Urban Development & Public Sentiment Analytics Project Report

Project Overview

The objective of this project was to develop a comprehensive **urban analytics dashboard** that integrates various data sources to help municipal governments optimize resource allocation, improve public services, and better understand citizen sentiment. This project utilized **311 service requests**, **public transport data**, and **real-time social media sentiment** to provide actionable insights for city planners.

Problem Statement

City planners and municipal governments often rely on outdated census data and anecdotal complaints to allocate resources for public services like transport, sanitation, and infrastructure maintenance. This leads to inefficient spending and a disconnect from real-time citizen needs. The aim of this project was to:

1. Create an **urban analytics dashboard** that aggregates multiple data sources, including:
    - 311 service requests
    - Public transport usage data
    - Real-time public sentiment from social media
2. Provide actionable insights for city planners to:
    - Visualize problem hotspots
    - Understand public opinion on new projects
    - Allocate resources based on current, evidence-based needs

---

Data Sources Used

Synthetic dataset was used to generate and randomize data for this project via python

```python
import pandas as pd
import numpy as np
import random
from datetime import datetime, timedelta

# Set seed for reproducibility
np.random.seed(42)

# ----- Shared Reference -----
neighborhoods = [
    'Downtown', 'Uptown', 'Harlem', 'Midtown', 'Brooklyn Heights',
    'Queens Village', 'Bronx Park', 'SoHo', 'Chinatown', 'Chelsea'
]
boroughs = ['Manhattan', 'Brooklyn', 'Queens', 'Bronx', 'Staten Island']

start_date = datetime(2025, 1, 1)

# Simulated baseline "satisfaction index" for neighborhoods (to drive correlations)
neighborhood_sentiment_index = {
    n: np.random.uniform(-0.3, 0.3) for n in neighborhoods
}

# ------------------------------------------------------------
# 1 311 SERVICE REQUESTS
# ------------------------------------------------------------
complaints = ['Pothole', 'Streetlight Out', 'Garbage Overflow', 'Noise Complaint',
              'Water Leakage', 'Tree Fallen', 'Graffiti', 'Broken Sidewalk']
statuses = ['Open', 'Closed', 'In Progress']
```

The below tables were generated via the above prompts

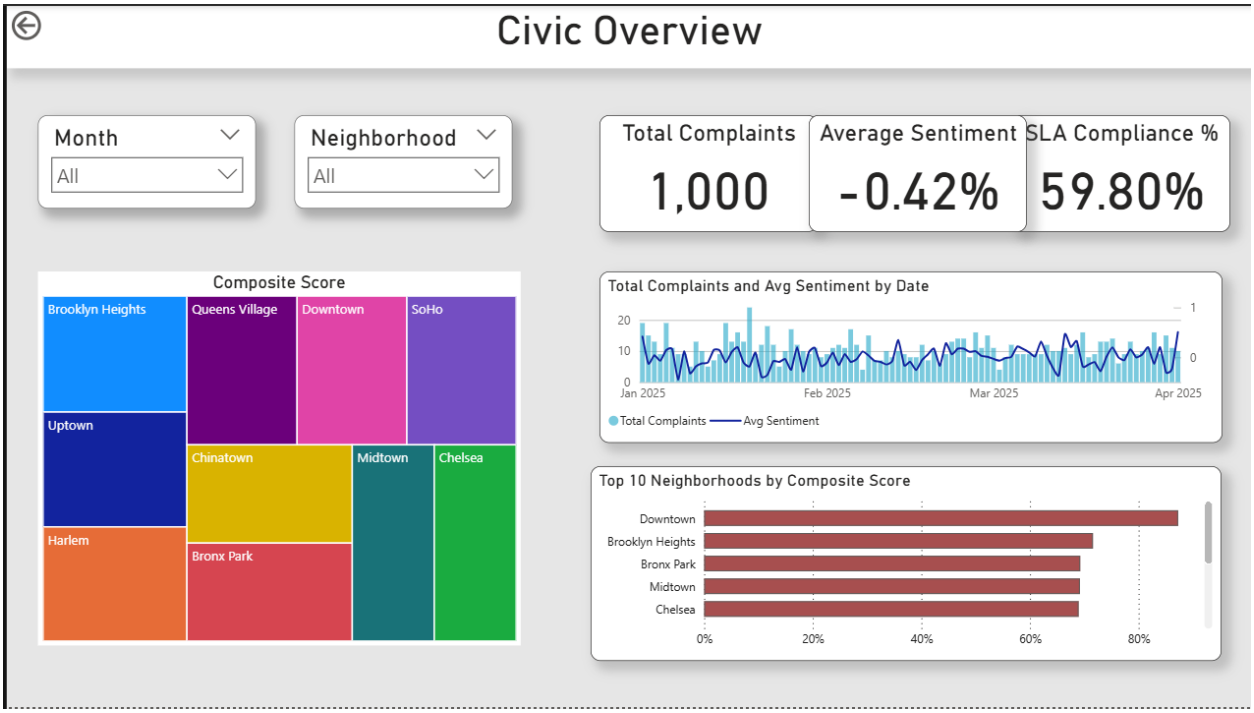
1. **311_ServiceRequests**:
    - This dataset contained all non-emergency service requests from the city, with fields such as Request_ID, Complaint_Type, Latitude, Longitude, Priority, Status, etc.
2. **SentimentData**:
    - Public sentiment data derived from social media scraping, containing fields like Sentiment_Score, Post_ID, Text, and Topic. This allowed us to evaluate the public mood surrounding various service requests.
3. **PublicTransport**:
    - Data about public transit ridership, delays, and modes of transport, including fields like Delay_Minutes, Mode, and Ridership.
4. **UrbanAnalysis**:
    - A summary of urban metrics, aggregating data on complaints, sentiments, and average statistics across neighborhoods.
5. **Date**:
    - A dedicated date table used for time-based slicers and filters, ensuring all data is tied to a consistent time dimension.

