

# Ecologia

## Caso Prático 10



Grupo 10, CDA1

### **Docentes**

Madalena Ramos

Luís Junqueira

André Silvestre N°104532 | Diogo Catarino N°104745

Eduardo Silva N°104943 | Francisco Gomes N°104944


# CP10 Ecologia | Passo a Passo das Tarefas

## I. Excel

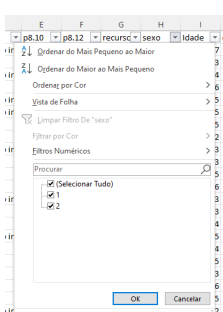
Antes de iniciar as tarefas propostas, de forma a salvaguardar e potenciar a análise estatística dos dados, decidimos efetuar algumas etapas de boas práticas no Excel, entre as quais, duplicámos o ficheiro de dados, de forma a trabalhar sobre a cópia e guardar sempre a fonte de dados original; retirámos as informações de contexto que podem estar a enquadrar as tabelas de dados/base de dados; identificámos na primeira linha o nome das variáveis de trabalho, e substituímo-las pelas etiquetas respetivas; fixámos a visualização da primeira linha, para ser mais prático ver os dados ao longo da folha; e aplicámos o filtro a todas as variáveis.

**1º.** Aplicar o filtro a todas as variáveis e apresentar a primeira linha fixa;

Selecionar a 1ª Linha – Base > Ordenar e Filtrar > Filtrar

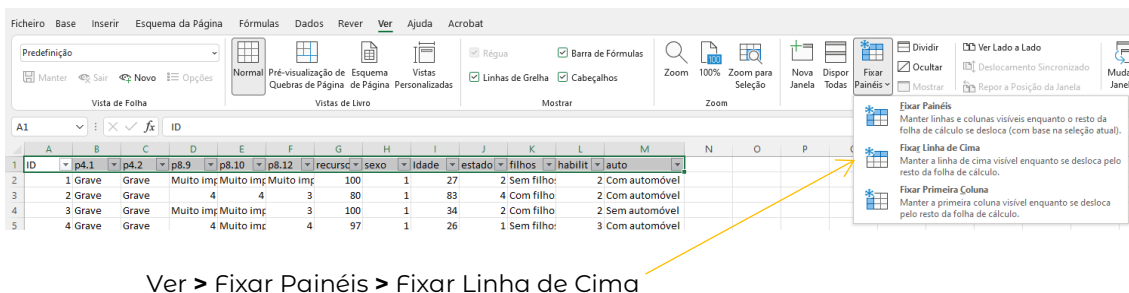


Em cada variável fica ativa uma seta que permite consultar e selecionar as categorias/valores diferentes existentes. Quando existe texto ou código para casos omissos, também aparece nesta lista.

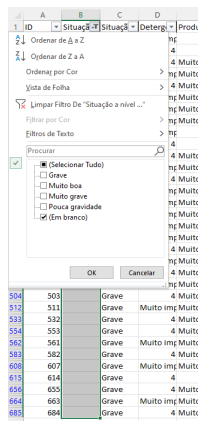


De modo a fixar a primeira linha:

Ver > Fixar Painéis > Fixar Linha de Cima

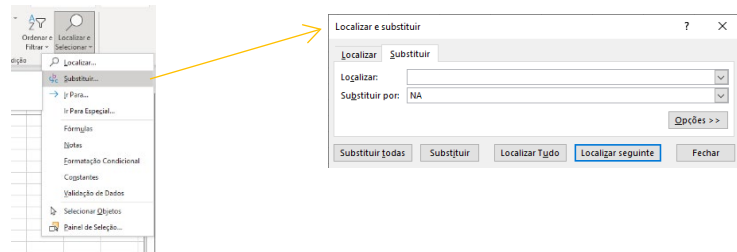


## 2º. Atribuir um código de não resposta a uma variável com valores omissos;

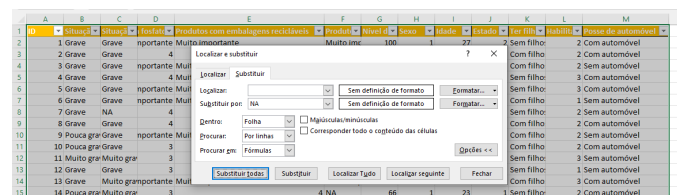


De modo a substituir os valores omissos de uma única variável, começamos por filtrar a coluna, selecionando apenas os valores “(Em branco)”.

Posteriormente: Base > Localizar e seleccionar > Substituir



A fim de facilitar o uso deste Excel no R, decidimos substituir os valores omissos em todas as variáveis. Para isso seleccionamos o conjunto de dados e repetimos o procedimento anterior.



## 3º. Atribuir texto às categorias de uma variável que tenha códigos;

Nas variáveis “Estado Civil”, “Habilitações Literárias” e “Sexo” estão presentes os códigos e não os nomes respetivos e, uma vez que existe uma tabela de codificação de correspondência na folha “Variáveis e Códigos”, é possível proceder à descodificação.

