

NUOVO CODICE:

```
require(ggplot2)
require(dplyr)
require(datasets)
data(ToothGrowth)

mydfs_VC <- list ()
mydfs_OJ <- list ()

k <- 1:10
for (i in 1:3) {
  mydfs_VC[[i]] <- ToothGrowth[k,]
  names(mydfs_VC[[i]]) <- c(paste("len", "VC", i, sep=""), paste("supp", "VC", i, sep=""),
    paste("dose", "VC", i, sep=""))
  k <- k+10
}
myToothGrowth_VC <- bind_cols(mydfs_VC)

j <- 31:40
for (i in 1:3) {
  mydfs_OJ[[i]] <- ToothGrowth[j,]
  names(mydfs_OJ[[i]]) <- c(paste("len", "OJ", i, sep=""), paste("supp", "OJ", i, sep=""),
    paste("dose", "OJ", i, sep=""))
  j <- j+10
}
myToothGrowth_OJ <- bind_cols(mydfs_OJ)

#t test su dose 1 e dose 3 VC
test1 <- t.test(x = myToothGrowth_VC$lenVC3, y = myToothGrowth_VC$lenVC1, paired = TRUE,
  alternative = "greater")

#t test su dose 1 e dose 2 VC
t.test(x = myToothGrowth_VC$lenVC2, y = myToothGrowth_VC$lenVC1, paired = TRUE,
  alternative = "greater")

#t test su dose 3 e dose 2 VC
t.test(x = myToothGrowth_VC$lenVC3, y = myToothGrowth_VC$lenVC2, paired = TRUE,
  alternative = "greater")

#t test su dose 1 e dose 3 OJ
t.test(x = myToothGrowth_OJ$lenOJ3, y = myToothGrowth_OJ$lenOJ1, paired = TRUE, alternative
  = "greater")

#t test su dose 1 e dose 2 OJ
t.test(x = myToothGrowth_OJ$lenOJ2, y = myToothGrowth_OJ$lenOJ1, paired = TRUE, alternative
  = "greater")
```

```
#t test su dose 3 e dose 2 OJ
t.test(x = myToothGrowth_OJ$lenOJ3, y = myToothGrowth_OJ$lenOJ2, paired = TRUE, alternative = "greater")
```

```
#t test su dose 1 e dose 1 differenti succhi
t.test(x = myToothGrowth_VC$lenVC1, y = myToothGrowth_OJ$lenOJ1, paired = TRUE, alternative = "two.sided")
```

```
#t test su dose 2 e dose 2 differenti succhi
t.test(x = myToothGrowth_VC$lenVC2, y = myToothGrowth_OJ$lenOJ2, paired = TRUE, alternative = "two.sided")
```

```
#t test su dose 3 e dose 3 differenti succhi
t.test(x = myToothGrowth_VC$lenVC3, y = myToothGrowth_OJ$lenOJ3, paired = TRUE, alternative = "two.sided")
```

```
#provo a fare un grafico del test 1
t1 <- data.frame(vals = rt(n = 100000, df = 9))
ggplot(t1, aes(x = vals)) +
  geom_density(size = 1) +
  geom_vline(xintercept = c(test1$statistic, qt(p = 0.975, df = 9, lower.tail = TRUE)), colour = c("blue", "red"), size = 1)
```

```
#qualcosa non mi quadra dal grafico, legato all'intervallo superiore infinito
#altra cosa: devo ragionare bene sul data wrangling, ci sono molti principi da applicare
```

```
#t test casereccio per verificare int di confidenza
tcrit <- qt(p = 0.975, df = 9)
mmedia_camp <- mean(myToothGrowth_VC$lenVC3-myToothGrowth_VC$lenVC1)
sd_camp <- sd(myToothGrowth_VC$lenVC3-myToothGrowth_VC$lenVC1)/sqrt(10)
```

```
diffVC3.1 <- myToothGrowth_VC$lenVC3 - myToothGrowth_VC$lenVC1
ggplot(data=myToothGrowth_VC, aes(y= diffVC3.1, x = 1)) +
  geom_boxplot()
```

```
diffVC3.2 <- myToothGrowth_VC$lenVC3 - myToothGrowth_VC$lenVC2
ggplot(data=myToothGrowth_VC, aes(y= diffVC3.2, x = 1)) +
  geom_boxplot()
```

```
diffVC2.1 <- myToothGrowth_VC$lenVC2 - myToothGrowth_VC$lenVC1
ggplot(data=myToothGrowth_VC, aes(y= diffVC2.1, x = 1)) +
  geom_boxplot()
```

Da: Americo Costantini

Inviato: lunedì 20 luglio 2015 14:03

A: 'ame.costantini@gmail.com'

