## Multiple Linear Regression and Correlation

### Anthony Sumter

Libraries

#install.packages("tidyverse","GGally","car","MASS", "gridExtra")  
library(tidyverse)

## -- Attaching packages ----------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1 v purrr 0.3.3  
## v tibble 2.1.3 v dplyr 0.8.3  
## v tidyr 1.0.0 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.4.0

## -- Conflicts -------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(GGally)

## Warning: package 'GGally' was built under R version 3.6.2

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

##   
## Attaching package: 'GGally'

## The following object is masked from 'package:dplyr':  
##   
## nasa

library(car)

## Warning: package 'car' was built under R version 3.6.2

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

## The following object is masked from 'package:purrr':  
##   
## some

library(MASS)

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

library(gridExtra)

## Warning: package 'gridExtra' was built under R version 3.6.2

##   
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':  
##   
## combine

#### Task 1

Read in the data from the “hour.csv” file into a data frame/tibble named “bike”. Convert “season” using the code.Convert “yr”, “mnth”, and “hr” to factors. You do NOT need to recode (rename) the levels of these factors.Convert the “holiday” variable to a factor and recode the levels from 0 to “NotHoliday” and 1 to “Holiday”.Convert “workingday” to a factor and recode the levels from 0 to “NotWorkingDay” and 1 to “WorkingDay”. Convert “weathersit” to a factor and recode the levels. Level 1 should be “NoPrecip”, 2 should become “Misty”, 3 should become “LightPrecip”, and 4 should become “HeavyPrecip”. Convert the “weekday” variable to a factor and recode the levels. Note that 6 is “Saturday” and 0 is “Sunday”. The rest of the days of the week are from 1 to 5, starting with “Monday”. Comment as to why we convert “yr”, “mnth”, and “hr” into factors? Why not just leave them as numbers?

**It is better to convert these variables into factors because it will be more efficient to to use in the regression. The model will fit better and probably deliver better predictions.**

bike = read\_csv("hour.csv")

## Parsed with column specification:  
## cols(  
## instant = col\_double(),  
## dteday = col\_date(format = ""),  
## season = col\_double(),  
## yr = col\_double(),  
## mnth = col\_double(),  
## hr = col\_double(),  
## holiday = col\_double(),  
## weekday = col\_double(),  
## workingday = col\_double(),  
## weathersit = col\_double(),  
## temp = col\_double(),  
## atemp = col\_double(),  
## hum = col\_double(),  
## windspeed = col\_double(),  
## casual = col\_double(),  
## registered = col\_double(),  
## count = col\_double()  
## )

bike = bike %>% mutate(season = as\_factor(as.character(season))) %>%  
mutate(season = fct\_recode(season,  
"Spring" = "1",  
"Summer" = "2",  
"Fall" = "3",  
"Winter" = "4"))  
bike = bike %>% mutate(yr = as\_factor(as.character(yr)), mnth = as\_factor (as.character (mnth)), hr = as\_factor(as.character (hr)), holiday = as\_factor(as.character (holiday)), workingday = as\_factor(as.character(workingday)), weathersit = as\_factor(as.character(weathersit)), weekday = as\_factor(as.character(weekday))) %>%  
mutate(holiday = fct\_recode(holiday,"NotHoliday" = "0","Holiday" = "1")) %>% mutate(workingday = fct\_recode(workingday,"NotWorkingDay" = "0","WorkingDay" = "1"))%>% mutate(weathersit = fct\_recode(weathersit,"NoPrecip" = "1","Misty" = "2", "LightPrecip"= "3", "HeavyPrecip" = "4"))%>% mutate(weekday = fct\_recode(weekday,"Sunday" = "0", "Monday" = "1", "Tuesday" = "2", "Wednesday" = "3", "Thursday" = "4", "Friday" = "5", "Saturday" = "6"))

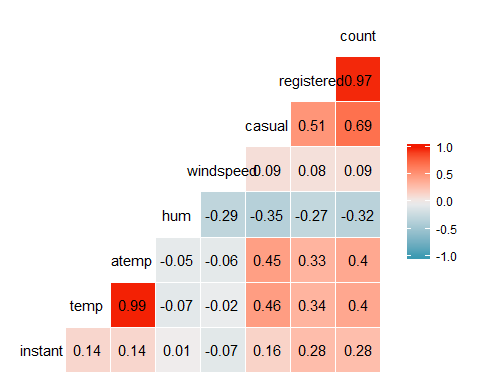
#### Task 2

Which of the quantitative variables appears to be best correlated with “count”?

