Solutions - Graphical Representations and Measures of Position and Variability

Sara Geremia

November 2, 2023

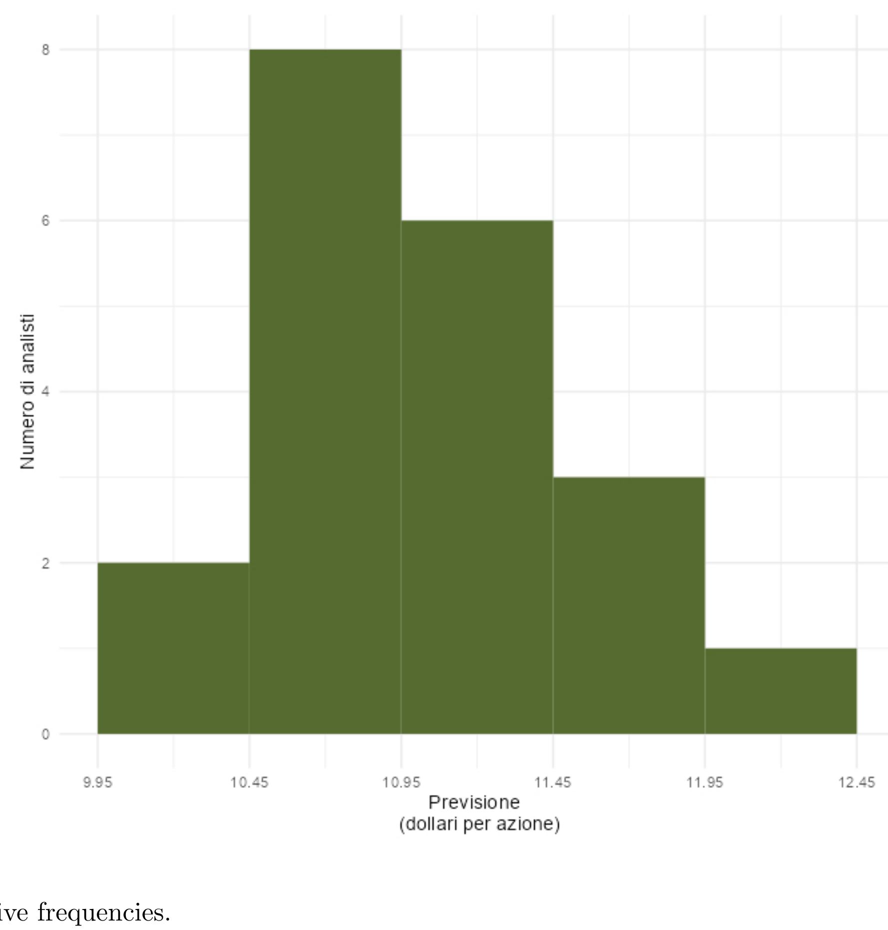
Chapter 2 Exercises: Graphical Representations

2.52

A sample of 20 financial analysts was asked to predict the earnings per share of a certain company (dollars per share) for the next year. The results are collected in the following table:

Forecast	**(9.95, 10.45]**	**(10.45, 10.95]**	**(10.95, 11.45]**	**(11.45, 11.95]**	**(11.95, 12.45]**
Number of analysts	2	8	6	3	1

Draw the histogram.



Find the relative frequencies.

Find the cumulative frequencies.

Find and interpret the cumulative percentage frequencies.

Forecast	**(9.95, 10.45]**	**(10.95, 10.95]**	**(10.95, 11.45]**	**(11.45, 11.95]**	**(11.95, 12.45]**
Number of analysts	2	8	6	3	1
Relative frequency	0.1	0.4	0.3	0.15	0.05
Cumulative frequency	0.1	0.5	0.80	0.95	1
Cumulative frequency (%)	10%	50%	80%	95%	100%

The cumulative percentage frequencies indicate the total percentage of data that lies at or before a certain value or interval in the distribution. These cumulative percentage values are useful for understanding how the data are distributed and for identifying significant cut-off points in the distribution. For example, 50% of analysts indicated that the earnings per share for the next year will be less than $10.95.

