|  |  |
| --- | --- |
| Escola Superior de Ciências Empresariais - ESCE/IPS | Mestrado Ciência de Dados para Empresas  Fundamentos Estatísticos  Profs. Sandra Nunes & Sandra Oliveira  2023/2024 |

CASO PRÁTICO 1

Estatística Descritiva

**Exercício 1:**

Classifique as seguintes variáveis:

|  |  |  |
| --- | --- | --- |
| Nacionalidade (Portuguesa/Outra) | Cor dos automóveis (Branco/Preto/Vermelho/Cinza/…) | Número de erros num ditado |
| Número de idas a um restaurante por mês | Idade de um indivíduo (anos) | Duração de uma bateria de telemóvel (horas) |
| Rendimento familiar mensal (Baixo/Médio/Alto) | Marcas de automóveis (A/B/C) | Peso de um indivíduo (kg) |
| Rendimento familiar mensal (euros) | Altura de um indivíduo (m) | Salário mensal (euros) |
| Escalão etário (menos de 18 anos/18-30 anos/mais de 30 anos) |  |  |

|  |
| --- |
| **Nominal:**  Nacionalidade (Portuguesa/Outra)  Cor dos automóveis (Branco/Preto/Vermelho/Cinza/...)  Marcas de automóveis (A/B/C)  Escalão etário (menos de 18 anos/18-30 anos/mais de 30 anos)  **Discreta:**  Número de erros num ditado  Número de idas a um restaurante por mês  Idade de um indivíduo (anos)  Duração de uma bateria de telemóvel (horas)  Peso de um indivíduo (kg)  Altura de um indivíduo (m)  **Ordinal:**  Rendimento familiar mensal (Baixo/Médio/Alto)  Rendimento familiar mensal (euros)  Salário mensal (euros) |

**Exercício 2:**

Numa empresa recolheu-se informação sobre algumas características de alguns dos seus colaboradores. Alguns dos outputs obtidos foram os seguintes.

Sexo (Feminino/Masculino)

Idade (anos)

2.1 Classifique as variáveis.

|  |
| --- |
| **Sexo:**  Nominal: uma ordem específica.  Proporção: A porcentagem associada a cada sexo é uma variável de proporção, indicando a distribuição relativa dos sexos na amostra.  **Idade:**  Discreta: valores contáveis (anos).  Frequência Absoluta e Relativa: A distribuição da idade é apresentada em termos de frequência absoluta (número de vezes que cada idade ocorre) e frequência relativa (porcentagem de cada idade em relação ao total). |

2.2 Analise os resultados.

|  |
| --- |
| **Sexo:**  A maioria dos colaboradores é do sexo feminino (84.2% feminino, 15.8% masculino).  **Idade:**  A distribuição das idades mostra uma predominância de colaboradores mais jovens, especialmente na faixa etária de 25 anos. |

**Exercício 3:**

Além das características indicadas no exercício anterior, também se recolheu informação sobre as seguintes variáveis:

Transporte utilizado pelos colaboradores nas deslocações casa-trabalho (Próprio/Transportes Públicos/Próprio e Transportes Públicos/ Outro)

Tempo semanal dedicado à prática de exercício físico (0-2/3-5/6-9/9-11/12-14/15 ou mais)

Número de irmãos

Uma imagem com texto, captura de ecrã, diagrama, Gráfico

Descrição gerada automaticamente Uma imagem com texto, diagrama, círculo, captura de ecrã

Descrição gerada automaticamente

Uma imagem com texto, diagrama, file, captura de ecrã

Descrição gerada automaticamente

3.1 Classifique as variáveis.

|  |
| --- |
| **Nominal:** As opções (Próprio/Transportes Públicos/Próprio e Transportes Públicos/Outro) são categorias sem uma ordem específica.  **Tempo Semanal Dedicado à Prática de Exercício Físico:**  Ordinal: As categorias (0-2/3-5/6-9/9-11/12-14/15 ou mais) representam uma ordem intrínseca, pois indicam intervalos crescentes de tempo dedicado à prática de exercício físico.  **Número de Irmãos:**  Discreta: O número de irmãos é uma variável discreta, pois representa uma contagem de valores inteiros. |

3.2 Analise os outputs.

|  |
| --- |
| **Transporte utilizado:** A maioria dos colaboradores utilizada o trasnporte publico, correspondendo 52,6% da amostra.  **Tempo semanal dedicado à prática de exercício físico:** A amostra indica que o periodo praticado com mais frequencia é 0-2h  **Irmãos:** A maioria dos colaboradores possui pelo menos 1 irmão. |

**Exercício 4:**

As notas dos alunos de uma turma são:

9, 15, 15, 17, 14, 16, 19, 10, 14, 15, 18, 16, 13, 17

Calcule a média, mediana e o desvio padrão e interprete os valores obtidos.

|  |
| --- |
| Média => 216/14 = 14,9  Desvio Padrão = 2.9  Mediana => ordenar os valores => 9,10,13,14,14,15,15,15,16,16,17,17,18,19 = 15  Avaliação: A média corresponde bem a amostra pois não há valores expressivos para influenciar no resultado.  O desvio padrão é baixo, indicando que a os valores da amostra são proximos da média. |

