Analisi statistica di un insieme di valori e istogramma delle relative classi di frequenza

Excel mette a disposizione lo strumento **Analisi statistica** e lo strumento **Istogramma** con i quali si calcolano rispettivamente le principali funzioni statistiche dell'analisi di dati e determina le frequenze degli stessi dati suddivisi in classi o sottoinsiemi predefiniti; di queste frequenze Excel disegna automaticamente l'istogramma.

Vogliamo eseguire l'analisi statistica e la suddivisione in classi di 50 numeri predefiniti. Nel *Foglio1* del file di Excel "Analisi statistiche" inseriamo una casella di testo con il titolo del progetto: "ANALISI STATISTICA E CLASSI DI FREQUENZA".

Selezioniamo la cella A6 e scriviamo "TABELLA DEI VALORI". Dalla cella A7 alla cella A57 inseriamo cinquanta numeri interi, casualmente, da 0 a 99.

Di questi numeri eseguiamo l'analisi statistica mediante lo strumento **Statistica descrittiva** contenuto negli **Strumenti di analisi**; per accedere a questi strumenti occorre fare clic su **Analisi dati** del gruppo **Analisi** della scheda **Dati**. Se il comando **Analisi dati** non è disponibile nella scheda **Dati** occorre caricarlo.

Per caricare lo strumento Analisi:

- 1. Nel menu File fare clic su Opzioni e quindi su Componenti aggiuntivi.
- 2. Nel menu **Gestisci** scegliere **Componenti aggiuntivi di Excel** e quindi fare clic sul pulsante **Vai**.
- 3. Nella finestra **Componenti aggiuntivi** mettere il segno di spunta nella casella **Strumenti di analisi** e fare clic su **OK**.

Reso disponibile il comando **Analisi dati**, facciamo clic sulla sua icona del gruppo **Analisi** della scheda **Dati**: si apre la finestra **Analisi dati**; in essa scegliamo **Statistica descrittiva** e quindi **OK**. Si apre la relativa finestra:

- 1. nella casella Intervallo di input inseriamo l'intervallo da A6 a A57;
- 2. lasciamo l'opzione Dati raggruppati per: colonne;
- 3. mettiamo il segno di spunta nella casella Etichette nella prima riga;
- 4. inseriamo l'indirizzo D6 nella casella Intervallo di output;
- 5. mettiamo il segno di spunta nella casella Riepilogo statistiche;
- 6. confermiamo con **OK**.

| nput | | | | OK |
|-------------------------------------|--|-----------|----|---------|
| Intervallo di input: | \$A\$ | 6:\$A\$57 | UN | |
| Dati raggruppati per: | pati per: <u>C</u>olonne | | | Annulla |
| | R | ighe | | 2 |
| 🔽 Etichette nella prima riga | | | | |
| Opzioni di output | | | | |
| Intervallo di <u>o</u> utput: | \$D\$ | 6 | | |
| 🔘 Nuovo <u>f</u> oglio di lavoro: | | | | |
| 🔘 Nuova car <u>t</u> ella di lavoro | | | | |
| Riepilogo statistiche | | | | |
| 📃 Livello di confidenza per med | dia: | 95 | % | |
| K-esimo più gr <u>a</u> nde: | 1 | | | |
| K-esimo più piccolo: | 1 | | | |

Excel inserisce automaticamente il simbolo di riferimento assoluto agli indirizzi delle celle. Dalla cella D6 Excel inserisce la tabella dell'analisi statistica (allarghiamo opportunamente le colonne) con i vari indici statistici:

| | D | E | | | | | | | | |
|----|----------------------|--------------|--|--|--|--|--|--|--|--|
| 5 | | | | | | | | | | |
| 6 | TABELLA DEI VALORI | | | | | | | | | |
| 7 | | | | | | | | | | |
| 8 | Media | 43,80392157 | | | | | | | | |
| 9 | Errore standard | 4,801471964 | | | | | | | | |
| 10 | Mediana | 41 | | | | | | | | |
| 11 | Moda | 93 | | | | | | | | |
| 12 | Deviazione standard | 34,28936839 | | | | | | | | |
| 13 | Varianza campionaria | 1175,760784 | | | | | | | | |
| 14 | Curtosi | -1,426259343 | | | | | | | | |
| 15 | Asimmetria | 0,316274406 | | | | | | | | |
| 16 | Intervallo | 98 | | | | | | | | |
| 17 | Minimo | 1 | | | | | | | | |
| 18 | Massimo | 99 | | | | | | | | |
| 19 | Somma | 2234 | | | | | | | | |
| 20 | Conteggio | 51 | | | | | | | | |

Vogliamo ora determinare le frequenze con cui si distribuiscono i valori raggruppandoli in classi predefinite, ossia vogliamo sapere quanti valori si trovano, per esempio, tra 0 e 25, quanti tra 26 e 50, poi tra 51 e 75 e infine tra 76 e 99; abbiamo stabilito quindi quattro classi di frequenze. Nella cella G6 scriviamo 25, che è il limite superiore della prima classe, in G7 scriviamo 50 (limite superiore della seconda classe), in G8 il valore 75.

