

1. OVERVIEW OF THE ANALYSIS

In this module, I landed a job as a data analyst for PyBer, a Python based ride-sharing app company. My first assignment will be to perform an exploratory analysis on data in some very large csv files. To aid this process, I will create several types of visualizations to tell a compelling story about the data. I will write Python scripts using Panda's libraries, the Jupyter notebook, and Matplotlib to create a variety of charts that showcase the relationship between the type of city, and the number of drivers and riders, as well as the percentage of total fares, riders and rivers by type of city.

PURPOSE: Help PyBer improve access to ride-sharing services and determine affordability for underserved neighborhoods.

2. RESULTS OF THE ANALYSIS

1. KPI analysis as a function of city type:

From **Table 2.1**. PyBer Summary Percentage Change we can conclude that our business is highly concentrated in urban areas, being ~68% of our business and ~80% of our drivers, bringing ~63% of our revenue, despite having more public transport options and distances being smaller (thus the lowest average fare per ride). We have managed to settle on urban areas (demonstrated by the lowest average fare per ride), reaching economies of scale.

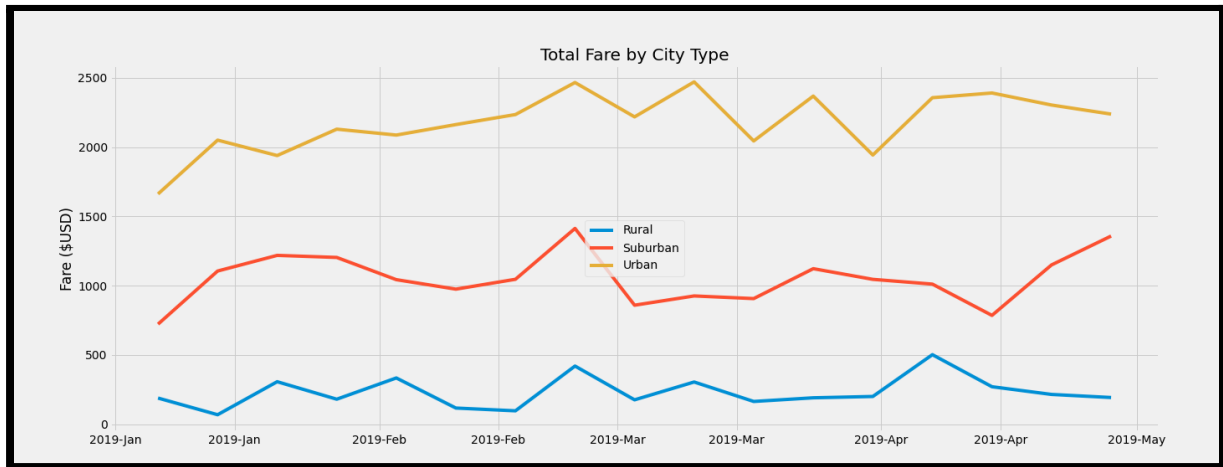
Table 2.1. PyBer Summary Percentage Change

	total_rides	total_drivers	total_amount	average_fare_per_ride	average_fare_by_driver
Rural	125	78	\$4327.93	\$34.62	\$55.49
Suburban	625	490	\$19356.33	\$30.97	\$39.50
Urban	1625	2405	\$39854.38	\$24.53	\$16.57

2. Fare analysis as a function of time:

From **Image 2.1**. Summary of Fares by City Type as a Function of Time we can conclude that the trend more or less remains equal no matter the city type. Peaks, and valleys are more noticeable in urban cities and less noticeable in rural cities. Speaking of our original purpose, affordability PyBer hasn't been necessarily successful at making it's services cheaper, as there is no clear downwards trend. Nevertheless, it has been successful to remain "constant" with the prices it offers.

Image 2.1. Summary of Fares by City Type as a Function of Time



3. SUMMARY OF THE ANALYSIS

Based on the results, provide three business recommendations to the CEO for addressing any disparities among the city types:

With the obtained results, it seems to me that our past strategy has been focused on urban areas. Nevertheless, if we want to improve access to ride-sharing services and improve affordability for underserved neighborhoods I would suggest our CEO to:

1. Determine cost per kilometer in three city types to determine whether or not we are aligned with our original purpose to offer cheap services to underserved communities. Average fare in rural areas is the highest but because distances tend to be longer.
2. There is a great potential to tackle suburban and rural areas where our business has not matured in these geographies. With 20% of urban drivers, suburban drivers generate half of the revenue, which may suggest that there is still unexploited potential in this market.
3. There may also be potential in rural areas as there are less public transport options and distances tend to be further apart (which brings higher revenue). I would consider the market size in some of these rural areas and incentivize drivers to jump into our platform, thus assessing our ride-sharing access concerns.