

I grafici parlano...

Livello scolastico: 2° biennio

| Abilità interessate | Conoscenze | Nuclei coinvolti | Collegamenti esterni |
|--|--|---|-----------------------------|
| Selezionare, produrre ed usare appropriate rappresentazioni grafiche delle distribuzioni doppie. | Principali rappresentazioni grafiche per le distribuzioni doppie rispetto a caratteri di qualsiasi natura. | <u>Dati e previsioni</u> Risolvere e porsi problemi Laboratorio di matematica | |

Contesto

Sociale.

Il contesto è di tipo matematico ed extramatematico; in particolare, per il contesto matematico, l'attività riguarda l'ambito statistico (distribuzioni doppie e rappresentazioni grafiche).

L'attività prende spunto da un'indagine effettuata dagli studenti, nell'ambito dell'area di progetto, dell'I.T.I.S. "C. Zuccante" di Mestre (Ve) sui lavori svolti dagli stessi studenti nel periodo estivo. La zona offre possibilità di lavoro estivo in diversi settori, grazie ad un buono sviluppo industriale e commerciale. L'istituto stesso, inoltre, organizza stages di formazione nel periodo estivo.

La Presidenza dell'Istituto avvertiva l'esigenza di esplorare più a fondo il fenomeno, di sapere se e quanti studenti erano interessati, a conoscere i settori produttivi di maggiore interesse. Per tale motivo ha commissionato un'area di progetto ad una classe terza dello stesso Istituto. Gli studenti hanno provveduto a predisporre un opportuno questionario, a farlo compilare, a codificarlo e a costruire il corrispondente data base relativo a 499 studenti (unità statistiche) rispetto ai 13 caratteri rilevati.

Descrizione dell'attività

Questa unità utilizza alcuni dei risultati dell'analisi per mostrare rappresentazioni grafiche particolarmente utili per sintetizzare le tabelle e coglierne il significato. Ciò consente quindi sia avere una idea più immediata e generale dell'andamento di un carattere, sia la possibilità di confrontare distribuzioni di caratteri diversi. Con l'avvertenza, sempre valida nella rappresentazione per immagini, che una stessa distribuzione presentata in un modo o in un altro sembra cambiare profondamente.

Prima fase

Una prima domanda riguarda l'"Età" degli studenti che lavorano, le "Classi" e i "Corsi di appartenenza": elettronica e telecomunicazioni (ET) ed informatica (IS).

L'insegnante mette in evidenza come il modo in cui le frequenze si dispongono segnala l'esistenza di una connessione fra le età e le classi. Qual è l'età minima e l'età massima degli studenti per ogni classe? La distribuzione per età degli "elettronici" e la distribuzione per età degli "informatici" sono analoghe nelle classi corrispondenti?

Studenti che lavorano d'estate per età e classe di appartenenza (dati assoluti)

| Classe | Età in anni compiuti | | | | | | Totale |
|--------|----------------------|----|----|----|----|----|--------|
| | 16 | 17 | 18 | 19 | 20 | 21 | |
| 3ET | 23 | 8 | 3 | | | | 34 |
| 4ET | | 21 | 22 | 6 | 2 | | 51 |
| 5ET | | | 28 | 10 | 4 | | 42 |
| 3IS | 23 | 9 | 6 | 2 | | | 40 |
| 4IS | | 26 | 12 | 6 | 2 | | 46 |
| 5IS | | | 28 | 15 | 2 | 1 | 46 |
| Totale | 46 | 64 | 99 | 39 | 10 | 1 | 259 |

Tabella 1

L'insegnante, a questo punto, invita gli studenti a rappresentare graficamente la tabella. (*In questa attività faremo riferimento ai colori anche se, per motivi tipografici, i grafici sono riportati in bianco e nero.*) Ecco l'elaborato degli studenti:

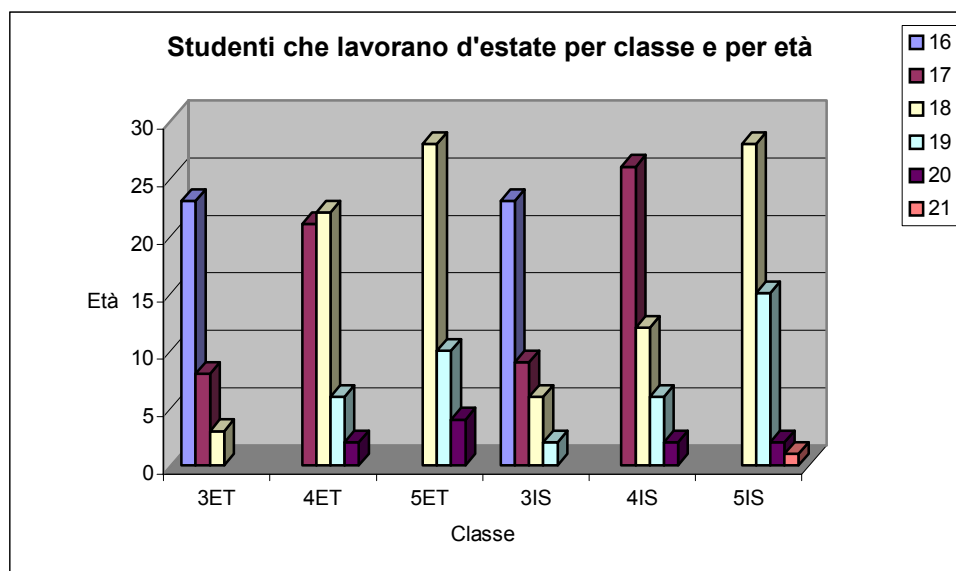


Figura 1

L'insegnante conduce gli studenti ad interpretare il grafico a colonne accostate invitandoli ad interpretare l'andamento delle età al variare delle classi, così come è evidenziato dai colori dei prismi. Un grafico a colonne giustapposte, anziché a prismi giustapposti sarebbe stato più o meno corretto?

Seconda fase

Studenti e insegnanti sono curiosi di sapere se i lavori estivi sono retribuiti e se ciò è condizionato dalla posizione lavorativa. Il data base consente di estrarre la seguente tabella a doppia entrata:

Studenti del triennio che hanno svolto attività
classificati per posizione e per compenso ricevuto (dati assoluti)

| Posizione | Compenso | | Totale |
|-------------|----------|----|--------|
| | Sì | No | |
| Impiegato | 16 | 1 | 17 |
| Operaio | 67 | 2 | 69 |
| Apprendista | 109 | 2 | 111 |
| Stage | 24 | 5 | 29 |
| In Proprio | 14 | 2 | 16 |
| Socio | 2 | | 2 |
| Coadiuvante | 2 | | 2 |
| Volontario | 5 | 8 | 13 |
| Totale | 239 | 20 | 259 |

Tabella 2

Sono confrontabili fra loro le otto distribuzioni del compenso rispetto alla posizione?
La risposta è ovviamente negativa ed è perciò necessario passare alla tabella con i dati percentuali, nella quale tutte le righe hanno per somma cento:

Studenti del triennio che hanno svolto attività
suddivisi per posizione e per compenso ricevuto (dati percentuali).

| Posizione | Compenso | | Totale |
|-------------|----------|------|--------|
| | Sì | No | |
| Impiegato | 94,1 | 5,9 | 100 |
| Operaio | 97,1 | 2,9 | 100 |
| Apprendista | 98,2 | 1,8 | 100 |
| Stage | 82,8 | 17,2 | 100 |
| In Proprio | 87,5 | 12,5 | 100 |
| Socio | 100 | | 100 |
| Coadiuvante | 100 | | 100 |
| Volontario | 38,5 | 61,5 | 100 |
| Totale | 92,3 | 7,7 | 100 |

Tabella 3

L'insegnante chiede agli studenti il significato della distribuzione che compare nell'ultima riga. Il valore 92,3 come si giustifica? Dopo aver condotto ad osservare che esso è compreso tra 82,8 e 100, l'insegnante fa riflettere gli studenti sull'uguaglianza:

$$\frac{0,941 \cdot 17 + 0,971 \cdot 69 + 0,982 \cdot 111 + 0,828 \cdot 29 + 0,875 \cdot 16 + 1 \cdot 2 + 1 \cdot 2 + 0,385 \cdot 13}{259} = \frac{239,015}{259} \cong 0,923$$

Ritornando al problema delle rappresentazioni grafiche, qual è il grafico più opportuno per rappresentare contemporaneamente tutte le distribuzioni condizionate del "Compenso" rispetto alla "Posizione"? Ha influenza il fatto che entrambi i caratteri siano qualitativi?

L'insegnante conduce gli studenti a costruire il grafico della Figura 2.

Come ha influenzato il grafico il fatto che tutte le distribuzioni condizionate danno per somma cento?

Come è stato possibile evidenziare la percentuale di studenti che ha ricevuto un compenso condizionatamente alla posizione lavorativa?

L'insegnante guida gli studenti verso le risposte: dal grafico è immediatamente visibile il fatto che il totale è cento e il doppio colore permette di visualizzare le differenti percentuali di coloro che hanno o non hanno ricevuto un compenso.

Naturalmente non era necessario ricorrere ad un grafico tridimensionale che tuttavia è quello scelto dagli studenti.

