



FACOLTA' DI PSICOLOGIA

Scienze della Mente

TESINA DI INFORMATICA

Fossataro Carlotta

Di Rienzo Ariana

Basso Alessandra

Marchetto Helena

SPSS

INDICE

INTRODUZIONE	3
FINESTRE	4
EDITOR DEI DATI	4
OUTPUT	11
SINTASSI.....	12
<i>Comandi principali</i>	13
<u>1. OPERAZIONI SULLE MATRICI</u>	13
<u>2. OPERAZIONI SULLE VARIABILI</u>	13
<u>3. OPERAZIONI SUI CASI</u>	15

Introduzione

SPSS, acronimo di Statistical Package for Social Science, è un software di statistica, la cui prima versione è stata realizzata nel 1968.

Oggi è uno dei programmi più utilizzati, perché è un valido sistema per l'analisi statistica e la gestione dei dati in un ambiente grafico in cui è possibile utilizzare menu descrittivi e semplici finestre di dialogo per eseguire automaticamente numerose operazioni.

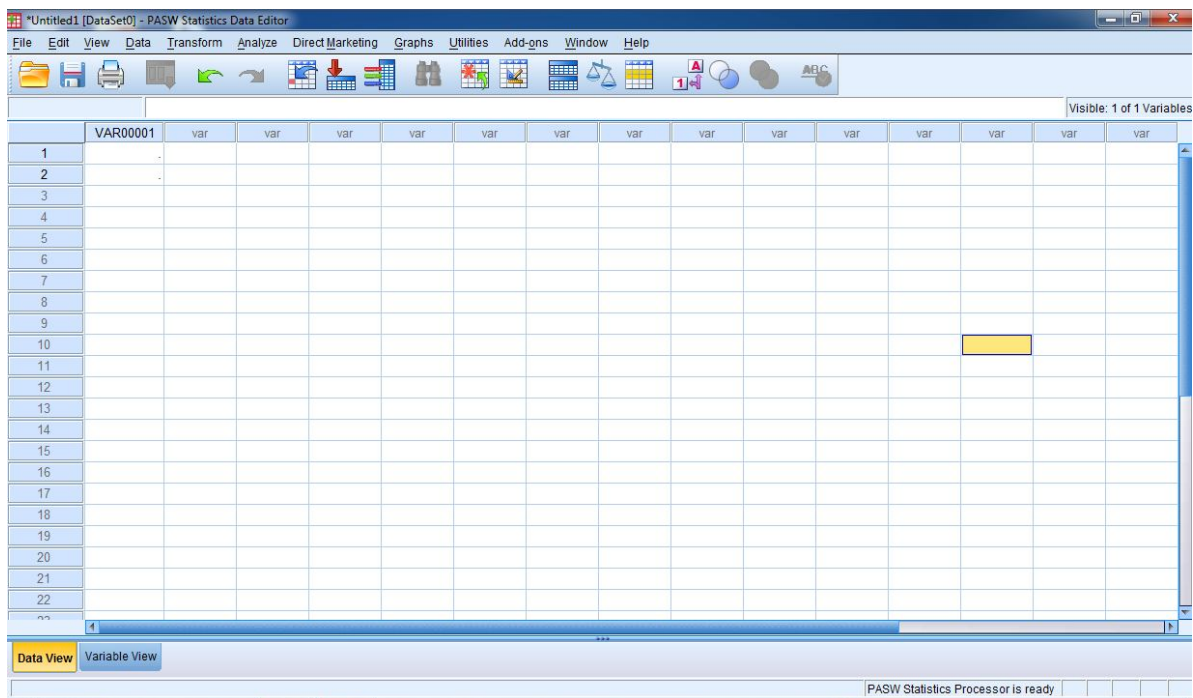


Figura 1: avvio SPSS

SPSS è un programma statistico studiato per soddisfare le esigenze di elaborazione dati nei seguenti settori applicativi:

- Sociologia e scienza politica
- Analisi/ricerche di mercato
- Psicologia
- Medicina/biologia

Esso può eseguire le elaborazioni dati riguardanti le seguenti tecniche di analisi:

- analisi monovariata
- analisi bivariata
- analisi multivariata
- test statistici parametrici e non parametrici
- analisi di serie temporali
- funzioni di sopravvivenza
- modelli lineari
- e molte altre

Inoltre può tracciare grafici e costruire tabelle a singola e doppia entrata, Spss è facile da usare perché dispone di un'interfaccia a finestre, oltre ad un vero e proprio linguaggio di programmazione interno.

Finestre

Il programma SPSS nella versione più completa è strutturato su tre finestre, una finestra di editor dei dati, una di output ed infine una di sintassi.

Editor dei dati

Nell'Editor dei dati viene visualizzato il contenuto del file di dati.

