Group 5: NYC Citibike Data Analysis

Anusha Muddapati, Harshith Sesham, Deekshit Vedula, Tejeshwine Viswanathan, Loren Young

4/21/2022

I. Introduction

- This project analyzes the correlation and influences that weather has on public bicycling in the New York City (NYC) metropolitan area.
- Using weather data taken from the National Oceanic and Atmospheric Administration (NOAA) for the month of January, 2022, these influences were examined on bicycle riders under the Citi BikeShare community.
- The following analysis allows for the understanding and interpretation of how weather impacts commuting via bicycle for a large American city in contemporary times.

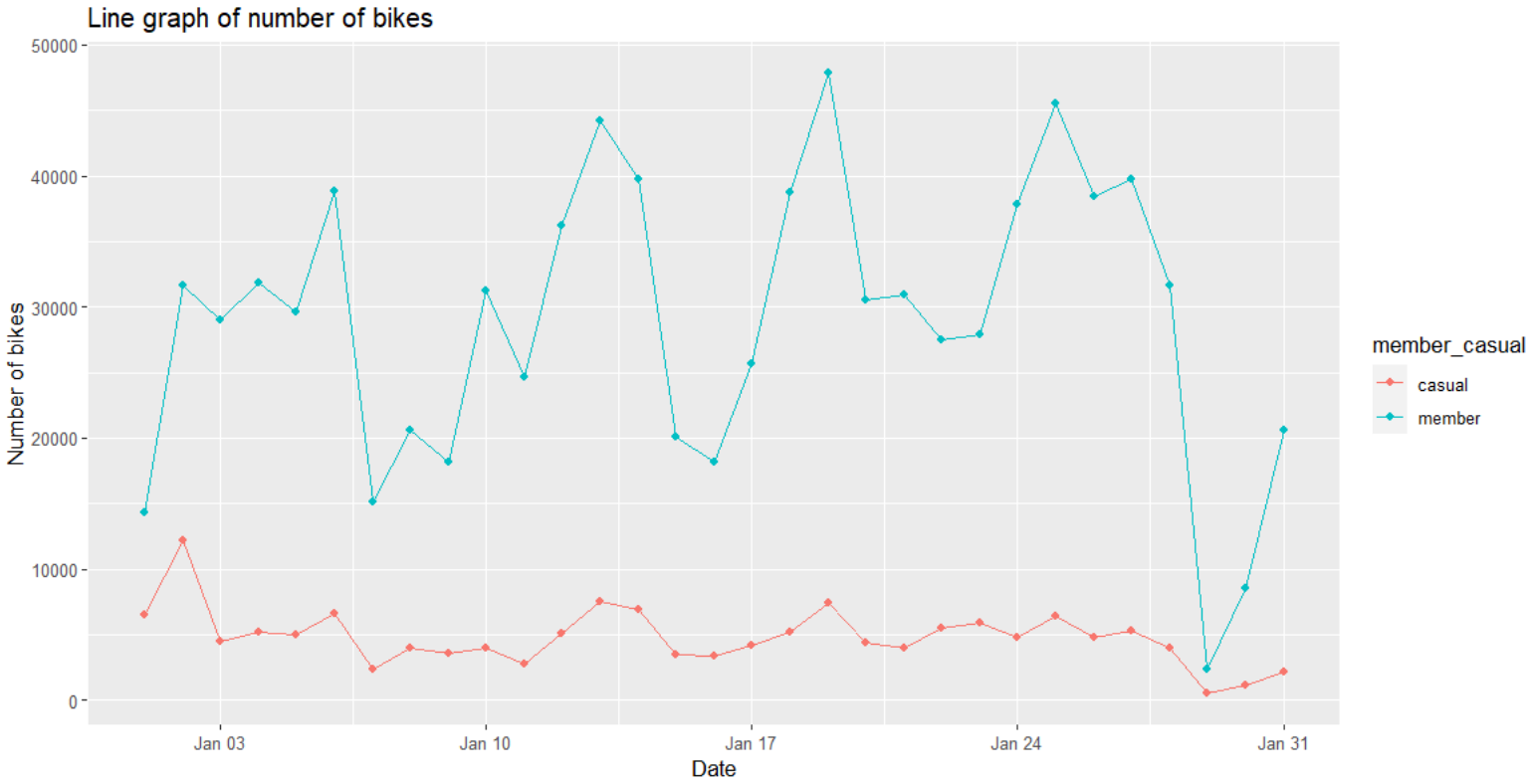
II. Data-set

- Citi BikeShare is a rental company which allows users to rent bicycles. A rider unlocks a bike and pays per time or distance that they use the bike. After a ride is ended at a designated BikeShare location, thus ending the ride, the data is recorded in the BikeShare database.
- Parameters recorded include the type of bike used, the start and end times, and the start and end geographical coordinates and the start and end stations.
- The weather data includes the type of cloud cover, precipitation, and temperature for different times of day during the month of January, 2022.
- Both data-sets thus align in their respective time-frames, with the BikeShare data-set recording over one million events for the month.

