Shared Scooters - a Fad or the Future?

An Investigation of their Usage & Viability

Asher N. Meyers and Cesar Luciano Espitia
CUNY School of Professional Studies
Capstone Project: Data 698 Spring 2019
May 15, 2019

Contents

Introduction

- Shared Scooters as New Transport Mode
- A Viable Solution?
- Austin, Texas Dataset

Methodology / Data Collection

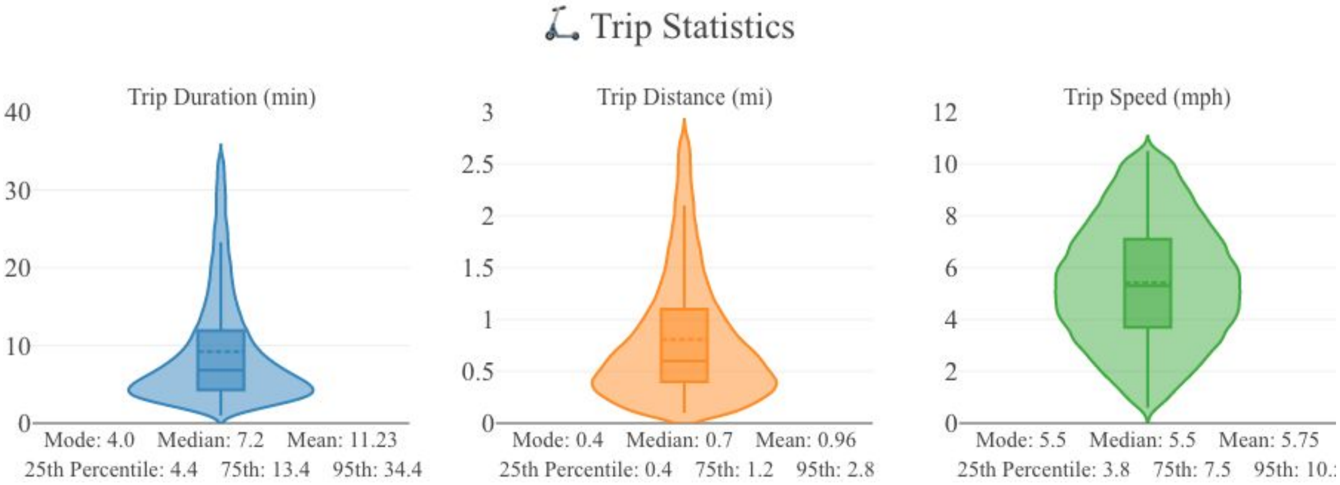
City of Austin Data Portal

- Data Set: Dockless Vehicle Trips
- oData calls to obtain a formatted CSV to import into Python (JupyterLab)
- (https://data.austintexas.gov/Transportation-and-Mobility/Dockless-Vehicle-Trips/7d8e-dm7r)

There were 3.7 million (from an original 4.4 million) filtered trip records over 13 months beginning in April 2018. The dataset contained each trip's origin, destination, duration, distance, and timestamp.

- Trips must be at least 1 minute, 0.1 miles, and 1 mile per hour
- less than 90 minutes, 25 miles, and 25 mph in average speed
    - Real world limit for range of 20 to 25 miles on a full charge

