

Module4__Russo

Matt Russo

2/24/2021

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:

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3    v purrr  0.3.4
## v tibble  3.0.6    v dplyr  1.0.4
## v tidyr   1.1.2    v stringr 1.4.0
## v readr   1.4.0    v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
```

```
df = read_csv('/Users/mattrusso/Desktop/goedert2020imputed (1).csv')
```

```
##
```

```
## -- Column specification -----
```

```
## cols(
##   id = col_character(),
##   frontal = col_character(),
##   prisms = col_character(),
##   session = col_character(),
##   cbsTotal = col_double(),
##   daysPostStroke = col_double(),
##   sex = col_character(),
##   admFIM = col_double(),
##   lesion_vol = col_double(),
##   age = col_double(),
##   mmse = col_double()
## )
```

#8. Graph the recovery trajectories on the cbsTotal of the 17 patients over the 6 weeks (sessions) as a function of prism treatment and frontal damage. Hint: facet on prisms and frontal using code similar to that used by Healy on p.77 without the smoothing geom.

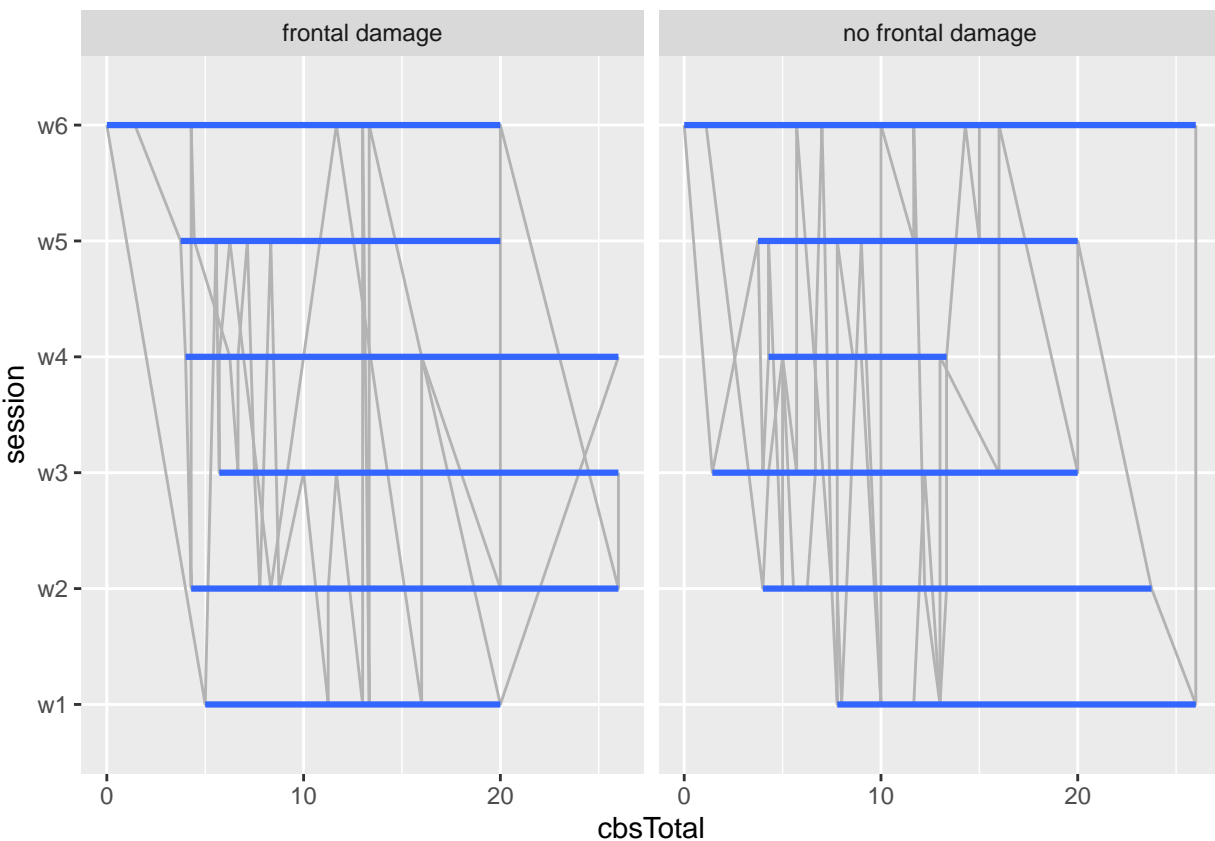
```
p <- ggplot(data=df, mapping = aes(x=cbsTotal, y = session))
p +geom_line(color="gray70", aes(group = prisms))+geom_smooth(size = 1.1, method = 'loess', se= FALSE)+

