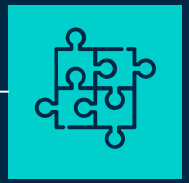
SHARED DOCKLESS MOBILITY: PREDICTING USAGE

Gillian Foster, Saranya Nagarajan, Katherine Wroble,
Malik Ouda, Mounika Tarigopula, Nadia Florez

TABLE OF CONTENTS



01

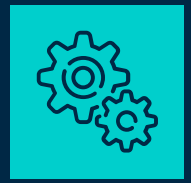
BACKGROUND +
DATA DESCRIPTION

02

VISUALIZATIONS

03

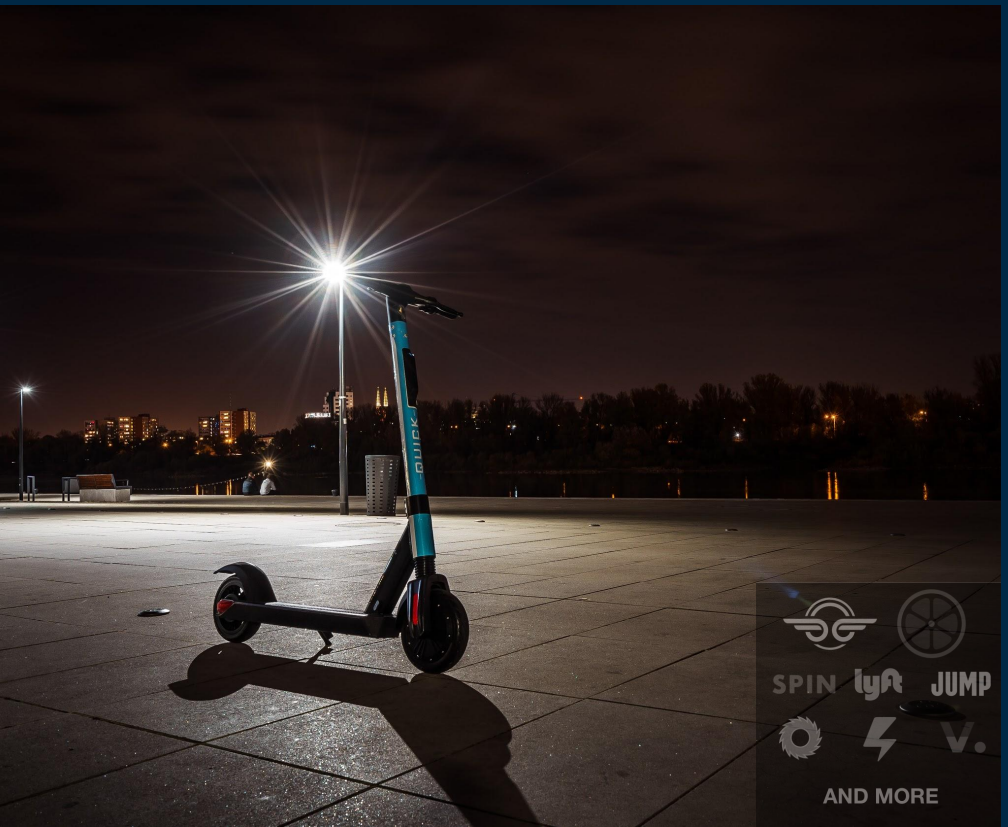
MODELING



04

REFLECTIONS

BACKGROUND

