Visualized-project.R

yilinji

2020-11-02

```
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────
──────────── tidyverse 1.3.0 ──
```

```
## ✓ ggplot2 3.3.2     ✓ purrr   0.3.4
## ✓ tibble  3.0.3     ✓ dplyr   1.0.2
## ✓ tidyr   1.1.2     ✓ stringr 1.4.0
## ✓ readr   1.3.1     ✓ forcats 0.5.0
```

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

```
library(psych)
```

```
##
## Attaching package: 'psych'
```

```
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```
#Import data "Wong" into R and created a new data frame "IQ"
IQ <- read_csv(file = "~/Desktop/2020-2021/PSYC 3031/R Project/Data/Won
g.csv")
```

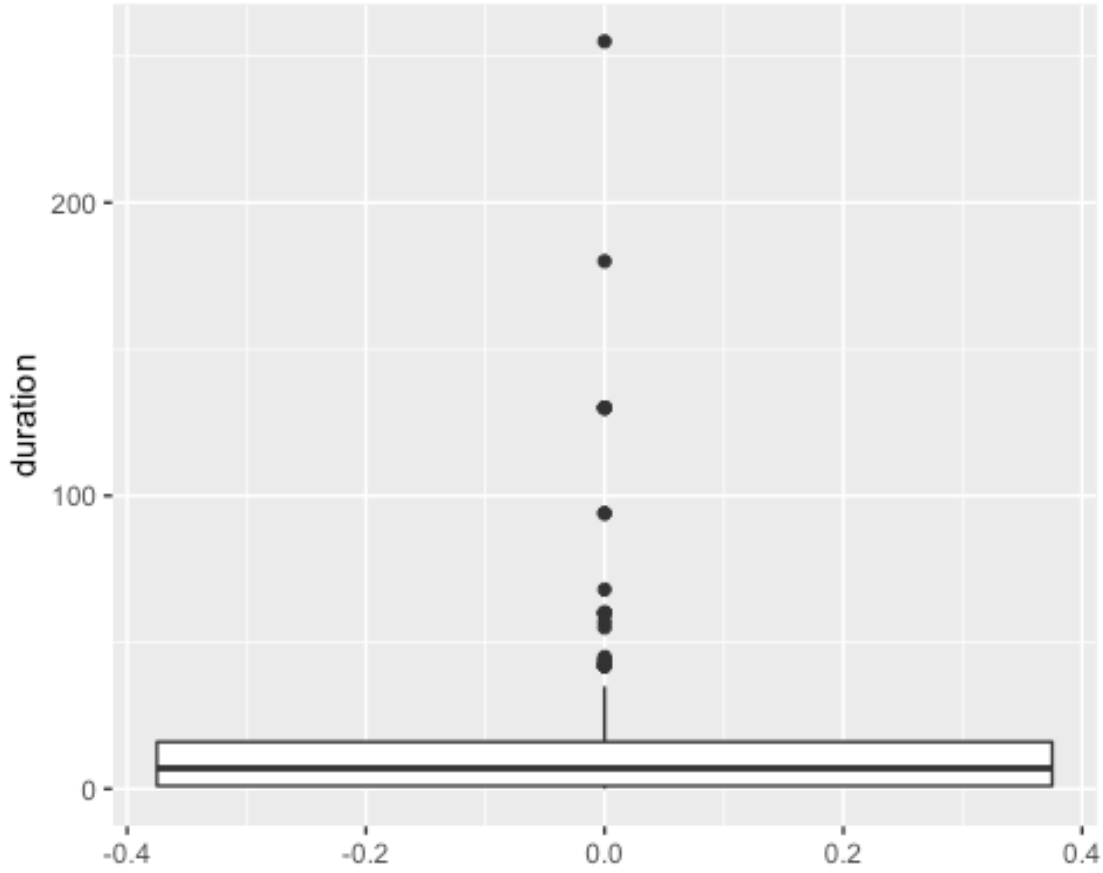
```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   X1 = col_double(),
##   id = col_double(),
##   days = col_double(),
##   duration = col_double(),
##   sex = col_character(),
##   age = col_double(),
##   piq = col_double(),
##   viq = col_double()
## )
```

