Stat. 651 Quiz 1

Lydia Gibson

Instructions: This is an open book, open notes, and open Google/internet test. You may use R on your own computer. You may use a calculator.

Type your answers to the questions into an Quarto/R Notebook. Answer the questions in order. Answer each the question that is asked above your R code chunks. You must write a sentence containing your answer.

Your files should have a name in the usual form for the class **lastname_firstname_Stat651_Quiz1.qmd** and **lastname_firstname_Stat651_Quiz1.docx** or **lastname_firstname_Stat651_Quiz1.pdf**. Submit **both** files in Canvas.

Academic Honesty: As a student at CSU East Bay you are held to the standards stated in the Academic Dishonesty Policy. Copying another student's work or allowing another student to copy your work is academically dishonest. I expect you to be academically honest while taking the test.

These question is related to the homework from Chapter 3.

For the *RailTrail* dataset from the *mosaicData* R package answer the following questions. The data is from Northampton, MA.

```
suppressPackageStartupMessages(library(tidyverse))
suppressPackageStartupMessages(library(mosaic))
```

```
head(RailTrail)
```

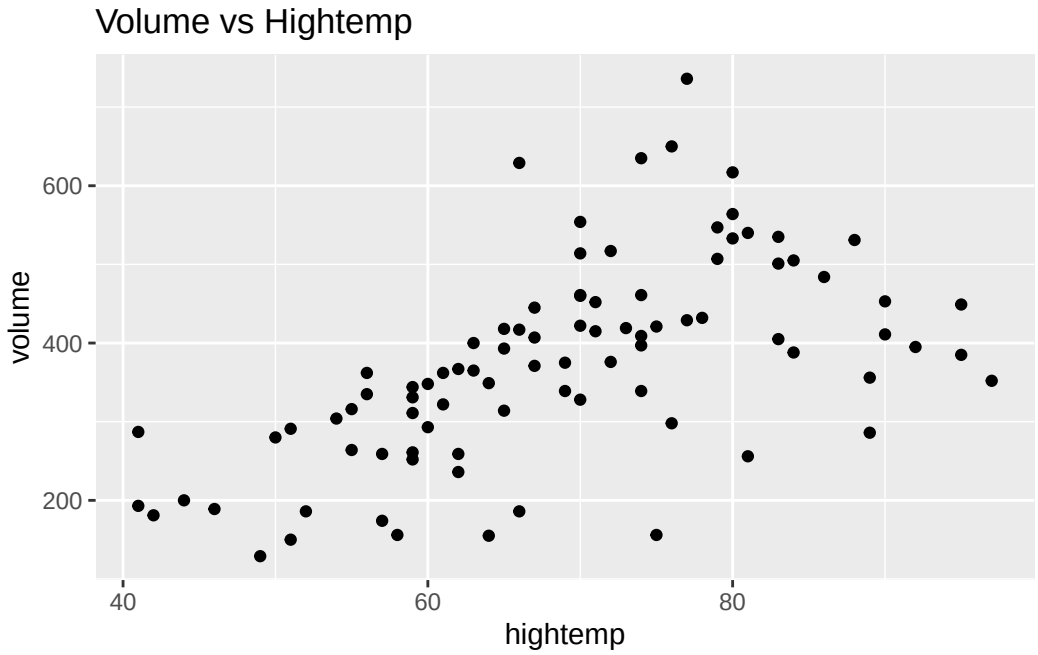
```
  hightemp lowtemp avgtemp spring summer fall cloudcover precip volume weekday
1       83      50    66.5      0      1    0        7.6   0.00    501    TRUE
2       73      49    61.0      0      1    0        6.3   0.29    419    TRUE
3       74      52    63.0      1      0    0        7.5   0.32    397    TRUE
4       95      61    78.0      0      1    0        2.6   0.00    385   FALSE
5       44      52    48.0      1      0    0       10.0   0.14    200    TRUE
6       69      54    61.5      1      0    0        6.6   0.02    375    TRUE
```

```
  dayType
1 weekday
2 weekday
3 weekday
4 weekend
5 weekday
6 weekday
```

