

# Citi Bike - A Deep Sea Dive

Team: Data Science For the Schwinn



# Outline

- Background
- Business Growth and Operations
- City Comparisons
- Revenue - Time Series Forecasting (CoVID-19 Considerations)
- Sponsorship and Advertisement
- Business Recommendations and Conclusion

Background

# Objective

Extract insight from Citi Bike business operations by leveraging publicly available data to inform the creation of a successful bike share program in another city.



# Milestones & Facts about Citi Bike NYC

**May 2013** - Citi Bike launches with 6,000 bikes and plans for rapid expansion. In first year, there were more than 100,000 registered members.

**October 2014** - Reorganized company (under Motivate), increased funding (10-year extension of Citigroup sponsorship, Mastercard), plans for expansion of stations and bike fleet with

**August 2015** - 140 dock stations added

**August 2016** - 140 dock stations added, increase number of bikes to 10,000.

**September 2017** - 142 stations added, increase number of bikes to 12,000

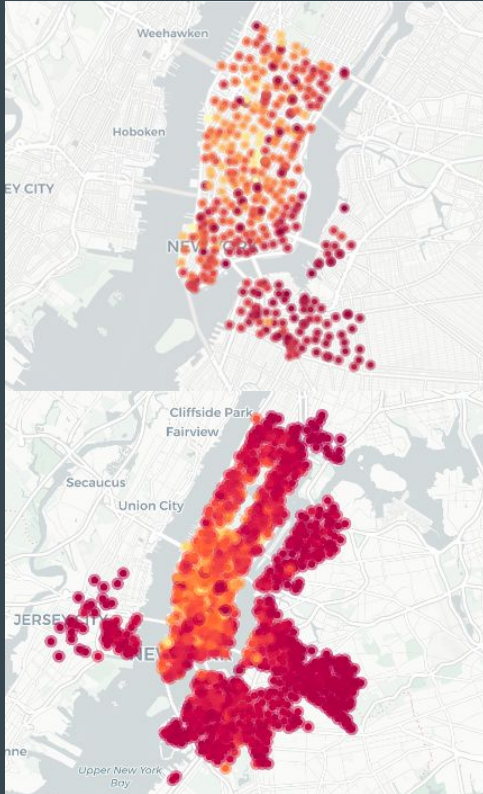
**July 2018** - Lyft acquires Motivate, and announces a five-year, US\$100 million plan to expand the bike-share system

# Datasets

- Daily Trip Data
  - Trip duration, rider type, start station, end station, etc.
  - 99 million observations
- Open Bus Data
  - How empty or full a station is
- Monthly Operating Reports
  - High-level estimates of revenue and operations
- Other Bikeshare Data
  - Washington D.C., San Francisco

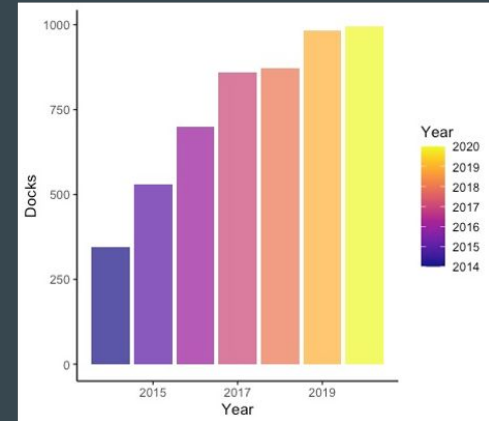
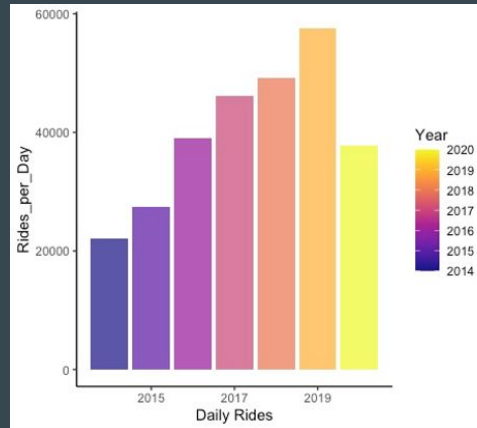
# Business Growth and Operations

# Business Growth Overview



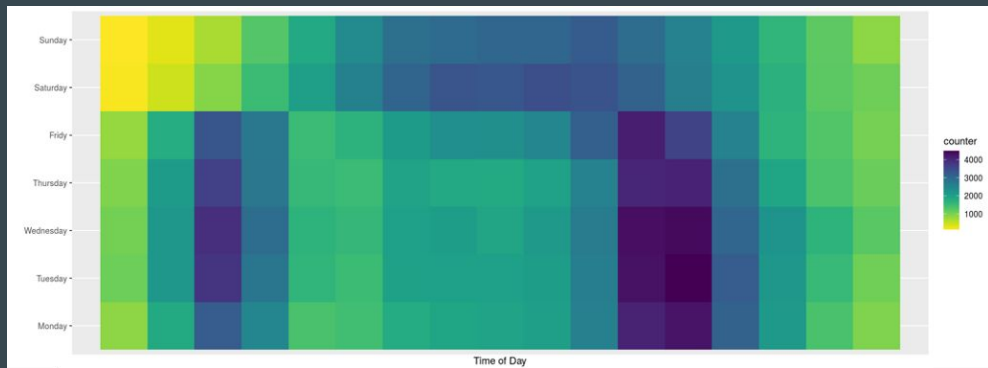
Citi Bike has expanded greatly in the city of New York since starting in 2013.

- Doubled number of docks by 2016
- Doubled average monthly ridership by 2019
- Locations: Manhattan, Brooklyn, Queens, Jersey City, Bronx





# Hourly Demand



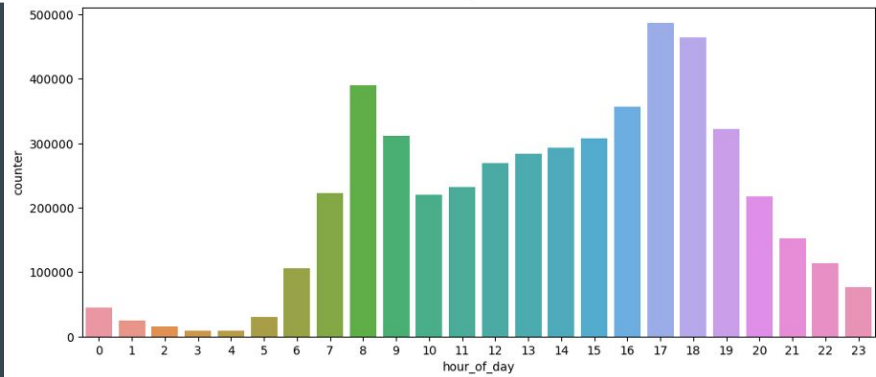
Most popular times:

Monday through Friday:

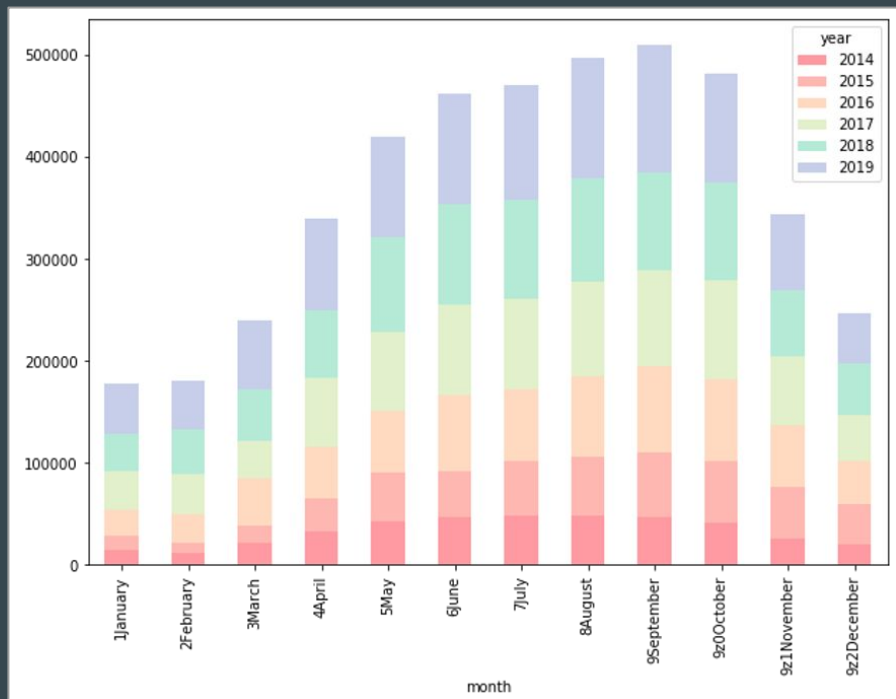
- 7 - 9 AM (Financial District/Midtown)
- 4 - 8 PM (increased demand near Central Park)

Saturday and Sunday:

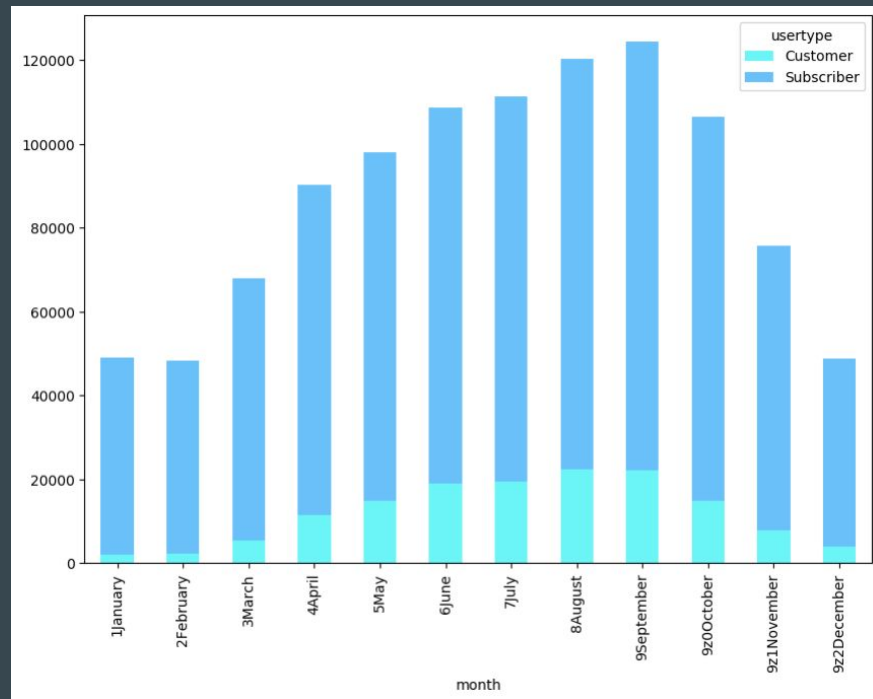
- Afternoons



# Ride Trends by Month

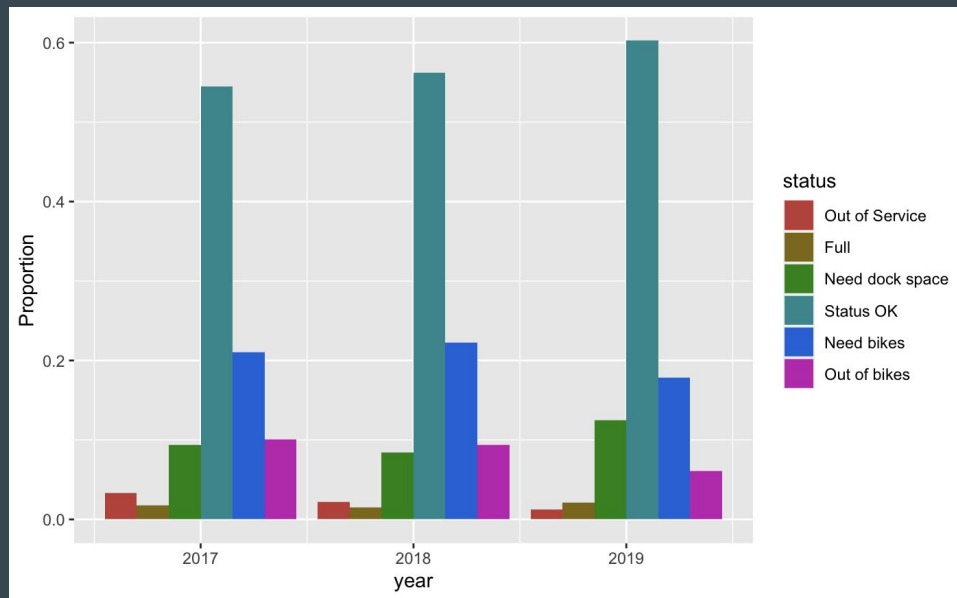
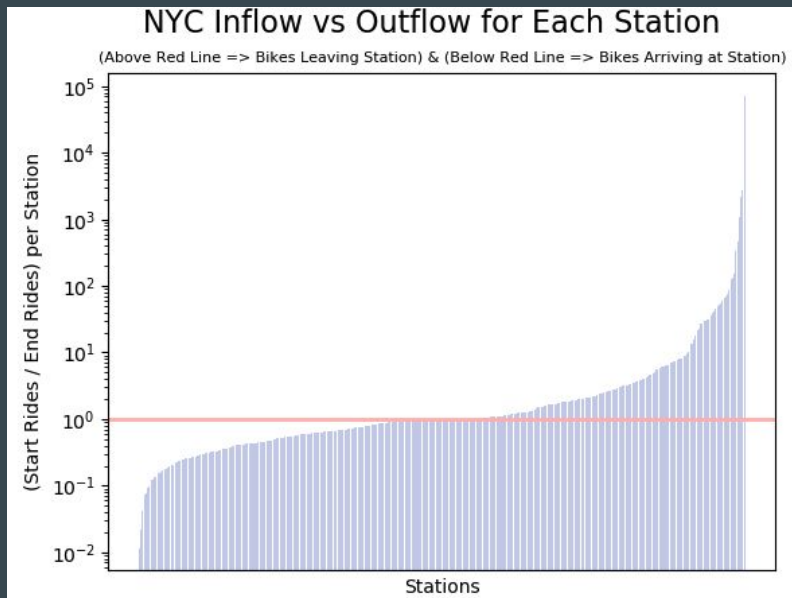


Annual Ridership Trend by Month



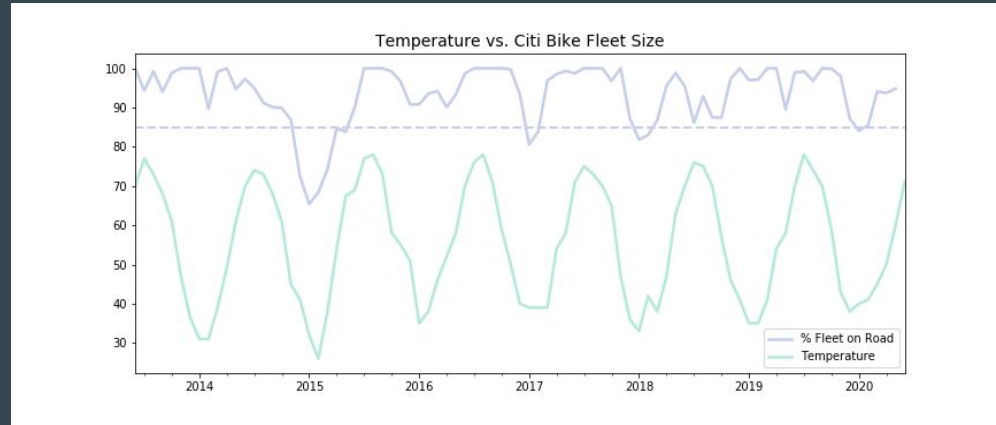
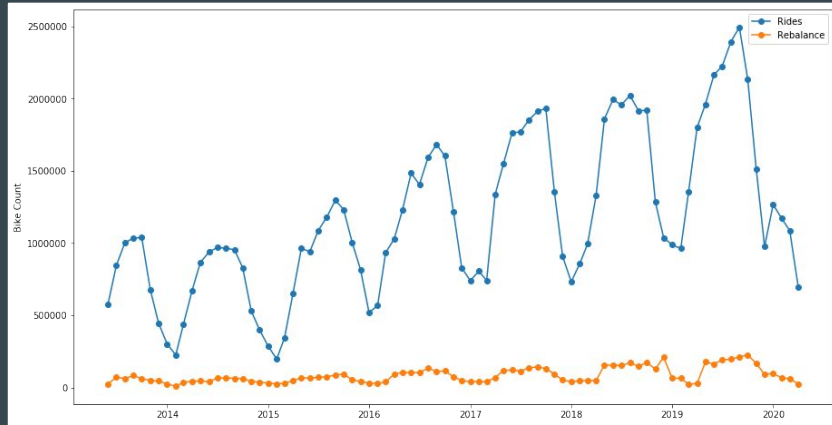
User Type by Month (2019)

# Necessity of Bike Rebalancing



- 97.7% of unique rides do not end at the same station
- Rides starting from the most popular stations (10th percentile) account for 45% of all rides

# Total Number of Rides v. Number of Rebalancing



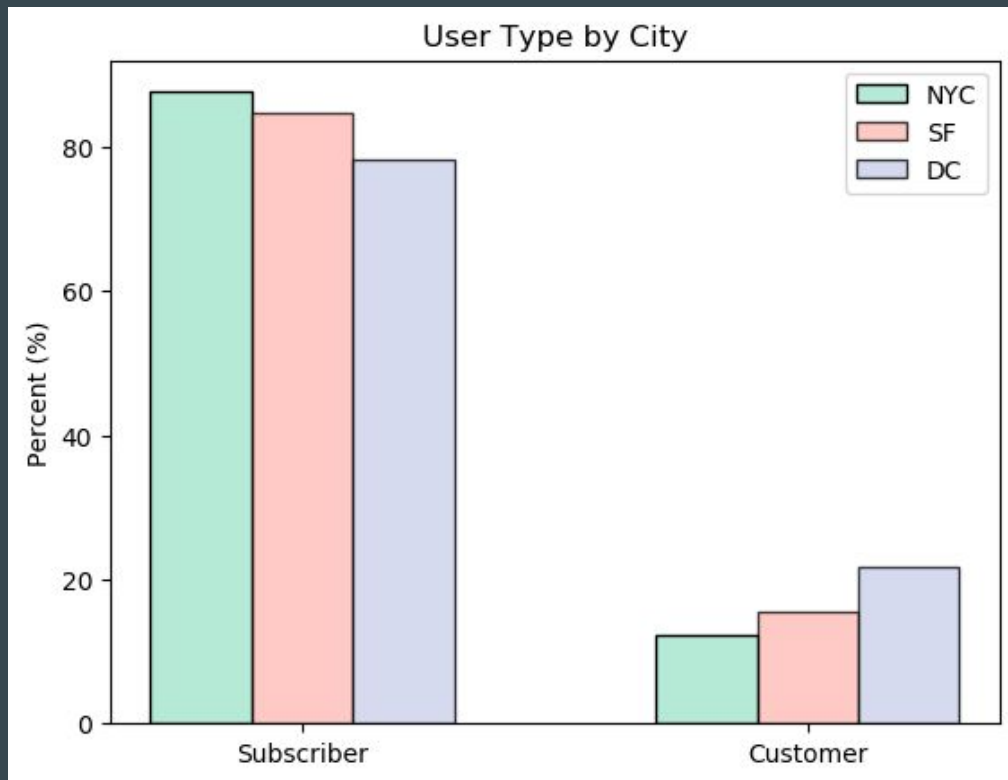
- On average, rebalancing was performed on around 8% of total bike rides during the month
- Decrease in rides correspond to a decrease in bike utilization, and therefore the decrease in bike fleet size

