

Giovanni Di Franco, *L'analisi dei dati con SPSS. Guida alla programmazione e alla sintassi dei comandi*, FrancoAngeli, Milano, Marzo 2009, 176 pagine; 22,00 Euro

Ricordo, fin da quando ero bambino, che mio padre, eccellente professore di lettere di liceo, non si occupava mai di quei lavori manuali che spesso capita di fare in casa, forse per non combinare guai.

All'occorrenza veniva in nostro aiuto un carissimo amico di famiglia al quale piaceva lavorare soprattutto con il trapano elettrico (è stato uno dei primi a possederne uno), anche quando questa soluzione risultava assolutamente spropositata; e, qualche volta, il danno era garantito.

Questa riflessione mi è tornata alla mente leggendo l'utilissimo libro di Giovanni Di Franco *L'analisi dei dati con SPSS* per la collana "La cassetta degli attrezzi" di Franco Angeli.

Spesso mi è capitato di sentire di persone, anche sufficientemente erudite, utilizzare questo software, principalmente adatto all'analisi statistica dei dati, per gestire una banca dati, oppure per la realizzazione grafica di una qual-

che procedura di gestione dati; insomma, SPSS come la soluzione finale per tutto.

Invece SPSS è uno tra gli strumenti in mano al ricercatore che deve essere usato con raziocinio e competenza; le sue enormi potenzialità devono essere dosate in base alle esigenze e al contesto nel quale si deve agire.

Continuando con la similitudine, a differenza di come il nostro caro amico di famiglia procedeva, va chiarito prima di tutto quale sia il lavoro che devo fare e cosa devo ottenere; poi e solo dopo mi devo chiedere se l'utilizzo del trapano elettrico sia necessario ai miei scopi. Quindi, secondo la superficie da trattare, il secondo passo è scegliere quali sono le punte più idonee da utilizzare, tra quelle a disposizione, adatte a perforare il ferro o il legno e quelle adatte per muri e pareti; infine la misura della punta da utilizzare (di solito, in un primo momento, è consigliabile andare per gradi, utilizzandone una di spessore minore al risultato atteso). Inoltre devo tener conto anche della velocità di perforazione da utilizzare e... giustamente l'autore propone questo patto con il lettore (pagg. 11-12) per evitare che gli utilizzatori di SPSS

Rassegna italiana di valutazione, a. XIII, n. 43/44, 2009

“più che usare un programma ne [siano] usati. Ciò è vero soprattutto per i programmi applicativi [...] che richiedono agli utenti competenze specialistiche al di là delle più o meno complesse competenze informatiche. [...] Per una corretta utilizzazione del [programma] è comunque necessaria una preparazione di base sia statistica sia di metodologia e tecnica della ricerca”.

La prima parte l'autore la spende per illustrare la storia e l'evoluzione del pacchetto SPSS e le condizioni minime per il suo funzionamento sia da parte della macchina dove è installato sia da parte dell'utilizzatore che deve avere quelle competenze minime d'informatica.

Utile anche la veloce carrellata sulle finestre e sottofinestre windows del programma; la distinzione tra comandi operativi, di definizione, manipolazione e trasformazione dei dati e i comandi di procedura statistica e grafica (cuore del pacchetto).

Durante ogni sessione SPSS genera diversi tipi di file, gli essenziali, per l'utente finale, sono: i *file dati*, vale a dire il file d'origine da cui scaturiscono le elaborazioni; *il file di sintassi* nel quale sono registrati tutti i comandi e *il file di output*, il file dei risultati ottenuti dalle elaborazioni che possono essere tabelle, grafici o entrambi.

L'autore avverte, se si vuole rispettare il patto iniziale, di prestare particolare attenzione al file di sintassi: il file dei comandi; imparare a conoscere i comandi risulta fondamentale per cominciare a rendersi autonomi rispetto al programma e tornare ad essere protagonista delle elaborazioni che si vanno a compiere e delle scelte che si fanno. È particolarmente utile il fatto di salvare tale file per poter ripetere e riu-

tilizzare, anche a distanza di tempo, i comandi impartiti anche in nuove situazioni con altri dati, naturalmente con gli opportuni aggiustamenti “personalizzati” e “personalizzabili” che si sarà ormai in grado di compiere.

Viene fornita anche una serie di suggerimenti su come creare la matrice dei dati, come definire le variabili ecc. in modo da far risultare più facile, poi, l'utilizzo nelle elaborazioni.

Naturalmente SPSS risulta compatibile con tutte le principali basi dati (Excel prima di tutto): per l'importazione è sufficiente rispettare delle semplici regole (che nel testo vengono richiamate e illustrate).

Nella prima parte l'autore, una volta descritto il programma, l'ambiente, le regole e forniti alcuni comandi di sintassi e disponendo di una matrice di dati, “autorizza” ad un immediato utilizzo coloro che vogliono subito iniziare a generare qualcosa con SPSS: per esempio l'analisi delle frequenze di una certa variabile presente nella matrice dei dati.

