

Matedit

*Un software per l'elaborazione delle matrici nell'ambito dell'analisi
multivariata dei dati*

Introduzione

WinMatedit è un ambiente software integrato, sviluppato per agevolare la manipolazione di dati in forma matriciale e per effettuare la loro elaborazione attraverso l'applicazione di strumenti di analisi statistica multivariata. Esso costituisce la trasposizione e l'ulteriore evoluzione in ambiente Microsoft Windows del pacchetto "Matedit" per DOS (Burba et al., 1992).

Il suo utilizzo consente di elaborare matrici di notevoli dimensioni utilizzando una serie completa di operatori elementari (combinando i quali si possono applicare le metodologie dell'analisi multivariata) e di automatizzarne l'esecuzione attraverso un opportuno linguaggio di "scripting".

Le principali elaborazioni possibili sono: classificazioni, ordinamenti, analisi delle componenti principali (PCA) e riordinamenti di tabelle. In particolare è possibile effettuare interattivamente la produzione di "Fuzzy sets" usando dendrogrammi di classificazioni gerarchiche colorando direttamente i loro rami. Questa tecnica innovativa permette di operare direttamente a diversi livelli di scala e di sintesi della conoscenza della struttura dei dati.

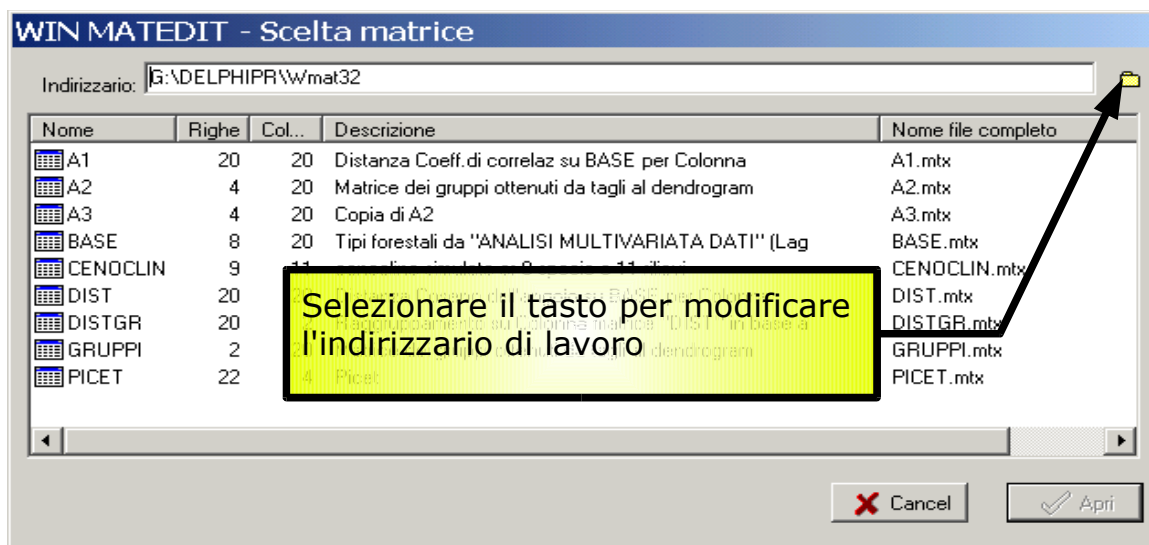
Per fornire ampie possibilità di impiego, sono disponibili molti tipi di calcolo della distanza o della somiglianza da poter applicare alle righe o alle colonne delle tabelle da analizzare, alcuni dei quali sono ad esempio: il coefficiente di correlazione, il coefficiente di Gower (per tabelle con dati misti o mancanti) ed il coefficiente di Bray-Curtis.

MATEDIT consente di testare ipotesi attraverso simulazioni e test delle permutazioni, di importare ed esportare in maniera semplice ed interattiva (cliccando su un bottone) i dati di input e di output su OpenOffice Calc o Microsoft Excel. Inoltre è attiva un'interfaccia per l'interscambio dei dati con il software SNNS (Stuttgart Neural Network Simulator) che consente l'integrazione tra MATEDIT e le applicazioni delle reti neurali. Questo semplifica notevolmente non solo le operazioni di scambio dati, ma anche la creazione degli insiemi di "training" per le reti neurali, evitando di utilizzare filtri di importazione ed esportazione.

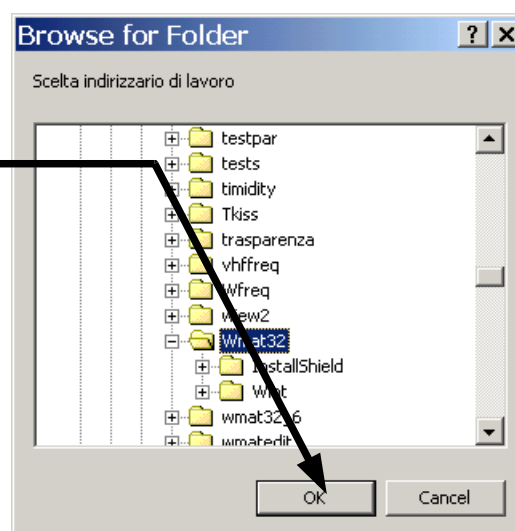
Tutto ciò fa sì che questo strumento, ideato inizialmente per agevolare la gestione e l'elaborazione di tabelle fitosociologiche, possa essere utilizzato in vari campi della ricerca ed abbia applicazione in tutti i settori che prevedono pianificazioni e programmi di piani di sviluppo.

1. Impostazione dell'indirizzario di lavoro

WinMatedit memorizza le matrici in files con suffisso MTX. La funzione di apertura punta ad una directory di lavoro che viene impostata come segue:



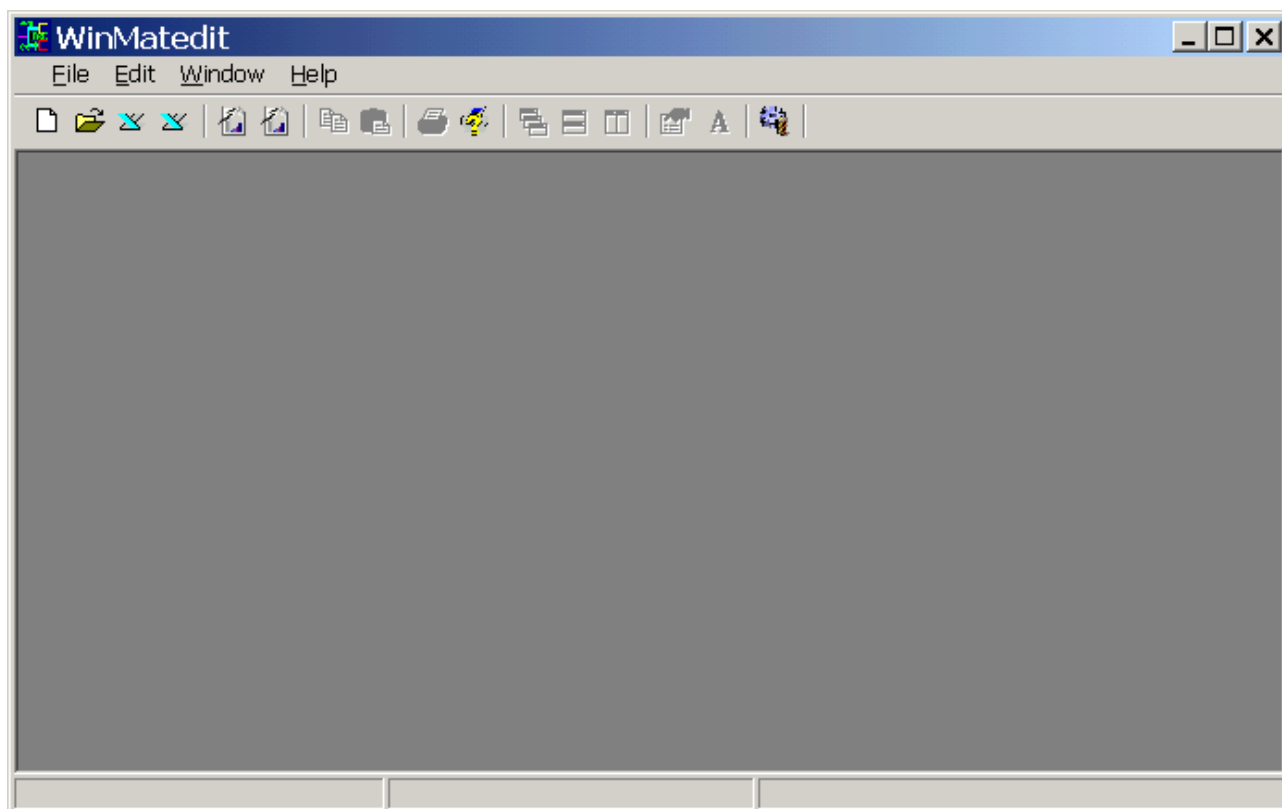
Selezionare il tasto OK dopo aver scelto l'indirizzario di lavoro



Poiché generalmente durante le elaborazioni viene utilizzato un grosso numero di matrici, è opportuno in fase di impostazione di un progetto, creare un indirizzario dedicato a tale progetto, in modo da facilitare l'individuazione dei files delle matrici.

2. Spazio di lavoro (desktop)

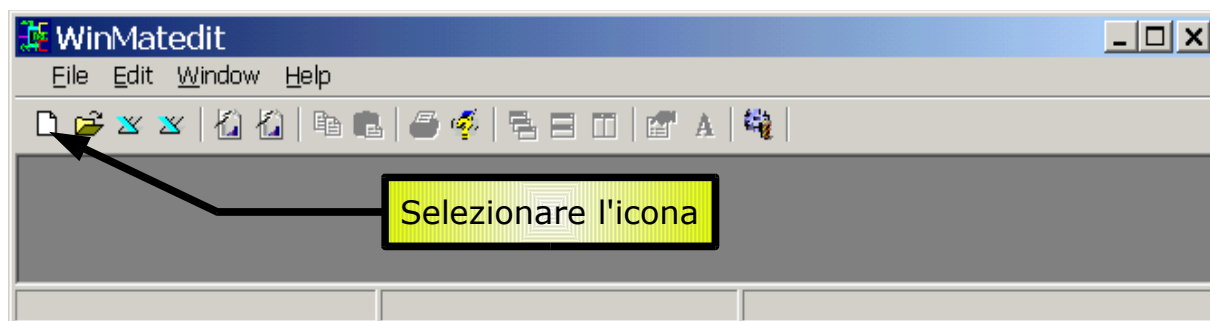
Appena viene messo in esecuzione WinMatedit presenta il suo spazio di lavoro.



Dopo aver selezionato l'indirizzario, per poter operare è necessario portare nello spazio di lavoro una o più matrici. Questo si ottiene creando una nuova matrice oppure aprendone una già esistente.

2.1 Creazione di una nuova matrice

Per creare una nuova matrice si devono effettuare i seguenti passaggi:



Selezionata l'icona come da figura, apparirà la seguente finestra:

The dialog box 'Creazione di una nuova matrice' contains the following fields and controls:

- Nome:** A text field containing 'G:\DELPHIPR\Wmat32'.
- Descrizione:** A text field containing 'Matrice di prova'.
- Dimensioni:** Two numeric input fields: 'Numero righe' with value '8' and 'Numero colonne' with value '20'.
- Tipologia:** Four radio buttons: 'Carattere', 'Intero', 'Intero piccolo', and 'Reale' (which is selected).
- Buttons:** 'Cancel' (with a red X icon) and 'OK' (with a green checkmark icon).

Labels on the left with arrows point to these elements:

- Nome** points to the 'Nome' field.
- Descrizione** points to the 'Descrizione' field.
- Numero colonne** points to the 'Numero colonne' field.
- Numero righe** points to the 'Numero righe' field.
- Tipo dati** points to the 'Tipologia' radio buttons.
- Conferma** points to the 'OK' button.

Dopo aver digitato nome, descrizione, numero colonne e righe, tipo di dati, premendo il tasto "OK" si otterrà la nuova matrice, ovviamente vuota, sullo spazio di lavoro.

The WinMatedit application window displays a new matrix titled 'NUOVA - Matrice di prova'. The matrix is an 8x8 grid of cells, each containing the value '0.00'. The first cell is highlighted with a blue border.