No caso do “Estado Civil”, como a tabela de correspondência está na vertical, a função que usámos foi a **PROCV()**, enquanto que nas “Habilitações Literárias” e do “Sexo” a tabela de correspondência encontra-se na horizontal pelo o que usámos a função **PROCH()**.

Para tal, inserimos uma coluna à direita de cada uma das variáveis, clicámos na primeira célula e inserimos a fórmula com a função correspondente, sendo que os atributos colocados em ambas foram (célula com o código que se pretende substituir; seleção da tabela de correspondências sem títulos e com as células fixas ; nº da coluna da tabela onde deve ser lida a correspondência; FALSO -correspondência exata)

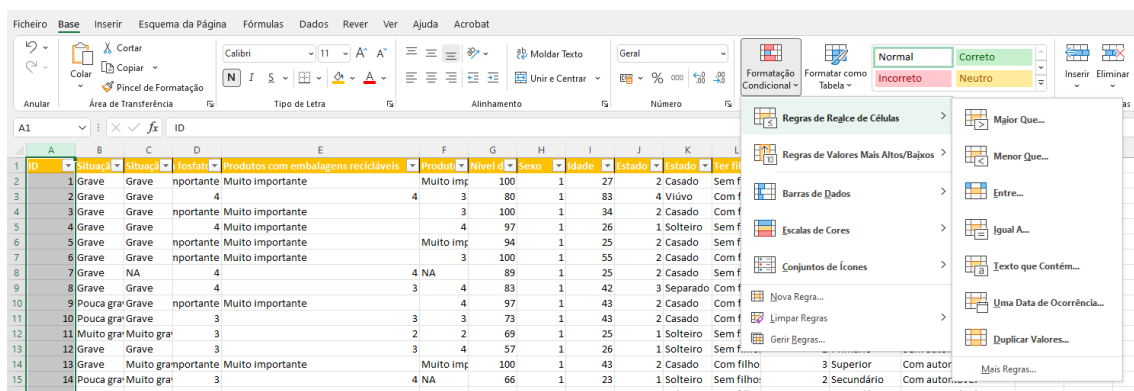
De seguida, copiamos a fórmula para as restantes células da coluna e atribuímos o novo nome à nova variável (coluna).

J	K	L	M	N	O
Estado	Estado	Ter filho	Habilitação	Habilitaç	Posse de automóvel
27	2	=PROCV(J2;Variáveis e códigos!\$D\$7:\$E\$10;2;FALSO)			
33	4	=PROCV(valor_pro; matriz_tabela; núm_índice_coluna; [procurar_intervalo])			
34	2	Com filho	2	Sem automóvel	

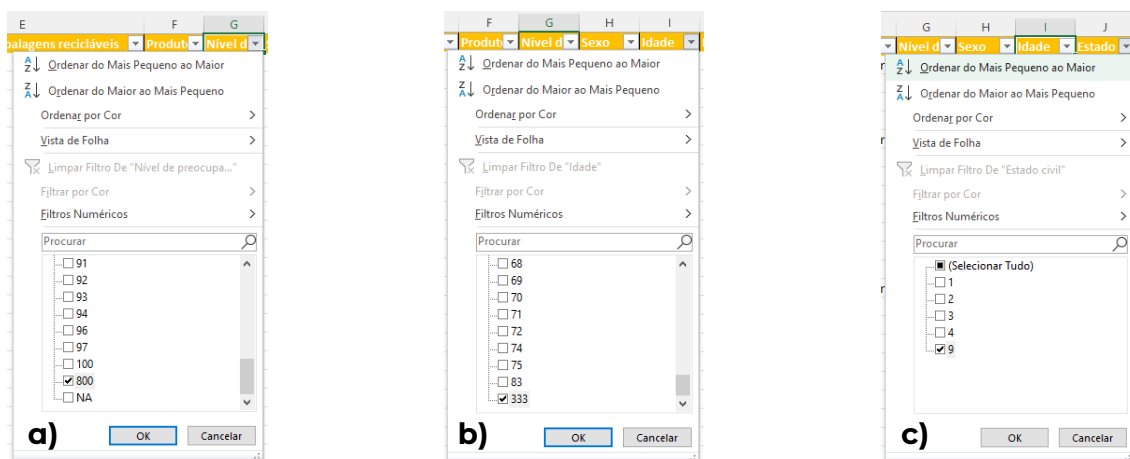
M	N	O	P	Q	R
Habilitação	Habilitações L	Posse de automóvel			
2	=PROCH(M2;Variáveis e Códigos!\$D\$19:\$F\$20;2;FALSO)				
2	=PROCH(valor_pro; matriz_tabela; núm_índice_linha; [procurar_intervalo])				
2	Secundário	Sem automóvel			

H	I	J	K	L
Sexo	Sexo_Texto	Idade	Estado civil	Estado C
1	=PROCH(H2;Variáveis e Códigos!\$E\$3:\$F\$4;2;FALSO)			
1	=PROCH(valor_pro; matriz_tabela; núm_índice_linha; [procurar_intervalo])			

#### 4º. Corrigir possíveis erros;



Inicialmente de forma a verificar que todos os inquéritos respondidos são únicos, averiguamos que não havia linhas duplicadas uma vez que nenhum dos ID's se repetia. Para constatar este facto, seleccionamos a coluna dos ID's e usamos a Regra de Formatação Condicional de Realce de Células “Duplicar Valores...”.



