

Tablelle e grafici di frequenza

Organizzare e visualizzare i dati



Esempio

Numero di protozoi contati in un CCS di 33 unità campionarie prelevate da uno stagno.

171 169 165 171 166 172 166 166 168 168 172
175 174 168 169 174 169 171 175 169 171 168
163 169 165 169 169 171 166 171 165 172 169

Il valore dei dati varia da 163 a 175 (13 valori distinti)



Tabella delle frequenze

Numero di protozoi	Frequenza
163	1
164	0
165	3
166	4
167	0
168	4
169	8
170	0
171	6
172	3
173	0
174	2
175	2

Somma delle frequenze: 33



Tabelle e grafici di frequenza

Servono a descrivere un insieme di dati che hanno un numero relativamente basso di valori distinti.

Ad ogni valore distinto dei dati è associata la rispettiva frequenza di occorrenza (numero di volte in cui compare il dato)



Tabella delle frequenze

Numero di protozoi	Frequenza
163	1
164	0
165	3
166	4
167	0
168	4
169	8
170	0
171	6
172	3
173	0
174	2
175	2

Quante unità campionarie hanno meno di 166 protozoi?



Dati ricavabili dalla tabella delle frequenze

- Quante unità campionarie hanno almeno 165 protozoi?
- Quante unità campionarie hanno da 167 a 170 protozoi?
- Quante unità campionarie hanno più di 170 protozoi?
- La metà delle unità campionarie ha più di 170 protozoi. Vero o falso?



Grafico a bastoncini o digramma a segmenti

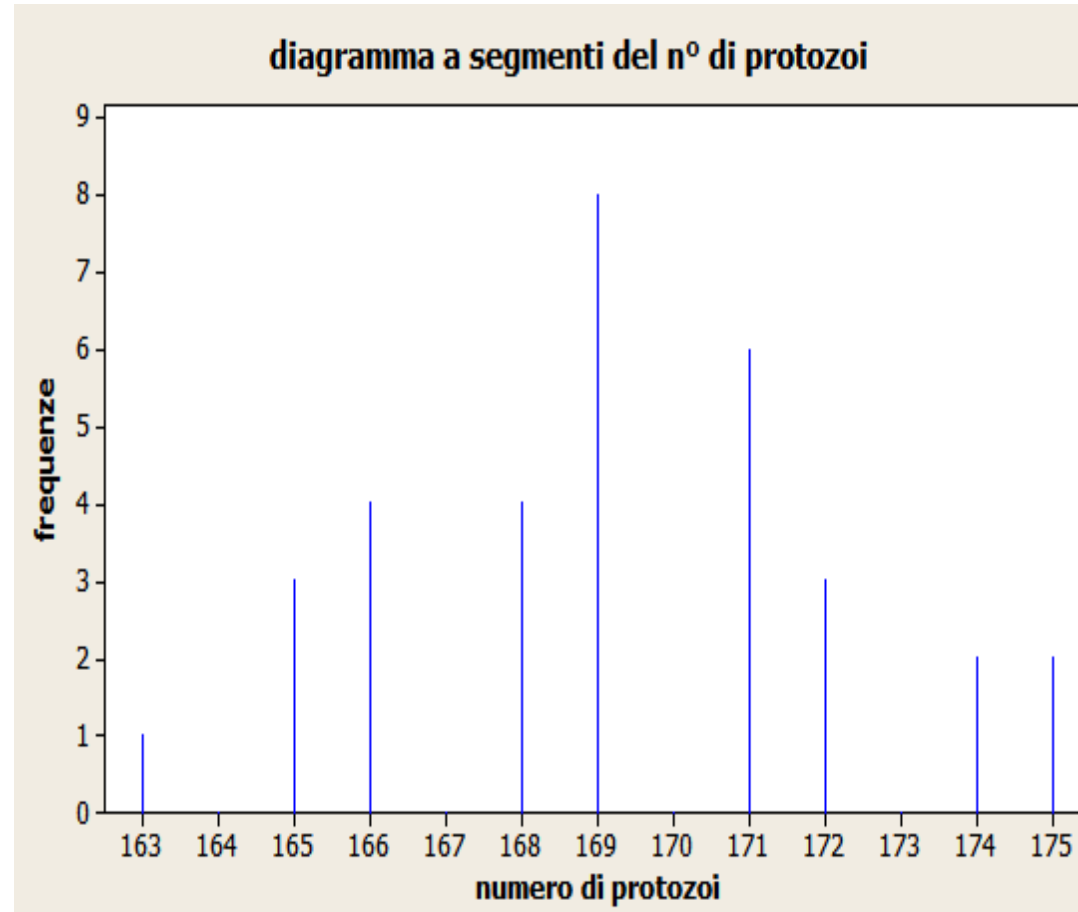
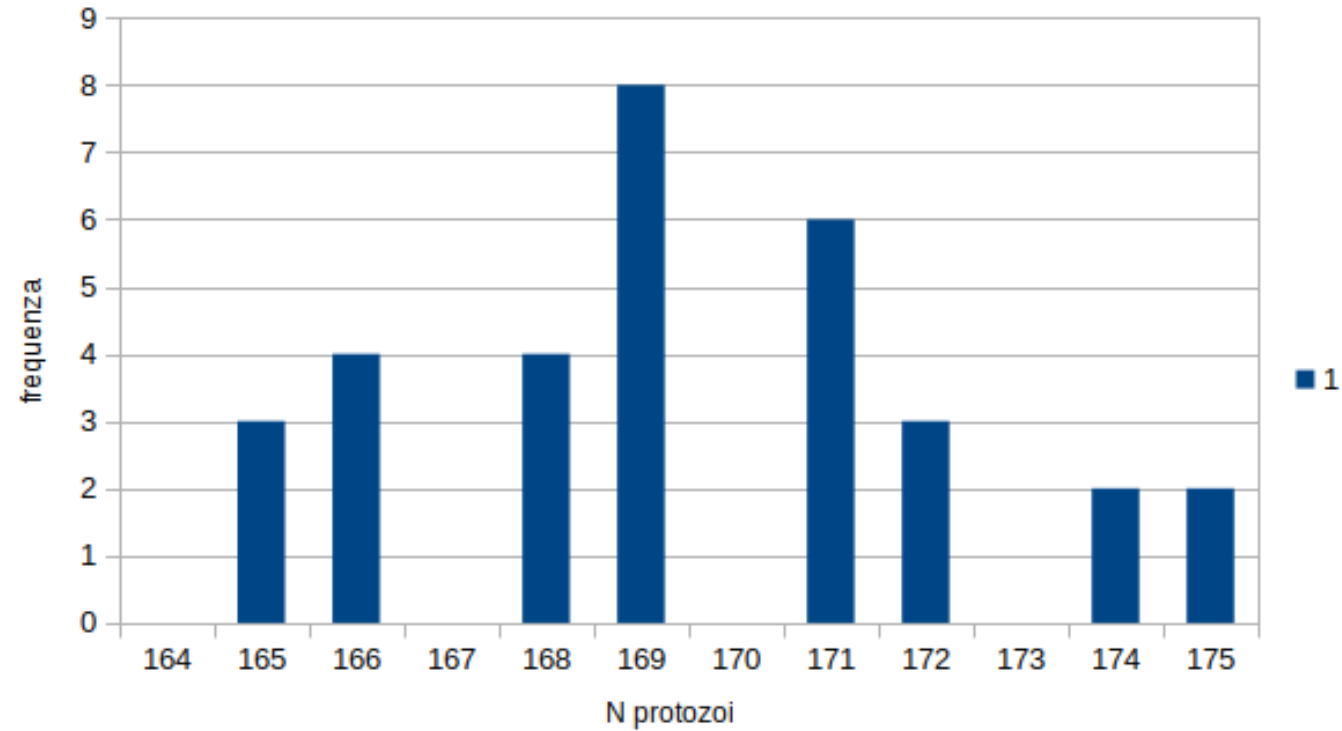
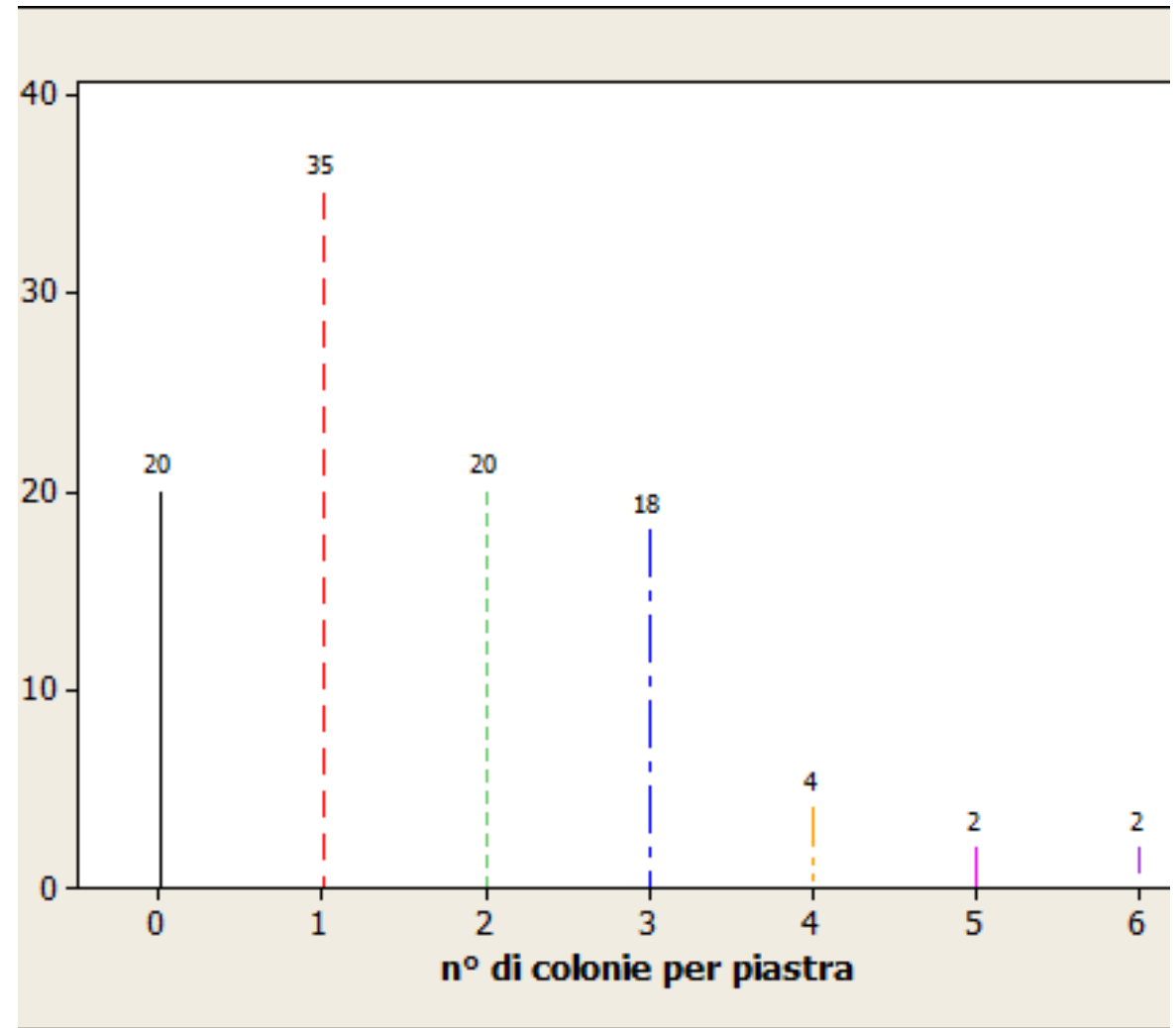


