****PYBER  
 Data Analysis****

****KAVYA SHABNAVEEs****

****HOME ASSIGNMENT – 5****

**INTRODUCTION:**

As part of analyzing trends from a dataset containing ride data for multiple cities have been provided, certain observations have been explained further in this document.

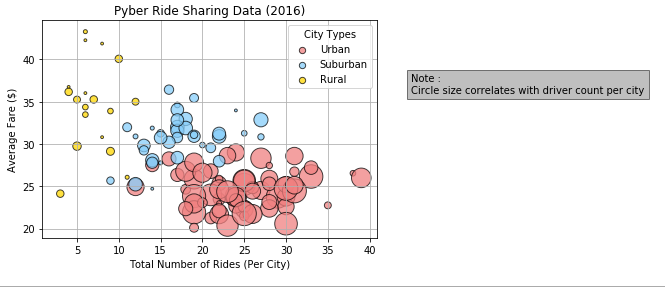
**CONTENTS:**

* Github/iPython Notebook file link
* Observations/Conclusions
* Limitations
* **GitHub/I Python Notebook file link:**
* **Observations/Conclusions**

1. The plot below shows total number of rides and average fare

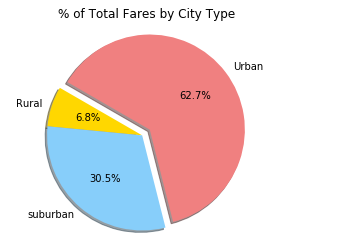
For each city.

* We could observe a major distribution of rides in urban areas compared to other areas, which explains demand for rides in urban cities.
* We could also notice that rural cities have managed to record highest average fare with low number of rides compared to other two, which explains higher ride fares in rural cities.
* Also, most of the suburban cities have maintained fare averages in range $30-$35



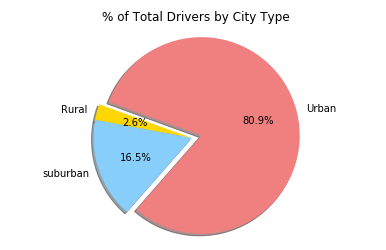
2. Below pie graph shows % of Total fares contributed by urban, suburban and rural areas.

* Urban cities have together generated 63% of revenue which is much higher than combined percentages of rural and suburban.



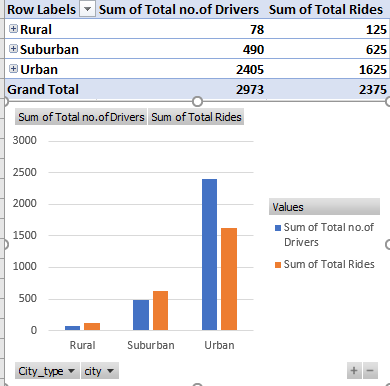
3. Below chart shows percentage of total drivers per each city type.

* Urban cities are observed to have more numbers of drivers than other type of cities.



4. The table with a bar chart below helps us analyze further on supply and demand related data for each type of city.

* It is clearly noticeable that drivers count is less than the number of rides in rural and urban cities whereas it’s a vice versa in case urban cities.



* **CONCLUSIONS:**
* If, we could further try and keep balance in drivers count based on the recent demand trend in respective type of cities, profits could be expected.
* We could extend market to some more cities in rural and urban as the average fare for these cities is higher than urban.
* From the provided set of data, urban cities are found to be high profit yielding cities. We could closely monitor on demand of specific cities in urban areas and distribute drivers accordingly could be more useful.
* Also reducing fares in rural areas could attract more customers.
* **LIMITATIONS:**
* The current dataset did not provide miles travelled for ride data. Calculating fare per mile would have provided another dimension of better insight to dataset.
* For example, if the fare per mile is low and miles travelled is high, this could hinder our analysis with rural and suburban areas.
* Likewise, if fare price is moderate and miles travelled is low, our urban data analysis would have proved our predictions wrong.