Indice

1	L'Aı	nalisi dei dati con SPSS	1
	1.1	Introduzione	1
	1.2	Importazione di un file di dati	3
	1.3	Preparazione di un file dati	8
	1.4	"German Credit Database": descrizione dei dati	10
	1.5	Sintassi	13
		1.5.1 Data List	14
		1.5.2 Variables Lables	14
		1.5.3 Values Labels	15
		1.5.4 Missing Value	15

Capitolo 1

L'Analisi dei dati con SPSS

1.1 Introduzione

*Spss*¹ è uno dei più diffusi software per l'elaborazione statistica dei dati permettendo di gestire i dati in un ambiente grafico; esso utilizza menù descrittivi e semplici finestre di dialogo per la maggior parte delle procedure. In SPSS ci sono diverse finestre di dialogo:

- Edtor dei dati: questa finestra visualizza il contenuto del file di dati. Possono essere creati nuovi file di dati o modificarne già esistenti. La finestra Data Editor si apre automaticamente con l'avvio di SPSS e può visualizzare solo un set di dati alla volta;
- Viewer: questa finestra visualizza tutti i risultati dell'analisi e può essere salvata per un successivo utilizzo. La visualizzazione di questa finestra avviene in modo automatico ogni qualvolta si esgue la procedura che genera l'*out put*.
- **Bozza**: questa finestra visualizza l'*out put* come semplice testo invece che come tabelle pivot interattive;
- Editor di Tabelle Pivot: questa finetsra permette di modificare la visualizzazione dell'*out put*; ad esempio permette di modivicare il posizionamento delle variabili nelle tabella (colonne, righe..);
- Editor dei grafici: questa finestra permette di modificare colori, dimesioni, caratteri, gli assi di ungrafico, far ruotare i grafici 3-D e3 anche cambiare il tipo di tabella.
- Editor di output testuale: questa finestra permette di modificare quella perte di testo relativa all'*out put* non visualizzata nella finestra pivot table editor.
- Editor della sintassi: questa finestra visualizza in forma di comandi di sintassi tutti le procedure che possono essere selezionate nelle finestre di dialogo durante l'analisi dei dati. In questo modo è possibile utilizzare particolari caratteristiche di SPSS non ottenibili attraverso le finestre di dialogo; ovviamente questi comandi possono essere salvati ed utilizzati per successive analisi.

¹Statistical Package for the Social Sciences.

• Editor degli script: questa finestra permette di personalizzare e automatizzare molte procedure in SPSS.

I dati utilizzati con Spss possono derivare da varie fonti: questionari, dati già raccolti, statistiche ufficiali; esso importa file provenienti anche da altri programmi quali ad esempio *Excel* o *Dbase*. Ogni matrice dati, di per sè non significa nulla, se non comunichiamo al programma di che tipo di dati si tratta, come sono organizzati nella matrice. Quetso lavoro viene fatto costruendo un file, il DATA LIST, che tramite la sintassi di SPSS comunica al software tutte le informazioni necessarie per l'elaborazione e la comprensione dei risultati². Nella figura seguente sono illustrati i passaggi logici per arrivare dalla matrice di dati al system file SPSS:



Figura 1.1: Passaggi logici

Creato il file di sistema di SPSS, si lavora sempre con lo stesso per svolgere le opportune elaborazioni statistiche. Il programma si apre con la seguente schermata che chiede che cosa si vuole fare:



Figura 1.2: Schermata Iniziale

²Questa procedura pu`o essere eseguita anche in modo automatico senza la visualizzazione della sintassi (vedi paragrafo 1.2).

si può aprire una sorgente di file già esistente o aprire un altro tipo di file e quindi eventualmente procedere con l'mportazione guidata di un file.