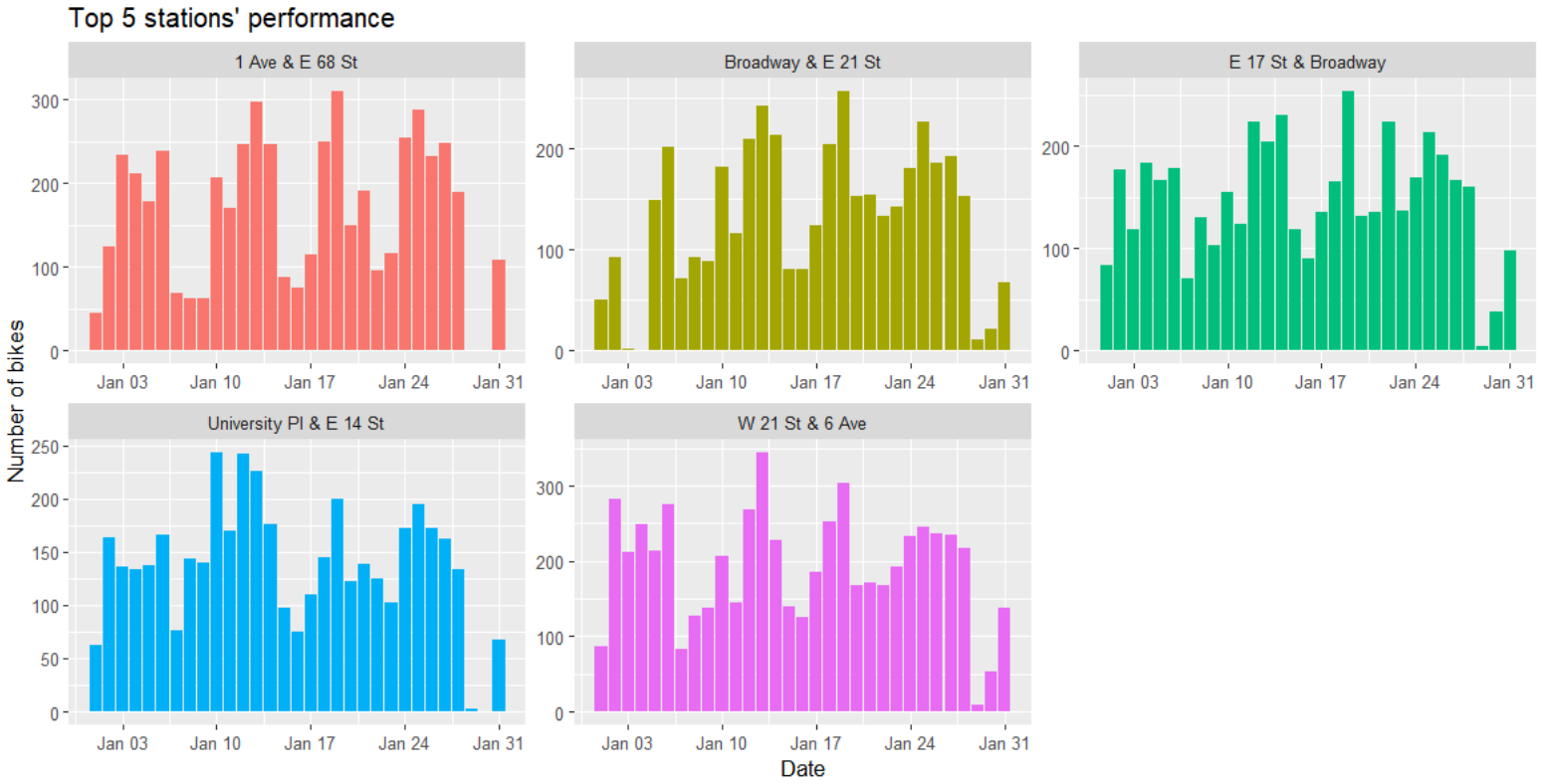
III. Analysis

- The premise of the analysis is to understand the correlation between several weather variables and BikeShare riders for the NYC area.
- Both data-sets were accumulated into a SQL database for easier accessibility of each parameter.
- Several weather related parameters were plotted against the types of bikes and number of riders for each day of the month.
- Regression analysis was employed to determine the correlations between parameters in the two data-sets to observe trends in the data.