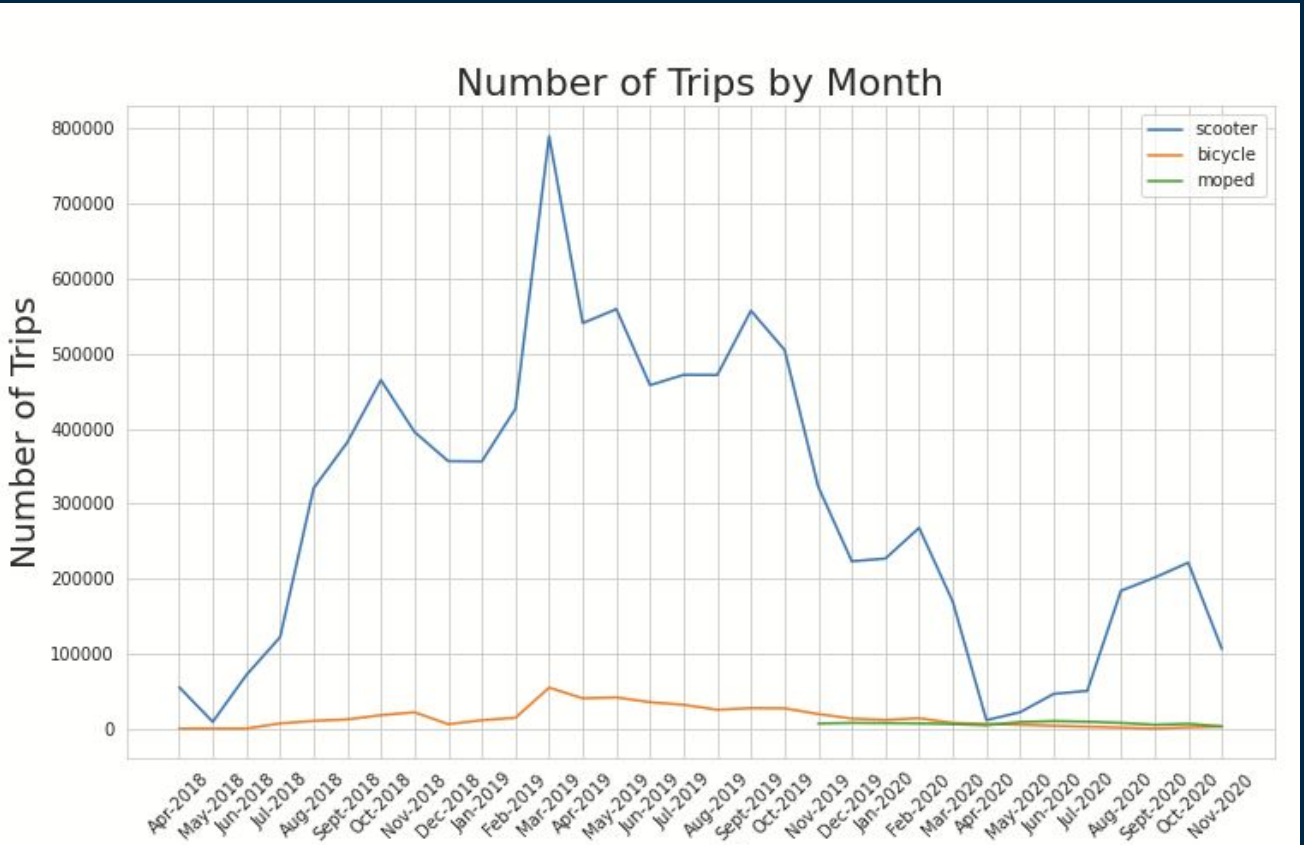


- A new form of transportation: shared dockless mobility
- We use advanced analytics to predict number of trips from shared dockless mobility services by leveraging historical weather and shared dockless mobility data
- Our model will help service provides better meet demand, limit the number of idle scooters and maximize customer satisfaction

