Modelos lineares aula 1

Danilo de Paula Santos

17/10/2019

# install.packages("compareGroups")  
# install.packages("ggplot2")  
# install.packages("data.table")  
  
  
library("compareGroups")  
library("ggplot2")  
library("haven")

# Estuda da relação entre infecção por acilóstomo e perda de sangue. Tailândia 1970

suwit <- read\_sav("C:/Users/00265021/Documents/Cadeiras/Modelos lineares/Suwit.sav")

compare\_suwit <- compareGroups( ~ ., data = suwit)  
  
summary(compare\_suwit)

##   
## --- Descriptives of each row-variable ---  
##   
## -------------------   
## row-variable: Identificação   
##   
## N mean sd lower upper   
## [ALL] 15 8 4.472136 5.523414 10.47659  
##   
## -------------------   
## row-variable: número de vermes   
##   
## N mean sd lower upper   
## [ALL] 15 552.4 513.9007 267.8113 836.9887  
##   
## -------------------   
## row-variable: Perda de sangue por dia   
##   
## N mean sd lower upper   
## [ALL] 15 33.45267 24.85249 19.68982 47.21551

tabela\_suwit <- createTable(compare\_suwit)

# Estudo Fat\_dat “Fitting Percentage of Body Fat to Simple Body Measurements”

fat\_dat <- read\_sav("C:/Users/00265021/Documents/Cadeiras/Modelos lineares/fat\_dat.sav")

Corrigindo os erros de digitação:

* The body densities for cases 48, 76, and 96, for instance, each seem to have one digit in error as can be seen from the two body fat percentage values.
* Case 42) over 200 pounds in weight who is less than 3 feet tall (the height should presumably be 69.5 inches, not 29.5 inches)!
* The percent body fat estimates are truncated to zero when negative (case 182)

