Assignment 1

Zening Ye

9/22/2021

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

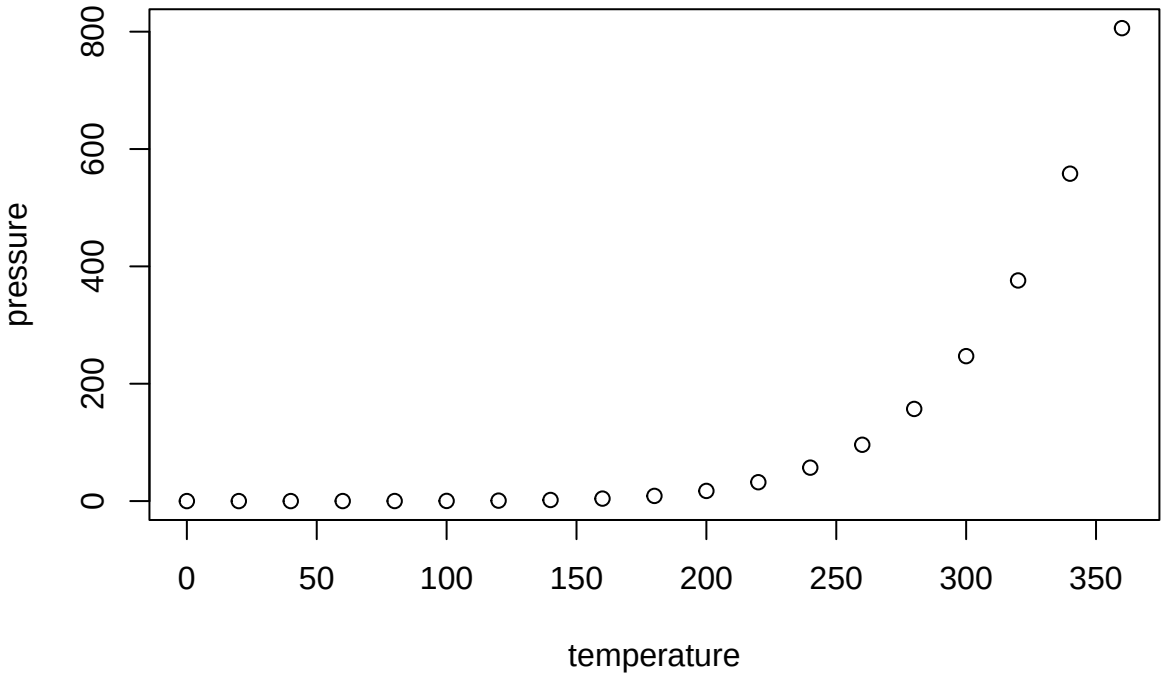
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

```
library(tidyverse) #Pull up the function we will use
```

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

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1
```

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

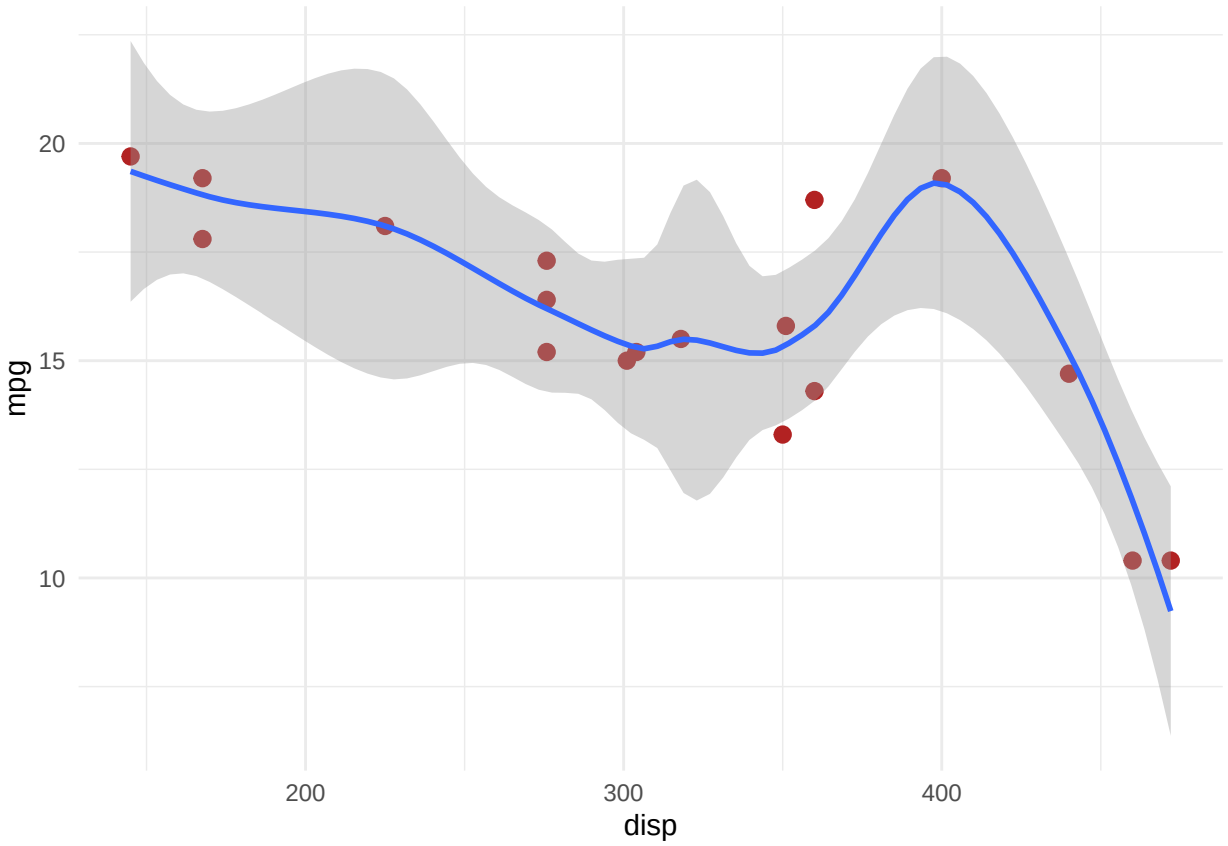
```
data(mtcars) # Select the data
mtcars_mpg2 <- mtcars[mtcars$mpg <20,] # Select the data for mpg < 20

