

---

## **“Weather” to Rent a Bike**

CHLOE CZERWINSKI  
SAMPLE NUMBER: 131  
WISCONSIN SCHOOL OF BUSINESS  
DECEMBER 2018

## Executive Summary

This report provides an analysis and recommendations for PhillyCycle to help expand the company. Methods of analysis include focusing on two groups: registered and casual users, and also how the time of day and weather effects the demand for bike rentals.

First, since 81.52% of all users are registered with PhillyCycle, their marketing campaign should be focused on registered users. Secondly, hourly average rentals in the summer is 230, which fall and spring fall shortly behind, meaning the demand for bike rentals is the greatest during summer. There is a strong relationship between temperature and average hourly rentals. Since average hourly rentals increase as temperature rises, PhillyCycle can invest in new warm cities, but have to be cautious of certain weather variables such as humidity.

As for days of the week, there is a greater demand for rentals on weekends for casual users and a greater demand during the week for registered users. Casual users rent bikes on the weekend for leisure and fun exercise and registered users rent bikes during the week to get to and from work. This was concluded since the highest peak in demand for registered users during the day was 9 a.m. and 6 p.m. on weekdays.

Variables that affect all users are not as significant when a regression was made because of how different casual and registered users are. With this report, probability is always uncertain and should be used with caution. There were very few duplicates and erroneous records meaning the data was of high quality, and isolating factors are much more valuable.

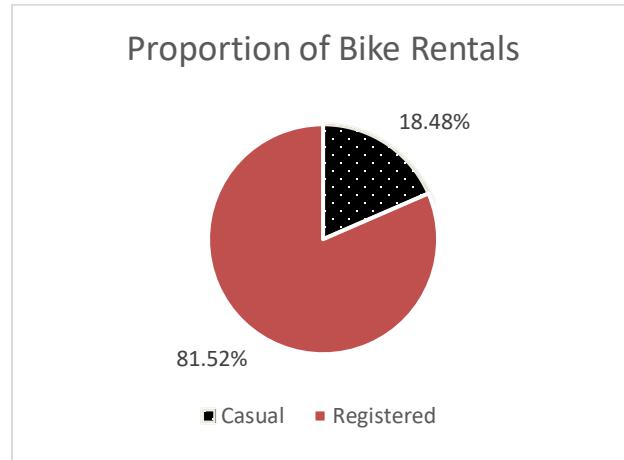
## Analysis

### Bike Rental Patterns

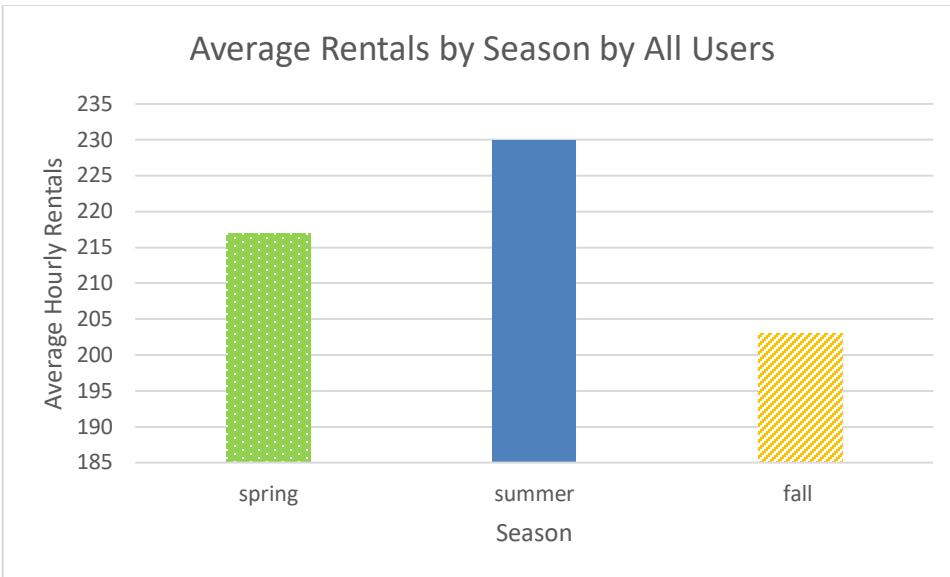
| Casual  | Users |
|---------|-------|
| Average | 35.56 |
| Median  | 17    |
| Maximum | 354   |

| Registered | Users  |
|------------|--------|
| Average    | 156.87 |
| Median     | 115    |
| Maximum    | 876    |