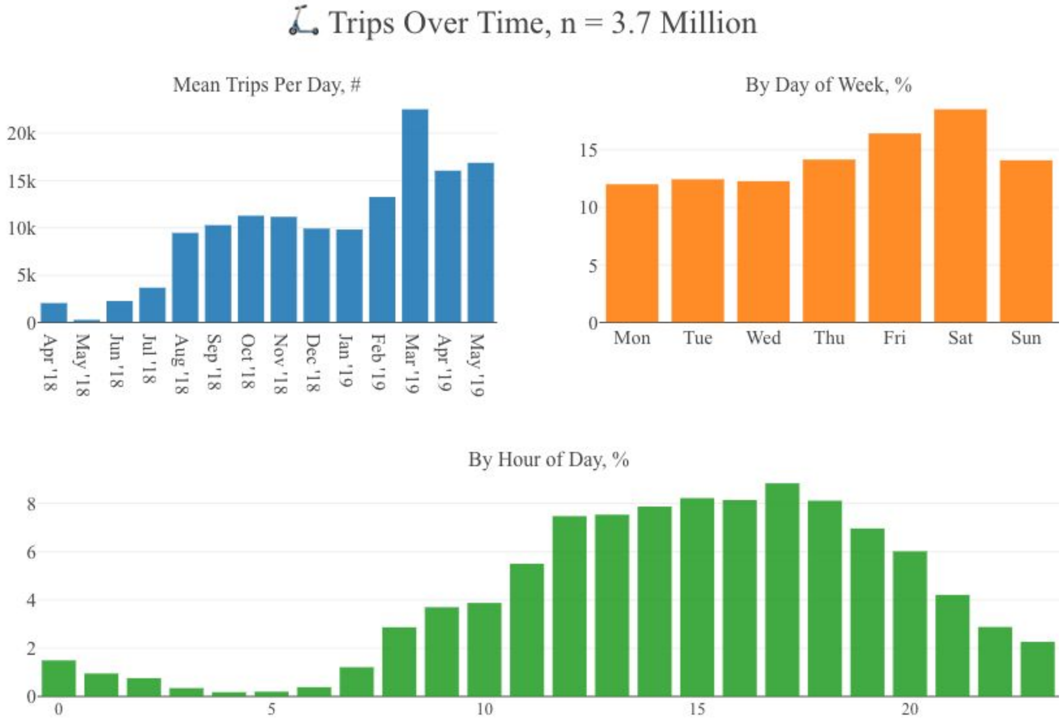
Trip Characteristics



- Trips range about 5-15 minutes long, 0.5-2 miles, and 4-8 miles per hour.
- Based on the typical industry pricing, the average trip would cost $2.80 for one mile, approximately seven times the average cost per mile of a car

Rides over Time

- After an initial low vehicle deployment, trips reached ~10,000 rides daily in the fall, and 16,000 in April 2019; a spike in March 2019 occurred thanks to a large festival
- Distribution of rides over time suggests non-commuter pattern - volume is highest during mid-day and on weekends



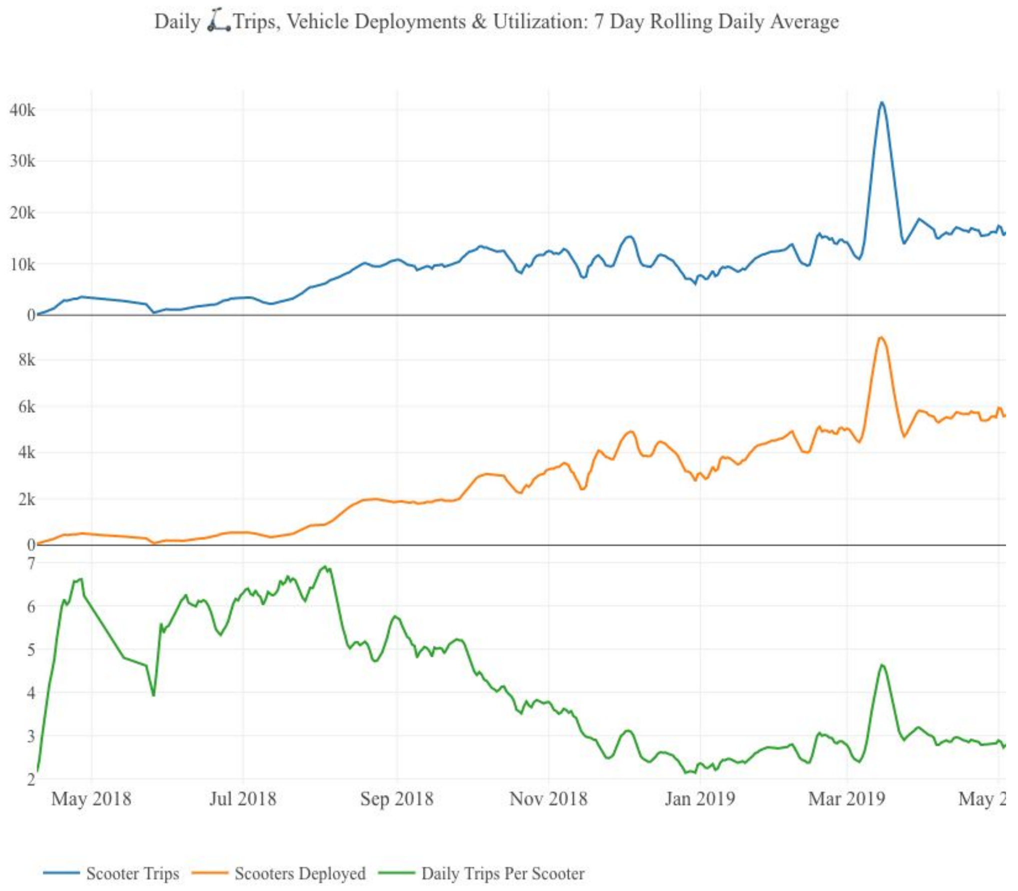
Deployment & Utilization

- The number of vehicles deployed (orange) has steadily increased at the cost of plummeting utilization* (green)

- Saturation of demand through abundant supply reached in August 2018

Note: data on vehicles available only if 1 or more trips per day were recorded

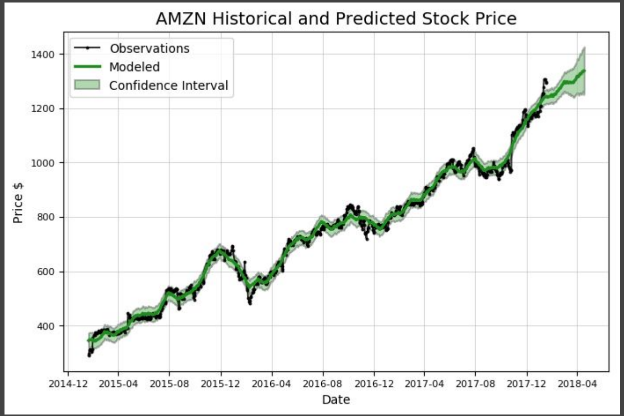
** Utilization: number of trips per vehicle per day*



Models: Forecasting

- Goal: help operators determine the daily supply needs based on demand
- Method: Basic neural Long Short-Term Model
- Packages: Keras module in Scikit

Powerful as it uses stored past information as the usage from yesterday is crucial to predicting the usage tomorrow.



Models: Forecasting

Model 1: Univariate Time-Series

- Data was aggregated to the day in order to be able to predict
- Used completed cases only
- Used days that had at least 100 trips
- Data was not normalized (time only input)

Model 2: Multivariate Time-Series

- Three (3) additional variables beyond the time input were used in second model
- Normalized data to account for order of magnitude in variables

Analysis Model 1: Univariate

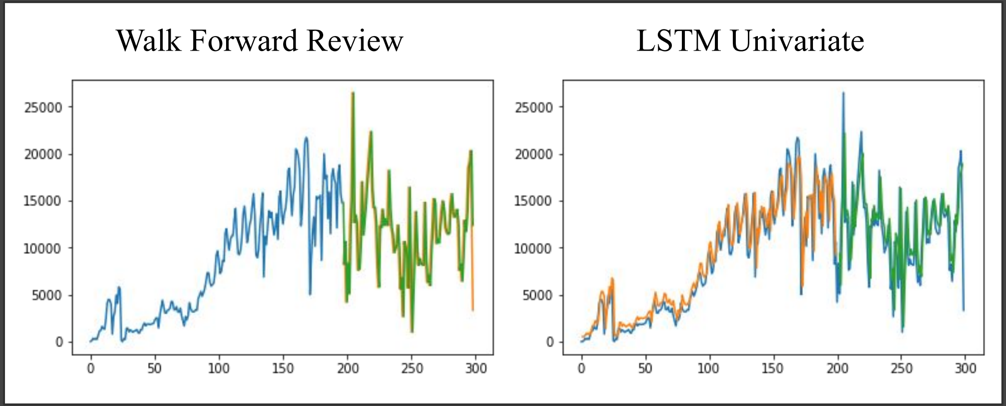
Model 1 was generated using a univariate LSTM model using the mean average method predict the trips one time unit forward.

All variables were excluded except for time (X) and Count of Trips (Y) as this is a time-series method that is similar to stock exchange prediction.

- model = Sequential()
  model.add(LSTM(4, input_shape=(1, look_back)))
  model.add(Dense(1))
  model.compile(loss='mean_squared_error', optimizer='adam')
  model.fit(trainX, trainY, epochs=100, batch_size=1, verbose=2)

- RMSE score of 3,617.91

- The prediction scores were in line with the actual scooter trips taken



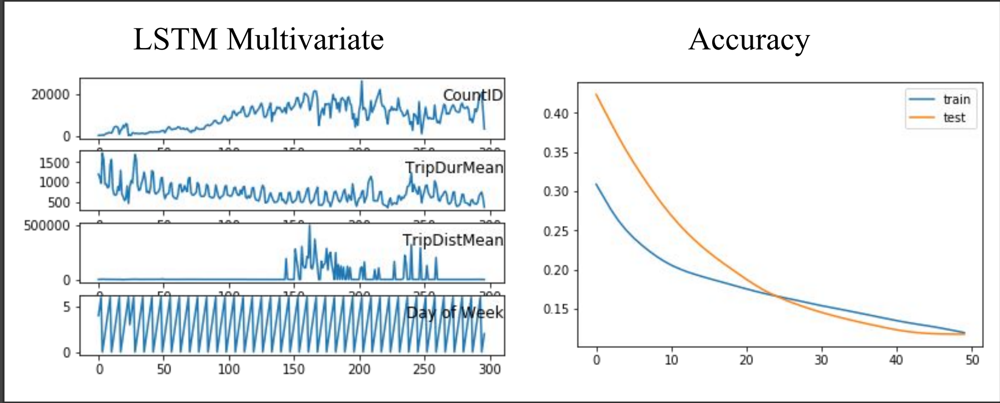
Metric	Step 298	Step 299	Step 300
Actual	15,752	15,047	9,244
Predicted	17,722	18,114	18,870
Delta (# / %)	1,970 / 13%	3,067 / 20%	9,626 / 104%

Analysis Model 2: Multivariate

MODEL 2 was generated using a multivariate LSTM model using the mean average error method to predict the trips one time unit forward.

Unlike Model 1, Model 2 has three (3) variables Trip Duration mean, Trip Distance mean and Day of Week, in addition to time and CountID (Count of Trips).

- model = Sequential()
  model2.add(LSTM(50, input_shape=(train_X.shape[1], train_X.shape[2])))
  model2.add(Dense(1))
  model2.compile(loss='mae', optimizer='adam')
  # fit network
  history = model2.fit(train_X, train_y, epochs=50, batch_size=72, validation_data=(test_X, test_y), verbose=2, shuffle=False)
- RMSE score of 3,919.99
- The prediction scores were in line with the actual scooter trips taken with less deltas than Model 1



Metric	Step 298	Step 299	Step 300
Actual	0.720	0.766	0.465
Predicted	0.558	0.427	0.342
Delta (# / %)	-0.162 / -0.22%	-0.338 / -0.44%	-0.123 / -0.26%

Model Selection

Model 2 had a better RMSE score vs Model 1 (3919 vs 3617)

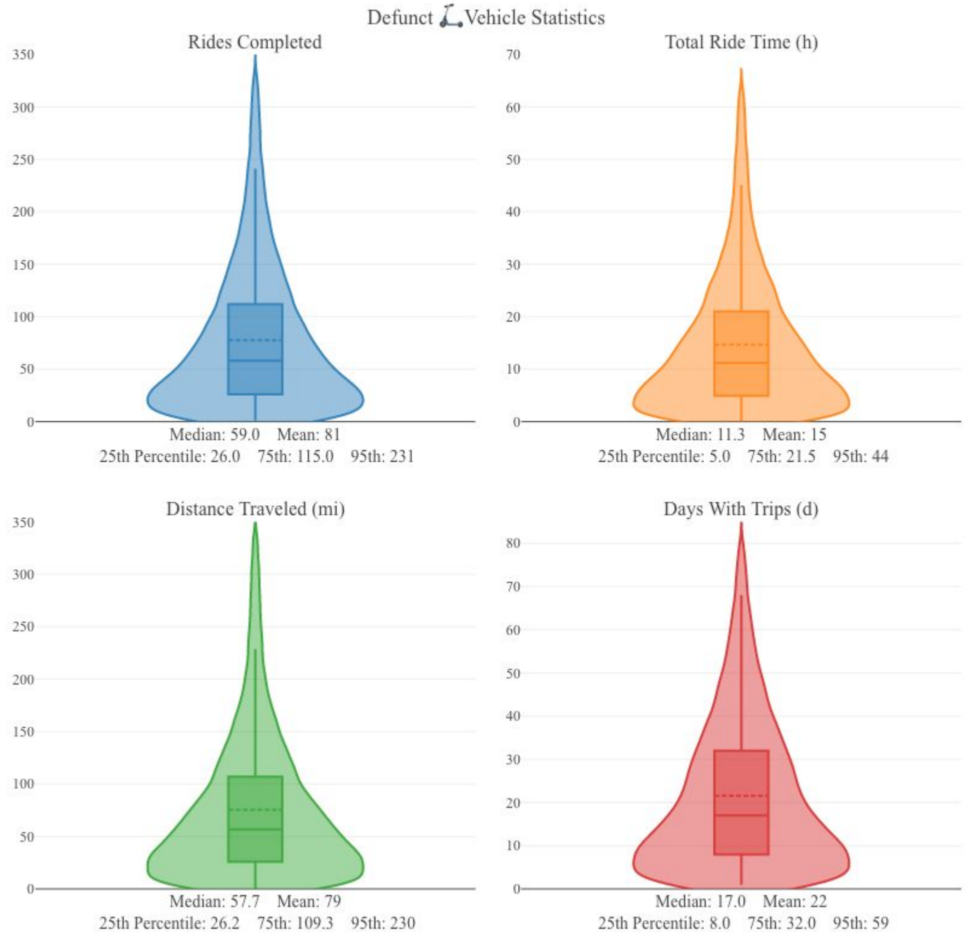
Model 2 displayed issues with correlation seen in the Loss Variance analysis

Model 1 chosen as the best model to predict daily scooter needs

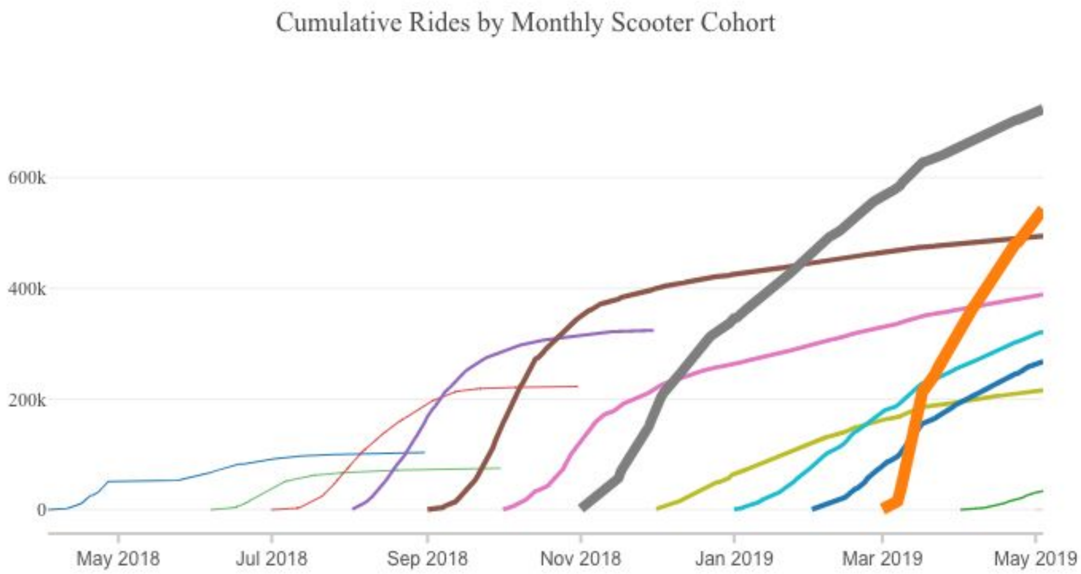
Useful for operators to determine the number of scooters to have available daily

Vehicle Longevity

- Scooters have a median longevity of 59 rides, 81 rides on average
- Days with 1 or more trips: Median of 17 days, mean of 22

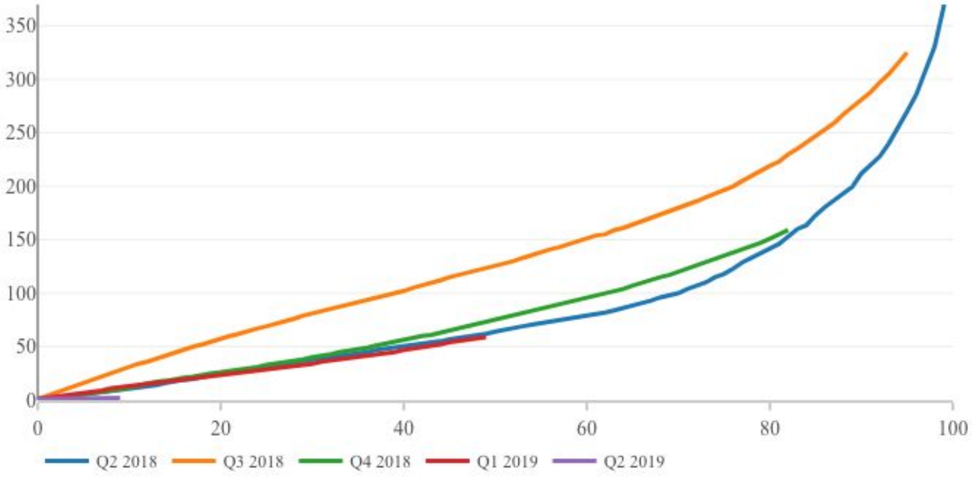


Vehicle Longevity: Cumulative Rides by Cohort



- Vehicles from April through August 2018 provided rides for less than four months each
- Subsequent cohorts, grouped by first month of deployment, have provided rides for longer
- Line thickness corresponds to the number of unique vehicles deployed (November ‘18 and March ‘19 saw deployments of 7,500 and 9,000 respectively)

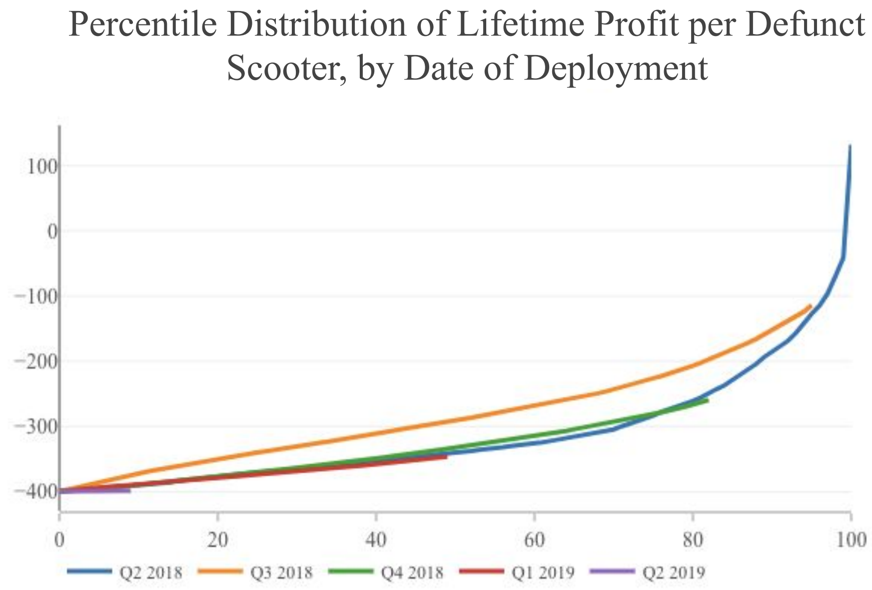
Percentile Distribution of Lifetime Rides Per Scooter by Cohort



- No significant improvement in longevity observed over time (defunct scooters shown)

