

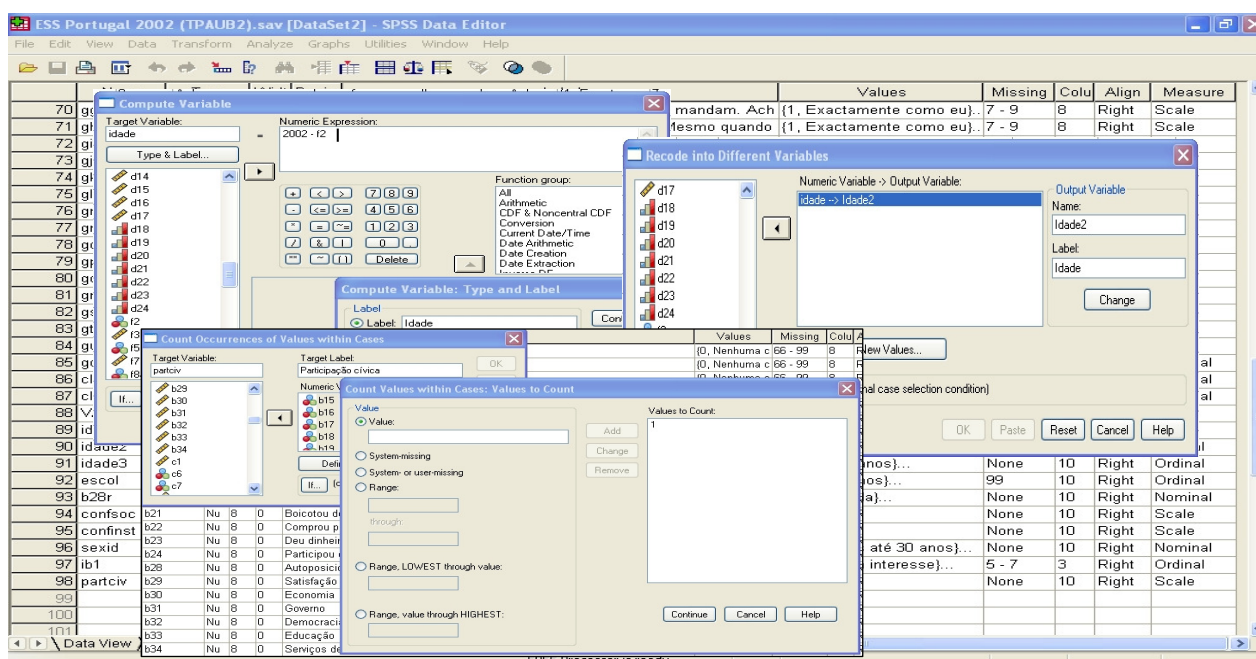


Instituto Superior de Economia e Gestão
UNIVERSIDADE TÉCNICA DE LISBOA

Mestrado em Marketing

Análise Quantitativa de Dados em Marketing

Breve introdução ao SPSS® Tratamento Preliminar de Dados (o essencial)



Docente: **Rui Brites**
e-mail: rui.brites@iscte.pt
Telef.: 217903485/969073534
Skype Name: britesrui

Setembro 2010

ÍNDICE

1. Introdução ao SPSS	1
1.1. Editor de dados	2
1.2. Outputs do SPSS (<i>Output Viewer</i>)	4
1.3. Menus do SPSS	5
1.4. Criação de bases de dados	8
2. Tratamento preliminar de dados	9
2.1. Transformação de variáveis	9
2.1.1. Recodificação de variáveis	9
2.1.2. Inversão da escala de uma variável	14
2.1.3. Transformação lógica de variáveis.....	15
3. Criação de Índices sintéticos	16
3.1. Fiabilidade dos índices	16
3.2. Média dos indicadores	17
3.3. Contagem de ocorrências (adição)	19
4. Selecção de casos	20
4.1. Selecção de uma sub-amostra	20
4.2. Selecção de uma amostra aleatória simples	21
4.3. Seleccionar uma amostra aleatória com nº de casos pré-definidos	21
5. Separar a análise por sub-amostras	22
Links úteis relacionados com a matéria deste manual	23
Anexo 1: Construção de um ponderador para a variável <i>sexo</i>	24
Anexo 2: Extracto do Questionário do exemplo (<i>ESS, round 1, 2002</i>)	29

Breve Introdução ao SPSS®

1. Introdução ao SPSS

O *package* estatístico *SPSS para Windows* (**Statistical Package for Social Sciences**)¹ é um poderoso sistema de análises estatísticas e manuseamento de dados, em que a utilização mais frequente, para a maioria das análises a efectuar, se resume à selecção das respectivas opções em menus e caixas de diálogo

É comercializado em Portugal pela PSE: <http://www.pse.pt/>, sendo possível obter uma Versão DEMO, com duração de 14 dias em <http://www.spss.com/>.

1.1. Editor de dados do SPSS

O editor de dados do SPSS (*Data Editor*) é composto por duas janelas sobrepostas: *Data View* e *Variable View*. A função da primeira – *Data View* – é a de introduzir os dados e da segunda - *Variable View* – é criar a estrutura da base de dados.

Muda-se de uma para outra clicando no respectivo separador.

O *Data Editor* do SPSS é um programa do tipo de folha de cálculo que permite facilmente criar ou editar ficheiros de dados. Abre automaticamente quando se entra no SPSS.

O seu aspecto é o seguinte:

¹ Tendo sido recentemente adquirido pela IBM, foi comercializado nas versões 17 e 18 como *PASW* (*Predictive Analyse Software*), tendo na versão 18 retomado a designação anterior precedida do nome da empresa: *IBM SPSS*

Janela *Data View*:

	n	pondsexo	pondsexo2	regiao	a8	a9	a10	b1	b2	b3
1	35787	1,160	7937	3	3	4	3	4	5	1
2	35794	1,160	7937	3	3	6	2	4	4	1
3	35795	1,160	7937	3	7	7	6	4	5	1
4	35797	1,160	7937	3	3	8	5	4	5	1
5	35800	1,160	7937	3	3	5	7	3	4	1
6	35805	1,160	7937	3	3	3	5	3	4	1
7	35811	1,160	7937	3	6	7	7	2	5	2
8	35814	1,160	7937	3	5	6	7	2	3	2
9	35815	1,160	7937	3	2	3	2	2	5	2
10	35851	1,160	7937	3	5	8	3	3	4	2
11	35853	1,160	7937	3	5	4	4	1	5	4
12	35872	1,160	7937	3	6	6	5	3	1	1
13	35883	1,160	7937	3	5	10	8	4	5	2
14	35893	1,160	7937	3	8	5	5	2	3	4
15	35896	1,160	7937	3	4	7	4	4	4	1
16	35908	1,160	7937	3	0	0	8	3	3	1
17	35914	1,160	7937	3	6	4	4	2	3	1
18	35941	1,160	7937	3	5	6	5	2	3	2
19	35944	1,160	7937	3	6	6	5	2	3	2

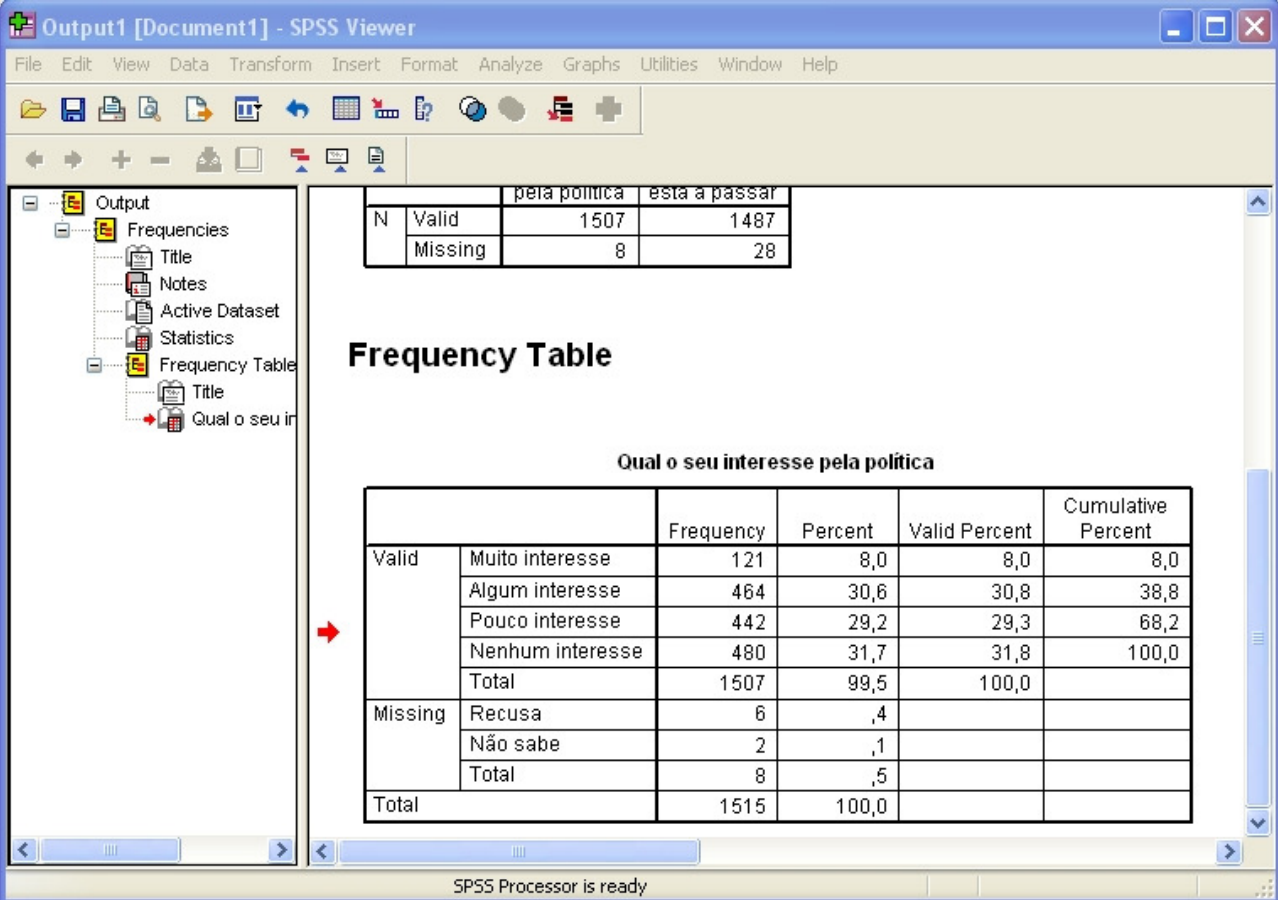
Janela *Variable View*:

	Name	Type	Width	Deci	Label	Values	Missing	Colu	A	Measure
1	n	Numeric	8	0		None	None	8		R Scale
2	pondsexo	Numeric	8	3	Ponderador para corrigir a amostra por sexo	None	None	10		R Scale
3	pondsexo2	Numeric	8	0	Ponderador para corrigir a amostra por sexo e	None	None	11		R Scale
4	regiao	Numeric	8	0	Região	{1, Norte}...	99	8		R Nominal
5	a8	Numeric	2	0	Confiança interpessoal	{0, Todo o cuidad	66 - 99	8		R Scale
6	a9	Numeric	2	0	Confiança na honestidade dos outros	{0, Tentam aprobe	66 - 99	8		R Scale
7	a10	Numeric	2	0	Confiança no altruísmo dos outros	{0, As pessoas e	66 - 99	8		R Scale
8	b1	Numeric	1	0	Qual o seu interesse pela política	{1, Muito interess	6 - 9	8		R Ordinal
9	b2	Numeric	8	0	Com que frequência a política lhe parece tão c	{1, Nunca}...	6 - 9	8		R Ordinal
10	b3	Numeric	8	0	Acha que podia participar num grupo dedicado	{1, De certeza qu	6 - 9	8		R Ordinal
11	b4	Numeric	8	0	Grau de dificuldade que sente em tomar uma	{1, É muito difícil	6 - 9	8		R Ordinal
12	b5	Numeric	8	0	Acha que os políticos, em geral, se interessa	{1, Quase nenhun	6 - 9	8		R Ordinal
13	b6	Numeric	8	0	Os políticos estão mais interessados em ganh	{1, Quase todos e	6 - 9	8		R Ordinal
14	b7	Numeric	8	0	Confiança no Parlamento	{0, Nenhuma conf	66 - 99	8		R Scale
15	b8	Numeric	8	0	Confiança no Sistema Jurídico	{0, Nenhuma conf	66 - 99	8		R Scale
16	b9	Numeric	8	0	Confiança na Polícia	{0, Nenhuma conf	66 - 99	8		R Scale
17	b10	Numeric	8	0	Confiança nos Políticos	{0, Nenhuma conf	66 - 99	8		R Scale
18	b11	Numeric	8	0	Confiança no Parlamento Europeu	{0, Nenhuma conf	66 - 99	8		R Scale
19	b12	Numeric	8	0	Confiança nas Nações Unidas	{0, Nenhuma conf	66 - 99	8		R Scale
20	b15	Numeric	8	0	Contactou um político ou representante do gov	{1, Sim}...	6 - 9	8		R Nominal

1.2. Output do SPSS (*Output Viewer*)

É nesta janela que são apresentados todos os resultados estatísticos. Abre automaticamente sempre que um determinado procedimento gera resultados. É possível editar as tabelas e gráficos produzidos, clicando duas vezes com a tecla esquerda do rato e modificar a sua aparência. Para copiar os outputs para o Word, basta seleccionar e fazer copy e colar (paste) no Word².

