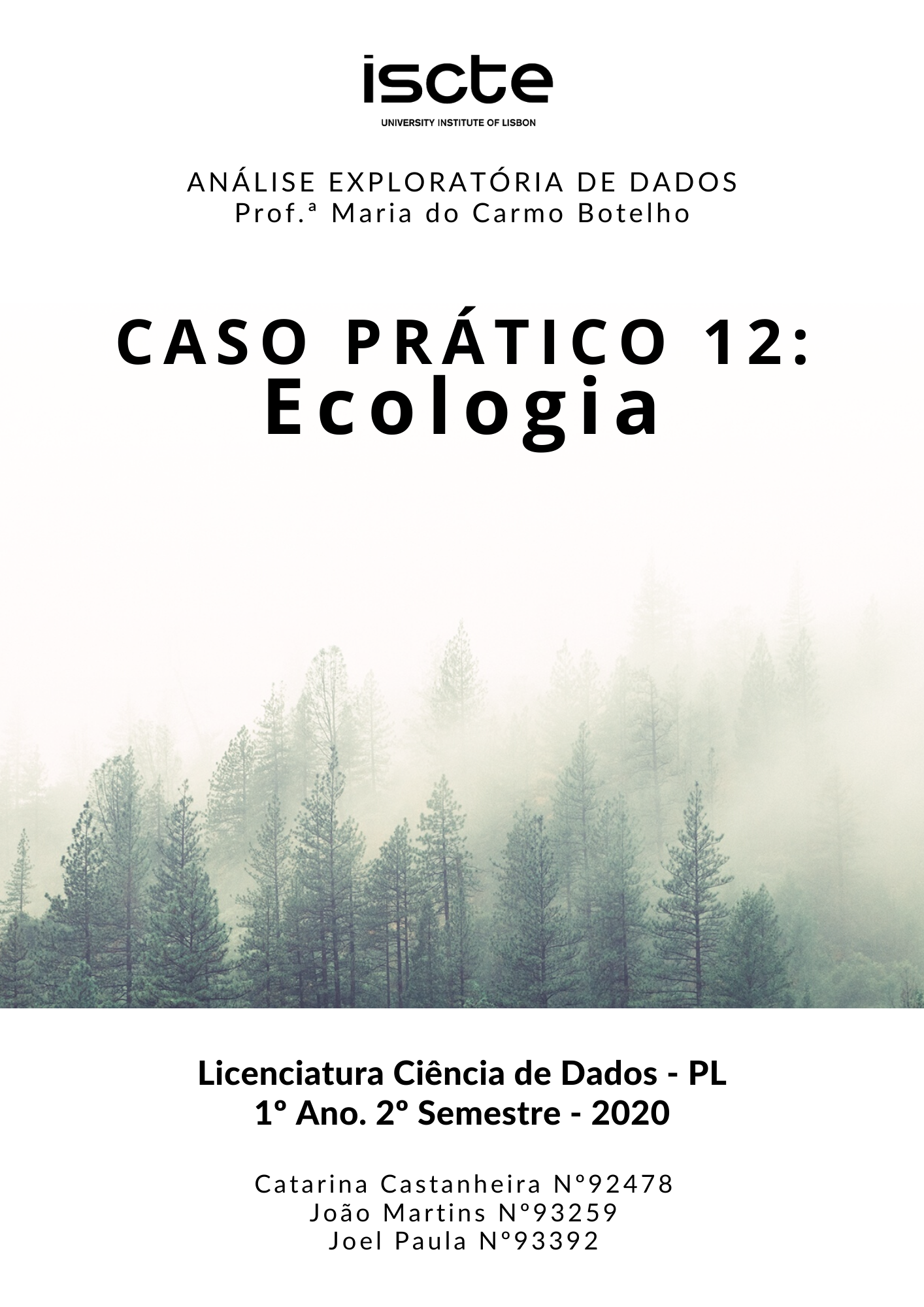
****

**Z**

Índice

[Introdução 3](#_Toc41834913)

[I – Excel 3](#_Toc41834914)

[II – Jamovi 7](#_Toc41834915)

# Introdução

No *dataset* fornecido para este caso prático, temos a informação relativa à resposta de uma amostra de 354 inquiridos, residentes em Portugal, a um questionário que tinha por objetivos principais avaliar a importância atribuída a diversos problemas ambientais, bem como aferir o nível de preocupação ambiental.

