**MTA Project Proposal**

The Metropolitan Transportation Authority(MTA) is North America's largest transportation network, serving a population of 15.3 million people across a 5,000-square-mile travel area surrounding New York City through Long Island, southeastern New York State, and Connecticut. Today most of the big company use data analysis techniques to help the company make the right decision. I’m interested to use MTA data to do the analysis and clean the data using different methodology. I want to present accurate and reliable data. As far as possible, avoid statistical errors, and find a way to deal with everyday challenges like outliers, missing data, data altering and developing graphical representation.

MTA shows there is many shutdowns, using the data set will help find the problem and try to solve it. As I mentioned, I would like to use these data set to optimize the important of the business and understand problems facing MTA , and to explore data in meaningful ways. In this project I will use Python and different tools such as, Pandas, Matplotlip and seaborn. I will use these tools to find the analyze and clean the data.