Janela ***Output Viewer***:



The screenshot shows the SPSS Output Viewer window titled "Output1 [Document1] - SPSS Viewer". The window contains a tree view on the left with "Output" expanded, showing "Frequencies", "Title", "Notes", "Active Dataset", "Statistics", "Frequency Table", "Title", and "Qual o seu in". The main area displays a frequency table for "Qual o seu interesse pela política".

		pela política	esta a passar
N	Valid	1507	1487
	Missing	8	28

Frequency Table

Qual o seu interesse pela política

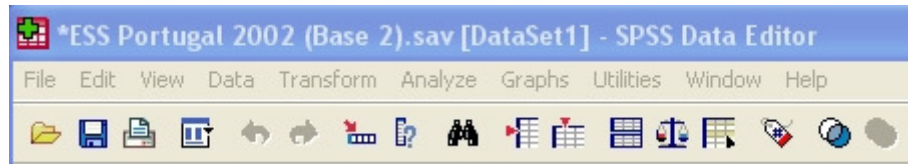
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Muito interesse	121	8,0	8,0	8,0
	Algum interesse	464	30,6	30,8	38,8
	Pouco interesse	442	29,2	29,3	68,2
	Nenhum interesse	480	31,7	31,8	100,0
	Total	1507	99,5	100,0	
Missing	Recusa	6	,4		
	Não sabe	2	,1		
	Total	8	,5		
Total		1515	100,0		

SPSS Processor is ready

² Até à versão 15, inclusive, para além da opção "copy" (texto simples), está também disponível a opção "copy objects" que permite copiar os quadros em formato imagem, que é possível redimensionar no Word (em diagonal, para manter a mesma relação entre altura e largura). Desde a versão 16 só está disponível a opção "copy". Quando os quadros são maiores do que as dimensões da página de texto, podem-se dimensionar na opção "Table/AutoFit/AutoFit to Window".

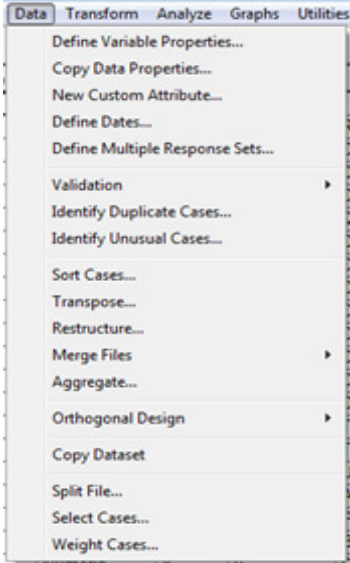
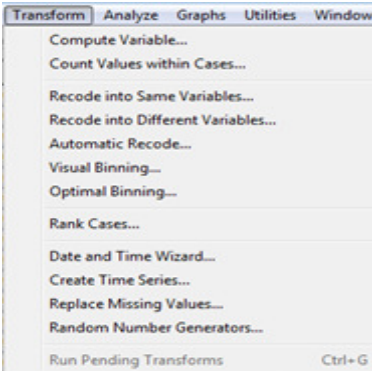
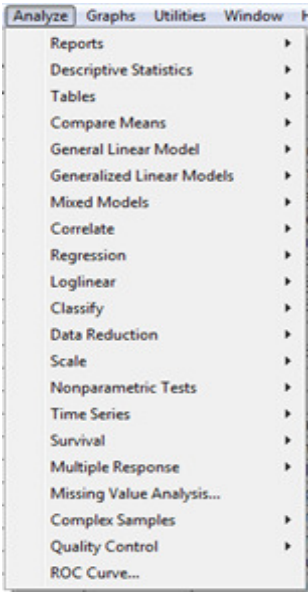
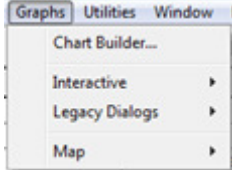
1.3. Menus do SPSS

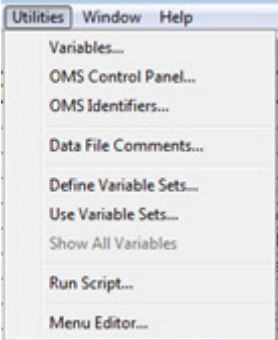
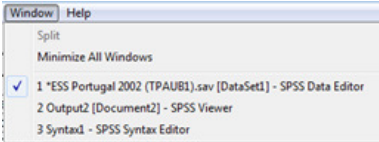
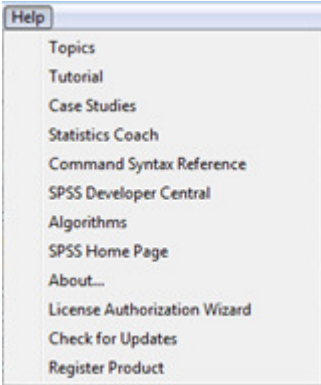
Os menus das janelas principais – *Data Editor* e *Outpur Viewer* - são idênticos e têm o seguinte aspecto visual:



Principais funcionalidades dos menus da versão 15.0

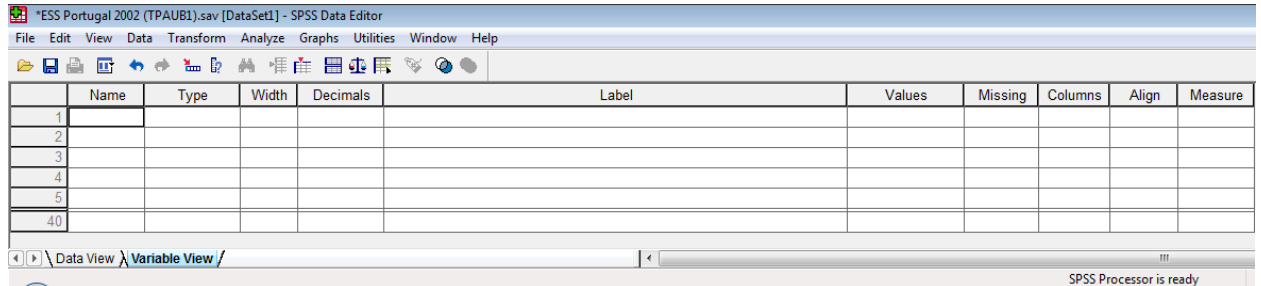
<p>File</p>	<p>The image shows the 'File' menu with options: New, Open, Open Database, Read Text Data..., Close (Ctrl+F4), Save (Ctrl+S), Save As..., Save All Data, Export to Database..., Mark File Read Only, Rename Dataset..., Display Data File Information, Cache Data..., Stop Processor (Ctrl+.), Switch Server..., Connect To Repository..., Print Preview, Print... (Ctrl+P), Recently Used Data, Recently Used Files, and Exit.</p>	<p>Criar, abrir, ler, exportar, gravar e imprimir ficheiros.</p>
<p>Edit</p>	<p>The image shows the 'Edit' menu with options: New, Open, Open Database, Read Text Data..., Close (Ctrl+F4), Save (Ctrl+S), Save As..., Save All Data, Export to Database..., Mark File Read Only, Rename Dataset..., Display Data File Information, Cache Data..., Stop Processor (Ctrl+.), Switch Server..., Connect To Repository..., Print Preview, Print... (Ctrl+P), Recently Used Data, Recently Used Files, and Exit.</p>	<p>Configuração/parametrização do SPSS (<i>Options</i>), inserir novas variáveis e novos casos.</p>
<p>View</p>	<p>The image shows the 'View' menu with options: Status Bar, Toolbars..., Fonts..., Grid Lines, Value Labels, Display Custom Attributes..., and Data (Ctrl+T).</p>	<p>Activar/desactivar barras de comandos, fontes, grelha, barra de status e mostrar etiquetas (<i>labels</i>) definidas.</p>

<p>Data</p>	 <p>The screenshot shows the 'Data' menu in SPSS. The menu items are: Define Variable Properties..., Copy Data Properties..., New Custom Attribute..., Define Dates..., Define Multiple Response Sets..., Validation (with a submenu arrow), Identify Duplicate Cases..., Identify Unusual Cases..., Sort Cases..., Transpose..., Restructure..., Merge Files (with a submenu arrow), Aggregate..., Orthogonal Design (with a submenu arrow), Copy Dataset, Split File..., Select Cases..., and Weight Cases...</p>	<p>Alteração global dos dados; Ordenar a base; Juntar ficheiros (<i>Merge Files</i>); Dividir a análise por grupos (<i>Split File</i>); Criar subconjuntos de casos para análise (<i>Select Cases</i>); Activar ponderadores (<i>Weight Cases</i>).</p>
<p>Transform</p>	 <p>The screenshot shows the 'Transform' menu in SPSS. The menu items are: Compute Variable..., Count Values within Cases..., Recode into Same Variables..., Recode into Different Variables..., Automatic Recode..., Visual Binning..., Optimal Binning..., Rank Cases..., Date and Time Wizard..., Create Time Series..., Replace Missing Values..., Random Number Generators..., and Run Pending Transforms (with a keyboard shortcut Ctrl+G).</p>	<p>Criar novas variáveis com base nas variáveis originais; Recodificar variáveis.</p>
<p>Analyse</p>	 <p>The screenshot shows the 'Analyze' menu in SPSS. The menu items are: Reports (with a submenu arrow), Descriptive Statistics (with a submenu arrow), Tables (with a submenu arrow), Compare Means (with a submenu arrow), General Linear Model (with a submenu arrow), Generalized Linear Models (with a submenu arrow), Mixed Models (with a submenu arrow), Correlate (with a submenu arrow), Regression (with a submenu arrow), Loglinear (with a submenu arrow), Classify (with a submenu arrow), Data Reduction (with a submenu arrow), Scale (with a submenu arrow), Nonparametric Tests (with a submenu arrow), Time Series (with a submenu arrow), Survival (with a submenu arrow), Multiple Response (with a submenu arrow), Missing Value Analysis..., Complex Samples (with a submenu arrow), Quality Control (with a submenu arrow), and ROC Curve...</p>	<p>Procedimentos de análise estatística.</p>
<p>Graphs</p>	 <p>The screenshot shows the 'Graphs' menu in SPSS. The menu items are: Chart Builder..., Interactive (with a submenu arrow), Legacy Dialogs (with a submenu arrow), and Map (with a submenu arrow).</p>	<p>Criar gráficos.</p>

Utilities		Informação sobre as variáveis.
Window		Comutar entre janelas; Minimizar janelas
Help		Ajuda em linha; Tutorial.

