The Battle between Uber and Lyft: Pricing Strategies in Boston

Soohyun Hwang, Yangzhuopeng Yi, Shengqi Zhou, Ghadeer Alqassab

4/26/20

ISTM 6209

Dr. Wendy Duan

Table of Contents

Introduction

Entering the age of smartphones and mobile apps, ridesharing service became one of the most recognizable and innovative ventures of the decade. Promoting the concept of "share economy", the ridesharing, or ride-hailing, industry became a disruptive force in the passenger transportation industry and forced the incumbent industries, such as the taxi industry, into near-extinction. In New York City, one of the biggest markets in the U.S., taxi companies serviced about 500,000 rides in 2015 (Schneider). By 2020, taxi companies only service around 200,000 rides. Unfortunately, in 2015, the taxi industry was valued at $12 billion which is less than a third of the value of the ridesharing industry (Ferenstein).

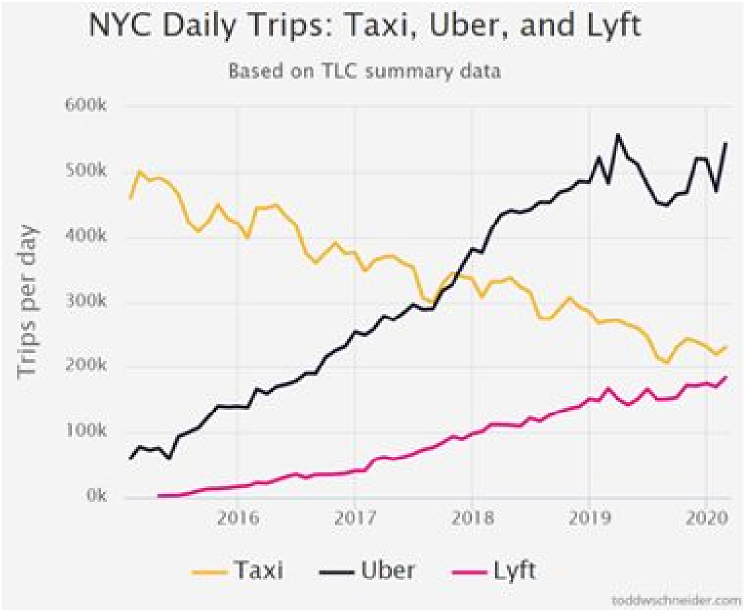

Fig. 1. Todd W. Schneider, *NYC Daily Trips: Taxi, Uber, and Lyft*, 2020.

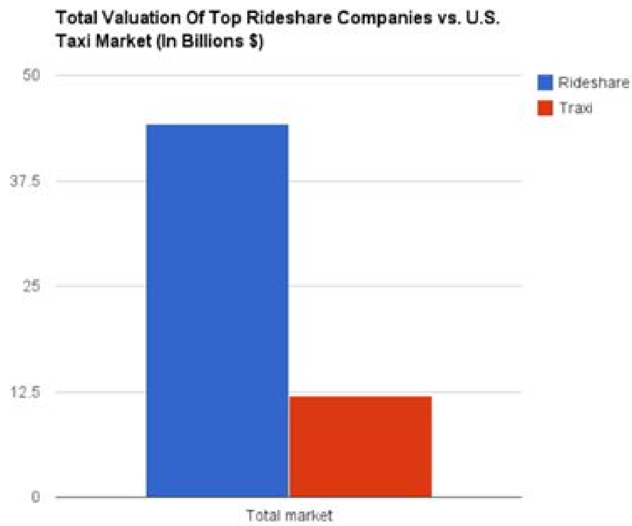

Fig. 2. Greg Ferenstein, *Total Valuation of Top Rideshare Companies vs. U.S. Taxi Market*, 2015.

From customers' perspectives, calling a ride through ridesharing companies is very simple. After downloading a ridesharing app, like Uber or Lyft, customers only have to submit their current locations and their destinations; consequently, ride requests are sent to drivers near them. Finally, if a driver accepts the request, the driver will arrive at the pick-up location and drive the passengers to their requested destinations. As a result, it is not a surprise that customers prefer ridesharing apps over traditional taxis. Among Americans between 18 and 29 years of age, the generation most accustomed to smartphones and mobile apps, more than half them have downloaded and used Uber and Lyft apps (Marshall). The ridesharing industry is currently valued at $61.3 billion and is expected to reach $218 billion by 2025 (Curley).

The Two Giants

Currently, there are two prominent players in the ridesharing industry: Uber and Lyft. Uber has nearly 70% market share in the U.S while Lyft has about 30% market share; it is not wrong to assume the ridesharing industry is essentially a duopoly. However, Uber has a significant lead over Lyft. Aside from the market share, Uber's revenue, assets, and equity are $14.14 billion, $31.76 billion, $10.33 billion, respectively. In comparison, Lyft's revenue, assets, and equity are $3.61 billion, $5.69 billion, $1.83 billion, respectively. Furthermore, in New York City, Uber services 462,113 rides per day, which measures much higher than Lyft's 149,142 rides (Iqbal).

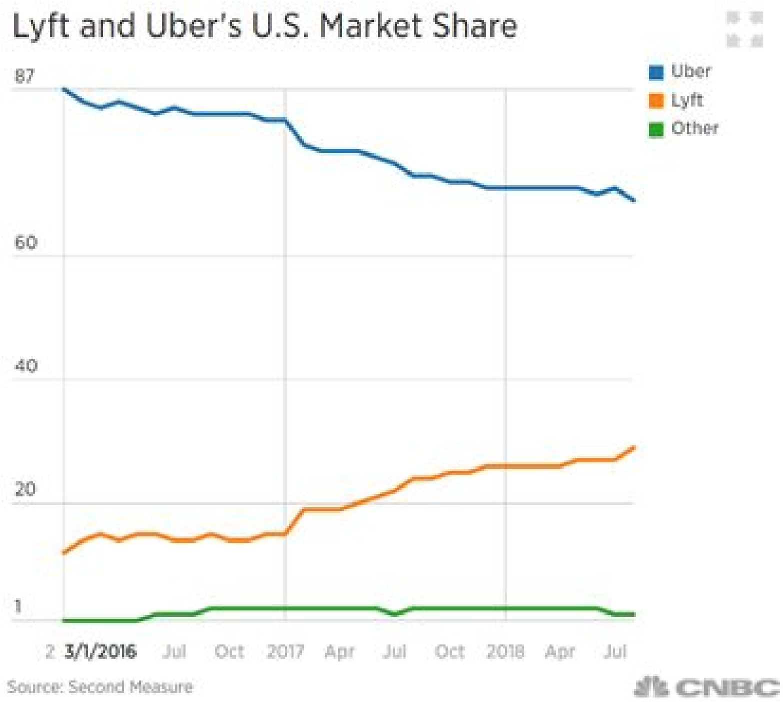

Fig. 3. Paayal Zaveri, *Lyft and Uber's U.S. Market Share*, 2018.

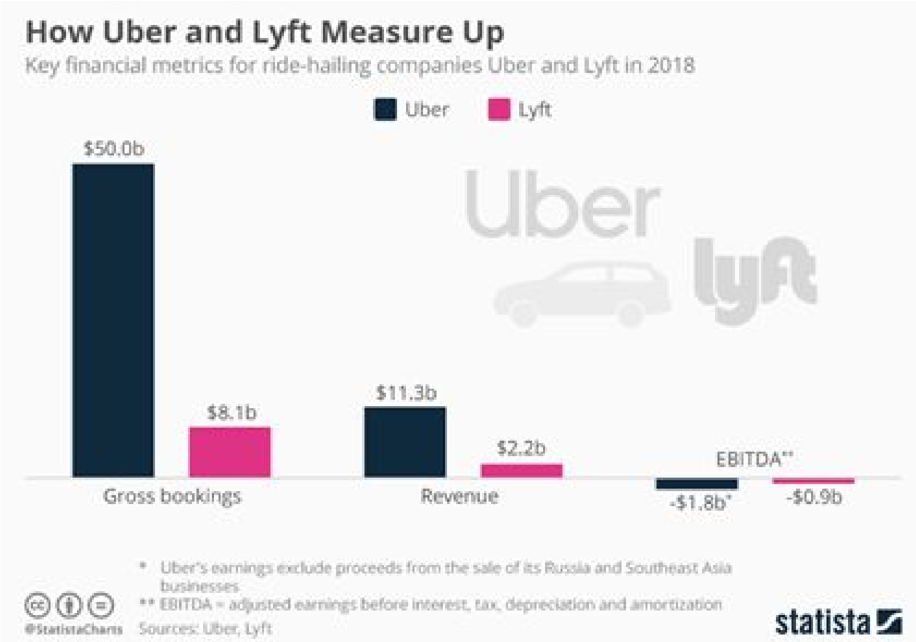

Fig. 4. Felix Richter, *How Uber and Lyft Measure Up*, 2019.

