
LABORATÓRIO DE ESTATÍSTICA



Matemática Aplicada à Economia e à Gestão
1º Ano / 1º Semestre
2026/2027

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 8: EXERCÍCIOS 3.1 (EXCEL) E 4.1 (EXCEL)

EXERCÍCIO 3.1 (JÁ FEITO NA AULA 5, E AGORA FAREMOS NO EXCEL)

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

- 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.
 - Represente graficamente os dados.
-
- Apresente a função cumulativa de frequências para os limites das classes e represente o correspondente polígono integral.
 - Calcule as medidas de tendência central para estes dados.



EXERCÍCIO 3.1 A): SOLUÇÃO



Answers:

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

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

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

f_x	=MÍNIMO(B1:B41)		
	D	E	F
a)			
	Min	Max	Amplitude de Classe
	35	155	30

f_x	=MÁXIMO(B2:B41)		
	D	E	F
a)			
	Min	Max	Amplitude de Classe
	35	155	30

f_x	=(E3-D3)/4		
	D	E	F
a)			
	Min	Max	Amplitude de Classe
	35	155	30

EXERCÍCIO 3.1 A): SOLUÇÃO

f_x $=\text{CONTAR.SE.S}(\$B\$1:\$B\$41;">=35";\$B\$1:\$B\$41;"<=65")$

D	E	F	G	H
Tabela de Frequências				
Classes	Frequência Absoluta	Frequência Relativa	Frequência Absoluta Acumulada	Frequência Relativa Acumulada
[35,65]	14	0,35	14	0,35
(65, 95]	17	0,425	31	0,775
(95, 125]	5	0,125	36	0,9
(125, 155]	4	0,1	40	1
Total	40	1		



Ar

a) $n = 40$, m

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

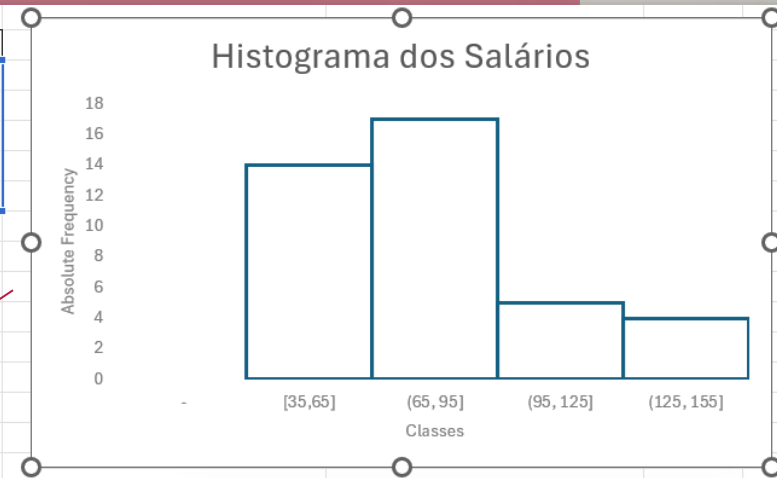
Salário (x)	x'_j	n_j	f_j	$N(x)$	F^*_j
[35, 65]	50	14	0.35	14	0.35
]65, 95]	80	17	0.425	31	0.775
]95, 125]	110	5	0.125	36	0.9
]125, 155]	140	4	0.1	40	1

EXERCÍCIO 3.1 B): SOLUÇÃO



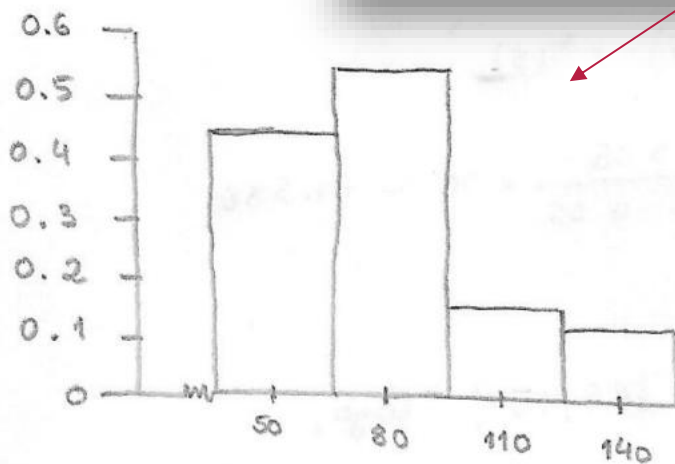
Ar

b)		
Histograma		
	Classes	Frequência Absoluta
	-	0
	[35,65]	14
	(65, 95]	17
	(95, 125]	5
	(125, 155]	4



