Koala Drinking Behavior

Sruthi Srikantan

5/6/2022

```
# loading packages
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

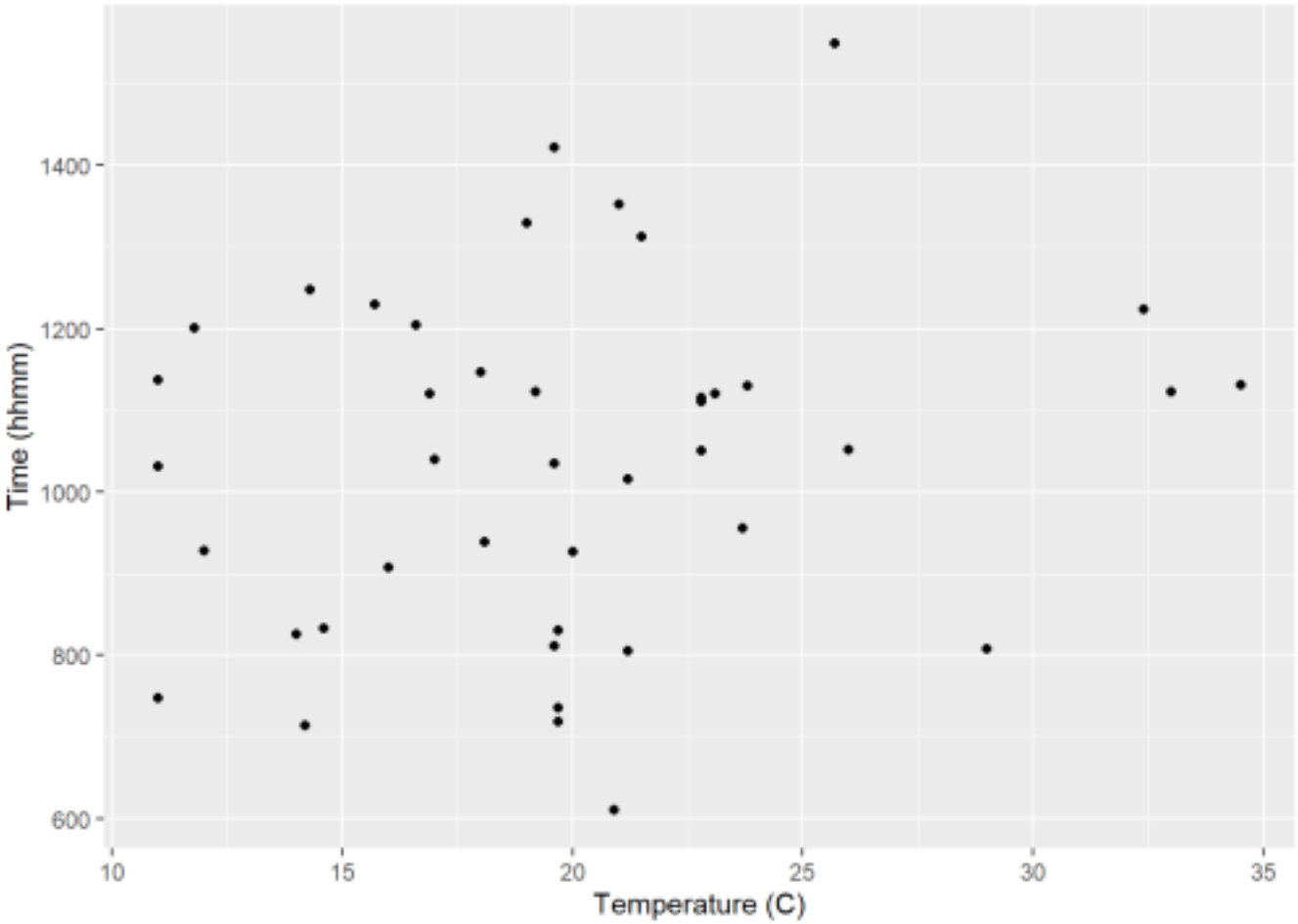
```
## v tibble 3.1.6 v dplyr 1.0.7
## v tidyr 1.2.0 v stringr 1.4.0
## v readr 2.1.2 v forcats 0.5.1
## v purrr 0.3.4
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()
```

```
library(dplyr)
```

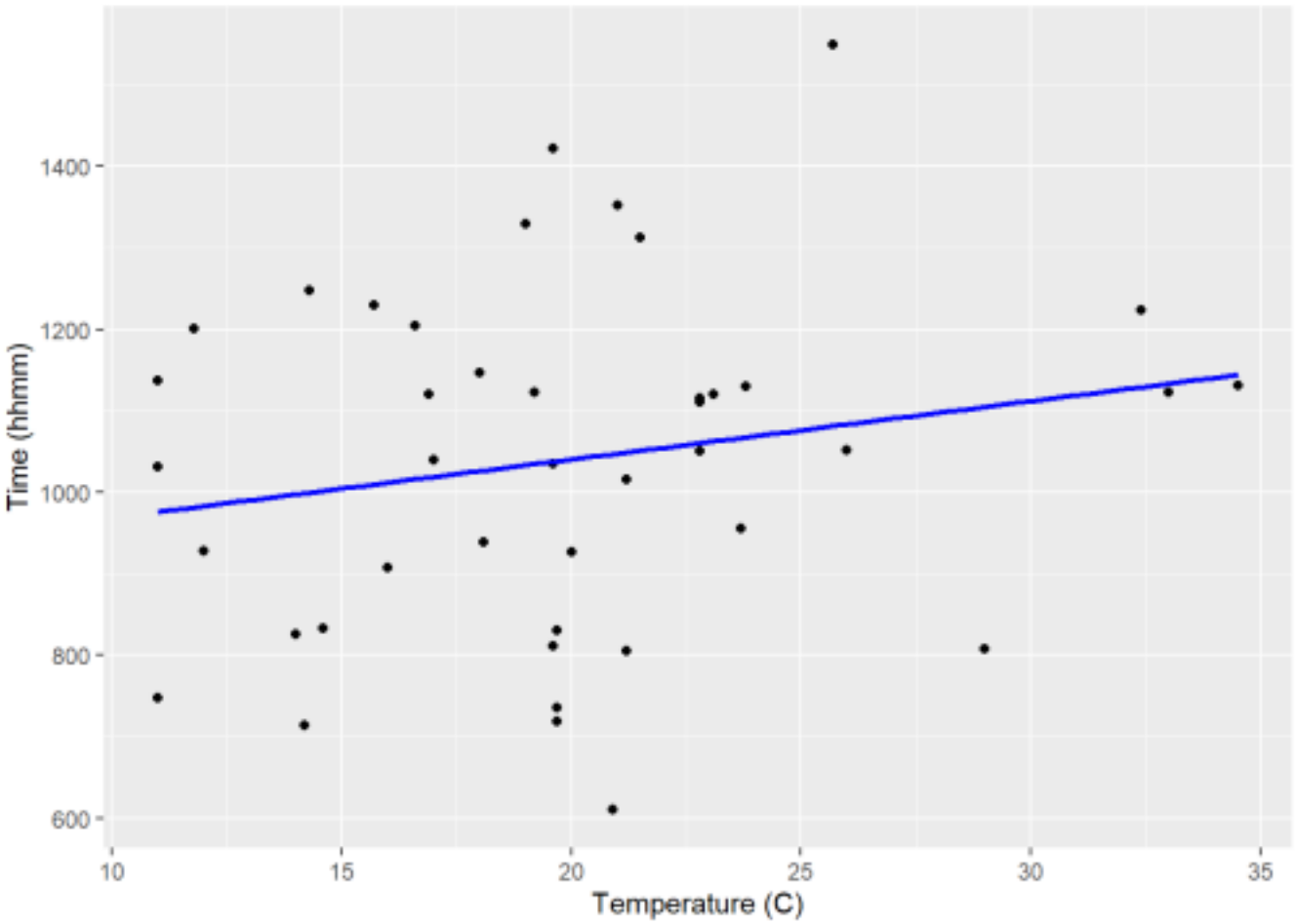
```
# This data is from the paper "An insight into natural koala drinking behavior"
Koalastand <- read.csv("Koala.csv")
```