Despite the vast differences, the competition is still fierce. Particularly in one element, the fare price, Uber and Lyft both invest hefty resources to derive pricing algorithms that not only provide attractive prices to customers, but also allow the companies to be financially sustainable. We already know both companies have the same goal: calculating the best fare price. However, we discovered Uber and Lyft have different approaches to meet the goal. This paper examines how Uber and Lyft determine their fare prices in Boston, Massachusetts.

The Price Model

How fares are calculated by Uber and Lyft are based on 5 criteria:

1. Base fare is a flat fee charged for every ride. However, the fee changes depending on the "ride class". For example, UberX is the standard ride that can accommodate up to four passengers. UberPool is a cheaper service that only allows at most two passengers per request and shares the car with other passengers with similar destinations. On the other hand, UberBlack is an upgraded ride service that summons high-end brand cars for the trip. As a result, premium ride types demand higher base fare. Correspondingly, Lyft have the same system: Lyft, Shared, and Lux.

2. Fee per minute is charged for each minute inside the vehicle.

3. Fee per mile is charged for every mile of the trip.

4. Booking Fee (Uber) or Trust and Safety Fee (Lyft) is a flat fee to cover the operating costs.

5. Surge Pricing is a "hidden" fee based on supply and demand. For example, when there are more customers than the number of drivers, the total fare is multiplied

by a “surge price multiplier”. Surge Pricing incentivizes drivers to be present on the road during peak hours.

Data

The data we analyzed, *Uber and Lyft Dataset Boston, MA* from www.kaggle.com, records 693,071 Uber and Lyft trips from November 26, 2018 to December 18, 2018. Important columns are: Hour of the day, Day, Month, Source (Pick-up location), Destination, Cab Type (Lyft or Uber), Product ID and Name (Ride Class), Distance, and Price. Out of 693,071 trips, Uber made 385,663 trips and Lyft made 307,408 trips. This trip volume corresponds with our initial knowledge because Uber has a bigger market share than Lyft in the U.S.

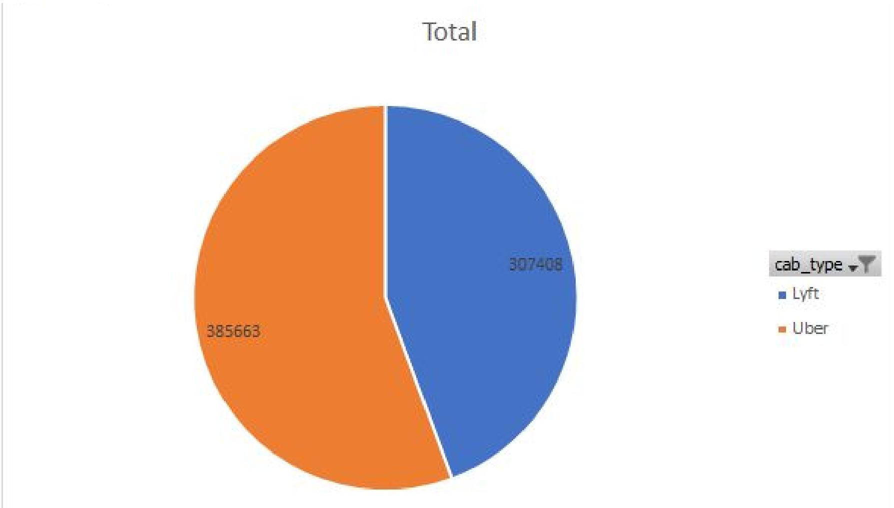

Fig. 5. Number of Rides Between Uber and Lyft

Fig. 6. Ride Types for Uber and Lyft

One interesting component of our dataset is that it also provides data on weather (Temperature, Humidity, Wind Speed, Precipitation, etc.) during the rides. We predict the weather also plays a part when Uber and Lyft compute fare prices because the weather can change customers' behaviors on requesting a ride. For instance, pedestrians will likely call for a ride on days with high temperature and high humidity to avoid heat exhaustion. During winter, demands for rides will also increase because the weather is too cold to walk long distances.

Analysis

Average of price	Column Labels		
Row Labels	Lyft	Uber	Grand Total
Clear	$17.31	$15.72	$16.49
Drizzle	$17.35	$15.77	$16.50
Foggy	$17.51	$15.64	$16.54
Light Rain	$17.34	$15.84	$16.56
Mostly Cloudy	$17.41	$15.84	$16.59
Overcast	$17.32	$15.80	$16.54
Partly Cloudy	$17.38	$15.77	$16.55
Possible Drizzle	$17.24	$15.81	$16.48
Rain	$17.31	$15.83	$16.54
Grand Total	**$17.35**	**$15.80**	**$16.55**

Fig. 7. *The effect of weather on price of rides*

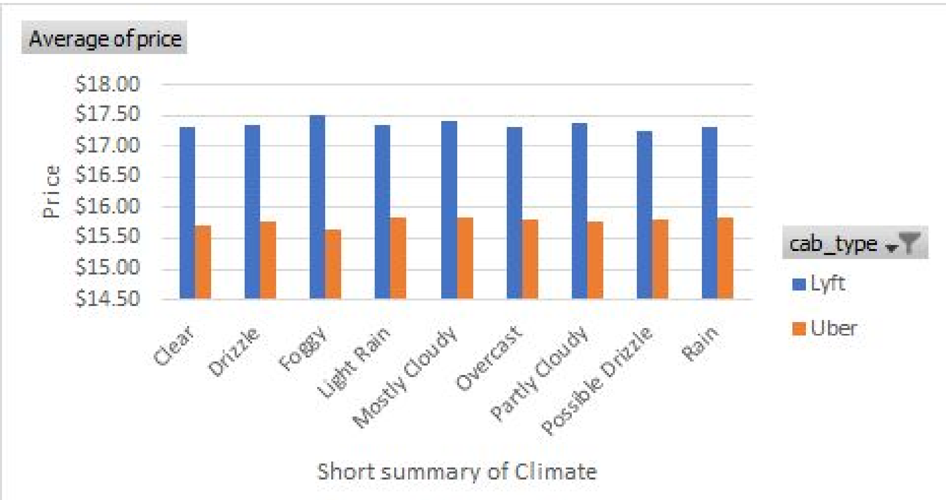

Fig. 8. The effect of weather on price of rides

From the bar chart above, we can find some differences between different weather conditions. For Lyft, when the day is Foggy, it has the highest average price, on the contrary, Uber has the lowest average price. For Uber, when the weather is light rain or mostly cloudy, its average price will be the highest.

