The More Cost-Effective Ride Share App - Uber or Lyft?

The goal of this analysis is to examine a sample data set consisting of simulated Uber and Lyft ride data from the Boston area to see if one ride share app is more cost-effective than the other. The data was retrieved from Kaggle.com (see link below) and is from a 3 ½ week period in 2018. I start off by running some exploratory queries to find out more information about the data set. Then I run some additional queries that help to answer more targeted questions and whether Uber or Lyft is more cost effective.

Note about this data set: Uber and Lyft do not make their data public so to pull this data together the creator of this data set used a custom application to query actual Uber and Lyft data every 5 minutes. This data set consists of simulated rides with what prices would have been if these rides were taken. For a full explanation on how this data was created, see the data source link below.

Data Source: <https://www.kaggle.com/ravi72munde/uber-lyft-cab-prices>

Programs Used: Microsoft SQL Server, Tableau Public

**Query 1: Record Count and Average Price**

SQL Code:

Select cab\_type, Count(\*) AS RecordCount, ROUND(AVG (price),2) AS AvgPrice

From [Cab Rides]

Group by cab\_type

ORDER BY AvgPrice DESC;

Output:

Table

Description automatically generated

Description: This query gives you a quick idea of the data we are working with. It returns the number of records for both Lyft and Uber as well as the average price for all rides.

**Query 2: Average Price and Distance by Rideshare App**

SQL Code:

Select source, destination, cab\_type, ROUND(AVG (price),2) AS AvgPrice,

ROUND(AVG(distance),2) AS AvgDistance, ROUND(AVG(Price/Distance),2) AS CostPerMile

From [Cab Rides]

Group by destination, source, cab\_type

Order by AVG (price) DESC;

Output:

A picture containing graphical user interface

Description automatically generated

Description: This returns the source (or Pickup spot), destination, cab type (Uber or Lyft), average price, average distance, and the average cost per mile from the Cab Rides table. The results are grouped by the destination, source and the cab type and are ordered by the average price in descending order.

Key Takeaways: Here you can see that 7 out of the top 10 most expensive routes are with Lyfts. When comparing routes with the same source and destination between Lyft and Uber, (Financial District to Boston University for example) the average price and cost per mile are higher for Lyfts. (See Boston University to Financial district for another example). One should keep in mind that not all trips with the same source and destination are the same, however by looking at the averages you can see the overall trends that are occurring.

**Query 3: Rides over $50**

SQL Code:

Select cab\_type, Count(\*) AS RecordCount, source, name, ROUND(AVG (price),2) AS AvgPrice,

ROUND(AVG(distance),2) AS AvgDistance, ROUND(AVG(Price/Distance),2) AS CostPerMile

From [Cab Rides]

Where Price > 50

Group by cab\_type, source, name

Order by RecordCount DESC;

Output:

Graphical user interface, application

Description automatically generated

Description: This query returns the cab type, record count, source, name, average price, average distance, and the average cost per mile for all cab rides costing over $50. The results are grouped by the cab type, source and name and ordered by the record count in descending order.

Key Takeaways: With this query I wanted to see which cab type had the most records where the price was greater than $50. What is interesting is that out of the top 10 source/name combinations, 9 of them are Lyft’s. Also, the cost per mile varies greatly, which is to be expected knowing the average distance varies greatly as well.

**Query 4: Rideshare App Comparison by Weekday and Hour**

SQL Code:

Select DATENAME(WEEKDAY, [Date-Time (EST)]) AS Weekday,

DATEPART (hour, [Date-Time (EST)]) [hour], cab\_type,

ROUND(AVG (price),2) AS AvgPrice, Count(\*) AS RecordCount

From [Cab Rides]

WHERE cab\_type = 'Lyft'

Group by DATEPART(hour,[Date-Time (EST)]), cab\_type,

DATENAME(WEEKDAY, [Date-Time (EST)])

ORDER BY AvgPrice DESC;

Output:

Table

Description automatically generated

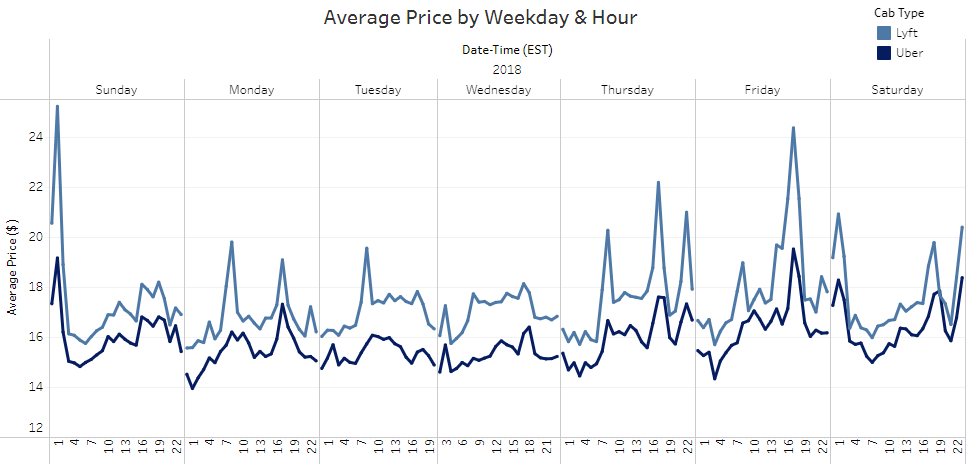
Description: This query returns the weekday, hour, cab type, average price, and the record count. The data is sorted by average price starting with the most expensive rides. Below are the results when the query is run for Lyft (Left) and for Uber (right).

Key Takeaways: With this data we can see that unsurprisingly the most expensive rides occur between Thursday and Sunday, due to there most likely being a higher demand these days. Also, the peak times are:

* + Thurs/Fri from 4pm – 6pm (8 out of the top 20)
  + Sat/Sun from 11pm – 2am (8 out of the top 20)
  + Sat from 6pm – 7pm (2 out of top 20)

Again, Lyft came out as clearly having the higher average price during these peak times. With the next few questions, we will look further into this price disparity.

The graph below gives you an idea of when the prices typically spike, and again you can see that the average ride cost is higher with Lyft.



**Query 5: Cost per Mile among Ride Types**

SQL Code:

Select cab\_type, name, ROUND(AVG(Price/Distance),2) AS CostPerMile,

ROUND(AVG(Price),2) AS AvgPrice, MAX (price) AS MaxPrice, MIN (price) AS MinPrice,

(MAX(Price) - MIN(Price)) AS MaxDifference

From [Cab Rides]

Group by cab\_type, name

Order by CostPerMile DESC;

Output:

Table

Description automatically generated

Description: This query returns the cab type, name, cost per mile, average price, maximum price, minimum price, and the maximum difference between the highest and lowest priced ride. The results are grouped by cab type and the name (vehicle type) and ordered by cost per mile.

Key Takeaways: From previous queries we saw that Lyft has had the consistently higher prices when compared to Uber. However, when digging deeper and comparing Lyft and Uber’s ride options on a cost per mile basis, you see some varied results. With the table above you see that both of the Lyft XL options (Rows 2, 7) are cheaper than the equivalent options at uber (Rows 1, 6) on a cost per mile basis.

Additionally, when comparing the standard and shared options for both apps, Lyft is $0.57 and $2.04 cheaper than an Uber on a cost per mile basis. We can see that Lyft is the more cost-effective option for these four ride types based on the average cost per mile, which is a better indicator of the true cost than the average price. The two remaining ride options for both apps are the Uber Black, Uber WAV, Lyft Lux Black, and the Lyft Lux options. Uber comes away as the more cost-effective app for these two ride options. When you put this all together you can see that Lyft has the edge when using the cost per mile metric.

As we recall from the first few queries, Lyfts consistently had the higher priced trips. What I want to investigate now is how that could be the case if Lyfts have a lower cost per mile for four out of the six ride options.

**Query 6: Overall Cost per Mile by Rideshare App**

SQL Code:

Select cab\_type, Count(\*) AS RecordCount, ROUND(AVG (price),2) AS AvgPrice,

ROUND(AVG(Price)/AVG(distance),2) AS CostPerMile,

ROUND(AVG(distance),3) AS AvgDistance

From [Cab Rides]

Group by cab\_type

ORDER BY AvgPrice DESC;

Output:

Application, table

Description automatically generated

Description: This query returns the cab type, record count, average price, cost per mile and the average distance for all records. The results are grouped by the cab type.