Nel caso in cui si seleziona la voce annulla per aprire un file sarà necessario selezionare dal menù file la voce apri per accedere ad un file già esistente, od eventualmente la voce nuovo per generare un nuovo file (vedi Figura 1.3)in ogni caso le finestre che si possono aprire sono le seguenti:

Nome della finestra	Contenuto	Estensione del file
DATA	il system file spss	.sav
SYNTAX	Sintassi SPSS, DATA LIST, comandi	.sps
OUTPUT	Output dei comandi	.lst

File	Modifica Visualizza E	oati Trasfo	rma Analiz:	za (
N	uovo			
A	pri		Dati	
A	pri database		 Sintass 	i –
Le	eggi dati testuali		Output	
S	alva alva con nome	Ctrl+S	Script Altro	
			-	59,0
Vi In	isualizza informazioni dati. Iiserisci dati nella cache	148 1		77,0
		V25002507-1		76,0
S	tampa	Ctrl+P		54,0
A	nteprima di stampa		-	87,0
С	ambia server		3	69,0
Ir	iterrompi processore	Ctrl+.		63,0
D	ati recenti		•	30.0
Fi	le recenti		•	35,0
E	sci			78 N

Figura 1.3: Menù

1.2 Importazione di un file di dati

La procedura da seguire per la creazione del file system è la seguente dal menù file si selezioni la voce apri quindi la voce dati si accede cos[']i alle finestre per l'importazione guidata del file. Questa procedura prevede sei passi qui di seguito descritti:

• nel primo passo il file viene visualizzato in una finestra di anteprima; è possibile scegliere di applicare un formato predefinito³ o eseguire i passi della procedura guidata per definire il metodo di lettura dei dati (vedi Figura 1.4);

³Il formato devessere stato salvato precedentemente.

840 1 81 28.5 2400 0 73 40 33 10200 0 83 31.08 870 0 93 31.17 17400 83 41.91	Benvenu Questa ir di testo e	ti nella Importazion mportazione guida I la specífica delle	ne guidata di testo ta semplifica la letti informazioni sulle ∨	ıra dei dati del file ariabili.
var1 var2 var3	- Il file di	testo è in un forma	to predefinito?	
3		si		Stoglia
		No		
e di testo: C:\Documer Q	nts and Settings\giusep	pe\Desktop\carme	la\icu.dat	
di testo: C\Documer 914 14	nts and Settings∖giusep _i 020 1	oe\Desktop\carme 3,0 8 7	ila\icu.dat 40 1	<u>50</u>
e di testo: C:\Documer <u> </u>	nts and Settings\giusep; 02,0 1 0	5e\Desktop\carme 30 87 27	la∖icu.dat 401 _1	<u>50</u>
e di testo: C\Documen 1 4 3 12	nts and Settings\giusepp 02.0 1 0 0	0e\Desktop\carma 30 87 27 59	la∖icu.dat 401 0	<u>50</u> 1 1 1
di testo: C\Documei	nts and Settings\giusep 020 0 0 0 0	28\Desktop\carms 30	la∖icu.dat 401 1 0 0	1 1 1 1 1
di testo: C\Documen 4 8 12 14	nts and Settings\giusep 020 1 0 0 0	00)Desktop\carma 30	la\icu.dat <u>40</u> 1 1 0 0	<u>50</u> 1 1 1 1 •

Figura 1.4: Importazione guidata di testo:passo 1

- nel secondo passo si devono fornire le informazioni sulle variabili. Per leggere correttamente i dati, durante l'importazione guidata si deve specificare dove finiscono i valori dei dati di una variabile e iniziano i valori di quella successiva. La disposizione delle variabili determina il metodo utilizzato per separare una variabile dalla successiva). I dati possono essere:
 - Delimitati: separati da spazi, virgole, tabulatori o altri caratteri . Le variabili di ciascun caso vengono registrate nello stesso ordine, ma non sempre nelle stesse posizioni di colonna;
 - Larghezza fissa: ciascuna variabile viene registrata nella stessa posizione della colonna e nella stesso record (riga) per ciascun caso del file di dati. Non è richiesto nessun delimitatore tra le variabili. In molti file di dati di testo generati automaticamente potrebbe sembrare che i valori dei dati siano mischiati insieme senza alcun simbolo di divisione; la posizione della colonna determina la variabile che verrà letta;

Inoltre è necessario specificare se la prima riga del file di dati contiene etichette descrittive che si vogliono utilizzare come nomi di variabili. Le etichette che includono più di otto caratteri verranno troncate. Se i primi otto caratteri non formano un nome univoco di variabile, il nome della variabile verrà modificato come necessario (vedi Figura 1.5);

