Graphical Review of Test #1

NPFL 054 lab session (Hladká and Holub, 2017)

Reading the data

```
test1 = read.csv("test1.anonymous.csv", header=T, sep="\t", na.strings=c("", "?"))

str(test1)
```

```
# 'data.frame':    65 obs. of  8 variables:
#  $ evaluation: Factor w/ 5 levels "good","moderate",..: 3 3 3 3 3 3 3 3 3 1 ...
#  $ pts1      : int  5 5 5 5 5 5 5 5 5 5 ...
#  $ pts2      : int  7 8 8 8 8 8 8 7 7 6 ...
#  $ pts3      : int  8 7 7 7 7 7 7 7 7 7 ...
#  $ score     : int  20 20 20 20 20 20 20 19 19 18 ...
#  $ reviewer  : Factor w/ 2 levels "BH","MH": 2 2 2 1 1 1 1 2 1 2 ...
#  $ mff       : logi  FALSE FALSE TRUE TRUE FALSE TRUE ...
#  $ form      : Factor w/ 2 levels "CZ","EN": 2 2 1 1 2 1 1 1 2 1 ...
```

How many students

```
# there are 65 students in the data set
# however, only 42 students attended the test

table(test1$reviewer)

# BH MH
# 21 21

# take only students who attended the test
test1 = test1[!is.na(test1$reviewer), ]

# nrow(test1)
# [1] 42
```

Reordering evaluation levels

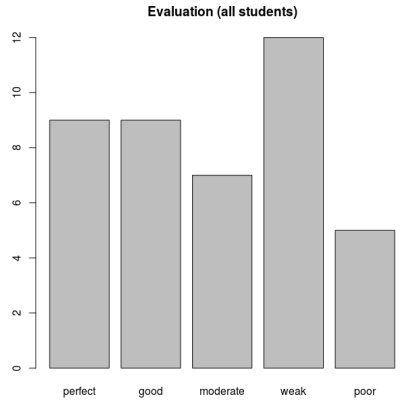
```
# by default, levels are ordered alphabetically

plot(test1$evaluation)

# > levels(test1$evaluation)
# [1] "good"     "moderate" "perfect"  "poor"     "weak"

# reordering
test1$evaluation = factor(test1$evaluation, levels(test1$evaluation)[c(3,1,2,5,4)])

# > levels(test1$evaluation)
# [1] "perfect"  "good"     "moderate" "weak"     "poor"
```

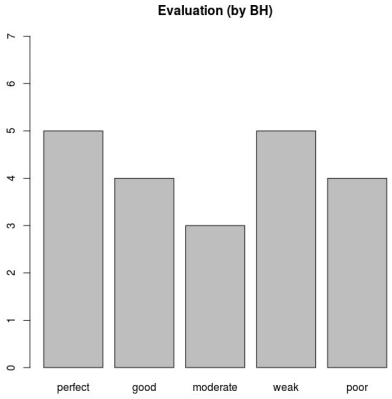
Evaluation summary

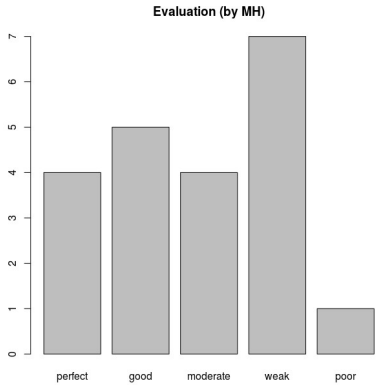
```
table(test1$evaluation)

#  perfect      good       moderate     weak     poor
#        9         9              7       12        5

table(test1$score)

#  0  3  4  5  6  7  8  9 10 12 13 14 15 16 18 19 20
#  1  3  1  1  2  4  3  2  2  2  2  1  6  2  1  2  7
```

Distribution of the evaluation category

```
plot(test1$evaluation, main="Evaluation (all students)")

plot(test1$evaluation[test1$reviewer == "BH"],
      main="Evaluation (by BH)", ylim=c(0,7))
plot(test1$evaluation[test1$reviewer == "MH"],
      main="Evaluation (by MH)")
```







Distribution of the score

```
plot(as.factor(test1$score), main="Score distribution (all students who attended the test)")

plot(as.factor(test1$score[test1$reviewer == "BH"]),
      main="Score distribution (by BH)")
plot(as.factor(test1$score[test1$reviewer == "MH"]),
      main="Score distribution (by MH)", ylim=c(0,4))
```