E' possibile creare nuovi file di dati o modificare quelli esistenti.

Se ci sono più file di dati aperti, ciascun file di dati viene visualizzato in una finestra a parte dell'Editor. Tutti i file aperti nell'editor hanno estensione *.sav*

La finestra di editor presenta al suo interno due fogli, uno di "visualizzazione dati" dove viene visualizzata la matrice dei dati, l'altro di "visualizzazione variabili" dove è possibile controllare le etichette discorsive attribuite ad ogni variabile e ai codici assegnati ad ogni livello della variabile.

*mpilata17.sav [InsiemeData] - PASW Statistics Data Editor

File Modifica Visualizza Dati Trasforma Analizza Grafici Utilità Finestra Aiuto

Visibili: 12 variabili di 12

	Berlusconi	Vendola	Casini	Tremonti	Bindi	Bonino	Dipietro	Maroni	Rutelli	Fini	Bossi	Bersani
19	7,00	3,00	6,00	5,00	3,00	3,00	,0	6,00	3,00	7,00	7,00	3,00
20	2,00	7,00	6,00	4,00	7,00	7,00	6,00	,0	6,00	4,00	2,00	6,00
21	6,00	6,00	5,00	5,00	4,00	3,00	3,00	6,00	,0	5,00	6,00	3,00
22	3,00	7,00	3,00	4,00	7,00	6,00	7,00	4,00	5,00	,0	6,00	6,00
23	2,00	8,00	6,00	5,00	7,00	7,00	7,00	2,00	6,00	6,00	,0	7,00
24	7,00	3,00	4,00	6,00	3,00	3,00	3,00	6,00	3,00	6,00	7,00	6,00
25	,0	8,00	3,00	2,00	8,00	8,00	8,00	2,00	6,00	3,00	1,00	7,00
26	8,00	,0	5,00	7,00	2,00	2,00	2,00	8,00	4,00	6,00	8,00	3,00
27	3,00	5,00	,0	5,00	5,00	6,00	5,00	4,00	5,00	1,00	6,00	3,00
28	2,00	7,00	5,00	,0	7,00	7,00	7,00	2,00	6,00	4,00	2,00	6,00
29	8,00	2,00	5,00	7,00	,0	2,00	2,00	7,00	3,00	7,00	8,00	2,00
30	8,00	2,00	6,00	7,00	2,00	,0	2,00	8,00	3,00	6,00	8,00	2,00
31	8,00	2,00	5,00	7,00	2,00	2,00	,0	7,00	3,00	6,00	8,00	3,00
32	2,00	8,00	4,00	2,00	7,00	8,00	7,00	,0	6,00	5,00	2,00	7,00
33	6,00	4,00	5,00	6,00	3,00	3,00	3,00	6,00	,0	5,00	7,00	3,00
34	3,00	6,00	1,00	4,00	7,00	6,00	6,00	5,00	5,00	,0	3,00	5,00
35	1,00	8,00	6,00	7,00	8,00	8,00	8,00	7,00	7,00	3,00	,0	7,00
36	7,00	3,00	3,00	6,00	2,00	2,00	3,00	7,00	3,00	5,00	7,00	6,00
37	,0	8,00	5,00	4,00	8,00	8,00	8,00	6,00	2,00	6,00	5,00	2,00
38	8,00	,0	4,00	8,00	4,00	2,00	2,00	7,00	6,00	4,00	4,00	7,00
39	5,00	4,00	,0	3,00	2,00	6,00	2,00	4,00	2,00	2,00	7,00	4,00
40	4,00	8,00	3,00	,0	6,00	4,00	7,00	2,00	6,00	2,00	4,00	6,00
41	8,00	4,00	2,00	6,00	,0	4,00	5,00	8,00	2,00	4,00	6,00	6,00
42	8,00	2,00	6,00	4,00	4,00	,0	4,00	7,00	5,00	4,00	3,00	4,00
43	6,00	2,00	2,00	7,00	5,00	4,00	,0	5,00	4,00	6,00	6,00	6,00
44	2,00	7,00	4,00	2,00	8,00	7,00	5,00	,0	2,00	6,00	1,00	4,00
45	6,00	6,00	2,00	6,00	2,00	5,00	4,00	2,00	,0	6,00	4,00	5,00
46	5,00	4,00	2,00	2,00	4,00	4,00	6,00	6,00	6,00	,0	4,00	4,00
47	1,00	4,00	7,00	4,00	6,00	3,00	6,00	1,00	4,00	4,00	,0	4,00
48	2,00	7,00	4,00	6,00	6,00	4,00	6,00	4,00	5,00	4,00	4,00	6,00
49	,0	8,00	6,00	4,00	7,00	8,00	8,00	4,00	6,00	3,00	4,00	7,00
50	8,00	,0	7,00	6,00	4,00	4,00	5,00	8,00	6,00	8,00	8,00	6,00
51	6,00	7,00	,0	6,00	5,00	7,00	6,00	6,00	5,00	5,00	5,00	3,00
52	4,00	6,00	6,00	,0	6,00	6,00	6,00	4,00	5,00	5,00	5,00	6,00
53	7,00	4,00	5,00	6,00	,0	4,00	4,00	7,00	6,00	7,00	8,00	4,00
54	8,00	4,00	7,00	6,00	4,00	,0	5,00	8,00	6,00	7,00	8,00	5,00

