
Ficha de Trabalho 1

1. Em relação às seguintes frases identifique e classifique as variáveis estatísticas:
 - (a) Temperatura em Setúbal às 12h do dia.
 - (b) Programa televisivo preferido pelos alunos de uma turma.
 - (c) Intensidade de um treino desportivo (baixa, moderada, elevada).
 - (d) Formas de pagamento numa loja.
 - (e) Quantidade de calorias por pacote de batatas fritas.
 - (f) Número de golos marcados por uma equipa de futebol durante o campeonato.
 - (g) Grau de satisfação de clientes com relação ao serviço prestado por uma empresa (valores de 0 a 5, onde 0 indica totalmente insatisfeito e 5 totalmente satisfeito).

2. Considere a tabela seguinte onde foram registadas informações de alguns hotéis na Europa:

Nome do Hotel	País	Preço do quarto	Número de quartos	Pontuação (0 – 10)
Hotel Ronda	Espanha	\$\$	18	8.4
Villad’Este	Itália	\$\$\$\$	166	8.6
Hotel Lisboa	Portugal	\$	81	8.5
Hotel Prem	Alemanha	\$	54	7.7
Hotel d’Europa	França	\$\$	47	7.6
Palace Luzern	França	\$\$	326	8.1
Hotel Palace	Portugal	\$\$\$\$	185	9.5
Hotel Arts	Espanha	\$\$\$	45	7.3
Hotel Sacher	Alemanha	\$\$\$	120	8.5
Duc de Bourgogne	França	\$	10	7.6
Villa Gallici	França	\$\$	22	9.0
Hotel Vila	Portugal	\$\$	233	9.1

- (a) Identifique a População, a Amostra, a unidade estatística, as variáveis estatísticas e os dados estatísticos, classificando-os.
- (b) Introduza a tabela no R.
- (c) Qual a dimensão da amostra recolhida?
- (d) Construa as tabelas de frequências das variáveis qualitativas.
- (e) Quantos hotéis ficam na Península Ibérica?
- (f) Qual a percentagem de hotéis com o preço mais elevado?

3. O ficheiro "stroke.txt" corresponde a uma amostra recolhida dos 829 casos de acidente vascular cerebral que ocorreram na Estónia durante o período 1991-1993. Os dados recolhidos apresentam os seguintes campos:

- sex: género do paciente (0 = masculino, 1 = feminino)
 - age: idade do paciente no momento do acidente vascular cerebral
 - dgn: diagnóstico com níveis: ICH (hemorragia intracraniana), INF (enfarte isquémico), SAH (hemorragia subaracnóidea), ID (não identificado).
 - coma: indica se o paciente entrou em coma após o AVC (1 = Não, 2 = Sim).
 - diab: se o paciente tem histórico de diabetes (No = Não, Yes = Sim).
 - dead: se o paciente morreu durante o estudo (TRUE = Sim, FALSE = Não).
 - obsmonths: tempo de observação do paciente em meses (definido como 0.1 para pacientes que morrem no mesmo dia do AVC)
- (a) Identifique a População, a unidade estatística e as variáveis estatísticas.
- (b) Caso seja possível, indique a dimensão da população e da amostra.
- (c) Classifique as variáveis estatísticas.
- (d) Construa as tabelas de frequências das variáveis qualitativas.
- (e) Quantos pacientes tiveram diagnóstico não identificado?
- (f) Qual a percentagem de pacientes que morreu durante o estudo?
- (g) Quantos pacientes entraram em coma e tinham histórico de diabetes?
- (h) Quantos pacientes morreram no dia do AVC?

4. No ficheiro EXCEL "Obesidade" tem as respostas a um inquérito efetuado num estudo sobre obesidade a um grupo de indivíduos obesos. Os dados recolhidos referem-se às respostas dadas às seguintes questões:

- Género: Feminino, Masculino;
- Idade: em anos;
- Altura: em metros;
- Peso: em quilogramas;
- FAVC = Se come alimentos altamente calóricos habitualmente: 1=Sim, 0=Não;
- FCVC = Se come habitualmente vegetais nas refeições: 1= Nunca, 2=Às vezes, 3=Sempre;
- NCP = Número de refeições principais que tem habitualmente: 0, 1, 2, 3, 4,...;
- CAEC = Se come entre as refeições principais: N=Não, S=Às vezes, F=frequentemente, A=sempre;
- Fumar = Se fuma: 1=Sim, 0=Não;
- CH2O = Quantidade de água que bebe diariamente: 1=menos de 1 litro, 2=entre 1 e 2 litros, 3=mais de 2 litros;
- FAF = Com que frequência pratica atividade física por semana: 0=não pratica, 1=1 ou 2 dias, 2=3 ou 4 dias, 3=mais de 4 dias;
- CALC = Com que frequência bebe bebidas alcoólicas: N=Nunca, S=Às vezes, F=frequentemente, A=Sempre;

- MTRANS = Que tipo de transporte utiliza habitualmente: Automóvel, Mota, Bicicleta, Transportes públicos, a pé;
- (a) Classifique as variáveis.
 - (b) Qual a dimensão da amostra recolhida?
 - (c) Qual a percentagem de pessoas inquiridas que come alimentos altamente calóricos habitualmente?
 - (d) Qual a percentagem de pessoas inquiridas que come entre as refeições principais?
 - (e) Qual a percentagem de pessoas inquiridas que utiliza habitualmente bicicleta ou transportes públicos?
 - (f) Construa as tabelas de frequências das variáveis quantitativas discretas.
 - (g) Em relação às variáveis quantitativas contínuas organize os dados em tabelas de frequência. Na construção das tabelas deve considerar o seguinte:
 - i. em relação à variável idade, considere 8 classes, com amplitude 6 anos e a primeira classe a começar nos 14 anos (considere intervalos abertos à esquerda e fechados à direita);
 - ii. em relação à variável altura, resuma a informação nas 3 classes: $[1.45, 1.60]$, $]1.60, 1.80]$ e $]1.80, 2.00]$;
 - iii. em relação à variável peso, recorra à regra de Sturges para definir as classes (considere intervalos fechados à esquerda e abertos à direita).
 - (h) Quantas pessoas inquiridas têm mais de 44 anos?
 - (i) Qual a percentagem de pessoas inquiridas com altura máxima de 1.80 metros?