DataVisualization

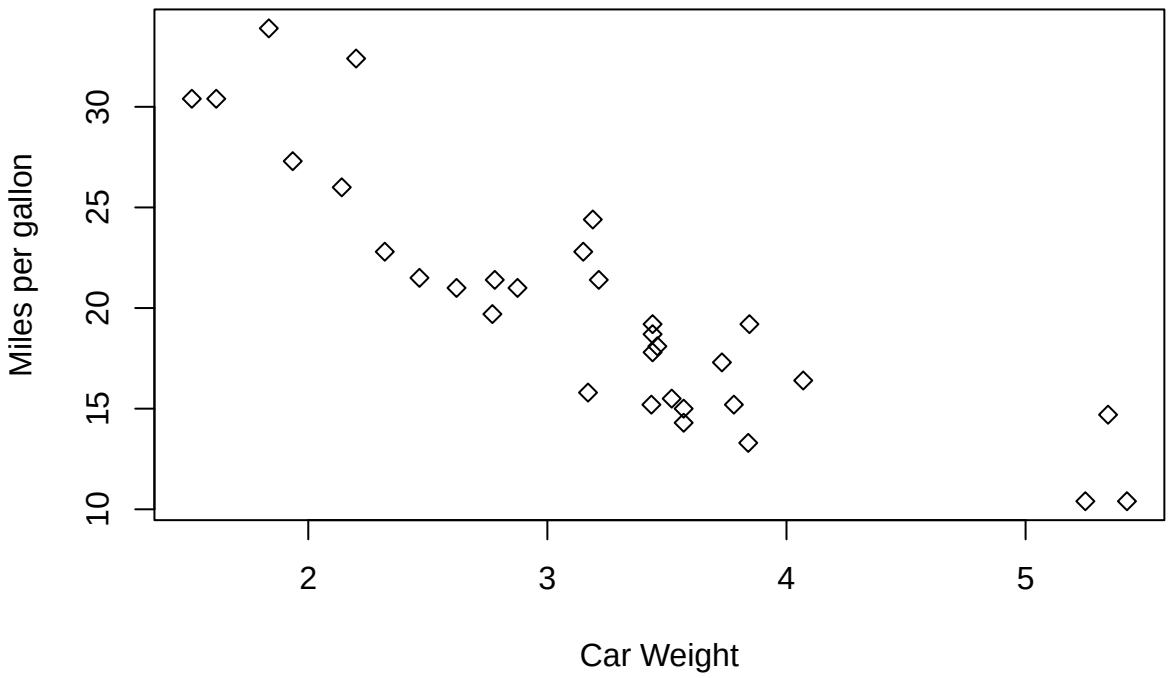
Arpan

2025-02-12

```
data("mtcars") #Read inbuilt dataset
str(mtcars) # Look at the structure of the data
```