2.54

A teacher wants to examine the possible relationship between students' scores on the entrance test and their scores (on a broader value scale) in the final exam. The results, for a random sample of 8 students, are collected in the following table:

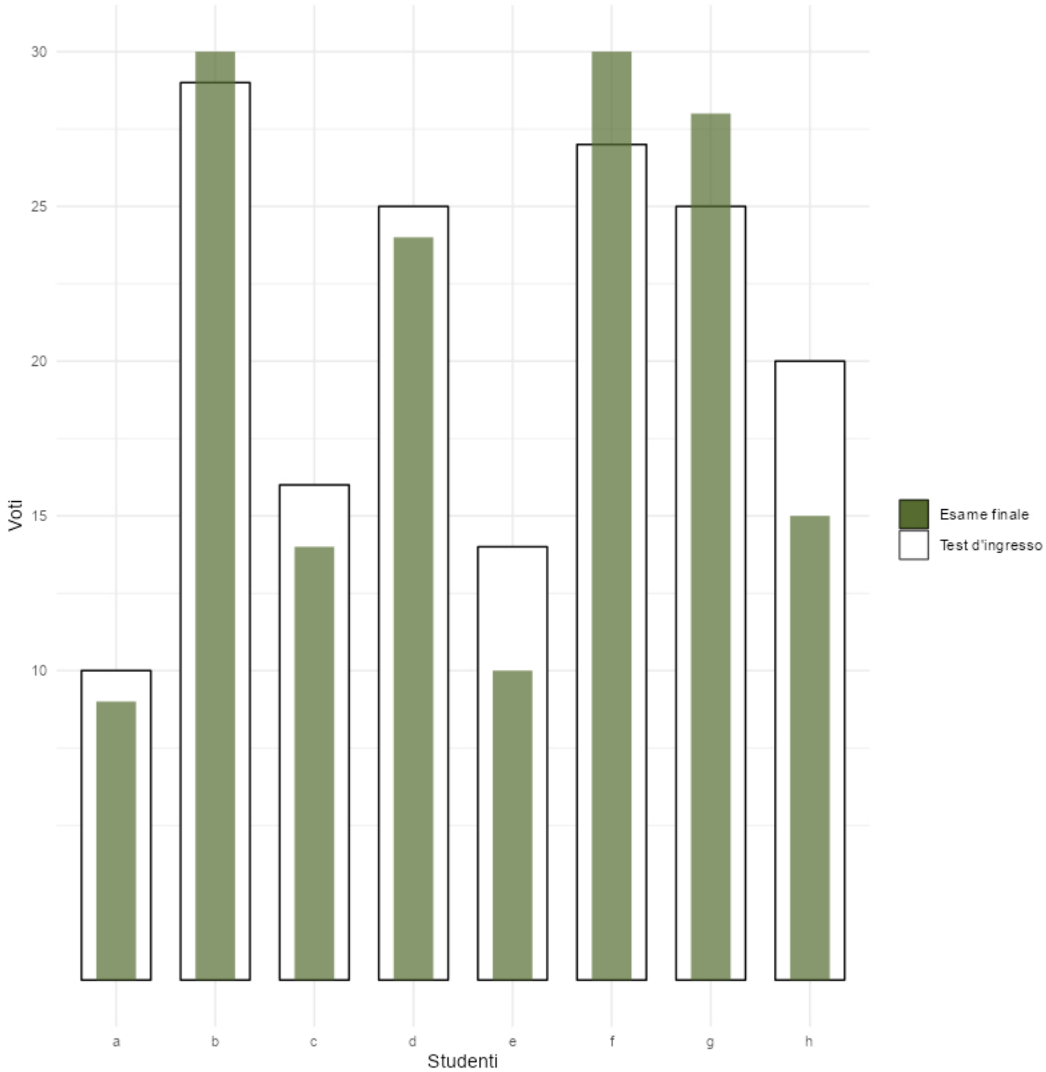
Entrance Test (X)	10	29	16	25	14	27	25	20
Final Exam (Y)	19	60	28	49	20	60	56	31

Represent the scores of the 8 students with overlapping bar charts.

Since the score scale for the final exam is broader, we assign the scores on a 0-30 scale to make them comparable with the entrance test scores.

