

# Tecniche statistiche di analisi del cambiamento

01-Ripasso statistica descrittiva

(v. 1.5, 3 ottobre 2018)

versione per stampa

Germano Rossi<sup>1</sup>

`germano.rossi@unimib.it`

<sup>1</sup>Dipartimento di Psicologia, Università di Milano-Bicocca

2018-19

# Link utili

Qualora avete dei dubbi e volete approfondire questi argomenti:

Ripassare i livelli di misura: [Elem01-Introduzione.pdf](#)

Ripassare le rappresentazioni grafiche:  
[Elem02-GraficiDistribuzioni.pdf](#)

Ripassare i percentili: [Elem03-Percentili.pdf](#)

Ripassare le statistiche della tendenza centrale:  
[Elem04-TendCentrale.pdf](#)

Ripassare le statistiche di variabilità  
[Elem05-IndiciVariabilita.pdf](#)

# Introduzione

In questo corso, basandomi sugli anni precedenti, ho notato che

- una parte degli studenti non ha fatto statistica alla triennale
- una parte degli studenti ha fatto statistica alla triennale, ma non la ricorda
- una parte degli studenti non ha idea di cosa sia SPSS o altro software per l'analisi dei dati

Per questo motivo iniziamo con un ripasso della statistica descrittiva, poi della verifica di ipotesi e delle principali tecniche di analisi dei dati che vengo affrontate normalmente nella triennale di questo Ateneo e successivamente, più in dettaglio, tecniche già fatte alla triennale, ma ampliate:

■ t-test

■ chi-quadro

■ correlazione

■ regressione

■ analisi della  
varianza

# Livelli di misura

Riassumiamo in una tabella alcune informazioni: le colonne contengono “variabili”; le righe, “casi statistici”; le celle contengono “valori”

| Id | Genere | corporatura | confezioni in offerta | spesa |
|----|--------|-------------|-----------------------|-------|
| a  | F      | grassa      | 12                    | 180   |
| b  | M      | media       | 2                     | 50    |
| c  | F      | mingherlina | 4                     | 40    |

- Il **genere** si limita a distinguere categorie fra loro diverse
- La **corporatura** oltre a distinguere le categorie, permette anche di metterle in ordine (da magro a grasso), ma non è possibile *quantificare* la differenza fra “mingherlina” e “normale” o fra “media” e “grassa”
- Le variabili **confezioni** e **spesa** usano anche un'*unità di misura* e permettono di **quantificare** le differenze fra i diversi valori (fra 2 e 4 confezioni c'è una differenza di 2)

# Livelli di misura: tipologie

- I diversi modi di misurare si chiamano **livelli di misura** o *scaie di misura*
  - I livelli di misura possono essere classificati con diverse modalità
- 
- **categorie**: tutte le forme di misurazione che possono non essere indicate con dei numeri (ad es. colore capelli, titolo di studio...)
  - **punteggi**: tutte le misurazioni che sono necessariamente indicate con numeri (ad es. età, anni di studio...)
  - **qualitative**: tutte le forme di misurazione che esprimono la qualità di ciò che viene misurato (titolo di studio)
  - **quantitative**: tutte le forme di misurazione che esprimono la quantità di ciò che viene misurato tramite un valore numerico (anno di nascita)

# Livelli di misura: tipologie

- I livelli di misura possono essere classificati con diverse modalità
- quantitative **discrete**: tutte le forme di misurazione che procedono a salti (ad es. numero di errori, numero di persone in una stanza...)
- quantitative **continue**: tutte le misurazioni che sono espresse come approssimazione (ad es. peso, altezza...)
- **nominali**: misurazioni che possono utilizzare solo etichette (capelli rossi)
- **ordinali**: misurazioni che permettono di ordinare in qualche modo i valori (titolo di studio)
- **intervallo**: misurazioni che utilizzano un'unità di misura senza un'origine fissa (anno di nascita)
- **a rapporto**: misurazioni che utilizzano un'unità di misura con origine fissa (età, peso)

# Riconoscere i livelli di misura

- Fra i livelli di misura, la tipologia più utilizzata è l'ultima.
  - Come possiamo riconoscere a quale livello stiamo misurando?
- 1 I valori di ciò che stiamo misurando permettono valori diversi?
    - No, allora è una costante. Non è una variabile!
  - 2 Prendiamo due qualunque valori possibili di quella variabile
  - 3 Possiamo dire che sono uguali o sono diversi?
    - Sì, allora siamo almeno a livello nominale
  - 4 Possiamo in qualche modo ordinare questi valori?
    - Sì, allora siamo almeno a livello ordinale
  - 5 Possiamo identificare un'unità di misura specifica?
    - Sì, allora siamo almeno a livello intervallo
  - 6 L'unità di misura ha un punto d'origine assoluto?
    - No, allora siamo a livello intervallo
    - Sì, allora siamo a livello a rapporto

# Livelli di misura: applicazione 1

Usando le variabili dell slide 4:

- **Genere:** selezionando 2 qualunque individui posso dire se sono entrambi dello stesso sesso o no, non posso metterli in ordine  $\Rightarrow$  *Nominale, Catoriale, Qualitativa*
- **Corporatura:** selezionando 2 qualunque individui diversi, posso metterli in ordine (dal mingherlino a grasso), ma non ho un'unità di misura  $\Rightarrow$  *Ordinale, Catoriale, Qualitativa*
- **Confezioni:** selezionando 2 qualunque individui diversi, posso metterli in ordine (di confezioni) e posso usare un'unità di misura che il punto origine assoluto, posso metterli in ordine (dal mingherlino a grasso), ma non ho un'unità di misura  $\Rightarrow$  *Rapporto, Punteggio, Quantitativa, Discreta*
- **Spesa:** idem come sopra  $\Rightarrow$  *Rapporto, Punteggio, Quantitativa, Continua*

## Livelli di misura: applicazione 2

Usando le variabili dell slide 5:

- **Colore capelli:** i colori sono delle semplici etichette, per cui  $\Rightarrow$  *Nominale, Catoriale, Qualitativa*
- **Titolo di studio:** sono delle etichette ma vanno acquisiti in ordine ben preciso, non c'è un'unità di misura  $\Rightarrow$  *Ordinale, Catoriale, Qualitativa*
- **Età:** in qualunque modo sia misurata (anni, mesi...)  $\Rightarrow$  *Rapporto, Punteggio, Quantitativa, Continua*
- **Anni di studio:** è simile al *Titolo di studio*, ma qui si considerano gli anni di studio  $\Rightarrow$  *Rapporto, Punteggio, Quantitativa, Discreta*
- **Anno di nascita:** l'anno di nascita non ha un'origine assoluta (l'anno 0 è arbitrariamente fissato)  $\Rightarrow$  *Intervallo, Punteggio, Quantitativa, Continua*

## Livelli di misura: applicazione 3

Usando le variabili dell slide 6:

- **Numero di errori:** l'errore è l'unità di misura; l'origine è assoluta  
⇒ *Rapporto, Punteggio, Quantitativa, Discreta*
- **Numero di persone in una stanza:** la persona è l'unità di misura  
⇒ *Rapporto, Punteggio, Quantitativa, Discreta*
- **Peso:** misurato in chili o in grammi c'è comunque un'unità di misura, con origine assoluta ⇒ *Rapporto, Punteggio, Quantitativa, Continua*
- **Altezza:** l'unità di misura può variare (metri o centimetri) ma tutte vengono espresse con la stessa unità ⇒ *Rapporto, Punteggio, Quantitativa, Continua*

Avendo un insieme di dati, possiamo fare delle analisi che hanno lo scopo di dare delle informazioni sintetiche su alcuni aspetti dei dati stessi.

