# Pyber Analysis

## Project Overview

The purpose of the project is to analyze ridesharing data accumulated at PyBer to determine whether actions can be taken to improve access to ride sharing services and better determine rideshare affordability for underserved neighborhoods.

The Pyber data includes total rides, total drivers, total fares, average fare per ride and driver, and total fares. The data is filterable by city type: Urban, Suburban, and Rural. As the data is proprietary, it is assumed to be accurate. It is also assumed that fare rates are the same for all city types.

## Results

Chart, line chart

Description automatically generated

The chart above illustrates that Urban Total Fares are greater than Suburban and Rural throughout the period reviewed. This is as expected and is not abundantly informative.

Analyzing fares with respect to the number of drivers and rides in each city type, we can determine the differences in ride sharing data among the different city types.

The pie charts below reflect the proportion of city types for Total Rides, Total Drivers, and Total Fares.

Chart, pie chart

Description automatically generated Chart, pie chart

Description automatically generatedChart, pie chart

Description automatically generated

Urban Rides and Fares comprise about two-thirds of their respective categories. However, Urban drivers account for a larger proportion (80%) of total drivers. This means that each Urban driver gets a thinner slice of their, albeit larger, ride and fare pies than their Suburban and Rural counterparts.

Table

Description automatically generated

As the pie charts suggested, Average Fares per Ride and Driver for Urban drivers are smaller than for Rural drivers. Suburban Average Fares fall in between. One might suspect that Rural drivers may choose the occupation for different reasons than Urban drivers.

Viewed as a bubble chart below, where circle size correlates with driver count per city, we readily see that many more Urban drivers are sharing a larger number of rides, but at lower fares than Rural drivers.

Chart, scatter chart

Description automatically generated

The three box-and-whisker charts below are used to more closely examine the statistical differences among city types.

Chart, box and whisker chart

Description automatically generated Chart, box and whisker chart

Description automatically generatedChart, box and whisker chart

Description automatically generated

If we compare the average number of rides between each city type, we observe that the average number of rides in the Rural cities is about 3.5 and 2.5 times lower than Urban and Suburban cities, respectively.

The average fare for rides in the Rural cities is about $11 and $5 more per ride than the Urban and Suburban cities, respectively.

The average number of drivers in Rural cities is nine to four times less than in Urban and Suburban cities, respectively.

Considering that Rural areas have nine time less drivers, but only 3.5 less rides with about 1.4 times larger fares, it would appear that Rural drivers generate more revenue per driver than Urban drivers – perhaps as much as three times as much.

## Summary

In Rural cities, there are less rides, but for greater distances when compared to Suburban, and more so Urban cities. This is likely because in Urban areas, the population significantly exceeds that of Suburban and Rural cities. Furthermore, it is likely that Urban riders are less prone to drive their own cars into the city, wanting to avoid parking fees and well as the stress of inner city driving. The Urban rider would thus be more likely to ride share for the shorter inner-city distances. Conversely, in the Rural cities, destinations are less clustered, and one might think that parking is much less of an issue than in Urban cities.

While that data provides insight on the differences between Urban, Suburban and Rural revenue generation, more information would enhance our ability to improve access to ride sharing services and better determine rideshare affordability for underserved neighborhoods.

We submit the following recommendations.

### Recommendations

1. Gather data on wait times for riders and reasons for ride. Wait time data should be readily available in PyBer’s database, however, reason for ride may be only attainable by survey. The current analysis is not greatly informative on shortcomings of riders’ access to rides. This information could assist in that effort.
2. Conduct a survey of riders on their satisfaction with access to and the cost of ride sharing. If the “reason for ride” information is not available in PyBer database, the questions could be combined in this survey. This information would be informative in improving access and service quality.
3. Gather data on the number of hours driven per ride and driver. This data should be readily available in PyBer’s database. Determining whether there are differences would provide insight on the income earned per hour by drivers by city type.
4. Conduct a survey of drivers on their reasons for choosing ride sharing as an occupation. This would provide insight on whether there are significant differences in the driver motivations between city types.

Why do you think there is such a big difference? By looking at the number of riders for each city, can you get a sense of the overall revenue?

By looking at the driver count data and fare data, can you get a sense of the overall revenue?

Chart, pie chart

Description automatically generatedChart, pie chart

Description automatically generatedChart, pie chart

Description automatically generated

Chart, line chart

Description automatically generated