Seguidamente, através da utilização dos filtros e da análise generalizada aos dados que nos foram cedidos, verificou-se se as categorias/valores fazem sentido de acordo com a variável, e detetámos 3 erros:

**a)** o valor 800 no “Nível de preocupação com a destruição de recursos naturais”, uma vez que a escala aplicada é de 0 a 100. Sendo assim o valor inserido não faz sentido, e consideramos que o valor correto seria 80;

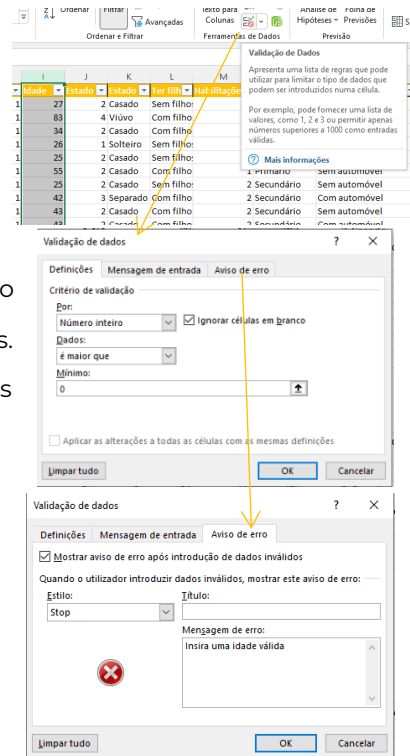
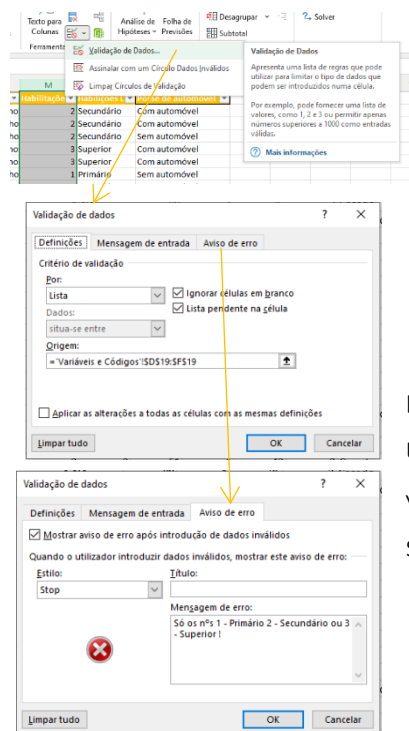
**b)** os valores 333 na “Idade”, e tal como no anterior, consideramos que o registo foi mal feito e que o valor correto seria o 33;

**c)** os valores 9 na variável no “Estado Civil”, sendo que os códigos apenas contemplam os números 1 a 4. Neste caso, optamos por converter os valores em NA (não respostas), dado que não é possível corrigir este valor.

## 5º. Criar regra de validação para duas variáveis: uma quantitativa e outra qualitativa;

Existe a possibilidade de limitar os valores/texto a inserir ou depois de inseridos, visualizar as situações de erro. Para tal, seleccionamos a coluna das variáveis escolhidas ("Idade" – variável quantitativa - e "Habilitações Literárias" – variável qualitativa).

Primeiro, seleccionámos a coluna da "Idade", fomos ao separador Dados > Validação de Dados > Definições. Limitámos os *inputs* a n°s maiores do que 0 e para valores inferiores um Aviso de Erro "Insira uma idade válida".



Já nas "Habilitações Literárias", realizámos o mesmo procedimento, mas limitámos as respostas aos valores 1 a 3 tal como é o código desta variável. Caso seja introduzido um valor diferente, apresenta o aviso: "Só os n°s 1 - Primário 2 - Secundário ou 3 - Superior !".

## 6º. Construir uma tabela de frequências absolutas com funções de contagem e cálculo de percentagens;

A fim de contruir uma tabela de frequências absolutas, criámos uma nova folha de cálculo denominada “Tabelas de Frequências”; copiámos a lista de categorias da folha "Variáveis e Códigos"; posteriormente, na célula à direita de cada categoria calculámos as contagens condicionadas, frequências absolutas com a função

A	B	C	D
<b>Género</b>	<b>n</b>		
Masculino	=CONTAR.SE(Dados!\$H\$2:\$H\$697;1)		
Feminino	CONTAR.SE(intervalo; critérios)		

**CONTAR.SE()**, fixámos as células correspondentes à coluna de dados a contar (“Sexo”) e copiámos a fórmula para as células restantes.

De seguida, criámos uma última linha com o Total – função **SOMA()**.

Para calcular as frequências absolutas, criámos uma coluna à direita com o cálculo da %, tal como ilustrado, e colocámos o estilo de percentagem na célula.

