

**Final Project: Analysis of Bike Rental Data**

**ALY6070 80809 Communicate/Visual Data Analysis SEC 01**

**Spring 2019 CPS**

**Group Members:**

* Chivatam Lakshmi Aishwarya
* Poulomi Chakraborty
* Shivam Chauhan
* Raviteja Vakkanthula
* Shakya Ashok Munghate

**Index**

* **Introduction**
* **Research Questions**
* **Analysis**
* **Conclusion**
* **Reference**

**Introduction**

In this project, we are have explored and analyzed a dataset that contains data about a bike rental company. Online bike rental companies have users who can be registered (registered in the company website) and casual (renting a bike as a guest in the company website). The bike rental companies offer their users rent a bike from a specific location and the bike can be returned to the same location or any other location. The number of customers renting bikes from the company website typically depends on several factors like whether it is a working day or a non-working day, whether it is a holiday or not a holiday, the time of the day, weather, temperature etc. Here, our goal is to analyze the dataset and then build the most useful research questions that can be answered by the data and will help the bike renting company to maximize their profit and then create graphs and charts to show our findings about the patterns and trends found in the dataset.

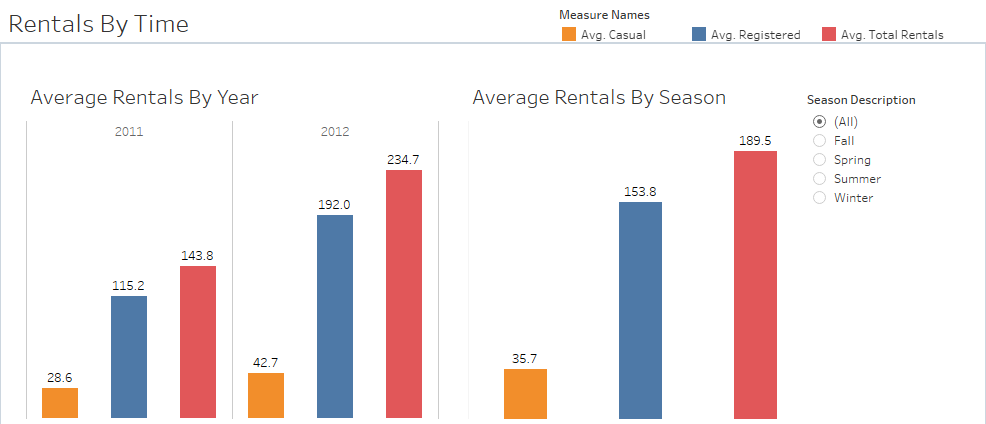
**Research Questions**

* How do the years and different seasons in a year impact the number of bike rentals?
* How does the number of bike rentals vary from month to month in a year?
* Does the average number of bike rentals differ on weekdays compared to weekends?
* Do we have any peak rental hours during the day?
* What is the effect of temperature and humidity on the number of bike rentals?
* What is the effect of weather on the number of bike rentals?

**Analysis**

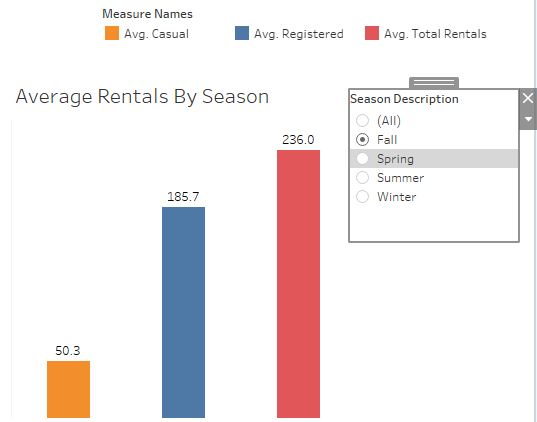
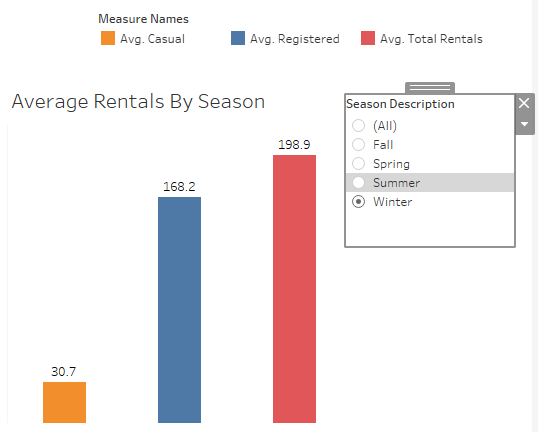
To get accurate outcomes, before starting the analysis we cleaned the dataset using R studio. We removed the Null and invalid values from the dataset. After that, we generated the below graphs and answered the research questions based on the clean dataset.

