

DISTRIBUZIONI DI FREQUENZA.

1. Calcolare le frequenze assolute della variabile “resistenza all’attacco da *Peronospora* in individui adulti di vite” a partire dal seguente dataset e rappresentare graficamente i dati:

Individuo	Resistenza
1	Alta
2	Molto alta
3	Alta
4	Media
5	Bassa
6	Bassa
7	Media
8	Nulla
9	Nulla
10	Alta
11	Media
12	Non rilevabile
13	Bassa
14	Media
15	Alta
16	Molto alta

Risoluzione

La variabile “Resistenza” è di tipo ordinale.
Passare da dati grezzi a serie ordinata:

Individuo	Resistenza
12	Non rilevabile
8	Nulla
9	Nulla
5	Bassa
6	Bassa
13	Bassa
4	Media
7	Media
11	Media
14	Media
1	Alta
3	Alta
10	Alta
15	Alta
2	Molto alta
16	Molto alta

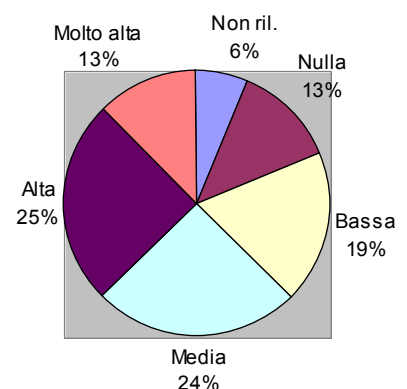
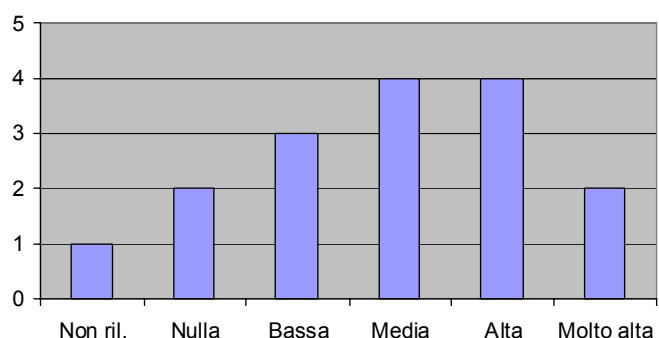
Distribuzione di frequenza:

Contare quante osservazioni ricadono in ciascuna classe.

Classe	Non ril.	Nulla	Bassa	Media	Alta	M.alta
Frequenza	1	2	3	4	4	2
N Totale						16

Rappresentazioni grafiche:

Istogramma; diagramma a torta.



N.B.: La percentuale è calcolata attraverso la proporzione:

$$\text{frequenza} : N \text{ totali} = x : 100\%$$

2. Calcolare le frequenze assolute, relative e cumulate della variabile “temperatura media annua dell’orizzonte organico in suoli agrari del Piemonte” a partire dal seguente dataset e rappresentare graficamente i dati:

Campione	Temperatura
Torino	17 °C
Pinerolo	16
Poirino	17
Biella	14
Ivrea	15
Asti	16
Alessandria	18
Cuneo	15
Demonte	12
Novara	17
Casale	18
Vercelli	17

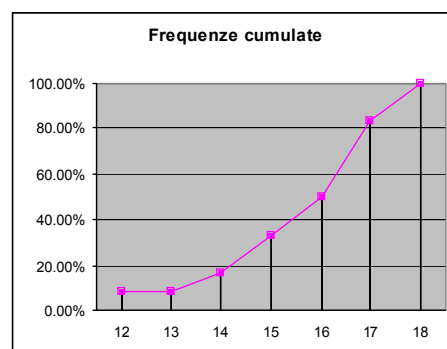
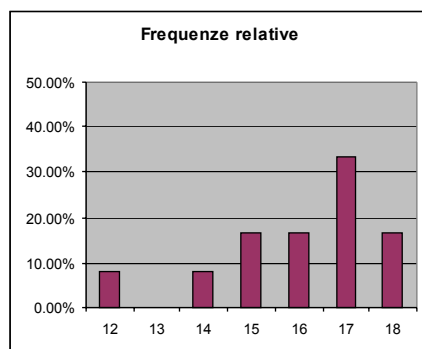
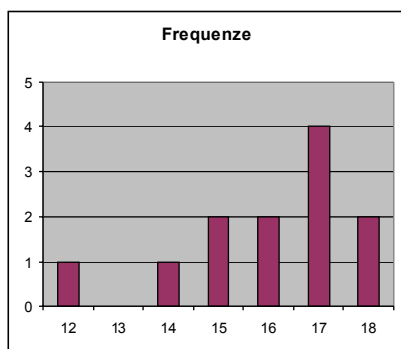
La variabile “Temperatura” è di tipo a intervalli (lo zero della scala è convenzionale).
 Passare da dati grezzi a serie ordinata:

Campione	Temperatura
Demonte	12
Biella	14
Ivrea	15
Cuneo	15
Pinerolo	16
Asti	16
Poirino	17
Novara	17
Vercelli	17
Torino	17
Alessandria	18
Casale	18

Distribuzione di frequenza:

Contare quante osservazioni ricadono in ciascuna classe. Le frequenze relative si ottengono calcolando la proporzione $\text{Freq. Assoluta} : N \text{ Totale} = x : 100$. Le frequenze cumulate di ciascuna classe si calcolano sommando la frequenza relativa della classe considerata con le frequenze relative di tutte le classi inferiori.

