



Engenharia de Dados Experimentais

SPSS - Comandos

Raquel Guiné



Conteúdo

1. Menu File
2. Menu Edit
3. Menu View
4. Menu Data
5. Menu Transform



IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Custom Utilities Add-ons Window Help

Rows are cases Columns are variables

Variable: 24 of 24 Variables

	ids	Rank	Gender	Athlete	Height	Weight
1	20183		Male	Non-athlete	65.92	192.65
2	20236	Freshman	Male	Athlete	88.11	
3	20243	Junior	Female	Non-athlete	65.99	128.40
4	20248	Freshman		Non-athlete	61.32	153.87
5	20255	Sophomore	Female	Non-athlete	65.75	
6	20278		Male	Non-athlete	73.66	179.20
7	20389		Male	Non-athlete	73.68	198.62

Data View Variable View

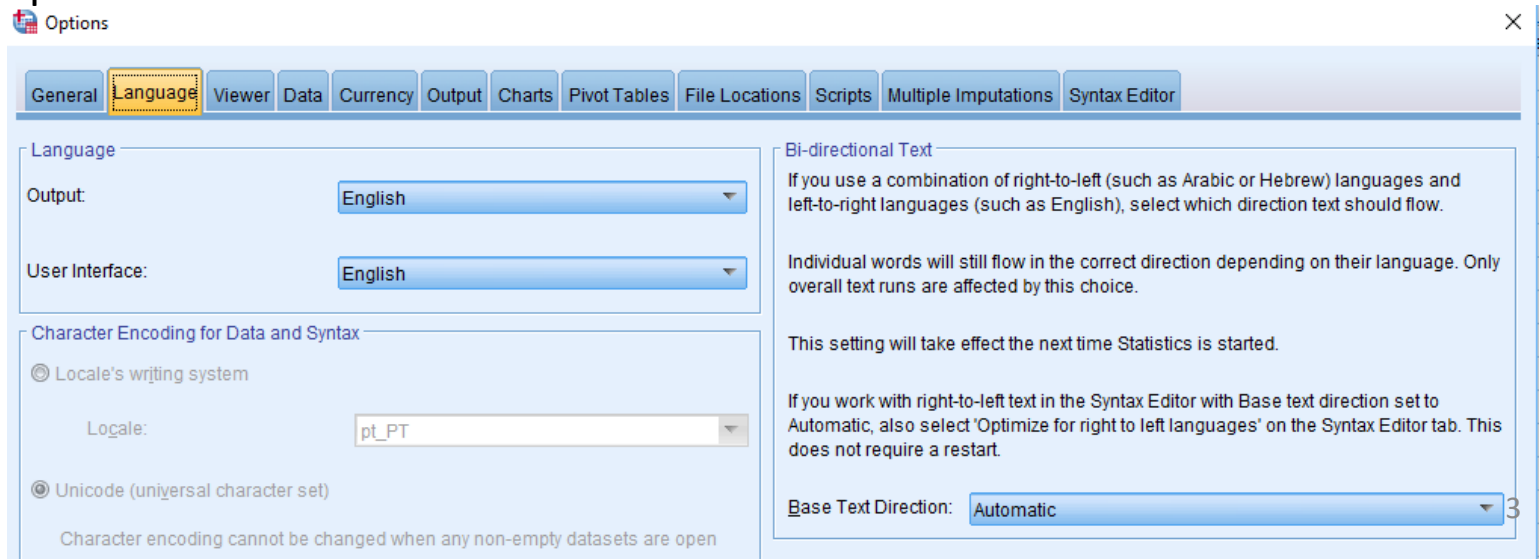
IBM SPSS Statistics Processor is ready. Cases: 190 Unicode: ON

NOTA:

Eu utilizo sempre o SPSS em Inglês, porque não gosto da tradução para Brasileiro, por isso todas as minhas referências são relativamente à versão Inglesa.

Nos vossos computadores devem instalar o SPSS e podem escolher trabalhar em PT ou EN:

- Abrem o SPSS
- Selcionam: Edit (ou Editar) / Options (opções)
- Depois escolhem o separador Language (Línguas) e colocam a língua da vossa preferência.



1. MENU FILE

- Comando **New**
 - ✓ **Data** – abre nova base de dados e para trabalhar com duas bases simultaneamente
 - ✓ **Sintax** – é um código informático (programa) – não vamos usar
 - ✓ **Output** – nova apresentação de resultados (quando fazemos alguma coisa num ficheiro data aberto, ele gera automaticamente um output)
- Comando **Open** – abre ficheiros que já estão gravados previamente.
- Se fizer **Open data** e não abrir **output**, ele cria um output novo que pode ser gravado.
- Os restantes comandos são básicos: close, print, save,...

2. MENU EDIT

- À partida alguns comandos estão desativados, mas ficam ativos se aplicáveis
Ex. No separador **Data View** seleccionar a linha 4 e com o rato (botão direito) fazer **Insert Cases** para introduzir uma nova linha.
Agora já podemos se quisermos fazer no menu **Edit** escolher o comando **Undo** para anular a introdução da linha.
Na vista de **Variable View** também posso introduzir novas variáveis com o mesmo procedimento explicado para a introdução de casos
- Podemos seleccionar uma linha e fazer **Edit copy** e depois **Edit paste** para a colar igual noutro sítio.
- Comando **Clear** – apaga uma linha ou coluna que esteja seleccionada

- Comando **Insert Variable** – insere uma nova variável

Ex: entre as V1.3 e V1.4 quero inserir uma nova variável: no modo **Data View** seleciono a coluna posterior (V1.4) e tenho duas alternativas: **Insert variable** com o botão direito do rato ou através do menu Edit.

Também se pode fazer inserindo linhas no modo **Variable View**.

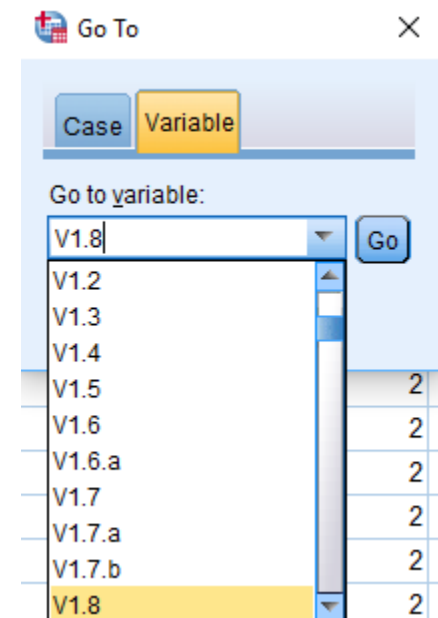
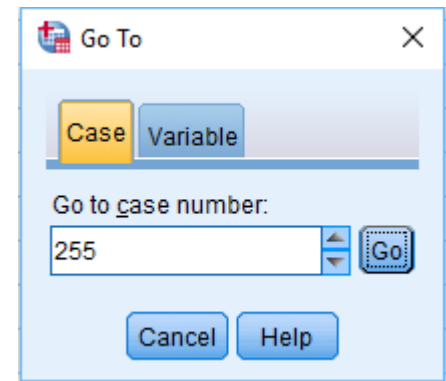
Nota: Se estiver no modo **Variable View** e clicar duas vezes sobre o número de uma determinada variável vai automaticamente para o separador **Data View**.

O inverso também acontece a passar de **Data View** Para **Variable View**.

(Tb se pode ir pelo separador ao fundo da tabela)

- Demonstração da importância do nº do questionário:
 - ✓ Seleciona-se a variável V1.1 (idade) porque se pretende ordenar por ordem crescente. Depois de selecionada a coluna idade, com o botão direito do rato escolher **Sort Ascending** – tudo se adaptou automaticamente e o nº do questionário é fundamental para se perceber os casos.
 - ✓ Nesse caso se não houvesse o nº do questionário já não sabíamos qual o inquérito a que correspondia cada caso.
 - ✓ Para voltar atrás: fazer ordenação ascendente pelo nº do questionário.
- Comando **Insert Case** – insere uma nova linha relativa a um caso, um sujeito, um questionário. Em alternativa ao menu **Edit**, seleciona-se a linha abaixo e com o botão direito do rato faz-se **Insert Case**.