Count of id	Column Labels		
Row Labels	Lyft	Uber	Grand Total
Clear	38653	48473	87126
Drizzle	3111	4185	7296
Foggy	4002	5058	9060
Light Rain	24328	30584	54912
Mostly Cloudy	64720	81490	146210
Overcast	97416	121479	218895
Partly Cloudy	56663	70561	127224
Possible Drizzle	8072	10564	18636
Rain	10443	13269	23712
Grand Total	**307408**	**385663**	**693071**

Fig. 9.The effect of weather on number of orders

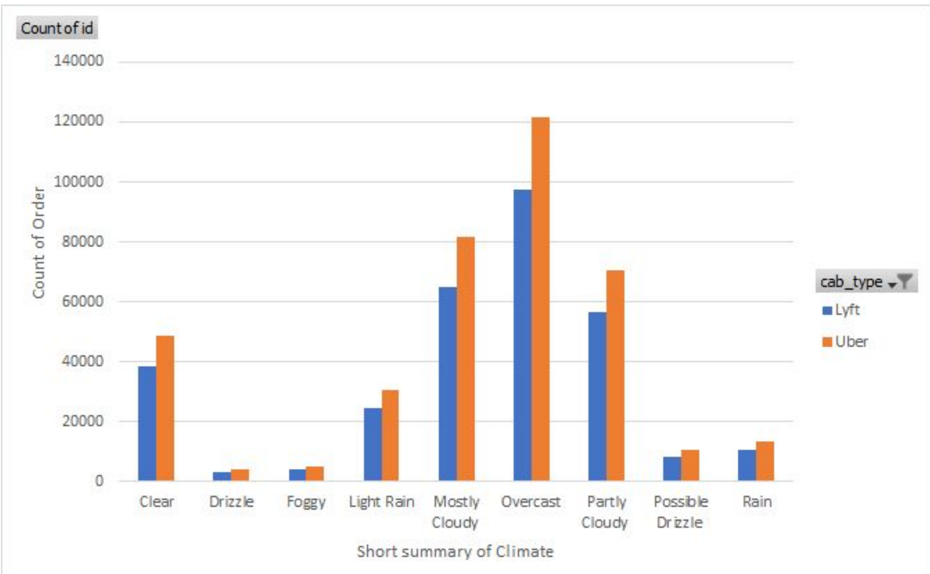

Fig. 10. The effect of weather on number of orders

With further research and based on the bar chart above, we can see that the most frequent weather in Boston between November and December is Cloudy.

Otherwise, we can find that Lyft's average price is about 10% higher than the price of Uber. This observation might be the reason why the number of ride orders for Lyft is always lower than Uber's at any time and any weather. Uber chooses to make money from their drivers instead of from their riders. Lyft takes 20% of each fare, plus the entire booking fee, while Uber takes 25% from each fare (Helling). This profit model helps Uber attract more riders and then with the demand increasing, more drivers are willing to drive for Uber. Even though the drivers might earn less money per order from Uber than what they can earn from Lyft, they still are willing to stay with Uber.

Comparing Uber And Lyft Prices in Different Hours

Average of price	Column Labels		
Row Labels	Lyft	Uber	Grand Total
0	$17.45	$15.73	$16.57
1	$17.31	$15.77	$16.51
2	$17.32	$15.87	$16.56
3	$17.35	$15.79	$16.54
4	$17.37	$15.86	$16.58
5	$17.33	$15.79	$16.54
6	$17.28	$15.83	$16.52
7	$17.32	$15.73	$16.49
8	$17.40	$15.87	$16.60
9	$17.26	$15.81	$16.51
10	$17.31	$15.76	$16.50
11	$17.50	$15.69	$16.56
12	$17.33	$15.69	$16.49
13	$17.35	$15.80	$16.55
14	$17.36	$15.81	$16.55
15	$17.34	$15.77	$16.52
16	$17.36	$15.74	$16.51
17	$17.38	$15.86	$16.61
18	$17.24	$15.85	$16.52
19	$17.38	$15.75	$16.55
20	$17.39	$15.88	$16.60
21	$17.33	$15.91	$16.60
22	$17.44	$15.78	$16.60
23	$17.30	$15.77	$16.50
Grand Total	**$17.35**	**$15.80**	**$16.55**

Fig. 11. The effect of time on average price of ride

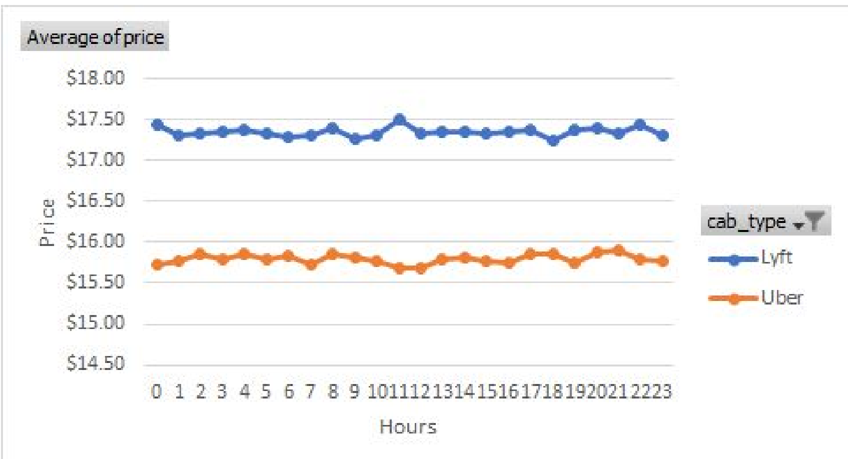

Fig. 12. The effect of time on average price of ride

The price of Lyft is overall higher than the price of Uber, but there is not a certain pattern difference between the price of Uber and Lyft.

Count of id	Column Labels		
Row Labels	Lyft	Uber	Grand Total
0	14657	17756	32413
1	12669	15879	28548
2	12548	16000	28548
3	12224	15591	27815
4	12502	15828	28330
5	11155	13853	25008
6	11935	15401	27336
7	11024	13912	24936
8	10700	13575	24275
9	12948	16023	28971
10	13325	17059	30384
11	13372	17012	30384
12	13633	16751	30384
13	13503	16881	30384
14	13306	17078	30384
15	13394	16989	30383
16	13349	17035	30384
17	13853	16531	30384
18	13404	16980	30384
19	12461	15094	27555
20	11715	15067	26782
21	12480	15252	27732
22	13247	16189	29436
23	14004	17927	31931
Grand Total	**307408**	**385663**	**693071**

Fig. 13. The number of order at different hours

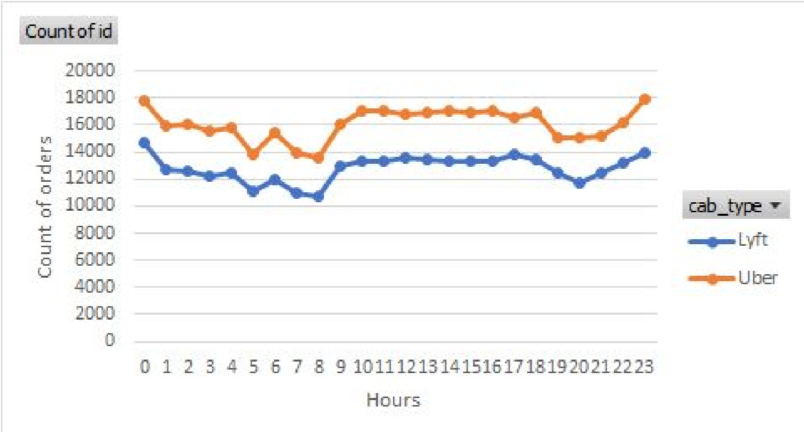

Fig. 14. The number of order at different hours

From the plot above, we can see that the count of orders is the highest during 23 p.m to 12 a.m. During 8 a.m in the early morning, the count of orders is the lowest. Furthermore, we can see that the count of order is lower during the time when people commute to work or school, and the overall count of order is much higher during daytime when people are at work or in school. Overall, Uber has more orders than Lyft.

