Lab 6

Lisa Strycker

October 21, 2019

Question 1.

Create a new R Markdown document and modify the YAML to:

- Include your name
- Change the syntax highlighting to any scheme but the default. The options are default, tango, pygments, kate, monochrome, espresso, zenburn, haddock, and textmate. You can also use NULL if you want no highlighting.
- Include the option to make it easy to modify the rendering between PDF and HTML.

Done. This is what the YAML code looks like:

title: "Lab 6"

author: "Lisa Strycker"

date: "October 21, 2019"

output: html_document: highlight: kate

pdf_document: highlight: tango.

Question 2.

Create a code chunk that:

- Loads all the packages you decide to use for the lab
- Sets a global chunk option to make all figures 6.5" wide and the height to a value that makes sense to you
- Does not display the code or any warnings, messages, etc. from the code, but evaluates every function/line of the code.

Done. See library and knitr code below.

```r
# tinytex::instal l_tinytex()
library(here)
```

```
## here() starts at U:/lisas/Hdrive-LisaS/odrive/R/joe nese class/Projects for Labs/lab 6 project
```

```r
library(rio)
library(tinytex)
library(tidyverse)
```

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

```
## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0


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

```r
library(janitor)
```

```
##
## Attaching package: 'janitor'

## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

```r
library(rmarkdown)
library(knitr)
knitr::opts_chunk$set(echo = FALSE, warning = FALSE, message = FALSE, eval = TRUE, error = FALSE, fig.wi
```
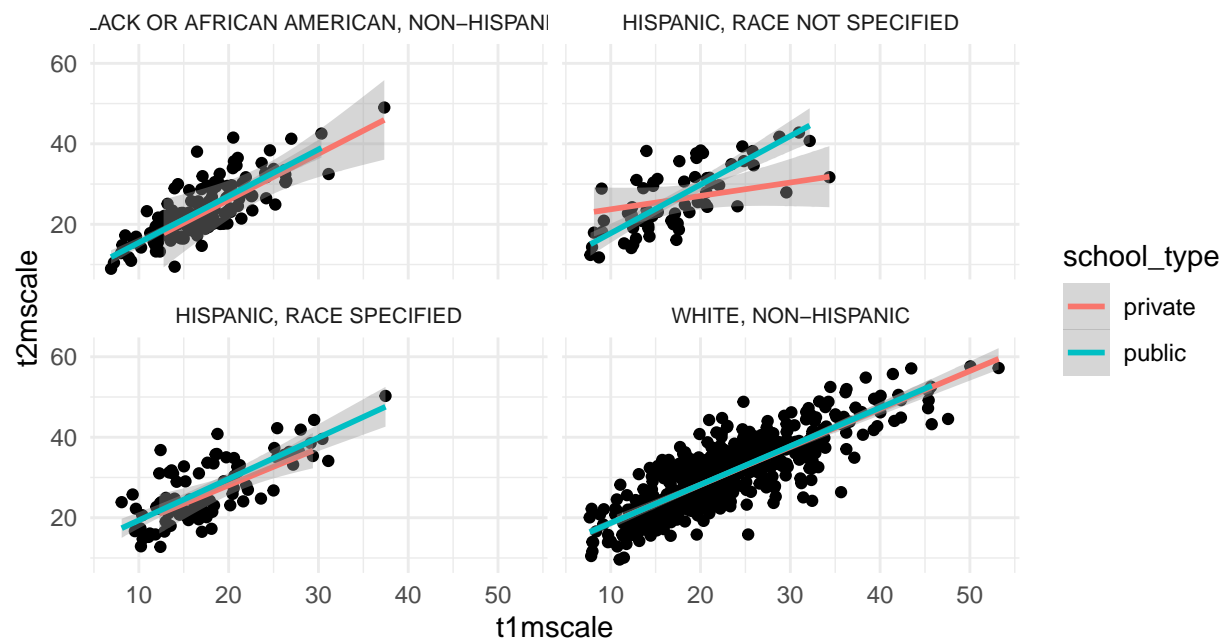
Question 3.

