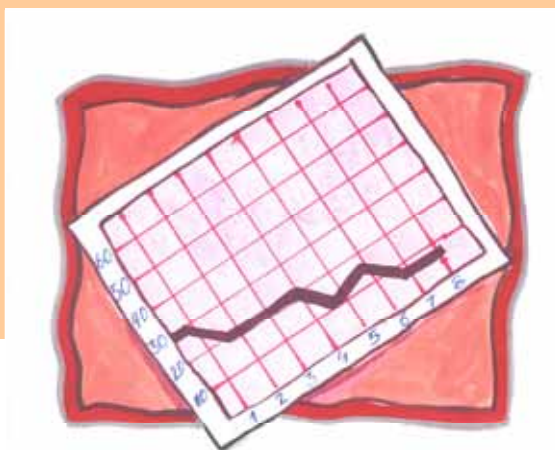




Dossiês Didáticos



IX – Representações Gráficas

Notas sobre a criação e apresentação de alguns tipos de gráficos

Ana Alexandrino da Silva

Nota Introdutória

O projecto ALEA - Acção Local de Estatística Aplicada - constitui-se como um contributo para a elaboração de novos suportes de disponibilização de instrumentos de apoio ao ensino da Estatística para os alunos e professores do Ensino Básico e Secundário. Este projecto nasceu de uma ideia conjunta da Escola Secundária Tomaz Pelayo e do Instituto Nacional de Estatística, assente nas necessidades e estruturas que os intervenientes possuem. Melhorar a literacia estatística é, assim, uma condição importante para garantir uma melhor prestação de um serviço de utilidade pública. O Ensino da Estatística no Ensino Básico e Secundário constitui um dos instrumentos mais importantes para cumprir esse objectivo. A página Internet do ALEA está no endereço: <http://alea-estp.ine.pt>.



A área Dossiês Didáticos foi concebida para apoiar a elaboração de materiais didáticos sobre temáticas variadas (População e Demografia, Inquéritos, Inflação e Preços, etc.). Apresentamos agora o **Dossiê IX – Representações Gráficas** – em formato PDF (Adobe Acrobat) e em papel.

Números anteriores:

Dossiê I - População e Demografia – Quantos Somos e Como somos.

Dossiê II – Ambiente e Recursos.

Dossiê III – A Inflação e o índice de preços no consumidor.

Dossiê IV – Estatística com Excel – Uma aplicação das noções.

Dossiê V – Censos 2001.- Tu também contas.

Dossiê VI – Notas sobre a História da Estatística.

Dossiê VII – Probabilidades com Excel.

Dossiê VIII – Números do Cinema

Índice

1.1. Introdução	3
História dos gráficos	3
Reflexões sobre a construção de gráficos	3
Formatação do gráfico	5
Estudos perceptivos	10
Elementos do gráfico	12
1.2. Gráficos de barras.....	18
Gráficos de barras simples (verticais ou horizontais).....	18
Algumas regras relacionadas com a construção dos gráficos de barras	20
Gráficos de barras agrupadas	23
Gráficos de barras empilhadas	25
Histograma	26
Pirâmide Etária	27
Séries temporais em Gráficos de barras.....	28
1.3. Gráficos de linhas.....	29
Gráficos de área	31
1.4. Gráficos circulares.....	31
1.5. Pictogramas	33
1.6. Ver também	34

Para mais informações sobre o tema contemplado neste dossiê contacte:

Ana Alexandrino da Silva

alea@ine.pt

1.1. Introdução

Os gráficos encontram-se presentes em quase todos os meios de divulgação de informação, designadamente nos jornais e revistas, nos manuais escolares, nas apresentações públicas e até os nossos relatórios individuais já não passam sem eles.

Contudo, fazer um gráfico ou um mapa que de facto informe e seja, simultaneamente, apelativo, legível e coerente com os dados não é tarefa fácil...

A grande vantagem dos gráficos reside na sua capacidade de contar uma história de

forma interessante e atractiva permitindo compreender rapidamente fenómenos que dificilmente seriam percebidos de outra forma. Contudo, tal não implica que este processo seja feito de forma simples, sendo necessário muito trabalho e cuidado.

Existem inúmeras formas de apresentar figurativamente a informação estatística e no caso particular dos gráficos são tantas as possibilidades que houve necessidade de restringir o objecto deste dossiê aos gráficos mais correntes e não proceder a uma abordagem exaustiva.

História dos gráficos

A história dos gráficos estatísticos é relativamente recente. O maior avanço deu-se apenas há cerca de 200 anos, em 1786, graças a William Playfair que inventou a maioria das formas gráficas que conhecemos hoje: o gráfico de barras, o gráfico de linhas baseado em dados económicos e o gráfico circular.