1. Create a scatterplot of the *volume* (number of crossing per day) against the *high temperature* that day.

Answer:

```
RailTrail %>%
  select(volume,hightemp) %>%
  ggplot(aes(y=volume, x=hightemp)) +
  geom_point() +
  ggtitle("Volume vs Hightemp")
```

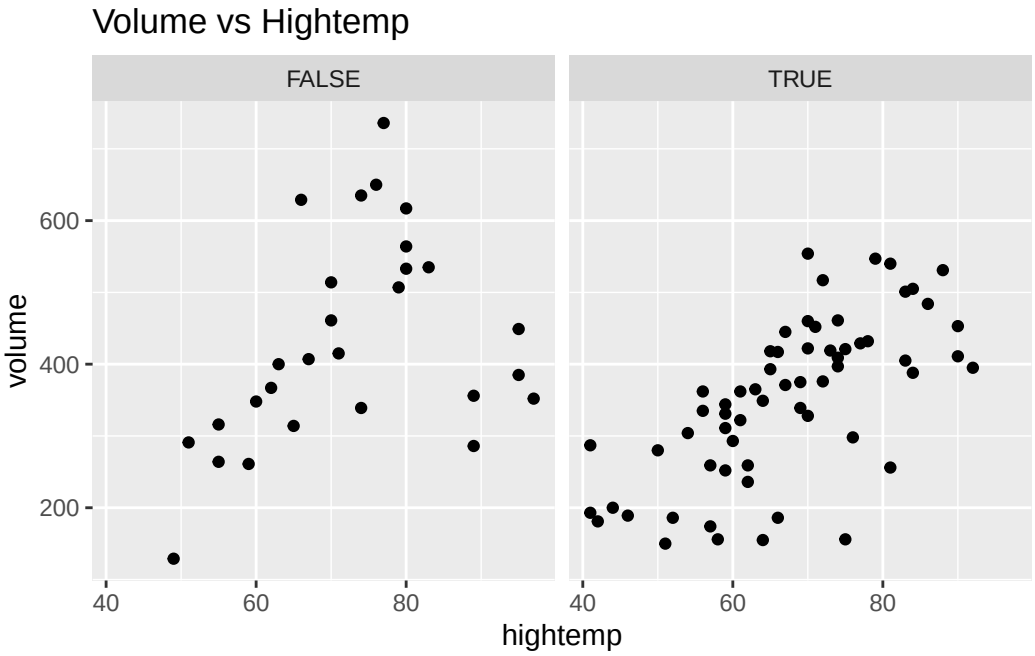


2

. Separate your previous plot into facets by *weekday*.

Answer:

```
RailTrail %>%
  select(volume,hightemp, weekday) %>%
  ggplot(aes(y=volume, x=hightemp)) +
  facet_wrap( ~weekday ) +
  geom_point() +
  ggtitle("Volume vs Hightemp")
```