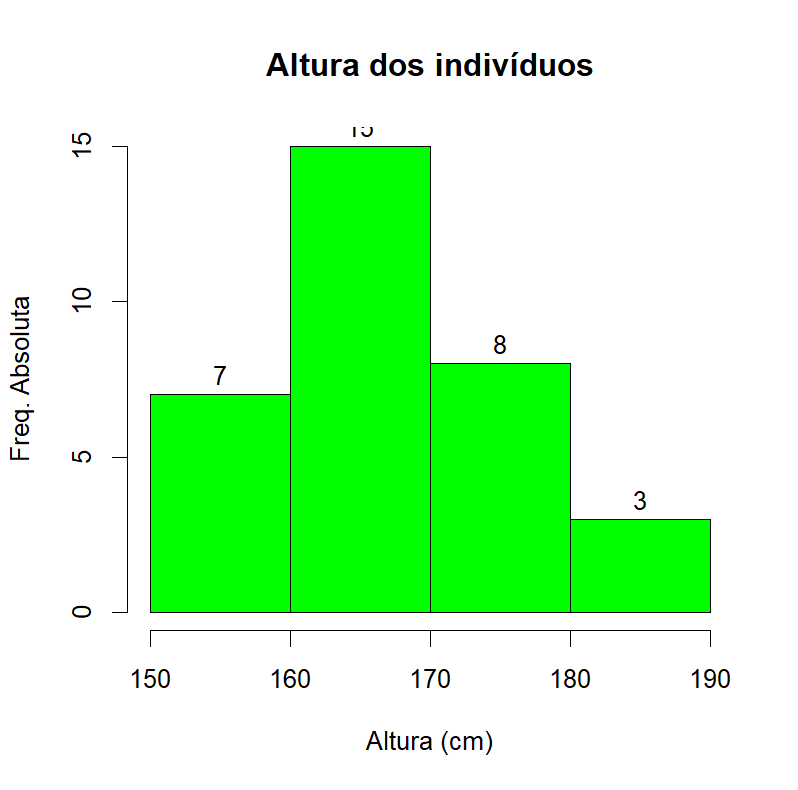
**Exercício 5:**

Analisando os seguintes *ouputs* caracterize o Peso (kg), a Altura (cm) e a Idade (anos) dos indivíduos que participaram num estudo.

**Peso**

N Min. 1st Qu. Median Mean 3rd Qu. Max. Sd

33 46.00 57.00 64.00 65.18 69.00 96.00 11.485



**Idade**

vars n mean sd median trimmed mad min max range skew kurtosis se

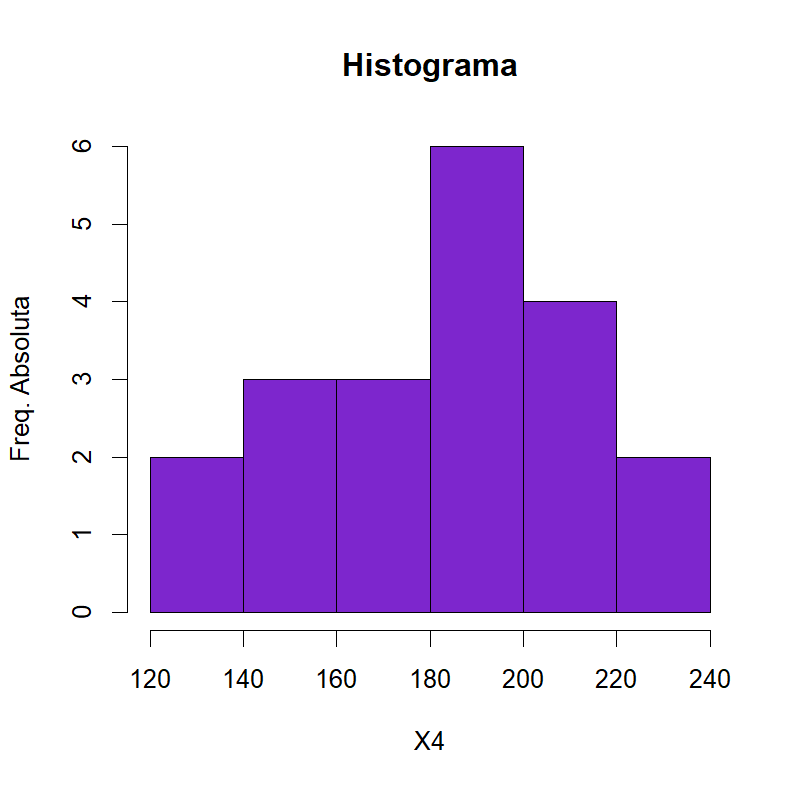
Idade 33 21.03 1.59 21 20.7 1.48 20 28 8 2.75 8.76 0.28

**Exercício 6:**

Que ferramenta(s) estatística(s) pode utilizar para analisar a simetria/assimetria da distribuição de frequências?

**Exercício 7:**

Analisando os outputs seguintes estude a simetria/assimetria dos dados.