- Comando **Go to Case** – permite identificar um caso para nos posicionarmos lá. Importante em bases muito grandes (Ex: 1000 pessoas/200 variáveis).
 - ✓ Abre uma janela que pede o nº do caso: ex: 255 e faz-se depois carregando **GO**.
- Comando **Go to Variable** – permite identificar uma variável para nos posicionarmos lá. Importante em bases muito grandes. Abre uma janela que pede o nº da variável: ex: V1.8 e faz-se depois **GO**.
- Comando **Options** – tem várias possibilidades, incluindo a mudança do idioma como já vimos.

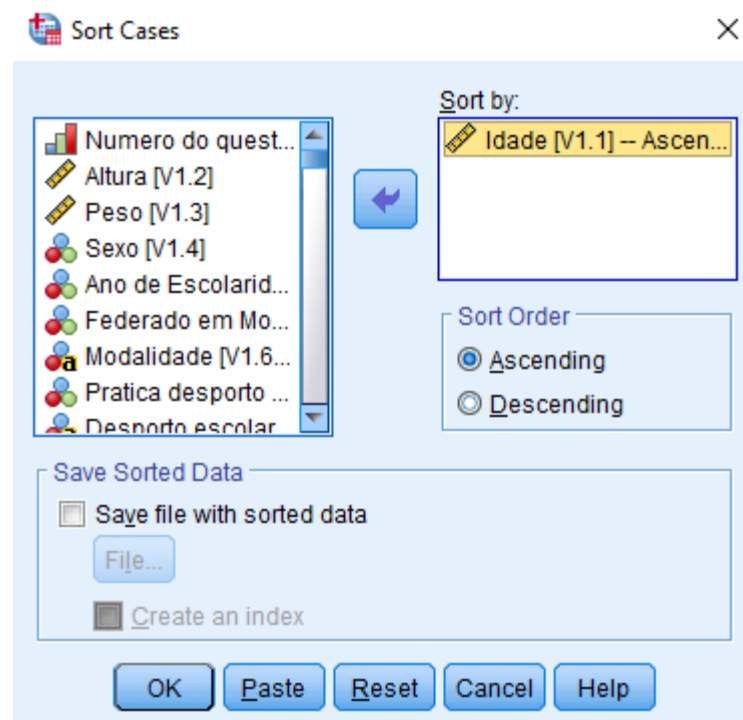


3. MENU VIEW

- Tem configurações de vista, incluindo variable e data view.

4. MENU DATA

- Comando **Sort Cases** – abra uma caixa, escolhe-se uma variável e manda-se com a seta para a caixa **sort by**: ordena ascendente ou descendente (igual ao que já vimos doutra forma)
- Comando **Sort Variables** – também ordena. Quando temos a base de dados construída podemos querer ordenar, p.ex., por ordem alfabética (não se usa normalmente)



- Comando **Select Cases** – seleciona casos.

Ex: só quero trabalhar com sexo Masculino.

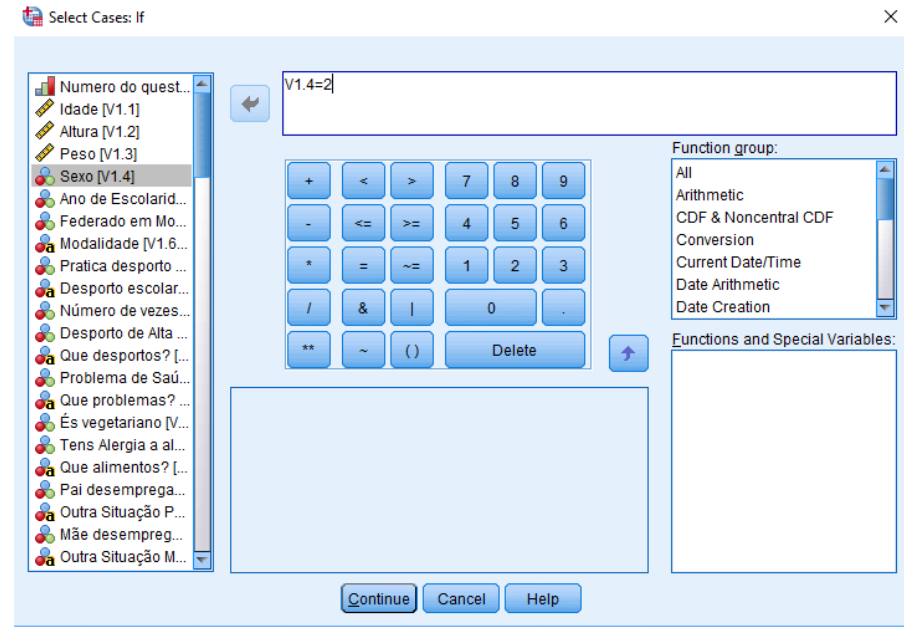
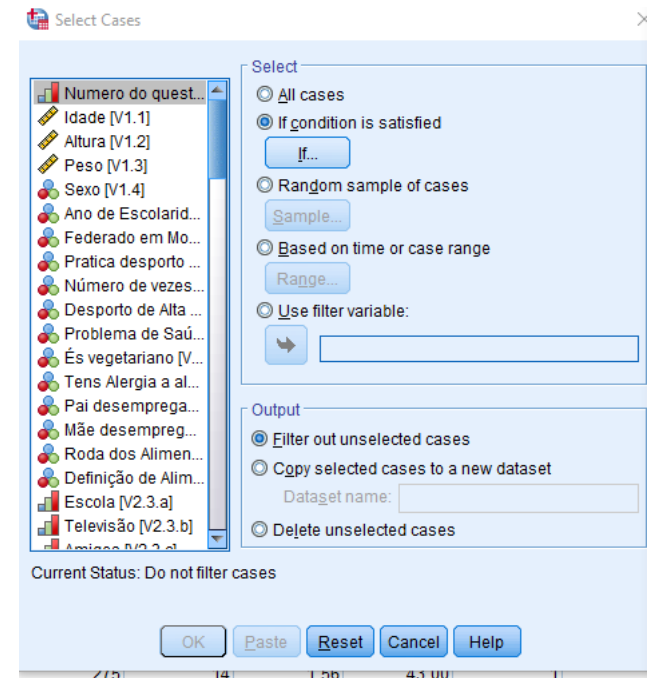
Na caixa escolho a variável sexo e escolho a 2ª opção e clico IF (por defeito estava selecionada a 1ª opção: all cases)

Escolho o valor da varável que quero, neste caso Masculino = 2 ou seja: $V1.4 = 2$

Fazer **continue** e **OK**

Na base os casos excluídos ficam com um traço.

55	55	13
56	56	13
57	57	13
58	58	12
59	59	12
60	60	12
61	61	12
62	62	13
63	63	12
64	64	11



- Escolho fazer **Analyze / Descriptive statistics / Frequencies**

✓ Escolho a variável **Ano de escolaridade** e envio para a caixa da direita e faço **OK**.

✓ No **OUTPUT** (ficheiro) aparece uma tabela com as frequências só para o sexo masculino – que são 425 casos.