Project Breakdown by Weeks

Week 1 – Data Integration & Geospatial Modeling

Tasks Completed:

- **Data Preparation**: Cleaned and aggregated the 311 service request data, ensuring it was ready for analysis.
- **Geospatial Data Integration**: Integrated neighborhood shapefiles into Power BI for visualizing spatial data. This allowed us to plot service requests as heatmaps on the map.
- **Mockup Dashboard Creation**: Developed an initial mockup of the dashboard for city planners, focusing on visualizing complaints and service request hotspots.



Key Insights:

- Visualized complaints geographically using a heatmap.
- Identified initial areas with high concentrations of service requests.

Week 2 – Service Request Dashboard & Sentiment Analysis

Tasks Completed:

- **Sentiment Analysis Pipeline**: Set up data collection for public sentiment from social media, using **Azure Cognitive Services** to generate sentiment scores.
- **Service Request Dashboard**: Built a dashboard that tracks service request volume and resolution times by category (e.g., potholes, broken sidewalks, etc.).
- **Sentiment Overlay**: Overlaid sentiment scores on the service request heatmap to identify how public sentiment correlates with service complaints.

Key Insights:

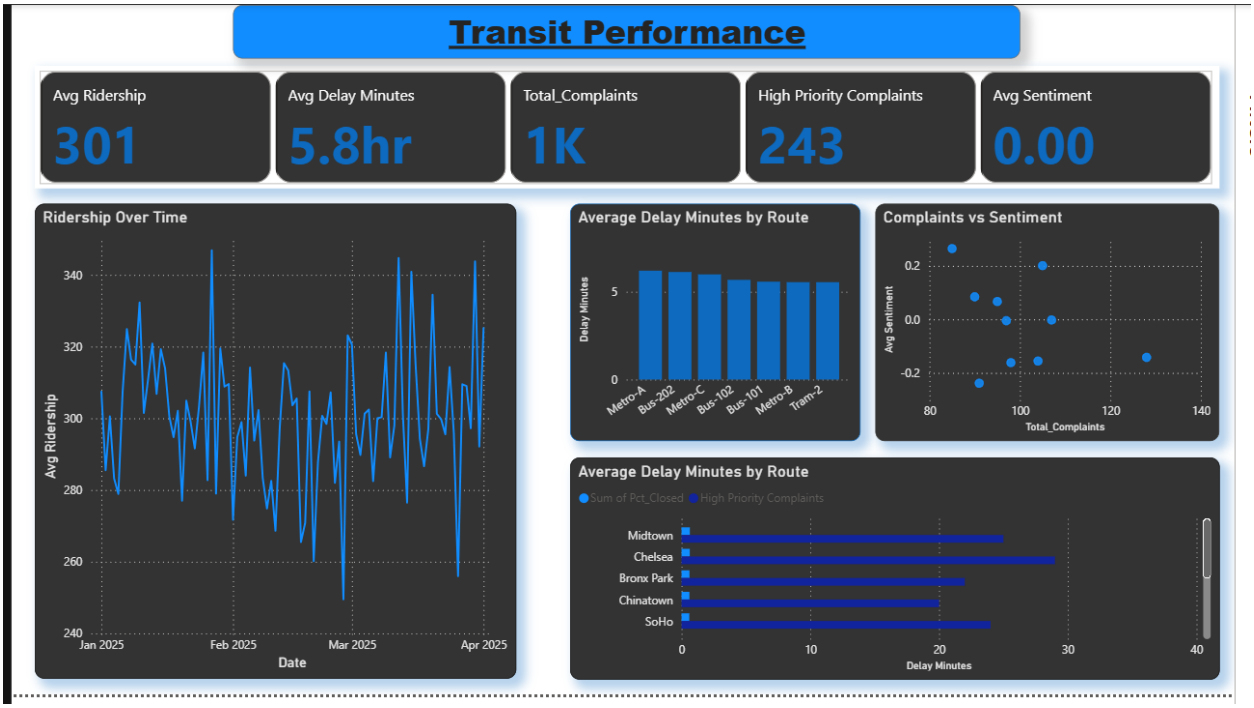
- Some neighborhoods showed significantly negative sentiment, which correlated with certain complaint types like potholes and broken infrastructure.
- The dashboard allowed planners to correlate service resolution times with sentiment data, providing insights into areas where improvements were most needed.

Week 3 – Transit Performance & Sentiment Correlation

Tasks Completed:

- **Transit Performance Dashboard**: Developed a dashboard to track public transport ridership, delays, and punctuality. This visualized key metrics such as **average ridership** and **average delay minutes**.
- **Sentiment vs. Ridership Correlation**: Correlated **service request volume** and **negative sentiment** with **transit performance**, focusing on neighborhoods where transit delays were causing public dissatisfaction.

- **Key Influencers Analysis**: Used Power BI's Key Influencers visual to identify the factors driving negative sentiment.



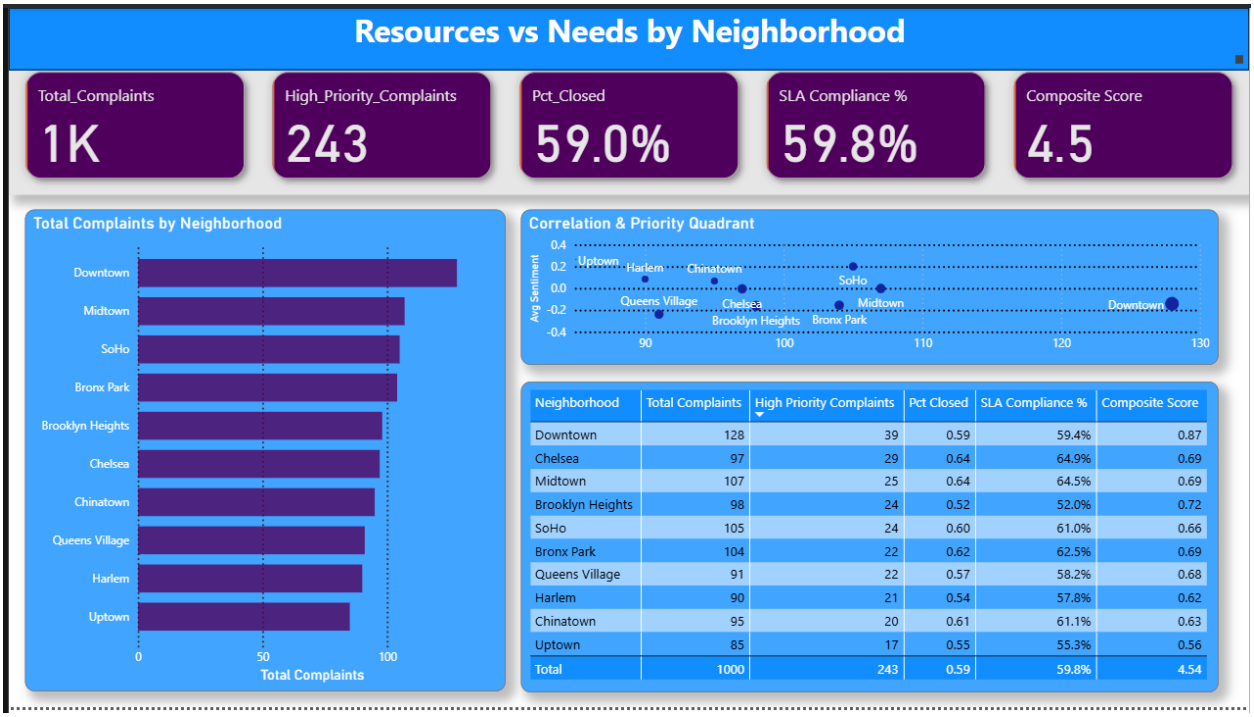
Key Insights:

- Negative sentiment was strongly linked to delays in public transport.
- There were clear patterns showing that neighborhoods with high complaints (especially about transit) also had lower public sentiment scores.

Week 4 – Resource Allocation & Needs Analysis

Tasks Completed:

- **Resource Allocation Dashboard**: Built a dashboard comparing available resources (e.g., public transport vehicles, maintenance teams) against complaint volume in each neighborhood.
- **Resource vs. Need Analysis**: Analyzed which neighborhoods were over-resourced or under-resourced. The dashboard prioritized areas with high complaint volumes but low resource allocation.
- **Heatmap of Resource Gaps**: Created a map showing resource gaps in neighborhoods, helping city planners prioritize areas for intervention.



Neighborhood	Total Complaints	High Priority Complaints	Pct Closed	SLA Compliance %	Composite Score
Downtown	128	39	0.59	59.4%	0.87
Chelsea	97	29	0.64	64.9%	0.69
Midtown	107	25	0.64	64.5%	0.69
Brooklyn Heights	98	24	0.52	52.0%	0.72
SoHo	105	24	0.60	61.0%	0.66
Bronx Park	104	22	0.62	62.5%	0.69
Queens Village	91	22	0.57	58.2%	0.68
Harlem	90	21	0.54	57.8%	0.62
Chinatown	95	20	0.61	61.1%	0.63
Uptown	85	17	0.55	55.3%	0.56
Total	1000	243	0.59	59.8%	4.54