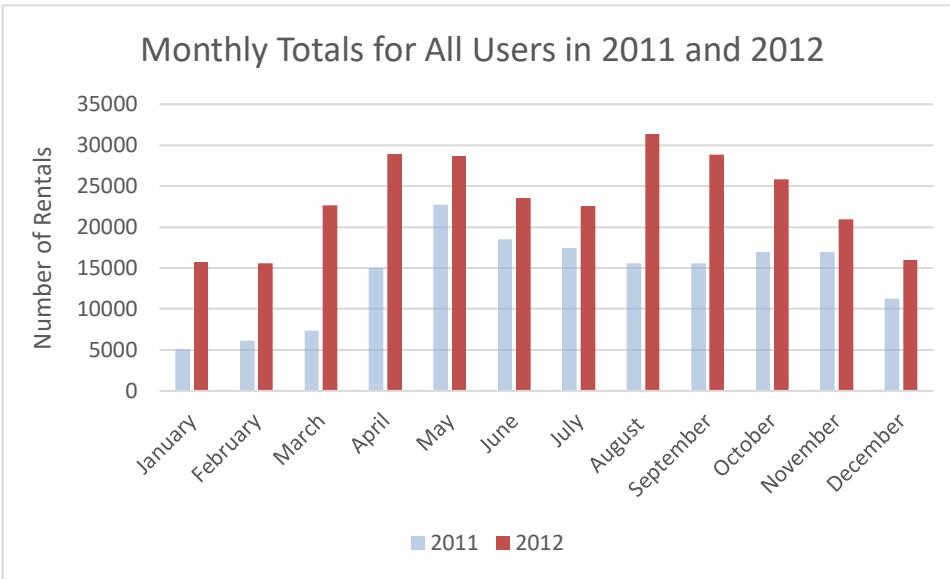
The average was chosen to present measurements for casual and registered users so PhillyCycle is aware of the central value in this set of data. Also included is the median and maximum to show the presence of outliers. Since the numbers go from 35.56 to 17 rentals per hour, this tells us the outliers will skew our data in the rest of the analysis.



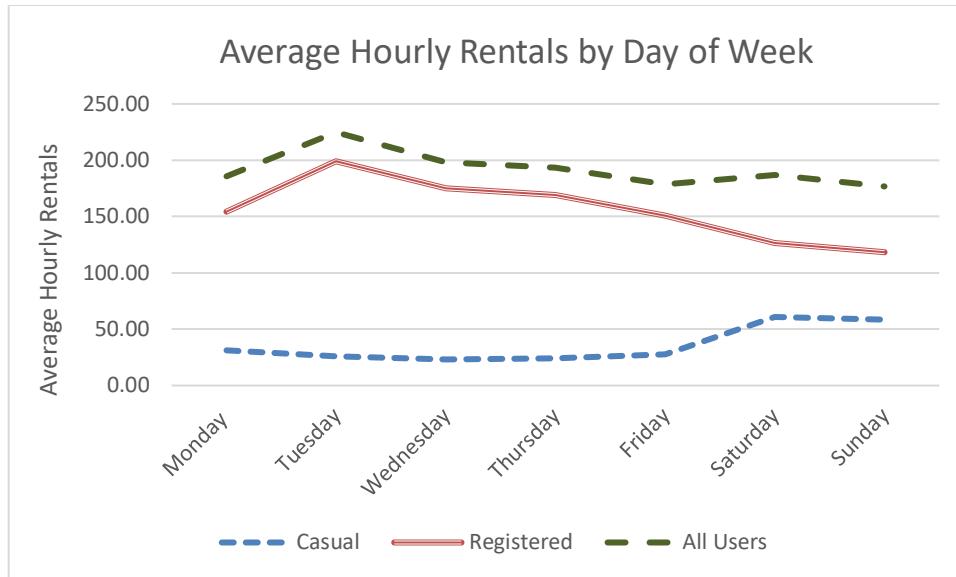
Registered users are overwhelmingly the majority of PhillyCycle's customer base at 81.52%. Casual users only make up 18.48%, and therefore, PhillyCycle should focus on registered users for the upcoming marketing campaign.



