Project Requirement:

- The data contains some duplicates and missing values that needs to be cleaned and prepared for data analysis.
- Insights to be fetched from the data
    - Which month have max. Uber pickups in New York City ?
    - What is the Hourly Rush in New York city on all days ?
    - Which Base_number has the most number of Active Vehicles ?
    - On what locations of New York City, we are getting rush ?
    - Examine the rush on Weekday with respect to each Hour

About Dataset

Uber TLC FOIL Response

This directory contains data on over 4.5 million Uber pickups in New York City from April to September 2014, and 14.3 million more Uber pickups from January to June 2015. Trip-level data on 10 other for-hire vehicle (FHV) companies, as well as aggregated data for 329 FHV companies, is also included. All the files are as they were received on August 3, Sept. 15 and Sept. 22, 2015.

FiveThirtyEight obtained the data from the NYC Taxi & Limousine Commission (TLC) by submitting a Freedom of Information Law request on July 20, 2015. The TLC has sent us the data in batches as it continues to review trip data Uber and other HFV companies have submitted to it. The TLC's correspondence with FiveThirtyEight is included in the files `TLC_letter.pdf`, `TLC_letter2.pdf` and `TLC_letter3.pdf`. TLC records requests can be made here.

This data was used for four FiveThirtyEight stories: Uber Is Serving New York's Outer Boroughs More Than Taxis Are, Public Transit Should Be Uber's New Best Friend, Uber Is Taking Millions Of Manhattan Rides Away From Taxis, and Is Uber Making NYC Rush-Hour Traffic Worse?.

The Data

The dataset contains, roughly, four groups of files:

- Uber trip data from 2014 (April - September), separated by month, with detailed location information

- Uber trip data from 2015 (January - June), with less fine-grained location information

- non-Uber FHV (For-Hire Vehicle) trips. The trip information varies by company, but can include day of trip, time of trip, pickup location, driver's for-hire license number, and vehicle's for-hire license number.

- aggregate ride and vehicle statistics for all FHV companies (and, occasionally, for taxi companies)

Uber trip data from 2014

There are six files of raw data on Uber pickups in New York City from April to September 2014. The files are separated by month and each has the following columns:

- `Date/Time` : The date and time of the Uber pickup
- `Lat` : The latitude of the Uber pickup
- `Lon` : The longitude of the Uber pickup
- `Base` : The TLC base company code affiliated with the Uber pickup

These files are named:

- `uber-raw-data-apr14.csv`
- `uber-raw-data-aug14.csv`
- `uber-raw-data-jul14.csv`
- `uber-raw-data-jun14.csv`
- `uber-raw-data-may14.csv`
- `uber-raw-data-sep14.csv`

Uber trip data from 2015

Also included is the file `uber-raw-data-janjune-15.csv` This file has the following columns:
- `Dispatching_base_num` : The TLC base company code of the base that dispatched the Uber
- `Pickup_date` : The date and time of the Uber pickup
- `Affiliated_base_num` : The TLC base company code affiliated with the Uber pickup
- `locationID` : The pickup location ID affiliated with the Uber pickup

The `Base` codes are for the following Uber bases:

B02512 : Unter

B02598 : Hinter

B02617 : Weiter

B02682 : Schmecken

B02764 : Danach-NY

B02765 : Grun

B02835 : Dreist

B02836 : Drinnen

For coarse-grained location information from these pickups, the file `taxi-zone-lookup.csv` shows the taxi `Zone` (essentially, neighborhood) and `Borough` for each `locationID`.

Non-Uber FLV trips

The dataset also contains 10 files of raw data on pickups from 10 for-hire vehicle (FHV) companies. The trip information varies by company, but can include day of trip, time of trip, pickup location, driver's for-hire license number, and vehicle's for-hire license number.

These files are named:

- `American_B01362.csv`
- `Diplo_B01196.csv`
- `Highclass_B01717.csv`
- `Skyline_B00111.csv`
- `Carmel_B00256.csv`
- `Federal_02216.csv`
- `Lyft_B02510.csv`
- `Dial7_B00887.csv`
- `Firstclass_B01536.csv`
- `Prestige_B01338.csv`

Aggregate Statistics

There is also a file `other-FHV-data-jan-aug-2015.csv` containing daily pickup data for 329 FHV companies from January 2015 through August 2015.

The file `Uber-Jan-Feb-FOIL.csv` contains aggregated daily Uber trip statistics in January and February 2015.