Understanding what drives prices of ride-hailing services

November 2023

Antonio Montilla

Database

- Dataset: Kaggle challenge (see Kaggle)
  - **693,000** request trips using Uber & Lyft APIs
  - 12 districts of the city of Boston
  - During the last week of November 2018

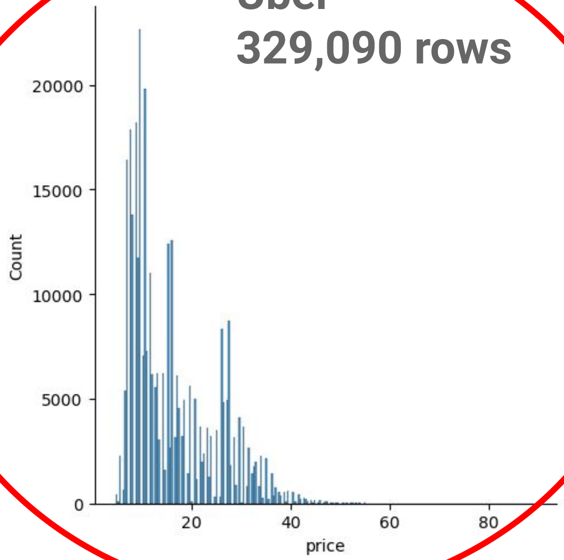
Main columns:

- Price (target variable)
- Distance
- Time: hour of the day, day of the week
- Location: destination & source
- Cab type
- Weather: rain, temperature, wind