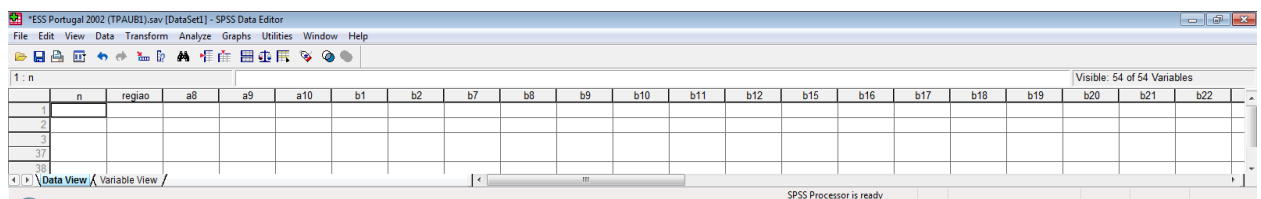
1.4. Criação de bases de dados

As bases de dados são criadas na janela *Variable View*, devendo a estrutura das variáveis obedecer às seguintes regras:



Name	<ul style="list-style-type: none"> - Máximo 64 caracteres (versões anteriores à 13, apenas 8); - Deve começar por uma letra; os restantes caracteres podem ser letras (maiúsculas ou minúsculas são iguais), algarismos, ou os símbolos @, #, _, \$. - Não se podem usar espaços em branco, nem os seguintes caracteres: !, ?, ' , " , * , + , - , % , vírgula, ponto e vírgula, \ , / , > , < - Evitar terminar o nome com o carácter _ (underscore); - Evitar usar caracteres acentuados ou com til.
Type	Por defeito é numérico, pode alterar-se para outro tipo, por exemplo carácter (<i>string</i>), data, etc
Width	Nº de caracteres do campo. Por defeito, 8. Pode ser aumentado – no caso das variáveis <i>string</i> , até 255.
Decimals	Define o número das casas decimais.
Label	Etiquetas dos nomes (<i>name</i>) das variáveis. Admite o máximo de 128 caracteres.
Values	Etiquetas dos valores (códigos) das variáveis nominais ou ordinais.
Missing	Define os códigos das respostas não válidas (não sabe, não responde, não tem que responder) que serão excluídas da análise.
Columns	Largura da coluna de introdução de dados. Por defeito, 8.
Align	Permite alinhar os dados à esquerda, centro ou direita.
Measures	Define o nível de medida das variáveis: nominal, ordinal ou scale.

A introdução dos dados processa-se na janela *Variable View* após ter sido criada a estrutura da base.



Tratamento preliminar de dados com SPSS®

2. Tratamento preliminar de dados

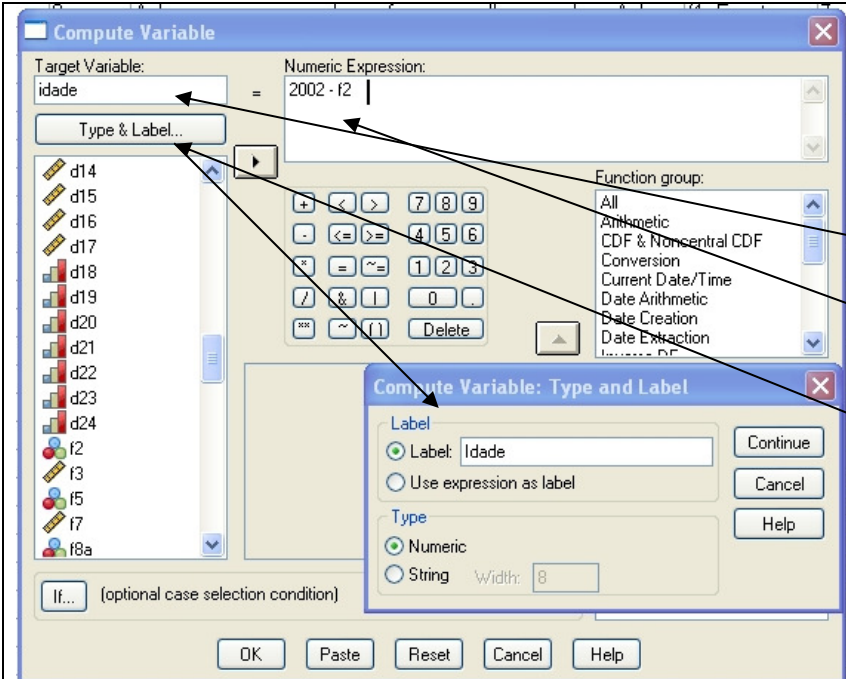
Nota: A base de dados que vamos utilizar é um extracto da base de dados do **European Social Survey (round 1)**³ com os resultados da aplicação do questionário em Portugal (ficheiro **ESS Portugal 2002 (TPAUB1).sav**)

2.1. Transformação de variáveis (comando *Transform*)

2.1.1. Recodificação de variáveis

Exemplo 1: Pretende-se criar duas novas variáveis: a variável **idade** a partir da variável **f3** (data de nascimento) e recodificá-la, criando uma nova variável - **idade2** - com 4 escalões: “até 30 anos”; “31 – 50 anos”; “51 – 65 anos” e “> 65 anos”:

1. Criação da variável **idade**:



- Nome da nova variável
- Expressão para criar a variável (2002 – o ano de nascimento)
- Label da nova variável

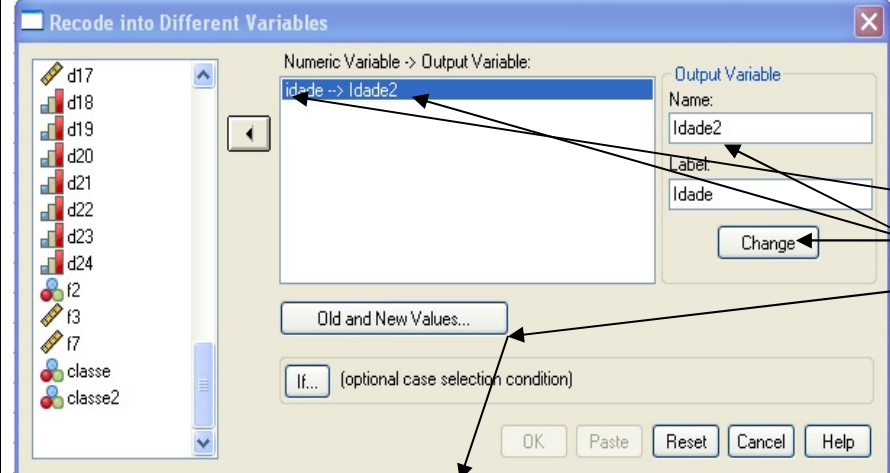
A variável **idade** acrescentou-se à base de dados:

idade	Numeric	8	0	Idade	None	None	1	Right	Scale
-------	---------	---	---	-------	------	------	---	-------	-------

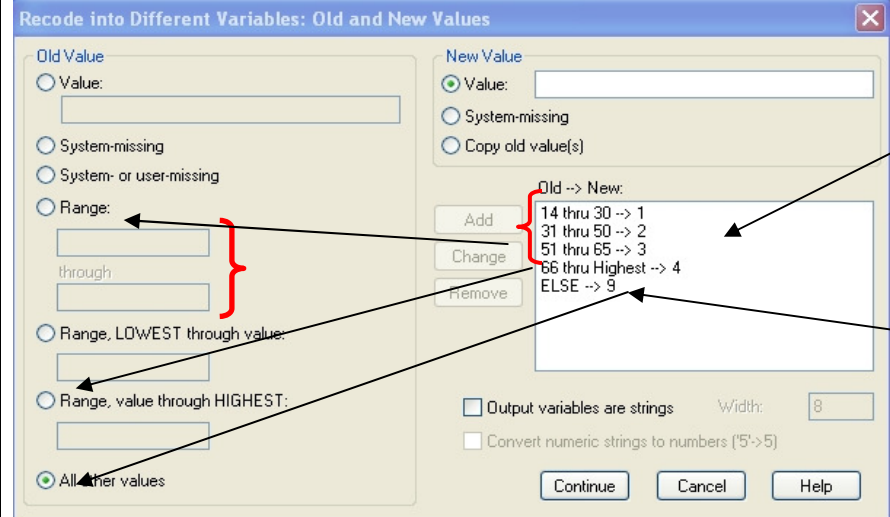
³ A base de dados original está disponível em <http://www.europeansocialsurvey.org/>.

2. Recodificação da variável *idade* e criação da variável *idade2*:

Vamos agora proceder à sua recodificação, criando uma nova variável – *idade2* – com 4 escalões⁴:



- Variável original
- Nova variável
- Clicar

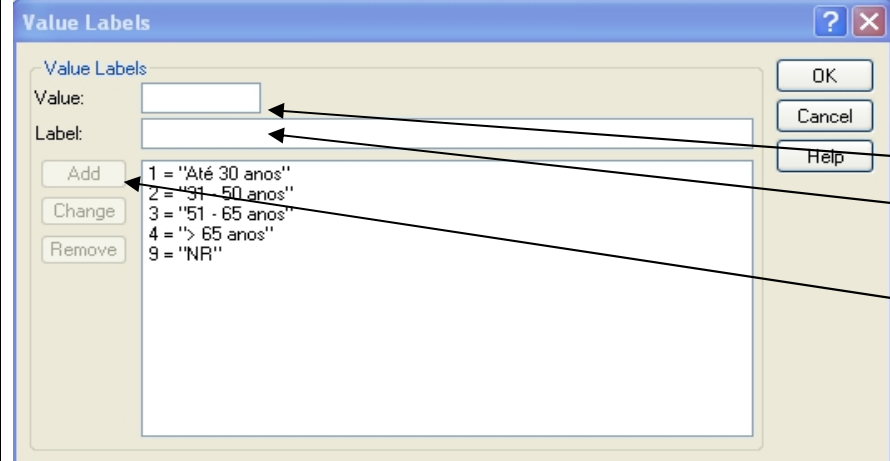


- Categorias
- Categoria para as "não respostas"

A variável *idade2* acrescentar-se-á à base e deverá ser completada com a alteração do nível de medida (*scale* para *ordinal*) e a definição dos respectivos *value labels*:

Idade2	Numeric	8	0	Idade	{1, Até 30 a ...	None	10	Right	Ordinal
--------	---------	---	---	-------	------------------	------	----	-------	---------

Clicar para criar as etiquetas das categorias (*Value labels*)



- Código da categoria
- Designação da categoria
- Clicar para adicionar

⁴ **Nota:** na recodificação de variáveis é **recomendável** manter as variáveis originais e criar novas variáveis recodificadas, escolhendo para o efeito a opção **Into different variable**.

Tratamento Preliminar de Dados

O resultado será o seguinte:

Idade					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Até 30 anos	341	22,6	22,6	22,6
	31 - 50 anos	505	33,4	33,4	56,0
	51 - 65 anos	315	20,8	20,8	76,8
	> 65 anos	350	23,2	23,2	100,0
	Total	1511	100,0	100,0	

Exemplo 2: Pretende-se recodificar a variável “escolaridade” (**f7**) criando uma nova variável (**escol**) com 3 escalões: “até 9 anos”; “10 – 12 anos” e “> 12 anos”:

Recode into Different Variables

Numeric Variable -> Output Variable

f7 -> escol

Old and New Values...

If... (optional case selection)

Old Value

☐ Value:

☐ System-missing

☐ System- or user-missing

☒ Range:

through

☐ Range, LOWEST through value:

☐ Range, value through HIGHEST:

☐ All other values

New Value

☒ Value:

☐ System-missing

☐ Copy old value(s)

Old -> New:

10 thru 12 -> 2

0 thru 9 -> 1

13 thru 25 -> 3

Add

Change

Remove

☒ Output variables are strings Width: 8

☐ Convert numeric strings to numbers ('5'>'5')

Continue Cancel Help

A variável **escol** acrescentar-se-á à base e deverá ser completada com a alteração do nível de medida (*scale* para *ordinal*) e a definição dos respectivos *value labels*.

O resultado será o seguinte:

escol	Numeric	8	0	Anos de escolaridade concluídos	{1, Até 9 anos}	99	1 R	Ordinal
-------	---------	---	---	---------------------------------	-----------------	----	-----	---------

Escolaridade					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Até 9 anos	1046	69.2	69.2	69.2
	10 - 12 anos	251	16.6	16.6	85.8
	> 12 anos	212	14.0	14.0	99.9
	NR	2	.1	.1	100.0
	Total	1511	100.0	100.0	

Exemplo 3: Pretende-se recodificar o “autoposicionamento político” (**b28**) criando uma nova variável (**b28r**) com 3 escalões: “esquerda”; “centro” e “direita”:

A variável **b28r** acrescentar-se-á à base e deverá ser completada com a alteração do nível de medida (*scale* para *ordinal*) e a definição dos respectivos *value labels*.