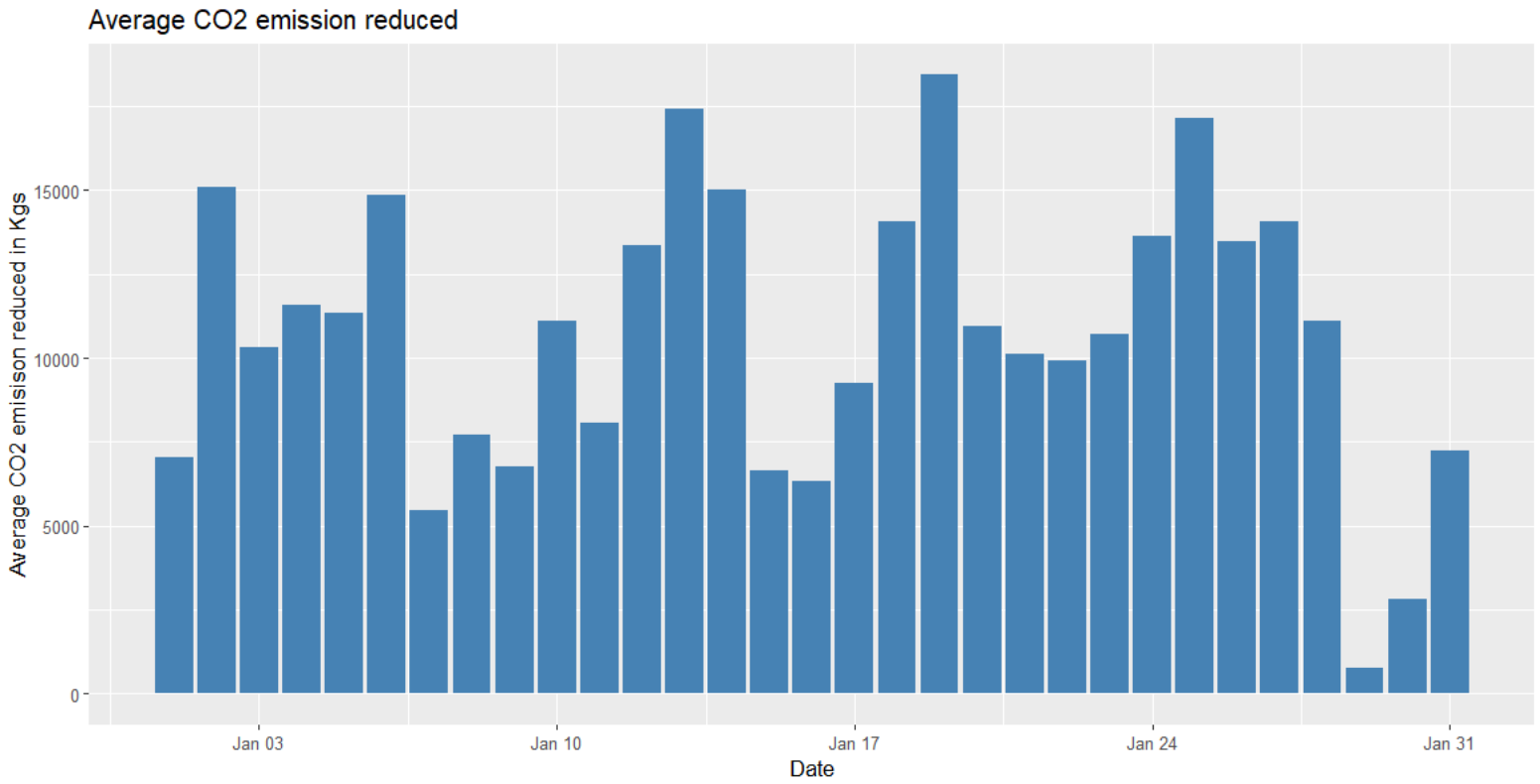
Number of bikers



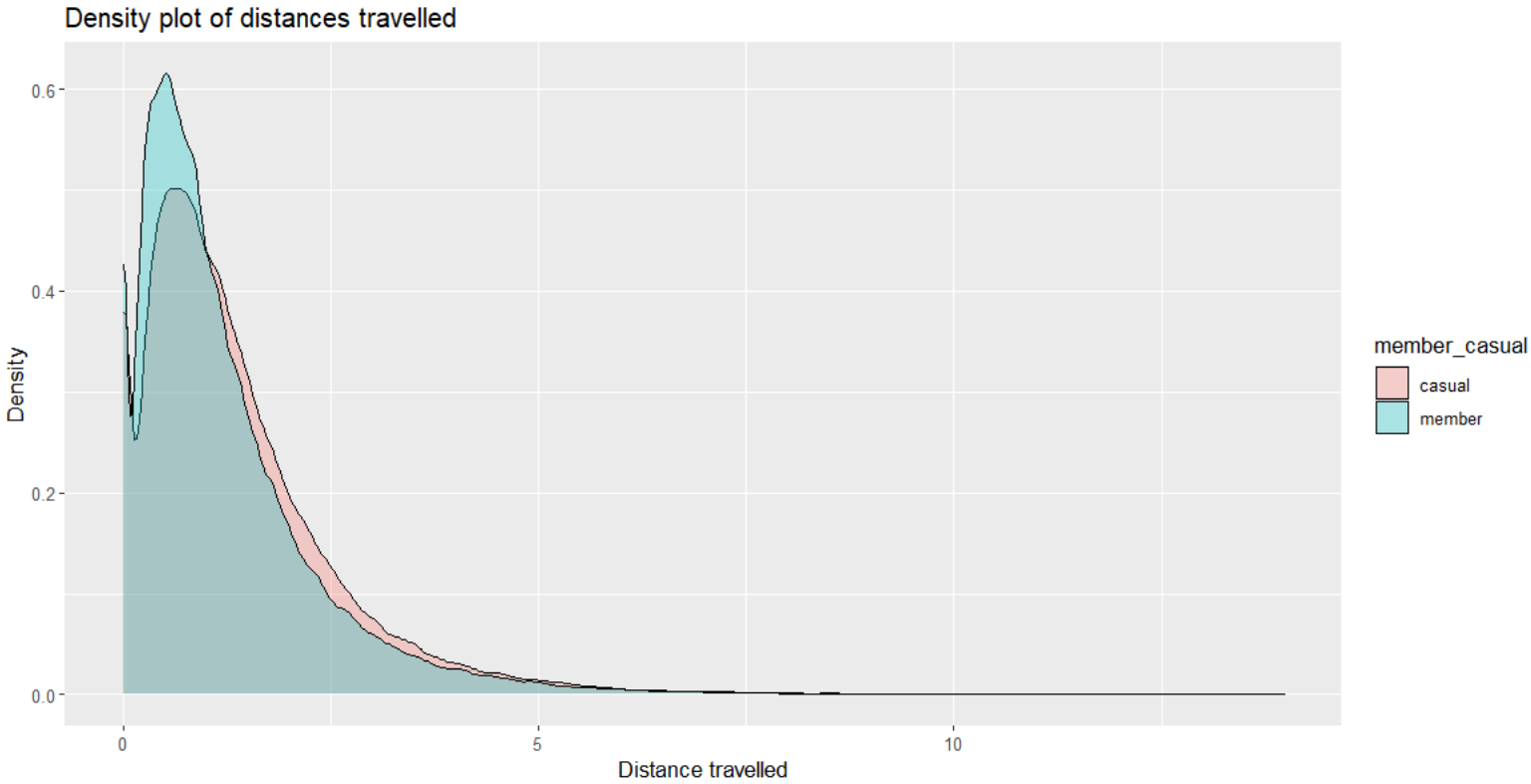
Top 5 best performing stations