O conjunto de dados contém um total de 17 variáveis, as quais podem ser vistas como pertencentes a dois grandes grupos: as variáveis que caracterizam a amostra de inquiridos – *sexo*, *idade*, *filhos* (ter filhos: Com filhos; Sem filhos), *habilit* (habilitações literárias: Primário, Secundário, Superior), *automovel* (posse de automóvel: Com automóvel; Sem automóvel), *rendimento* (sem rendimento, <400€, 400-600€, 600-1000€, >= 1000€); e as variáveis que retêm informação sobre a importância atribuída aos vários problemas ambientais: Os incêndios florestais (*p8.1)*, acidentes de trabalho (*p8.3*), poluição dos rios (*p8.6*), poluição das praias (*p8.9*), má qualidade do ar (*p8.10*), utilização de pesticidas (*p8.12*), efeito de estufa (*p8.13*), reciclagem dos lixos (*p8.14*), destruição de florestas tropicais (*p8.19*), nível de preocupação com o ambiente (*preoc\_ambiente*). Dentro deste segundo conjunto de variáveis, e exceptuando preoc\_ambiente (que tem opção de resposta entre valores 0 e 100), todas têm opção de resposta numa escala de Likert de 5 pontos (entre Nada Importante -1 - e Muito Importante - 5).

Ainda analisando as variáveis, é possível classificá-las pelo tipo, da seguinte forma:

* *sexo*, *filhos*, *automóvel*: variáveis qualitativas nominais;
* *habilit*, *rendimento*, conjunto de variáveis *p8.1* a *p8.19*: variáveis qualitativos ordinais;
* *idade, preoc\_ambiente*: variáveis quantitativas discretas.

Num primeiro momento, iremos fazer uma primeira limpeza do *dataset* fornecido, com recurso ao MS Excel e gerar estatísticas descritivas para as variáveis relativas a Incêndios Florestais, Efeito de Estufa, Preocupação Ambiental, Habilitações Literárias e Sexo.

Os passos para esta primeira fase estão definidos abaixo:

# I – Excel

1. **Aplicar o filtro a todas as variáveis e apresentar a primeira linha fixa:**
   1. Selecionar primeira linha 🡪 View 🡪 Freeze Panes 🡪 Freeze Panes
   2. Data 🡪 Filter
2. **Atribuir código de não resposta a uma variável – *Incêndios Florestais (p8.1)* – com valores omissos:**
   1. Selecionar coluna *p8.1* 🡪 Search & Replace
   2. Find: 🡪 Deixar campo vazio
   3. Replace: 🡪 Inserir número 99
3. **Atribuir texto às categorias de uma variável que tenha códigos; neste aspeto são trabalhadas as variáveis sexo, filhos e *habilit*:**
   1. Consultar folha “Variáveis e Códigos” 🡪 Verificar quais variáveis têm código 🡪 Selecionar coluna *sexo* na folha Dados\_Eco1 🡪 Criar cópia de coluna e atribuir nome *sexotexto* 🡪 Aplicar a todas as linhas da coluna: =HLOOKUP(Dados\_Eco1!L2;'Variáveis e códigos'!$D$3:$E$4;2;FALSE)
   2. Selecionar coluna *filhos* 🡪 Criar cópia de coluna e atribuir nome *filhostexto* 🡪 Aplicar a todas as linhas da coluna: =VLOOKUP(O2;'Variáveis e códigos'!$D$7:$E$8;2; FALSE)
   3. Selecionar coluna *filhos* à Search & Replace 🡪 Find: à deixar campo vazio 🡪 Replace: 🡪 introduzir número 99
   4. Selecionar coluna *habilit* 🡪 Criar cópia de coluna e atribuir o nome *habilitacoestexto* 🡪 Aplicar a todas as linhas da coluna: =HLOOKUP(Q2;'Variáveis e códigos'!$D$11:$F$12;2;FALSE)
4. **Corrigir possíveis erros, associados a todas as variáveis:**
   1. Com recurso aos filtros criados para todas as colunas, consultar cada uma por valores estranhos (fora da amplitude de resposta esperada) ou erros
      1. Na coluna da variável *preoc\_ambiente* 🡪 Selecionar no filtro os valores acima de 100 🡪 eliminar conteúdo das células respetivas
      2. Na coluna da variável *idade* 🡪 Selecionar no filtro valores acima de 100 🡪 eliminar conteúdo das células respetivas
      3. Na coluna da variável *automovel* 🡪 Selecionar no filtro o valor “Comm Automóvel” 🡪 corrigir célula respetiva com “Com Automóvel”
5. **Criar regra de validação para duas variáveis – uma quantitativa e outra qualitativa:**
   1. Para a Variável Quantitativa (Preocupação Ambiental):

