

SAS for Beginners

Giovanni Boscaino

giovanni.boscaino@unipa.it

Il Corso prevede la formazione in Programmazione SAS. Il SAS è una suite di applicativi prodotto dalla SAS Institute e distribuita in tutto il mondo. Sono molte le aziende che cercano personale capace di muoversi in ambiente di programmazione SAS e quindi la sua conoscenza risulta un elemento distintivo nel mercato del lavoro.

L'obiettivo del corso è quindi quello di conferire agli studenti le capacità e abilità nella programmazione finalizzata alla acquisizione, gestione e manipolazione dei dati e alla reportistica di analisi statistiche di base. Il contenuto del corso si sovrappone alle attività di e-learning erogate dal SAS di Programming Base 1 e 2.

Il completamento del corso è propedeutico all'ottenimento dell'open-badge rilasciato da SAS in seguito a un test di verifica condotto presso la SAS Institute (opzionale per lo studente).

Dettagli del corso

Erogazione: 26 aprile – 3 giugno 2022; 2 ore di lezione al giorno tutti i giorni (secondo il calendario delle lezioni di Statistica Sociale 1 C.I. del Corso di Laurea in Statistica per l'Analisi dei Dati) per una durata totale di 52 ore di lezione frontale.

Modalità di verifica: gli studenti dovranno dimostrare di aver acquisiti le competenze e abilità in programmazione attraverso una prova orale in cui risponderanno ai quesiti di natura teorica e pratica (utilizzando il software) che il docente porrà loro.

Crediti maturati: 6 CFU

Lingua del corso: italiano

Argomenti trattati:

Numero di ore	Argomenti
1	Introduzione al corso: obiettivi, metodo di verifica, materiale didattico
2	librerie permanenti e temporanee, step, statement, global statement, dataset sas, sottomissione di programma, log, errori, warnings, commenti, colori, dati strutturati e non strutturati, colonne, righe, attributi di colonna, date
8	Inserimento manuale dei dati, proc contents, libname, libname xlsx, options validvarname, libname clear. Proc import csv, tab e xlsx, dbms dlm, replace, guessingrows, inserimento dataset sas in excel, dataset sas permanente da excel con libname, Proc Print noobs, obs=, var. WHERE statement, and, or, like, is (not)missing, where also: Proc means, mean std, min, max, mode, q1, median, q3, percentili, scarti, format, class, nonobs, variance, ways, statement output e class Proc univariate, histogram, endpoints, midpoints; Proc freq, nocum, nopercnt, nofreq, nocol norow, nopercnt, crosstlist, list, or, relrisk, chisq, expected, fisher. Grafici con freq plot, plots=freqplot, opzioni scale orient type groupby twoway, esportazione dataset dei risultati della proc tables, out=, outpct. Macro variabile;
6	statement format formati per variabili numeriche: w.d, comma, dollar, eurox, datew, mmdyyw, ddmmyyww, monyyw, monname, weekdate, zw.; proc format per formati numerici, date e alfanumerici; opzione library= proc sort: noduprecs, nodupkey, dupout, all
8	Passo data: where, format, keep, drop; not like; Nuove variabili per assegnazione, funzioni numeriche e per variabili alfanumeriche, opzione "OF name1-nameN" Funzioni per date: month, year, day, weekday, qtr, today,

	<p>mdy, yrdif. Funzione SCAN. IF/THEN IF/THEN/ELSE e integrazione con AND/OR. Statement LENGTH. IF/THEN/DO creazione di piu' dataset in un unico passo data Title; Footnote; ods noproctitle; Label; BY per le proc; proc export; esportare in excel con libname; ODS CSVALL, RTF, EXCEL, POWERPOINT, PDF DATA STEP: compilazione, esecuzione, PDV, implicit return, implicit output, explicit output; Opzioni KEEP= e DROP=; PUTLOG statement: ALL, column=, "message". retain; creazione di variabili cumulate Elaborazione dati in gruppi: by statement in data step, first.column e last.column; if expression per output; il caso di piu' variabili di condizionamento funzioni: "of stringa:", RAND, LARGEST, ROUND, CEIL, FLOOR, INT, INTCK, INTNX, COMPBL, COMPRESS, STRIP, FIND, TRANWRD, CAT, CATS, CATX. PUT, INPUT, format, Informat, Rename= proc format CNTLIN</p>
7	<p>Concatenazione dataset; merging di tabelle: statement merge + by, 1 a 1, 1 a molti, righe non corrispondenti. Opzione IN= di merge; merge di tabelle con le stesse variabili; merge di piu' di due tabelle; DO LOOP e PDV in esecuzione; DO LOOP con etichette; posizione di output; DO LOOP nidificati e posizione dell'output; DO UNTIL; DO WHILE; Differenze DO LOOP e WHILE/UNTIL Tabelle wide e narrow; trasformare wide a narrow, e narrow a wide proc transpose; array</p>
20	<p>ESERCITAZIONI</p>