## MultLinearRegressionAssign\_Tidy

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.4 v dplyr 1.0.2  
## v tidyr 1.1.2 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.0

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

library(tidymodels)

## -- Attaching packages -------------------------------------- tidymodels 0.1.2 --

## v broom 0.7.2 v recipes 0.1.15  
## v dials 0.0.9 v rsample 0.0.8   
## v infer 0.5.4 v tune 0.1.2   
## v modeldata 0.1.0 v workflows 0.2.1   
## v parsnip 0.1.5 v yardstick 0.0.7

## -- Conflicts ----------------------------------------- tidymodels\_conflicts() --  
## x scales::discard() masks purrr::discard()  
## x dplyr::filter() masks stats::filter()  
## x recipes::fixed() masks stringr::fixed()  
## x dplyr::lag() masks stats::lag()  
## x yardstick::spec() masks readr::spec()  
## x recipes::step() masks stats::step()

library(glmnet)

## Loading required package: Matrix

##   
## Attaching package: 'Matrix'

## The following objects are masked from 'package:tidyr':  
##   
## expand, pack, unpack

## Loaded glmnet 4.1

library(GGally)

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

library(ggcorrplot)  
library(MASS)

##   
## Attaching package: 'MASS'

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

library(car)

## 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(lubridate)

##   
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':  
##   
## date, intersect, setdiff, union

library(lmtest)

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

bike\_cleaned <- read\_csv("C:/Users/Karibu/Desktop/UNCW/Spring 1 2021/BAN 502/Module 1/Mod2Assignment2/bike\_cleaned.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## instant = col\_double(),  
## dteday = col\_character(),  
## season = col\_character(),  
## mnth = col\_character(),  
## hr = col\_double(),  
## holiday = col\_character(),  
## weekday = col\_character(),  
## workingday = col\_character(),  
## weathersit = col\_character(),  
## temp = col\_double(),  
## atemp = col\_double(),  
## hum = col\_double(),  
## windspeed = col\_double(),  
## casual = col\_double(),  
## registered = col\_double(),  
## count = col\_double()  
## )

bike = bike\_cleaned

bike = bike%>% mutate(dteday = mdy(dteday))

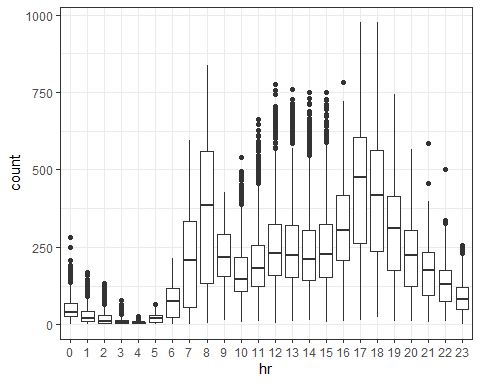
bike <- mutate\_if(bike, is.character, as.factor)

bike$hr <- as.factor(bike$hr)  
bike

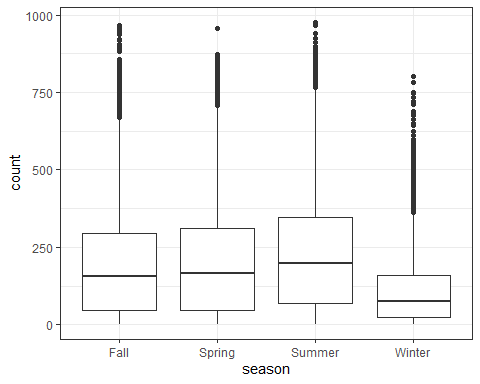
## # A tibble: 17,379 x 16  
## instant dteday season mnth hr holiday weekday workingday weathersit  
## <dbl> <date> <fct> <fct> <fct> <fct> <fct> <fct> <fct>   
## 1 1 2011-01-01 Winter Jan 0 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## 2 2 2011-01-01 Winter Jan 1 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## 3 3 2011-01-01 Winter Jan 2 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## 4 4 2011-01-01 Winter Jan 3 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## 5 5 2011-01-01 Winter Jan 4 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## 6 6 2011-01-01 Winter Jan 5 NotHol~ Saturd~ NotWorkin~ Misty   
## 7 7 2011-01-01 Winter Jan 6 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## 8 8 2011-01-01 Winter Jan 7 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## 9 9 2011-01-01 Winter Jan 8 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## 10 10 2011-01-01 Winter Jan 9 NotHol~ Saturd~ NotWorkin~ NoPrecip   
## # ... with 17,369 more rows, and 7 more variables: temp <dbl>, atemp <dbl>,  
## # hum <dbl>, windspeed <dbl>, casual <dbl>, registered <dbl>, count <dbl>

We coveretd hr into a factor to make utilize it as a categorical variable(levels) in multilinear regression analysis.