Enquanto no século XIX, se assistiu à criação e disseminação alargada dos gráficos estatísticos na comunidade científica, no século XX houve um aumento exponencial da sua utilização em documentos de divulgação alargada e acessíveis ao grande público.

Desde Playfair muito se avançou na divulgação dos gráficos estatísticos, usados agora um pouco por todo o lado - nas escolas, nos média, etc. mas a maioria dos gráficos actualmente em uso datam desse tempo (século XVIII/XIX).

Com o aparecimento dos computadores retomaram-se os estudos desenvolvidos na área dos gráficos sendo imperativo fazer referência a Edward TUKEY (1977) responsável pela invenção de gráficos indispensáveis na análise exploratória de dados, como sejam a caixa de bigodes e o diagrama de caule e folhas, entre outros.

Reflexões sobre a construção de gráficos

Com a tecnologia existente, a produção de gráficos está ao alcance de todos. Mas é importante ter alguns cuidados.

Neste dossiê serão compilados um conjunto de critérios subjacentes à criação de um gráfico. Este processo inicia-se no momento

em que se decide optar por um gráfico e só termina quando o resultado se considera satisfatório.

Com a enchente de gráficos que se vive nos dias de hoje, o leitor tornou-se exigente. A reacção a um gráfico demasiado ‘carregado’ de informação, pode ser o afastamento, e mesmo que lhe seja dedicado alguma atenção, poucas recordações subsistem. Este distanciamento também pode ser causado por um excesso de elementos gráficos não informativos, originando gráficos apelidados por TUFTE (1983) de lixo gráfico (*chart junk*).

Antes de mais, deve questionar-se a necessidade de mostrar os dados graficamente. De facto, em certos casos, não fará sentido recorrer a um gráfico quando o objectivo não é dar uma imagem, mas sim fornecer dados concretos, quer em situações em que apenas se detêm poucos valores como para os casos em que se pretendem divulgar muitos dados.

Outro dos problemas com que se debate quem produz gráficos é a restrição de espaço, obrigando à acumulação de informação num único gráfico ou a um dimensionamento reduzido das imagens, com consequências na sua leitura.

WALLGREN (1996) sintetiza esta fase preparatória em oito perguntas que não podem ser respondidas separadamente:

- Um gráfico é realmente a melhor opção?
- Qual é o público-alvo?
- Qual é o objectivo do gráfico?

- Que tipo de gráfico se deve usar?
- Como deve ser apresentado o gráfico?
- Qual deve ser o tamanho do gráfico?
- Deverá ser usado apenas um gráfico?
- A que meios técnicos se deve recorrer?

Após ter sido seleccionado o modelo de gráfico mais adequado ao contexto respectivo, inicia-se a construção do gráfico propriamente dita.

Quando finalmente se pensa ter obtido o gráfico pretendido, torna-se fundamental proceder a uma análise crítica, no sentido de compreender se esta é a forma mais eficaz de transmitir a mensagem inicial. Um gráfico mal compreendido pode provocar uma interpretação errada. Por outro lado, um gráfico visualmente desagradável pode afastar o leitor, em vez de o informar: “*Um mau gráfico é pior do que nenhum gráfico*” (WALLGREN, 1996, p. 89).

Na tentativa de encontrar a melhor imagem que satisfaça todos os requisitos iniciais, entra-se num processo iterativo que só termina quando se garante uma elevada legibilidade e pertinência. Por conseguinte, a adopção do gráfico apenas se pode consumir após serem formuladas, e convenientemente respondidas, as seguintes perguntas:

- O gráfico é fácil de ler?
- O gráfico pode ser mal interpretado?
- O gráfico tem o tamanho e a forma certa?
- O gráfico está localizado no sítio certo?
- O gráfico beneficia por ser a cores?
- A compreensão do gráfico foi testada com alguém?

Formatação do gráfico

A representação gráfica é um tema complexo onde se cruzam áreas tão diversas como a estatística, o desenho e a psicologia. Um gráfico pode representar correctamente as variáveis, conter todos os elementos necessários e não ser, nem atractivo, nem de fácil leitura.

É possível redesenhar um gráfico, através da modificação ou supressão de alguns

elementos gráficos, sem que haja perda de informação (TUFT, 1983). No entanto, muitos dos gráficos divulgados necessitam de uma certa sofisticação a este nível, sendo comum encontrar imagens visualmente semelhantes provenientes do *assistente de gráficos* do software *Excel*, que por serem imagens muito vistas, e portanto cansativas, não atraem o leitor.