O resultado será o seguinte:

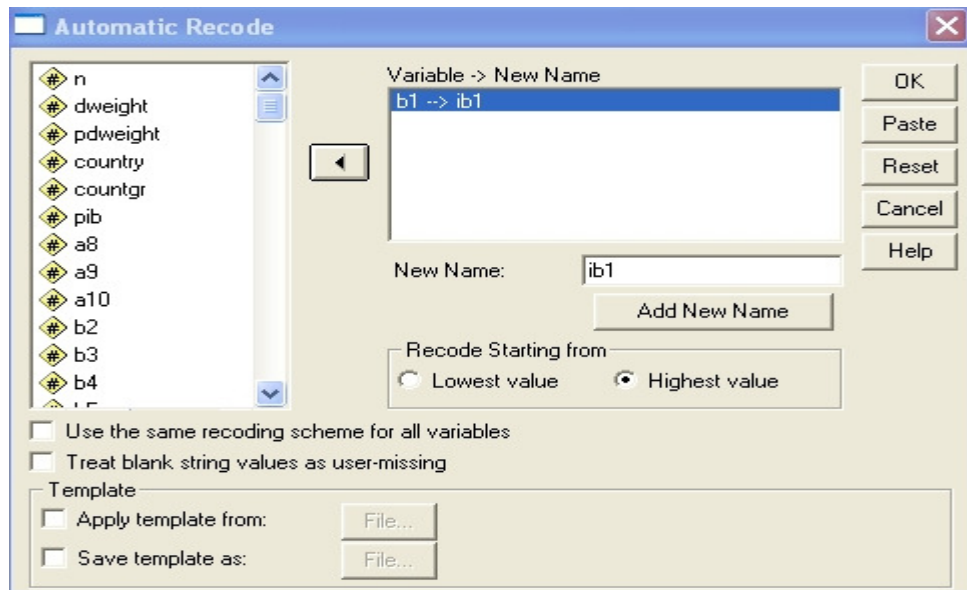
b28r	Numeric	8	0	Autoposicionamento político	{1, Esquerda}..None	1R	Nominal
------	---------	---	---	-----------------------------	---------------------	----	---------

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Esquerda	296	19.6	24.5	24.5
	Centro	620	41.0	51.2	75.7
	Direita	294	19.4	24.3	100.0
	Total	1211	79.9	100.0	
Missing	System	304	20.1		
Total		1515	100.0		

Tratamento Preliminar de Dados

2.1.2. Inversão da escala de uma variável

Pretende-se criar uma nova variável (**ib1**) com a inversão da escala da variável interesse pela política (**b1**) de modo a que 1 corresponda a “nenhum interesse” e 4 a “muito interesse”:



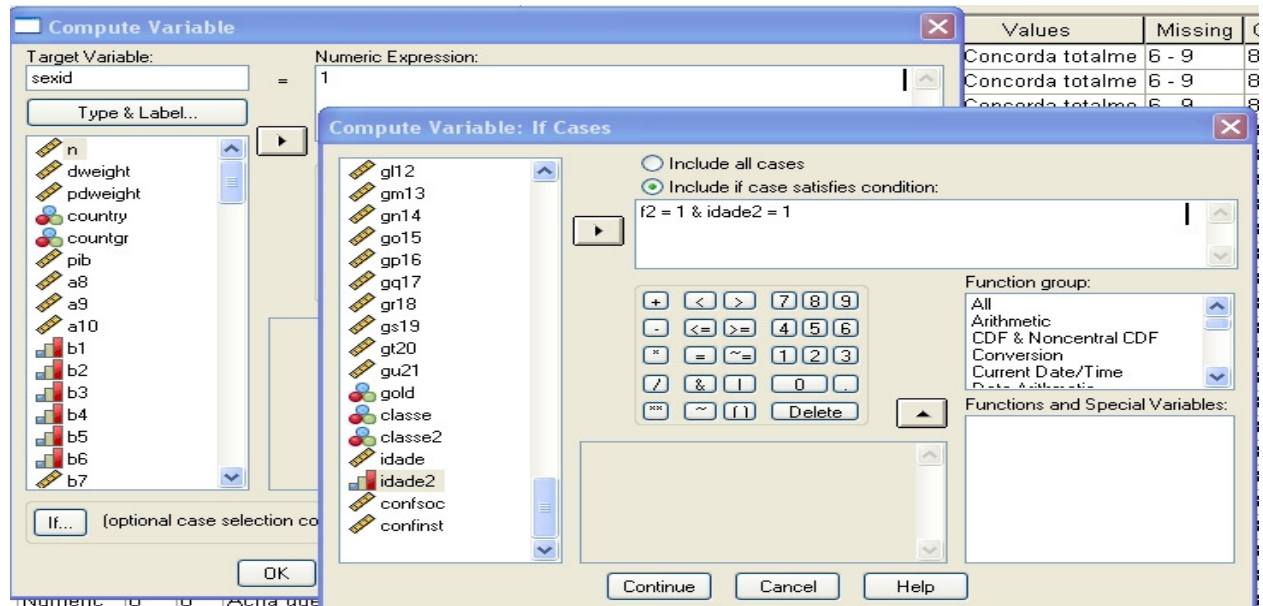
O resultado é o seguinte:

b1	Numeric	1	0	Qual o seu interesse pela política	{1, Muito interesse}...	6 - 9	8	Right	Ordinal
ib1	Numeric	1	0	Qual o seu interesse pela política	{1, Nenhum interesse}...	5 - 7	3	Right	Ordinal

b1						ib1					
Qual o seu interesse pela política						Qual o seu interesse pela política					
		Frequency	Percent	Valid Percent	Cumulative Percent			Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Muito interesse	117	7.7	7.8	7.8	Valid	Nenhum interesse	489	32.4	32.5	32.5
	Algum interesse	456	30.2	30.3	38.1		Pouco interesse	441	29.2	29.3	61.9
	Pouco interesse	441	29.2	29.3	67.5		Algum interesse	456	30.2	30.3	92.2
	Nenhum interesse	489	32.4	32.5	100.0		Muito interesse	117	7.7	7.8	100.0
	Total	1503	99.5	100.0			Total	1503	99.5	100.0	
Missing	Recusa	6	.4			Missing	Não sabe	2	.1		
	Não sabe	2	.1				Recusa	6	.4		
	Total	8	.5				Total	8	.5		
Total		1511	100.0			Total		1511	100.0		

2.1.3. Transformação lógica de variáveis

Exemplo: pretende-se criar uma variável - **sexid** - através da transformação lógica das variáveis **f2** e **idade2**, com 4 categorias: “homens até 30 anos”, “homens com mais de 30 anos”, “mulheres até 30 anos” e “mulheres com mais de 30 anos”.



Nota: repetir o comando para as restantes categorias, cujas expressões numéricas são as seguintes:

Categoria 2: $f2 = 1 \text{ \& } idade2 > 1$

Categoria 3: $f2 = 2 \text{ \& } idade2 = 1$

Categoria 4: $f2 = 2 \text{ \& } idade2 > 1$

A variável **sexid** acrescentar-se-á à base e deverá ser completada com a alteração do nível de medida (*scale* para *ordinal*) e a definição dos respectivos *label* e *value labels*. O resultado é o seguinte:

sexid	Nu	8	0	Sexo e Idade	{1, Homens at	None	10	Rig	Nominal
-------	----	---	---	--------------	---------------	------	----	-----	---------

Sexo e Idade					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Homens até 30 anos	158	10.5	10.5	10.5
	Homens com mais de 30 anos	472	31.2	31.2	41.7
	Mulheres até 30 anos	183	12.1	12.1	53.8
	Mulheres com mais de 30 anos	698	46.2	46.2	100.0
	Total	1511	100.0	100.0	

3. Criação de Índices sintéticos

Os índices sintéticos são compostos por um conjunto de indicadores agregados que medem um construto teórico, estando a sua agregação, por conseguinte, justificada teoricamente. Podem ser criados através de transformação algébrica – médias – quando os indicadores estão “medidos” numa escala tipo *Likert*, ou aditiva, quando se trata de indicadores dicotómicos (presença/ausência)⁵.

3.1. Fiabilidade dos índices

Os índices sintéticos devem ser, para além de teoricamente fundamentados, consistentes internamente⁶. Para o efeito, deve testar-se a sua fiabilidade (reliability) através do *Alpha de Cronbach*, que consiste numa fórmula estatística que pondera o rácio entre a variância total dos indicadores que compõem o índice e a variância de cada um dos indicadores.

Varia entre 0 e 1. Quanto mais perto de 1, maior a consistência da escala.⁷ A Tabela seguinte resume os níveis de fiabilidade recomendados por diversos autores⁸

<i>Crítérios de recomendação de Fiabilidade estimada pelo α de Cronbach (adaptado de Peterson, 1994)</i>		
Autor	Condição	α considerado aceitável
Davis, 1964, p. 24	Previsão individual	Acima de 0.75
	Previsão para grupos de 25-50 indivíduos	Acima de 0.5
Kaplan & Sacuzzo, 1982, p. 106	Investigação fundamental	0.7-0.8
	Investigação aplicada	0.95
Murphy & Davidsholder, 1988, p. 89	Fiabilidade inaceitável	<0.6
	Fiabilidade baixa	0.7
	Fiabilidade moderada a elevada	0.8-0.9
	Fiabilidade Elevada	>0.9
Nunnally, 1978, p. 245-246	Investigação preliminar	0.7
	Investigação fundamental	0.8
	Investigação aplicada	0.9-0.95

Outros autores, contudo, consideram aceitáveis, embora a fiabilidade seja pobre, valores de referência entre [0,05 e 0,06]⁹.

⁵ Os índices sintéticos assim construídos não são, obviamente, ponderados. A ponderação deve ser teórica ou estatisticamente fundamentada. Uma forma de ponderar estatisticamente os índices sintéticos construídos a partir de indicadores medidos numa escala tipo Likert, é através de uma Análise de Componentes Principais (ACP), com uma só componente, em que os escores standardizados são automaticamente ponderados pela importância relativa dos indicadores.

⁶ Testar a sua coerência interna, ou seja, verificar se as respostas dos inquiridos são devidamente congruentes entre si.

⁷ Sempre que o valor do *Alpha* seja negativo, tal dever-se-á, na maior parte dos casos, ao facto de se utilizarem itens (indicadores) cujas escalas variam em sentidos deferentes, sem que tenha havido o cuidado de proceder às respectivas inversões. Trata-se, por conseguinte, de um erro que deve ser corrigido, assegurando, através da respectiva recodificação dos indicadores, que as escalas são todas crescentes ou decrescentes (variam no mesmo sentido).

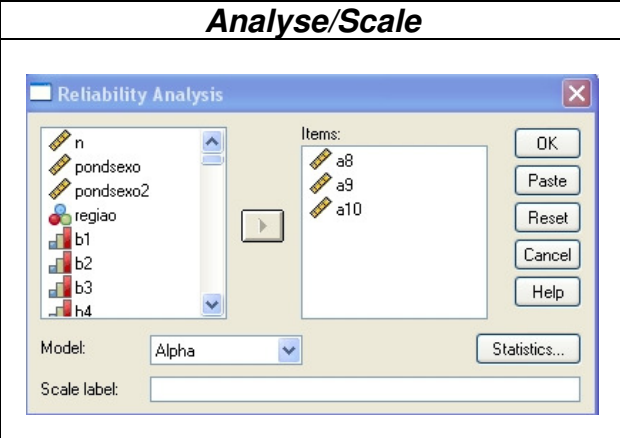
⁸ Cfr. Maroco, J. e T. Garcia-Marques: Qual a fiabilidade do alfa de Cronbach? *Questões antigas e soluções modernas?*, disponível em: [http://repositorio.ispa.pt/bitstream/10400.12/133/1/LP%204\(1\)%20-%2065-90.pdf](http://repositorio.ispa.pt/bitstream/10400.12/133/1/LP%204(1)%20-%2065-90.pdf), consultado em 08/09/2010.

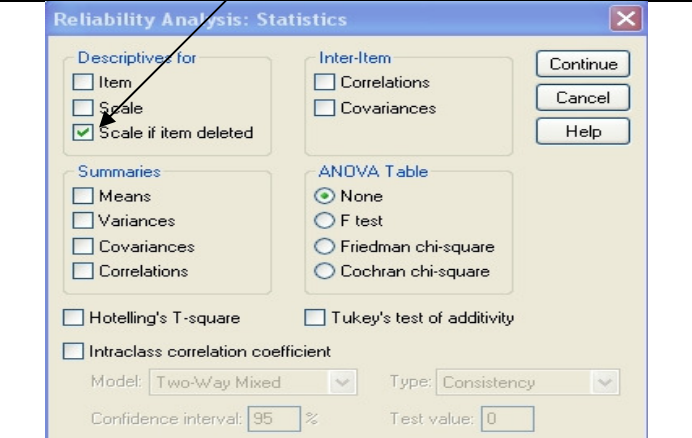
⁹ George D. e P. Mallery, *SPSS for Windows, step by step: A simple guide and reference. 11.0 Update*, Boston, Allyn & Bacon, 2003.

