The dataset consists of mixed data types (binary,categorical and numeric). The categorical variables include ‘weathersit’, ‘season’, ‘yr’ and ‘workingaday’ as a binary. The predictors ‘temp’, ‘hum’ and ‘windspeed’ are a class of continuous numeric data types. The response variables ‘casual’ and ‘registered’ are also continuous numeric class variables which are the number of casual and registered riders for a given day.

To plot correlations among the continuous and categorical (factor) variables, the two data types were split into data frames.

From the correlation matrices obtained for casual and registered riders with continuous predictors (temperature, windspeed and humidity), the variables are not significantly correlated. Multicollinearity does not affect the variables.

The analysis was determined by a Pearson Correlation test on the continuous predictors Temperature, Humidity and Windspeed. The following was concluded from the analysis.

*Temperature* shows a relatively high positive correlation coefficient *(0.54)* with the number of riders in both groups. It also shows a low negative correlation coefficient to *Windspeed (-0.157)* suggesting that on a windy day the temperature is expected to be low. *Humidity* shows a low positive correlation with *Temperature (0.1269)*.

*Humidity* and *Windspeed* are both negatively correlated with the number of casual *(-0.077* and *-0.167)* and registered riders *(-0.091* and *-0.21)*.

It is sufficiently clear from the analysis that *Temperature* is the major factor influencing ridership in both casual and registered groups of riders.