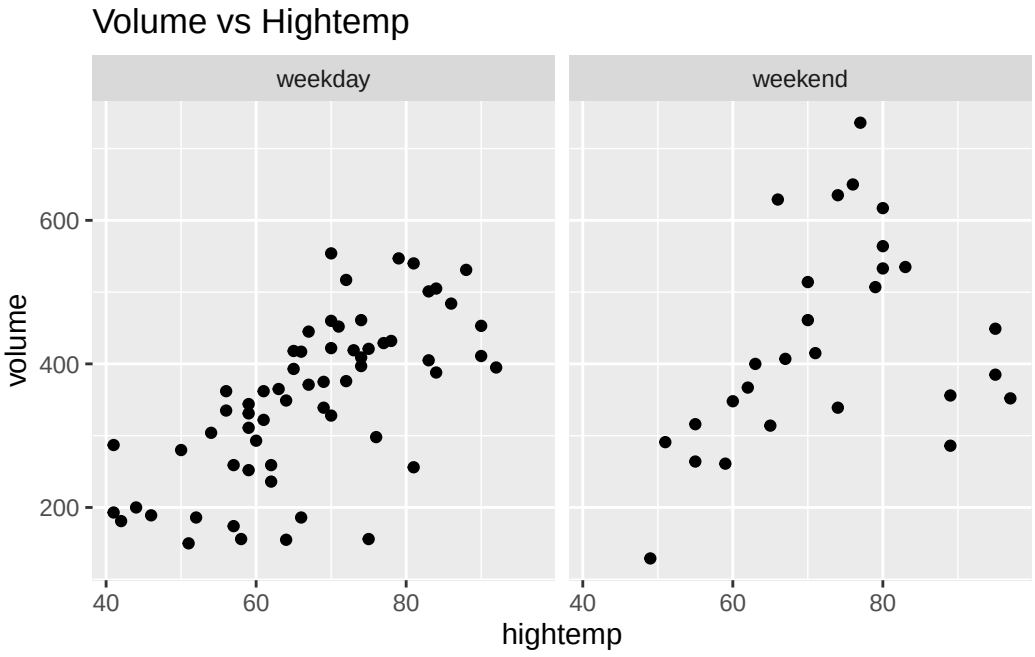
3

. Examine the plot you have created in the previous question, is there anything about the plot that is unclear? Suggest a way to fix the issue you have described. Make a improved plot.

Faceting by `weekday`, the resulting plots are labeled with the logical responses rather than character responses expressing which facet is a weekday and which is a weekend. By faceting by `dayType`, the new facet labels clearly say which plot is the weekdays and which is the weekends.

Answer:

```
RailTrail %>%
  select(volume,hightemp, dayType) %>%
  ggplot(aes(y=volume, x=hightemp)) +
  facet_wrap( ~dayType ) +
  geom_point() +
  ggtitle("Volume vs Hightemp")
```

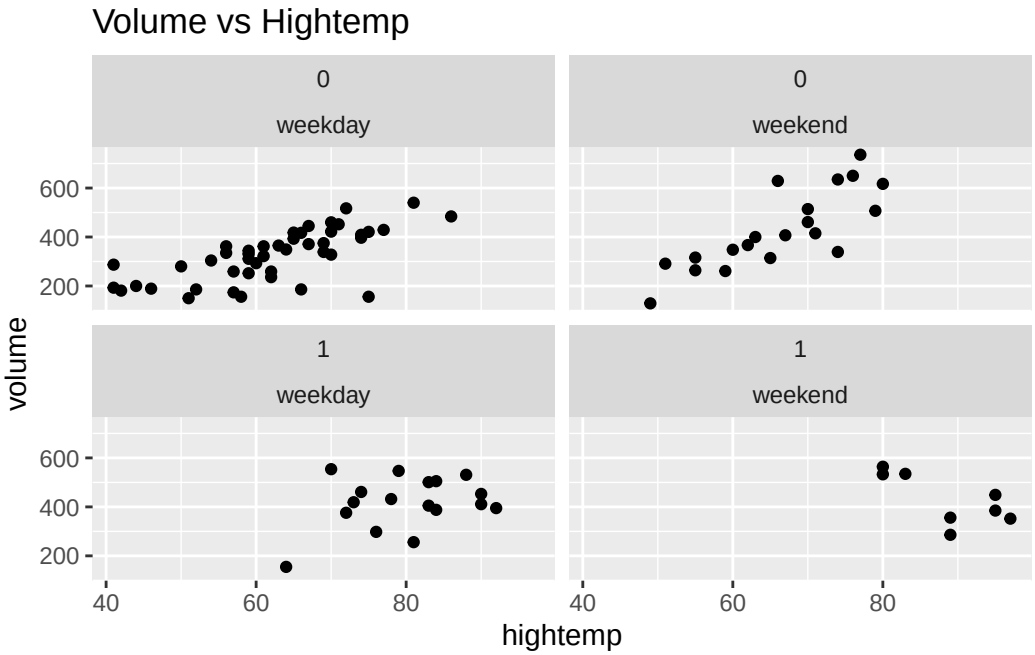


4

. Separate your plot into facets by summer and weekday. The summer variable used for the rows and the weekday variable for the columns.

