

# CEsA

*Centro de Estudos sobre África e do Desenvolvimento*  
Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa

Colecção

*Documentos de Trabalho*

*nº 85*

Carlos Sangreman  
Nuno Cunha  
Bruno Damásio

## **MANUAL BÁSICO DO PASW**

Lisboa  
2010

*O CEa não confirma nem infirma  
quaisquer opiniões expressas pelos autores  
nos documentos que edita.*

### **Resumo**

Pretende-se explicar um modo rápido e eficaz de se utilizar o PASW Statistics 18 (ex-SPSS) no manejo e análise de dados. Ou seja, a forma como a abordagem de dados pode ser realizada no domínio das ciências sociais utilizando a estatística descritiva.

Salienta-se a melhor forma de contornar diversas vicissitudes intrínsecas ao processo de elaboração de uma base de dados no programa em questão. Frisa-se não só a introdução de conceitos estatísticos elementares, bem como a sua aplicação no âmbito dos procedimentos inerentes ao tratamento de informação recolhida em contexto de inquérito.

**Palavras-chave:** base de dados, célula, gráfico, inquérito, matriz, PASW (ex-SPSS), tabela de frequências, variável

## **Preâmbulo**

O texto deste pequeno Manual começou a ser construído para uma formação na Guiné- Bissau, em 1999, no âmbito do projecto “Observatório do Bem-estar num bairro sub urbano de Bissau” financiado pelo Ministério do Trabalho e Solidariedade (actual Ministério do Trabalho e Segurança Social) e executado pela ONG guineense Associação para o Desenvolvimento (AD).

A versão, então preliminar, elaborada por mim, teve melhorias profundas da autoria do Dr. Nuno Cunha, nessa altura a trabalhar no MTS e hoje na OIT, para uma nova formação no mesmo país, num projecto similar, em 2000, mas alargado a outros dois bairros de Bissau, com financiamento da União Europeia, e execução de três ONG portuguesas – Instituto Marquês Vale Flor, Associação para a Cooperação entre Povos e Instituto de Estudos de Desenvolvimento - e três guineenses - AD, AIFA-PALOP e TINIGUENA.

O texto foi também distribuído aos investigadores universitários do projecto “*Urbanização Acelerada em Maputo e Luanda – Impacto da Guerra e das Transformações Sócio - Económicas (Décadas de 80 e 90)*”, com financiamento da Fundação para a Ciência e Tecnologia, e execução do CESA/ISEG e do CEA/ISCTE, entre 1998 e 2004.

Para uma formação no âmbito da cooperação em 2001-2002 com o então Ministério do Emprego e Formação e Integração Social de Cabo Verde a versão do texto foi mais uma vez alterada para incorporar a evolução de versões do programa, e com a amável cedência da empresa PSE, representante do programa PASW (ex - SPSS) em Portugal, de elementos tirados do manual que utilizam nos seus cursos de formação.

Em 2010, o CESA resolveu colocar este Manual on line de forma a poder servir de apoio aos mestrados e programas de doutoramento que se multiplicam com a reforma de Bolonha. Assim, foi solicitado ao Dr. Bruno Damásio que actualizasse o texto de acordo com a última versão do programa e fizesse uma revisão geral de todo o conteúdo.

Este Manual é assim uma obra viva, evolutiva, que foi expressamente construída para ensinar em África o apuramento estatístico de dados a pessoas com esse tipo de trabalho e com preocupações muito práticas, de forma a esses técnicos terem as suas capacidades alargadas de forma sustentada, sem aumentar a sua dependência do exterior.

Os autores,

Carlos Sangreman

Nuno Cunha

Bruno Damásio

# 1- Generalidades do PASW

De um ponto de vista genérico, podemos afirmar que a ambiência do PASW se estrutura em:

- 1) **Janelas**, salientando-se a importância das janelas *Data Editor* (que se subdivide nas janelas *Data View – Janela de Dados* e na *Variable View – Janela de Variáveis*) e *Output Viewer*.

A primeira apresenta o conteúdo dos ficheiros de dados, permitindo a sua edição e transformação e a segunda expõe-nos os resultados estatísticos, bem como as tabelas e os gráficos;

| ID | HH | ID  | Parentesco   | Sexo      | Idade | Nasc     | Tipo                   | Prof          | Pos_Prof  | For_Inf       | Instr      | Lingua                     | Rel             |           |
|----|----|-----|--------------|-----------|-------|----------|------------------------|---------------|-----------|---------------|------------|----------------------------|-----------------|-----------|
| 1  |    | 101 | chefe de fa. | feminino  | 72    | maputo   | domestica              | domestica     | domestica | não estudou   | makhua     | maometana                  |                 |           |
| 2  | 1  | 102 | sobrinho/a   | masculino | 46    | maputo   | serralheiro            |               | 99        | 99            | 4ª classe  | makhua                     | maometana       |           |
| 3  | 1  | 103 | noiva/genro  | feminino  | 33    | gaza     | domestica              | domestica     | domestica | 9º ano        | changana   | catolica                   |                 |           |
| 4  | 2  | 201 | chefe de fa. | masculino | 50    | zambúzia | carpinteiro            | conta propria | informal  | não estudou   | n tams     | catolica                   |                 |           |
| 5  | 2  | 202 | cônjuge      | feminino  | 39    | maputo   | domestica              | domestica     | domestica | não estudou   | ronga      | ziona                      |                 |           |
| 6  | 2  | 203 | filho/a      | masculino | 15    | maputo   | estudante              | crianca       | crianca   | até 4ª classe | ronga      | ziona                      |                 |           |
| 7  | 2  | 204 | filho/a      | feminino  | 3     | maputo   | crianca                | crianca       | crianca   | crianca       | crianca    | ronga                      | crianca         |           |
| 8  | 3  | 301 | chefe de fa. | feminino  | 48    | maputo   | vendedora/eira / co... | conta propria | informal  | até 4ª classe | ronga      | igreja belém de moçambi... |                 |           |
| 9  | 3  | 302 | filho/a      | feminino  | 24    | maputo   | vendedora/eira / co... | conta propria | informal  | até 4ª classe | ronga      |                            | 99              |           |
| 10 | 3  | 303 | filho/a      | feminino  | 19    | maputo   | domestico              | domestico     | domestico | 99            | ronga      | maometana                  |                 |           |
| 11 | 3  | 304 | filho/a      | feminino  | 22    | maputo   | vendedora/eira / co... | conta propria | informal  | até 4ª classe | ronga      | maometana                  |                 |           |
| 12 | 3  | 305 | neto/a       | masculino | 6     | maputo   | crianca                | crianca       | crianca   | crianca       | crianca    | crianca                    |                 |           |
| 13 | 4  | 401 | chefe de fa. | masculino | 99    | maputo   | guarda                 | conta alheia  | informal  | 99            | ronga      |                            | 99              |           |
| 14 | 4  | 402 | cônjuge      | feminino  | 56    | maputo   | vendedora/eira / co... |               | 99        | informal      | 99         | changana                   | 12 apóstolos    |           |
| 15 | 5  | 501 | chefe de fa. | feminino  | 43    | gaza     | vendedora/eira / co... |               | 99        | informal      | até 9º ano | changana                   | velho apóstolos |           |
| 16 | 5  | 502 | filho/a      | feminino  | 18    | maputo   | estudante              | crianca       | crianca   | crianca       | até 9º ano | portugues                  | IURD            |           |
| 17 | 5  | 503 | filho/a      | feminino  | 16    | maputo   | estudante              | crianca       | crianca   | crianca       | até 9º ano | portugues                  | nazare          |           |
| 18 | 5  | 504 | filho/a      | feminino  | 6     | maputo   | crianca                | crianca       | crianca   | crianca       | crianca    | portugues                  | crianca         |           |
| 19 | 6  | 601 | chefe de fa. | feminino  | 36    | maputo   | vendedora/eira / co... |               | 99        | informal      | até 9º ano | portugues                  | maometana       |           |
| 20 | 6  | 602 | irmão        | masculino | 29    | maputo   | pedreiro               | conta propria | informal  | 4ª classe     | portugues  |                            | 99              |           |
| 21 | 6  | 603 | irmão        | feminino  | 30    | maputo   | vendedora/eira / co... | conta propria | informal  | 4ª classe     | portugues  |                            | maometana       |           |
| 22 | 6  | 604 | irmão        | feminino  | 27    | maputo   |                        |               | 99        | 99            | 99         | até 4ª classe              | portugues       | 99        |
| 23 | 6  | 605 | irmão        | masculino | 22    | maputo   | pedreiro               | conta propria | informal  | até 9º ano    | portugues  |                            | maometana       |           |
| 24 | 6  | 606 | filho/a      | feminino  | 18    | maputo   |                        |               | 99        | 99            | 99         | até 9º ano                 | portugues       | maometana |
| 25 | 6  | 607 | filho/a      | feminino  | 13    | maputo   | estudante              | crianca       | crianca   | crianca       | 4ª classe  | portugues                  |                 | maometana |

← Data Editor

```

GET
FILE=C:\Users\BDemasio\Desktop\CEA\II_maputo_final.sav*.
DATASET NAME DataSet1 WINDOW=PSORT.
DESCRIPTIVES VARIABLES=ID_HH ID_ID II_Parentesco II_Sexo II_Idade II_Nasc II_Tipo_Prof II_Pos_Prof II_For_Inf
/STATISTICS=MEAN STDDEV MIN MAX.
    
```

| Identificação Agregada                    | N    | Minimum | Maximum | Mean     | Std. Deviation |
|---|------|---------|---------|----------|----------------|
| Identificação Agregada                    | 4620 | 1       | 723     | 357,38   | 207,448        |
| Identificação Individual                  | 4620 | 101     | 72305   | 35677,41 | 20763,975      |
| Parentesco em relação ao chefe de família | 4620 | 1       | 18      | 4,16     | 3,271          |
| Idade                                     | 4620 | 1       | 97      | 23,13    | 16,949         |
| Provincia onde nasceu                     | 4620 | 1       | 18      | 1,83     | 1,786          |
| Tipo de profissão                         | 4529 | 1       | 111     | 23,67    | 30,140         |
| Posição Profissional                      | 4491 | 1       | 111     | 47,07    | 37,591         |
| Sector formal ou informal                 | 4434 | 1       | 111     | 47,24    | 37,529         |
| Valid N (listwise)                        | 4337 |         |         |          |                |

← Output Viewer

- 2) **Caixas de Diálogo**, relevando-se que permitem a selecção de variáveis e opção de análise, são um dos elementos principais no manejo do PASW e são utilizadas, segundo a imagem abaixo, são utilizadas para seleccionar variáveis e opções para análise;

Variáveis Fonte → [Identificação Agregada, Identificação Individual, Parentesco em relação ao chefe de família, Sexo [II\_Sexo], Idade [II\_Idade], Provincia onde nasceu, Tipo de profissão [II\_Tipo\_Prof], Posição Profissional, Sector formal ou informal]

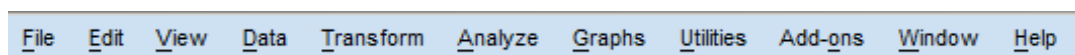
Variable(s): [Empty box]

Buttons: Statistics..., Charts..., Format..., Display frequency tables, OK, Paste, Reset, Cancel, Help

← Variáveis de Destino

3) **Menus**, correspondendo cada um deles a determinadas funcionalidades, de acordo com a imagem. Tais como *File* (criar / abrir / ler / imprimir ficheiros), *Edit* (modificar / copiar texto; parametrizar opções diversas), *View* (funções de activação), *Data* (alteração global dos dados), *Transform* (alteração / (cálculo de variáveis), *Analyze* (selecção de procedimentos estatísticos), *Graphs* (geração de gráficos), *Utilities* (obtenção de informação acerca das variáveis), *Window* (funções relativas à visualização das janelas), *Help* (obtenção de ajuda sob determinadas formas);

Menu da janela de dados



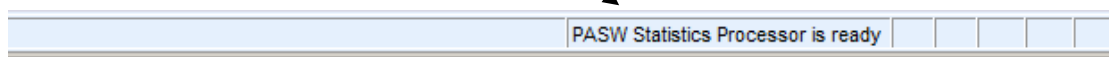
4) **Barra de Ferramentas**, facilitando o acesso às tarefas mais comumente utilizadas;

Barra de ferramentas da janela da edição de dados



5) **Barra de Estado**, informando em que estado se encontram os comandos, as variáveis filtradas e os ficheiros particionados.