3.2. Média dos indicadores

Exemplo: Pretende-se criar dois índices sintéticos: **Confiança social** (variáveis $a8+a9+a10$) e **Confiança institucional** ($b7+b8+b9+b10$), através das médias algébricas dos respectivos indicadores:

1. Testar a fiabilidade dos índices:





Reliability Statistics

Cronbach's Alpha	N of Items
,720	3

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Confiança interpessoal	8,95	12,423	,600	,556
Confiança na honestidade dos outros	7,79	14,070	,551	,620
Confiança no altruísmo dos outros	9,15	15,245	,476	,706

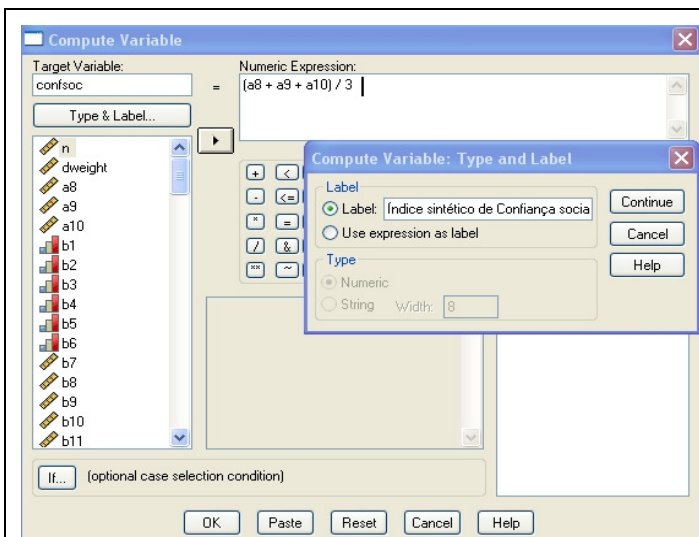
Reliability Statistics

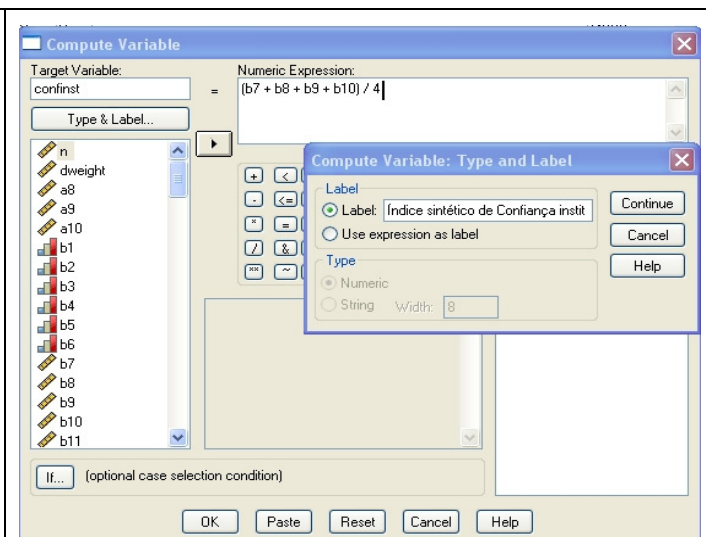
Cronbach's Alpha	N of Items
,797	4

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Confiança no Parlamento	12,07	28,527	,630	,736
Confiança no Sistema Jurídico	12,14	28,179	,673	,714
Confiança na Polícia	11,41	30,663	,531	,786
Confiança nos Políticos	13,55	31,718	,611	,748

2. Criar os Índices sintéticos:



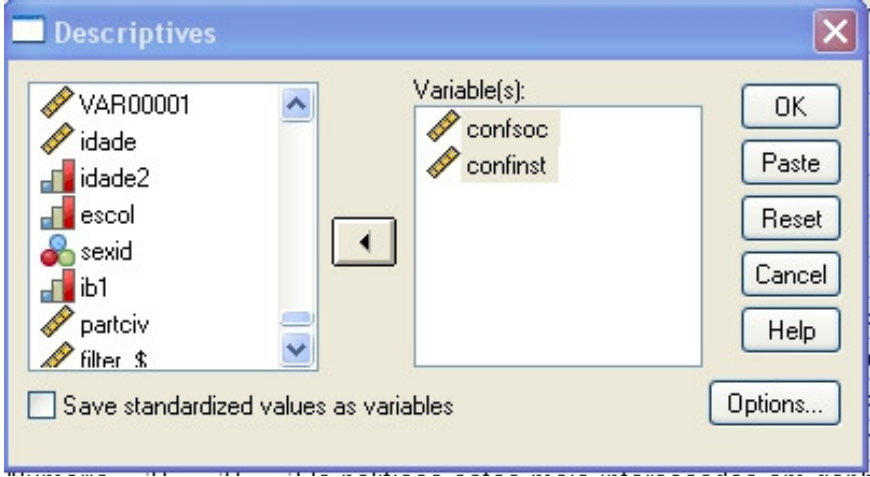


Tratamento Preliminar de Dados

As 2 novas variáveis (índices) acrescentaram-se à base¹⁰:

confsoc	Numeric	8	1	Índice sintético de Confiança social	None	None	1	Right	Scale
confinst	Numeric	8	1	Índice sintético de Confiança institucional	None	None	1	Right	Scale

Os resultados são os seguintes:



The image shows the SPSS 'Descriptives' dialog box. On the left, a list of variables includes VAR00001, idade, idade2, escol, sexid, ib1, partciv, and filter \$. On the right, under 'Variable(s):', the variables 'confsoc' and 'confinst' are listed. At the bottom, there is a checkbox for 'Save standardized values as variables' which is currently unchecked. Buttons for OK, Paste, Reset, Cancel, Help, and Options... are on the right side of the dialog.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Índice sintético de Confiança social	1480	.0	10.0	4.316	1.7469
Índice sintético de Confiança institucional	1338	.0	9.3	4.097	1.7603
Valid N (listwise)	1319				

¹⁰ Nota: tratando-se de variáveis rácio, deverão ter casas decimais (1 ou 2).

Tratamento Preliminar de Dados

3.3. Contagem de ocorrências (adição)

Exemplo: pretende-se criar uma variável – **partciv** – que traduza o **Índice de participação cívica**, que integre (conte) apenas os inquiridos que responderam **sim** (1) às questões **b15** a **b24**.

1. Testar a fiabilidade do índice:

Reliability Statistics		Item-Total Statistics				
Cronbach's Alpha	N of Items		Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
,769	10	Contactou um político ou representante do governo/poder local	17,60	1,185	,400	,763
		Trabalhou para um partido político/movimento cívico	17,53	1,269	,566	,734
		Trabalhou numa organização ou associação de outro tipo	17,54	1,253	,502	,740
		Usou um emblema auto-colante de campanha/movimento	17,55	1,213	,535	,733
		Assinou uma petição	17,55	1,194	,546	,733
		Participou numa manifestação	17,53	1,319	,434	,750
		Boicotou determinados produtos	17,52	1,372	,375	,751
		Comprou produtos por razões de ordem política, ética ou ambiental	17,56	1,250	,400	,750
		Deu dinheiro a um grupo ou organização política	17,52	1,328	,452	,746
		Participou em acções de protesto ilegais	17,49	1,526	,199	,774

2. Criar o índice:

A variável **partciv** acrescentar-se-á à base. O resultado é o seguinte:

partciv	Nu	8	0	Participação cívica	None	None	10	Rig Scale
---------	----	---	---	---------------------	------	------	----	-----------

Índice sintético de Participação cívica

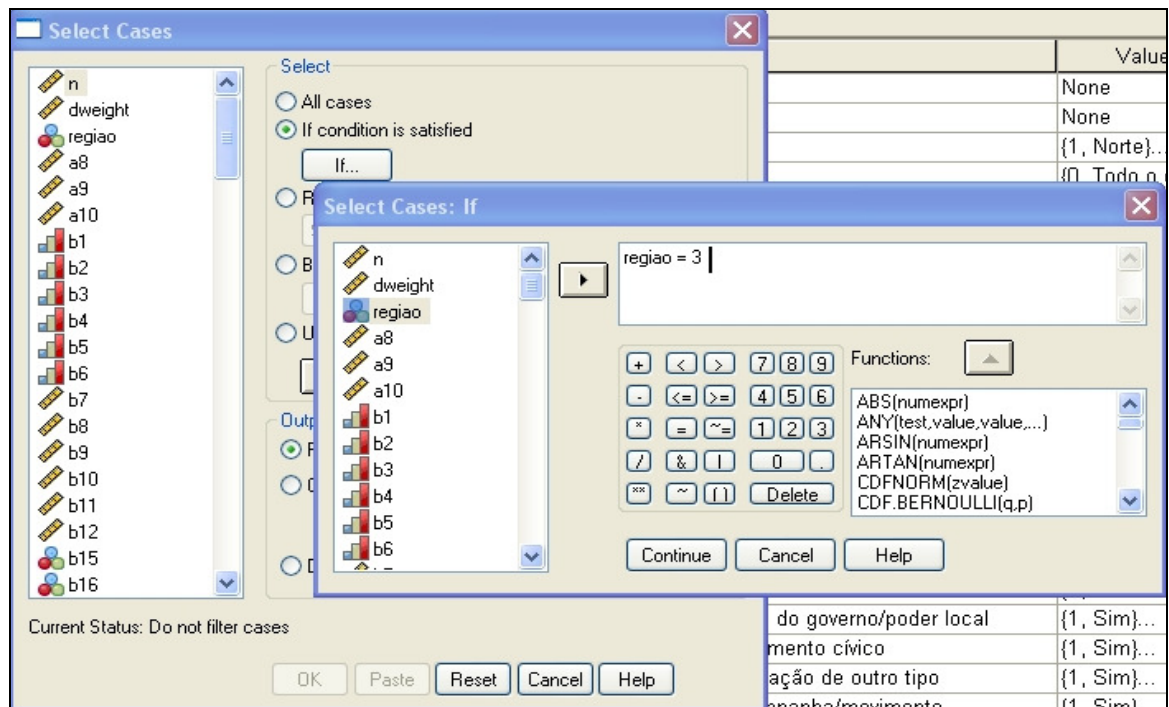
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	1174	77.7	77.7	77.7
	1	145	9.6	9.6	87.3
	2	85	5.6	5.6	92.9
	3	39	2.6	2.6	95.5
	4	30	2.0	2.0	97.5
	5	16	1.1	1.1	98.5
	6	9	.6	.6	99.1
	7	11	.7	.7	99.9
	9	1	.1	.1	99.9
	10	1	.1	.1	100.0
	Total	1511	100.0	100.0	

A interpretação é a seguinte: 77,7% (1174) inquiridos não assinalaram nenhum indicador, 9,6% (145) assinalaram apenas 1, 5,6% (85) assinalaram 2, etc...

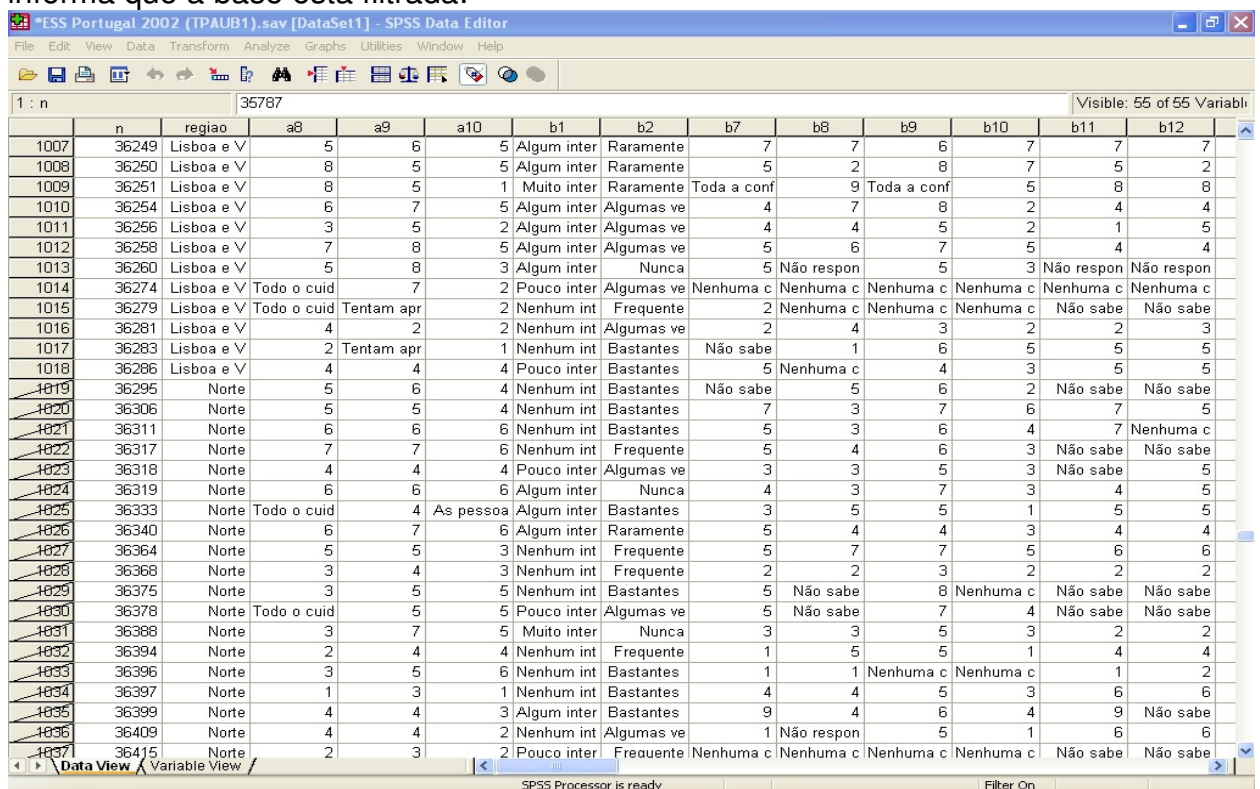
4. Selecção de casos

4.1. Seleccionar uma sub-amostra

Exemplo: Pretende-se seleccionar (filtrar) apenas os inquiridos da região de Lisboa e Vale do Tejo (*regiao*=3).