Grafico a barre



Esempio 2. Frequenza del n° di colonie di microrganismi su 90 piastre: diagramma a segmenti



Domande

- Quante piastre non hanno nessuna colonia?
- Quante piastre hanno almeno una colonia?
- Qual'è il numero massimo di colonie per piastra?
- Quante piastre hanno più di 3 colonie?
- Quale percentuale di piastre ha un numero di colonie compreso tra 2 e 3?
- Più della metà delle piastre ha un numero di colonie compreso tra 1 e 2 (V/F)



Frequenze relative

f: frequenza di occorrenza del valore x

n: numero totale di osservazioni

Si definisce la frequenza relativa il rapporto f/n

Osservazioni:

la somma delle frequenze è uguale al numero totale di osservazioni

La somma delle frequenze relative è uguale a 1

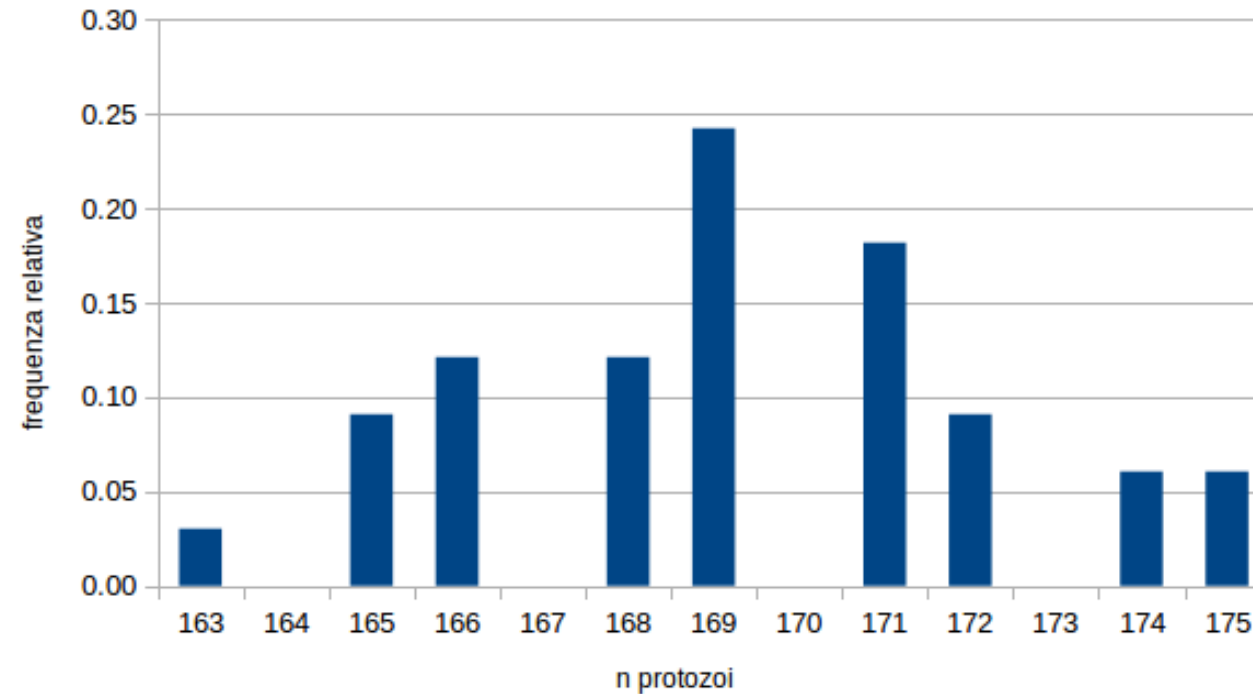


Esempio 1: numero di protozoi in un CCS (33 unità)

Numero di protozoi	Frequenza	Frequenza relativa
163	1	$1/33 = 0.03$
164	0	$0/33 = 0$
165	3	$3/33 = 0.09$
166	4	0.12
167	0	0
168	4	0.12
169	8	0.24
170	0	0
171	6	0.18
172	3	0.09
173	0	0
174	2	0.06
175	2	0.06



Grafico a barre delle frequenze relative



Esempio 1: numero di protozoi in un CCS (33 unità)

Numero di protozoi	Frequenza	Frequenza relativa
163	1	$1/33=0.03$
164	0	$0/33=0$
165	3	$3/33=0.09$
166	4	0.12
167	0	0
168	4	0.12
169	8	0.24
170	0	0
171	6	0.18
172	3	0.09
173	0	0
174	2	0.06
175	2	0.06


Qual è la percentuale di unità campionarie con un numero di protozoi minore di 166?