DATA DESCRIPTION

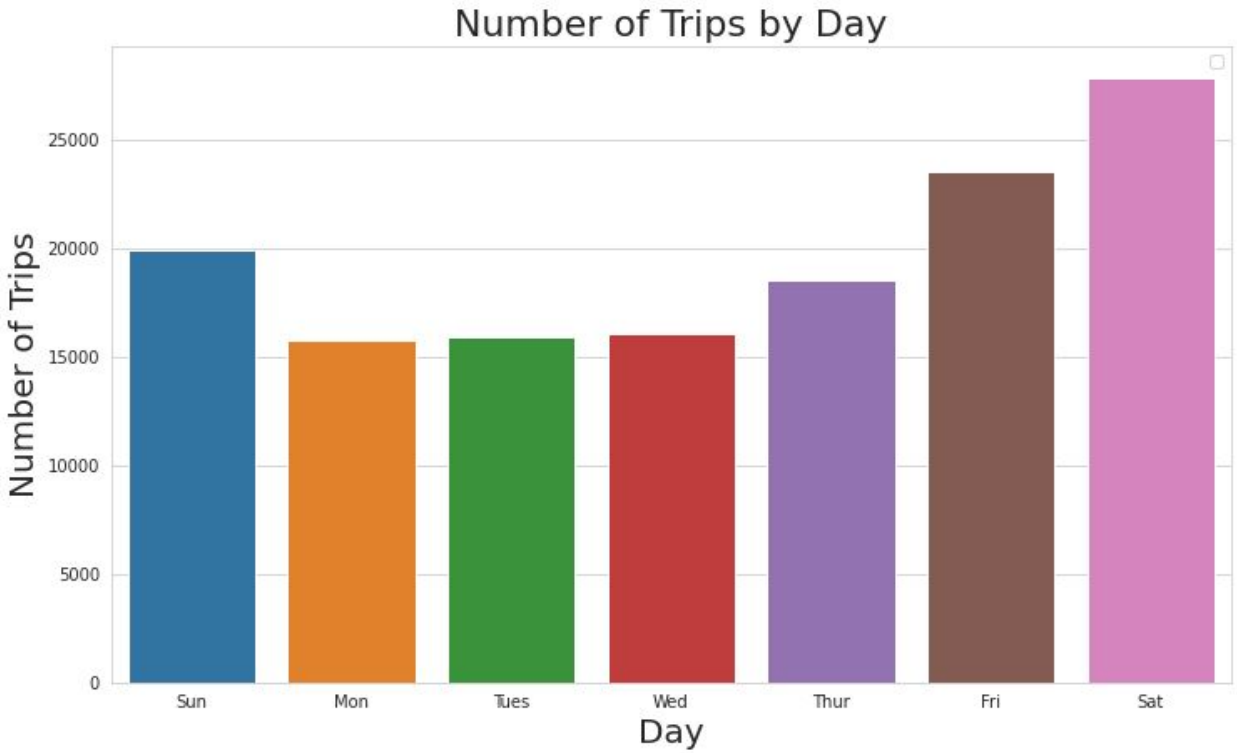
- Data preprocessing
  - Original mobility dataset: ~10MM rows of individual trip information
  - Cleaned and merged with weather data
- Data description
  - April 4, 2018 - November 16, 2020
  - 1450 rows, 23 columns
  - Merged weather + mobility data
    - Mobility data: Number of scooter/bicycle/moped trips, average trip duration and distance, date information
    - Weather data: Temperature, humidity, dew point, wind speed, pressure, precipitation