Visualizzazione dati Visualizzazione variabili

PASW Statistics Il processore è pronto

Figura 2: editor dei dati, visualizzazione dati

*impilata17.sav [InsiemeData] - PASW Statistics Data Editor

File Modifica Visualizza Dati Trasforma Analizza Grafici Utilità Finestra Aiuto

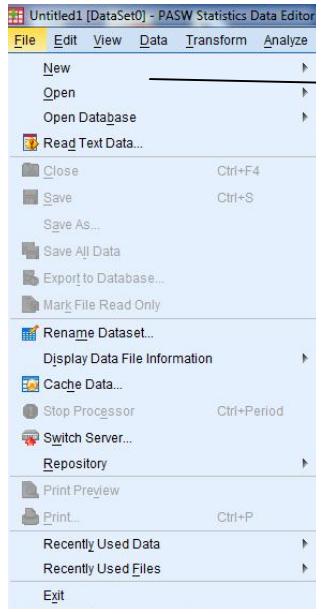
	Nome	Tipo	Larghezza	Decimali	Etichetta	Valori	Mancante/i	Colonne	Allinea	Misura	Ruolo
1	Berlusconi	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
2	Vendola	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
3	Casini	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
4	Tremonti	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
5	Bindi	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
6	Bonino	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
7	Dipietro	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
8	Maroni	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
9	Rutelli	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
10	Fini	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
11	Bossi	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
12	Bersani	Numerico	8	2		Nessuno/a	Nessuno/a	10	Destra	Nominale	Input
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											
28											
29											
30											
31											
32											
33											
34											
35											
36											
37											

Visualizzazione dati Visualizzazione variabili

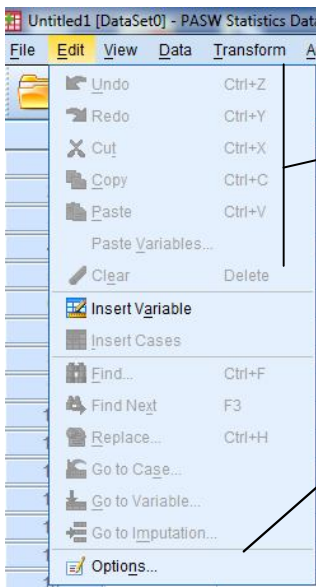
PASW Statistics Il processore è pronto

Figura 3: editor dei dati, visualizzazione variabili

Nella finestra di editor dei dati è possibile richiamare diverse operazioni dai vari menù a tendina. Vengono descritte di seguito le più importanti.