Key Takeaways: In the table above, you can see that for all rides the overall average cost per mile is $0.72 higher for Lyfts than it is for Ubers. This is why on average Lyft rides tend to be higher priced. Does this mean that Uber should be the preferred choice when price is the primary concern? Let’s take a deeper look into the numbers.

**Query 7: Cost per Mile Side by Side**

SQL Code:

Select cab\_type, name, Count(\*) AS RecordCount,

ROUND(AVG(Price/Distance),2) AS CostPerMile

From [Cab Rides]

WHERE cab\_type = 'Lyft'

Group by cab\_type, name

Order by CostPerMile DESC;

Output:

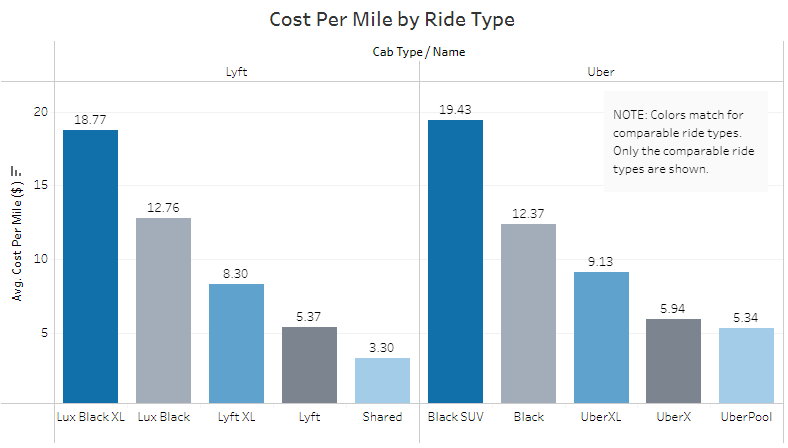
A screenshot of a computer

Description automatically generated with medium confidence

Description: Similar to query 5, this query returns the cab type, name, and the cost per mile as well as the record count. The Lyft data is on the left, the Uber data is on the right, and both are ordered by the cost per mile.

Key Takeaways: When you look at the table above, you will notice that for four out of the five comparable ride options highlighted in blue, Lyft actually has a lower cost per mile. For the last ride option highlighted in yellow, Uber has the lower cost per mile (Lyft Lux vs Uber WAV).

Also, see the bar chart below for a visual of all five comparable ride types and their average costs per mile.



It is worth pointing out that the only ride options that are not comparable between the two apps are the Lyft Lux and the Uber WAV (Wheelchair accessible Vehicle) options. These have two very different purposes and account for a good portion of the difference in the overall costs per mile. These two differing ride options are the major reason we see that Lyft has the higher prices on average, but when looking at the ride options individually, Lyft more often than not has the lower cost per mile. The query below shows what happens when I filter out the Lyft Lux and Uber WAV records. Now the overall cost per mile is much closer.

**Query 8: Overall Cost per Mile excluding “WAV” and “Lux”**

SQL Code:

Select cab\_type, Count(\*) AS RecordCount, ROUND(AVG (price),2) AS AvgPrice,

ROUND(AVG(Price)/AVG(distance),2) AS CostPerMile,

ROUND(AVG(distance),3) AS AvgDistance

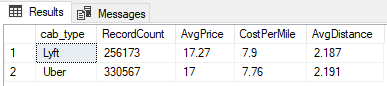
From [Cab Rides]

WHERE name <> 'WAV' AND name <>'Lux'

Group by cab\_type

ORDER BY AvgPrice DESC;

Output:



Description: This query returns the cab type, record count, average price, cost per mile and the average distance. All records where the name equals Lyft Lux and Uber WAV were filtered out.

Key Takeaways: Initially from the first few queries it seemed like Uber was consistently the cheaper rideshare app when looking at average prices. After doing some further investigation and looking at the cost per mile metric, we saw that on average, Lyft was the better priced app for four out of the six ride options. Lyft has a lower cost per mile for the two XL ride options, the standard economy option and the shared or pool option. Uber only had the edge with the luxury black and the WAV options. The more cost-effective app really depends on which ride option you choose but given that Lyft had a lower cost per mile for a majority of the options and the cost per mile across all comparable options was very close, I give the slight edge to Lyft as the more price friendly option.