Answer:

```
p1 <- RailTrail %>%
  select(volume,hightemp, dayType, summer) %>%
  ggplot(aes(y=volume, x=hightemp)) +
  facet_wrap( summer ~ dayType ) +
  geom_point() +
  ggtitle("Volume vs Hightemp")

p1
```



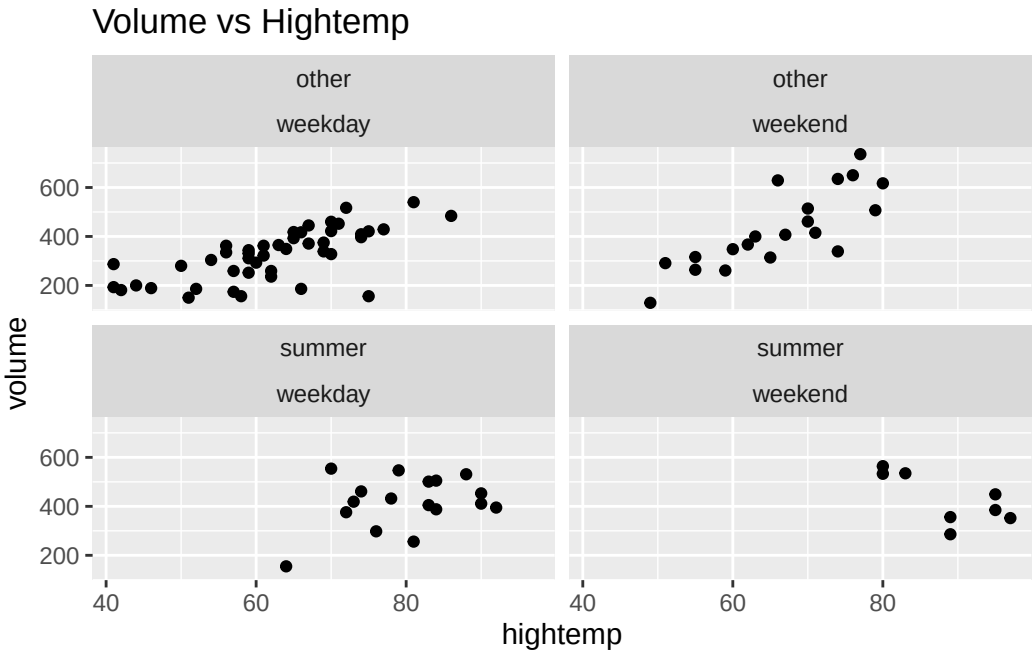
5

. Examine the plot you have created in the previous question, is there anything about the plot that is unclear? Suggest a way to fix the issue you have described. Make a improved plot. Hint: Change the *summer* variable to a factor and use %+% to replace the data in the original plot p.

Faceting by `summer`, the resulting plots are labeled with the logical responses rather than character responses expressing which facet is summer and which is not. By setting summer as a factor, the new facet labels clearly say which plot is the summer and which is other.

Answer:

```
RailTrail2 <- RailTrail %>%
  mutate ( summer = factor( summer, labels = c("other", "summer")))

#RailTrail2

p1 %+% RailTrail2
```