Comparing Distance and Average Price

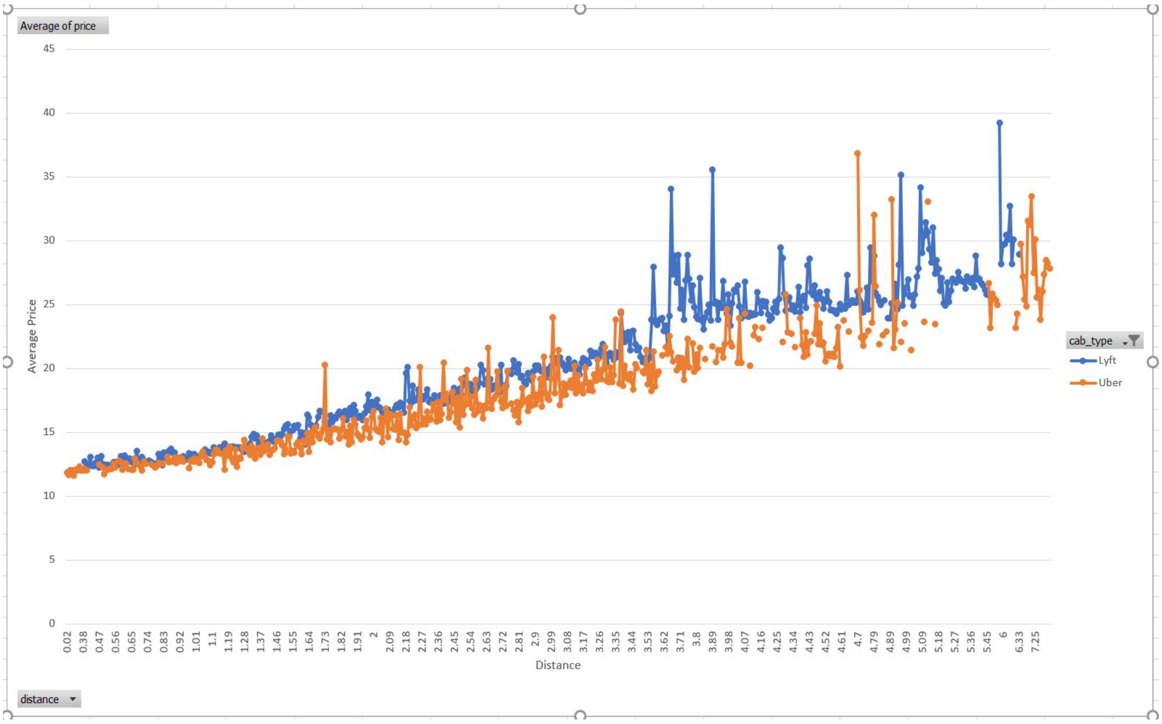

Fig. 15. Average price in different distance

From the distance plot, we can see that in Boston, the average Lyft prices are slightly higher than Uber prices. However, for long distance trips, Lyft and Uber price difference will be slightly higher than low distance trips. If Lyft set their riding price lower than Uber, then Lyft would have more customers, but also Lyft would have fewer drivers.

What is happening in this ride-sharing market is that this industry is the type of industry with huge overcapacity, which means price wars will continue until someone is annihilated, and the last one standing wins enough pricing power to cover their costs.

Comparing Price Per Distance and Distance

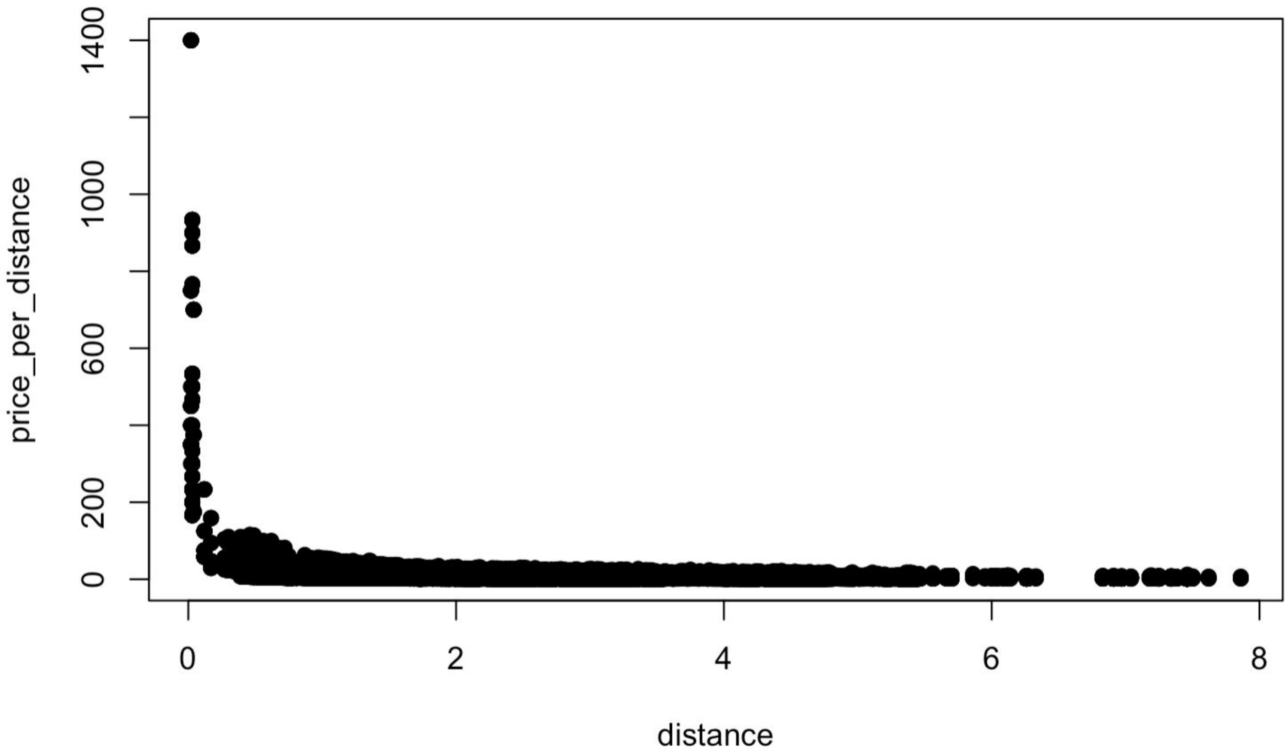

Fig. 16. Price per distance vs distance

We can see from this graph that the price per distance is really high for short distance trips. In order to examine the plot more thoroughly, we see that after the distance exceeds 6 miles, the price does not vary much.

We can see from this plot that the price per distance is really high for short distances.

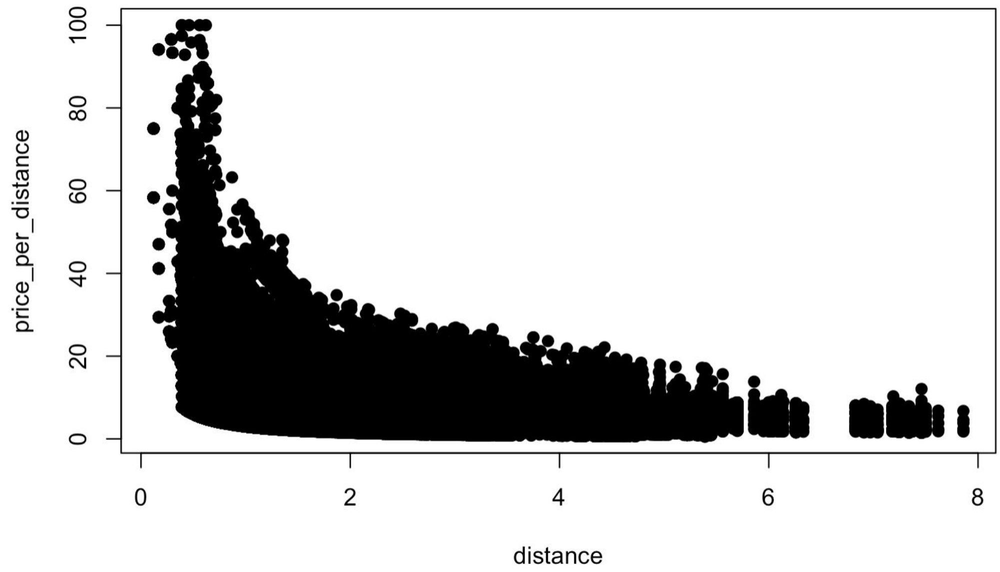

Fig. 17. Price per distance vs distance

(Note: This is a zoom in part of the plot of Price per distance vs Distance, so we can see the trend clearly.)