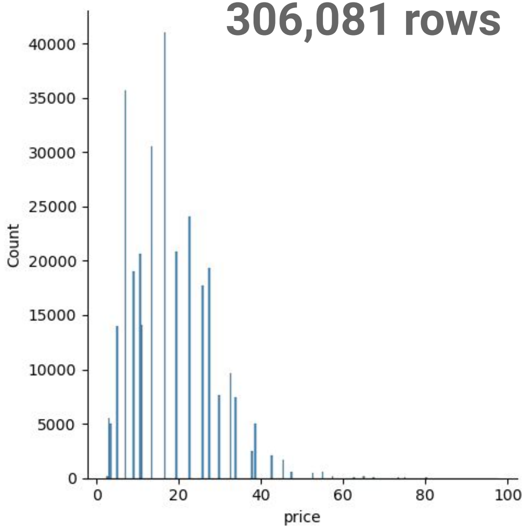
Database: downshifting the problem

- Price histograms of rides from Uber vs. Lyft

Uber
329,090 rows

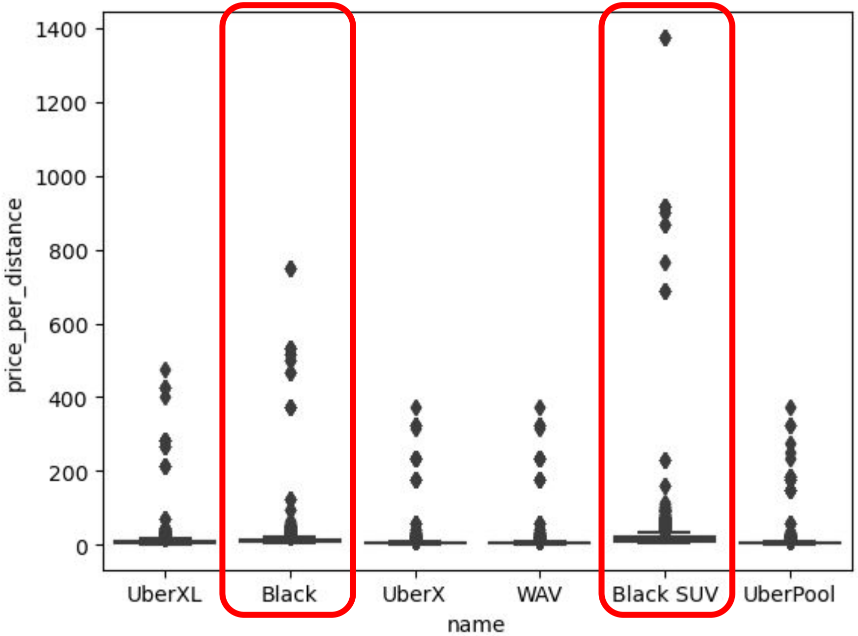


Lyft
306,081 rows



Database: downshifting the problem

- Price per mile of Uber by type of service

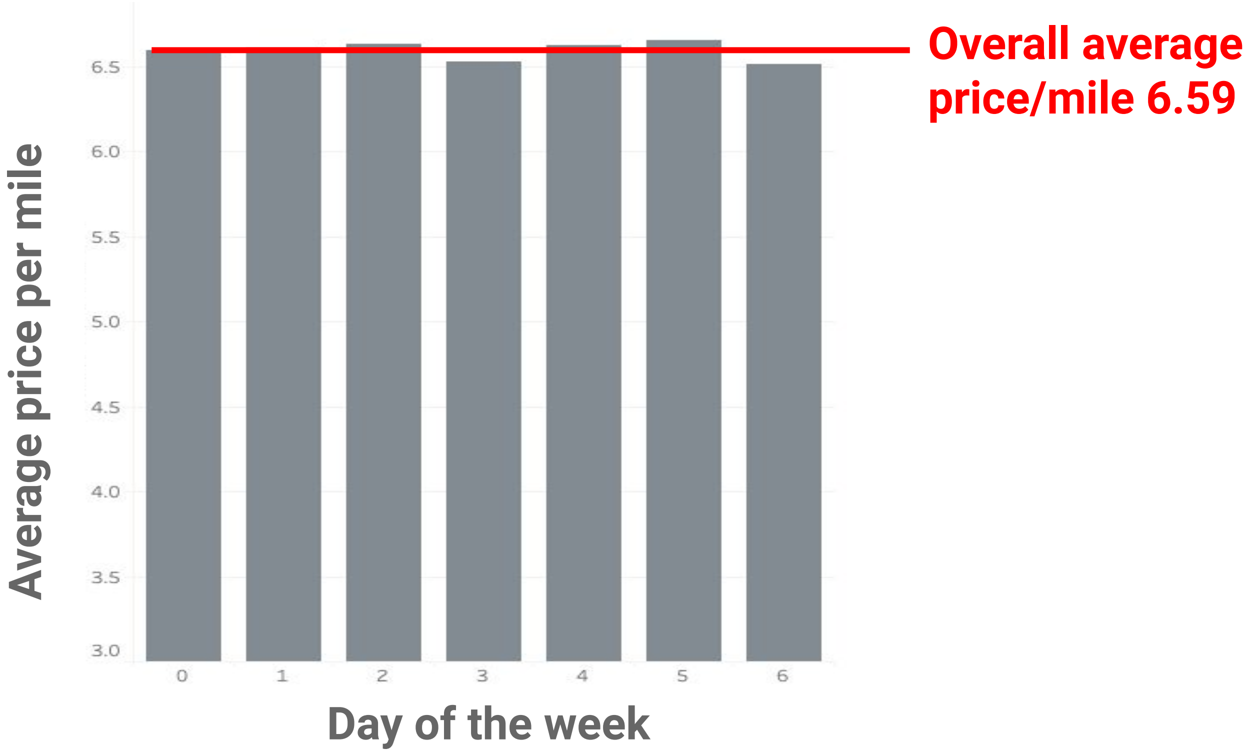


Decision:

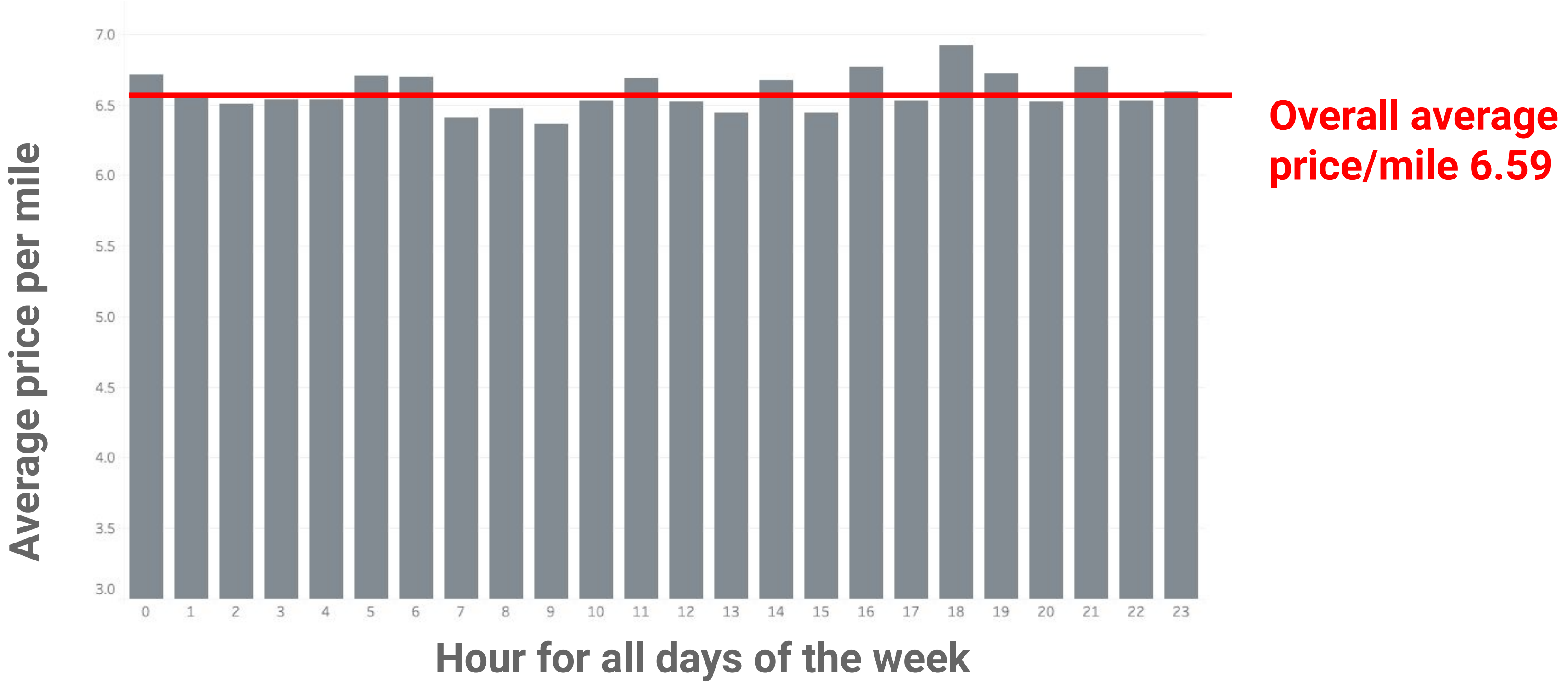
- Exclude Black SUV
- Exclude Black

Final rows: 219,407

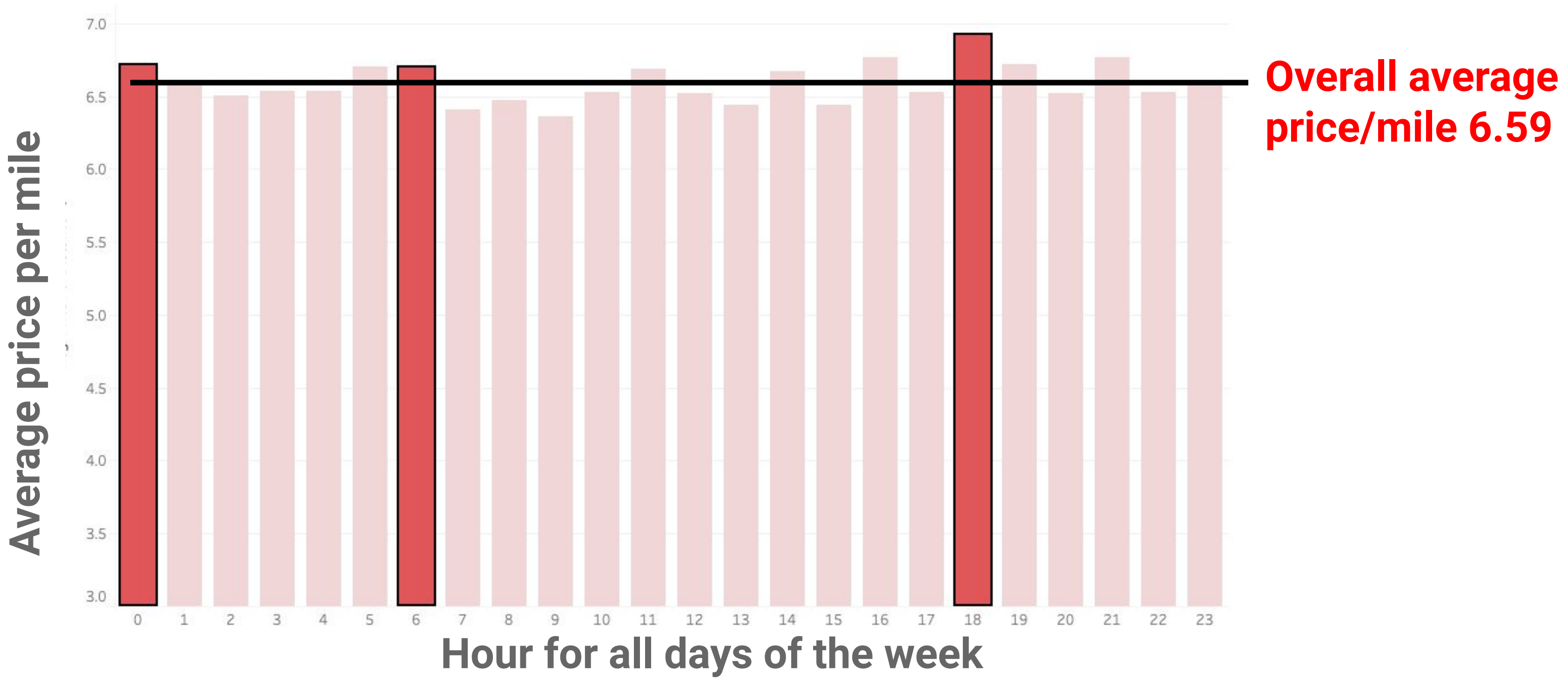
First: ¿how average prices change per day?



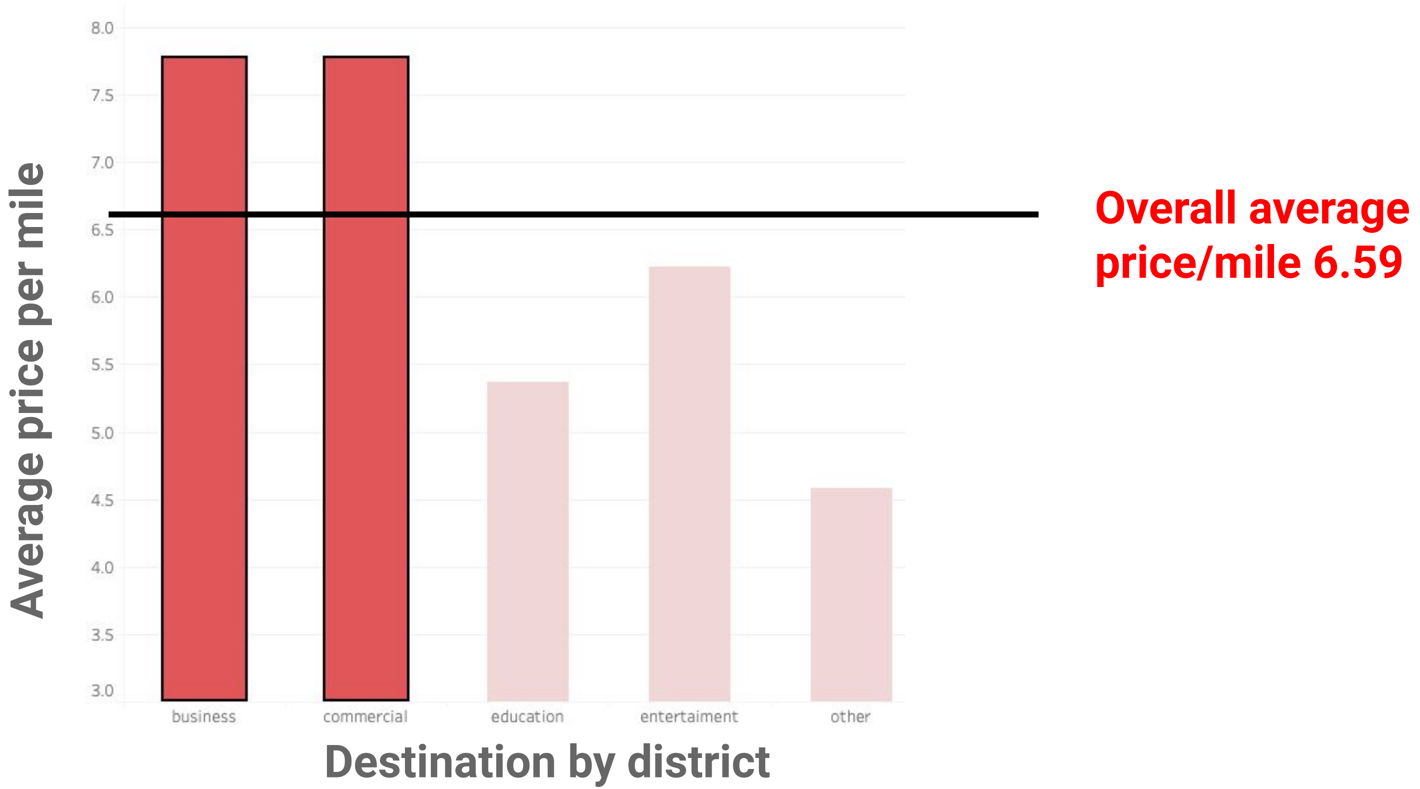
¿What about by hour of the day?