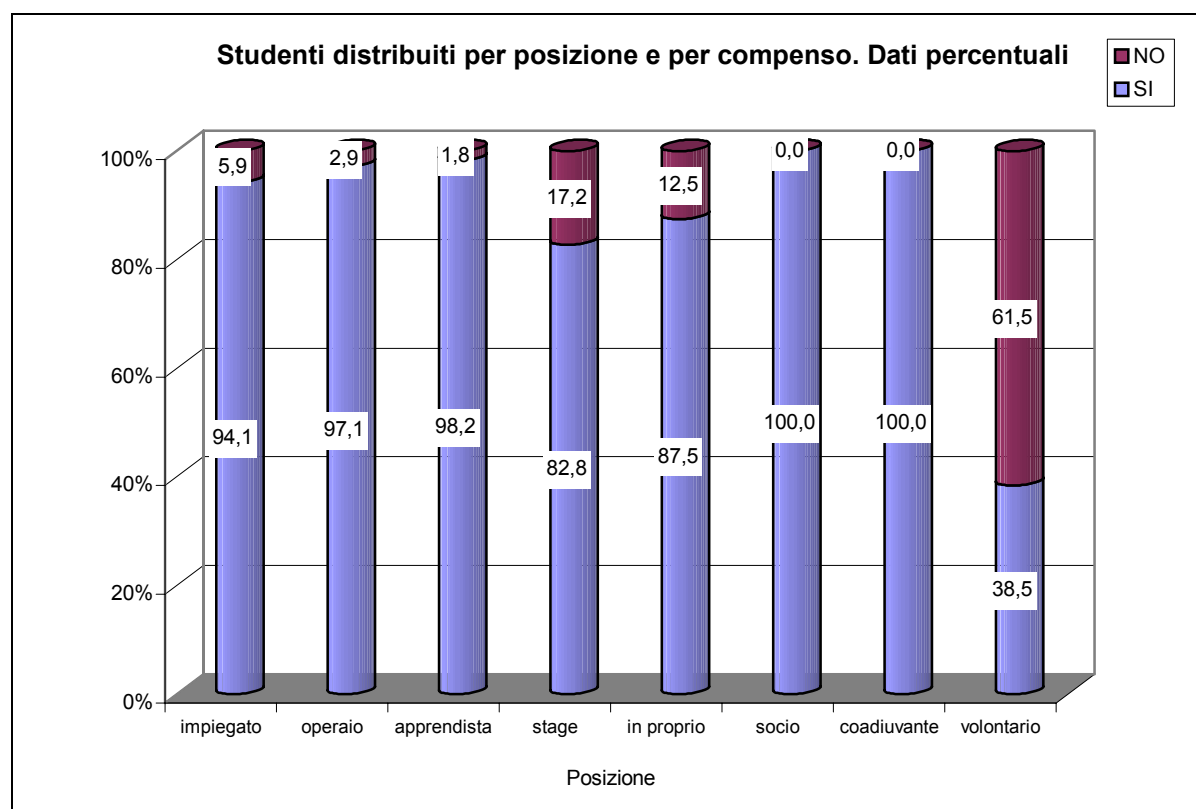


Figura 2

Come si poteva prevedere la percentuale più alta di studenti senza compenso ha operato nel volontariato, mentre chi lavora come socio o come coadiuvante riceve in ogni caso una retribuzione. Attenzione a trarre conclusioni affrettate: in ciascuna di queste due posizioni hanno operato soltanto due studenti!

Terza fase

L'attività proposta costituisce un approfondimento opzionale seppur molto suggestivo per l'interpretazione delle relazioni interne ad una distribuzione doppia. Essa fa perno sui concetti di primo quartile, mediana e terzo quartile di una distribuzione di un carattere almeno ordinato rettilineo.

Una sintesi grafica della tabella doppia, quando almeno uno dei caratteri è ordinato, è fornita dall'insieme di più diagrammi a scatola ("Boxplot"). Nel "boxplot" l'escursione del carattere si valuta in ordinata. Il primo ed il terzo quartile sono i valori in ordinata che delimitano la scatola stessa, mentre la mediana è l'altezza relativa al simbolo corrispondente interno alla scatola. A

partire da ciascun lato orizzontale della scatola sino all'ultima osservazione di coda (valore minimo, valore massimo) si disegna poi un segmento verticale.

Quando, come in una distribuzione doppia con almeno un carattere ordinale, si confrontano giustapponendoli più "boxplot" tutti relativi a tale carattere, in ascissa si pone un asse categorico che serve per spaziare i grafici corrispondenti alle diverse modalità dell'altro carattere.