A	B	C
<b>Género</b>	<b>n</b>	
Masculino	390	
Feminino	306	
NA	0	
<b>Total</b>	=SOMA(B2:B4)	

A	B	C
<b>Género</b>	<b>n</b>	<b>%</b>
Masculino	390	=B2/\$B\$5
Feminino	306	44%
NA	0	0
<b>Total</b>	<b>696</b>	<b>100%</b>

Após calcular, e de modo a tornar a tabela de fácil leitura decidimos formatar os cabeçalhos e linhas envolventes de forma automática com a função de “Formatar como Tabela”.

The screenshot shows the Excel interface with the 'Formatar como Tabela' dialog box open. The dialog box has a 'Criar Tabela' button and a 'Formatar como Tabela' button. The 'Formatar como Tabela' button is highlighted, and a dropdown menu is visible showing the 'Estilo de Percentagem' (Ctrl+Shift+%) option. The dropdown menu also includes the text 'Formatar como percentagem.'.

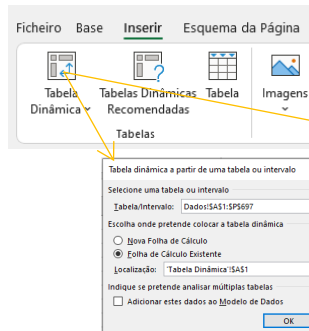
Repetimos estes procedimentos para as restantes variáveis que refletem a caracterização dos inquiridos da base de dados em estudo.

A	B	C	D	E	F	G	H	I
<b>Género</b>	<b>n</b>	<b>%</b>			<b>Habilitações Literár</b>	<b>n</b>	<b>%</b>	<b>% acum</b>
Masculino	390	56%			Primário	264	38%	38%
Feminino	306	44%			Secundário	257	37%	75%
NA	0	0			Superior	175	25%	100%
<b>Total</b>	<b>696</b>	<b>100%</b>			NA	0	0	100%
					<b>Total</b>	<b>696</b>	<b>100%</b>	
<b>Filhos</b>	<b>n</b>	<b>%</b>						
Sim	379	54%			<b>Estado Civil</b>	<b>n</b>	<b>%</b>	<b>% acum</b>
Não	303	44%			Solteiro	254	36%	36%
NA	14	2%			Casado	374	54%	90%
<b>Total</b>	<b>696</b>	<b>100%</b>			Separado	34	5%	95%
					Viúvo	15	2%	97%
<b>Automóvel</b>	<b>n</b>	<b>%</b>			NA	19	3%	100%
Sim	390	56%			<b>Total</b>	<b>696</b>	<b>100%</b>	
Não	306	44%						
NA	0	0%						
<b>Total</b>	<b>696</b>	<b>100%</b>						

## 7º. Construir uma tabela de frequências dinâmica;

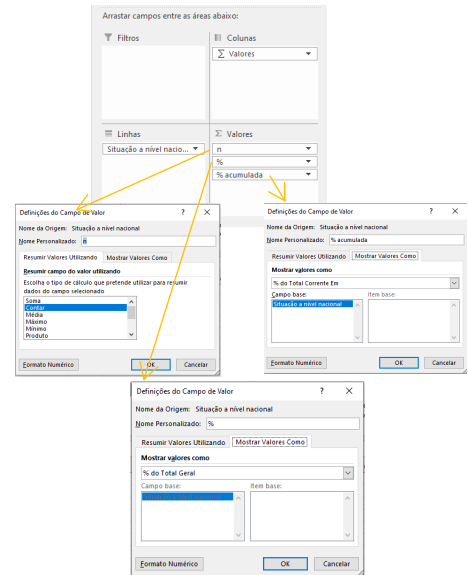
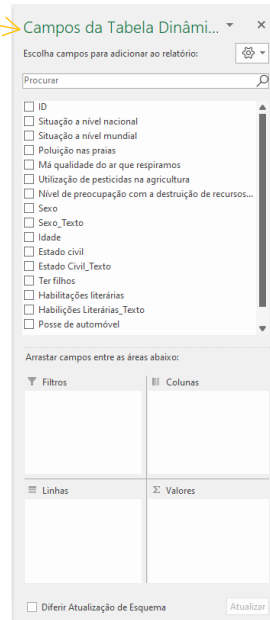
Para construir uma tabela dinâmica começamos por criar uma nova folha de cálculo designada “Tabelas Dinâmicas”.

**Tabela Dinâmica**



De seguida, Menu Inserir > Tabela Dinâmica > Seleccionar toda a tabela da folha “Dados”

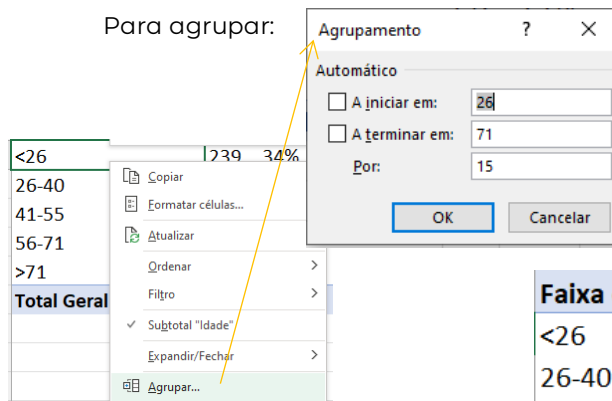
Seleccionámos a variável que queremos fazer a tabela (“Situação a Nível Nacional”); acrescentámos as colunas de frequências absolutas (%) e de % acumulada; e substituímos os títulos das colunas.