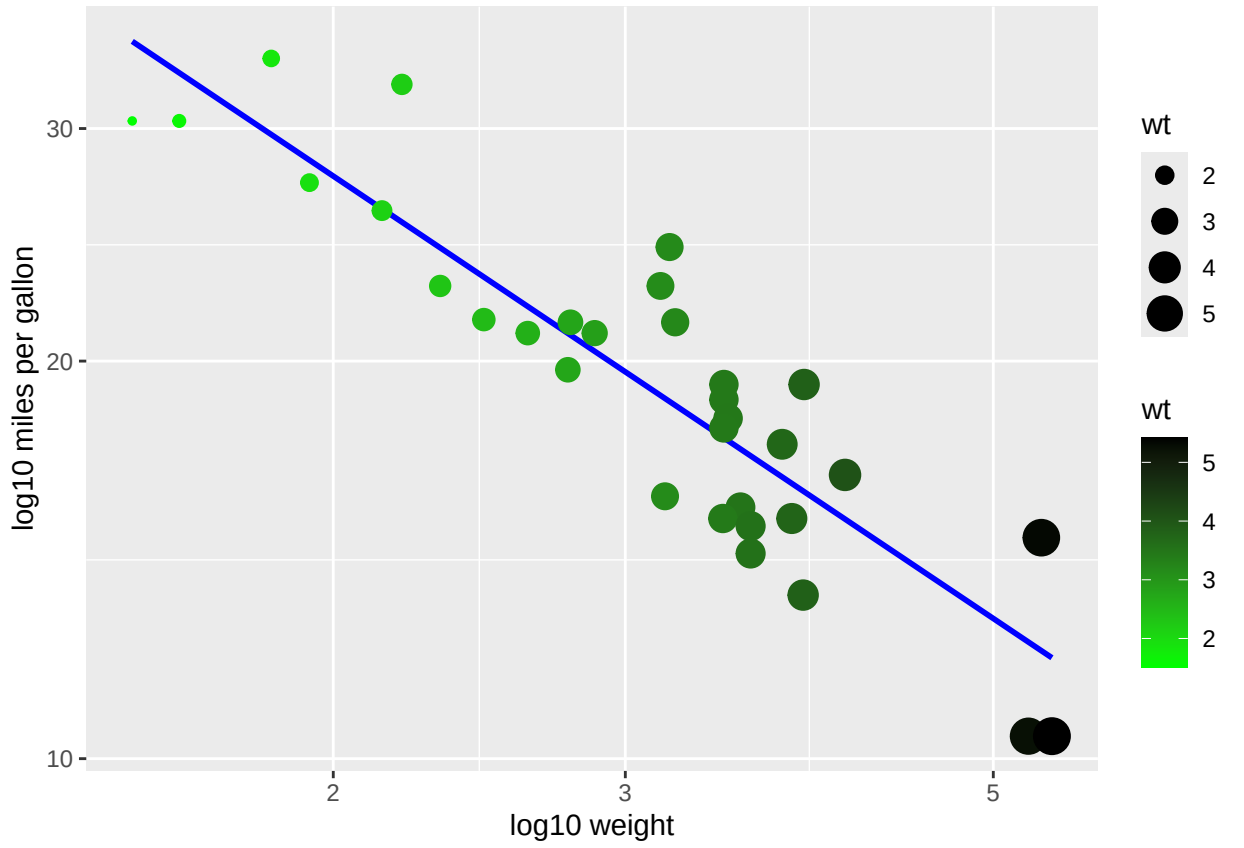
```
## 'data.frame':    32 obs. of  11 variables:
##  $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
##  $ cyl : num  6 6 4 6 8 6 8 4 4 6 ...
##  $ disp: num  160 160 108 258 360 ...
##  $ hp  : num  110 110 93 110 175 105 245 62 95 123 ...
##  $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
##  $ wt  : num  2.62 2.88 2.32 3.21 3.44 ...
##  $ qsec: num  16.5 17 18.6 19.4 17 ...
##  $ vs  : num  0 0 1 1 0 1 0 1 1 1 ...
##  $ am  : num  1 1 1 0 0 0 0 0 0 0 ...
##  $ gear: num  4 4 4 3 3 3 3 4 4 4 ...
##  $ carb: num  4 4 1 1 2 1 4 2 2 4 ...
```

```
#### plot using base R function
plot(mtcars$wt,mtcars$mpg,
     xlab="Car Weight", # add X label
     ylab="Miles per gallon",# add Y label
     font.lab=6,#change font