		Dime				
O Delimit	tate -	- Variabili delimitat	e da caratteri spe	cifici (ad es. virgole	e o tabulatori)	
C Larghe	zzafissa -	- Variabili allineate	in colonne a largh	iezza fissa		
omi delle v	ariabili compaic	ono all'inizio del file	۶			
C SI						
No						
e di testo: í	Nocuments at	nd Settings\diusen	ne\Deskton\carm	ala\iru dat		
e di testo: (Q	C\Documents ar	nd Settings\giusep	pe\Desktop\carm	ela\icu.dat 40		
e di testo: (Q	C\Documents ar	nd Settings\giusep 20 1	pe\Desktop\carm 30 87	ələ\icu.dət 40 1	<u>50</u> 1	•
e di testo: (0 1 2	C\Documents ar 1_0 4 8	nd Settings\giusep _i 2 <u>10</u> 1 0	pe\Desktop\carm 30 87 27	ela\icu.dat 401 _1	<u>5,0</u> 1 1	
e di testo: (<u>Q</u> 1 2 3	C\Documents ar 10 4 8 12	nd Settings\giusep 	pe\Desktop\carm 30 87 27 59	ela\icu.dat <u>,40</u> 1 1 0	<u>50</u> 1 1 1	-
e di testo: (0 1 2 3 4	2\Documents ar 1_0 4 8 12 14	nd Settings\giusep 2 1 0 0 0 0	pe\Desktop\carm 30 87 87 27 59 59 77	ela\icu.dat 40 1 1 0 0		•

Figura 1.5: Importazione guidata di testo:passo 2

• nel terzo passo devono essere fornite informazioni sui casi a, in particolare si devono specificare le seguenti voci (vedi Figura 1.6):

	a cui inizia il primo	caso di dati	0	<u></u>		
appresente	zione dei casi —					
💿 Ogni ri	ga rappresenta un	caso				
🔿 Un nur	nero di variabili rap	opresenta un cas	o: 1	•		
lumero di ca	asi da importare —					
 Tutti i d 	casi					
C Lorimi	1000	casi				
C Davas	i i e e e e e e e e e e e e e e e e e e			• •		
(Percei	nuale casuale di c	asi (approssimai	ivaj. jiu	· /*		
nteprima da	ati					
Q		1	87	1	1 🔺	
<u>Q</u>	4		0.		1	
1 2	4 8	ō	27	1		
9 1 2 3	4 8 12	0 0	27 59	1 0	1	
Q	4 8 12 14	0	27 59 77	1 0 0		
Q 1 2 3 4 ▲	4 8 12 14	0 0 0	27 59 77	1 0 0		

Figura 1.6: Importazione guidata di testo:passo 3

- Numero di riga a cui inizia il primo caso dati: indica la prima riga del file di dati che contiene valori. Se le prime righe del file di dati contengono etichette descrittive o testo di altro tipo che non rappresenta un valore, questa non sarà la riga 1;
- Numero di righe che rappresentano un caso: nel caso di file a larghezza fissa controlla il modo in cui si determina ogni caso e l'inizio del successivo. Ogni riga è definita dal relativo numero di riga all'interno del caso e della posizione della colonna; è necessario specificare il numero di righe per ogni caso per poter leggere correttamente i dati;
- Numero di casi da importare: è possibile importare tutti i casi del file di dati, solo i
 primi n casi o un campione casuale determinato in base a una percentuale specificata.

