Project Title and Summary

Streets & Safety: A Neighborhood-Level View of Pavement Condition and Fire Hydrant Coverage in Syracuse

This project will produce an interactive dashboard and analytical report that combine Syracuse's pavement condition ratings and fire hydrant locations to provide a clear, neighborhood-level view of basic infrastructure quality and emergency readiness. By aggregating two underutilized but high-value datasets, the project helps residents, community organizations, and city officials understand where street conditions are deteriorating, where fire protection assets are concentrated or sparse, and where mismatches between infrastructure quality and safety coverage may exist.

Problem Statement

Syracuse residents regularly experience the consequences of aging infrastructure: rough streets, delayed maintenance, and concerns about emergency response readiness. While the City of Syracuse publishes detailed open data on pavement conditions and fire hydrant locations, these datasets exist in isolation and are difficult for non-technical users to interpret or connect meaningfully.

Currently, there is no public, easy-to-use tool that allows stakeholders to view street quality and fire safety infrastructure together at a neighborhood scale. This limits the ability of residents to understand conditions in their communities, community organizations to advocate for targeted improvements, and city officials to communicate infrastructure priorities transparently.

The absence of a combined, neighborhood-level view leads to fragmented understanding and anecdote-driven discussions rather than data-informed decision-making. This project addresses that gap by synthesizing multiple datasets into a single, accessible analytical product.

Data Sources

Primary Syracuse Open Data datasets:

- **Syracuse Pavement Ratings (2022 & 2023)**
    - Street segment–level condition ratings
    - Strengths: structured, moderate size, directly tied to infrastructure investment

    - Limitations: no neighborhood field; represents inspection snapshots, not continuous conditions

- **Syracuse Fire Hydrants**
    - Point locations of hydrants maintained by the city
    - Strengths: clear geometry, public safety relevance
    - Limitations: no neighborhood field; hydrant presence does not directly measure response capacity

Supporting dataset:

- **Syracuse Neighborhood Boundaries**
    - Used to assign both pavement segments and hydrants to a shared aggregation level

Key limitations acknowledged upfront:

- No direct join key exists between datasets
- Hydrant density is a proxy, not a guarantee, of fire response effectiveness
- Pavement condition does not capture all street usability factors (e.g., snow removal, traffic volume)

Technical Approach

The project will use a **SQL + Python (pandas) → Power BI** pipeline designed to remain transparent, reproducible, and accessible to city staff.

Data preparation and analysis:

1. Load pavement ratings and fire hydrant datasets into Python and/or a SQL database.
2. Clean and validate fields (remove null geometries, standardize condition scores, flag outliers).
3. Perform exploratory data analysis:
    - Distribution of pavement condition scores
    - Spatial concentration of hydrants
    - Year-to-year changes in pavement quality

4. Use Power BI's built-in spatial capabilities to perform a **one-time spatial assignment** of both datasets to Syracuse neighborhood boundaries.
5. Aggregate metrics at the neighborhood level:
    - Average pavement condition score
    - Percentage of street segments rated "poor"
    - Hydrant count per neighborhood
    - Hydrant density normalized by street mileage (where feasible)

LLM augmentation (limited and controlled):

- Generate plain-language summaries of neighborhood-level patterns (e.g., "Pavement conditions declined while hydrant coverage remained stable").
- Assist with documentation and data dictionary explanations.
- All LLM outputs will be grounded in computed metrics and manually reviewed.

Validation strategies:

- Cross-check aggregates against raw counts
- Ensure consistent neighborhood assignment across datasets
- Explicitly label proxies and assumptions in both visuals and documentation

---

Deliverable Description

The final deliverables will include:

1. **Interactive Power BI Dashboard**
    - Neighborhood-level map showing pavement condition and hydrant density
    - Comparative bar charts highlighting infrastructure disparities
    - Trend visuals for pavement ratings across years
    - Filters for neighborhood and year
2. **Analytical Report (PDF or web-formatted)**
    - Methodology and data limitations
    - Key findings and patterns
    - Guidance on responsible interpretation

3. **Technical Documentation**
    - Data pipeline description
    - Field definitions
    - Reproducibility notes for future updates

Success Criteria

- Stakeholders can identify neighborhoods with poor pavement and low hydrant coverage within two minutes of using the dashboard.
- Aggregated metrics are reproducible and match raw data checks.
- The project clearly documents limitations and avoids causal claims.
- The dashboard can be updated with new pavement rating releases with minimal rework.
- City staff or community users can articulate at least one practical use case (planning, advocacy, communication).

Timeline

Week	Activities
1	Dataset acquisition, schema review, initial EDA
2	Data cleaning, normalization, exploratory analysis
3	Neighborhood spatial assignment in Power BI
4	Aggregation and SQL modeling
5	Dashboard design and implementation
6	Report writing, validation, documentation, final presentation

---

Risks and Mitigations

- **Risk:** Misinterpretation of hydrant density as fire response quality
  Mitigation: Clear labeling and disclaimers; no causal language
- **Risk:** Spatial assignment errors
  Mitigation: Spot-check neighborhood assignments; document method

- **Risk:** Over-simplified rankings stigmatize neighborhoods
  Mitigation: Default views focus on patterns and comparisons, not rankings
- **Risk:** Dataset updates change schemas
  Mitigation: Modular pipeline and documented assumptions