     pch=23 #Choose the shape for the data points in plot
     )
```



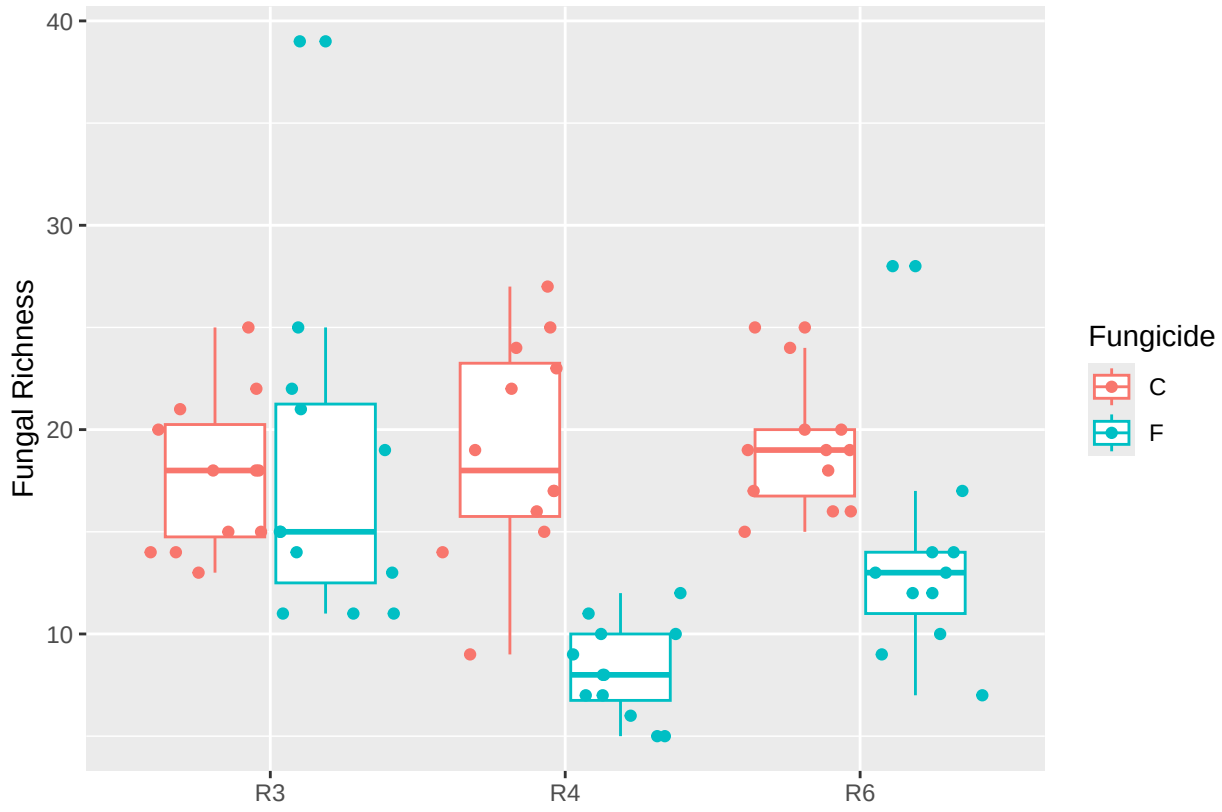

```
#####GGPLOT2####
library(ggplot2)
ggplot(mtcars,aes(x=wt, y=mpg))+ #size=wt))+#### using ggplot, it takes data,(x and y), and then we add
    geom_smooth(method=lm,se=FALSE,color="blue")+ #add a smooth line, showing linear relationship accor
  geom_point(aes(size=wt,color=wt))+ #adding layer where we show our data points, and put colour of poi
  xlab("log10 weight")+#add x label
  ylab("log10 miles per gallon")+# add y label
  scale_color_gradient(low="green",high="black")+#scaling by different colours gradient
  scale_x_log10()+
  scale_y_log10()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

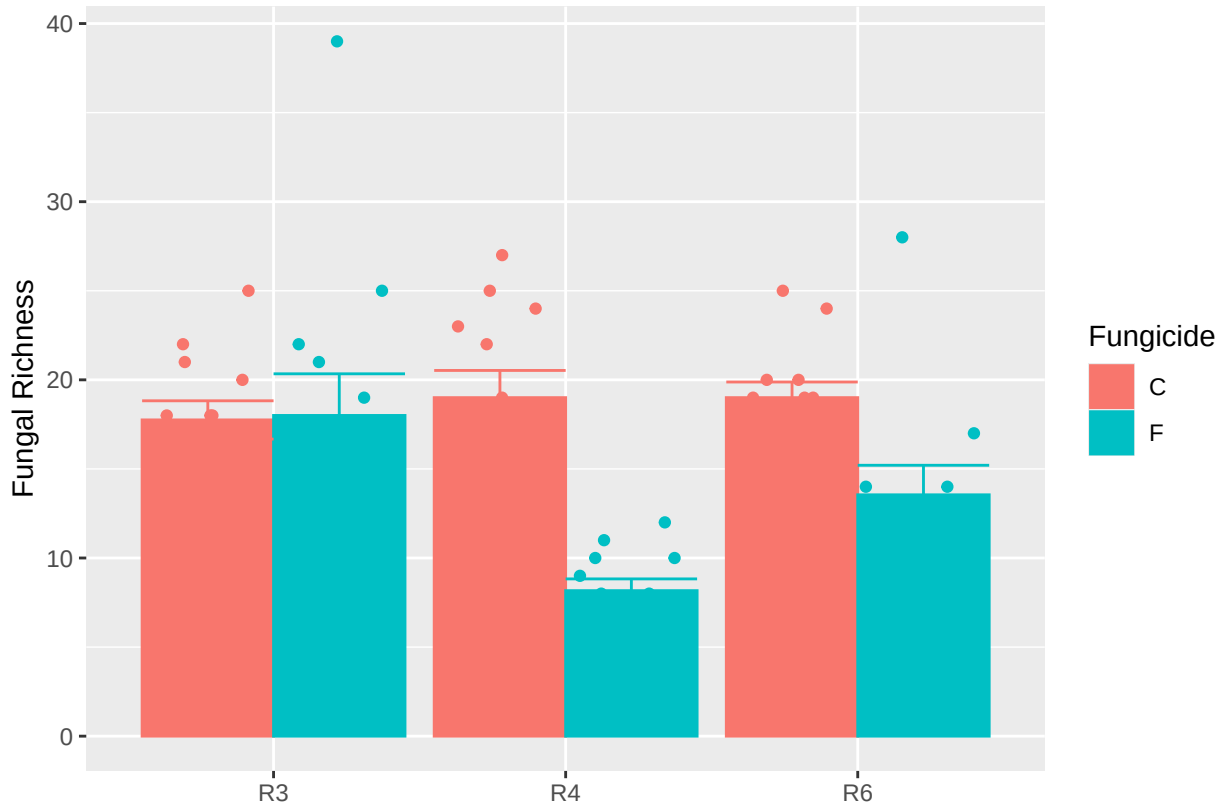



```
  #scale_y_continuous(labels=scales::percent)# transform proportion into percentage
