

Formulario di R per FSSB, AA 2016/17, Epifani

- **getwd()** : restituisce la cartella di lavoro
- **dir()** : mostra il contenuto della cartella di lavoro
- **ls()** : mostra le variabili presenti nell'area di lavoro
- **rm(list=ls())** : ripulisce area di lavoro
- **d = read.table("nomedataframe.txt", header=TRUE)**: legge il file di dati nomedataframe.txt considerando l'intestazione (header) e salvandolo nella variabile d come un data frame
- **write.table(d, "nomedataframe.txt")**: crea un file di testo che contiene il dataframe (o la matrice o il vettore) e lo salva nella cartella di lavoro corrente.
- **data(USArrests)**: carica dataset (nell'esempio USArrest) o liste di dataset
- **model.matrix(object)**: crea una design matrix
- **attach(d)**: consente di accedere direttamente alle variabili contenute nel dataframe d (nomevariabile invece di d\$nomevariabile).
- **detach(d)**: impedisce l'accesso diretto alle variabili contenute nel dataframe d.
- **names(d)**: restituisce il nome delle variabili contenute in d.
- **is.na(x)**: restituisce un vettore logico: valori TRUE in corrispondenza degli NA (valori non disponibili) contenuti in x.
- **length(x)**: restituisce il numero di elementi contenuti in x.
- **dim(d)**: dimensioni (num. righe \times num. colonne) del dataframe (o della matrice) d.
- **str(d)**: permette di ottenere informazioni sul contenuto del dataframe (o matrice o vettore) d.
- **as.numeric(object)** Crea o costringe oggetti di tipo "numerico"
- **cat = as.factor(x)**: trasforma il vettore x in una variabile categorica chiamata cat.
- **cat = factor(x = character(), levels, labels = levels)**: codifica un vettore come un fattore (categoria)
- **levels(cat)**: mostra quanti livelli (categorie) sono presenti all'interno della variabile cat
- **tapply(x,y,function)**: divide i dati in x in sottogruppi ottenuti sulla base dei valori in y e applica la funzione function a ciascun sottogruppo.
- **apply(Data,MARGIN,function)**: applica la funzione function a ciascun sottoinsieme della tabella (o matrice) Data individuato da MARGIN
- **c(1,2)** o **c(x,y)**: permette di concatenare gli elementi 1 e 2 o i vettori x e y in un singolo vettore.
- **rep(n,m)**: (n e m scalari) costruisce un vettore formato dall'elemento n ripetuto m volte.
- **rep(v, u)**: (v e u vettori della stessa lunghezza) costruisce un vettore formato dalla ripetizione di ogni elemento di v un numero di volte pari al corrispondente elemento di u.
- **seq(a, b, by=passo)**: costruisce un vettore di elementi equispaziati da "a" a "b" con passo "by".
- **seq(a, b, length=L)**: costruisce un vettore di elementi equispaziati di lunghezza L.
- **unique(x)**: estrae i valori distinti di un vettore x.
- **matrix(dati, nrow=n, ncol=m, byrow=F)**: costruisce una matrice $n \times m$ riempiendola per colonne con i valori contenuti nel vettore dati.
- **cbind(w,v,u)**: costruisce una matrice unendo le colonne definite dai vettori w,v,u.
- **rbind(w,v,u)**: costruisce una matrice unendo le righe definite dai vettori w,v,u.
- **sort(x)**ordina un vettore x in ordine ascendente o discendente
- **table(cat)**: calcola le frequenze assolute dei dati in x
- **prop.table(table(cat))**: calcola le frequenze relative dei dati in x.
- **sum(x)**: somma tutte le componenti del vettore (o matrice) x.
- **prod(x)**: esegue il prodotto tra componenti del vettore (o matrice) x.
- **diff(x)**: esegue la differenza tra componenti del vettore x.
- **colSums(m)**: esegue la somma degli elementi contenuti in ciascuna colonna di una matrice m.
- **rowSums(m)**: esegue la somma degli elementi contenuti in ciascuna riga di una matrice m.
- **mean(x)**: media dei dati contenuti in x.
- **median(x)**: mediana dei dati contenuti in x.
- **min(x)**: minimo dei dati contenuti in x.
- **max(x)**: massimo dei dati contenuti in x.