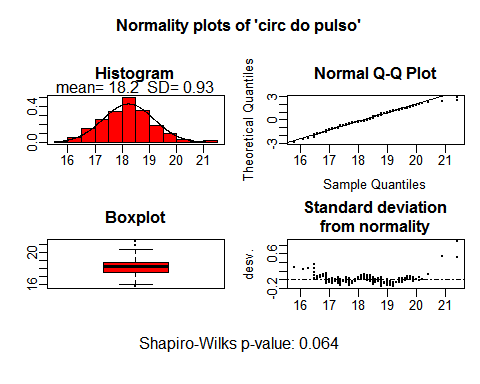
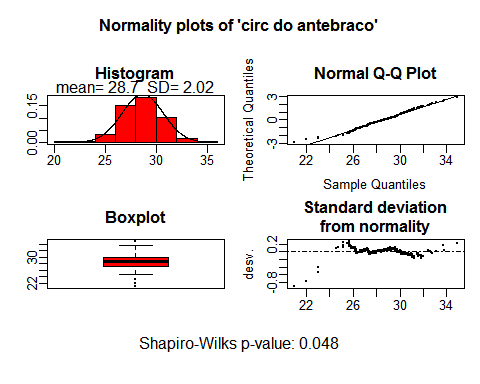
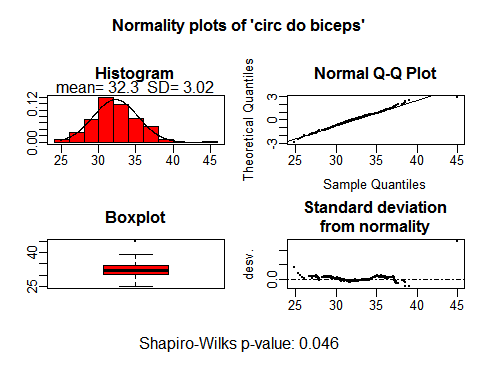
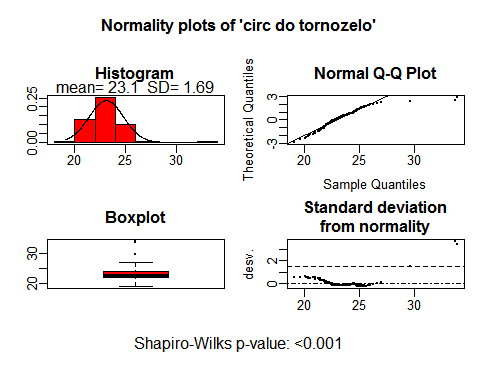
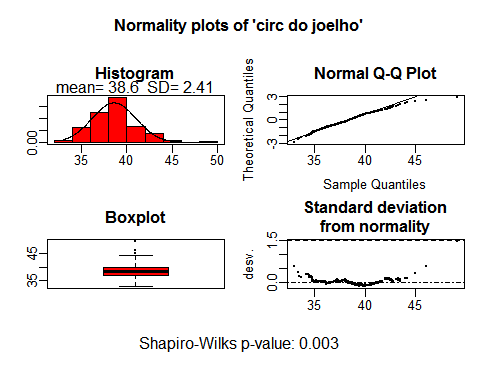
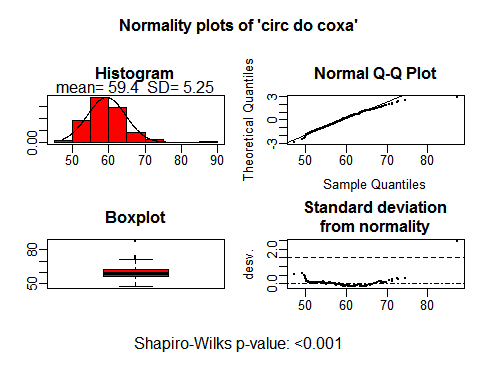
