DEMAND FOR BIKE RENTALS: A report to the management of Capital Bike Share

Group 3: Carter Che, Brendan Griffin, Dani Herzberg, Nicole Kuker, Anthony Long, Carl Thomas

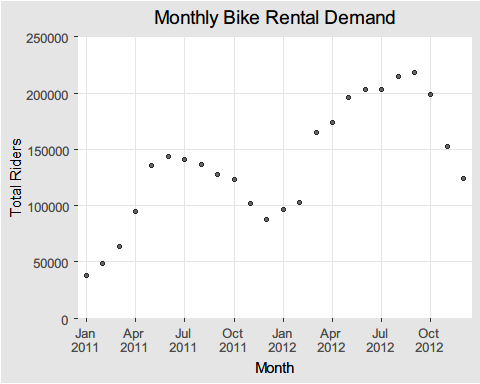
2 March 2021

# Introduction

To be written at later date

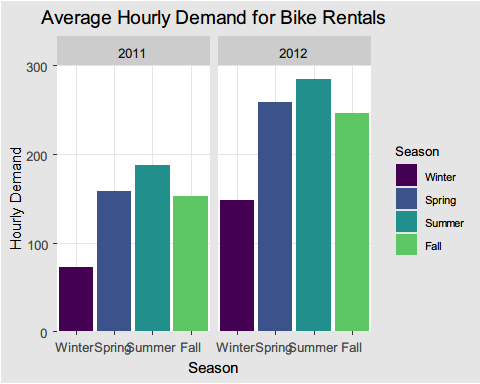
# Overall Trends in Bike Rental Demand

Question 1 - Anthony & Carl General trend - positive over the years

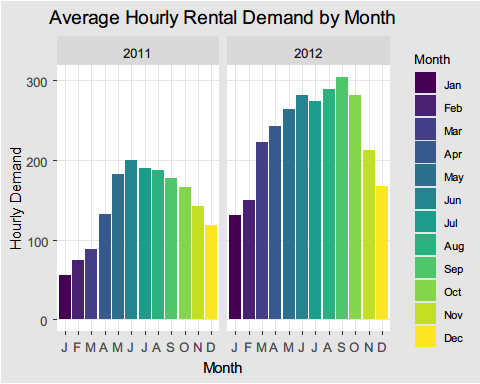


# Trends by Season

Looking at the overall seasonal impact on hourly demand, it is clear that the Winter is a time of depressed demand likely related to the cold temperatures, which will be addressed later. The Spring and Fall are quite similar in demand terms only trailing summer by a small margin on an hourly basis. The seasonal trends in hourly demand are consistent between the two years of data, which indicates it is a continuing trend.

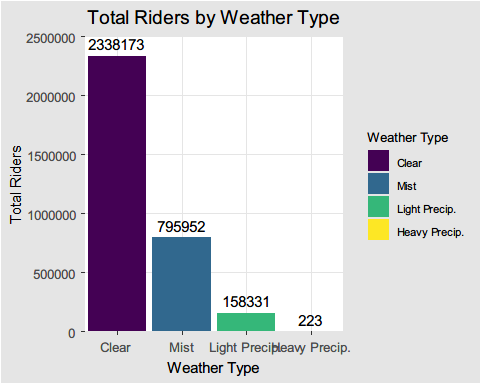


When the seasons are split into the individual months, there is a more gradual transition into and out of the Winter depressed demand period. The overall trend of seasonality affecting demand holds and January’s demand being 40% of that in June, the peak month. Once again the monthly trend holds consistent between the two years in the dataset.

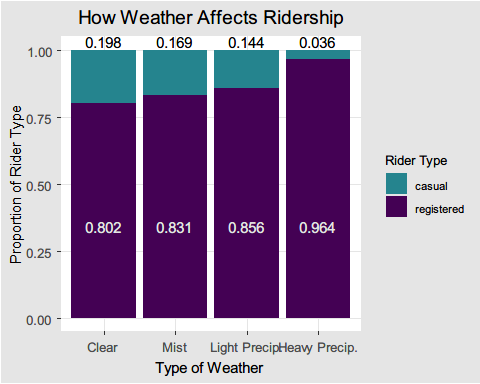


# Further Trends: Looking at Weather and Rider Type

Hourly weather conditions were documented within the dataset and categorized as Clear, Mist, Light Precipitation, and Heavy Precipitation. As weather conditions worsen, there is a steep decrease in overall ridership. Though there are more registered riders than casual riders, this decrease in demand holds true when looking at both subsets of rider type.