# City Comparisons

# Bike Share Comparison - Subscriber %

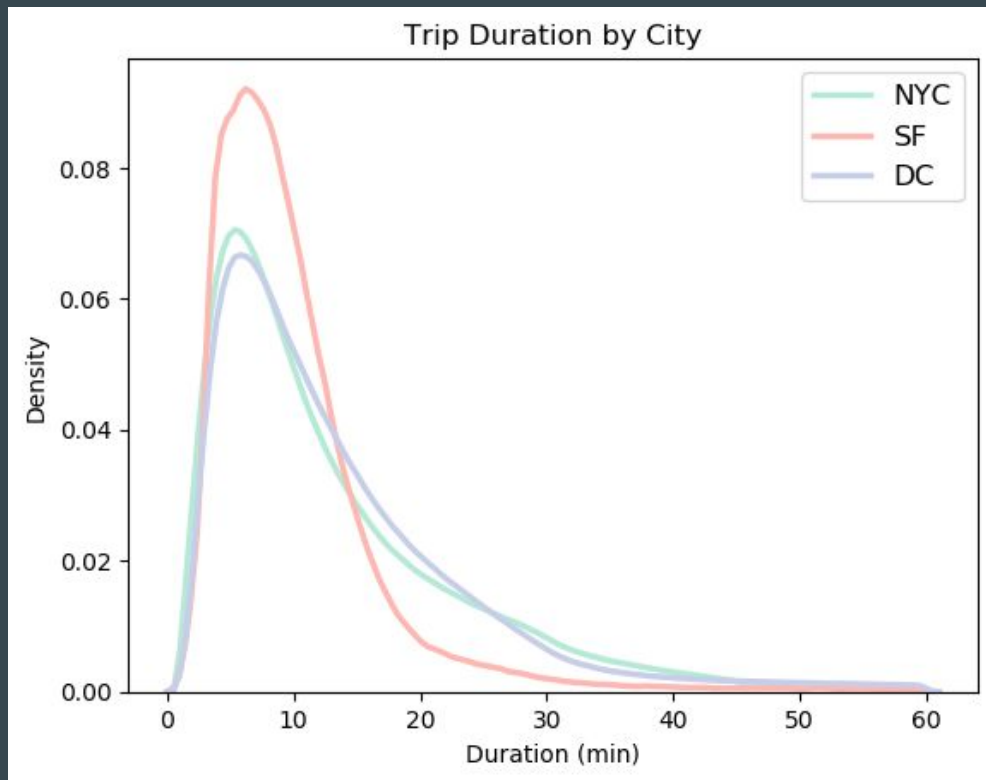
## Takeaway

- Similar amt. of subscribers
- Differences
  - Population Density (people per square mile)
    - NYC - 27,000
    - SF - 17,000
    - DC - 9,800



# Bike Share Comparison - Ride Duration

- Short trips are common
- SF has least spread
  - Hilliness
    - 6th most hilly city
    - DC = 49th
    - NYC = 58th
  - Pilot Program
    - “four successful years as a small scale proof of concept”
    - Started with “350 bikes and 35 stations”



# Bike Share Comparison - Scale of Business Model

- NYC has the most amount of stations and rides
- More fair comparison is
  - Number of rides per year per person
  - Proxy for how “built out” the business is



	Number of Stations	Number of Rides	Rides per Year per Capita
City			
New York City	1,230	99,026,236	0.75
Washington DC	488	19,117,643	0.51
San Francisco	74	669,959	0.1



# Revenue

# Sources of Revenue

- **Annual Membership**

- Unlimited 45-minute rides for a full year ( \$169/yr )

- **Sponsorship (9 in total)**

- Citi Bank is largest (\$41M/2012 & \$70.5/2014-2024)

- **Casual Membership**

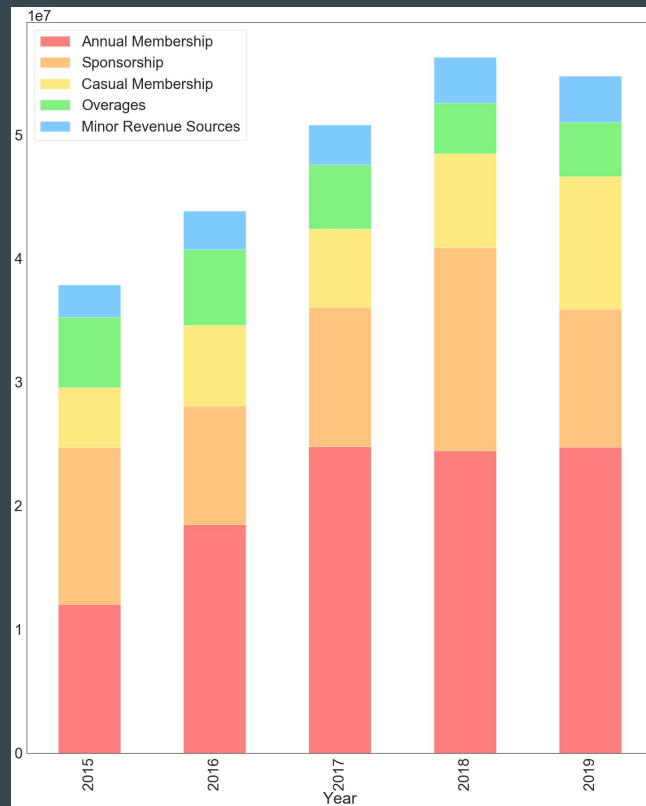
- Single ride, day pass, etc. ( \$3/trip or \$12/day )

- **Overages**

- penalization for extra usage of bike ( \$0.15/min or \$4/15 min )

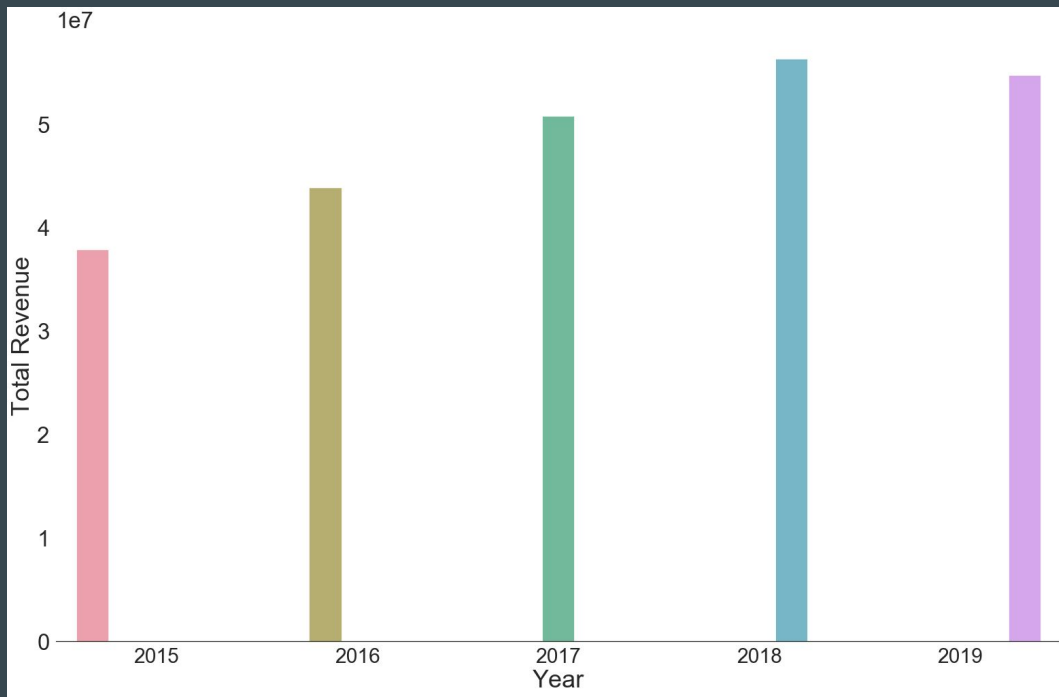
- **Minor Revenue Sources**

- 'other revenue' and 'tax'

