

CONTACTO

Professora: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt



<https://doity.com.br/estatistica-aplicada-a-nutricao>



<https://basiccode.com.br/produto/informatica-basica/>

PROGRAMA



1. Conceitos
Fundamentais de
Estatística



2. Análise
Exploratória de
Dados



3. Redução e
Sintetização de
Dados



4. Associação e
Relações entre
Variáveis



5. Números
Índices



6. Sucessões
Cronológicas

AULA PRÁTICA 5: EXERCÍCIOS 2.5 E 3.1

EXERCÍCIO 2.5

2.5 As notas de Laboratório de Estatística (LE) e de Álgebra Linear (AL) obtidas por 10 alunos constam do quadro que se segue:

LE	15	12	18	14	13	10	15	16	17	17
AL	14	15	17	13	15	10	14	15	18	17

- a)** Represente estes dados graficamente.
- b)** Analise a possibilidade de existência de relação linear entre as notas destas duas uc's.

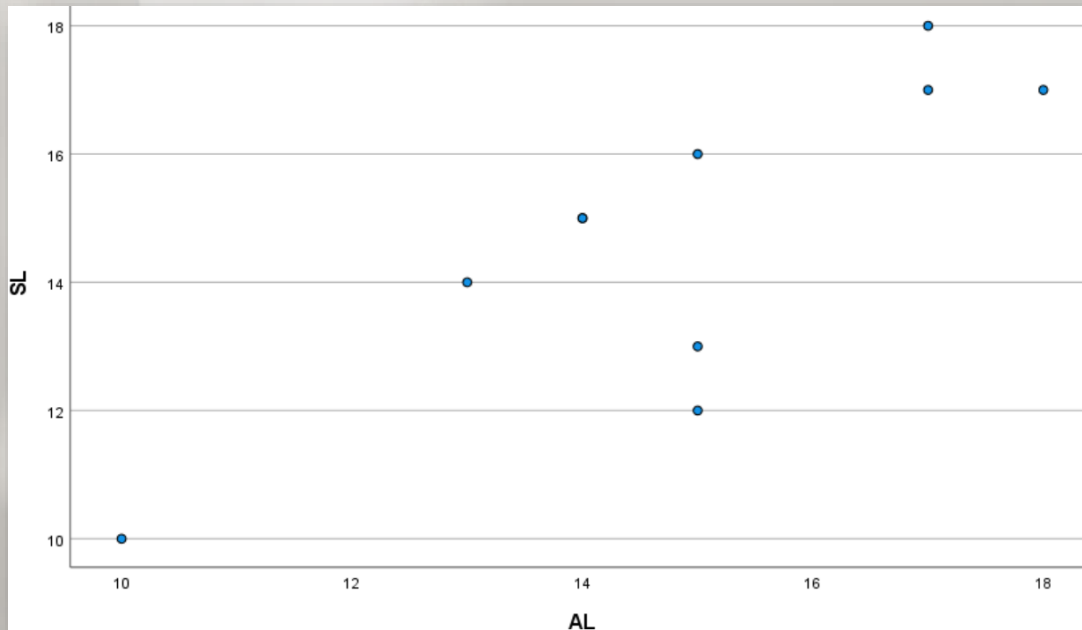


EXERCÍCIO 2.5 A): SOLUÇÃO



Answer:

Diagrama de Dispersão
ou Scatterplot



This scattetplot shows the general trend of the data and help visualize any linear relationship.