Import the ecls-k_samp.sav dataset (stored on Canvas in Files -> data), and produce the plot below. Do
not show the code you used (colors, themes, etc. don't matter here).
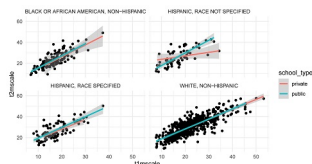
Done. See plot below. Note that the code I used is not shown.

```
## [1] "U:/lisas/Hdrive-LisaS/odrive/R/joe nese class/Projects for Labs/lab 6 project"
```

Question 4.

Run the following lines of code to store the mean and standard deviation of t1mscale.

- t1mean <- mean(eclskclean$t1mscale, na.rm = TRUE)
- t1sd <- sd(eclskclean$t1mscale, na.rm = TRUE)

Extend this code to calculate (in the same code chunk) the mean and standard deviation of t2mscale. Using the values you calculated, use an inline code evaluation below to report the means/SDs for the two time points (t1mscale and t2mscale). Also report the difference between the means (i.e., the average gain).

Done. See code below.

```
t1mean <- round(mean(eclskclean$t1mscale, na.rm = TRUE), digits=2)
t1sd <- round(sd(eclskclean$t1mscale, na.rm = TRUE), digits=2)
t2mean <- round(mean(eclskclean$t2mscale, na.rm = TRUE), digits=2)
t2sd <- round(sd(eclskclean$t2mscale, na.rm = TRUE), digits=2)
```

Check mean, SD, and difference score values.

```
t1mean
```

```
## [1] 20.51
```

```
t1sd
```

```
## [1] 7.37
```

```
t2mean
```

```
## [1] 28.57
```

```
t2sd
```

```
## [1] 8.78
```

```
t2mean - t1mean
```

```
## [1] 8.06
```

See sentence below with inline code.

The mean (SD) was 20.51 (7.37) at time 1 and 28.57 (8.78) at time 2. Average gain was 8.06.

Question 5.

Pretend you are trying to teach somebody how to load data. Describe the process below that we've discussed in class, including why it helps reproducibility, and echo chunks of code as necessary without actually evaluating any of it.

It helps reproducibility to load data into an R project directory so that data are always accessible, and the R code always works (and analyses can be reproduced), for anyone at any time on any computer.

To load data:

* Create a new R project directory.

* Within this directory create "data" and "scripts" folders.

* Within the scripts folder, create a new R markdown file to contain all the code needed for a certain set of analyses.

* Paste the raw data that you will be importing into the data folder.

* Use the here() function to locate this dataset in the R markdown file.

For instance, to import the ecls-k_samp.sav SPSS file, Use code like this (for this code, echo = TRUE but eval = FALSE so that you can see the code but it is not actually evaluated):

```
eclsk <- import(here("data", "ecls-k_samp.sav"))
```

Then clean the imported data, using the characterize() code to keep value labels and the clean_names() code to fix variable names (lower case, underscores instead of spaces, etc.). Use the "View" function to see and check the imported dataset. The code looks like this (again, for this code, echo = TRUE but eval = FALSE so that you can see the code but it is not actually evaluated):

```
eclskclean <- eclsk %>%
  characterize() %>%
  clean_names()
View(eclskclean)
```

Does it work? Let's run the code and view the imported dataset. This time, echo = FALSE and eval = TRUE so that you do not see the code but the code is actually evaluated. Yes!