($0.03 + 0 + 0.09 = 0.12$ Risposta 12 %)



Domande

- Qual è la percentuale di unità campionarie con un numero di protozoi minore di 170?
- Qual è la percentuale di unità campionarie con un numero di protozoi maggiore o uguale a 172?
- Più del 20% delle unità campionarie ha un numero di protozoi uguale a 169 (V/F).



Esempio 3: tabella delle frequenze per la variabile numero di orchidee in un CCS di 50 quadrati

n° orchidee	frequenza	frequenza relativa
0	21	0.42
1	15	0.30
2	6	0.12
3	3	0.06
4	2	0.04
5	1	0.02
8	1	0.02
12	1	0.02
totali	50	1.00



Grafico a barre delle frequenze

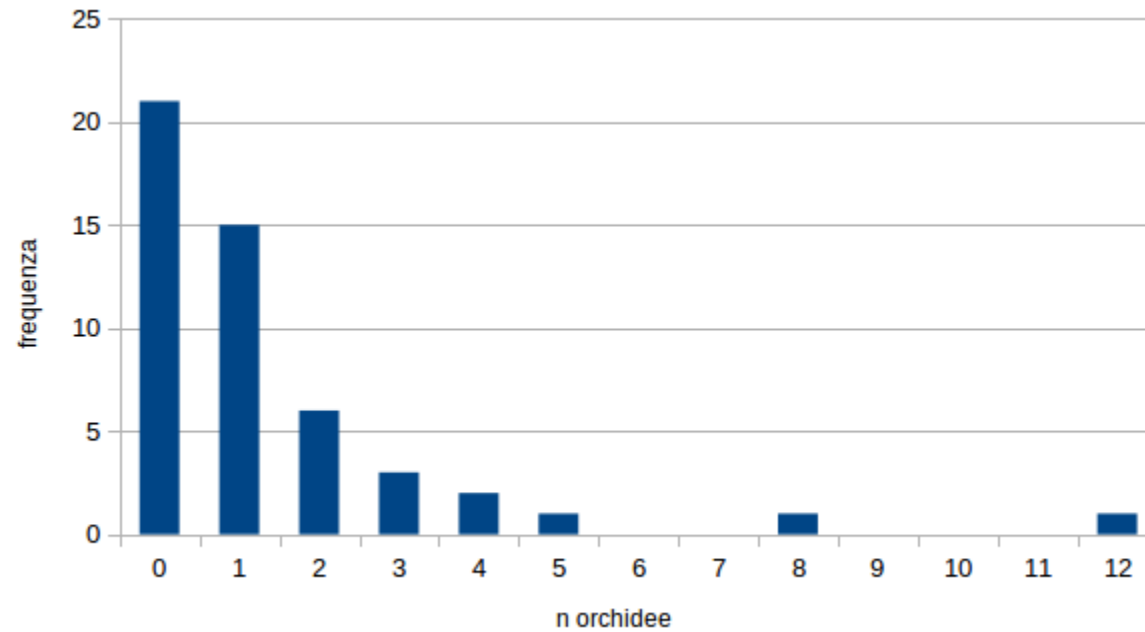
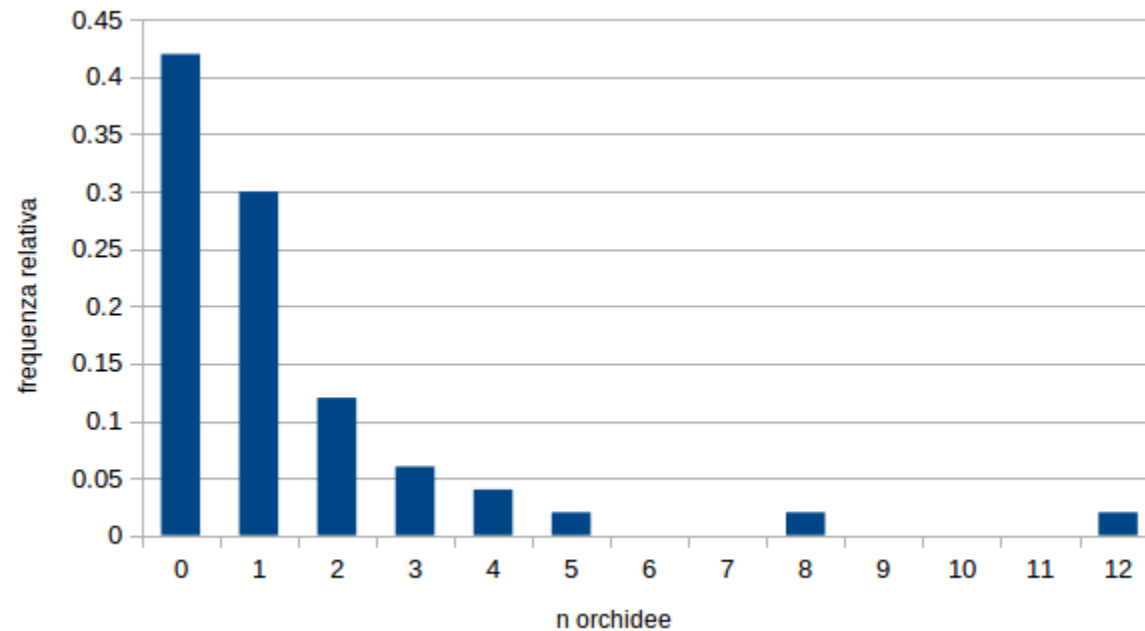


Grafico a barre delle frequenze relative



Qual è la frazione di quadrati che non ha orchidee?
($21/50 = 0,42$)

Qual è la percentuale di quadrati con al più una orchidea?
(frazione: $(21+15)/50 = 0,42 + 0,30 = 0,72$. In percentuale 72 %)

Qual è probabilità, scegliendo a caso un quadrato, di trovare una sola orchidea?
($15/50 = 0,30$)

n° orchidee	frequenza	frequenza relativa
0	21	0.42
1	15	0.30
2	6	0.12
3	3	0.06
4	2	0.04
5	1	0.02
8	1	0.02
12	1	0.02
totali	50	1.00



Domande

Qual è la frazione di quadrati che ha orchidee?

Qual è la percentuale di quadrati che ha almeno una orchidea?

Qual è la percentuale di quadrati con più di 5 orchidee?

Scegliendo a caso un quadrato, qual è la probabilità che abbia 3 orchidee?

E che ne abbia zero?



Colore degli occhi degli alunni di una prima elementare (20 alunni)

Colore degli occhi	Frequenza	Frequenza relativa
blu	5	0.25
verdi	3	0.15
castani	8	0.40
neri	4	0.20