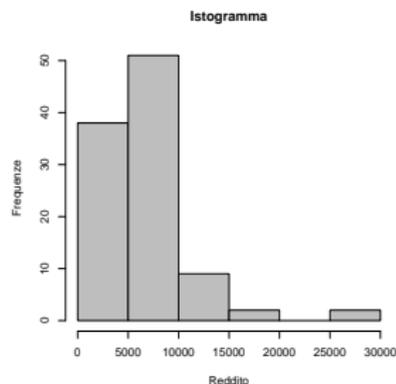
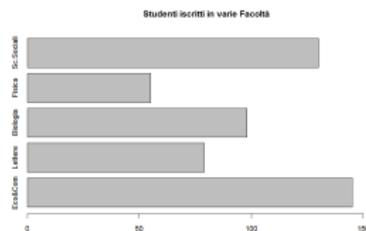
I tipi di statistiche descrittive possibili sono:

- Rappresentazioni riassuntive (sia grafiche sia testuali)
- Statistiche della **tendenza centrale**: danno indicazioni sulla prevalenza dei dati
- Statistiche di **posizione**: associano certi valori con la posizione che occupano nella distribuzione e viceversa
- Statistiche di **variabilità**: come i dati sono dispersi

# Rappresentazioni riassuntive

- Nate per esprimere a colpo d'occhio l'insieme dei dati
  - **Tabelle di frequenza:** riportano i valori della variabile misurata con le relative frequenze e percentuali
  - **Grafici a barre:** rappresentano graficamente la tabella di frequenza con variabili nominali/ordinali
  - **Istogrammi:** rappresentano graficamente la tabella di frequenza con variabili intervallo/rapporto

|     | f  | %    |
|-----|----|------|
| F   | 13 | 43,3 |
| M   | 17 | 56,7 |
| Tot | 30 |      |



# Statistiche di posizione

- Pensate per variabili almeno a livello ordinale
  - i valori vengono ordinati dal più piccolo al più grande
  - Si suddivide l'intera distribuzione in  $n$  parti uguali
  - si usa la posizione centrale o delle due metà per avere informazioni sui dati
- Le suddivisioni più utilizzate sono:
  - in 4 parti: quartili (Q1, Q2, Q3)
  - in 10 parti: decili (D1, D2... D9)
  - in 100 parti: centili o percentili (P1, P2, ... P99)
  - in 3 parti: terzili (T1, T2)

la formula generica per trovare la  
posizione dove cade l' $n$ -tile è:

$$\frac{N}{\text{parti}}$$

# Statistiche della tendenza centrale

È un'indicazione generica di come sta andando la distribuzione della variabile

- **Nominale: Moda** ( $M_o$ ) cioè il *valore* più frequente
- **Ordinale: Mediana** ( $M_{dn}$ ) è il *valore in posizione centrale* della distribuzione (pressapoco  $N/2$ )
- **Intervallo/Rapporto: Media** aritmetica, ( $\bar{X}$ ,  $M_d$ ,  $M$ ) è la somma ( $\sum$ ) di tutti i valori di una distribuzione, divisa per la numerosità ( $N$  o  $N-1$ )

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N} = \frac{\sum X}{N} \text{ oppure } \frac{\sum X}{N-1}$$

# Statistiche di variabilità

Danno indicazioni su quanto sono dispersi i dati attorno alla tendenza centrale

- **Nominale:** Numero delle classi di equivalenza (non usata)
- **Ordinale:** Primo e terzo quartile (Q1 e Q3)
- **Intervallo/Rapporto:** Campo di variazione o gamma (di oscillazione) o *range* o intervallo (Max - Min)
- **Intervallo/Rapporto:** Differenza interquartilica (DI o IQR), Q3-Q1
- **Intervallo/Rapporto:** Varianza ( $s^2$  e deviazione standard ( $s$ ))

$$s^2 = \frac{\sum (X_i - \bar{X})^2}{N - 1} \quad s = \sqrt{s^2}$$

# Punteggi standardizzati

- I punteggi grezzi (quindi a livello I/R) possono essere espressi in punteggi standardizzati
- I punteggi standardizzati (o punti z) sono espressi usando la deviazione standard come unità di misura

$$z = \frac{X - \bar{X}}{s}$$

- Il punto z risponde alla domanda “Il punteggio X a quante deviazioni standard dista dalla media?”
- Per cui i punti z hanno media 0 e deviazione standard 1
- Valori negativi indicano punteggi inferiori alla media
- Valori positivi, punteggi sopra la media

# Punteggi standardizzati

- Dal momento che qualunque valore di qualunque variabile I/R può essere trasformata in un punto z...
- i punti z sono una trasformazione che permette di confrontare fra loro variabili diverse

## Esempio

Andrea è stato valutato da uno psicologo con un test di abilità matematiche e uno di abilità verbali. I suoi punteggi sono stati trasformati in punti z e corrispondono a:

matematiche -0.21; verbali +1.75

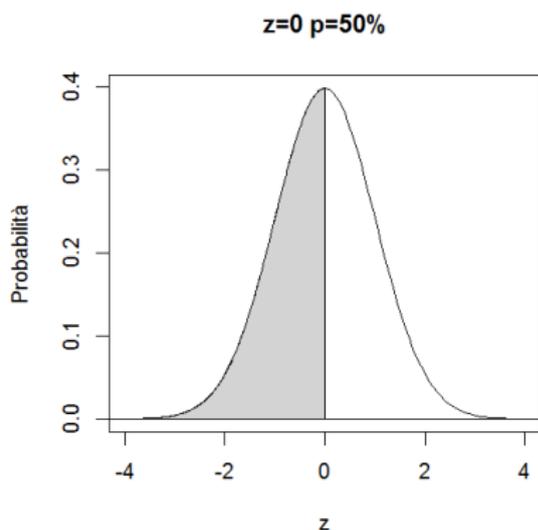
Andrea ha abilità matematiche che sono poco inferiori a quelli della maggior parte della gente (il 42% delle persone ha un punteggio inferiore al suo), mentre ha capacità verbali che sono superiori a quelle del 96% delle persone.