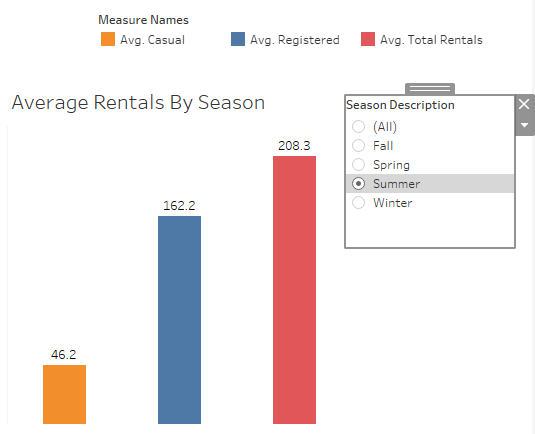
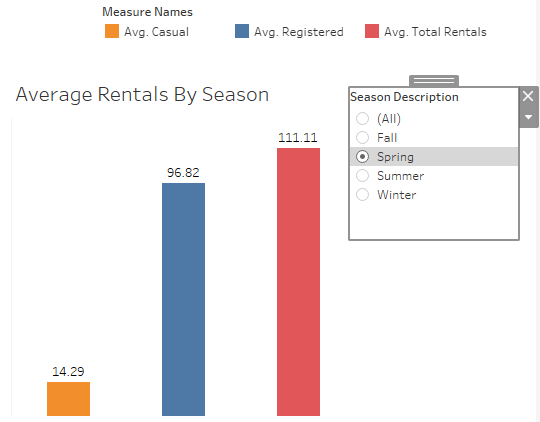
**Graph:1:**



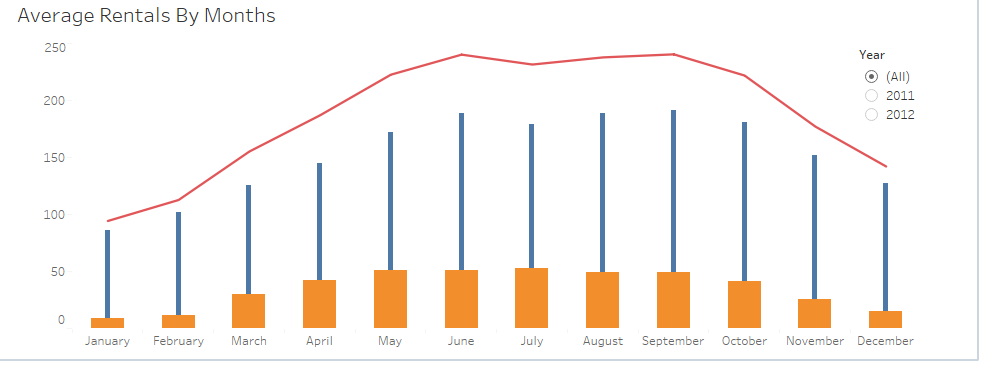
We have created a graph to understand the effect of the time on the number of bike rentals. In our dataset, we have the year in which a bike was rented as well as we have the season of the specific bike rental. So here, we have created two graphs to check if there is any significant impact of both the year and the season on the number of bike rentals. The first graph shows how many bikes were rented in each year (2011 and 2012) by both casual users and registered users as well as the average number of total rentals in mentioned for each year and the second graph shows how the number of bike rentals is impacted by the each of the four seasons – Fall, Spring, Summer and Winter for both the casual and registered users as well as the total average rental for the specific season is mentioned.

In the first graph we can see that the number of bike rentals in 2011 is significantly higher than the number of bike rentals in 2012 for casual and registered users as well as the total average of bike rentals. In the second graph, we can see that the number of bike rentals is highest over the fall season compared to any other season in the year. On the other hand, the number of bike rentals are lowest in the Spring. The outcomes found for each season are shown below.