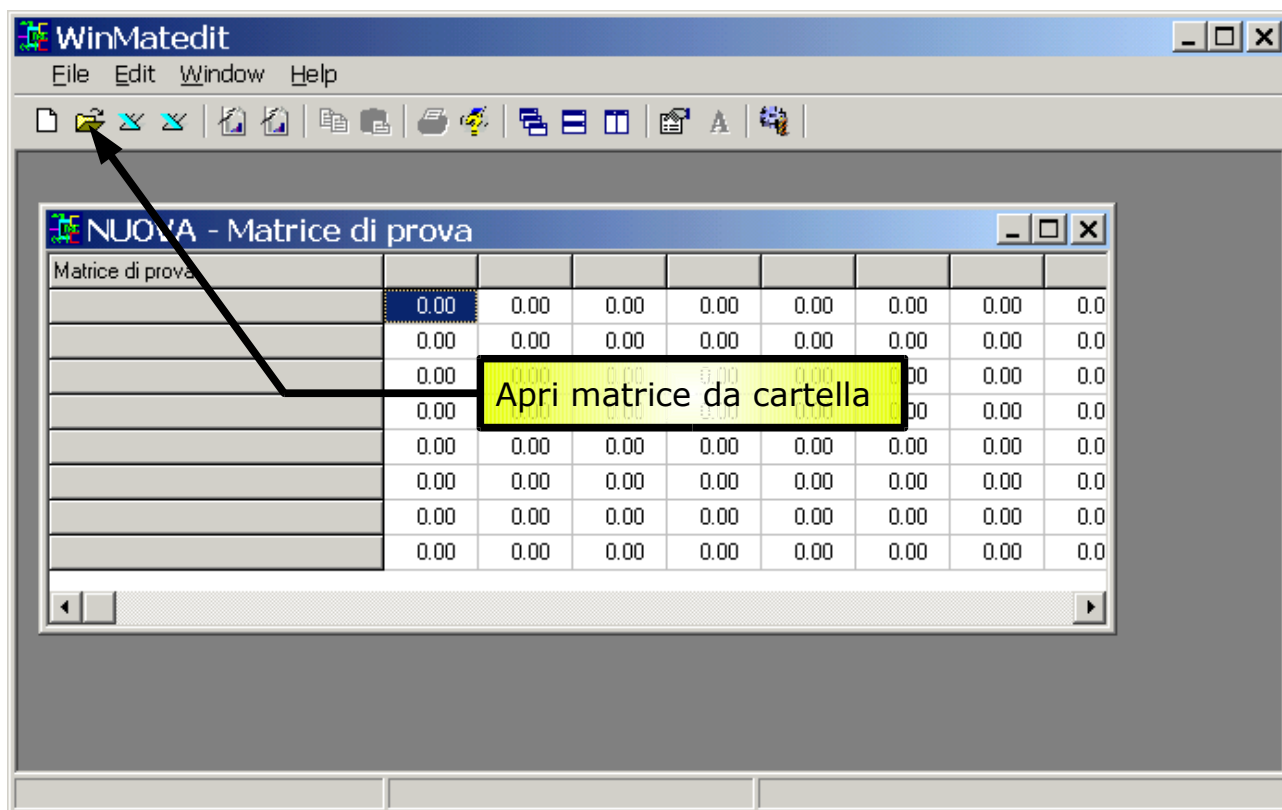
Matrice di prova								
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0

Naturalmente in tale matrice dovranno essere inseriti i dati da elaborare. L'inserimento dei dati può essere eseguito manualmente oppure tramite il

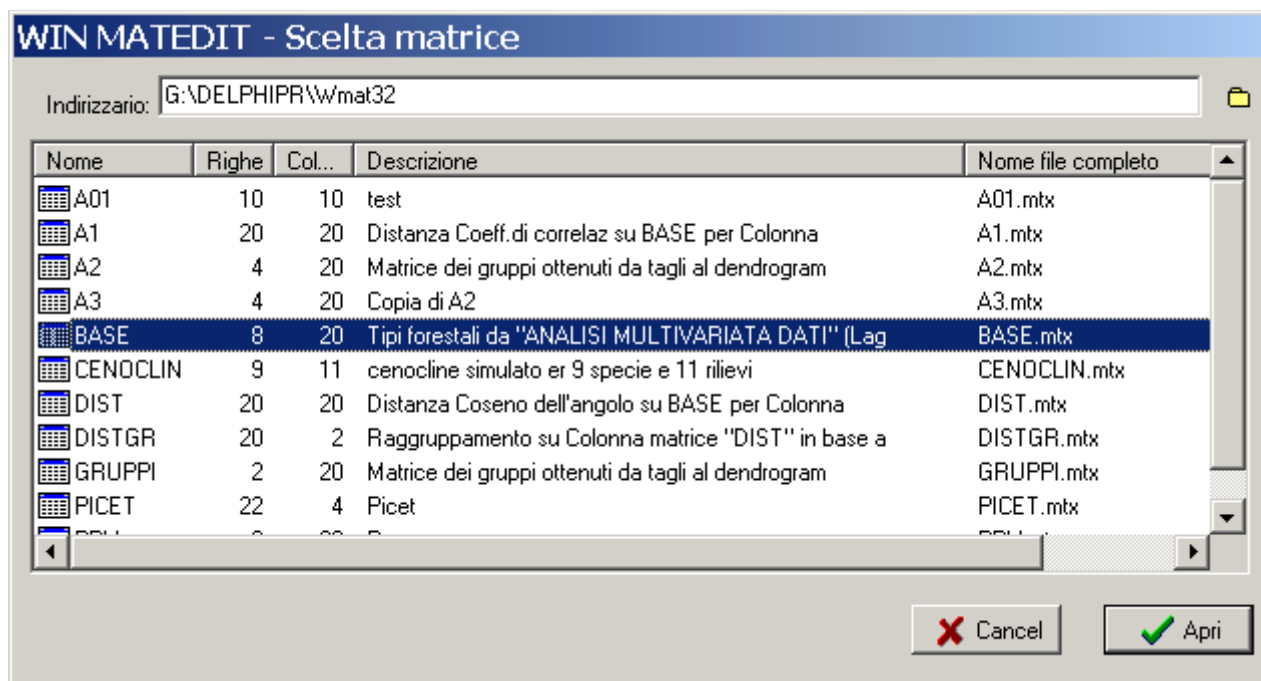
caricamento automatico di tabelle Excel o OpenOffice già pronte (vedi capitolo 3).

2.2 Apertura di una matrice già esistente

Bisogna procedere come segue:

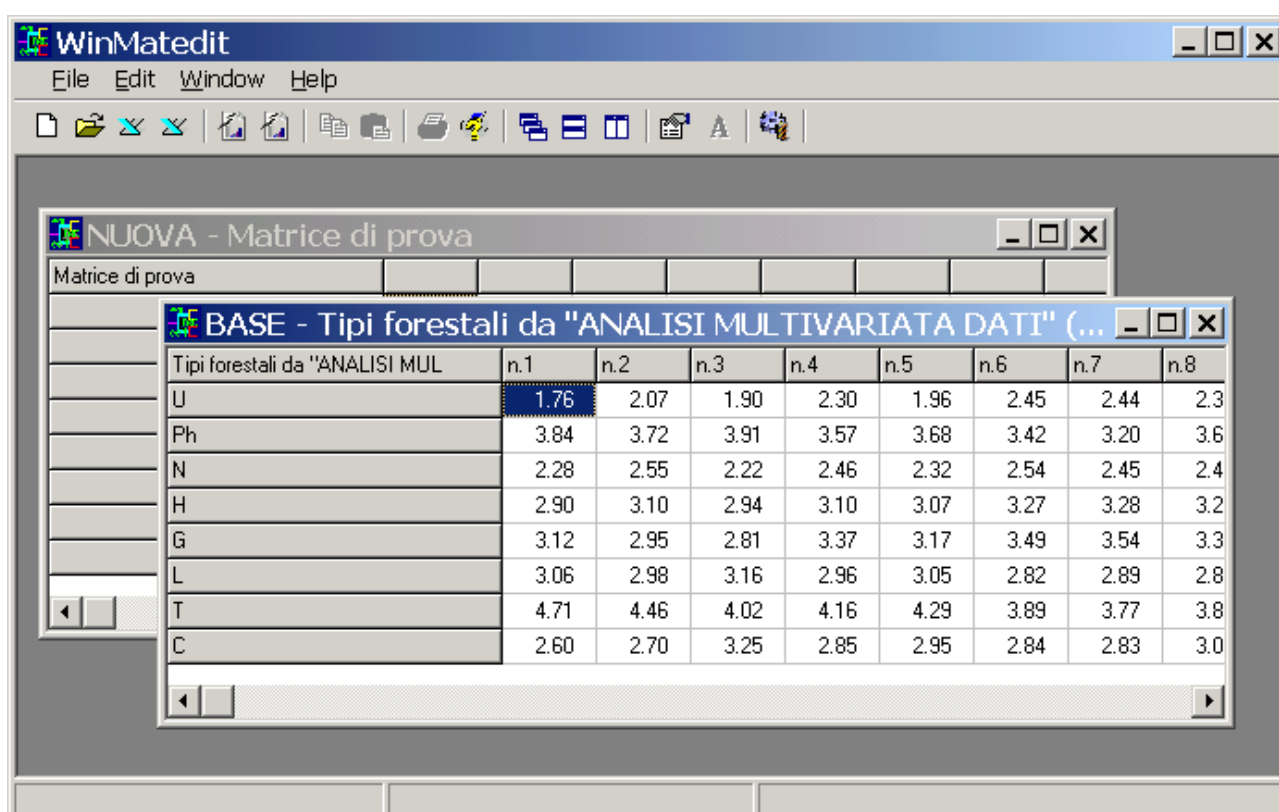


Verrà visualizzata la seguente finestra:



Qui si può scegliere la matrice (evidenziandola) e confermarne l'apertura, portandola nello spazio di lavoro.

Si otterrà:

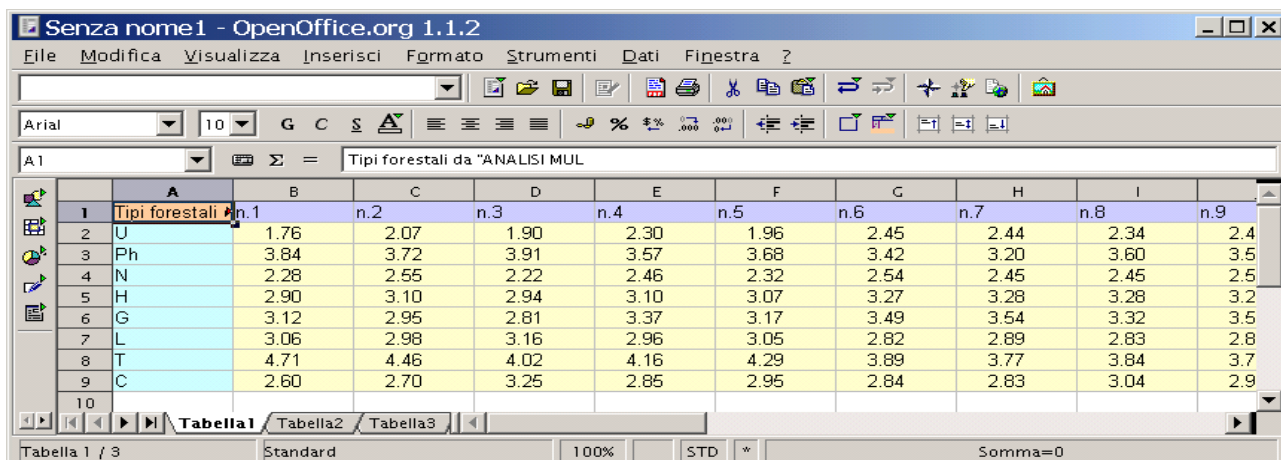


3. Caricamento automatico dei dati

Il caricamento automatico dei dati in una tabella vuota, avviene importando gli stessi da fogli elettronici (Microsoft Excel o "OpenOffice Calc"). Per un corretto caricamento i fogli devono avere le seguenti caratteristiche:

- la prima riga deve contenere le intestazioni delle colonne;
- la prima colonna deve contenere le intestazioni delle righe;
- la cella A1 può contenere una semplice descrizione dei dati della matrice.

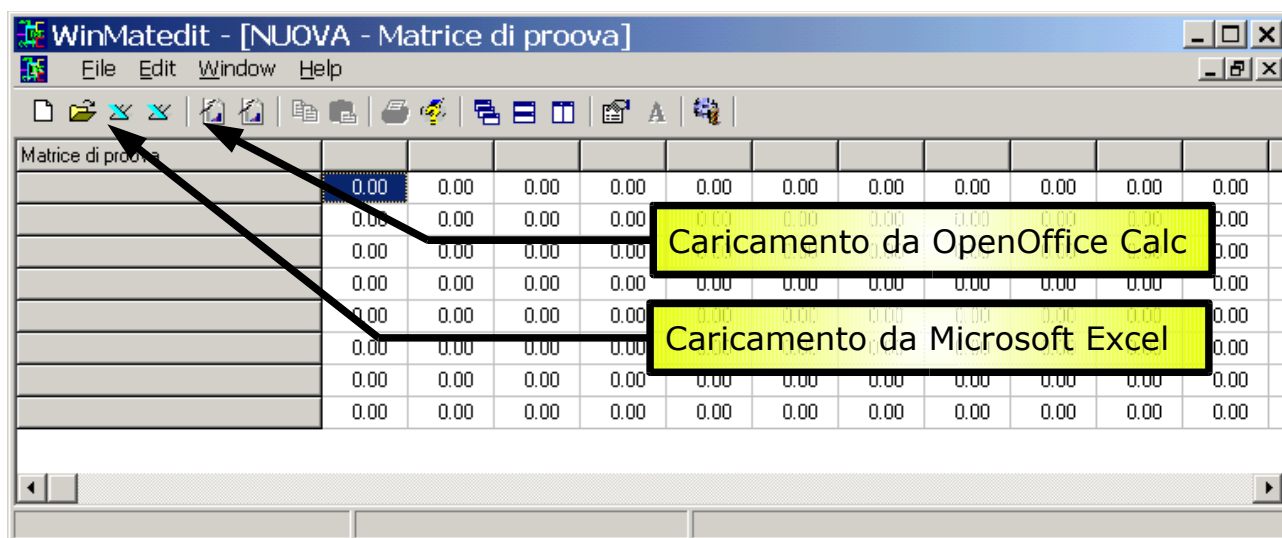
Ad esempio:



The screenshot shows the OpenOffice.org 1.1.2 interface. The spreadsheet has a title bar "Senza nome1 - OpenOffice.org 1.1.2". The menu bar includes File, Modifica, Visualizza, Inserisci, Formato, Strumenti, Dati, Finestra, and ?. The toolbar contains various icons for file operations and editing. The spreadsheet itself has a grid with columns labeled A through I and rows numbered 1 through 10. The first row (row 1) contains headers: "Tipi forestali" in column A, and "n.1" through "n.9" in columns B through I. The first column (column A) contains labels: "U", "Ph", "N", "H", "G", "L", "T", and "C" in rows 2 through 9. The cells from B2 to I9 contain numerical values. The status bar at the bottom shows "Tabella 1 / 3", "Standard", "100%", "STD", and "Somma=0".