✓ Sabemos que a maior parte dos rapazes frequentam o 7º ano (N = 101 casos)

- Posso fazer o mesmo depois só para mulheres com a opção select cases também. Dá resultados para 426 mulheres, das quais uma não identificou o ano de escolaridade

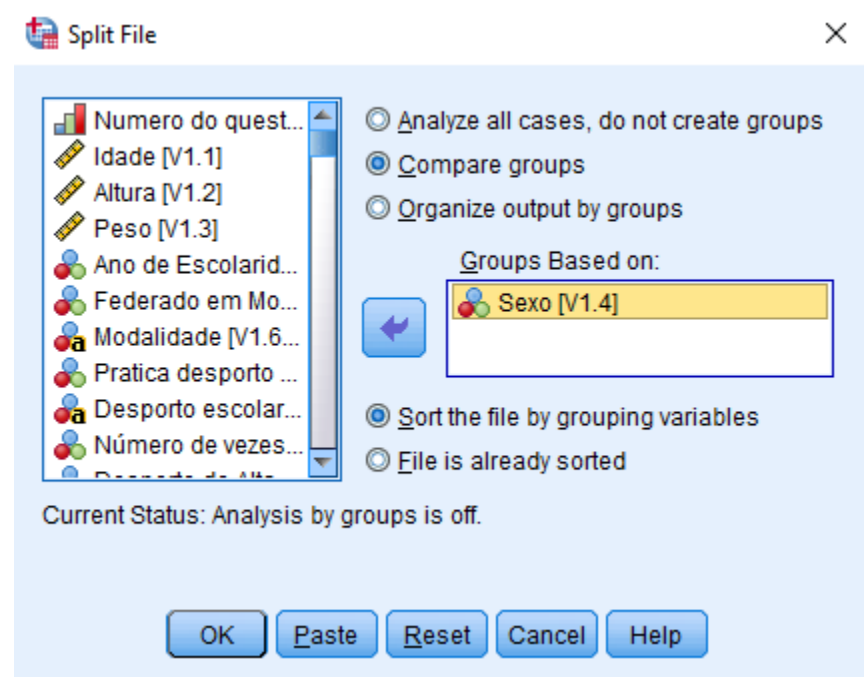


Ano de Escolaridade					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5	96	22,6	22,6	22,6
	6	85	20,0	20,0	42,6
	7	101	23,8	23,8	66,4
	8	62	14,6	14,6	80,9
	9	81	19,1	19,1	100,0
	Total	425	100,0	100,0	

- NOTA:** com o comando **Select Cases** faço o tratamento de cada grupo separadamente

- Comando **Split File** – trabalho ao mesmo tempo os diferentes grupos de forma separada

- ✓ Voltar a Data / Select cases / all cases
- ✓ Fazer comando **Data / Split file**
- ✓ Escolher a segunda opção: **Compare groups**
- ✓ Pegar a variável **sexo** e enviar com a seta para a caixa à direita, e fazer **OK**.



- Depois fazer: **Analyze / Descriptive statistics / Frequencies**
- Escolho a variável **V1.5 = Ano de Escolaridade** para a caixa da direita e faço **OK**.

- Aparece no **OUTPUT** só uma tabela mas com os dois grupos separados (Masc/Fem).

Ano de Escolaridade

Sexo			Frequency	Percent	Valid Percent	Cumulative Percent
Feminino	Missing	System	1	100,0		
	Valid	5	91	21,4	21,4	21,4
		6	100	23,5	23,5	44,9
		7	84	19,7	19,8	64,7
		8	74	17,4	17,4	82,1
		9	76	17,8	17,9	100,0
		Total	425	99,8	100,0	
	Missing	System	1	,2		
	Total		426	100,0		
Masculino	Valid	5	96	22,6	22,6	22,6
		6	85	20,0	20,0	42,6
		7	101	23,8	23,8	66,4
		8	62	14,6	14,6	80,9
		9	81	19,1	19,1	100,0
		Total	425	100,0	100,0	

NOTA:

Select Cases dá os dados separadamente e temos de os analisar em separado.

Split File dá os dados separadamente, mas são tratados em conjunto.

- Aparece no **OUTPUT** só uma tabela mas com os dois grupos separados (Masc/Fem).

Ano de Escolaridade

Sexo			Frequency	Percent	Valid Percent	Cumulative Percent
Feminino	Missing	System	1	100,0		
	Valid	5	91	21,4	21,4	21,4
		6	100	23,5	23,5	44,9
		7	84	19,7	19,8	64,7
		8	74	17,4	17,4	82,1
		9	76	17,8	17,9	100,0
		Total	425	99,8	100,0	
	Missing	System	1	,2		
	Total		426	100,0		
Masculino	Valid	5	96	22,6	22,6	22,6
		6	85	20,0	20,0	42,6
		7	101	23,8	23,8	66,4
		8	62	14,6	14,6	80,9
		9	81	19,1	19,1	100,0
		Total	425	100,0	100,0	

NOTA:

Select Cases dá os dados separadamente e temos de os analisar em separado.

Split File dá os dados separadamente, mas são tratados em conjunto.


- Interpretação do **Valid Percent**
- Ex. Fazer **Analyze / Descriptive Statistics / Frequencies** à variável **V1.8 (Desporto de alta competição)**

Desporto de Alta Competição

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Sim	159	18,7	18,8	18,8
	Não	688	80,8	81,2	100,0
	Total	847	99,4	100,0	
Missing	System	5	,6		
Total		852	100,0		

- Verifica-se que há 5 *missings*, que representam 0,6%.
- Às vezes, como neste exemplo, há casos omissos, e nesse casos temos de considerar a coluna do **Valid Percent**, muito embora seja importante dar a informação sobre os casos omissos.
- Na coluna do **Valid Percent** estão os casos usados efetivamente, sem os *missings* (casos não respondidos e deixados em branco na folha).
- Para a interpretação, na escrita de um trabalho devo dizer no texto que houve 0,6% de casos omissos, mas depois trabalho com a coluna do **Valid Percent**.
- Na apresentação dos dados em tabela num trabalho deve-se usar as colunas: **Frequency** e **Valid Percent** e normalmente não uso **Percent** nem **Cumulative Percent**.

1. MENU TRANSFORM

- Comando **Recode into Same Variables**: faz a recodificação dentro da própria variável, i.e., substitui os valores anteriores pelos novos.
- Comando **Recode into Different Variables**: faz a recodificação para uma nova variável, que pode estar já criada ou ser criada no momento . A variável original permanece lá


Eu prefiro fazer sempre este procedimento
- Numa escala de Likert, por exemplo, temos os níveis 1 a 5 (e normalmente estão na ordem ascendente). Mas pode ter havido algumas perguntas que foram colocadas de forma inversa – por conveniência, por exemplo para só ter um quadro.

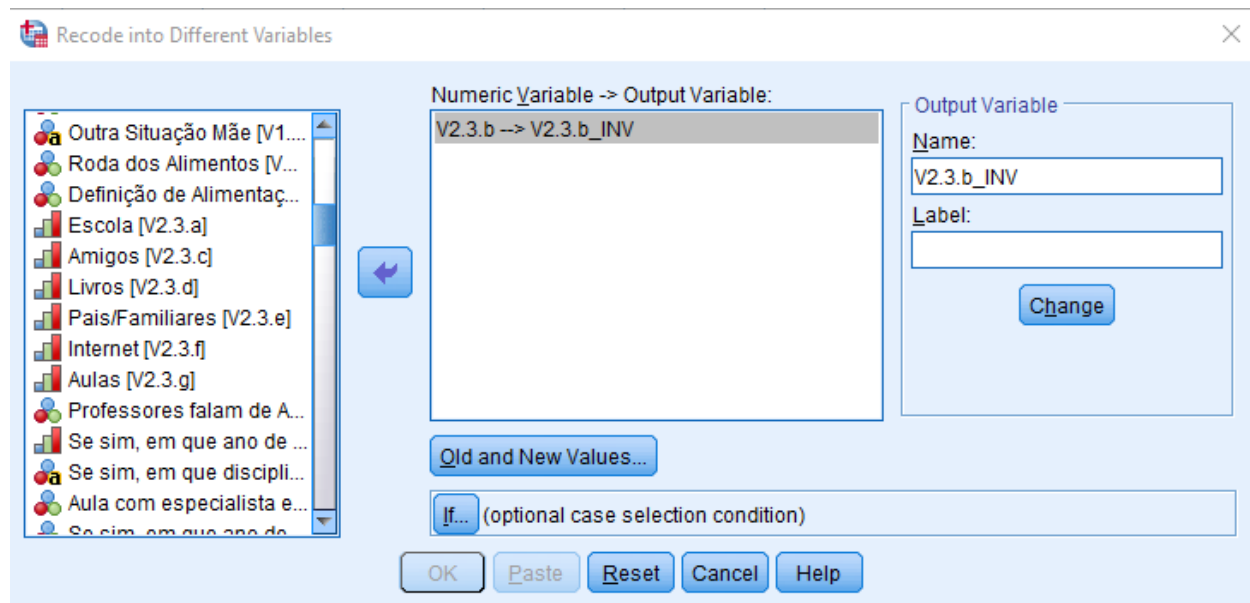
- Ex. vamos supor que para a variável **2.3 - opção Televisão**, queríamos inverter os valores:

Atual: 1 2 3 4 5
 Novo: 5 4 3 2 1

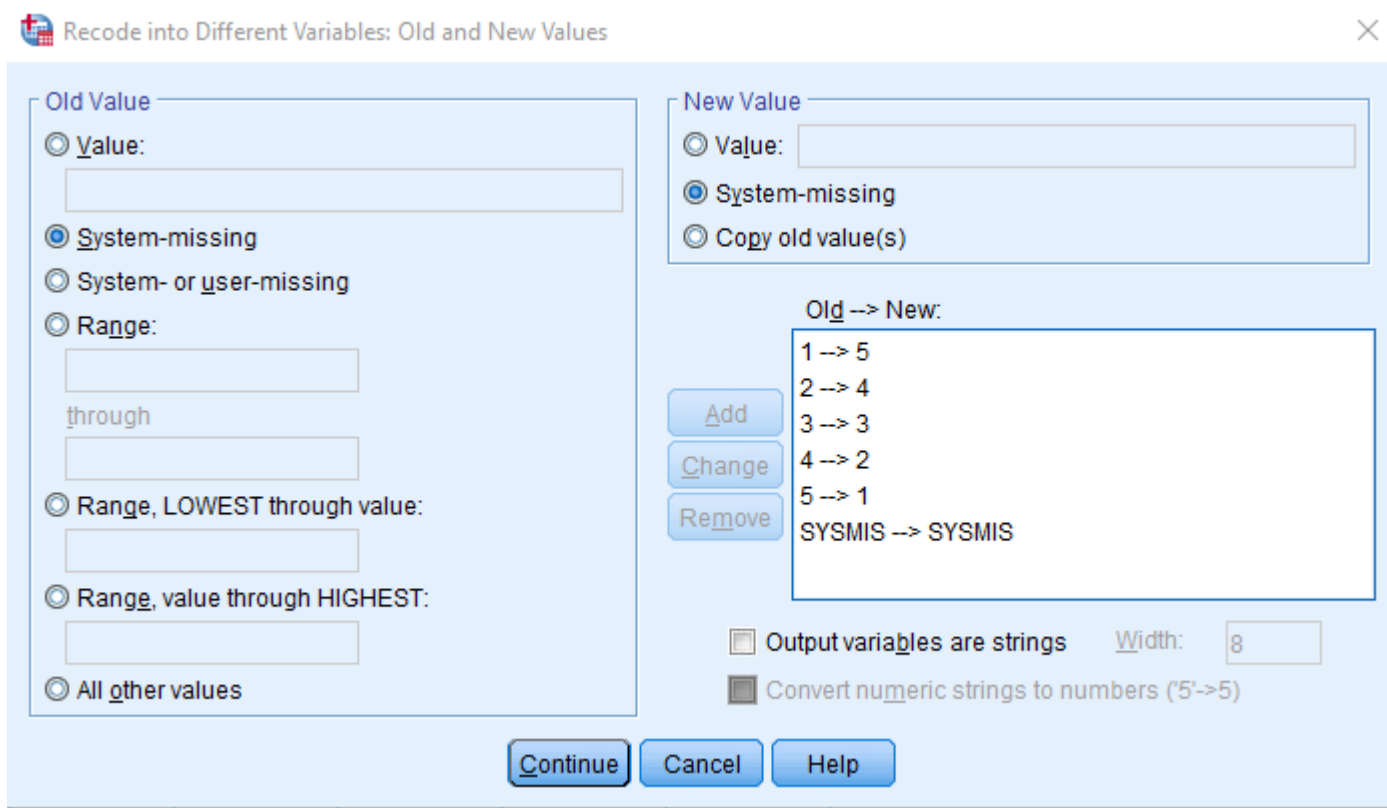
2.3. Onde lêes ou ouves falar em alimentação saudável?

	Nunca 1	2	3	4	Sempre 5
Escola					
Televisão					

- Faço: **Transform / Recode into Different Variables**
- Para a caixa das variáveis envio a variável em causa: **V2.3.b**
- Atribuo um nome à nova variável a criar: **V2.3.b_INV** e carrego em **Change**



- Selecciono a opção **Old and New Values** e faço a correspondência:
- Para cada um Old e New faço **Add** para ir completando o quadro.
- Lembrar que pode haver *missing* e nesse caso mantém-se *missing*.
- Depois faço **Continue** e **OK**
- Aparece no final a nova variável que é preciso “arranjar”, por exemplo tirar decimais e codificar o significado dos valores 1 a 5



Recode into Different Variables: Old and New Values

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:

1 --> 5

2 --> 4

3 --> 3

4 --> 2

5 --> 1

SYSMIS --> SYSMIS

☒ Output variables are strings Width: 8

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

Continue **Cancel** **Help**

- Outro exemplo:
 - ✓ Em relação à variável idade quero definir classes, dividir a idade em grupos etários de amplitude 2.
 - ✓ Para tal tenho primeiro de saber os valores da idade dos casos na amostra:
 - ✓ Faço: **Analyze / Descriptive Statistics / Explore**
 - ✓ Escolho var **V1.1=Idade**.
 - ✓ No **OUTPUT** dá informação sobre a idade mínima e máxima (10 e 18, respetivamente), pelo que vou ter 4 classes

- ✓ Classe 1: [10 – 12[(valores 10 a 11)
- ✓ Classe 2: [12 a 14[(valores 12 a 13)
- ✓ Classe 3: [14 a 16[(valores 14 a 15)
- ✓ Classe 4. [16 a 18] (valores 16 a 18)

No SPSS trabalha-se sempre em classes de limites reais e não em classes de limites aparentes.

Descriptives			
		Statistic	Std. Error
Idade	Mean	12,47	,056
	95% Confidence Interval for Mean	Lower Bound	12,36
		Upper Bound	12,58
	5% Trimmed Mean	12,43	
	Median	12,00	
	Variance	2,668	
	Std. Deviation	1,633	
	Minimum	10	
	Maximum	18	
	Range	8	
	Interquartile Range	3	
	Skewness	,226	,084
	Kurtosis	-,707	1,167

- Vai-se fazer a recodificação da variável:

- ✓ **Transform / Recode into Different Variables**

- ✓ Selecciono a var **V1.1=idade** para o quadro
 - ✓ À direita no **Output Variable** tenho de escrever o nome: **GRidade** (Gr significa em grupos)
 - ✓ Carrego **Change** e no quadro central da variável já foi atualizado.
 - ✓ Depois tenho de escolher a opção em baixo Old and New Values, e aparece mais um quadro, onde defino as gamas de variação (na opção **Range** defino os valores dos intervalos e faço **Add**).
 - ✓ Depois faço **Continue** e **OK**.

