PyBer Analysis:

Note: I know I am off with some of my data, particularly the driver counts per city. I spent hours on figuring this out; I played around with different types of merging, tried changing the s value (did totdrivers\_urban\*0.5, etc), tried using urban.groupby(["city"])["driver\_count"].count() and urban.groupby(["city"])["driver\_count"].value\_counts(), and even used difflib in case some of the cities had typos in them. But nothing worked. And when I view pd.DataFrame(urban.groupby(['city']), I notice that the contents of the cities are not actually grouped together. Literally have no idea what to do. Regardless, I can tell that there are some obvious trends…

1. There are obviously more rides in urban cities, which could be attributed to the significantly higher population sizes in urban areas compared to that of rural or suburban areas. Also, there are fewer people in urban areas than in urban or suburban areas who have a car due to costs of maintenance and inconvenience. It would also follow that there are more drivers in urban areas (although my numbers are off) since there is greater demand for rides.
2. The average fare is, on a whole, lower in urban areas than rural areas. This could be attributed to the idea that people in rural areas need to driver longer/greater distance to reach their destination. Rural areas tend to be more spread out while urban areas are more concentrated/compact.
3. A majority of PyBer’s sales comes from urban rides.