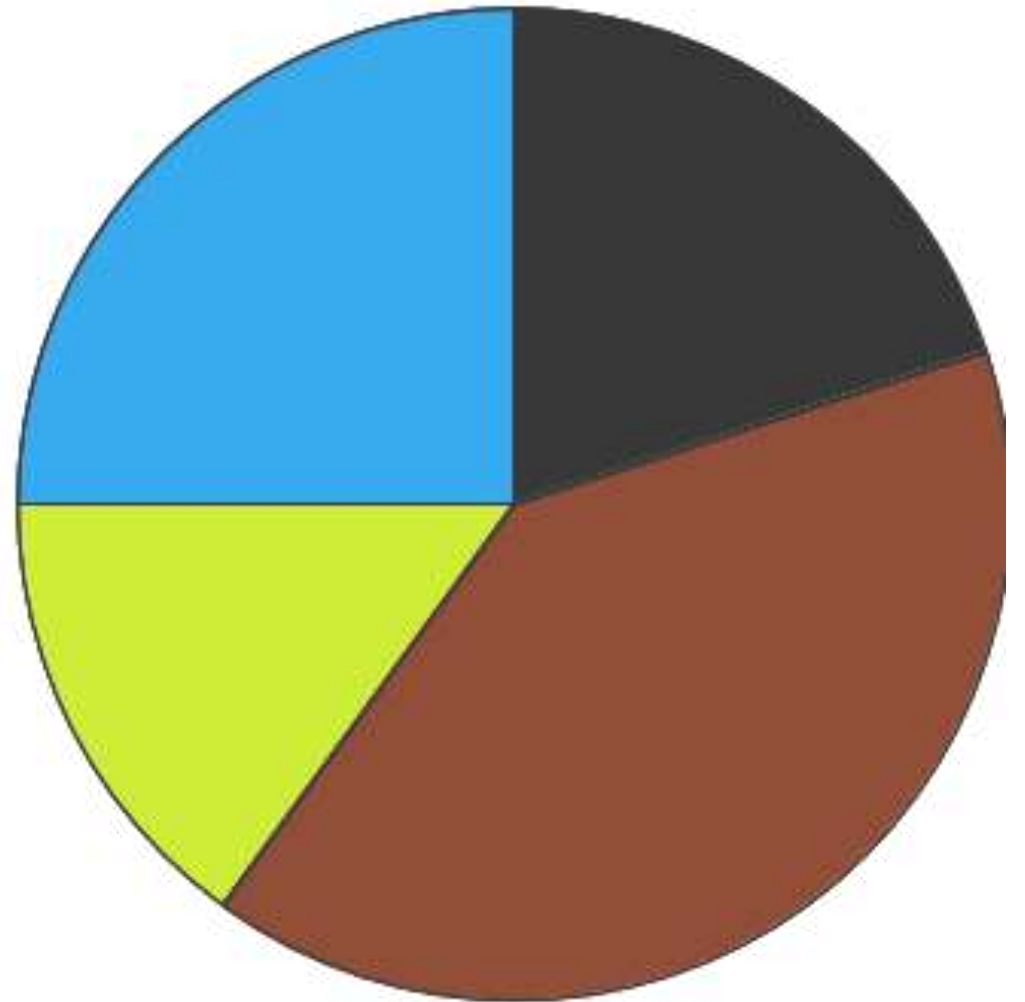


Grafico a torta

colore degli occhi “blu”
– settore blu

freq. rel del colore blu
= 0,25

frazione di area del
settore blu = 0,25
(25 % dell'area totale)



Esempio: gruppi sanguigni

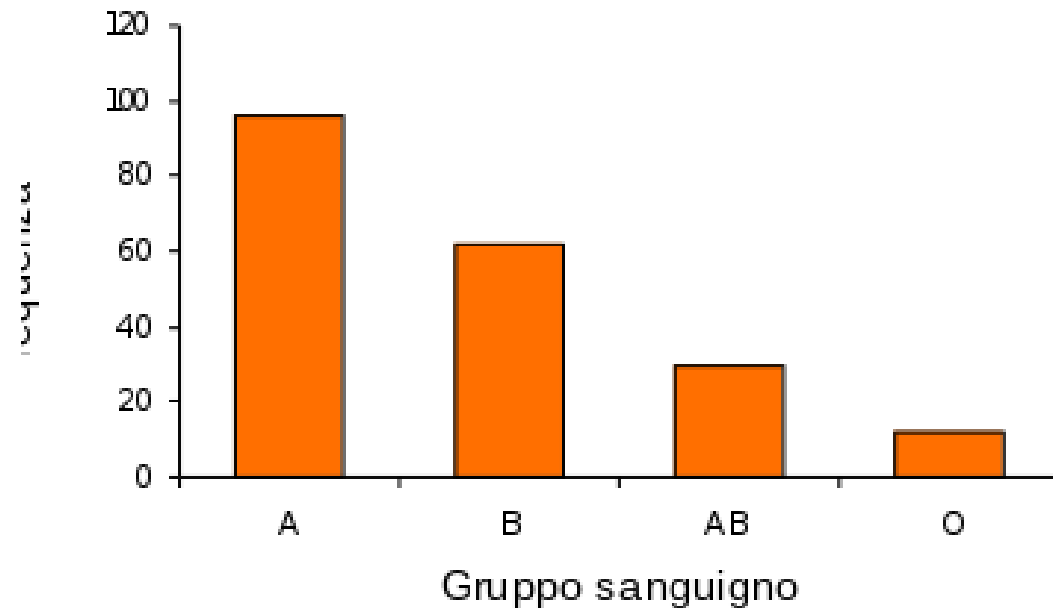
In un laboratorio sono state eseguite 200 analisi e sono stati osservati i gruppi sanguigni

Gruppo	Frequenza	Frequenza relativa	Freq. relativa %
O	96	0,48	48%
A	62	0,31	31%
B	30	0,15	15%
AB	12	0,06	6%

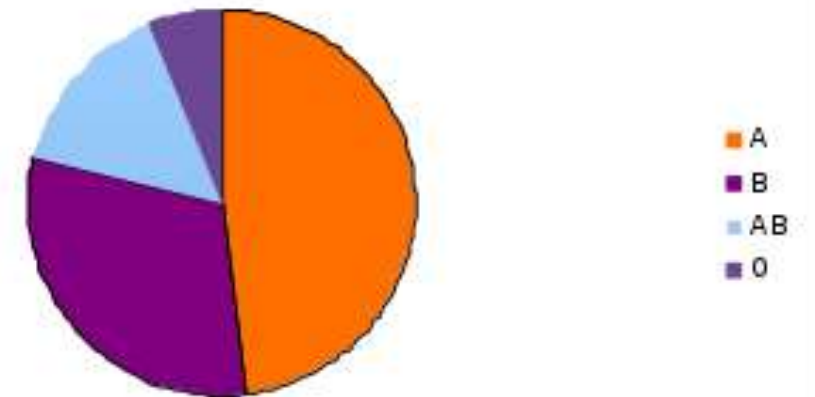


Grafici a barre e a torta

Gruppi sanguigni



Gruppi sanguigni



Raggruppamenti dei dati e istogrammi

Se il nostro insieme di dati ha troppi valori distinti

- suddividiamo i valori in gruppi o classi
- determiniamo il numero di valori dei dati che cadono in ciascuna classe.
- Rappresentiamo i dati con un istogramma

Istogramma: grafico a barre con le colonne adiacenti l'una all'altra



Esempio 1

Le lunghezze (cm) dei coyote femmine (CCS con $n = 40$)

$n = 40$

93.0 97.0 92.0 101.5 93.0 84.5 102.5 97.8 91.0 98.0 93.5 91.7
90.2 91.5 80.0 86.4 91.4 83.5 88.0 71.0 81.3 88.5 86.5 90.0
84.0 89.5 84.0 85.0 87.0 88.0 86.5 96.0 87.0 93.5 93.5 90.0
85.0 97.0 86.0 73.7

Valore più piccolo: 71.0 cm

Valore più grande: 102.5 cm

Campo di variazione (range): [71.0, 102.5]



Campione di 40 coyote femmine

divido il campo di variazione (range) delle osservazioni in classi di uguale ampiezza, in modo da includere tutti i dati

es. scelgo classi di ampiezza 5 cm

70.0-75.0

classe di **estremi** 70.0 e 75.0

$70.0 \text{ cm} \leq \text{lunghezze} < 75.0 \text{ cm}$

(convenzione di inclusione a sinistra; il valore 75.0 non appartiene alla classe 70.0-75.0)

7 classi:

70.0-75.0, 75.0-80.0, 80.0-85.0, 85.0-90.0, 90.0-95.0, 95.0-100.0, 100.0-105.0

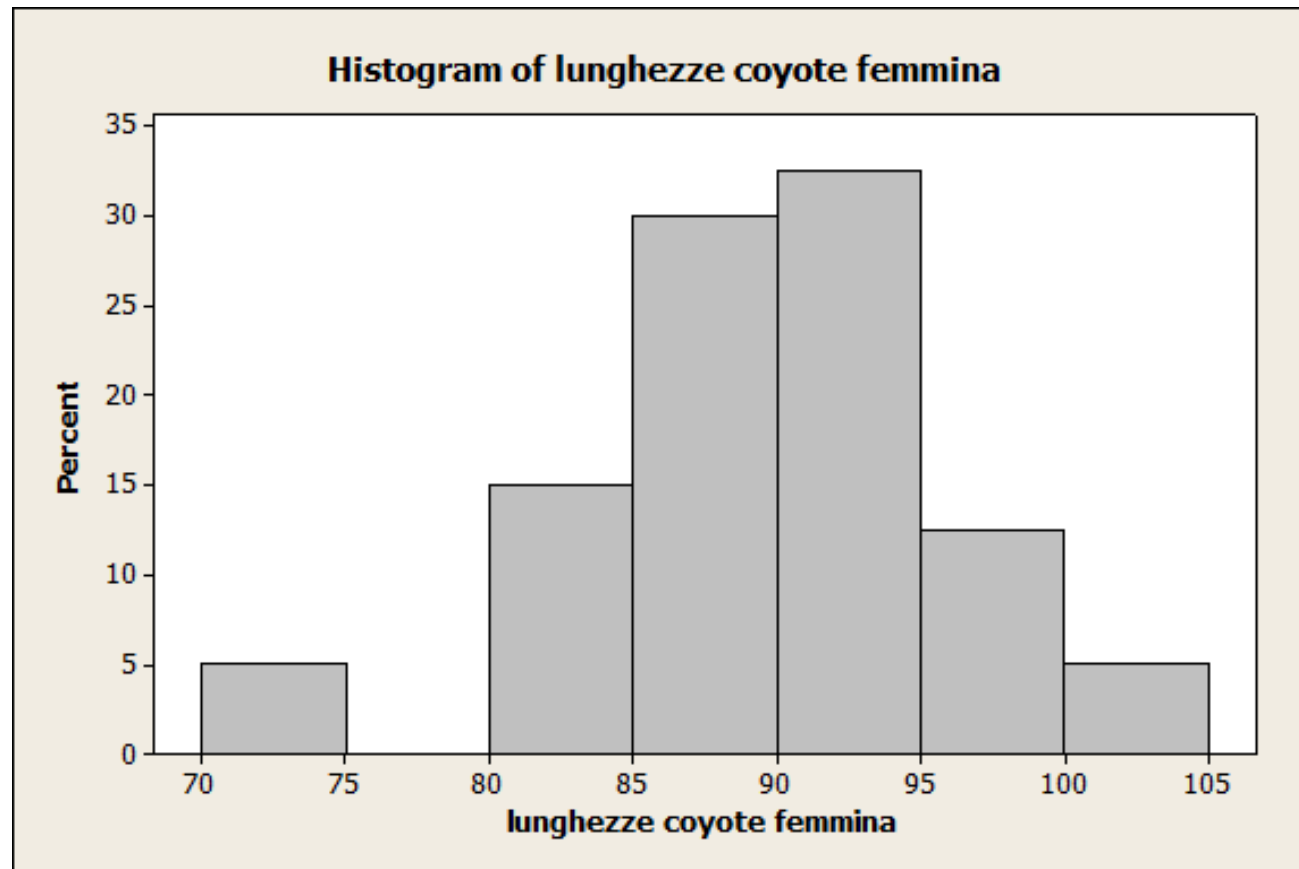


Tabella delle frequenze e delle frequenze relative

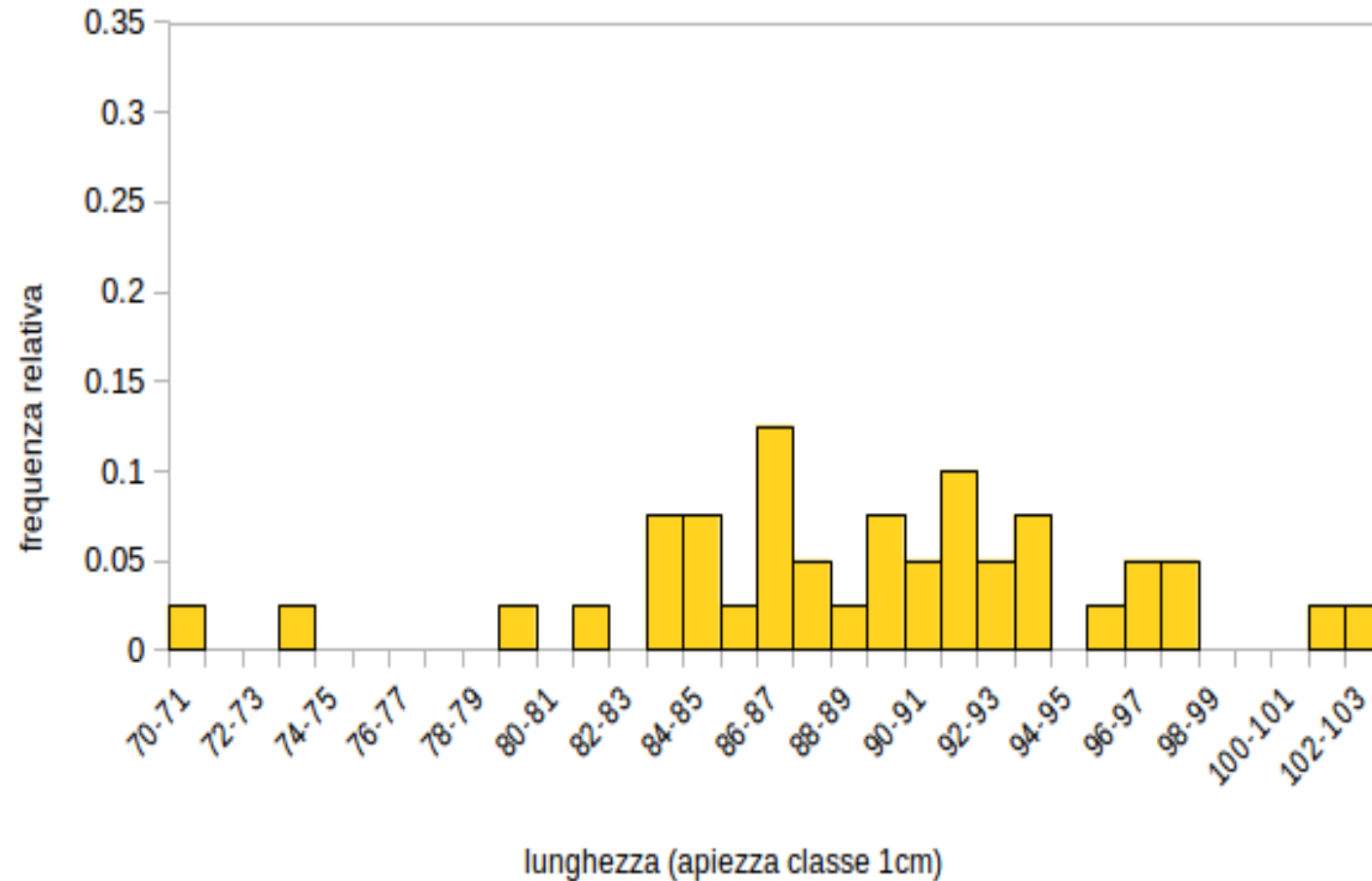
Classe	Frequenza	Freq. Rel.	Freq. Rel. %
70- 75	2	0.05	5
75- 80	0	0	0
80- 85	6	0.15	15
85- 90	12	0.3	30
90- 95	13	0.325	32.5
95-100	5	0.125	12.5
100-105	2	0.05	5
Totale	n=40	1.00	100