```
# Here we're creating a basic scatter plot of our Koala data looking at the environmental temper
ature and time spent drinking
ggplot(data = Koalastand, aes(x = Temperature.max.C., y = Time.hhmm.)) + geom_point() + labs(x=
"Temperature (C)",y="Time (hhmm)")
```



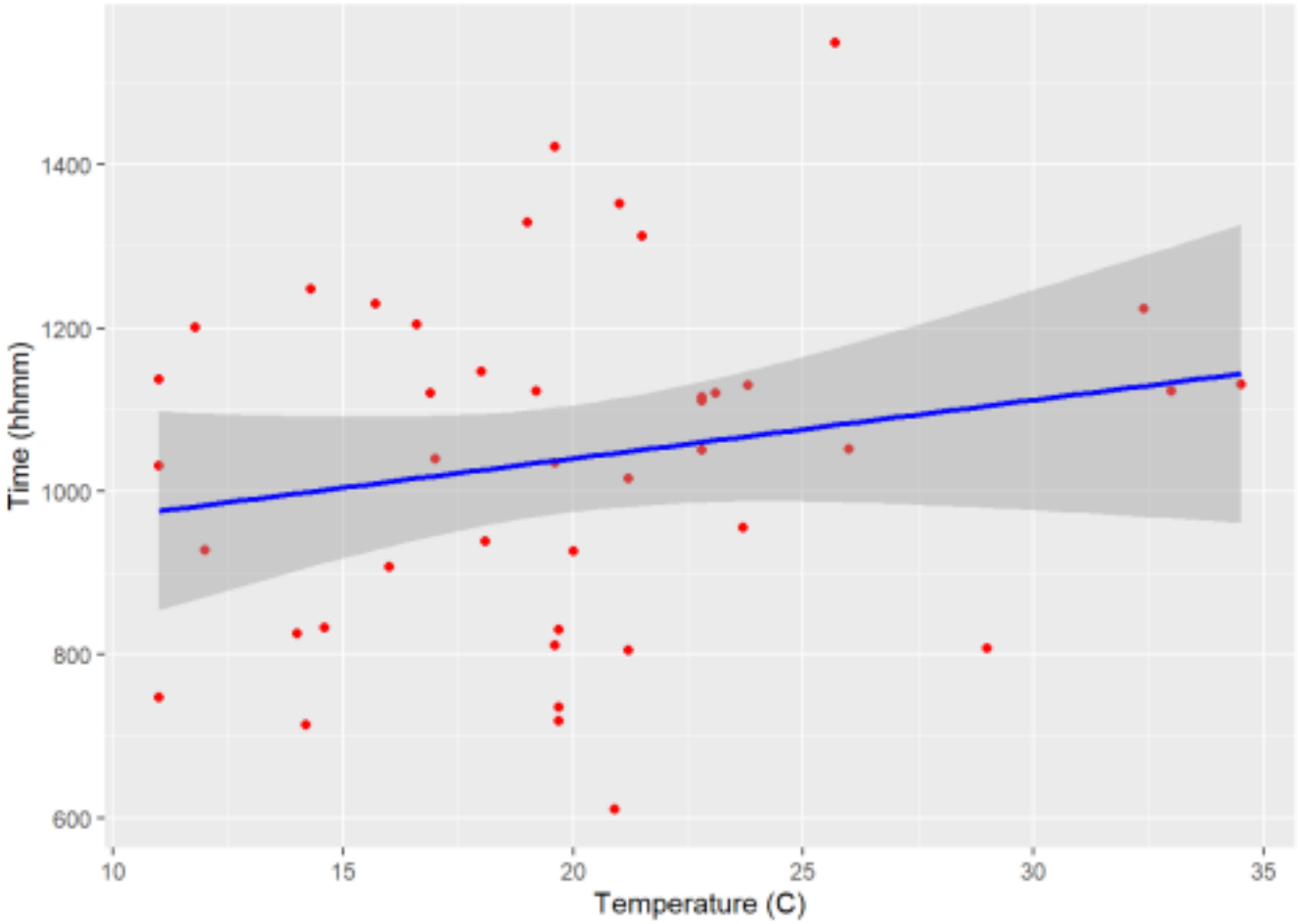

```
#Let's add a regression line
ggplot(data = Koalastand, aes(x = Temperature.max.C., y = Time.hhmm.)) + geom_point() + geom_smo
oth(method=lm, color="blue", se=FALSE) + labs(x="Temperature (C)",y="Time (hhmm)")
```

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



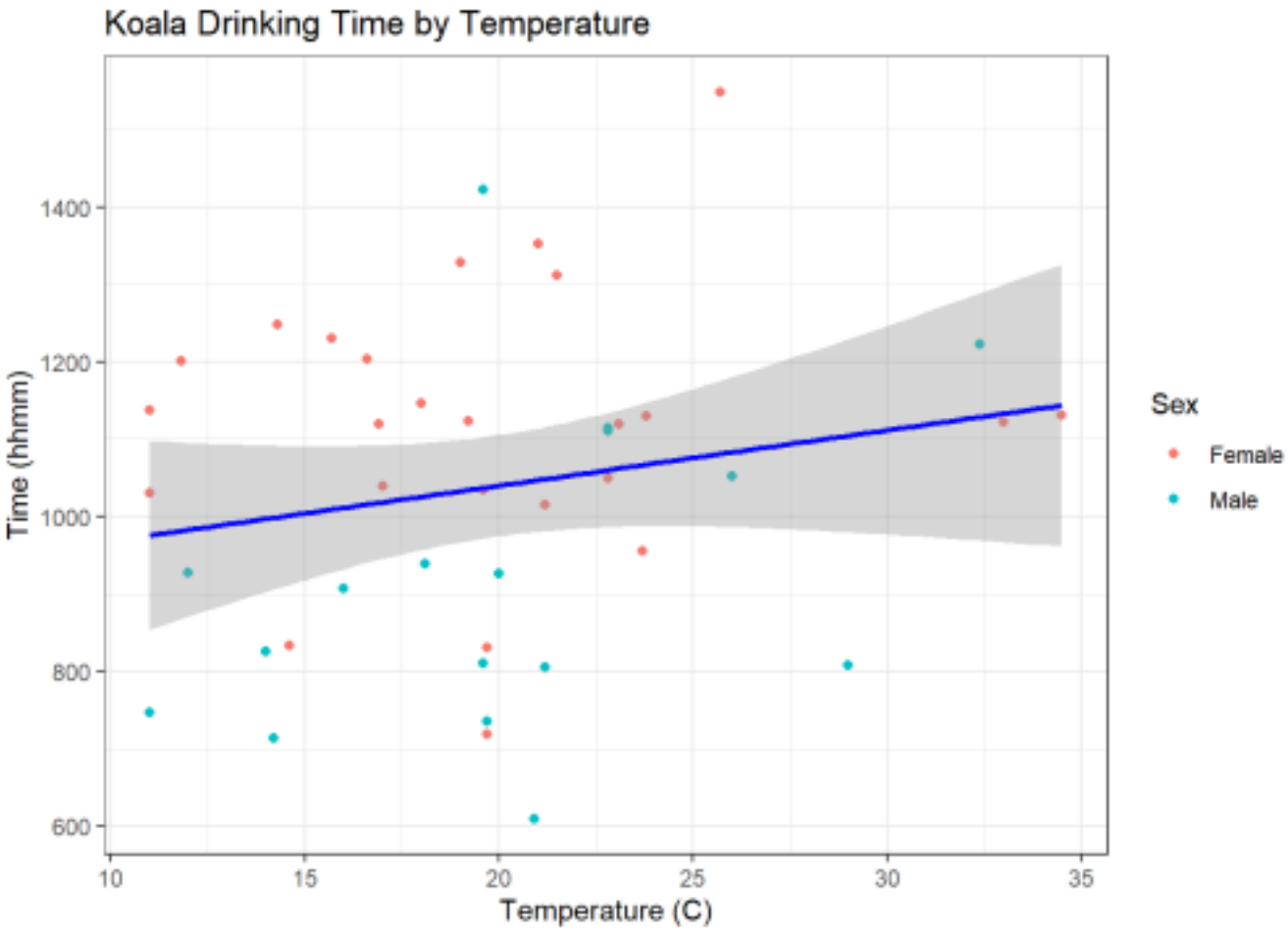