VISUALIZATION: MOBILITY DATA



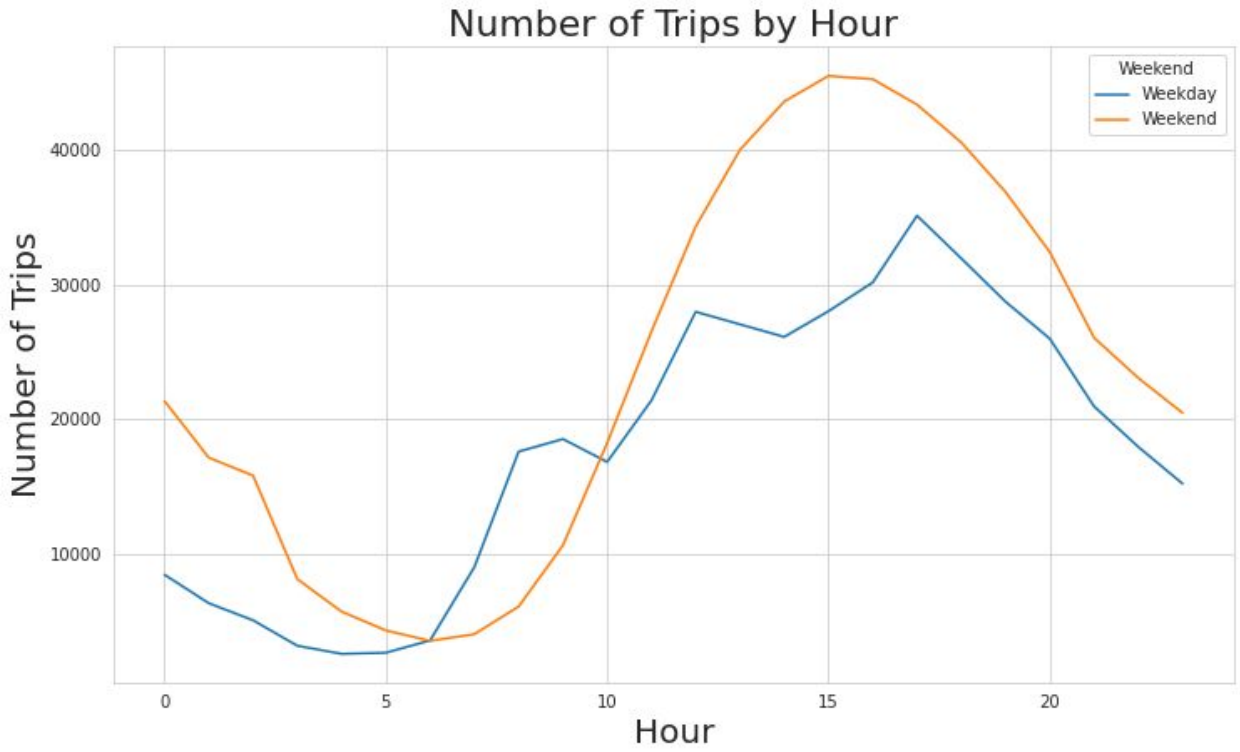
- Mopeds were introduced in late 2019
- Scooter rides are by far the most popular mode
- Seasonality effects are present
- Covid-19 sharply reduced number of trips from starting in March 2020