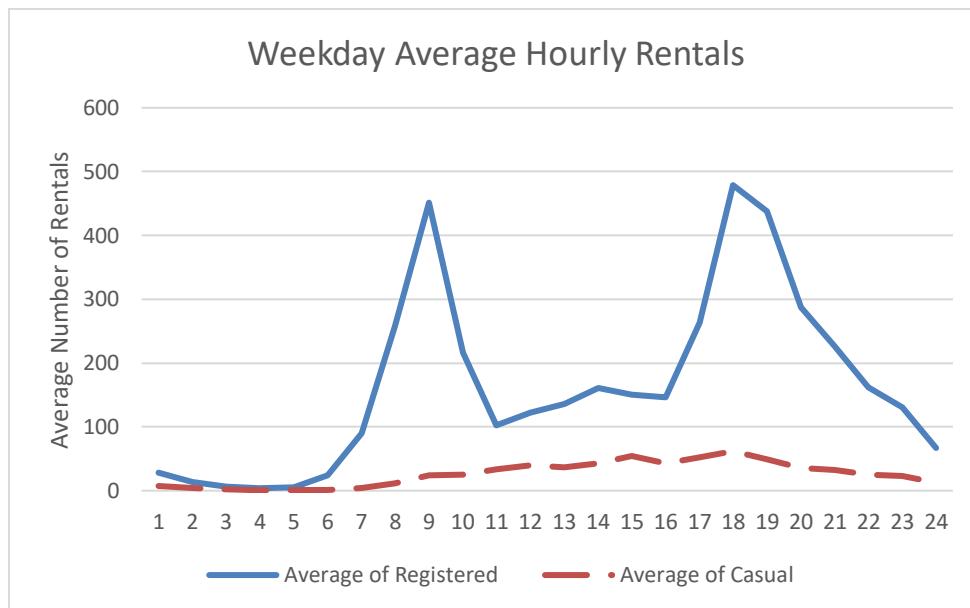
PhillyCycle can use this column chart to adjust pricing, inventory, and marketing strategies for each season. Summer has the highest number of hourly average rentals at 230, whereas spring has 217 average rentals and 203 in the fall. It seems PhillyCycle has 200 regular users, and almost 15 people have a preference by season (since the difference between spring and summer is 13 and the difference between spring and fall is 14).

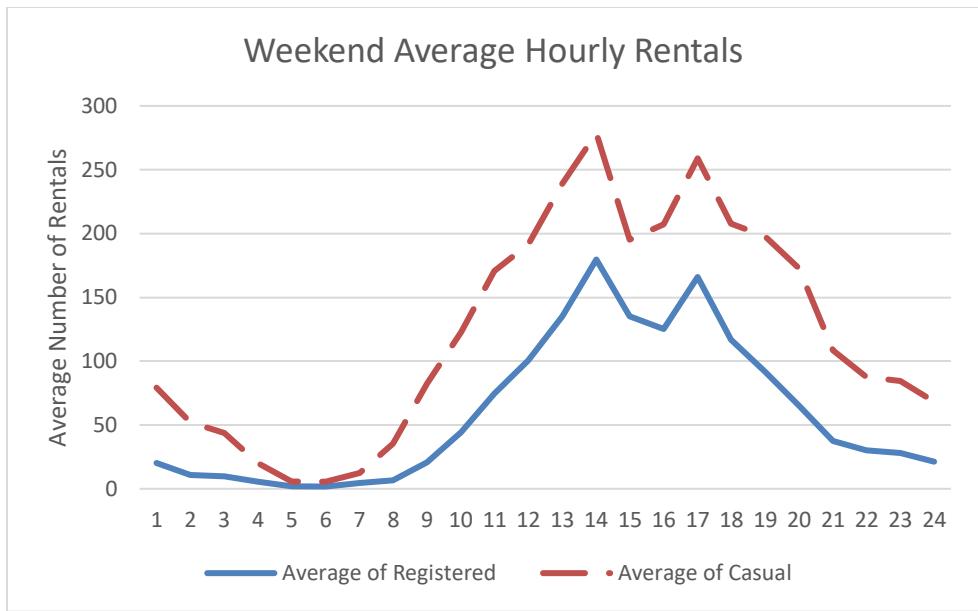


In the winter months there is a significant decrease in demand for bike rentals, however, the summer months have the greatest demand. In 2011, May had the highest number of bike rentals at 22,698, but in 2012 August had the highest number selling 31,380 rentals. Overall, there are significantly more bike rentals for each month in 2012 than 2011.

