Median middle number in dataset

Mode most often occurring

Mean sum / n

House prices - median

Salaries - median

Goals - mean

Girls names - mode

User satisfaction - mode

# read in the athletes.csv file

ath <- read.csv('C:\\student\\X00098007\\athletes.csv',header=T)

ath\_f <- ath[ath$Sex==1,]

ath\_m <- ath[ath$Sex==0,]

ath\_fht <- ath\_f$Ht

ath\_mht <- ath\_m$Ht

# open a new window

# hist\_male <- hist(ath\_mht, col = "blue", labels = TRUE)

#x11()

#hist\_female <- hist(ath\_fht, col = "pink", labels = TRUE)

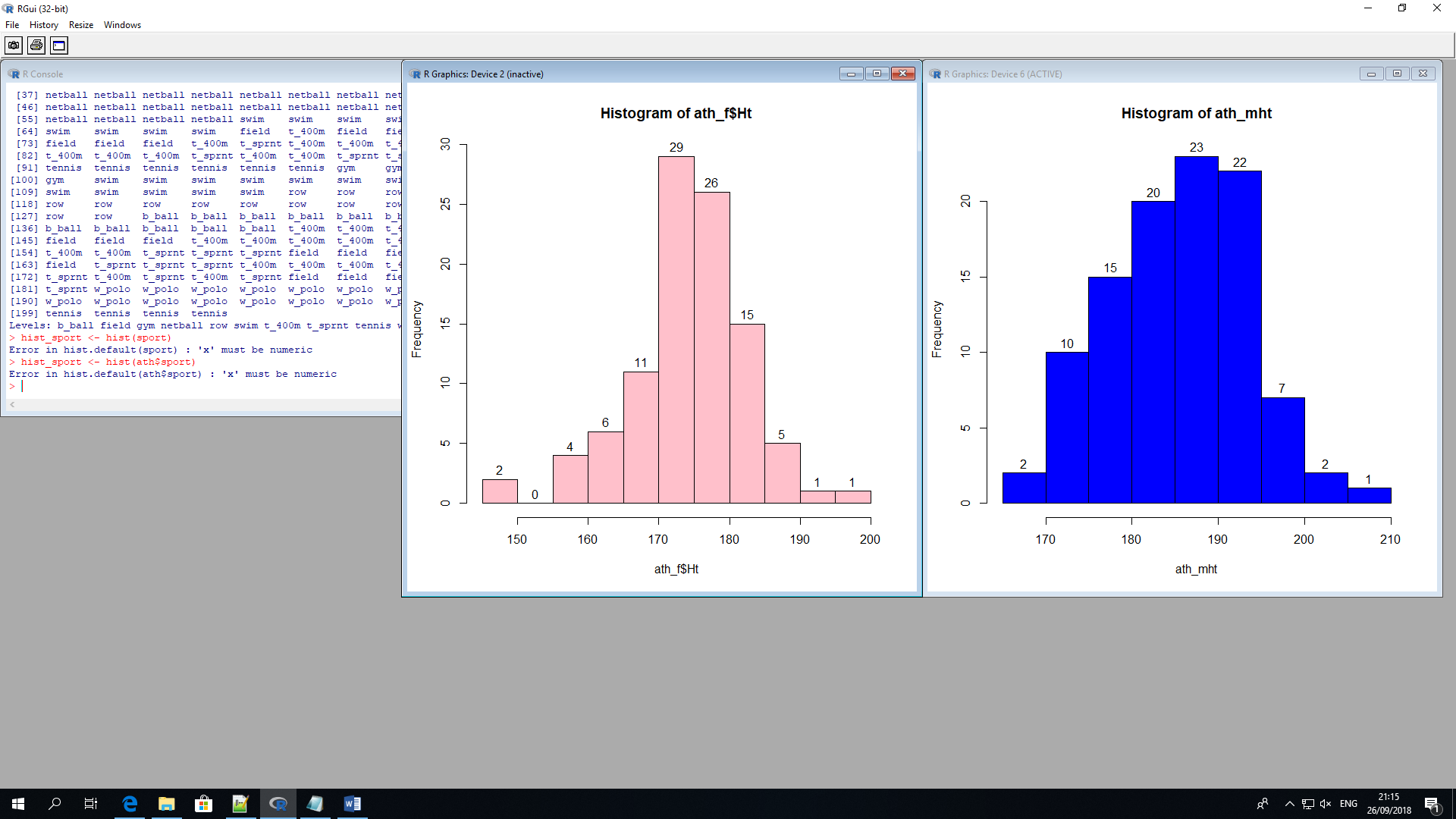
hist(ath\_fht, col = "pink", labels = TRUE)

x11()

hist(ath\_mht, col = "blue", labels = TRUE)

hist\_sport <- hist(ath$sport)

boxplot(ath\_f)



