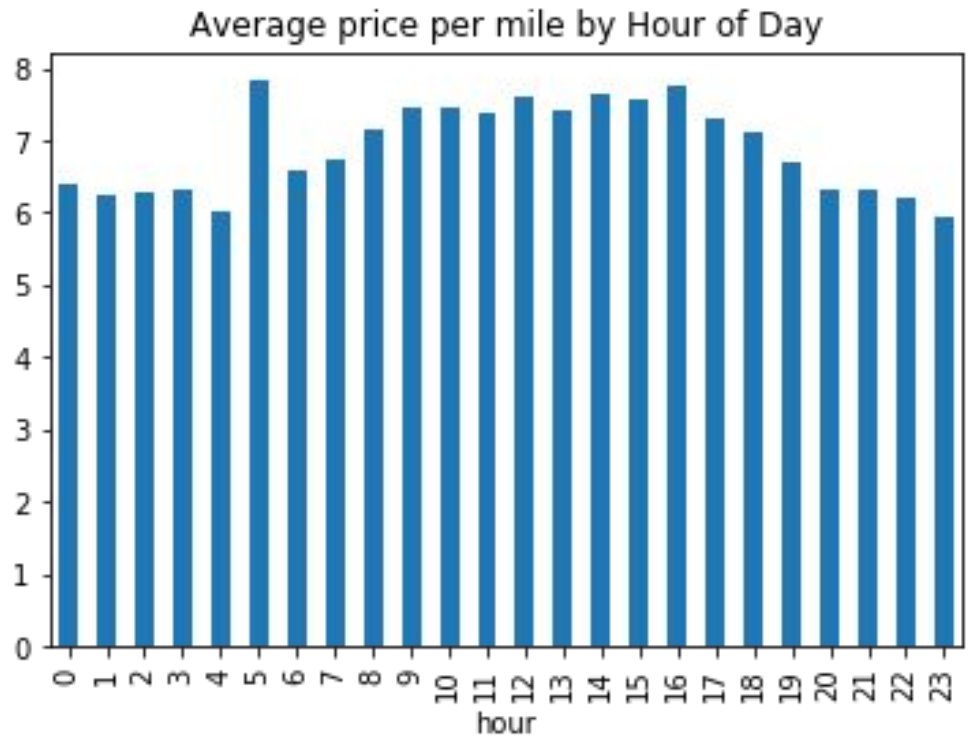

# Being competitive from the perspectives of company and driver

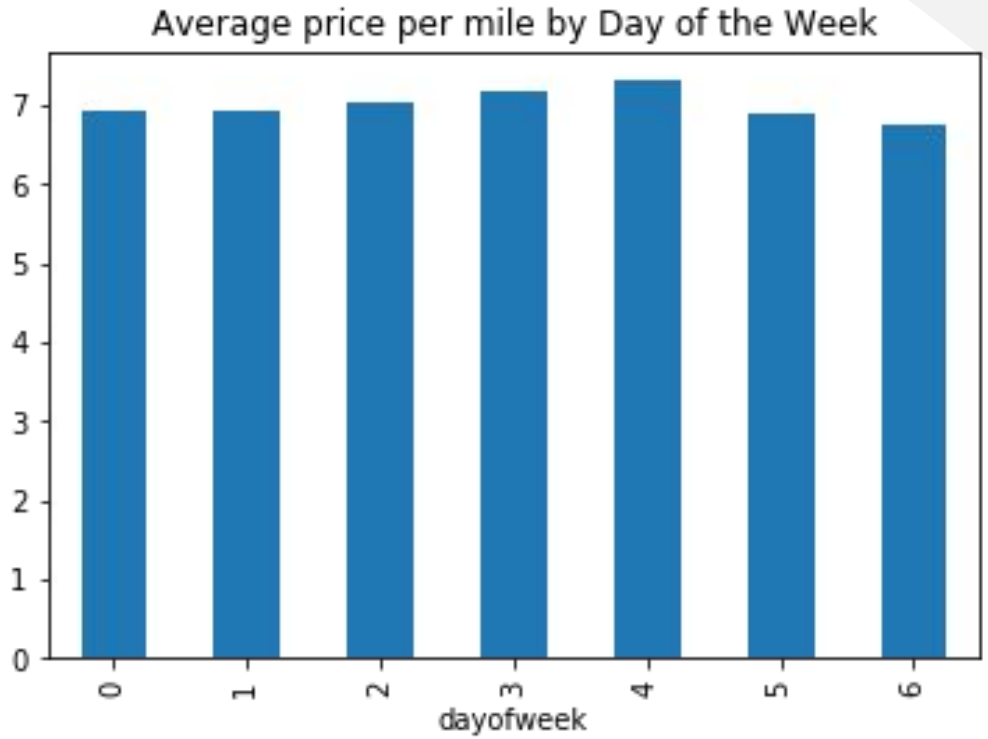
DS501 Case 3 Group 9
Jack Charbonneau, Jannik Hass, Weixiao Huang, Ye Hong