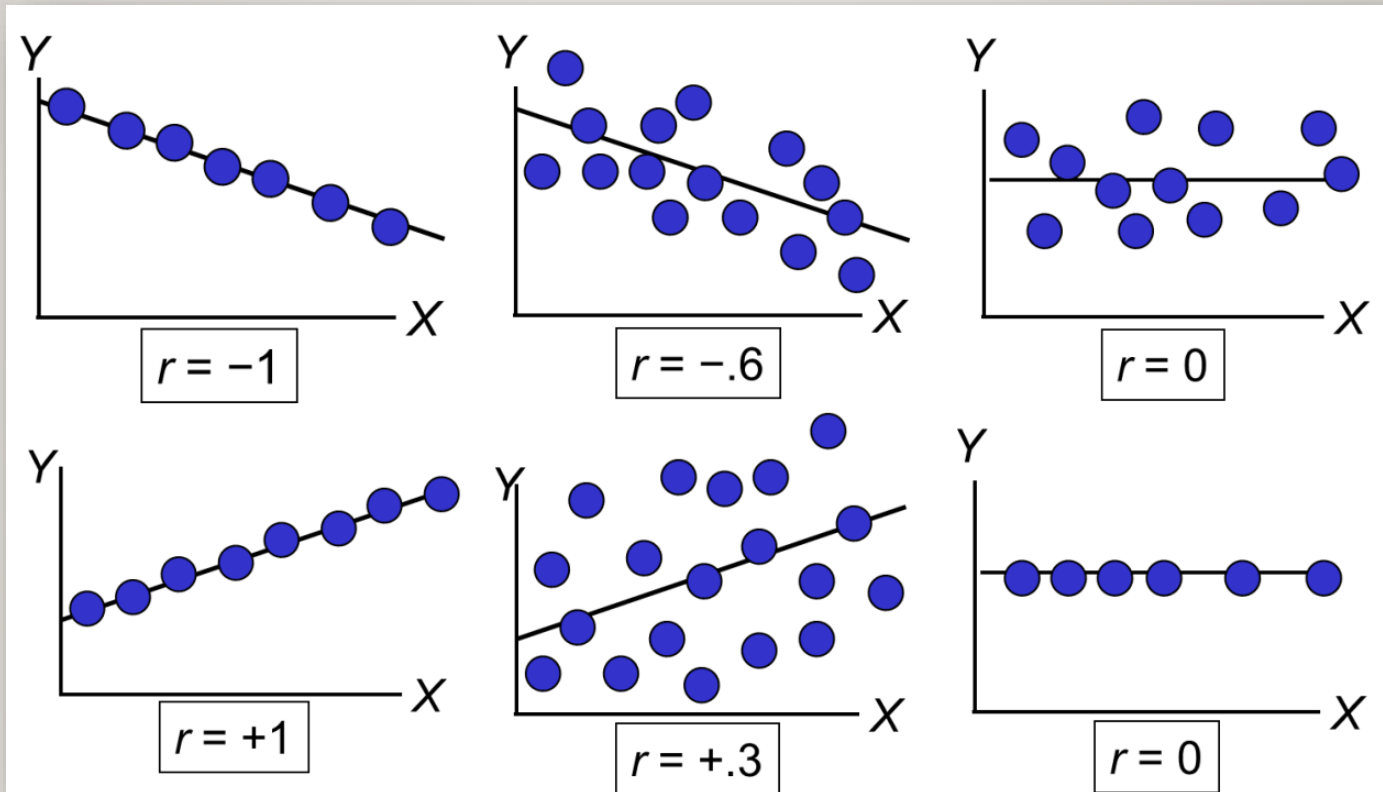
EXERCÍCIO 2.5 B): SOLUÇÃO



Answer:

b) O diagrama de dispersão construído na alínea anterior evidencia uma forte possibilidade de existência de uma relação linear positiva entre as notas pois as maiores notas de uma UC estão associadas tendencialmente maiores notas da outra UC. Também se pode verificar que a associação entre as notas é considerável porque os pontos encontram-se muito próximos de uma reta hipotética de declive positivo.

SCATTER PLOTS OF DATA WITH VARIOUS CORRELATION COEFFICIENTS



EXERCÍCIO 3.1

3.1 Os salários (em u.m.) de 40 trabalhadores fabris constam do quadro que se segue:

50	55	85	62
35	70	75	68
90	38	72	70
45	120	50	75
80	50	48	83
38	35	45	92
42	125	150	85
110	115	70	140
100	95	75	145
60	90	155	90

- a) Agrupe os dados em 4 classes de amplitude constante com o limite inferior da primeira classe igual a 35 u.m. e obtenha o quadro de frequências.
 - b) Represente graficamente os dados.
-
- c) Apresente a função cumulativa de frequências para os limites das classes e represente o correspondente polígono integral.
 - d) Calcule as medidas de tendência central para estes dados.