Gravidade da Situação	n	%	% acumulada
Muito boa	6	1%	1%
Pouca gravidade	147	21%	22%
Grave	368	53%	75%
Muito grave	142	20%	95%
NA	33	5%	100%
<b>Total Geral</b>	<b>696</b>	<b>100%</b>	

Seguidamente, realizámos de novo todos os passos anteriormente referidos, mas para construir uma tabela com a variável “Idade” de modo a agrupá-la, permitindo assim uma melhor análise desta variável de caracterização social dos inquiridos.

Para agrupar:

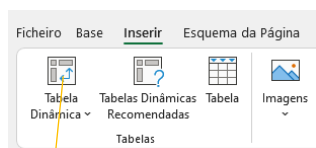


Faixa etária	n	%	% acumulada
<26	239	34%	34%
26-40	132	19%	53%
41-55	147	21%	74%
56-71	163	23%	98%
>71	15	2%	100%
<b>Total Geral</b>	<b>696</b>	<b>100%</b>	

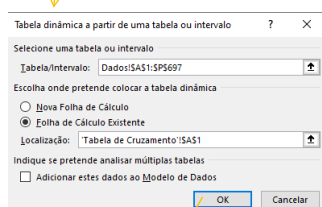
## 8º. Construir uma tabela de cruzamentos dinâmica;

Tal como na tarefa anterior, começámos por criar uma nova folha de cálculo designada “Tabela de Cruzamento”.

### Tabela de Cruzamento

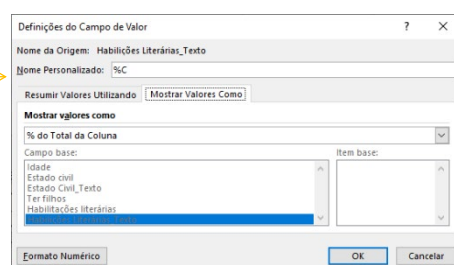
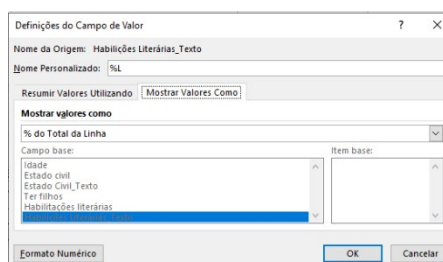
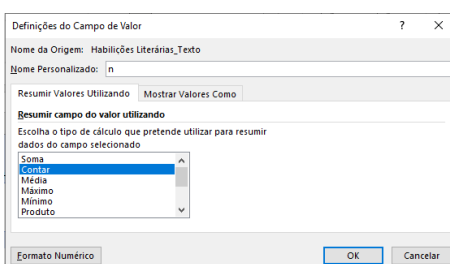
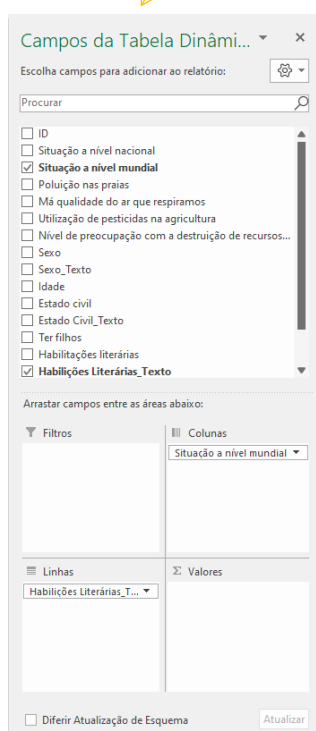


De seguida, Menu Inserir > Tabela Dinâmica > Selecionar toda a tabela da folha “Dados”



Decidimos as variáveis que queríamos cruzar ( Variável em linha – “Habilitações Literárias”; Variável em coluna – “Situação a Nível Mundial ”) e no campo “ Valores” o n, % em linha e % em coluna.