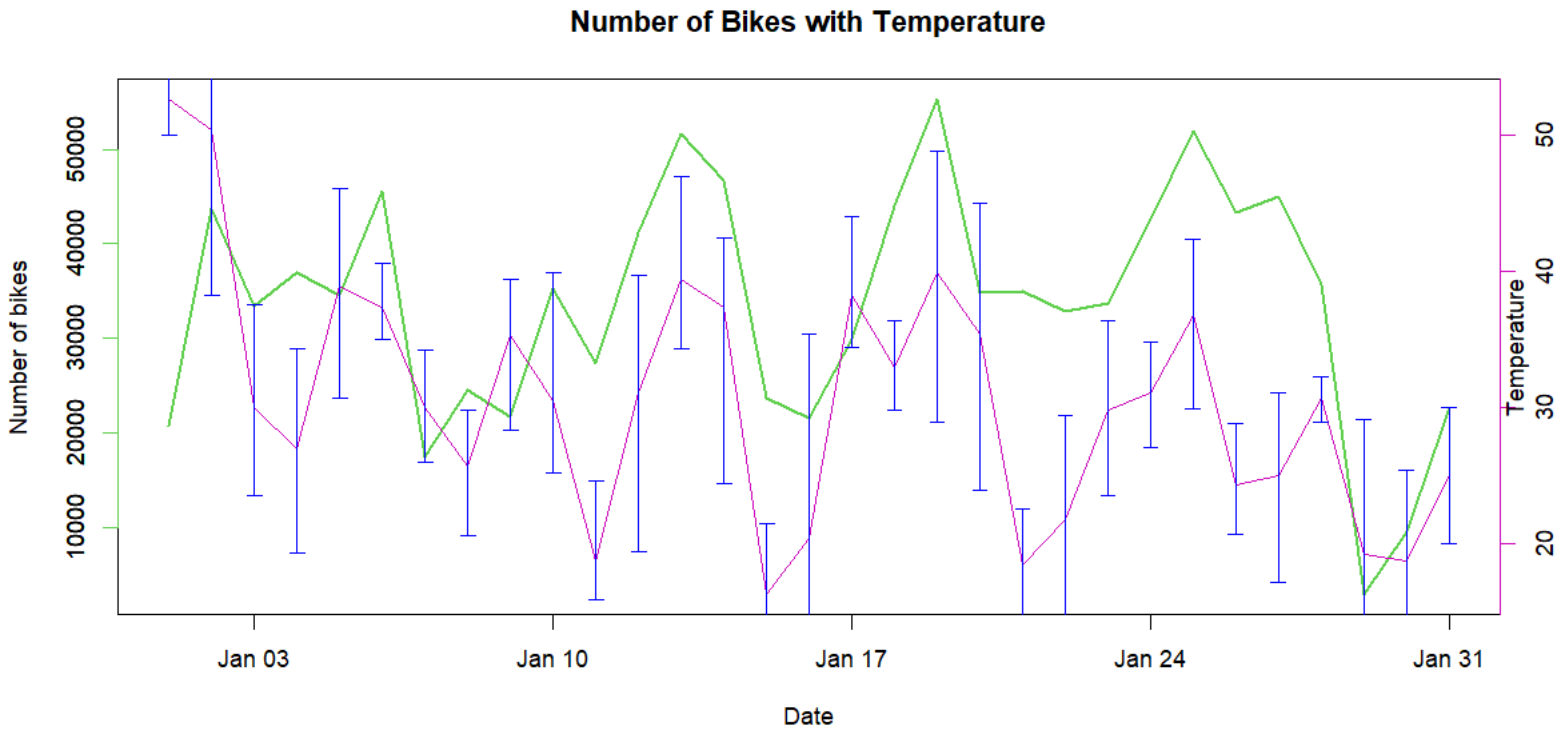
Average CO2 emission reduced



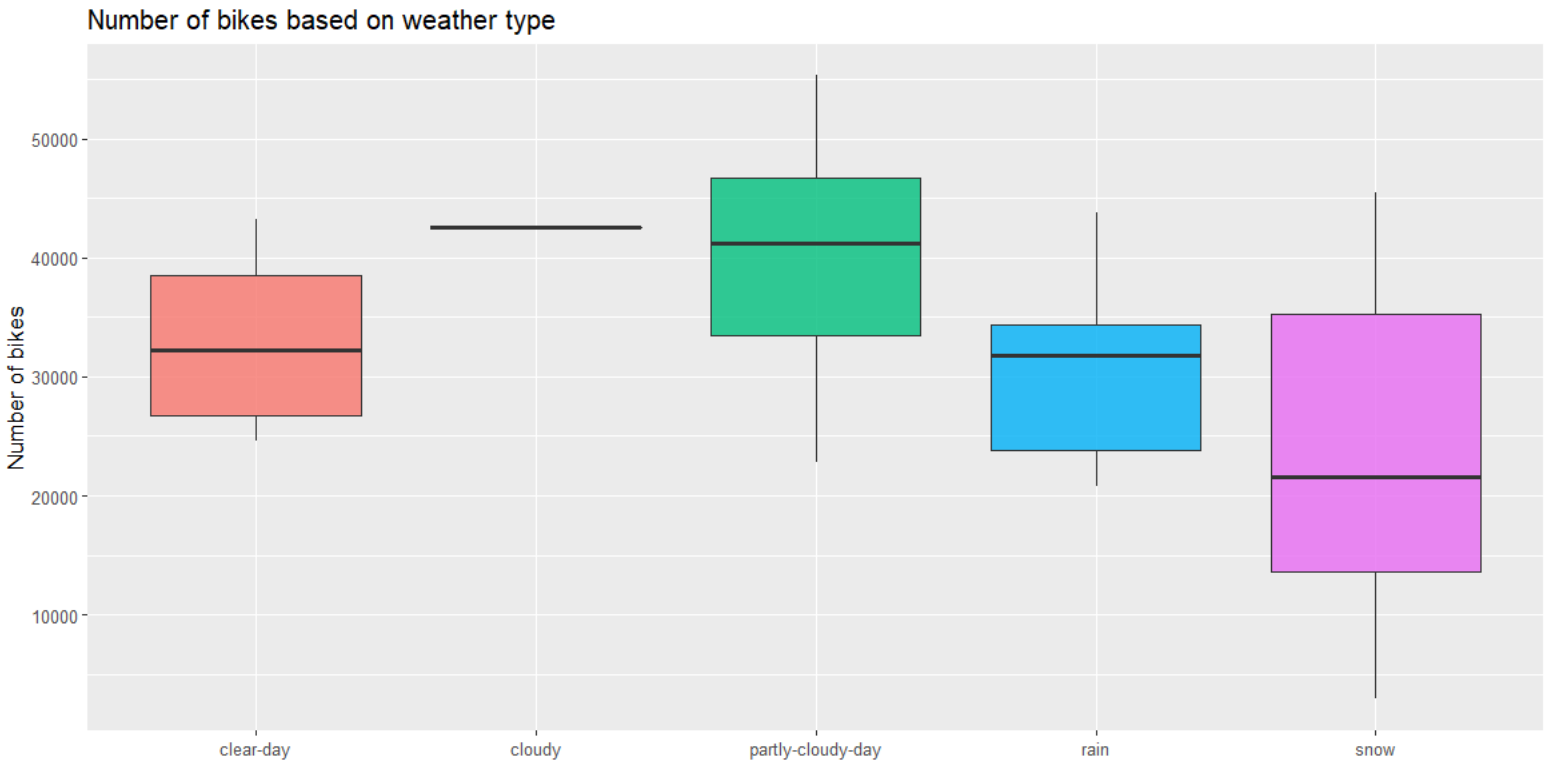