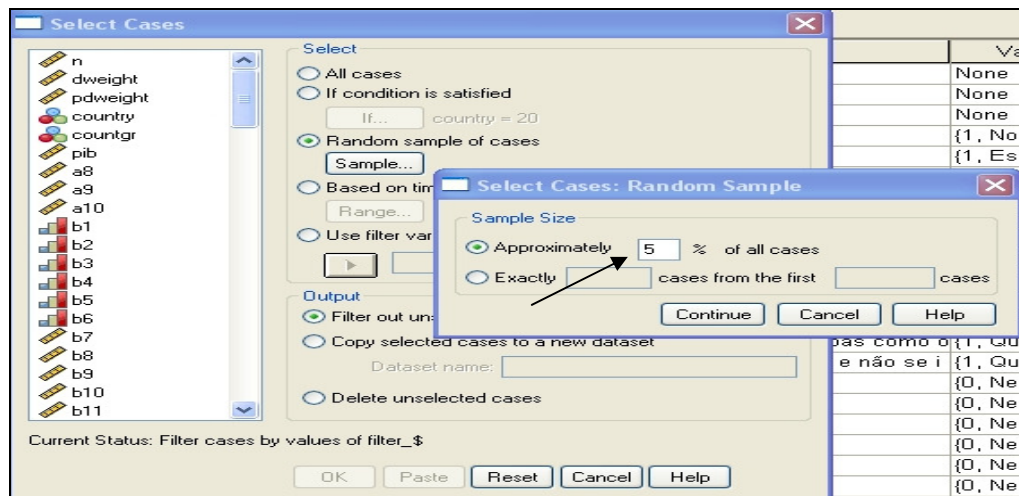


Na base de dados (*Data View*) os registos não seleccionados (filtrados) aparecem tracejados, mantendo-se assim até que se anule a selecção (filtro). A barra de *status* informa que a base está filtrada:



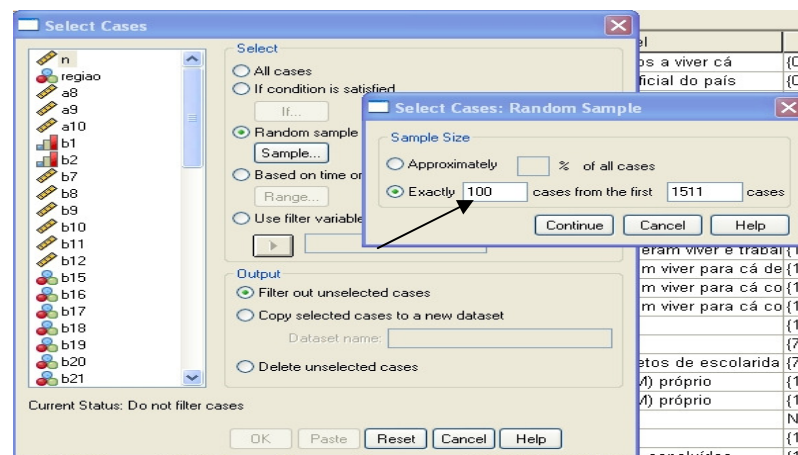
4.2. Seleccionar uma amostra aleatória simples

Exemplo: Pretende-se seleccionar aleatoriamente uma amostra de cerca de 5% dos inquiridos:

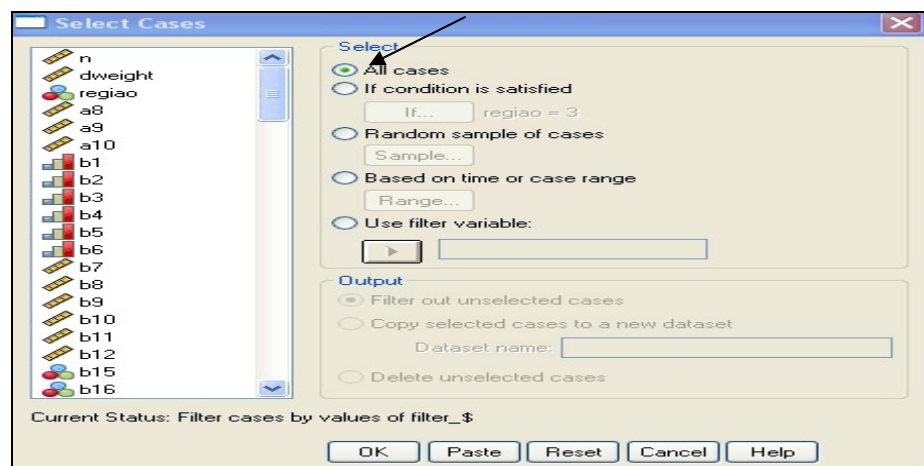


4.3. Seleccionar uma amostra aleatória com nº de casos pré-definidos

Exemplo: Pretende-se seleccionar aleatoriamente uma amostra de 100 inquiridos:

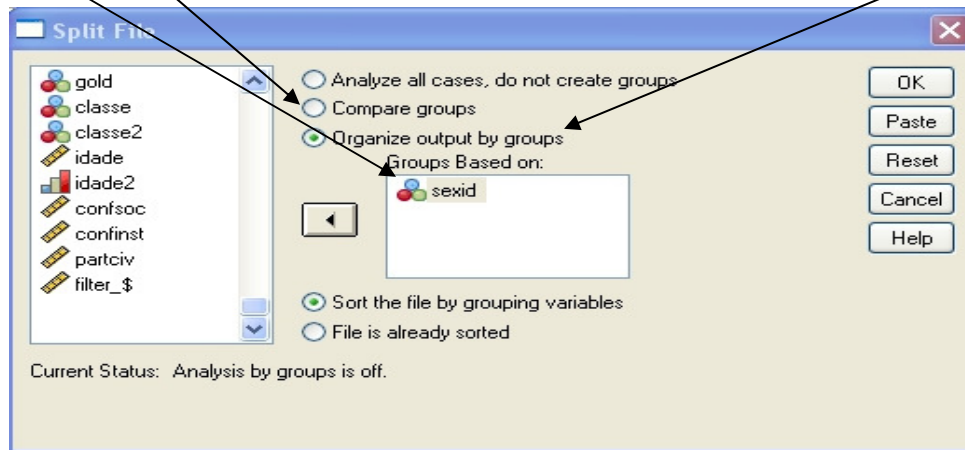


Nota importante: não esquecer de desactivar o filtro quando não for necessário:



5. Separar a análise por sub-amostras

Exemplo: Pretende-se separar a análise pelas 4 categorias (sub-amostras) da variável **sexid** (sexo e idade). É possível obter os resultados na mesma tabela, seleccionando a opção **Compare groups** ou em tabelas diferentes, com a opção **Organize output by groups**:



No primeiro caso, os resultados são os seguintes:

Descriptive Statistics						
Sexo e Idade		N	Minimum	Maximum	Mean	Std. Deviation
Homens até 30 anos	Índice sintético de Confiança social	156	1,0	9,0	4,859	1,4670
	Índice sintético de Confiança institucional	147	,0	8,5	4,248	1,7428
	Valid N (listwise)	145				
Homens com mais de 30 anos	Índice sintético de Confiança social	456	,0	10,0	4,259	1,8100
	Índice sintético de Confiança institucional	440	,0	9,3	4,066	1,8441
	Valid N (listwise)	430				
Mulheres até 30 anos	Índice sintético de Confiança social	182	,0	9,3	4,604	1,6277
	Índice sintético de Confiança institucional	168	,0	8,0	4,116	1,6234
	Valid N (listwise)	167				
Mulheres com mais de 30 anos	Índice sintético de Confiança social	686	,0	10,0	4,153	1,7629
	Índice sintético de Confiança institucional	583	,0	9,3	4,075	1,7401
	Valid N (listwise)	577				

No segundo caso seriam produzidas 4 tabelas, uma por cada categoria da variável **split**:

Descriptive Statistics ^a					
	N	Minimum	Maximum	Mean	Std. Deviation
Índice sintético de Confiança social	156	1,0	9,0	4,859	1,4670
Índice sintético de Confiança institucional	147	,0	8,5	4,248	1,7428
Valid N (listwise)	145				
a. Sexo e Idade = Homens até 30 anos					

Descriptive Statistics ^a					
	N	Minimum	Maximum	Mean	Std. Deviation
Índice sintético de Confiança social	456	,0	10,0	4,259	1,8100
Índice sintético de Confiança institucional	440	,0	9,3	4,066	1,8441
Valid N (listwise)	430				
a. Sexo e Idade = Homens com mais de 30 anos					

Descriptive Statistics ^a					
	N	Minimum	Maximum	Mean	Std. Deviation
Índice sintético de Confiança social	182	,0	9,3	4,604	1,6277
Índice sintético de Confiança institucional	168	,0	8,0	4,116	1,6234
Valid N (listwise)	167				
a. Sexo e Idade = Mulheres até 30 anos					

Descriptive Statistics ^a					
	N	Minimum	Maximum	Mean	Std. Deviation
Índice sintético de Confiança social	686	,0	10,0	4,153	1,7629
Índice sintético de Confiança institucional	583	,0	9,3	4,075	1,7401
Valid N (listwise)	577				
a. Sexo e Idade = Mulheres com mais de 30 anos					

Links úteis relacionados com a matéria deste manual

<http://support.spss.com/ProductsExt/SPSS/ESD/17/Download/User%20Manuals/English/SPSS%20Statistics%20Brief%20Guide%2017.0.pdf>

<http://stat.med.up.pt/ch1index.html>

<http://www2.dce.ua.pt/leies/pacgi/Folhaapoio1.pdf>

<http://www2.uca.es/serv/ai/formacion/spss/Inicio.pdf>

<http://www.stat.tamu.edu/spss.php/>

<http://ritco.biz/files/Statistical%20Analyses%20using%20SPSS.pdf>

<http://calcnet.mth.cmich.edu/org/spss/>

ANEXO 1

Construção de um ponderador

Construção de um ponderador

A distribuição da variável sexo na base de dados é a seguinte:

Sexo

		Frequency	Percent
Valid	Masculino	630	41,7
	Feminino	881	58,3
	Total	1511	100,0

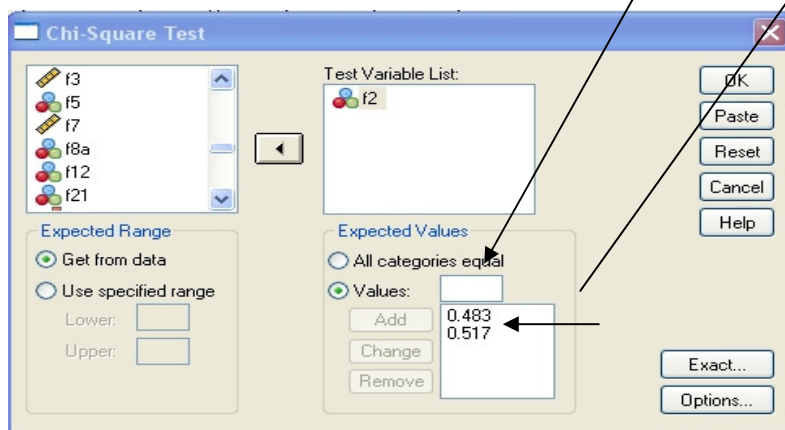
O Censo de 2001 apresentava a seguinte distribuição:

População residente por sexo			
Zona Geográfica	População Residente		
	Homens + Mulheres	Homens	Mulheres
T: Portugal	10 356 117	5 000 141	5 355 976
%		48.3	51.7

Fonte: INE, Recenseamento Geral da População e Habitação - 2001 (Resultados Definitivos)

Torna-se necessário, por conseguinte, criar um ponderador para ponderar a base.