# Punteggi standardizzati

Come siamo passati da un punto  $z$  ad una %?

- Una variabile con  $N$  molto grande che si distribuisce normalmente, trasformata in punti  $z$  tende ad avvicinarsi alla curva normale
- Esiste una tabella dei punti  $z$ , costruita su una particolare curva normale: quella che ha  $M=0$  e  $s=1$
- questa tabella serve per trovare l'area della curva normale corrispondente ad un determinato punto  $z$
- Le tabelle possono riportare informazioni diverse che fanno riferimento alla stessa cosa.

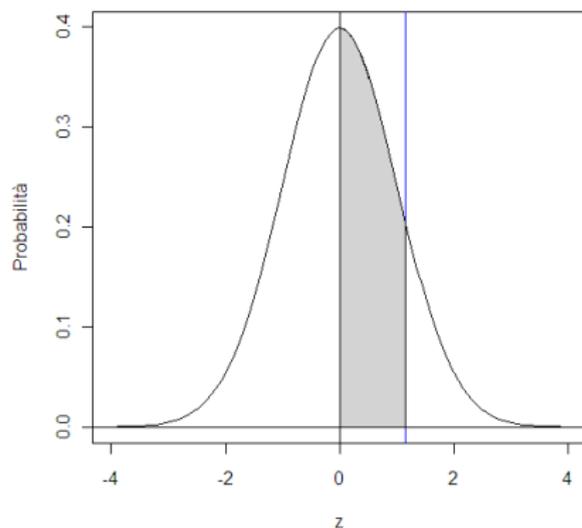
# Tavole della distr. normale



- La curva normale è simmetrica
- Quindi ogni metà è il 50%
- Allontanandoci da  $z=0$  abbiamo aree simili per  $z$  simili
- $z=-0.10$  -> 3.98% sotto la media
- $z=0.10$  -> 3.98% sopra la media

# Esempio di Tavola della distr. normale

- Welkowitz, Cohen, Ewen Tavola A (p. 473 ss.)
- La tavola riporta le proporzioni di area sottese alla curva normale calcolate a partire dalla media (ricordarsi che l'intera area è simmetrica)



- Per ogni punto  $z$  viene indicata l'area fra  $z=0$  e il punto  $z$  stesso (area in grigio)
- La proporzione di area è indicata come percentuale (34,13) con due decimali
- Si può trasformare facilmente in proporzione (.3413)

# Tavole della distr. normale

## Tavole statistiche

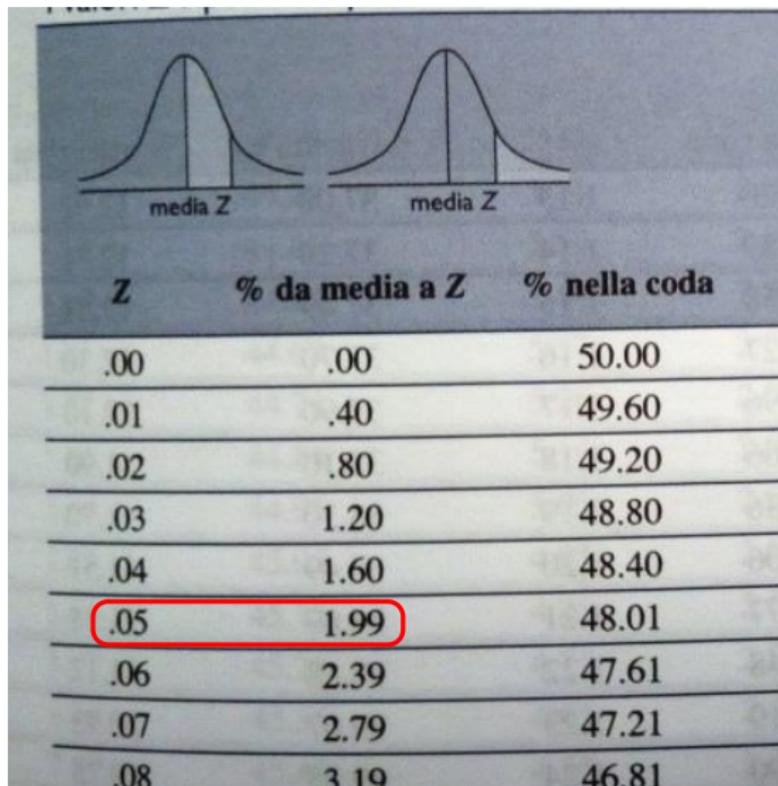
Tabella A

Percentuale di area sottesa dalla curva normale tra la media e un valore  $z$  dato

| $z$ | 0,00  | 0,01  | 0,02  | 0,03  | 0,04  | 0,05  | 0,06  | 0,07  | 0,08  | 0,09  |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0,0 | 00,00 | 00,40 | 00,80 | 01,20 | 01,60 | 01,99 | 02,39 | 02,79 | 03,19 | 03,59 |
| 0,1 | 03,98 | 04,38 | 04,78 | 05,17 | 05,57 | 05,96 | 06,36 | 06,75 | 07,14 | 07,53 |
| 0,2 | 07,93 | 08,32 | 08,71 | 09,10 | 09,48 | 09,87 | 10,26 | 10,64 | 11,03 | 11,41 |
| 0,3 | 11,79 | 12,17 | 12,55 | 12,93 | 13,31 | 13,68 | 14,06 | 14,43 | 14,80 | 15,17 |
| 0,4 | 15,54 | 15,91 | 16,28 | 16,64 | 17,00 | 17,36 | 17,72 | 18,08 | 18,44 | 18,79 |

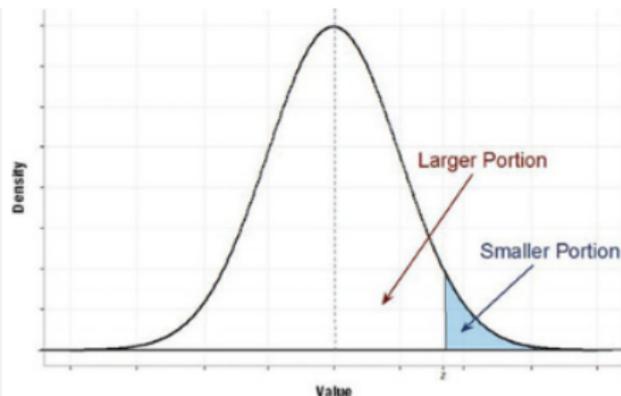
- La prima colonna indica il primo decimale del punto  $z$ , ogni colonna successiva indica il secondo decimale
- All'incrocio fra una riga (ad es. 0,3) e una colonna (0,05) troviamo l'area corrispondente (espressa in %)
- Es. l'area fra  $z=0,35$  e 0 è pari a 13,68
- Es. l'area fra  $z=0,05$  e 0 è pari a 1,99