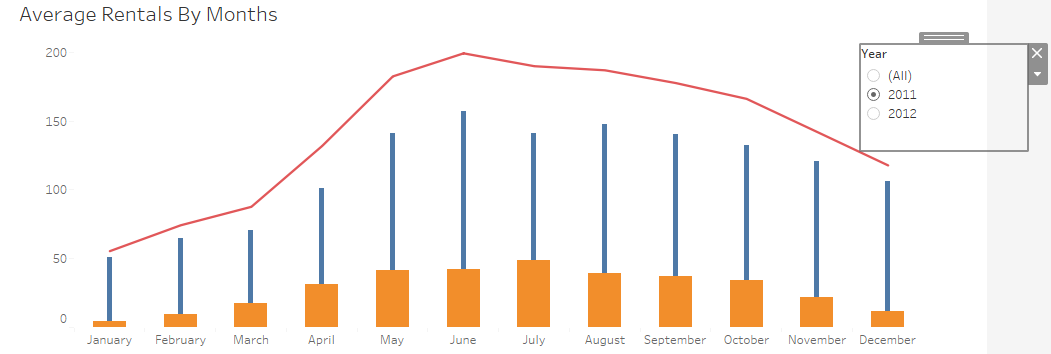
 

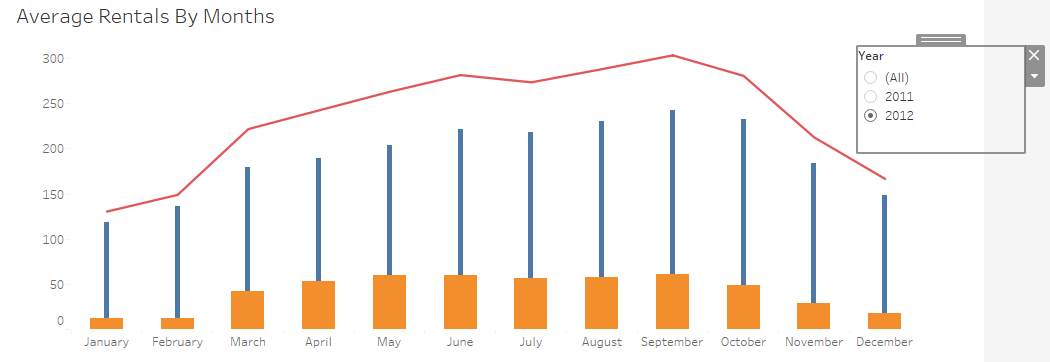


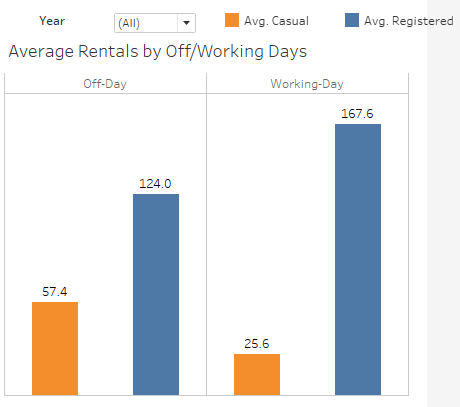
**Graph:2:**



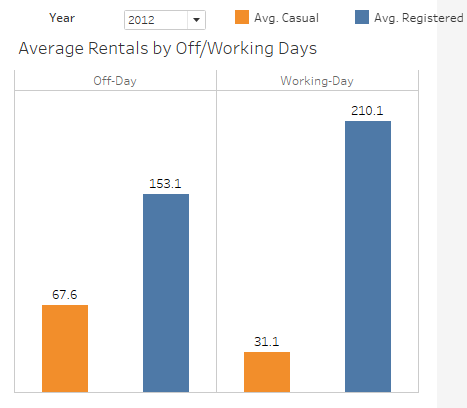
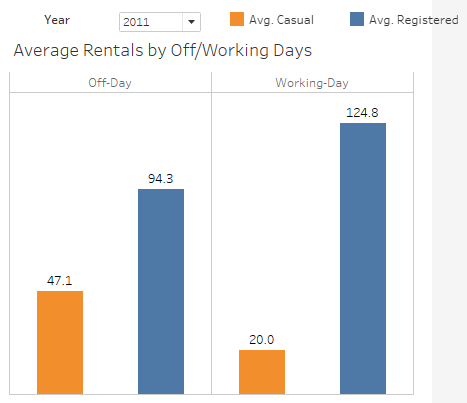
We have created this graph to see how the number of bike rentals vary from month to month. In our dataset, you have records for two years – 2011 and 2012. So, here we have created the graph in such a way that we can check the overall rentals for two years together as well as we can the individual outcome for each year. From the above graph, we can see that bike rentals are highest between the months of May and October. On the other hand, the bike rentals are significantly low in number in the months of January, February, November and December. But the results slightly vary when we refer to the outcomes separately both each year i.e. number