Σ Valores
n
%L
%C



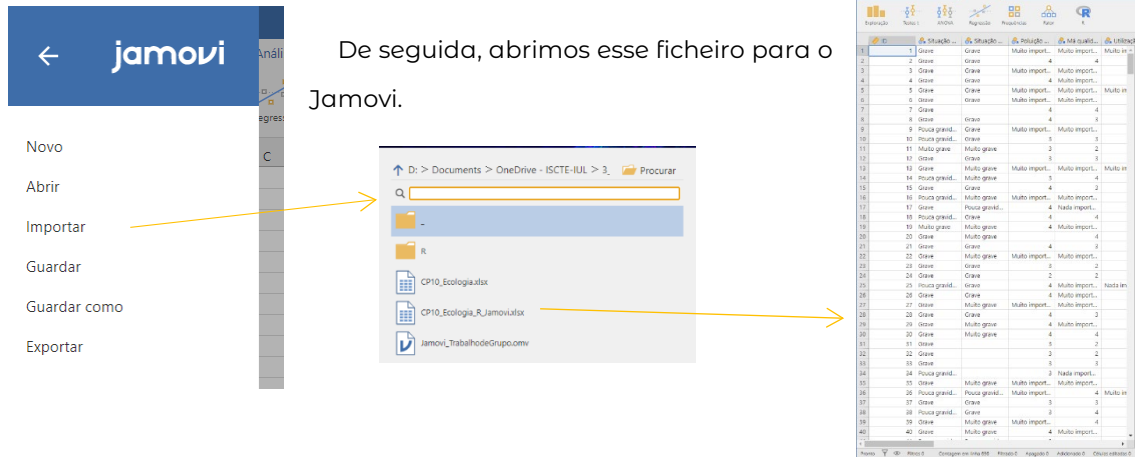
Alterámos os Rótulos das Colunas e de Linhas para os nomes das variáveis em estudo e colocámos 1 casa decimal nas %.

Situação a Nível Mundial														Total		
Muito grave		Grave		Pouca gravidade		NA										
Habilitações Literárias	n	%L	%C	n	%L	%C	n	%L	%C	n	%L	%C	n	Total n	Total %L	Total %C
Primário	138	52,3%	42,7%	113	42,8%	35,0%		0,0%	0,0%	13	4,9%	43,3%	264	100,0%	37,9%	
Secundário	91	35,4%	28,2%	137	53,3%	42,4%	12	4,7%	60,0%	17	6,6%	56,7%	257	100,0%	36,9%	
Superior	94	53,7%	29,1%	73	41,7%	22,6%	8	4,6%	40,0%	0,0%	0,0%	0,0%	175	100,0%	25,1%	
<b>Total Geral</b>	<b>323</b>	<b>46,4%</b>	<b>100,0%</b>	<b>323</b>	<b>46,4%</b>	<b>100,0%</b>	<b>20</b>	<b>2,9%</b>	<b>100,0%</b>	<b>30</b>	<b>4,3%</b>	<b>100,0%</b>	<b>696</b>	<b>100,0%</b>	<b>100,0%</b>	

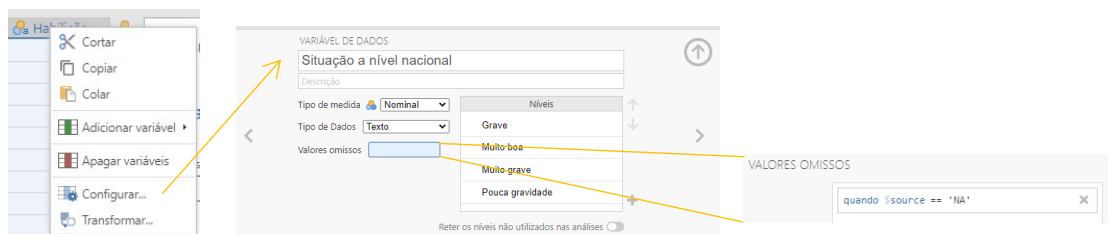


## II. Jamovi

Para realizar as tarefas propostas no *software* Jamovi e posteriormente no R começámos por criar um Excel exclusivamente com a 1ª folha de cálculo dos dados anteriormente trabalhados.



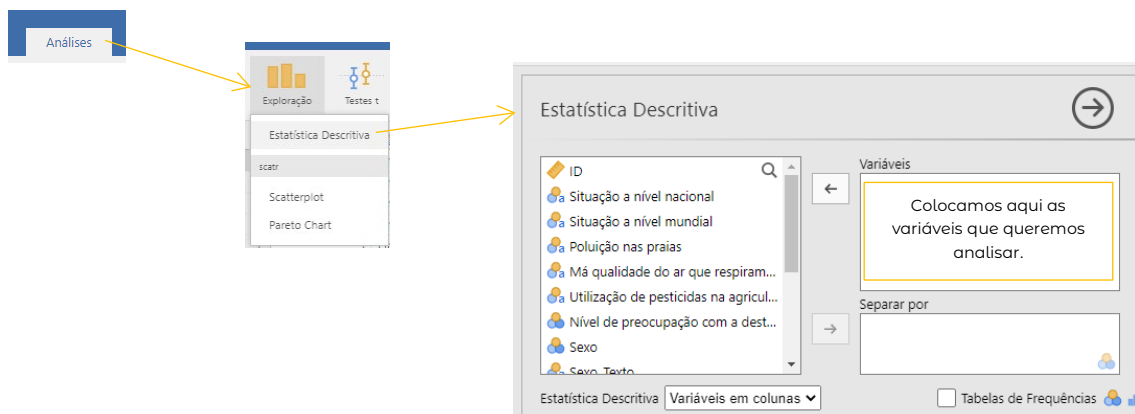
Uma vez que os valores omissos no Jamovi são considerados automaticamente caso a célula esteja em branco, mas no Excel criado anteriormente as não respostas estão como “NA”, tivemos de configurar todas as colunas para detetarem as não respostas