# Tavole della distr. normale



- La prima colonna indica il punto z (con 2 decimali)
- La seconda colonna riporta l'area fra il punto z e la media (espressa come %)
- Es. l'area fra  $z=0,05$  e 0 è pari a 1,99
- La terza colonna riporta l'area fra il punto z e la coda positiva (espressa come %)
- Es. l'area fra  $z=0,05$  e la coda positiva è pari a  $50-1,99=48.01$

# Tavole della distr. normale



| z   | Larger Portion | Smaller Portion | y     |
|-----|----------------|-----------------|-------|
| .00 | .50000         | .50000          | .3989 |
| .01 | .50399         | .49601          | .3989 |
| .02 | .50798         | .49202          | .3989 |
| .03 | .51197         | .48803          | .3988 |
| .04 | .51595         | .48405          | .3986 |
| .05 | .51994         | .48006          | .3984 |

- La prima colonna indica il punto  $z$  (con 2 decimali)
- La seconda colonna riporta l'area fra il punto  $z$  e la coda negativa
- Es. l'area fra  $z=0,05$  e la coda negativa è pari a  $50+1,99=51.99$
- La terza colonna riporta l'area fra il punto  $z$  e la coda positiva
- La quarta colonna riporta il valore dell'ordinata corrispondente al punto  $z$

# Scale derivate dai punti z

- Ci sono scale standardizzate utilizzate comunemente in psicologia (specie per i test) che derivano dai punti z
- **punteggi T**: hanno media 50 e  $ds=10$ . Si ottengono con  
 $T = 10z + 50$
- **QI o IQ**: la maggior parte dei test d'intelligenza (come il WAIS) utilizza una media di 100 e  $ds=15$ . Si ottengono con  
 $QI = 15z + 100$
- **QI o IQ**: il test d'intelligenza Stanford-Binet utilizza una media di 100 e  $ds=16$ . Si ottengono con  $QI = 16z + 100$

# Cos'è SPSS

**SPSS** significa **Statistical Package for Social Sciences** (Pacchetto statistico per le scienze sociali) ed è un programma per gestire dati e calcolare statistiche. Ma adesso si chiama **IBM Spss**

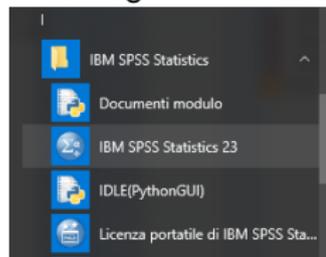
- Permette di inserire, nominare e gestire le misurazioni statistiche
- Permette di manipolare le variabili (in modo analogo ai fogli elettronici)
- Permette di calcolare le statistiche descrittive, di fare grafici
- Permette di fare l'analisi dei dati

Esiste un clone “open source” chiamato **PSPP** gratuito (ma non ancora completo)

Esiste anche un software che imita SPSS solo nell'aspetto, ma usa  $\mathbb{R}$  per fare i calcoli: **Jamovi**, gratuito (ma non ancora completo) che userò qualche volta in classe

# SPSS: Primo avvio

## ■ Per eseguire SPSS

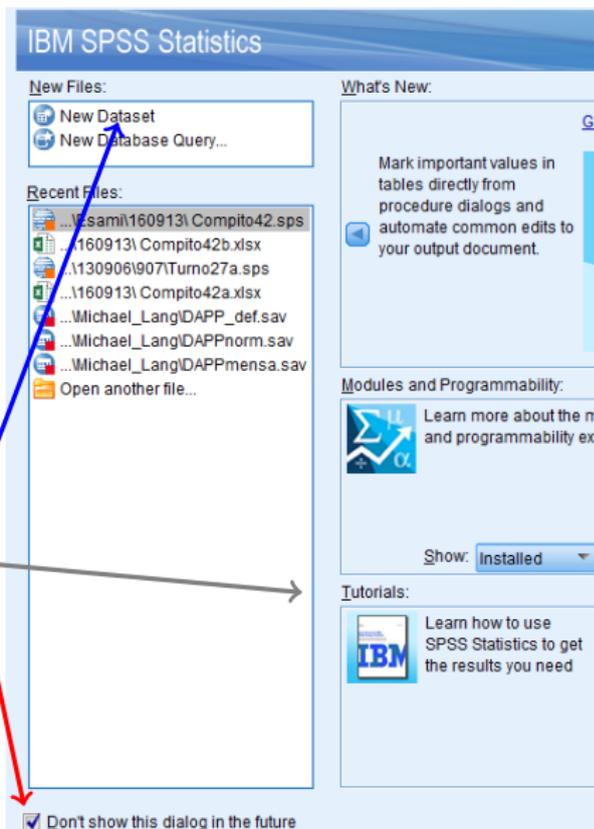


■ Alla prima esecuzione compare una finestra di dialogo che permette di effettuare alcune scelte. Per disattivarla, click-ate qui

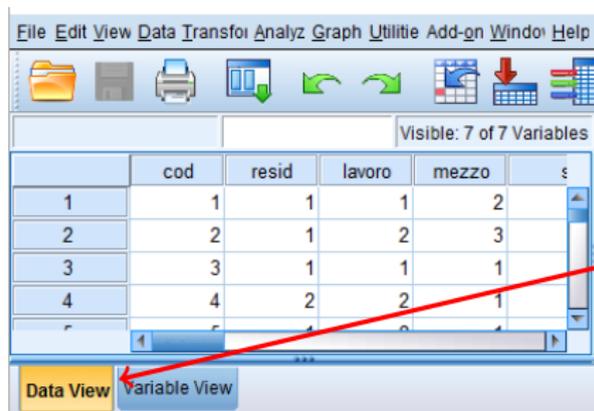
■ Potete selezionare “Tutorials”

■ Potete selezionare “New Dataset”

■ poi



# SPSS: Finestra di base



File Edit View Data Transform Analyze Graph Utilities Add-on Windows Help

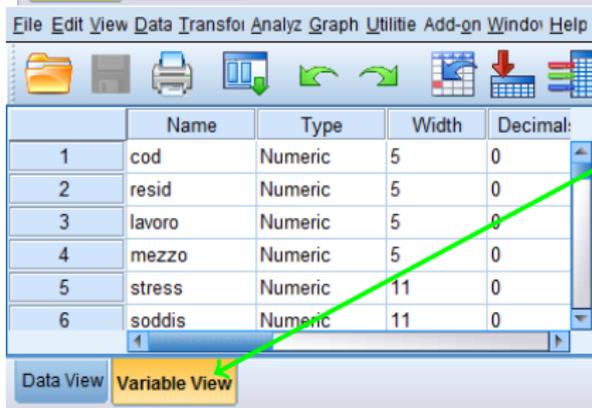
Visible: 7 of 7 Variables

|   | cod | resid | lavoro | mezzo | s |
|---|-----|-------|--------|-------|---|
| 1 | 1   | 1     | 1      | 2     |   |
| 2 | 2   | 1     | 2      | 3     |   |
| 3 | 3   | 1     | 1      | 1     |   |
| 4 | 4   | 2     | 2      | 1     |   |
| 5 | 5   | 1     | 1      | 1     |   |

Data View Variable View

A red arrow points from the 'Variable View' tab to the 'Variable View' window below.