VISUALIZATION: MOBILITY DATA



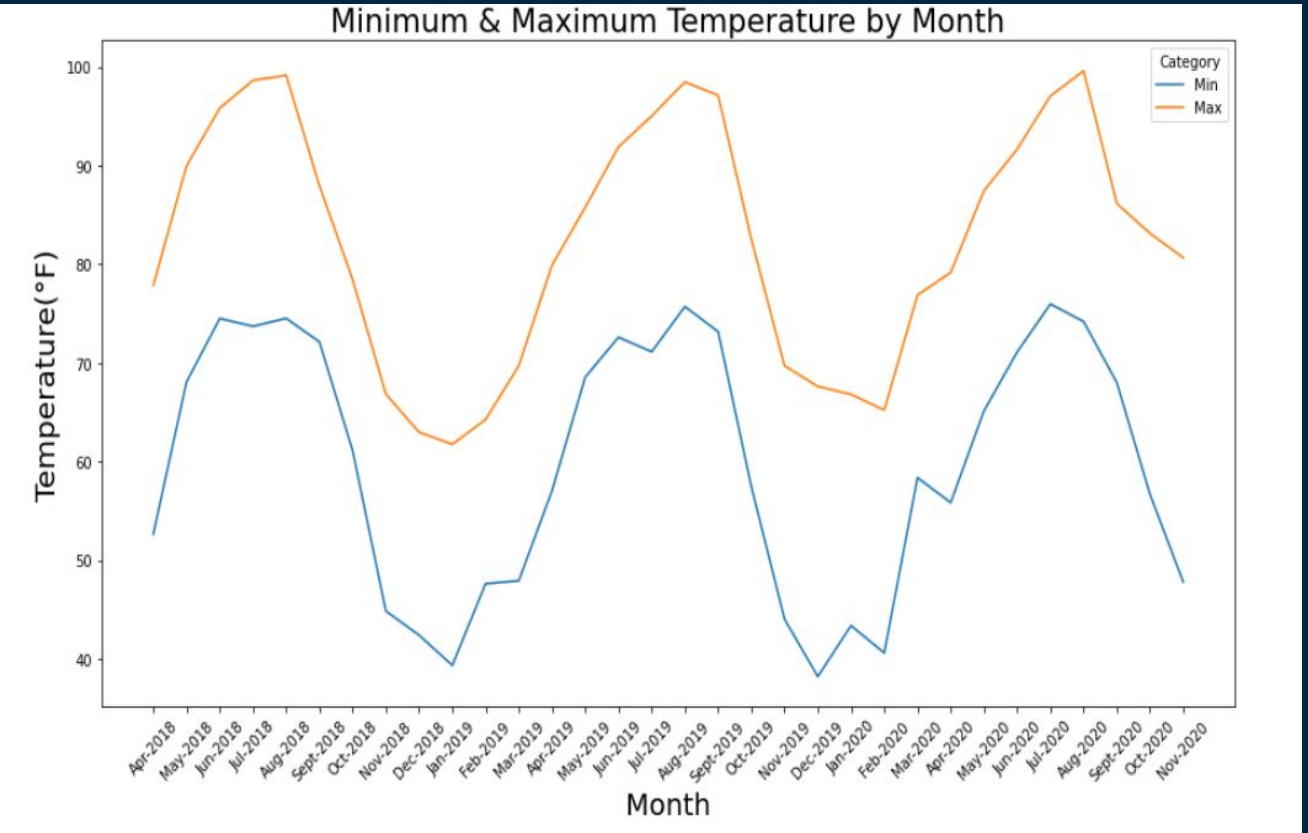
- The plot shows that shared mobility trips occur mostly on Friday, Saturday and Sunday. Particularly, Saturday shows the highest amount of trips