b) Histograma

- Altura das ba



Formatar Série de Dados

Opções de Série



Opções de Série

Desenhar Série No

☒ Eixo Principal

☐ Eixo Secundário

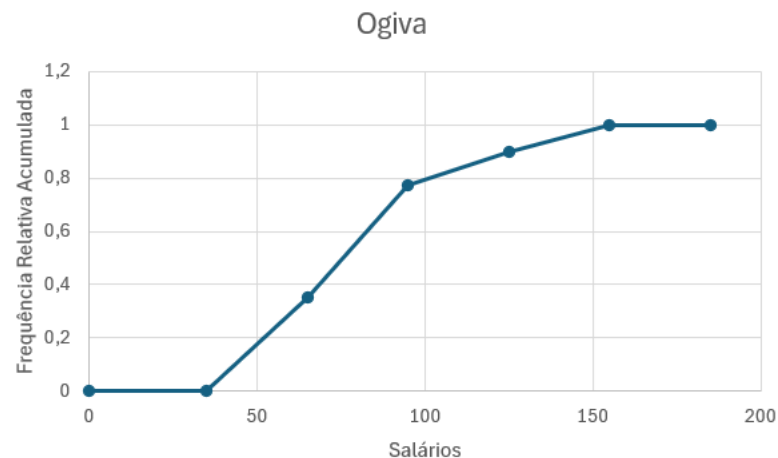
Sobreposição de Séries ☐ -35%

Largura do Intervalo ☐ 0%

EXERCÍCIO 3.1 C): SOLUÇÃO

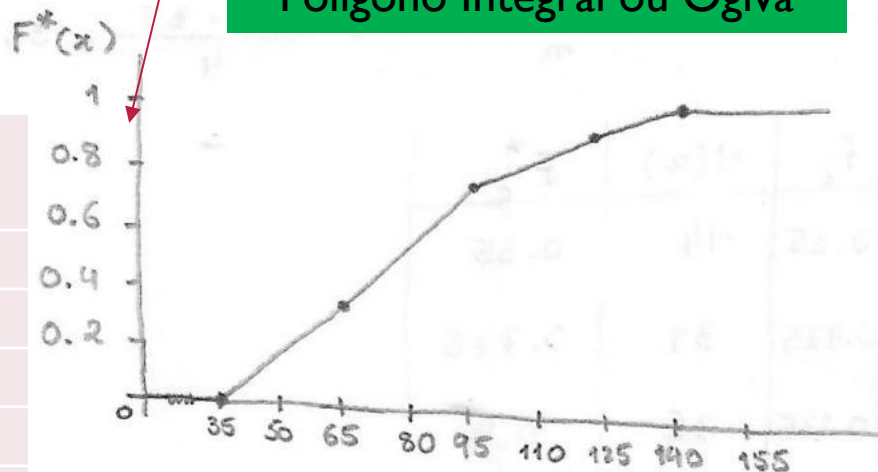
c)
Ogiva

Salários	Frequência Relativa Acumulada
0	0
35	0
65	0,35
95	0,775
125	0,9
155	1
185	1



Polígono Integral ou Ogiva

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



Formatar Série de Dados

Opções de Série



Linha Marcador

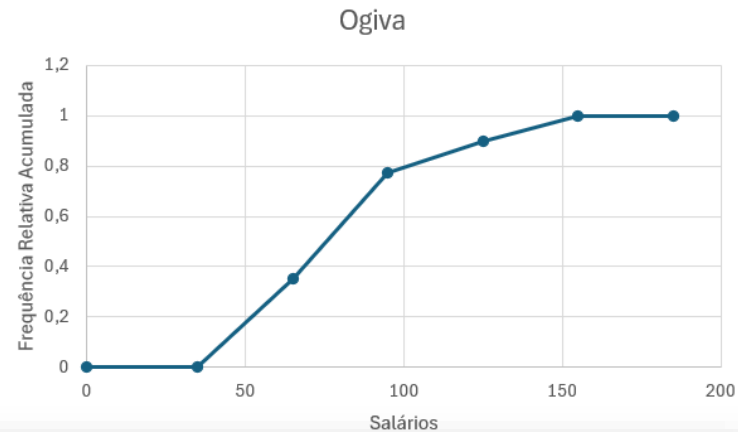
Linha

- ☐ Sem linha
- ☐ Linha contínua
- ☐ Linha de gradação
- ☒ Automático

EXERCÍCIO 3.1 C): SOLUÇÃO

c)
Ogiva

Salários	Frequência Relativa Acumulada
0	0
35	0
65	0,35
95	0,775
125	0,9
155	1
185	1