**It appears that the variables best correlated with count are temp and atemp.**

ggcorr(bike,label = TRUE,label\_round = 2)

## Warning in ggcorr(bike, label = TRUE, label\_round = 2): data in column(s)  
## 'dteday', 'season', 'yr', 'mnth', 'hr', 'holiday', 'weekday', 'workingday',  
## 'weathersit' are not numeric and were ignored

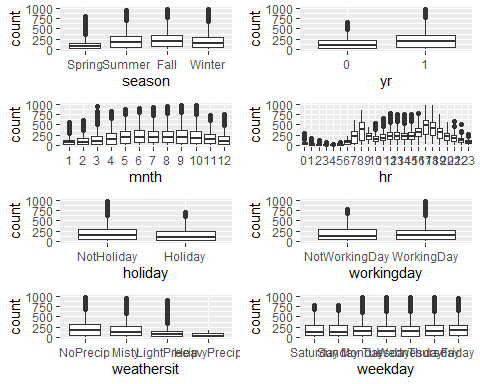


#### Task 3

We cannot use correlation to assess the relationship between a categorical predictor variable and our response variable. One option is to visualize the relationship between the categorical and response variables via a boxplot (or similar visualization). Conduct a boxplot-based analysis for each of the categorical variables.Which variables appear to affect “count”? Provide a brief explanation as to why you believe that each variable does or does not affect “count” (use your intution to help you answer this question). I strongly suggest using grid.arrange to reduce the space needed for your plots.

**The season variable does affect bike sharing because the summer and fall months tend to produce better weather making people more likely to ride a bike. The Year doesn’t affect it because you are only evaluating two years worth of data (very limited). Months affect bike sharing as, just like the season variable,the count rises in the nicer weather months. Hours affect bike sharing as the most busiest rides take place during work hours to get to work and to get home from work. Holiday affects bike sharing because on non holidays people will more likely need to use the service becuase of having to work, if its a holiday people aren’t working so no need for a bike. Working Day affects bike sharing for the same reasons as previously mentioned for holidays; if you have to work you need to find a way to get there, if your not working theres not a strong need. The weather affects bike sharing because if it’s bad weather conditions it becomes dangeourous to ride and people are less likely to take the risk of getting into an accident due to poor weather. Weekday affects the count because there is a greater need to use a bike as Monday through Friday is generally accepted as a typical work week; Saturday and Sunday are not always factored into the work week lessening the need to use a bike sharing service.**

p1 = ggplot(bike, aes(x=season,y=count)) + geom\_boxplot()  
p2 = ggplot(bike, aes(x=yr,y=count)) + geom\_boxplot()  
p3 = ggplot(bike, aes(x=mnth,y=count)) + geom\_boxplot()  
p4 = ggplot(bike, aes(x=hr,y=count)) + geom\_boxplot()  
p5 = ggplot(bike, aes(x=holiday,y=count)) + geom\_boxplot()  
p6 = ggplot(bike, aes(x=workingday,y=count)) + geom\_boxplot()  
p7 = ggplot(bike, aes(x=weathersit,y=count)) + geom\_boxplot()  
p8 = ggplot(bike, aes(x=weekday,y=count)) + geom\_boxplot()  
grid.arrange(p1,p2,p3,p4, p5,p6,p7,p8, ncol = 2)



#### Task 4

Use forward stepwise regression to build a multiple linear regression model to predict “count”.What variables are included in your forward model? Comment on the quality of the model. Does the modelmatch our intuition/common sense? Is there evidence of multicollinearity?

