# prepare csv file from excel: choose format (eg import data: open excel>data>from text>csv), simplify column names and replace commas by dots (in document options, not using search & replace function!)

# read .csv file with read\_csv (or read\_delim)

> Comp1 <- read\_delim("Comp1.csv", col\_names = TRUE)

# specify column types if not recognized by read\_csv

Comp1 <- read\_csv("Comp1", col\_names = TRUE, col\_types = cols(code = col\_character(), genotype = col\_character(),barcode = col\_character(), NPPC = col\_integer(), location = col\_integer(), REP = col\_integer(), plant\_height = col\_double(), branch\_number = col\_double(), pod\_weight = col\_double()))

# check column specifications

> spec(Comp1)

# eingelesene Tabelle in Fenster ansehen

> view(Comp1)

# subsetting the data set

> comparison\_1A <- Comp1[Comp1$code %in% c("PH01", "PH06", "PH17"), ]

> view(comparison\_1A)

# pipe and summarize (descriptive statistics)

# na.rm=TRUE means remove not available values

# `range´ produces extra line (doubled) in output table

> DStatsComp\_1A <- comparison\_1A %>%

+ group\_by(code) %>%

+ summarize(medianheight = median(plant\_height, na.rm = TRUE),

+ meanheight = mean(plant\_height, na.rm = TRUE),

+ sdheight = sd(plant\_height, na.rm = TRUE),

+ minheight = min(plant\_height, na.rm = TRUE),

+ maxheight = max(plant\_height, na.rm = TRUE),

+ medianbranchno = median(branchno, na.rm = TRUE),

+ meanbranchno = mean(branchno, na.rm = TRUE),

+ sdbranchno = sd(branchno, na.rm = TRUE),

+ minbranchno = min(branchno, na.rm = TRUE),

+ maxbranchno = max(branchno, na.rm = TRUE),

+ medianpodweight = median(pod\_weight, na.rm = TRUE),

+ meanpodweight = mean(pod\_weight, na.rm = TRUE),

+ sdpodweight = sd(pod\_weight, na.rm = TRUE),

+ minpodweight = min(pod\_weight, na.rm = TRUE),

+ maxpodweight = max(pod\_weight, na.rm = TRUE))

> view(DStatsComp\_1A)

# create boxplot

# options: ylim/xlim = range covered by axis, yla/xlab=label axis

> library(ggplot2)

> ggplot(comparison\_1A, aes(x=genotype, y=plant\_height)) +

+ geom\_boxplot()