Analisi del Data set #1

Ad un campione di 30 studenti delle scuole superiori, intenzionati a proseguire gli studi iscrivendosi ad un corso di laurea triennale, è stato chiesto a quale tipo di corso intendono iscriversi scegliendo fra i seguenti indirizzi: Medico Sanitario (MS), Economico (Ec), Industriale (In), Scienze di Base (SB), Altri (Al). Le risposte date dagli studenti sono nella Tabella 1.

**Tabella 1**.Risposte al questionario proposto

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| In | Al | SB | Al | Al | MS | Al | Ec | In | Ec |
| Al | Al | Al | Ec | In | Ec | Ec | In | Al | Al |
| MS | In | Al | MS | MS | Ec | SB | Al | Al | MS |

Le risposte al questionario rappresentano un **raw** **data set** di tipo **qualitativo nominale**; le risposte al questionario date da ciascun allo studente (variabile casuale dell’esperimento) rappresentano le **categorie** del data set. Per interpretare le risposte è necessario costruire la **tabella della distribuzione delle frequenze** dove accanto a ciascuna categoria viene indicata la frequenza di ciascuna risposta. Utilizzare le funzioni predefinite di Excel; le risposte sono contenute nel file di testo DS#1.txt.

**Preparazione del Raw Data Set.**

1. Importare il file DS#1.txt nel foglio Excel nelle (A1:J3)
2. Spostare i dati in basso e a destra per inserire un titolo
   1. Seleziona/Trascina; Taglia/Incolla.
3. Selezionare le celle contenenti le risposte e impostare il Formato Celle come Testo.
4. Impostare la Larghezza delle Colonne uguale a 4 mm
   1. Formato🡺Colonna 🡺Larghezza (inserire il valore 8 nella apposita finestra)
5. Unificare le celle contenenti il titolo ed evidenziarle
   1. Selezionare le celle contenenti il titolo
   2. Premere il pulsante “unisci-centra” sulla barra dei formati (in alternativa si imposti il Formato🡺Celle🡺Allineamento🡺unione celle) da menu Formato.
   3. Inserire un colore di riempimento alla cella
6. Selezionare i dati per introdurre un Formato Condizionale alle celle che contengono la categoria Ingegneria (In).
   1. Selezionare i dati
   2. Formato🡺Formattazione Condizionale
      1. scegliere la Condizione
      2. definire il Formato

L’aspetto Finale del DataSet#1 deve essere il seguente

**Tabella 2.** Raw Data set con le risposte al questionario

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **In** | Al | SB | Al | Al | MS | Al | Ec | **In** | Ec |
| Al | Al | Al | Ec | **In** | Ec | Ec | **In** | Al | Al |
| MS | **In** | Al | MS | MS | Ec | SB | Al | Al | MS |

**Costruzione della Tabella con La distribuzione delle Frequenze.**

L’analisi di un data set qualitativo nominale permette di definire per ciascuna categoria i valori di frequenza, frequenza relativa e della probabilità. Di seguito vengono indicati le istruzioni in ambiente Excel che consentono di costruire la tabella

1. Delimitazione di un area del foglio elettronico destinata alla tabella
2. Definizione delle categorie
3. Calcolo delle **frequenze** *f* di ciascuna categoria
   1. Inserimento della Funzione Predefinita Statistica CONTA.SE(*Data set*; *categoria*) per a fianco di ciascuna categoria.
4. Calcolo delle **frequenze** **relative** *f*i e delle **probabilità**
   1. Calcolo della **frequenza totale** (*f*t)come somma delle frequenze *f*
   2. Calcolo della **frequenza** **relativa** di ciascuna categoria come rapporto *f*ir = *f*i / *ft*
   3. Calcolo della **probabilità** come prodotto *f*ir\*100

L’aspetto Finale della Tabella con la distribuzione delle frequenze deve essere il seguente.

# Tabella 3. Corsi di Laurea scelti da 30 studenti sceDistribuzione delle Frequenze

|  |  |  |  |
| --- | --- | --- | --- |
| Categorie | Frequenza | Fr. Relativa | Probabilità (%) |
| Altri | 12 | 0.39 | 39 |
| Economia | 6 | 0.20 | 20 |
| Ingegneria | 5 | 0.17 | 17 |
| Medico Sanitario | 5 | 0.17 | 17 |
| Scienze di base | 2 | 0.07 | 7 |
| Somma | 30 | 1.00 | 100 |

**Rappresentazione grafica dell’informazione contenuta in un data set qualitativo.**

L’informazione presente in un data set di tipo qualitativo può essere rappresentata in modo grafico utilizzando il **Grafico a Barre** per illustrare la Distribuzione delle Frequenze e la **Pie Chart** per illustrare la distribuzione delle probabilità delle diverse categorie.

|  |
| --- |
| **Grafico a Barre.** Diagramma cartesiano (x, y) dove in ascissa sono elencate le categorie e in ordinata la frequenza. La **distribuzione delle frequenze** è rappresentata con una serie di barre verticali; la frequenza di ciascuna categoria è indicata dalla altezza della barra corrispondente. |

|  |
| --- |
| **Pie Chart.** La Pie Chart o grafico a torta è costituito da un cerchio la cui area è suddivisa in settori la cui area è proporzionale alla **frequenza relativa** di ciascuna categoria. L’area di ciascun settore è pari a *f*ir\*360 e la somma dei settori è pari all’area del cerchio. |

Il Grafico a Barre ed il grafico Tipo Torta si ricavano entrambi con la procedura guidata Inserisci Grafico di Excel.

**Grafico a Barre della distribuzione delle Frequenze**

**Pie Chart della distribuzione delle probabilità**