	A	B	C	D	E	F	G	H	I	
1	Tipi forestali	n.1	n.2	n.3	n.4	n.5	n.6	n.7	n.8	n.9
2	U	1.76	2.07	1.90	2.30	1.96	2.45	2.44	2.34	2.4
3	Ph	3.84	3.72	3.91	3.57	3.68	3.42	3.20	3.60	3.5
4	N	2.28	2.55	2.22	2.46	2.32	2.54	2.45	2.45	2.5
5	H	2.90	3.10	2.94	3.10	3.07	3.27	3.28	3.28	3.2
6	G	3.12	2.95	2.81	3.37	3.17	3.49	3.54	3.32	3.5
7	L	3.06	2.98	3.16	2.96	3.05	2.82	2.89	2.83	2.8
8	T	4.71	4.46	4.02	4.16	4.29	3.89	3.77	3.84	3.7
9	C	2.60	2.70	3.25	2.85	2.95	2.84	2.83	3.04	2.9
10										

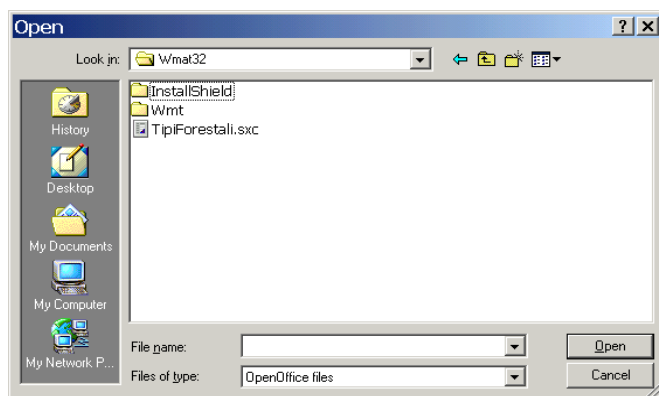
Per importare una matrice bisogna inizialmente creare una matrice vuota impostando le dimensioni volute e quindi utilizzare le icone predisposte per il caricamento:



A questo punto le operazioni da effettuare differiscono leggermente a seconda che venga utilizzato OpenOffice Calc oppure Microsoft Excel.

- OpenOffice Calc : dopo aver selezionato l'icona del caricamento indicata nella

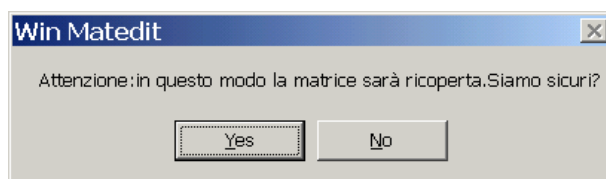
figura precedente, si aprirà un dialogo nel quale si deve indicare il nome del



file OpenOffice da caricare:

- Microsoft Excel : per prima cosa bisogna aprire con Microsoft Excel il foglio da importare, senza chiudere l'applicazione WinMatedit; quindi, dopo aver nuovamente focalizzato WinMatedit, si deve selezionare l'icona di importazione da Microsoft Excel.

In entrambi i casi apparirà la seguente richiesta di conferma:



Confermando otterremo l'importazione dei dati nella nostra matrice:

WinMatedit - [NUOVA - Matrice di prova]											
File Edit Window Help											
Tipi forestali da "ANALISI MUL	n.1	n.2	n.3	n.4	n.5	n.6	n.7	n.8	n.9	n.10	n.11
U	1.76	2.07	1.90	2.30	1.96	2.45	2.44	2.34	2.47	2.61	
Ph	3.84	3.72	3.91	3.57	3.68	3.42	3.20	3.60	3.52	3.33	
N	2.28	2.55	2.22	2.46	2.32	2.54	2.45	2.45	2.54	2.62	
H	2.90	3.10	2.94	3.10	3.07	3.27	3.28	3.28	3.26	3.37	
G	3.12	2.95	2.81	3.37	3.17	3.49	3.54	3.32	3.53	3.58	
L	3.06	2.98	3.16	2.96	3.05	2.82	2.89	2.83	2.80	2.66	
T	4.71	4.46	4.02	4.16	4.29	3.89	3.77	3.84	3.78	3.73	
C	2.60	2.70	3.25	2.85	2.95	2.84	2.83	3.04	2.90	2.84	

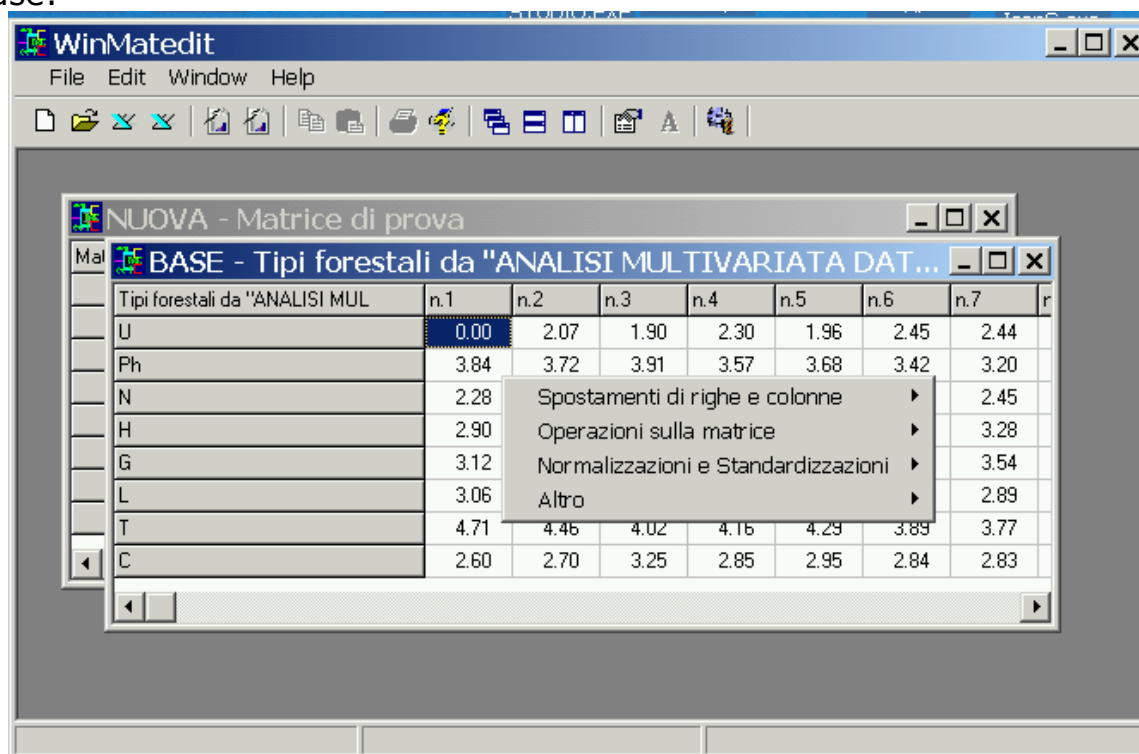
L'operazione di caricamento è così completata e la matrice è pronta per essere elaborata con le funzioni di WinMatedit.

Gli elementi delle matrici di tipo "Reale" possono essere valorizzati con "NULL" per rappresentare la mancanza del dato e possono contenere una lettera (A..Z e a..z) per descrivere un dato di tipo qualitativo (es: colore degli occhi B=Blu, V=Verdi, M=Marrone...); in questo modo possono essere trattate matrici "miste" per il calcolo della matrice di somiglianza di Gower.

4. Funzioni di MATEDIT (operazioni disponibili)

Per poter effettuare operazioni su una matrice, è necessario individuarla tra quelle aperte sullo spazio di lavoro (quella attiva presenta la striscia superiore di colore blu invece che grigio), oppure aprirla nel modo descritto in precedenza (vedi paragrafo 2.1).

Per visualizzare il menu che consente di accedere alle varie operazioni, ci si posiziona con il cursore sulla matrice aperta e si preme il tasto *destro* del mouse:



Le operazioni in elenco si attivano localizzando il puntatore sulla funzione desiderata e premendo il tasto destro del mouse. Appare un menù a più livelli le cui opzioni sono descritte di seguito in dettaglio.

Molte delle operazioni disponibili sono applicabili sia sulle righe che sulle colonne: in tal caso ci si limiterà a descrivere l'operazione in una delle due situazioni (righe o colonne), essendo l'opzione duale facilmente deducibile

applicando il ragionamento alla trasposta della matrice di partenza.

- Spostamenti di righe e di colonne
 - Permuta le righe della matrice
 - Permuta le colonne della matrice
 - Ordinamento alfabetico delle righe
 - Ordinamento alfabetico delle colonne
 - Esporta in formato BASIC

➤ Operazioni sulla matrice

→ Distanza

- ♦ Per righe / Per colonne

Viene proposta la scelta del tipo di distanza da calcolare:

Scelta tipo di distanza

Tipi di distanza disponibili

- ☒ Euclidea
- ☐ Simil. ratio Copert
- ☐ Simil ratio Binaria
- ☐ Assoluta
- ☐ Coeff. di correlaz
- ☐ Yager
- ☐ Coseno dell'angolo
- ☐ Coeff. di Bray-Curtis

☐ Modalita' veloce

Nome nuova matrice:

☒ OK

Le formule utilizzate per il calcolo sono le seguenti:

■ Distanza Euclidea

$$E_{i,j} = \sqrt{\sum_{r=1}^R (x_{r,i} - x_{r,j})^2}$$

$E_{i,j}$ = distanza euclidea tra i vettori colonna i-esimo e j-esimo

$x_{r,i}$ = generico elemento del vettore colonna i-esimo

$x_{r,j}$ = generico elemento del vettore colonna j-esimo

r = generica riga della matrice

R = numero totale di righe della matrice

■ Similarity ratio su valori di copertura

$$S_{i,j} = \frac{\sum_{r=1}^R x_{r,i} x_{r,i}}{\sum_{r=1}^R x_{r,i}^2 + \sum_{r=1}^R x_{r,j}^2 - \sum_{r=1}^R x_{r,i} x_{r,j}}$$

$S_{i,j}$ = Similarity Ratio tra i vettori colonna i-esimo e j-esimo

$x_{r,i}$ = generico elemento del vettore colonna i-esimo

$x_{r,j}$ = generico elemento del vettore colonna j-esimo

r = generica riga della matrice

R = numero totale di righe della matrice

■ Similarity Ratio in base a Presenza/Assenza (binaria)

$$S_{i,j} = \frac{\sum_{r=1}^R \alpha_{r,i} \alpha_{r,j}}{\sum_{r=1}^R \alpha_{r,i}^2 + \sum_{r=1}^R \alpha_{r,j}^2 - \sum_{r=1}^R \alpha_{r,i} \alpha_{r,j}}$$

ove $[x_{x,y} \neq 0] \Rightarrow [\alpha_{x,y} = 1]$ e $[x_{x,y} = 0] \Rightarrow [\alpha_{x,y} = 0]$

$S_{i,j}$ = Similarity Ratio in base a valori di presenza/assenza tra i vettori colonna i-esimo e j-esimo

$\alpha_{r,i}$ = generico elemento del vettore colonna i-esimo

$\alpha_{r,j}$ = generico elemento del vettore colonna j-esimo

r = generica riga della matrice

R = numero totale di righe della matrice

■ Distanza Assoluta

$$A_{i,j} = \sum_{r=1}^R |x_{r,i} - x_{r,j}|$$

$A_{i,j}$ = distanza assoluta tra i vettori colonna i-esimo e j-esimo

$x_{r,i}$ = generico elemento del vettore colonna i-esimo

$x_{r,j}$ = generico elemento del vettore colonna j-esimo

r = generica riga della matrice

R = numero totale di righe della matrice

■ Coefficiente di Correlazione

$$R_{i,j} = \frac{\sum_{r=1}^R (x_{r,i} - \bar{x}_i)(x_{r,j} - \bar{x}_j)}{\sqrt{\sum_{r=1}^R (x_{r,i} - \bar{x}_i)^2 \sum_{r=1}^R (x_{r,j} - \bar{x}_j)^2}}$$

$R_{i,j}$ = coefficiente di correlazione tra i vettori colonna i-esimo e j-esimo

$x_{r,i}$ = generico elemento del vettore colonna i-esimo

$x_{r,j}$ = generico elemento del vettore colonna j-esimo

r = generica riga della matrice

R = numero totale di righe della matrice

$$\bar{x}_i = \frac{\sum_{r=1}^R x_{r,i}}{R}$$

media del vettore colonna i-esimo

$$\bar{x}_j = \frac{\sum_{r=1}^R x_{r,j}}{R}$$

media del vettore colonna j-esimo

■ Somiglianza basata sulla formula di Yager

$$Y_{i,j} = 1 - \frac{\sum_{r=1}^R (x_{r,i} - x_{r,j})^2}{\sqrt{R}}$$

$Y_{i,j}$ = somiglianza di Yager tra i vettori colonna i-esimo e j-esimo

$x_{r,i}$ = generico elemento del vettore colonna i-esimo

$x_{r,j}$ = generico elemento del vettore colonna j-esimo

r = generica riga della matrice

R = numero totale di righe della matrice

■ Coseno dell'angolo

$$\cos \alpha_{i,j} = \frac{\sum_{r=1}^R x_{r,i} x_{r,j}}{\sqrt{\sum_{r=1}^R x_{r,i}^2 \sum_{r=1}^R x_{r,j}^2}}$$