```
#Adding a confidence interval to see how well points fit line
ggplot(data = Koalastand, aes(x = Temperature.max.C., y = Time.hhmm.)) + geom_point(color="red")
+ geom_smooth(method=lm, color="blue") + labs(x="Temperature (C)",y="Time (hhmm)")
```

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

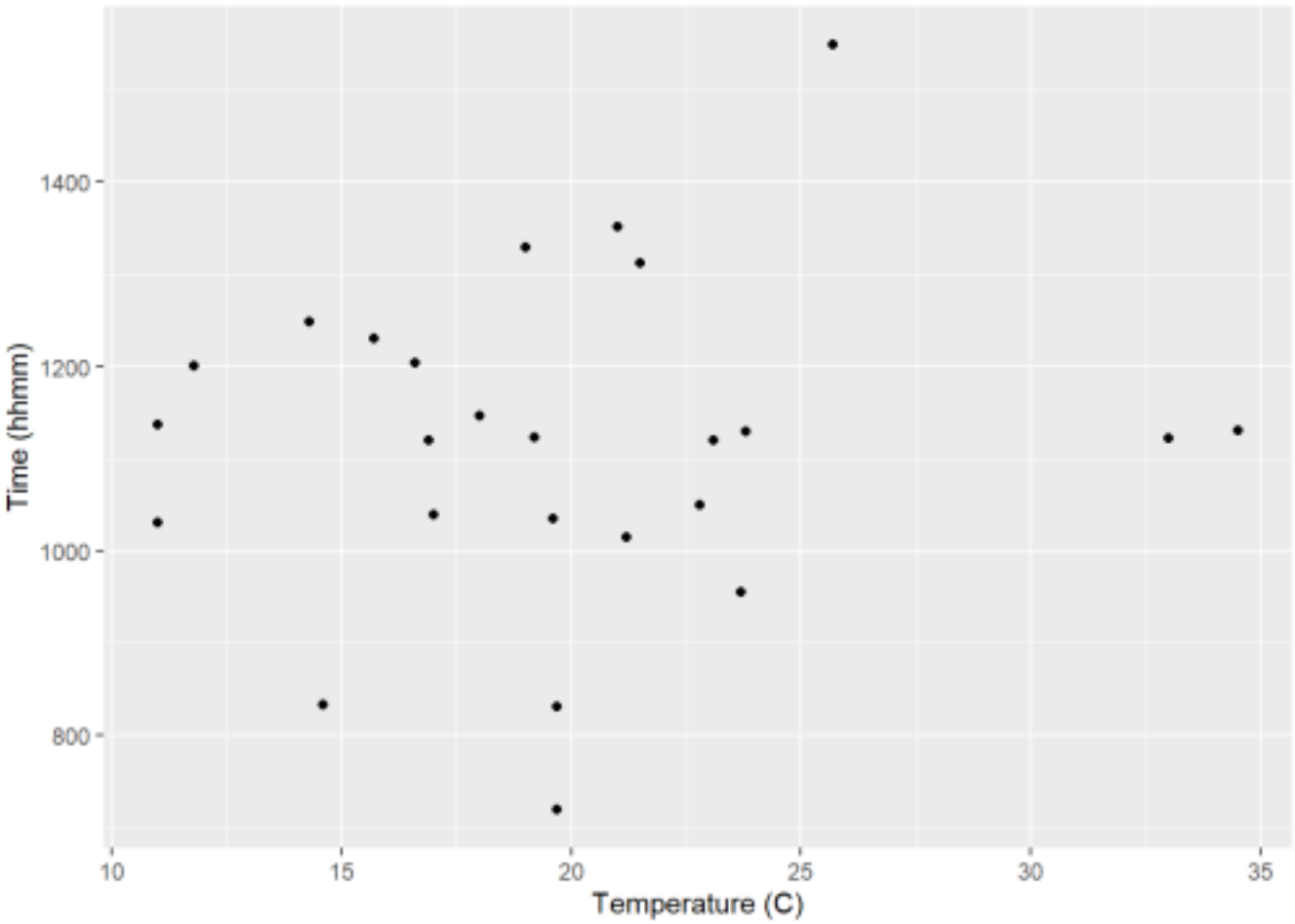



```
#Let's differentiate points by sex
ggplot(data = Koalastand, aes(x = Temperature.max.C., y = Time.hhmm., color=Sex)) + geom_point()
+ geom_smooth(method=lm, color="blue") + labs(x="Temperature (C)",y="Time (hhmm)", color="Sex",
title="Koala Drinking Time by Temperature")+theme_bw()
```

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



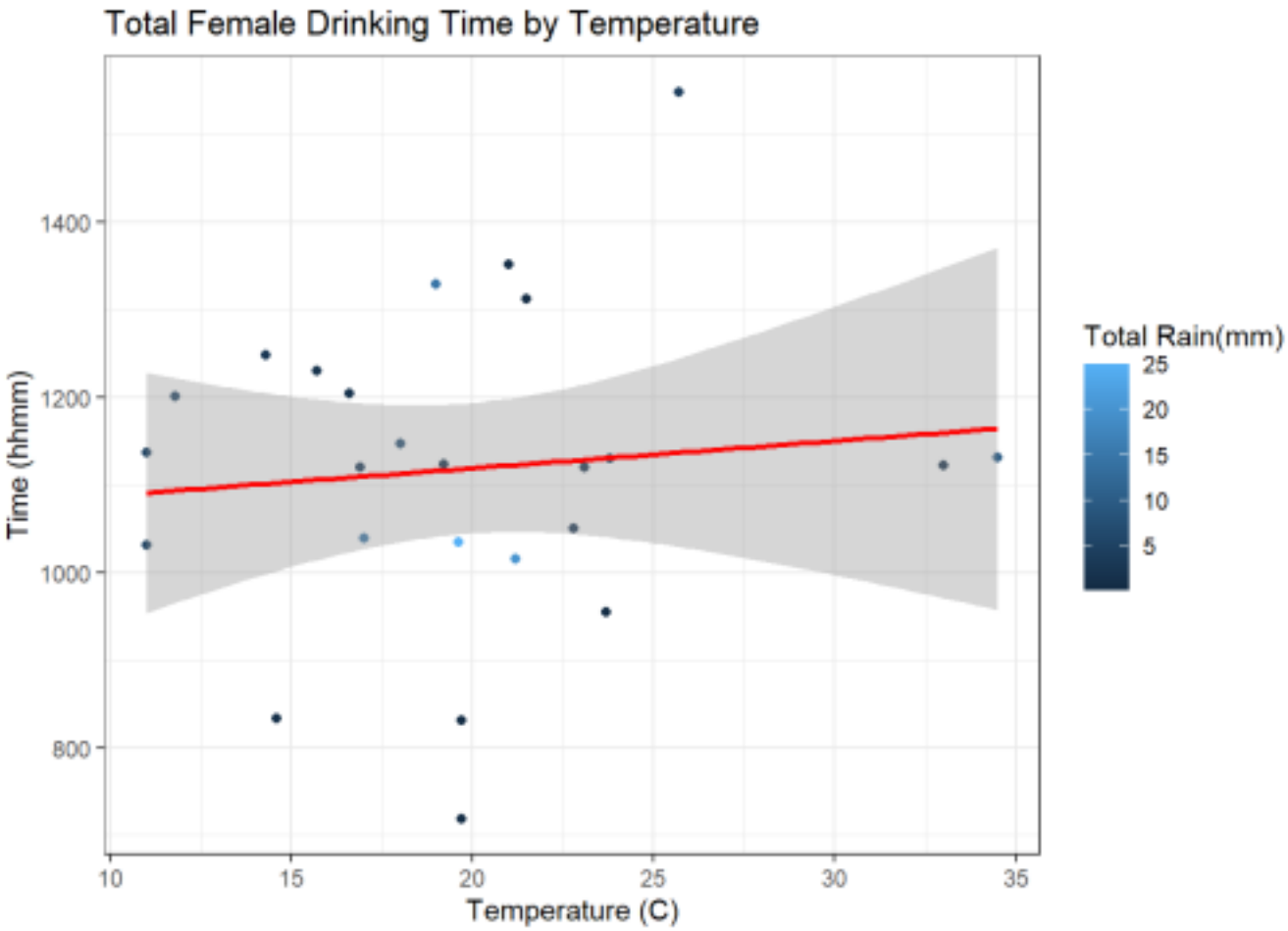

```
# Here we're creating another scatter plot to look at the drinking behaviors of females more clo
sely.
koalafem <- read.csv("Koala female.csv")
ggplot(data = koalafem, aes(x = Temperature.max.C., y = Time.hhmm.)) + geom_point() + labs(x="Te
mperature (C)",y="Time (hhmm)")
```



#Let's add a regression line and differentiate points by amount of rain. Our graph shows that th e less rain, the more time spent drinking in warmer temperatures.