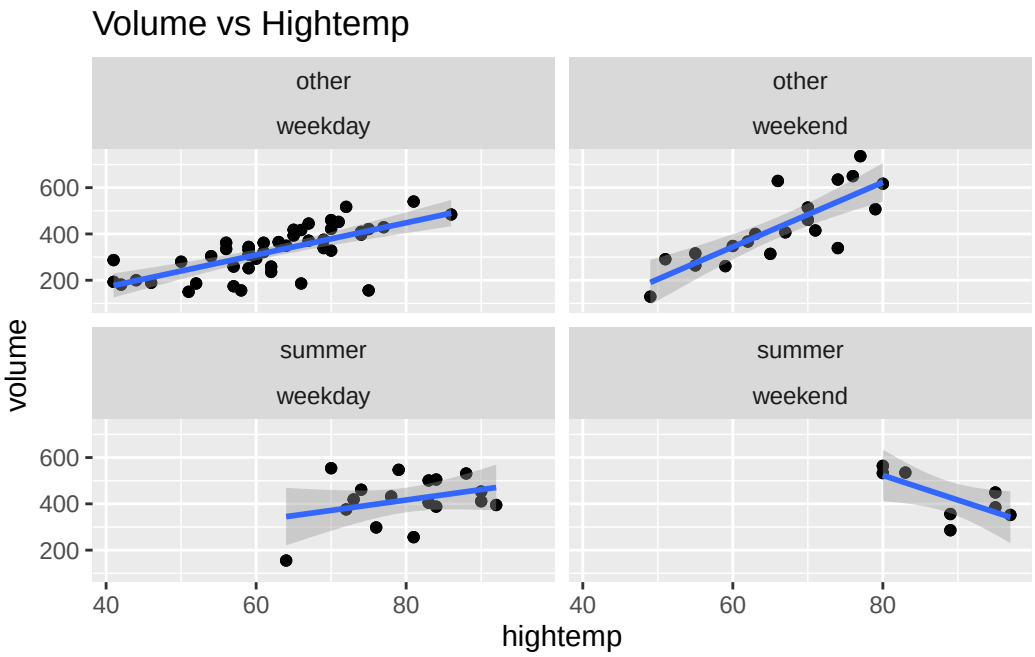
6

. Add regression lines to the four facets. When does the relationship between *volume* and *hightemp* change?

The relationship between `volume` and `hightemp` changes above 80 degrees Fahrenheit.

Answer:

```
p2 <-RailTrail2 %>% select(hightemp, volume, summer, dayType) %>%
  ggplot(aes(y=volume, x=hightemp)) +
  geom_point() +
  facet_wrap( summer ~ dayType ) +
  geom_point() +
  geom_smooth(method = "lm") +
  ggtitle("Volume vs Hightemp")


p2
```

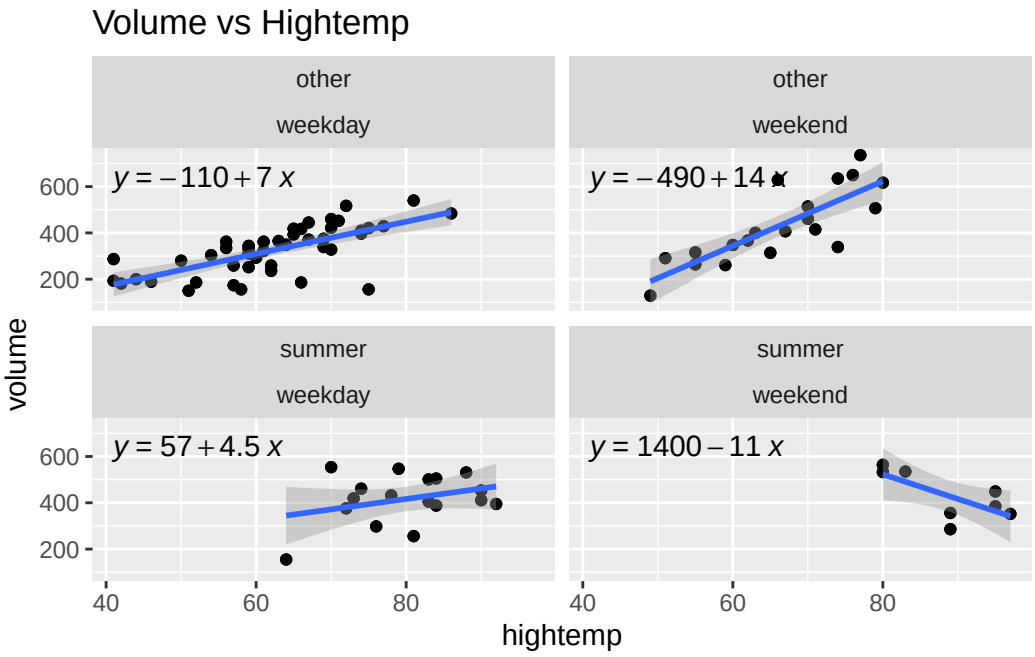


7

. (Extra Credit) Compute the slope of each regresion line. Hint: Use the map function from the **purrr** R package.