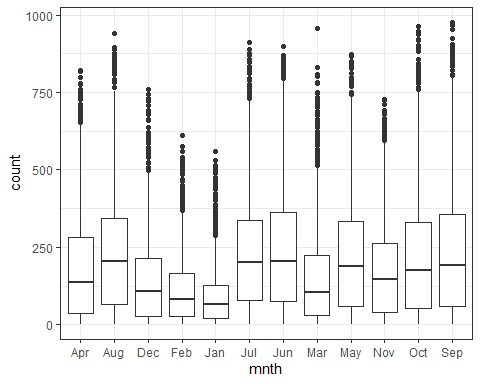
ggplot(bike,aes(x=hr,y=count))+ geom\_boxplot()+ theme\_bw()



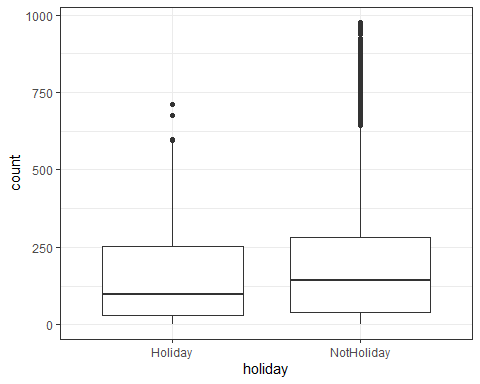
ggplot(bike,aes(x=season,y=count))+ geom\_boxplot()+ theme\_bw()



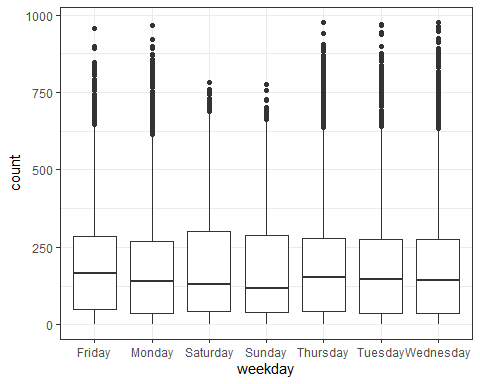
ggplot(bike,aes(x=mnth,y=count))+ geom\_boxplot()+ theme\_bw()



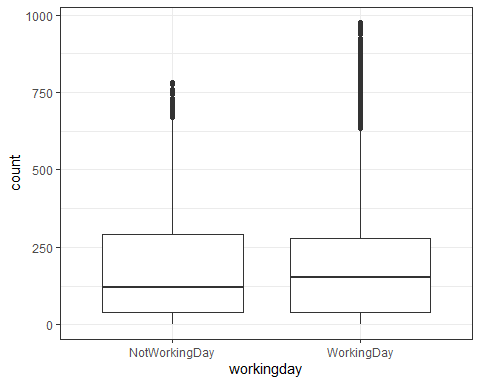
ggplot(bike,aes(x=holiday,y=count))+ geom\_boxplot()+ theme\_bw()



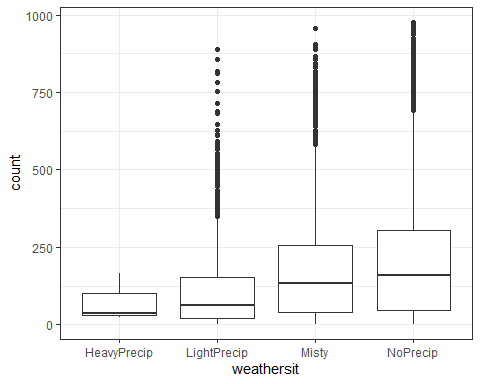
ggplot(bike,aes(x=weekday,y=count))+ geom\_boxplot()+ theme\_bw()



ggplot(bike,aes(x=workingday,y=count))+ geom\_boxplot()+ theme\_bw()



ggplot(bike,aes(x=weathersit,y=count))+ geom\_boxplot()+ theme\_bw()

 From the above boxplots, hr of the day(commute), weather(more bikes with no precipitation) , month and season (warmer seasons more bikes) affects the count of bicycles in the city, while holiday, working day are not good predictor variables.