Exercise RMarkdown and reproductibility course

The EDB RMarkdown superstar learning team

14/12/2022

Contents

1 Introduction

This exercise is the best way to practice `RMarkdown` (Salinas 2020) now that you have had an overview of how it works. The principle is that you want to produce a copy of this `pdf` file. We provide most of the `R code`, because the scope is to work on the `RMarkdown` (Xie, Allaire, and Grolemund 2018) integration, not on the `R code`.

2 Starters

You want to create a new `Rmarkdown` file and save it in the `Exercise` folder.

3 Load packages

What packages will I need? Well I can fill that in later anyway.

BTW: To add a new code chunk the keyboard shortcut is `Ctrl` + `Alt` + `i`. Or use the insert button on the top left side of this window.

4 Research compendium

Let's organize our project. We create a folder called `data` and list the files in the current working folder.

```
system("mkdir -p ../data")
list.files(path = ".")
```

```
## [1] "Exercise_files" "Exercise.html"  "Exercise.md"    "Exercise.pdf"
## [5] "Exercise.Rmd"
```

```
list.dirs(path = "..", full.names = TRUE, recursive = TRUE)
```

```
## [1] ".."
## [2] "../analysis"
## [3] "../analysis/Exercise_files"
## [4] "../analysis/Exercise_files/figure-html"
## [5] "../analysis/Exercise_files/figure-latex"
## [6] "../data"
```

We can now copy the `datasets.xlsx` file (from `data/` in the main folder to the `Exercise/data/` folder).

```
file.copy("../../data/datasets.xlsx", "../data")
```

If you're running an Unix system, you can also use a bash command (`mv ../../data/datasets.xlsx ../data/`). This is the occasion to try to include a code chunk with another language (i.e. bash)!