- La finestra principale è formata da due pannelli
- uno per la visualizzazione dei dati
  - Ogni colonna è una variabile
  - Ogni riga un caso statistico
- uno per la descrizione delle variabili
  - dopo il nome, informazioni sul tipo (numerico, stringa, data...), l'ampiezza, i decimali (se numero)...



File Edit View Data Transform Analyze Graph Utilities Add-on Windows Help

|   | Name   | Type    | Width | Decimal: |
|---|--------|---------|-------|----------|
| 1 | cod    | Numeric | 5     | 0        |
| 2 | resid  | Numeric | 5     | 0        |
| 3 | lavoro | Numeric | 5     | 0        |
| 4 | mezzo  | Numeric | 5     | 0        |
| 5 | stress | Numeric | 11    | 0        |
| 6 | soddis | Numeric | 11    | 0        |

Data View Variable View

A green arrow points from the 'Variable View' tab to the 'Variable View' window above.

# SPSS: Finestra di base



Le voci di menù più utilizzate sono:

- **Dati**: permette di agire sul file dei dati (ordinare, selezionare, filtrare...)
- **Trasforma**: permette di manipolare le variabili (calcolare nuove variabili, ricodificare, contare valori...)
- **Analizza**: È il menù più utilizzato perché contiene tutte le procedure statistiche disponibili

La prima volta che usate SPSS vi conviene fare l'**esercitazione** disponibile nell'Help (Guida).

# SPSS: Per i principianti

In tutti i lucidi, per indicare un percorso di menù, userò questa convenzione

voce principale | sottomenù | sotto-sottomenù:  
ad es. Guida | Esercitazione

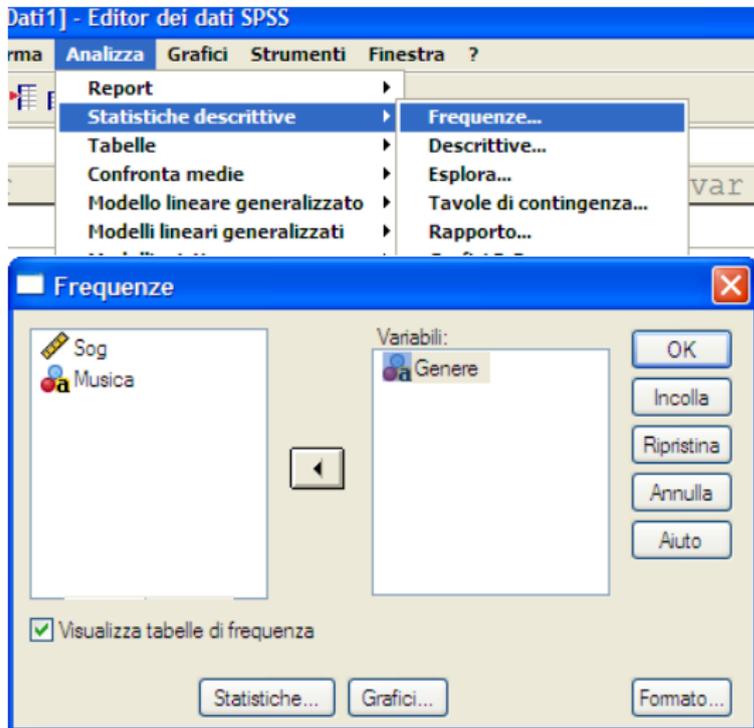
- Guida per installare SPSS (dopo averlo scaricato): <http://www.germanorossi.it/mi/file/InstallareSPSS25.pdf>
- Breve introduzione (pratica) a SPSS:  
<http://www.germanorossi.it/mi/file/Esercitazione%20SPSS%201.pdf>

# Distribuzione e grafici in SPSS

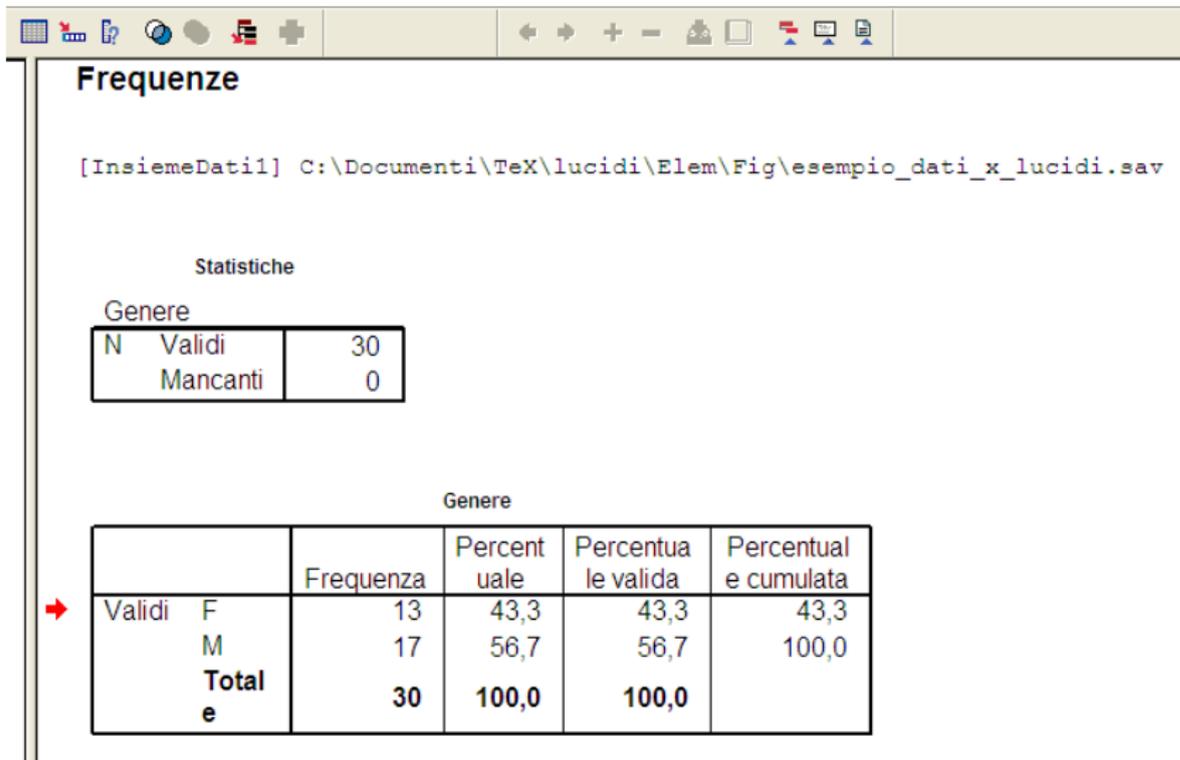
- La maggior parte delle statistiche che abbiamo visto e che vedremo e i relativi grafici, si ottengono in SPSS tramite tre comandi
  - `Analizza | Statistiche descrittive | Frequenze...`
  - `Analizza | Statistiche descrittive | Descrittive...`
  - `Analizza | Statistiche descrittive | Esplora...`
- In Spss si può fare solo quello che è previsto dal software; i grafici non sono bellissimi
- In Excel si possono fare molti più grafici (anche più “belli”)
- In R (usato in questi lucidi) si può fare praticamente tutto, ma è molto più complicato