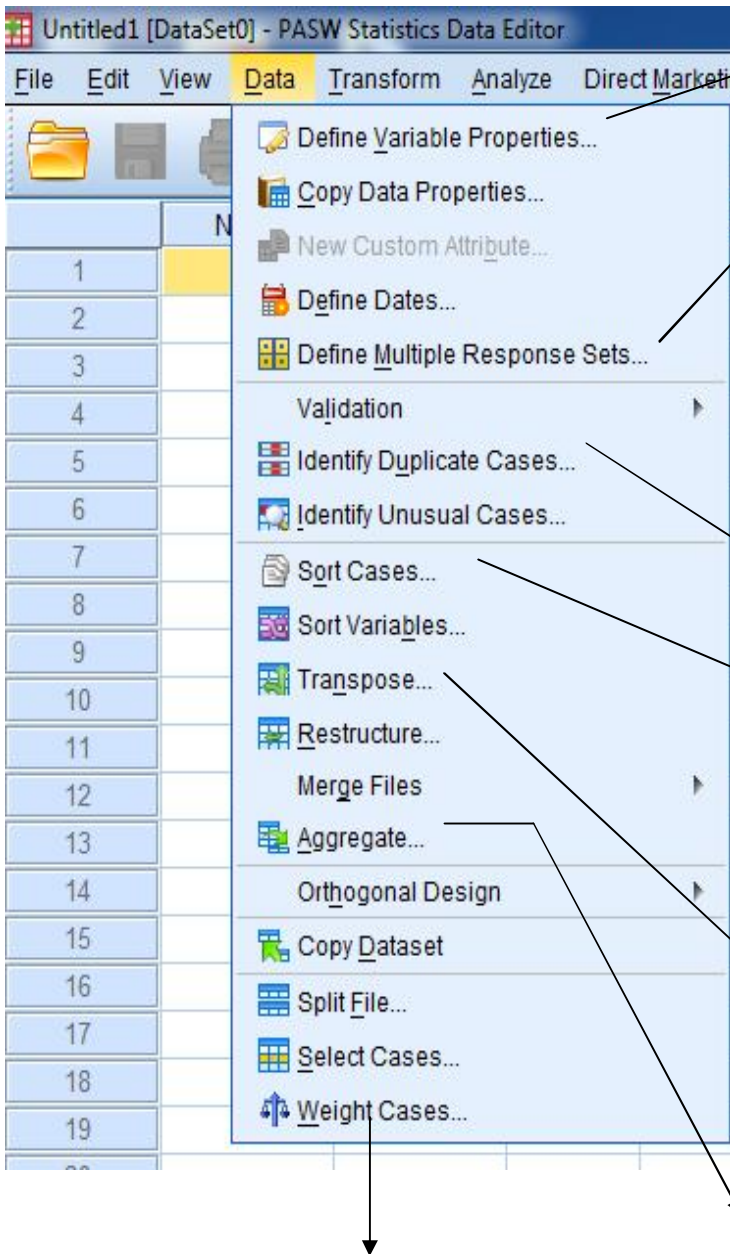


New ed *Open* consentono di aprire una sintassi, una matrice dati nuove o già presenti sul computer.



Varie voci che permettono di gestire ogni nostra operazione: consentono di *tagliare*, *copiare*, *incollare*, ogni dato da noi inserito.

Dalla voce *Options* possiamo richiamare l'operazione "Aprire una sola matrice alla volta" per evitare che il programma apra eventualmente più



Permette di definire le proprietà formali delle variabili che ho inserito

Definisce gli insiemi a risposta multipla: solitamente si hanno tante colonne quante sono le risposte multiple. Ad esempio in un questionario in cui viene chiesto al soggetto di indicare quali sono i due valori (famiglia, amore, lavoro, amicizia,...) più importanti se i soggetti sono 100 avrò 100 colonne. Se però sono interessato a conoscere in quanti hanno risposto

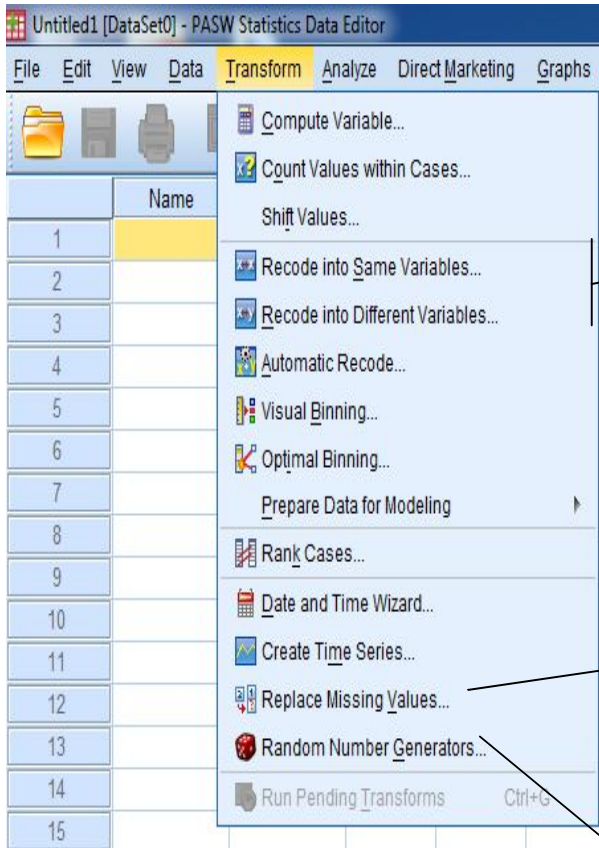
E' una routine utile per controllare se due righe si somigliano troppo. Questo perchè nelle ricerche può capitare che si generino dati uguali o simili. volutamente o

Operazione che consente di ordinare i casi secondo una particolare proprietà, ad esempio in base al genere è possibile cambiare l'ordine delle righe. Azione utile per controllare se un algoritmo è sensibile alla permutazione dei dati, implementabile

Operazione che permette l'inversione del ruolo di righe e di colonne.