Nel gruppo **Analisi** della scheda **Dati** facciamo clic su **Analisi dati**; nell'elenco scegliamo **Istogramma** e confermiamo con **OK**.

Nella finestra Istogramma

- inseriamo l'Intervallo di input: da A7 a A57;
- inseriamo l'Intervallo della classe: da G6 a G8;
- scriviamo per l'Intervallo di output: H6;
- mettiamo il segno di spunta nella casella Grafico in output;
- confermiamo con OK.

| Input | | | - |
|-------------------------------------|----------------|----------|--------------------|
| Intervallo di input: | \$A\$7:\$A\$57 | 1 | OK |
| I <u>n</u> tervallo della classe: | \$G\$6:\$G\$8 | 1 | OK Annulli 2 |
| Etichette | | | 2 |
| Opzioni di output | | | |
| Intervallo di <u>o</u> utput: | \$H\$6 | | |
| 🔿 Nuovo <u>f</u> oglio di lavoro: | | | |
| 🔊 Nuova car <u>t</u> ella di lavoro | | | |
| 📃 Pareto (istogramma ordi | nato) | | |
| Pe <u>r</u> centuale cumulativa | | | |
| Crafico in output | | | |

Excel mostra in una tabella la distribuzione delle frequenze cercate che visualizza in un grafico a istogramma. A questo punto possiamo:

- cancellare le celle da G6 a G8,
- rendere più leggibile la tabella delle frequenze,
- rendere più chiaro il grafico delle stesse frequenze.

| | A | В | С | D | E | F | G | Н | 1 | J | К | L | M | Ν | 0 | р | Q |
|----|-----------|-----------|------|----------------------|-----------------|--------|------|---------|-----------|-----|------|--------|-----------|------|--------|---------|---|
| 1 | | 1 | | | and an an an an | | | | | | | | | | | | |
| 2 | | A | NALI | SI STATISTICA E | CLASSI D | I FREC | QUEN | ZA | | | | | | | | | |
| 3 | | | | | | | | | | | | | | | | | |
| 4 | | | | | | | | | | | | | | | | | |
| 5 | | | | | | | | | | | | | | | | | |
| 6 | TABELLA D | EI VALORI | | TABELLA DEI VALORI | | | | Classe | Frequenza | | | | | | | | |
| 7 | 98 | | | | | | | 0 - 25 | 21 | | | CLA | SSI DI FR | EQU | ENZA | | |
| 8 | 12 | | | Media | 43,80392157 | | | 26 - 50 | 10 | | 25 - | | | | | | |
| 9 | 93 | | | Errore standard | 4,801471964 | | | 56 - 75 | 7 | | 2604 | | | | | | |
| 10 | 1 | | | Mediana | 41 | | | 76 - 99 | 13 | | | - | | | | | |
| 11 | 93 | | | Moda | 93 | | | | | | 20 - | | | | | | |
| 12 | 1 | | | Deviazione standard | 34,28936839 | | | | | | | | | | | | |
| 13 | 45 | | | Varianza campionaria | 1175,760784 | | | | | ze | 15 - | | | | | | |
| 14 | 18 | | | Curtosi | -1,426259343 | | | | | e | | | | | | - | |
| 15 | 93 | | | Asimmetria | 0,316274406 | | | | | 망 | | | | | | | |
| 16 | 17 | | | Intervallo | 98 | | | | | e l | 10 - | | | | | | |
| 17 | 68 | | | Minimo | 1 | | | | | | | | | | | | |
| 18 | 69 | | | Massimo | 99 | | | | | | 5 - | | | | | | |
| 19 | 9 | | | Somma | 2234 | | | | | | 5 | | | | | | |
| 20 | 41 | | | Conteggio | 51 | | | | | | - | | | | | | |
| 21 | 71 | | | | | | | | | | 0 + | | , | - te | | | |
| 22 | 7 | | | | | | | | | | | 0 - 25 | 26 - 50 | | 56 -75 | 76 - 99 | |
| 23 | 50 | | | | | | | | | | | Classi | | | | | |
| 24 | 89 | | | | | | | | | | | | 23 | | | | |
| 25 | 93 | | | | | | | | | | | | | | | | |
| 26 | 29 | | | | | | | | | | | | | | | | |

Rinominiamo il Foglio1 con "Analisi statistica" e salviamo il lavoro.

3