# Spss: frequenze

- Calcoliamo le frequenze con il comando  
Analizza | Statistiche descrittive | Frequenze...
- Poi spostiamo una variabile nominale/ordinale fra le Variabili
- E premiamo OK



# Spss: frequenze



The screenshot shows the SPSS 'Frequenze' (Frequencies) output window. At the top, the file path is displayed: [InsiemeDati1] C:\Documenti\TeX\lucidi\Elem\Fig\esempio\_dati\_x\_lucidi.sav. Below this, the 'Statistiche' (Statistics) section is shown for the variable 'Genere'. A small table indicates that there are 30 valid cases and 0 missing cases. The main part of the output is a larger table titled 'Genere' which provides a detailed breakdown of the frequency distribution, including percentages and cumulative percentages for each gender category (F and M), as well as a total row.

**Frequenze**

[InsiemeDati1] C:\Documenti\TeX\lucidi\Elem\Fig\esempio\_dati\_x\_lucidi.sav

**Statistiche**

Genere

|   |          |    |
|---|----------|----|
| N | Validi   | 30 |
|   | Mancanti | 0  |

**Genere**

|              |          | Frequenza | Percentuale  | Percentuale valida | Percentuale cumulata |
|--------------|----------|-----------|--------------|--------------------|----------------------|
| Validi       | F        | 13        | 43,3         | 43,3               | 43,3                 |
|              | M        | 17        | 56,7         | 56,7               | 100,0                |
| <b>Total</b> | <b>e</b> | <b>30</b> | <b>100,0</b> | <b>100,0</b>       |                      |

# Spss: istogramma (Frequenze)

- Analizza | Statistiche descrittive | Frequenze...

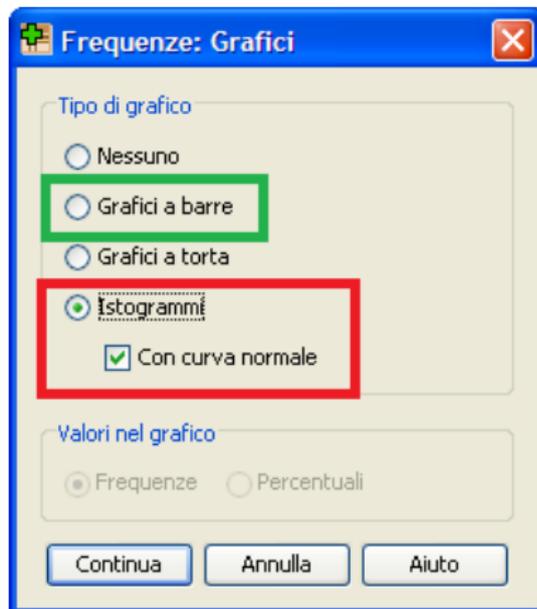
■ Pulsante

■ Selezioniamo Istogramma

■ eventualmente anche Con curva normale

■ Pulsante

■ e



# Spss: istogramma (Esplora)

- Analizza | Statistiche descrittive | Esplora...

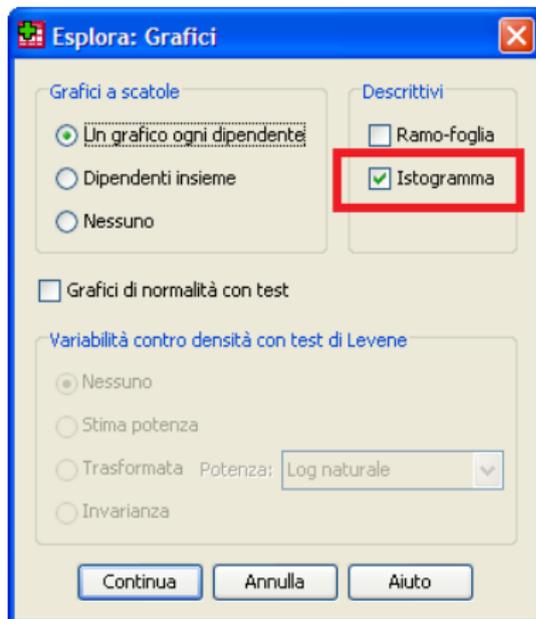
- Pulsante

- Selezioniamo Istogramma

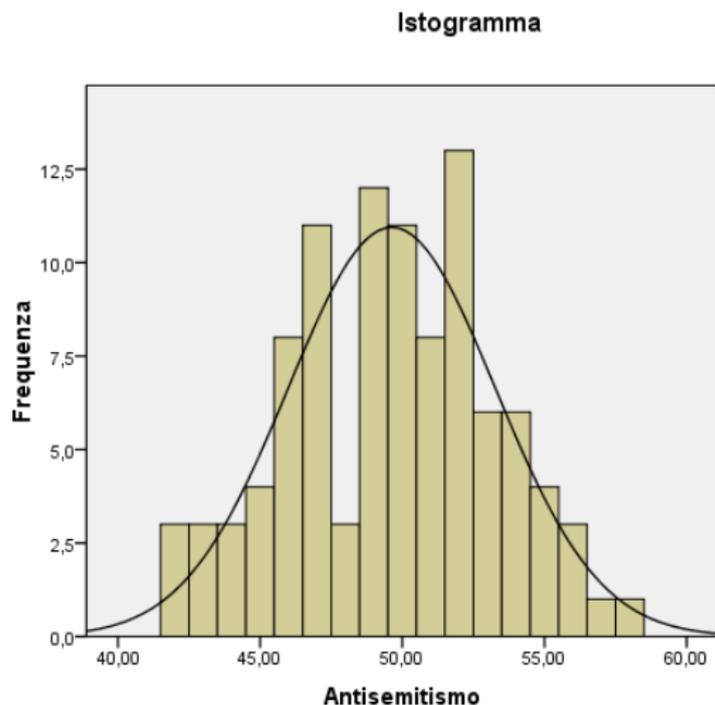
- (non è possibile avere la curva normale)

