**Critical Nodes and Community Detection in California's Road System**

**Team members**

Kun Peng, Zules Zou, Yan zhang

**Backgroud (motivation)**

The California road network plays a vital role in supporting daily life and economic activities across the state. To improve traffic management and network resilience, it's essential to understand its key components and vulnerabilities. This research aims to identify critical traffic hubs and distinct functional zones within the network. By doing so, we can gain insights into the system's strengths and weaknesses, leading to practical solutions for optimizing traffic flow and enhancing overall infrastructure resilience.

**Research Questions**

1. Which nodes within the California road network are the most critical in terms of maintaining connectivity and preventing network fragmentation?
2. How can community detection methods be applied to identify functional zones within the network, such as residential, commercial, or industrial areas?
3. How does the failure of key nodes impact the network's overall connectivity and the integration of different functional zones?
4. What are the potential traffic management and infrastructure improvement suggestions to enhance network resilience and connectivity？

**Data materials**

* **California Road Network Dataset:** This dataset contains information on intersections and endpoints represented as nodes, and roads connecting these points represented as undirected edges.
* **Source:** J. Leskovec, K. Lang, A. Dasgupta, M. Mahoney. [Community Structure in Large Networks: Natural Cluster Sizes and the Absence of Large Well-Defined Clusters](http://arxiv.org/abs/0810.1355). Internet Mathematics 6(1) 29--123, 2009.

**Planned network metrics and tools**

* **Metrics:** Degree centrality and betweenness centrality, Clustering Coefficient,
* **Community Detection Algorithm:** The Louvain algorithm
* **Gephi or NetworkX** for visualization and computation of network metrics, **Python** for data preprocessing and analysis

**Timelines**

* Week 1-2: Preprocess the dataset and calculate network centrality metrics.
* Week 3-4: Apply community detection algorithms and analyze results.
* Week 5-6: Simulate node failures and measure the impact on network connectivity.
* Week 7-8: Analyze final results, generate visualizations, and prepare the final report/presentation.