Recode into Different Variables: Old and New Values

Old Value

☒ Value:

☐ System-missing

☐ System- or user-missing

☒ Range:

16 through 18

☐ Range, LOWEST through value:

☐ Range, value through HIGHEST:

☐ All other values

New Value

☒ Value:

☐ System-missing

☐ Copy old value(s)

Old -> New:

SYSMIS -> SYSMIS
10 thru 11 -> 1
12 thru 13 -> 2
14 thru 15 -> 3

Add Change Remove

☐ Output variables are strings Width: 8

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

Continue Cancel Help

- Foi criada no fim a nova variável **GRidade** que tem de ser “arranjada”:
 - ✓ Definir significado dos valores
 - ✓ Retirar casas decimais
 - ✓ Variável como ordinal
 - ✓ Etc...
- No OUTPUT parece um gráfico que é o **Diagrama de Bigodes**
- Neste caso o digrama é simples, mas há casos mais complexos, por exemplo com *outliers*

Value Labels

Value: 4

Label: 16 a 18 anos

1 = "10 a 11 anos"
2 = "12 a 13 anos"
3 = "14 a 15 anos"

Add Change Remove

Spelling...

OK Cancel Help

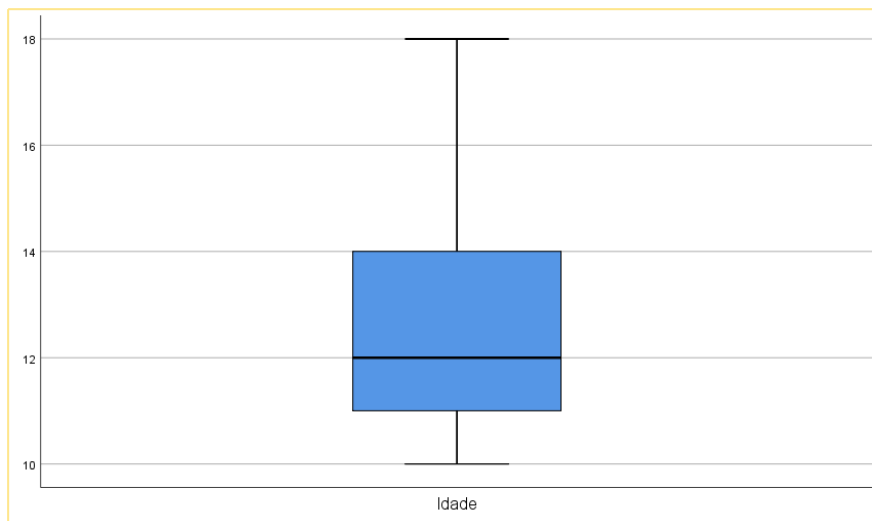
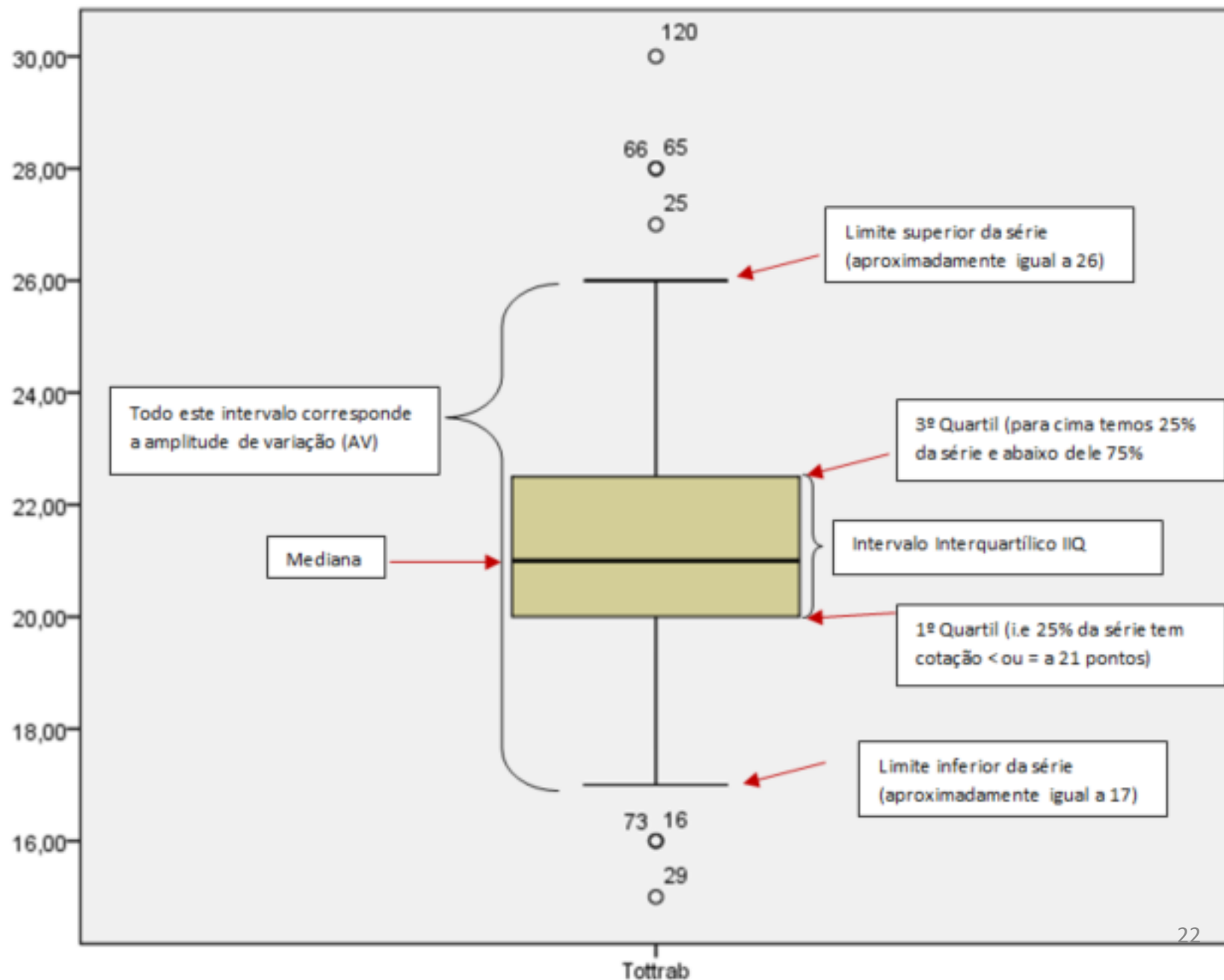


Diagrama de Bigodes



Interpretação do Diagrama de Bigodes

- Outliers: Símbolos (bolas) com os números 120, 66, 65... número de questionário correspondente;
 - Nos outliers aparecem bolas o ou asteriscos *.
 - Quando o outlier aparece com o: é um outlier moderado
 - Quando o outlier aparece com *: é um outlier severo
 - Quando há muitos outliers (ultrapassam uma determinada %) é necessário eliminá-los das análises – ir case select e não considerar esses casos? Em vez de os apagar da folha?
- A mediana divide a série em 2 partes iguais (50% dos casos estão acima da linha a grosso e 50% dos casos abaixo dessa linha);
- A amplitude de variação corresponde à diferença entre o limite superior e inferior da série;
- A diferença entre o 3º e o 1º quartil designa-se por intervalo interquartilico;
- Abaixo do 1º quartil estão 25% dos dados;

- Comando **Compute Variable**:
 - i. Se quiser combinar duas variáveis qualitativas
 - ii. Se quiser somar,
 - iii.

- Também faz a aglutinação de variáveis.

Ex: Calcular uma
média entre as
variáveis do **grupo 3.1**

	Nunca 1	2	3	4	Sempre 5
Na cantina da escola há informação afixada sobre hábitos de alimentação saudável.					
Vejo regularmente programas de televisão que me incentivam a praticar uma alimentação mais equilibrada.					
Converso com os meus amigos sobre o que comer, porque me preocupo com a minha aparência física.					
Gosto de ler livros que me ajudem a escolher o que comer, para me alimentar de forma mais saudável.					
Os meus pais/familiares preocupam-se em preparar-me refeições equilibradas.					
Procuo na internet os problemas de saúde provocados por uma alimentação incorreta.					
Quando o assunto da aula é alimentação, tento fazer o maior número de perguntas para me manter informado.					