**The variables that are included in the forward model are hr, temp, atemp, hum, mnth, season, yr, weathersit, windspeed, holiday, workingday, and weekday. The model quality is pretty good based on the r squared value being .69 and the p value being less than .05. The model does happen to match common sense even though there is evidence of multicollinearity as proof by various variables having negative coefficients.**

bike2 = bike %>% dplyr::select(-c(instant, dteday, registered, casual))  
  
allmod = lm(count ~., bike2)   
  
emptymod = lm(count ~1, bike2)   
  
forwardmod = stepAIC(emptymod, direction = "forward", scope=list(upper=allmod,lower=emptymod),trace = TRUE)

## Start: AIC=180764.7  
## count ~ 1  
##   
## Df Sum of Sq RSS AIC  
## + hr 23 286734681 285026910 168713  
## + temp 1 93677759 478083832 177657  
## + atemp 1 91907421 479854170 177721  
## + hum 1 59618351 512143240 178853  
## + mnth 11 42909976 528851615 179431  
## + season 3 37729358 534032233 179584  
## + yr 1 35876722 535884870 179641  
## + weathersit 3 12285030 559476561 180393  
## + windspeed 1 4970060 566791531 180615  
## + holiday 1 546889 571214702 180750  
## + workingday 1 524387 571237204 180751  
## + weekday 6 687929 571073662 180756  
## <none> 571761591 180765  
##   
## Step: AIC=168712.5  
## count ~ hr  
##   
## Df Sum of Sq RSS AIC  
## + atemp 1 50518941 234507969 165324  
## + temp 1 50101685 234925225 165355  
## + mnth 11 44822160 240204750 165761  
## + season 3 39619754 245407156 166117  
## + yr 1 36875130 248151780 166307  
## + weathersit 3 13766672 271260238 167858  
## + hum 1 4924310 280102600 168412  
## + windspeed 1 1476211 283550699 168624  
## + holiday 1 561784 284465126 168680  
## + weekday 6 719530 284307380 168681  
## + workingday 1 485366 284541544 168685  
## <none> 285026910 168713  
##   
## Step: AIC=165324  
## count ~ hr + atemp  
##   
## Df Sum of Sq RSS AIC  
## + yr 1 33463769 201044200 162650  
## + weathersit 3 9227265 225280704 164632  
## + hum 1 7008684 227499285 164799  
## + season 3 6580442 227927527 164835  
## + mnth 11 5854560 228653409 164907  
## + weekday 6 607638 233900331 165291  
## + holiday 1 274006 234233963 165306  
## + temp 1 152153 234355816 165315  
## + windspeed 1 120557 234387412 165317  
## + workingday 1 90170 234417799 165319  
## <none> 234507969 165324  
##   
## Step: AIC=162650.2  
## count ~ hr + atemp + yr  
##   
## Df Sum of Sq RSS AIC  
## + weathersit 3 8408358 192635842 161914  
## + season 3 7190305 193853896 162023  
## + mnth 11 6486062 194558138 162102  
## + hum 1 4341837 196702363 162273  
## + weekday 6 641648 200402552 162607  
## + holiday 1 324763 200719438 162624  
## + windspeed 1 109311 200934889 162643  
## + workingday 1 106404 200937797 162643  
## + temp 1 91735 200952465 162644  
## <none> 201044200 162650  
##   
## Step: AIC=161913.7  
## count ~ hr + atemp + yr + weathersit  
##   
## Df Sum of Sq RSS AIC  
## + season 3 7771024 184864818 161204  
## + mnth 11 7464989 185170852 161249  
## + hum 1 805099 191830743 161843  
## + weekday 6 686172 191949670 161864  
## + holiday 1 413536 192222305 161878  
## + workingday 1 212428 192423414 161897  
## + temp 1 134482 192501360 161904  
## + windspeed 1 44407 192591435 161912  
## <none> 192635842 161914  
##   
## Step: AIC=161204.1  
## count ~ hr + atemp + yr + weathersit + season  
##   
## Df Sum of Sq RSS AIC  
## + mnth 11 2051323 182813495 161032  
## + hum 1 1810161 183054657 161035  
## + weekday 6 704303 184160515 161150  
## + holiday 1 392702 184472116 161169  
## + temp 1 352584 184512234 161173  
## + workingday 1 214973 184649845 161186  
## <none> 184864818 161204  
## + windspeed 1 158 184864660 161206  
##   
## Step: AIC=161032.2  
## count ~ hr + atemp + yr + weathersit + season + mnth  
##   
## Df Sum of Sq RSS AIC  
## + hum 1 2356411 180457084 160809  
## + weekday 6 692672 182120823 160978  
## + holiday 1 312321 182501174 161004  
## + temp 1 233052 182580443 161012  
## + workingday 1 203953 182609542 161015  
## <none> 182813495 161032  
## + windspeed 1 68 182813428 161034  
##   
## Step: AIC=160808.7  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum  
##   
## Df Sum of Sq RSS AIC  
## + weekday 6 581105 179875980 160765  
## + holiday 1 322997 180134087 160780  
## + workingday 1 194139 180262945 160792  
## + windspeed 1 114287 180342797 160800  
## + temp 1 100025 180357059 160801  
## <none> 180457084 160809  
##   
## Step: AIC=160764.7  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum +   
## weekday  
##   
## Df Sum of Sq RSS AIC  
## + holiday 1 274717 179601263 160740  
## + workingday 1 274717 179601263 160740  
## + windspeed 1 112085 179763895 160756  
## + temp 1 77171 179798809 160759  
## <none> 179875980 160765  
##   
## Step: AIC=160740.1  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum +   
## weekday + holiday  
##   
## Df Sum of Sq RSS AIC  
## + windspeed 1 111562 179489701 160731  
## + temp 1 95460 179505803 160733  
## <none> 179601263 160740  
##   
## Step: AIC=160731.3  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum +   
## weekday + holiday + windspeed  
##   
## Df Sum of Sq RSS AIC  
## + temp 1 160954 179328746 160718  
## <none> 179489701 160731  
##   
## Step: AIC=160717.7  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum +   
## weekday + holiday + windspeed + temp  
##   
## Df Sum of Sq RSS AIC  
## <none> 179328746 160718