Indicador de estado

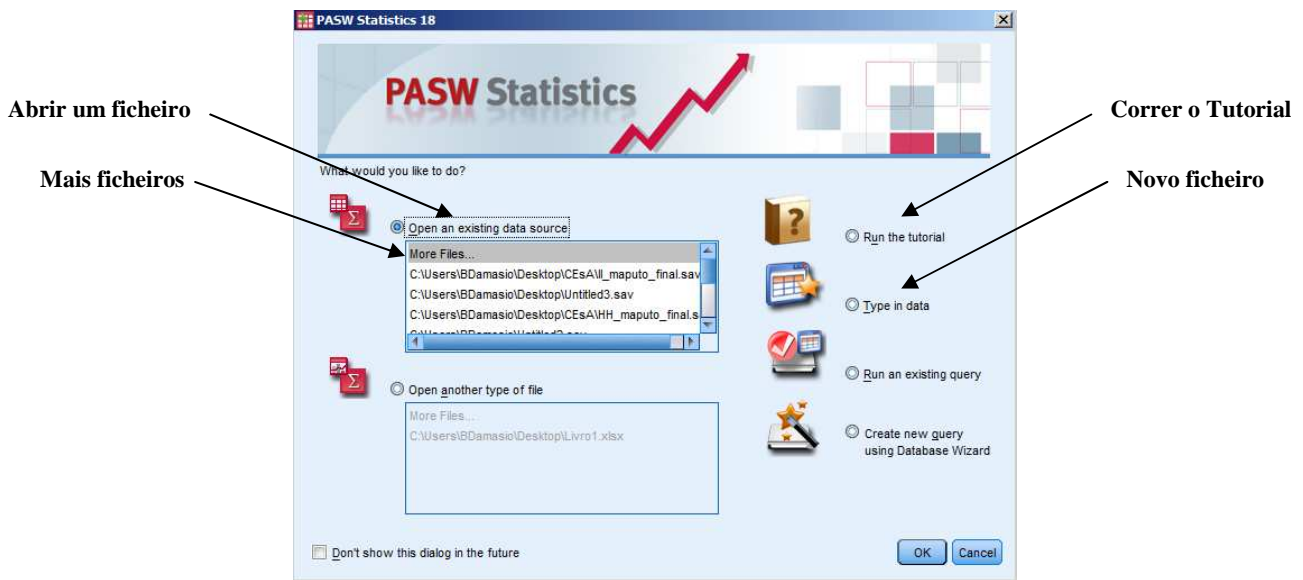


## 2- Comandos e Operações Básicos

### 2.1- Abrir um Ficheiro de Dados

Depois de iniciarmos o nosso programa (clicando, por exemplo, no menu *Start* e acedendo depois ao *PASW Statistics 18*) seremos confrontados automaticamente com um menu onde podemos abrir o ficheiro sobre o qual desejamos trabalhar. Basta para tal procurar o ficheiro na janela sob o título *Open an existing file*. No caso do ficheiro procurado não se encontrar na lista que nos é apresentada seleccionamos a opção *More Files* (Mais Ficheiros).

Naturalmente, como nos restantes programas que utilizam a plataforma Windows também podemos recorrer na barra de menus ao comando *File / Open / Data*.



## 2.2 - Iniciar um Novo Ficheiro

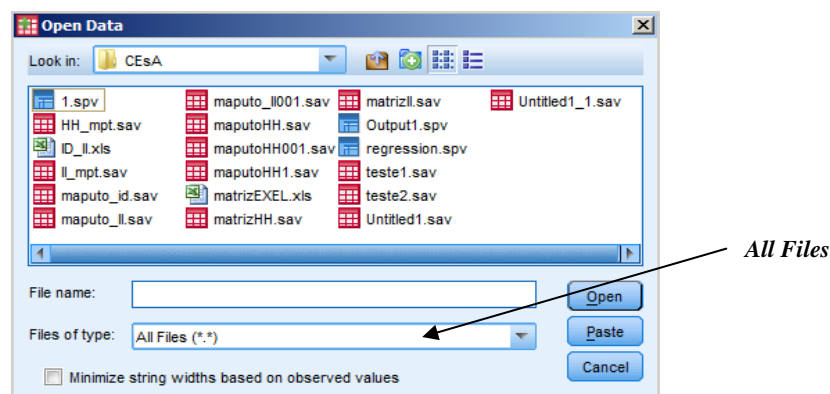
Caso desejemos **iniciar um novo ficheiro**, podemos seleccionar a opção *Type in data*, ou, numa situação em que já nos encontramos no interior do programa, recorrer na barra de menus ao comando *File / New*.

## 2.3- Salvar um Ficheiro de Dados

A **gravação dos dados** é feita recorrendo aos comandos habituais: *Save* no caso de uma gravação normal, *Save as* no caso de pretendermos realizar a gravação com um nome diferente do original.

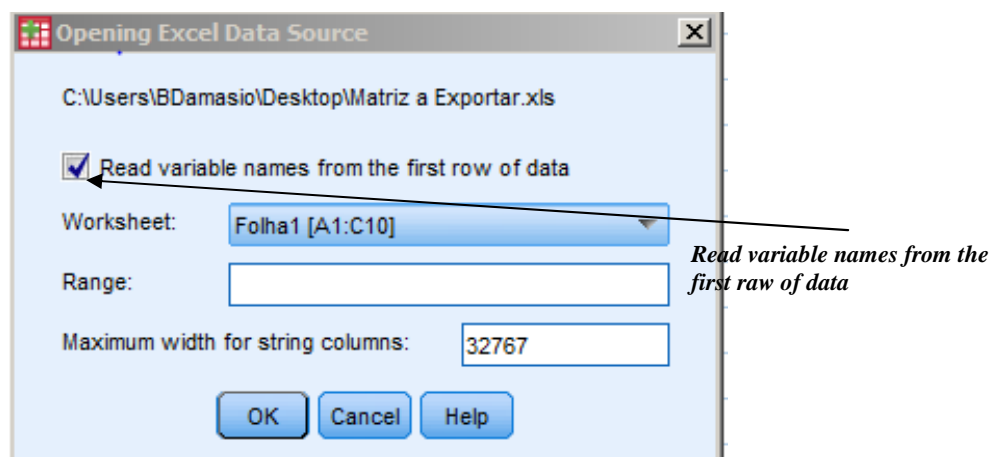
## 2.4- Abrir um Ficheiro de Dados Criado em Excel

O *PASW* consagra igualmente a possibilidade da **abertura de dados criados noutras aplicações**, sendo o *Excel* uma delas<sup>1</sup>. Para o efeito, devemos na janela de dados seleccionar a opção *File / Open / Data*. Selecciona-se *Files of type*, e depois deve ser assinalada a opção *All files (\*.\*)*.



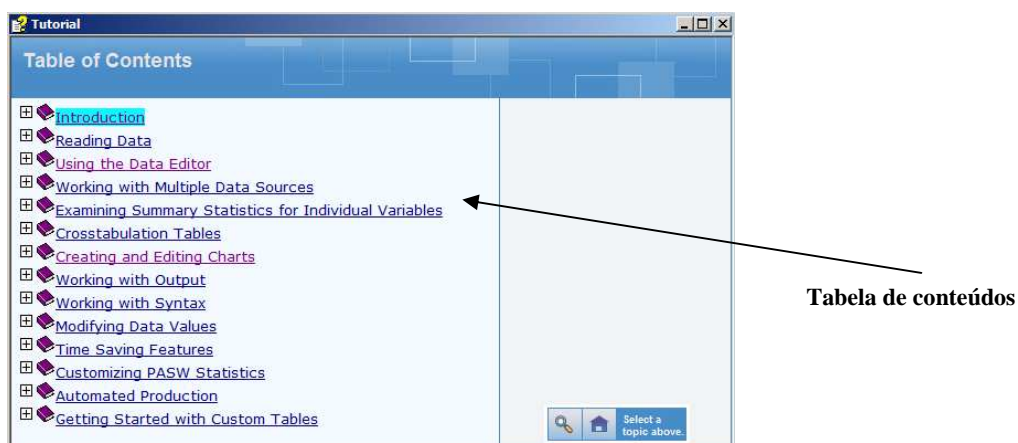
<sup>1</sup> Para mais informações acerca da abertura de dados criados em Access, ver no tutorial do *PASW*, secção *Reading Data*, subsecção *Reading Data from a Database*

Querendo que os dados inseridos na linha 1 da folha de *Excel* a exportar passem a ser as variáveis no nosso ficheiro do *PASW*, basta seleccionar a opção *Read variable names from the first row of data* (normalmente, activa por defeito).



## 2.5- Obtenção de Ajuda

Por sua vez, ainda no contexto da primeira janela, após iniciarmos o programa, reclamando por socorro, temos a hipótese de, escolhendo a opção *Run Tutorial*, usufruir da **ajuda** que o *PASW* nos providencia. Deste modo, deparamo-nos com o índice detalhado da **tabela de conteúdos** (*Table of Contents*), do **tutorial** do programa ao qual podemos permanentemente recorrer em caso de dúvida ou incerteza.



O *PASW* promove **ajuda** de outras formas, nomeadamente, através do menu *Help* que se distingue na barra de menus, sendo o mesmo válido para todas as restantes janelas, bem como para as caixas de diálogo do *PASW*, uma vez que a hipótese de pedirmos ajuda está sempre presente, materializada no símbolo *Help*.



## 2.6- Definição de Variáveis

Uma etapa fundamental, do trabalho com o *PASW*, é a da **definição de variáveis**. O programa permite-nos definir, de uma forma geral, as seguintes características de uma variável:

- 1) o seu **nome**;
- 2) o seu **tipo** (alfabética, numérica, moeda, data, etc);
- 3) a sua **descrição** e dos seus **valores**;
- 4) os códigos especiais para os **valores em falta**.

Na folha *Variable View* cada linha representa uma **variável**, identificada na **coluna Name**. Essa identificação não pode ter mais de 8 letras ou números. À partida, o *PASW* assume por defeito as variáveis como **variáveis numéricas**. No caso de estarmos perante uma **variável alfabética** deveremos na coluna *Type* escolher a opção *String*.

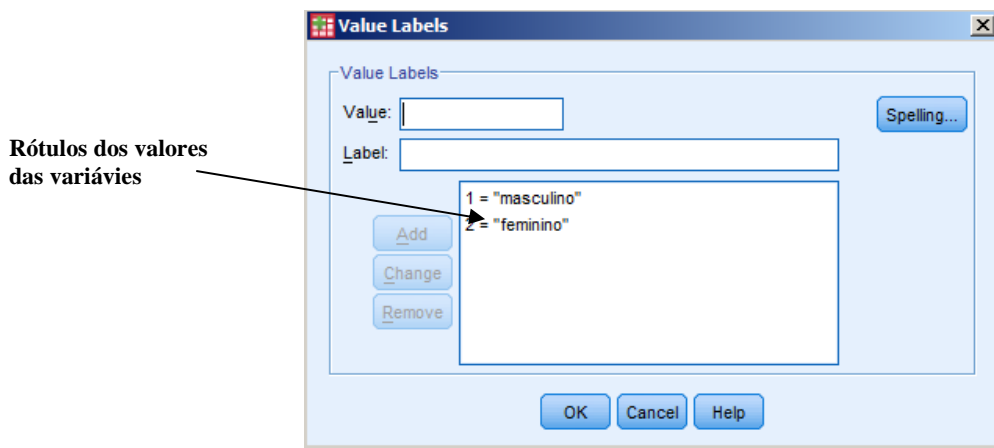
|    | Name       | Type    | Width | Decimals | Label              | Values           | Missing | Columns | Align | Measure |
|----|------------|---------|-------|----------|--------------------|------------------|---------|---------|-------|---------|
| 65 | CnPCFTalk  | Numeric | 8     | 0        | Se a casa em ...   | {1, destruiu ... | 99      | 7       | Right | Nominal |
| 66 | TipoAnter  | Numeric | 8     | 0        | De que tipo era... | {1, precária}... | 99      | 8       | Right | Nominal |
| 67 | ActCond    | Numeric | 8     | 0        | Quais as condi...  | {1, alugada}...  | 99      | 8       | Right | Nominal |
| 68 | Ren_QT     | Custom  | 8     | 0        | Se casa é de r...  | None             | 99      | 8       | Right | Scale   |
| 69 | Reind_Quem | Numeric | 8     | 0        | Se casa é de r...  | {1, Administr... | 99      | 8       | Right | Nominal |
| 70 | LicConstr  | Numeric | 8     | 0        | Se construiu te... | {1, sim}...      | 99      | 8       | Right | Nominal |
| 71 | InspConstr | Numeric | 8     | 0        | Onde se inspiro... | {1, modelo t...  | 99      | 22      | Right | Nominal |
| 72 | DesCasa    | Numeric | 8     | 0        | Fez desenho d...   | {1, sim}...      | 99      | 8       | Right | Nominal |
| 73 | QuemDes    | Numeric | 8     | 0        | Quem desenhou?     | {1, chefe de...  | 99      | 8       | Right | Nominal |
| 74 | DataCstr   | Numeric | 9     | 0        | Qual a data de ... | None             | 0       | 8       | Right | Scale   |


A coluna *Width* diz respeito ao espaço reservado à visualização dos valores da variável, enquanto que a coluna *Decimals* controla o número de casas decimais visíveis.

Como o nome identificativo da variável pode ser insuficiente para uma boa percepção da variável é possível na coluna *Label* adicionar um **rótulo** à variável, isto é, um pequeno texto com uma informação concisa sobre a variável. Imaginemos uma variável respeitante ao estado civil, que dificilmente se pode explicitar com apenas oito caracteres. O *Label* serviria neste caso para contornar esta questão.

Outra importante função é a de **rotular os valores das variáveis**, função à qual se acede através da coluna *Values*.

Quando nos socorremos de códigos numéricos para fazer a identificação das variáveis podemos através desta função dar os nomes que pretendemos aos diferentes códigos das variáveis.