Density plot of distance travelled (by membership)



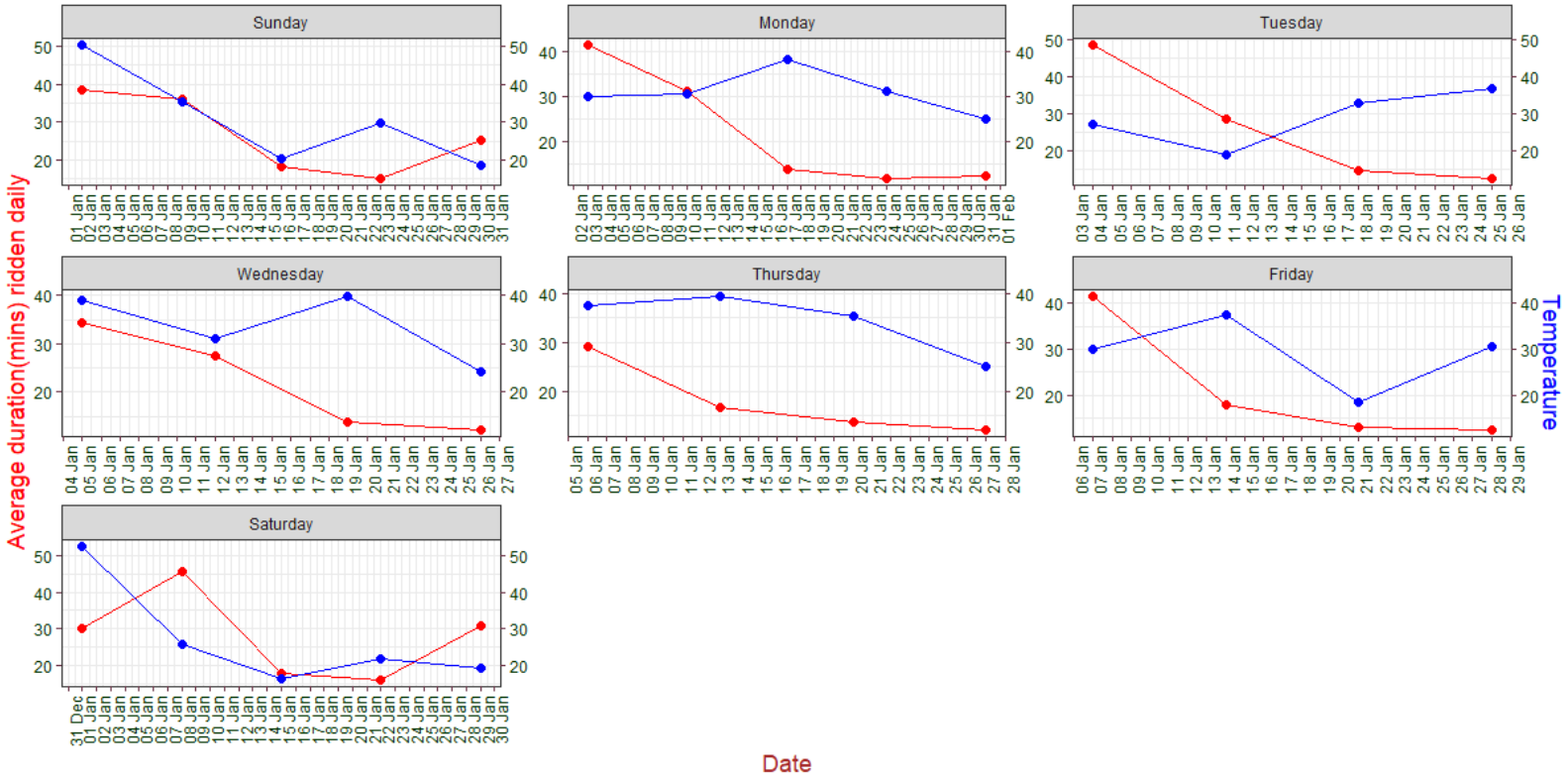
Number of bikes with Temperature (Min and Max)