```
ggplot(data = koalafem, aes(x = Temperature.max.C., y = Time.hhmm., color=Rain.mm.)) + geom_poin t()+ geom_smooth(method=lm, color="red") + labs(x="Temperature (C)", color="Total Rain(mm)",y="T ime (hhmm)", title="Total Female Drinking Time by Temperature")+theme_bw()
```

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



#Let's add the r and p values to observe the relationship of the data. P value tells us the probability of the data while the R tells us the trend of the data.

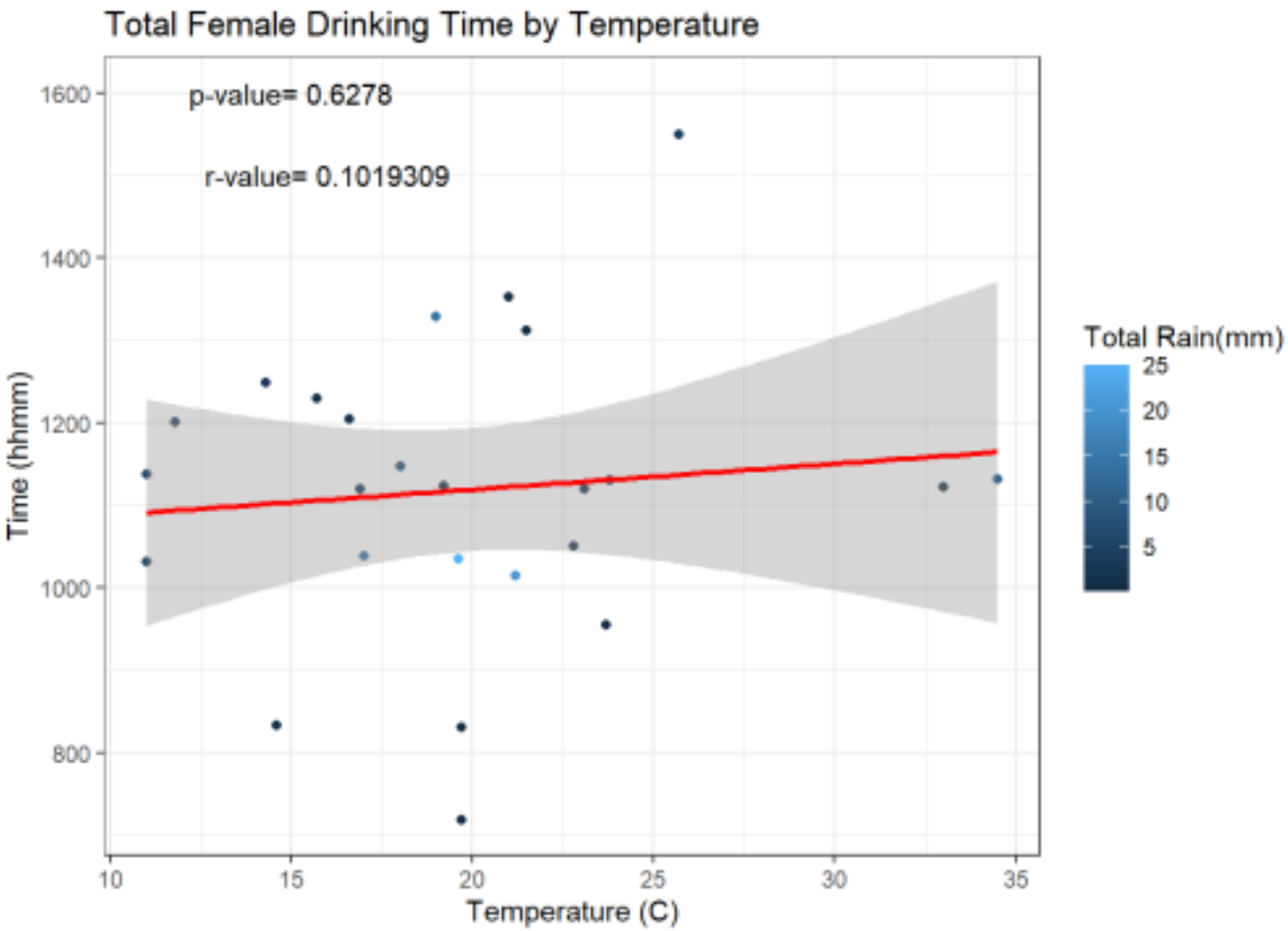
#Finding r and p-values using cor.test

```
cor.test(koalafem$Temperature.max.C.,koalafem$Time.hhmm.)
```

```
##
## Pearson's product-moment correlation
##
## data: koalafem$Temperature.max.C. and koalafem$Time.hhmm.
## t = 0.4914, df = 23, p-value = 0.6278
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.3055045 0.4778172
## sample estimates:
## cor
## 0.1019309
```

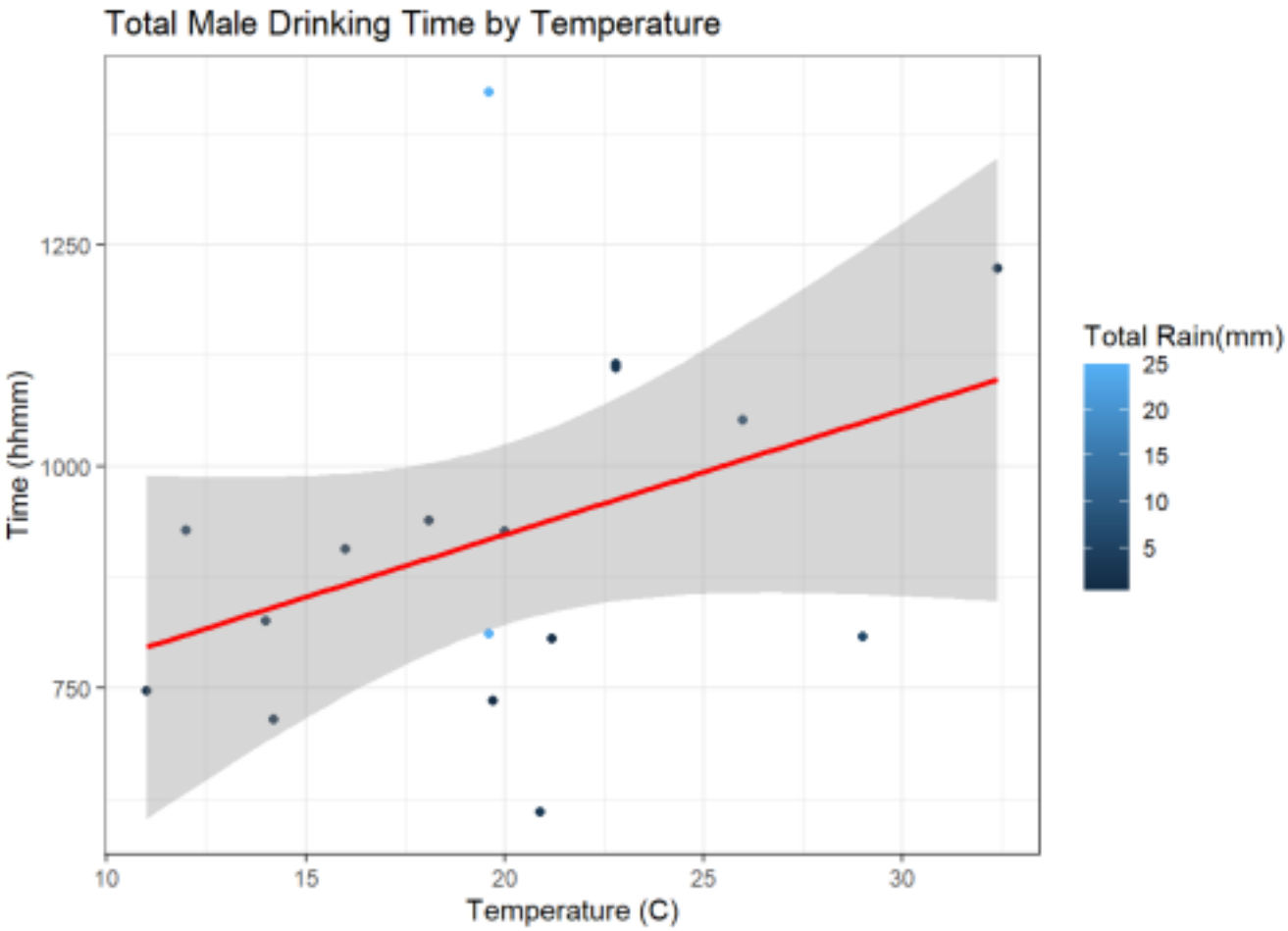
#Adding p and r value to graph using annotate function