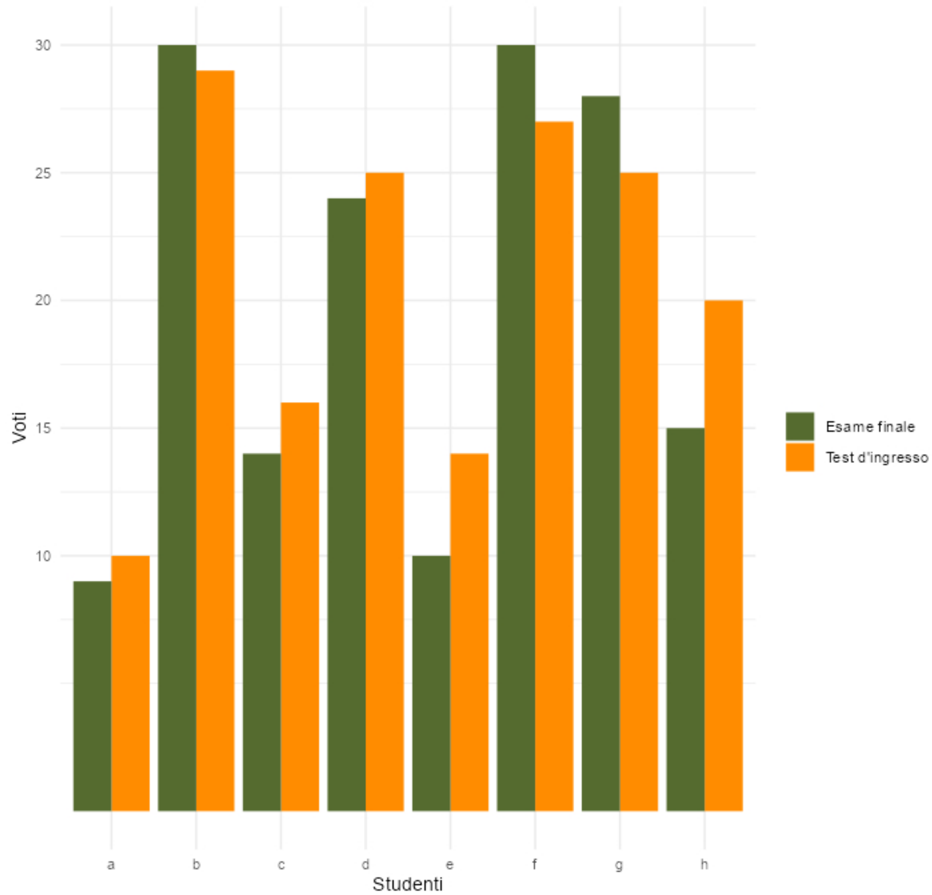
$19 : 60 = x : 30 \rightarrow x = \frac{19*30}{60} = 9.5$

Entrance Test (X)	10	29	16	25	14	27	25	20
Final Exam - 0-30 scale (Y)	9.5	30	14	24.5	10	30	28	15.5



30
25
20
15
10
Voti
a
b
c
d
e
f
g
h
Studenti
Esame finale
Test d'ingresso

Represent the scores of the 8 students with adjacent bar charts.



Find the minimum, first quartile, median, third quartile, and maximum of the distributions of variables X and Y.

The median is obtained from the ordered disaggregated distribution, identifying the central value among the 8 values of the distributions.

Entrance test (X)	10	14	16	20	25	25	27	29
Final exam (Y)	19	20	28	31	49	56	60	60
Final exam - 0-30 range (Y)	9.5	10	14	15.5	24.5	28	30	30

Since $n = 8$ is even (for both distributions), we know that the median falls between the 4^{th} and 5^{th} observation (in the ordered data) for all three distributions.

$$Me_X = \frac{20+25}{2} = 22.5$$
$$Me_Y = \frac{15.5+24.5}{2} = 20$$

The first quartile is always found in the ordered (disaggregated) distribution, but taking the 4 values to the left of the median (excluding the median itself). Therefore, it is located between the 2^{nd} and 3^{rd} observations in the first half of the data:

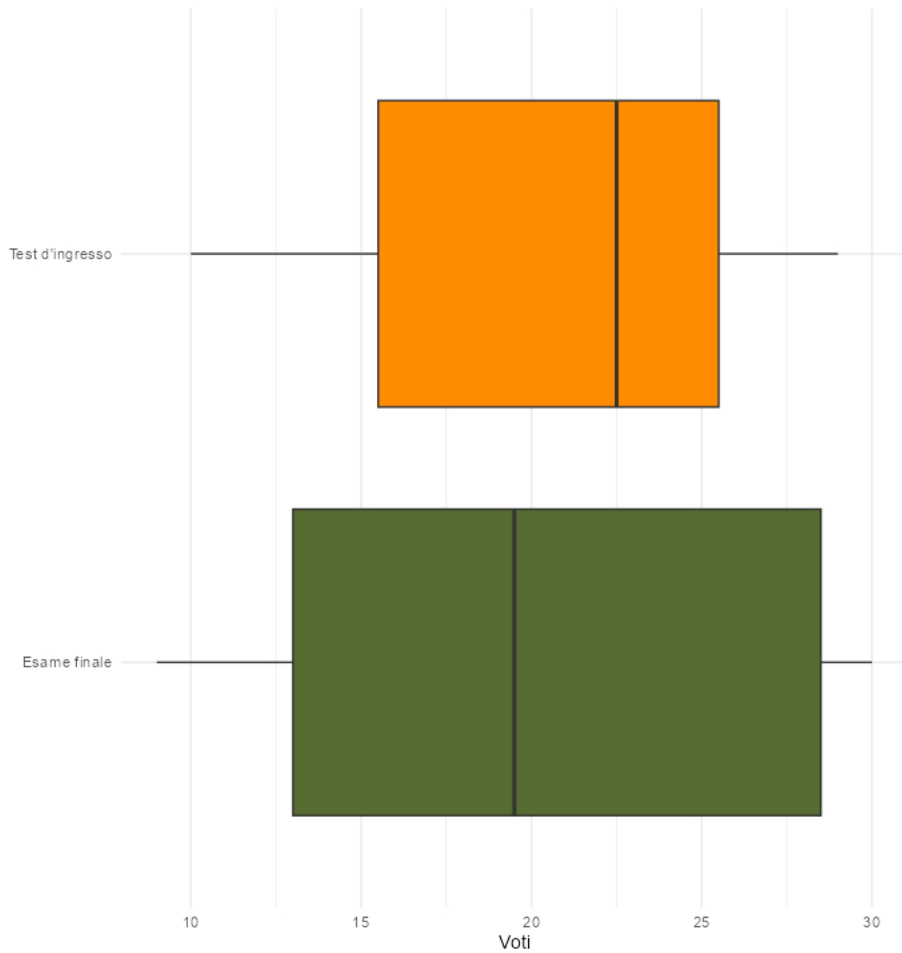
$$Q1_X = \frac{14+16}{2} = 15$$
$$Q1_Y = \frac{10+14}{2} = 12$$

The third quartile is always found in the ordered (disaggregated) distribution but taking the 4 values to the right of the median (excluding the median itself). Therefore, it is located between the 6^{th} and 7^{th} observations in the second half of the data:

$$Q3_X = \frac{25+27}{2} = 26$$
$$Q3_Y = \frac{28+30}{2} = 29$$

Quantiles	Min	I Quartile	Median	III Quartile	Max
Test d'ingresso (X)	10	15	22.5	26	29
Entrance test (Y)	19	24	40	58	60
Final exam - 0-30 range (Y)	9.5	12	20	29	30

Build the boxplot.