EXERCÍCIO 3.1 A): SOLUÇÃO



Answers:

a) $n = 40$, $m = 4$ (classes)

Amplitude de classe : $h = \frac{\max(x_i) - \min(x_i)}{m} = \frac{155 - 35}{4} = 30$

Wages	x'_j	n_j	f_j	$N(x)$	F^*_j
[35, 65]	50	14	0.35	14	0.35
[65, 95]	80	17	0.425	31	0.775
[95, 125]	110	5	0.125	36	0.9
[125, 155]	140	4	0.1	40	1

EXERCÍCIO 3.1 B): SOLUÇÃO

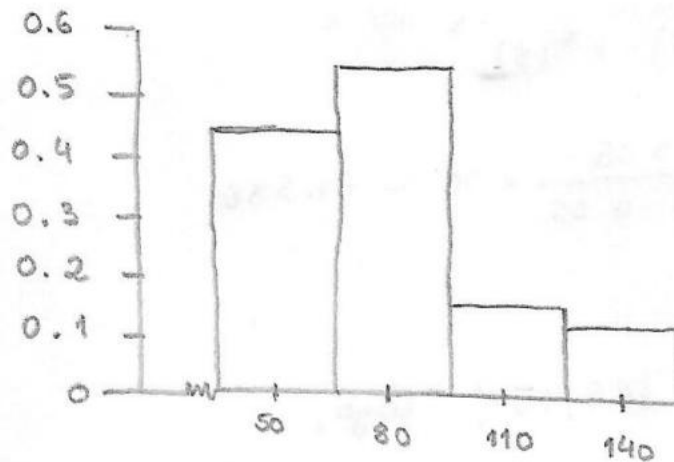


Answer:

b)

Histograma

• Altura das barras: $\left(\frac{14}{30} = 0.4(6), \frac{17}{30} = 0.5(6), \frac{5}{30} = 0.1(6), \frac{4}{30} = 0.1(3) \right)$



EXERCÍCIO 3.1 C): SOLUÇÃO



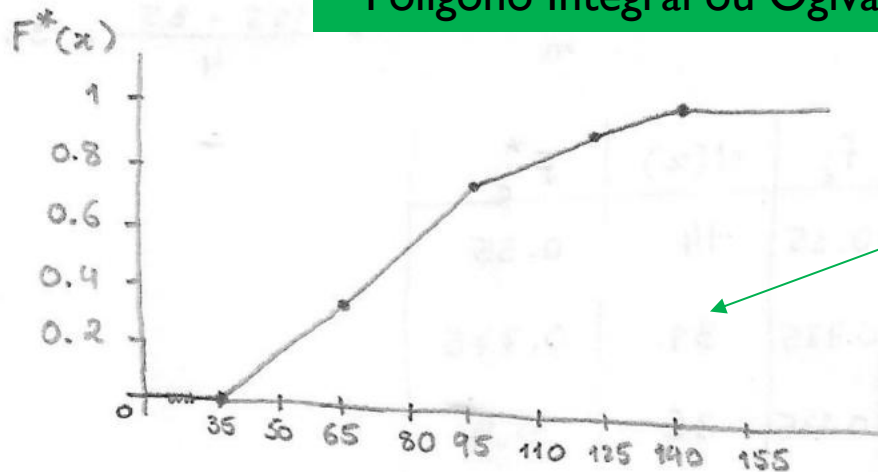
Answer:

Frequências Relativas Acumuladas

Wages	x_j	n_j	f_j	$N(x)$	F_j^*
[35, 65]	50	14	0.35	14	0.35
]65, 95]	80	17	0.425	31	0.775
]95, 125]	110	5	0.125	36	0.9
]125, 155]	140	4	0.1	40	1

c) $F^*(35) = 0$, $F^*(65) = 0.35$, $F^*(95) = 0.35 + 0.425 = 0.775$,
 $F^*(125) = 0.35 + 0.425 + 0.125 = 0.9$, $F^*(155) = 0.35 + \dots + 0.125 + 0.1 = 1$

