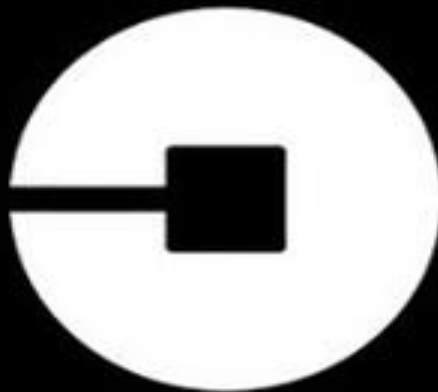


Uber Pickups



Leif Waldner, Parth Mishra

Project Description

We want to explore an Uber dataset that outlines Uber pickups in New York City. This dataset encompasses location, time, dispatcher, and date. We want to look at trends over time of spacial and temporal uber ride distribution in New York City. If possible, we also want to integrate with other demographic information to get more nuanced insights.

Prior Work

FiveThirtyEight has provided the dataset and done some research in this area:

- [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](#)
-

List of Datasets

<https://www.kaggle.com/fivethirtyeight/uber-pickups-in-new-york-city>

- Cross-reference with demographic information of NYC
-

Proposed Work - What do we need to do?

Cleaning: Ensure validity and accuracy of data set

Integration: Possibly integrate with demographic datasets for more complex queries

Preprocessing: Depending on integration, may need to process data for consistency in data types

Visualization: Create heat map

Tools

Python - Pandas, SKlearn for any model building

R - Exploratory data analysis

D3 - Visualization

Evaluation

Inferential Statistical tests to validate/invalidate hypotheses of dataset