of bike rentals in July in 2011 was significantly higher than the number of bike rental in July in 2012 but again, for March and April the number of rentals increased in 2012 compared to 2011. The individual outcomes for each year (2011 and 2012) are shown below.



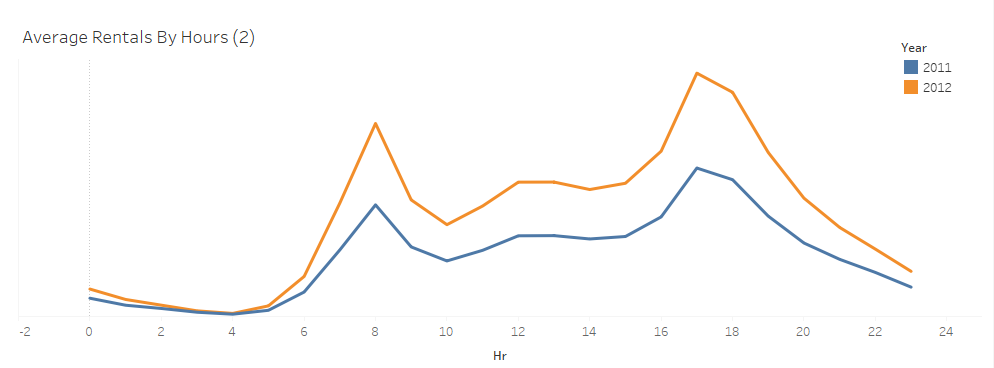


**Graph:3:****:**

We have created a graph to see if there is any difference between the number of average bike rentals on the weekdays and the same on weekends. From the above graph, we can see that the registered users are more likely to rent a bike on weekdays whereas the casual users are more likely to rent a bike on weekends. We have checked this trend for the two years 2011 and 2012 together as well as individually for each year. Even from the individual results, we can see that the trend is same. The results for individual years are shown below.

