

Aluno: _____

nº USP: _____

Cada linha da Tabela 1 abaixo contém valores de quatro variáveis. Considere os valores das quatro variáveis para todas as vinte observações e desenvolva os itens de 1 a 3

Tabela 1. Quatro variáveis				
Observação	X1	X2	X3	X4
1	18	2	-7	128
2	26	5	-17	52
3	79	17	-61	118
4	62	11	-42	14
5	93	16	-51	88
6	22	4	-13	79
7	37	7	-27	23
8	20	3	-17	98
9	48	8	-34	15
10	26	4	-2	75
11	47	9	-28	11
12	48	8	-18	16
13	31	5	-9	58
14	102	15	-49	62
15	138	19	-58	178
16	96	14	-51	49
17	98	18	-60	143
18	53	13	-42	31
19	61	12	-29	18
20	51	10	-23	7

1. Estude as correlações lineares entre todas as variáveis, X1, X2, X3 e X4:

1.a Determine o valor do coeficiente de correlação linear de Pearson para cada par de variáveis e apresente o resultado com **cinco** casas decimais, na forma matricial, ou seja, apresente a matriz de correlação entre as variáveis.

1.b Indique qual é o par de variáveis que apresenta o maior coeficiente de correlação.

1.c Para o par de variáveis indicado no item 1.b, apresente um gráfico de dispersão.

2. Estude as covariâncias entre as variáveis:

2.a Obtenha a covariância amostral entre as variáveis, apresente o resultado com **duas** casas decimais no formato matricial, ou seja, apresente a matriz de covariâncias entre as variáveis.

3. Considere que a variável X4 seja explicada pela variável X2 de acordo com um modelo de regressão linear simples, ou seja:

$$X4 = b_0 + b_1 X2$$

3.a Obtenha as estimativas para o intercepto b_0 e para inclinação b_1 .

3.b Informe o percentual da variação em X4 que pode ser explicada pela variabilidade em X2.

3.c Calcule os resíduos para cada uma das 20 observações.

3.d Faça um gráfico dos resíduos (para todas as observações).

Considere neste gráfico a variável X2 no eixo-x; e os valores de resíduos correspondentes no eixo-y.

Tabela 2. PL e AC de empresas

Empresa	PL	AC
1	60.000	28.000
2	90.000	54.000
3	90.000	36.000
4	80.000	30.000
5	106.000	60.000
6	66.000	26.000
7	85.000	60.000
8	70.000	30.000
9	64.000	40.000

4. Faça uma análise da correlação linear entre PL e AC. (Tabela 2)

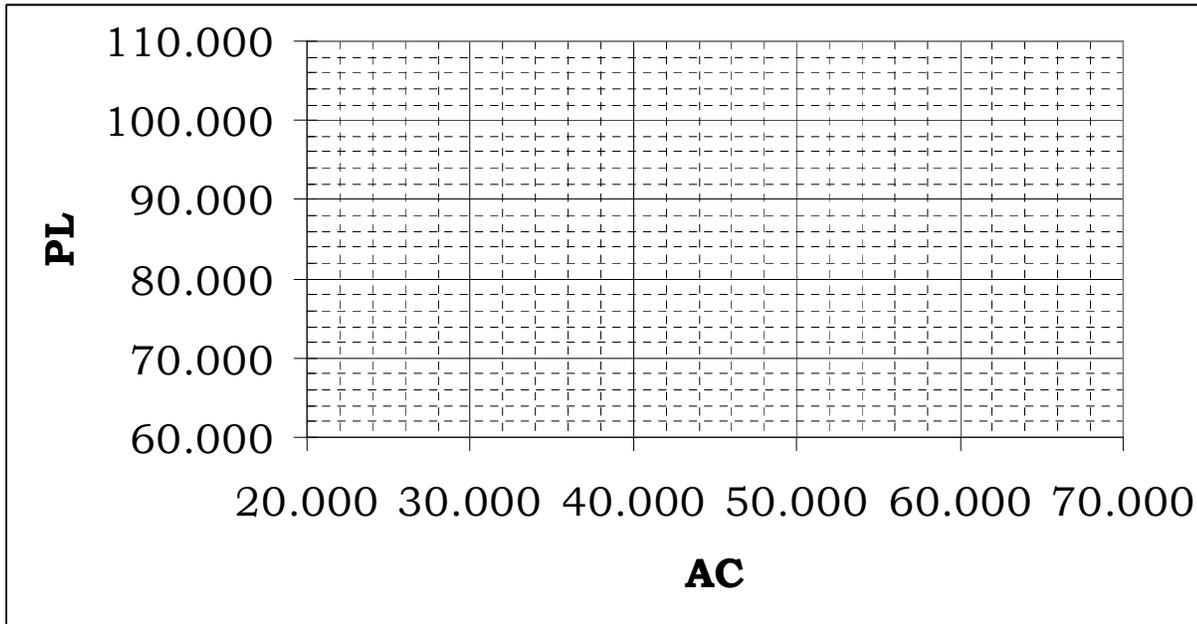
4.a Faça um gráfico de dispersão (se preferir utilize a Figura 1 na próxima página) e comente o que é sugerido pelo gráfico sobre a correlação populacional entre PL e AC.