O *Excel* permite alguma manipulação visual no leque de gráficos que apresenta. Seguidamente, é apresentado um exemplo de como se pode melhorar a leitura, modificando o aspecto do gráfico.

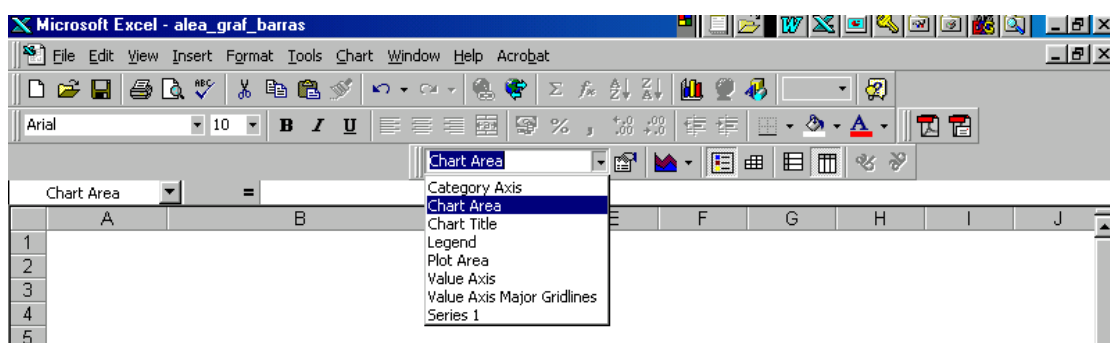
A primeira coisa a ter em conta quando se pretende elaborar um gráfico é a organização dos dados. O tipo de gráfico seleccionado é influenciado pela forma como estão dispostos os dados. A melhor forma é dispor os dados numa tabela, com as respectivas identificações, para que estes possam ser utilizados como títulos e legendas do gráfico.

A tabela dos dados:

Qualificação académica da população dos 15-64 anos

Qualificação académica	Sexo	
	Masculino	Feminino
Nenhum	7,5%	11,3%
Obrigatório	69,3%	61,5%
Secundário	15,7%	16,7%
Superior	7,5%	10,5%

Possibilidades de formatação de gráficos com o Excel



1. Área do gráfico (*chart area*)
2. Legenda (*legend*)
3. Eixo das categorias (*category axis*)
4. Área do desenho (*plot area*),
5. Eixo de valores (*value axis*),
6. Linhas de grelha (*gridlines*)
7. Série de dados (*series*)

Descrição do processo de formatação

Partindo do critério de que pelo menos dois terços da área do gráfico devem ser afectados às barras ou, genericamente, à área do desenho, (SCHMID, 1992), ampliou-se o espaço preenchido por estas.

No eixo dos valores foram retiradas as casas decimais e suprimidos alguns valores, apesar de se terem mantido as respectivas linhas de grelha. Poder-se-ia ter deixado apenas o sinal de % junto ao último valor, retirando os sinais de % nos valores 0 e 40. Foi também retirada a linha do eixo e as marcas dos eixos, para além de se ter encurtado a amplitude do intervalo de valores dado que a maior das barras não ultrapassava os 80%.

A linha do eixo das categorias apresenta um maior peso visual do que as restantes linhas auxiliares, estando as designações orientadas horizontalmente para facilitar a leitura.

Foram retiradas as molduras do gráfico, da legenda e das barras por se considerar não existir qualquer vantagem em mantê-las, sobrecarregando desnecessariamente a construção gráfica, e posicionou-se a legenda no interior do gráfico para diminuir a distância percorrida pelos olhos entre as componentes e as suas designações. Mudaram-se as cores das barras, aumentou-se a sua grossura e simultaneamente diminuiu-se o espaço entre grupos de barras.

A figura da direita não é mais do que a figura da esquerda depois de transformada recorrendo às potencialidades do software.