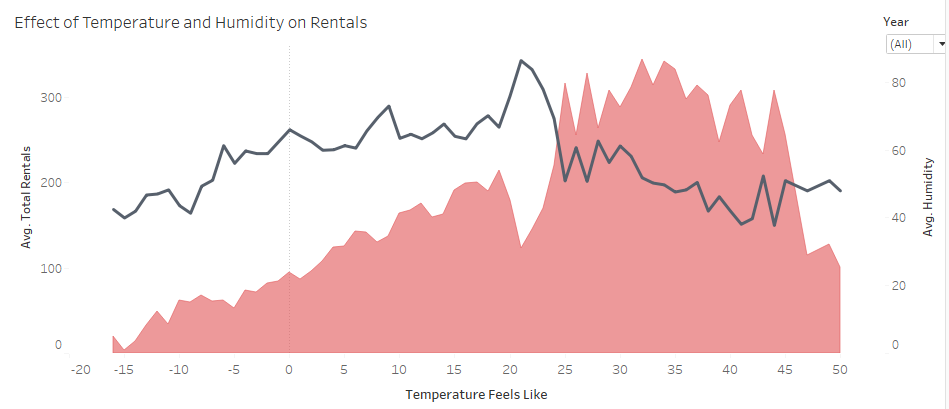


**Graph:4:**

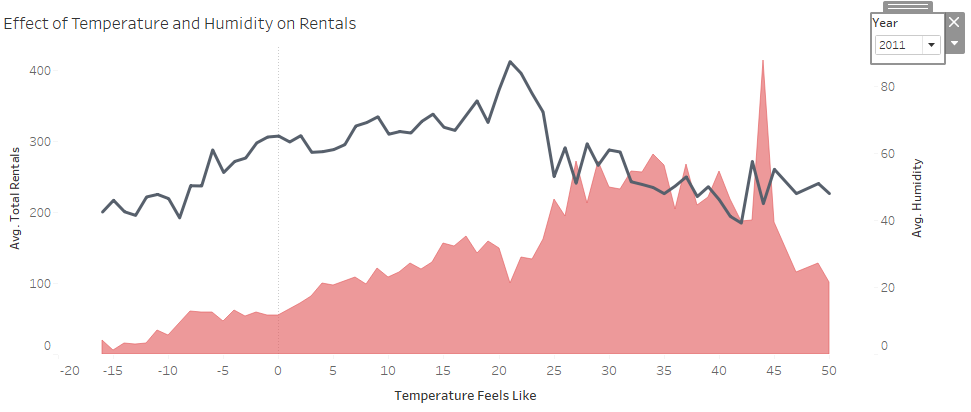


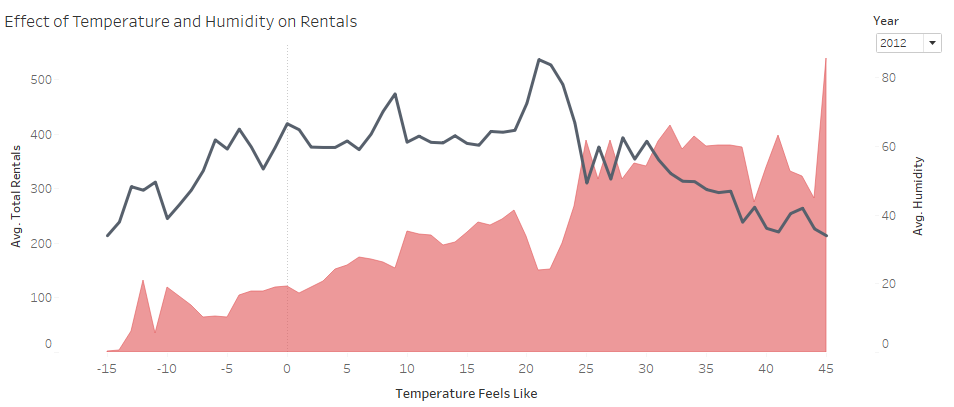
We also checked how bike rentals vary during the hours of the day. The above graph shows that the peak hours for bike rentals are 8:00 AM and between 5:00 PM and 6:30 PM. In our graph, we have shown the rental pattern depending on hours for both the years – 2011 and 2012 and we have found that the peak hours are same for both the years. This can give the idea to the bike rental owners that more bikes should be made available so that the rental demand is met.

**Graph:5:**

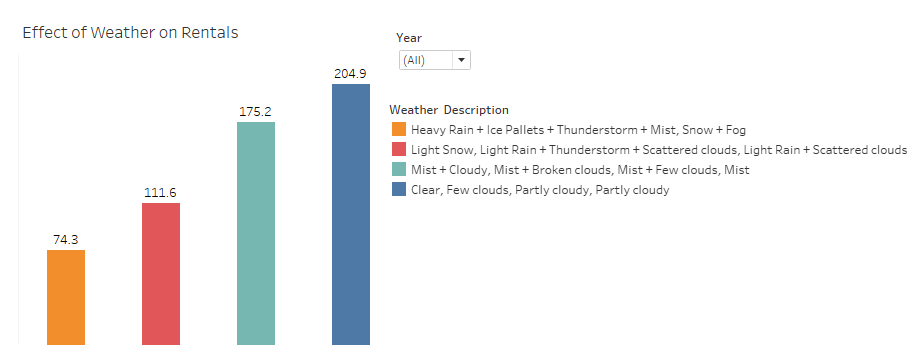


We have created the above graph to check the effect of temperature and humidity on the number of bike rentals throughout the year. In our dataset, we have records for two years -2011 and 2012, so, we have plotted the graph for two years together as well as separately for each year. We can see from the above graph that, the number of rentals is maximum when temperature is between 20 and 40 degrees and the humidity is below average. Even when we checked separately for each year we found that the pattern dependency of the number of rentals on the temperature and humidity is almost same in both the years. The individual graphs for each year are shown below-