summary(forwardmod)

##   
## Call:  
## lm(formula = count ~ hr + atemp + yr + weathersit + season +   
## mnth + hum + weekday + holiday + windspeed + temp, data = bike2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -393.87 -60.66 -7.96 51.31 439.18   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -67.542 6.612 -10.216 < 2e-16 \*\*\*  
## hr1 -17.294 5.345 -3.236 0.00122 \*\*   
## hr2 -26.369 5.364 -4.916 8.91e-07 \*\*\*  
## hr3 -37.112 5.403 -6.869 6.67e-12 \*\*\*  
## hr4 -40.263 5.408 -7.445 1.01e-13 \*\*\*  
## hr5 -23.501 5.373 -4.374 1.23e-05 \*\*\*  
## hr6 35.393 5.359 6.605 4.10e-11 \*\*\*  
## hr7 170.418 5.348 31.864 < 2e-16 \*\*\*  
## hr8 310.801 5.342 58.183 < 2e-16 \*\*\*  
## hr9 163.101 5.347 30.501 < 2e-16 \*\*\*  
## hr10 108.444 5.370 20.196 < 2e-16 \*\*\*  
## hr11 133.843 5.409 24.742 < 2e-16 \*\*\*  
## hr12 173.142 5.456 31.735 < 2e-16 \*\*\*  
## hr13 168.102 5.494 30.600 < 2e-16 \*\*\*  
## hr14 152.249 5.525 27.558 < 2e-16 \*\*\*  
## hr15 161.707 5.535 29.213 < 2e-16 \*\*\*  
## hr16 223.834 5.524 40.522 < 2e-16 \*\*\*  
## hr17 377.535 5.491 68.750 < 2e-16 \*\*\*  
## hr18 345.587 5.455 63.350 < 2e-16 \*\*\*  
## hr19 236.919 5.404 43.841 < 2e-16 \*\*\*  
## hr20 157.293 5.375 29.266 < 2e-16 \*\*\*  
## hr21 107.840 5.353 20.147 < 2e-16 \*\*\*  
## hr22 70.907 5.343 13.272 < 2e-16 \*\*\*  
## hr23 32.112 5.338 6.015 1.83e-09 \*\*\*  
## atemp 127.975 30.624 4.179 2.94e-05 \*\*\*  
## yr1 85.431 1.563 54.658 < 2e-16 \*\*\*  
## weathersitMisty -10.409 1.920 -5.421 6.00e-08 \*\*\*  
## weathersitLightPrecip -65.189 3.236 -20.145 < 2e-16 \*\*\*  
## weathersitHeavyPrecip -62.580 58.893 -1.063 0.28797   
## seasonSummer 38.178 4.856 7.862 4.00e-15 \*\*\*  
## seasonFall 32.055 5.749 5.575 2.51e-08 \*\*\*  
## seasonWinter 67.994 4.882 13.928 < 2e-16 \*\*\*  
## mnth2 3.426 3.920 0.874 0.38219   
## mnth3 14.299 4.407 3.244 0.00118 \*\*   
## mnth4 6.230 6.548 0.951 0.34144   
## mnth5 20.657 7.007 2.948 0.00320 \*\*   
## mnth6 6.238 7.205 0.866 0.38662   
## mnth7 -13.269 8.082 -1.642 0.10065   
## mnth8 7.897 7.879 1.002 0.31622   
## mnth9 32.269 7.001 4.609 4.07e-06 \*\*\*  
## mnth10 15.843 6.483 2.444 0.01455 \*   
## mnth11 -9.840 6.238 -1.577 0.11474   
## mnth12 -6.256 4.954 -1.263 0.20672   
## hum -82.802 5.554 -14.909 < 2e-16 \*\*\*  
## weekdaySunday -16.089 2.878 -5.591 2.30e-08 \*\*\*  
## weekdayMonday -6.814 2.970 -2.294 0.02180 \*   
## weekdayTuesday -5.240 2.899 -1.807 0.07071 .   
## weekdayWednesday -2.464 2.894 -0.851 0.39469   
## weekdayThursday -2.940 2.892 -1.016 0.30947   
## weekdayFriday 1.356 2.885 0.470 0.63823   
## holidayHoliday -26.228 4.881 -5.374 7.81e-08 \*\*\*  
## windspeed -29.167 7.052 -4.136 3.55e-05 \*\*\*  
## temp 116.384 29.513 3.943 8.06e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 101.7 on 17326 degrees of freedom  
## Multiple R-squared: 0.6864, Adjusted R-squared: 0.6854   
## F-statistic: 729.1 on 52 and 17326 DF, p-value: < 2.2e-16