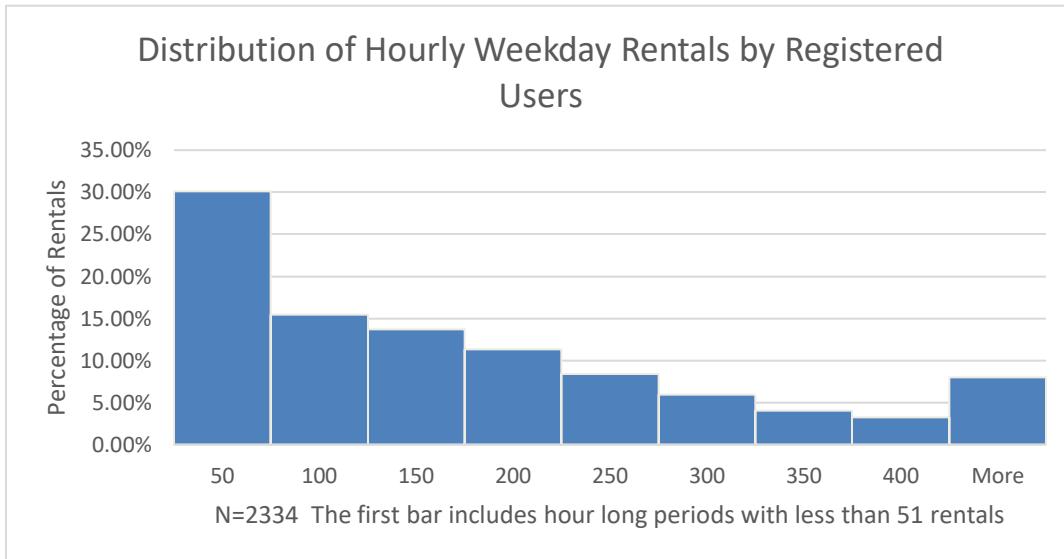


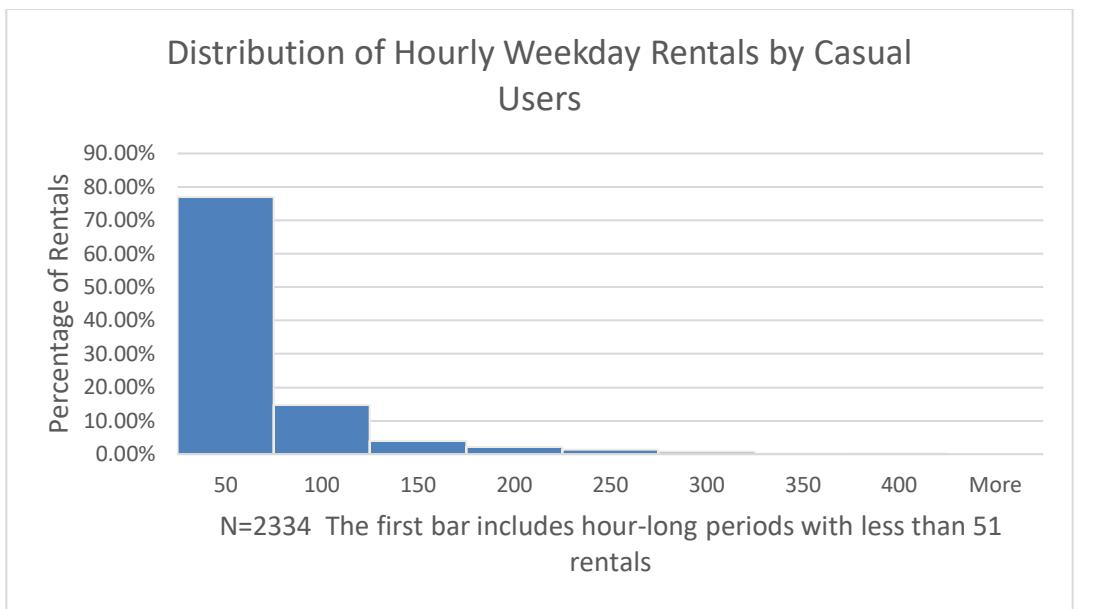
Based on this line graph, registered users and casual users are inversely related. Tuesday has the greatest demand for registered users at 200 rentals and casual users peak on Saturday and Sunday, at about 61 rentals. What this tells PhillyCycle is that registered users are consistent during the week because they use bikes as a means for transportation, whereas causal users use bike rentals on the weekend possibly for leisure or a fun exercise. This also poses the question of whether there is enough bikes at each station during the week since registered users are getting to work every day, and there might not be any left for casual users.