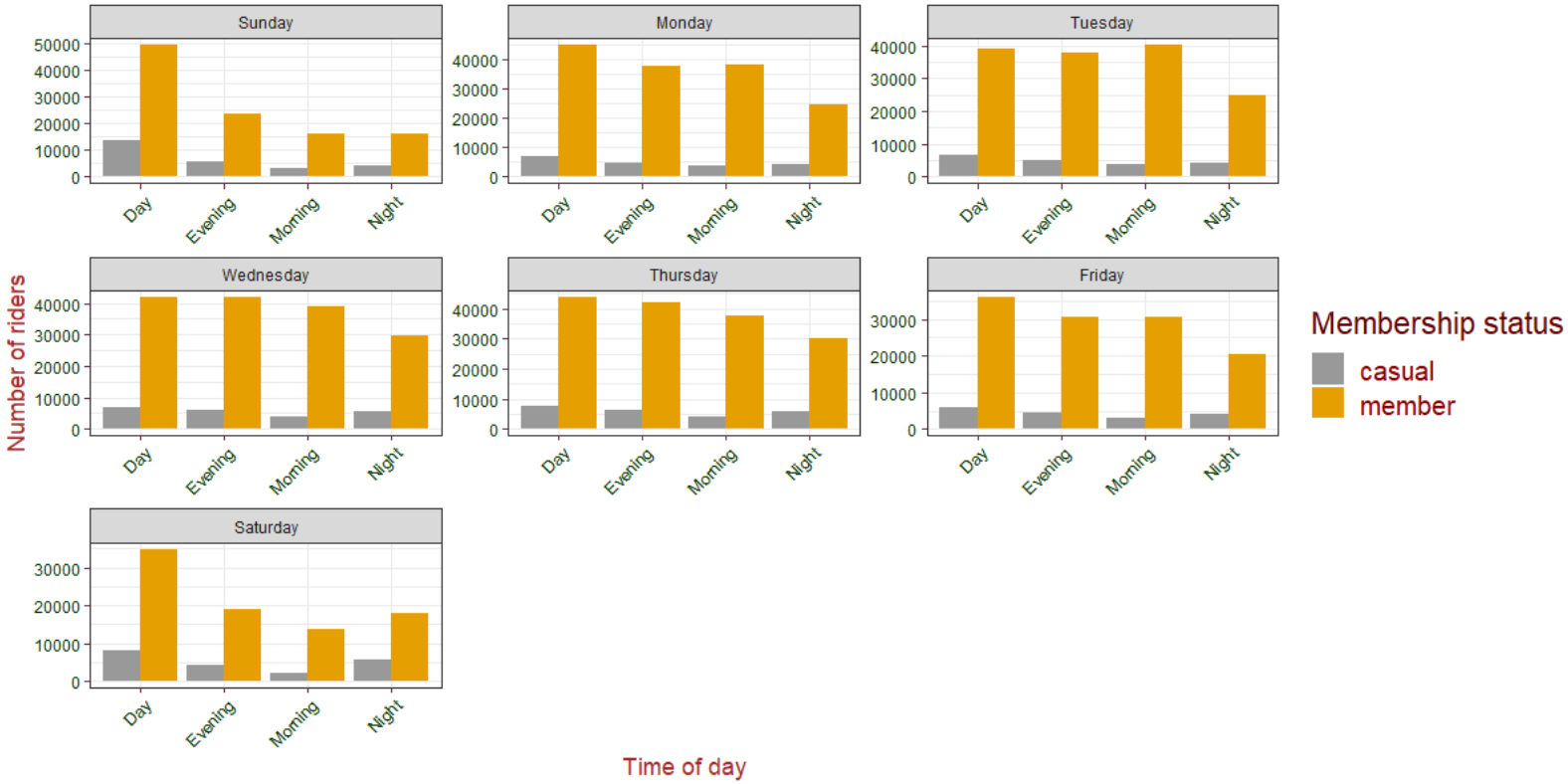
Number of bikers by weather type



Average duration of rides daily against Temperature



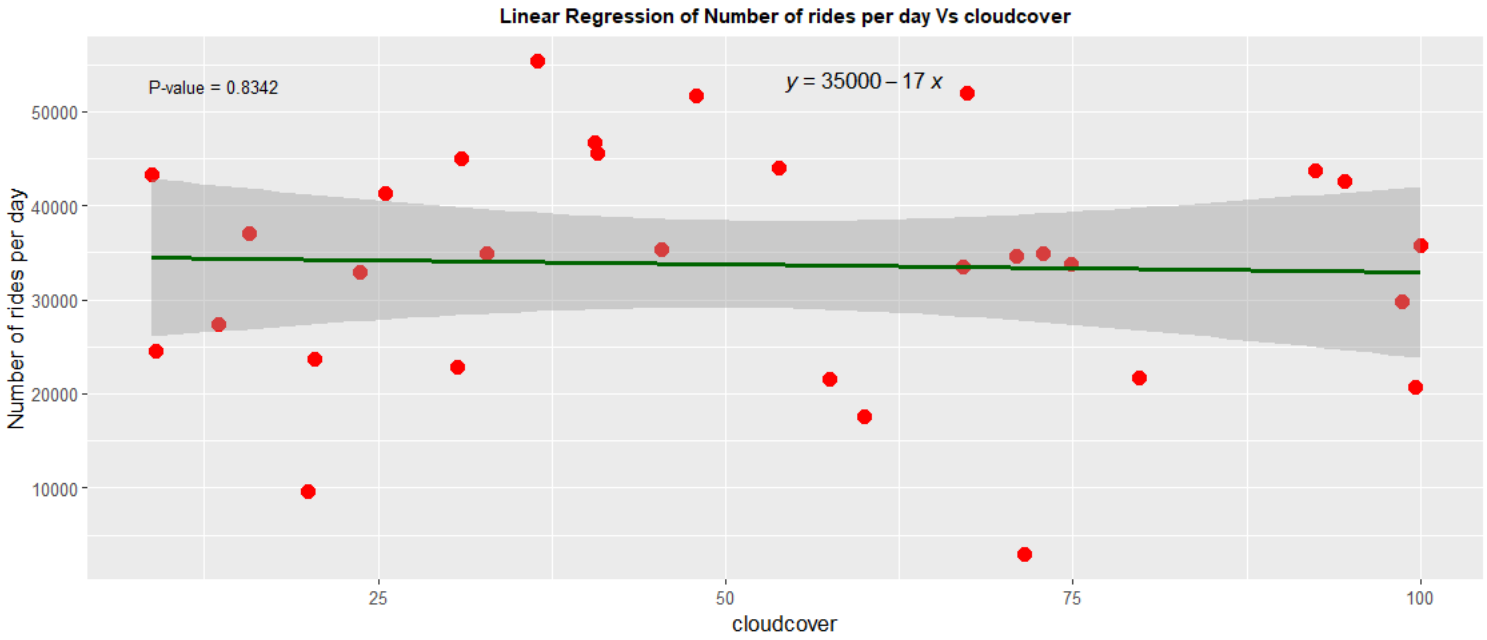