VISUALIZATION: MOBILITY DATA



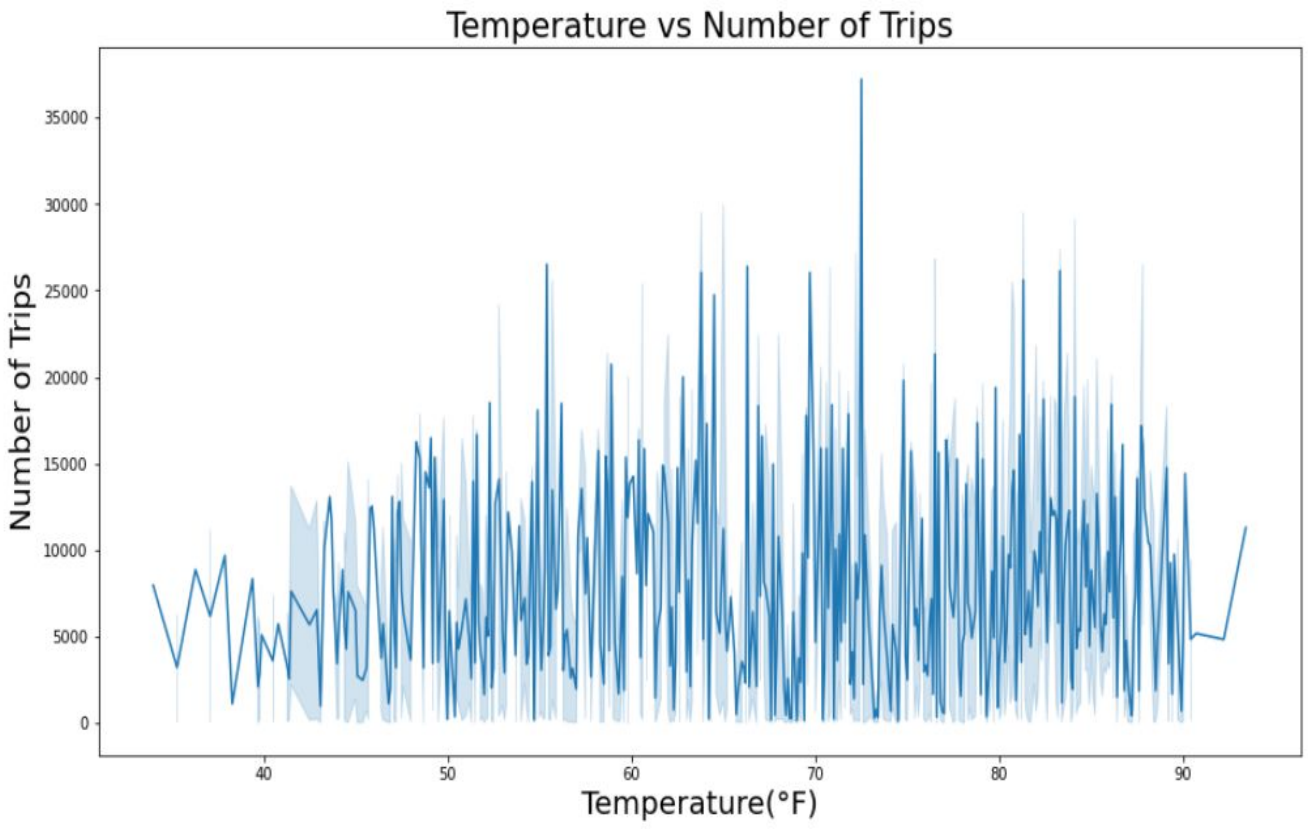
- On average, weekends have a greater number of trips than weekdays
- Weekends see a steady rise in number of trips throughout morning and early afternoon hours, peak around 3pm and decrease after.
- For weekdays, we see three peaks

VISUALIZATION: WEATHER



- Gathered daily weather data for the same period as the mobility data
- Highest temperature observed during June/July
- Lowest temperature during December/ January

VISUALIZATION: WEATHER VS. TRIPS



- Number of trips at a lower range for 40F<temp>90F
- No distinct correlation between temperature and number of trips for a normal temperature range

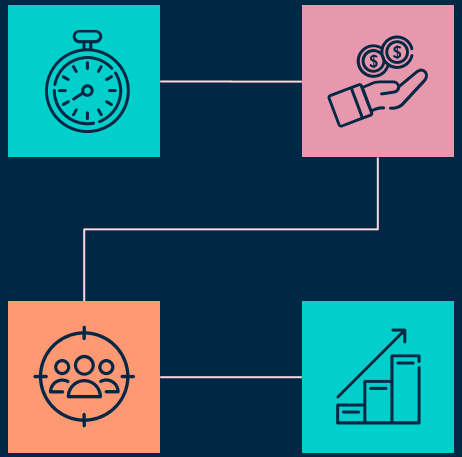
MODELING - What Didn't Work

CatBoost

Good for categorical values

Random Forests

Performs well for multi-class object detection



XGBoost

Performs well with unbalanced data

Decision Tree

Good when simplicity is important

MODELING - What Didn't Work

MODELING- What Worked

Linear Regression

Tuned with L1 regularization and limited weather data

One-Hot Encoding

Creating numerical features from weekdays and months

Time Series

Inclusion of the previous day's trip count was our most significant improvement

24,237,005

Final Regression MSE

3,771

Final Regression MAE

REFLECTIONS

- Interesting data insights and visualizations
- Achieved minimally accurate model for predicting vehicle usage
- Usefulness of the model
  - Can be used to predict usage and increase or decrease supply accordingly
  - Minimize cost of idle scooters
- Using some weather data improved our model
  - Holiday and event (ACL and SXSW) data could improve our model

THANK YOU!

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik