Shared Scooters - a Fad or the Future? An Investigation of their Usage & Viability in the Cities of North America

Asher Meyers and Cesar Espitia

Data 698. Capstone Project.
Advisor: Nasrin Khansari
March 26, 2019

[THIS PAGE LEFT INTENTIONALLY BLANK]

Table of Contents

List of Images and Tables

Abstract

This capstone project focused on the newest trend in urban transport – scooters. An open data set obtained from the city of Austin, Texas contains approximately 1 million records spanning 9 months, from 2017 to 2018, with one row for each trip taken. Variables of note include origin, destination, timestamps, distance traveled, and trip duration. The purpose of this capstone is to forecast future ridership, industry revenue and losses, and the aggregate consumption of vehicle hardware. The overall conclusion states

Keywords: scooters, data698, austin, capstone

PURPOSE

BACKGROUND

In the summer of 2017, a former Uber employee started a new transport service. Instead of people using their own cars, the company would rent electric kick scooters to the public. The scooters would be left in public view, and accessed by an app.

Travis VanderZanden started his scooter service in Santa Monica and Venice, two beaches on the doorstep of America's second biggest city. The scooters caught on, allowing VanderZanden to raise tens and then hundreds of millions of dollars to expand the service, known as Bird, to other cities. Other rivals, have arisen, including Lime, which appears to have eclipsed Bird slightly, in ridership.

The scooters have elicited both contempt and joy in great measure - much like any new form of transport. The charges against them range from putting pedestrians in harm's way, to injuring riders with unsafe hardware.

More poignantly, the shared bike services on which Bird et al. are based, have mostly disappeared from American cities, because scooters have proven more popular.

With Lime alleged to be spending $23 million monthly [citation needed], it's worth considering if rented scooters are financially viable, and if not, whether they can become so.

The scooters generate copious data, most of which is held in private by operators. However, the city of Austin, Texas has released trip data which can shed light on the performance of scooters to date. However, it should be noted that future vehicle designs and service features may ameliorate the issues presented here.

DATA EXPLORATION

Data Cleaning and Preparation: Filtering Out Questionable Trips

The dataset contains over 3.5 million trips. Many of those trips were found to be questionable, as either physically impossible, unlikely, or not indicative of a real life trip taken.

In reviewing the various trip statistics, such as duration and distance, bounds were set for the minimum and maximum acceptable values. Using interquartile range (IQR) ratios to identify outliers was considered, but transportation-specific thresholds were deemed more precise and reliable.

Trips must be at least 1 minute, 0.1 miles, and 1 mile per hour. They also must be less than 90 minutes, 25 miles, and 25 mph. The upper bounds correspond to the maximum limits of the scooters, which are capped at 15-18 mph on flat ground, 20-25 miles of range when fully charged. The lower bounds were more restrictive, resulting in 15% of raw trips removed, while an additional 1% of the original trips were removed for passing the

upper thresholds. The high volume of lower outliers (15%, approximately 600 thousand) is likely due to sensors detecting a trip where none occurred, or due to riders unlocking, but not quite riding, the scooter.

SUMMARY STATISTICS

The distribution of trip duration, distance, revenue and speed, from the filtered trips, are shown below. The statistics shown in text are from all filtered trips, while the graphic is from a 50,000 unit sample of those trips. (A sample was drawn to avoid long processing times for generating graphs.)

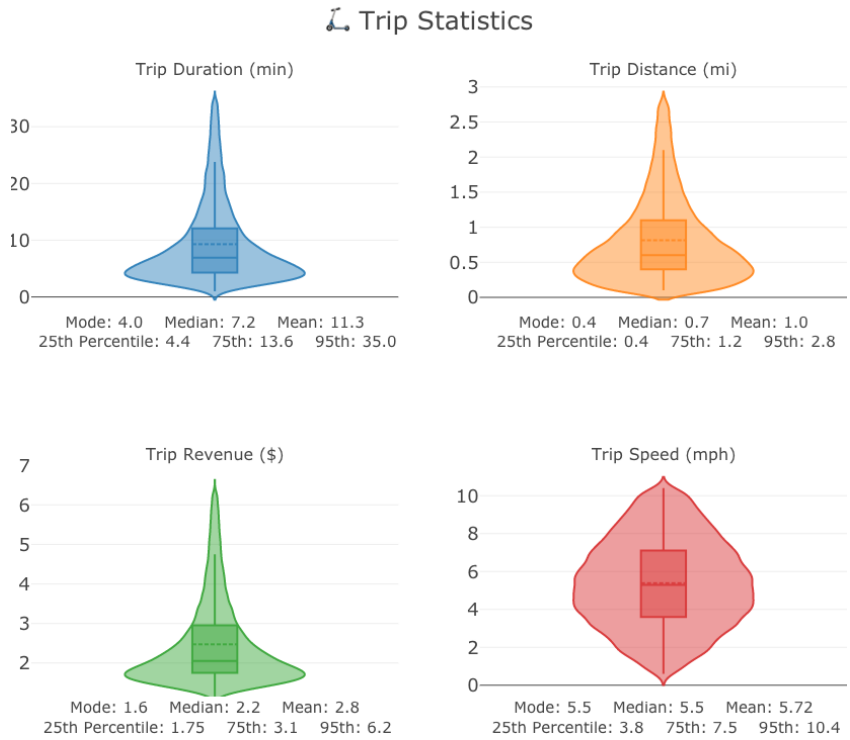


Image 1: Trip Statistics for Duration, Distance, Revenue and Speed.

Since scooters usually have the same pricing structure across operators, of $1 per ride + $0.15 a minute[1], the median ride fare amounts to $2.20, the average, $2.70. At a median $3.14 per mile, that is approximately seven times the average cost per mile for driving.[2] Some riders have received ~50% discounts, for having low incomes, but the data available does not identify the share of trips attributable to these riders.

We can also see how scooter use varies over time. For the initial pilot month of April in 2018, there were ~50k rides; scooters were removed for most of May, and then returned by June. By August, scooters were seeing ~300-350k rides a month, until February, when there were 365k rides, and then March, which saw 680k rides (estimated).

[1] Lime, Bird, Spin, Lyft, Razor have these rates; JUMP, which primarily offers ebikes, does not charge an unlock fee, but $0.15 per minute.
[2] $0.43 per mile; $8,427 average spent on cars per household, 2016; 53.8 miles traveled per household per day, p12

[Trips over time by month graph here]

Ridership increased modestly over the Fall months of September and October, 8.6% and 13% respectively, then falling 4.2%, likely due to colder weather, and 9.0% in December, due to finals and the academic semester ending, as well as even colder weather; January saw ridership similar to December, for similar reasons. By February, ridership increased 21% month over month, and 86% in March (estimated).

Notably, Austin is home to the University of Texas and its 50k students, whose classes run from the end of August until mid-December, and again from late January until mid-May.

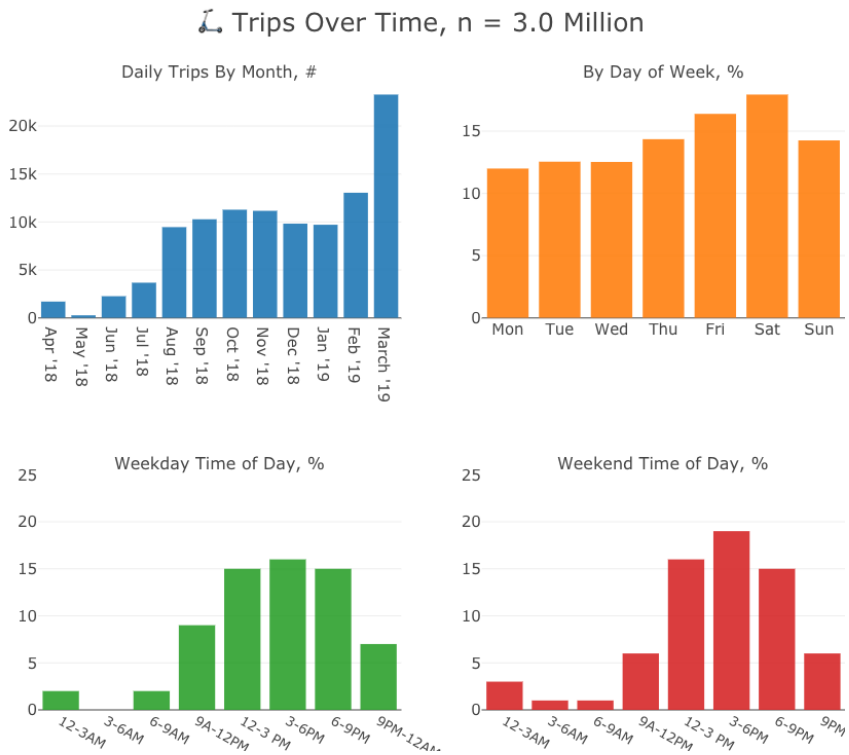


Image 2: Trips Over Time Last 12 Months.

Ridership is not only a weekend affair - on Monday through Wednesday, 12-13% of rides occur. Then rides increase sequentially from 14.4% on Thursday to 16.4% on Friday to a Saturday peak of 17.8% of rides on Saturday, then 14% on Sunday.

This suggests a baseline level of daily ridership, consisting of people using them to get to work on weekdays and/or personal matters throughout the week, plus recreational riding during the day and around nightlife.

During the work week[3], there were 11.6 rides per day for every trip at night, while for the weekend, there were 9.16 trips during the day for each scooter trip at night. (Austin's university students often do not have classes on Friday, and thus Thursday night is the beginning of the weekend.)

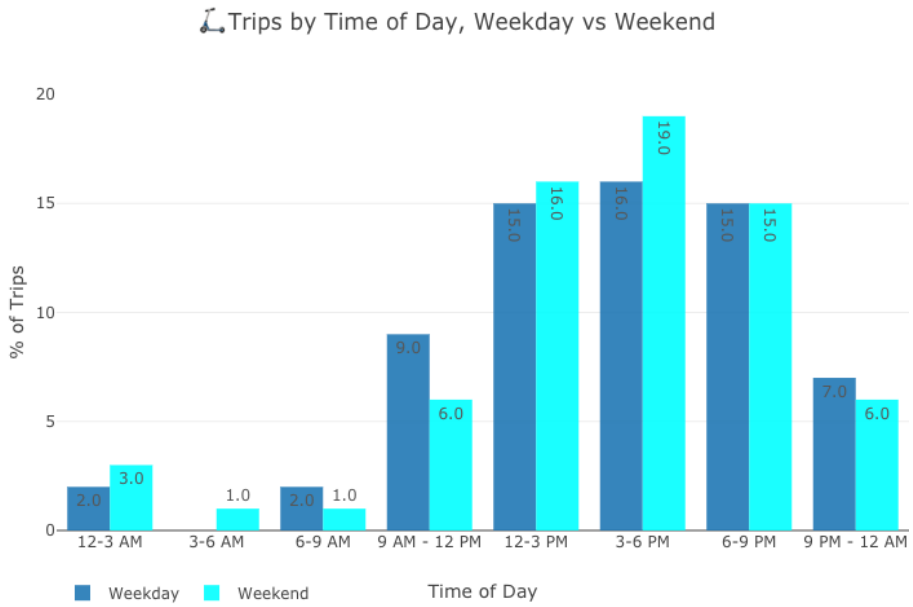


Image 3: Trips by Time of Day, Weekday vs Weekend.

NB: Would like to redefine weekends from Sat-Sun to Fri 6 PM - Sun 6 PM

Dataset Limitations: Estimating Daily Vehicle Deployments & Utilization

One limitation of the data is that vehicle counts can only be estimated, since a vehicle will not appear in the dataset on a given date, if it had no trips that day; the dataset is a record of trips, not of vehicles and their deployment status.

Thus, gleaning the number of vehicles out on a given day is necessarily an underestimate, because some number of vehicles will have zero rides that day.

Similarly, estimates are made of vehicle utilization - the number of rides per scooter per day; but these estimates are necessarily an overestimate, since they do not include scooters that had zero trips on a given date.

[3] The work week runs from Sunday 6 PM through Friday 6 PM, since, weekend behavior begins on Friday evening and ends on Sunday evening, as people prepare for work the next day.

Daily Trips, Deployments, and Utilization

Using a rolling average, with a window of 7 days, to smooth out variation between weekdays and weekends, scooter trips climbed steadily until August 2018, to about 10k daily. There was modest growth through the fall, with 10-12k trips daily (on a rolling basis); a sharp rise in trips occurred during Austin's major festival South by Southwest; in the festival's aftermath, daily trips has returned to trend, of about 12-13k trips a day.