- Pulsante

- e



# Spss: istogramma (Frequenze/Esplora)



- Spss usa automaticamente gli intervalli di classe
- non è possibile modificare l'ampiezza di classe
- La curva normale non c'è con Esplora

# Spss: n-tili (Frequenze)

- In **Statistiche...**
- **Quartili** calcola i quartili
- **Punti di divisione** divide in n parti uguali
- **Percentili**: scrivete il percentile che volete e aggiungete

Fasce di età

|            |             |      |
|------------|-------------|------|
| N          | Validi      | 100  |
|            | Mancanti    | 0    |
| Percentili | 25          | 1,00 |
|            | 33,33333333 | 2,00 |
|            | 45          | 2,00 |
|            | 50          | 2,00 |
|            | 66,66666667 | 3,00 |
|            | 75          | 3,00 |

Valori percentili

Quartili

Punti di divisione per:  gruppi uguali

Percentili:

Aggiungi

Cambia

Rimuovi

Qui abbiamo chiesto: i quartili, i terzili e il 45esimo percentile  
← Qui i risultati espressi come percentili

# Spss: Esplora...

- Dopo aver scelto le variabili, metterle in Variabili dipendenti



- poi

- click-are su

Statistiche...

e

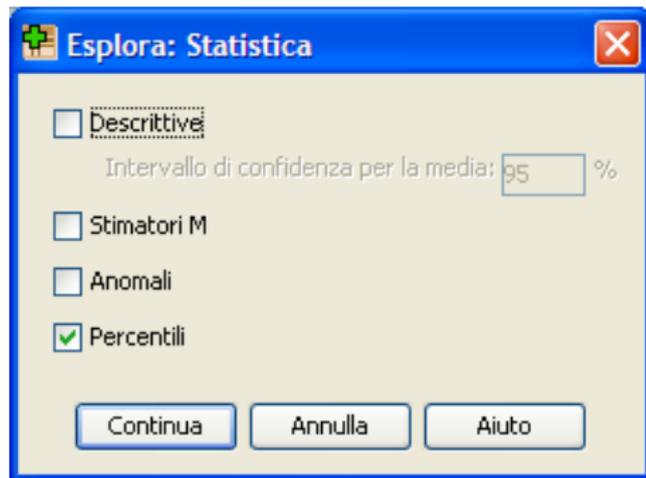
scegliere Percentili

- Quindi, click-are su

Continua

- Poi

OK



Esplora non permette di scegliere, ma fornisce alcuni n-tili notevoli

Percentili

|                                    |              | Percentili |      |      |      |      |      |      |
|------------------------------------|--------------|------------|------|------|------|------|------|------|
|                                    |              | 5          | 10   | 25   | 50   | 75   | 90   | 95   |
| Media ponderata<br>(Definizione 1) | Fasce di età | 1,00       | 1,00 | 1,00 | 2,00 | 3,00 | 3,00 | 3,00 |
| Cardini di Tukey                   | Fasce di età |            |      | 1,00 | 2,00 | 3,00 |      |      |

I tre quartili e i valori corrispondenti al 5% e 10% su entrambi i lati. Questi valori hanno particolarmente senso con variabili “normali” (capiremo più avanti) e sono spesso riportati nei manuali di standardizzazione dei test psicologici.

# Spss: moda

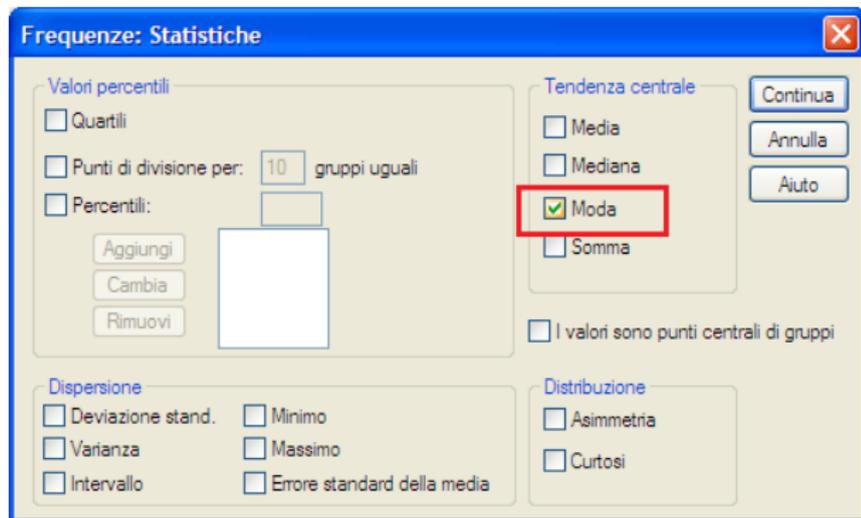
## ■ Tramite

Analizza |  
Statistiche descrittive |  
Frequenze...  
pulsante

Statistiche

,  
possiamo far  
calcolare la moda.

## ■ poi Continua e OK



| Statistiche |          |    |
|-------------|----------|----|
| Genere      |          |    |
| → N         | Validi   | 30 |
|             | Mancanti | 0  |
|             | Moda     | 1  |

La media viene visualizzata da Spss in molte procedure. Quelle specifiche sono:

- `Analizza | Statistiche descrittive | Frequenze...` (fra le varie statistiche che è possibile stampare vi è anche la media)
- `Analizza | Statistiche descrittive | Descrittive...` (è la procedura specifica per le statistiche descrittive)
- `Analizza | Statistiche descrittive | Esplora...` (stampa la media come una delle diverse statistiche per capire l'andamento e la distribuzione di una variabile)

# Spss: Media con Frequenze...

- Dopo aver scelto le variabili, click-are su  e selezionare Media
- Quindi, click-are su
- Con variabili quantitative conviene de-selezionare anche  Visualizza tabelle di frequenza
- oppure in 
  - Sopprimi le tabelle con più di n categorie
  - Numero massimo di categorie:
- Poi

Tendenza centrale

Media

Mediana

Moda

Somma

# Spss: Media con Descrittive...

- Dopo aver scelto le variabili, click-are su **Opzioni...**
- Normalmente **Media** è già selezionato
- Potete ordinare i risultati in vari modi
- Poi **OK**