Then we exam the graph separately for Uber and Lyft.

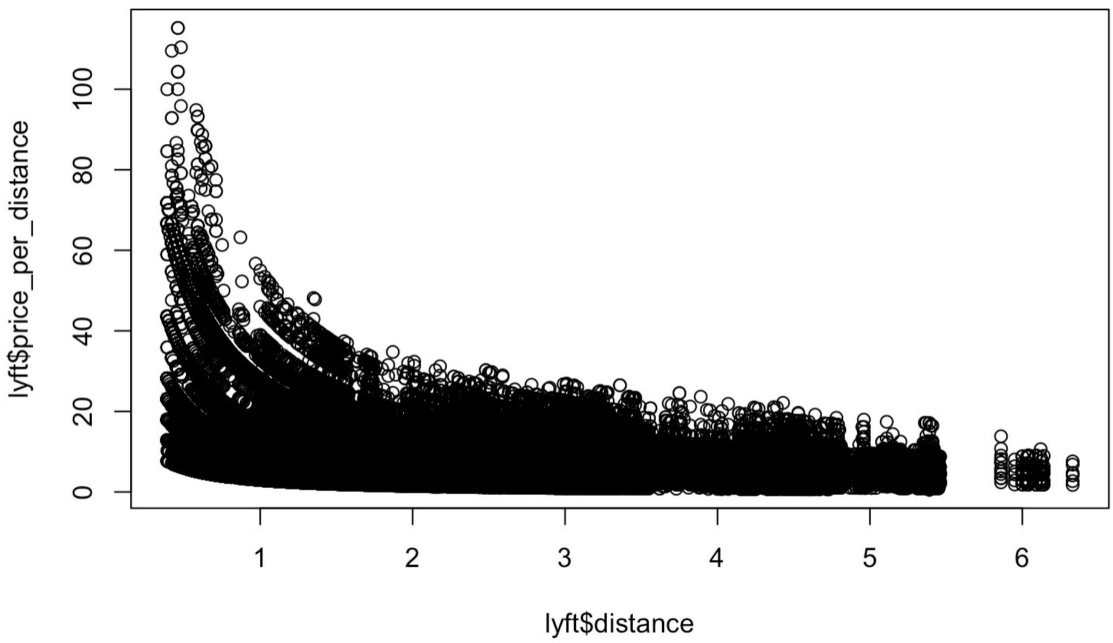

Fig. 18. Price per distance vs distance for Lyft

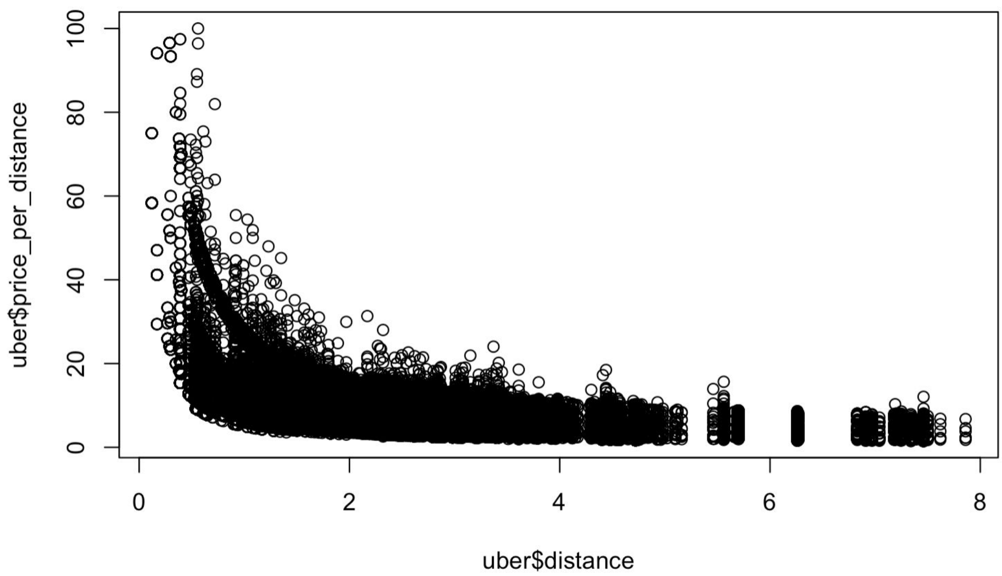

Fig. 19. Price per distance vs distance for Uber

From the plot, we can see that for Uber and Lyft, the trend is pretty much the same. With shorter distances, comes higher price per distance.

Comparing Uber And Lyft Prices in Different Ride Types

Row Labels	Count of id
Lyft	**307408**
Lux	51235
Lux Black	51235
Lux Black XL	51235
Lyft	51235
Lyft XL	51235
Shared	51233
Uber	**330568**
Black	55095
Black SUV	55096
UberPool	55091
UberX	55094
UberXL	55096
WAV	55096
Grand Total	**637976**

Fig. 20. Orders of different products

Row Labels	Average of price
Lyft	**17.35139613**
Lux	17.77124036
Lux Black	23.06246804
Lux Black XL	32.32408607
Lyft	9.610884747
Lyft XL	15.30936274
Shared	6.029892843
Uber	**15.79534317**
Black	20.52378619
Black SUV	30.2867631
UberPool	8.752500408
UberX	9.765074237
UberXL	15.6781436
WAV	9.765019239
Grand Total	**16.54512549**

Fig. 21. Average price of different products

The plots above illustrate the total number of rides and average price of each ride for different types of rides offered by Lyft and Uber. From the right graph, we can see that Lyft has a product called Lux, it is different from Lux Black and Lux Black XL. Lyft Lux's description on the offical website is: High-end, four seats, the perfect way to add a little luxury and arrive in style. For Lux Black: Premium black car service limited to only the most luxurious makes and models. When checking the list of Uber's products, we can find that it only offers, Black and Black SUV - two ride types with luxury automobiles. When Lyft's Luxury products compete with Uber's Luxury products, its advantage is offering one more option than Uber. We also can find that Uber offers WAV, a product that provides affordable rides in wheelchair-accessible vehicles. However, Lyft does not have such ride service.

The Surge Multiplier Variable

The Surge Multiplier variable has seven categories as showing on the table below:

Surge Multiplier	**Count**	**Uber**	**Lyft**
1.00	**672096**	**385663**	**286433**
1.25	**11085**	**0**	**11085**
1.50	**5065**	**0**	**5065**
1.75	**2420**	**0**	**2420**
2.00	**2239**	**0**	**2239**
2.50	**154**	**0**	**154**
3.00	**12**	**0**	**12**

Fig. 22. Surge Multiplier distribution

As the table illustrates, all Uber rides in this dataset does not have any surge multiplier unlike Lyft rides, where 6.8% of its rides have surge multipliers. It is noticeable that there are 12 Lyft rides that have multiplier of 3. In order to justify why the price has increased in those 12 rides, we have analyzed these 12 rides separately and found the below:

· The distance of these 12 trips is between 1.84 and 4.64 miles and the average is 3.11 miles. Therefore, the distance is not the main reason for price surges.

· The temperature was between 31 and 44 fahrenheit and the average was 39 fahrenheit

· 50% of the 12 rides were taken between 12 pm to 4 pm.
· 58% of the 12 rides were taken on 27th and 28th of November, during Thanksgiving Day, which could explain why the price surged.
· The below scatterplot shows the days and the price of the trips in November and December.