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

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 13

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 3

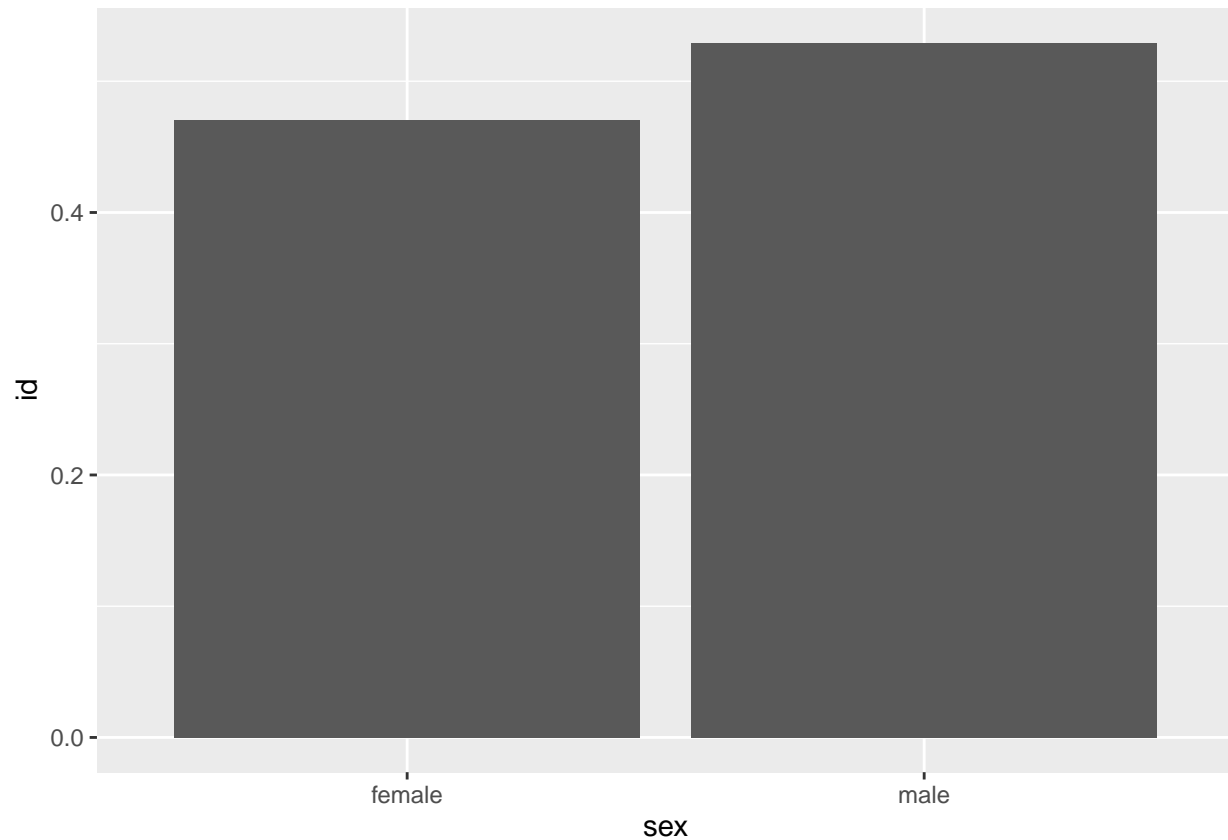
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```



Including Plots

You can also embed plots, for example: #9. Get a graph of the proportion of men and women in the sample. However, because each subject is represented in multiple rows in the dataset, you cannot get this information from the large dataframe. Use the follow code (subbing the dataframe or variable names you are using) to create a smaller dataframe that you can use for graphing the proportion of men and women in the sample, using the geom_bar function, with that proportion represented on the y-axis. Here is the code:

```
sexDF <- df %>%
  select(id,sex)
uniqueSex <- unique(sexDF)
p <- ggplot(data=uniqueSex, mapping = aes(x=sex, y = id))
p+geom_bar(mapping = aes(y=..prop..,group=1))
```



#10. Working with the original dataframe, use piping and the necessary dplyr functions to create summary table of the mean and sd for the cbsTotal score by treatment condition (prisms vs no prisms). Healy describes how to do this in the first part of Chapter 5. You can also review the <https://r4ds.had.co.nz/book-section-18> on Pipes and search the dplyr::summarise function in the R help.

```
library(dplyr)
mean_df <- df %>%
  group_by(prisms) %>%
  summarize(mean = mean(cbsTotal),sd = sd (cbsTotal))
view(mean_df)
```