- Como fazer: **Transform / Compute Variable**
- Aparece uma caixa – temos de definir o Target Variable (variável final que vai armazenar a média) – chamei **MED3.1**.
- Depois é necessário definir a fórmula do cálculo - **Numerical Expression**: da coluna da direita seleciono a var. 3.1ª e com a seta envio para a caixa OU se clicar duas vezes rápido nessa var. ela também vai para a caixa. Depois escrevo o sinal + a partir das teclas ou com o teclado azuis e repito para as restantes variáveis. Colocar os parêntesis e o sinal divisão. No final fazer **OK**.

Numeric Expression:

$(V3.1.a + V3.1.b + V3.1.c + V3.1.d + V3.1.e + V3.1.f + V3.1.g) / 7$

- Em alternativa posso usar a função Média (MEAN): Vou a **Function Group** e escolho **All**, e encontro **MEAN**, faço duplo click para selecionar e aparece MEAN(?,?). No local dos ?? coloco a 1ª e a última variáveis (têm de estar seguidas!). No final fazer **OK**.

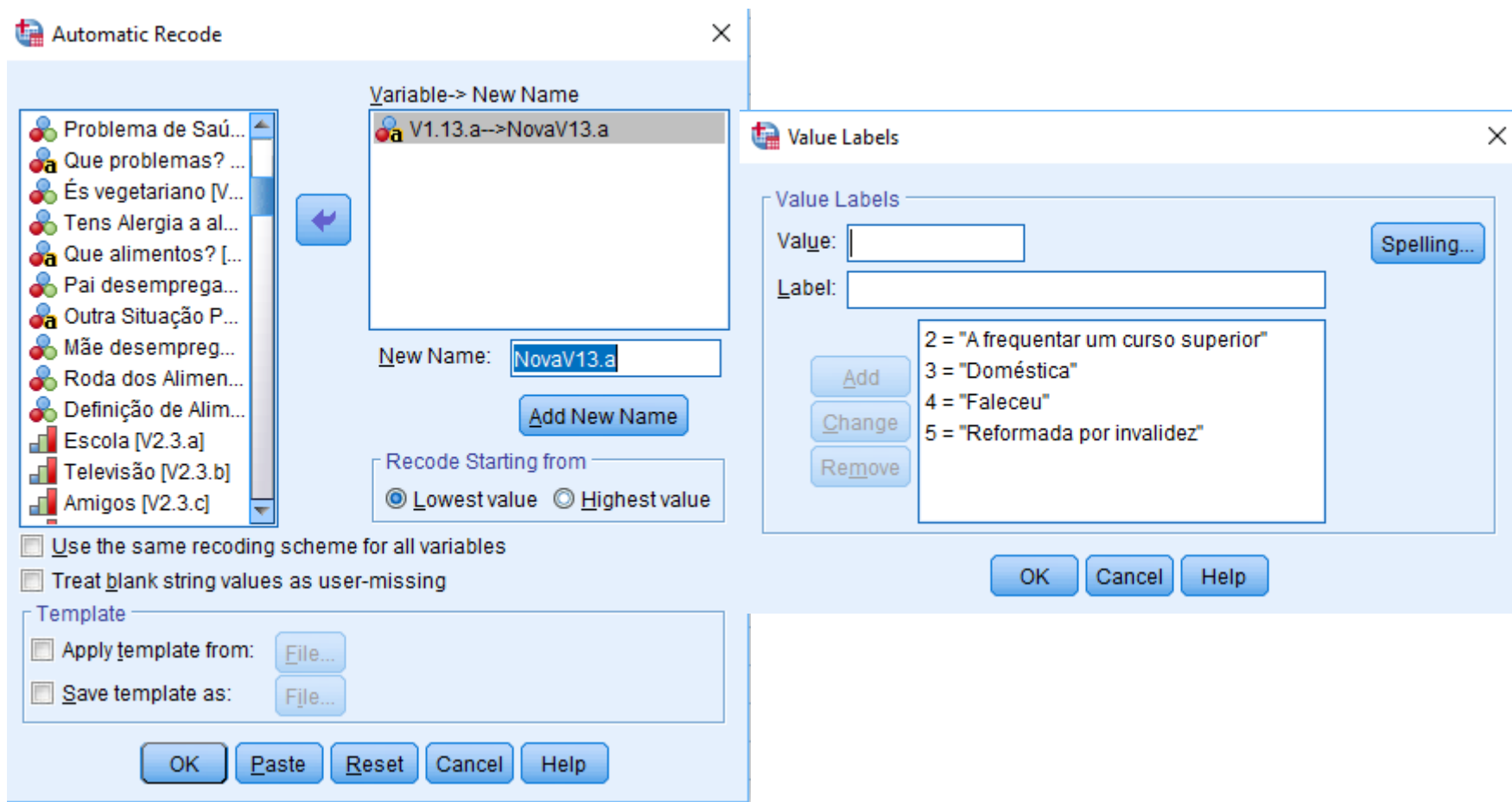
Numeric Expression:

MEAN(V3.1.a, V3.1.g)

- Comando **Automatic Recode**
- Faz a recodificação automática de variáveis.
- Se verificarmos no modo **Variable View**, há algumas variáveis que são string (um conjunto de caracteres que se escreve: ex. para descrever a profissão)
- Se quero recodificar a variável string em numérica – uso o automatic recode.
- Faço: **Transform / Automatic Recode**
- Na lista das variáveis à esquerda aparecem algumas com um a – essas são todas string
- Seleciono a variável **V13.a=Outra Situação da Mãe** e movo para a caixa **Variable -> New Name**
- E depois escrevo na caixa de New Name: **NovaV13.a**
- **Add New Name** e **OK**.

Isto só é possível em variáveis string em que o nº de códigos diferentes não sejam demasiado elevados.

- Neste caso a variável **NovaV13.a** tem 4 códigos diferentes – pode-se ver na coluna correspondente depois de criada.
- Automaticamente na vista Variable View já foram introduzidos os códigos (Values)

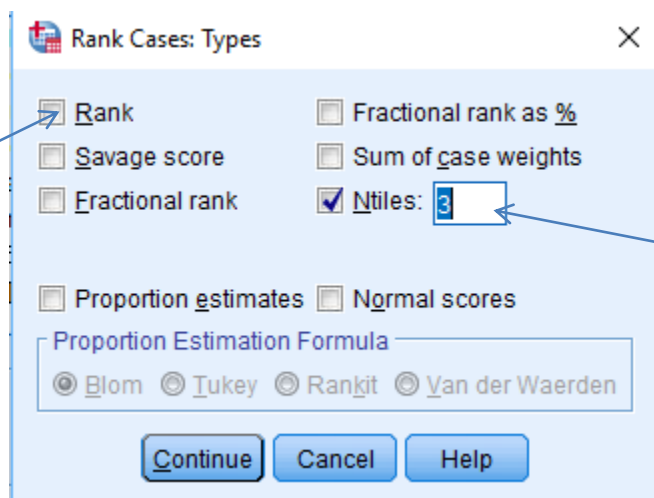


- Comando **Rank Cases**: só para variáveis quantitativas
- Classe com frequências homogéneas – são criadas no SPSS. Interessa o tamanho de cada um dos grupos.
- Por exemplo, quero criar 3 grupos homogéneos na variável **V.2.6.1.b (Copos de água)**, que varia entre 0 e 50:

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Copos de Água	831	0	50	6,48	3,753
Valid N (listwise)	831				

- Faço: **Transform / Rank Cases**
- Escolho **Rank Types** e abre uma nova janela:
- Continue** e **OK**.

