Personal Assessment

The database design fulfills the requirement of the project because it displays traffic data for the interstates. I constricted the amount of data to be displayed to only high-volume areas which will give regulators the best idea of where toll gantries (Toll gantry – automated tolling by license plate or electronic pass, limiting the chances of serious accidents, eliminating the need for hiring toll workers, and maximizing public safety.) should be placed. I however was not able to find enough data on roads such as Route 2, Route 11, and Route 8 which are routes according to my research that Connecticut would toll if tolling was reintroduced(https://www.nbcconnecticut.com/news/local/CTDOT-Releases-Preliminary-Report-on-Tolls-500715881.html). In addition, I could not find enough data to suggest that tolling those roads would help at all.

From the reports that I read Connecticut's plan in addressing congestion mostly centers around performing much needed repairs to I-84 especially to the viaduct in Hartford. An idea of tolling I-84 to help control congestion and pay for improvements is an idea. If Connecticut's goal to control congestion is with tolling without enough data to support is unwise.

I had great difficulty in finding new data. Most data for state roads came from the Department of Transportation; https://www.ct.gov/dot/cwp/view.asp?a=3532&q=330402 . One difficulty was interpreting the data. In many towns they placed multiple stations to collect traffic data along the same route. The numbers differently greatly and in did not appear to be any meaningful order. I did the best I could to select a data point that appeared to be as close to the average as possible. For highways since there were fewer data points I averaged them together. The data for highways can be found on data.gov.

There was ample information on the majority of Connecticut State Routes that I added to my data base except for Routes 2,8, 9, and 11, roads that I was keen on finding data on. All data I used was collected from the CT DOT (Department of Transportation). The CT DOT is responsible for maintaining the roads and collects all traffic data. They collect the data with traffic monitors (machines that have little tubes across the road and count the number of cars that run them over). From this ADT (annual daily traffic) is calculated. The average daily traffic references the average amount of traffic a road sees in both directions per day for the year that that data was taken.

Interstate data can be found on Data.gov which is where I got data on all the Interstates that run through CT. Statistics were collected in multiple locations in a town, thus I averaged all the numbers. I did find some errors in the data, for example in the data they had I-384 going through Madison, I know for a fact that is not true, since it is a road I drive on every day. I ended up having to reference the Interstates on Wikipedia which interestingly was more accurate in what towns the highways went through then the data I found from the government. I partly suspect that this is because the data is older, dated in 2011; in addition, the tax codes could have been different then. But that is unlikely since the tax codes are assigned to towns in alphabetic order starting from 001 through 169.

Lastly while writing the data base I chose to use varchars for every data point. Which for route numbers is acceptable, although for the number of cars a road sees on average daily throughout the year was not the greatest idea. While writing the queries I realized it would have been better if I had written them as integers as all data was rounded up to the nearest whole number. However I did not have enough time

to rewrite the database to use integers for the NumCar value, I cast varchar into an integer value that I used in the tables AverageDailyTraffic and AnnualAverageDailyTraffic.

In the data I found on data.gov there is more thorough numbers on Connecticut routes then what I found on the CT DOT's website. If I had more time to work on the assignment, I would add the additional information to the data I found the DOT's website. I also wanted to make use of the different exits on the highways, as there was data on some of those as well that could help to narrow down toll placement even further. I also would have liked to compile accident data because law makers in Connecticut are concerned about the risk of accidents involving tolls. The overlaying cause of toll removals was due to the deadly accident that occurred on Jan 20th, 1983 involving failure with a tractor trailer trucks' breaks leading to an explosion of several vehicles.

I would have created a separate table for accidents that contained an identification for the accident, which would have been used as the primary key. Then I would have included in the table, the tax code of the town where the accident took place, the route name, and who was involved, and if there were any fatalities. Knowing the location and frequency of accidents may also help to plan toll placement but also could be used in future planning for highway development, especially for I-84 where assessment for upgrades is already underway.

Also, in the Interstates table I included the number of exits on each highway in Connecticut. While that value is important to the Interstate, I never used it in the database. Looking back if I had more time to develop the database, I would have broken out the exits into their own table, maybe even included which mile marker they were at. This would help during the construction of tolls to prevent motorists from avoiding tolls.

My ER diagram is in the third normal form, because with Highways all items are only related to the roads. All the tables have items in them that are always related to the table. None of the columns with the tables have multiple items, as each road in each town only has one overall all ADT (average daily traffic) Though populations in the Towns table may not be 100% related to the primary key, but I think that is up to interpretation. Since I used the tax code for each town as the primary key, I believe the ratio of population to area is relevant to the Towns table.

The grade I believe I earned for the final project would be an A, since I have created the charts in the business section demonstrating the difference of traffic between various interstates. My overall goal would have been to include the accident data although time did not permit. So, a B may be more appropriate.