correlations

Code ▼

```
Hide
library(boot)
library(ggm)
library(Hmisc)
library(ggplot2)
library(polycor)
                                                                              Hide
# we have cor() , cor.test() which are part of base R system, while rcorr()
which is a part of Hmisc package.
data<-read.delim('/home/atrides/Desktop/R/statistics_with_R/06/Data_Files/Ex
am Anxiety.dat',header=TRUE)
head(data)
                                                                              Hide
is.factor(data$Gender)
                                                                              Hide
data$Gender<-factor(data$Gender)</pre>
data$GenderX<-as.numeric(data$Gender)</pre>
is.factor(data$Gender)
                                                                              Hide
# Using cor()
```

```
# Using cor()
# pearson with complete.obs
cor(data[,cbind('Code','Revise', 'Exam','Anxiety','GenderX')], use='complet
e.obs', method='pearson')
```

Hide

```
# kendall with complete.obs
cor(data[,cbind('Code','Revise', 'Exam','Anxiety','GenderX')], use='complet
e.obs', method='kendall')
```

Hide

1 of 2 24/08/20, 4:58 pm

```
# kendall with pairwise.complete.obs
cor(data[,cbind('Code','Revise', 'Exam','Anxiety','GenderX')], use='pairwis
e.complete.obs', method='kendall')
                                                                            Hide
# using rcorr(), from Hmisc . This function does not work on dataframes, con
vert it into a matrix
# Note: rcorr() works only on matrix and gives, pearson and spearman only an
d use pairwise exclusion of missing values
dm <- data.matrix(data[,cbind('Code','Revise', 'Exam','Anxiety', 'Gender')])</pre>
rcorr(dm, type=c('pearson'))
                                                                            Hide
# using cor.test(), only does one pair of variables at a time
cor.test(data$Exam, data$Anxiety, alternative = 'two.sided', method='pearson
', conf.level = 0.95)
# Note:
# use pearson r for parametric
# use spearman rho for non-parametric
# use tendall tau for non-parametric and small datasets
                                                                           Hide
# Now , if data is non-normal , we would use bootstrapping correaltion
library(boot)
data<-read.delim('/home/atrides/Desktop/R/statistics with R/06/Data Files/Th
e Biggest Liar.dat',header=TRUE)
                                                                            Hide
bootTau<-function(data, i)cor(data$Position[i], data$Creativity[i], use='com
plete.obs', method='kendall')
boot kendall<-boot(data, bootTau, 2000)</pre>
boot kendall
                                                                           Hide
# 95% confidence interval
boot.ci(boot kendall)
```

2 of 2 24/08/20, 4:58 pm