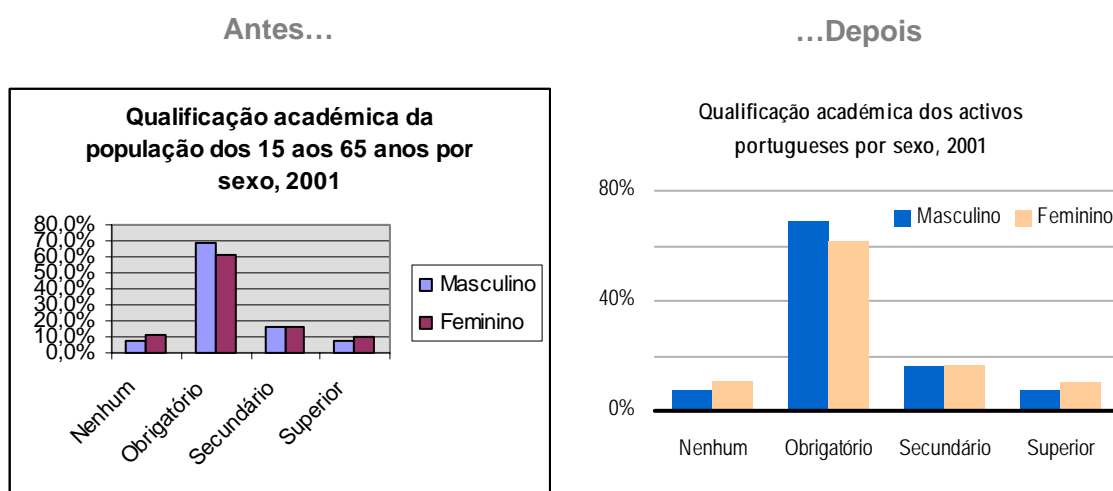


Figura 1 – Gráfico de barras antes e depois de ser modificado através do Excel

(Re)desenho do gráfico através do Excel

1 – Área do gráfico

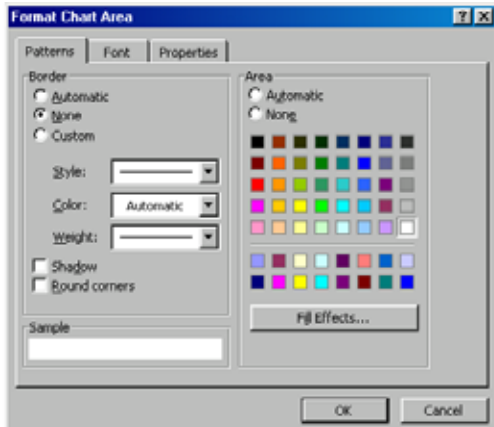


Gráfico sem moldura e com área a branco...

Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001

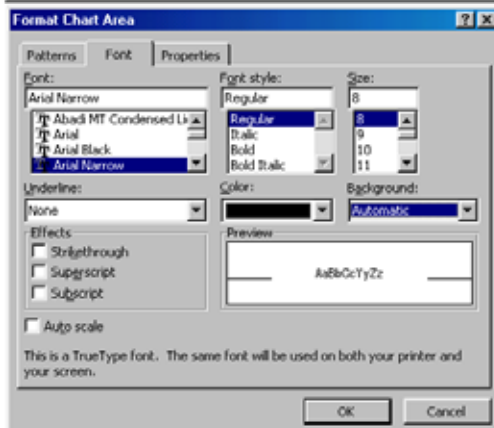
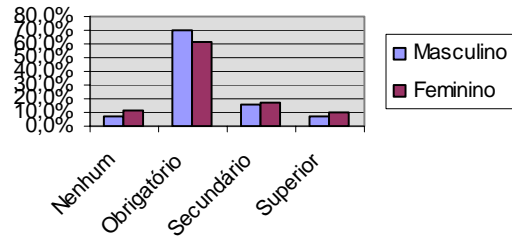
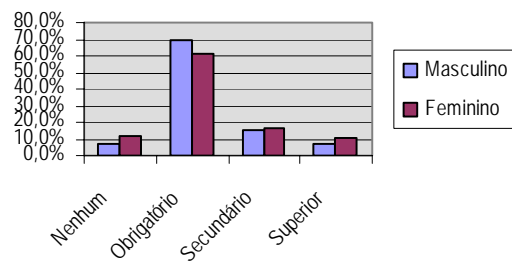


Gráfico com tipo de letra Arial narrow, tamanho 8...

Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001



2 – Legenda

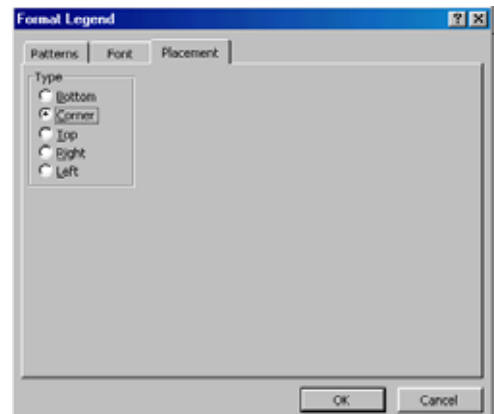


Gráfico com legenda no canto superior direito...

Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001

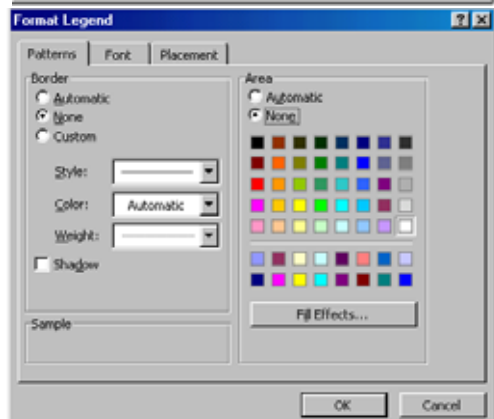
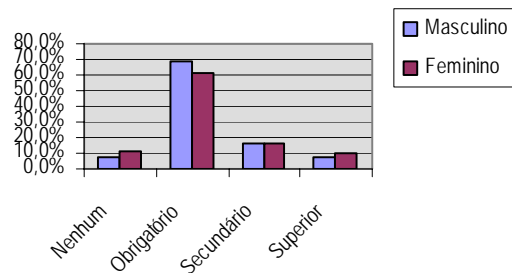
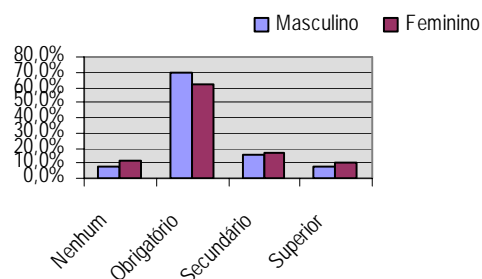


Gráfico com legenda sem moldura, fundo e símbolos na horizontal...

Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001



3 - Eixo das categorias

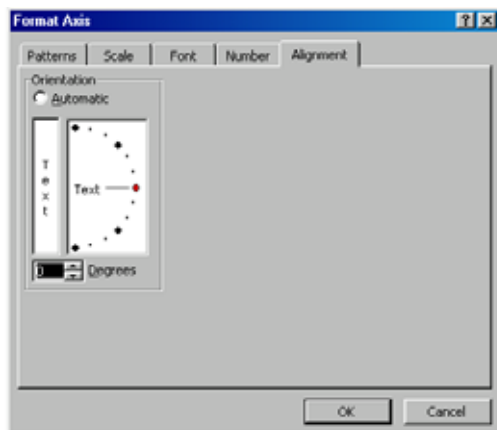
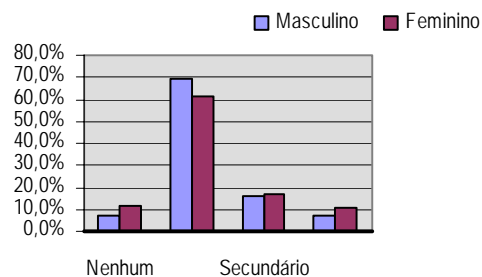


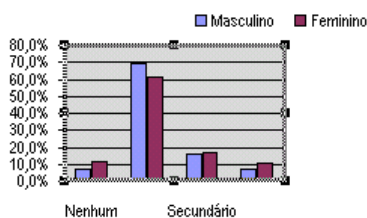
Gráfico com identificações das categorias na horizontal...

Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001

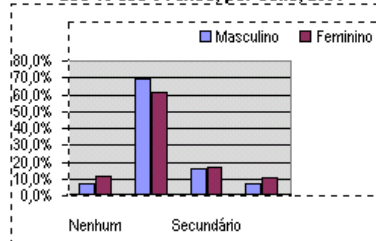


Modificação de elementos dentro do gráfico - É possível aumentar a área do desenho para que todas as designações apareçam.

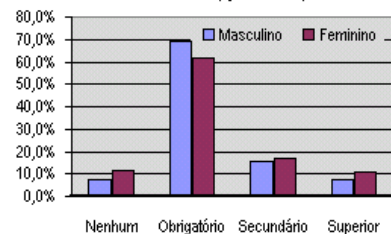
Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001



Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001



Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001



4 - Eixo dos valores



Gráfico sem linha e *tick marks* no eixo dos valores...

Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001

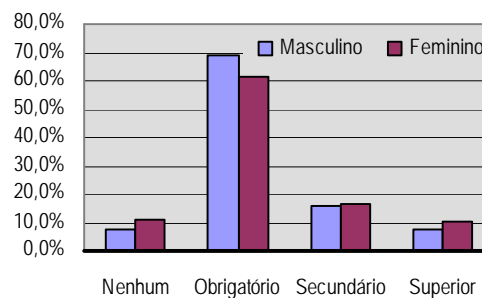
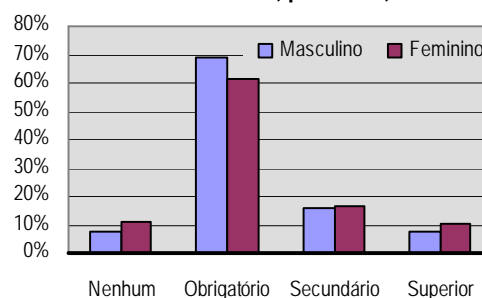


Gráfico com eixo de valores sem casas decimais...