¿What about by hour of the day?

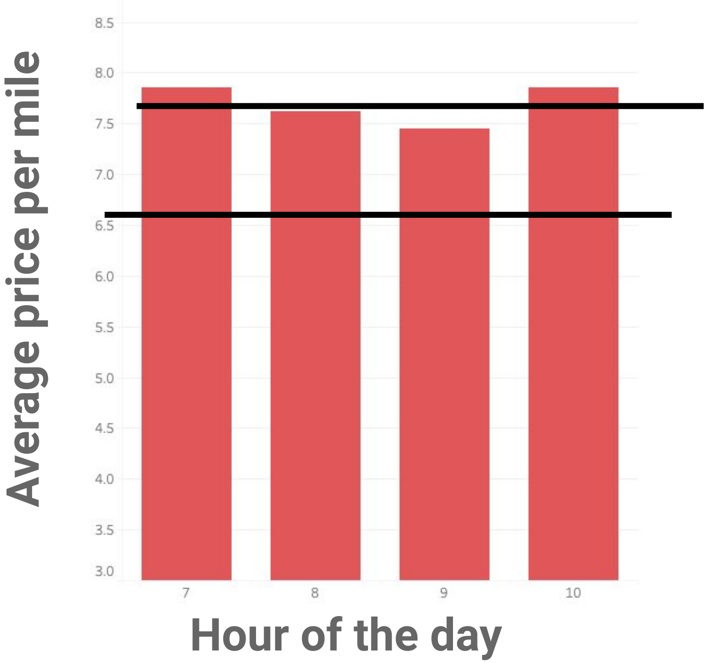


¿What about by the destination of the ride?



Hyp. 1: prices are higher in morning rush hours?

Filter 1 (day): Monday-Thursday
Filter 2 (destination): “financial” district



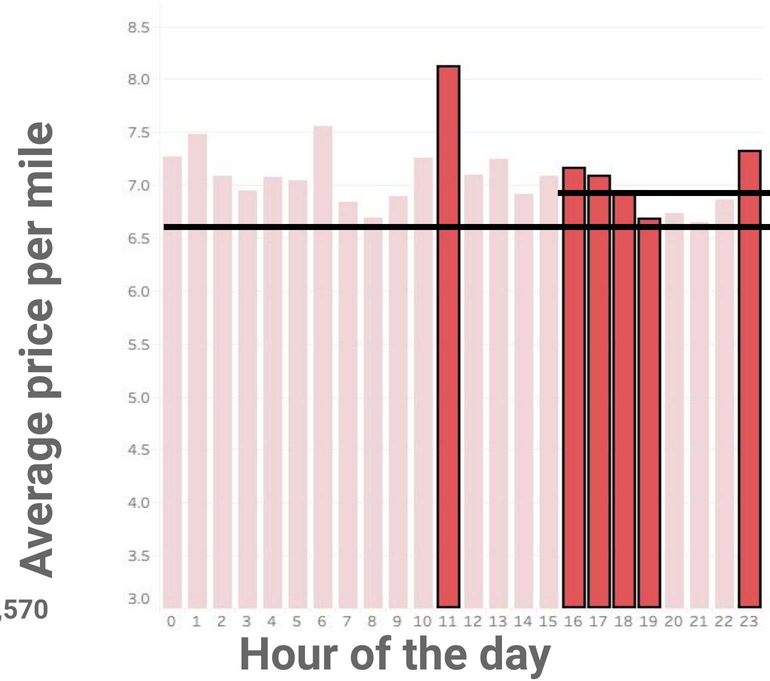
Morning rush average price/mile 7.62

Overall average price/mile 6.59

Rows: 3,424

Hyp. 2: prices are higher in afternoon rush hours?