```
ggplot(data = koalafem, aes(x = Temperature.max.C., y = Time.hhmm., color=Rain.mm.)) + geom_poin
t()+ geom_smooth(method=lm, color="red") + labs(x="Temperature (C)", color="Total Rain(mm)",y="T
ime (hhmm)", title="Total Female Drinking Time by Temperature")+theme_bw()+ annotate("text",x=15
,y=1600,label="p-value= 0.6278")+annotate("text",x=16,y=1500,label="r-value= 0.1019309")
## `geom_smooth()` using formula 'y ~ x'
```




```
#Let's look at male drinking patterns
maledata <-Koalastand[Koalastand$Sex %in% "Male", ]
ggplot(data = maledata, aes(x = Temperature.max.C., y = Time.hhmm., color=Rain.mm.)) + geom_poin
t()+ geom_smooth(method=lm, color="red") + labs(x="Temperature (C)", y="Time (hhmm)",color="Tota
l Rain(mm)", title="Total Male Drinking Time by Temperature")+theme_bw()
```

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

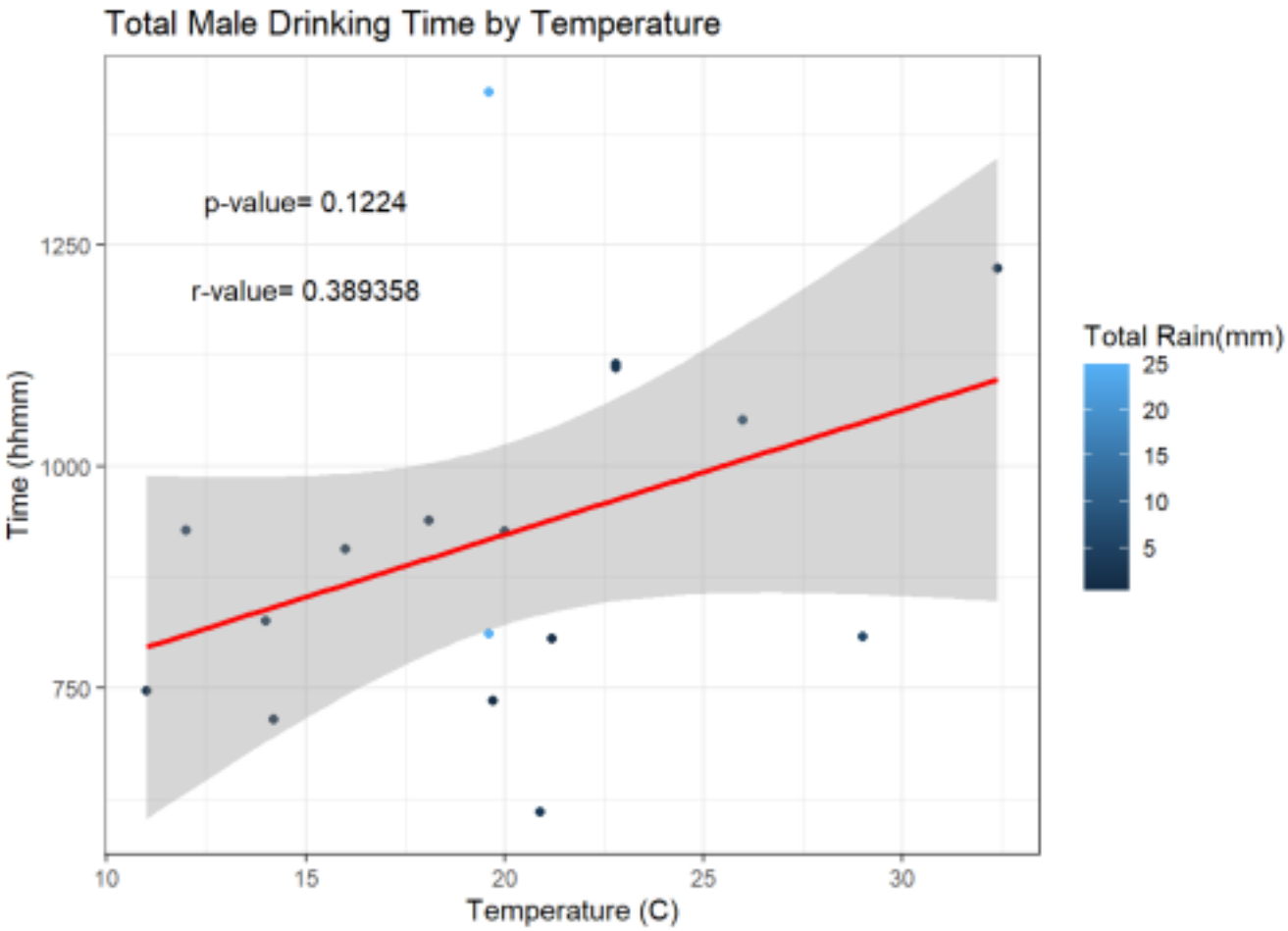



```
# finding r and p values
cor.test(maledata$Temperature.max.C.,maledata$Time.hhmm.)
```

```
##
## Pearson's product-moment correlation
##
## data: maledata$Temperature.max.C. and maledata$Time.hhmm.
## t = 1.6372, df = 15, p-value = 0.1224
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.1123036 0.7328543
## sample estimates:
## cor
## 0.389358
```

```
ggplot(data = maledata, aes(x = Temperature.max.C., y = Time.hhmm., color=Rain.mm.)) + geom_poin
t()+ geom_smooth(method=lm, color="red") + labs(x="Temperature (C)", y="Time (hhmm)",color="Tota
l Rain(mm)", title="Total Male Drinking Time by Temperature")+theme_bw()+ annotate("text",x=15,y
=1300,label="p-value= 0.1224")+annotate("text",x=15,y=1200,label="r-value= 0.389358")
```

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

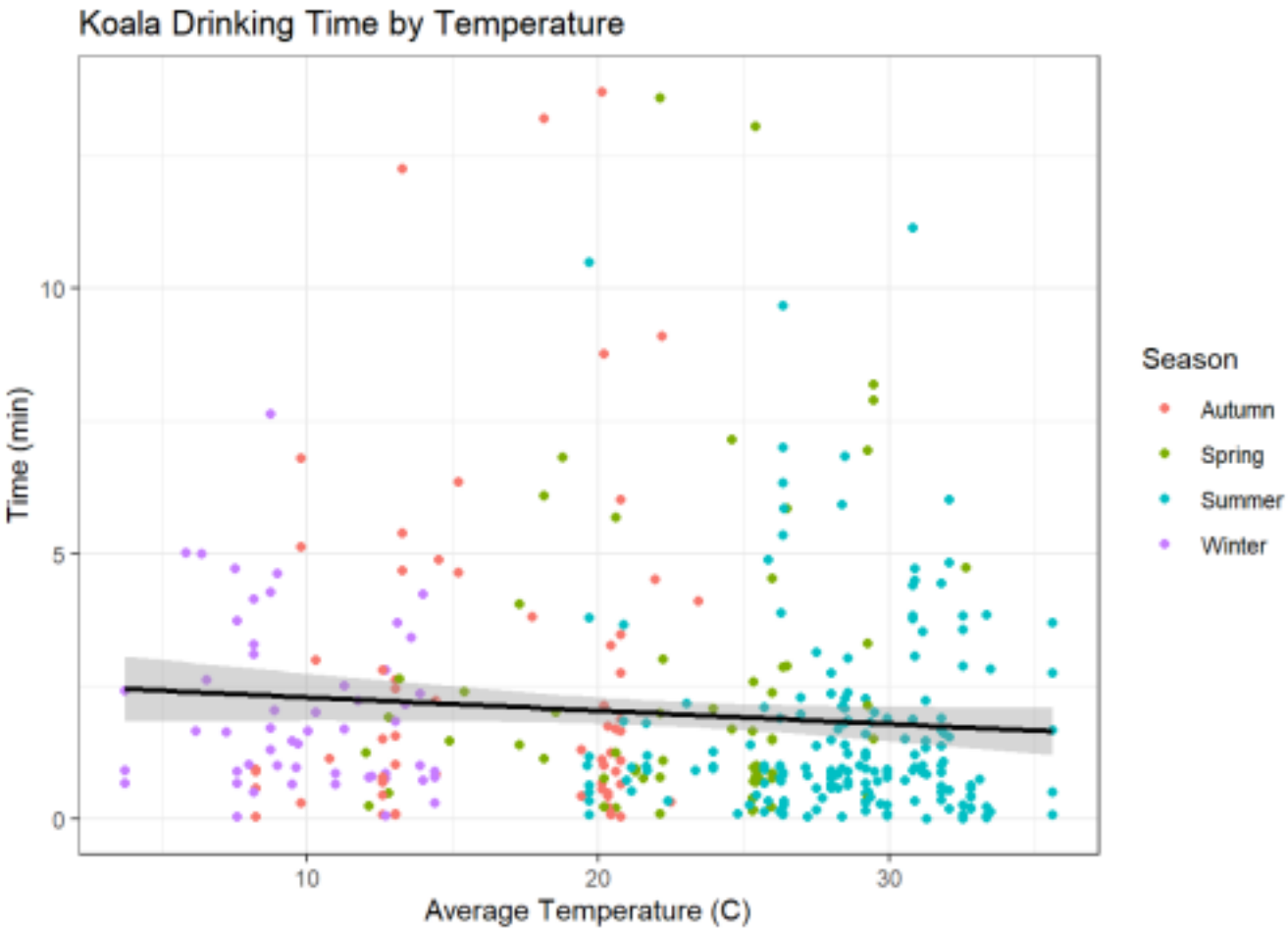