Imaginemos uma variável com o nome *Sexo*, que tem **duas categorias** *Masculino* e *Feminino*, com os códigos **1** e **2**. Assim, na coluna *Values*, seleccionamos a linha sobre a qual desejamos proceder à definição dos códigos. Seguidamente fazemos clique no botão . Surge-nos então a caixa de diálogo *Value Labels*. Na caixa de texto *Value* digitamos o código (p.ex. 1). Depois, no caixa de texto *Label* escrevemos o nome respeitante a esse código (no nosso caso *Masculino*), e premimos o botão *Add* para adicionar o rótulo à listagem de rótulos. De seguida repete-se o mesmo processo para os restantes códigos. Se pretendemos **alterar o rótulo do código** temos de premir *Change* e se pretendemos removê-lo temos de premir *Remove*.

O facto de utilizarmos a função de rotulagem diminui em grande parte a possibilidade de cometermos


|    | II_Sexo   |
|----|-----------|
| 13 | masculino |
| 14 | feminino  |
| 15 | feminino  |
| 16 | feminino  |
| 17 | feminino  |
| 18 | 22        |
| 19 | feminino  |
| 20 | masculino |
| 21 | feminino  |

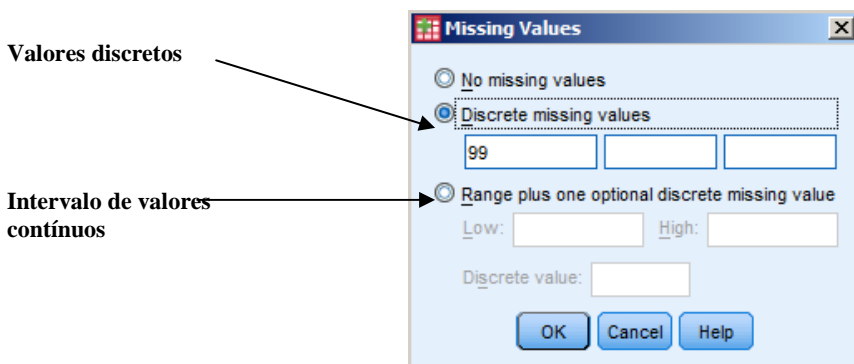
erros, já que um simples erro na introdução de um número, que poderia passar despercebido, torna-se muito mais perceptível com a existência dos rótulos. Imaginemos a situação atrás descrita [1 – Masculino, 2- Feminino] e que por lapso quem introduz o dado digita 22. O facto de este número não se transformar num rótulo desperta automaticamente a atenção de quem está a introduzir os dados diminuindo consideravelmente a possibilidade de este género de erros ocorrer.

Para que esta função esteja activa, temos de verificar na barra de menus, *View*, se a função *View labels* está seleccionada. De salientar que a barra de

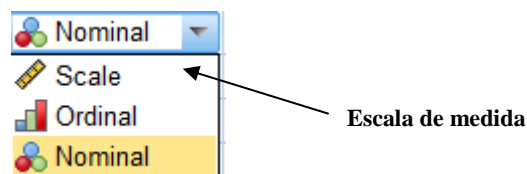
ferramentas no ícon  realiza a dita operação. Do mesmo modo, estes rótulos irão surgir no *Output* de análises estatísticas que incluam essa variável.

Campos de dados em branco, ou valores não válidos, em variáveis numéricas são automaticamente convertidos em *missing values* (valores em falta) pelo *PASW*, surgindo com um (.) na janela do editor de dados. Desta forma, estes valores são **excluídos dos cálculos estatísticos**. Por vezes nos inquéritos existe a possibilidade de surgirem respostas como *Não sabe / Não responde*. Se introduzimos códigos

para estas respostas podemos desejar que elas não façam parte das tabelas e outras análises estatísticas. Para que o PASW trate essas variáveis como *Missing* temos de os declarar *Missing Values*. As não respostas são definidas na caixa de diálogo *Missing Values* que pode ser acedida clicando no botão  das células da coluna *Missing* na folha *Variable View*. Na caixa de texto *Discrete missing values* introduzimos os **valores discretos** que pretendemos tornar *missing values*, enquanto que na caixa *Range plus one optional discrete missing value* podemos definir um **intervalo de valores contínuos** e um valor discreto como *missing values*. Podemos declarar até um máximo de três códigos como *missing value*, sendo que dois deles podem definir um intervalo contínuo de *missing values*. Mais adiante no manual faremos uma descrição mais detalhada sobre a vantagem estatística da utilização de *missing values*.



Uma função extremamente útil na folha *Variable View* é a possibilidade de **copiarmos as definições** que fazemos numa variável para outras variáveis. Imaginemos que tínhamos uma variável respeitante à região onde nasceu. Fazemos a codificação das diversas regiões com o processo já descrito de rotulagem dos valores. Seguidamente temos uma variável respeitante à região onde habita. Para não termos de repetir todo o processo de codificação seleccionamos a célula com essa definição, premimos as teclas *CTRL+C* ou escolhemos na barra de menus *Edit / Copy*. Depois seleccionamos com o rato a célula respeitante à variável para a qual pretendemos efectuar a cópia e premimos as teclas *CTRL+V* ou escolhemos na barra de menus: *Edit / Paste*. Por outro lado, a coluna *Measure* remete para a **Escala de Medida da Variável**. Desta forma, podemos escolher as seguintes opções:



1) *Nominal*, para **variáveis categóricas qualitativas**, cujos valores atribuídos representam categorias não hierarquizáveis, por exemplo, os idiomas falados numa determinada província, ou o tipo de profissões exercidas;

2) *Ordinal*, no caso de estarmos na presença de **variáveis categóricas qualitativas**, cujos valores são hierarquizáveis, objectivando criar uma relação de preferência, na prática, dada uma escala, A é preferido a B que, por seu turno, é preferido a C, exemplificando, grau de satisfação com o nível de vida num determinado espaço (muito, pouco ou nada);

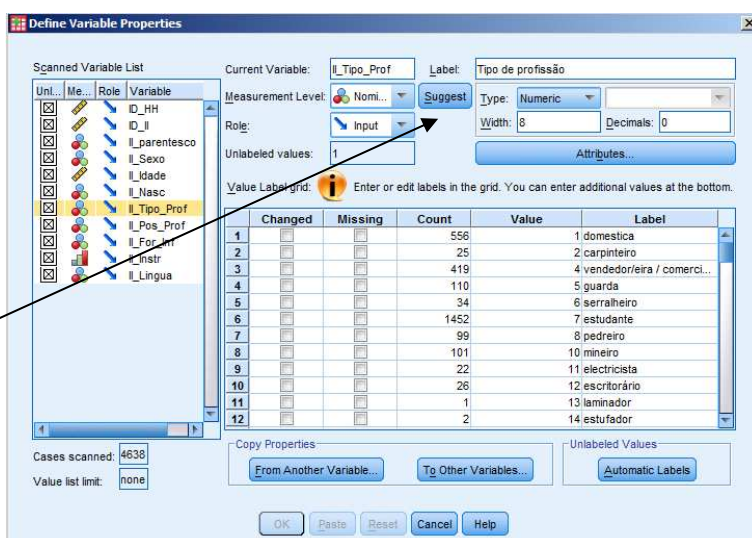
3) *Scale*, remete para **variáveis quantitativas** cujos atributos, contínuos ou em escadas de intervalos, resultam de características de informação susceptíveis de serem medidas e quantificadas, com diferentes intensidades (por exemplo idade em anos ou rendimento em milhares de euros).

Torna-se inevitável apontar, de momento, que o *PASW* consagra a possibilidade de, de forma prática, conseguir, por um lado **analisar os rótulos das variáveis** e demais propriedades que lhes atribuímos por outro **aclarar o processo de definição** das mesmas.

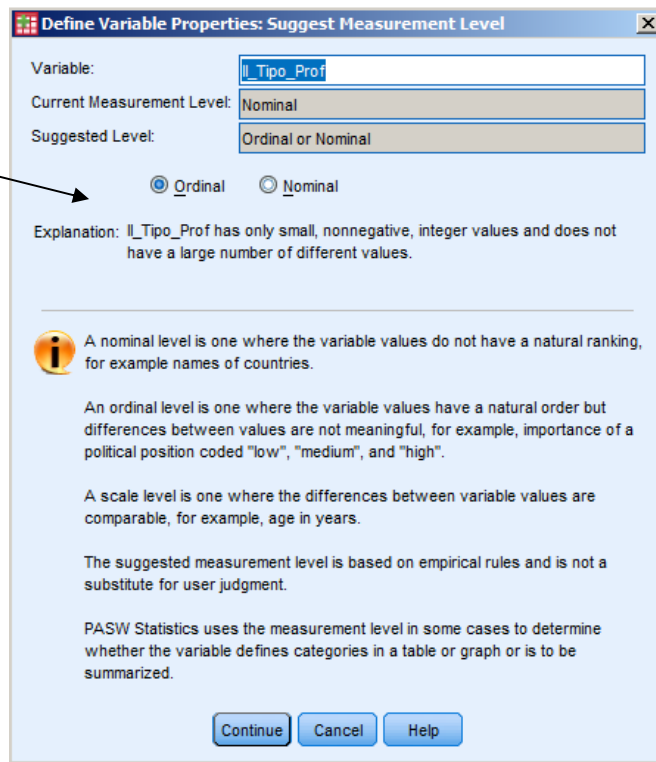
Fazendo *Data / Define Variable Properties* surge-nos a janela de selecção das variáveis a indicar, consoante com a imagem, assim, após clicarmos em *Continue*, acedemos a uma panorâmica detalhada das ditas, suas respectivas propriedades, qualidades e rótulos.

Desta feita, a janela referida, constitui um precioso auxiliar no que concerne ao processo de definição das variáveis, dado que, não só permite **analisar os valores efectivos dos dados**, identificando valores sem rótulo, como **fornece sugestões de rótulos e de medidas para as variáveis** e, igualmente, concede a possibilidade de listarmos os valores e propriedades de dados respectivos a cada variável seleccionada. Logo, a opção *Data / Define Variable Properties*, fornece-nos a capacidade de copiar os *Value labels* (e outros atributos) já definidos, de uma variável para a variável seleccionada ou, pelo contrário, da variável seleccionada para múltiplas variáveis adicionais. Segundo o exemplo que se segue:

Sugestão de escala de medida da variável

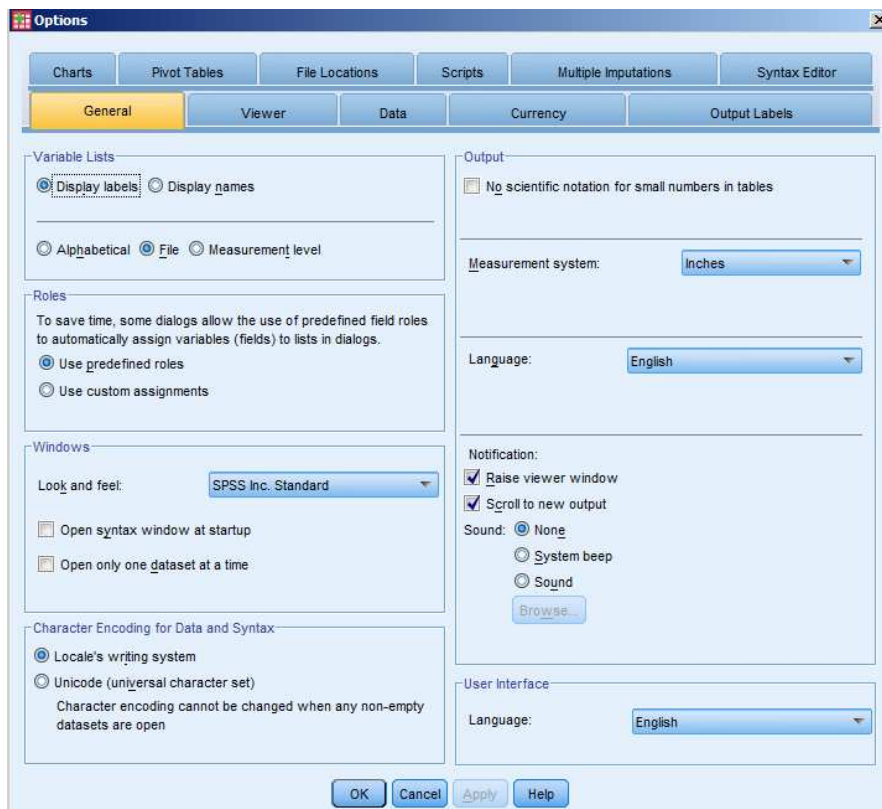


Explicação da sugestão



## 2.7- Alteração das Opções de Ambiente

Se ambicionarmos proceder a **alterações** ao nível **do aspecto do ambiente de trabalho** basta escolher na barra de menus a opção *Edit / Options*. Podem, do mesmo modo, ser modificadas as opções que remetem para o formato das variáveis, bem como as relativas à visualização dos diversos tipos de janelas.





### 3- Introdução de Dados e Cálculos Simples

#### 3.1- Introdução de Dados

Quando estamos perante um ficheiro de dados, na janela do *Data editor*, anteriormente apresentada no presente texto, podemos alternar entre duas janelas. Uma, a **janela de edição de dados** (*Data View*), permite-nos proceder à introdução e alteração dos nossos dados. Outra, a da **definição das variáveis** (*Variable View*), permite-nos aceder a diferentes informações sobre as diversas variáveis. Para alterar a janela sobre a qual desejamos trabalhar temos dois botões no canto inferior esquerdo.