Per rispondere alla domanda se il tempo lavorativo in giorni si differenzia a seconda del settore d'impiego, una soluzione potrebbe essere quella di costruire un insieme di "boxplot" utilizzando i dati riportati nella Tabella 4. In essa sono indicati, per ciascuno dei settori più importanti, il minimo (min), il primo quartile (q1), la mediana, il terzo quartile (q3) ed il massimo (max).

Alcuni valori caratteristici per principali
settori d'impiego

| Valori caratteristici | Settore | | | |
|--------------------------|-------------|----------------|--------------|-------------|
| | elettronico | manifatturiero | ristorazione | informatico |
| min | 14 | 15 | 2 | 14 |
| q1 | 30 | 40 | 42 | 15 |
| mediana | 59 | 55 | 70 | 30 |
| q3 | 70 | 60 | 90 | 53 |
| max | 120 | 170 | 180 | 90 |

Tabella 4

Il "boxplot" della durata in giorni del "Rapporto di lavoro" condizionato al "Settore d'impiego" è riportato nella Figura 3.

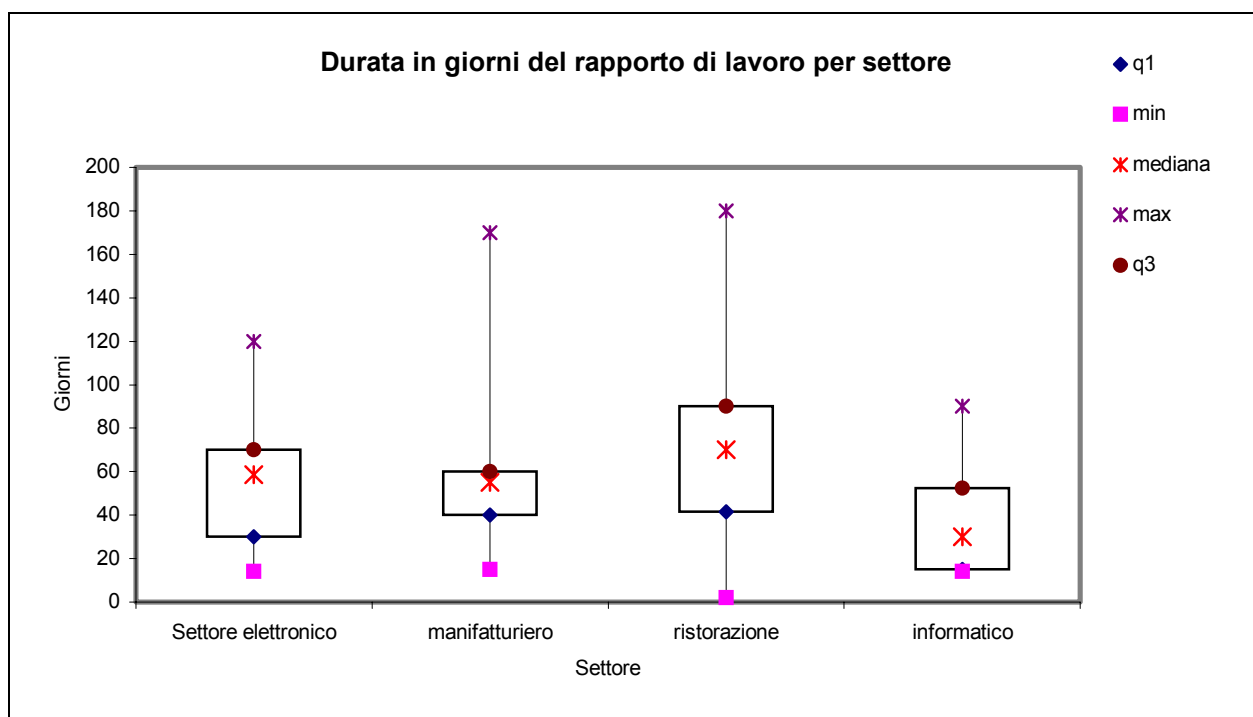


Figura 3

Qual è il settore in cui si manifesta il minor campo di variazione rispetto alla durata del rapporto di lavoro? E quello con maggiore variabilità? Può incidere il contesto economico del territorio ad alta densità turistica e manifatturiera?

Poiché all'interno della scatola cade circa il 50% del collettivo, l'insegnante conduce gli studenti ad osservare che, per esempio, il 50% circa degli studenti del settore manifatturiero ha "resistito" fra 40 e 60 giorni; nel settore elettronico la durata del rapporto di lavoro del 50% circa degli studenti che occupano la parte centrale della distribuzione ordinata dei giorni di lavoro è compresa tra 30 e 70 giorni.

Cosa succede negli altri settori? Nel settore informatico, qual è il numero massimo dei giorni di lavoro del 75% degli studenti che lavorano di meno? L'insegnante ricorda il significato del terzo quartile e porta gli studenti alla risposta.