AIT614 Team 1 Project Topic

Cross Zeigler, Suchada Hapikul, Ewin Hong, Cheng Yuan Lee, Sagar Deepakgiri Goswami

2/27/2022

* Free to pick your own topic but it needs to get approved by the instructor.
* Turn in one or two paragraph description(s) of your proposed topic before you start to work on your proposal.

What we may want to include for the project topic paper:

* Description of the project
* Our hypothesis/question
* Tools/techniques we may use
* Domain of the project

Chosen Topic: NYC traffic dataset for traffic prediction (Multiple Linear Regression Models, K-means Clustering, Geo-Spatial Analysis (if possible))

Finding insights to improve the NYC traffic system

Choke points?

Potential points/regions to improve upon

Residents can use it to know when there will be high volumes of traffic so they know when to leave for optimal commute times

Emergency services can use it to know what roads to travel to improve their efficiency

Stakeholders in transportation sector could use outcomes to provide better services, such as taxis, to increase profits

Assist Department of Transportation in framing better policies and provide them with insights for better infrastructure planning.

Link to dataset: <https://data.cityofnewyork.us/Transportation/DOT-Traffic-Speeds-NBE/i4gi-tjb9>

Use R to clean up some null value.

Excel power query?

**Description**

New York City (NYC) suffers heavily from traffic jams and congestion. Knowing about them in advance with accurate predictions can help a lot in planning trips ahead of time, thereby saving on time otherwise lost in traffic during the commute. Our project will utilize datasets from NYC open data and focus on traffic in the city to discover patterns of transportation in order to resolve congestion amongst other traffic issues within NYC. The primary goal of the project is to provide valuable insights into the city’s traffic system to improve transportation within NYC for increased satisfaction, safety, and decreased travel times. Some potential focus areas will be on certain points and regions within NYC, specifically those dealing with high volumes of traffic at specific times of day, and giving valuable advice to certain audiences to improve their relationship with traffic in the area. Our audiences include residents to give them optimal schedules for commutes to the Department of Transportation to assist them in framing better policies for improved infrastructure planning.

**Domain**

Our

**Questions and Tools**

Our main question is “How can we improve traffic within NYC”, followed by “To what degree can we improve traffic within New York City”. This gives us a clear focus that avoids being too simple, as we have two goals that can’t be answered by just looking up the question and that require nuance to understand completely. We can focus on certain areas that will likely have an impact on other areas as traffic improves in one place, giving us another point of consideration in the project, i.e. how do certain parts of NYC affect each other in terms of congestion and traffic overall.

Other smaller questions include “What causes such bad traffic within NYC to begin with” and “How do people currently navigate traffic”. These secondary questions help to prop up the primary ones by getting to the underlying issues of traffic within NYC and asking how people optimize their travel personally. This can help us understand how residents see themselves within the system and how they seek to improve their commute, and if there are any other steps we can take to provide optimal personal planning and improve infrastructure based on residents’ responses.

We’ll utilize several tools to complete the project. One of them will be Python or R in order to clean the data amongst any analyses we may run Another tool may be Power Query to explore the data and provide certain insights into traffic within NYC. Currently we’re debating on what NoSQL database and Big Data Engine to run the project on, but we’re considering the ones given in class and will make a decision upon further research.