The screenshot shows the PASW Statistics Data Editor interface. The main window displays a data table with the following columns: ID\_HH, ID\_II, II\_parentesco, II\_Sexo, II\_Idade, II\_Nasc, II\_Tipo\_Prof, II\_Pos\_Prof, II\_For\_Inf, II\_Instr, II\_Lingua, and II\_Rel. The table contains 25 rows of data. Annotations include:

- Caso**: Points to the first row (ID\_HH 1).
- Variável**: Points to the 'II\_Sexo' column header.
- Célula**: Points to the cell containing 'feminino' in the 'II\_Sexo' column for the first row.
- Data View**: Points to the 'Data View' button at the bottom left.
- Variable View**: Points to the 'Variable View' button at the bottom left.

| ID_HH | ID_II | II_parentesco  | II_Sexo        | II_Idade | II_Nasc  | II_Tipo_Prof          | II_Pos_Prof   | II_For_Inf | II_Instr      | II_Lingua | II_Rel                     |
|-------|-------|----------------|----------------|----------|----------|-----------------------|---------------|------------|---------------|-----------|----------------------------|
| 1     | 1     | chefe de fa... | feminino       | 72       | maputo   | domestica             | doméstica     | doméstica  | não estudou   | makhua    | maometana                  |
| 2     | 1     | sobrinho/a     | masculino      | 46       | maputo   | serralheiro           | 99            | 99         | 4ª classe     | makhua    | maometana                  |
| 3     | 1     | nora/genro     | feminino       | 33       | gaza     | domestica             | doméstica     | doméstica  | 9º ano        | changana  | católica                   |
| 4     | 2     | chefe de fa... | masculino      | 60       | zambézia | carpinteiro           | conta propria | informal   | não estudou   | n kami    | católica                   |
| 5     | 2     | cônjuge        | feminino       | 39       | maputo   | domestica             | doméstica     | doméstica  | não estudou   | ronga     | ziona                      |
| 6     | 2     | filho/a        | masculino      | 15       | maputo   | estudante             | criança       | criança    | até 4ª classe | ronga     | ziona                      |
| 7     | 2     | filho/a        | feminino       | 3        | maputo   | criança               | criança       | criança    | criança       | ronga     | criança                    |
| 8     | 3     | chefe de fa... | feminino       | 48       | maputo   | vendedor/eira / co... | conta propria | informal   | até 4ª classe | ronga     | igreja belém de moçambi... |
| 9     | 3     | filho/a        | feminino       | 24       | maputo   | vendedor/eira / co... | conta propria | informal   | 4ª classe     | ronga     | 99                         |
| 10    | 3     | 303            | filho/a        | 19       | maputo   | domestica             | doméstica     | doméstica  | 99            | ronga     | maometana                  |
| 11    | 3     | 304            | filho/a        | 22       | maputo   | vendedor/eira / co... | conta propria | informal   | até 4ª classe | ronga     | maometana                  |
| 12    | 3     | 305            | neto/a         | 6        | maputo   | criança               | criança       | criança    | criança       | ronga     | criança                    |
| 13    | 4     | 401            | chefe de fa... | 99       | maputo   | guarda                | conta alheia  | informal   | 99            | ronga     | 99                         |
| 14    | 4     | 402            | cônjuge        | 56       | maputo   | vendedor/eira / co... | 99            | informal   | 99            | changana  | 12 apóstolos               |
| 15    | 5     | 501            | chefe de fa... | 43       | gaza     | vendedor/eira / co... | 99            | informal   | até 9º ano    | changana  | velho apóstolos            |
| 16    | 5     | 502            | filho/a        | 18       | maputo   | estudante             | criança       | criança    | até 9º ano    | portugues | IURD                       |
| 17    | 5     | 503            | filho/a        | 16       | maputo   | estudante             | criança       | criança    | até 9º ano    | portugues | nazare                     |
| 18    | 5     | 504            | filho/a        | 6        | maputo   | criança               | criança       | criança    | criança       | portugues | criança                    |
| 19    | 6     | 601            | chefe de fa... | 36       | maputo   | vendedor/eira / co... | 99            | informal   | até 9º ano    | portugues | maometana                  |
| 20    | 6     | 602            | irmão          | 29       | maputo   | pedreiro              | conta propria | informal   | 4ª classe     | portugues | 99                         |
| 21    | 6     | 603            | irmão          | 30       | maputo   | vendedor/eira / co... | conta propria | informal   | 4ª classe     | portugues | maometana                  |
| 22    | 6     | 604            | irmão          | 27       | maputo   | 99                    | 99            | 99         | até 4ª classe | portugues | 99                         |
| 23    | 6     | 605            | irmão          | 22       | maputo   | pedreiro              | conta propria | informal   | até 9º ano    | portugues | maometana                  |
| 24    | 6     | 606            | filho/a        | 18       | maputo   | 99                    | 99            | 99         | até 9º ano    | portugues | maometana                  |
| 25    | 6     | 607            | filho/a        | 13       | maputo   | estudante             | criança       | criança    | 4ª classe     | portugues | maometana                  |

Abordaremos inicialmente a janela de edição de dados. Esta surge-nos no ecrã sobre a forma de uma matriz tipo folha quadriculada, em que cada linha é um **caso** (no programa case) e em que cada coluna é uma **variável**.

Os casos são os diferentes **elementos sobre os quais vamos preencher a informação**, podendo, por exemplo, num inquérito às famílias, ser os vários elementos do agregado familiar.

As variáveis (colunas) são os diferentes **itens de informação que foram recolhidos para os diversos casos**. Podem ser por exemplo a data de nascimento, a profissão, etc.

A intersecção da linha com a coluna é chamada de **célula**. É sobre cada célula que nós **introduzimos os nossos dados**. Para tal basta escrever o dado no *Editor de Células*, e carregar em *Enter* ou passar para outra célula. Se em vez de *Enter* se carregar em *Esc* o valor original da célula continua inalterado.

É possível **apagar ou copiar valores** individuais de células ou grupos de valores. Para apagar, basta seleccionar com o rato a célula ou células a apagar.

Com a área seleccionada, premir a tecla *Delete*, ou as teclas *CTRL+X* (ou, premindo o botão direito do rato, seleccionar a opção *Cut*) no caso de desejarmos **transferir o conteúdo para outra ou outras células**.

Para copiar a fase da selecção é igual, só que após a área estar seleccionada têm de se premir as teclas *CTRL+C* (ou, premindo o botão direito do rato, seleccionar a opção *Copy*) ou escolher na barra de menus *Edit / Copy*. A seguir selecciona-se com o rato a célula, ou canto superior esquerdo do conjunto de células, para onde se pretende efectuar a cópia. Escolhida a área é altura de premir as teclas *CTRL+V* (ou, premindo o botão direito do rato, seleccionar a opção *Paste*) ou escolher na barra de menus *Edit / Copy*.

Para **inserir um novo caso** (linha) existem três hipóteses. A primeira é a de seleccionamos com o rato uma célula abaixo da qual se pretende inserir o novo caso, bem como na barra de menus a opção *Edit / Insert case*; a segunda é a de, com o botão direito do rato, escolhermos, igualmente, uma célula abaixo, a opção *Insert Case*; a terceira é a de optarmos por, na barra de ferramentas, realçada a linha abaixo da qual queremos inserir o novo caso, premir o ícone



Ícone da barra de ferramentas

Para **inserir uma nova variável** (coluna) o procedimento é em tudo idêntico ao da inserção de casos, porém, ao invés de selectarmos a opção *Insert case*, devemos escolher *Insert variable*, ou, em alternativa, o ícone



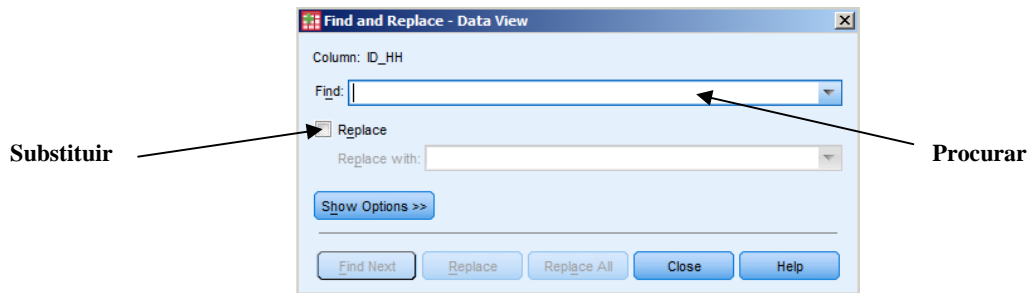
Ícone da barra de ferramentas

### 3.2- Procura de Dados

Por vezes temos necessidade de **encontrar dados** específicos, o que de uma forma manual implicaria uma perda de tempo desnecessária. O *PASW* tem uma função que nos permite realizar essa tarefa de uma forma mais rápida e eficiente.

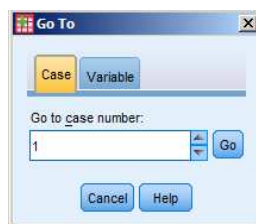
No caso de desejarmos **procurar um dado** específico (numérico ou texto) devemos, antes de tudo, seleccionar com o rato uma célula da coluna onde pretendemos efectuar a procura. Já com a coluna seleccionada, na barra de menus escolhe-se *Edit / Find* (ou *CTRL+F*). Na caixa *Find and Replace – Data*

View, o subtítulo será *Column: "Nome da Variável"*, neste caso introduz-se a informação que desejamos procurar e depois preme-se *Find Next*.




A janela referida contém outras duas opções fundamentais. A saber, a opção *Replace* (**Substituir**) e, consequentemente, a opção *Replace with* (**Substituir por**) que é bastante útil no caso de desejarmos substituir um conjunto de valores por outros. A outra opção é a de, premindo *Show Options* (Mostrar Opções), se desejarmos levar em conta a diferença entre maiúsculas e minúsculas, considerando que o programa ignora automaticamente tal diferença, deverá ser seleccionada a opção *Match case*.

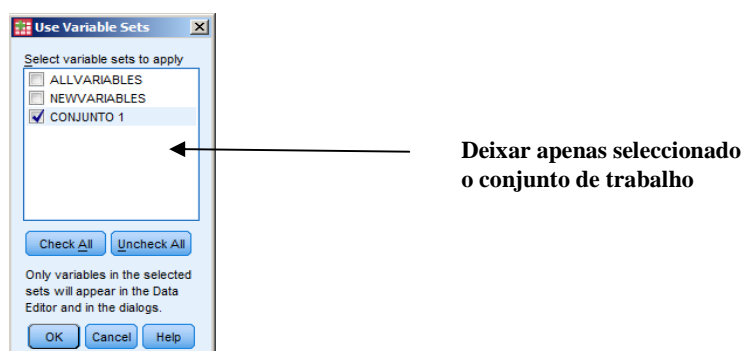
No caso de desejarmos procurar antes um caso, apenas temos de na barra de menus escolher *Edit / Go to Case*, e introduzir na caixa *Case Number* o número da linha pretendida.



### 3.3- Filtrar, Esconder e Exibir Conjuntos de Variáveis

Existe a possibilidade de se **seleccionar conjuntos de variáveis** com os quais pretendemos trabalhar. Para o efeito é suficiente ir à barra de menus opção *Utilities / Define variable sets* e escolher, dentro das variáveis disponíveis na *caixa de diálogo* que aparece, as variáveis específicas que desejamos utilizar. Para finalizar o processo é ainda necessário notar o novo conjunto de variáveis (*Set*) escolhido de uma designação própria, para tal basta escrever na caixa de *Set name* o nome do novo conjunto e depois seleccionar a opção *Add Set*.

Uma vez definido o agrupamento das únicas variáveis que queremos que apareçam na nossa janela *Data Editor*, é necessário, para que possamos visualizá-lo, **aplicar o filtro** referido fazendo *Utilities / Use variable sets* (ou em alternativa, na barra de ferramentas, premir o ícone ) e **eleger apenas o nome do nosso conjunto de variáveis recém criado**, no exemplo abaixo reproduzido, *Conjunto 1*.



