week9_pdf

Mathi Manavalan

3/21/2020

R Studio API

Libraries

```
library(tidyverse)
library(lubridate)
library(ggplot2)
```

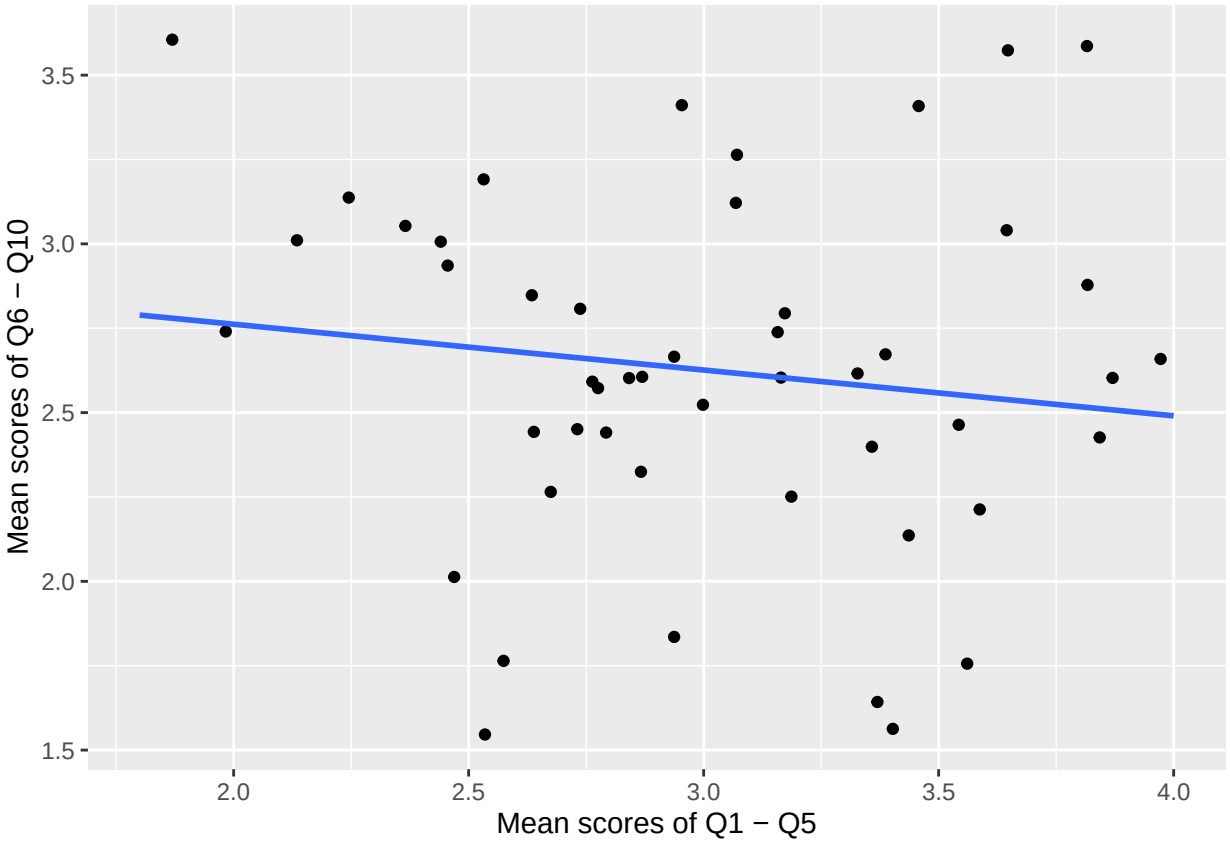
Data Import

```
week9_tbl <- as.tibble(read_csv(file = "../data/week3.csv")) %>%
  mutate(timeStart = ymd_hms(timeStart), timeEnd = ymd_hms(timeEnd)) %>%
  mutate(condition = factor(condition, levels = c("A", "B", "C"),
                            labels = c("Block A", "Block B", "Control")),
         gender = factor(gender, levels = c("M", "F"), labels = c("Male", "Female"))) %>%
  mutate(avg15 = (q1 + q2 + q3 + q4 + q5)/5, avg610 = (q6 + q7 + q8 + q9 + q10)/5)
```

Cleaning steps done

- imported data as a tibble
- formatted date and time into appropriate variable types
- formatted condition and gender into appropriately labeled factors
- created two new columns of data
  - column of average of scores on questions Q1-Q5
  - column of average of score on questions Q6-Q10

Visualization



The above plot displays a scatterplot of the average scores of participants in questions Q1 through Q5 and the average scores of same participants in questions Q6 through Q10.

Analysis

```
analysis <- summary(lm(week9_tbl$avg610 ~ week9_tbl$avg15))
analysis
```

```
##
## Call:
## lm(formula = week9_tbl$avg610 ~ week9_tbl$avg15)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0804 -0.2533  0.0010  0.2924  1.0824
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3.0334     0.4392   6.906 1.13e-08 ***
## week9_tbl$avg15  -0.1358     0.1437  -0.945     0.35
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.5196 on 47 degrees of freedom
## Multiple R-squared:  0.01863,    Adjusted R-squared:  -0.002248
## F-statistic: 0.8923 on 1 and 47 DF,  p-value: 0.3497
```

The correlation was 0.0186322 ($p<.05$), which is not statistically significant.