MUSA 550

2020

Fall Final Project

The project will conduct a research of bikeshare data in Bay Area. The dataset will cover the past few years to the current year. This research will find out how COVID-19 affected bikeshare in Bay Area. In this project, l will compare pre-COVID data with Post-COVID data.The entire dataset will surpass 1,000,000 rows. The dataset includes start point, end point, duration, time, and start and end station. The following steps will be conducted:

• Download the dataset, merge these data and slice some outliers.

> Create function to read bike csv files by months. Do some data cleaning for time, create columns for day of week, month, and hour. Create column for hour categories, including rush hour, other, evening and daytime. Divide the dataset before March 17 and after March 17. Rename the columns names in order to merge dataset for PreCovid and PostCovid, as well as 2019 and 2020.

• What is the difference between each day, week or month?

> Groupby different elements. Create graphic for ridership by day of week, by month, by hours.

• Map out the start and end points. Find out the duration distribution of these points.

> Please read the Github posts for more information

• Groupby station, find out which stations are more popular or less popular. A heat map or clustering analysis could help us to visualize these data.

> Please read the Github posts for more information

• Map out these data based on the census tract in Bay area. Find out which census tract is more welcome?

> Please read the Github posts for more information

• Do the same steps for Post-Covid. Is there any change between pre-Covid and postCovid? Do people change their destinations? How does Covid affect the ridership?

> Please read the Github posts for more information

• Use Openstreetmap, create station index, heat map for bike stations

Use OpenStreetmap to import street network. Find out the nearest street for each bike station. Calculate the bike station count for each street segment. Map out the Bike station based on street network. Also map out the heat map for bike stations.

For more information, please read Github post and Jupyter notebook.

• conclusion and recommendation

Please read the Github posts for more information