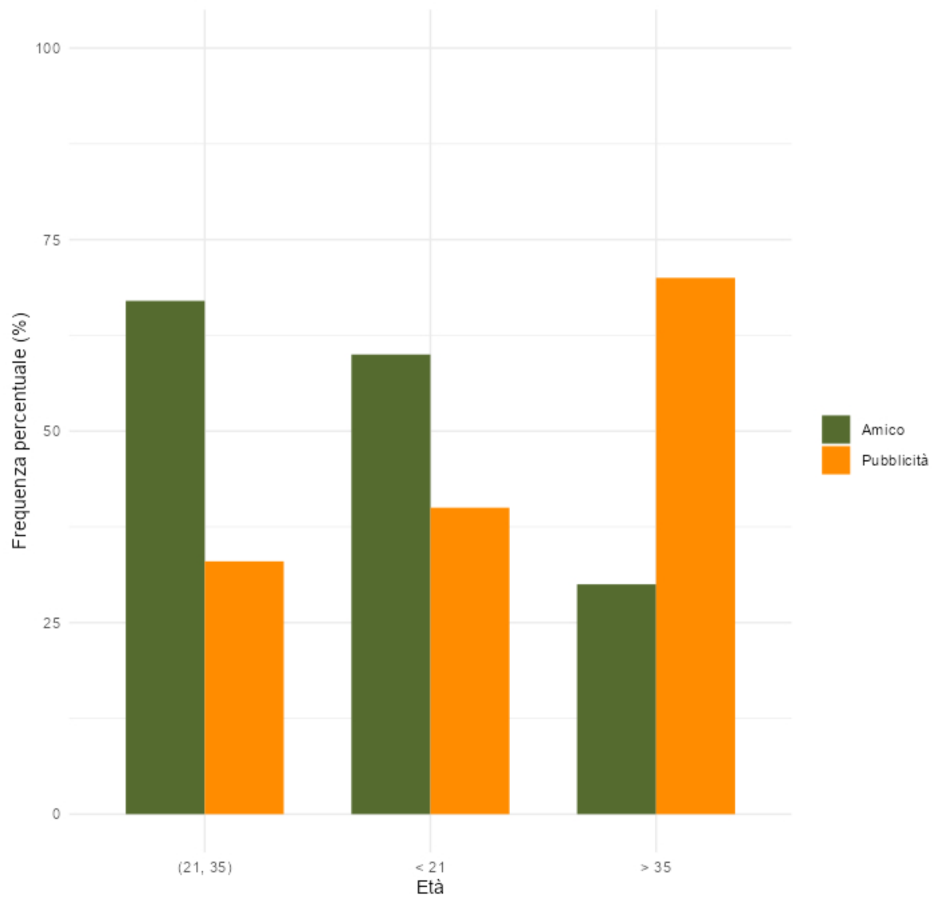
2.64

How do people become aware of a new product? A random sample of 200 customers from a certain store were asked, in addition to their age, whether they became aware of the new product from a friend or through advertising in local newspapers. The results indicated that 50 people were under 21 years old, 90 were between 21 and 35 years old, and 60 were over 35 years old. Among those under 21, 30 became aware of the new product from a friend and the rest through an advertisement in the local newspaper. One third of those in the age group from 21 to 35 had learned about the product from the same newspaper ad, while the rest from a friend. In the category of respondents over 35 years old, 30% had heard about the product from a friend and the rest from the advertisement in the local newspaper.

(a) Describe the data with a contingency table.

	Friend	**Advertisement**
< 21	30	20
(21, 35]	60	30
> 35	18	42

(b) Describe the data graphically.



The bar graph compares the percentage of customers who became aware of the product from a friend with the percentage of customers who became aware of the product through advertising in newspapers, by age group.

Chapter 3 Exercises: Measures of Location and Variation

3.46

A major airport recently commissioned an external consultant to study problems related to flight delays. The consultant collected the number of minutes of delay for a sample of flights, the results of which are collected in the following table:

Minutes of Delay	(0, 10]	(10, 20]	(20, 30]	(30, 40]	(40, 50]	(50, 60]
Number of Flights	30	25	13	6	5	4

Minutes of Delay	(0, 10]	(10, 20]	(20, 30]	(30, 40]	(40, 50]	(50, 60]	Total
Central Value (C)	5	15	25	35	45	55	
Number of Flights (P)	30	25	13	6	5	4	83
C*P	150	375	325	210	225	220	1505

(a) Evaluate the average number of minutes of delay.

$Mean = \frac{1505}{83} = 18.13$

(b) Calculate the sample variance and standard deviation.

$Variance = \frac{(5-18.13)^2*30+(15-18.13)^2*25+(25-18.13)^2*13+(35-18.13)^2*6+(45-18.13)^2*5+(55-18.13)^2*4}{83} = 202.2355$

$Standard\,Deviation = \sqrt{202.2355} = 14.22095$

3.53

Consider the following four populations.

- 1, 2, 3, 4, 5, 6, 7, 8
- 1, 1, 1, 1, 8, 8, 8, 8
- 1, 1, 4, 4, 5, 5, 8, 8
- -6, -3, 0, 3, 6, 9, 12, 15

They all have the same mean. Without performing calculations, arrange the populations based on the magnitude of their variances, from smallest to largest, and then actually calculate the variances.

- 1, 2, 3, 4, 5, 6, 7, 8
- 1, 1, 4, 4, 5, 5, 8, 8
- 1, 1, 1, 1, 8, 8, 8, 8
- -6, -3, 0, 3, 6, 9, 12, 15

Mean $= \frac{1+2+3+4+5+6+7+8}{8} = 4.5$

Variances:

- $\frac{(1-4.5)^2+(2-4.5)^2+(3-4.5)^2+(4-4.5)^2+(5-4.5)^2+(6-4.5)^2+(7-4.5)^2+(8-4.5)^2}{8} = 5.25$
- $\frac{(1-4.5)^2*2+(4-4.5)^2*2+5*2+(8-4.5)^2*2}{8} = 7.4375$
- $\frac{(1-4.5)^2*4+(8-4.5)^2*4}{8} = 12.25$
- $\frac{(-6-4.5)^2+(-3-4.5)^2+(0-4.5)^2+(3-4.5)^2+(6-4.5)^2+(9-4.5)^2+(12-4.5)^2+(15-4.5)^2}{8} = 47.25$