Por defeito está
Rank (ordenar)
Desmarco essa
opção



Marco a opção
NTiles,
E escolho o nº de classes:
neste caso quero 3 classes

- Foi agora criada uma variável automaticamente NV2.6.1 (que contém 3 classes).
NOTA: ao criar ficou com um ponto no final por ter a letra b à frente, mas teve de ser apagado pois dava problemas
- Se quiser confirmar, coloco o rato em cima da variável no cabeçalho da coluna e com o botão direito do rato faço Sort Descending.
- Criei efetivamente 3 classes, mas é agora necessário identificá-las.
- Para tal, preciso de saber quem faz parte de cada grupo:
- Faço: **Analyze / Descriptive Statistics / Frequencies**
- Selecciono a variável **V2.6.1.b + NV2.6.1** a última variável : Percentile....
- Vão as duas para a caixa da direita e faço **OK**.
- Vamos agora ver o **OUTPUT**:
 - ✓ A Tabela mais pequena no final dá o resumo e podemos daí saber cada classe.
 - ✓ Para isso temos de ver a percentagem acumulada (Cumulative percent)

Grupo 1 20,9%

Grupo 2 66,2%

Grupo 3 100%

Percentile Group of V2.6.1.b					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	174	20,4	20,9	20,9
	2	376	44,1	45,2	66,2
	3	281	33,0	33,8	100,0
	Total	831	97,5	100,0	
Missing	System	21	2,5		
Total		852	100,0		

- Depois é necessário saber quem pertence a cada classe etária através da tabela maior (ver página seguinte).

Verifica-se assim que:

Cumulative % 20,9 corresponde a 4 copos	⇒	1ª Classe vai desde 0 até 4
Cumulative % 66,2 corresponde a 6 copos	⇒	2ª Classe vai desde 5 até 6
Cumulative % 100,0 corresponde a 50 copos	⇒	3ª Classe vai desde 7 até 50

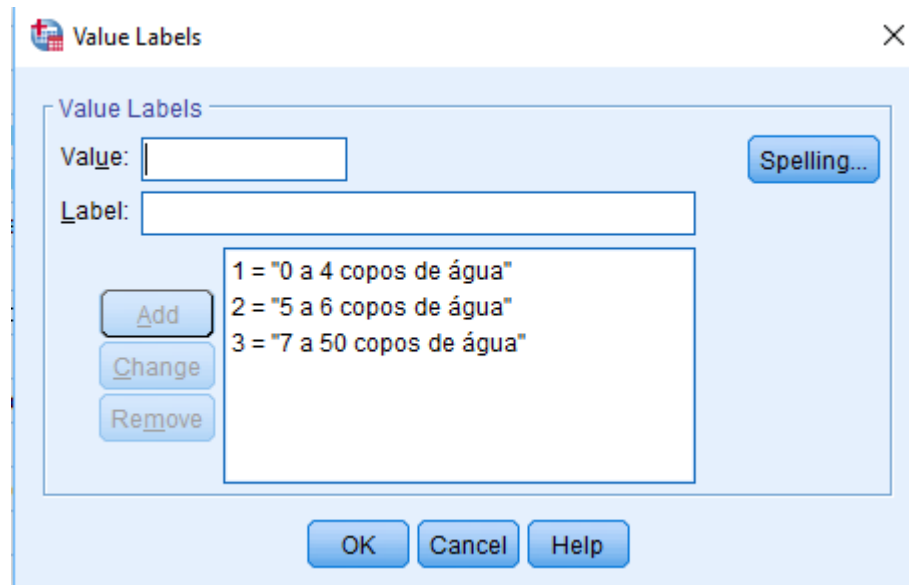
- Agora é necessário voltar à base de dados na vista Variable View e fazer:
- Renomear a variável NV2.6.1 para GRCopAg
- Retirar as casas decimais
- Apagar o label (para não gerar confusão) e
- Especificar os valores.

Copos de Água					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	1	,1	,1	,1
	1	3	,4	,4	,5
	2	26	3,1	3,1	3,6
	3	54	6,3	6,5	10,1
	4	90	10,6	10,8	20,9
	5	273	32,0	32,9	53,8
	6	103	12,1	12,4	66,2
	7	62	7,3	7,5	73,6
	8	62	7,3	7,5	81,1
	9	8	,9	1,0	82,1
	10	103	12,1	12,4	94,5
	11	3	,4	,4	94,8
	12	11	1,3	1,3	96,1
	13	2	,2	,2	96,4
	14	1	,1	,1	96,5
	15	11	1,3	1,3	97,8
	17	1	,1	,1	98,0
	19	1	,1	,1	98,1
	20	11	1,3	1,3	99,4
	30	4	,5	,5	99,9
	50	1	,1	,1	100,0
Total		831	97,5	100,0	
Missing	System	21	2,5		
Total		852	100,0		31

- Especificar os valores na variável

GRCopAg:

- 1ª Classe vai desde 0 até 4
- 2ª Classe vai desde 5 até 6
- 3ª Classe vai desde 7 até 50



Value Labels

Value:

Label:

Add

Change

Remove

1 = "0 a 4 copos de água"

2 = "5 a 6 copos de água"

3 = "7 a 50 copos de água"

Spelling...

OK Cancel Help

- Agora fazer: **Analyse / Descriptive Statistics / Frequencies**
- Tem lá duas variáveis em memória: V2.6.1.b e NV2.6.1 ,
- Faço: **Reset** e Selecionar variável: **GRCopAg**

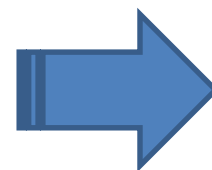
- Ver no **OUTPUT** (página seguinte):

✓ Grupo maior tem N = 376 (veja na coluna Frequency)