```

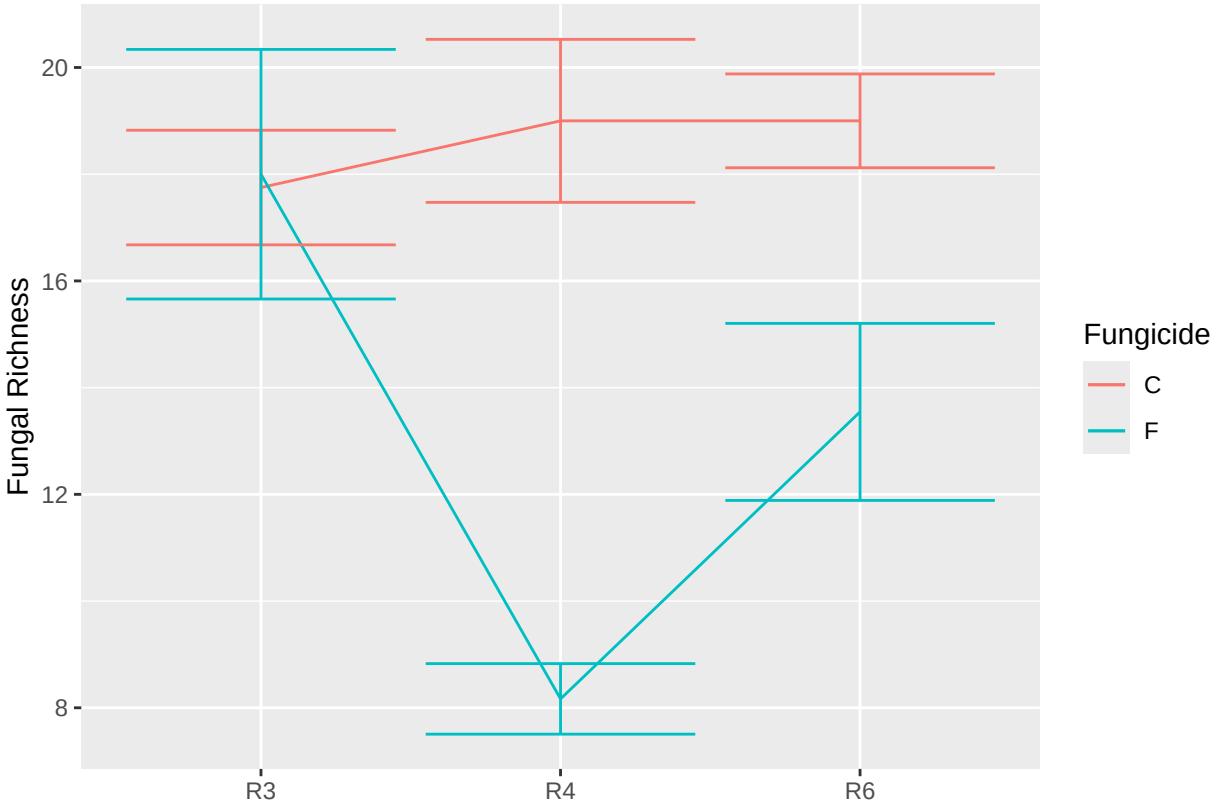
```
######ggplot with categorical X and numeric Y variables
bull.richness<-read.csv("bull_richness.csv")
#Subsetting the soybean data under no till condition
bull.richness.soy.no.till<-bull.richness[bull.richness$Crop=="Soy"&
                                           bull.richness$Treatment=="No-till",]
ggplot(bull.richness.soy.no.till,aes(x=GrowthStage,y=richness,colour=Fungicide))+
  geom_boxplot()+
  xlab("")+
  ylab("Fungal Richness")+
  #geom_point(position=position_dodge(width=0.9)) #width determines how far you want the points to dodg
  geom_point(position=position_jitterdodge(dodge.width=0.9))#jitterdodge ensures they are not overlappi
```



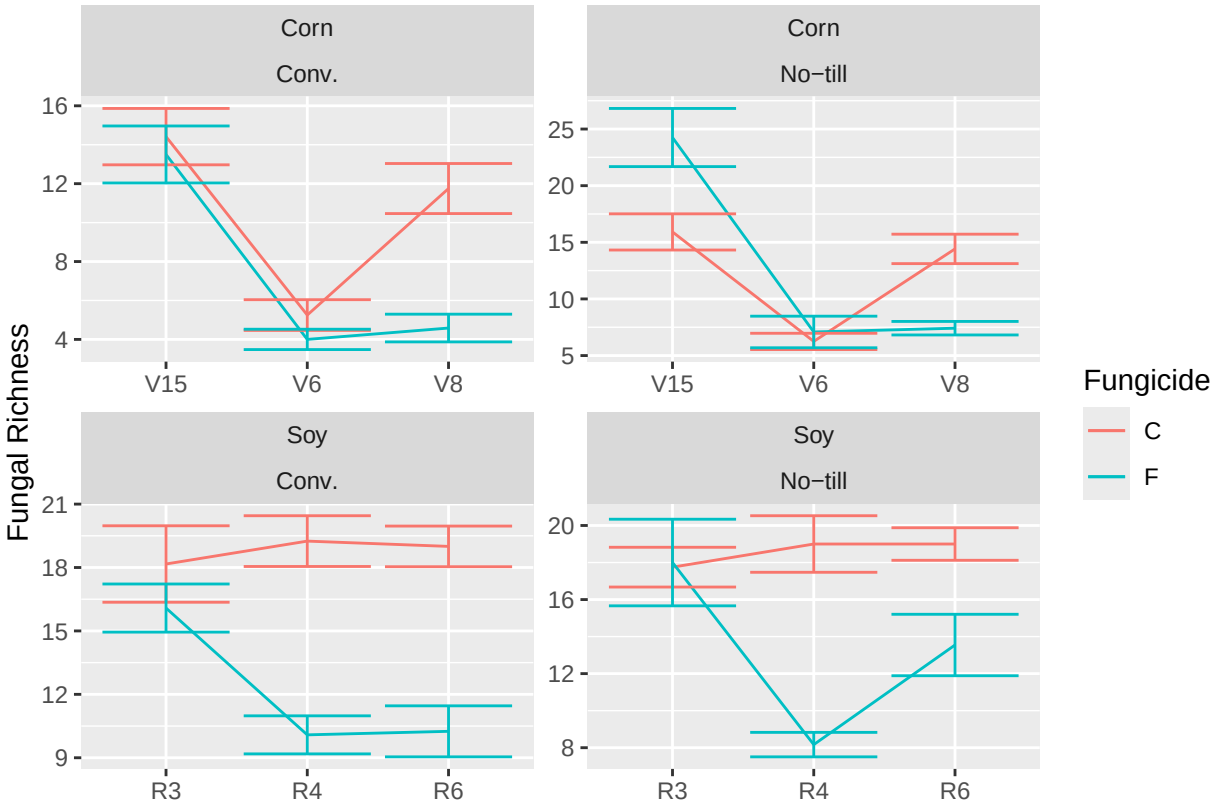

```
#####barcharts