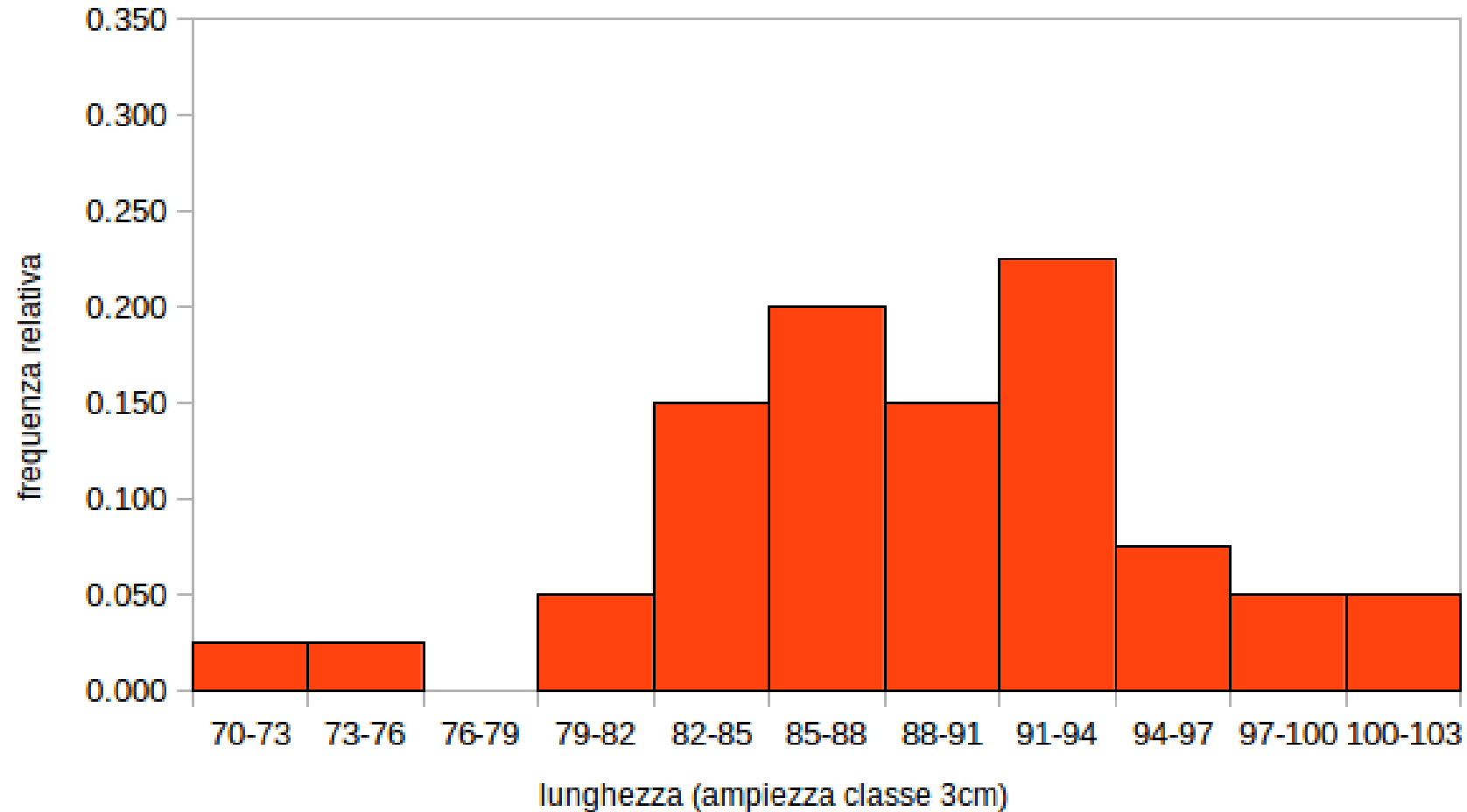
Istogramma delle lunghezze (cm) dei coyote femmina



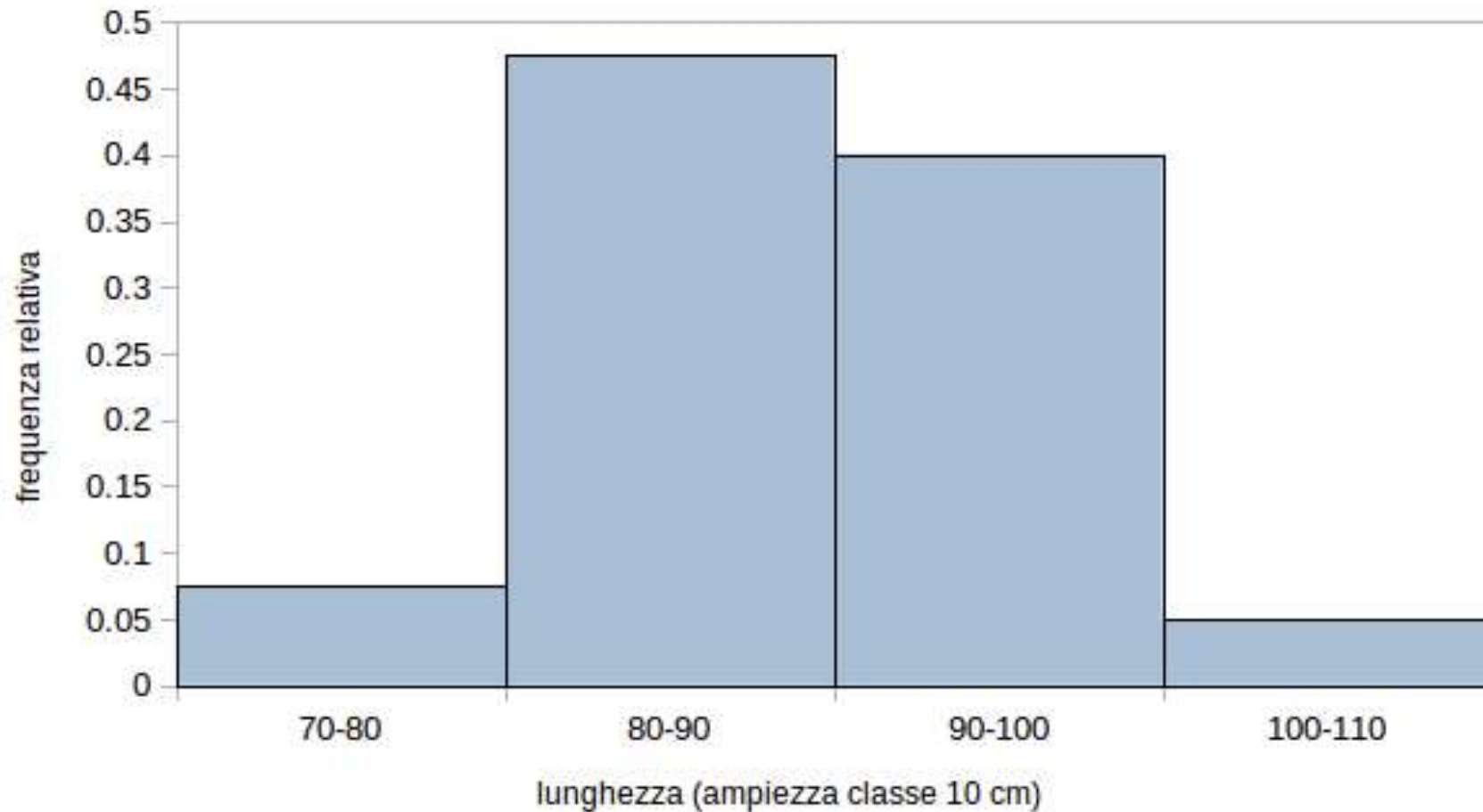
Istogramma – 33 classi di ampiezza 1cm



Istogramma – 11 classi di ampiezza 3cm



Istogramma- 4 classi di ampiezza 10 cm



Suggerimenti

Un istogramma appropriato dovrebbe avere tra le 7 e le 20 classi

- Numero di valori distinti ≤ 100 , numero di classi 7 -10
- Numero di valori distinti 101–200, numero di classi 10–15
- Numero di valori distinti >200 , numero di classi 14-20



Informazioni da un istogramma

- Grado di simmetria dei dati
- Grado di dispersione dei dati
- Eventuale presenza di intervalli con un'alta concentrazione
- Eventuale presenza di vuoti
- Eventuale distacco di alcuni dati



Distanza (mm) tra le orbite di 15 piccioni domestici

Lunghezze (mm)