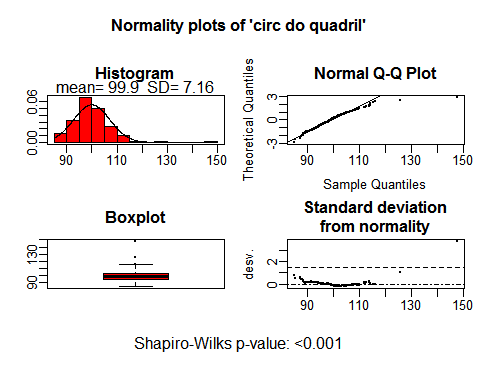
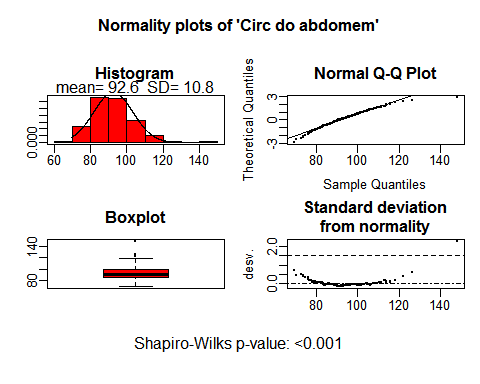
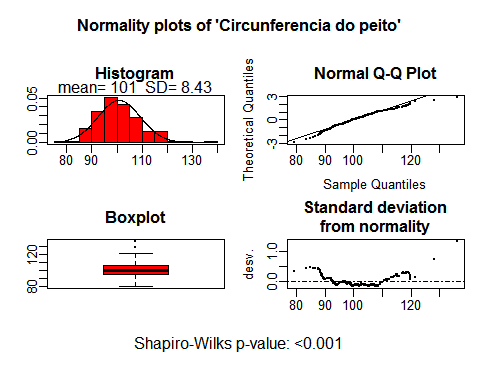
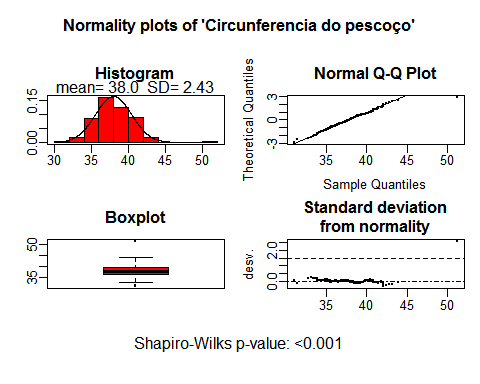
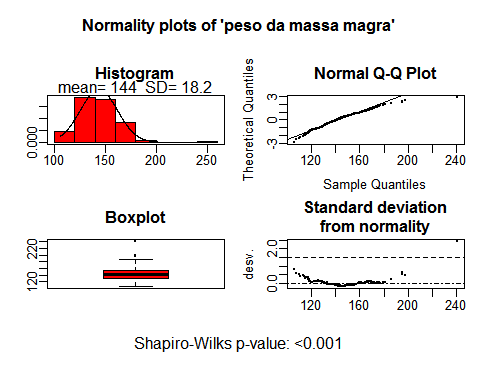
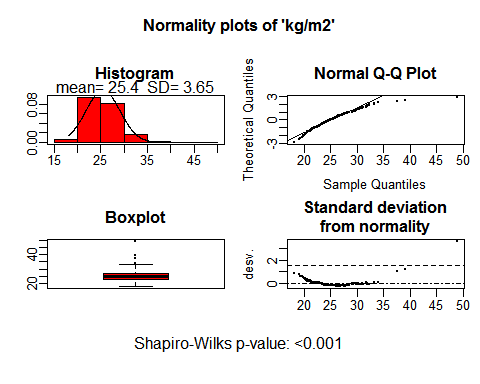
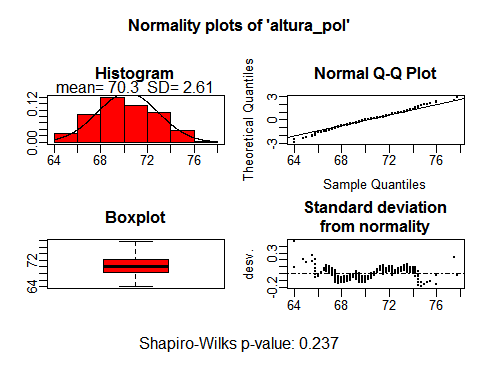