ggplot(bull.richness.soy.no.till,aes(x=GrowthStage,y=richness,color=Fungicide,fill=Fungicide))+# using
   #geom_point(position=position_dodge(width=0.9)) #width determines how far you want the points to dod
   geom_point(position=position_jitterdodge(dodge.width=0.9))+#jitterdodge ensures they are not overlap
stat_summary(fun=mean,geom="bar",position="dodge")+ #plot the data by mean and dodge it
  stat_summary(fun.data=mean_se,geom="errorbar", position="dodge")+
  xlab("")+
  ylab("Fungal Richness")
```




```
######Line plot connecting means (change geom to line and take out dodge)
ggplot(bull.richness.soy.no.till,aes(x=GrowthStage,y=richness,group=Fungicide,color=Fungicide))+# using
   #geom_point(position=position_dodge(width=0.9)) #width determines how far you want the points to dod
   #geom_point(position=position_jitterdodge(dodge.width=0.9))+#jitterdodge ensures they are not overla
stat_summary(fun=mean,geom="line")+ #plot the data by mean and dodge it
  stat_summary(fun.data=mean_se,geom="errorbar")+
  xlab("")+
  ylab("Fungal Richness")
```




```
#####Faceting
ggplot(bull.richness, aes(x=GrowthStage,y=richness, group=Fungicide, colour=Fungicide))+
  stat_summary(fun=mean,geom="line")+ #plot the data by mean and dodge it
  stat_summary(fun.data=mean_se,geom="errorbar")+
  xlab("")+
  ylab("Fungal Richness")+
  facet_wrap(~Crop*Treatment,scales="free") # facetwrap splits the data, we can use * to split it by mu
```


Corn
Conv.
Corn
No−till
Soy
Conv.
Soy
No−till
Fungal Richness
16
12
8
4
25
20
15
10
5
21
18
15
12
9
20
16
12
8
V15
V6
V8
R3
R4
R6
Fungicide
C
F