10,2

10,4

10,4

10,7

10,8

11,3

11,3

11,5

11,6

11,8

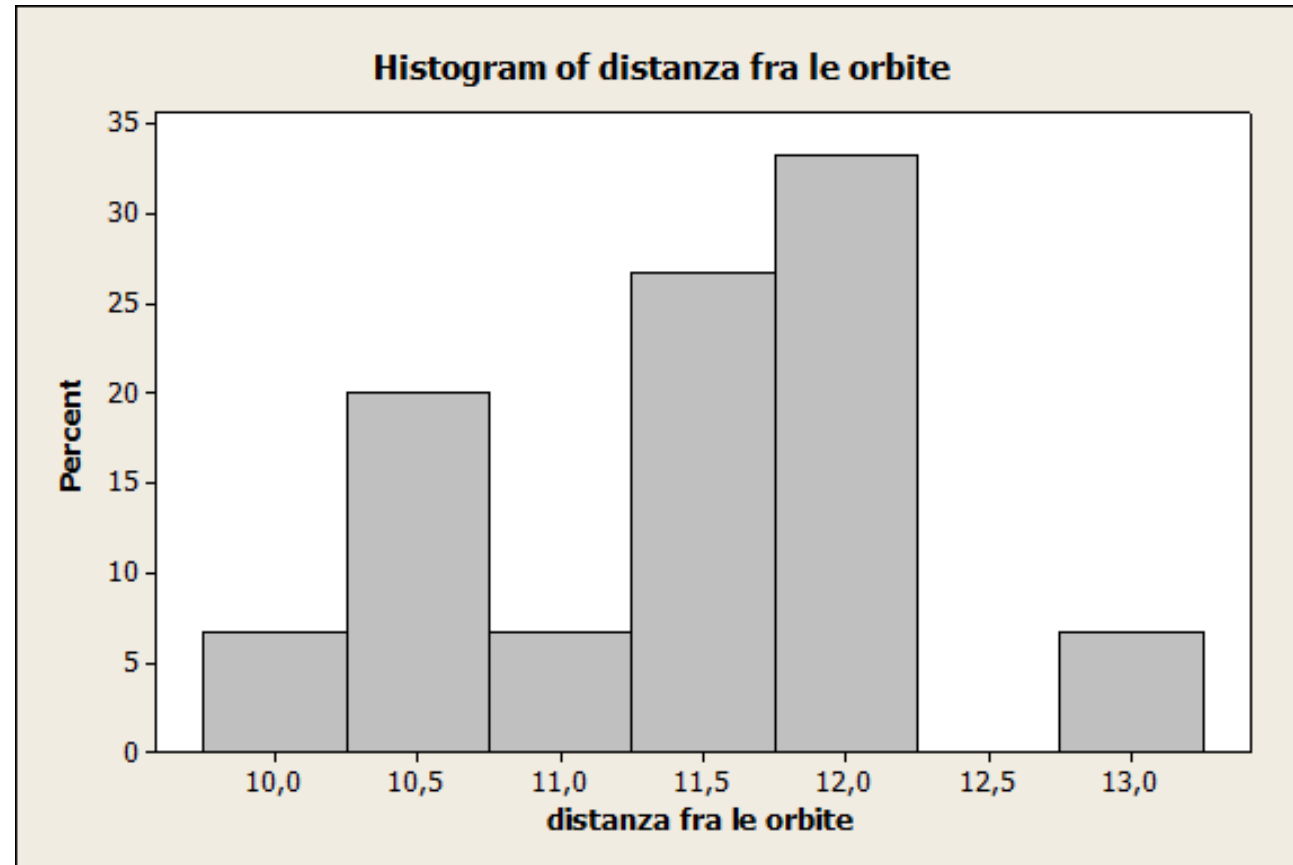
11,9

11,9

12,1

12,2

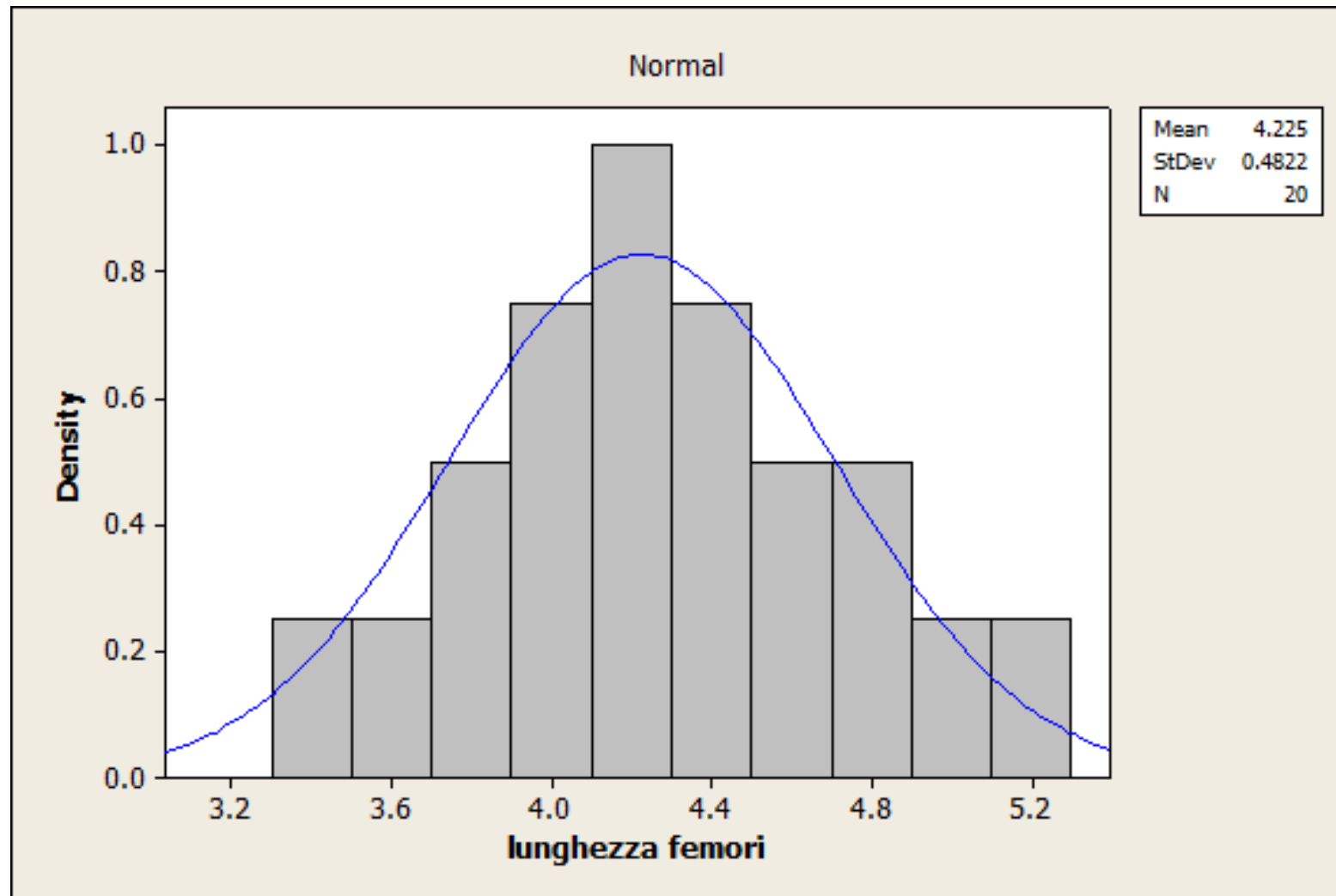
12,9



7 classi di ampiezza 0,5mm



Lunghezza (mm x 10⁻¹) di 20 femori dell'afide Pemphigus



Moda Campionaria

- La classe modale è la classe di frequenza massima
- La moda campionaria di un insieme di dati, se esiste, è l'unico valore che ha frequenza massima
- Una distribuzione di frequenze è detta unimodale se cresce fino a suo valore massimo (raggiunto nella moda) e poi decresce



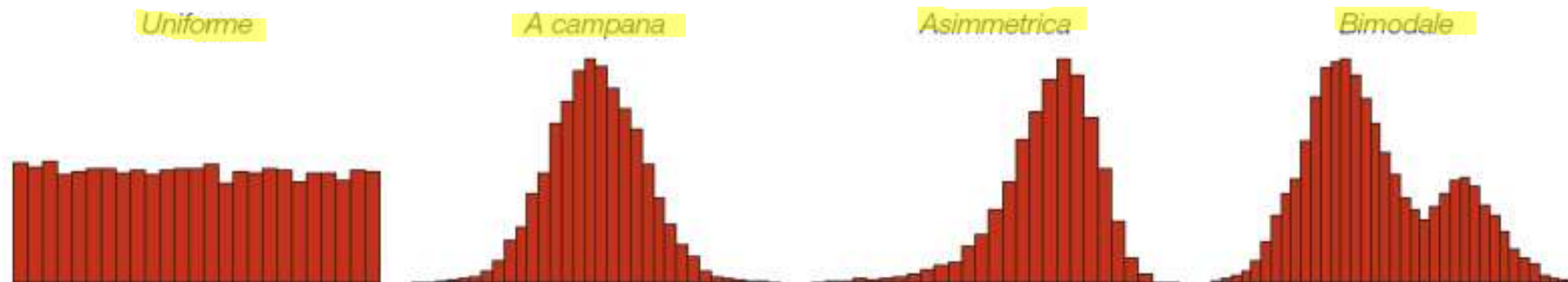


Figura 2.1-3
Alcune possibili forme di distribuzioni di frequenza.

