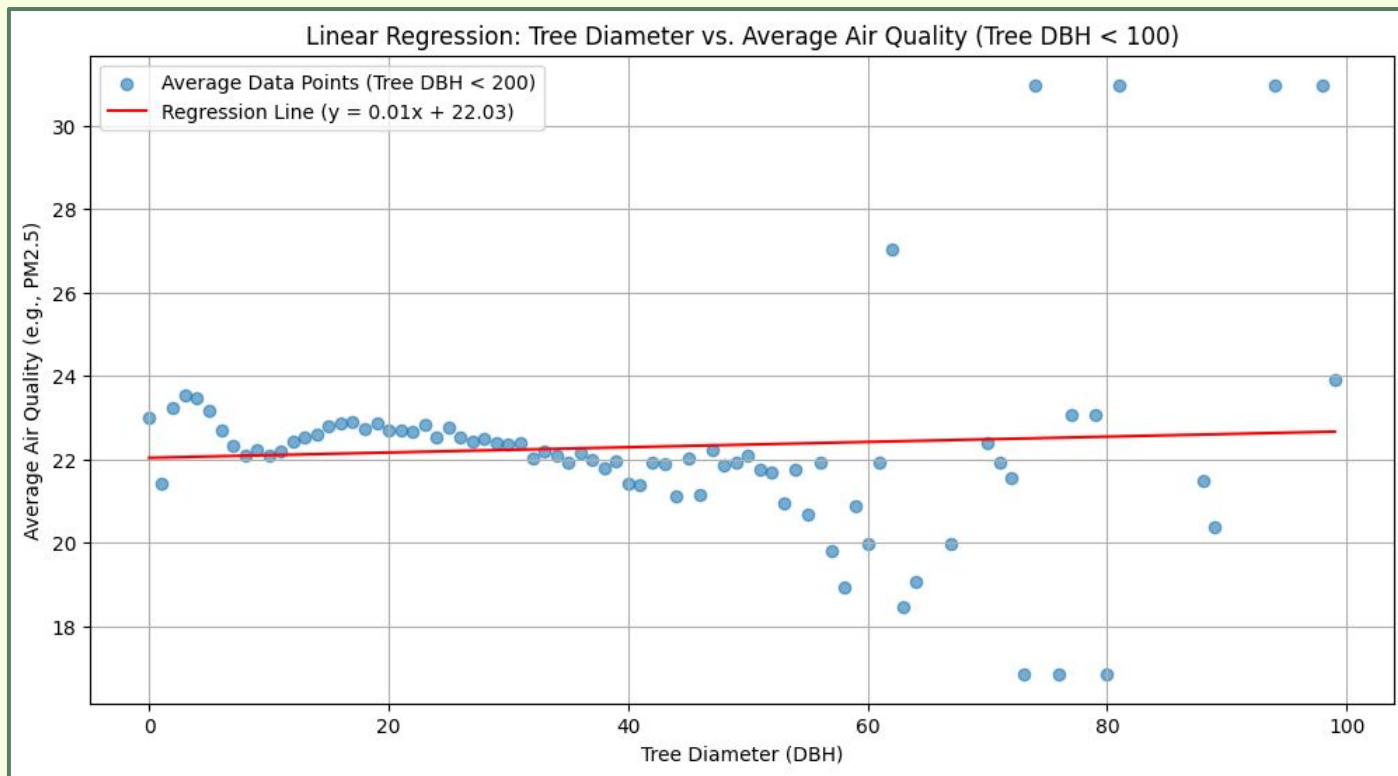
STATISTICAL MODELING AND RELATIONSHIPS



Linear Regression Visualization:

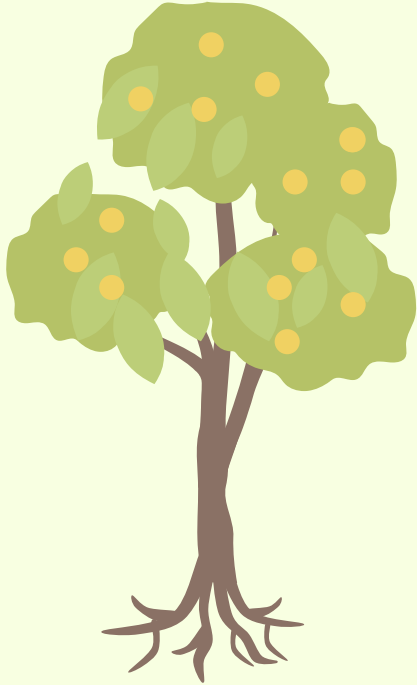
- X-axis: Tree Diameter (DBH) in inches with clear scale
- Y-axis: PM2.5 concentration ($\mu\text{g}/\text{m}^3$) with standard units

LINEAR REGRESSION VISUALIZATION (RESIDUALS)



EMERGING INSIGHTS AND NEXT STEPS

IMPORTANT FINDINGS AND DIRECTION OF PROJECT



Key Findings:

- Harlem and Bronx pollution patterns
- Population density impacts
- Vehicle traffic correlation
- Building emissions effects

Health Implications:

- Asthma ED visits correlation
- Environmental justice considerations

Next Steps:

- Deeper statistical analysis
- Integration of demographic data
- Funding allocation analysis
- Policy recommendation development

THE END