Selecionar coluna da variável *preoc\_ambiente* 🡪 Data Validation 🡪 Settings 🡪 Selecionar números inteiros de 0 a 100 🡪 Error Alert 🡪 Definir título para o erro e texto associado

* 1. Para a Variável Qualitativa (Posse de Automóvel):

Selecionar coluna da variável automovel 🡪 Data Validation 🡪 Settings 🡪 List 🡪 Source: ='Variáveis e códigos'!$E$16:$E$17 🡪 Error Alert 🡪 Definir título para o erro e texto associado; Definir “blank” como permitido

1. **Construir uma tabela de frequências absolutas com funções de contagem e cálculo de percentagens – Efeito de Estufa:**
   1. Criar página *Tabela de Frequências*, para análise de frequências sobre a variável *p8.13 (Efeito de estufa)*;
   2. Coluna C, criar linhas com respostas possíveis (Nada Importante, 2, 3, 4, Muito Importante) das linhas 3 a 7;
   3. Na coluna D criar a fórmula de n que apenas conta as linhas cujo valor da coluna de resposta corresponde ao da resposta possível =COUNTIF(Dados\_Eco1!H:H;'Tabela de Frequências'!C3) e copiar para as restantes linhas;
   4. Criar um total de respostas, na coluna D, linha 8 (célula D8) =SUM(D3:D7);
   5. Criar a coluna de percentagens de respostas na coluna E, com a fórmula = =(D3/$D$8) e copiar para as restantes linhas da tabela.
2. **Construir uma tabela de frequências dinâmica – Incêndios Florestais:**
   1. Confirmar que filtros das variáveis estão com todos os campos de resposta ativos;
   2. Insert 🡪 Pivot Table;
   3. Nomear nova folha com “Tabela de Frequências Dinâmicas”;
   4. Selecionar *p8.1* em Pivot Table Fields;
   5. Arrastar *p.8.1* para Values e garantir que *p8.1* está também em Rows;
   6. Clicar duas vezes no título da nova coluna gerada 🡪 Value Field Settings 🡪 Mudar o nome da coluna para *n*;
   7. Arrastar *p.8.1* para Values novamente;
   8. Clicar duas vezes no título da nova coluna gerada 🡪 Value Field Settings 🡪 Mudar o nome da coluna para *%* 🡪 Show Value As 🡪 % of Column Total;
   9. Com recurso ao filtro na tabela, retirar as respostas “99”;
   10. Ordenar as categorias de respostas, arrastando as células de acordo;
   11. Format as Table 🡪 Escolher uma visualização sem cores alternadas em linhas.
3. **Construir uma tabela de cruzamentos dinâmica (Habilitações Literárias, Sexo e Preocupação Ambiental):**
   1. Confirmar que filtros das variáveis estão com todos os campos de resposta ativos;
   2. Insert 🡪 Pivot Table;
   3. Nomear nova folha com “Tabela de Frequências Dinâmicas”;
   4. Selecionar *habilitacoestexto*, *sexotexto* e *preoc\_ambiente* em Pivot Table Fields;
   5. Arrastar *habilitacoestexto* para Columns para Values, *sexotexto* para Rows e *proc\_ambiente* para Values
   6. Mudar nome das colunas da tabela;
   7. Com recurso ao filtro na tabela, retirar as respostas “blank”;
   8. Renomear *Grand Total* para *Total* - Format as Table;
   9. Format as Table 🡪 Escolher uma visualização sem cores alternadas em linhas.