5 Load the project data

We'll use the `mtcars` dataset that we'll load from the `datasets.xlsx` file. Here we do not provide the code because it is part of the exercise.

6 Analyze the data

6.1 Get a look at the data

```
my_mtcars2 <- within(my_mtcars, {
  vs <- factor(vs, labels = c("V", "S"))
  am <- factor(am, labels = c("automatic", "manual"))
  cyl  <- ordered(cyl)
```

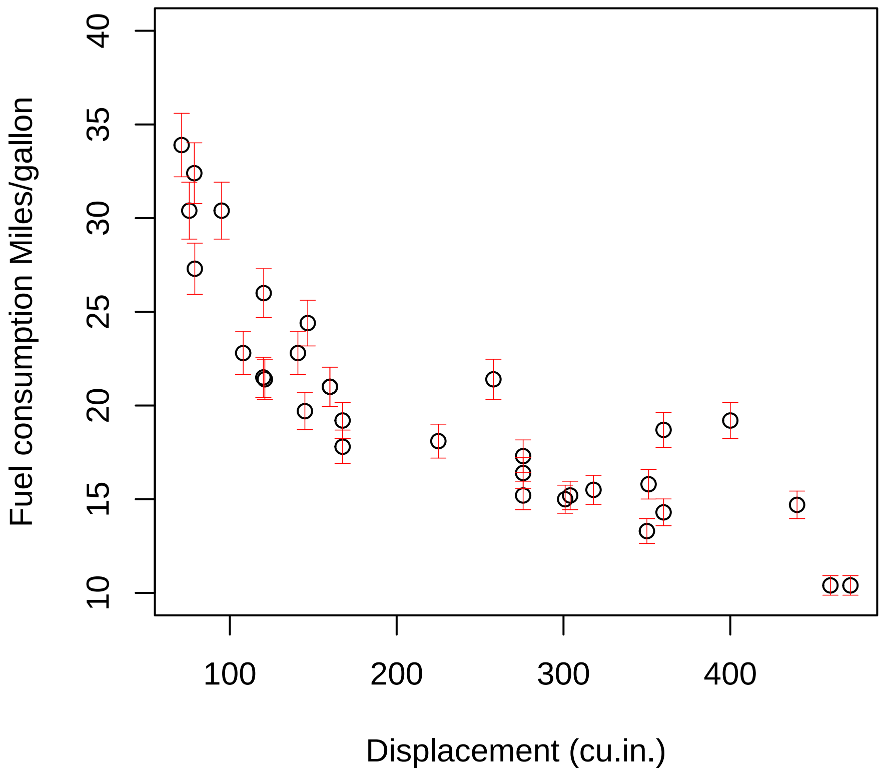
```r
  gear <- ordered(gear)
  carb <- ordered(carb)
})
summary(my_mtcars2)
```

```
##       mpg         cyl         disp             hp             drat      
##  Min.   :10.40   4:11   Min.   : 71.1   Min.   : 52.0   Min.   :2.760  
##  1st Qu.:15.43   6: 7   1st Qu.:120.8   1st Qu.: 96.5   1st Qu.:3.080  
##  Median :19.20   8:14   Median :196.3   Median :123.0   Median :3.695  
##  Mean   :20.09          Mean   :230.7   Mean   :146.7   Mean   :3.597  
##  3rd Qu.:22.80          3rd Qu.:326.0   3rd Qu.:180.0   3rd Qu.:3.920  
##  Max.   :33.90          Max.   :472.0   Max.   :335.0   Max.   :4.930  
##        wt             qsec        vs             am     gear   carb  
##  Min.   :1.513   Min.   :14.50   V:18   automatic:19   3:15   1: 7  
##  1st Qu.:2.581   1st Qu.:16.89   S:14   manual   :13   4:12   2:10  
##  Median :3.325   Median :17.71                         5: 5   3: 3  
##  Mean   :3.217   Mean   :17.85                                4:10  
##  3rd Qu.:3.610   3rd Qu.:18.90                                6: 1  
##  Max.   :5.424   Max.   :22.90                                8: 1  
```