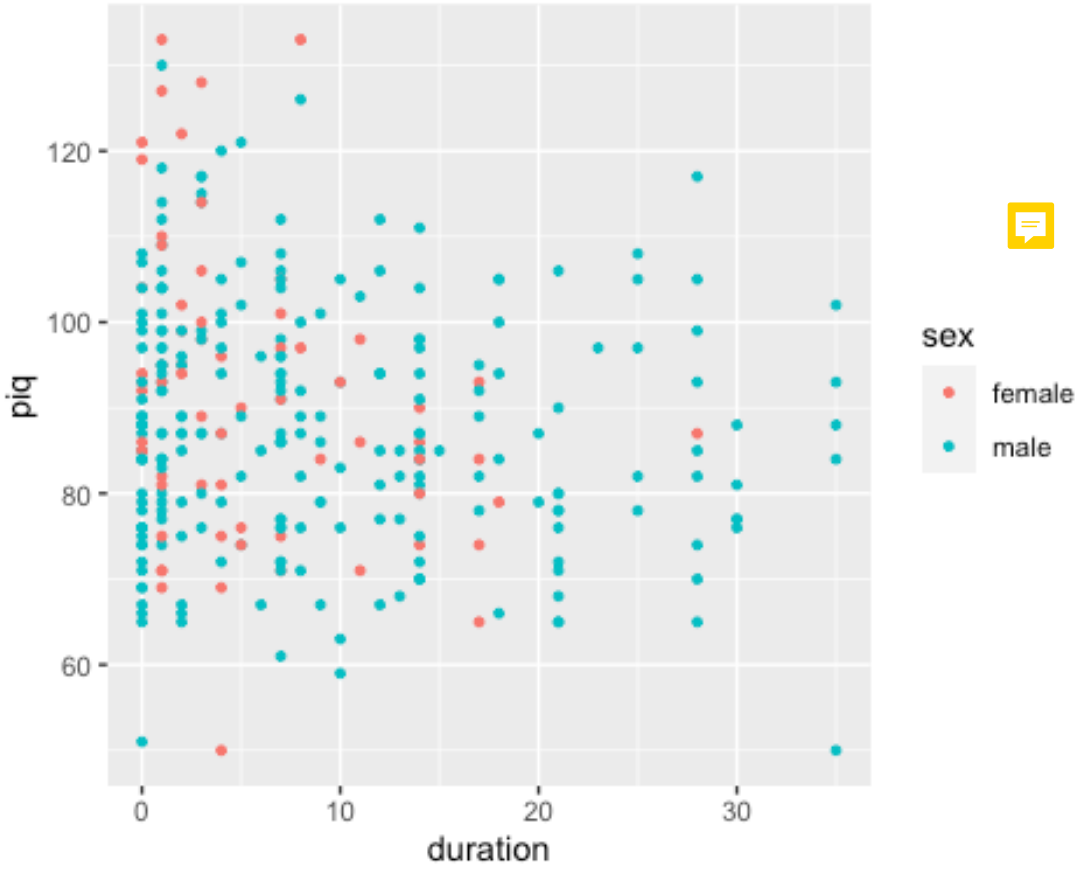
```r
#Using "select" function to abstract variables that we need
IQ <- select(IQ, sex, duration, piq, viq)

#Useing "mutate" function to convert duration as numerical
IQ <- mutate(IQ, duration = as.numeric(duration))

#Using "mutate" function to record the "sex" column
IQ <- mutate(IQ, sex = fct_recode(sex, "male" = "Male", "female" = "Fem
ale"))

#create a new data frame "description to obtain a general description o
f ...
#variables we need
description <- describe(select(IQ, duration, piq, viq))
description <- round(description,digits = 2)

#Export description from R
write_csv(description, "Genernal Description.csv")

#Using "select" function to abstract "sex" and using "summary" function
s...
#... to get a description
IQ %>%
  select(sex) %>%
  summary()

##      sex
##  female: 71
##  male  :260

#Using "describeBy" function to get descriptions of column "duration",
"piq"...
#... and "viq" but group by "sex"
IQ %>%
  select(duration, piq, viq) %>%
  describeBy(group = IQ$sex)

##
##  Descriptive statistics by group
## group: female
##          vars  n  mean    sd median trimmed   mad min max range skew
 kurtosis
## duration    1 71  9.27 13.18      4    6.11  4.45   0  68    68 2.36
     5.65
## piq         2 71 89.18 18.00     87   87.96 17.79  50 133    83 0.48
    -0.06
## viq         3 71 94.35 14.25     92   94.54 17.79  64 131    67 0.03
    -0.81
##            se
## duration 1.56
```

```
## piq      2.14
## viq      1.69
## ------------------------------------------------------------
## group: male
##          vars   n  mean    sd median trimmed   mad min max range ske
w kurtosis
## duration    1 260 15.67 28.43      7    9.62  8.90   0 255   255 4.5
2    26.29
## piq         2 260 87.11 14.26     87   86.84 14.83  50 130    80 0.1
9    -0.14
## viq         3 260 95.13 14.02     94   94.62 14.83  64 132    68 0.3
0    -0.48
##            se
## duration 1.76
## piq      0.88
## viq      0.87
```