#### 

#### Task 5

Repeat Task 4, but for backward stepwise. Does this model differ from the forward model? If so,how?

**The model does differ from the previous model as it takes into consideration only 11 variables versus 12 in the forward stepwise.Even with that being the case it still procduces the same result of the model quality being pretty good based on the r squared value being .69 and the p value being less than .05 and shows evidence of multicollinearity with various variables having negative coefficients.**

backmod = stepAIC(allmod, direction = "backward", trace = TRUE)

## Start: AIC=160717.7  
## count ~ season + yr + mnth + hr + holiday + weekday + workingday +   
## weathersit + temp + atemp + hum + windspeed  
##   
##   
## Step: AIC=160717.7  
## count ~ season + yr + mnth + hr + holiday + weekday + weathersit +   
## temp + atemp + hum + windspeed  
##   
## Df Sum of Sq RSS AIC  
## <none> 179328746 160718  
## - temp 1 160954 179489701 160731  
## - windspeed 1 177057 179505803 160733  
## - atemp 1 180751 179509498 160733  
## - holiday 1 298893 179627639 160745  
## - weekday 6 498795 179827541 160754  
## - mnth 11 2426171 181754917 160929  
## - hum 1 2300667 181629413 160937  
## - season 3 2398467 181727213 160943  
## - weathersit 3 4208731 183537478 161115  
## - yr 1 30920851 210249597 163480  
## - hr 23 196741474 376070220 173542

summary(backmod)

