Final Project Goals

For our final project, our group has decided to use the Pennsylvania Road Map dataset, found on the Stanford Dataset Collection website. This data set is based off of the road map network within the state of Pennsylvania, and is in the following format:

Nodes: *Number of Nodes*
Edges: *Number of Edges*
FromNodeId: *ID of Incoming Node*
ToNodeId: *ID of Outgoing Node*

As this data set is a .txt file, we plan to parse each entry to convert the data to a usable format. In this dataset, there are around 1088092 nodes and 3083796 edges, where a node is an intersection of roads or endpoints and the edges are the roads that connect the intersection/endpoints.

The traversal we plan to use is BFS. Using breadth first search on the dataset would allow us to search each path from the starting node and branch out from each connecting path. Additionally, we want to use the shortest path algorithm (specifically Dijkstra's Algorithm) to determine the shortest distance between two nodes. In a broader sense, the shortest distance between two nodes would correlate to the shortest path that one would take for driving in the state. For the uncovered algorithm, we plan to use A* search to determine the shortest distance. The A* search is an extension of Dijkstra's project that has a lower cost, thus it will be interesting to see how the two algorithms compare. The traversals and algorithms listed will allow us to fully utilize the dataset and implement a graph that provides further insight into the data.

Our ultimate goal is to create an executable that allows the user to enter starting and ending node indices and get the shortest distance between the two nodes. Additionally, the user can specific the algorithm they want to use:

./executable_name [start node index] [end node index] --[algorithm name]

EXAMPLE: ./final 3 39 --dijkstras