**UBER (New York City) project:**

Link of data [Link](https://www.kaggle.com/datasets/shuhengmo/uber-nyc-forhire-vehicles-trip-data-2021)

**Purpose of project:**

* **Data Analysis:**
  + **EDA and figure out underlying trip patterns in 2021.**
  + **Try exploring Uber's user portrait in NYC (which orders are urgent and what kind of users should be given higher priorities?)**
* **Data Science:**
  + **Based on fhvhv\_tripdata\_2021 and weather data, build predict model to predict the peak footfall.**

**Break down in Data analysis step:**

1. **Collection data:**
   1. **Download data from Kaggle with /\*.csv format and insert this data in DataBase to be like data warehouse to be easy for ETL at many software I used like python or power Bi because don’t load in cache memory.**
   2. **We have large data, so I select specific data depending on main business problem**
2. **Solving Business Problem:**
   1. **figure out underlying trip patterns in 2021**
   2. **Try exploring Uber's user portrait in NYC (which orders are urgent and what kind of users should be given higher priorities?)**
3. **Manipulation and Cleaning Data**
4. **Create Dashboard**