Filter 1 (day): Monday-Thursday
Filter 2 (source): “financial” district



Rows: 4,570

Afternoon rush average price/mile 7.04

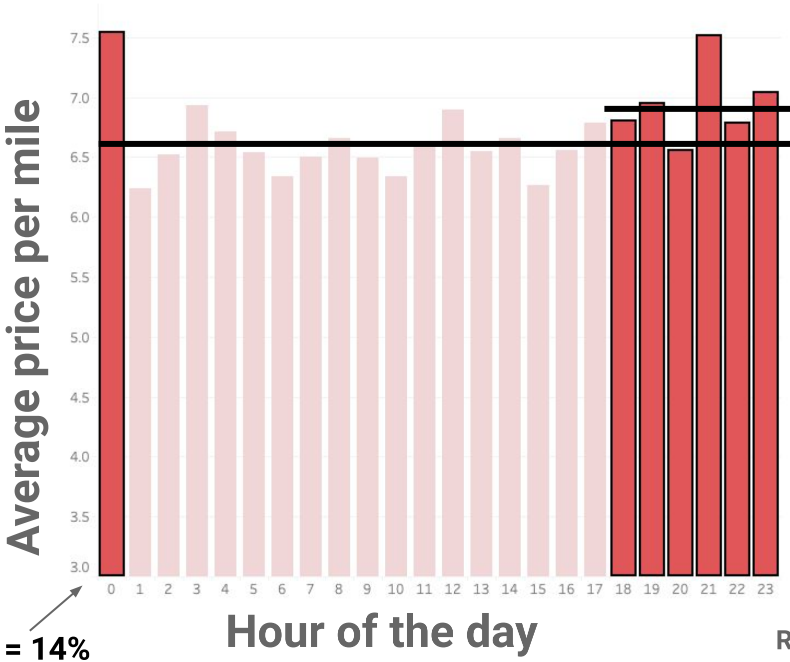
Overall average price/mile 6.59

At 11pm, increase = 11%

7.0%

Hyp. 3: prices are higher in weekend rush hours?

Filter 1 (day): Friday-Saturday
Filter 2 (destination): “entertainment” & “commercial” districts