Operazione utile nei casi in cui si ottiene una matrice "sbilanciata" in cui il campione non è rappresentativo della popolazione e si vuole calcolare un vettore di pesi. Ad esempio se si hanno 54 femmine ma ne servono 52 quest'operazione fa in modo che ogni femmina valga un po' meno di 1 e i maschi poco di più. (Riproporzionamento a posteriori che deve essere fatto per non alterare la validità dei dati)

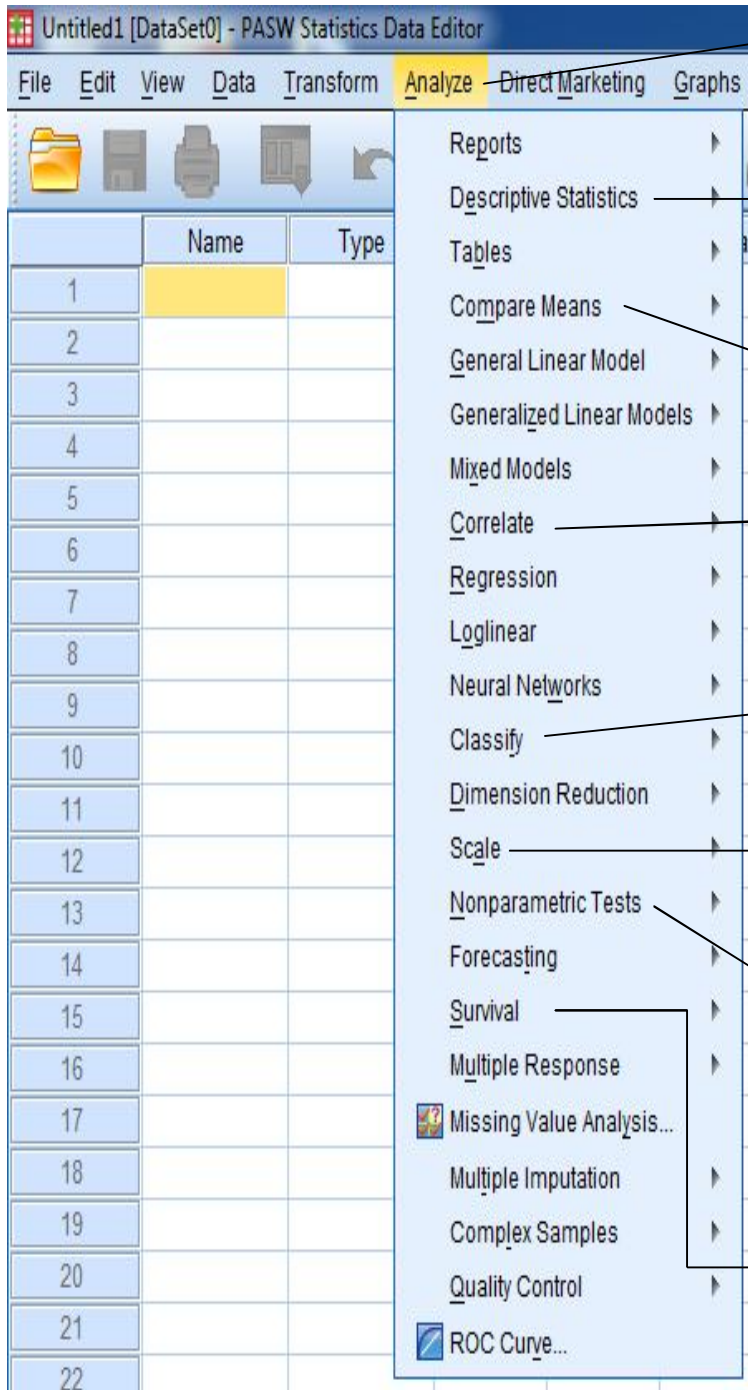
Operazione che permette di far collapsare dei risultati. Ad esempio se intervisto 2000 soggetti su scala nazionale e sono interessato a conoscere i risultati su scala regionale passerò da 2000 righe a 20



E' possibile applicare l'operazione *Ricodifica nelle stesse variabili* (ad esempio calcolo l'età in mesi e non in anni) o *in variabili differenti* (facendo attenzione a non produrre un doppione).

Operazione che permette di sostituire i valori mancanti. Può capitare che dopo aver somministrato un questionario si ottengano dei vuoti d'informazione, quindi si può decidere se riempire arbitrariamente questi vuoti oppure ignorare quelle righe senza.

Operazione che partendo da un seme di inizio consente di generare numeri casuali per riempire le celle con valori mancanti in matrice.



Analizza richiama tutti gli algoritmi di analisi dei dati.

Statistica descrittiva monovariata (moda, media, mediana) e bivariata (Q di Yule, K di Kohen).

Permette di confrontare le varie media (T test, Anova univariata).

Permette di stimare le correlazioni tra diverse variabili (R di Pearson).