Poichè viene eseguito un processo indipendente di decisione pseudo-casuale per ciascun caso, l'equivalenza tra la percentuale di casi selezionati e la percentuale specificata può essere solo approssimativa. Maggiore è il numero di casi iclusi nel data file e maggiore sarà l'approssimazione della percentuale di casi selezionati rispetto alla percentuale specificata;

 nel quarto passo, è illustarta la soluzione più opportuna per leggere un file di dati e consente di modificare il metodo impiegato per leggere le variabili dal file di dati. Quando il file è a larghezza fissa le linee verticali nella finestra di anteprima indicano le posizioni in cui la procedura ipotizza possa iniziare una variabile. Si possono inserire, spostare ed eliminare linee di interruzione delle variabili in base alle necesità. Quando il file è delimitato è necessario specificare i delimitatori tra le variabili (spazi, virgole,...)e se sono presenti più delimitatori consecutivi, questi sono considerati valori mancanti; inoltre è necessario specificare il qualificatore di testo (vedi Figura 1.7);

Delimitatori tra le variabili		✓ Spazio ✓ Punto e virgo	✓ Spazio Punto e virgola		Nessuna Virgolette doppie Virgolette singole Altro:			
Altro:								
Altro: Anteprima dati	V2	V3	V4		V6	1	V7	
Anteprima dati	V2	87 87	1 1	V5	1 1	0	V7	
Anteprima dati	V2	87 27 59	V4	V5	1 0	0	V7	
Anteprima dati	1 0 0	V3 87 27 59 77	V4	V5 1 1 1	1 0 0		V7	

Figura 1.7: Importazione guidata di testo:passo 4

• nel quinto passo si controlla il nome della variabile e il formato dei dati usati in modo da leggere le variabili e scegliere quelle che saranno incluse nel file di dati finale (vedi figura 1.8);

Nome var	iahile:							
√1								
-								
Formato d	latı:		-					
Numericc	0							
teorima da	sti							
teprima da	ati	1 10	1	1	1 10			
teprima da V1	ati V2	V3	V4	V5	V6		<u>V7</u>	
teprima da V1	ati V2	87 87		1 1	1 1	0	<u>V7</u>	
teprima da V1	ati 1 0	87 27 29	1 1 1	V5	V6	0	<u>V7</u>	
teprima da V1	ati 1 0 0	V3 87 27 59 77	1 1 0	1 1 1 1	V6 1 0 0	0	<u>V7</u>	
teprima da V1	ati V2 1 0 0 0	V3 87 27 59 77	V4 1 0 0	V5 1 1 1 1	V6 1 0 0 1		<u>V7</u>	

Figura 1.8: Importazione guidata di testo:passo 5

• nel sesto ed ultimo passo è possibile salvare le impostazioni in un file da utilizzare per l'importazione di file di dati di testo simili. Inoltre si può incollare la sintassi generata e quindi poterla personalizzare e salvare in modo da poterle utilizzare in altre sessioni o durante i processi di produzione (vedi figura 1.9).

wart x 1 628 2 630	ar2 9ar3 9 840 1 2400 0	Se	Ivare il formato C Sì	di file per usi fu	ituri?	Salva :	on nom	ie
3 632 4 633	870 0 •	lr	icollare la sinta (^S)	ssi?		Copia i dati in lo	cale	
			No	Una.co	pia cache dei c	lati è una copia	dell'int	ero file
Anteprima da	ati	Pren	nere Fine per c	omple file di dati, i tempor file di d	memorizzata in aneo. La creazi ati può migliora	uno spazio dis one di una cop re le prestazior	co ia cach il	e del
Anteprima da V1	ati	Pren	nere Fine per c	omple file di dati, tempor file di d	memorizzata in aneo. La creazi ati può migliora V6	uno spazio dis one di una cop re le prestazior V7	co ia cachi i.	e del
Anteprima da V1 4	ati V2	Pren	nere Fine per c	binditi, i tempor file di di V5	memorizzata in aneo. La creazi ati può migliora V6	uno spazio disi one di una cop re le prestazior V7	co ia cachi i.	e del
Anteprima da V1 4	ati V2 1 0	Pren	nere Fine per c	di dati, i tempor file di d V5	memorizzata in aneo. La creazi ati può migliora V6 1 0.	uno spazio dis one di una cop re le prestazion 0	ia cachi ii.	e del
unteprima da V1 4 B 12	ati 1 0 0	Pren 87 27 59	nere Fine per c	di dati, i tempor file di d V5	memorizzata in aneo. La creazi ati può migliora V6 1 0 0	uno spazio disi one di una cop re le prestazion 0 0 0	ia cach i. 0 0	e del
unteprima da V1 4 8 12 14	ati 1 0 0 0	Pren 87 27 59 77	nere Fine per c V4 1 0 0	V5	memorizzata in aneo. La creazi ati può migliora 1 0 0 1	uno spazio dis: one di una cop re le prestazion V7 0 0 0 0 0	ia cach i. 0 0 0	e del
Anteprima de V1 4 8 12 14 77	ati V2 1 0 0 0	Pren 87 27 59 77 77 77	Nere Fine per c V4 1 0 0	di dati, tempor file di d file di d 1 1 1 1 1 1 1 1 1 1 1	memorizzata in aneo. La creazi ati può migliora 1 0 0 1 1	uno spezio dis: one di una cop re le prestazion V7 0 0 0 0	ia cach i. 0 0 0 0	e del
Anteprima da V1 4 8 12 14 27 <	ati 1 0 0 0 1 1 1	Pren 87 27 59 77 77 72	V4 1 1 0 0	di dati, tempor file di d file di d 1 1 1 1 1 1 1 1 1 1 1 1 1	memorizzata in aneo. La creazi ati può migliora 1 0 0 1 1	uno spezio dis: ne di una cop re le prestazion V7 0 0 0 0	ia cachi ii 0 0 0	e del