Qualificação académica da população dos 15 aos 64 anos, por sexo, 2001



5 - Área do desenho

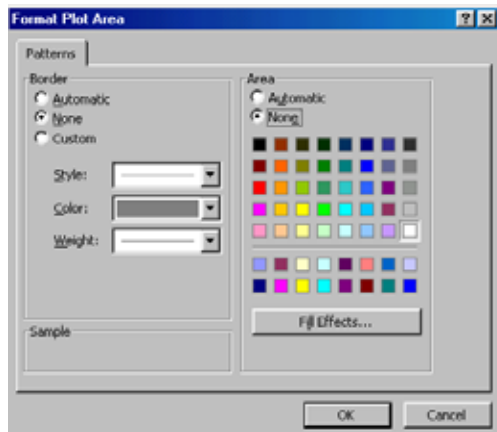
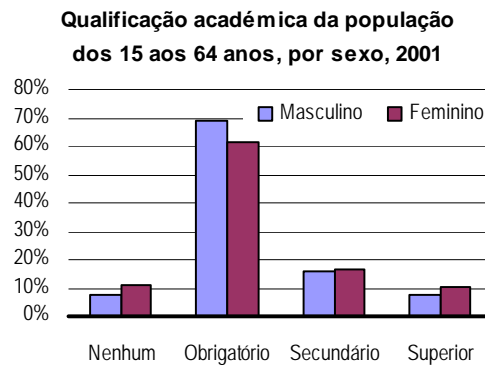


Gráfico com área de desenho a branco e sem moldura...



6 - Linhas de grelha

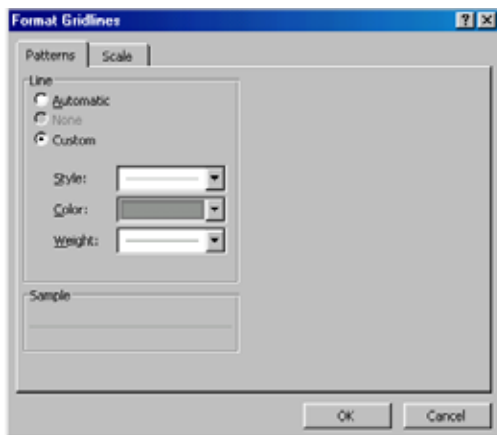


Gráfico com linhas de grelha a cinzento...

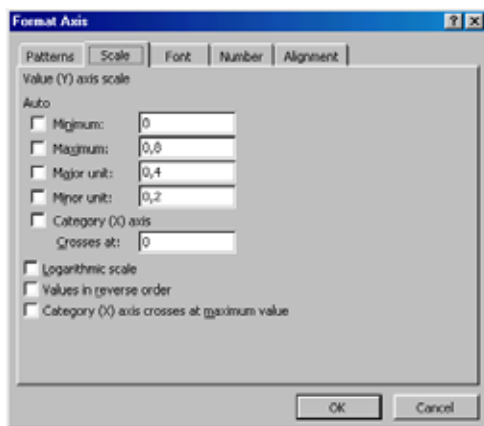
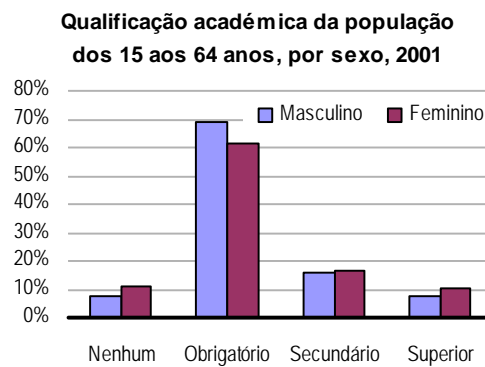
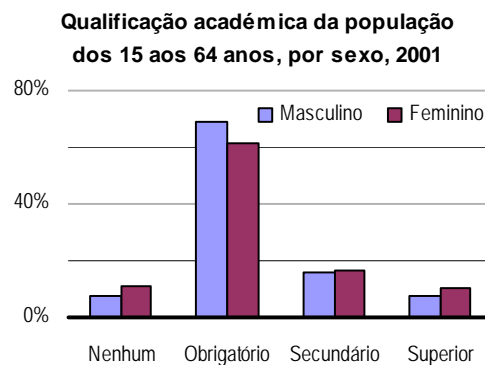
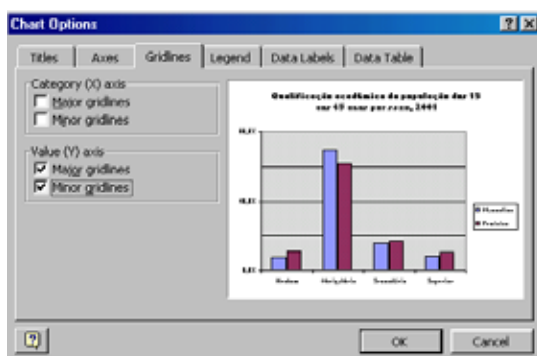
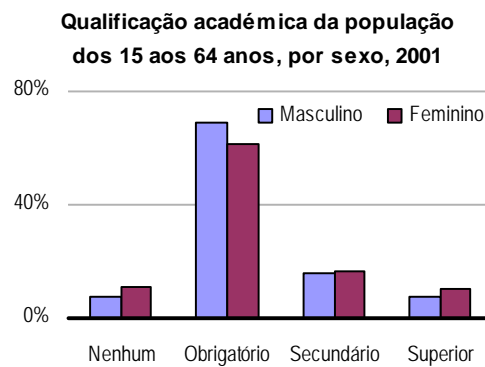