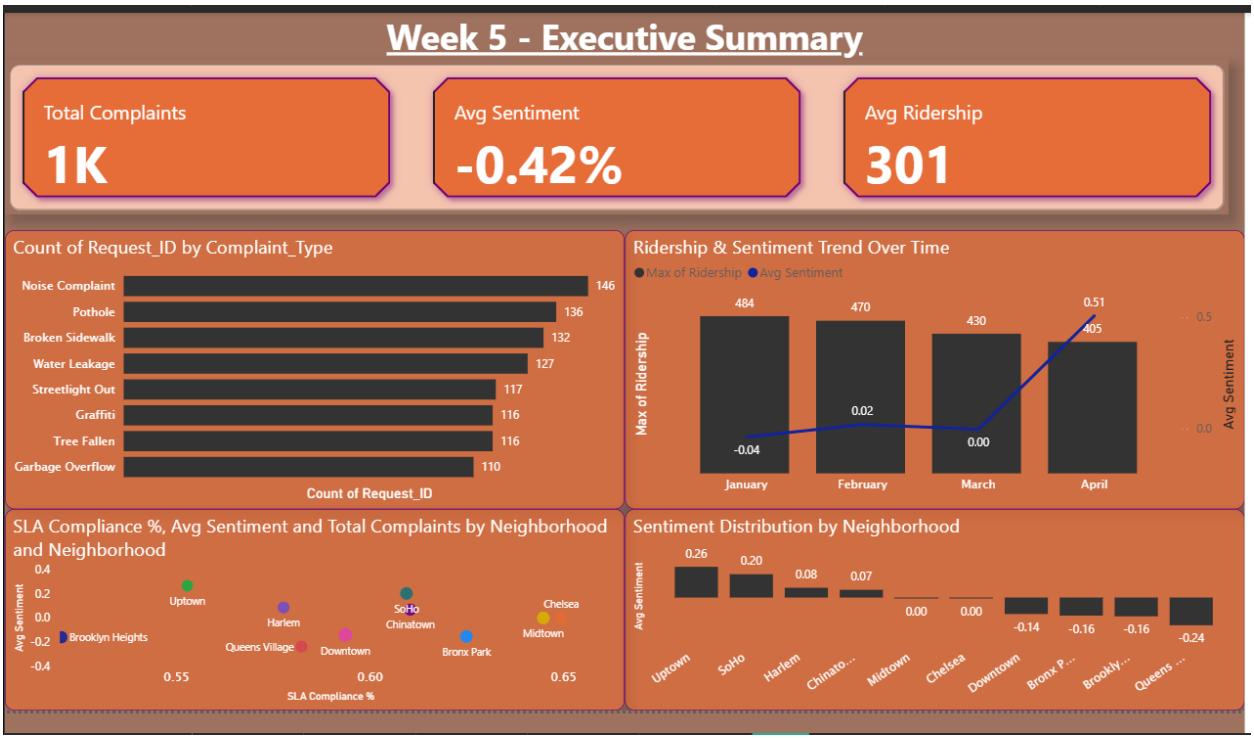
Key Insights:

- Neighborhoods like **Downtown** and **Midtown** had high complaint volumes but lacked sufficient resources to address these issues.
- The resource allocation dashboard provided planners with clear, actionable insights on where to allocate additional resources.

Week 5 – Executive Summary Dashboard

Tasks Completed:

- **KPI Summary Dashboard**: The final executive summary dashboard aggregated key metrics and insights across the first four weeks of the project. It was designed to provide policymakers with a high-level overview of the most important data, including:
  - **Total Complaints**
  - **Avg Sentiment**
  - **Avg Ridership**
- **Key Visuals**: Added visuals that highlighted complaint distribution by type, sentiment trends over time, and resources vs. needs by neighborhood.
- **Final Report**: A one-page, printable summary of the entire project, including insights and recommendations.



Key Insights:

- **Complaints Distribution**: A large portion of complaints came from **noise** and **potholes**.
- **Ridership and Sentiment**: Transit ridership increased in certain months, while sentiment improved in others, reflecting improved satisfaction after delays were addressed.
- **Resource Allocation Gaps**: Areas like **Uptown** and **SoHo** were underserved in terms of resources despite higher complaint volumes.
- The **Executive Summary** dashboard provided an **easily digestible format** for stakeholders, enabling quick decision-making.

Conclusion & Recommendations

1. **Data-Driven Decision-Making**: The integration of 311 service request data, public transport data, and sentiment analysis allows municipal planners to move from reactive to proactive management, making evidence-based decisions about where resources should be allocated.
2. **Hotspot Identification**: Key insights showed that neighborhoods like **Downtown** and **Midtown** have high complaint volumes and negative sentiment. These areas should be prioritized for immediate intervention.
3. **Transit & Complaints Correlation**: There is a clear correlation between public transport delays and negative sentiment. Improving transit reliability could enhance public satisfaction.

4. **Resource Allocation Gaps**: Some neighborhoods were underserved in terms of resources, even though they had high complaint volumes. These neighborhoods should be targeted for additional resources.

Final Recommendations for City Planners:

- Focus on **high complaint/negative sentiment areas** (Downtown, Midtown) by improving service resolution times and allocating more resources.
- **Enhance public transit reliability** to improve public sentiment and reduce complaints.
- Ensure that resource allocation matches neighborhood needs, especially in areas with high complaints but low resources.