Figura 1.9: Importazione guidata di testo:passo 6

Completate le operazioni di esportazione il file di dati si presenterà come in figura 1.11.

🧰 Senza tit	olo - Editor	dei dati SPSS						
File Modifica	Visualizza D	ati Trasforma	Analizza Gra	fici Strumenti	Finestra ?			
2 8 8 0		1 12 14		5 90				
1:v1		4						
	v1	v2	V3	v4	v5	v6	v7	vE_
1	4,00	1	87,0	1	1	1	0	
2	8,00	0	27,0	1	1	0	0	•
Visua	alizzazione d	l ati (Visualiz:	zazione varia	ibili 🖌 📢		li.	le.	
		SP	SS Processore	pronto				

Figura 1.10: Editor dei dati.

1.3 Preparazione di un file dati

Come si può notare dalla figura 1.11, l'editor dei dati, è un file costituito da due fogli in uno sono visualizzati i dati e nell'altro le varibili (vedi figura 1.10).

🛄 Senza	a titolo ·	- Editor de	i dati SPSS								×
File Mod	difica Visi	ualizza Dati	Trasforma /	Analizza Gr	afici Strum	ienti Fin	estra ?				
	s 🔍 -		一的州相	ř 🛛 🗗	 						
	Nome	Tipo	Lunghezza	Decimali	Etichetta	Valori	Mancante	Colonne	Allinea	Misura	
	1 v1	Numerica	3 🛟	2	01	Nessun	Nessuno	8	Destra	Scala	
	2 v2	Numerica	1	0	-	Nessun	Nessuno	8	Destra	Nominale	
	3 V3	Numerica	2	1	-	Nessun	Nessuno	8	Destra	Scala	1
< > \ \	alua /isualizz	azione dati	√Visualizza	in zione vari	iabili /	Moccup	Maccupa	Ĵ•	Doctro	Nominala	
			SPSS	Processore	pronto	j.					1

Figura 1.11: Editor dei dati.

Il foglio relativo alle variabili è costituito da dieci campi che devono essere impostati per ogni variabile, ogni record corrisponde ad una variabile. Gli attributi da definire per ogni variabile coinvolta nell'analisi sono i seguenti:

- nome: in questo campo deve essere inserito il nome della variabile se non è stato già inserito durante l'importazione guidata del testo.
- tipo: in questo campo si specifica il tipo di variabile numerica, data... La selezione di questa caratteristica avviene attraverso la finestra di dialogo Tipo di varibile che è automaticamente aperta quando si clicca sulla cella (vedi Figure 1.12).

Fipo di variabile		?
 Numerica Virgola Punto In notazione scientifica Data Dollaro Valuta personalizzata Stringa 	Lunghezza: 3 Decimali: 2	OK Annulla Aiuto

Figura 1.12: *Tipo di variabile*.