##   
## Call:  
## lm(formula = count ~ season + yr + mnth + hr + holiday + weekday +   
## weathersit + temp + atemp + hum + windspeed, data = bike2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -393.87 -60.66 -7.96 51.31 439.18   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -67.542 6.612 -10.216 < 2e-16 \*\*\*  
## seasonSummer 38.178 4.856 7.862 4.00e-15 \*\*\*  
## seasonFall 32.055 5.749 5.575 2.51e-08 \*\*\*  
## seasonWinter 67.994 4.882 13.928 < 2e-16 \*\*\*  
## yr1 85.431 1.563 54.658 < 2e-16 \*\*\*  
## mnth2 3.426 3.920 0.874 0.38219   
## mnth3 14.299 4.407 3.244 0.00118 \*\*   
## mnth4 6.230 6.548 0.951 0.34144   
## mnth5 20.657 7.007 2.948 0.00320 \*\*   
## mnth6 6.238 7.205 0.866 0.38662   
## mnth7 -13.269 8.082 -1.642 0.10065   
## mnth8 7.897 7.879 1.002 0.31622   
## mnth9 32.269 7.001 4.609 4.07e-06 \*\*\*  
## mnth10 15.843 6.483 2.444 0.01455 \*   
## mnth11 -9.840 6.238 -1.577 0.11474   
## mnth12 -6.256 4.954 -1.263 0.20672   
## hr1 -17.294 5.345 -3.236 0.00122 \*\*   
## hr2 -26.369 5.364 -4.916 8.91e-07 \*\*\*  
## hr3 -37.112 5.403 -6.869 6.67e-12 \*\*\*  
## hr4 -40.263 5.408 -7.445 1.01e-13 \*\*\*  
## hr5 -23.501 5.373 -4.374 1.23e-05 \*\*\*  
## hr6 35.393 5.359 6.605 4.10e-11 \*\*\*  
## hr7 170.418 5.348 31.864 < 2e-16 \*\*\*  
## hr8 310.801 5.342 58.183 < 2e-16 \*\*\*  
## hr9 163.101 5.347 30.501 < 2e-16 \*\*\*  
## hr10 108.444 5.370 20.196 < 2e-16 \*\*\*  
## hr11 133.843 5.409 24.742 < 2e-16 \*\*\*  
## hr12 173.142 5.456 31.735 < 2e-16 \*\*\*  
## hr13 168.102 5.494 30.600 < 2e-16 \*\*\*  
## hr14 152.249 5.525 27.558 < 2e-16 \*\*\*  
## hr15 161.707 5.535 29.213 < 2e-16 \*\*\*  
## hr16 223.834 5.524 40.522 < 2e-16 \*\*\*  
## hr17 377.535 5.491 68.750 < 2e-16 \*\*\*  
## hr18 345.587 5.455 63.350 < 2e-16 \*\*\*  
## hr19 236.919 5.404 43.841 < 2e-16 \*\*\*  
## hr20 157.293 5.375 29.266 < 2e-16 \*\*\*  
## hr21 107.840 5.353 20.147 < 2e-16 \*\*\*  
## hr22 70.907 5.343 13.272 < 2e-16 \*\*\*  
## hr23 32.112 5.338 6.015 1.83e-09 \*\*\*  
## holidayHoliday -26.228 4.881 -5.374 7.81e-08 \*\*\*  
## weekdaySunday -16.089 2.878 -5.591 2.30e-08 \*\*\*  
## weekdayMonday -6.814 2.970 -2.294 0.02180 \*   
## weekdayTuesday -5.240 2.899 -1.807 0.07071 .   
## weekdayWednesday -2.464 2.894 -0.851 0.39469   
## weekdayThursday -2.940 2.892 -1.016 0.30947   
## weekdayFriday 1.356 2.885 0.470 0.63823   
## weathersitMisty -10.409 1.920 -5.421 6.00e-08 \*\*\*  
## weathersitLightPrecip -65.189 3.236 -20.145 < 2e-16 \*\*\*  
## weathersitHeavyPrecip -62.580 58.893 -1.063 0.28797   
## temp 116.384 29.513 3.943 8.06e-05 \*\*\*  
## atemp 127.975 30.624 4.179 2.94e-05 \*\*\*  
## hum -82.802 5.554 -14.909 < 2e-16 \*\*\*  
## windspeed -29.167 7.052 -4.136 3.55e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 101.7 on 17326 degrees of freedom  
## Multiple R-squared: 0.6864, Adjusted R-squared: 0.6854   
## F-statistic: 729.1 on 52 and 17326 DF, p-value: < 2.2e-16

#### Task 6

If you look carefully, you will notice that the coefficients and p value for “workingday” in the model with all of the predictors (the model used to begin the backward stepwise approach) are listed as “NA”. This is typically a sign that that variable is perfectly correlated with another variable and is, thus, being “kicked out” of the model. Describe how “workingday” is represented in the model via other variables.

**Working day is represented in the model via other variables such as weekday (Sunday-Friday) and hours (1-23). The working day is conducted during the weekday and work is done betwen those specific hours.**

#### Task 7

Comment on the usability of this model. Any cautions concerning its potential use?

**The usuability of this model is pretty good as being supported by a decent r squared value and a significant p value. The cautions concerning the potential use is the presence of multicolinearity because it can become difficult to determine how each predictor affects the respone of bike sharing.**