Gráfico com escala de valores para os dois tipos de linhas de grelha...



7 - Série de dados

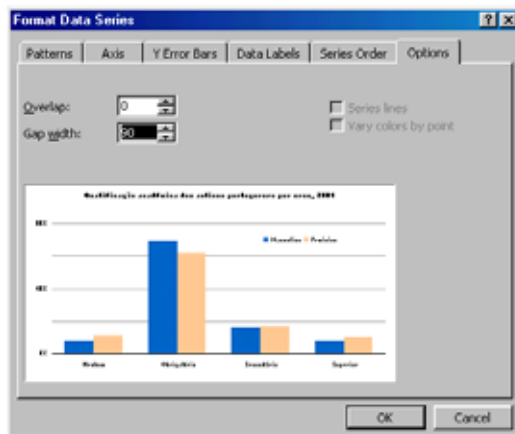
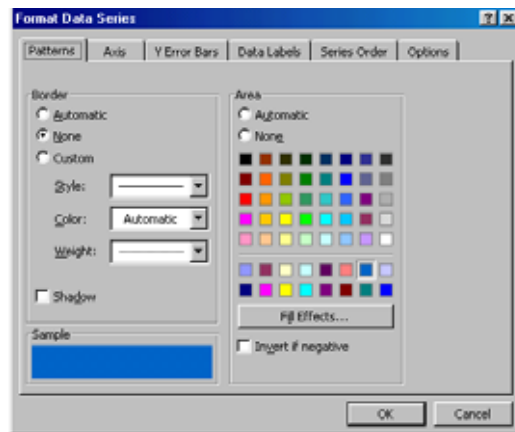


Gráfico com barras de cor diferente e sem moldura...

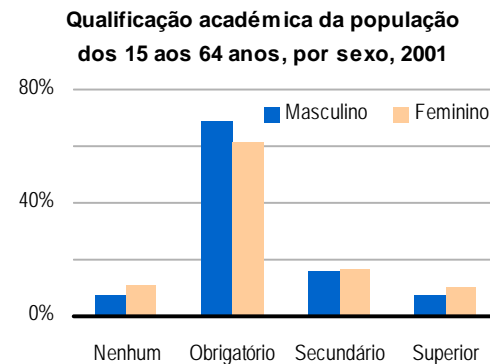
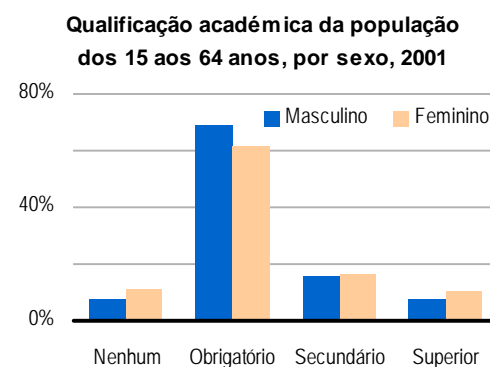


Gráfico com espaço entre as barras alterado...



Estudos perceptivos

A percepção gráfica é um dos elementos mais importantes a ter em conta quando se elabora um gráfico, porque permite dar uma fundamentação científica à construção gráfica e sustentar a escolha de uma forma gráfica em detrimento de outra. A leitura das imagens pode ser condicionada pela dificuldade em estimar correctamente os dados representados.