Financial Viability

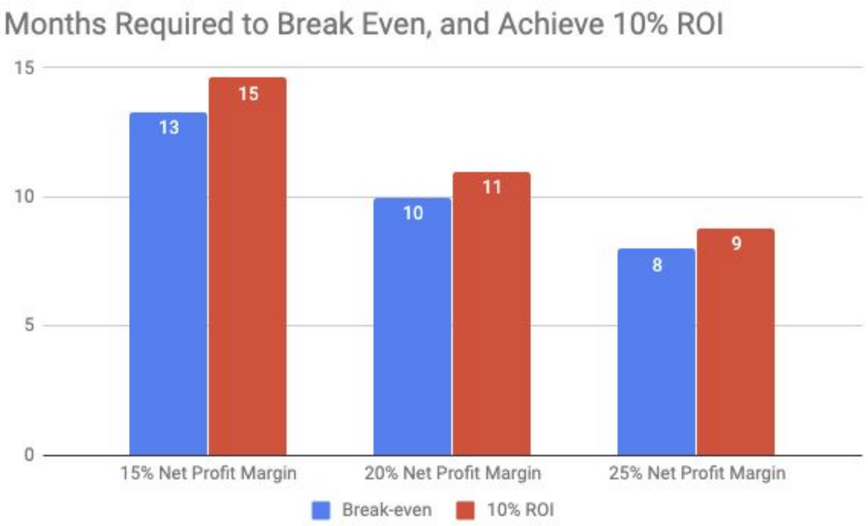
- Assumptions: $400 acquisition cost per scooter, a 33% gross unit profit margin, and $1 per ride + $0.15 per minute
- Lifecycle profit per vehicle had a median value of negative $340
- Weighted average loss per ride: $4, to date



Threshold of Viability

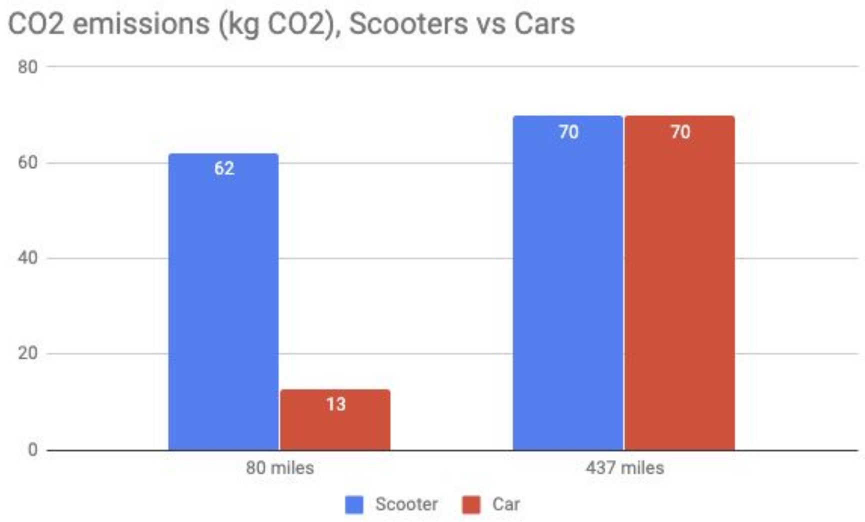
- Various scenarios of profitability were considered, to identify the longevity required to make scooters financially viable

- At prevailing prices, scooters would require a longevity of 8-13 months to break even, and 9-15 months to reach a 10% return on the scooter acquisition cost



Is it Green?

- An estimate of CO_2 emissions from scooter manufacturing and operation suggests emissions per mile from scooters are 5x higher, due to low vehicle longevity
- Scooters would have to reach 437 miles of lifespan mileage to attain parity in CO_2 emissions with a typical combustion engine car
- Most of the scooter emissions are due to the lithium-ion battery



Conclusion

Viability & Sustainability

- With a median lifespan under 100 miles, the scooters are yielding large financial losses
- The limited lifespans lead to more, not less, carbon dioxide emissions per mile of driving replaced

From a future ridership perspective, two (2) models were developed using the Long Short-Term Memory predictive models:

- a univariate model that only focuses on the output
- a multivariate model that takes into account take trip characteristics netted in overfitting
- After creating both models, it is clear that the univariate model is the best model for this exercise and reinforces the mantra that simplicity is best

Thank You