Final Project Idea Assignment

- Project Title: Optimizing Urban Commute Efficiency Through Network Analysis of Dallas Traffic Flows

- Team Members:

    - Rahul Juluru – 113686572 – Rahul.juluru-1@ou.edu
    - Mason Cacurak – 113561612 – mason.cacurak@ou.edu
    - Saketh Reddy Aredla – 113682526 – saketh.reddy.aredla-1@ou.edu

- Motivation and Problem Statement:

    - Effective traffic control is critical in cutting down congestion, time on the road, and pollution in city centrality. In this project we will analyze a real-world traffic dataset that is comprised of floating car data across thousands of road segments in Dallas, TX.
    - We will model this dataset as a directed weighted transportation network where each road segment is an edge between intersections. In looking at Dallas's traffic as a complex network, we will aim to identify key road segments and congestion communities with the goal of optimizing travel times.
    - Note: As of now, we have not been able to find a sufficient dataset containing temporal data allowing us to center the project around optimizing commute duration based on the time of day.
    - This work examines the questions: "How does traffic congestion spread through a network of highways". "Which intersections are most critical for maintaining smooth traffic flow". "Which road segments act as high centrality connectors that most effectively reduce overall commute times"? "Can we identify the main congestion bottlenecks in the traffic network based on the flow and speed data"?

- Data Description:
    - Data Source: A unified and validated traffic dataset for 20 U.S. cities
    - Type of Data: Public floating car dataset with average speeds, estimated traffic flow, and geo coordinates for road segments in Dallas, TX
    - Type of Network: Directed and weighted transportation network built from road segment connections
    - Edges: Road segments linking the intersection nodes and weighted by average traffic speed or flow
    - Approximate Size: Several thousand nodes and edges comprised of primary roads in Dallas

- Network Science Concepts and Methods:

- Centrality Analysis: Compute degree, betweenness, and eigenvector centrality to identify high-traffic or influential road segments.
- Community Detection: Use Louvain and Girvan–Newman algorithms to detect communities of road segments or time periods showing similar congestion patterns.
- Network Robustness / Resilience: Simulate random and targeted removal of nodes (e.g., closure of key segments) to measure the network's tolerance to disruptions.
- Diffusion Modeling: Represent congestion spread using SIR-style contagion models or threshold diffusion to simulate how traffic jams propagate spatially and temporally.
- Temporal Analysis: Build evolving snapshots of the network over hours/days and analyze changing connectivity metrics.
- Libraries: Networkx, Numpy, Matplotlib,Pandas, Community-louvain

- Expected Outcomes: What will you be learning, measuring, or showing?

We hope to discover how traffic congestion propagates and spreads across the Dallas road network by analyzing real floating car dataattempt to quantify structural properties and resilience of the Dallas, TX urban road network via quantitative indicators and show how network science analysis can provide insightful results useful for decision-making on the part of urban transportation systems.

1. Quantitative Findings

Calculate centrality rankings (degree, betweenness, eigenvector) to recognize the highest influencing or high-impact road segments that are liable for sustaining network flow.

Identify community structures by Louvain and Girvan–Newman algorithms to discover groups of road segments or time intervals that have common congestion patterns.

Model diffusion of congestion in contagion-based models to estimate traffic congestion spread among interrelated nodes.

Quantify how improving flow on high impact segments changes average shortest path lengths and overall network efficiency

2. Qualitative Findings

Find key chokepoints that seriously impact traffic flow.

Extract policy-level inferences on traffic control—e.g., special observation or rerouting

Detect behavioral commuting habits like peak periods and weekend drops.

Provide recommendations for smart-city traffic optimization and infrastructure planning.

Identify reoccurring congestion patterns across time to help predict and mitigate delays.

- Deliverables:

  - Written Report: Explaining dataset transformation, network construction, methods, and results.
  - Visualizations: Network maps, Community-Louvain plots
  - Code Repository: Well-documented Python notebook using NetworkX and visualization tools.
  - Presentation Slides: Summarizing findings, visual graphs, and interpretations.

- References:

  - Xu, Xiaotong; Zheng, Zhenjie; Hu, Zijian; Feng, Kairui; Ma, Wei (2023). A unified and validated traffic dataset for 20 U.S. cities. figshare. Dataset. https://doi.org/10.6084/m9.figshare.24235696
  - NetworkX Developers. (2024). NetworkX – Network Analysis in Python. Version 3.x. Python Software Foundation. https://networkx.org
  - NumPy Developers. (2024). NumPy: Fundamental package for scientific computing with Python. Version 1.x. https://numpy.org
  - Pandas Development Team. (2024). Pandas – Python Data Analysis Library. Version 2.x. https://pandas.pydata.org
  - Matplotlib Development Team. (2024). Matplotlib: Visualization with Python. Version 3.x. https://matplotlib.org