At 12am increase = 14%

Rows: 4,050

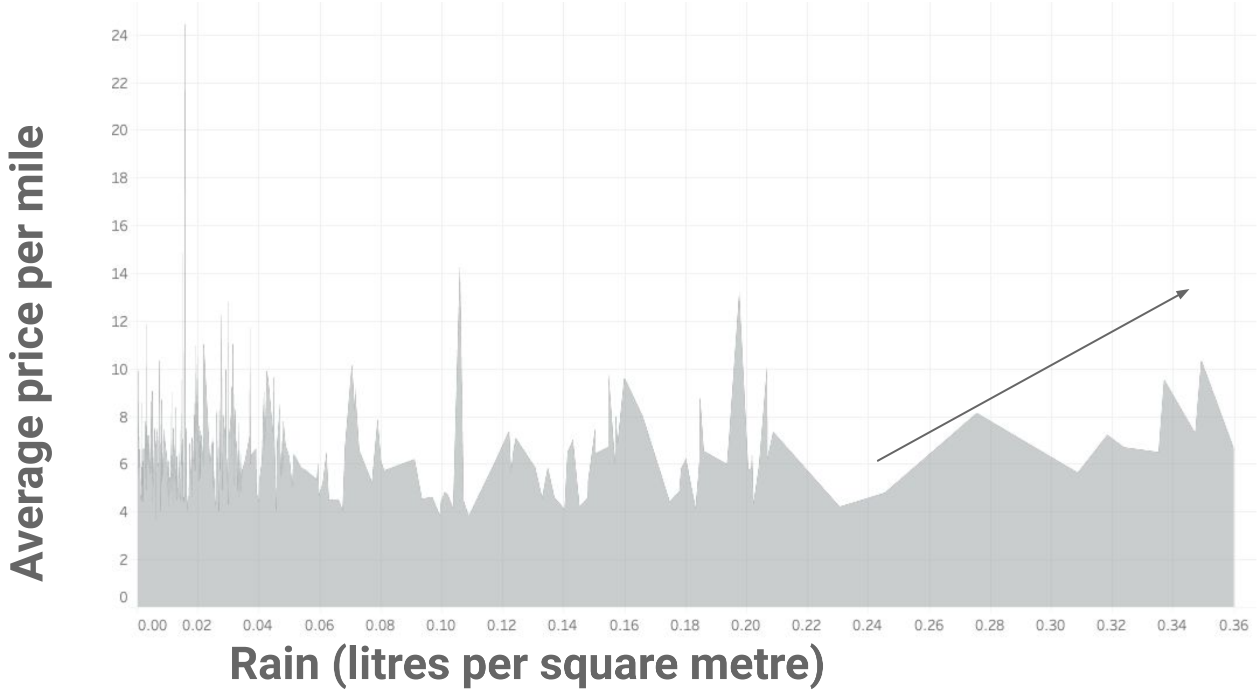
Weekend rush average price/mile 6.95

Overall average price/mile 6.59

5.5%

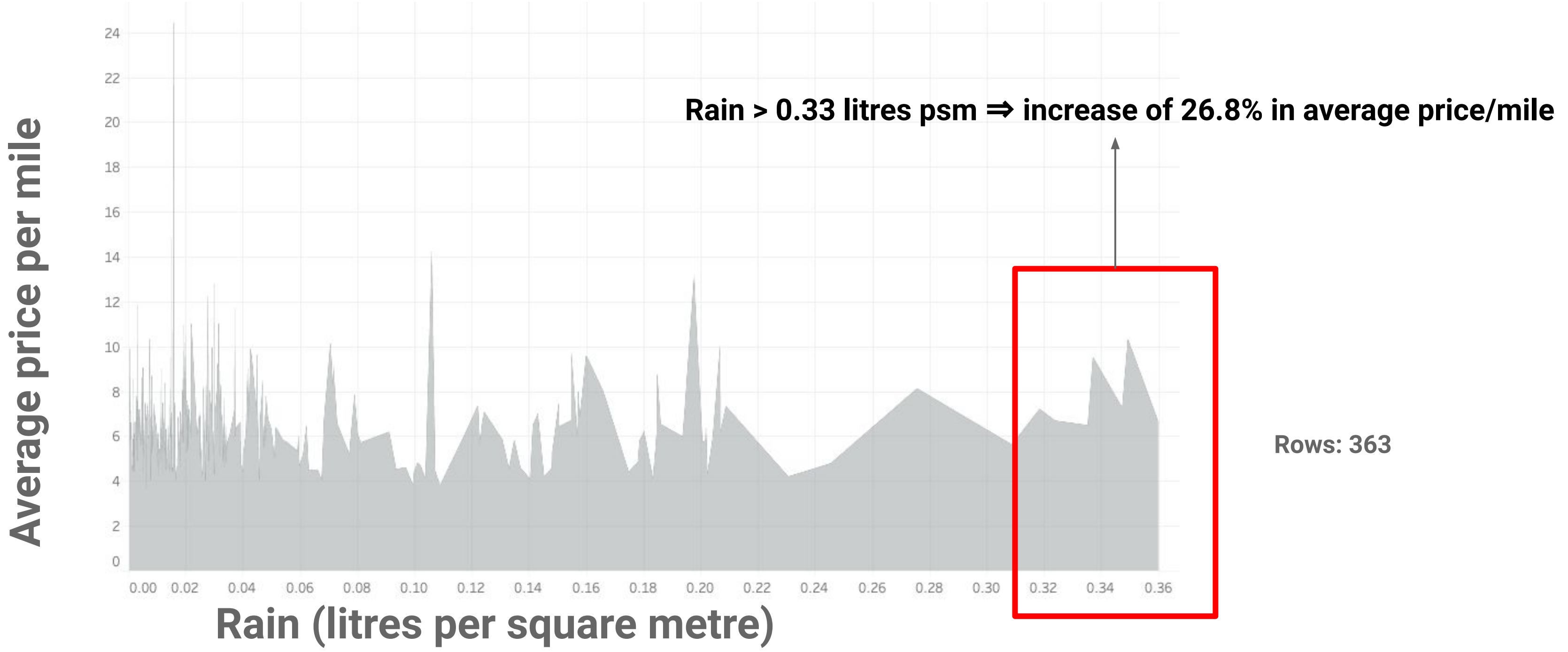
Hyp. 4: prices are impacted by weather conditions?

Average price/mile per hour versus rain (litres per square metre)



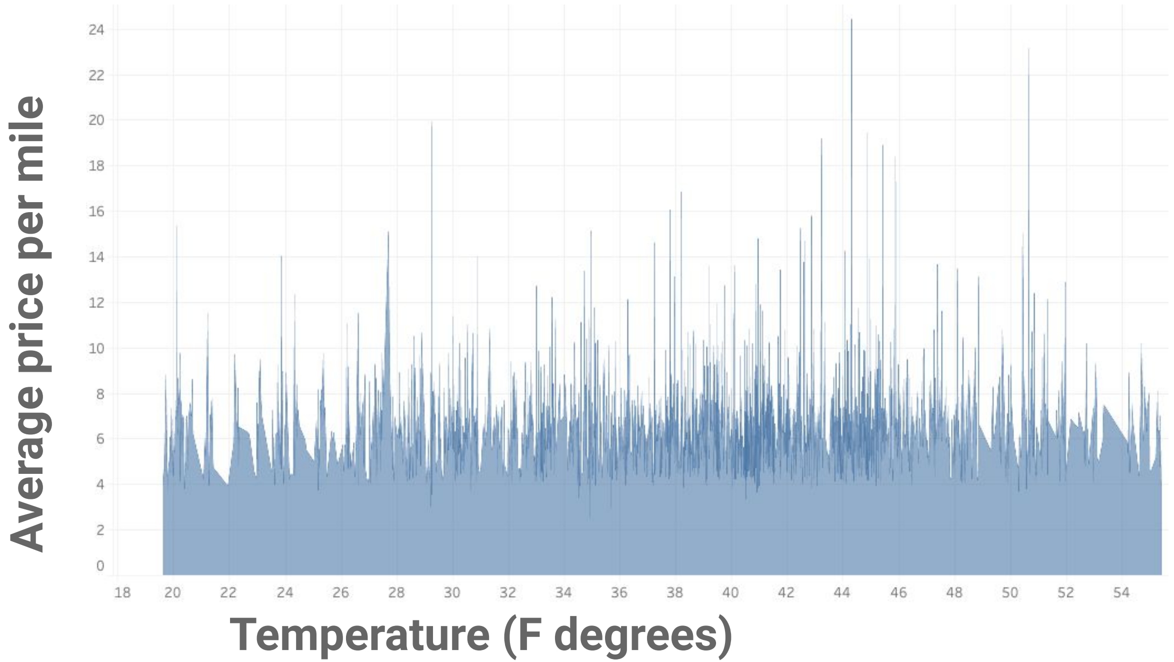
Hyp. 4: prices are impacted by weather conditions?