As weather conditions worsen, casual riders make up a decreasing proportion of total ridership. This suggests that casual riders may be more influenced by the weather, and are more likely to choose other modes of transportation when weather conditions are unpleasant or uncomfortable.



# Predicting Total Bike Rental Demand

### Daily Temperature

To further analyze the relationship between bike demand and weather conditions, we ran two different regression analyses. The first analysis attempted to define the relationship between normalized temperature and total bike demand. The equation of the linear model generated is:

Expected Demand = -0.04 + (381.30 \* Temp)

The linear model is statistically significant given the p-value is below the alpha of 0.05 (it is also below an alpha of 0.01). The coefficient t-value is a measure of how many standard deviations our coefficient estimate is far away from 0. In the temperature linear model, the t-statistic value is relatively far away from zero and is large relative to the standard error, which could indicate a relationship exists.

However, the R squared value is only 0.1638, which indicates that the linear model is not a strong fit to the data and that the relationship between temperature and demand is not particularly strong. This R squared means roughly 16.4% of the variance found in overall demand can be explained by the actual temperature.

The scatterplot indicates that there is a wide range of demand at almost all temperatures. The demand range only tightens at the extremes of the temperature scale.

### Daily Feeling Temperature

Question 6 - Fit another linear model predicting total daily bike rentals from daily feeling temperature. Write the linear model, interpret the slope, etc.

The equation of this linear model is: Expected Demand = -11.88 + (423.18 \* Feels Like Temp) (where feels like temperature is the normalized feels like temperature in the data set)

The linear model is statistically significant given the p-value is below the alpha of 0.05 (it is also below an alpha of 0.01). The p-value tests the null hypothesis that there is no correlation between the feels like temperature and overall demand and we can reject the null hypothesis.

However, the R squared value is only 0.1607, which indicates that the linear model is not a strong fit to the data and that the relationship between the feels like temperature and demand is not particularly strong. Only 16.1% of the variance in the overall demand can be explained by the feels like temperature.

The slope of the linear model is steeper than in the actual temperature model indicating a one degree change in the feels like temperature has a bigger impact on demand than a one degree change in the actual temperature.

The scatterplot indicates that there is a wide range of demand at almost all feels like temperatures. The demand range only tightens at the extremes of the temperature scale.

### Which is the better predictor?

Part of Question 6: Is the temperature or feeling temperature a better predictor of bike rentals?

The actual temperature is a better predictor of overall demand for bike rentals. The higher R squared and lower standard error in the temperature linear model indicates a marginally better fit than the feels like temperature model. That being said, the R squared is low for both models and the standard error high. Neither model is a particularly good fit for the data and does not explain a large portion in the demand for bike rentals.

# Correlation between Rental Demand and Typical Office Hours

### Is there a “rush hour” for bike rentals?

Hypothesis: There must be high demand during office timings. Early morning and late evening can have different trends (cyclist) and low demand from 10:00pm to 4:00am. Do you agree?

\*\*Add distrubtion of demand over hours to final report

To assess whether there is higher demand during office hours than during the overall daytime hours, a hypothesis test was run. The null hypothesis being tested was stated as “Office Hour Demand” (or the sample mean) is greater than or equal to the “Daytime Demand” (or the population mean). The result of the hypothesis test was that we did not reject the null hypothesis and we have evidence that “Office Hour Demand” is greater than the “Daytime Demand”. We ran this same test for “Office Hour Demand” being greater than or equal to the “Trough Daytime Demand” (or the hours outside office timings) and, again, did not reject the null hypothesis and there is evidence that Office Hours have more demand than the remaining Daytime Hours.

\*\*Need to do early morning versus late evening?

The overnight hours (20:00 to 04:00) had a hypothesis that their demand was lower than the Daytime Demand and so another hypothesis test was run. The null hypothesis was stated as “Overnight Demand” is less than or equal to “Daytime Demand”. Again, the result was to not reject the null hypothesis as the test statistic was not in the rejection range. So, there is evidence that the overnight hours are low demand periods.

### Workdays vs. Weekends/Holidays

Hypothesis: Registered users demand more bike on weekdays compared to the weekend or holiday. Do you agree?

# Recommendations and Conclusions

Question 7 - What is your overall recommendation about the demand for bike rentals for the Capital Bike Share management? Share your insights with the company based on your analysis.

* Overnight discounts or peak commute surcharges?
* Lower staffing and inventory in winter? Good time for maintenance?

# Appendix