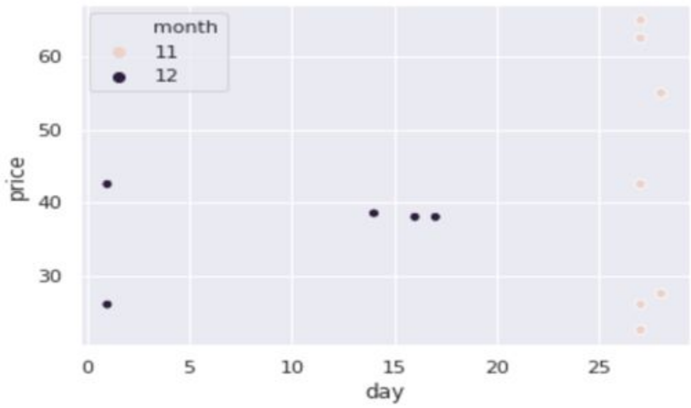

Fig. 23. The 12 Rides with Highest Surge Multiplier

The Fare Price of Uber and Lyft

	Overall	Uber	Lyft
Maximum Price	**97.5**	**89.5**	**97.5**
Minimum Price	**2.5**	**4.5**	**2.5**
Average Price	**16.4**	**15.7**	**17.3**
Maximum Distance	**7.8**	**7.86**	**6.33**
Minimum Distance	**0.02**	**0.02**	**0.39**
Average Distance	**2.1**	**2.19**	**2.18**

The table illustrates that the average price ride of Uber is lower than that of Lyft by almost $1.60. However, if we compare the average price of Uber rides, which does not have any surge multiplier, with the average price of Lyft rides, we notice the average price of Lyft dropped to $16.50. Therefore, the actual average difference between Uber and Lyft is only $0.80.

Price and Distance

According to the table above, it seems there is no significant difference between Uber and Lyft prices when it comes to average distance, even though Uber seems to have trips with longer distances.

The plot below shows that the distance of rides are between 0.02 and 7.8 miles, while the vast majority of the rides are less than 5 miles.

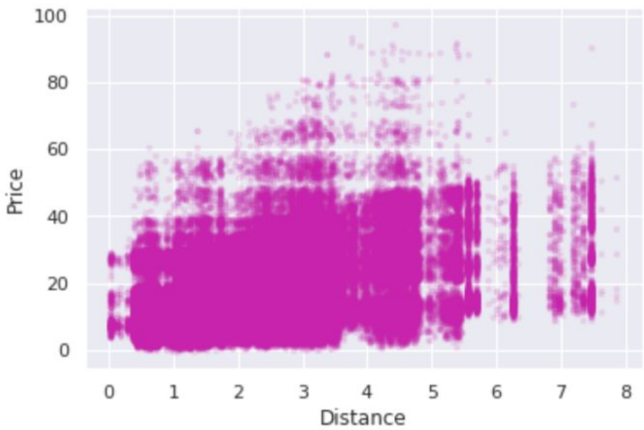

Fig. 24. Rides per Distance

From the plot below, we can see that most of the long distance rides were Uber trips while Lyft trips significantly dominate rides that cost $60 or more.

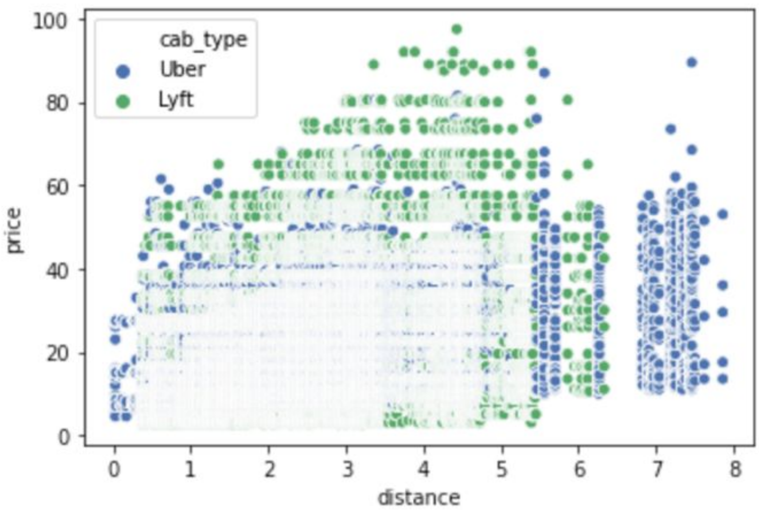

Fig. 25. Ride per Distance by Cab Type

Regression Analysis

Multiple regression models were prepared to test which factors most likely influence fare prices:

First Model:

OLS Regression Results

Dep. Variable:	price	R-squared:	0.000
Model:	OLS	Adj. R-squared:	-0.000
Method:	Least Squares	F-statistic:	0.7541
Date:	Sun, 26 Apr 2020	Prob (F-statistic):	0.687
Time:	03:26:16	Log-Likelihood:	-2.3296e+06
No. Observations:	637976	AIC:	4.659e+06
Df Residuals:	637964	BIC:	4.659e+06
Df Model:	11		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	193.4535	324.145	0.597	0.551	-441.860	828.767
humidity	-0.0252	0.158	-0.159	0.874	-0.335	0.285
pressure	0.0005	0.002	0.268	0.788	-0.003	0.004
visibility	0.0109	0.007	1.528	0.127	-0.003	0.025
windSpeed	0.0066	0.006	1.064	0.287	-0.006	0.019
temperature	-0.0008	0.004	-0.213	0.831	-0.008	0.006
windBearing	-0.0002	0.000	-1.488	0.137	-0.001	7.86e-05
temperatureMax	0.0015	0.005	0.297	0.766	-0.008	0.011
temperatureMin	-0.0011	0.004	-0.242	0.809	-0.010	0.008
sunsetTime	-0.0002	0.000	-0.517	0.605	-0.001	0.000
sunriseTime	0.0002	0.000	0.517	0.605	-0.000	0.001
cloudCover	0.0604	0.044	1.362	0.173	-0.027	0.147

The purpose of the first model is to test how the weather is influencing the price. In order to verify this effect, we have selected the following variables for this regression model: Humidity, Pressure, Visibility, Wind Speed, Temperature, Wind Bearing, Temperature Max, Temperature Min, Sunset Time, Sunrise Time, and Cloud Cover.

After reviewing the P-Value of the selected variables, we conclude that all the weather variables, as they display in the snapshot, are not statistically significant in determining the price of the ride. However, this might not reflect the reality since the dataset was taken for only two months (November and December). Furthermore, the weather during these two months was similar; therefore, we need more data with different climates in order to make sure that this result is not misleading.