Polígono Integral ou Ogiva



Interval Endpoint	$F^*(x)$
35	0
65	0,35
95	0,775
125	0,9
155	1

EXERCÍCIO 3.1 D): SOLUÇÃO



Answer:

$$\bar{x} = \frac{\sum_{j=1}^m n_j \cdot x'_j}{n} = \sum_{j=1}^m f_j \cdot x'_j$$

$$Me = l + \frac{0.5 - F^*(l)}{F^*(L) - F^*(l)} \cdot h$$

d) Média do salário: $\bar{x} \approx \frac{1}{n} \sum_{j=1}^m x'_j \cdot n_j = \frac{1}{40} (50 \times 14 + 80 \times 17 + 110 \times 5 + 140 \times 4) = 79.25$

Mediana do salário: a classe mediana é $[65, 95]$. Logo,

$$m_e \approx 65 + \frac{0.5 - F^*(65)}{F^*(95) - F^*(65)} \times 30 = 65 + \frac{0.5 - 0.35}{0.775 - 0.35} \times 30 \approx 75.588$$

Salário (x)	x'_j	n_j	f_j	$N(x)$	F^*_j
[35, 65]	50	14	0.35	14	0.35
[65, 95]	80	17	0.425	31	0.775
[95, 125]	110	5	0.125	36	0.9
[125, 155]	140	4	0.1	40	1

EXERCÍCIO 3.1 D): SOLUÇÃO



Answer:

$$Mo = l + \frac{f^{**}}{f^{*} + f^{**}} \cdot h$$

Moda do salário: a classe modal é $]65, 95]$. Logo,

$$m_o \approx 65 + \frac{0.125}{0.35 + 0.125} \times 30 \approx 72.895 \quad (\text{fórmula de King})$$

$$m_o \approx 65 + \frac{0.425 - 0.35}{(0.425 - 0.35) + (0.425 - 0.125)} \times 30 = 71$$

Salário (x)	x_j^*	n_j	f_j	$N(x)$	F_j^*
[35, 65]	50	14	0.35	14	0.35
]65, 95]	80	17	0.425	31	0.775
]95, 125]	110	5	0.125	36	0.9
]125, 155]	140	4	0.1	40	1

MÉDIA PARA DADOS CLASSIFICADOS

Definition:

The **mean** for grouped data is the weighted average of the class midpoints, weighted by the frequencies.

Formula (Grouped Data – Silvestre, 2007):

$$\bar{x} = \frac{\sum_{j=1}^m n_j \cdot x'_j}{n} = \sum_{j=1}^m f_j \cdot x'_j$$

Legend / Notation:

- $j = 1, 2, \dots, m \rightarrow$ class index
- $n_j =$ absolute frequency of class j
- $f_j =$ relative frequency of class j
- $x'_j =$ midpoint of class j
- $m =$ number of classes
- $n = \sum_{j=1}^m n_j \rightarrow$ total number of observations

Silvestre (2007)

MEDIANA PARA DADOS CLASSIFICADOS

Definition:

The **median class** is the class interval that contains the **middle value** (0.5 in cumulative relative frequency) of the distribution.

Formula (Grouped Data)

$$Me = l + \frac{0.5 - F^*(l)}{F^*(L) - F^*(l)} \cdot h$$

Legend / Notation:

- l = lower boundary of the median class
- L = upper boundary of the median class
- $h = L - l$ = class width
- $F^*(l)$ = cumulative relative frequency **before the median class**
- $F^*(L)$ = cumulative relative frequency **at the upper boundary of the median class**

Silvestre (2007)

MODA PARA DADOS CLASSIFICADOS

Definition:

The **modal class** is the class interval with the **highest frequency** in a grouped frequency distribution

King's Formula (Grouped Data):

$$Mo = l + \frac{f^{**}}{f^{*} + f^{**}} \cdot h$$

Legend / Notation:

- l = lower boundary of the modal class
- h = class width
- f^{*} = relative frequency of the class before the modal class
- f^{**} = relative frequency of the class after the modal class

Silvestre (2007)

OBRIGADA!

Questões?