Answer:

```
suppressPackageStartupMessages(library(ggpubr))
p2 + stat_regline_equation()
```

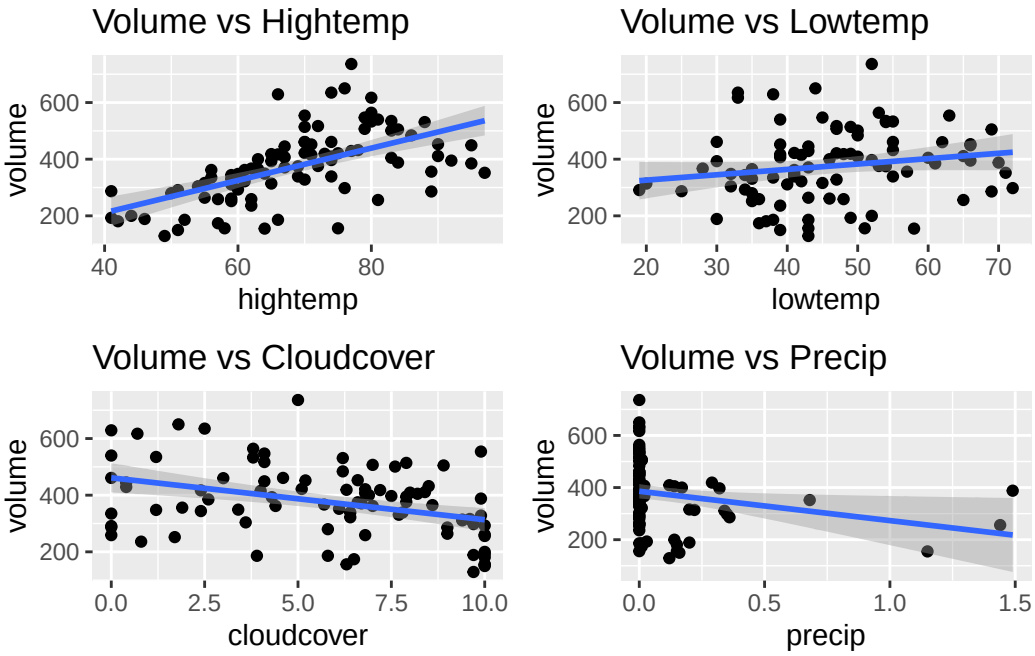


8

. Plot *volume* versus each of the following variables: *hightemp*, *lowtemp*, *cloudcover*, *precip*. Add regression lines. Put the 4 plots into one ggplot using a function from the **cowplot** R package.

Answer:

```
suppressPackageStartupMessages(library(cowplot))

p3 <- RailTrail %>%
  select(volume, hightemp) %>%
  ggplot(aes(y = volume, x = hightemp)) +
  geom_point() +
  geom_smooth(method = "lm") +
  ggtitle("Volume vs Hightemp")

p4 <- RailTrail %>%
  select(volume, lowtemp) %>%
  ggplot(aes(y = volume, x = lowtemp)) +
  geom_point() +
  geom_smooth(method = "lm") +
  ggtitle("Volume vs Lowtemp")

p5 <- RailTrail %>%
  select(volume, cloudcover) %>%
  ggplot(aes(y = volume, x = cloudcover)) +
  geom_point() +
  geom_smooth(method = "lm") +
  ggtitle("Volume vs Cloudcover")

p6 <- RailTrail %>%
  select(volume , precip) %>%
  ggplot(aes(y = volume, x = precip)) +
  geom_point() +
  geom_smooth(method = "lm") +
  ggtitle("Volume vs Precip")

plot_grid(p3, p4, p5, p6)
```



9. Export the RailTrail data to a .csv file and load it into the Tableau.

```
write_csv(RailTrail, "RailTrail.csv")
```

10. Make the same plots in Tableau. Arrange them into a Story. Export the Story as a .pptx file.

Answer:

```
#See pptx file
```