4

```
## [1] "U:/lisas/Hdrive-LisaS/odrive/R/joe nese class/Projects for Labs/lab 6 project"

##    child_id         teacher_id         school_id
## Length:984        Length:984         Length:984
## Class :character  Class :character   Class :character
## Mode  :character  Mode  :character   Mode  :character
##
##
##
##     k_type          school_type          sex
## Length:984        Length:984         Length:984
## Class :character  Class :character   Class :character
## Mode  :character  Mode  :character   Mode  :character
##
##
##
##     ethnic          famtype            numsibs          ses_cont
## Length:984        Length:984         Min.   : 0.000   Min.   :-4.4700
## Class :character  Class :character   1st Qu.: 1.000   1st Qu.:-0.4325
## Mode  :character  Mode  :character   Median : 1.000   Median :-0.0200
##                                      Mean   : 1.446   Mean   : 0.1048
##                                      3rd Qu.: 2.000   3rd Qu.: 0.6300
##                                      Max.   :10.000   Max.   : 2.5600
##    ses_cat           age            t1rscale         t1mscale
## Length:984        Min.   :54.00   Min.   :10.67   Min.   : 6.903
## Class :character  1st Qu.:65.73   1st Qu.:16.95   1st Qu.:15.173
## Mode  :character  Median :68.98   Median :21.46   Median :19.353
##                   Mean   :68.94   Mean   :23.37   Mean   :20.506
##                   3rd Qu.:72.12   3rd Qu.:27.68   3rd Qu.:24.338
##                   Max.   :79.00   Max.   :66.69   Max.   :53.200
##    t1gscale         t2rscale         t2mscale         t2gscale
## Min.   : 7.695   Min.   :12.28   Min.   : 8.951   Min.   : 8.59
## 1st Qu.:17.656   1st Qu.:25.19   1st Qu.:21.879   1st Qu.:22.42
## Median :22.805   Median :32.27   Median :28.363   Median :28.70
## Mean   :23.228   Mean   :32.95   Mean   :28.567   Mean   :28.32
## 3rd Qu.:28.504   3rd Qu.:38.29   3rd Qu.:34.115   3rd Qu.:33.67
## Max.   :42.808   Max.   :69.91   Max.   :57.614   Max.   :47.47
##  ir_treadgain     ir_tmathgain     ir_tgkgain         t1arslit
## Min.   :-10.732  Min.   :-9.469  Min.   :-12.002  Min.   :1.000
## 1st Qu.: 5.122  1st Qu.: 4.473  1st Qu.: 2.510  1st Qu.:2.230
## Median : 8.808  Median : 7.613  Median : 5.143  Median :2.680
## Mean   : 9.576  Mean   : 8.061  Mean   : 5.097  Mean   :2.697
## 3rd Qu.: 13.135  3rd Qu.:11.452  3rd Qu.: 7.641  3rd Qu.:3.190
## Max.   : 33.868  Max.   :28.185  Max.   : 17.915  Max.   :5.000
##    t1arsmat         t1arsgen         t2arslit         t2arsmat
## Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
## 1st Qu.:2.110   1st Qu.:2.010   1st Qu.:2.990   1st Qu.:3.010
## Median :2.560   Median :2.690   Median :3.430   Median :3.660
## Mean   :2.646   Mean   :2.755   Mean   :3.469   Mean   :3.611
## 3rd Qu.:3.230   3rd Qu.:3.480   3rd Qu.:3.917   3rd Qu.:4.250
## Max.   :5.000   Max.   :5.000   Max.   :5.000   Max.   :5.000
##    t2arsgen        ar_slitgain       ar_smathgain       ar_sgkgain
## Min.   :1.000   Min.   :-1.7900   Min.   :-1.2300   Min.   :-1.7900
## 1st Qu.:2.980   1st Qu.: 0.3300   1st Qu.: 0.4400   1st Qu.: 0.3275
```

```
##  Median :3.950   Median : 0.7100   Median : 0.9550   Median : 0.9700
##  Mean   :3.721   Mean   : 0.7715   Mean   : 0.9653   Mean   : 0.9658
##  3rd Qu.:4.520   3rd Qu.: 1.1525   3rd Qu.: 1.4300   3rd Qu.: 1.5200
##  Max.   :5.000   Max.   : 3.0500   Max.   : 3.6700   Max.   : 4.0000
##    testdate1          testdate2            elapse
##  Min.   :1998-09-12  Min.   :1999-03-11  Min.   :129.0
##  1st Qu.:1998-10-20  1st Qu.:1999-04-22  1st Qu.:170.0
##  Median :1998-11-03  Median :1999-05-04  Median :184.0
##  Mean   :1998-10-31  Mean   :1999-05-04  Mean   :184.7
##  3rd Qu.:1998-11-12  3rd Qu.:1999-05-17  3rd Qu.:199.2
##  Max.   :1998-12-17  Max.   :1999-06-17  Max.   :250.0
```