Além das tarefas acima descritas, procedemos à verificação de possíveis valores duplicados, através de:

Selecionar coluna de ID 🡪 Home 🡪 Conditional Formatting 🡪 Highlight Cell Rules 🡪 Duplicate Values

Contudo, não foram encontrados registos duplicados.

# II – Jamovi

Neste segmento, temos por objetivo analisar descritivamente, com recurso ao software Jamovi, as variáveis Preocupação Ambiental e Efeito de Estufa, relacionando também as duas – neste caso, detalhando a distribuição de inquiridos por categoria de resposta a Efeito de Estufa.

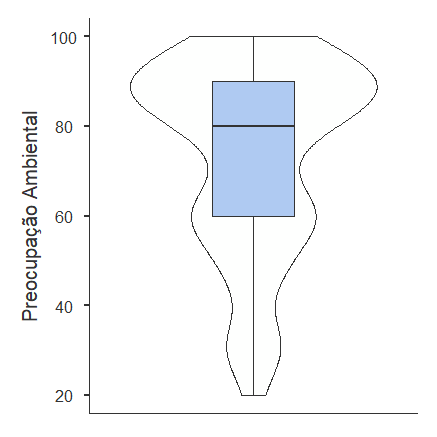
Inicialmente, começamos por gerar uma tabela com estatísticas descritivas da amostra e também das respostas à variável Preocupação Ambiental.

Para isto, as tarefas concretizadas foram:

1. Data 🡪 Selecionar coluna com variável *preoc\_ambiente*, com botão direito do rato 🡪 Setup 🡪Alterar título da variável e definir para *continuous* (Nota: a variável não é de origem contínua, mas sim quantitativa discreta; contudo, para permitir a representação com *boxplot* no programa, é necessário defini-la daquela forma);
2. Analysis 🡪 Exploration 🡪 Descriptives 🡪 Selecionar variável preoc\_ambiente e arrastar para Variables;
3. Selecionar as opções: N, Missing, Mean, Median, Mode, Std. Deviation, Variance, Minimum, Maximum, Quartiles, Boxplot e Violin (estas duas últimas para gerar um gráfico com a distribuição das observações).

Obtemos a tabela e gráfico seguintes:

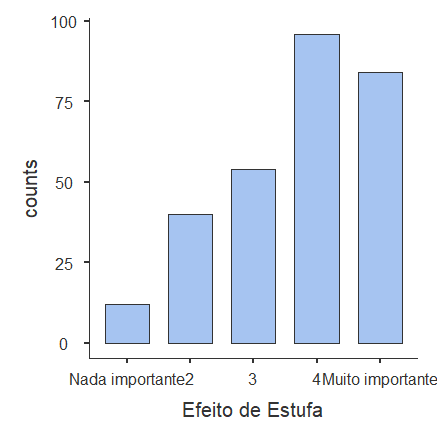
| Descritivos | | | |
| --- | --- | --- | --- |
|  |  |  |  |
|  | | **Preocupação Ambiental** | |
| N |  | 352 |  |
| Missing |  | 2 |  |
| Mean |  | 73.4 |  |
| Median |  | 80.0 |  |
| Mode |  | 90.0 |  |
| Standard deviation |  | 21.5 |  |
| Variance |  | 463 |  |
| Minimum |  | 20 |  |
| Maximum |  | 100 |  |
| 25th percentile |  | 60.0 |  |
| 50th percentile |  | 80.0 |  |
| 75th percentile |  | 90.0 |  |
|  | | | |