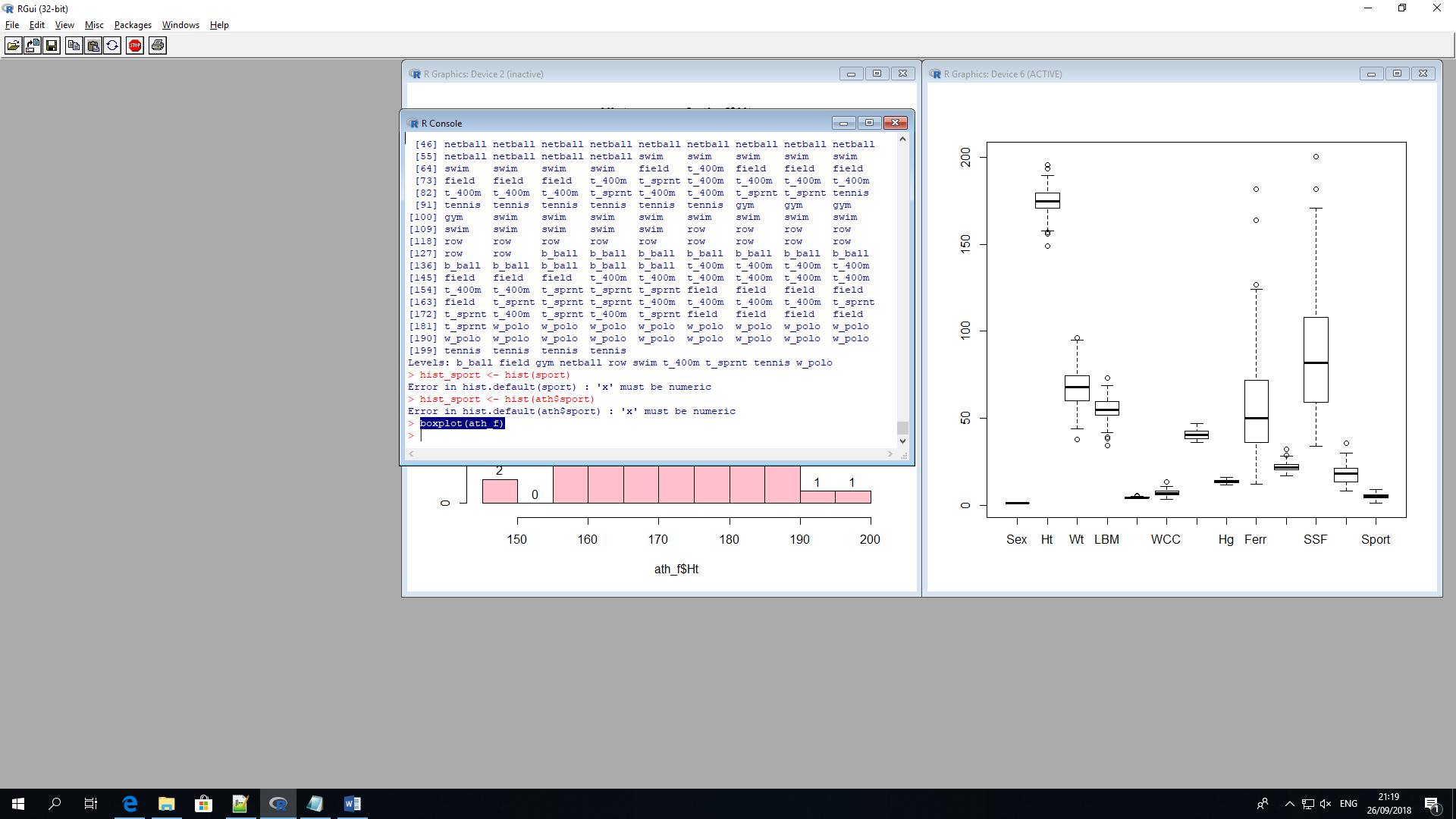
Outliers are of interest in the financial industry as they can cause incorrect results in monetary calculations,