Il programma nella prima finestra del menù a tendina propone la scelta dei *grafici* da produrre (ma qual è quello più adatto al dato che devo andare a rappresentare?); le *statistiche* (ma quali sono quelle di cui ho bisogno?); la scelta dei *valori percentili*; *la tendenza centrale* (cosa rappresenta nel mio caso?); *la dispersione* (quale è più adatta ai dati che sto analizzando?); *la distribuzione* (e meglio la curtosi o l'asimmetrica?).

Torniamo al punto di partenza e alle domande iniziali: ho tutti gli elementi e le competenze necessarie per fare le scelte giuste?

Con il pacchetto SPSS sono forniti voluminosissimi manuali per ogni

argomento, anche in italiano, ma ormai banalizzati e concentrati sulle illustrazioni dettagliate delle operazioni piuttosto che sulla teoria statistica. Invece, per esempio, i vecchi manuali della versione 4 (siamo ormai alla versione n.17), che era ancora in DOS, restano, anche se solo in inglese, degli ottimi trattati e testi di statistica (se possibile, se ne consiglia la lettura).

Naturalmente si può lavorare anche con i menù a tendina e con le finestre di dialogo ma alcune opzioni e alcuni comandi sono disponibili solo attraverso l'uso della sintassi a cui l'autore dedica tutto l'intero capitolo 3.

È consigliabile, a questo punto, aprire il pacchetto SPSS con dei propri dati e provare a scrivere alcune righe di sintassi, salvando, di volta in volta, le soluzioni che più soddisfano le nostre esigenze. Oltre a fare una utilissima pratica dello strumento ci ritroveremo dei mattoni di procedura di elaborazione dati che potremmo tranquillamente riutilizzare (riadattandoli) in altre e successive occasioni semplicemente rispettando le regole di scrittura (vedi pagg. 73-74).

Nella parte successiva è affrontato il problema di come preparare i dati per meglio utilizzarli con SPSS. Ciò è vero soprattutto, quando si importano i dati da un'altra applicazione e/o quando questi sono stati predisposti per fini poco compatibili alle esigenze del motivo del nuovo trattamento.

Per esempio, una buona assegnazione delle etichette alle variabili e alle modalità (breve ma nello stesso tempo chiaramente esplicative) risulta fondamentale per ottenere tabelle e grafici immediatamente leggibili. È un lavoro a volte lungo e in parte noioso, ma è essenziale; va fatto una sola volta

per tutte e ripaga nelle future elaborazioni.

Vengono presentati diversi comandi e sottocomandi di sintassi: mi permetto di suggerire che non sarebbe male prevedere come servizio aggiuntivo, come guida alla programmazione e alla sintassi con SPSS, che questo volumetto presentasse alla fine un indice analitico dei comandi e degli argomenti trattati in modo da rintracciare facilmente, come in un normale manuale, le giuste definizioni adottate e i comandi di sintassi chiave da utilizzare e da ricordare.

Nell'ultima parte l'autore presenta una carrellata dei comandi rintracciabili nel menù dedicato all'analisi statistiche (cuore del pacchetto SPSS); la descrizione procede passo dopo passo alle varie opportunità offerte dal menù. Una volta che ci si è abituati ad usare la sintassi dei comandi, si riuscirà, con relativa facilità, a provare, secondo le proprie finalità di analisi, a sfruttare tutta la batteria di possibilità offerte dal software; la lettura dei passi successivi del volume diventa molto produttiva se, contemporaneamente, come già proposto, abbiamo aperto sul nostro personal computer una sessione di SPSS: provando i comandi di sintassi proposti e, naturalmente, salvando i file di output e di sintassi che abbiamo generato.

A pagina 118-119 del libro sono presentati un gran numero di statistiche di controllo e come presentarle convenientemente nelle tabelle.

In seguito, molto utile, un paragrafo è dedicato alle diverse procedure per comparare i gruppi; quindi, i successivi, alle procedure di analisi della correlazione, della regressione e dell'analisi log-lineare. A questo punto però, la conoscenza delle nozioni di statistica

risulta indispensabile per la comprensione delle varie statistiche di calcolo presentate.

Infine vengono trattati, molto esaurientemente, le procedure di Cluster Analysis per il raggruppamento dei dati della matrice in gruppi; le procedure di riduzione dimensionale dei dati, per l'analisi delle componenti principali; e le procedure di scaling che consentono di studiare e valutare, attraverso una serie di coefficienti, l'attendibilità complessiva di una scala e le relazioni esistenti nel suo complesso. È perfettamente inutile ricordare che le nozioni di statistica sono, anche e soprattutto qui, indispensabili.

Visto l'uso improprio di strumenti di questo genere, come ricordato in apertura, è indispensabile sottolineare come anche in valutazione non basta avere "elaborato i dati con SPSS" per avere fatto una buona analisi; SPSS è uno strumento eccezionale di elaborazione dei dati da usare *se serve* e, qualora servisse, da utilizzare con criterio e intelligenza. Il volume di Di Franco consente di avvicinarsi con facilità a questo *software* anche per chi non l'ha mai maneggiato precedentemente a patto, come già ricordato, di possedere alcune competenze statistiche di base.

Enrico Todini