$\cos \alpha_{ij}$ = coseno dell'angolo tra i vettori colonna i-esimo e j-esimo

$x_{r,i}$ = generico elemento del vettore colonna i-esimo
 $x_{r,j}$ = generico elemento del vettore colonna j-esimo
 r = generica riga della matrice
 R = numero totale di righe della matrice

■ Formula di Bray-Curtis

$$S_{i,j} = \frac{\sum_{r=1}^R 2\min(x_{r,i}, x_{r,j})}{\sum_{r=1}^R (x_{r,i} + x_{r,j})}$$

$S_{i,j}$ = Coefficiente di Bray-Curtis tra i vettori colonna i-esimo e j-esimo

$x_{r,i}$ = generico elemento del vettore colonna i-esimo
 $x_{r,j}$ = generico elemento del vettore colonna j-esimo
 r = generica riga della matrice
 R = numero totale di righe della matrice

$\min(x_{r,i}, x_{r,j})$ = il minore tra i generici elementi dei vettori colonna i-esimo e j-esimo

■ Indice di Gower (per tabelle di dati misti e mancanti)

$$S_{i,j} = \frac{\sum_{r=1}^R 1 - \frac{|x_{r,i} - x_{r,j}|}{x_{r_{max}} - x_{r_{min}}}}{\sum_{r=1}^R P_r}$$

$S_{i,j}$ = Coefficiente di Gower tra i vettori colonna i-esimo e j-esimo

$x_{r,i}$ = generico elemento del vettore colonna i-esimo
 $x_{r,j}$ = generico elemento del vettore colonna j-esimo
 r = generica riga della matrice
 R = numero totale di righe della matrice

P_r = indicatore di presenza della coppia di variabili ove:

$P_r = 1$ se $x_{r,i}$ o $x_{r,j}$ sono entrambe definite

$P_r = 1$ se $x_{r,i}$ o $x_{r,j}$ sono entrambe definite

→ Raggruppamento

- Raggruppamento righe/colonne della matrice

La funzione di raggruppamento righe/colonne della matrice consente di raggruppare le righe/colonne della matrice selezionata utilizzando una matrice "traccia" che deve avere le seguenti caratteristiche:

- deve essere di tipo intero piccolo;
- deve avere un numero di righe pari al numero dei gruppi di righe/colonne che si vogliono ottenere;
- deve avere un numero di colonne pari al numero di righe/colonne della matrice originaria;
- ogni riga deve contenere gli ordinali (da 1 ad N con N=numero di righe/colonne della matrice originaria) di riga/colonna che compongono il gruppo corrispondente: sulle colonne eccedenti il numero di elementi di ciascun gruppo verrà riportato 0
- poiché il programma non effettua alcun controllo sulla presenza di tutte le righe/colonne della matrice originaria in uno dei gruppi è anche possibile escludere alcune righe/colonne dalla matrice risultante (semplicemente omettendo di inserirne i relativi ordinali nel gruppo).

Esempio:

L'applicazione di questa matrice di gruppi ad una matrice di 10 righe/colonne produrrà una matrice di 3 righe/colonne i cui valori saranno calcolati con le opportune funzioni di seguito descritte a partire, rispettivamente, dalle righe/colonne 1,4,9, dalle righe/colonne 2,6,8,10 e dalle righe/colonne 3,5,7 della matrice originale.

Le funzioni di raggruppamento disponibili sono le seguenti (per comodità si fa riferimento al raggruppamento di righe, rappresentando il c-esimo elemento della i-esima riga della matrice risultante); il raggruppamento per colonne si comporta in maniera analoga partendo dalla matrice trasposta della originale.

- Somma:

$$S_{i,c} = \sum_{j=1}^{N, y_{i,j} \neq 0} x_{y_{i,j},c}$$

$S_{i,c}$ = c-esimo elemento della i-esima riga

i = numero del gruppo $1 \leq i \leq N_g$ con N_g Numero dei gruppi ovvero numero delle righe della matrice dei gruppi

c = $1 \leq c \leq N_c$ con N_c Numero delle colonne della matrice

$y_{i,j}$ = elemento della matrice di raggruppamento.

$X_{y,c}$ = elemento della matrice originale

■ Contingenza:

$$C_{i,c} = \sum_{j=1}^{N, y_{i,j} \neq 0} \alpha_{y_{i,j},c} \quad x_{y_{i,j},c} > 0 \Rightarrow \alpha_{y_{i,j},c} = 1 \quad x_{y_{i,j},c} = 0 \Rightarrow \alpha_{y_{i,j},c} = 0$$

$C_{i,c}$ = i-esimo elemento della m-esima riga

i = numero del gruppo $1 \leq i \leq N_g$ con N_g Numero dei gruppi ovvero numero delle righe della matrice dei gruppi

c = $1 \leq c \leq N_c$ con N_c Numero delle colonne della matrice

$y_{i,j}$ = elemento della matrice di raggruppamento.

$X_{y,c}$ = elemento della matrice originale

■ Media su presenti

$$m_{i,c} = \frac{S_{i,c}}{C_{i,c}}$$

$S_{i,c}$ = Somma degli elementi delle righe gruppo i colonna c

$C_{i,c}$ = Contingenza degli elementi delle righe gruppo i colonna c

■ Media

$$M_{i,c} = \frac{\sum_{j=1}^{N_{G_i}} x_{y_{i,j},c}}{N_{G_i}}$$

$M_{i,c}$ = i-esimo elemento della m-esima riga

N_{G_i} = numero dei componenti del gruppo i-esimo

i = numero del gruppo $1 \leq i \leq N_g$ con N_g Numero dei gruppi ovvero numero delle righe della matrice dei gruppi

c = $1 \leq c \leq N_c$ con N_c Numero delle colonne della matrice

$y_{i,j}$ = elemento della matrice di raggruppamento.

$X_{y,c}$ = elemento della matrice originale

■ Sigma 2 (varianza)

$$\sigma_{i,c}^2 = \frac{\sum_{j=1}^{N_{G_i}} (x_{y_{i,j},c} - M_{i,c})^2}{N_{G_i} - 1}$$

- $\sigma_{i,c}^2$ = varianza degli elementi dell'iesimo gruppo di righe sulla colonna c
 $M_{i,c}$ = media degli elementi dell'iesimo gruppo di righe sulla colonna c
 N_{G_i} = numero dei componenti del gruppo i-esimo
 i = numero del gruppo $1 \leq i \leq N_g$ con N_g Numero dei gruppi ovvero numero delle righe della matrice dei gruppi
 c = $1 \leq c \leq N_c$ con N_c Numero delle colonne della matrice
 $y_{i,j}$ = elemento della matrice di raggruppamento.
 $X_{y,c}$ = elemento della matrice originale

■ Baricentro

$$b_{i,c} = \frac{\sum_{j=1}^{N_{G_i}} x_{y_{i,j},c} j}{S_{i,c}}$$

- $b_{i,c}$ = baricentro degli elementi dell'iesimo gruppo di righe sulla colonna c
 $S_{i,c}$ = somma degli elementi dell'iesimo gruppo di righe sulla colonna c
 N_{G_i} = numero dei componenti del gruppo i-esimo
 i = numero del gruppo $1 \leq i \leq N_g$ con N_g Numero dei gruppi ovvero numero delle righe della matrice dei gruppi
 c = $1 \leq c \leq N_c$ con N_c Numero delle colonne della matrice
 $y_{i,j}$ = elemento della matrice di raggruppamento.
 $X_{y,c}$ = elemento della matrice originale

■ Chi Quadrato

$$\chi_{i,c}^2 = \frac{\left(\frac{N_{G_i} k_c}{R} - C_{i,c}\right)^2}{\frac{N_{G_i} k_c}{R}} + \frac{\left\{\frac{[N_{G_i}(R-k_c)]}{R} - (n_i - C_{c,i})\right\}^2}{\frac{N_{G_i}(R-k_c)}{R}}$$

- $\chi_{i,c}^2$ = chi quadrato degli elementi dell'iesimo gruppo di righe sulla colonna c
 k_c = numero di presenze (elementi diversi da 0) della

		colonna c della matrice di partenza
R	=	numero totale delle righe della matrice di partenza
$C_{i,c}$	=	Contingenza degli elementi delle righe gruppo i colonna c
N_{G_i}	=	numero dei componenti del gruppo i-esimo

- Creazione matrice TRASPOSTA
- Copia Matrice
- Numera automaticamente la matrice vettore
- Prodotto DESTRO con un'altra matrice
- Mostra il grafo di una classificazione fatta su questa matrice delle distanze
- Calcola il rango delle variabili sulle RIGHE
- Merge di una matrice con un'altra
- Calcola Autovettori ed Autovalori della matrice
- Normalizzazioni e standardizzazioni
 - Proiezione nell'intervallo 0...1 di tutti i valori della matrice
 - Tra 0 ed 1 per vettore
 - Somma ad 1 per vettore
 - Normalizzazione per vettore
 - Standardizzazione
 - Matrice delle varianze
 - Complemento
 - Operazioni sulla matrice
- Grafici
 - Dendrogramma
 - Crea uno scattergram
- Altro
 - Visione della matrice
 - Analisi della classificazione

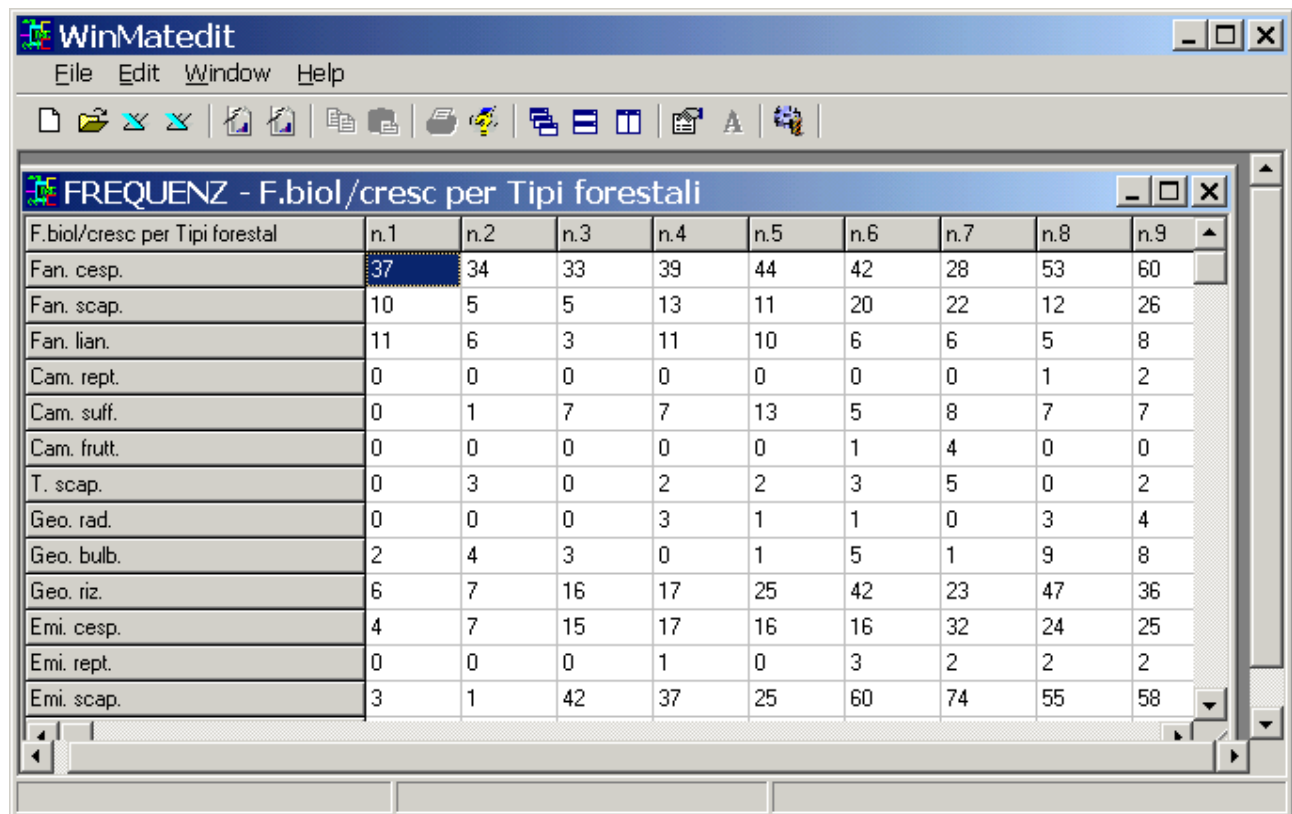
5. Principali modalità di utilizzo

(descrizione di classificazione, ordinamento, PCA, fuzzy, rango variabili come script aggiunti di default al pacchetto.

WinMatedit è stato sviluppato per agevolare le operazioni sulle matrici richieste

dalla teoria dell'analisi multivariata dei dati.

Elenchiamo di seguito i processi che più comunemente vengono utilizzati per ottenere una classificazione degli elementi di una tabella. Negli esempi che seguono, considereremo un esempio nel quale le righe rappresentano forme biologiche / di crescita e le colonne tipi forestali: l'obiettivo è classificare i tipi forestali in gruppi distinti e omogenei.