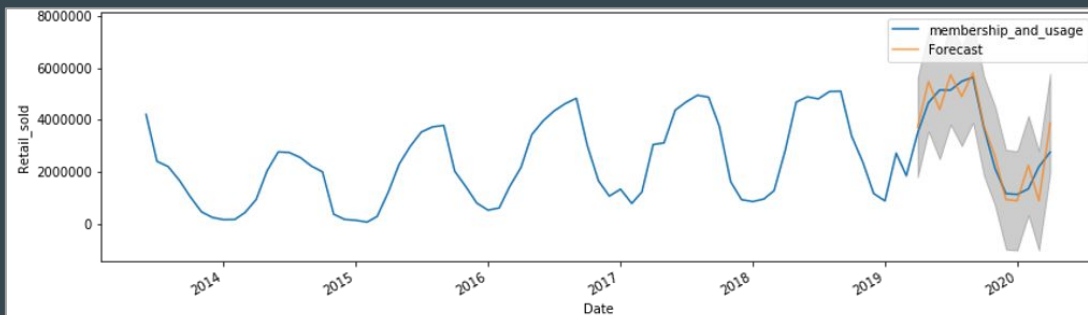
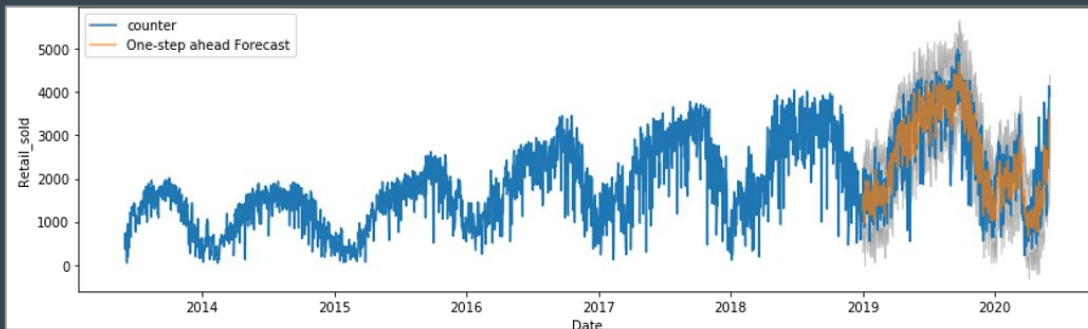


# Revenue Growth

- \$5 million annual revenue growth
- 2013 - 2017:
  - Average +13.8%
- 2018 and 2019:
  - 9.8% and -2.8%
  -
- Only 3 of top 10 US Companies have 2019 revenue growth > 13.8%



# Seasonal ARIMA Forecasting



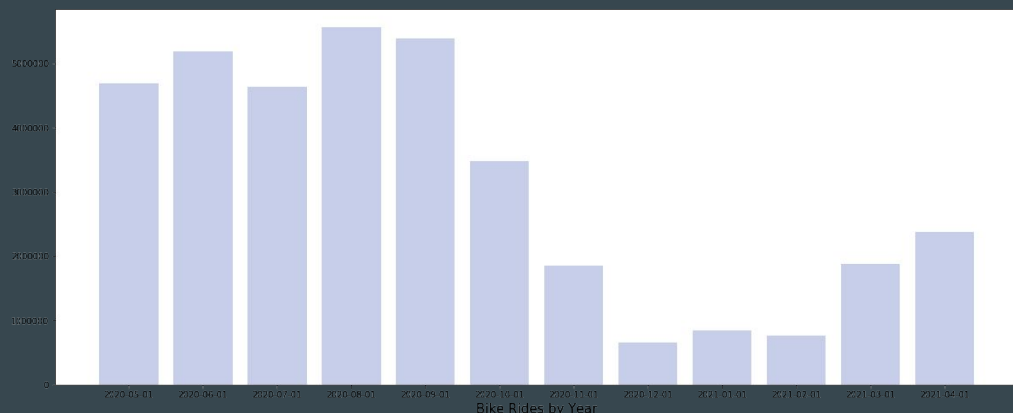
Forecasted:

- Daily ridership
- Monthly ridership
- Monthly revenue

Seasonal Trends

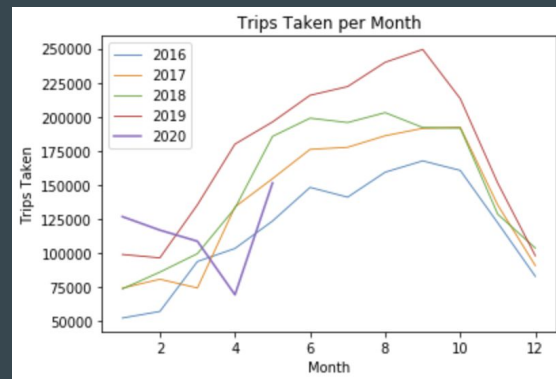
- Daily rides:  
weekly/seasonal trend
- Revenue: seasonal trend

# Revenue Projections



	2020 (Projected)	2020	2019	2018
April	\$3,751,805	\$2,751,308	\$3,484,582	\$2,838,209
May	\$5,188,371	-	\$4,660,579	\$4,681,279
June	\$4,634,427	-	\$5,147,428	\$4,885,485
July	\$5,567,771	-	\$5,142,237	\$4,804,056

- Sharp drop in April revenue from projected revenue due to CoVID-19 lockdown
- April to May increase in demand:
  - 2019: 40%
  - 2020: 119%



# Revenue

## Business Recommendations:

- Increase subscriber base
  - *Drop price of casual rides to bring new clientele to business in hopes of converting them to subscribers*
  - *Doing so may help to break through subscriber plateau of last 3 years*
- Sponsorship
  - *Encourage sponsorship contracts for long term giving to avoid apparent dives in revenue*
    - *or*
  - *Eliminate sponsors and make bike fleet into travelling billboards*

# Sponsorship and Advertisement

# Sponsorship vs. Advertising

## Basic Idea

1. Bikes are moving advertisements
2. Sponsors pay premium for their logo on the bikes
3. Advertising could generate more revenue



<https://indyweek.com/news/raleigh-s-bike-share-program-back-agenda-maybe/>

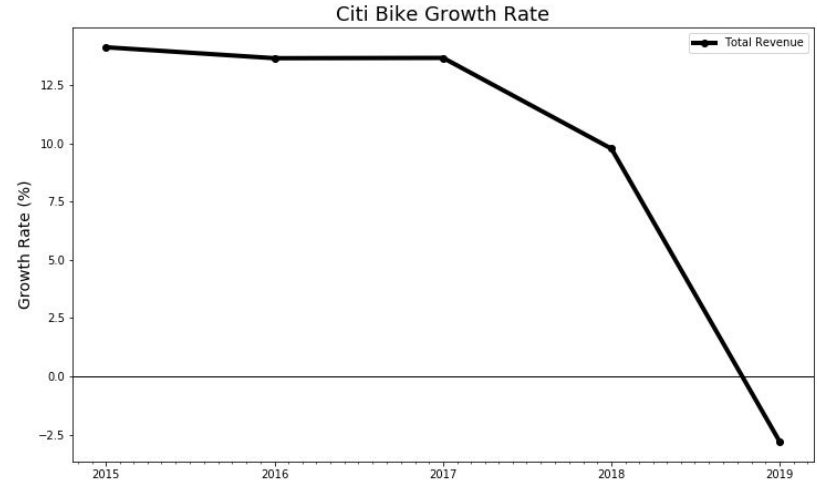
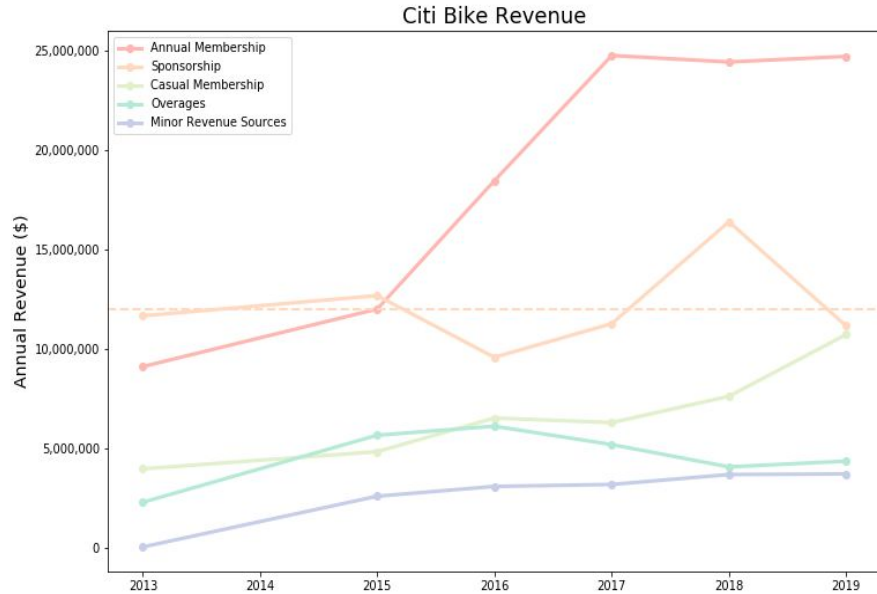


<https://www.post-gazette.com/news/transportation/2018/12/02/Healthy-Ride-Pittsburgh-bike-sharing-program-adding-stations-new-neighborhoods/stories/201812020119>

<https://tinyurl.com/y9s346jr>



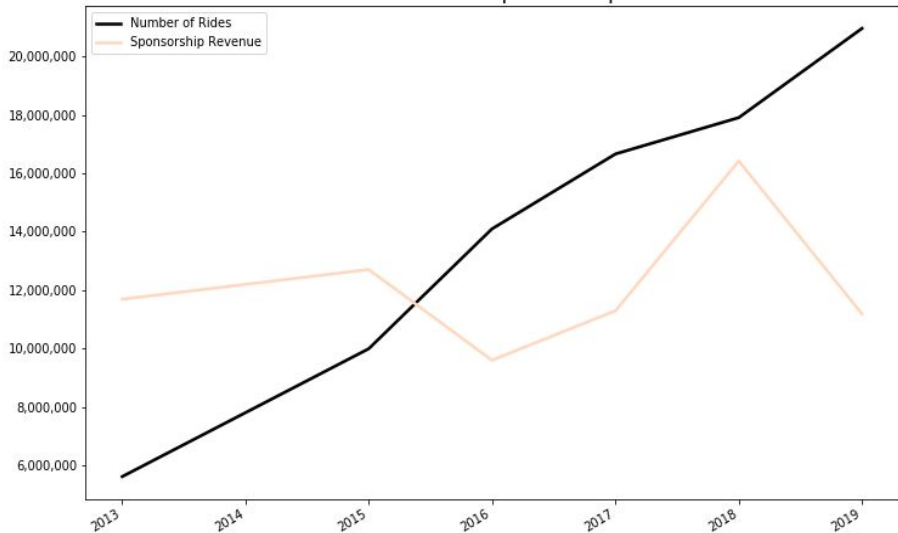
# Sponsorship vs. Advertising



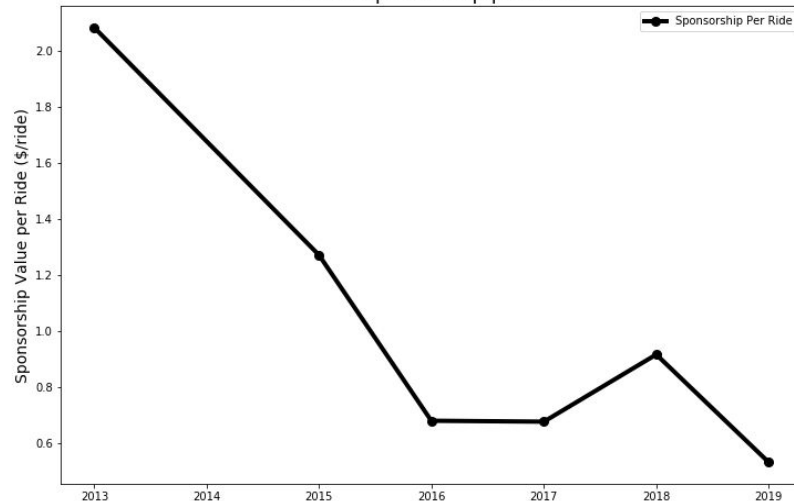
- Advertising must generate over \$12 Million / yr
- Fluctuation has large impact on business

# Sponsorship Cost per Ride

Number of Rides vs. Sponsorship Revenue



Annual Sponsorship per Ride



- Proxy for cost of logo on bike
- Advertising has become cheaper for Citibank
- ~ \$1.00/ride is the going rate for bike share sponsorship

# Sponsorship vs. Advertising

How cheap or expensive is sponsorship in relation to traditional advertising?

- Citi Bike - \$1.03/ride
- Taxis - \$1.75/thousand impressions [[source](#)]



# Sponsorship vs. Advertising

How many bicycle impressions are necessary to equate the two?

- Citi Bike - \$1.03/ride
- Taxis - \$1.75/thousand impressions

$$\frac{\$1.75}{1000 \text{ impr.}} = \left( \frac{\$1.03}{1 \text{ ride}} \right) * \left( \frac{1 \text{ ride}}{N \text{ impr.}} \right)$$
$$\Rightarrow N \text{ impr} = 590$$

*On each Citi Bike ride, 590 onlookers are necessary for sponsorship to be just as cheap as taxi cab advertising*

Scenario	Reality	Super Expensive	Super Cheap
<b>Sponsorship Cost per Ride</b>	\$1.03 / ride	\$1,750 / ride	\$0.0175 / ride
<b>Needed Impressions</b>	590 people	1,000,000 people	10 people

# Business Recommendations and Conclusions - ML Draft

- Sponsorship
  - Use sponsorship until program is popular enough to warrant advertising
  - Encourage either:
    - flat rate sponsorship per ride
    - long term contracts so revenue does not appear erratic
- Costs
  - In Northeast expect to move ~15% of bike fleet into storage in winter
  - Factor in bike rebalancing through modeling or lobby for dockless bikes
- Subscriptions
  - Decrease casual rider cost to attract new customers
  - Offer generous subscription rules (>2 hrs) knowing that most rides are < 30 minutes
  - Doing so will help break through subscription plateau

# Given more time...

- May and June Operations Monthly
- Cost Estimation
  - Bike Maintenance
  - Rebalancing cost

# Bike Share Comparison

## Business Recommendations

- Subscriber percentage
  - *Focus on maximizing cost of subscription rather than cost of 1-time rentals*
  - *Decrease 1-time rental cost to get people comfortable with the program*
    - *Similar to grocery stores decreasing turkey price around Thanksgiving*
- Ride duration
  - *Offer generous subscription rules (unlimited 2 hour rides) with the knowledge that the common subscriber will not take advantage of this*
  - *Generous subscription rules will help justify high premium of subscription for potential customers*

# Business Recommendations and Conclusions

- Effective bike rebalancing to meet growing demand
- Flat rate sponsorship per ride
-



# Questions?

...

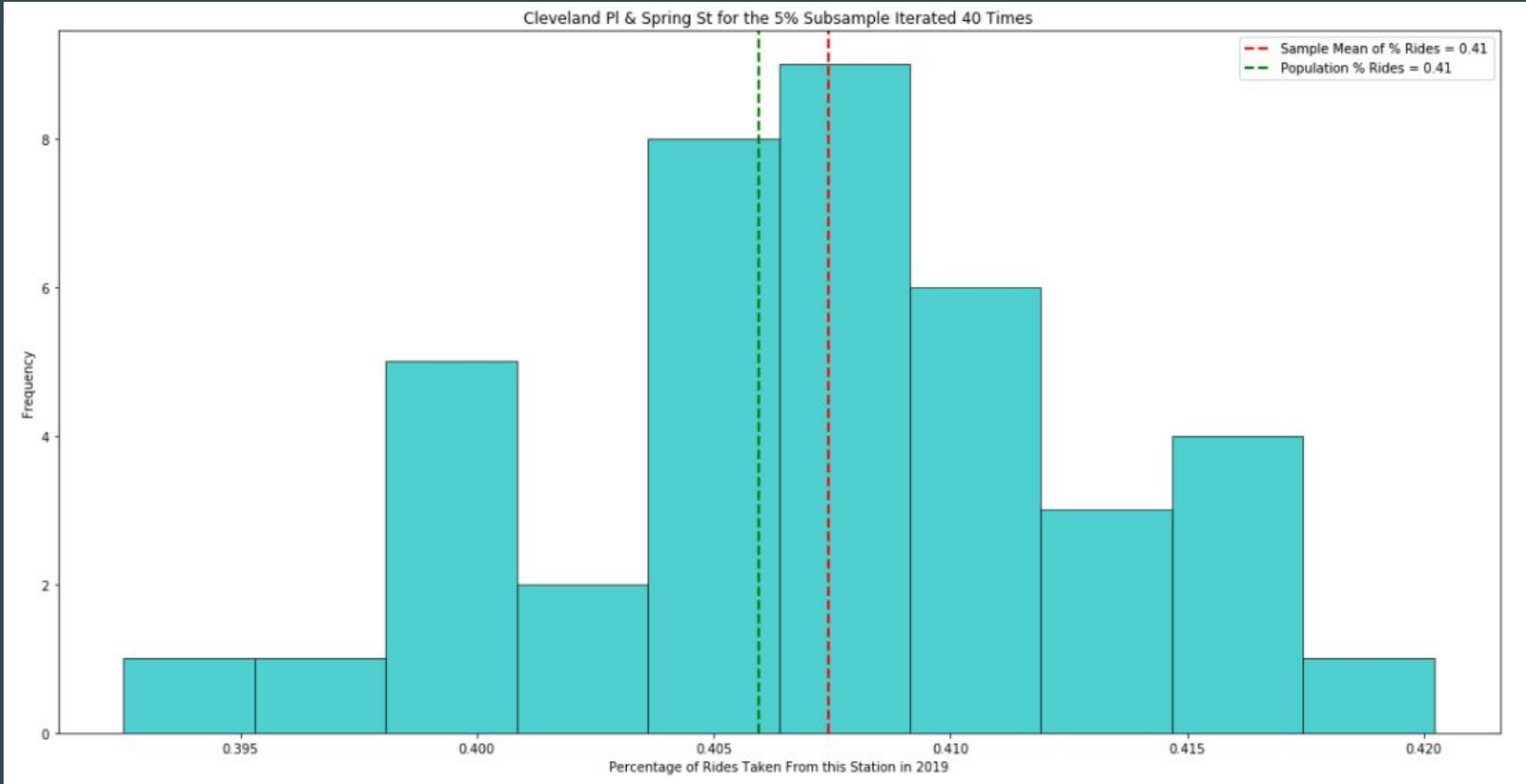


# Slides as Backup

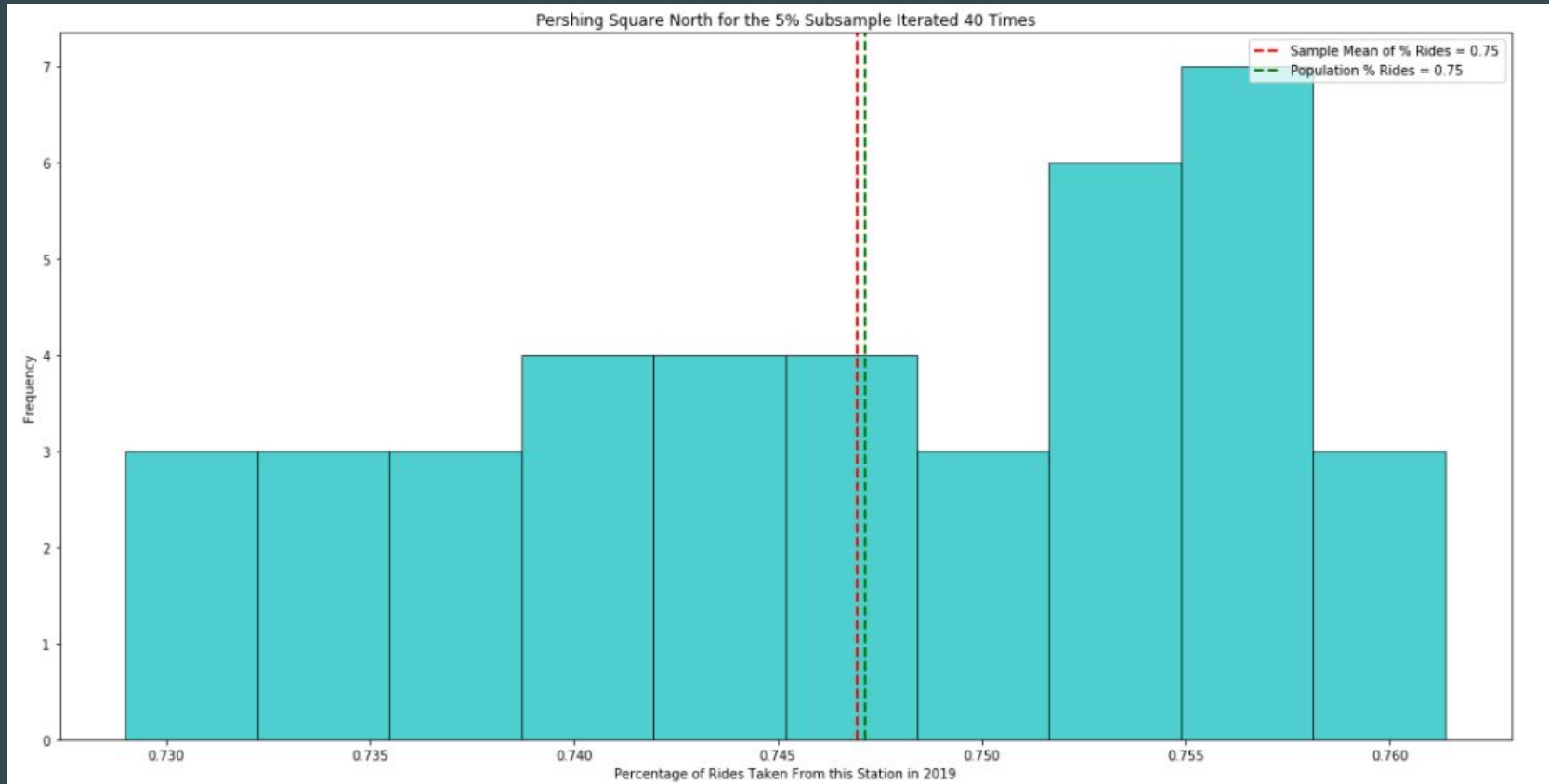
# Data Cleaning, Validation, and Storage

- 140 csv's converted to SQLite database (db)
- Query number of rides from each station
- Calculate proportion of rides from each station
- Randomly sampled from db
  - 1% sample 40 times
  - 5% sample 40 times
  - 10% sample 40 times
- Compared histogram of sample proportions to global proportions
- 5% sample used for high-level analysis
- SQLite db used to search full dataset and not break Jupyter notebook

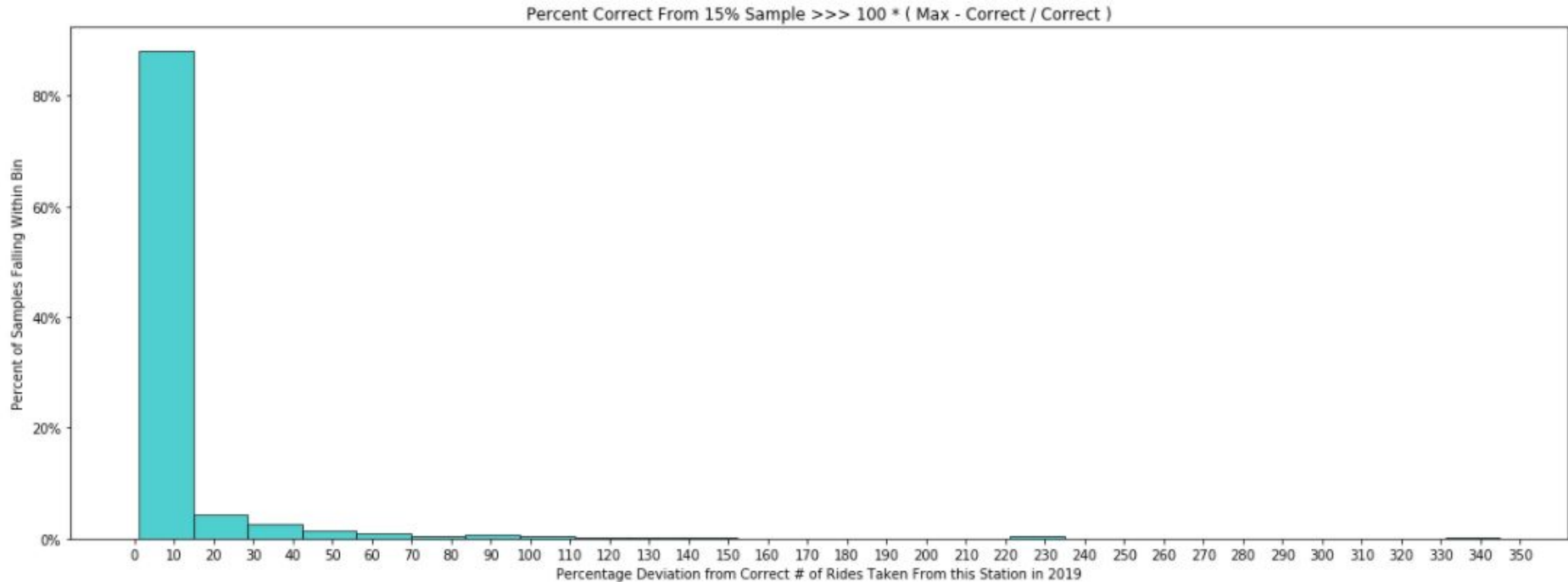
# % Rides from Station A - Distribution of Random Sample (Good)



# % Rides from Station A - Distribution of Random Sample (Bad)



# Percent Deviation From Correct Answer When Sampling

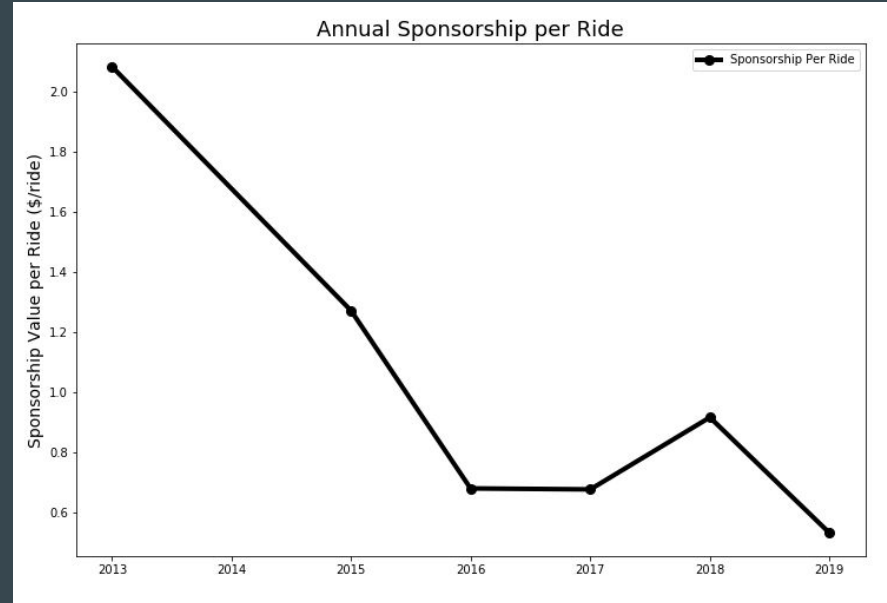




# Sponsorship vs. Advertising

## Business Recommendations

- Use sponsorships and not advertising at beginning of program
  - Reassess later if \$/ride continues to decrease
- Suggest flat rate sponsorship







# Sponsorship vs. Advertising

How many bicycle impressions are necessary to equate the two?

- Citi Bike - \$1.03/ride
- Taxis - \$1.75/thousand impressions

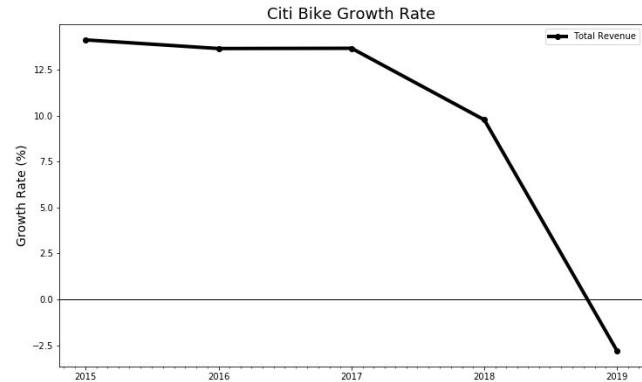
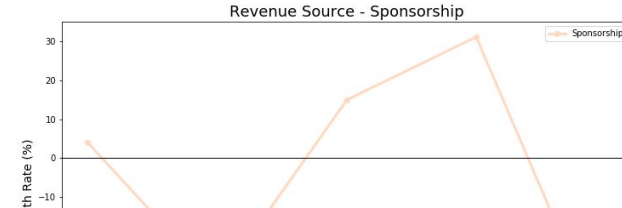
$$\frac{\$1.75}{1000 \text{ impr.}} = \left( \frac{\$1.03}{1 \text{ ride}} \right) * \left( \frac{1 \text{ ride}}{N \text{ impr.}} \right)$$

$$\Rightarrow N \text{ impr} = 590$$

# Revenue - Impact of Sponsorship - JW VERSION, i think a bar plot would be more effective, also i hate the black plot - SORRY!

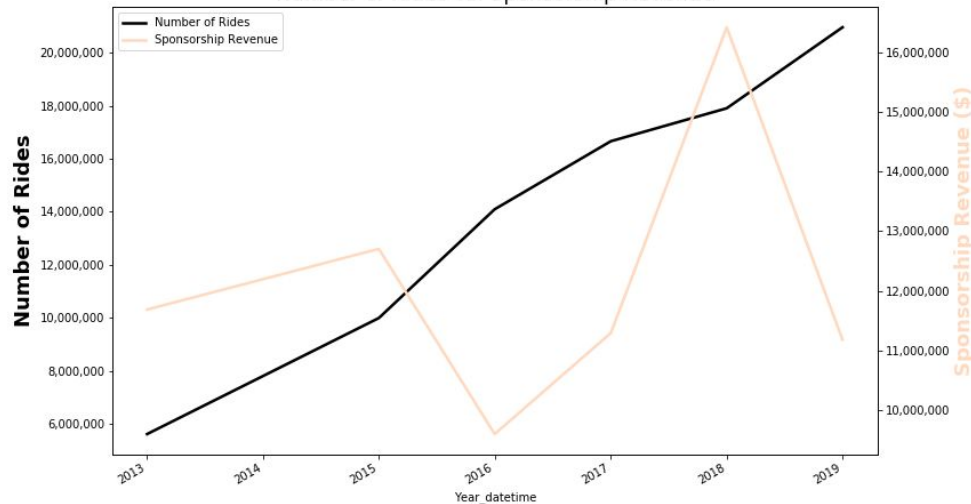
Importance:

- Sponsorship is 2nd largest revenue
- *Fluctuation has large impact on business*

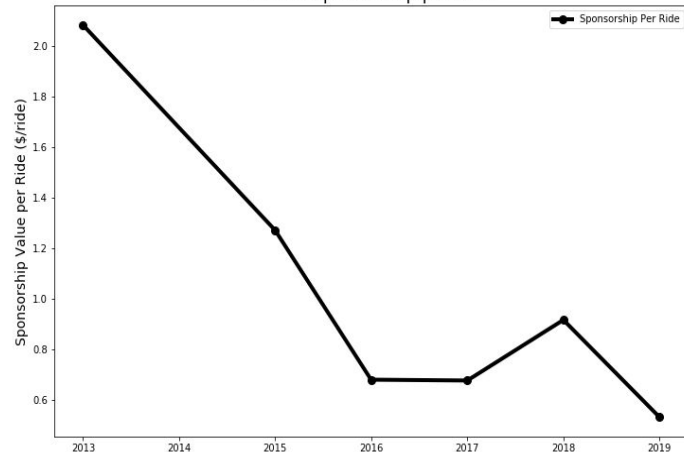


# Sponsorship Cost per Ride

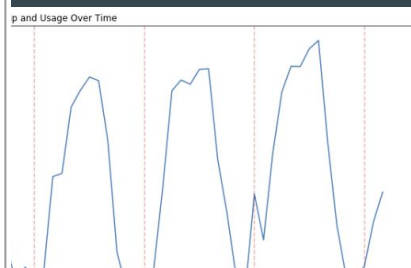
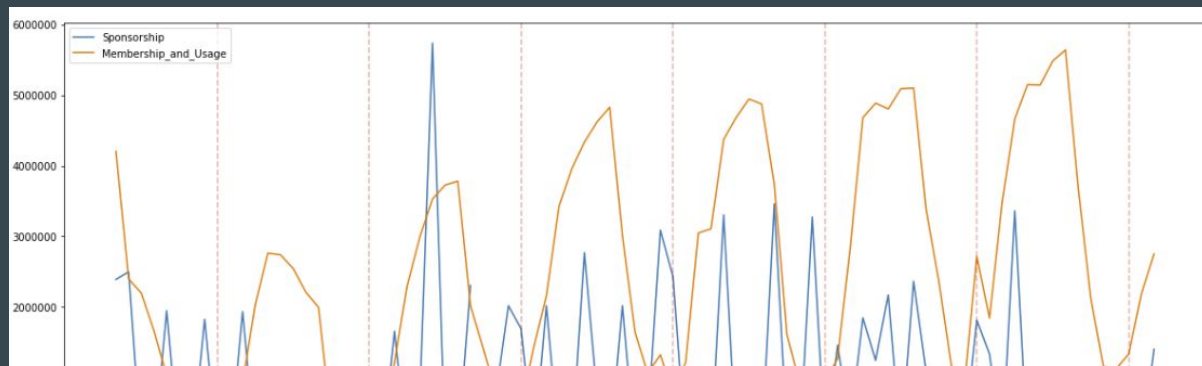
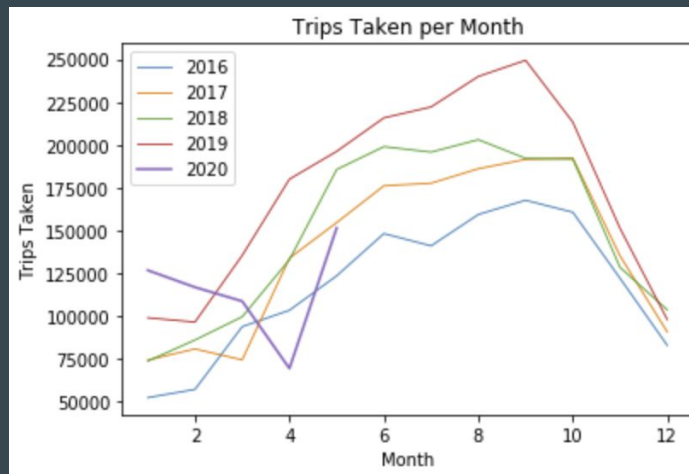
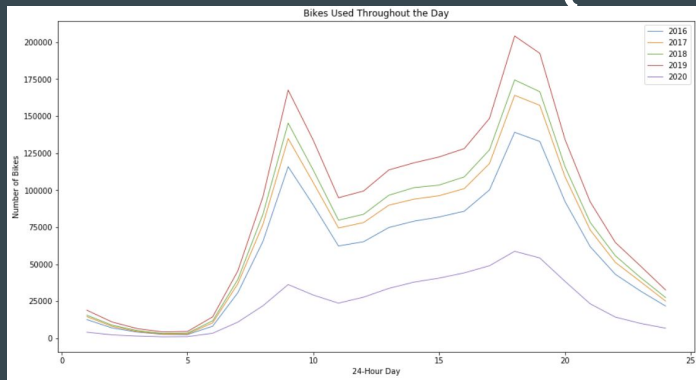
Number of Rides vs. Sponsorship Revenue



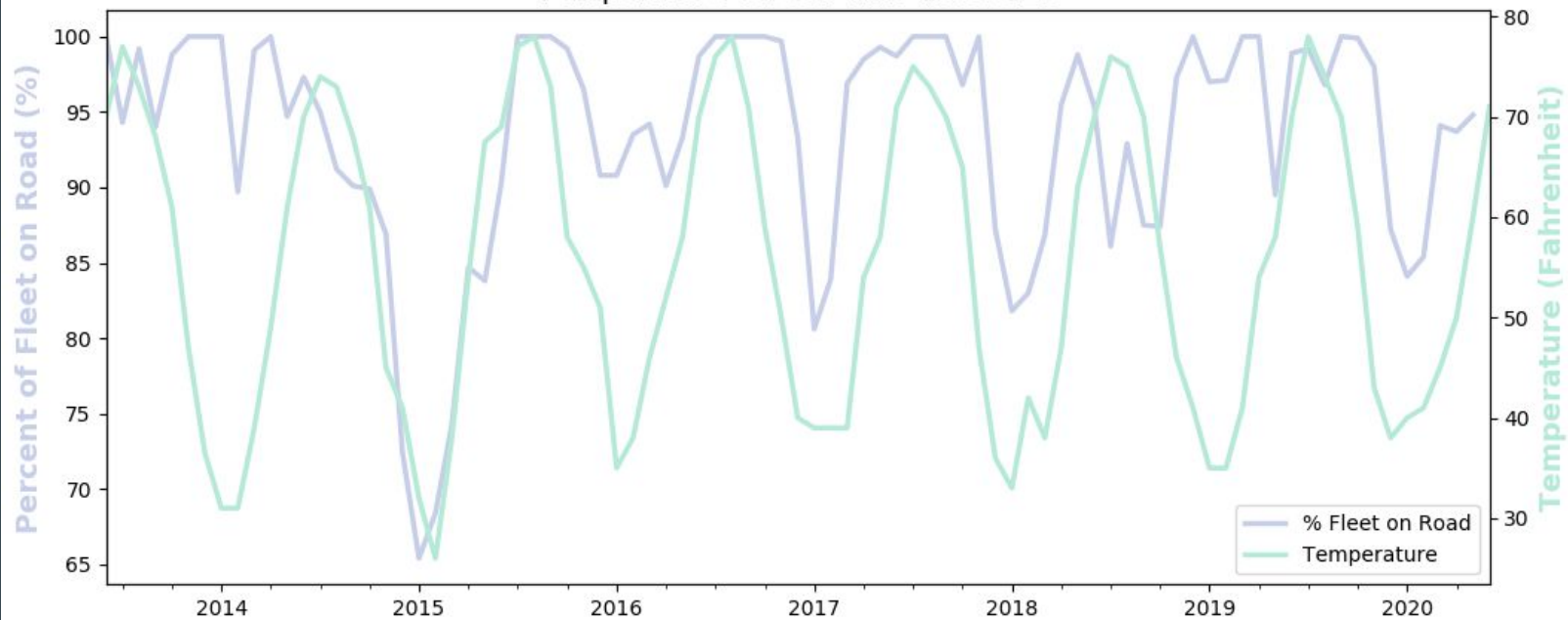
Annual Sponsorship per Ride



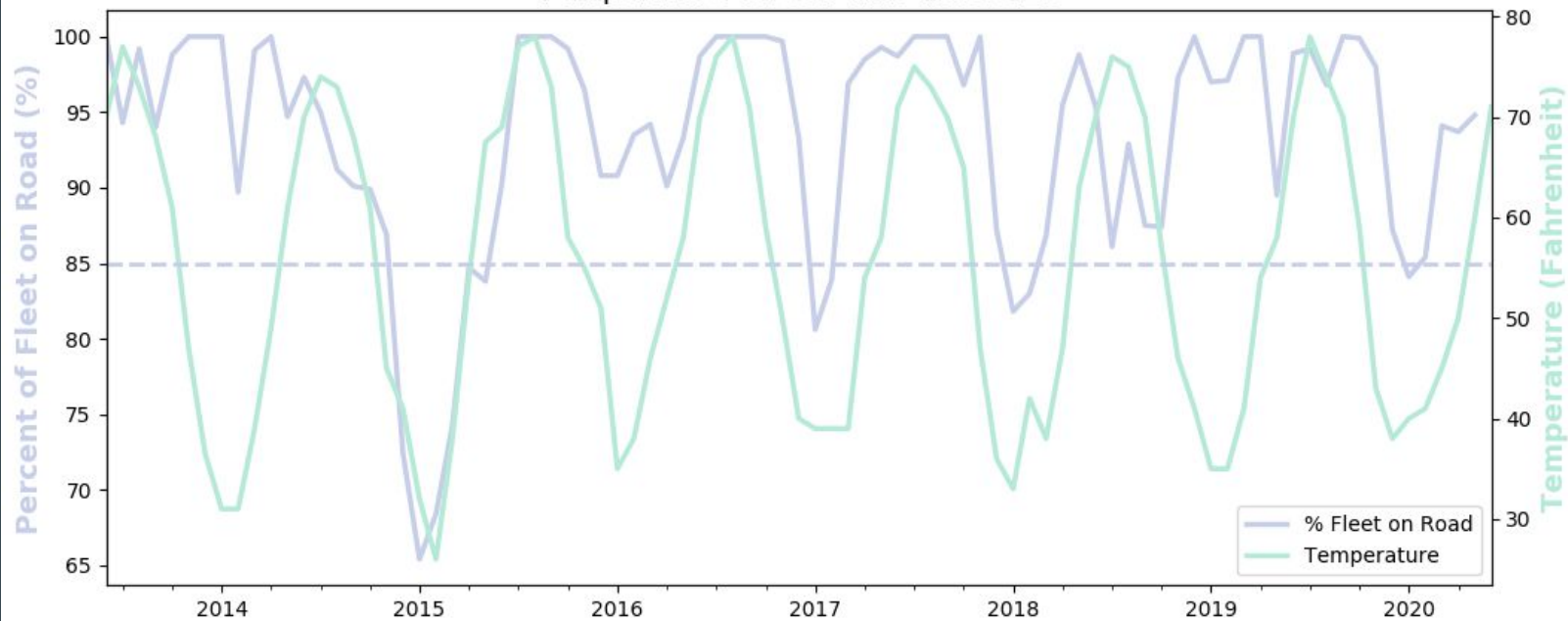
# Effect of COVID 19 (Rider Behavior)



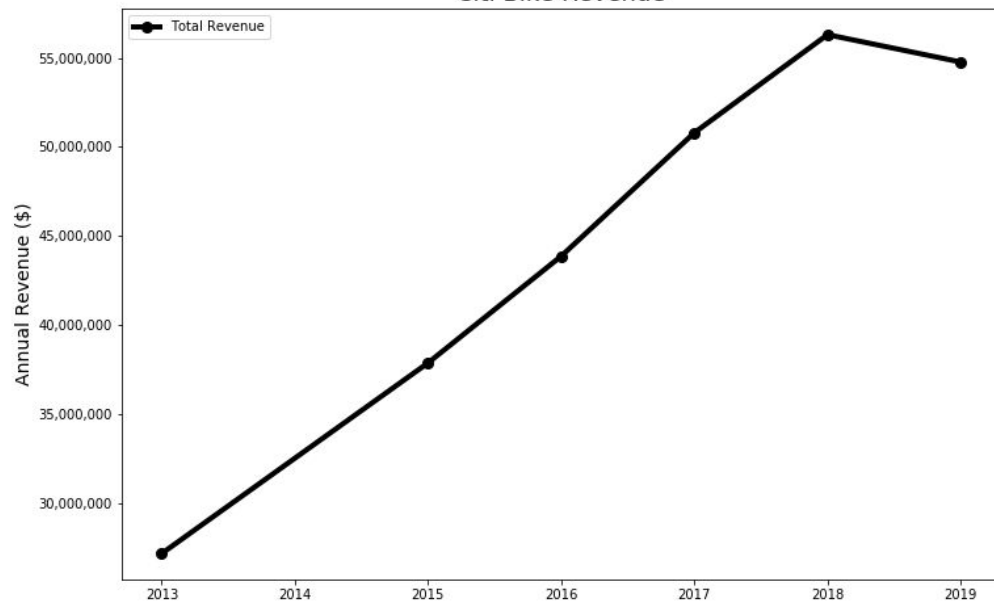
Temperature vs. Citi Bike Fleet Size



Temperature vs. Citi Bike Fleet Size



Citi Bike Revenue





# Business Recommendations and Conclusions - ML Draft

- Costs
  - In Northeast expect to move ~15% of bike fleet into storage in winter
  - Factor in bike rebalancing through modeling or lobby for dockless bikes
- Subscriptions
  - Decrease casual rider cost to attract new customers
  - Offer generous subscription rules (>2 hrs) knowing that most rides are < 30 minutes
  - Doing so will help break through subscription plateau
- Sponsorship
  - Use sponsorship until program is popular enough to warrant advertising
  - Until then charge flat rate per ride
  - Encourage sponsorship contracts so revenue does not appear erratic

# Revenue

## Business Recommendations:

- Increase subscriber base
  - *Drop price of casual rides to bring new clientele to business in hopes of converting them to subscribers*
  - *Doing so may help to break through subscriber plateau of last 3 years*
- Sponsorship
  - *Encourage sponsorship contracts for long term giving to avoid apparent dives in revenue*
    - *or*
  - *Eliminate sponsors and make bike fleet into travelling billboards*

# Bike Share Comparison

## Business Recommendations

- Subscriber percentage
  - *Focus on maximizing cost of subscription rather than cost of 1-time rentals*
  - *Decrease 1-time rental cost to get people comfortable with the program*
    - *Similar to grocery stores decreasing turkey price around Thanksgiving*
- Ride duration
  - *Offer generous subscription rules (unlimited 2 hour rides) with the knowledge that the common subscriber will not take advantage of this*
  - *Generous subscription rules will help justify high premium of subscription for potential customers*