

1. Responda as perguntas a seguir, considerando a tabela abaixo representa os salários dos funcionários de uma empresa de Informática:

Tabela -1								
Classe	Salário (R\$)		Número de funcionários					
	Limite inf	Limite sup						
1	500	1000	3	3	750,00	2.250,00	1.334.025,00	4002075
2	1000	1500	11	14	1.250,00	13.750,00	429.025,00	4719275
3	1500	2000	40	54	1.750,00	70.000,00	24.025,00	961000
4	2000	2500	44	98	2.250,00	99.000,00	119.025,00	5237100
5	2500	3000	2	100	2.750,00	5.500,00	714.025,00	1428050
			100				190.500,00	16347500

média	R\$ 1.905,00	0,5
mediana	R\$ 1.950,00	0,5
moda	R\$ 2.043,48	0,5
variância	R\$ 163.475,00	0,5
desvio padrão	R\$ 404,32	0,5
Zona de normali	R\$ 1.500,68	0,5
	R\$ 2.309,32	
Porcentagem de funcionários na Zona de normalidade	67,2%	0,5

- a) Qual é o valor da **média** dos salários dos funcionários?
- a) R\$ 1.500,68 b) R\$ 1.905,00 c) R\$ 1.950,00 d) R\$ 2.043,48 e) R\$ 2.309,32
- b) Qual é o valor da **mediana** dos salários dos funcionários?
- a) R\$ 1.500,68 b) R\$ 1.905,00 c) R\$ 1.950,00 d) R\$ 2.043,48 e) R\$ 2.309,32
- c) Qual é o valor da **moda** dos salários dos funcionários?
- a) R\$ 1.500,68 b) R\$ 1.905,00 c) R\$ 1.950,00 d) R\$ 2.043,48 e) R\$ 2.309,32
- d) Qual é o valor do **variância** dos salários dos funcionários?
- a) R\$ 158.900,00 b) R\$ 156.215,00 c) R\$ 163.475,00 d) R\$ 126.125,00 e) R\$ 159.546,00
- e) Qual é o valor do **desvio padrão** dos salários dos funcionários?
- a) R\$ 344,15 b) R\$ 371,04 c) R\$ 398,62 d) R\$ 404,32 e) R\$ 468,24
- f) Quais são os valores (**superior e inferior**) da **primeira Zona de Normalidade**?
- a) R\$ 1.500,68 b) R\$ 1.905,00 c) R\$ 1.950,00 d) R\$ 2.043,48 e) R\$ 2.309,32
- g) Qual a **porcentagem** de funcionários na zona de normalidade?
- a) 49 % b) 56 % c) 67 % d) 78 % e) 89 %

2. Num grupo de 7 pessoas da FATEC foram pesquisados os anos de estudos e a renda familiar média (em salários mínimos). Estes dados encontram-se na Tabela - 2 (ao lado). Quais alunos estão na zona de normalidade de anos de Estudo? Existe uma correlação entre os anos de estudos e a renda familiar. Qual é esta correlação? Determine a reta que melhor ajusta os dados e responda qual é a expectativa de renda familiar média no caso de termos uma família analfabeta? E de egressos da FATEC?

Tabela - 2					
Aluno	Anos de estudos (X)	Renda familiar (Y)	$\sum_{i=1}^{n=7} x_i^2$	$\sum_{i=1}^{n=7} y_i^2$	$\sum_{i=1}^{n=7} x_i y_i$
A	13	4	169,0	16,0	52,0
B	16	5	256,0	25,0	80,0
C	22	7	484,0	49,0	154,0
D	9	2	81,0	4,0	18,0
E	11	3	121,0	9,0	33,0
F	14	4	196,0	16,0	56,0
G	10	3	100,0	9,0	30,0
	95	28	1407	128	423,0

Médias	13,6	4,0
Variância	16,8	2,3
Dp's	4,1	1,5
Zona de Normal Anos Estudos	9,5	17,7
Alunos na Zona de Normalidade	A, B, E, F, G	1,0
Covariância	6,14	
Correlação	0,99	0,5
Intercepto (a)	-0,958	0,5
Inclinação (b)	0,4	
Para o caso de x=0 teremos que o salário médio familiar de :	-0,958	0,5
Para o caso de x=15 teremos que o salário médio familiar de :	4,522	0,5
Para o caso de x=16 teremos que o salário médio familiar de :	4,887	

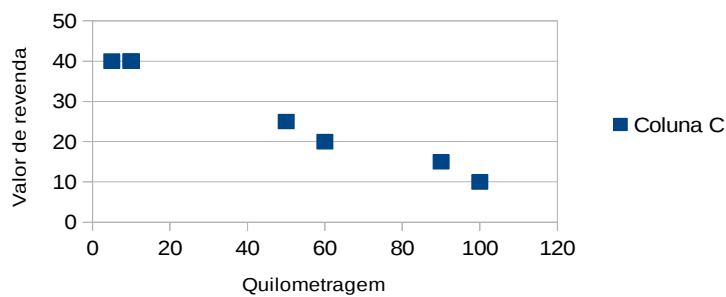
3. Uma pesquisa sobre venda de carros, relacionando o valor de revenda (em milhares de reais) e quilometragem do carro (em milhares de quilômetros) estão na Tabela - 2.
- Qual deles tem maior variabilidade (coeficiente de variação).
 - Faça um gráfico de dispersão.
 - Determine o coeficiente de correlação quilometragem e o valor de revenda.
 - Determine a reta que melhor ajusta e pontos (usar método dos mínimos quadrados).
 - Considerando a reta determinada no item anterior, determine qual seria a estimativa de um carro zero?
 - Depois de qual quilometragem, teríamos que pagar para nos desfazer do carro?