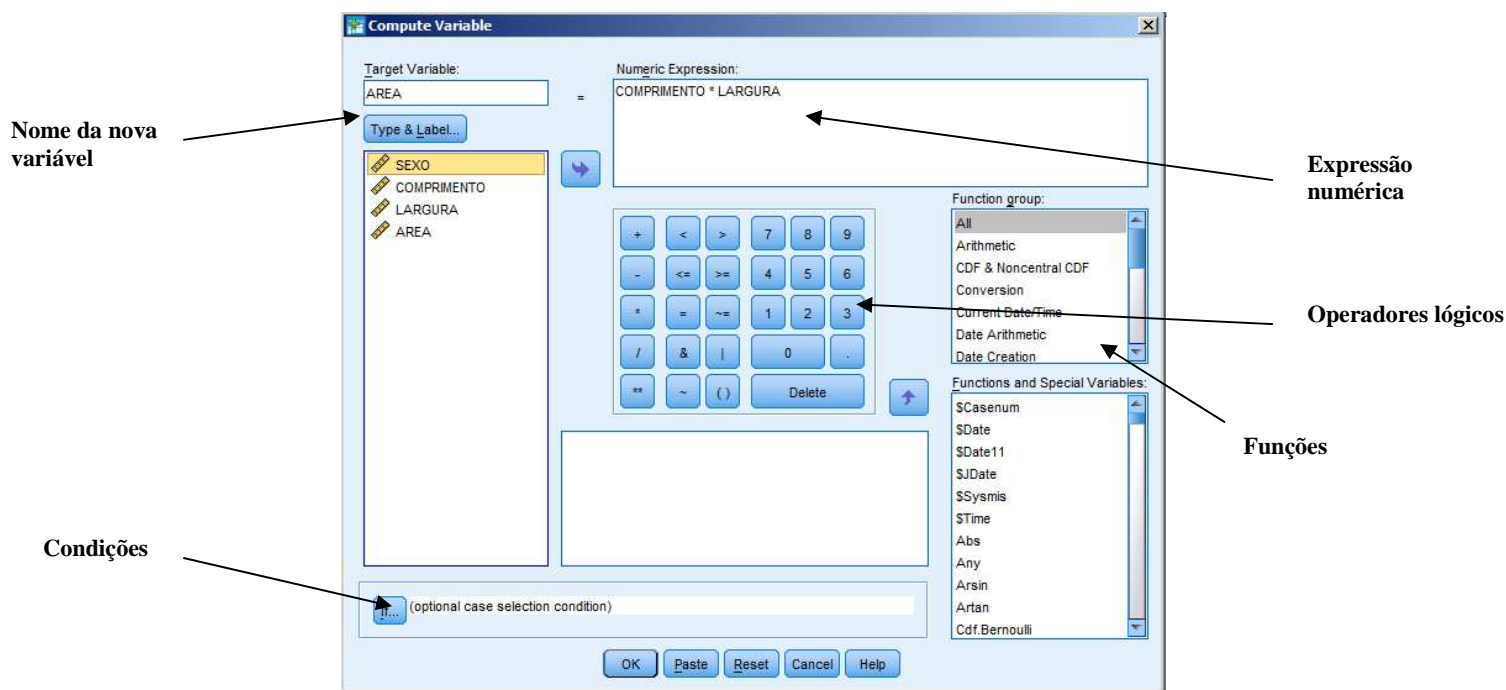



## 4- Transformação de Dados

O PASW permite a **criação / alteração e recodificação de uma variável** a partir de operações de transformações sobre outras variáveis.

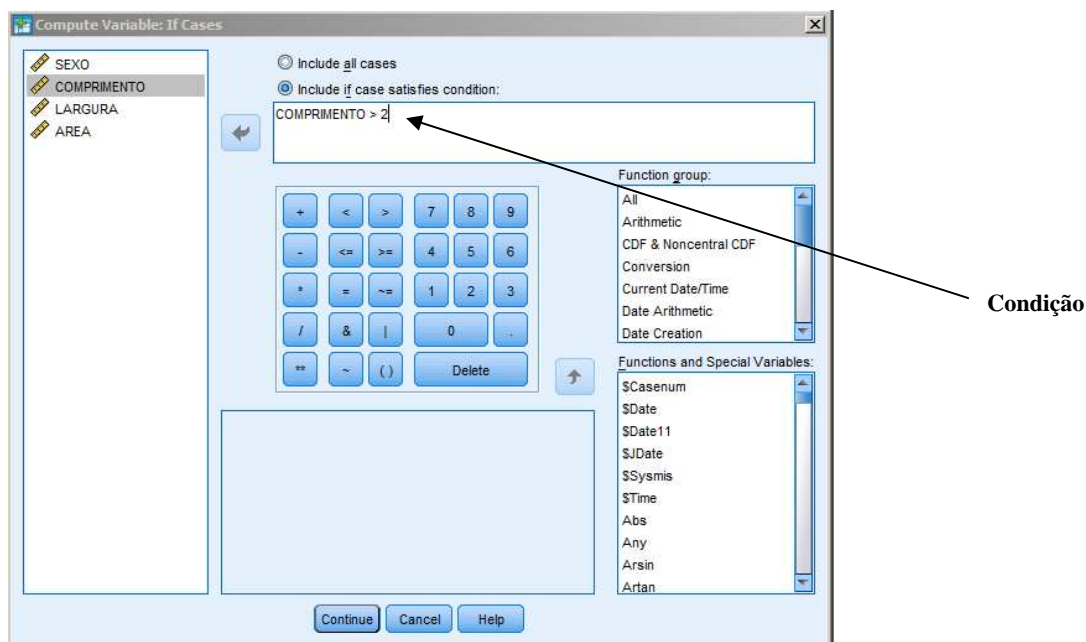
### 4.1- Cálculo de uma Nova Variável com Base em Variáveis Existentes

Uma das possibilidades é a **criação de uma nova variável a partir da realização de cálculos sobre outras variáveis**. Por exemplo encontrar uma determinada área a partir do comprimento (*compr*) e da largura (*largura*) de um talhão. Na barra de menus escolhemos *Transform / Compute Variable*. Na caixa *Target Variable* introduzimos o nome da variável de destino, que pode já existir, ou se assim o desejarmos, ser uma variável nova. Neste caso o seu nome poderia ser *Area*, sendo também possível definir o seu *Label (Rótulo)* e o tipo de variável acedendo ao botão *Type & Label*.

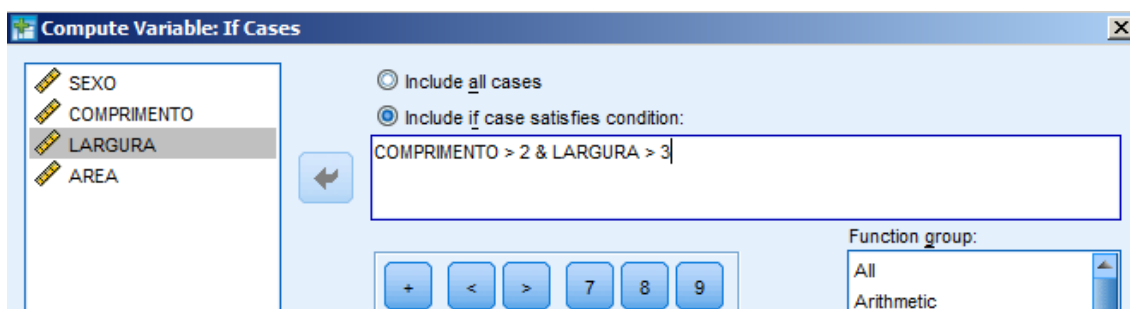


Depois, na caixa *Numeric Expression* construímos a expressão para gerar o valor da variável de destino. Podemos recorrer a uma lista de funções presente em baixo. No caso da área seria *compr\*largura*. Para tal escolheríamos, na caixa da esquerda, a variável *compr*, pressionando depois a caixa com a seta  para passar o valor para a caixa da direita. Seguidamente na lista de operadores à direita dessa janela premiamos sobre o operador *\**, voltando depois a efectuar o mesmo procedimento com a variável *largura*. O PASW permite também efectuar este tipo de **operações** (assim como outras que veremos mais adiante) para **subconjuntos de casos**. Imaginemos que no mesmo exemplo da área apenas nos interessava efectuar este cálculo para talhões de largura superior a 2 metros. Teríamos então de premir na caixa *Compute Variable* o botão *If*, depois teríamos de escolher a opção *Include if case satisfies condition*, e na

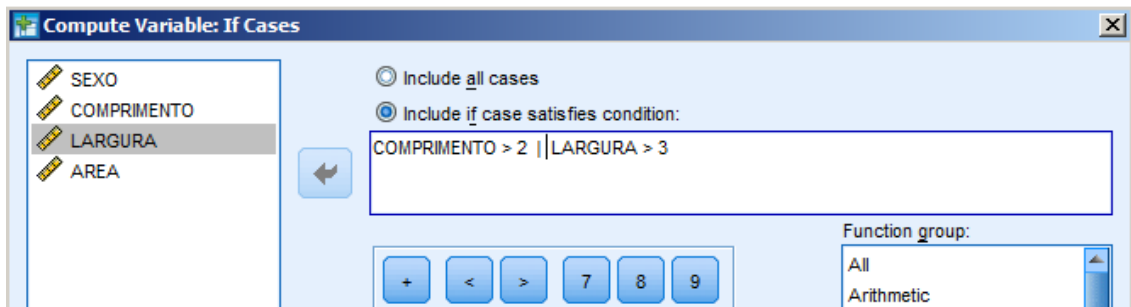
janela de baixo introduzir a condição. Para tal, efectuaríamos o mesmo procedimento que utilizámos anteriormente para passar a variável largura para a direita, e na caixa de operadores escolheríamos o operador  $>$  digitando depois na expressão o valor 2. No fim seria necessário pressionar o botão *Continue*.


















No âmbito da opção *if*, caixa de diálogo *Recode into different variables*, existem mais dois procedimentos para selecção de subconjuntos de casos que consideramos relevantes. Imaginemos que desejávamos efectuar o cálculo referido para casos que verificassem simultaneamente duas condições (exemplo talhões com largura superior a 2 metros e com comprimento superior a 3 metros). Por outro lado, supondo que ambicionávamos efectuar o cálculo referido para casos que verificassem uma de duas condições (exemplo talhões com largura superior a 2 metros ou com comprimento superior a 3 metros). No primeiro caso teríamos de, após seleccionada a variável *largura* para a direita e escrita a expressão *largura > 2*, digitar  $&$  (ou *and*) *compr > 3*.



Por seu turno, no que ao segundo caso respeita, em vez de  $&$ , seria  $|$  (ou *or*), ou seja, a expressão *largura > 2 | compr > 3*, como nos mostra a figura abaixo.

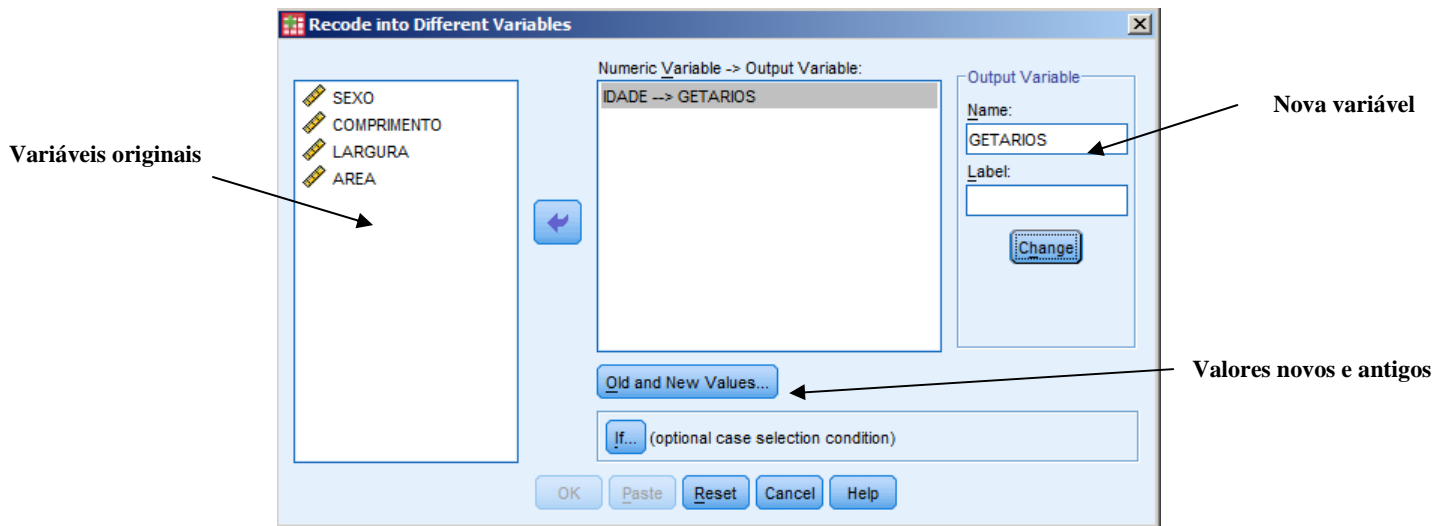


Operadores utilizáveis:

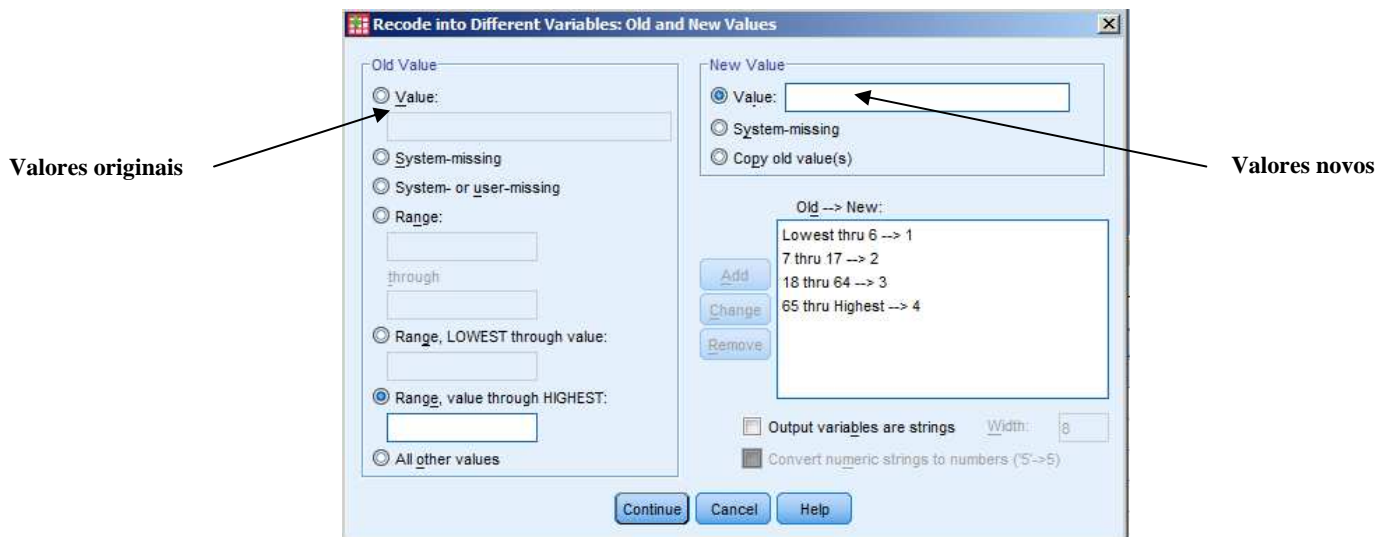
|   |                |   |                |  |                |
|---|----------------|---|----------------|--|----------------|
|  | Somas          |  | Menor          |  | Maior          |
|  | Subtrações     |  | Menor ou igual |  | Maior ou igual |
|  | Multiplicações |  | Igual          |  | Diferente      |
|  | Divisões       |  | E lógico       |  | Ou lógico      |
|  | Exponenciação  |  | Negação lógica |  | Parêntesis     |

#### 4.2- Recodificação - Variáveis Diferentes

Outra função do PASW é a **recodificação de variáveis**. Isto é a **alteração dos códigos das variáveis**. Uma das opções na alteração da codificação das variáveis é a **criação de uma nova variável** para receber essas transformações. Esta opção permite-nos **manter a primeira variável com os valores iniciais**, não se correndo assim o risco de perda de informação. Imaginemos uma situação prática em que temos uma variável (*Idade*), que contém a idade de todos os membros da população estudada, mas que para o nosso estudo se torna mais simples utilizar essa informação agrupada nos seguintes grupos etários: *Crianças [0,6]*; *Jovens [7,18]*; *Adultos [19,65]*; *Idosos [+ de 65]*. Para tal temos de agrupar essa informação, mas com o cuidado de não perdermos a informação inicial, pois poderá ser útil para recolher outro género de informação. Para o efeito, devemos, na barra de menus, escolher *Transform / Recode into different variables*, o que nos levará à caixa de diálogo *Recode into Different Variables*. Na caixa da esquerda escolhemos a variável que pretendemos transformar, passando-a para a caixa *Input Variable -> Output Variable* (no caso do nosso exemplo a variável *Idade*). Na caixa *Output Variable* podemos dar o nome à variável (no espaço *Name* e mesmo uma explicação mais detalhada na janela *Label*).



Premimos então o botão *Old and New Values*, onde iremos **especificar a recodificação**. Como pretendemos recodificar intervalos contínuos, isto é, conjuntos de valores entre os quais a idade pode variar, carregamos em *Range*, escrevemos na janela da esquerda o valor mais baixo que queremos recodificar e na janela da direita o valor mais alto. Por exemplo, no caso do grupo dos *Jovens* escrevemos 7 no lado esquerdo e 18 no lado direito. Introduzimos então o novo código no lado direito – *New Value* – na caixa *Value*, premindo depois o botão *Add*. Não nos podemos esquecer que este valor tem de ser igual ao código que depois vamos atribuir nessa variável. O mesmo faremos para o grupo dos *Adultos*. O mesmo já não podemos fazer para o grupo dos idosos, já que não existe um limite superior. Temos então de carregar em *Range*, mas onde por baixo está uma caixa onde no lado direito está escrito *Through highest*. Escreveremos então o valor nessa janela e depois teremos de fazer o mesmo que fizemos para os outros grupos aquando da recodificação no lado direito. No fim de fazermos isto para todos os grupos premimos o botão *Continue*. No caso das Crianças teremos de fazer o oposto escolhendo a caixa onde está escrito *Lowest through*. Existe também a possibilidade de transformar a variável numa variável alfabética. Para tal teremos de seleccionar a opção *Output variables are strings* podendo nesta situação inserir directamente os *labels* sem recurso a codificações. De qualquer forma esta função não é muito aconselhável pois é sempre mais fácil proceder a eventuais alterações nas variáveis quando as mesmas utilizam um sistema de codificação das variáveis.



Se já tivermos finalizado as alterações, voltando à caixa, premimos o botão *Change* e depois *OK*. Surgirá então, à direita da última variável, uma nova variável com o nome que nós lhe tivermos dado. Devemos então introduzir os códigos novos da variável. Por exemplo 1 para *Crianças*, 2 para *Jovens*, 3 para *Adultos* e 4 para *Idosos*.

#### 4.3- Recodificação - Mesmas Variáveis

Podemos efectuar essa mesma alteração **dentro das mesmas variáveis**, embora esta operação não seja muito aconselhável na prática, já que apresenta **riscos em termos de perda de informação**, pois após a alteração não é possível aceder aos valores anteriormente digitados. Por esta razão deve-se procurar evitar a utilização desta função. De qualquer forma deixamos aqui um exemplo deste género de recodificação, em que utilizaremos a variável família. No caso do recurso a esta função **devemos proceder a uma gravação constante do ficheiro** sempre que uma alteração for bem-sucedida. Na barra de menus escolhemos *Trasform / Recode into same variables*. Na janela da esquerda seleccionamos a variável que pretendemos transformar passando-a para a janela *Variables* (No caso do nosso exemplo a variável *Família*). Premimos então o botão *Old and New Values*, onde iremos especificar a recodificação. Imaginemos que tenhamos as seguintes codificações iniciais: 1 – *Chefe de Família*; 2 – *Cônjuge*; 3 – *Filho*, e que pretendemos ficar apenas com a seguinte codificação: 1 – *Chefe de Família*; 2 – *Outros*. Para tal introduzimos o código da variável que pretendemos alterar na caixa *Value* no lado esquerdo - *Old Value* (por exemplo 3). Introduzimos então o novo código no lado direito – *New Value* – na caixa *Value* premindo depois o botão *Add*. Se pretendermos retirar essa alteração devemos premir o botão *Remove* e caso pretendamos alterar os valores temos de premir o botão *Change*. Depois de premir *Continue* nesta janela, premir *OK* na janela *Recode into Same Variables*. Seguidamente, temos de alterar a codificação da variável *Família*. Para tal temos de refazer o processo de rotulagem das variáveis para esta variável como foi descrito anteriormente para a variável *Sexo*.

### 5- Outro tipo de Comandos e Operações

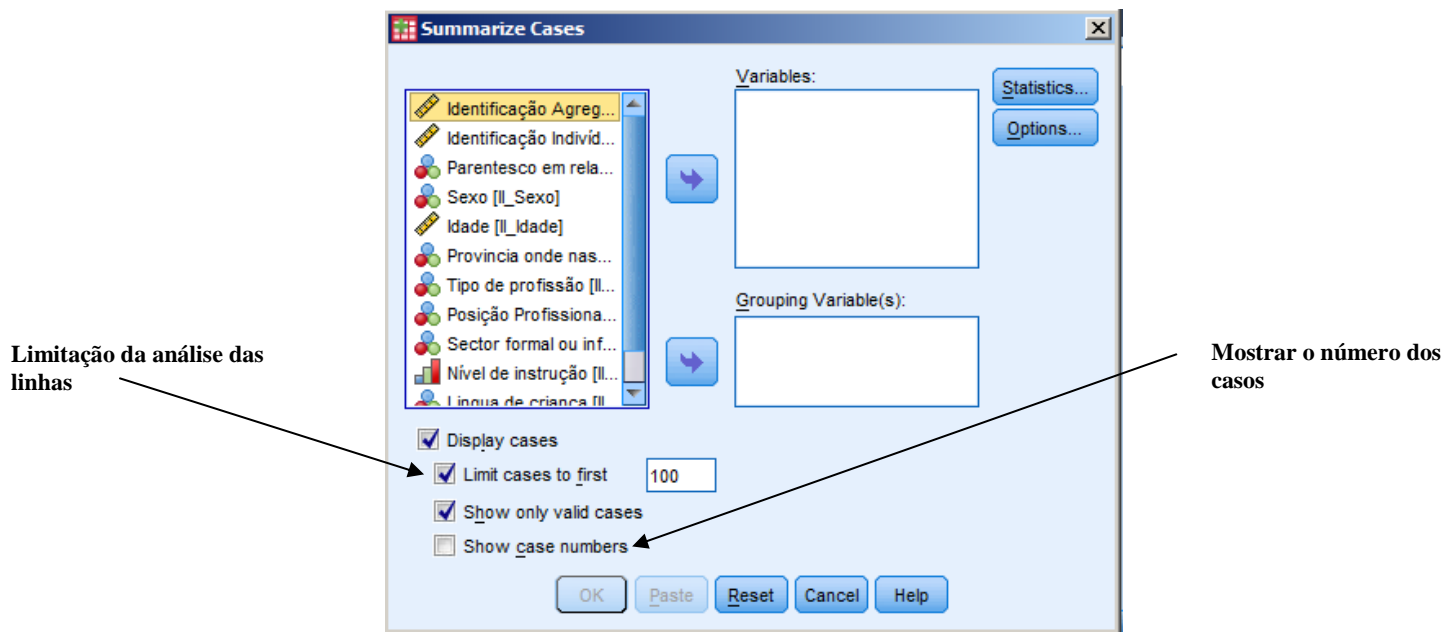
#### 5.1- Detecção de Erros

Antes de entrarmos na fase de análise de dados propriamente dita um dos passos fundamentais para o desenvolvimento correcto de um processo de análise estatística é a **detecção de possíveis erros** na introdução de dados numa matriz. Para tal utilizaremos uma função do *PASW*, que elabora uma espécie de **relatório** que nos possibilita a identificação das linhas em que determinado erro acontece. Uma situação clássica de erro é a da situação em que uma criança com menos de 10 anos de idade aparece como pedreiro na variável profissão. Em primeiro lugar, para descobrir a existência desses erros teremos de proceder à selecção dessas situações, recorrendo na barra de menus à função *Data Select Cases*<sup>2</sup>. Usando essa função estabeleceremos uma condição que leve à selecção de indivíduos que na variável

---

<sup>2</sup> Para perceber o funcionamento desta função ver o capítulo que se encontra mais adiante sobre **SELECCÃO DE CASOS**

profissão tenham a opção pedreiro e na variável idade apresentem valores inferiores a 10. Chegou o momento de **pedir a elaboração do relatório**. Para tal na barra de menus escolhemos a função *Analyze / Reports / Case*. Ficamos então perante a janela *Summarize Cases*. Escolhemos então as variáveis que desejamos analisar, normalmente as que têm a possibilidade de erro. Por defeito, o *PASW* limita a análise aos primeiros 100 casos. Se desejarmos realizar a análise para a totalidade da amostra teremos de desactivar a função *Limit cases to first*. Como o nosso objectivo é a detecção das linhas em que se encontram potenciais erros interessa que o resultado da pesquisa nos indique o número da linha. Para tal teremos de activar a função *Show case numbers*.



Como resultado deste processo iremos obter uma tabela com o seguinte formato:


Case Summaries

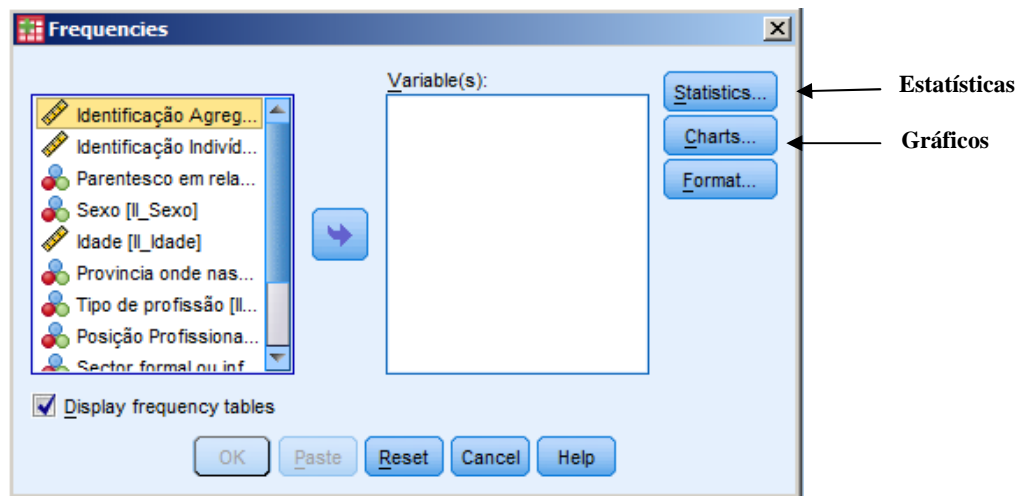
|       | Case Number | Tipo de profissão | Idade |
|-------|-------------|-------------------|-------|
| 1     | 1128        | Pedreiro          | 8     |
| 2     | 1313        | Pedreiro          | 7     |
| 3     | 1408        | Pedreiro          | 6     |
| 4     | 1483        | Pedreiro          | 5     |
| 5     | 1522        | Pedreiro          | 5     |
| Total | N           |                   | 5     |