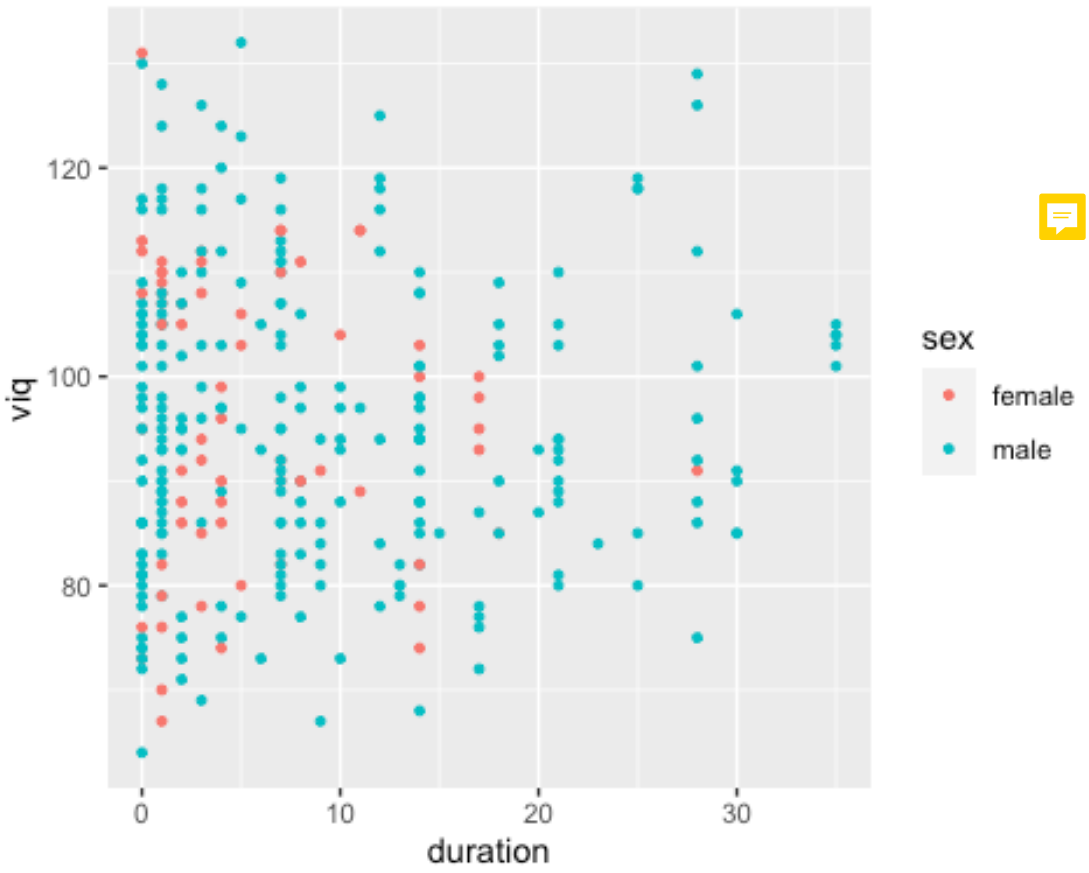
```r
#Using "ggplot" and "geom_boxplot" function to obtain the boxplot diagr
am of...
#... column "duration"
ggplot(data = IQ,
       mapping = aes(y = duration)) +
  geom_boxplot()
```



```r
#Printing the outliers out and assigned to "Outlier"
Outlier <- boxplot.stats(IQ$duration)$out

#Using "-which" function to give all of the FALSE indices of...
#"IQ$duration %in% Outlier" to remove all of the outliers in the datase
t
IQ<- IQ[-which(IQ$duration %in% Outlier),]

#Using "ggplot" and "geom_point" functions to conduct a scatterplot dia
gram...
#which x-axise is "duration", y-axies is "piq" and seprated by color fr
om "sex"
ggplot(IQ, aes(x = duration, y = piq, color = sex))  +
  geom_point(size = 1)
```




```r
#Using "ggplot" and "geom_point" functions to conduct a scatterplot dia
gram...
#which x-axise is "duration", y-axies is "viq" and seprated by color fr
om "sex"
ggplot(IQ, aes(x = duration, y = viq, color = sex))  +
  geom_point(size = 1)
```


viq
120
100
80
0
10
20
30
duration
sex
female
male