Peer Review for Ziyi Yang’s Capstone Project

Yebei Yao

**Questions: Is the question clearly stated and well motivated?**

The question I grabbed from the report is to analysis imbalance scooter demand pattern across Chicago, and identify the districts that can’t get access to convenient Scooters within walking distance.

I think it’s a innovative topic since unlike taxi or bike, scooter has more random and distribution pattern and appropriate distribution can be helpful

**Data Analysis: does the data analysis help you understand the context?**

As for now, learning from the report, I ‘m not quite familiar the trip data you use, maybe it will be helpful to make some data set exploratory analysis to make audience now better about the scooter distribution pattern currently before the feature engineering.

**Methods: Are the methods clearly defined?**

The method is to predict the scooter trips across the unit of census tract, the way you define the demand count the trip of each census tract is clear. And because you choose the subset data in September,

Considering the model. I will borrow the advice Elizebeth gave me that, besides choosing the model, it will be helpful introduce the reason why you choose the model briefly.

**Other Opinions:**

For the literature review part, the paper you mentioned has great vision of the cluster spatial analysis, the prediction model about long and short-term memory, and the method of creating buffer zone, why don’t you apply those method into your analysis. Especially, maybe the buffer would be much helpful to define the walking distance.

I like the way you put socio-economic features into account, and as for built environment, I have a tiny opinion that, since the way you measure the geometry density by the percentage of the object in specific tract, but considering each census tract has various area, I did some research and I think the kernel density might be more helpful for feature engineering the density, here’s the vedio that might be helpful.

https://www.youtube.com/watch?v=PBZVTjmhl74

I really like the output of the correlation matrix, which suggests the feature you choose has great correlation, which is either has higher correlation nor little correlation!

I'm looking for the future process, I’d like to learn more about how the results of the model can really help the scooter company to put the enough amount of scooter in the right place across time.