Oggetto: codice r t test e data wrangling

```

require(ggplot2)
require(dplyr)
require(datasets)
data(ToothGrowth)

mydfs_VC <- list ()
mydfs_OJ <- list ()

k <- 1:10
for (i in 1:3) {
  mydfs_VC[[i]] <- ToothGrowth[k,]
  names(mydfs_VC[[i]]) <- c(paste("len","VC", i,sep=""), paste("supp","VC", i,sep=""),
    paste("dose","VC", i,sep=""))
  k <- k+10
}
myToothGrowth_VC <- bind_cols(mydfs_VC)

j <- 31:40
for (i in 1:3) {
  mydfs_OJ[[i]] <- ToothGrowth[j,]
  names(mydfs_OJ[[i]]) <- c(paste("len","OJ", i,sep=""), paste("supp","OJ", i,sep=""),
    paste("dose","OJ", i,sep=""))
  j <- j+10
}
myToothGrowth_OJ <- bind_cols(mydfs_OJ)

#t test su dose 1 e dose 3 VC
test1 <- t.test(x = myToothGrowth_VC$lenVC3, y = myToothGrowth_VC$lenVC1, paired = TRUE,
  alternative = "greater")

#t test su dose 1 e dose 2 VC
t.test(x = myToothGrowth_VC$lenVC2, y = myToothGrowth_VC$lenVC1, paired = TRUE,
  alternative = "greater")

#t test su dose 3 e dose 2 VC
t.test(x = myToothGrowth_VC$lenVC3, y = myToothGrowth_VC$lenVC2, paired = TRUE,
  alternative = "greater")

#t test su dose 1 e dose 3 OJ
t.test(x = myToothGrowth_OJ$lenOJ3, y = myToothGrowth_OJ$lenOJ1, paired = TRUE, alternative
  = "greater")

#t test su dose 1 e dose 2 OJ
t.test(x = myToothGrowth_OJ$lenOJ2, y = myToothGrowth_OJ$lenOJ1, paired = TRUE, alternative
  = "greater")

#t test su dose 3 e dose 2 OJ
t.test(x = myToothGrowth_OJ$lenOJ3, y = myToothGrowth_OJ$lenOJ2, paired = TRUE, alternative
  = "greater")

#t test su dose 1 e dose 1 differenti succhi
t.test(x = myToothGrowth_VC$lenVC1, y = myToothGrowth_OJ$lenOJ1, paired = TRUE,
  alternative = "two.sided")

```

```
#t test su dose 2 e dose 2 differenti succhi
t.test(x = myToothGrowth_VC$lenVC2, y = myToothGrowth_OJ$lenOJ2, paired = TRUE,
alternative = "two.sided")

#t test su dose 3 e dose 3 differenti succhi
t.test(x = myToothGrowth_VC$lenVC3, y = myToothGrowth_OJ$lenOJ3, paired = TRUE,
alternative = "two.sided")

#provo a fare un grafico del test 1
t1 <- data.frame(vals = rt(n = 100000, df = 8))
ggplot(t1, aes(x = vals)) +
  geom_density() +
  geom_vline(xintercept = c(test1$statistic, qt(p = 0.975, df = 8, lower.tail = TRUE)), colour =
c("blue", "red"))

#qualcosa non mi quadra dal grafico, legato all'intervallo superiore infinito
#altra cosa: devo ragionare bene sul data wrangling, ci sono molti principi da applicare
```

Americo Costantini

ConTe.it - EUI Limited

Antifraud Department

Via Della Bufalotta, 374/F

00139 ROMA

americo.costantini@conte.it

Tel. [+39.06.834.60.169](tel:+39.06.834.60.169)

Le informazioni contenute in questo messaggio e-mail sono destinate esclusivamente alla persona o entità cui è stato indirizzato e può contenere informazioni riservate. Ogni reinvio, diffusione o altro utilizzo o qualunque azione basata sul presente messaggio da parte di persone o entità diverse dal destinatario, se non espressamente autorizzate dal mittente, sono proibiti. Se avete ricevuto questo messaggio per errore, siete pregati di contattare immediatamente il mittente via e-mail e di cancellarlo da ogni vostro computer.

The information transmitted in this e-mail message is intended only for the person or entity to which it is addressed and may contain confidential information. Any retransmission, dissemination or other use of, or taking of any action in reliance upon, this information by person or entity other than the intended recipient, if not clearly authorized by the sender, is prohibited. If you have received this communication in error, please notify the sender immediately by e-mail and delete the message from any computer.

This email has been scanned by the Symantec Email Security.cloud service.
For more information please visit <http://www.symanteccloud.com>
