

# Predicting hourly customer footfall for Uber

## Problem statement

In this project, I want to improve Uber's ability to meet demand by predicting the number of Uber requests for each NY taxi zone based on date, hour, and weather conditions.

## Context

Demand for taxi services depends on many factors. I'm assuming that some of the important ones are:

- Time of the day: during rush hours (morning & evening) the demand may increase due to the need to commute to and from the office;
- Day of the week: on the weekends and public holidays the demand may be higher due to the increased number of tourists;
- Location: downtown and areas with places of cultural interest may have higher demand;
- Weather conditions: people may prefer taking an Uber instead of driving themselves (or walking, or biking) when it's cold, snowing, or raining.

The ability to meet the demand when it surges, and manage drivers' expectations when the demand decreases yields higher satisfaction among customers and drivers, higher operational efficiency, and increased revenue.

## The value of being able to predict customer footfall:

1. It would allow Uber to implement measures to meet the surge in demand (for example, by incentivizing more drivers to work during peak hours) and receive higher revenues.
2. Alternatively, it could shape Uber's pricing policy to charge higher prices when demand increases.
3. Understanding what factors affect demand also allows Uber to manage passengers' expectations about the wait time and drivers' expectations about the number of customers.

## The clients for this project would be:

- Chief Revenue Officer - since the project's success could lead to increased revenue.
- Potentially, the Operations team, since knowing the factors that affect wait time may result in more efficient driver allocations.

## Criteria for success

Increased revenue by xx%.

## Scope of solution space

Creating a model to predict demand for Uber rides based on the date, time, location and weather conditions.

Deliverables: a GitHub repo with notebooks for data wrangling, EDA, pre-processing, modeling, project report, and slide deck.

## Data sources

New York for-hire vehicles trip data for 2021 from Taxi and Limousine Commission:  
[https://www.kaggle.com/datasets/shuhengmo/uber-nyc-forhire-vehicles-trip-data-2021/data?select=fhvhv\\_tripdata\\_2021-02.parquet](https://www.kaggle.com/datasets/shuhengmo/uber-nyc-forhire-vehicles-trip-data-2021/data?select=fhvhv_tripdata_2021-02.parquet)

**Target feature:** Count of Uber requests by zone and hour.