Uma imagem com texto, diagrama, captura de ecrã, Gráfico

Descrição gerada automaticamente

vars n mean sd median trimmed mad min max range skew kurtosis

X4 20 186 31.23 192.5 188.44 25.95 120 230 110 -0.52 -0.78

se

X4 6.98

vars n mean sd median trimmed mad min max range skew kurtosis

X9 24 218.33 40.45 212.5 216.25 37.06 160 310 150 0.52 -0.74

se

X9 8.26

**Exercício 8:**

1. Considere a seguinte amostra

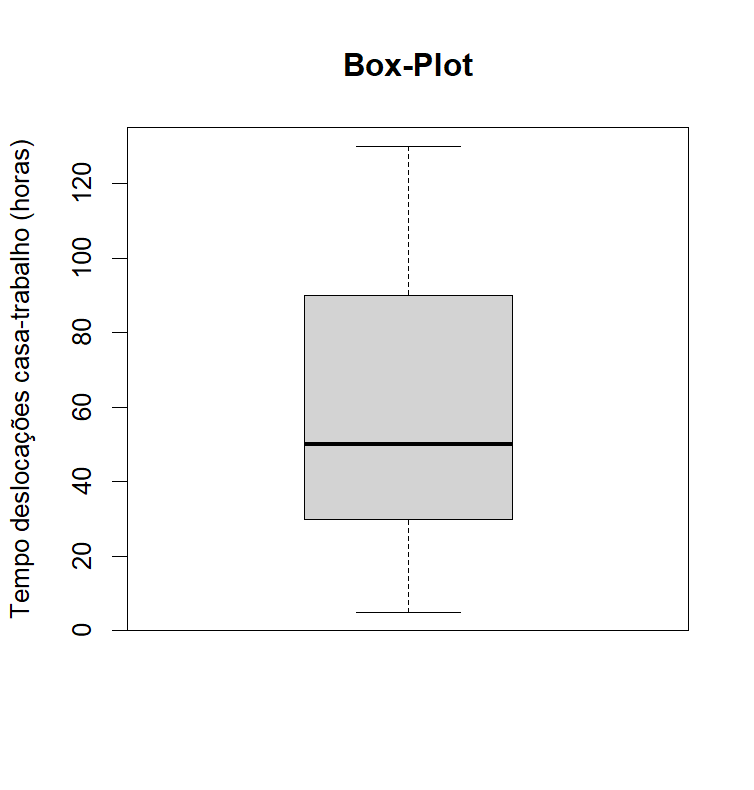
12, 24, 37, 42, 55, 62, 72, 77, 246, 592

Qual(ais) a(s) medida(s) de tendência central e de dispersão adequadas para caracterizar a amostra? Justifique.

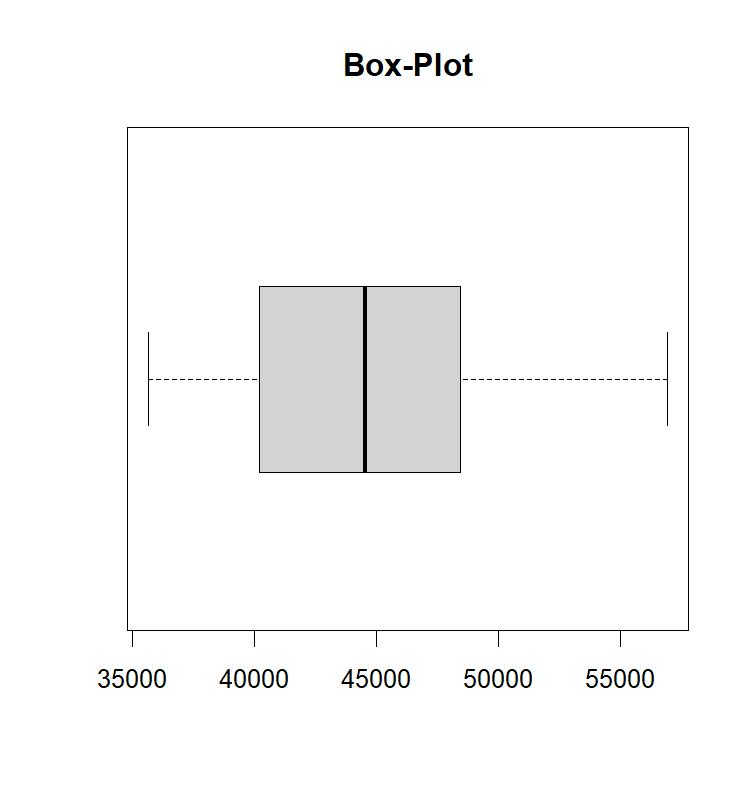
1. Que medida(s) é(são) adequada(s) para caracterizar as variáveis Sexo, Curso e Estado civil?

**Exercício 9:**

Considere as box-plot seguintes e identifique (aproximadamente) os 5 números.

Uma imagem com texto, captura de ecrã, diagrama, Retângulo

Descrição gerada automaticamente

 Uma imagem com texto, diagrama, captura de ecrã, Retângulo

Descrição gerada automaticamente