Utilizaremos então a coluna *Case Number* para sabermos **quais são as linhas em que existem os erros**, podendo depois procurar a origem dos erros e corrigir a informação. No caso de não ser possível detectar a origem dos erros é mais correcto alterarmos a informação para *Não sabe/Não responde*, de forma a que não existam incoerências estatísticas na nossa base de dados

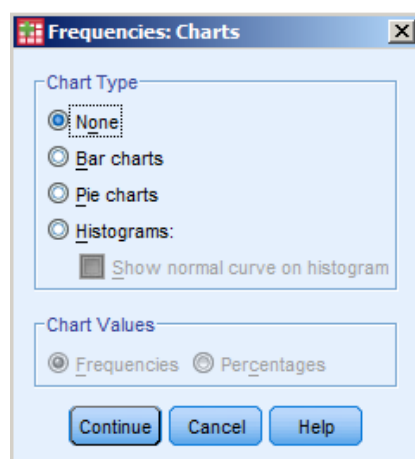
## 5.2- Análise de Dados: Tabela de frequências

A análise de dados mais imediata é a **análise das frequências**. Por frequência entende-se **o número de vezes que determinada possibilidade ocorre numa variável**, ou, de uma forma mais simples, **o número de pessoas (casos) que nos deram essa resposta**. Por exemplo, na variável *Sexo*, as frequências dão-nos o número de respostas para o sexo masculino e para o sexo feminino. Se desejarmos poderão dar-

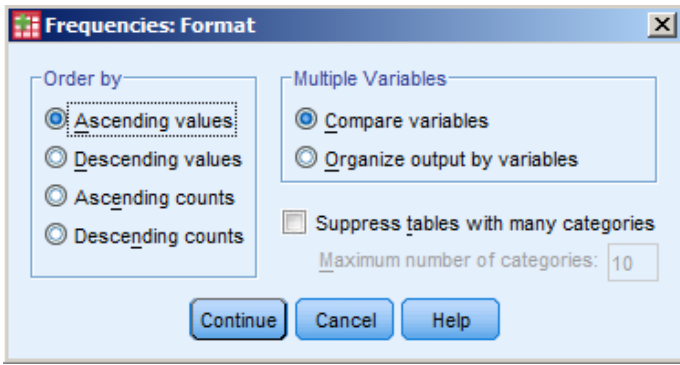
nos esses valores também em termos percentuais. Para fazermos uma **tabela de frequências** temos de escolher na barra de menus *Analyze / Descriptive Statistics / Frequencies*. Abre-se então uma caixa de diálogo com duas janelas. Na do lado esquerdo temos os nomes das variáveis da nossa matriz de trabalho. É nessa janela que teremos de escolher a variável (ou as variáveis, pois é permitido fazer a frequência de quantas variáveis desejamos) que pretendemos analisar. Para tal podemos fazer um *duplo click* sobre a variável, ou fazer apenas um click e premir depois o botão  ao lado com uma seta para deslocar a variável escolhida para a janela da direita, onde ficarão os nomes das variáveis sobre as quais desejamos trabalhar. Pegando no nosso exemplo, em primeiro lugar passamos a variável *Sexo* para a direita e depois premimos o botão *OK*. Existe também a possibilidade de trabalharmos com **gráficos** utilizando para tal o botão *Charts* e a possibilidade de trabalhar com outras estatísticas (botão *Statistics*), possibilidade que estudaremos mais adiante.



Premindo o botão *Charts* surge-nos a janela *Frequencies: Charts*. Aqui podemos escolher se desejamos obter um gráfico de barras, um gráfico circular ou um histograma. Em qualquer um dos casos teremos de escolher se os valores incorporados nos gráficos deverão ser as frequências *per si* ou as respectivas percentagens.







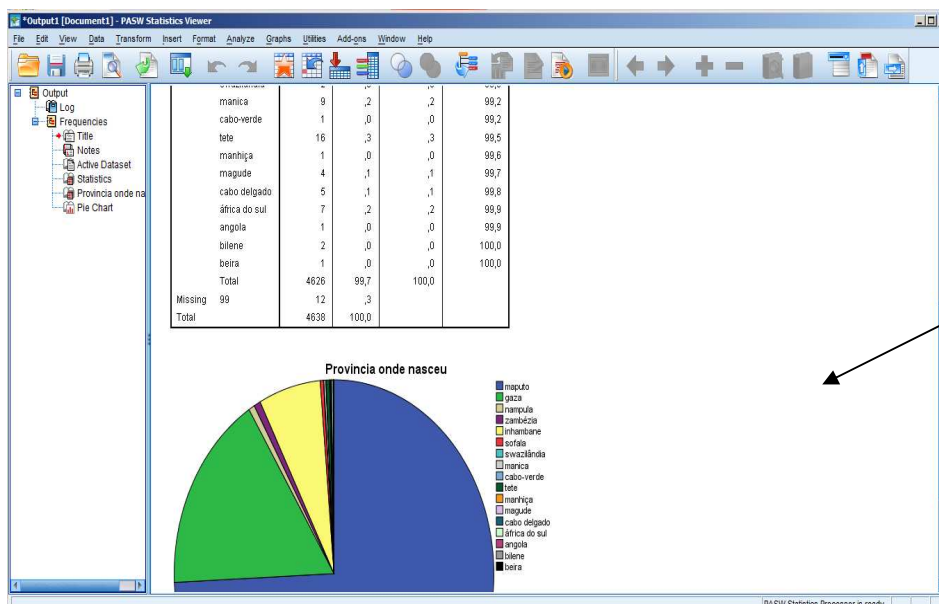
Outro botão que se encontra ainda na janela *Frequencies* é o botão *Format*, que nos permite, entre outras coisas, decidir qual a **ordem** sobre a qual desejamos que **as respostas se ordenem nas tabelas**. Esta função encontra-se no lado esquerdo por baixo de *Order by*. Assim, podemos **ordenar as linhas por ordem ascendente** do número dos códigos (*Ascending Values*), por **ordem descendente** do número dos códigos (*Descending Values*), por **ordem ascendente** em relação ao número de vezes que a resposta se repete (*Ascending Count*) ou por **ordem descendente** em relação ao número de vezes que a resposta sucede (*Descending Count*).

Ainda no âmbito da função *Frequencies* pode-se trabalhar directamente com gráficos ou com outras estatísticas que não as frequências. Para tal, se o nosso objectivo não for a realização de tabelas de frequências, podemos sempre desactivar a função *Display frequency tables* e realizar apenas os gráficos ou as outras tabelas estatísticas que pretendemos. Quando premimos o botão *OK*, surge-nos uma janela, embora já apresentada, ainda não devidamente tratada, **a janela do Output**.

Ainda no âmbito da função *Frequencies* pode-se trabalhar directamente com gráficos ou com outras estatísticas que não as frequências. Para tal, se o nosso objectivo não for a realização de tabelas de frequências, podemos sempre desactivar a função *Display frequency tables* e realizar apenas os gráficos ou as outras tabelas estatísticas que pretendemos. Quando premimos o botão *OK*, surge-nos uma janela, embora já apresentada, ainda não devidamente tratada, **a janela do Output**.

### 5.3- A Janela do Output

Nesta janela o ecrã divide-se em duas partes. O lado esquerdo funciona como um índice da informação que podemos encontrar no lado direito, quer esta esteja visível ou não nesse lado. O lado direito contém o *output* (resultados) em si.

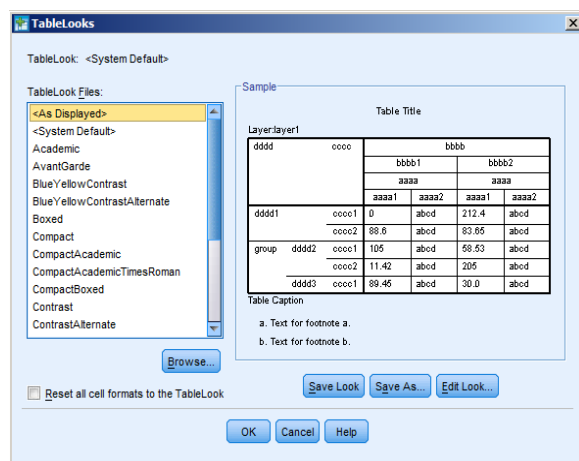


Exemplo de Output



A cada variável corresponde um pequeno ícone, neste caso um pequeno livro amarelo, que nos permite, pressionando sobre o botão correspondente, que o ecrã da direita nos apresente automaticamente os resultados correspondentes a essa variável.

Estes ícones têm também a função de nos permitir esconder informação que não nos interessa que apareça no ecrã da direita (até porque é este ecrã que nos vai servir de base para a impressão), mas sem fazer com que essa informação seja perdida. Para tal basta fazer um duplo clique sobre a informação que não queremos visualizar no ecrã da direita. Em cada ícone existe um pequeno livro aberto. Quando a informação é escondida esse pequeno livro passa a estar fechado. No caso do livro amarelo quando a informação se encontra escondida este aparece apenas com uma pequena linha vermelha, enquanto que quando a informação é visível essa linha é acompanhada por mais duas linhas vermelhas, um pouco mais claras. Esta espécie de índice permite-nos também fazer arranjos na visualização, já que seleccionando o ícone que desejamos com o botão direito do rato, e sem o soltar, podemos deslocar essa informação para onde desejamos, bastando para tal soltar o botão do rato quando já tivermos a informação no local que desejamos. É possível também apagar informação bastando para tal carregar na tecla *Delete* no teclado quando tivermos seleccionado a variável que desejamos apagar. Na parte direita do ecrã (vamos chamar-lhe janela de visualização do *output*) também é possível mover a informação que desejamos visualizar, utilizando as teclas *PgUp* (para mover para cima) e *PgDn* (para mover para baixo) do teclado, ou a barra que se encontra no lado direito. No caso das frequências, nesta janela podemos visualizar o título, que podemos alterar, bastando para tal fazer sobre ele um duplo clique e escrever depois o novo título. Podemos também transformar a visualização desse título com uma barra que nos aparece em cima (podemos alterar por exemplo o tamanho ou a côm das letras). Temos depois a tabela em si mesma, onde aparecem os valores que vamos analisar. Também na tabela existe um título, que à partida é igual ao nome da variável que já vem da edição dos dados, mas que também é possível alterar. Para alterar as características da tabela (quer seja o título, quer seja a parte gráfica da tabela) temos de seleccionar a tabela com um duplo clique, carregar no botão direito do rato e seleccionar a função *Table Properties* (propriedades da tabela). Para a mesma função podemos também seleccionar na barra de menus a função *Format / Table properties*. Aí podem-se definir uma série de características da tabela, como o formato das linhas, dos caracteres, etc. Podemos também alterar o visual propriamente dito da tabela, recorrendo por exemplo na base de menus à função *Format / TableLooks*. Nesta função existem uma série de visuais já definidos pelo computador.



Passemos agora à análise do conteúdo propriamente dito da tabela. A tabela de frequências surge-nos

como uma grelha. Nas colunas encontram-se as estatísticas que vamos analisar. Na primeira coluna encontra-se uma indicação de estarmos na presença de dados válidos (*Valid*) ou inválidos (*Missing*). Esta indicação refere-se à hipótese de os valores não respondidos serem considerados valores em falta ou inválidos (*missing values*). Na coluna à esquerda encontram-se os nomes das variáveis sobre as quais vamos analisar as respostas.

