Proposta comandi per il Report

Descrittiva

Descrittiva Univariata

* Distribuzioni di frequenza e indici di posizione
* table(X) -> Frequenze assolute
* table(X)/length(X)\*100 -> Frequenze relative [1]
* mean(X)
* median(X)
* sd(X)
* Grafici
* pie(table(X), col = “”) -> Grafico a torta
* barplot(table(X)/length(X)\*100, col= “”) -> Grafico a barre [1]
* box plot (X)
* hist (X)

Dopo aver scelto alcune variabili da approfondire, con i comandi esposti sopra si possono ricavare le distribuzioni di frequenza (per confronti usare le relative) e gli indici di posizione. Quanto ottenuto può essere sintetizzato in tabelle e raffigurato con qualche grafico.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variabile | Frequenze (%) | Media | Mediana | … |
| X | … | … | … |  |

Esempio di sintesi univariata.

Descrittiva Bivariata

* Distribuzioni di frequenza e indici di posizione
* table(X, Y) -> Frequenze assolute delle due variabili[2.1]
* tabella.XY <-table(X,Y) -> indicizzo 2.1 [2.2]
* prop.table(tabella.XY, margin=2)\*100 -> Distribuzione condizionata con le F. Relative [2.3]
* tapply(X, Y, mean)
* tapply(X, Y, median)
* Primo sottoinsieme <-subset(a,subset = X == "x1") [3] \*
* Secondo sottoinsieme <-subset(a,subset = X == "x2") [4]
* Grafici
* barplot(prop.table(tabella.XY, 2), legend = T, beside = T, ylim= c(0,1.),col=””) -> Grafico [2.3]
* box plot (X ~ Y, data = a, xlab = "", ylab = "")
* barplot(table(factor(Primo sottoinsieme $ Y, levels= “0:10”))/sum(table(Primo sottoinsieme $ Y)) \*100) [3] \*\*
* barplot(table(factor(Secondo sottoinsieme $ Y, levels= “0:10”))/sum(table(Secondo sottoinsieme $ Y)) \*100) [4]

Nella Bivariata si possono incronciare due variabili in modo da ottenere distribuzioni condizionate da poter sintetizzare in tabelle o raffigurare tramite grafici a barre/box plot.

Si può anche utilizzare il comando “subset” per indagare la distribuzione tra una variabile (Y) e la modalità di un’altra variabile (x1, x2)

Descrittiva Multivariata

* tapply (X, list(Y,Z), mean)
* tapply (X, list(Y,Z), median)

Per indagare tre variabili si può utilizzare il comando tapply.

I dati ottenuti possono essere sintetizzati in matrici

Legenda

a -> Matrice dati utilizzata per il report

X -> Variabile

Y -> Variabile

x1-> Prima modalità variabile X

x2-> Seconda modalità variabile X

\*-> In questo caso “seleziono” la modalità (x1) della variabile (X)

\*\*-> In questo esempio Y è una variabile quantitativa avente 10 valori (0:10)