The screenshot shows the WinMatedit software window. The title bar reads 'WinMatedit'. The menu bar includes 'File', 'Edit', 'Window', and 'Help'. The toolbar contains various icons for file operations and editing. The main window displays a table titled 'FREQUENZ - F.biol/cresc per Tipi forestali'. The table has 10 columns: 'F.biol/cresc per Tipi forestal', 'n.1', 'n.2', 'n.3', 'n.4', 'n.5', 'n.6', 'n.7', 'n.8', and 'n.9'. The rows represent different biological forms: 'Fan. cesp.', 'Fan. scap.', 'Fan. lian.', 'Cam. rept.', 'Cam. suff.', 'Cam. frutt.', 'T. scap.', 'Geo. rad.', 'Geo. bulb.', 'Geo. riz.', 'Emi. cesp.', 'Emi. rept.', and 'Emi. scap.'. The cell at the intersection of 'Fan. cesp.' and 'n.1' is highlighted with a blue background and contains the value 37.

F.biol/cresc per Tipi forestal	n.1	n.2	n.3	n.4	n.5	n.6	n.7	n.8	n.9
Fan. cesp.	37	34	33	39	44	42	28	53	60
Fan. scap.	10	5	5	13	11	20	22	12	26
Fan. lian.	11	6	3	11	10	6	6	5	8
Cam. rept.	0	0	0	0	0	0	0	1	2
Cam. suff.	0	1	7	7	13	5	8	7	7
Cam. frutt.	0	0	0	0	0	1	4	0	0
T. scap.	0	3	0	2	2	3	5	0	2
Geo. rad.	0	0	0	3	1	1	0	3	4
Geo. bulb.	2	4	3	0	1	5	1	9	8
Geo. riz.	6	7	16	17	25	42	23	47	36
Emi. cesp.	4	7	15	17	16	16	32	24	25
Emi. rept.	0	0	0	1	0	3	2	2	2
Emi. scap.	3	1	42	37	25	60	74	55	58

- Dendrogramma e ordinamento della matrice

Per ottenere un dendrogramma dobbiamo calcolare la matrice delle distanze o delle somiglianze tra le colonne. Otterremo una matrice simmetrica con 1 sulla diagonale se utilizzeremo una funzione di somiglianza, 0 se utilizzeremo una funzione di distanza.

Chiudendo la finestra del dendrogramma, ci viene chiesto di generare una matrice "dei gruppi": tale matrice conterrà una unica riga che rappresenta la permutazione che, applicata alla matrice originale, ci permette di ottenere la matrice ordinata.

Generazione matrice dei gruppi da dendrogramma

Nome e indirizzario della matrice da creare
 G:\DELPHIPR\Wmat32 **FREQGR**

Descrizione
 Matrice dei gruppi ottenuti da tagli al dendrogramma

Dimensioni
 Numero righe: 1 Numero colonne: 20

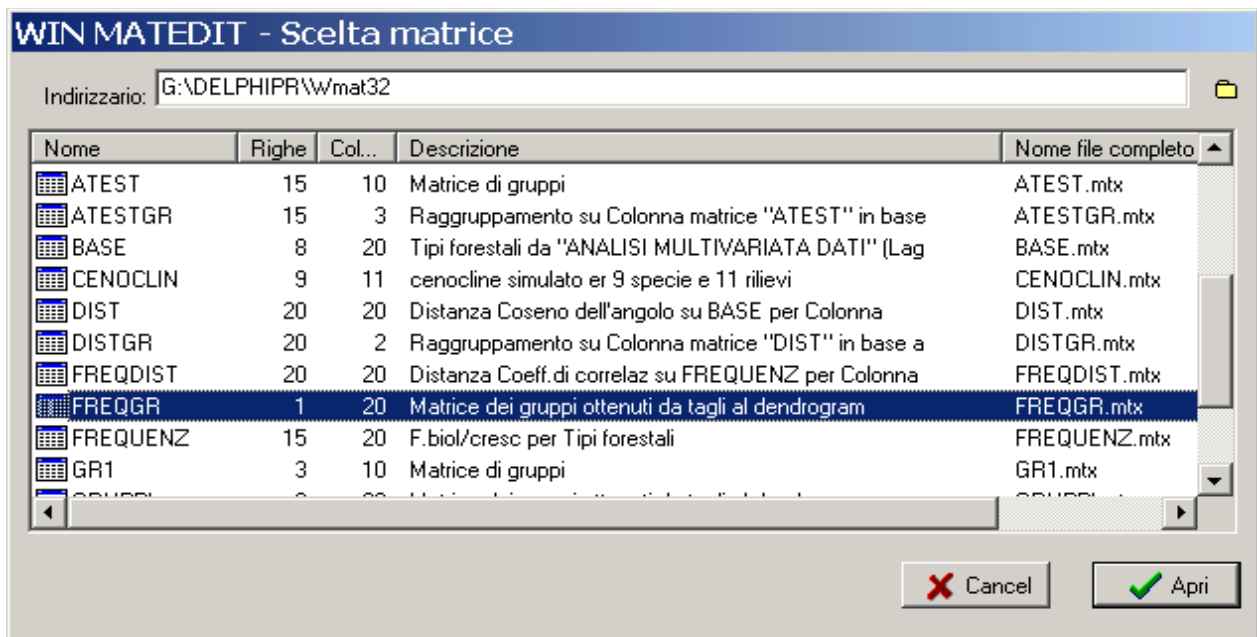
Tipologia
☐ Carattere ☐ Intero ☒ Intero piccolo ☐ Reale

Cancel OK

Aprendo tale matrice otteniamo:

Matrice dei gruppi ottenuti da										
1	2	5	3	16	18	19	20	4	9	

Se ora riapriamo **una copia** della matrice originale (per evitare di modificarla), con la funzione "Spostamenti di righe e colonne" -> "Permutazione delle colonne della matrice" potremo ottenere la matrice ordinata. Per ottenere ciò dovremo selezionare la matrice ottenuta con la chiusura del dendrogramma:



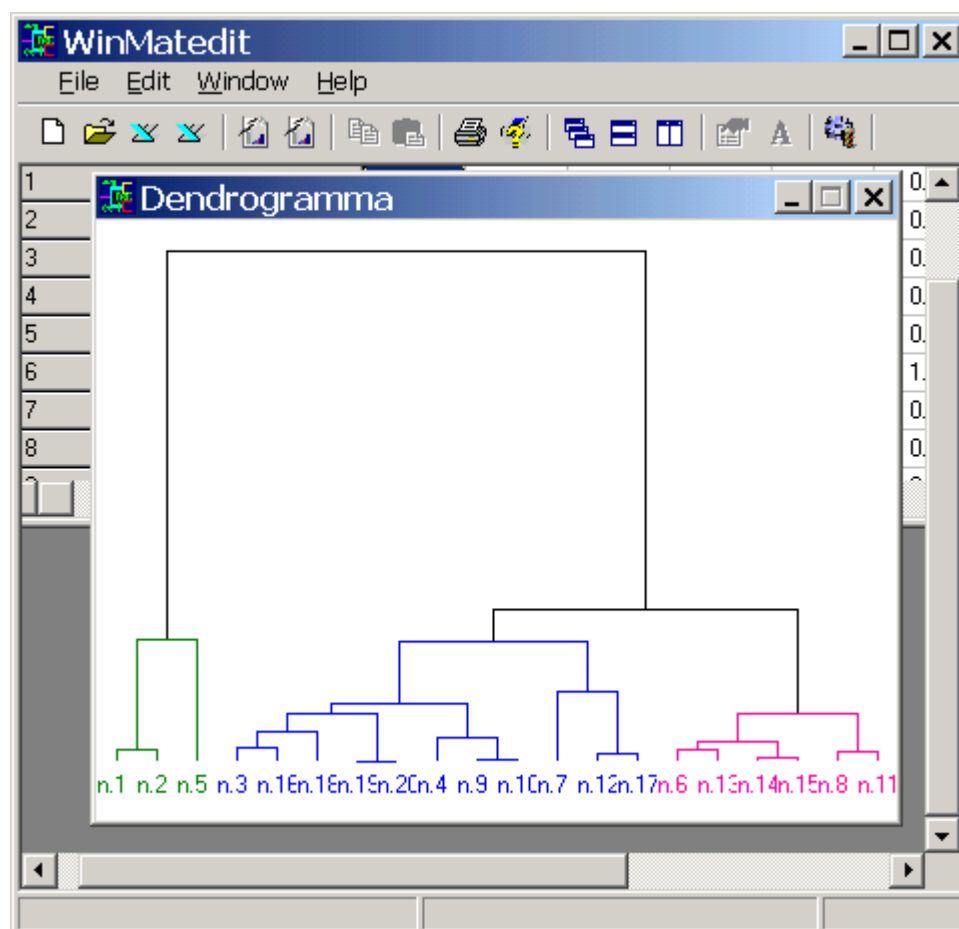
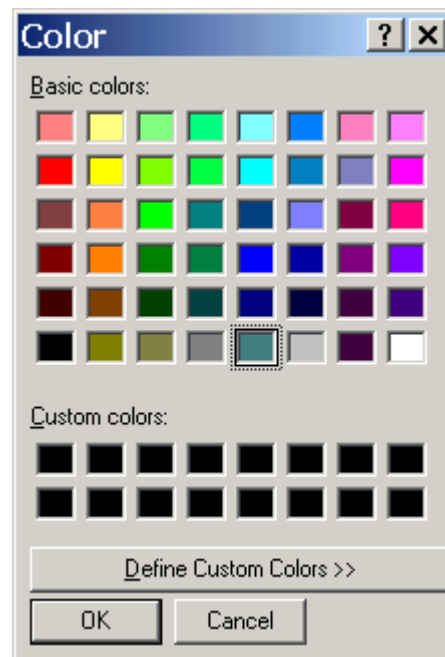
Otterremo:

WinMatedit - [FREQORD - Copia di FREQUENZ]												
Copia di FREQUENZ	n.1	n.2	n.5	n.3	n.16	n.18	n.19	n.20	n.4	n.9	n.10	n.7
Fan. cesp.	37	34	44	33	46	27	40	36	39	60	51	28
Fan. scap.	10	5	11	5	13	10	18	18	13	26	26	22
Fan. lian.	11	6	10	3	4	4	3	0	11	8	8	6
Cam. rept.	0	0	0	0	0	0	1	0	0	2	1	0
Cam. suff.	0	1	13	7	2	2	20	20	7	7	3	8
Cam. frutt.	0	0	0	0	4	2	5	5	0	0	1	4
T. scap.	0	3	2	0	0	0	3	1	2	2	2	5
Geo. rad.	0	0	1	0	0	3	0	0	3	4	3	0
Geo. bulb.	2	4	1	3	4	7	6	8	0	8	5	1
Geo. riz.	6	7	25	16	35	11	20	21	17	36	32	23
Emi. cesp.	4	7	16	15	17	20	22	26	17	25	25	32
Emi. rept.	0	0	0	0	1	0	1	2	1	2	5	2
Emi. scap.	3	1	25	42	65	41	48	48	37	58	58	74
Emi. ros.	1	0	6	6	2	1	5	5	10	8	8	15
NF	16	15	16	10	14	10	15	14	14	18	11	4

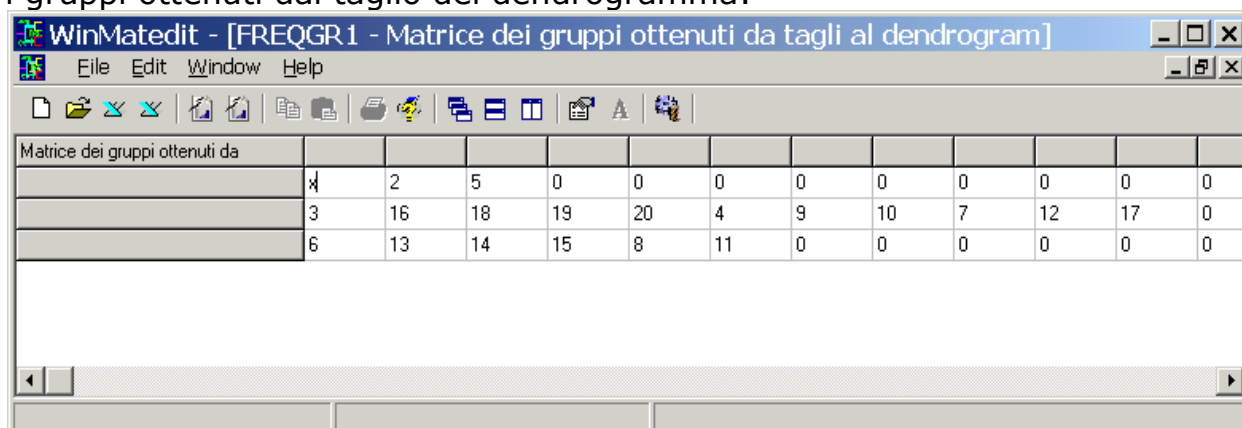
- Raggruppamento delle colonne

Riprendiamo il dendrogramma ottenuto a partire dalla matrice delle distanze: cliccando con il mouse sui tratti verticali, otterremo una finestra di richiesta colore:

Selezionando opportunamente i colori (la scelta dei colori non ha influenza sul risultato dell'elaborazione, ma serve solo per evidenziare i tagli del dendrogramma), potremo impostare i tagli del dendrogramma:



Chiudendo ora la finestra del dendrogramma, ci verrà chiesto di creare una nuova matrice dei gruppi: questa volta essa conterrà tante righe quanti sono i gruppi ottenuti dal taglio del dendrogramma:



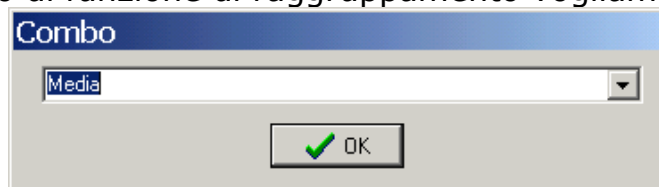
WinMatedit - [FREQGR1 - Matrice dei gruppi ottenuti da tagli al dendrogram]