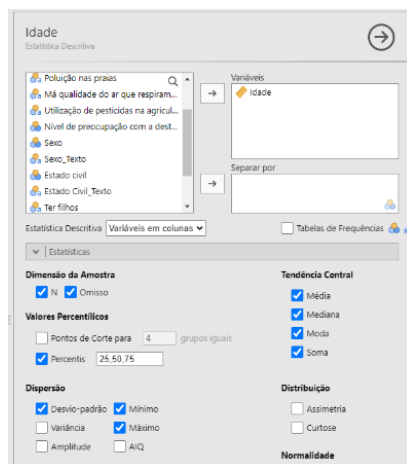
❖ Construir de forma automática e com recurso ao módulo R:

- 3 tabelas com medidas descritivas (com 1 e com 2 variáveis)

Para construir as tabelas Análises > Exploração > Estatísticas Descritivas



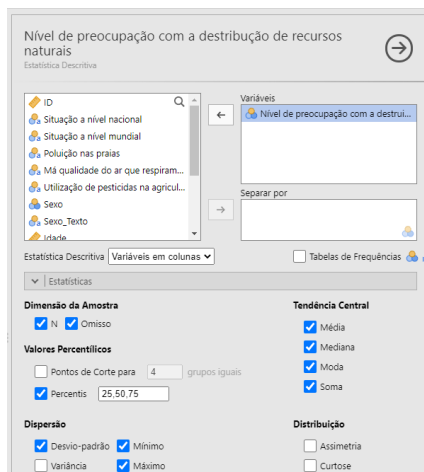
i. Tabela das Medidas Descritivas da “**Idade**”



**Idade**

Estadística Descritiva	
Idade	
N	696
Omisso	0
Média	39.8
Mediana	37.0
Moda	24.0
Soma	27705
Desvio-padrão	17.2
Mínimo	18
Máximo	83
25º percentil	24.0
50º percentil	37.0
75º percentil	56.0

ii. Tabela das Medidas Descritivas do “**Nível de preocupação com a destruição de recursos naturais (0-100)**”



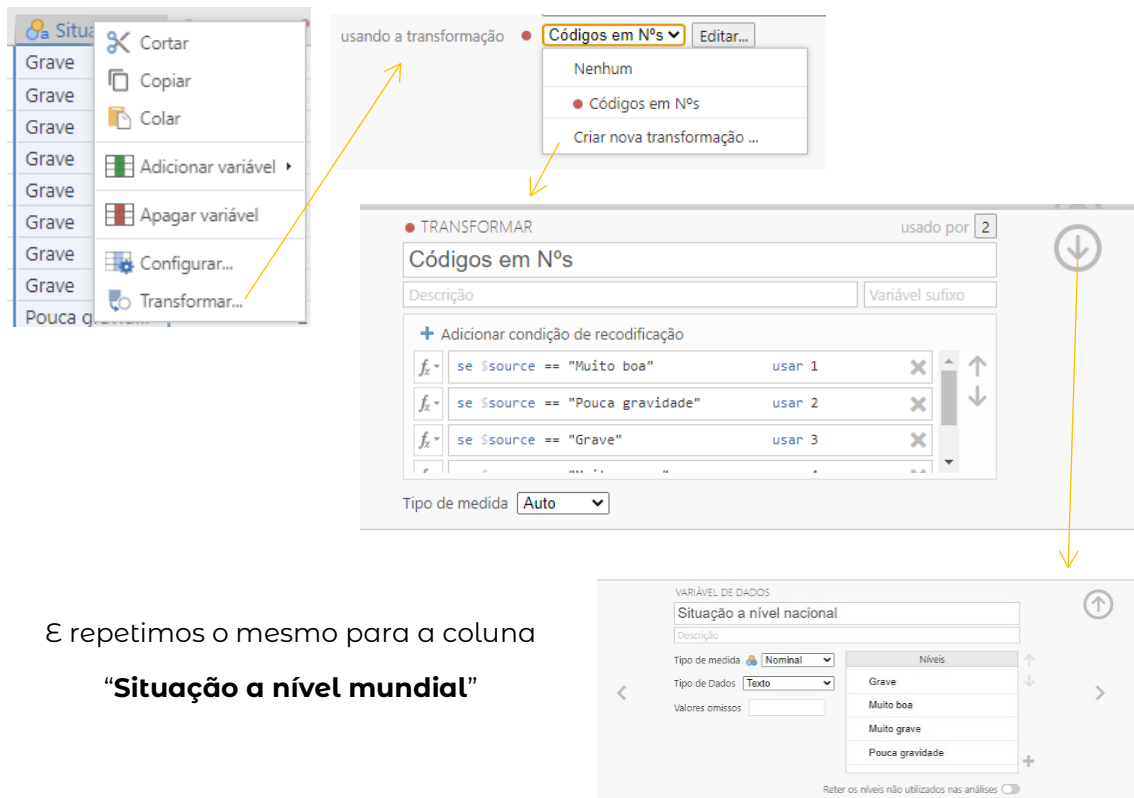
**Nível de preocupação com a destruição de recursos naturais**