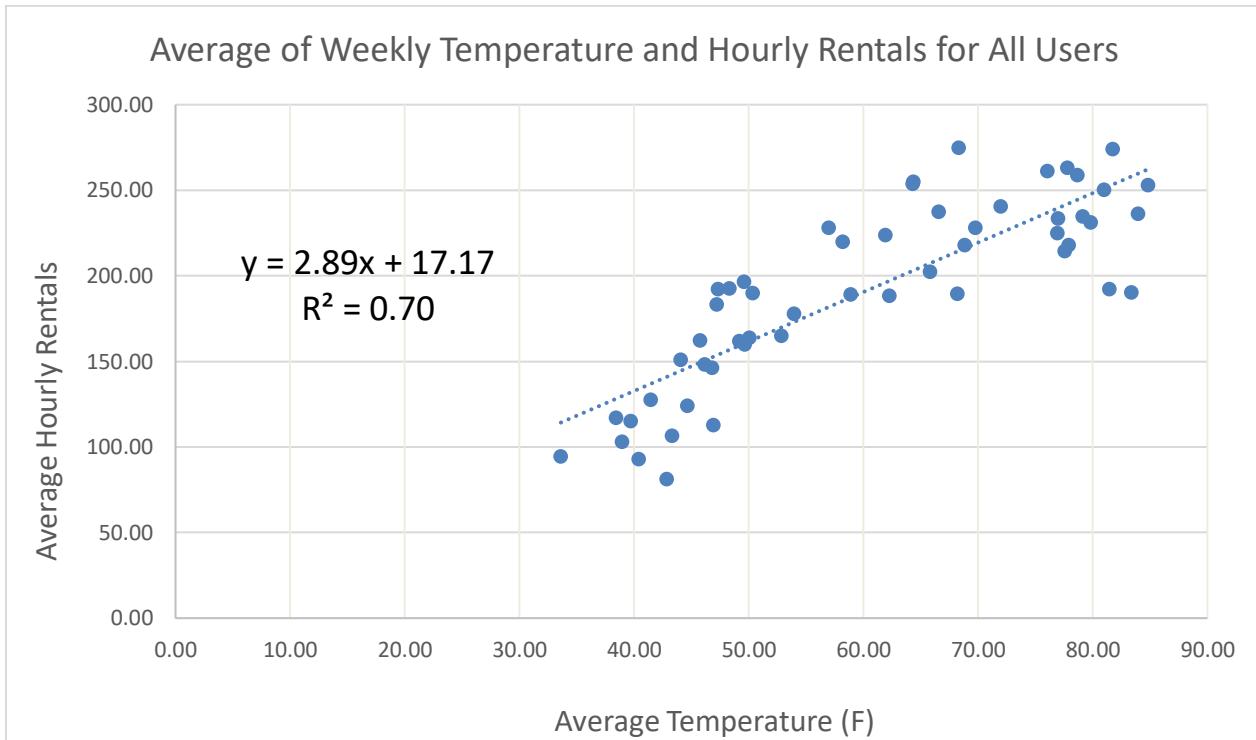


Registered users greatest demand for rentals is at 9 a.m. and 6 p.m. This tells PhillyCycle again that registered users could use bikes as a means for transportation to get to work. This also lets PhillyCycle know that the registered user base is loyal since they have to get to work every day. In contrast, casual users peak hours are 2 p.m. and 5 p.m. This could be for leisure or a fun exercise.





Looking at registered users, 30% of the rentals show that 50 or less bikes are rented in an hour. However, casual users use 50 or less rentals in 80% of their rentals. There are usually 50-100 rentals hourly for both registered and casual users, but registered users are more evenly distributed. Overall, these two histograms show that on average there are fewer rentals per hour for casual users than registered users. Also, since trends and tendencies are more prevalent for registered users during the week, PhillyCycle can capitalize on this.