- lunghezza:
- decimali: in questo campo si definiscono inumeri decimali per le variabili numeriche;
- etichetta: in questo campo si assegnano le etichette alle variabili, se ne da cioè una descrizione;
- valori: in questo campo si possono assegnare valori descrittivi per ogni valore assunto da una variabile. Questo campo risulta particolarmente utile quando si lavora con file con codici numerici per variabili non numeriche; per accedere alla finestra di dialogo Etichette dei dati si procede come mel caso del campo tipo (vedi Figura 1.13);

Etichette dei valori	? 🔀
Etichette valori	ок
Valore:	Annulla
Acciunal	Aiuto
Cambia	
Rimuovi	

Figura 1.13: Etichette dei dati.

- mancante: in questo campo si possono definire specificati valori dei dati come valori mancanti
- colonne: in questo campo si specifica il numero di caratteri per la larghezza della colonna;
- allinea: in questo campo si imposta l'allineamento delle variabili nel foglio dei dati, di defoult l'allineamento è a destra;

• misura: in questo campo si definisce la scala della variabile che si seleziona cliccanndo sulla freccia verso il basso che è visualizzata nel momento che si seleziona la cella di interesse (vedi Figura 1.14);



Figura 1.14: Misuar della variabile.

1.4 "German Credit Database": descrizione dei dati

Si consideri il set di dati qui di seguito descritto e si importi e prepari il file system di SPSS con le procedure appena decsritte. Il file di dati german.dat è composto da 1000 osservazioni che individuano soggetti che hanno richiesto un credito; 700 individui si sono rivelati solventi buoni pagatori, mentre 300 si sono rivelati insolventi cattivi pagatori. I primi costituiscono il gruppo "*good*" contrassegnato con 1, i secondi il gruppo "*bad*" contrassegnato con 2. I crediti in questione sono di piccolo importo (importo massimo circa pari a 16000 DM) e breve durata (durata massima 60 mesi). Le informazioni relative ai richiedenti credito sono descritte da 20 variabili (attributi), 7 continune e 13 categoriche.

Le varibili sono così definite:

- Variabile X_1 : Saldo attuale del conto corrente, variabile categorica ordinale a 4 modalità:
 - $-A_{11}: X_1 < 0DM$
 - $A_{12}: 0 \le X_1 < 200 DM$
 - A_{13} : $X_1 \ge 200 DM$ / salario ricevuto per almeno un anno
 - A_{14} : nessun conto corrente
- Variabile X_2 : Durata del prestito in mesi, variabile quantitativa
- Variabile X_3 : Storia del credito, variabile categorica nominale a 5 modalità
 - A_{30} : nessun credito richiesto/tutti i crediti pagati in tempo
 - A_{31} : tutti i crediti richiesti a questa banca pagati in tempo
 - A_{32} : crediti attuali fin ora pagati in tempo
 - A_{33} : ritardi nei pagamenti in impegni passati
 - A_{34} : conto critico/altri crediti esistenti (non in questa banca)
- Variabile X_4 : Scopo della richiesta di credito, variabile categorica nominale a 11 modalità
 - A_{40} : auto (nuova)
 - A_{41} : auto (usata)

- A_{42} : mobilio/attrezzature
- A_{43} : radio/televisione
- A_{44} : elettrodomestici
- A_{45} : riparazioni
- A_{46} : istruzione
- A_{47} : vacanza
- A_{48} : riqualificazione
- A_{49} : affari
- A_{410} : altro
- Variabile X_5 : Ammontare del credito richiesto, variabile quantitativa
- Variabile X_6 : Tipologia conto di risparmio/obbligazioni, variabile categorica ordinale a 5 modalità
 - $-A_{61}: X_6 < 100DM$
 - $A_{62}: 100 \le X_6 < 500 DM$
 - $A_{63}: 500 \le X_6 < 1000 DM$
 - $-A_{64}: X_6 \ge 1000DM$
 - A₆₅: nessun conto di risaprmio/sconosciuto
- Variabile X_7 : Anni di assunzione dell'attuale impiego, variabile categorica ordinale a 5 modalità
 - A_{71} : disoccupato
 - $-A_{72}: X_7 < 1$ anno
 - A_{73} : $1 \le X_7 < 4$ anni
 - A_{74} : $4 \le X_7 \le 7$ anni
 - $-A_{75}: X_7 \ge 7 \text{ anni}$
- Variabile X_8 : Stima della rata in percentuale all'entrata disponibile, variabile quantitativa
- Variabile X_9 : Stato civile e sesso, variabile categorica nominale a 5 modalità
 - A₉₁: maschio: divorziato/separato
 - A₉₂: femmina: divorziata/separata/sposata
 - A_{93} : maschio: single
 - A_{94} : maschio: sposato/vedovo
 - A_{95} : femmina: single