```
# Now we'll observe data from our second resource "Needing a drink: Rainfall and temperature dri
ve the use of free water by a threatened arboreal folivore"
# We'll first read our csv file
Kdrink<-read.csv("Koaladrinks2.csv")

# We want to visualize the drinking time versus temperature as we did with the previous source.
We will be utilizing a scatter plot as above.

ggplot(data = Kdrink, aes(x=Tmean,y=DrinkT, color=Season)) + geom_point()+ geom_smooth(method=l
m, color="black") + labs(x="Average Temperature (C)", y="Time (min)",color="Season", title="Koal
a Drinking Time by Temperature")+theme_bw()


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

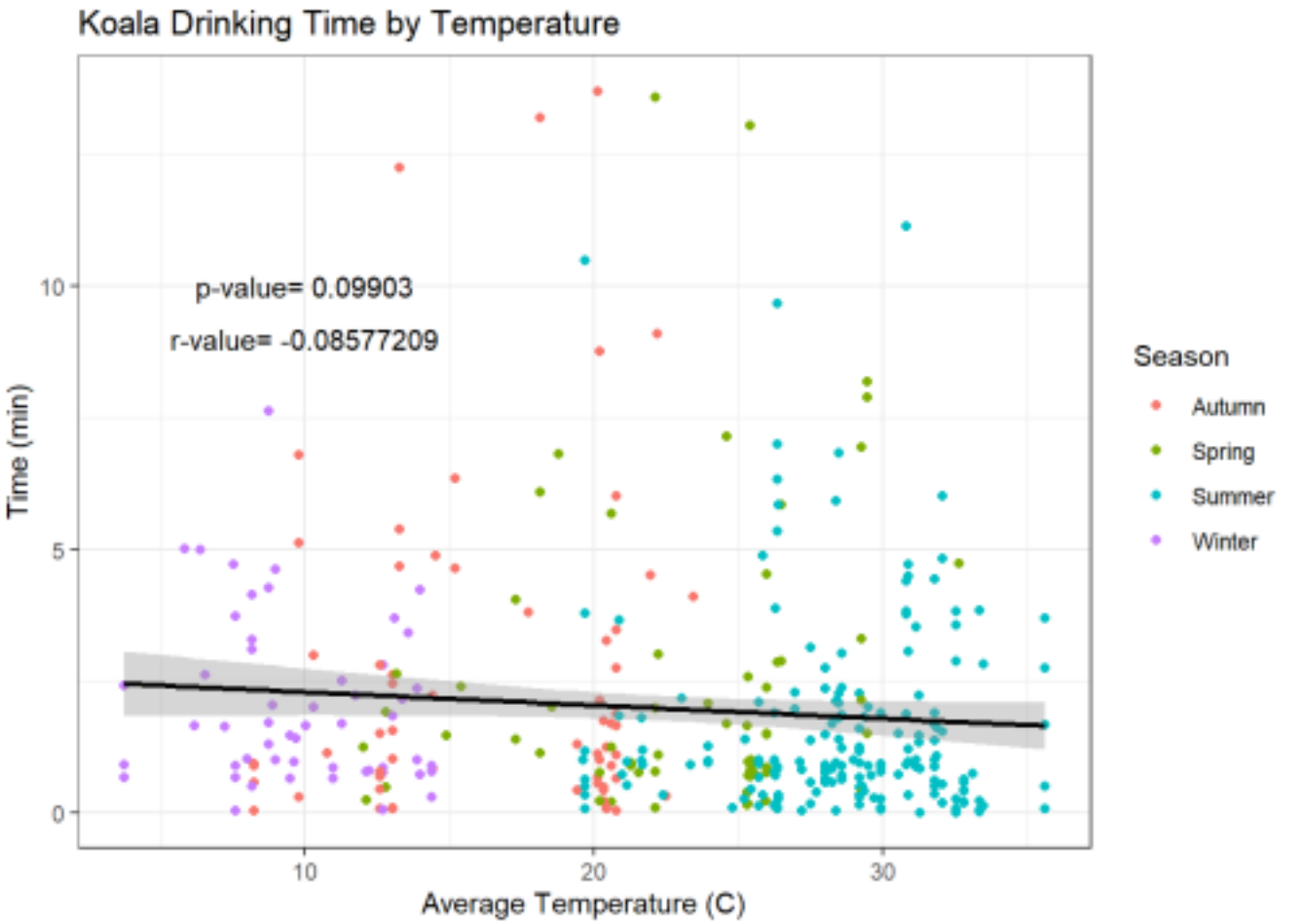



```
# Finding r and p values
cor.test(Kdrink$Tmean,Kdrink$DrinkT)
```

```
##
## Pearson's product-moment correlation
##
## data: Kdrink$Tmean and Kdrink$DrinkT
## t = -1.6537, df = 369, p-value = 0.09903
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.18596422 0.01618541
## sample estimates:
## cor
## -0.08577209
```