Since the R<sup>2</sup> value of the trendline is 0.70, this means the data and regression line are close (because R-squared is between 0-100%). Temperature and average hourly rentals are also highly correlated. I found this by taking the square root of 0.70 to get 0.84. The regression equation tells PhillyCycle that for each degree increase, rentals only increase by 2.89 (or 3). The two variables appear to be strongly correlated which aligns with the analysis on seasons as well (Higher temperatures in the summer, higher demand for bike rentals. Lower temperatures in the winter, lower demand for bike rentals).

PhillyCycle can adjust their pricing techniques if there is a statistical significance between average hourly rentals and weekday versus weekend between casual, registered, and all users. After conducting a hypothesis test to determine if there is statistical significance at the 95% confidence level, PhillyCycle can conclude that weekend/weekday does have an effect on average hourly rentals for both casual and registered users. Using the sample size I was given, sample means and standard deviations, here are the results:

### *All Users*

Null Hypothesis: The average hourly rentals during the weekend is different from average hourly rentals for weekdays.

|                    | Weekend | Weekday |
|--------------------|---------|---------|
| Mean               | 182.29  | 196.28  |
| Standard Deviation | 176.79  | 186.40  |
| Sample Size        | 1691    | 643     |

The effect on average hourly rentals for all users for weekends and weekdays is not statistically significant. It is not significant because the confidence interval contains 0. Therefore, it is not useful to apply weekend and weekday to registered and casual users combined.

### *Casual Users*

Null Hypothesis: The average hourly rentals during the weekend is more than the average hourly rentals for weekdays.

|                    | Weekend | Weekday |
|--------------------|---------|---------|
| Mean               | 59.70   | 26.37   |
| Standard Deviation | 75.90   | 31.28   |
| Sample Size        | 1691    | 643     |

Testing this hypothesis at a 95% significance level gives a confidence interval of (28.98, 37.68). This means that 95% of the time the weekend will increase average

hourly rentals by 29 casual users.

### *Registered Users*

Null Hypothesis: The average hourly rentals during the weekend is less than the average hourly rentals for weekdays.

|                    | Weekend | Weekday |
|--------------------|---------|---------|
| Mean               | 122.59  | 169.91  |
| Standard Deviation | 106.82  | 167.40  |
| Sample Size        | 1691    | 643     |

Testing this hypothesis at a 95% significance level gives a confidence interval of (-61.22, -33.42). This means that 95% of the time, the weekend will decrease the average hourly rentals of registered users by 33-61.

### **Regression Analysis**

To determine the significance of all individual factors for all users, a regression analysis was conducted. These are the following factors in the regression:

- Temperature (F)
- Weather
- Season
- Year 2012
- Weekend

| <i>Regression Statistics</i> |         |
|------------------------------|---------|
| Multiple R                   | 0.71    |
| R Square                     | 0.50    |
| Adjusted R Square            | 0.50    |
| Standard Error               | 130.41  |
| Observations                 | 2334.00 |

|               | <i>Coefficients</i> | <i>Standard Error</i> | <i>t Stat</i> | <i>P-value</i> | <i>Lower 95%</i> | <i>Upper 95%</i> | <i>Lower 99.0%</i> | <i>Upper 99.0%</i> |
|---------------|---------------------|-----------------------|---------------|----------------|------------------|------------------|--------------------|--------------------|
| Intercept     | -110.33             | 18.73                 | -5.89         | 0.00           | -147.05          | -73.60           | -158.61            | -62.05             |
| Temp (F)      | 3.70                | 0.26                  | 14.38         | 0.00           | 3.20             | 4.21             | 3.04               | 4.36               |
| humidity      | -1.82               | 0.17                  | -10.58        | 0.00           | -2.16            | -1.49            | -2.27              | -1.38              |
| windspeed     | -0.27               | 0.36                  | -0.75         | 0.45           | -0.98            | 0.43             | -1.20              | 0.66               |
| Spring        | 27.94               | 9.55                  | 2.93          | 0.00           | 9.21             | 46.67            | 3.32               | 52.56              |
| Summer        | 2.10                | 12.36                 | 0.17          | 0.86           | -22.14           | 26.34            | -29.76             | 33.97              |
| Fall          | 60.42               | 8.45                  | 7.15          | 0.00           | 43.85            | 77.00            | 38.63              | 82.21              |
| Weekend       | 182.17              | 6.77                  | 26.92         | 0.00           | 168.90           | 195.45           | 164.73             | 199.62             |
| Year 2012     | 82.48               | 5.46                  | 15.10         | 0.00           | 71.77            | 93.19            | 68.40              | 96.56              |
| Mist          | -2.74               | 6.63                  | -0.41         | 0.68           | -15.75           | 10.27            | -19.85             | 14.36              |
| Precipitation | -39.43              | 10.82                 | -3.64         | 0.00           | -60.65           | -18.21           | -67.33             | -11.54             |