File Edit Window Help

Matrice dei gruppi ottenuti da

	1	2	5	0	0	0	0	0	0	0	0	0
3	16	18	19	20	4	9	10	7	12	17	0	0
6	13	14	15	8	11	0	0	0	0	0	0	0

Con la funzione "Operazioni sulla matrice"-> "Raggruppamento" -> "Raggruppa le colonne", a partire dalla matrice originale (attenzione: la matrice deve essere quella di partenza, non quella eventualmente riordinata con un'altra elaborazione). Durante quest'operazione ci verrà chiesto di scegliere che tipo di funzione di raggruppamento vogliamo adottare:

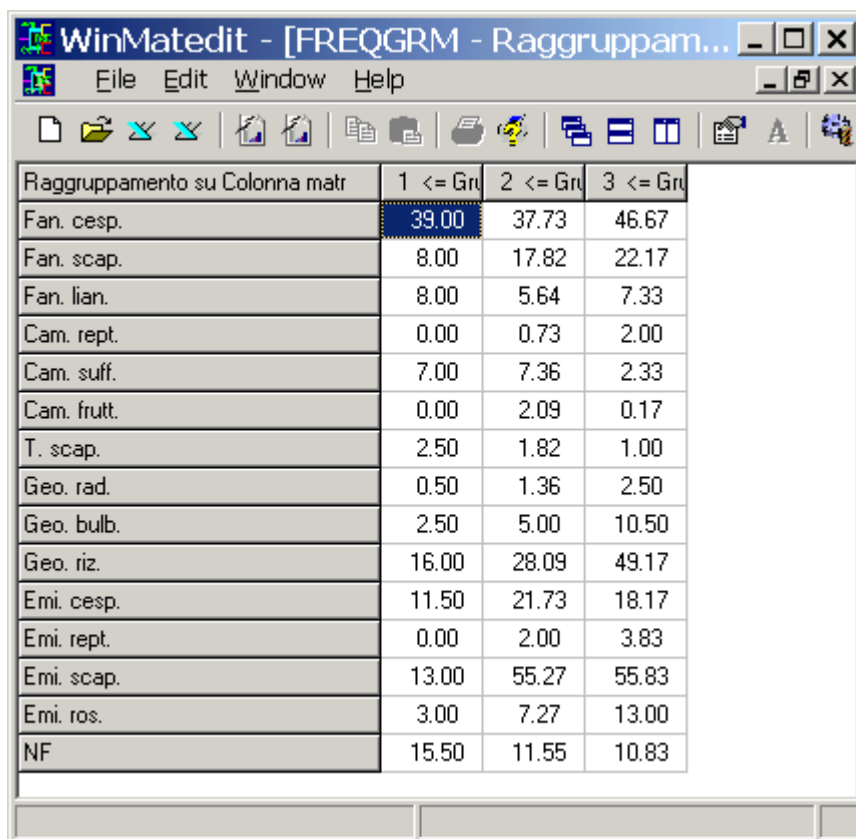


Combo

Media

OK

Otterremo la seguente matrice, con tre colonne, una per ciascun gruppo individuato:



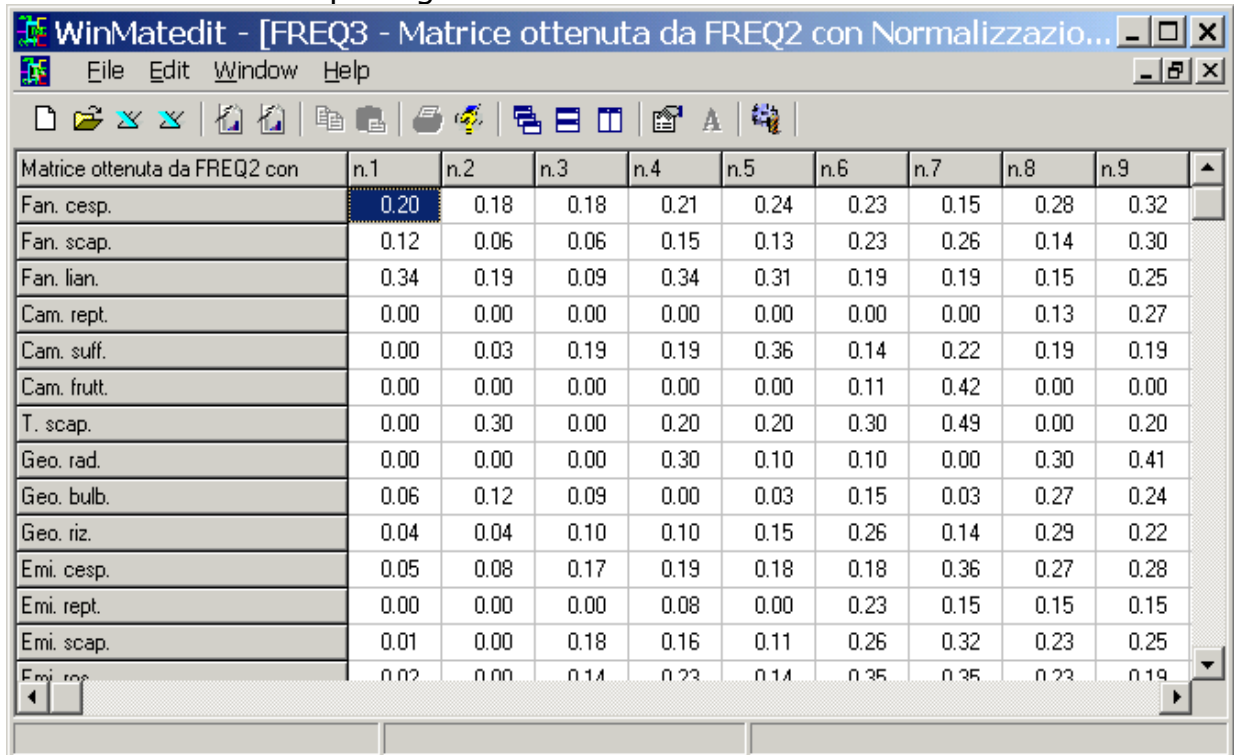
WinMatedit - [FREQGRM - Raggruppam...]

File Edit Window Help

Raggruppamento su Colonna matr	1 <= Gr	2 <= Gr	3 <= Gr
Fan. cesp.	39.00	37.73	46.67
Fan. scap.	8.00	17.82	22.17
Fan. lian.	8.00	5.64	7.33
Cam. rept.	0.00	0.73	2.00
Cam. suff.	7.00	7.36	2.33
Cam. frutt.	0.00	2.09	0.17
T. scap.	2.50	1.82	1.00
Geo. rad.	0.50	1.36	2.50
Geo. bulb.	2.50	5.00	10.50
Geo. riz.	16.00	28.09	49.17
Emi. cesp.	11.50	21.73	18.17
Emi. rept.	0.00	2.00	3.83
Emi. scap.	13.00	55.27	55.83
Emi. ros.	3.00	7.27	13.00
NF	15.50	11.55	10.83

- Analisi delle componenti principali

Dobbiamo come prima cosa normalizzare le righe della matrice di partenza con "Normalizzazioni e Standardizzazioni" -> "Normalizzazione per vettore" -> "Normalizzazione per riga":



Matrice ottenuta da FREQ2 con	n.1	n.2	n.3	n.4	n.5	n.6	n.7	n.8	n.9
Fan. cesp.	0.20	0.18	0.18	0.21	0.24	0.23	0.15	0.28	0.32
Fan. scap.	0.12	0.06	0.06	0.15	0.13	0.23	0.26	0.14	0.30
Fan. lian.	0.34	0.19	0.09	0.34	0.31	0.19	0.19	0.15	0.25
Cam. rept.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.27
Cam. suff.	0.00	0.03	0.19	0.19	0.36	0.14	0.22	0.19	0.19
Cam. frutt.	0.00	0.00	0.00	0.00	0.00	0.11	0.42	0.00	0.00
T. scap.	0.00	0.30	0.00	0.20	0.20	0.30	0.49	0.00	0.20
Geo. rad.	0.00	0.00	0.00	0.30	0.10	0.10	0.00	0.30	0.41
Geo. bulb.	0.06	0.12	0.09	0.00	0.03	0.15	0.03	0.27	0.24
Geo. riz.	0.04	0.04	0.10	0.10	0.15	0.26	0.14	0.29	0.22
Emi. cesp.	0.05	0.08	0.17	0.19	0.18	0.18	0.36	0.27	0.28
Emi. rept.	0.00	0.00	0.00	0.08	0.00	0.23	0.15	0.15	0.15
Emi. scap.	0.01	0.00	0.18	0.16	0.11	0.26	0.32	0.23	0.25
Emi. ros.	0.02	0.00	0.14	0.23	0.14	0.35	0.35	0.23	0.19

Ora con la funzione "Operazioni sulla matrice" ->"Distanza"-> "Crea la matrice delle distanze per colonne", creiamo la matrice delle correlazioni tra le colonne:

WinMatedit - [FREQ10 - Distanza Coeff.di correlaz su FREQ3 per Colo...]

File Edit Window Help

	n.1	n.2	n.3	n.4	n.5	n.6	n.7	n.8	n.9
n.1	1.00	0.63	0.36	0.52	0.57	0.04	-0.23	0.18	0.35
n.2	0.63	1.00	0.17	0.38	0.54	0.21	0.10	-0.12	0.24
n.3	0.36	0.17	1.00	0.36	0.65	0.29	0.05	0.58	0.25
n.4	0.52	0.38	0.36	1.00	0.74	0.32	0.05	0.30	0.53
n.5	0.57	0.54	0.65	0.74	1.00	0.23	0.10	0.20	0.30
n.6	0.04	0.21	0.29	0.32	0.23	1.00	0.54	0.06	-0.10
n.7	-0.23	0.10	0.05	0.05	0.10	0.54	1.00	-0.56	-0.56
n.8	0.18	-0.12	0.58	0.30	0.20	0.06	-0.56	1.00	0.66
n.9	0.35	0.24	0.25	0.53	0.30	-0.10	-0.56	0.66	1.00
n.10	0.17	-0.00	-0.12	0.34	-0.11	0.32	-0.09	0.26	0.43
n.11	0.05	0.19	-0.15	0.31	-0.13	0.21	-0.34	0.52	0.65
n.12	-0.34	0.14	-0.54	-0.17	-0.32	0.12	0.17	-0.42	0.02
n.13	0.02	-0.29	-0.20	-0.28	-0.46	-0.30	-0.72	0.34	0.33
n.14	0.15	-0.25	-0.06	-0.15	-0.37	0.19	-0.49	0.44	0.26

Dalla matrice ottenuta eseguiamo "Operazioni sulla matrice"->"Calcola Autovalori e Autovettori della matrice" : otterremo una matrice contenente gli autovalori sulla diagonale:

WinMatedit - [FREQVAL - Autovalori]

File Edit Window Help

Autovalori	Percentua	Percentua	Percentua	Percentua	Percentua	Percentua	Percentua	Percentua	Percentua
n.1	6.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
n.2	0.00	3.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00
n.3	0.00	0.00	2.79	0.00	0.00	0.00	0.00	0.00	0.00
n.4	0.00	0.00	0.00	2.40	0.00	0.00	0.00	0.00	0.00
n.5	0.00	0.00	0.00	0.00	1.59	0.00	0.00	0.00	0.00
n.6	0.00	0.00	0.00	0.00	0.00	1.23	0.00	0.00	0.00
n.7	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00
n.8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.00
n.9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40
n.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
n.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
n.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
n.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
n.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

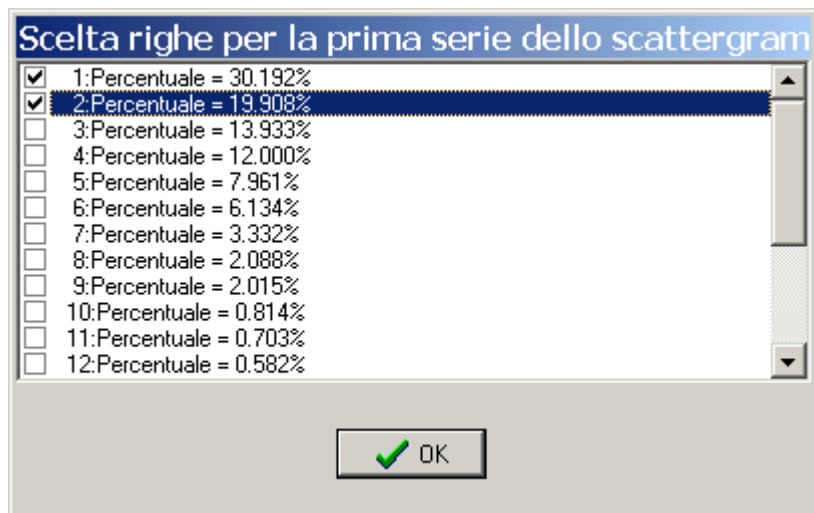
ed una matrice degli autovettori:

WinMatedit - [FREQVEC - autovettori]							
File Edit Window Help							
autovettori	Percentuale = 30.192%	Percentuale = 19.90%	Percentua	Percentua	Percentua	Percentua	
n.1	0.06	0.35	0.03	0.08	-0.26	0.53	
n.2	-0.04	0.27	-0.29	0.09	-0.16	0.42	
n.3	-0.05	0.35	0.25	0.15	-0.18	-0.31	
n.4	0.05	0.40	-0.24	0.06	0.04	-0.12	
n.5	-0.11	0.41	-0.13	0.10	-0.30	-0.13	
n.6	0.03	0.11	-0.09	0.55	0.20	-0.21	
n.7	-0.25	-0.04	-0.15	0.38	0.26	-0.07	
n.8	0.22	0.27	0.25	-0.12	0.04	-0.35	
n.9	0.25	0.28	-0.10	-0.22	-0.02	-0.10	
n.10	0.24	0.09	-0.07	0.15	0.40	0.11	
n.11	0.28	0.12	-0.18	-0.12	0.35	-0.05	
n.12	0.12	-0.26	-0.40	0.13	-0.05	0.02	
n.13	0.33	-0.17	0.12	-0.13	-0.27	0.03	

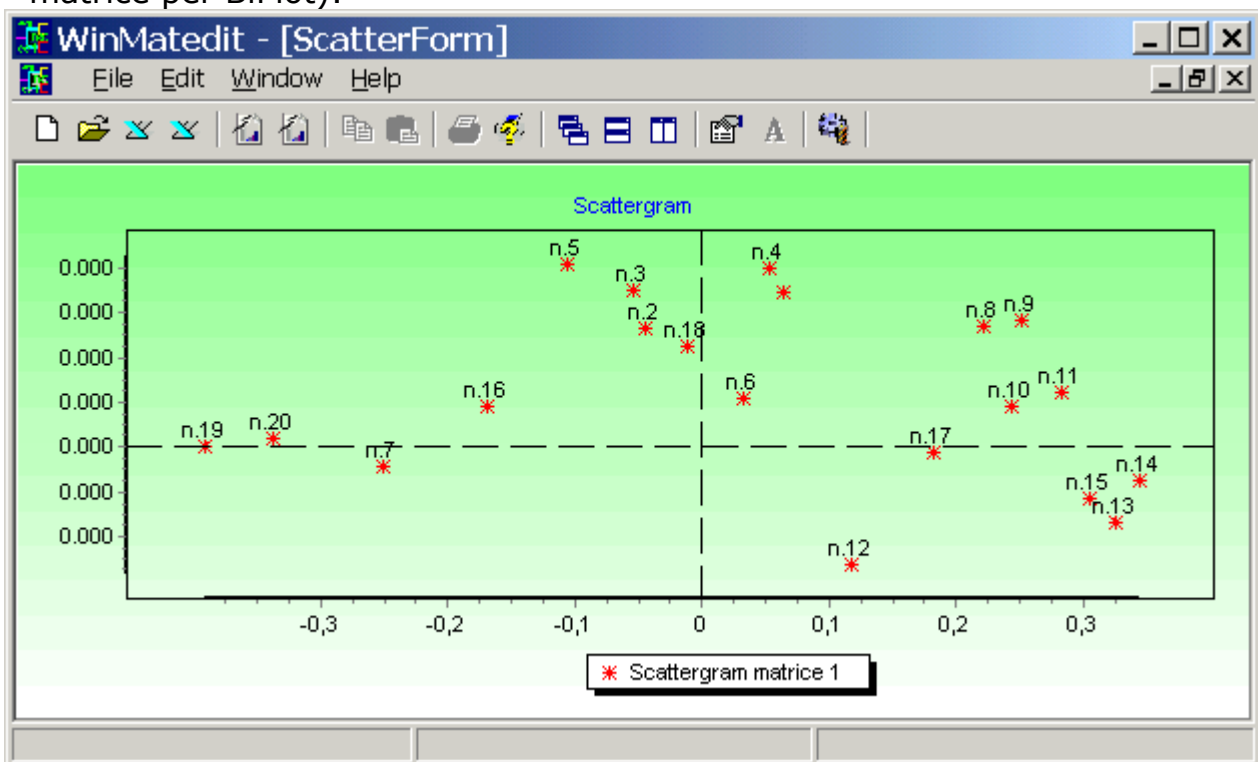
Trasponiamo la matrice degli autovettori con la funzione "Operazioni sulla matrice" -> "Trasposta" e otterremo:

WinMatedit - [FREQVETT - Matrice trasposta di "FREQVEC"]									
File Edit Window Help									
Matrice trasposta di "FREQVEC"	n.1	n.2	n.3	n.4	n.5	n.6	n.7	n.8	n.
Percentuale = 30.192%	0.06	-0.04	-0.05	0.05	-0.11	0.03	-0.25	0.22	
Percentuale = 19.908%	0.35	0.27	0.35	0.40	0.41	0.11	-0.04	0.27	
Percentuale = 13.933%	0.03	-0.29	0.25	-0.24	-0.13	-0.09	-0.15	0.25	
Percentuale = 12.000%	0.08	0.09	0.15	0.06	0.10	0.55	0.38	-0.12	
Percentuale = 7.961%	-0.26	-0.16	-0.18	0.04	-0.30	0.20	0.26	0.04	
Percentuale = 6.134%	0.53	0.42	-0.31	-0.12	-0.13	-0.21	-0.07	-0.35	
Percentuale = 3.332%	-0.14	0.49	0.21	-0.43	-0.14	0.19	0.03	0.21	
Percentuale = 2.088%	-0.04	0.24	0.05	-0.29	0.05	0.02	-0.25	0.05	
Percentuale = 2.015%	-0.11	-0.06	0.26	-0.06	0.01	-0.32	0.21	-0.04	
Percentuale = 0.814%	0.21	0.02	0.38	0.09	-0.18	-0.10	0.40	-0.10	
Percentuale = 0.703%	0.19	-0.07	-0.10	-0.05	-0.16	0.24	0.20	-0.30	
Percentuale = 0.582%	0.07	-0.14	-0.22	-0.08	0.37	-0.12	-0.00	-0.14	
Percentuale = 0.200%	0.08	0.08	-0.23	-0.19	0.27	-0.39	0.44	0.36	

Utilizzando la funzione "Dendrogramma" -> "Crea uno scattergram", ci viene chiesto di impostare la riga e la colonna da rappresentare sullo scattergram:



Confermando otterremo (dopo aver selezionato Cancel alla richiesta della matrice per BiPlot):

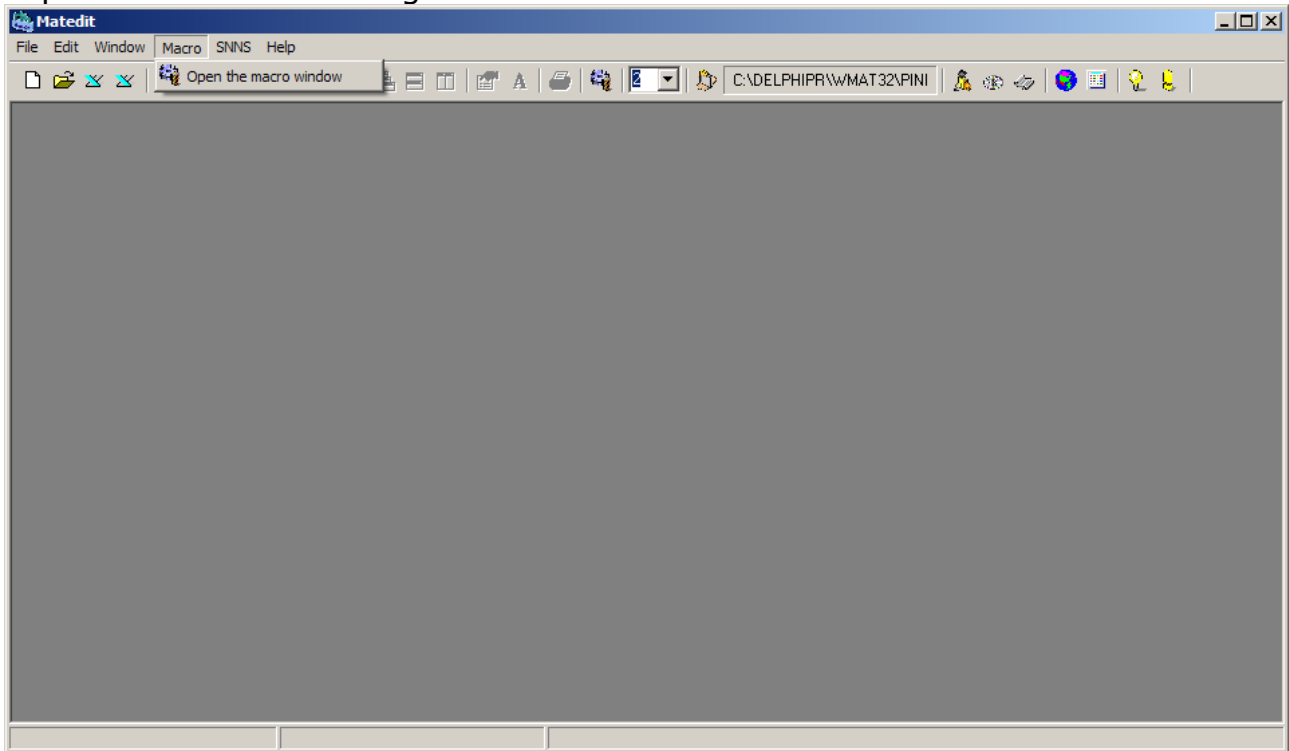


Si noti che per ottenere la visualizzazione delle etichette è necessario sfiorare l'elemento con il cursore del mouse tenendo premuto il tasto maiuscolo della tastiera. Questo per permettere, in caso di grafici molto affollati, di visualizzare solo alcune etichette e non tutte.

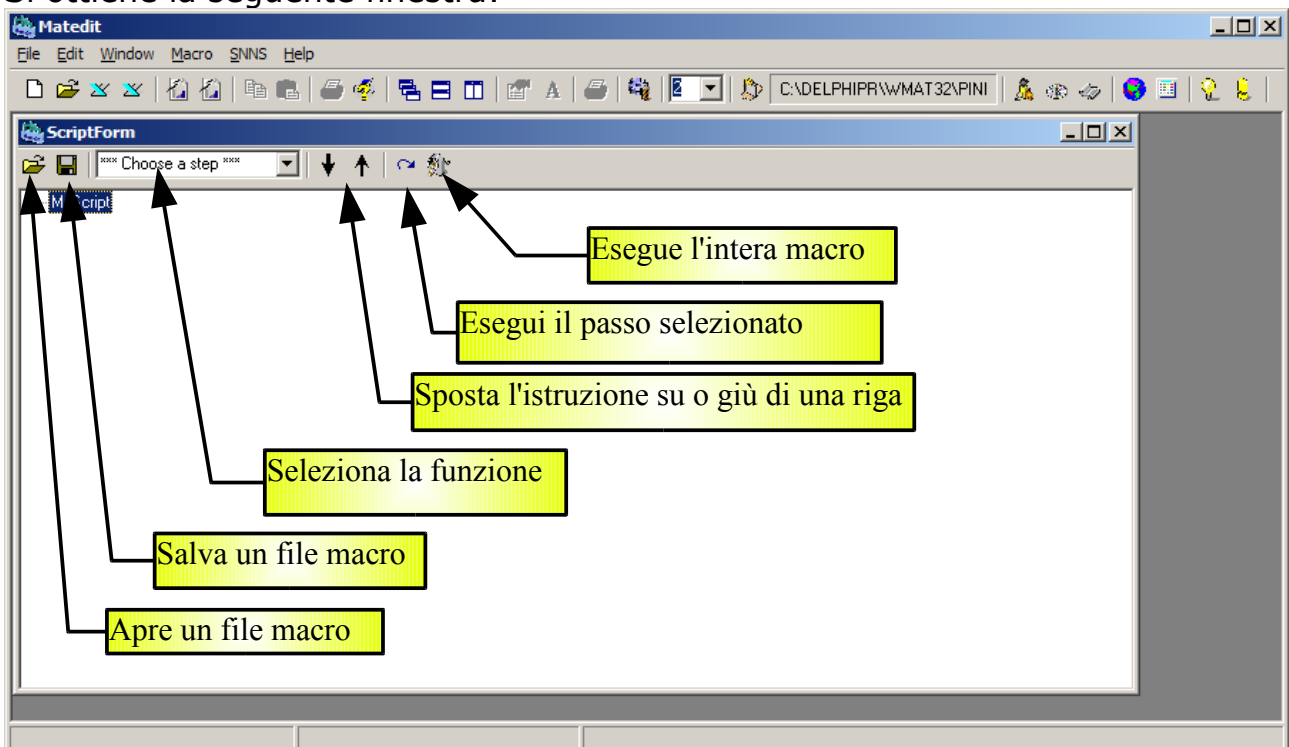
6. Memorizzazione di procedure (macro)

I passi elementari che Matedit è in grado di effettuare possono essere memorizzati in semplici macro che semplificano l'esecuzione ripetitiva di elaborazioni.

Per accedere alle funzioni di gestione delle macro, deve essere selezionata l'opzione di menu che segue:



Si ottiene la seguente finestra:



Nella figura sono indicate le funzioni espletate dai vari tasti.

Ogni singolo passo di elaborazione deve essere effettuato su una matrice che è stata preventivamente aperta. L'apertura di una matrice avviene con la

funzione "openmatrix". Questa funzione permette di indicare se deve essere aperta una matrice prefissata oppure se l'applicazione deve chiedere all'utente quale matrice aprire (generalmente utilizzata per selezionare la matrice di partenza). Essa associa inoltre un "identificativo" ad ogni matrice aperta; tale identificativo verrà utilizzato dalle altre funzioni per indicare su quale matrice operare.