Para conhecermos a diferença entre a amostra observada e a amostra esperada recorreremos ao teste de aderência do χ^2 :



Os resultados são os seguintes:

Sexo				Test Statistics	
	Observed N	Expected N	Residual		Sexo
Masculino	630	729.8	-99.8	Chi-Square ^a	26.404
Feminino	881	781.2	99.8	df	1
Total	1511			Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 729.8.

As diferenças entre a amostra observada e a amostra esperada são estatisticamente significativas ($\chi^2(1)$ 26,4; $p=0.000$).

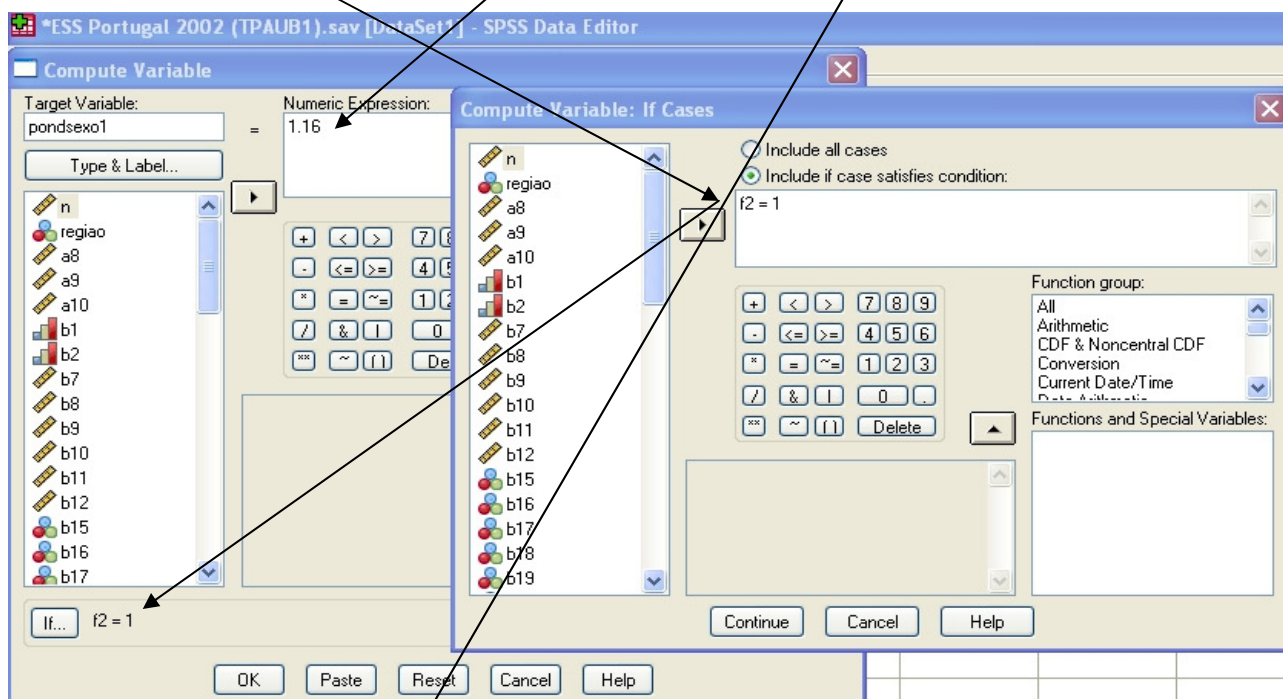
ANEXO 1: Construção de um ponderador

Exemplo 1: Cálculo do ponderador para a base não extrapolada:

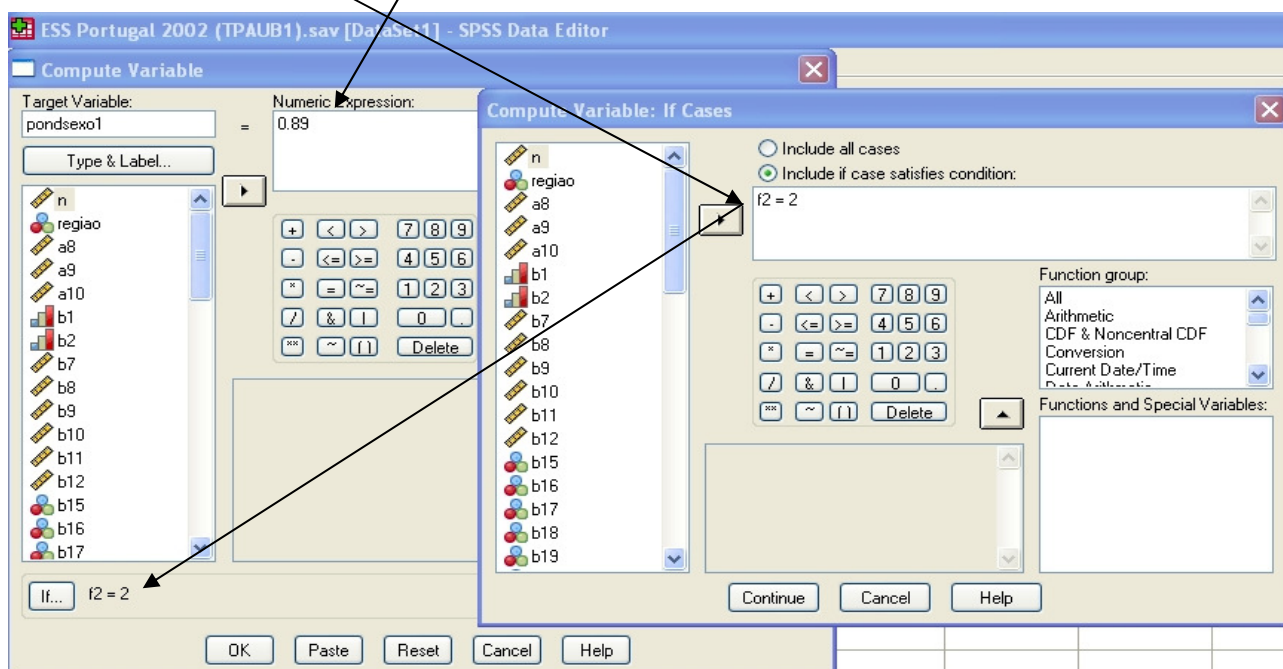
	Homens	Mulheres
Amostra observada	630	881
Amostra esperada	730	781
Diferença	+ 100	- 100
Ponderador	$(730 \cdot 100 / 630) / 100$	$(781 \cdot 100 / 881) / 100$
	1.16	0.89

Construção do ponderador

Sexo masculino:



Sexo feminino:



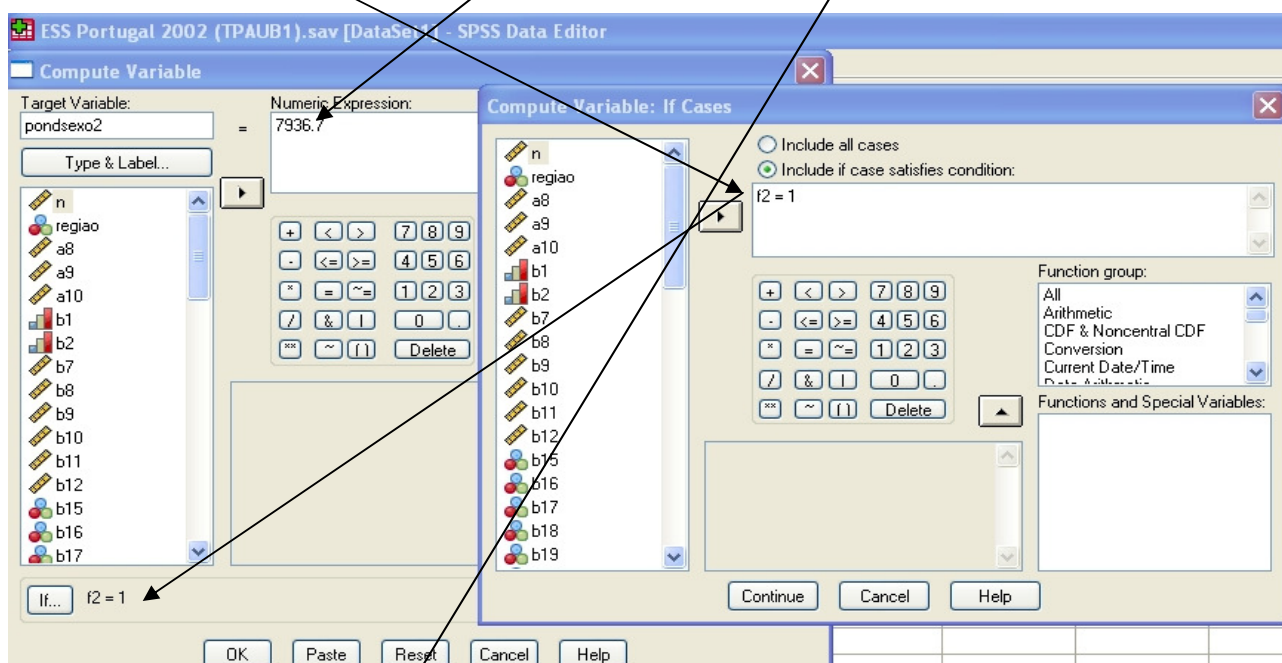
ANEXO 1: Construção de um ponderador

Exemplo 2: Cálculo do ponderador para a base extrapolada:

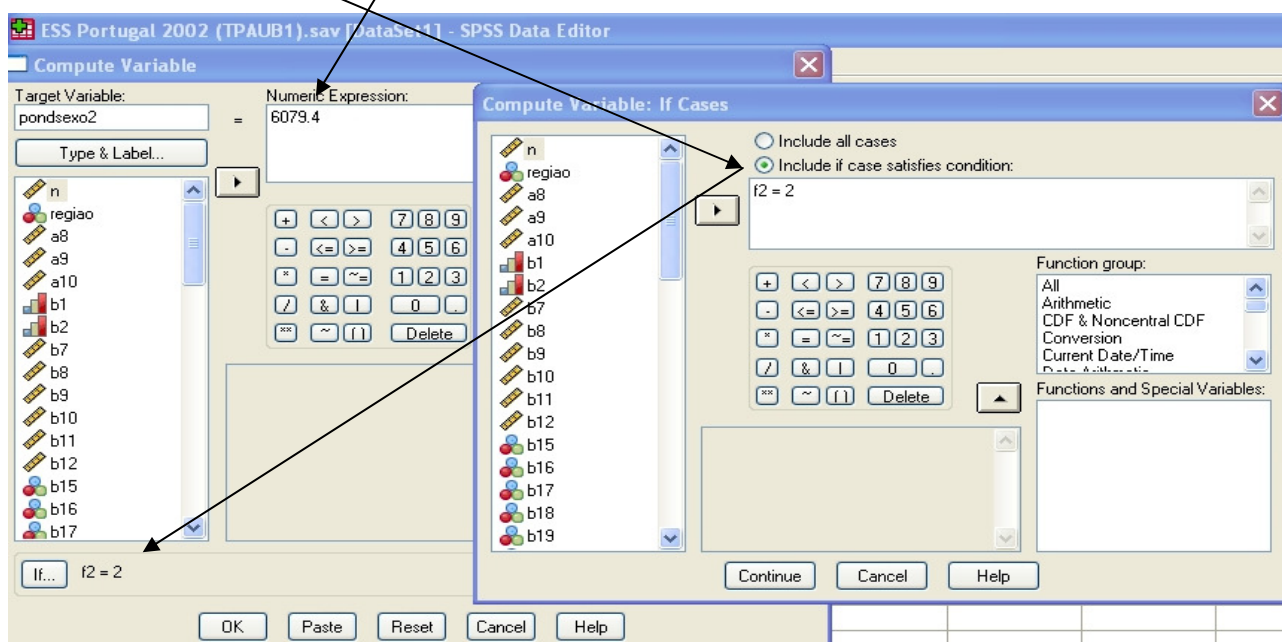
	Homens	Mulheres
Amostra observada	630	881
Total Portugal	5000141	5355976
Ponderador	$(5000141/630)$	$(5355976/881)$
	7936,7	6079,4

Construção do ponderador:

Sexo masculino:

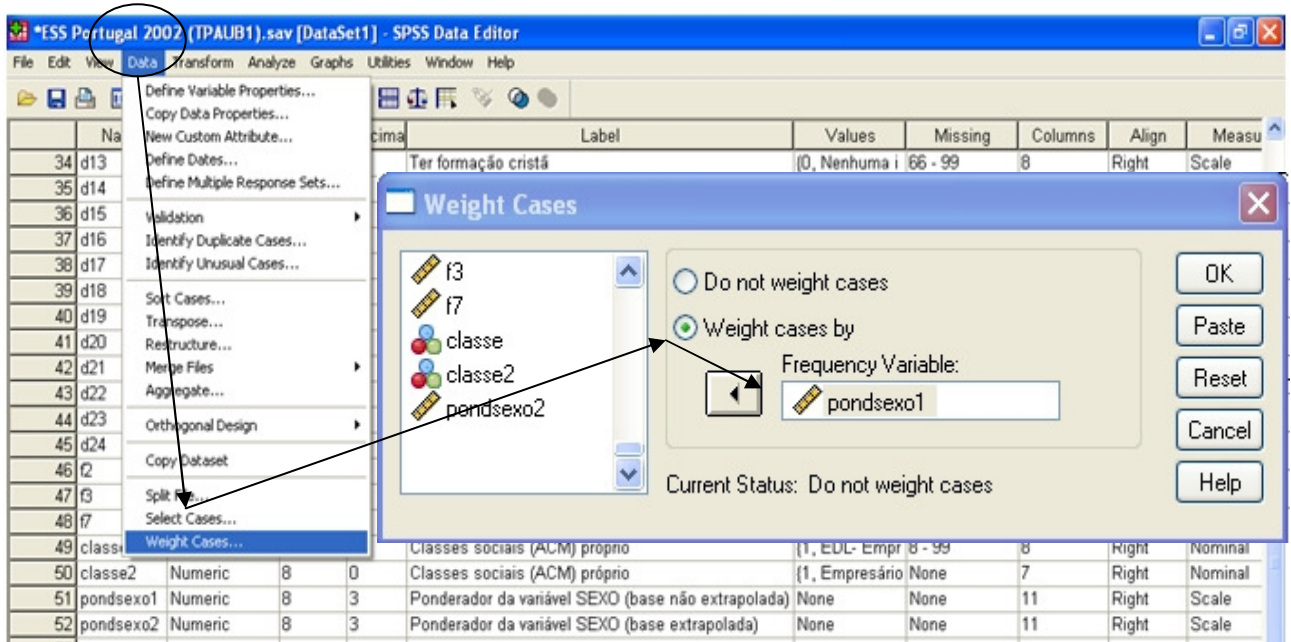


Sexo feminino:



ANEXO 1: Construção de um ponderador

Activação de um ponderador



Os resultados são os seguintes:

Base não extrapolada				Base extrapolada			
		N	%			N	%
Sexo	Masculino	731	48,2	Sexo	Masculino	5000121	48,3
	Feminino	784	51,8		Feminino	5355951	51,7
	Total	1515	100,0		Total	10356072	100,0
		N	%			N	%
Região	Norte	580	38,3	Região	Norte	3967582	38,3
	Centro	268	17,7		Centro	1831898	17,7
	Lisboa e Vale do Tejo	550	36,3		Lisboa e Vale do Tejo	3762232	36,3
	Alentejo	68	4,5		Alentejo	465404	4,5
	Algarve	48	3,2		Algarve	328957	3,2
	Total	1515	100,0		Total	10356072	100,0

ANEXO 2

European Social Survey ***(round 1 - 2002)***

Extracto do Questionário

Versão completa em:

<http://www.europeansocialsurvey.org/>

European Social Survey

(round 1 - 2002)

Extracto do Questionário adaptado

Questionário integral disponível em: <http://www.europeansocialsurvey.org/>

O INQUÉRITO DE 2002

O objectivo central do ESS é o de desenvolver e conduzir um estudo sistemático centrado na mudança de valores, atitudes, atributos e padrões de comportamento entre os europeus, de modo a compreender a sua distribuição e variação, bem como o sentido e a intensidade da mudança, dentro de cada país e entre países.

O questionário, cuja aplicação é feita de dois em dois anos através de entrevista pessoal com cerca de uma hora de duração, consiste num módulo permanente, e em dois módulos rotativos, que serão repetidos em intervalos, cada um deles dedicado a um tema ou tópico específico. Assim, enquanto os módulos rotativos têm como objectivo permitir um conhecimento detalhado das respectivas temáticas, com interesse académico ou para o estabelecimento de medidas políticas, o módulo permanente permitirá monitorar mudanças e continuidades numa vasta gama de variáveis sócio-económicas, sócio-políticas, sócio-psicológicas e sócio-demográficas.

A8: De uma forma geral, acha que todo o cuidado é pouco quando se lida com as pessoas ou acha que se pode confiar na maioria das pessoas?

Todo o cuidado é pouco										A maioria das pessoas é de confiança			
										(Recusa)	(NS)	(NR)	
0	1	2	3	4	5	6	7	8	9	10	77	88	99

A9: Acha que a maior parte das pessoas tentam aproveitar-se de si sempre que podem, ou pensa que a maior parte das pessoas são honestas?

Tentam aproveitar-se de mim										São honestas			
										(Recusa)	(NS)	(NR)	
0	1	2	3	4	5	6	7	8	9	10	77	88	99

A10: Acha que, na maior parte das vezes, as pessoas estão preocupadas com elas próprias ou acha que tentam ajudar os outros?

As pessoas estão preocupadas com elas próprias										As pessoas tentam ajudar os outros			
										(Recusa)	(NS)	(NR)	
0	1	2	3	4	5	6	7	8	9	10	77	88	99

ANEXO 2: Questionário ESS

VOU AGORA FAZER-LHE ALGUMAS PERGUNTAS SOBRE POLÍTICA

B1: De um modo geral, qual o seu interesse pela política?

Muito interesse,	1
Algum interesse,	2
Pouco interesse,	3
ou, Nenhum interesse?	4
(Recusa)	7
(Não sabe)	8
(Não responde)	9

Diga-me, por favor, qual a confiança pessoal que tem em cada uma das instituições que lhe vou dizer:

		Nenhuma confiança											Toda a confiança	(Recusa)	(NS)	(NR)
B7	Parlamento Nacional?	0	1	2	3	4	5	6	7	8	9	10	77	88	99	
B8	Sistema Jurídico?	0	1	2	3	4	5	6	7	8	9	10	77	88	99	
B9	Polícia?	0	1	2	3	4	5	6	7	8	9	10	77	88	99	
B10	Políticos?	0	1	2	3	4	5	6	7	8	9	10	77	88	99	
B11	Parlamento Europeu?	0	1	2	3	4	5	6	7	8	9	10	77	88	99	
B12	Nações Unidas?	0	1	2	3	4	5	6	7	8	9	10	77	88	99	

Há várias acções que se podem desenvolver para melhorar as coisas em Portugal ou para evitar que corram mal. Durante os últimos 12 meses, fez alguma das seguintes coisas?

		Sim	Não	(Recusa)	(NS)	(NR)
B15	Contactou um político, um representante do governo central ou um representante do poder local	1	2	7	8	9
B16	Trabalhou para um partido político ou movimento cívico.	1	2	7	8	9
B17	Trabalhou numa organização ou associação de outro tipo	1	2	7	8	9
B18	Usou um emblema auto-colante de campanha/movimento	1	2	7	8	9
B19	Assinou uma petição	1	2	7	8	9
B20	Participou numa manifestação.	1	2	7	8	9
B21	Boicotou determinados produtos	1	2	7	8	9
B22	Comprou de propósito determinados produtos por razões de ordem política, ética ou ambiental.	1	2	7	8	9
B23	Deu dinheiro a um grupo ou organização política	1	2	7	8	9
B24	Participou em acções de protesto ilegais (corte de estrada, boicote eleitoral).	1	2	7	8	9

B28: Em política é costume falar-se de esquerda e direita. Como é que se posicionaria nesta escala, em que 0 representa a posição mais à esquerda e 10 a posição mais à direita?

Esquerda										Direita	(Recusa)	(NS)	(NR)
0	1	2	3	4	5	6	7	8	9	10	77	88	99

ANEXO 2: Questionário ESS

B29: Tudo somado, qual é o seu grau de satisfação com a vida em geral?

B30: De um modo geral qual o seu grau de satisfação com o estado actual da economia portuguesa?

B31: Qual é o seu grau de satisfação com a forma como o Governo está a actuar?

B32: Qual o seu grau de satisfação com o funcionamento da democracia em Portugal?

	Extrema- mente insatisfeito										Extrema- mente satisfeito	(Recusa)	(NS)	(NR)
B29	0	1	2	3	4	5	6	7	8	9	10	77	88	99
B30	0	1	2	3	4	5	6	7	8	9	10	77	88	99
B31	0	1	2	3	4	5	6	7	8	9	10	77	88	99
B32	0	1	2	3	4	5	6	7	8	9	10	77	88	99

B33: Utilizando a seguinte escala, diga, por favor, como avalia, no geral, o estado da Educação em Portugal, hoje em dia?

Extremamente mau											Extremamente bom				(Recusa)	(NS)	(NR)
0	1	2	3	4	5	6	7	8	9	10	77	88	99				

B34: E relativamente aos serviços de saúde em geral? Qual o seu grau de satisfação com os Serviços de Saúde em Portugal hoje em dia?

Extremamente mau											Extremamente bom				(Recusa)	(NS)	(NR)
0	1	2	3	4	5	6	7	8	9	10	77	88	99				

VOU AGORA FAZER-LHE ALGUMAS PERGUNTAS SOBRE MIGRAÇÕES

Da seguinte lista de aspectos diga qual a importância que deve ser dada a cada um deles para deixar vir para cá alguém que nasceu, foi educado e viveu fora de Portugal.

		Não deve ser dada importância nenhuma										Deve ser dada muita importância	(Recusa)	(NS)	(NR)
D10	ter boas qualificações académicas	0	1	2	3	4	5	6	7	8	9	10	77	88	99
D11	ter familiares próximos a viver cá	0	1	2	3	4	5	6	7	8	9	10	77	88	99
D12	saber falar Português	0	1	2	3	4	5	6	7	8	9	10	77	88	99
D13	ter formação cristã	0	1	2	3	4	5	6	7	8	9	10	77	88	99
D14	ser branco	0	1	2	3	4	5	6	7	8	9	10	77	88	99
D15	ser rico	0	1	2	3	4	5	6	7	8	9	10	77	88	99
D16	ter qualificações profissionais de que Portugal precisa	0	1	2	3	4	5	6	7	8	9	10	77	88	99
D17	Querer adaptar- se ao mesmo modo de vida dos Portugueses	0	1	2	3	4	5	6	7	8	9	10	77	88	99

ANEXO 2: Questionário ESS

Diga, por favor, em que medida concorda ou discorda com cada uma das seguintes afirmações.

		Concorda totalmente	Concorda	Nem concorda nem discorda	Discorda	Discorda totalmente	(Recusa)	(NS)	(NR)
D18	As pessoas que vêm viver e trabalhar para cá fazem com que os salários baixem.	1	2	3	4	5	7	8	9
D19	As pessoas que vêm viver e trabalhar para cá, em regra, prejudicam mais as expectativas económicas dos pobres do que dos ricos.	1	2	3	4	5	7	8	9
D20	As pessoas que vêm viver e trabalhar para cá ajudam a preencher lugares em que há falta de trabalhadores.	1	2	3	4	5	7	8	9
D21	Se as pessoas que vieram viver e trabalhar para cá estiverem desempregadas por muito tempo deviam ser obrigadas a ir embora.	1	2	3	4	5	7	8	9
D22	As pessoas que vieram viver para cá devem ter os mesmos direitos do que todas as outras pessoas.	1	2	3	4	5	7	8	9
D23	Se as pessoas que vieram viver para cá cometerem um crime grave , devem ser obrigadas a ir embora.	1	2	3	4	5	7	8	9
D24	Se as pessoas que vieram viver para cá cometerem qualquer crime , devem ser obrigadas a ir embora.	1	2	3	4	5	7	8	9

FINALMENTE GOSTARIA DE LHE FAZER ALGUMAS PERGUNTAS SOBRE SI E AS OUTRAS PESSOAS DO SEU AGREGADO FAMILIAR

F2: CODIFIQUE O SEXO []

F3: E em que ano nasceu?

--	--	--	--

Recusa 77 // Não sabe 88 // Não Responde 99

F7: Quantos anos completos de escolaridade terminou?

ESCREVER O NÚMERO:

--	--

Recusa 77
Não sabe 88
Não responde 99