Second Model:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  price   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                 -0.000
Method:                 Least Squares   F-statistic:                    0.7857
Date:                Sun, 26 Apr 2020   Prob (F-statistic):              0.828
Time:                        03:29:16   Log-Likelihood:            -2.3296e+06
No. Observations:              637976   AIC:                         4.659e+06
Df Residuals:                  637936   BIC:                         4.660e+06
Df Model:                          39
Covariance Type:            nonrobust
==================================================================================
                     coef    std err          t      P>|t|      [0.025      0.975]
----------------------------------------------------------------------------------
Intercept         14.1954      0.048    294.340      0.000      14.101      14.290
C(month)[T.12]     2.4040      0.042     56.906      0.000       2.321       2.487
C(day)[T.2]       -0.0237      0.065     -0.366      0.714      -0.151       0.103
C(day)[T.3]       -0.1075      0.065     -1.660      0.097      -0.234       0.019
C(day)[T.4]        0.0856      0.100      0.859      0.391      -0.110       0.281
C(day)[T.9]        0.2786      0.249      1.119      0.263      -0.209       0.767
C(day)[T.10]      -0.1696      0.195     -0.871      0.384      -0.551       0.212
C(day)[T.13]       0.0173      0.068      0.256      0.798      -0.115       0.150
C(day)[T.14]      -0.0124      0.065     -0.192      0.848      -0.139       0.114
C(day)[T.15]      -0.0117      0.065     -0.181      0.857      -0.139       0.115
C(day)[T.16]       0.0144      0.065      0.223      0.824      -0.112       0.141
C(day)[T.17]      -0.0748      0.065     -1.156      0.248      -0.202       0.052
C(day)[T.18]       0.0035      0.069      0.050      0.960      -0.131       0.138
C(day)[T.26]       2.3669      0.050     47.251      0.000       2.269       2.465
C(day)[T.27]       2.4217      0.034     70.799      0.000       2.355       2.489
C(day)[T.28]       2.3577      0.035     66.456      0.000       2.288       2.427
C(day)[T.29]       2.3951      0.038     63.640      0.000       2.321       2.469
C(day)[T.30]       2.2500      0.042     53.185      0.000       2.167       2.333
C(hour)[T.1]      -0.0655      0.079     -0.830      0.407      -0.220       0.089
C(hour)[T.2]      -0.0096      0.079     -0.121      0.903      -0.165       0.146
C(hour)[T.3]      -0.0420      0.080     -0.526      0.599      -0.198       0.114
C(hour)[T.4]      -0.0023      0.079     -0.030      0.976      -0.158       0.153
C(hour)[T.5]      -0.0434      0.082     -0.527      0.598      -0.205       0.118
C(hour)[T.6]      -0.0603      0.080     -0.751      0.453      -0.218       0.097
C(hour)[T.7]      -0.0809      0.082     -0.981      0.327      -0.243       0.081
C(hour)[T.8]       0.0308      0.083      0.371      0.711      -0.132       0.194
C(hour)[T.9]      -0.0577      0.079     -0.729      0.466      -0.213       0.098
C(hour)[T.10]     -0.0685      0.078     -0.875      0.382      -0.222       0.085
```

The purpose of the second model is to test whether certain times of the day are influencing fare prices. In order to verify, we selected the following variables and coded them as categorical variables for this regression model: Hour of the day, Day of the month, Months.

After reviewing the regression result, we conclude that the hour of the day is not statistically significant in determining the price. On the other hand, the following factors are the most important factors. Since they have very small P-Values, they are statistically significant in relation to the price:

o Months: As illustrated above, months have a very strong positive correlation (2.4) in relation to price, meaning the average ride price difference between November and December is $2.40.

o Days: As illustrated above, not all days are statistically significant. However, we can easily notice that Days 26, 27, 28, 29, and 30 are statistically significant. Therefore, we can conclude that the average price at the end of month is more expensive than any other days in a month and the price is highly influenced by the last days of the month.

Third Model:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  price   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                 -0.000
Method:                 Least Squares   F-statistic:                    0.7857
Date:                Sun, 26 Apr 2020   Prob (F-statistic):              0.828
Time:                        03:29:16   Log-Likelihood:            -2.3296e+06
No. Observations:              637976   AIC:                         4.659e+06
Df Residuals:                  637936   BIC:                         4.660e+06
Df Model:                          39
Covariance Type:            nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
Intercept        14.1954      0.048    294.340      0.000      14.101      14.290
C(month)[T.12]    2.4040      0.042     56.906      0.000       2.321       2.487
C(day)[T.2]      -0.0237      0.065     -0.366      0.714      -0.151       0.103
C(day)[T.3]      -0.1075      0.065     -1.660      0.097      -0.234       0.019
C(day)[T.4]       0.0856      0.100      0.859      0.391      -0.110       0.281
C(day)[T.9]       0.2786      0.249      1.119      0.263      -0.209       0.767
C(day)[T.10]     -0.1696      0.195     -0.871      0.384      -0.551       0.212
C(day)[T.13]      0.0173      0.068      0.256      0.798      -0.115       0.150
C(day)[T.14]     -0.0124      0.065     -0.192      0.848      -0.139       0.114
C(day)[T.15]     -0.0117      0.065     -0.181      0.857      -0.139       0.115
C(day)[T.16]      0.0144      0.065      0.223      0.824      -0.112       0.141
C(day)[T.17]     -0.0748      0.065     -1.156      0.248      -0.202       0.052
C(day)[T.18]      0.0035      0.069      0.050      0.960      -0.131       0.138
C(day)[T.26]      2.3669      0.050     47.251      0.000       2.269       2.465
C(day)[T.27]      2.4217      0.034     70.799      0.000       2.355       2.489
C(day)[T.28]      2.3577      0.035     66.456      0.000       2.288       2.427
C(day)[T.29]      2.3951      0.038     63.640      0.000       2.321       2.469
C(day)[T.30]      2.2500      0.042     53.185      0.000       2.167       2.333
C(hour)[T.1]     -0.0655      0.079     -0.830      0.407      -0.220       0.089
C(hour)[T.2]     -0.0096      0.079     -0.121      0.903      -0.165       0.146
C(hour)[T.3]     -0.0420      0.080     -0.526      0.599      -0.198       0.114
C(hour)[T.4]     -0.0023      0.079     -0.030      0.976      -0.158       0.153
C(hour)[T.5]     -0.0434      0.082     -0.527      0.598      -0.205       0.118
C(hour)[T.6]     -0.0603      0.080     -0.751      0.453      -0.218       0.097
C(hour)[T.7]     -0.0809      0.082     -0.981      0.327      -0.243       0.081
C(hour)[T.8]      0.0308      0.083      0.371      0.711      -0.132       0.194
C(hour)[T.9]     -0.0577      0.079     -0.729      0.466      -0.213       0.098
C(hour)[T.10]    -0.0685      0.078     -0.875      0.382      -0.222       0.085
```

The purpose of the third model is to determine if choosing Uber or Lyft has a significant difference on the ride price. We also want to measure how the ride type and the surge multiplier are impacting the price. In order to verify these effects, we have selected the following factors and coded them as categorical variables for this regression model: Cab Type, Product ID, Surge Multiplier.

As illustrated in the regression result, we conclude that all the above factors are statistically significant, and we can establish the facts about the price below:

- Choosing Lyft over Uber will in average increase the price by almost \$4.

- Choosing Lyft Lux SUV will in average increase the price by almost \$17.

- When Lyft rides have a Surge Multiplier of 1.25, the average price of the ride will increase by \$4.80 compared to rides that do not have any Surge Multiplier. Having Surge Multiplier of 1.5, 1.75, 2, 2.5, and 3 will raise the price on average by \$10, \$14, \$19, \$18, and \$28, respectively.