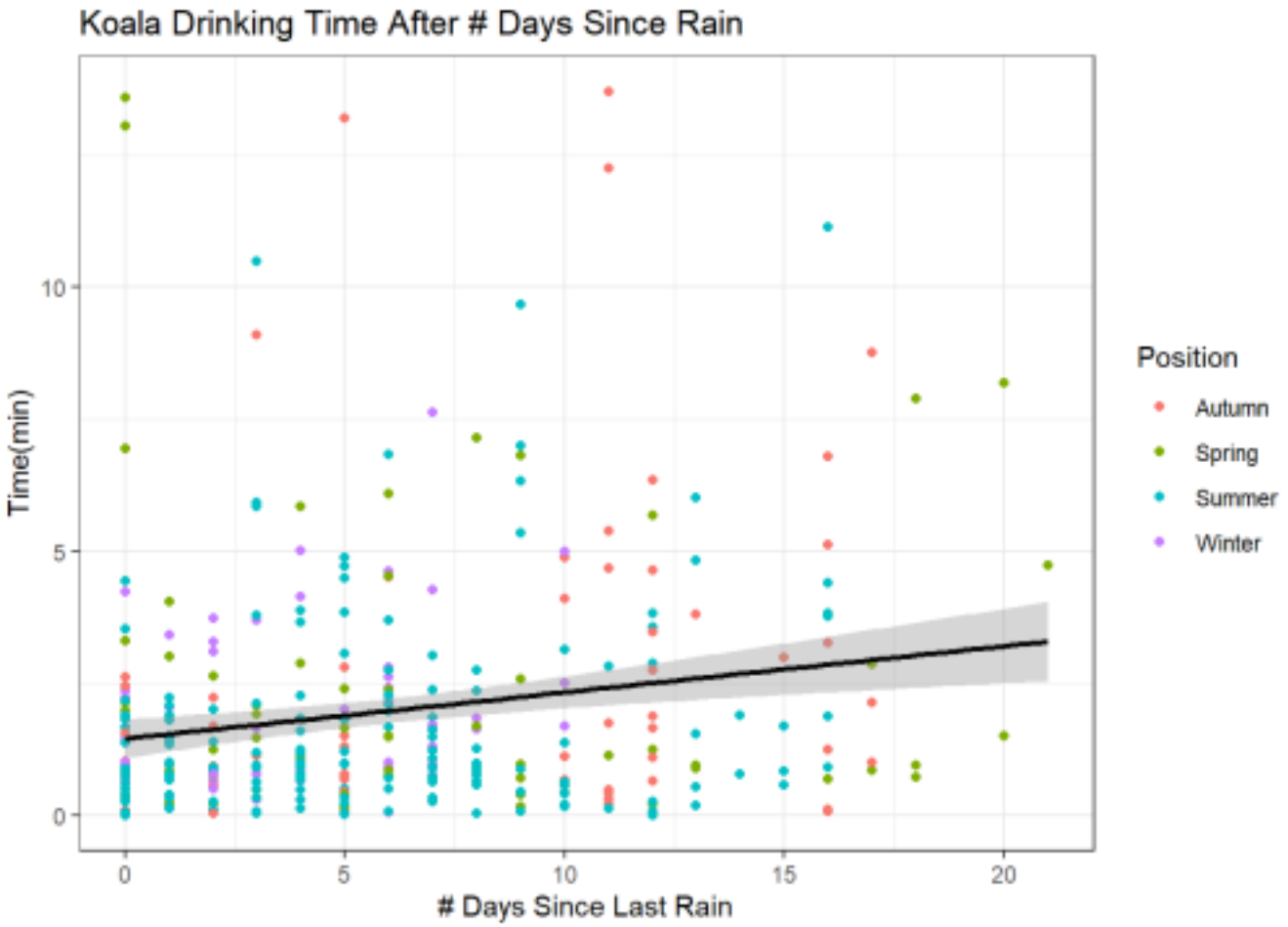
```
ggplot(data = Kdrink, aes(x=Tmean,y=DrinkT, color=Season)) + geom_point()+ geom_smooth(method=l
m, color="black") + labs(x="Average Temperature (C)", y="Time (min)",color="Season", title="Koal
a Drinking Time by Temperature")+theme_bw()+ annotate("text",x=10,y=10,label="p-value= 0.09903")
+annotate("text",x=10,y=9,label="r-value= -0.08577209")
```

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




```
#Let's see how the days since the last rain affects how long Koala's drink for ggplot(data = Kdrink, aes(x=DRain,y=DrinkT, color=Season)) + geom_point()+ geom_smooth(method=l m, color="black")+labs(x="# Days Since Last Rain", y="Time(min)",color="Position", title="Koala Drinking Time After # Days Since Rain")+theme_bw()
```

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

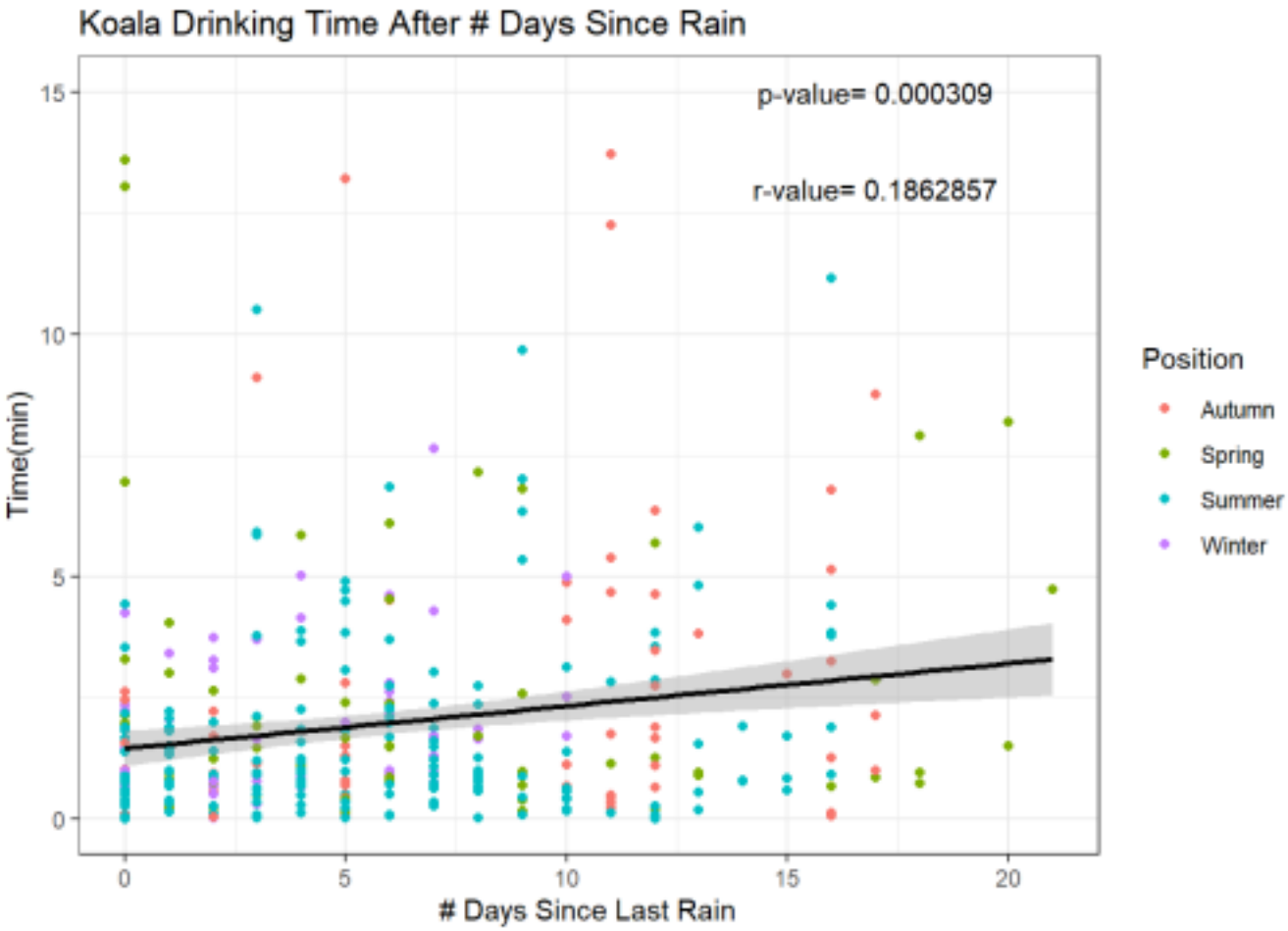



```
cor.test(Kdrink$DRain,Kdrink$DrinkT)
```

```
##
## Pearson's product-moment correlation
##
## data: Kdrink$DRain and Kdrink$DrinkT
## t = 3.6422, df = 369, p-value = 0.000309
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.08610263 0.28273915
## sample estimates:
## cor
## 0.1862857
```

```
#Adding r and p values to graph
ggplot(data = Kdrink, aes(x=DRain,y=DrinkT, color=Season)) + geom_point()+ geom_smooth(method=l
m, color="black")+labs(x="# Days Since Last Rain", y="Time(min)",color="Position", title="Koala
Drinking Time After # Days Since Rain")+theme_bw()+ annotate("text",x=17,y=15,label="p-value=
0.000309")+annotate("text",x=17,y=13,label="r-value= 0.1862857")
```