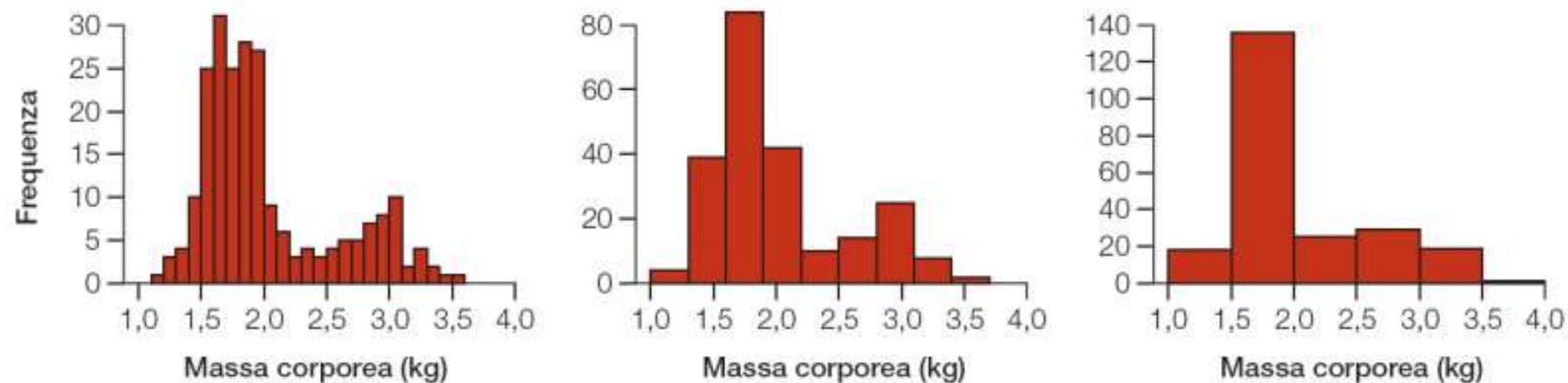


Figura 2.1-4
Massa corporea di 228 femmine di salmone rosso campionate dal Pick Creek in Alaska (Hendry *et al.*, 1999). In ciascun caso sono presentati gli stessi dati, ma le ampiezze degli intervalli sono diverse: 0,1 kg (a sinistra), 0,3 kg (in centro) e 0,5 kg (a destra).



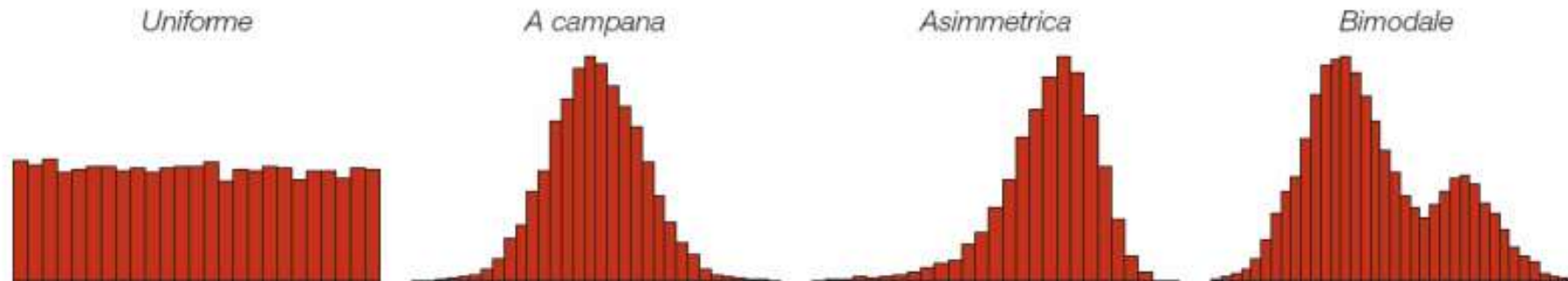


Figura 2.1-3
Alcune possibili forme di distribuzioni di frequenza.

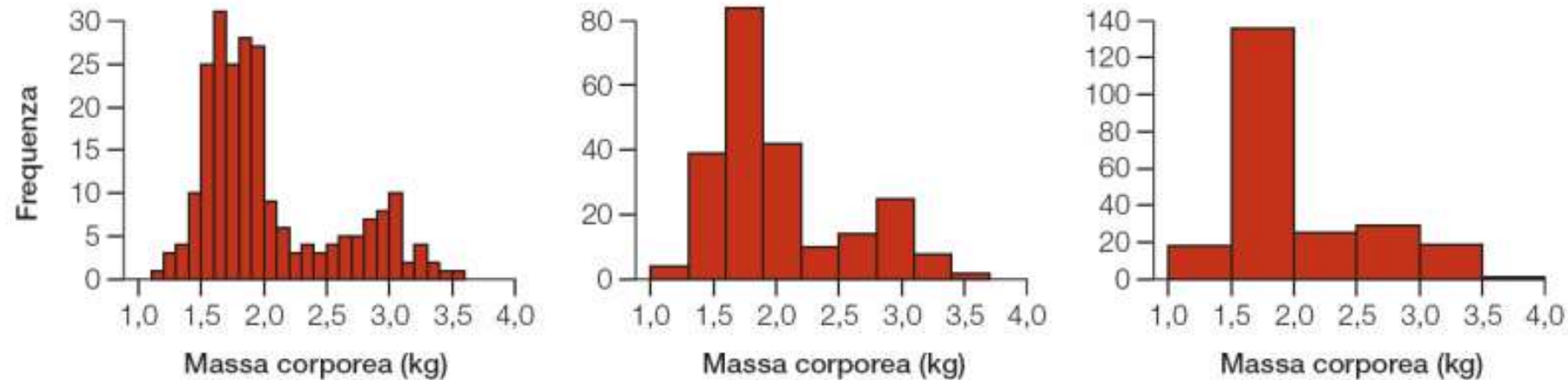


Figura 2.1-4
Massa corporea di 228 femmine di salmone rosso campionate dal Pick Creek in Alaska (Hendry *et al.*, 1999). In ciascun caso sono presentati gli stessi dati, ma le ampiezze degli intervalli sono diverse: 0,1 kg (a sinistra), 0,3 kg (in centro) e 0,5 kg (a destra).



Diagrammi ramo-foglia (stem and leaf)

Ogni valore dei dati è diviso in due parti, una più significativa (ramo) e una meno significativa (foglia)

Esempio: lunghezza (mm) dell'ala destra di 10 passeri

59 64 68 71 73 75 75 77 80 80

Ramo	Foglia
5	9
6	4 8
7	1 3 5 5 7
8	0 0



Esempio

Età di 30 pazienti ammessi al pronto soccorso di un certo ospedale

23, 18, 31, 79, 44, 51, 24, 19, 17 25, 27 ,14, 67, 35, 74, 20, 40, 55, 16, 21, 55, 45, 53,
27, 15, 30, 20, 65, 18, 59

ramo	foglia
1	4, 5, 6, 7, 8, 9
2	0, 0, 1, 3, 4, 5, 7, 7
3	0, 1, 5, 8
4	0, 4, 5
5	1, 3, 5, 5, 9
6	5, 7,
7	4, 9

