Esercizio Statistica

Il dataset seguente include 20 studenti diplomati in Italia nell’anno scolastico 2021/22, di cui sono state raccolti le seguenti informazioni:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Macro Regione Diploma | Sesso | Score Test | Tipo Diploma | Titolo Studio Padre | Voto Diploma |
| Sud | M | 17,8 | Ist. Professionale | Licenza media | 68 |
| Nord | F | 21,2 | Liceo | Diploma | 97 |
| Sud | M | 16,3 | Ist. Tecnico | Diploma | 73 |
| Nord | F | 21,0 | Liceo | Laurea | 100 |
| Centro | M | 19,7 | Ist. Professionale | Diploma | 83 |
| Sud | F | 18,4 | Liceo | Laurea | 101 |
| Centro | F | 15,7 | Liceo | Laurea | 98 |
| Nord | M | 17,7 | Ist. Tecnico | Diploma | 90 |
| Nord | F | 17,8 | Ist. Tecnico | Diploma | 62 |
| Nord | M | 14,8 | Liceo | Laurea | 71 |
| Sud | M | 17,4 | Ist. Tecnico | Laurea | 70 |
| Nord | M | 23,3 | Liceo | Licenza media | 62 |
| Sud | F | 16,5 | Liceo | Laurea | 101 |
| Centro | M | 17,5 | Ist. Professionale | Licenza media | 70 |
| Sud | F | 26,0 | Liceo | Licenza media | 100 |
| Nord | M | 18,5 | Ist. Tecnico | Licenza media | 68 |
| Nord | F | 21,0 | Liceo | Laurea | 78 |
| Centro | F | 15,3 | Ist. Tecnico | Licenza media | 82 |
| Sud | M | 23,3 | Liceo | Diploma | 87 |
| Centro | M | 21,5 | Liceo | Diploma | 101 |

1. Definisci la natura di ciascuna variabile (n.b.: non basta dire solo se è quantitativa o qualitativa).
   1. Qual è la differenza tra lo score e il voto di diploma?
2. Costruisci la distribuzione di frequenze (assolute, relative e percentuali, anche le cumulate) per il titolo di studio del padre.
   1. Qual è la percentuale di studenti per cui il padre non ha conseguito il titolo di laurea?
   2. Rappresenta graficamente le frequenze relative.
   3. Qual è la classe mediana? Utilizza le frequenze assolute per individuarla.
   4. Qual è la moda?
3. Dividi la variabile *voto diploma* in quattro classi di uguale ampiezza e costruisci la distribuzione di frequenze (anche le cumulate).
   1. Qual è la percentuale di studenti che ha avuto più di 70?
   2. Rappresenta graficamente le frequenze assolute.
   3. Qual è la classe mediana? Utilizza le frequenze percentuali per individuarla.
   4. Senza considerare la distribuzione in classi, indica il valore mediano.
4. Individua la classe mediana per il tipo di diploma.
5. Sono stati raccolti i dati relativi ai voti di diploma **(x)** di un campione di studenti provenienti da licei, istituti tecnici e professionali.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Liceo** | | **Tecnico** | | **Professionale** | |
| **x1** | **n1** | **x2** | **n2** | **x3** | **n3** |
| 91 | 12 | 73 | 8 | 68 | 8 |
| 75 | 5 | 68 | 6 | 70 | 10 |
| 82 | 7 | 90 | 7 | 83 | 4 |
| 95 | 8 | 82 | 10 | 91 | 2 |
| 88 | 9 | 70 | 4 | 86 | 3 |

* 1. Calcola la media e la mediana per ciascun gruppo.
  2. Calcola la media ponderata tra i tre gruppi

1. Considera la distribuzione dei punteggi al test, distinguendo gli studenti tra liceali e non liceali.

|  |  |
| --- | --- |
| Liceo | Non Liceo |
| 21,2 | 17,8 |
| 21,8 | 17,5 |
| 18,4 | 19,7 |
| 15,7 | 16,3 |
| 14,8 | 18,5 |
| 23,3 | 17,7 |
| 16,5 | 22,8 |
| 26,0 | 17,4 |
| 21,0 | 17,8 |
| 23,3 |  |
| 21,5 |  |

* 1. Calcola la media e la mediana per ciascun gruppo.
  2. Calcola la media ponderata tra i due gruppi.
  3. Calcola la varianza e la deviazione standard dei due gruppi e indica quale è più variabile.
  4. Dai risultati ottenuti, indica per entrambe le distribuzioni se c’è simmetria, asimmetria positiva o negativa.