



## Lista de Exercícios 2

- 2.1) As notas de uma sala com nove estudantes em um relatório de meio de curso ( $x$ ) e em um exame final ( $y$ ) são as seguintes:

$x$	77	50	71	72	81	94	96	99	67
$y$	82	66	78	34	47	85	99	99	68

- (a) Faça o gráfico de dispersão.
  - (b) Determine o coeficiente de correlação e explique.
  - (c) Estime a reta de regressão linear.
  - (d) Estime a nota do exame final de um aluno que teve uma nota 85 no relatório de meio do curso.
- 2.2) Um engenheiro de *Machine Learning* está investigando o impacto de um hiperparâmetro crucial (Taxa de Aprendizado) na precisão de um modelo de classificação. Para simplificar a análise, a taxa de aprendizado foi codificada (valores de 1.0 a 2.0). O objetivo é encontrar a relação entre essa taxa e a precisão alcançada pelo modelo em testes de validação. Os dados registrados são os seguintes:

Taxa de Aprendizado ( $x$ )	Precisão do modelo ( $\%$ , $y$ )
1,0	8,1
1,1	7,8
1,2	8,5
1,3	9,8
1,4	9,5
1,5	8,9
1,6	8,6
1,7	10,2
1,8	9,3
1,9	9,2
2,0	10,5

- (a) Faça um diagrama de dispersão para esses dados.
- (b) Obtenha o coeficiente de correlação linear e interprete-o no contexto do problema.
- (c) Estime a reta de regressão linear que descreve a relação entre a taxa de aprendizado e a precisão do modelo.

- (d) Estime a média da precisão do modelo quando a taxa de aprendizado codificada é 1,75.
- 2.3) Em um projeto de otimização de Big Data, uma equipe de Ciência de Dados está investigando a relação entre a complexidade teórica de um algoritmo de ordenação (representada por uma métrica codificada,  $x$ ) e o tempo real de execução (em milissegundos - ms) para processar um conjunto de dados de 1 milhão de registros. O objetivo é prever o tempo de execução com base na complexidade. Os dados coletados de diferentes configurações e otimizações do algoritmo são:

Complexidade do Algoritmo ( $x$ )	Tempo de Execução ( $y$ )
4,3	126
4,5	121
5,9	116
5,6	118
6,1	114
5,2	118
3,8	132
2,1	141
7,5	108

- (a) Faça um diagrama de dispersão para esses dados.
- (b) Determine o coeficiente de correlação linear e interprete-o no contexto do desempenho do algoritmo.
- (c) Determine a equação da reta de regressão para prever o tempo de execução a partir da complexidade do algoritmo.
- (d) Estime o tempo de execução quando a complexidade do algoritmo codificada é  $x = 4,8$  unidades.

- 2.1) (b) 0,5610  
(c)  $\hat{y} = 12,0623 + 0,7771x$   
(d) 78,1158
- 2.2) (b) 0,7070  
(c)  $\hat{y} = 6,4136 + 1,8091x$   
(d) 9,5795
- 2.3) (b)  $-0,9787$   
(c)  $\hat{y} = 153,1755 - 6,3240x$   
(d) 122,8203