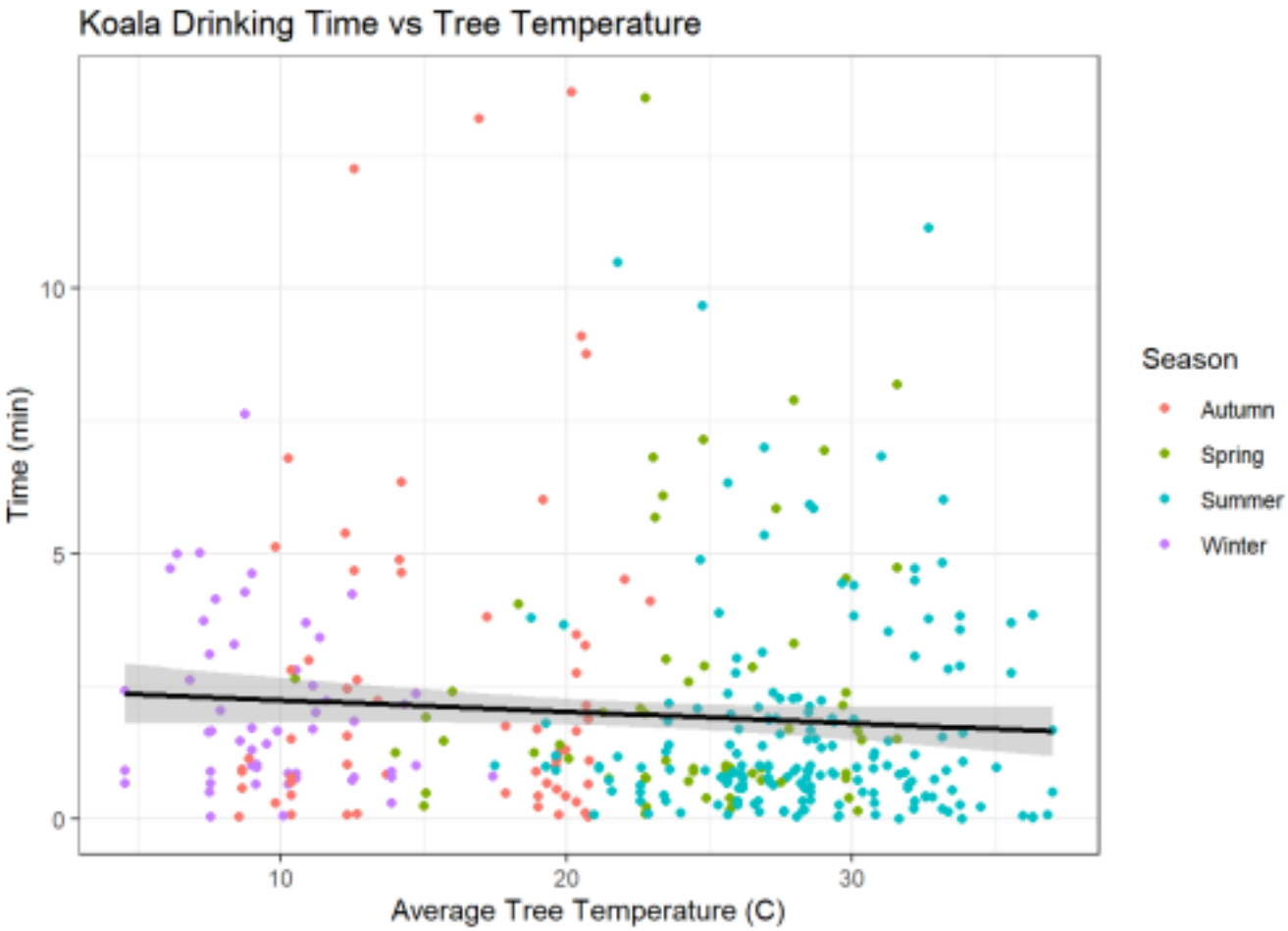
```
## `geom_smooth()` using formula 'y ~ x'
```




```
# How does tree temperature affect drinking behavior
Kdrink2 <-read.csv("Koaladrinks3.csv")

ggplot(data = Kdrink2, aes(x=TTree,y=DrinkT, color=Season)) + geom_point()+geom_smooth(method=l
m, color="black")+labs(x="Average Tree Temperature (C)", y="Time (min)",color="Season", title="K
oala Drinking Time vs Tree Temperature")+theme_bw()


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

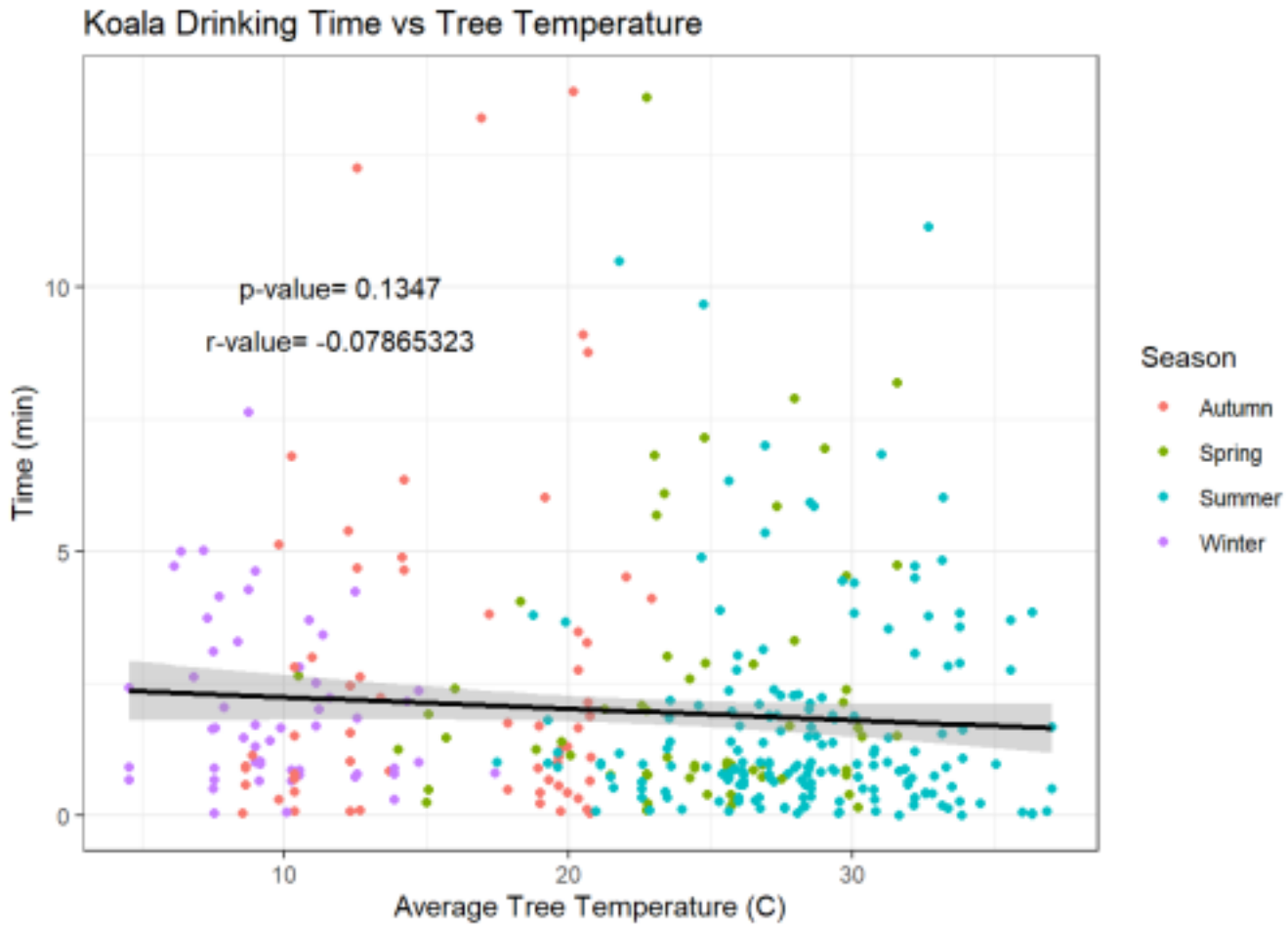



```
#Finding r and p values
cor.test(Kdrink2$TTree,Kdrink2$DrinkT)
```

```
##
## Pearson's product-moment correlation
##
## data: Kdrink2$TTree and Kdrink2$DrinkT
## t = -1.4991, df = 361, p-value = 0.1347
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.18012821 0.02447826
## sample estimates:
## cor
## -0.07865323
```

```
ggplot(data = Kdrink2, aes(x=TTree,y=DrinkT, color=Season)) + geom_point()+geom_smooth(method=l
m, color="black")+labs(x="Average Tree Temperature (C)", y="Time (min)",color="Season", title="K
oala Drinking Time vs Tree Temperature")+theme_bw()+ annotate("text",x=12,y=10,label="p-value=
0.1347")+annotate("text",x=12,y=9,label="r-value= -0.07865323")
```

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


Koala Drinking Time vs Tree Temperature
p-value= 0.1347
r-value= -0.07865323
Time (min)
0
5
10
10
20
30
Average Tree Temperature (C)
Season
Autumn
Spring
Summer
Winter