Comprende vari metodi di raggruppamento (Cluster Analysis).

Permette di valutare l'attendibilità di uno strumento (Analisi affidabilità).

Comprende diversi test utilizzabili quando si hanno pochi casi e/o vengono violati alcuni modelli, ad esempio se si lavora sulla mediana e

Permette di valutare i cosiddetti falsi positivi e falsi negativi (usato particolarmente in medicina).

Output

Tutti i risultati statistici, le tabelle e i grafici vengono visualizzati nella finestra di output che contiene i risultati delle elaborazioni dei dati e si apre automaticamente al termine di ogni elaborazione. È possibile modificare l'output e salvarlo per un successivo utilizzo. I file generati hanno estensione *.spo* e nel caso si voglia salvare singolarmente ogni grafico il file avrà estensione *.cht*

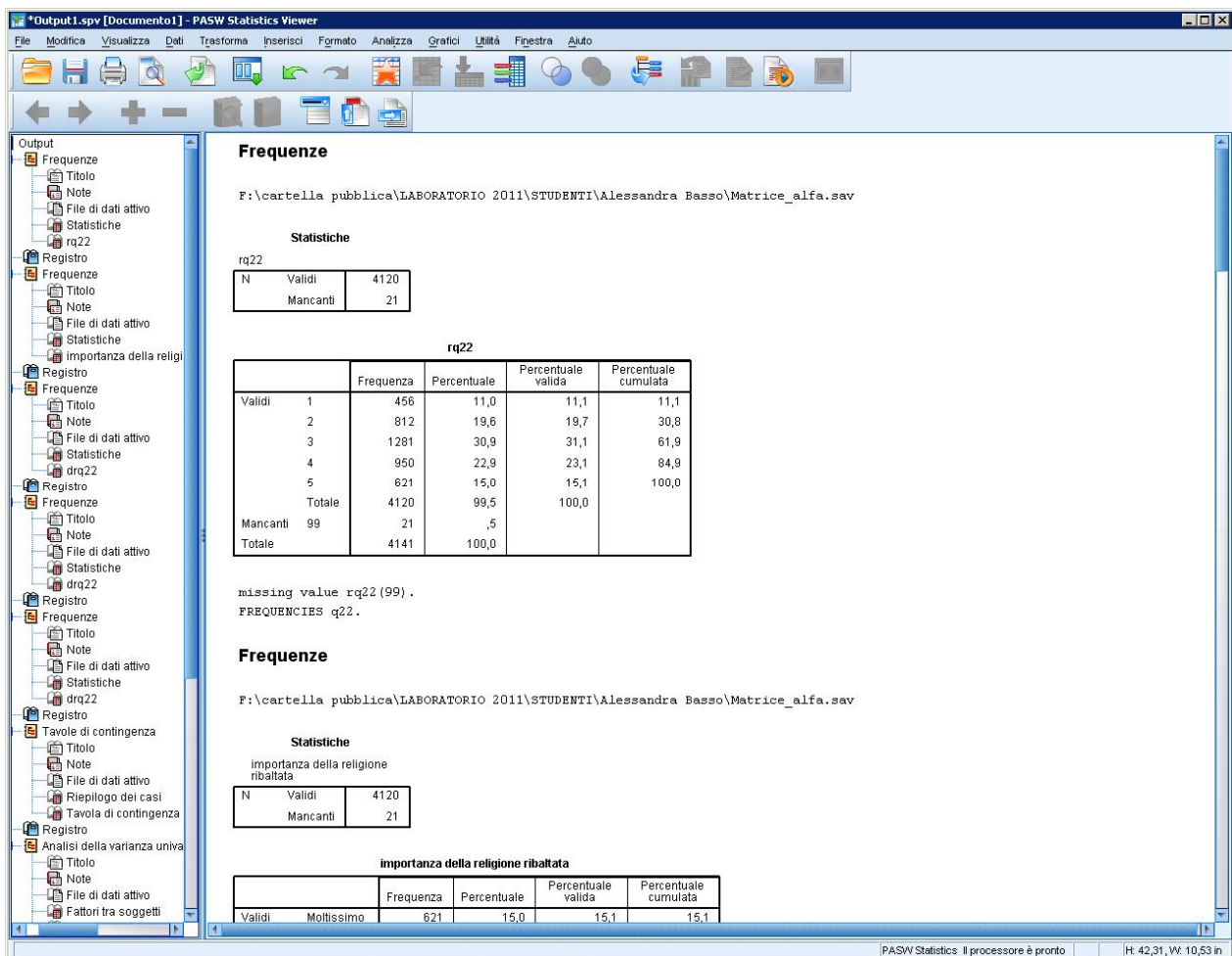



Figura 4: finestra di output.