Select

Opens a matrix for further elaboration

Matrix file name

Select

If the user must choose the matrix during execution

no

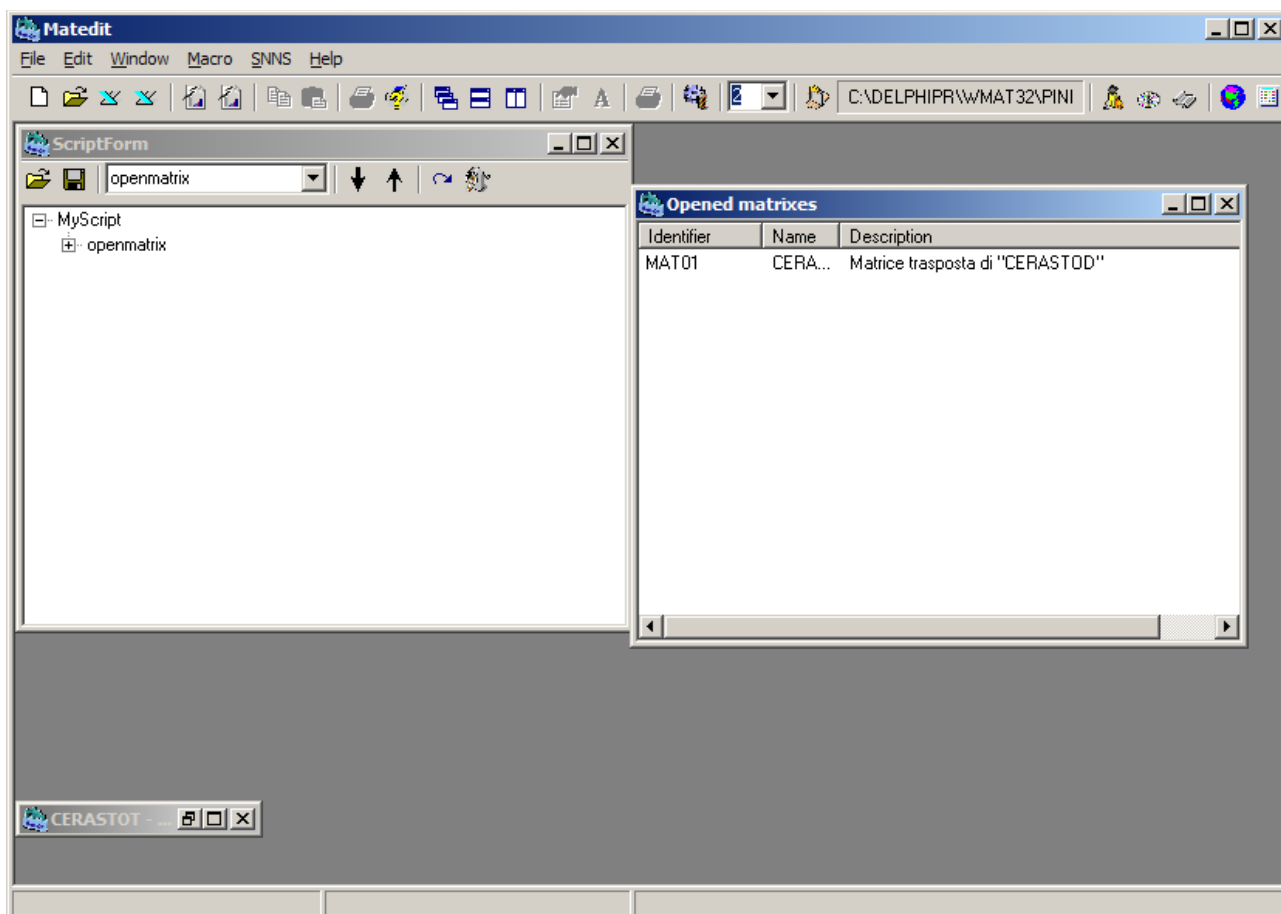
Matrix identifier

MAT01

Comment

Abort OK

Quando si attiva l'elaborazione della macro, si apre una finestra che elenca le matrici aperte dalla procedura:



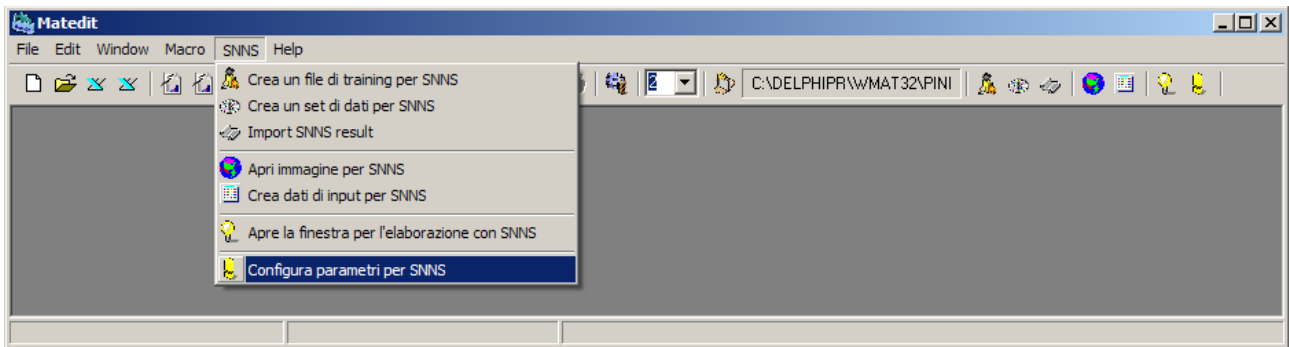
Matedit è corredato da alcune macro che effettuano le principali procedure di calcolo dell'analisi multivariata (PCA, FUZZY sets...).

Le macro vengono memorizzate su files in formato XML; è possibile editare tali files con un editor XML esterno, anche se si consiglia di utilizzare la funzione macro di Matedit.

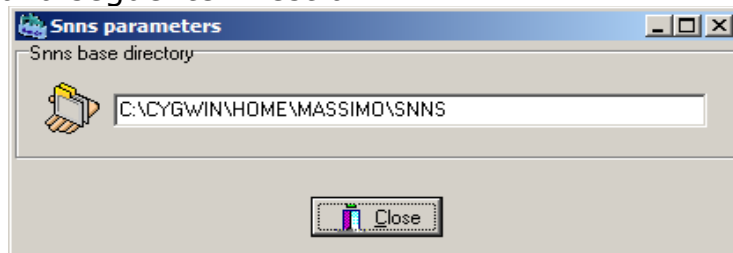
7. Integrazione con SNNS (Stuttgart Neural Network Simulator)

Al fine di poter fornire un ampio spettro di strumenti di analisi, Matedit è fornisce l'interfacciamento con SNNS, applicativo sviluppato da alcune università tedesche, che consente di effettuare elaborazioni utilizzando le peculiarità delle reti neurali. Sono possibili due tipi di interfacciamento: uno per l'analisi di matrici tradizionali ed uno per utilizzare SNNS per tecniche di "remote sensing". In questo manuale descriveremo solamente il primo tipo; per il secondo si rimanda ad una successiva versione.

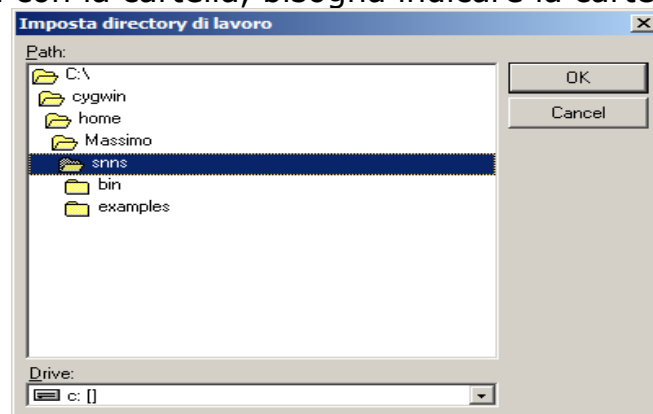
Per poter utilizzare SNNS con Matedit è necessario procurarsi la versione di SNNS per Windows, e decomprimerla in un indirizzario. Bisogna quindi configurare Matedit per impostare tale indirizzario.



Viene visualizzata la seguente finestra:

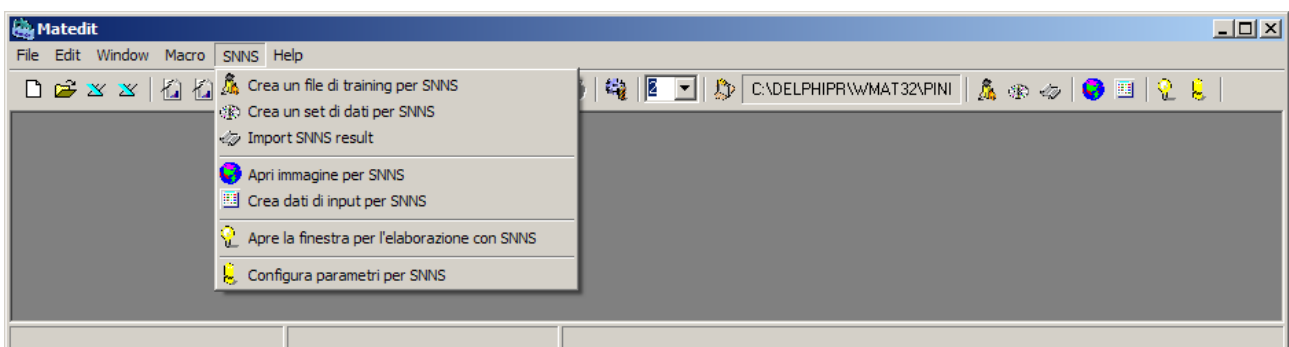


Puntando sull'icona con la cartella, bisogna indicare la cartella base di SNNS:



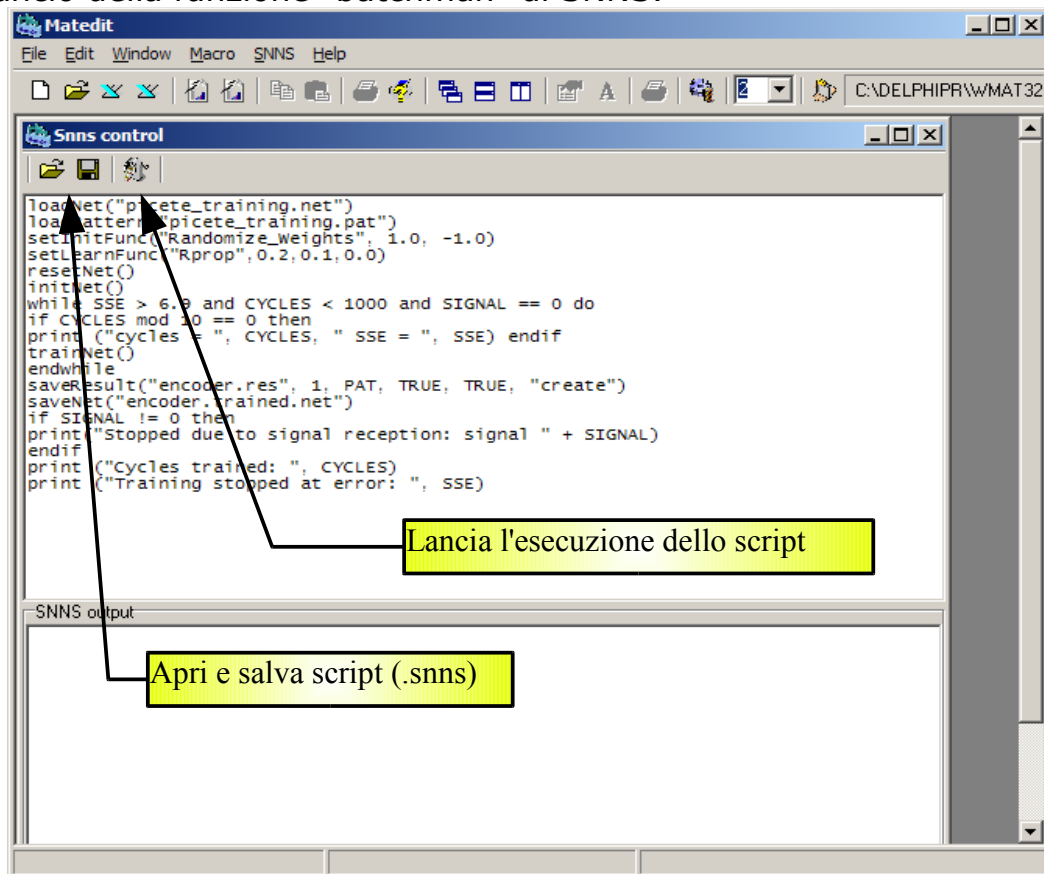
La directory da scegliere è generalmente quella che contiene del sottodirectory "bin" ed "examples", come indicato nell'esempio.

Le funzioni disponibili sono:



- *Crea un file di training per SNNS* : Provvede a creare un file di training per SNNS (.pat) e contemporaneamente crea un file .net che rappresenta una possibile rete di partenza per il training con SNNS. La matrice di partenza deve essere costituita da n vettori colonna, ciascuno dei quali rappresenta un pattern noto.

- *Crea un set di dati per SNNS* : Crea un file di dati (.pat); i dati devono essere costituiti da m vettori colonna della stessa dimensioni di quelli della matrice di training (il numero di righe deve essere uguale a quello della matrice di training).
- *Importa un file risultato di SNNS* : Crea automaticamente una matrice di m (uno per ogni vettore di input) vettori colonna, in cui gli elementi di ciascun vettore rappresenta con un valore da 0 a 1 la sua associazione ad uno degli n pattern noti.
- *Apri la finestra con l'elaborazione con SNNS* : apre la finestra per il lancio della funzione "batchman" di SNNS.



Per la sintassi delle funzioni SNNS si rimanda al capitolo 12 (pag.235) del manuale di SNNS.