Interpretação da ogiva

- Até 65: 35% dos indivíduos têm salários inferiores ou iguais a 65.
- Até 95: 77,5% dos indivíduos têm salários inferiores ou iguais a 95.
- Até 125: 90% dos indivíduos têm salários inferiores ou iguais a 125.
- Até 155: 100% dos indivíduos têm salários inferiores ou iguais a 155.

O que podemos concluir?

1. 77,5% dos indivíduos recebem um salário igual ou inferior a 95.
2. 90% dos indivíduos recebem um salário igual ou inferior a 125.
3. A totalidade dos indivíduos tem salários iguais ou inferiores a 155.
4. O maior aumento da frequência relativa acumulada ocorre entre os salários de 65 e 95, passando de 35% para 77,5%. Isto indica que existe uma concentração significativa de indivíduos nesse intervalo salarial.

EXERCÍCIO 3.1 D): SOLUÇÃO



Answer:

=MÉDIA(B2:B41)			
C	D	E	F
	d)	Média	79,45
		Mediana	75
		Moda	50

=MED(B2:B41)			
D	E	F	
		Média	79,45
		Mediana	75
		Moda	50

=MODA(B2:B41)		
D	E	F
	Média	79,45
	Mediana	75
	Moda	50

EXERCISE 4.1

4.1 Os gastos em publicidade e as vendas de uma determinada empresa (em u.m.) constam do quadro que se segue:

Ano	1	2	3	4	5	6	7	8	9	10
Vendas	120	150	160	200	210	250	300	360	450	550
Gastos	10	12	14	20	22	25	30	38	50	60

- a)** Analise a relação entre as vendas e os gastos em publicidade através de uma representação gráfica que considere adequada.
- b)** Estime a reta de regressão dos gastos em publicidade em função das vendas. Interprete o resultado obtido.
- c)** Avalie a qualidade do modelo estimado a partir do cálculo de uma medida adequada para o efeito.



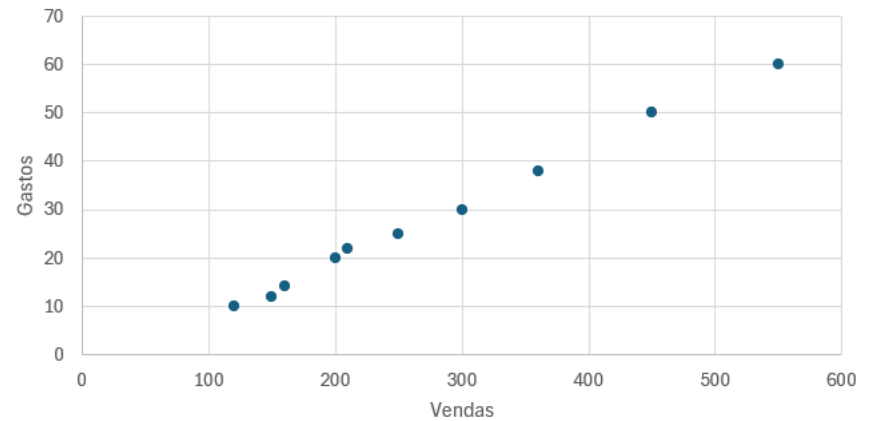
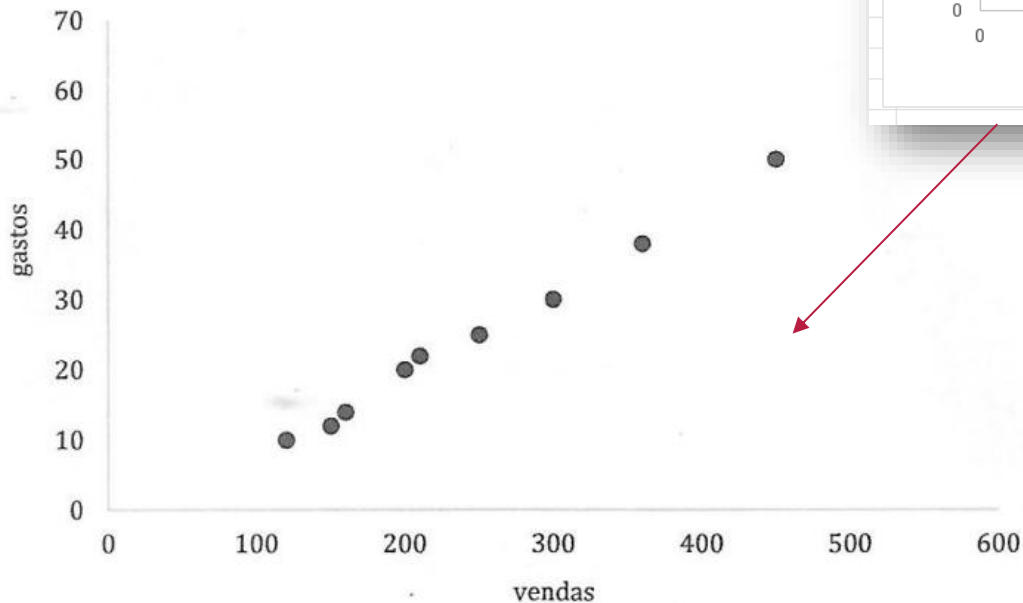
EXERCÍCIO 4.1 A): SOLUÇÃO