Here is the regression equation as follows:

$$\text{All rentals} = -110.33 + 3.70 \text{ (Temperature)} + -1.82 \text{ (Humidity)} + -0.27 \text{ (Windspeed)} + 27.94 \text{ (Spring)} + 2.10 \text{ (Summer)} + 60.42 \text{ (fall)} + 182.17 \text{ (Weekend)} + 82.48 \text{ (Year 2012)} + -2.74 \text{ (Mist)} + -39.43 \text{ (Precipitation)}$$

This regression is important because it shows the relationship between rentals and the variables. For example, for every degree temperature rises, temperature will increase hourly rentals by 3.70, holding all other variables constant. According to this regression, windspeed, summer, and mist are not significant since their p value is greater than 0.05, and the confidence interval includes 0.

The factors are not strongly correlated because R-squared is 0.50, This could mean there are other factors not shown that affect the amount of rentals that could be significant. Or, the divergence in behavior between casual and registered users could cause all users combined to not be strongly correlated.

## Conclusion

First, in order to maximize revenue and have a clear, strategic marketing plan, PhillyCycle should focus on registered and casual users separately. All users data is not very significant and the two groups have an inverted relationship with certain variables such as weekend and weekday. This calls for separate, different marketing plans and dynamic pricing.

One recommendation is to use dynamic pricing for registered and casual users on weekdays versus weekends. In order to capitalize profits, PhillyCycle should lower prices on weekdays for casual users since they prefer to ride for leisure or fun exercise on weekends. In contrast, registered users overwhelmingly rented bikes during the week to transport themselves to work. Therefore, the prices should be lowered for registered users on weekends to encourage them to rent on the weekends as well.

Another recommendation for PhillyCycle is to mainly focus on registered users for their upcoming marketing campaign. This makes sense since registered users make up over 80% of the dataset. PhillyCycle can use different marketing strategies for different seasons for both causal and registered users because summer had the highest demand whereas winter had the lowest demand. Prices can be adjusted too by season to encourage them to rent more bikes in the winter and fall.

It would benefit PhillyCycle to put more stations in populated areas (such as Philadelphia) with lots of business around the area. Since registered users rent the most bikes at the times 9 a.m. and 6 p.m., putting more stations around businesses would encourage more workers around the area to use bikes as a means of transportation and sign up. As for casual users, since they have different reasons for renting, it would be beneficial to put more bike stations by scenic routes such as a park or near an attraction.

When considering temperature, if PhillyCycle wants to expand and open more stations, they should consider expanding in warmer locations with a lower humidity because most bike rentals occurred in the summer, followed by spring, fall, and then winter. Also, humidity was the only statistically significant variable for weather. As humidity increased by 1, bike rentals decreased by 1.82. More research would have to be done before expanding to new warmer locations such as Phoenix.

In conclusion, PhillyCycle can use variables such as weekend/weekday, temperature, hour, and season to focus their pricing and create an efficient marketing strategy to maximize profits and expand their business.

## **Appendix**

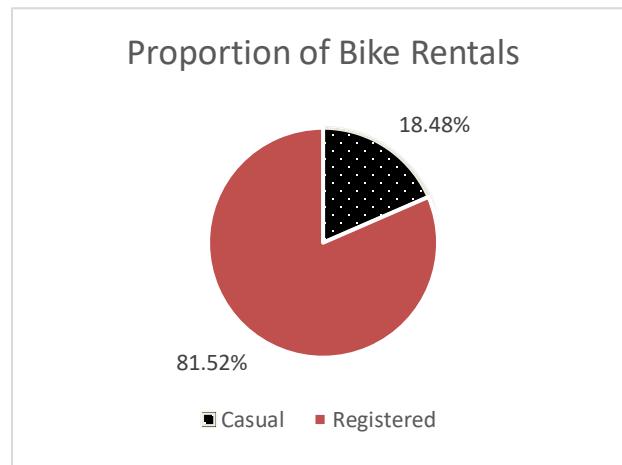
### **Notes on Data Preparation**

First, I removed any duplicates. Since there were only two duplicates, this tells me the data set is of high quality. Then, I removed any temperature 50 degrees Celsius or above since this is equal to 122 degrees Fahrenheit, and these outliers would shift the data. After removing the outliers, I added variables such as season, temperature in Fahrenheit, weather, hour, weekend, year, day of week, and month. By using the "if" function in excel, I was able to add dummies by using 0 and 1 for weather and seasons. Lastly, I made pivot tables and regressions to explain the relationships between the variables added.