# Business Problem

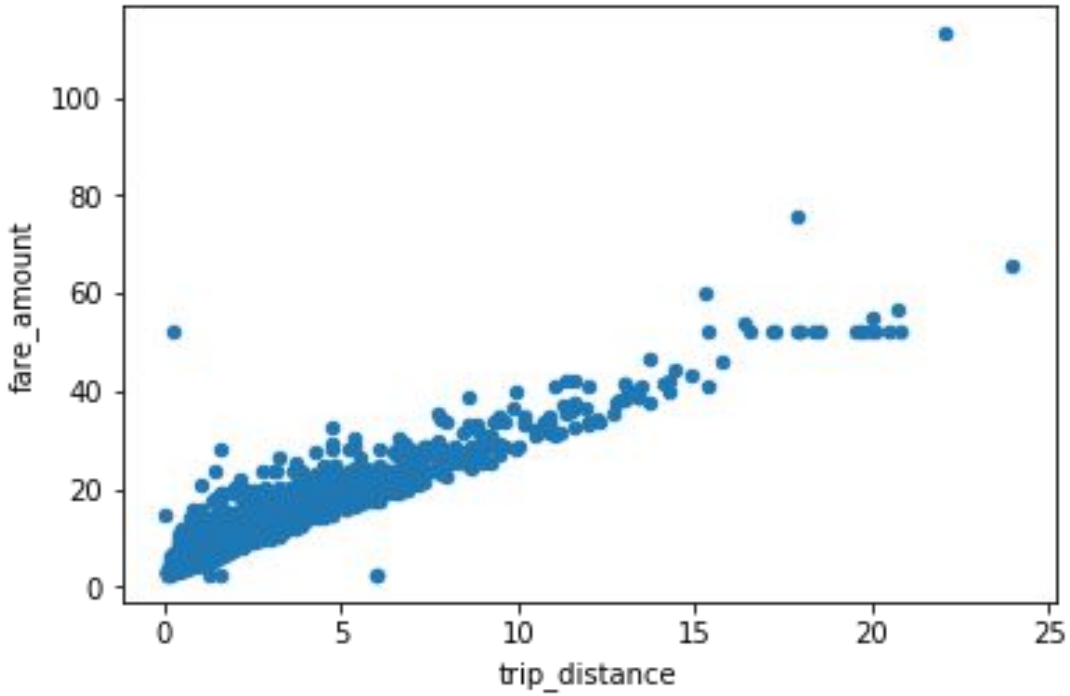
- Predicting trip price to remain competitive amongst ride-share companies
- Estimating Uber drivers upcoming trip duration to help them better manage their schedule