Resposta:

a)

Scatterplot/Diagrama de Dispersão



Interpretação:

O diagrama de dispersão mostra uma **relação positiva forte entre as vendas e as despesas com publicidade**, uma vez que os pontos se encontram muito próximos de uma linha reta com declive positivo.

EXERCÍCIO 4.1 B): SOLUÇÃO



Resposta:

X – Vendas
Y – Gastos

A	B	C	D
T	Anos	Vendas (X)	Gastos (Y)
1	1990	120	10
2	1991	150	12
3	1992	160	14
4	1993	200	20
5	1994	210	22
6	1995	250	25
7	1996	300	30
8	1997	360	38
9	1998	450	50
10	1999	550	60

$$\hat{y} = a + bx, \text{ onde: } b = \frac{\frac{1}{10} \sum_{i=1}^{10} x_i y_i - \bar{x} \bar{y}}{\frac{1}{10} \sum_{i=1}^{10} x_i^2 - \bar{x}^2} = \frac{s_{xy}}{s_x^2} \quad \text{e} \quad a = \bar{y} - b \bar{x}$$

$$\bar{x} = \frac{1}{10} \times 2750 = 275$$

$$\bar{y} = \frac{1}{10} \times 281 = 28.1$$

$$\hat{y} = -4.46 + 0.1184 x$$

x_i	y_i	$x_i y_i$	x_i^2
120	10	1200	14400
150	12	1800	22500
160	14	2240	25600
200	20	4000	40000
210	22	4620	44100
250	25	6250	62500
300	30	9000	90000
360	38	13680	129600
450	50	22500	202500
550	60	33000	302500
2750	281	98290	933700

=INTERCETAR(D2:D11;C2:C11)

D	E	F	G	H	I
			b)		
			Reta de Regressão	Gastos = a+b*Vendas	
			a	-4,46762	
			b	0,11843	

=DECLIVE(D2:D11;C2:C11)

D	E	F	G	H	I
			b)		
			Reta de Regressão	Gastos = a+b*Vendas	
			a	-4,46762	
			b	0,11843	

$$b = \frac{\frac{1}{10} \times 98290 - 275 \times 28.1}{\frac{1}{10} \times 933700 - 275^2} = \frac{2101.5}{17745} \approx 0.1184 \quad \text{e} \quad a \approx 28.1 - 0.1184 \times 275 = -4.46$$

MÉTODO DOS MÍNIMOS QUADRADOS

The Least Squares Method minimizes the sum of squared errors:

$$\min_{a,b} Q(a,b) = \min_{a,b} \sum_{i=1}^n (y_i - a - bx_i)^2$$

- Estimador do Parâmetro b (declive)

$$b = \frac{\frac{1}{n} \sum_{i=1}^n y_i x_i - \bar{x} \bar{y}}{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2} = \frac{s_{yx}}{s_x^2}$$