- **quantile(x, probs=p)**: quantile di ordine p dei dati contenuti in x .
- **summary(object)**: mostra i parametri fondamentali dell'oggetto x dato come argomento.
- **range(x)**: restituisce un vettore di due elementi che rappresentano il minimo e il massimo di x .
- **sd(x)**: deviazione standard dei dati contenuti in x .
- **var(x)**: varianza dei dati contenuti in x .
- **cor(x,y)**: coefficiente campionario di correlazione di x e y .
- **cov(x,y)**: covarianza campionaria di x,y .
- **par (mfrow=c(nrighes,ncolonne))**: permette di creare un'interfaccia grafica specificando il numero di righe e di colonne (quindi il numero di grafici che si vogliono fare).
- **dev.off()**: ripristina l'interfaccia grafica a quella di default.
- **barplot(table(cat))**: costruisce il diagramma a barre dei dati in cat (frequenze assolute).
- **barplot(prop.table(table(cat)))**: costruisce il diagramma a barre dei dati in cat (frequenze relative).
- **pie(table(cat))**: costruisce il diagramma a torta dei dati in cat .
- **plot(x,y)**: rappresenta i punti le cui coordinate sono contenute nei vettori x e y .
- **pairs(dataframe)**: produce una matrice di tutti gli scatterplot in $dataframe$.
- **plot(x,y,type='l')**: costruisce un grafico collegando i punti (le cui coordinate sono contenute nei vettori x e y) con delle linee.
- **points(x,y)**: sovrappone il nuovo grafico dei punti (x,y) ad un grafico precedente (senza cancellarlo).
- **abline(v=2)**: sovrappone ad un grafico precedente la retta verticale $x=2$.
- **abline(h=-1)**: sovrappone ad un grafico precedente la retta orizzontale $y=-1$.
- **abline(a=3,b=-0.5)**: sovrappone ad un grafico precedente la retta $y=-0.5x+3$.
- **hist(x,prob=TRUE)**: realizza un istogramma con le frequenze relative dei dati in x .
- **boxplot(x)**: realizza un boxplot dei dati in x .
- **boxplot(x ~ y)**: realizza un boxplot dei dati in x distinti a seconda delle categorie di y .
- **qqnorm(x)**: costruisce il qqplot con i quantili empirici dei dati in x e i quantili teorici della fdr normale standard.
- **qqline(x)**: aggiunge una linea a un quantile-quantile plot normale.
- **image(object)**: crea una griglia di rettangoli colorati in scala in base ai valori dei quantili z . È utile per rappresentare dati 3D.
- **regress = lm(y ~ x)**: realizza una regressione lineare di y su x (modello $y = \alpha + \beta x + e$) e salva i risultati nella variabile `regress`.
- **abline(regress)** : sovrappone a un grafico precedentemente realizzato la retta ottenuta tramite la regressione lineare.
- **update(regress)**: aggiorna e rifitta il modello di regressione `regress`.
- **anova(object)**: esegue l'Analisi della Varianza per uno o più modelli fittati, seleziona il modello in base al p -value.
- **step(object, direction = "backward")**: seleziona il modello in base all'AIC.
- **stdres(object)** {package MASS}: crea residui standardizzati.
- **shapiro.test()**: calcola un test di Shapiro di normalità.
- **hatvalues(model)**: calcola alcune diagnostiche della regressione (leave-one-out deletion) lineare e generalizzata.
- **predict(regress, x.new, interval="confidence", level = 0.91)**: calcola l'intervallo di confidenza (di livello 91%) per la media della previsione del nuovo dato contenuto in $x.new$ secondo il modello definito in `regress`.
- **round(x,c)**: approssima x alla c -esima cifra decimale.
- **which(x=="c")**: seleziona gli indici di posizione del vettore x in corrispondenza dei valori di x uguali a c .
- **IQR(x)**: calcola il range interquartile di x .