✓ Grupo menor tem n = 174

- Dividindo: $376 / 174 = 2,16 > 1,5$

Os grupos **NÃO** são homogêneos



Não devo usar
testes
paramétricos

Statistics

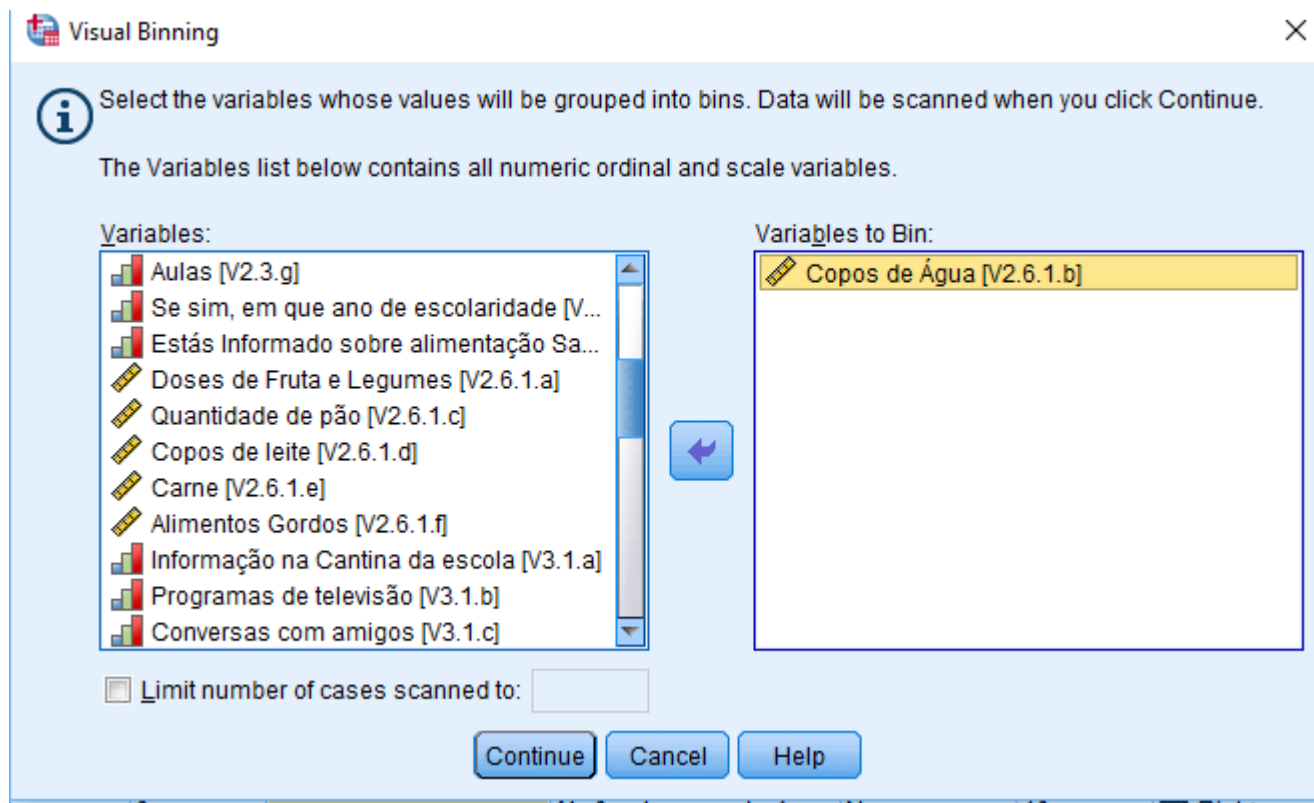
GRCopAg

N	Valid	831
	Missing	21

GRCopAg

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 a 4 copos de água	174	20,4	20,9	20,9
	5 a 6 copos de água	376	44,1	45,2	66,2
	7 a 50 copos de água	281	33,0	33,8	100,0
	Total	831	97,5	100,0	
Missing	System	21	2,5		
Total		852	100,0		

- Comando **Visual Binning**: cria classes de limites aparentes automaticamente com intervalos iguais: cria um histograma
- Faço: **Transform / Visual Binning** (permite dividir uma série em grupos com intervalos iguais)
- Aparece uma janela em que estão disponíveis só as variáveis quantitativas
- Ex. seleciono a variável: **V2.6.1.b = Copos de água**
- Continue**



- Abre uma nova janela com um histograma.
- Tenho de designar uma outra variável: Binned Variable = **GRTCoAgua**
- Por defeito, já lá aparece o valor mínimo: 0 e o valor máximo: 50

Visual Binning

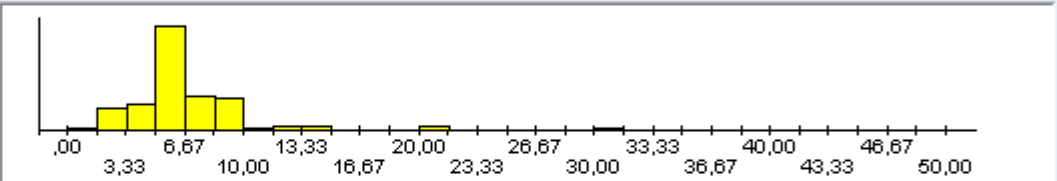
Scanned Variable List:

- Copos de Água [V2.6....]

Name: V2.6.1.b Label: Copos de Água

Binned Variable: GRTCoAgua Label: Copos de Água (Binned)

Minimum: 0 Nonmissing Values Maximum: 50



Enter interval cutpoints or click Make Cutpoints for automatic intervals. A cutpoint value of 10, for example, defines an interval starting above the previous interval and ending at 10.

Grid:

	Value	Label
1		HIGH
2		

Cases Scanned: 852

Missing Values: 21

Copy Bins

- From Another Variable...
- To Other Variables...

Upper Endpoints

- ☒ Included (<=)
- ☐ Excluded (<)

Make Cutpoints...

Make Labels

☐ Reverse scale

OK Paste Reset Cancel Help

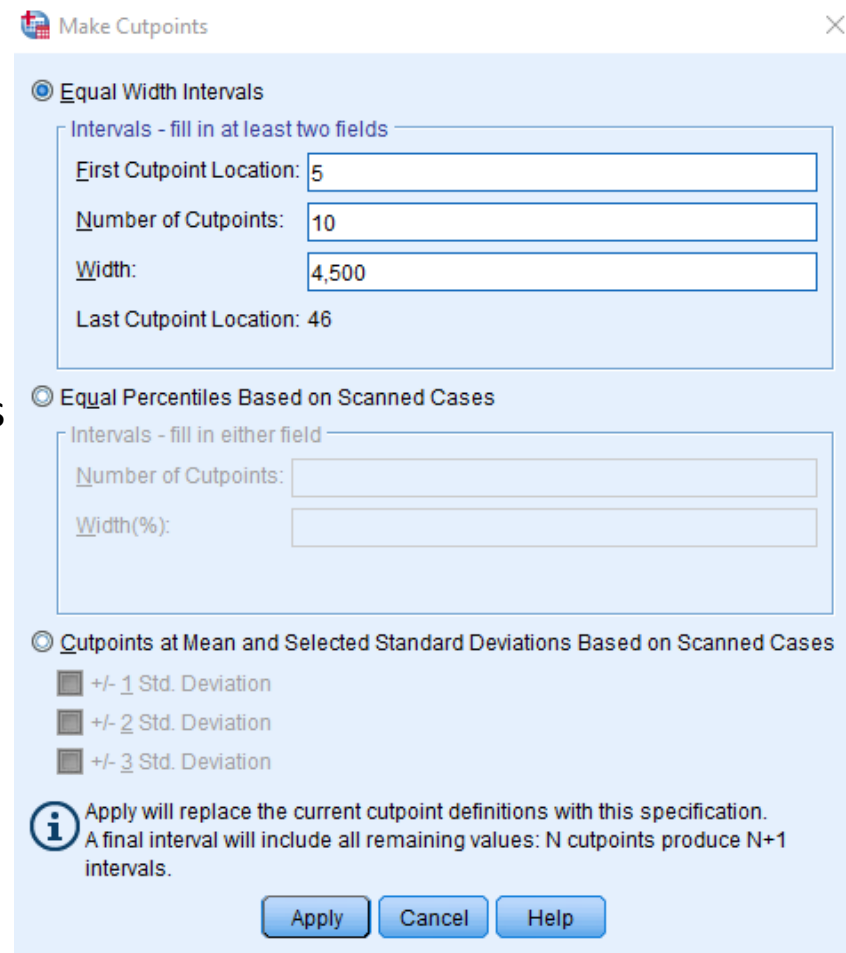
- Selecciono a opção **Make Cutpoints** (à direita em baixo)
- Aparece mais uma janela onde tenho de especificar o que quero:

- First Cutpoint location: 5

- Number of cutpoints: 10 → dá 10 classes

Depois de inserir estes dois, se clicar com o rato na caixa de With, já aparece automaticamente 4,500

- Carregar **Apply**
- Escolher **Make Labels** e ele dá os nomes automaticamente
- **OK.**
- Dá mensagem a dizer que se vai criar uma variável nova.
- **OK.**



Make Cutpoints

☒ **Equal Width Intervals**

Intervals - fill in at least two fields

First Cutpoint Location: 5

Number of Cutpoints: 10

Width: 4,500

Last Cutpoint Location: 46

☐ **Equal Percentiles Based on Scanned Cases**

Intervals - fill in either field

Number of Cutpoints:

Width(%):

☒ **Cutpoints at Mean and Selected Standard Deviations Based on Scanned Cases**

☐ +/- 1 Std. Deviation

☐ +/- 2 Std. Deviation

☐ +/- 3 Std. Deviation

i Apply will replace the current cutpoint definitions with this specification. A final interval will include all remaining values: N cutpoints produce N+1 intervals.

Apply Cancel Help

FIM