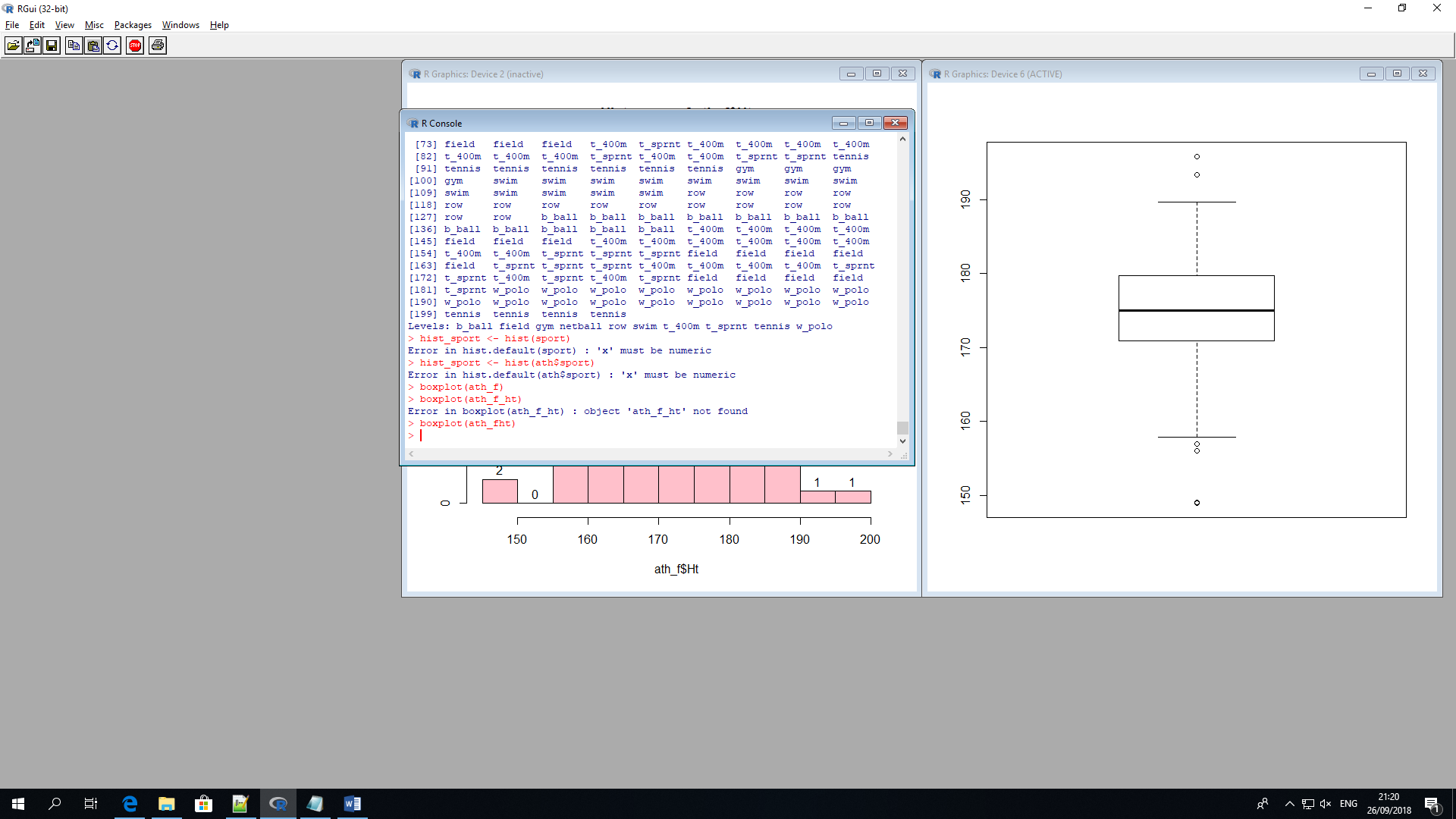
such as interest earned, deposits made

Large deposit, money laundering...

Boxplot

boxplot(ath\_f)





boxplot(ath\_mht)

