**Source Data Modification Comments:**

When loading the data, there are 2375 rows in ride\_data.csv, so I wondered why did I have 2407 rows when I merged the city and ride data? The answer was that there were two records for Port James in city\_data.csv, one with 3 drivers and the other with 15 drivers. Since, I was going to use this data set in future processes, I made the decision to manually remove one record and change the driver count to 18

for the remaining record. This change resulted reasonable records counts when merging, summing,

or manipulating the data sets.

**Data Trend Analysis:**

1. It might seem obvious but there are more people using Pyber in an urban setting. It makes sense that people in a city may either not have a car or don't want to use it in the city due to traffic or parking and using a ride service can be convenient and cost effective. On the other hand, people in rural and suburban environments live there because they want to be in an non-urban environment, and since public transportation services are limited to non-existent, inconvenient, and expensive outside of cities, these people would be mostly self-sufficient in the transportation department, so they would use the service less. Drawing on my experience and knowledge, people in rural and suburban areas would use ride services in special situations like going to the airport where you might not want to leave your car, and since services that older people use are in urban to suburban areas, they would in many case have the option to use a ride service to get to those services. And the difference between suburban and rural usage could be related to how close suburbanites are to a city and its amenities, greater availability of public transportation, and ride services like Pyber ride share.
2. Cost surprised me a little because there is a pretty good distribution of rides to cost across the three types. There were some outliers, but they weren't numerous and didn't seem a concern. I would have guessed that urban users would do more and shorter trips resulting in more rides and more at a lower cost than other groups, and that was displayed in the data. I never have used a Pyber type ride service, so I don't know about the costs, but I have to think that especially rural and suburban users, cost would impact the number of rides they take which supports my assumption that usage would mainly be for special occasions or services. Another possible cost savings and incentive for urban people would cost sharing that is possible in that setting. If we added a few more data points, we could have a lot of interesting information, and information that might be helpful in advertising for business growth.
3. As for the number of drivers, I was surprised by the size of some the gold bubbles because these are related rural cities, and I would expect fewer drivers but that wasn't the case. It seems from looking at the color and size of the data that there are an adequete number of drivers in a lot of cases for rural drivers. Certainly, the suburban users had enough drivers to choose from, and usage correlates or drives(supply and demand) the need for drivers which is also shown by the urban bubbles. I would expect more bubbles and bigger bubbles in a city and the bubble chart supports this expectation. From television, not experience, I have an idea that it is hard for people to get a taxi, so if you look at the bubbles there is certainly more demand and supply as the environment moves from rural to urban. Is the real or perceived lack of taxis driving the demand for a Pyber like service? It might be good to look further into why people choose a Pyber type service over a taxi. In any case, the number of drivers looks adequete for all environments.
4. In conclusion, this was interesting data, and I can think of other information that might make it more useful and informative, like male/female usage, reason for usage, other alternatives to Pyber, satisfaction/dissatisfaction with the service, and crime statistics just to name a few.

**Pie Charts:**

1. Total Fares by City-Type:

The pie chart for Total Fares by City-Type shows what I would have expected in that Urban people use the ride service the most, followed by Suburban, and Rural. The percentage of fares to city type is reasonable and expected so no outlier situations are obvious. The percentages were created by summing the number of fares for each city type, and then dividing each sum by the total of all fares. This is also known as: part / whole = %.

1. Total Rides by City-Type:

Again the data displayed results that were expected, and not a whole lot of other information. Pie charts seem like big picture tools, so they are doing their job in this case. There were more rides in Urban areas than Suburban areas and the same with Rural areas. This isn't a surprise, but could lead to a need to look at the data or data collected differently. It is a good start. . The percentages were created by counting the number of rides for each city type, and then dividing each count by the total of all rides. This is also known as: part / whole which is the %.

1. Total Drivers by City-Type:

The pie chart percentages reflect expectations, other charts, and knowledge gained from doing this project. The number of drivers for Urban areas is the greatest with Suburban next, and Rural bringing up the rear. The data makes sense and reflects supply and demand. What isn't visible in this chart but can be seen in a bubble chart is that there is a surprising number of drivers available to Rural users of the service, although that would depend on where they live. The percentages were created by counting the number of drivers from the city\_data.csv because this has one driver count per city without doing any work. The driver counts were grouped by city, by type and each group count was dived by the total driver count. This is also known as: part / whole which is the %.