Number of Riders by Membership status and Time of day



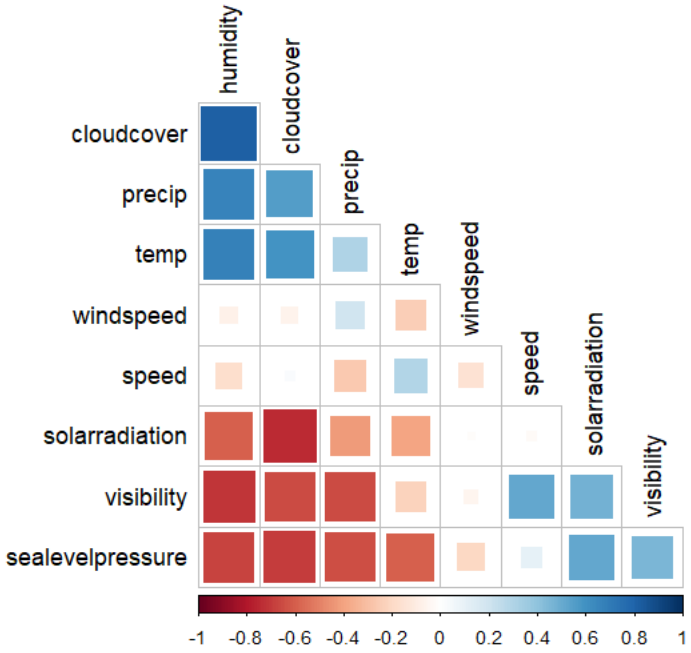
Linear Regression: Number of Rides ~ Weather conditions