Average price/mile per hour versus rain (litres per square metre)



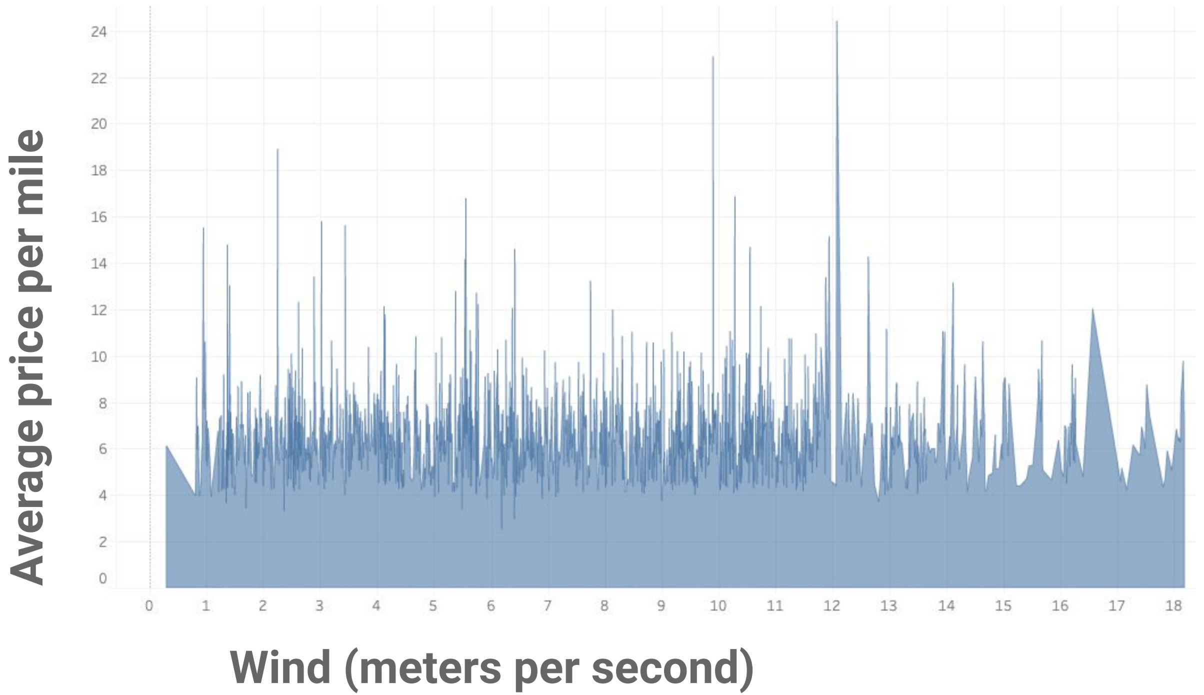
Hyp. 4: prices are impacted by weather conditions?

Average price/mile per hour versus temperature (F degrees)



Hyp. 4: prices are impacted by weather conditions?

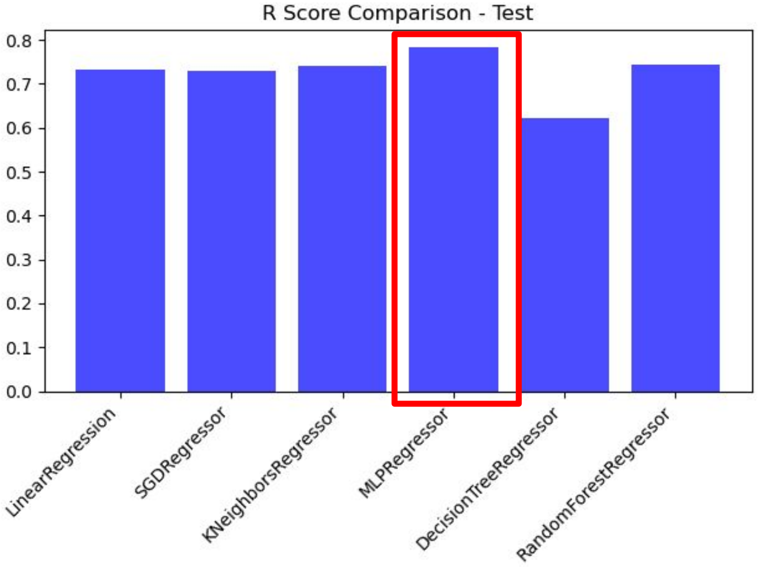
Average price/mile per hour versus wind (meters per second)