De seguida, geramos uma tabela com a distribuição de frequências (absoluta, relativa, e relativa acumulada) sobre a variável Efeito de Estufa, assim como um gráfico de barras com distribuição de frequências absolutas:

1. Data 🡪 Selecionar coluna com variável *p8.13*, com botão direito do rato 🡪 Setup 🡪Alterar título da variável e organizar por ordem as categorias de resposta em Levels;
2. Exploration 🡪Descriptives 🡪 Selecionar variável Efeito de Estufa e arrastar para Variables;
3. Selecionar as opções: Frequency Tables e Bar Plot.

| Frequencies of Efeito de Estufa | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| **Levels** | | **Counts** | | **% of Total** | | **Cumulative %** | |
| Nada importante |  | 12 |  | 4.2 % |  | 4.2 % |  |
| 2 |  | 40 |  | 14.0 % |  | 18.2 % |  |
| 3 |  | 54 |  | 18.9 % |  | 37.1 % |  |
| 4 |  | 96 |  | 33.6 % |  | 70.6 % |  |
| Muito importante |  | 84 |  | 29.4 % |  | 100.0 % |  |
|  | | | | | | | |



Finalmente, é interessante analisar a distribuição de respostas à variável Preocupação Ambiental, estruturando-as por categoria de resposta na variável Efeito de Estufa. Desta forma, é possível perceber, por exemplo, das pessoas que classificam o Efeito Estufa como Muito Importante como é que estas caracterizam o seu nível de Preocupação Ambiental.

1. Analysis 🡪 Exploration 🡪 Descriptives 🡪 Sselecionar variável Preocupação Ambiental e arrastar para Variables; Selecionar variável Efeito de Estufa e arrastar para Split By;
2. Selecionar as opções: N, Mean, Mode, Minimum, Maximum, Quartiles, Boxplot e Data.

Obtemos então o seguinte output:

| Descriptives | | | | | |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  | | **Efeito de Estufa** | | **Preocupação Ambiental** | |
| N |  | Nada importante |  | 12 |  |
|  |  | 2 |  | 40 |  |
|  |  | 3 |  | 54 |  |
|  |  | 4 |  | 95 |  |
|  |  | Muito importante |  | 84 |  |
| Mean |  | Nada importante |  | 64.6 |  |
|  |  | 2 |  | 79.8 |  |
|  |  | 3 |  | 71.8 |  |
|  |  | 4 |  | 73.4 |  |
|  |  | Muito importante |  | 73.3 |  |
| Mode |  | Nada importante |  | 30.0 |  |
|  |  | 2 |  | 100 |  |
|  |  | 3 |  | 85.0 |  |
|  |  | 4 |  | 90.0 |  |
|  |  | Muito importante |  | 90.0 |  |
| Minimum |  | Nada importante |  | 20 |  |
|  |  | 2 |  | 30 |  |
|  |  | 3 |  | 20 |  |
|  |  | 4 |  | 20 |  |
|  |  | Muito importante |  | 20 |  |
| Maximum |  | Nada importante |  | 100 |  |
|  |  | 2 |  | 100 |  |
|  |  | 3 |  | 100 |  |
|  |  | 4 |  | 100 |  |
|  |  | Muito importante |  | 100 |  |
| 25th percentile |  | Nada importante |  | 30.0 |  |
|  |  | 2 |  | 65.0 |  |
|  |  | 3 |  | 60.0 |  |
|  |  | 4 |  | 60.0 |  |
|  |  | Muito importante |  | 60.0 |  |
| 50th percentile |  | Nada importante |  | 77.5 |  |
|  |  | 2 |  | 90.0 |  |
|  |  | 3 |  | 77.5 |  |
|  |  | 4 |  | 80.0 |  |
|  |  | Muito importante |  | 80.0 |  |
| 75th percentile |  | Nada importante |  | 90.0 |  |
|  |  | 2 |  | 96.3 |  |
|  |  | 3 |  | 85.0 |  |
|  |  | 4 |  | 90.0 |  |
|  |  | Muito importante |  | 90.0 |  |
|  | | | | | |