6.2 First analysis, fuel consumption

I want to assess the motors consumption as a function of the displacement. Engine displacement is the measure of the cylinder volume swept by all of the pistons of a piston engine, excluding the combustion chambers.

```r
plot(mpg ~ disp,
     data = my_mtcars,
     xlab="Displacement (cu.in.)",
     ylab="Fuel consumption Miles/gallon",
     ylim=c(10,40))
arrows(x0 = my_mtcars$disp,
       y0 = my_mtcars$mpg * 0.95,
       x1 = my_mtcars$disp,
       y1 = my_mtcars$mpg * 1.05,
       angle = 90,
       code = 3,
       length = 0.04,
       lwd = 0.4,
       col="red")
```



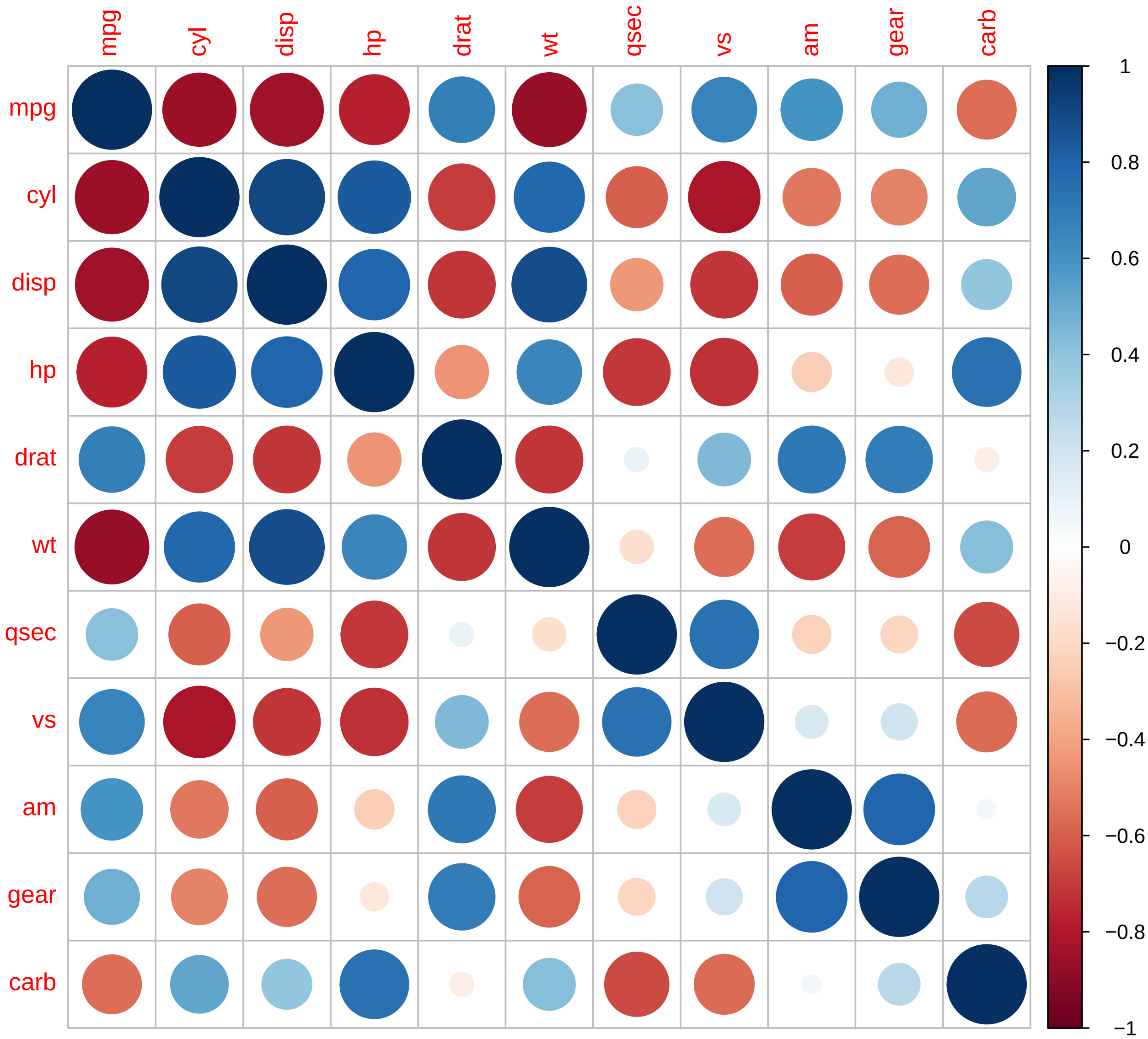
Whoo, that was **interesting**. Can we conclude that bigger motors consume more fuel?

Let's move on to the next analysis.

6.3 Second analysis, is that all?

To assess whether other variables could also explain the consumption I will plot the correlation among each pair of variables in my data.

```
mcor <- cor(my_mtcars)
corrplot(mcor)
```



Whoohoohooo, that plot looks great!

It looks like the **number of cylinders** (cyl), **Gross horsepower** (hp) and the **Weight** (wt. 1000 lbs), could play a role too.

If you are looking for the variables meaning, you can find it here.

6.4 Third analysis, we like camembert, they like pie!

Let's produce a pie chart showing the proportion of cars that have different carb values.

```
#concat the names for the legend
names = as.factor(paste(unique(my_mtcars$carb),"carbs"))
#get the percentages
percent=100*table(my_mtcars$carb)/length(my_mtcars$carb)
pie(x=percent, label=paste(percent, "%"),  col=rainbow(length(names)),
    main="Percentage of cars per number of carburators" )
```

```r
#add legend
legend("right",legend = names, fill = rainbow(length(names)), cex=0.8)
```

