# PyBer Ride Share Analysis

## Analysis Description

Pyber is a startup company in the ride sharing space. They operate in about 120 different cities in the U.S. The CEO is preparing to meet with PyBer's investors and present progress on growth of the company, services being provided, and plans for expansion to gain more market share.

The CEO, (a data scientist) has asked the PyBer Analytics dept to analyze the last 4 months of ride data, look for trends and correlations between fares, city types, and other insights as the data may reveal. This analysis presents those findings.

## Resources

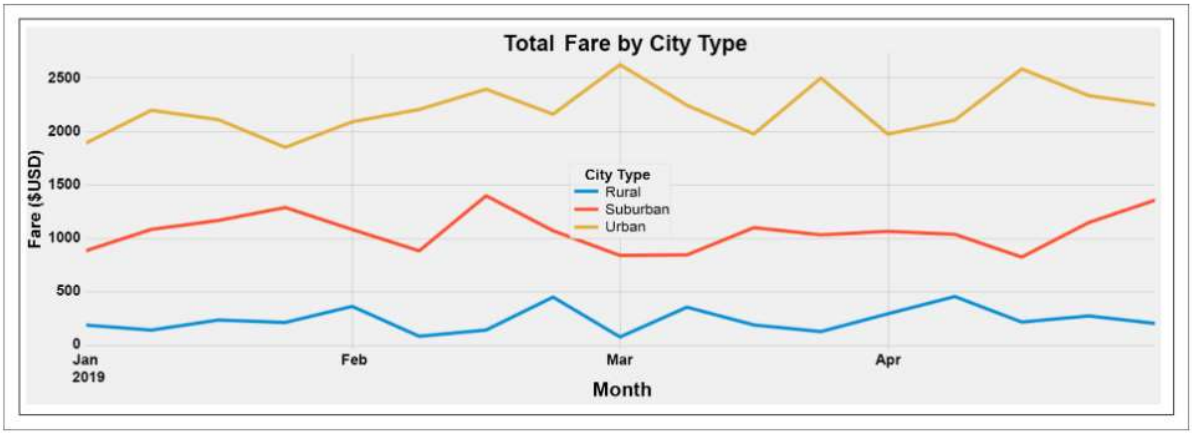
- Python 3.6
- Import Libraries Pandas, Numby, Matplotlib, SciPy,
- Data file ride_data.cvs
- Data file city_data.cvs

Summary Analysis

| | Total Rides | Total Drivers | Total Fares | Average Fares per Ride | Average Fares per Driver |
|---|---|---|---|---|---|
| **Rural** | 125 | 78 | $4,327.93 | $34.62 | $55.49 |
| **Suburban** | 625 | 490 | $19,356.33 | $30.97 | $39.50 |
| **Urban** | 1625 | 2405 | $39,854.38 | $24.53 | $16.57 |

From the data above, we can draw the following conclusions:

- Total Rides and Revenue increases considerably as we move from sparsely populated Rural areas to Suburban to densely populated Urban areas.
- However, the rise in revenue is not linear – that is, when ridership doubles lets say, we do not see a corresponding doubling of Fare revenue.
- That is because Avg Fares go down as Total Rides goes up. In addition, Drivers do not make as much per ride in Urban areas, either. This is troublesome for Drivers trying to earn enough money in Urban areas, where their cost of living is higher – than say in a Suburban areas.
- We might infer that it may be more difficult to recruit, train, and retain top driving talent in the more densely populated areas. With so much news attention on the dissatisfaction of drivers in this space – due to the high cost of living coupled with low earnings – PyBer Management may need to brainstorm additional ideas or programs for retaining top talent.




Multiple-Line Chart

The total Fare by City Type, shows the sum of the Fares – in weekly increments – over a 4 month period. Each line represents a different city type – that is – Rural, Suburban, and Urban. From this data we can draw the following conclusions:

- City Types are not tightly correlated with each other, from the standpoint that peaks in one type are not in sync timewise with peaks in other City Types.
- Some natural variations seem to occur (fares are up and down from week to week) in all City Types.
- Avg Revenue is basically flat over the 4 month period – ending about the same as when it started in January. For a startup, this is not good, as investors generally want to see significant percentage increases in revenue and marketshare.
- The natural tendency would be to expand where revenue and ridership is the highest – that is in the Urban centers. However, it is important to not just look at revenue and ridership – but who and where are the highest areas of “profitability.” Since cost data was not provided to us for this analysis, it would be wise to proactively ask the CEO for access to this data so a more comprehensive analysis can be completed.