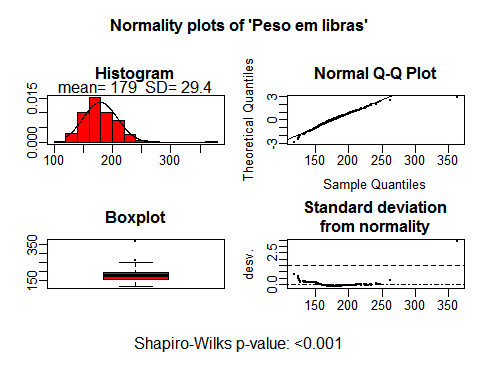
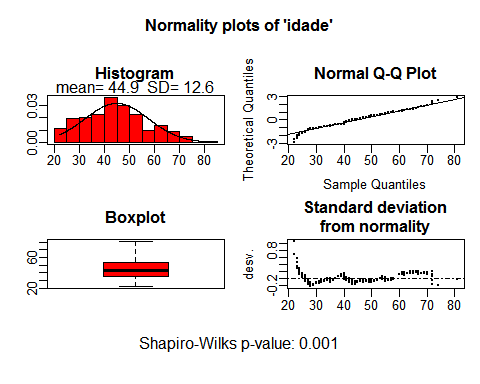
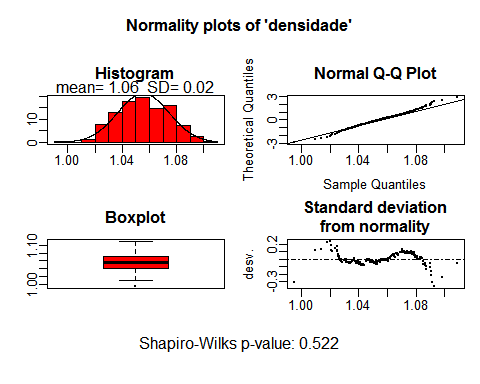
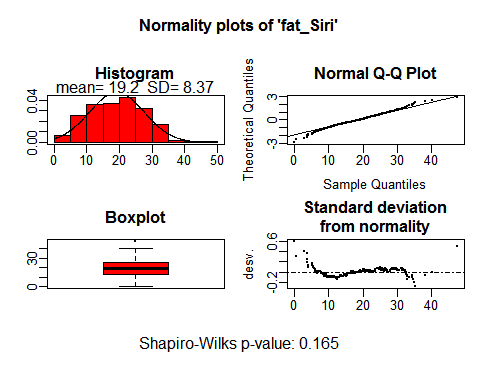
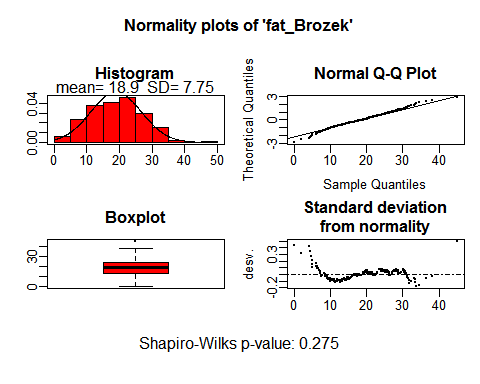
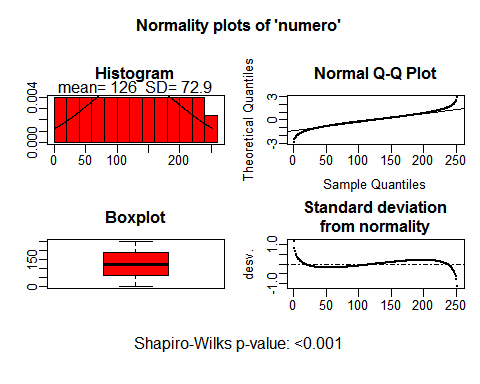
# Altura  
  
