

**Project Title:****CityPulse: Urban Sentiment & Mobility Dashboard****Goal:**

Build an integrated dashboard that visualizes public sentiment and mobility patterns in Chicago, combining social media sentiment with official city transportation and 311 service data to reveal civic trends and commuter experiences.

**Core Tools:**

- **Languages:** Python (pandas, numpy, matplotlib)
  - **Libraries:** snsrape, TextBlob, VADER, requests
  - **Visualization:** Tableau or Power BI
  - **Data Sources:**
    - Chicago 311 Service Requests API
    - Chicago Transit Authority (CTA) Ridership Data ([data.cityofchicago.org](https://data.cityofchicago.org))
    - Twitter/X (via snsrape for hashtags like #ChicagoCommute, #CTA, #RushHour)
- 

**4-Week Timeline****Week 1: Data Collection & Cleaning**

**Objective:** Assemble and clean data from all sources.

**Tasks:**

1. Use the **Chicago 311 API** to extract complaints related to transit (e.g., “street light,” “pothole,” “transit delay”).
2. Download **CTA ridership datasets** (bus and train) for recent months.
3. Scrape ~1,000–2,000 tweets from Chicago-related hashtags using **snsrape**.
4. Clean and preprocess all data:
  - Handle missing values and duplicates.
  - Normalize timestamps and locations (convert to datetime, latitude/longitude).
  - Save clean datasets to CSV for repeatable runs.

**Deliverables:**

- Cleaned CSVs (311, CTA, tweets).
  - Data dictionary describing variables.
- 

**Week 2: Sentiment Analysis & Integration**

**Objective:** Extract sentiment and merge datasets for comparison.

**Tasks:**

1. Use **VADER or TextBlob** to compute sentiment scores (polarity and subjectivity) for each tweet.
2. Categorize sentiment into three bins: *positive*, *neutral*, *negative*.
3. Aggregate tweets by **day and sentiment category**.
4. Merge the sentiment summary with **CTA ridership** and **311 complaint trends** by date.
5. Validate that timestamps align across datasets (daily granularity).

**Deliverables:**

- Python notebook with sentiment scoring workflow.
  - Combined dataset (sentiment + mobility metrics).
- 

**Week 3: Correlation & Visualization**

**Objective:** Identify relationships between sentiment, service data, and mobility metrics.

**Tasks:**

1. Calculate correlations between:
  - Sentiment polarity and CTA ridership.
  - 311 complaints and tweet negativity.
2. Generate exploratory visualizations:
  - Line chart: sentiment trends vs. daily ridership.

- Heatmap: neighborhood-level complaint density vs. negative sentiment.
- 3. Build prototype visuals in Tableau or Power BI:
  - Time series overview.
  - Interactive filters (date, neighborhood, complaint type).

**Deliverables:**

- Set of analytical plots and correlation tables.
  - Preliminary Tableau/Power BI dashboard draft.
- 

## **Week 4: Dashboard Finalization & Storytelling**

**Objective:** Build a polished dashboard and synthesize key findings.

**Tasks:**

1. Finalize the Tableau/Power BI dashboard:
  - Add map layers for sentiment and complaints.
  - Include daily trend charts and top hashtags.
  - Create tooltips for hover insights.
2. Write a **1-page data story** summarizing:
  - Trends observed (e.g., “Sentiment drops during Monday rush hours”).
  - Insights for civic planners or CTA operations.
  - Future extensions (real-time updates, anomaly alerts).
3. Prepare a short presentation or GitHub README outlining:
  - Objectives, data, workflow, visuals, and learnings.

**Deliverables:**

- Completed Tableau/Power BI dashboard.
  - 1-page insight brief or GitHub README.
- 

## **Expected Outcomes**

- A functional and interactive **CityPulse dashboard** showing how Chicagoans’ moods correlate with daily mobility challenges.
- Demonstration of applied NLP, API integration, data wrangling, and visualization—aligned with Mukul’s machine learning and Tableau expertise

Mukul Ved CV

- A strong portfolio piece combining urban data science and public sentiment analytics.