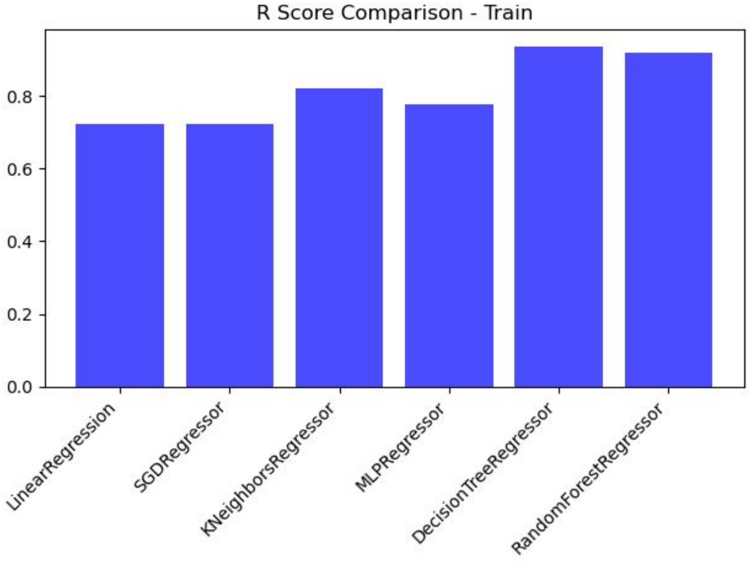


Weather seems to be less relevant

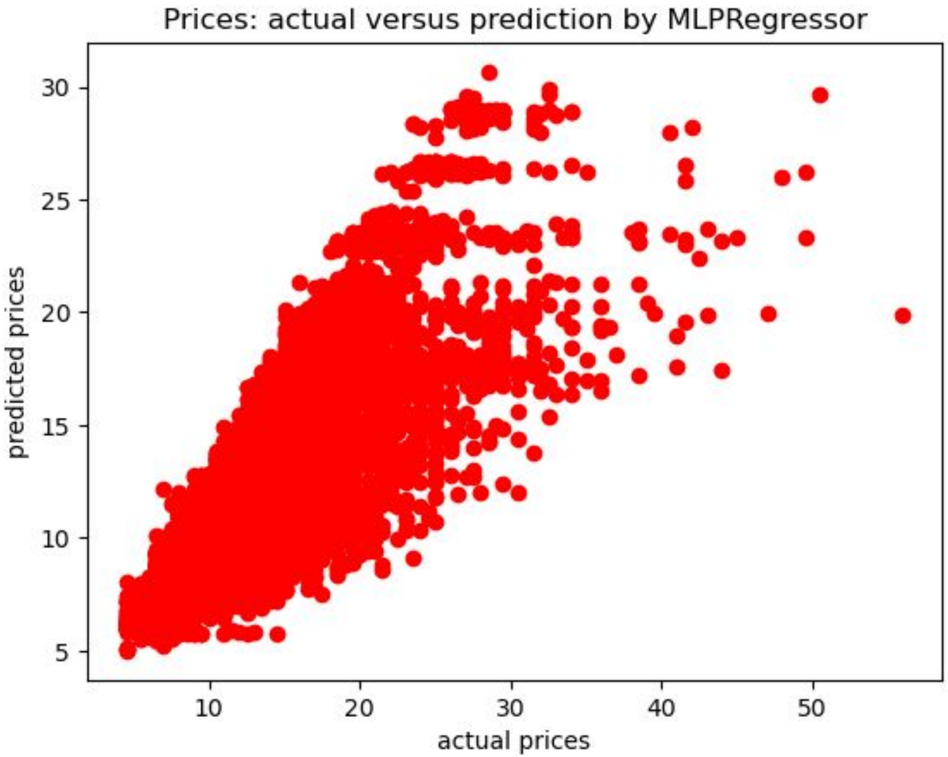
Objective 2: a ML regression model to predict prices

- Construct ML regression models to see effect of X variables in explaining prices
- Data wrangling, data transformations, X-Y split, etc…
- …testing different models:



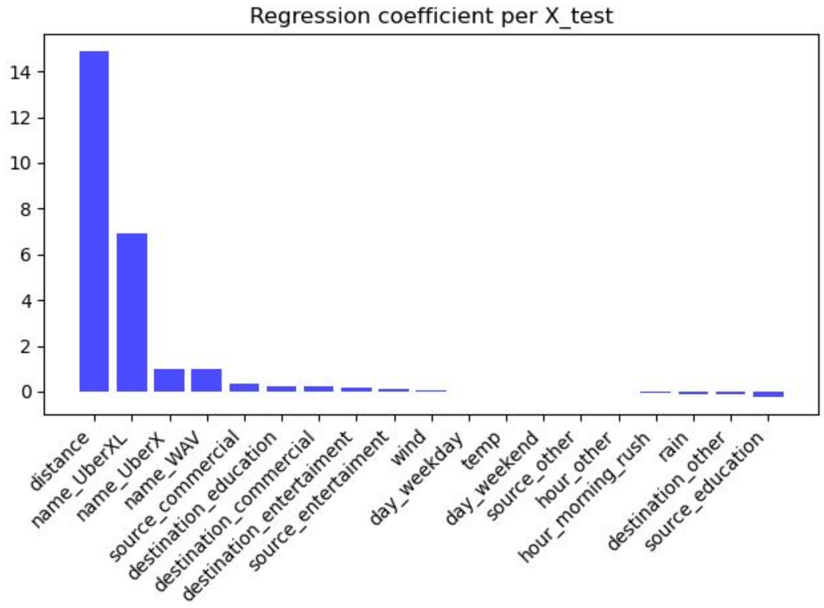


MLP Regressor: Y_pred versus Y_test



BUT: which factors are more relevant for prices?

- Results based on a Linear Regression model, as MLPRegressor does not allow to see coefficients for each X_train columns.



- **No surprise:** distance explains most of the variability in prices.
- **Other factors**: type of Uber, e.g. if XL or if WAV.
- **Other columns seem very irrelevant** in explaining prices.
- In fact, a separate model only with ‘distance’ and cab type delivers same R_score

Final thoughts, caveats

- **Distance** is by far the most single factor influencing the price set by ride-hailing firms.
- Other factors that have some relevance in prices are also the **type of service** chosen in the App.
- But, **dynamic pricing is at play** (e.g. morning rush hours, late nights, extreme weather conditions)...
- …only that for a **small share of the rides** (less than 10%).
- The **main caveats** for these insights are the specification of the data:
    - Timing: 1 week in late-November 2018
    - Location: only 12 districts of Boston.
    - User: same requests.