

Exercícios

Ano letivo 2024/25

1. Correlação

1.1. O departamento de qualidade de um determinado hospital, pretende fazer um estudo sobre o tempo que os utentes encaminhados para cirurgia pelos seus médicos de família demoram a ser efetivamente operados. Estes utentes têm que primeiro fazer uma consulta de referência no hospital (x dias após o encaminhamento) e só posteriormente são operados (y dias após a consulta de referência). Escolheram-se aleatoriamente 15 utentes nestas condições, tendo sido reportados os seguintes tempos de espera, também disponíveis em *tempos.csv* (fonte: Afonso e Nunes, 2019):

x	69	76	51	34	62	13	40	7	64	41	64	26	40	44	48
y	28	64	7	26	38	18	40	20	44	32	31	32	36	25	73

- a) Através da análise gráfica, parece-lhe existir relação linear entre as duas variáveis?
- b) Calcule e interprete o valor do coeficiente de correlação linear de Pearson.

1.2. (Excel) Observou-se uma amostra de 26 modelos de *machine learning* tendo-se registado para cada um deles o tempo de execução (em segundos) e o número de parâmetros, tendo-se obtido os seguintes resultados, que também estão disponíveis em *ML.csv*:

Modelo	N.º parâmetros	Tempo	Modelo	N.º parâmetros	Tempo
1	49	71,66	14	21	34,88
2	15	20,3	15	22	35,23
3	1	1,46	16	39	61,27
4	35	52,2	17	17	21,35
5	7	10,24	18	6	9,14
6	11	18,41	19	43	65,95
7	9	10,37	20	1	4,95
8	43	64,23	21	39	57,82
9	32	49,5	22	45	67,29
10	50	75,15	23	28	44,98
11	49	77,32	24	14	24,88
12	19	28,4	25	25	44,56
13	19	24,87	26	17	23,6

- a) Considera que o tempo de execução está relacionado com o número de parâmetros? Justifique com auxílio de um gráfico.
- b) Calcule a covariância entre o tempo de execução e o número de parâmetros.
- c) Calcule o coeficiente de correlação de Pearson entre o tempo de execução e o número de parâmetros.
- d) Repita as alíneas anteriores, considerando o tempo em minutos. Pronuncie-se sobre as diferenças obtidas.

1.3. (■) Os dados apresentados na tabela seguinte foram extraídos do Relatório do Desenvolvimento Humano, de 2023/24, publicado pelo Programa das Nações Unidas para o Desenvolvimento (link: [Relatório do Desenvolvimento Humano \(RDH\) 2023-2024 | United Nations Development Programme](#)). Para os 27 países da União Europeia foram calculados: o Índice de Desenvolvimento Humano (IDH), o IDH ajustado às desigualdades (IDHAjD), o Índice de Desenvolvimento do Género (IDG), o Índice de Desigualdade de Género (IDiG), e o IDH ajustado à pressão planetária (IDHAjPP). Os dados estão disponíveis no ficheiro *RDH2324.xlsx*.

País	IDH	IDHAjD	IDG	IDiG	IDHAjPP
Alemanha	0,950	0,881	0,966	0,071	0,833
Áustria	0,926	0,859	0,972	0,048	0,789
Bélgica	0,942	0,878	0,975	0,044	0,803
Bulgária	0,799	0,703	0,995	0,206	0,72
Chéquia	0,895	0,848	0,988	0,113	0,782
Chipre	0,907	0,827	0,977	0,253	0,803
Croácia	0,878	0,817	0,993	0,087	0,807
Dinamarca	0,952	0,898	0,981	0,009	0,839
Eslováquia	0,855	0,808	1,002	0,184	0,776
Eslovénia	0,926	0,882	0,999	0,049	0,832
Espanha	0,911	0,796	0,988	0,059	0,839
Estónia	0,899	0,835	1,022	0,093	0,766
Finlândia	0,942	0,886	0,989	0,032	0,787
França	0,910	0,820	0,986	0,084	0,823
Grécia	0,893	0,801	0,969	0,120	0,809
Hungria	0,851	0,800	0,989	0,230	0,769
Irlanda	0,950	0,886	0,991	0,072	0,814
Itália	0,906	0,802	0,969	0,057	0,825
Letónia	0,879	0,802	1,022	0,142	0,782
Lituânia	0,879	0,795	1,028	0,098	0,748
Luxemburgo	0,927	0,839	0,993	0,043	0,685
Malta	0,915	0,837	0,980	0,117	0,806
Países Baixos	0,946	0,885	0,960	0,025	0,796
Polónia	0,881	0,797	1,009	0,105	0,78
Portugal	0,874	0,774	0,998	0,076	0,807
Roménia	0,827	0,739	0,981	0,230	0,759
Suécia	0,952	0,878	0,983	0,023	0,839

- Quais os índices que lhe parecem estar mais correlacionados? Confirme o seu palpite calculando os respetivos coeficientes de correlação.
- Se recorresse ao coeficiente de Spearman em vez do coeficiente de Pearson, parece-lhe que a avaliação da associação entre as variáveis se iria alterar substancialmente? Justifique.