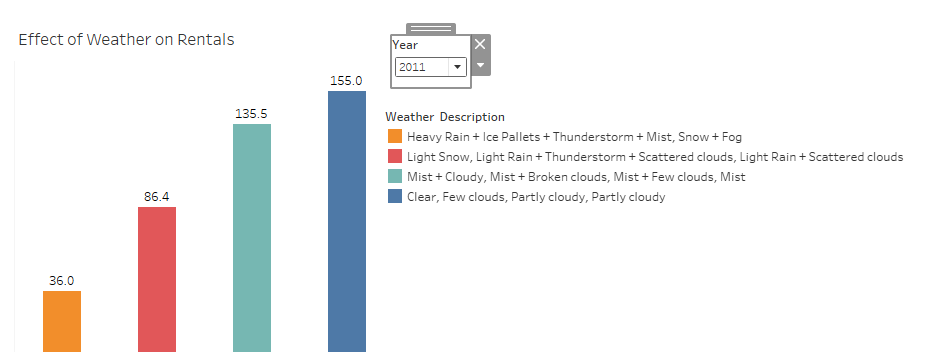


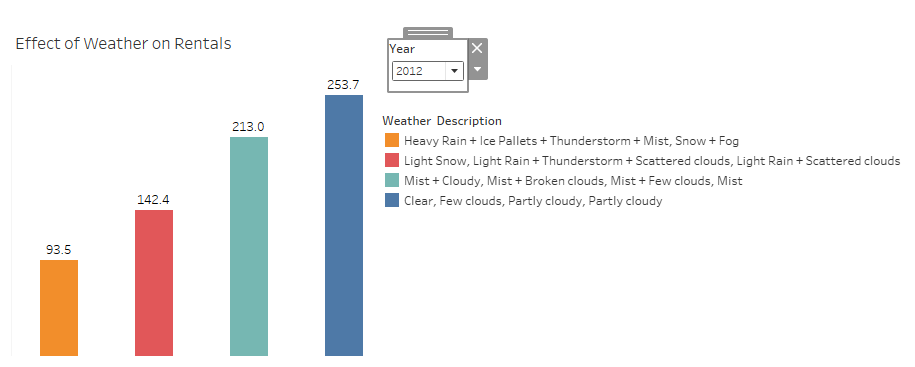


**Graph:6:**



Further, we have how the different weathers impact the number of bike rentals. In our dataset, we have four kind of weather mentioned which are mentioned in the graph. Here, we have plotted the graph for two years (2011 and 2012) together as well as separately for each year. From the above graph, we can see bike rentals are maximum when the weather is clear or partly cloudy and on the other hand, the bike rentals minimum when there is heavy rain or thunderstorms or there is snow or fog. Even from the individual graphs for each year, we can see that the outcome is same. The individual graphs for 2011 and 2012 are shown below.





**Conclusion**

From the above analysis we have concluded the below few points –

* From year 2011 to 2012, Average usage of bike rentals has been increased by 63%.
* Average bike rentals are maximum in fall compared to other seasons.
* June and September months have the highest number of bike rentals.
* Registered users rented bikes mostly on weekdays while casuals users rented more bikes on weekends.
* About 35% more bikes were rented on working days for registered users while we observe a 55% decline in number of bikes rented by casual users.
* Peak rental hours are 8:00AM and 5:00PM.
* We observed higher average of bike rentals when the temperature is between 25-45 oC.
* We have also observed that average rentals decrease when there is high humidity.
* We observe that people prefer to use bike rental services more in clear and cloudy weather than rainy and snow days.

**Reference**

The 10 best data visualization blogs to follow, The 10 best data visualization blogs to follow. (2019). Tableau Software. Retrieved 15 May 2019, from <https://www.tableau.com/learn/articles/best-data-visualization-blogs>

Data source: Bike Sharing in Washington D.C. Dataset, Bike Sharing in Washington D.C. Dataset. (2019). Kaggle.com. Retrieved 18 May 2019, from <https://www.kaggle.com/marklvl/bike-sharing-dataset>