Project Group -

Members: 5

Student numbers: Zhenlin Xu(5721679), Luyang Cao(5685915), Yuchen Qi(5695392), Xinghao Lou(5698715), Yimen Xu(5696925)

Research Objective

Requires data modeling and quantitative research in Transport, Infrastructure & Logistics

- Investigate the distribution and hotspot of the origin and destination of the travel by taxi in New York City.
- Explore the relationship between OD hotspot and city land use and its change over time (before and after COVID-19)
- Forecast the future hotspot of the origin and destination of the travel by taxi in New York City
- Analyze the characteristics of trip counts per day, weekday average trips, etc.

Contribution Statement

Be specific. Some of the tasks can be coding (expect everyone to do this), background research, conceptualisation, visualisation, data analysis, data modelling

```
**Author 1**:

**Author 2**:

**Author 3**:
```

Data Used

The data used in this group project were collected and provided to the NYC Taxi and Limousine Commission (TLC). Raw data is available at: TLC Trip Record Data - TLC (nyc.gov).

We also use the NYC's Zoning and Land Use data provided by https://zola.planning.nyc.gov/about/#9.72/40.7125/-73.733 (https://zola.planning.nyc.gov/about/#9.72/40.7125/-73.733).

Data Pipeline

1.Background research: reading taxi travel-related literature about study of Origin-Destination pair.

2.Data collecting: downloading taxi travel data collected from 2018-2022 provided by TLC Trip Record Data - TLC (nyc.gov).

3.Data checking: using Numpy and Pandas to clean and filter the data.

4.Statistical explanatory: using Pandas to calculate several traffic parameters and Matplotlib to visualize the results.

5.Origin and destination clustering: using the clustering algorithms from Scikit-learn to calculate cluster centers of taxi travel origins and destinations.

6.Origin and destination visualization: using Plotly to visualize the spatial-temporal evolution of the New York City taxi travel origin and destination hotspots (before and after COVID-19).