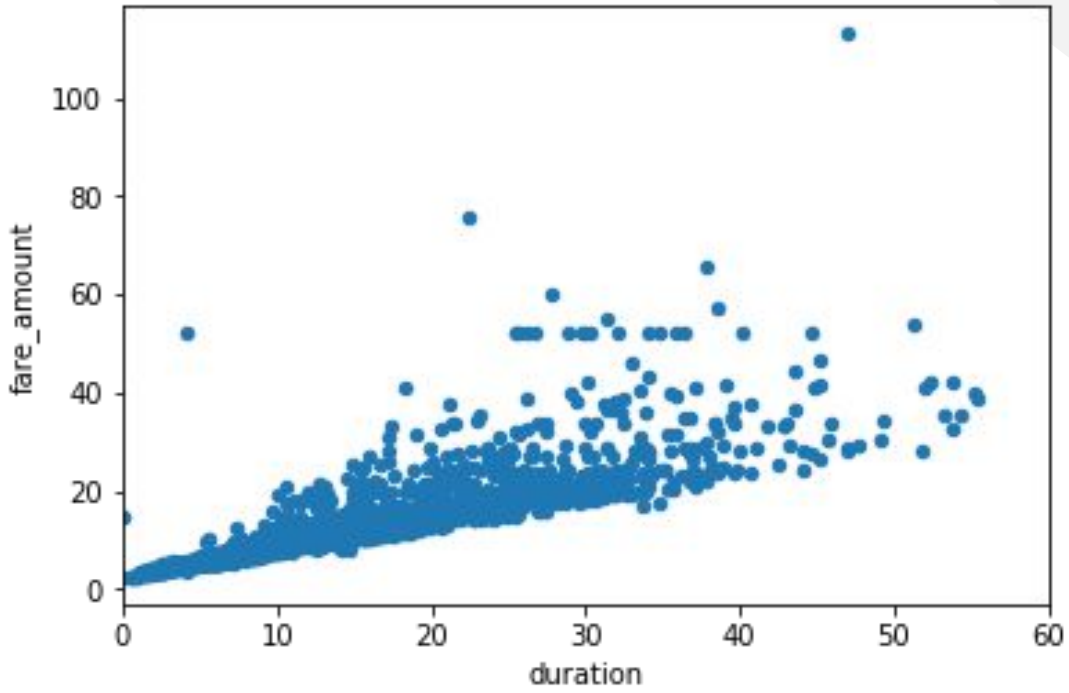
# Analyzing variability of price



# Main factors affecting price



# Solution

## Multi-Linear Regression Model

- Input features
  - Day of the Week
  - Hour of the Day
  - Trip Distance
  - Trip Duration
- Output
  - Fare Amount
- Training Data
  - January
- Testing Data
  - February

# Results

- Mean absolute error: 7.6e-7
- Mean square error: 1.2e-6

# Trip Duration Estimation

- **Datasets**: Weather and Trip Data
- **Period:** January 2018 - December 2018
- Merge two datasets based on **Date**
- Calculate **Duration** for each trip
- Retrieve **Time Details** into 3 new features
- **8759874** Trips in January 2018

| MonthDayNum | DayOfWeekNum | pickup_time |
|---|---|---|
| 1 | 0 | 00:21:05 |
| | Sunday | |
| 1 | 0 | 00:44:55 |
| 1 | 0 | 00:08:26 |
| 1 | 0 | 00:20:22 |
| 1 | 0 | 00:09:18 |

# Multi-Linear Regression Model

- **Training Feature**
  - Trip_distance, Passenger_count
  - Weather (AWND, PRCP,  SNOW, SNWD, TMAX, TMIN)
- **Target Feature**
  - Duration
- **Training dataset** - January 2018 / **Testing dataset** - February 2018
- **Mean Absolute Error (MAE)**
  - 7.353581824993934e-05
- **Root Mean Square Error (RMSE)**
  - 0.00044689055700656904

# Estimate Duration Result

```
# Customize ['trip_distance','passenger_count','DayOfWeekNum','AWND','PRCP','SNOW','TMAX','TMIN']
x_1 = [[5, 3.0, 6,7.83, 5.0, 3.0, 22.0, 13.0]]
y_1 = model.predict(x_1)
print(y_1)
```