Tabela – 2					
Carro	Quilometragem (X)	Valor de revenda (Y)			
A	5	40	25,0	1600,0	200,0
B	50	25	2500,0	625,0	1250,0
C	10	40	100,0	1600,0	400,0
D	100	10	10000,0	100,0	1000,0
E	60	20	3600,0	400,0	1200,0
F	90	15	8100,0	225,0	1350,0
G	10	40	100,0	1600,0	400,0
	325	190	24425	6150	5800,0

Médias	46,4	27,1	
Variância	1333,7	141,8	
Dp's	36,5	11,9	
Coeficiente de Variação	0,8	0,4	0,5
Covariância	-431,6		
Correlação	-0,99	0,5	
Inclinação	-0,324	0,5	
Intercepto	42,17		
Valor do carro x=0	42,17	0,5	
Valor da quilometragem y=0	130,30	0,5	

0,5

Gráfico de Dispersão



4. O tempo de CPU e disco I/O's (entrada/saída) de sete programas são apresentados na Tabela 1. Faça um gráfico de dispersão relacionando o tempo de CPU (x_i) e tempo de disco I/O's (y_i) (ambos em milissegundos). Considere a seguinte Tabela - 3
- Qual deles tem maior variabilidade (coeficiente de variação).
 - Determine o coeficiente de correlação entre o tempo de CPU (x_i) e tempo de disco I/O's (y_i).
 - Determine a reta que melhor ajusta e pontos (usar método dos mínimos quadrados)
 - Usando o exercício anterior, determine qual seria o tempo de I/O's para um programa que tenha tempo de CPU de 30 milissegundos.

$$\sum_{i=1}^{n=7} x_i = 67 \quad \sum_{i=1}^{n=7} y_i = 291 \quad \sum_{i=1}^{n=7} x_i y_i = 4.782 \quad \sum_{i=1}^{n=7} x_i^2 = 999 \quad \sum_{i=1}^{n=7} y_i^2 = 24.461$$

Tabela - 3

Programa	Tempo de CPU (x)	Time I/O's (y)	$\sum_{i=1}^{n=7} x_i^2$	$\sum_{i=1}^{n=7} y_i^2$	$\sum_{i=1}^{n=7} x_i y_i$
A	24	113	576	12769	2712
B	15	100	225	10000	1500
C	9	27	81	729	243
D	8	27	64	729	216
E	6	12	36	144	72
F	4	9	16	81	36
G	1	3	1	9	3
	67	291	999	24461	4782

Média	9,6	42	51,102041	1766,245
Desvio Padrão	7,1	42	7,1485691	42,02672
Coef de variação	0,75	1,01	0,5	
Covariância	285			
Coef de correlação	0,9495	0,5		285,2
Intercepto (a)	-11,9	0,5		
Inclinação (b)	5,6			4,6231
io de I/O's para x=30:	155,6	0,5		

Formulário

A fórmula para dados no rol

Variância da variável X:

$$Var(X) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

$$Var(X) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

Desvio padrão da variável X:

$$\sigma_x = \sqrt{Var(X)}$$

Coefficiente de variação da variável X:

$$Coefv(X) = \frac{\sigma_x}{\bar{x}}$$

Covariância entre as variáveis X e Y:

$$Cov(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

$$Cov(X, Y) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}$$

Coefficiente de correlação entre as variáveis X e Y:

$$Corr(X, Y) = \frac{Cov(X, Y)}{(\sigma_x \sigma_y)}$$

Regressão linear: $y = a + bx$

Coefficiente de inclinação:

$$b = \frac{Cov(X, Y)}{(\sigma_x)^2}$$

Coefficiente de intercepção:

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

A fórmula para dados agrupados

$$Mediana = l_i + \frac{\frac{n}{2} - Fac_{anterior}}{f_i} \cdot h$$

$$Moda = l_i + \frac{\Delta_{anterior}}{\Delta_{anterior} + \Delta_{posterior}}$$

Onde temos:

- l_i limite inferior da classe i
- f_i frequência da classe i
- h amplitude dos intervalos de cada classe
- n número de dados analisados
- $Fac_{anterior}$ frequência acumulada da classe anterior a classe da mediana
- $\Delta_{anterior}$ variação das frequências da classe da moda e da classe anterior;
- $\Delta_{posterior}$ variação das frequências da classe da moda e da classe posterior;;