Classe	12	13	14	15	16	17	18	
Frequenza	1	0	1	2	2	4	2	N Totale = 12
Freq. relativa	8.33%	0	8.33%	16.67%	16.67%	33.33%	16.67%	Totale = 100%
Freq. cumulata	8.33%	8.33%	16.66%	33.33%	50%	83.33%	100%	



3. Rappresentare graficamente il seguente dataset relativo alla lunghezza dell'apice vegetativo (cm) in individui di Pino strobo in vivaio. Quanti Pini hanno l'apice più lungo di 19 cm?

12	8	7	15	14	16	6	14
11	5	17	21	17	13	19	10
3	15	16	16	20	17	13	9
4	20	16	18	19	19	18	17
23	8	9	9	25	26	16	14

La variabile "Lunghezza dell'apice" è di tipo a rapporti (può assumere qualsiasi valore, lo zero della scala è reale).
Passare da dati grezzi a serie ordinata:

3	8	10	14	16	17	18	20
4	8	11	14	16	17	19	21
5	9	12	14	16	17	19	23
6	9	13	15	16	17	19	25
7	9	13	15	16	18	20	26

Distribuzione di frequenza:

Non è possibile utilizzare una classe per ogni centimetro, altrimenti si otterrebbero troppe classi. Per avere un'INDICAZIONE del numero di classi (C) da utilizzare si può fare ricorso alla formula di Sturges:

$$S = 1 + 10/3 * \log_{10} (N)$$

$$S = 1 + 10/3 * \log_{10} (40) = 6.34 \approx 6 \text{ classi}$$

Per ottenere l'ampiezza di ogni classe (in cm) è possibile suddividere il RANGE di dati (dato maggiore – dato minore) per il numero di classi:

$$(26-3) : 6 = 3.83 \approx 4 \text{ cm}$$

Poiché la totalità dei dati del campione è riportata nella tabella precedente, il limite inferiore della prima classe può essere 3 (non c'è nessun dato inferiore a questo valore), ma anche un limite inferiore pari a 0 (origine reale della scala di misura) è corretto (ma in questo caso l'ampiezza delle classi va calcolata secondo la formula $(26-0)/6$). Con una precisione di 0.01 cm, le 6 classi quindi sono:

3.00 - 6.99

7.00 – 10.99

11.00 – 14.99

15.00 – 18.99

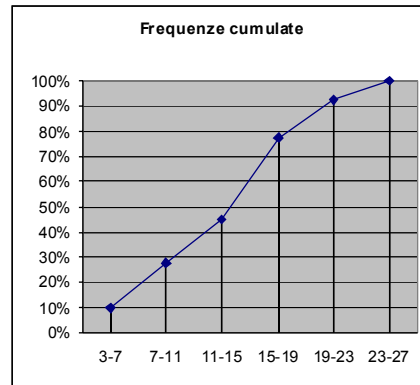
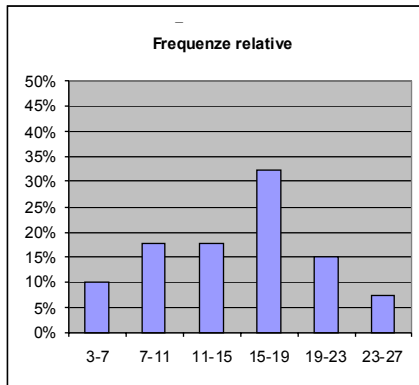
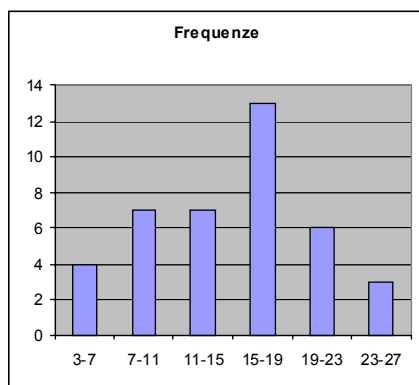
19.00 – 22.99

23.00 – 26.99

Per comodità, queste classi potranno essere scritte come 3-7, 7-10, 10-15 ecc., ma avendo ben presente i confini reali di ogni classe (non devono esserci valori attribuibili a più di una classe). Con una precisione di 1 cm, le classi invece sarebbero 3-6, 7-10, 11-14, 15-18, 19-22, 23-26. È comunque preferibile ricorrere alla prima alternativa, in cui appare più evidente che l'ampiezza di ogni classe è pari a 4 cm.

Contare quante osservazioni ricadono in ciascuna classe. Le frequenze relative si ottengono calcolando la proporzione Freq. Assoluta : N Totale = x : 100. Le frequenze cumulate di ciascuna classe si calcolano sommando la frequenza relativa della classe considerata con le frequenze relative di tutte le classi inferiori.

Classe	3-7	7-11	11-15	15-19	19-23	23-27	
Frequenza	4	7	7	13	6	3	N Totale = 40
Freq. relativa	10%	17.5%	17.5%	32.5%	15%	7.5%	Totale = 100%
Freq. cumulata	10%	27.5%	45%	77.5%	92.5%	100%	



Per rispondere alla domanda “Quanti Pini hanno l’apice più lungo di 19 cm” occorre calcolare le frequenze cumulate “dall’alto”, cioè sommando le frequenze relative a partire dalla classe superiore e procedendo in senso inverso:

Classe	3-7	7-11	11-15	15-19	19-23	23-27	
Frequenza	4	7	7	13	6	3	N Totale = 40
Freq. relativa	10%	17.5%	17.5%	32.5%	15%	7.5%	Totale = 100%
Freq. Cumulate “dall’alto”	100%	90%	72.5%	55%	22.5%	7.5%	



Le classi corrispondenti ad una lunghezza dell’apice uguale o superiore a 19 cm, cioè le ultime due, assommano una frequenza cumulata del 22.5%.