mtcars_mpg2 <- mtcars_mpg2[, c(1,2,3,4,10)] # Select column from mtcars_mpg2 in 1, 2, 3, 4, 10

source(file = "hand_functions.R", echo = TRUE) # Use the function create by R. Scrip
```

```
##
## > sum_special <- function(df_x) {
## +     try(if (!is.data.frame(df_x))
## +         stop("Input data must be a data frame."))
## +     sp_means <- apply(df_ .... [TRUNCATED]
```

```
sp_out <- sum_special(mtcars_mpg2) # Name the data in different way
```

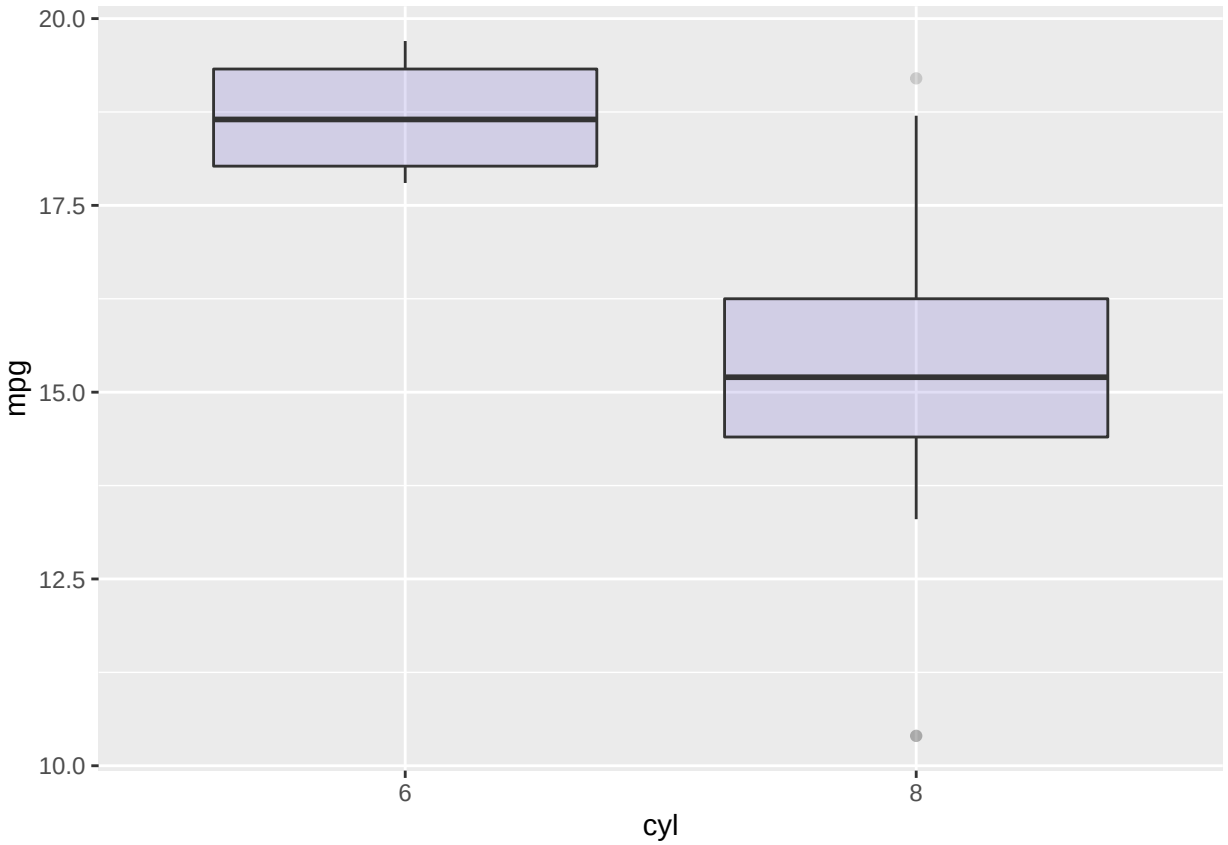
```
#### Plot the data with x = disp, y = mpg by using geom_point() and create a smooth line for the entire
ggplot(mtcars_mpg2) +
  aes(x = disp, y = mpg) +
  geom_point(shape = "bullet", size = 4L, colour = "#B22222") +
  geom_smooth(span = 0.5) +
  theme_minimal()
```

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



note that this boxplot cannot be made with esquisse() unless the data is adjusted. What adjustment is needed?

```
ggplot(mtcars_mpg2, aes(x=as.factor(cyl), y=mpg)) +
  geom_boxplot(fill="slateblue", alpha=0.2) +
  xlab("cyl")
```



By factoring cyl we change the data from a vector to a factor, then we can use factor data to create a boxlpot.

What I learned from this class is to use esquisser(data = mtcars_mpg2, viewer = "browser") to create a plot and get in touch with tidyverse. In addition, to understand how to change shape in ggplot function.