**Statistiche descrittive**

|                   | N   | Media   |
|-------------------|-----|---------|
| gruppo            | 100 | ,44     |
| Antisemitismo     | 100 | 49,6300 |
| Validi (listwise) | 100 |         |

**Descrittive: Opzioni**

Media  Somma

**Dispersione**

Deviazione stand.  Minimo

Varianza  Massimo

Intervallo  Errore standard della media

**Distribuzione**

Curtosi  Asimmetria

**Ordine di visualizzazione**

Variabili nel file

Alfabetico

Medie crescenti

Medie decrescenti

Continua Annulla Aiuto

# Spss: Variabilità

Anche varianza e dev. st. sono visualizzate da Spss (sempre N-1) in molte procedure. Quelle specifiche sono:

- `Analizza | Statistiche descrittive | Frequenze...` (fra le varie statistiche vi è anche quelle di variabilità)
- `Analizza | Statistiche descrittive | Descrittive...` (è la procedura specifica per le statistiche descrittive)
- `Analizza | Statistiche descrittive | Esplora...` (stampa le misure di variabilità come parte delle diverse statistiche per capire l'andamento e la distribuzione di una variabile)

# Spss: Variabilità con Frequenze...

- Dopo aver scelto le variabili, click-are su **Statistiche...** e selezionare quelle che servono

| Dispersione  | Distribuzione                       |
|--|-------------------------------------|
| <input type="checkbox"/> Deviazione stand.           | <input type="checkbox"/> Asimmetria |
| <input type="checkbox"/> Varianza                    | <input type="checkbox"/> Curtosi    |
| <input type="checkbox"/> Intervallo                  |                                     |
| <input type="checkbox"/> Minimo                      |                                     |
| <input type="checkbox"/> Massimo                     |                                     |
| <input type="checkbox"/> Errore standard della media |                                     |

- Quindi, click-are su **Continua**
- Con variabili quantitative conviene selezionare anche

|   |
|---|
| <input checked="" type="checkbox"/> Visualizza tabelle di frequenza |
| <input type="checkbox"/> Sopprimi le tabelle con più di n categorie |
| Numero massimo di categorie: <input type="text" value="10"/>        |

- oppure in **Formato...**

- Poi **OK**

# Spss: Variabilità con Descrittive...

- Dopo aver scelto le variabili, click-are su

Opzioni...

## Dispersione

- Deviazione stand.
- Minimo
- Varianza
- Massimo
- Intervallo
- Errore standard della media

## Distribuzione

- Curtosi
- Asimmetria

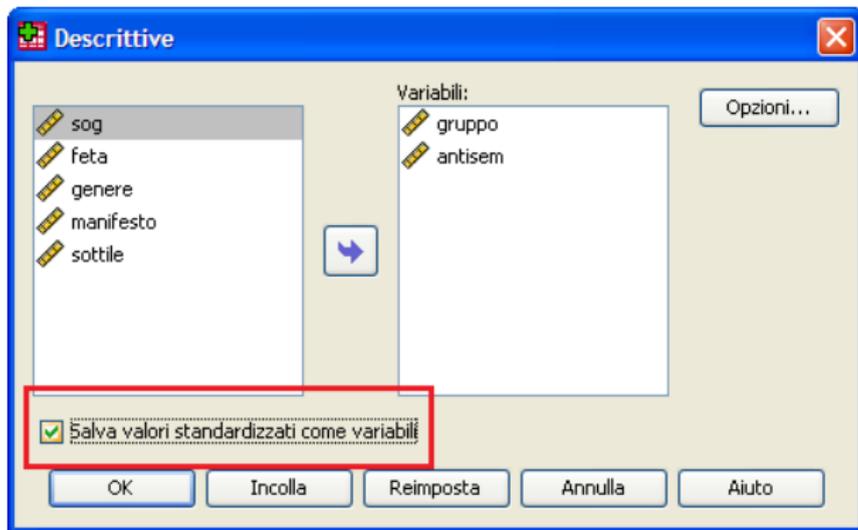
- Poi  e

Statistiche descrittive

|                   | N          | Intervallo | Minimo     | Massimo    | Deviazione std. | Varianza   | Asimmetria |            | Curtosi    |            |
|-------------------|------------|------------|------------|------------|-----------------|------------|------------|------------|------------|------------|
|                   | Statistica | Statistica | Statistica | Statistica | Statistica      | Statistica | Statistica | Errore std | Statistica | Errore std |
| gruppo            | 100        | 1          | 0          | 1          | ,499            | ,249       | ,245       | ,241       | -1,980     | ,478       |
| Antisemitismo     | 100        | 16,00      | 42,00      | 58,00      | 3,64499         | 13,286     | -,092      | ,241       | -,523      | ,478       |
| Validi (listwise) | 100        |            |            |            |                 |            |            |            |            |            |

# Spss: punti z

- Spss permette di calcolare i punti z di una variabile per ogni unità statistica, tramite **Analizza | Statistiche descrittive | Descrittive...** e attivando il flag **Salva valori standardizzati come variabili**



# Spss: punti z

- All'insieme delle variabili, viene aggiunta una variabile con il nome corrispondente preceduto da una "Z"
- Questa variabile può essere usata come qualsiasi altra



The screenshot shows the SPSS Data Editor window for a file named "esempio2.sav [InsiemeData1]". The window title bar includes the SPSS logo and the file name. Below the title bar is a menu bar with options: File, Modifica, Visualizza, Dati, Trasforma, Analizza, Grafici, and Strume. A toolbar with various icons is located below the menu bar. The main data grid shows a single row of data for subject "1 : sog". The columns are labeled "Zgruppo", "Zantsem", and "var". The values for "Zgruppo" and "Zantsem" are as follows:

|    | Zgruppo  | Zantsem  | var |
|----|----------|----------|-----|
| 9  | -0,88196 | 1,74760  |     |
| 10 | -0,88196 | 1,19891  |     |
| 11 | 1,12250  | 0,65021  |     |
| 12 | 1,12250  | -0,17284 |     |
| 13 | 1,12250  | 0,65021  |     |
| 14 | -0,88196 | -0,72154 |     |
| 15 | -0,88196 | 0,65021  |     |

# Autovalutazioni

Su elearning trovate 2 quiz di poche domande ciascuno, che potete fare e rifare tutte le volte che volete (ma tra un tentativo e l'altro dovete lasciar passare almeno 1 ora di tempo):

- Quiz intitolato: *Autovalutazione sulle scale di misura* (10 domande estratte casualmente da un database di 62, 15 minuti per rispondere)
- Quiz intitolato: *Autovalutazione sulle statistiche descrittive* (11 domande estratte casualmente da un database di 88, 20 minuti per rispondere)