Fourth Model:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  price   R-squared:                       0.119
Model:                            OLS   Adj. R-squared:                  0.119
Method:                 Least Squares   F-statistic:                 8.623e+04
Date:                Sun, 26 Apr 2020   Prob (F-statistic):               0.00
Time:                        04:38:24   Log-Likelihood:            -2.2892e+06
No. Observations:              637976   AIC:                         4.578e+06
Df Residuals:                  637974   BIC:                         4.578e+06
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     10.3413      0.024    434.527      0.000      10.295      10.388
distance       2.8337      0.010    293.647      0.000       2.815       2.853
==============================================================================
Omnibus:                    68336.762   Durbin-Watson:                   2.164
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            95053.297
Skew:                           0.855   Prob(JB):                         0.00
Kurtosis:                       3.806   Cond. No.                         6.07
==============================================================================
```

The purpose of the fourth model is to determine how the distance influences the price: the distance has a direct effect on the price. As the distance increased by one mile, the average price will increase by $2.80.

The Final Model

We have chosen only the significant factors to estimate the fare price, so our estimated regression line will include only the following factors (See Appendix for Model Results):

- Cab Type
- Product ID
- Surge Multiplier
- Distance
- Day of the month
- Months

Recommendation

After we analyze the influence of weather, we think that Lyft should use a flexible pricing method based on the weather. In Boston during November and December, we recommend that Lyft focuses on cloudy days. Our suggestions include:

- Offering discounts to riders if they choose Lyft on Cloudy days.
- Matching price with Uber on Cloudy days after customers can offer evidence to prove that Uber's price is lower.

Furthermore, we suggest Lyft should improve its ride type diversity:

- Following Uber and developing a similar product like Uber WAV.
- If Lyft can set the Lyft WAV's price to be lower than the price of Uber's WAV, it can help Lyft earn a good reputation and build brand awareness.

Besides improving product diversity, Lyft also needs to concentrate its resources to promote its breadwinning product, Lyft, by advertising on social media (instagram and facebook), and promoting Lyft's economical characteristics.

Lyft should learn Uber's operational strategy; they can try to decrease their rate. The average price is $16.55, but Lyft's price is $0.80 higher than the average. If Lyft can decrease its price to around $16.55, then it might be more acceptable for customers.

Conclusion

This report studies the pricing strategies of Uber and Lyft. First, we used the dataset to compare the ride price of Uber and Lyft under different weather conditions. Considering the weather conditions, we suggest Lyft drivers to be more active on the foggy days because the average price is higher.

Next, we compared Uber And Lyft's price in different hours and realized the ride request count is highest during midnight, while the count is lowest in the early morning. As a result, we suggest Lyft drivers be more active during late hours.

Then we compared the relationship between distance and price and saw that within a certain distance, the fare price of Uber is higher than the fare price of Lyft. The price per distance is very high for short distance trips.

In addition, we compared Uber and Lyft's prices in different ride services. We discovered that Lyft has a price advantage on its standard product: Lyft.

Finally, we explored the surge multiplier variable, compared Uber and Lyft prices, and built few models to explain ride prices. We can see that the fare price of Lyft has a higher variance but Uber, overall, has a lower price than Lyft. The biggest challenge is pricing their products too low can have a disastrous impact on the companies’ financial performances.

Work Cited

Curley, Robert. "Global Ride Sharing Industry Valued at More than $61 Billion." *Business Traveller*, 4 Jan. 2019, www.businesstraveller.com/business-travel/2019/01/04/value-of-global-ride-sharing-industry-estimated-at-more-than-61-billion/.

Ferenstein, Greg. "Cabs Crushed: Ridesharing Is Worth More than 3x Entire Taxi Market (In 1 Graph)." *Medium*, The Ferenstein Wire, 12 Mar. 2015, medium.com/the-ferenstein-wire/cabs-crushed-ridesharing-is-worth-more-than-3x-entire-taxi-market-in-1-graph-c5522d8e0e26.

Helling, Brett. "Uber Fees: How Much Does Uber Pay, Actually? (With Case Studies)." Ridester.com, 25 Mar. 2020, www.ridester.com/uber-fees/.

Iqbal, Mansoor. "Lyft Revenue and Usage Statistics (2020)." *Business of Apps*, 24 Apr. 2020, www.businessofapps.com/data/lyft-statistics/.

Marshall, Aarian. "A Third of Americans Use Ride-Hail. Uber and Lyft Need More." *Wired*, Conde Nast, 8 Jan. 2019, www.wired.com/story/uber-lyft-ride-hail-stats-pew-research/.

Richter, Felix. "Infographic: How Uber and Lyft Measure Up." *Statista Infographics*, 6 Mar. 2019, www.statista.com/chart/17261/lyft-vs-uber/.

Schneider, Todd. "Taxi, Uber, and Lyft Usage in New York City." *Toddwschneider.com*, 5 Apr. 2016, toddwschneider.com/posts/taxi-uber-lyft-usage-new-york-city/.

Zaveri, Paayal. "Lyft Has Now Delivered 1 Billion Rides." *CNBC*, CNBC, 18 Sept. 2018, www.cnbc.com/2018/09/18/lyft-hits-1-billion-rides.html.

Appendix

OLS Regression Results

```
==============================================================================
===========
Dep. Variable:                  price   R-squared:                       0.928
Model:                            OLS   Adj. R-squared:                  0.928
Method:                 Least Squares   F-statistic:                 2.404e+05
Date:                Sun, 26 Apr 2020   Prob (F-statistic):               0.00
Time:                        04:50:44   Log-Likelihood:            -1.4921e+06
No. Observations:              637976   AIC:                         2.984e+06
Df Residuals:                  637941   BIC:                         2.985e+06
Df Model:                          34
Covariance Type:            nonrobust
==============================================================================
==========================================================
                         coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------
-----
Intercept                      7.3773      0.006   1176.061      0.000       7.365
7.390
C(month)[T.12]                 1.2504      0.011    111.814      0.000       1.228
1.272
C(day)[T.2]                   -0.0368      0.017     -2.110      0.035      -0.071
-0.003
C(day)[T.3]                   -0.0318      0.017     -1.823      0.068      -0.066
0.002
C(day)[T.4]                    0.0180      0.026      0.683      0.495      -0.034
0.070
C(day)[T.9]                    0.0180      0.065      0.276      0.783      -0.110
0.146
C(day)[T.10]                   0.0542      0.051      1.055      0.291      -0.046
0.155
C(day)[T.13]                  -0.0394      0.018     -2.173      0.030      -0.075
-0.004
C(day)[T.14]                  -0.0369      0.017     -2.120      0.034      -0.071
-0.003
C(day)[T.15]                  -0.0151      0.017     -0.869      0.385      -0.049
0.019
```

C(day)[T.16]	-0.0340	0.017	-1.950	0.051	-0.068	0.000
C(day)[T.17]	-0.0096	0.017	-0.549	0.583	-0.044	0.025
C(day)[T.18]	-0.0105	0.018	-0.569	0.569	-0.047	0.026
C(day)[T.26]	1.2325	0.013	94.045	0.000	1.207	1.258
C(day)[T.27]	1.2185	0.009	134.838	0.000	1.201	1.236
C(day)[T.28]	1.2216	0.009	133.458	0.000	1.204	1.240
C(day)[T.29]	1.2344	0.010	124.152	0.000	1.215	1.254
C(day)[T.30]	1.2200	0.011	108.951	0.000	1.198	1.242
C(cab_type)[T.Uber]	-4.9632	0.010	-495.111	0.000	-4.983	-4.944
product_id[T.6c84fd89-3f11-4782-9b50-97c468b19529]	10.7587	0.015	711.737	0.000	10.729	10.788
product_id[T.6d318bcc-22a3-4af6-bddd-b409bfce1546]	20.5218	0.015	1357.605	0.000	20.492	20.551
product_id[T.6f72dfc5-27f1-42e8-84db-ccc7a75f6969]	5.9131	0.015	391.177	0.000	5.883	5.943
product_id[T.8cf7e821-f0d3-49c6-8eba-e679c0ebcf6a]	5.15e-15	2.52e-16	20.431	0.000	4.66e-15	5.64e-15
product_id[T.997acbb5-e102-41e1-b155-9df7de0a73f2]	-1.0125	0.015	-66.982	0.000	-1.042	-0.983
product_id[T.9a0e7b09-b92b-4c41-9779-2ad22b4d779d]	2.463e-05	0.015	0.002	0.999	-0.030	0.030
product_id[T.lyft]	-5.7985	0.010	-562.532	0.000	-5.819	-5.778
product_id[T.lyft_line]	-8.6861	0.010	-840.086	0.000	-8.706	-8.666
product_id[T.lyft_lux]	7.6516	0.010	742.315	0.000	7.631	7.672
product_id[T.lyft_luxsuv]	16.9132	0.010	1640.810	0.000	16.893	16.933
product_id[T.lyft_plus]	-0.1001	0.010	-9.706	0.000	-0.120	-0.080
product_id[T.lyft_premier]	2.3604	0.010	229.002	0.000	2.340	2.381

C(surge_multiplier)[T.1.25]	4.5943	0.024	188.394	0.000	4.546	4.642
C(surge_multiplier)[T.1.5]	9.3103	0.036	261.245	0.000	9.240	9.380
C(surge_multiplier)[T.1.75]	14.1276	0.051	275.565	0.000	14.027	14.228
C(surge_multiplier)[T.2.0]	19.0310	0.053	357.192	0.000	18.927	19.135
C(surge_multiplier)[T.2.5]	17.4404	0.202	86.197	0.000	17.044	17.837
C(surge_multiplier)[T.3.0]	25.9884	0.724	35.880	0.000	24.569	27.408
distance	2.7946	0.003	1009.733	0.000	2.789	2.800