Estadística Descritiva	
Nível de preocupação com a destruição de recursos naturais (0-100)	
N	690
Omisso	6
Média	85.4
Mediana	86.5
Moda	100
Soma	58931
Desvio-padrão	12.9
Mínimo	20
Máximo	100
25º percentil	77.0
50º percentil	86.5
75º percentil	97.0

iii. Tabela das Medidas Descritivas do “**Nível Nacional e Mundial**”

Para transformar esta variável qualitativa em quantitativa, de modo a calcular as Medidas Descritivas usamos as potencialidades do Jamovi.

Selecionar a coluna “Situação a nível nacional” > Transformar > Criar Usando a transformação > Criar Nova Transformação > Adicionar condição de Recodificação. Este último passo repete-se tantas vezes quantas designações diferentes a nossa escala possuir ou quantas substituições pretendermos fazer.



E repetimos o mesmo para a coluna  
**“Situação a nível mundial”**

Posteriormente, criámos duas tabelas com as duas variáveis em estudo, sendo a 1ª com os códigos em n°s e a 2ª com os códigos em texto.

### Situação a Nível Nacional e Mundial

#### Estatística Descritiva

	Situação a nível mundial	Situação a nível nacional
N	666	663
Omisso	30	33

### Situação a Nível Nacional e Mundial

#### Estatística Descritiva

	Situação a nível nacional - N°s	Situação a nível mundial - N°s
N	663	666
Omisso	33	30
Mediana	3	3.00
Moda	3.00	3.00 *
25º percentil	3.00	3.00
50º percentil	3.00	3.00
75º percentil	3.00	4.00

\* Existe mais de uma moda, apenas a primeira é apresentada

#### Frequências

##### Frequências de \$Key

Níveis	Contagens	% do Total	% acumulada
Grave	368	55.5 %	55.5 %
Muito boa	6	0.9 %	56.4 %
Muito grave	142	21.4 %	77.8 %
Pouca gravidade	147	22.2 %	100.0 %

##### Frequências de \$Key

Níveis	Contagens	% do Total	% acumulada
Grave	323	48.5 %	48.5 %
Muito grave	323	48.5 %	97.0 %
Pouca gravidade	20	3.0 %	100.0 %

- 1 gráfico com cruzamento de variáveis

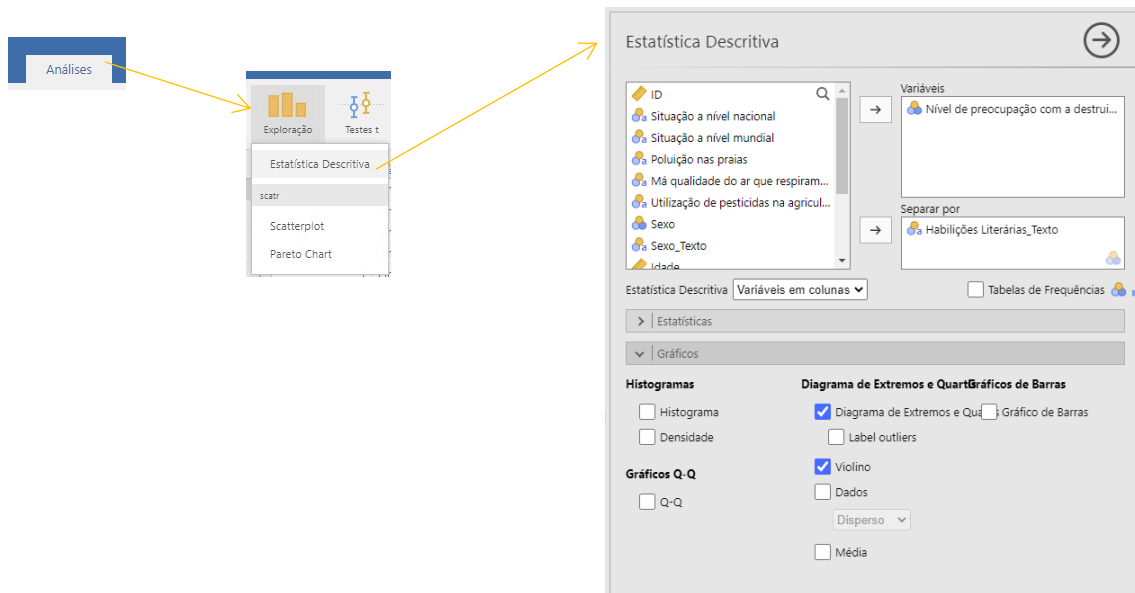


Gráfico de Violinos e de Extremos e Quartis do **“Habilidades Literárias”** em função do **“Nível de Preocupação com a Destruição de Recursos Naturais”**

Nível de preocupação com a destruição de recursos naturais (0-100)