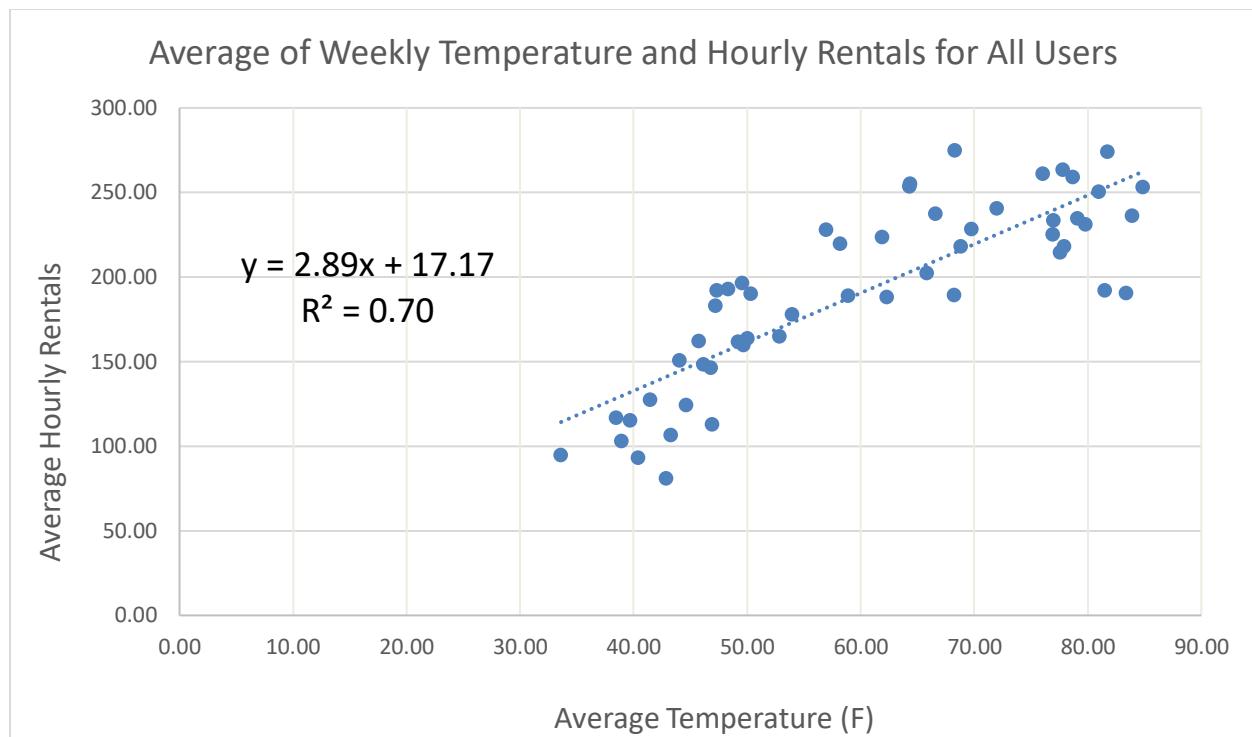
## Elevator Charts

| Casual Users | Registered Users |
|--------------|------------------|
| Average      | 35.56            |
| Median       | 17               |
| Maximum      | 354              |

This table shows the hourly rental measure of central tendency for causal and registered users. It exposes the differences in casual and registered users. Specifically, the difference between the median and average tell PhillyCycle that there are large volumes of rentals at a time on certain dates, which PhillyCycle can profit from.



Next, I included the proportion of bike rentals again because one of PhillyCycle's main goals is the upcoming marketing campaign. They should know where to spend most of their marketing budget and how.



Lastly, I included the scatterplot because this graph is most representative of the correlation between temperature and average hourly rentals. It is also extremely important since temperature can tie in the findings for season too. And, this graph can tell PhillyCycle where to locate their stations. It can tell us how many bikes to put at each station during each month, which it is very important for a bike rental business to control their inventory at certain times each year.