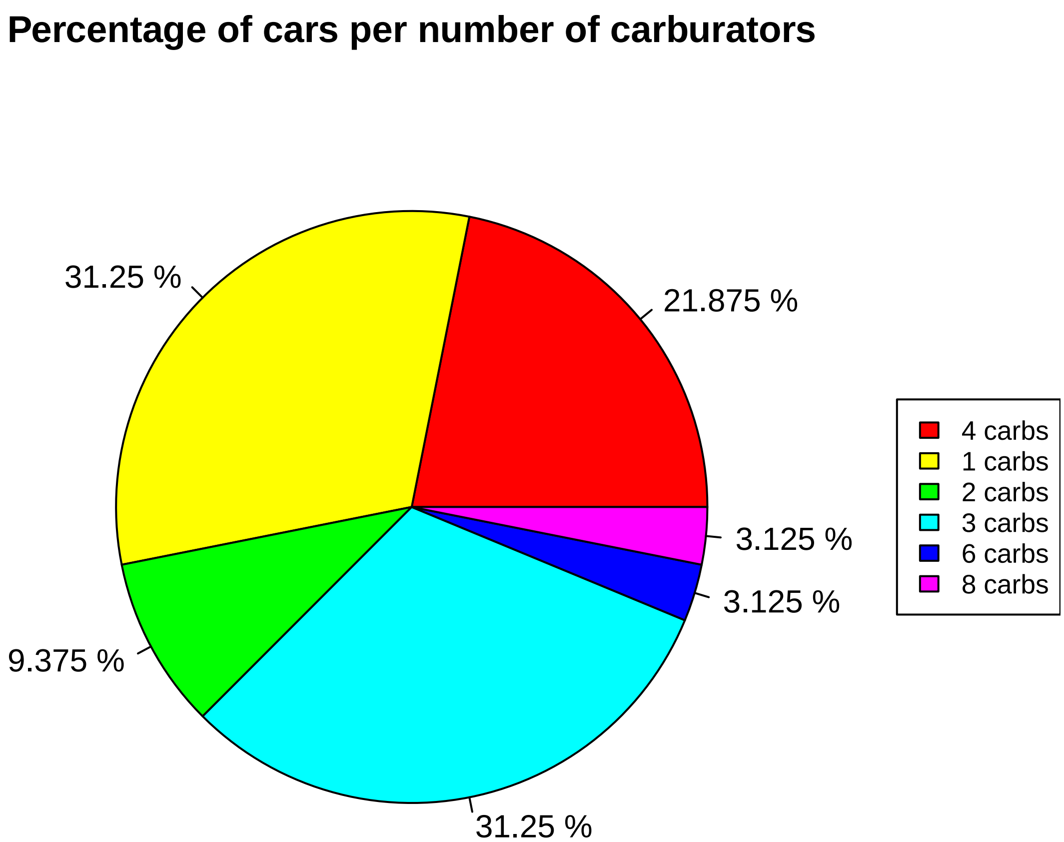


Figure 1: have you ever tried the camenbert pie?

7 Let's make sure that my figures have been saved!

```r
list.dirs(path = ".", full.names = TRUE, recursive = TRUE)
```

```
## [1] "."                              "./Exercise_files"
## [3] "./Exercise_files/figure-html"  "./Exercise_files/figure-latex"
```

```r
list.files(path = "./Exercise_files/figure-latex/")
```

```
## [1] "ho_mae-1.pdf"    "mpg_disp-1.pdf"  "mtcorplot-1.pdf"
```

8 Print a nice table of the `mtcars` dataset

Try to reproduce this table with `kable()` and the `kableExtra` functions.

Table 1: a wonderful table

mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
21.0	6	160	110	~~3.90~~	2.620	16.46	0	1	4	4
21.0	6	160	110	~~3.90~~	2.875	17.02	0	1	4	4
22.8	4	108	93	~~3.85~~	2.320	18.61	1	1	4	1
21.4	6	258	110	~~3.08~~	3.215	19.44	1	0	3	1
18.7	8	360	175	~~3.15~~	3.440	17.02	0	0	3	2
18.1	6	225	105	~~2.76~~	3.460	20.22	1	0	3	1

9 Insert an image

Insert the image of a cute kitten here :) Limit the width to 200px and center the picture!

```
knitr::include_graphics("../data/IMG_9769.JPG")
```

Figure 2: A cute kitten

```
# this is another way to include external images with knitr - it allows you to use
# the code chunk parameters to place/resize/etc the image
```

Bibliography

Salinas, Isaac Quintanilla. 2020. "R Markdown Guide."

Xie, Yihui, Joseph J Allaire, and Garrett Grolemund. 2018. *R Markdown: The Definitive Guide*. Chapman; Hall/CRC.