Sintassi

La finestra di sintassi consente di dare i comandi a SPSS nel suo linguaggio di programmazione, detto SPSS Syntax. E' possibile selezionare con il mouse una determinata sequenza di comandi e mandarla in esecuzione mediante il tasto “run” . Copiando i comandi con il copia-incolla e cambiando il nome delle variabili nel testo si evita di dover digitare molte volte le sequenze di comandi ripetitive. Chiudendo la finestra di sintassi, SPSS chiederà se si desidera salvare il contenuto per poterlo riutilizzare per successive elaborazioni.

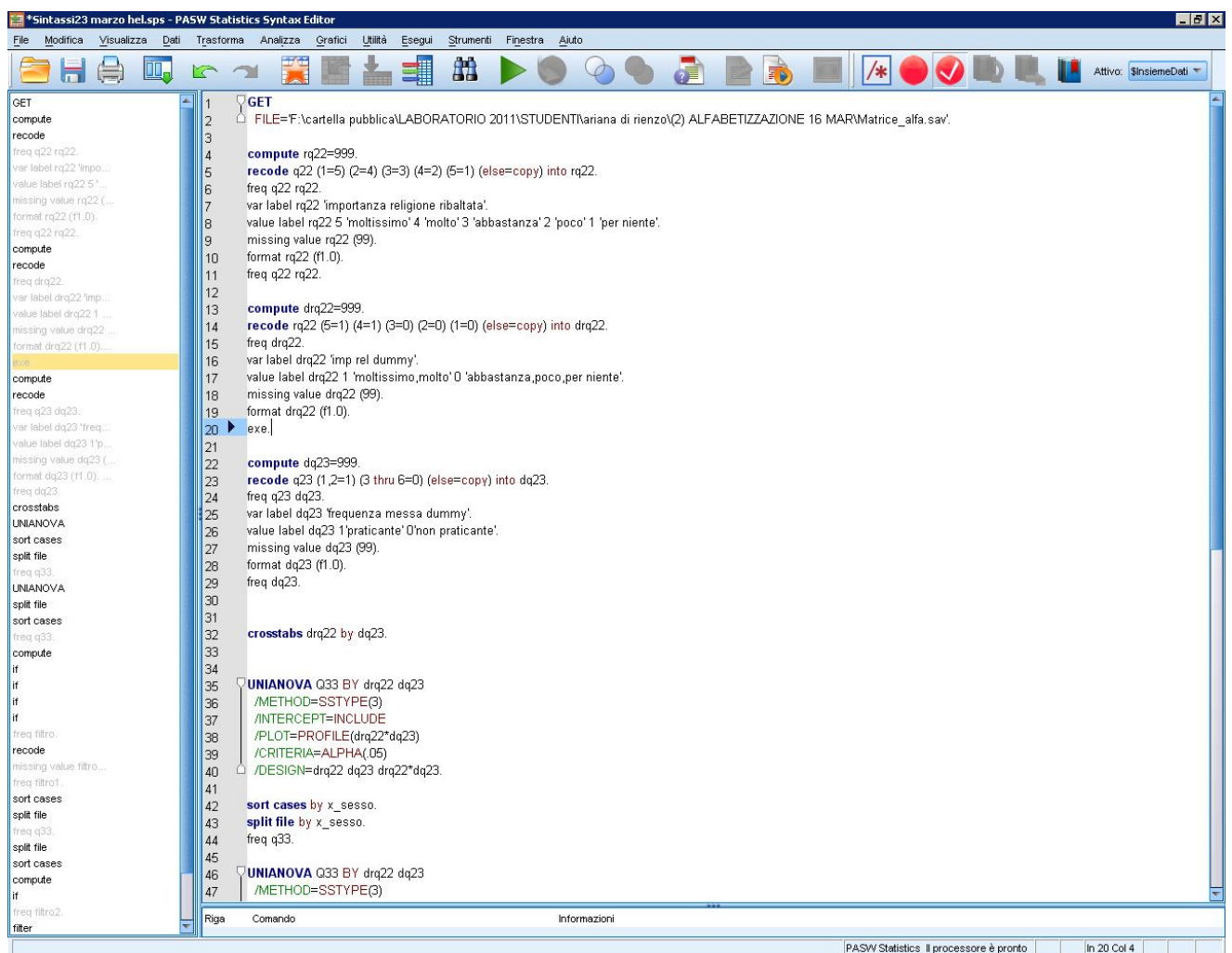


Figura 5: finestra di sintassi.

Come per le altre due finestre (editor e output) SPSS ci consente di salvare le sessioni in un file che ha estensione .sys, in modo da poter ripetere l'analisi successivamente. E' consigliabile

salvare il foglio di sintassi perché questo mi consente di avere memoria del lavoro svolto nelle sessioni precedenti.