- Variabile X_{10} : Altri debitori/garanti, variabile categorica nominale a 3 modalità

- A_{101} : nessuno
- A_{102} : co-richiedente
- A_{103} : garante
- Variabile X_{11} : Anni dell'attuale residenza, variabile quantitativa
- Variabile X_{12} : Proprietà, variabile categorica nominale, 4 modalità
 - A_{121} : beni immoboli
 - A_{122} : se non è A_{121} : società di credito edilizio/assicurazione sulla vita
 - A_{123} : se non è $A_{121}oA_{122}$: auto o altro, non presente nella variabile 6
 - A_{124} : sconosciute/nessuna proprietà
- Variabile X_{13} : Età, variabile quantitativa
- Variabile X_{14} : Altri schemi di pagamento, variabile categorica nominale a 3 modalità
 - A_{141} : banca
 - A_{142} : negozi
 - A_{143} : nessuno
- Variabile X_{15} : Casa di abitazione, variabile categorica a tre modalità
 - A_{151} : affitto
 - A₁₅₂: di proprità
 - A_{153} : ospite
- Variabile X_{16} : Numero di crediti attualmente in questa banca, variabile quantitativa
- Variabile X_{17} : Lavoro, variabile categorica nominale a 4 modalità
 - A_{171} : disoccupato/non-specializzato, non-residente
 - A_{172} : non specializzato residente
 - A173: impiegato specializzato/funzionario
 - A_{174} : dirigente/autonomo/impiego altamente qualificato/ufficiale
- Variabile X_{18} : Numero di persone a carico, variabile quantitativa
- <u>Variabile X_{19} </u>: Telefono, variabile categorica nominale a 2 modalità
 - $-A_{191}:$ no
 - $-A_{192}$: si
- Variabile X_{20} : Lavoratore straniero, variabile categorica nominale a 2 modalità
 - $-A_{201}$: si
 - $-A_{202}$: no.

1.5 Sintassi

Nel paragrafo 1.2 si è visto come contemporaneamente alla finestra di editor dei dati, attraverso la copia della sintassi, si può ottenere il file di sintassi; in figura 1.15 si osserva una finestra di editor di sintassi.

📓 Syntax1 - Editor della sintassi S	PSS				
File Modifica Visualizza Analizza Gra	fici Strumenti	Esegui	Finestra	?	
GET DATA /TYPE = TXT	Esegui corrente				^
$/FILE = 'C: \Lambda_{i}$					
/DELCASE = LINE					
/DELIMITERS = " "					
/ARRANGEMENT = DELIMITED					
/FIRSTCASE = 1					
/IMPORICASE = ALL					
VARIABLES =					
V1F3.2 V2 F1 0					
V2 F1.0 V2 F2 1					
V4 E1 0					
V5 F1 0					
V6 F1 0					~
Esegue il comando corrente	📍 SPSS Proces	ssore pro	onto		1

Figura 1.15: Editor di sintassi.

In SPSS sono presenti vari comandi e funzioni che possono essere suddivisi in tre principali categorie (vedi tabella):

- 1. **Comandi di definizione dei dati**, che forniscono informazioni sule variabili, il tipo e la loro locazione nel file di dati;
- 2. **Comandi di trasformazione dei dati**, che sono usati per creare nuove variabili, modificare quelle esistenti e selezionare un sottocampione di casi per l'analisi;
- 3. Comandi che attivano procedure, come le statistiche, i grafici, il salvataggio dei dati.

