# MECH 6317.001 Supplemental Project Document – Tanner Kogel

The following is a supplemental document to aid in understanding the presentation.

## Background

Amtrak is a national Passenger Railroad Corporation with over 21,000 miles of train routes in 46 U.S. states, 3 Canadian provinces, and the District of Colombia that operates over 300 trains every day with more than 500 destinations.

## Amtrak Rail Network

A node is defined as a rail station that Amtrak operates from and the edge between nodes represents the number of Amtrak routes that operate between those two stations. This network was found to have 530 nodes and 592 edges.

## Non-Induced Populous City Network

A node is defined as a city with a population above 100,000 according to the 2020 Census and the edge between nodes represents the number of Amtrak routes that travel between those two cities without stopping at another city defined in this graph. This network was found to have 331 nodes and163 edges.

## Induced Populous City Network

A node is defined as a city with a population above 100,000 according to the 2020 Census and the edge between nodes represents the number of Amtrak routes that travel directly between those two cities without stopping in between. This network was found to have 331 nodes and 74 edges.

## Graph Analysis

Multiple analyses were performed on the three graphs which showed that there was much correlation between the original Amtrak rail network and the non-induced city network, but not much correlation to the induced city network, which lost too many edges to accurately portray the original. Modularity and assortativity was found for the city networks only because the attributes that were used was only found for the defined cities, and despite the non-induced graph showing a strong correlation to the original, and the induced graph not showing much correlation, both

graphs returned similar answers of modularity and assortativity.

Community detection was also run on this network, but no relevant results were found, thus it was not included in the presentation. The files of communities are included in the .zip file, but the main focus of the results is that a very large number of communities are found with small numbers of nodes that are all very geographically close and are often neighboring stops on the different routes. This result was expected and did not seem to add any important information to the presentation and was thusly dropped.

**Conclusion**

The conclusions that can be drawn from the graph analysis is that Amtrak rail is not an efficient or effective means of traveling in North America. The graph is not fully connected, therefore in order to reach specific destination it is likely that the passenger would have to also take buses, planes, or cars to get to another station, making Amtrak unable to travel from any major city to another independently. Additionally, the diameter of the graph is very large, meaning there are a large amount of stops that the passenger will have to go through when taking a long journey on Amtrak, therefore making the trip much longer than alternatives. The large diameter could speak for a large availability of destinations allowing the accessibility of rail travel to be more widespread, but it is seen that of the 331 cities in the U.S. with populations above 100,000, 203 of them are represented as a node with a degree of 0, meaning that only 128 of the 331 most populous cities in the U.S. even have active Amtrak rail station (less than 40%). Therefore, Amtrak almost exclusively markets itself as a journey-oriented experience to view the beautiful North American continent, and not as a daily commuting option or an alternative to air travel.

The induced and non-induced sampling of the graph into cities yielded results that tended to show that the induced sample did not reflect the original graph almost at all, as most of the graphical features were missing from almost every single analysis method performed, while the non-induced sampling yielded results that seemed close to allude to the original in a way that the induced graph could not. Despite this, both graphs yielded very similar results in terms of modularity and assortativity, which raises the question of whether or not the results of modularity and assortativity were representative of the entire graph since it was unable to be computed.