4.b Calcule o coeficiente de correlação linear amostral r .

4.c Teste a afirmativa de que $\rho = 0$ (utilize o nível de significância $\alpha = 0,01$):

- escreva as hipóteses H_0 e H_1 ;
- obtenha o valor crítico;
- calcule a estatística teste;
- estabeleça uma conclusão com relação a H_0 ;
- faça uma conclusão com base na afirmativa original.

Figura 1. Figura para ser utilizada na análise de correlação entre AC e PL, exercício 4.a.



5. Obtenha um modelo de regressão linear simples que descreva PL como função de AC (utilize a Tabela 2).

5.a Obtenha a equação $PL = b_0 + b_1 AC$.

5.b Trace a reta de regressão no diagrama de dispersão.

5.c Obtenha os resíduos e faça um gráfico dos resíduos (eixo-y) com os valores de AC (eixo-x)

5.d Qual é o valor do coeficiente R^2 e o que significa?

Tabela 3. Preços médios da Kombi Furgão

Ano	Álcool	Gasolina
2006	26920	28335
2005	27529	27680
2004	24102	25295
2003	20286	24646
2002	20102	20963
2001	16102	18338
2000	16050	17026
1999	12834	13783
1998	12960	13173
1997	9375	10189
1996	7123	10879
1995	8331	8487

A Tabela 3 apresenta o valor médio do veículo Kombi Furgão a álcool (Álcool) e a gasolina (Gasolina) para os anos de 1995 a 2006. (Tabela modificada de A.L. Bruni, *Estatística Aplicada à gestão empresarial*, Editora Atlas, 2008)

6. Estude as correlações entre todas as variáveis. Para tanto:

6.a Construa gráficos de dispersão para cada par de variáveis, ou seja, para Ano-Álcool, Ano-Gasolina e Álcool-Gasolina.

6.b Calcule o coeficiente de correlação linear de Pearson entre cada par de variáveis e escreva a matriz de correlação 3x3. (valores com **quatro** casas decimais)

6.c Verifique se as correlações tem significância estatística ao nível ao nível $\alpha = 0,05$.

6.d Obtenha a matriz de covariância.

6.3 Com base nos itens anteriores, comente sobre a correlação entre as variáveis.

7. Considere um modelo de regressão linear simples para escrever o preço da Kombi a álcool (y) em termos do preço da Kombi a gasolina (x).

7.a Obtenha a equação linear (valores de b_0 e b_1) que descreva o preço da Kombi a álcool em termos do preço da Kombi a gasolina. Interprete o resultado obtido. (valores com **quatro** casas decimais)

7.b Calcule o coeficiente de determinação (R^2). Interprete o seu significado.

7.c Faça um gráfico com os pontos originais e com a reta obtida no item 7.a.

7.d Se o preço da Kombi a gasolina for igual a 18.000 qual é a estimativa do preço da Kombi a Álcool?

A Tabela 4 apresenta quatro registros da Quantidade vendida diária de carros em uma concessionária e também valores de uma variável denominada x_1 e valores de uma variável *dummy*, x_2 .

Tabela 4

Quantidade	x_1	x_2
5	2	0
4	1	0
0	0	0
2	1	1

8 Investigue a correlação linear entre as variáveis “Quantidade” e “ x_1 ” apresentadas na Tabela 4, para tanto:

8.a Faça um gráfico de dispersão com a variável “Quantidade” no eixo-y e a variável “ x_1 ” no eixo-x.

8.b Determine o coeficiente de correlação linear de Pearson entre as variáveis (utilize 4 casas decimais).

8.c Com base nos itens 8.a e 8.b, comente sobre a correlação linear entre as variáveis.

9 Suponha um modelo linear no qual a quantidade (Q) seja uma função de x_1 . A partir dos dados da Tabela 4 desenvolva um modelo de regressão linear simples:

9.a Escreva a equação para estimativa de Q : $\hat{Q} = a + b x_1$

ou seja, determine o intercepto “ a ” e a inclinação “ b ”. Escreva a equação linear.

9.b Interprete o significado do modelo. Como a variável x_1 influencia a Quantidade vendida diária?

10 Suponha que no dia correspondente à última linha da Tabela 4 tenha ocorrido um evento que pode ter influenciado a quantidade vendida. Investigue a influência deste evento na quantidade vendida através de um modelo de regressão linear múltipla com a variável x_1 e com a variável *dummy*, x_2 , utilizada para descrever a ocorrência do evento.

10.a Apresente o diagrama de dispersão análogo ao exercício 8.a, mas agora apresente símbolos diferentes para diferentes categorias da variável categórica x_2 .

10.b Escreva a matriz \mathbf{Z} e a matriz transposta \mathbf{Z}' . Obtenha o resultado $\mathbf{Z}'\mathbf{Z}$

10.c Obtenha a matriz inversa: $(\mathbf{Z}'\mathbf{Z})^{-1}$.

10.d Obtenha a matriz $\mathbf{Z}'\mathbf{Y}$.

10.e Obtenha os valores de b 's: $\boldsymbol{\beta} = (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{Y})$

10.f Escreva a equação para estimativa da quantidade: $\hat{Q} = b_0 + b_1 x_1 + b_2 x_2$

e comente o significado de b_0 , b_1 e b_2 .