Prima di lanciare un comando è necessario adottare alcune regole nella scrittura quali terminare ogni comando con un punto e separare i sottocomandi dai comandi principali con uno slash. In caso contrario nel foglio di output verrà esplicitata la tipologia di errore.

Comandi principali

I comandi principali che realizzo scrivendo direttamente nel file di Sintassi si possono condensare indicativamente sotto tre categorie:

1. OPERAZIONI SULLE MATRICI

- NEW FILE
- GET FILE= “ percorso nome file.sav”.
- GET FILE= “ percorso nome file.sav”
/ DROP x_ nome variabile.
- GET FILE= “ percorso nome file.sav”
/ KEEP x_ nome variabile.
- SAVE OUTFILE= “ percorso nome file.sav”.
- SAVE OUTFILE= “ percorso nome file.sav”
/ DROP x_ nome variabile.
- SAVE OUTFILE= “ percorso nome file.sav”
/ KEEP x_ nome variabile.

2. OPERAZIONI SULLE VARIABILI

Ovvero tutti i comandi utili per caricare i nostri dati nella maniera a noi più congeniale. Tra questi ci sono:

- VAR LABEL: comando utilizzato per dare un'etichetta discorsiva alla variabile e rendere intuitiva la sua lettura (in editor dei dati nel foglio “visualizzazione variabili”). L'etichetta può avere massimo otto caratteri.
Esempio: VAR LABEL x_ sesso “genere” nel caso si voglia ridefinire la nostra variabile x_ sesso in genere.

- **VALUE LABEL** : comando utilizzato per dare i codici alle variabili
Esempio: `VALUE LABEL x_sesso 1 "maschio" 2 "femmina"` nel caso si volesse dare l'etichetta maschio e femmina alle due modalità della variabile `x_sesso` che hanno codice 1 e 2.
- **FORMAT**: comando utilizzato per specificare il formato della nostra variabile.
Esempio: `FORMAT x_sesso (f 2.1)` nel caso ci fosse una variabile numerica che occupa al massimo due spazi di cui uno decimale.
Esempio: `FORMAT x_sesso (a 1.0)` nel caso ci fosse una variabile non numerica che occupa al massimo uno spazio.
- **MISSING VALUE**: comando utilizzato per specificare come sono stati codificati i valori mancanti.
Esempio: `MISSING VALUE x_sesso (99)` nel caso decidessi di assegnare il codice 99 a tutti i missing value. Generalmente si assegnano i valori 99, 98. Nel caso si lavorasse con la variabile ordinale "età" nella quale l' eventuale valore 99 o 98 potrebbe essere presente in matrice si utilizza 999 o qualsiasi altro numero che non può essere informativo dell'età.
- **COMPUTE**: comando utilizzato per creare una nuova variabile in matrice.
Esempio: `COMPUTE reddito = 999.`
- **RECODE**: comando utilizzato per ricodificare la mia variabile.
Esempio: `RECODE x_ts (1 2 3 = 1) (4 = 2) (5 = 3) (else = copy) into titolo studio.`
Nel caso si volesse ricodificare la variabile `x_ts` con soli tre codici, 1 2 3. In questo modo si condensano le prime modalità (1=nessun titolo; 2= elementari; 3=medie)in un solo livello.
- **CORR**: comando utilizzato per effettuare correlazioni fra due variabili cardinali.
Esempio: `CORR peso with altezza.`

3. OPERAZIONI SUI CASI

- **FILTER:** comando utilizzato se vogliamo fare una selezione dei casi da trattare in matrice. Esempio: FILTER by donne nel caso volessi lavorare solo su soggetti femmine. Una volta eseguito il comando tutti i casi maschi presenti sulle righe della mia matrice vengono sbarrati.
- **FILTER OFF:** comando utilizzato se vogliamo tornare a lavorare su tutti i casi della mia matrice.
- **SELECT:** comando utilizzato per effettuare la medesima operazione di FILTER però è irreversibile e quindi, una volta eseguito, i casi sbarrati vanno definitivamente persi.
- **SPLIT:** comando utilizzato per organizzare la matrice di partenza in tante sottomatrici quante sono le modalità della mia variabile split. Questo comando deve essere sempre preceduto da quello SORT CASES affinché la mia matrice sia riordinata in base alla variabile split.

Esempio: SORT CASES by x_ sesso.

SPLIT file by x_ sesso.

In questo caso ottengo due sottomatrici, una con tutti i casi maschio e una con tutti i femmina.

Per concludere vi vogliamo ricordare che questi appena elencati sono solo i principali comandi di SPSS, ma per chi avesse necessità di maggiori dettagli consigliamo di consultare la guida online.