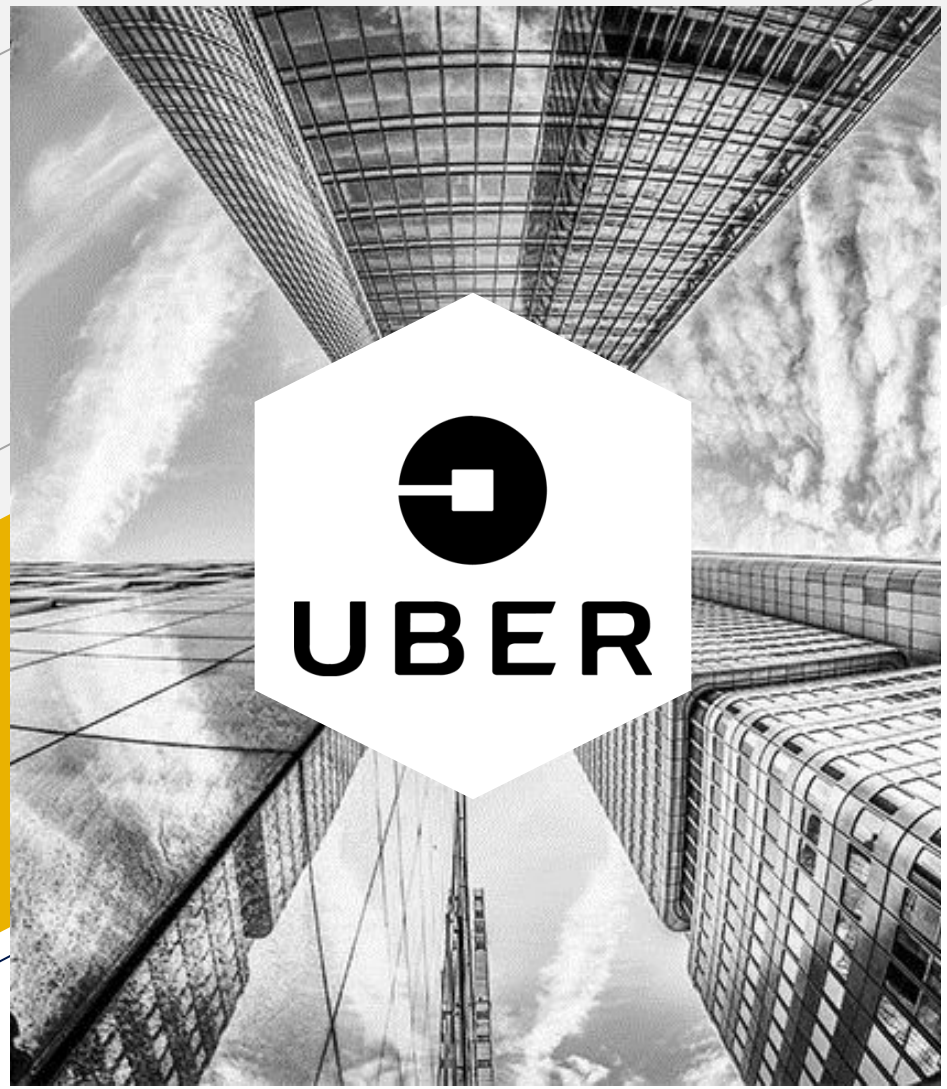
```
[[1237.54185227]]
```

```
# Customize ['trip_distance','passenger_count','DayOfWeekNum','AWND','PRCP','SNOW','TMAX','TMIN']
x_1 = [[10, 1.0, 3,7.83, 3.0, 1.0, 15.0, 4.0]]
y_1 = model.predict(x_1)
print(y_1)
```

```
[[1062.98795181]]
```

```
# Customize ['trip_distance','passenger_count','DayOfWeekNum','AWND','PRCP','SNOW','TMAX','TMIN']
x_1 = [[1.4, 3.0, 6,7.83, 0.0, 2.0, 22.0, 13.0]]
y_1 = model.predict(x_1)
print(y_1)
```

```
[[976.36505681]]
```




# Thank You