fat\_dat$altura\_pol <- ifelse( fat\_dat$numero == 42, 69.5, fat\_dat$altura\_pol)  
  
# Densidades  
  
fat\_dat$densidade <- ifelse( fat\_dat$numero == 48, 1.0865, fat\_dat$densidade)  
fat\_dat$densidade <- ifelse( fat\_dat$numero == 76, 1.0566, fat\_dat$densidade)  
fat\_dat$densidade <- ifelse( fat\_dat$numero == 96, 1.0591, fat\_dat$densidade)

compare\_fat <- compareGroups( ~ ., data = fat\_dat)  
  
summary(compare\_fat)

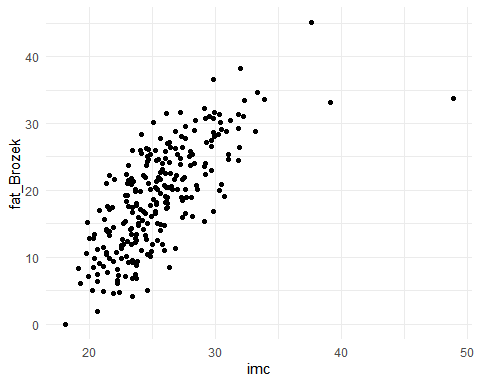
##   
## --- Descriptives of each row-variable ---  
##   
## -------------------   
## row-variable: numero   
##   
## N mean sd lower upper   
## [ALL] 252 126.5 72.89033 117.4569 135.5431  
##   
## -------------------   
## row-variable: fat\_Brozek   
##   
## N mean sd lower upper   
## [ALL] 252 18.93849 7.750856 17.97689 19.9001  
##   
## -------------------   
## row-variable: fat\_Siri   
##   
## N mean sd lower upper   
## [ALL] 252 19.15079 8.36874 18.11253 20.18906  
##   
## -------------------   
## row-variable: densidade   
##   
## N mean sd lower upper   
## [ALL] 252 1.055455 0.018909 1.053109 1.057801  
##   
## -------------------   
## row-variable: idade   
##   
## N mean sd lower upper   
## [ALL] 252 44.88492 12.60204 43.32146 46.44838  
##   
## -------------------   
## row-variable: Peso em libras   
##   
## N mean sd lower upper   
## [ALL] 252 178.9244 29.38916 175.2783 182.5706  
##   
## -------------------   
## row-variable: altura\_pol   
##   
## N mean sd lower upper   
## [ALL] 252 70.30754 2.609583 69.98378 70.6313  
##   
## -------------------   
## row-variable: kg/m2   
##   
## N mean sd lower upper   
## [ALL] 252 25.4369 3.648111 24.9843 25.88951  
##   
## -------------------   
## row-variable: peso da massa magra   
##   
## N mean sd lower upper   
## [ALL] 252 143.7139 18.23164 141.452 145.9758  
##   
## -------------------   
## row-variable: Circunferencia do pescoço   
##   
## N mean sd lower upper   
## [ALL] 252 37.99206 2.430913 37.69047 38.29365  
##   
## -------------------   
## row-variable: Circunferencia do peito   
##   
## N mean sd lower upper   
## [ALL] 252 100.8242 8.430476 99.77829 101.8701  
##   
## -------------------   
## row-variable: Circ do abdomem   
##   
## N mean sd lower upper   
## [ALL] 252 92.55595 10.78308 91.21816 93.89375  
##   
## -------------------   
## row-variable: circ do quadril   
##   
## N mean sd lower upper   
## [ALL] 252 99.90476 7.164058 99.01596 100.7936  
##   
## -------------------   
## row-variable: circ do coxa   
##   
## N mean sd lower upper   
## [ALL] 252 59.40595 5.249952 58.75462 60.05728  
##   
## -------------------   
## row-variable: circ do joelho   
##   
## N mean sd lower upper   
## [ALL] 252 38.59048 2.411805 38.29126 38.8897  
##   
## -------------------   
## row-variable: circ do tornozelo   
##   
## N mean sd lower upper   
## [ALL] 252 23.10238 1.694893 22.89211 23.31266  
##   
## -------------------   
## row-variable: circ do biceps   
##   
## N mean sd lower upper   
## [ALL] 252 32.27341 3.021274 31.89858 32.64825  
##   
## -------------------   
## row-variable: circ do antebraco   
##   
## N mean sd lower upper   
## [ALL] 252 28.66389 2.020691 28.41319 28.91458  
##   
## -------------------   
## row-variable: circ do pulso   
##   
## N mean sd lower upper   
## [ALL] 252 18.22976 0.933585 18.11394 18.34559

tabela\_fat <- createTable(compare\_fat)

plot(compare\_fat)



disp\_fat\_dat <- ggplot(fat\_dat, aes(x = imc, y = fat\_Brozek))+  
 geom\_point() +  
 theme\_minimal()  
  
disp\_fat\_dat



lm\_imc\_brozek <- lm(formula = fat\_Brozek ~ imc,  
 data = fat\_dat)  
  
summary(lm\_imc\_brozek)

##   
## Call:  
## lm(formula = fat\_Brozek ~ imc, data = fat\_dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -21.4292 -3.4478 0.2113 3.8663 11.7826   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -20.40508 2.36723 -8.62 7.78e-16 \*\*\*  
## imc 1.54671 0.09212 16.79 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.324 on 250 degrees of freedom  
## Multiple R-squared: 0.53, Adjusted R-squared: 0.5281   
## F-statistic: 281.9 on 1 and 250 DF, p-value: < 2.2e-16

confint(lm\_imc\_brozek)

## 2.5 % 97.5 %  
## (Intercept) -25.067331 -15.74283  
## imc 1.365275 1.72815

A cada aumento de 1 unidade no IMC ocorre aumento de 1.55% na gordura corporal. O intervalo de confiança é de 1.36 a 1.73.