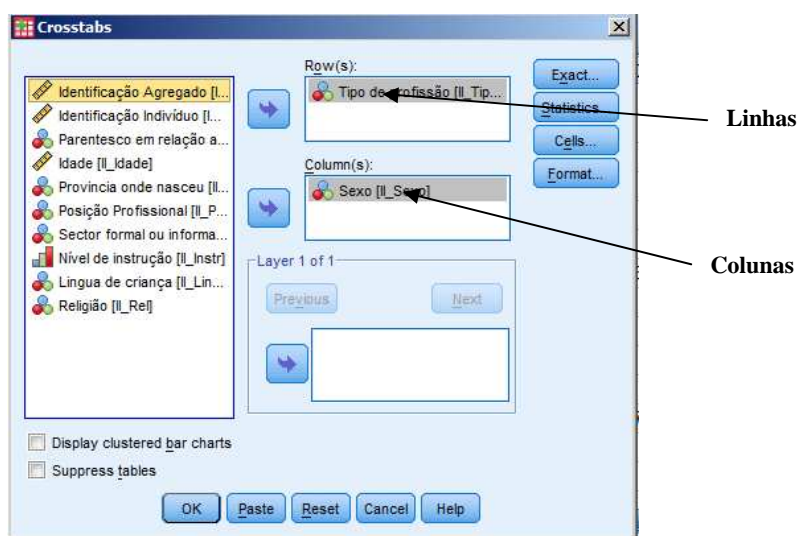
Na primeira coluna propriamente dita, sobre a qual se pode ler *Frequency*, podemos ver o **número de vezes que cada resposta foi dada em termos absolutos**. Na coluna *Percent* temos a **respectiva percentagem**, isto é, o número de vezes que cada resposta é dada em cada 100.

| Provincia onde nasceu |               |           |         |               |                    |
|-----------------------|---------------|-----------|---------|---------------|--------------------|
|                       |               | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid                 | maputo        | 3436      | 74,1    | 74,3          | 74,3               |
|                       | gaza          | 777       | 16,8    | 16,8          | 91,1               |
|                       | nampula       | 28        | ,6      | ,6            | 91,7               |
|                       | zambézia      | 31        | ,7      | ,7            | 92,3               |
|                       | inhambane     | 290       | 6,3     | 6,3           | 98,6               |
|                       | sofala        | 15        | ,3      | ,3            | 98,9               |
|                       | swazilândia   | 2         | ,0      | ,0            | 99,0               |
|                       | manica        | 9         | ,2      | ,2            | 99,2               |
|                       | cabo-verde    | 1         | ,0      | ,0            | 99,2               |
|                       | tete          | 16        | ,3      | ,3            | 99,5               |
|                       | manhiça       | 1         | ,0      | ,0            | 99,6               |
|                       | magude        | 4         | ,1      | ,1            | 99,7               |
|                       | cabo delgado  | 5         | ,1      | ,1            | 99,8               |
|                       | áfrica do sul | 7         | ,2      | ,2            | 99,9               |
|                       | angola        | 1         | ,0      | ,0            | 99,9               |
|                       | bilene        | 2         | ,0      | ,0            | 100,0              |
|                       | beira         | 1         | ,0      | ,0            | 100,0              |
| Total                 | 4626          | 99,7      | 100,0   |               |                    |
| Missing               | 99            | 12        | ,3      |               |                    |
| Total                 | 4638          | 100,0     |         |               |                    |

Para o cálculo desta coluna ainda são tidas em conta as respostas em falta. Só que por vezes essas não respostas alteram um pouco os valores da realidade. Imaginemos que estamos no caso de uma pergunta sobre a profissão de uma pessoa. Se num grupo de 100 pessoas, 50 pessoas respondessem *pedreiro*, 35 *comerciante*, 10 *carpinteiro*, e 5 não respondessem, as percentagens, contabilizando todas as respostas, seriam: 50% pedreiro, 35% comerciante, 10% carpinteiro, 5% não responderam. Se não contabilizarmos essas não respostas as percentagens seriam: 52,6% pedreiro, 36,8% comerciante e 10,5% carpinteiro. Ou seja, incluir essas não respostas não alteraria muito a nossa informação. Se, também num grupo de 100, pessoas 50 respondessem pedreiro, 20 comerciante e 30 não dessem resposta, as percentagens, contabilizando as não respostas, seriam: 50% pedreiro, 20% comerciante e 30% não responderam. Se pelo contrário não fossem contabilizadas essas não respostas, as percentagens seriam bem diferentes: 71,4% pedreiro, 28,6% comerciante. Neste caso talvez esta segunda hipótese seja bem mais realista. Daí a vantagem da utilização da hipótese de **tornar as não-respostas *missing values***. Assim, no lado direito surge outra coluna – *Valid Percent* – que calcula o **número de vezes que cada resposta é dada em cada 100 respostas válidas**, isto é, excluindo as não respostas. Mais à direita encontra-se a coluna *Cumulative Percent* que nos dá o valor das **percentagens acumuladas**. Isto é, o valor de cada linha é igual à soma da percentagem dessa variável com a soma das percentagens das variáveis que estão por cima. Naturalmente que a percentagem acumulada da última variável tem de ser igual a 100. Estes últimos valores serão úteis para algumas análises estatísticas que veremos mais à frente.

## 5.4- Cruzamento de Variáveis

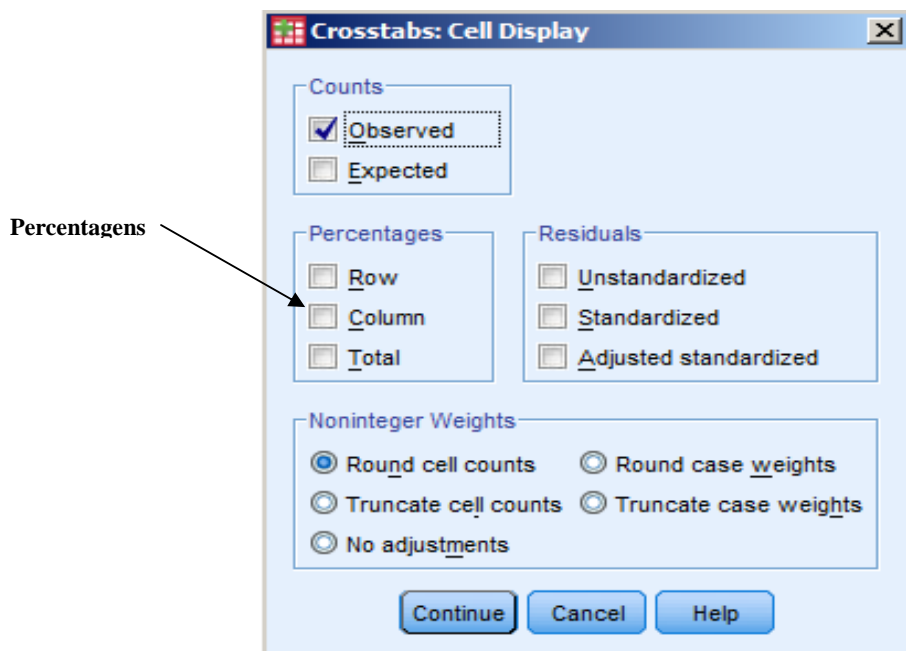
Por vezes a simples análise das frequências é insuficiente para compreendermos bem as respostas aos nossos inquiridos. Por exemplo, no caso dos dados referentes às profissões, podemos analisar uma tabela de frequências com as profissões, e ficarmos com a ideia de quais são as profissões mais frequentes nos nossos inquiridos. Mas, se desejarmos uma **análise mais profunda** pode ser interessante saber qual é a **distribuição dessas profissões pelos vários escalões etários ou até pelo sexo**. Por exemplo, podemos desejar verificar qual é a idade em que as pessoas se dedicam mais ao comércio e depois saber se quem se dedica mais a esse comércio são os homens ou as mulheres. Para realizar essas tarefas o PASW contém outros instrumentos que não as tabelas de frequências – **o cruzamento de variáveis**. Isto é, **uma função que cria uma tabela que faz a contagem de quantas vezes é que determinada combinação de valores de duas ou mais variáveis ocorrem**.



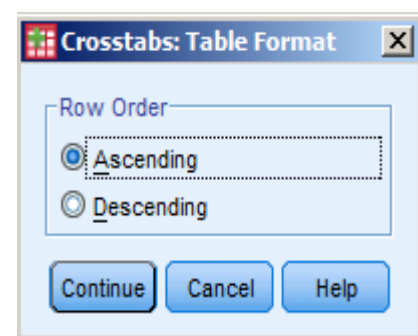
Uma das formas de fazer um quadro de cruzamento de variáveis é na barra de menus escolher *Análise / Descriptive Statistics / Crosstabs*. Como é habitual surge-nos no lado esquerdo da caixa *Crosstabs*. A novidade é a da **existência de duas caixas para receber essas variáveis**. Na de cima está escrito *Rows (linhas)* e é onde vamos colocar as variáveis que na tabela desejamos colocar nas linhas. Na caixa de baixo está escrito *Columns (colunas)* e, naturalmente, é onde vamos colocar as variáveis que na tabela desejamos colocar nas colunas.

Temos aqui um exemplo de uma tabela de cruzamento de duas variáveis (*Sexo* e *Profissão*) de um inquirido imaginário, num local hipotético, onde existiriam apenas duas profissões (carpinteiro e vendedor). Neste caso, quando estivéssemos na caixa *Crosstabs*, passávamos a variável *Profissão* para *Rows* e a variável *Sexo* para *Columns*. Para finalizar a operação é suficiente premir *OK*.

Podemos visualizar as percentagens das células. Para tal, na caixa de diálogo *Crosstabs*, apontando a opção *Cells*, devemos escolher, em *Percentages*, *Row*, *Column* ou *Total*.



Em princípio, na tabela as colunas surgirão pela ordem numérica dos seus códigos. Neste caso, em princípio *Masculino* tinha o código 1 e *Feminino* tinha o código 2. Mas nós podemos desejar alguma ordem diferente. Podemos desejar, por exemplo, que mais à esquerda surjam as colunas com maior número de respostas (ao contrário do que acontece na nossa tabela). Para tal, ainda na mesma caixa, na opção *Format*, devemos seleccionar a opção *Descending* – isto é os valores das colunas vão decrescendo da esquerda para a direita. Pelo contrário, se escolhermos *Ascending*, os valores das colunas vão crescendo da direita para a esquerda. Depois de realizadas estas operações premimos o botão *Continue*.



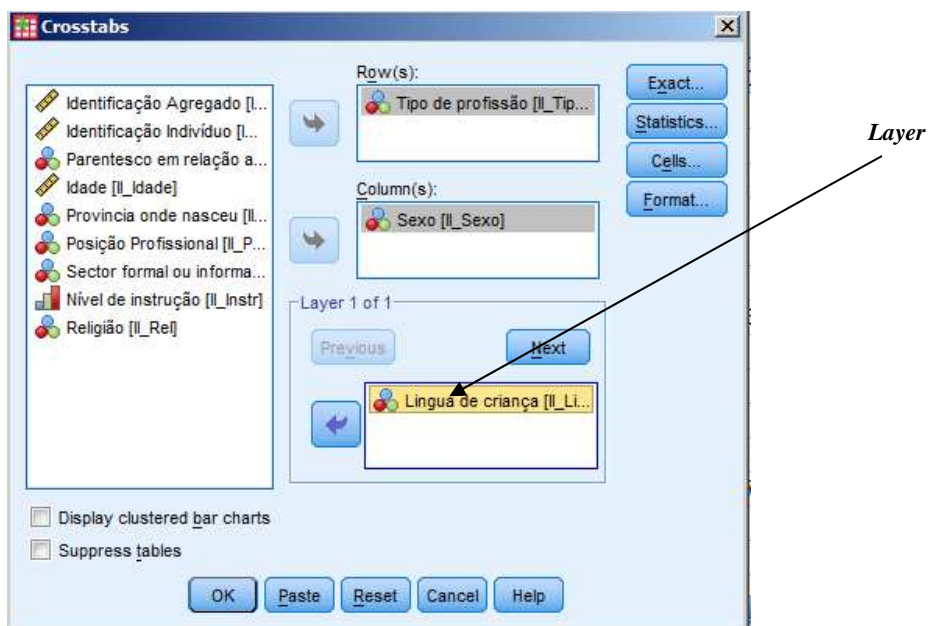
Em princípio, depois de realizados todos estes procedimentos bastará pressionar o botão *OK* para que nos surja a tabela na janela do *Output*.

DISTRIBUIÇÃO DAS PROFISSÕES POR SEXO


|           |             | sexo |       |     |       | Total |       |
|-----------|-------------|------|-------|-----|-------|-------|-------|
|           |             | Masc |       | Fem |       | N     | %     |
|           |             | N    | %     | N   | %     |       |       |
| Profissão | Carpinteiro | 4    | 40.0  | 8   | 66.7  | 12    | 54.5  |
|           | Vendedor    | 6    | 60.0  | 4   | 33.3  | 10    | 45.5  |
| Total     |             | 10   | 100.0 | 12  | 100.0 | 22    | 100.0 |

Depois de seleccionada a tabela poderemos ainda proceder a alterações directamente na tabela, consoante os nossos desejos. É possível também que este cruzamento de variáveis seja feito com mais de duas variáveis. Para o efeito é suficiente que, ainda no contexto da caixa de diálogo *Crosstabs*,

seleccionemos as variáveis a cruzar (qualquer uma da terceira em diante) para a caixa *Layer*, premindo sucessivamente o botão *Next*.



## 5.5- Variáveis de Resposta Múltipla

Por vezes nos inquéritos surgem questões em que é possível recolher mais que uma resposta, as chamadas **Respostas Múltiplas**. Para resolver essa questão o *PASW* possui uma função que permite criar variáveis para analisar essas respostas, as **Variáveis de Respostas Múltiplas (VRM)**. Este tipo de variáveis pode ser recolhido de duas formas. Vejamos uma situação em que o inquirido é confrontado com um campo com diversos itens, nas quais deve seleccionar aquelas que se aplicam à sua situação. Imaginemos uma situação ligada às razões de emigração, em que o indivíduo pode apresentar mais do que uma razão. No caso de um campo de selecção, isso significaria que existiria uma listagem com as várias opções, das quais os indivíduos seleccionaram as suas razões. As variáveis assim codificadas são designadas como questões de dicotomia múltipla, uma vez que cada item possui dois estados - seleccionado ou não. O primeiro passo para a análise deste género de variáveis é a **criação de uma variável para cada uma das opções. Todas as variáveis estarão codificadas com 1=sim, 2 =não**. Passaremos agora para a **fase de criação das VRM**. Na barra de menus escolheremos a opção *Analyze / Multiple Responses / Define Variable Sets*. Surgirá então a caixa de diálogo *Define Multiple Response Sets*. O primeiro passo é a **escolha das variáveis que desejamos que pertençam ao conjunto**. No lado direito teremos o conjunto total das variáveis da nossa matriz, no qual seleccionaremos aquelas que desejamos analisar como VRM e que passaremos para a janela da esquerda (*Variables in Set*) pressionando o botão  .