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

- Estimador do Parâmetro a (ordenada na origem)

$$a = \bar{y} - b\bar{x}$$

- Reta de Regressão

$$\hat{y} = a + bx$$

EXERCÍCIO 4.1 C): SOLUÇÃO



Resposta:

$$\bar{y} = \frac{1}{10} \times 281 = 28.1$$

$$\hat{y} = -4.46 + 0.1184 x$$

$$R^2 = 1 - \frac{\frac{1}{10} \times 8.141472}{\frac{1}{10} \times 2496.9} \approx 0.99674$$

=RQUAD(D2:D11;C2:C11)

	E	F	G	H
			c)	
			Coefficiente de Determinação	
			R^2	0,99674

$$R^2 = 1 - \frac{s_e^2}{s_y^2} = 1 - \frac{\frac{1}{n} \sum_{i=1}^n (e_i - \bar{e})^2}{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\frac{1}{n} \sum_{i=1}^n e_i^2}{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}$$

x_i	y_i	$\hat{y}_i$	e_i	e_i^2	$(y_i - \bar{y})^2$
120	10	9.748	0.252	0.063504	327.61
150	12	13.3	-1.3	1.69	259.21
160	14	14.484	-0.484	0.234256	198.81
200	20	19.22	0.78	0.6084	65.61
210	22	20.404	1.596	2.547216	37.21
250	25	25.14	-0.14	0.0196	9.61
300	30	31.06	-1.06	1.1236	3.61
360	38	38.164	-0.164	0.026896	98.01
450	50	48.82	1.18	1.3924	479.61
550	60	60.66	-0.66	0.4356	1017.61
2750	281	281	0	8.141472	2496.9

Interpretação:

- **99,67% da variação nas despesas com publicidade é explicada pelas vendas.**
- O modelo apresenta um **ajustamento excelente** e indica uma **relação linear positiva muito forte** entre as vendas e as despesas com publicidade.

AVALIAÇÃO DA QUALIDADE DO AJUSTAMENTO: COEFICIENTE DE DETERMINAÇÃO

- Parte da seguinte decomposição:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

Variação total = Variação residual + Variação explicada

- Dividindo por n obtém-se,

$$\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

$$s_y^2 = s_e^2 + s_{\hat{y}}^2$$

- Dividindo ambos os membros pela variância total tem-se,

$$1 = \frac{s_e^2}{s_y^2} + \frac{s_{\hat{y}}^2}{s_y^2}$$

Silvestre (2007)

- A partir desta expressão resulta o **coeficiente de determinação** dado por,

$$R^2 = 1 - \frac{s_e^2}{s_y^2} = \frac{s_{\hat{y}}^2}{s_y^2}$$

- Com

$$0 \leq R^2 \leq 1$$

quanto maior o seu valor tanto melhor é a qualidade do ajustamento.

Interpretação:

- $0 \leq R^2 \leq 1$.
- Quanto mais próximo R^2 estiver de 1, **melhor é o ajustamento do modelo aos dados**. O coeficiente de determinação R^2 representa a proporção da variabilidade da variável dependente Y que é explicada pelo modelo de regressão.

AVALIAÇÃO DA QUALIDADE DO AJUSTAMENTO: COEFICIENTE DE DETERMINAÇÃO

In simple linear regression, the total variation in the dependent variable y can be decomposed as:

$$\text{Total Variation (SST)} = \text{Explained Variation (SSR)} + \text{Residual Variation (SSE)}$$

1. Soma dos Quadrados da Variação Total

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2$$

- Measures the total variability of the observed values around their mean.

2.

Soma dos Quadrados da Variação Explicada pela Regressão

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

- Measures the variability explained by the regression model.

3. Soma dos Quadrados dos Erros

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- Measures the variability that the model **fails to explain** (errors).

- SSR — soma dos quadrados da variação explicada pela regressão.
- SST — soma dos quadrados da variação total.

Relationship:

$$SST = SSR + SSE$$

Coefficient of Determination (R^2):

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

- Represents the proportion of the total variation in y explained by the model.
- Higher $R^2 \rightarrow$ better fit.

$$R^2 = \frac{\hat{s}_y^2}{s_y^2} = 1 - \frac{s_e^2}{s_y^2}$$

Silvestre (2007)

COEFICIENTE DE CORRELAÇÃO

Definition:

The sample correlation coefficient measures the **degree and direction** of the linear relationship between two variables X and Y .

It is defined as

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

or equivalently,

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Silvestre (2007)

OBRIGADA!

Questões?