Categoria	Comandi
Definizione	data list, variable labels, value labels, missing value
Trasformazione	compute, recode, if, temporary, select if
Procedure	Frequencies, Crosstabs, Save outfile

In questo capitolo saranno trattati i principali comandi per la definizione di dati in modo da poter generare e preparare il file system di spss anche dall'editor della sintassi. Questa operazione richiede:

• l'impostazione del DATAT LIST, il comando, cioè che dice al programma dove i dati grezzi possono essere trovati e definisce per ciascuna variabile il nome e il numero iniziale e finale delle colonne occupate. Il DATA LIST è indispensabile: senza di esso il programma non riesce a distinguere tra loro le diverse variabili;

- la definizione delle VARIABLES LABELS, il comando, cioè, che permette di attribuire a ciascuna variabile un'etichetta esplicativa tale da permettere una migliore leggibilità dei risultati delle elaborazioni.
- la definizione delle VALUE LABELS, il comando che fornisce etichette descrittive per i valori delle variabili.

1.5.1 Data List

Esistono due tipi di DATA LIST: quello per file di dati in formato fisso e quello per file di dati in formato delimitato.

- *Data List Fixed* si usa quando i valori di ogni variabile sono registrati, nel file di dati grezzo, nella stessa posizione per ogni caso;
- *Data List Free* si usa quando le variabili sono registrate, nel file di dati grezzo, nello stesso ordine per ogni caso, ma non necessariamente nella stessa posizione .

Per una corretta esecuzione il comando deve essere scritto nel seguente ordine: nella prima riga deve essere digitato il nome del comando seguito dal *path* del file che contiene i dati grezzi e dal parametro FIXED o FREE in funzione del di formato dei dati grezzi. Nelle righe successive si assegnano i nomi alle variabili, il numero iniziale e finale delle colonne occupate ed il tipo di variabile. Più precisamente devono essere rispettate le seguenti condizioni:

- al nome delle variabile numeriche, che contengono cioè solo numeri, segue l'indicazione delle colonne occupate (se le colonne sono 2 o più è sufficiente indicare la prima e l'ultima separate da un trattino,-) ed eventualmente il numero di decimali tra parentesi;
- le variabili alfanumeriche o di tipo stringa, che possono contenere una combinazione di lettere e numeri, e altri caratteri sono contraddistinte da una (A) posta dopo l'indicazione;
- il punto, ., alla fine di questo set di istruzioni è obbligatorio.

In figura 1.15 si osserva un esempio di data list.

1.5.2 Variables Lables

Questo comando assegna un'etichetta descrittiva (label) ad ogni variabile. Le etichette possono comprendere spazi bianchi e qualsiasi altro carattere, ogni etichetta può raggiungere fino a 256 caratteri. Anche in questo caso il nome del comando deve essere digitato nella prima riga e poi seguito dai nomi delle variabili e le relative descrzioni, le uniche accortezze che devono essere poste sono le seguenti:

- i nomi delle variabili devono coincidere con quelli del DATA LIST;
- ogni variabile va separata dalla successiva con uno slash;
- un'etichetta non pu essere spezzata in pi righe;
- l'inizio e la fine delle label deve essere segnalata con le virgolette.

1.5.3 Values Labels

Dopo aver definito la lista delle variabili con il DATA LIST e le etichette (label) delle variabili con VARIABLE LABEL non resta che attribuire delle etichette a ciascun valore di ogni variabile con il comando VALUE LABELS.Anche in questo caso il nome del comando deve essere digitato nella prima riga e poi seguito dai valori da attribuire alle relative variabili, le uniche accortezze che devono essere poste sono le seguenti:

- ogni value labels non deve superare i 60 caratteri;
- il nome delle varaibile indicata deve coincidere con quello delle variabel labels e del data list;
- ogni variabile va separata dalla successiva con uno slash /;
- 1'inizio e la fine delle label segnalata con le virgolette.

1.5.4 Missing Value

Qualche volta l'informazione per un particolare caso non è utilizzabile si pu allora assegnare, con il comando MISSING VALUE, valori per identificare le mancate risposte. Ad esempio con il comando: MISSING VALUE PROVA (0) attribuisce al valore 0 della variabile PROVA il significato di valore mancante per l'utente.