South by Southwest (SXSW): Before, During, and After

In anticipation of swelling demand from the festival, operators put out more vehicles, going from 3-5k pre-festival, to a peak of 10k during the festival, down to 3-5k after festival-goers had left town.

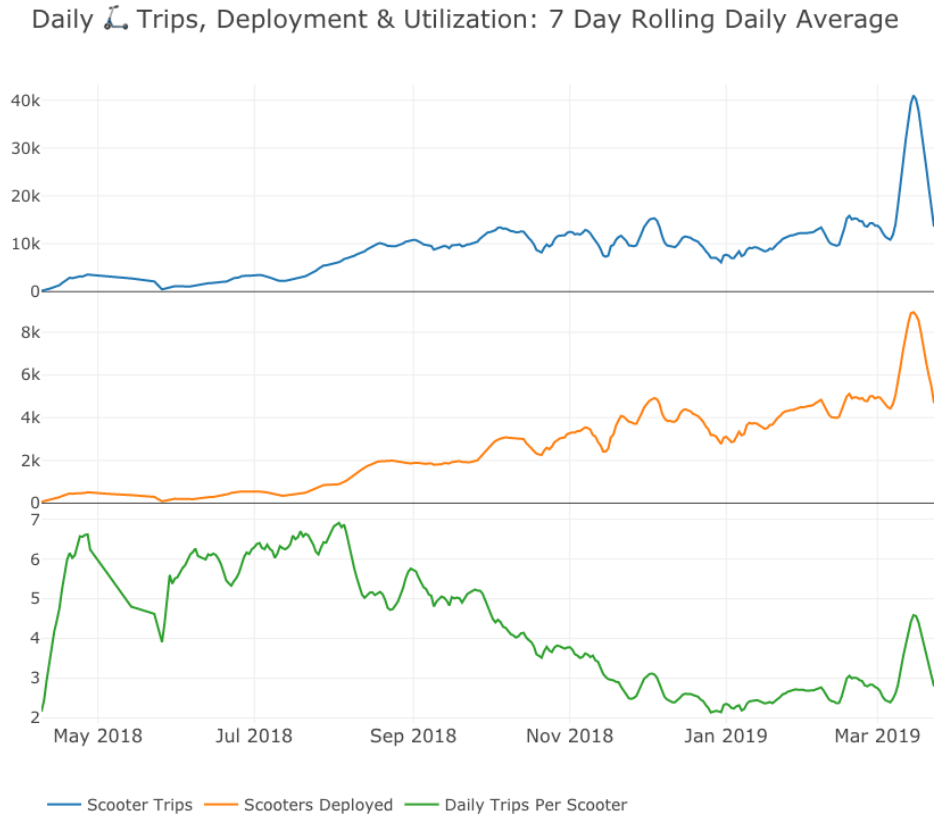


Image 4: Daily Scooter Trips, Deployment & Utilization. 7 Day Rolling Daily Average.

And indeed, trips did triple to quadruple, hitting 30-46k, or 39.7k rides per day. Utilization, or trips per scooter per day, increased 80%, from 2.5 in the week before to 4.5 for the SXSW period, March 8 - 16.

Performance per Vehicle Analysis

Beyond just analyzing trip statistics, we can also consider how vehicles performed, since the dataset includes persistent unique vehicle IDs, that allow us to track the vehicle from one trip to the next.

Over 95% of the time, if a scooter has not appeared in the trip dataset for two weeks, it will never return. Therefore, all scooters that have not been observed active within the last two weeks are assumed to be defunct. This pool of defunct scooters are then analyzed, to estimate the volume of usage a shared scooter will see over its life.

When viewed as violin plots, we see there are a number of vehicles failures early on in a vehicle's lifespan, as well as a number of scooters with very high longevity. This results in means that are quite a bit higher than the median, for given trip statistics.

The median scooter completes 63 rides over 11.9 hours of ride time, and 60 miles; it's trips are scattered over 18 days; for every 5 days where it has trips, it has 4 without, between its first and last recorded rides.

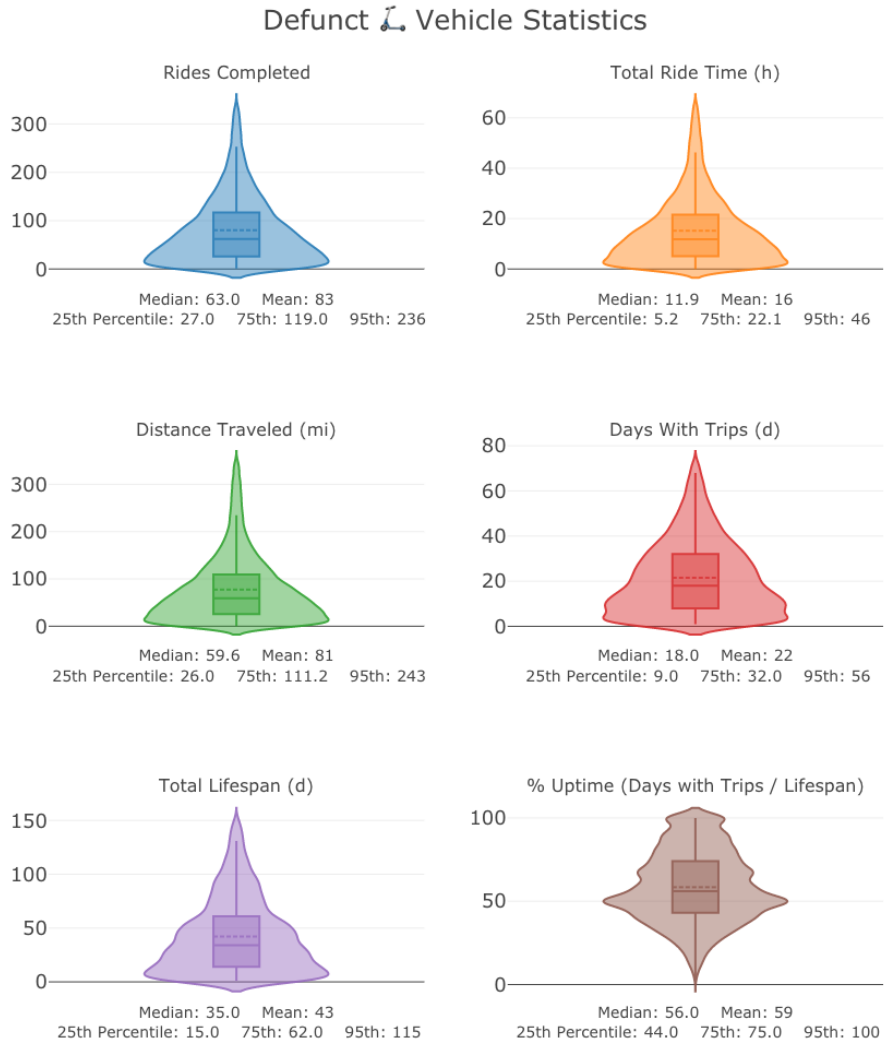


Image 5: Defunct Scooter Vehicle Statistics

We can also see how the population of scooters changes over time; in April through July, we see about 1,000 added per month, a net addition of a few hundred. August and September bring a net 1,500-1,600 each; aside

from November adding an astonishing net 5,000, the scooter population is stable until March, where over 7,000 scooters are added, on net, much of that due to the rollout for SXSW attendees. (According to the dataset, only four vehicles started in May).

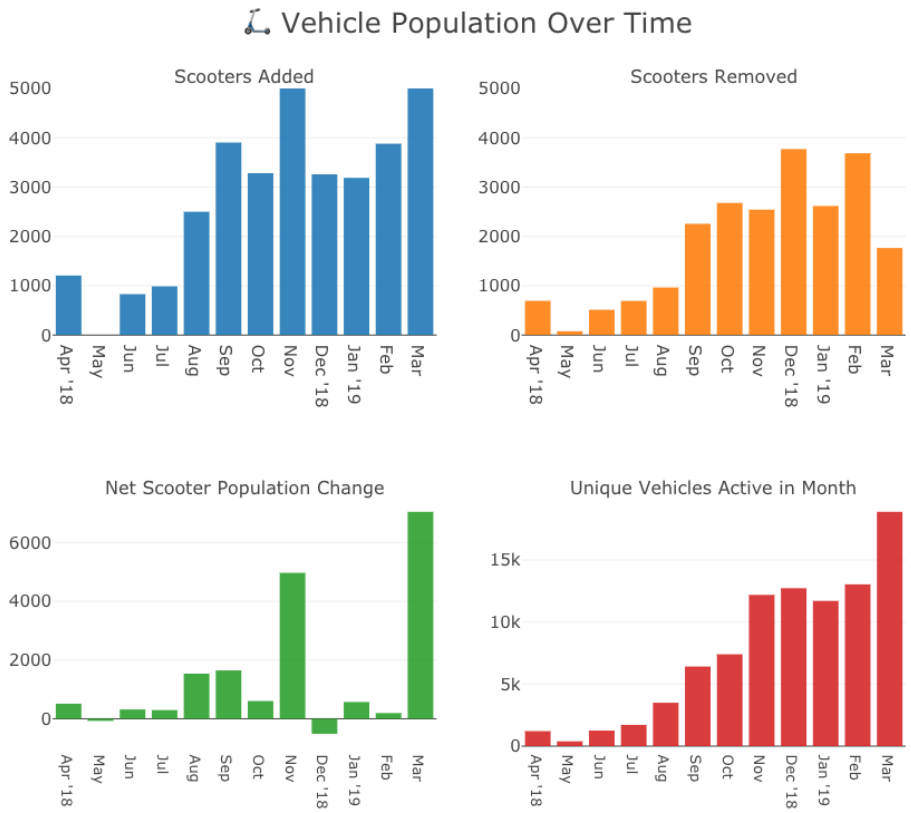


Image 6: Scooter Vehicle Population Over Time

Since we use median vehicle statistics to evaluate scooter lifecycle metrics, to evaluate changes over time in scooter performance, 50%+ of the vehicles must reach defunct status. The latest month to reach that threshold is December, though January is quite close, with 53% of vehicles still active - and the additional 3% are likely to be defunct anyhow, thanks to the 14-day grace period for classifying defunct vehicles.

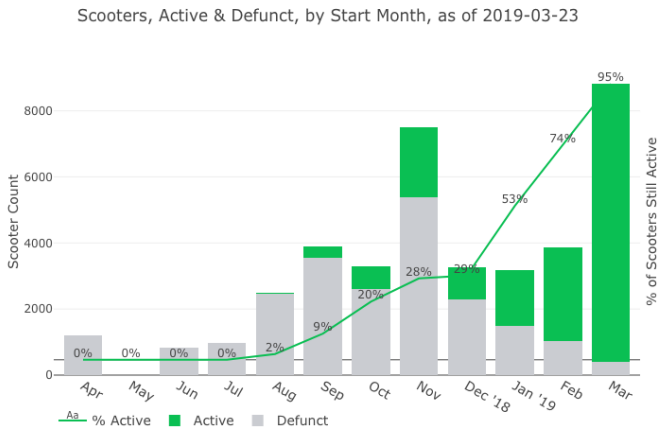


Image 7: Scooters, Active & Defunt by Start Month, Last 12 Months.

One might expect that with new vehicle technology, as kinks are worked out, vehicle longevity would increase over time. That has not occurred, with scooters lasting 25-30 days

What is causing scooters to only last 60 or so rides? There are two main contributors possible - the amount of time being ridden, and the number of days out on the streets. Put differently, is using the scooter wearing it down, or is putting the scooters out on the streets the issue, irrespective of use?

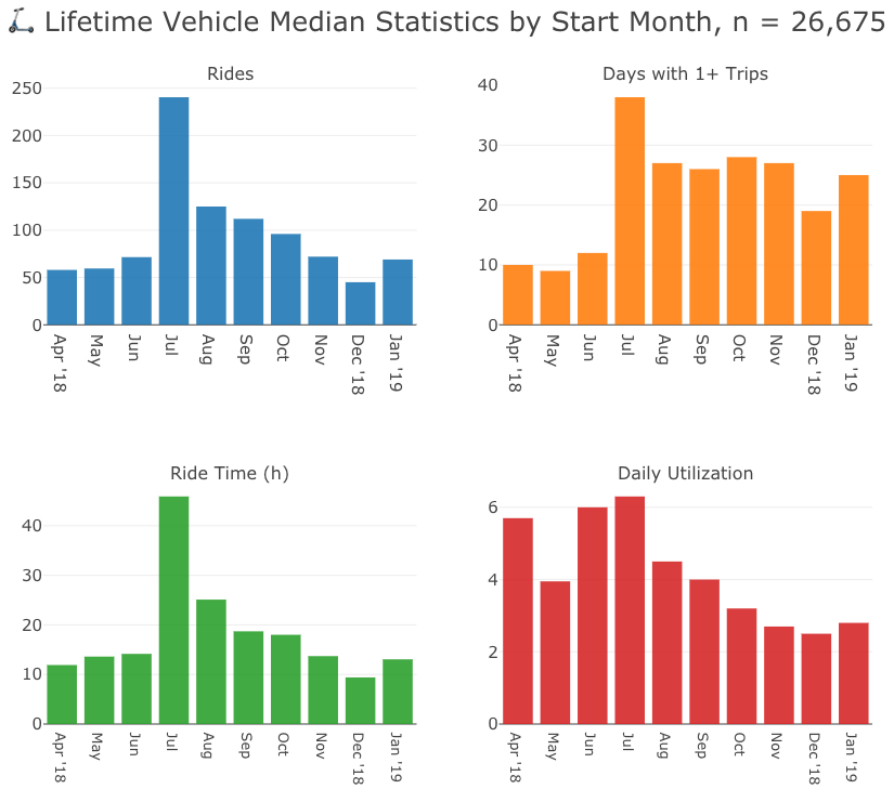


Image 8: Scooter LIfetime Vehicle Median Statistics by Start Month. Last 12 Months.

The data collected strongly suggests the latter. Scooters that were first deployed in July had by far the highest median lifetime rides of 241 per scooter. This was a dramatic departure from previous months seeing 50-72 median rides per scooter, and after, lifetime rides declining and bottoming out in December at 45 and rebounding to 69 for scooters 'born' in January.

At 38 days of active operation, July scooters had 13 more active days (+52%), than the average of 25 days in the ensuing months of August through January. A marked difference, to be sure, but much less than the +280% difference in rides (240.5 in July vs an average of 86.5 in August - December). Even as colder weather approached after the summer, the number of days active on the road held stable. That leaves utilization as the primary explanation, for why lifetime rides per scooter has plummeted.

One particular source of potential damage to scooters is the recharging process, which involves loosely shuttling the scooters in a vehicle, often a pickup truck or van, for recharging indoors. 'Gig workers' are often

responsible, and have little incentive to treat the vehicles with care, when they can earn more by working faster and hauling more vehicles.

However, utilization does not explain the first three months of April - June 2018, where utilization was high but lifetime rides were still low. A lack of operator experience, less mature hardware, as well as an initial wave of vandalism or reckless riding may be factors.

[Add in utilization only chart here]

July: The Golden Age of Longevity

July can be seen as a 'golden age' for Austin scooter longevity the favorable weather was kind to the vehicles and enticing to riders, and there was a high ratio of daily ridership to daily scooter deployment - 3,674 daily rides to 573 scooters deployed.

Weather likely played a factor as well - July scooters saw a median 38 days active, versus 20-30 days among scooters that came later. Notably, scooters before July were only active for a median of 9-12 days, though their utilization rates were approximately twice as high as scooters appearing in the fall.

Despite steadily increasing daily ride volumes, utilization has plummeted from 6 trips per vehicle per day, to under 3. And since vehicles apparently go defunct mostly from age and not usage (i.e. rides), declining utilization directly translates to lower lifetime rides per vehicle.

Models

This section describes the various types of models used for this capstone project. The models include 2 forecasting models to determine expected trips the following day and one financial model to determine the financial viability of this new transportation mode, in the hands of the private sector.

Forecasting

The purpose of this step is to take the modified dataset and begin exploring potential models that will be used to predict the expected trips in the City of Austin for the following day.

Time series forecasting is probably one of the more difficult models to create due to the nature of long-term short-term (LTSM) memory in the analysis. Unlike a moving average analysis, LTSM is highly valued in stock predictions where stockholders and trading companies can use these models to determine whether to buy or sell on any particular day. Similar to stocks, instead of determining whether to buy or sell, this model would help companies like Lime and Bird to be able to predict the number of units (scooters) to be deployed based upon the predicted demand.

A visual representation of the data is shown in the following image video (or copy the link into a browser) to the left to show the typical patterns over a day and the links visualized data to the right.

Origin Data Over Time Video

Link Visual

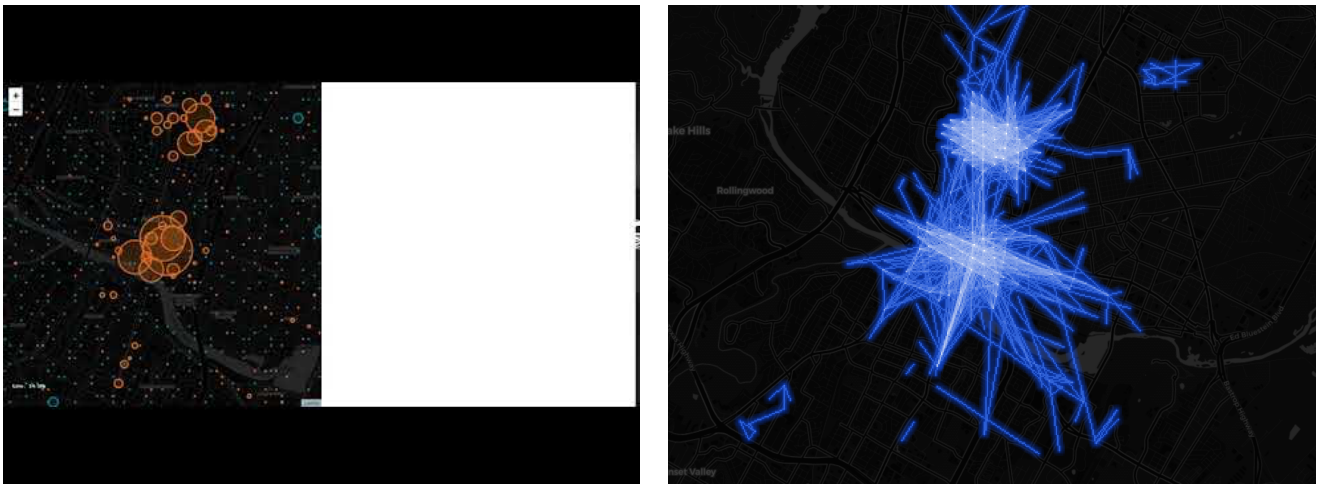


Image 9: Video Visualizer of Origin and Destination Pairs of Trip Data and Link Visual of Pairs (https://www.youtube.com/watch?v=VnsktKfSqUE)

As can be seen most of the trip is centralized around the main east-west strip and around the University of Texas at Austin. The following information describes the 3 models built for this step and the relevant analysis to provide reasons for model selection in the next step.

Model 1: Univariate LTSM Model

The perfect place to start is to use a is to use the univariate time step model which takes in the x variable of time normalized to units and the y variable is the number of scooter trips in a day.

The sample python script below that takes test and train data to make predictions is:

```
# create and fit the LSTM network
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)
```

The model was created using the LTSM model function within the keras package and had 1 shift in time step and was done over 100 simulations for each chunk of data. The full script can be found in Appendix B.

Image 9 below shows the results of the univariate model testing.

For Model 1, the train score was 2174.30 RMSE and the test score was 3617.91 RMSE. This means that the average RMSE was 50% higher in the test data set which is a good starting point for our predictions.

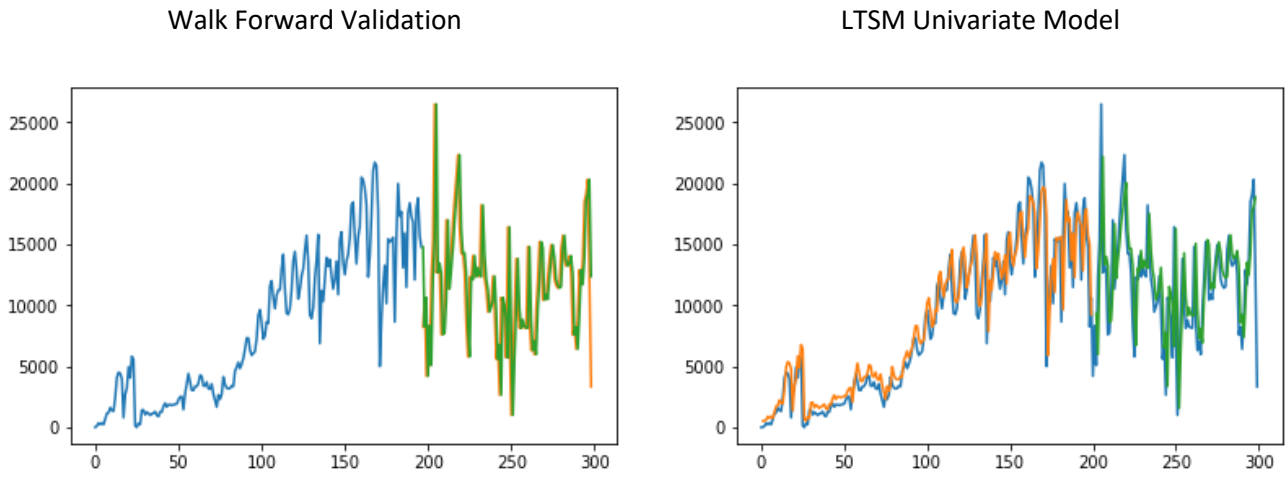


Image 10: LTSM Univariate Model.

Table 1: LTSM Last 3 Time Steps Comparison

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%

With the above graphs, RMSE and time step comparisons, we can tell that Model 1 is a good predictor for short predictions as the expected number of trips for the region are within range. This behavior is sensible as the prior days seem to have been a spike in season due to an event which the model would self regulate again once a few more days are in the short-term memory of the model.

Model 2: Multivariate LTSM Model

The next place to explore is to determine if more variables would allow the model lower its RMSE without overfitting due to collinearity or other typical traps in model creation. For Model 2, Y total variables will be

used in addition to time normalized to units for this analysis; the y variable stays the same as Model 1 and is the number of scooter trips in a day.

The sample python script below that takes test and train data to make predictions is:

```
# create and fit the LSTM network
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)
```

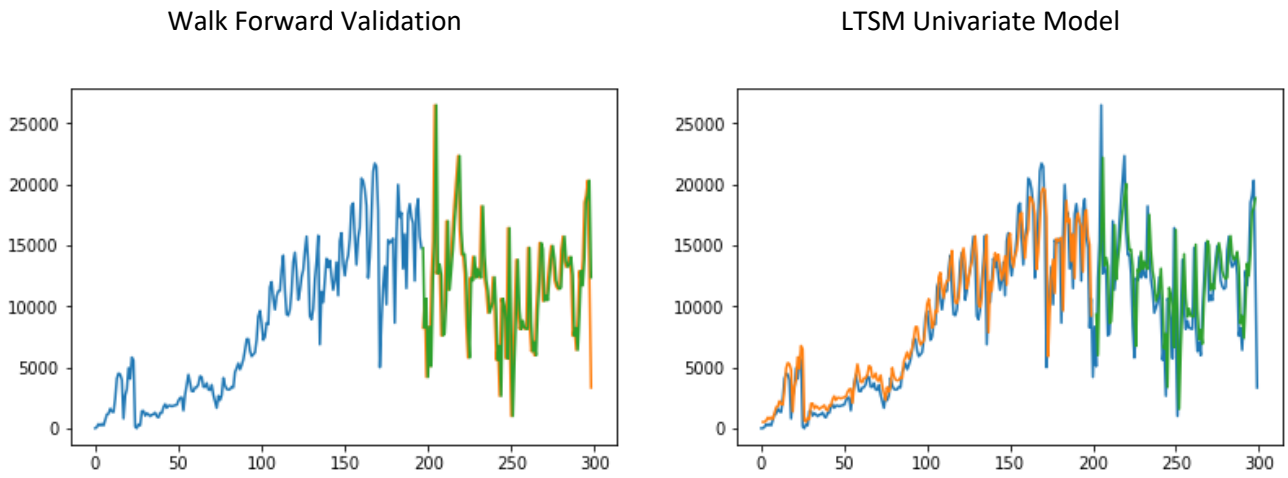


Image 11: LTSM Multivariate Model.

Table 2: LTSM Last 3 Time Steps Comparison

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%

With the above graphs, RMSE and time step comparisons, we can tell that Model 1 is a good predictor for short predictions as the expected number of trips for the region are within range. This behavior is sensible as the prior days seem to have been a spike in season due to an event which the model would self regulate again once a few more days are in the short-term memory of the model.

Financial Viability

According to an article in The Information [see Drive folder for copy], leading scooter sharing firm Bird optimistically projected in a May 2018 report, that 33% gross margins would soon be achieved, versus the 19% net margins it was already seeing. Gross margin refers to the revenue left after accounting for the 'cost of goods sold (COGS)', the direct costs associated with a ride such as charging, credit card processing, customer service and insurance.

Therefore, a $3 ride would have $1 in gross profit. The 33% margin figure was used for our estimates of costs per ride.

The Information also reported Bird vehicles cost $551 each; while the current price is apt to have changed - perhaps downwards thanks to increased orders obtaining bulk discounts, and perhaps upwards due to a new generation of sturdier, hardened vehicles - this figure was retained for estimating financial results and is used for all operators.

As before, results are apt to change over time, as new models are introduced, so we have chosen to display financial results by scooter start month, which will reflect improvements, if any, to vehicle financials.

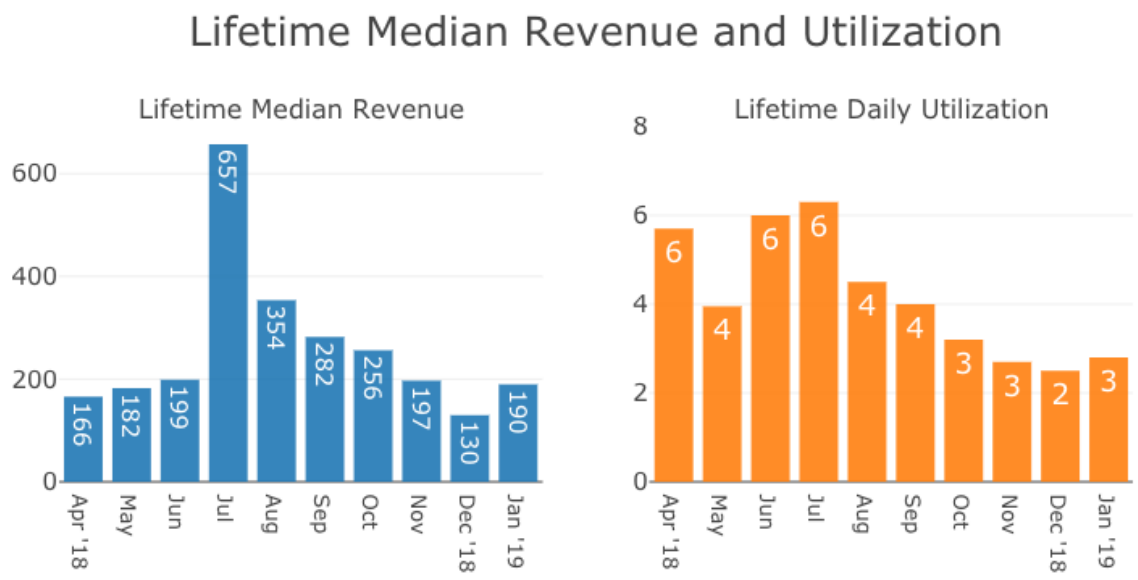


Image 10: Vehicle Financial Statistics by Scooter Start Month. Last 12 Months.

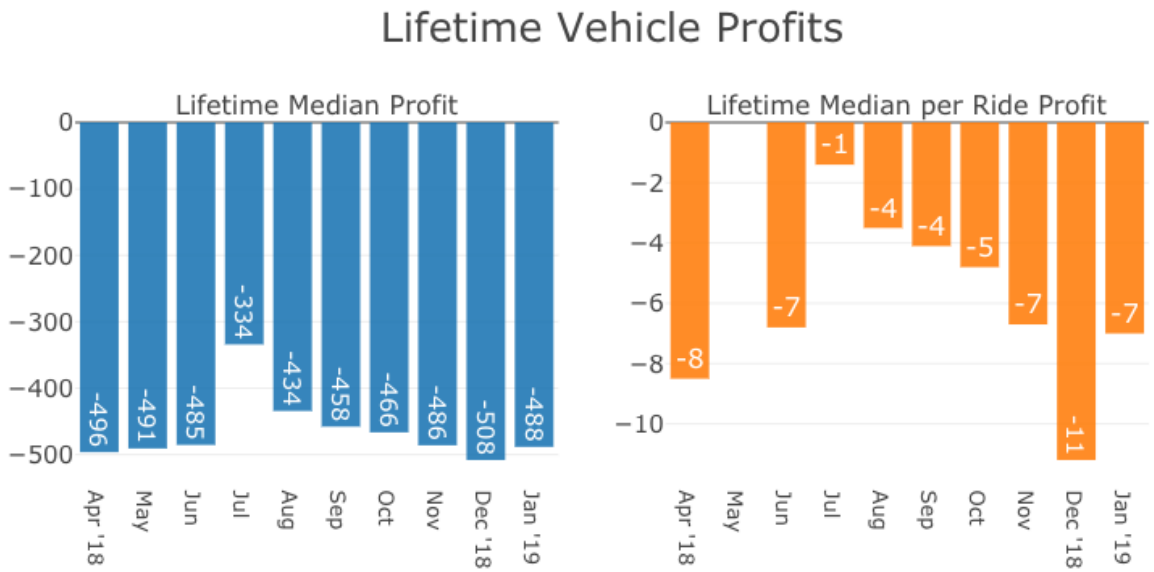
When considering scooters by start month, revenue has ranged from $130, for scooters first released in December, to $657 for July. Since July, there has been a swift decline in lifetime revenue. As discussed earlier, this is primarily due to sinking utilization, which went below 3 in November and has not recovered since, aside from a temporary respite during SXSW.

Losses, at Scale

With the assumed 33% net profit margins and a $551 acquisition cost per vehicle, vehicles are seeing losses of $300-500 lost per vehicle. Even in July, where vehicles saw their highest use by far, $334 was lost per vehicle. Across all defunct scooters, the average loss per ride is $5.74. Total losses to date are $10.8 m

The true profit levels are likely even lower: gross margin figures don't account for costs that aren't directly tied to a ride, such as transporting the scooters to market, developing the app, or staffing the company office in California, where most are based.[4]

Uber's JUMP does not charge a $1 unlock fee; all the operators are required to give discounts to low income users; for instance, Lime offers 50% off unlock and per minute fees, while Bird waives the $1 unlock fees. It is not known how many rides these users take, or the overall impact on revenue earned.



Fifty days apart, Lime announced that it had done 26 million, then 34 million rides, translating to an average of 160 thousand daily rides. At $5.74 a ride, that amounts to $918k in losses daily, or about $28M per month, without considering central staffing and other overhead costs.

If Lime comprises 40% of scooter rides among US based firms (but including markets abroad, where they are present), industry-wide losses would amount to $70m monthly.

The Path to Viability

With an average fare of $2.77 per trip and a gross profit margin of 33%, it would take just over 600 trips to earn enough gross profit to pay for the vehicle's $551 acquisition cost. With 3 trips per day, the scooters would have to survive 200 days of active use, about 6.5 months, nearly ten times the current median lifespan of 21 days among scooters that began operation in 2018.

[4] Lime, Lyft and Uber are based in San Francisco, while Bird and Razor are based in Southern California

However, the true longevity needed to break even is likely higher - some rides will be taken by low-income riders with discounted pricing, and general overhead, logistics expenses are not included in the 67% unit costs. And, a more ruggedized vehicle with greater longevity will likely cost more.

Assuming an additional 25% for other expenses, and a 25% premium for a ruggedized vehicle, translates to 300 days of active use required to break even. With a moderate return on investment target of 10%, vehicles would need 330 days, or 11 months, of active use.

To date, scooters have approximately 60% uptime - they had at least 1 ride per day, for 60% of the dates between their first and last recorded rides, inclusive.

By comparison, a first class ticket from New York to London, at $5,407 for a 3,440 mile flight, or $1.57 per mile[5]. Lime has said that 34% of riders reported income under $50,000, compared to a median US wage of $31k (2016) and $61k in household income (2017).

Given that the average American spends approximately 20% of household income on transport, and travels on the order of 10,000 miles a year, relying on scooters to provide even a moderate fraction of their total annual mileage would send them into penury.

Lime, for one, reported 3.4 lifetime trips per customer in February 2019[6]; their growth has slowed with the onset of winter, and their scooter service is approximately a year old. In cities like Waterloo[7] and Kansas City[8], trips have averaged one to two rides per user per month. In the most recent month available, December, Charlotte reports 8% of existing riders took 4-8 trips in the month, and <2% took more.

Scooters have been framed as an affordable alternative to a cab or ridehail (i.e. Uber or Lyft in the US), and indeed they are, but primarily for short trips taken alone - and these rival services are responsible for only 1.7% of urban travel nationwide. Replacing or outdoing cabs would still have only a modest impact on total urban car miles traveled.

There are slightly more trips proportionally on the weekends, with 14-18% of rides happening on Saturday and Sunday each, while 12-16% of rides happen Monday through Friday

[5] Hoffower, H. (2018, May 28). *8 Reasons first-class airfare is so expensive.* Retrieved from https://www.businessinsider.com/why-first-class-plane-tickets-are-expensive-2018-5

[6] Marinova, P. (2019, Feb 7). *Lime Investor Sarah Smith: It's 'Inevitable' That E-Scooters Are Coming to Every Major Market.* Retreived from http://fortune.com/2019/02/07/lime-funding-sarah-smith-bain-capital-ventures/

[7] Lam, P. (2018, Dec 4). *Lime e-scooter pilot in Waterloo sees more than 18,000 trips in 9 weeks. Retrieved from* https://www.cbc.ca/news/canada/kitchener-waterloo/lime-e-scooter-pilot-in-waterloo-sees-more-than-18-000-trips-in-9-weeks-1.4930584

[8] Harper, R. (2018, Dec 11). *Lime Scooters pulled off Kansas City streets (for now).* Retrieved from https://www.kctv5.com/news/lime-scooters-pulled-off-kansas-city-streets-for-now/article_8a87c11e-fdc1-11e8-adc8-e7783830346d.html

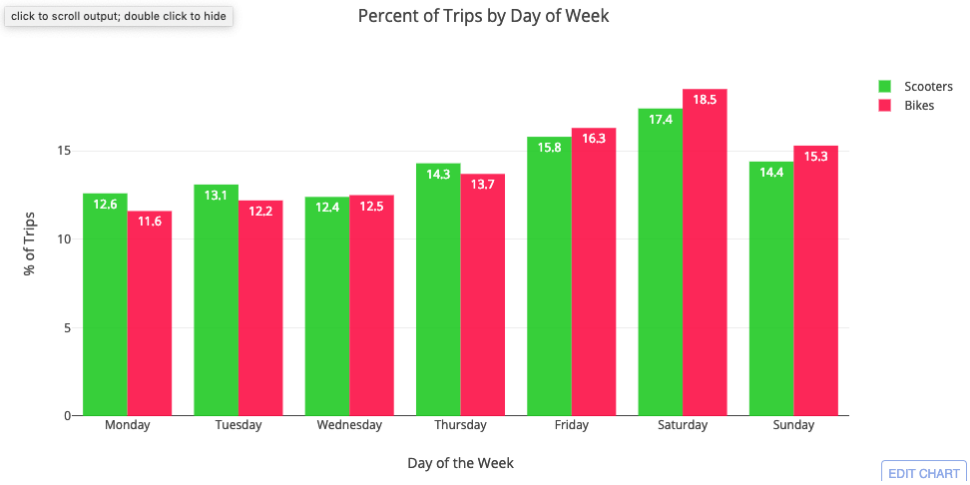


Image 11: Percent of Trips by Day of Week. Last 12 Months.

During the week, rides peak in the early afternoon to early evening, with a fair amount of rides happening in the morning and mid-evening; on the weekends, there are fewer morning rides and more afternoon rides, and a slight bump in late night rides.

Compared to an average trip distance of XXXX, riders don't show much tolerance for long scooter rides.

CONCLUSION

APPENDIX A

APPENDIX B

Forecasting Python Models .PY file