Select Weather Condition

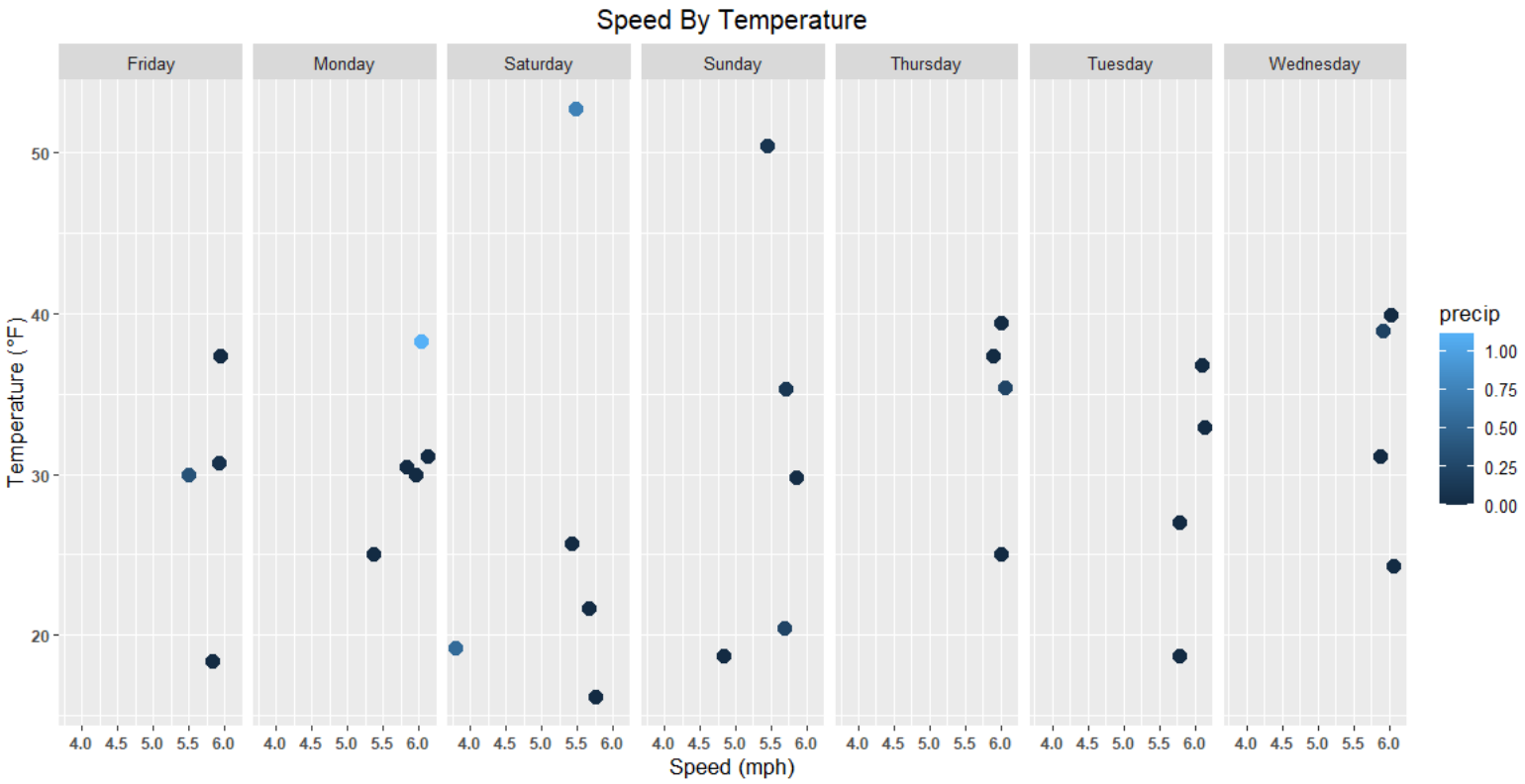
cloudcover



Correlation plot of Average Speed of Bikers per day against Weather Conditions



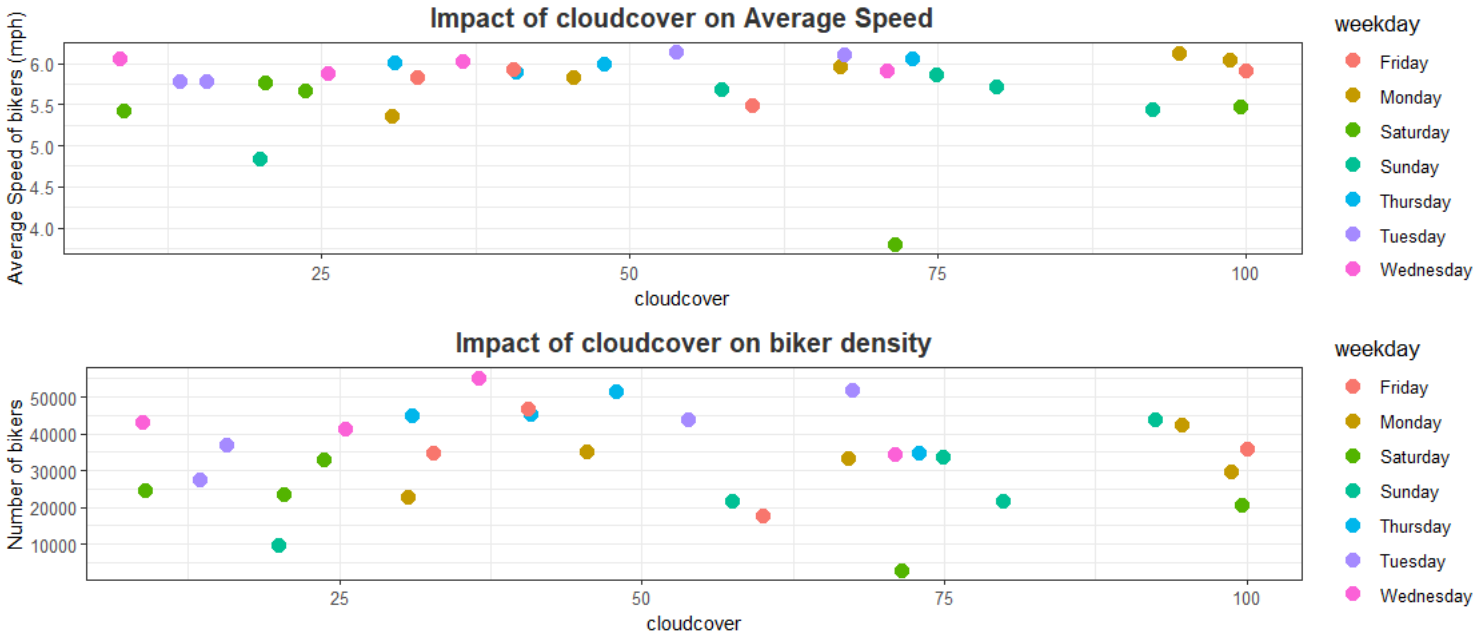
Speed By Temperature



Impact of Weather Conditions on Bike rides

Select Weather Condition

cloudcover



THANK YOU!!