Na fase da construção, a informação é codificada no gráfico através de símbolos, comprimentos, declives dos segmentos de recta, áreas, textura ou cor. Quando um gráfico é analisado, a informação codificada é

visualmente descodificada, sendo o processo de descodificação, denominado de percepção gráfica, um factor de controlo na capacidade de um gráfico transmitir informação (CLEVELAND, MCGILL, 1987).

A extracção de informação a partir dos gráficos envolve tarefas perceptivas realizadas pelo sistema visual olho-cérebro. No quadro seguinte, estas tarefas estão ordenadas segundo a sua precisão na extracção de informação quantitativa. Quanto menos precisa for a tarefa perceptiva maior o erro de leitura, ou seja, maior a diferença entre o valor percebido e o valor correcto.

Mais preciso ↓ Menos preciso	Posição numa escala comum		A
	Posição em escalas não alinhadas		B
	Tamanho		C
	Ângulo		D
	Declive		E
	Área		F
	Volume		G

Figura 2 – Avaliação de tarefas perceptivas ordenadas segundo a sua precisão
(adaptado de CLEVELAND, MCGILL, 1984, 1987)

Por exemplo, nos gráficos de barras agrupadas, o leitor estima os valores através da posição das barras na mesma escala ou em escalas separadas, consoante a forma de apresentação dos dados.

A comparação entre barras próximas (Figura 3 - A) é melhor do que a comparação entre barras mais afastadas (Figura 3 - B), ou seja, nesta última forma o leitor tem mais dificuldade em estimar os valores.

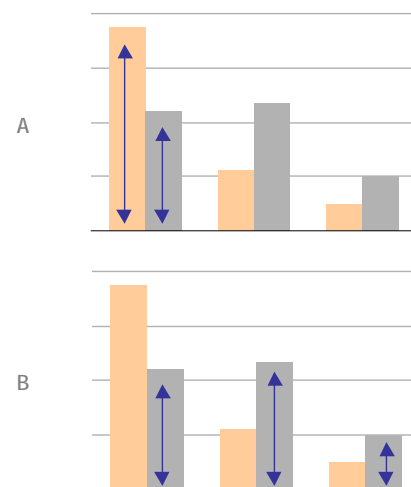


Figura 3 – Exemplos das tarefas A e B

Na comparação entre gráficos de barras e gráficos circulares, os primeiros revelaram-se perceptivamente mais adequados, dado que a estimação dos comprimentos demonstrou ser duas vezes mais precisa que a estimação de ângulos. Veja-se o caso da região Norte e da região de Lisboa e Vale do Tejo. No gráfico circular não se tem a percepção de qual é o maior. Pelo contrário, o gráfico de barras mostra claramente a diferença.

É comum encontrar gráficos a três dimensões em que a profundidade não descreve qualquer variável. Como o volume é o que maiores problemas traz em termos de percepção, não deve ser utilizado.

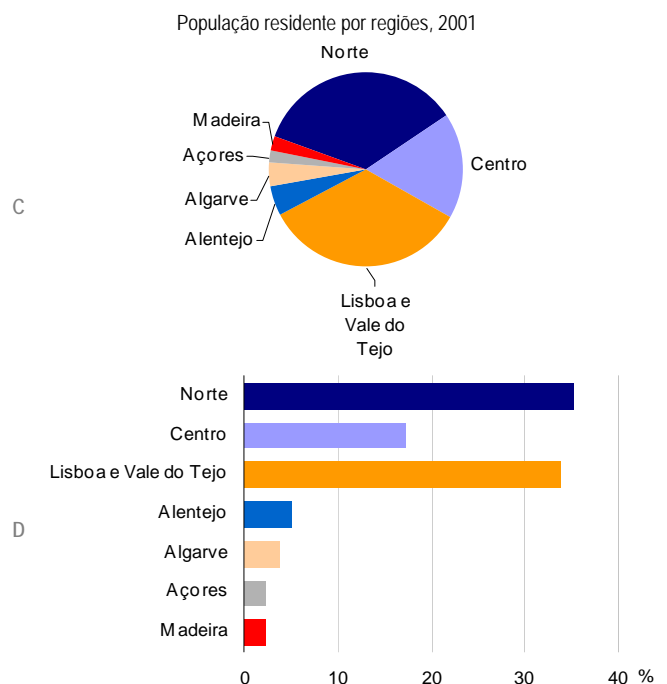


Figura 4 – Exemplos das tarefas C e D

Elementos do gráfico

Os gráficos incorporam o seguinte conjunto de elementos: o título, os eixos de valores e de categorias (baseados no sistema de coordenadas), a legenda, as identificações dos dados e as linhas auxiliares (Figura 5).

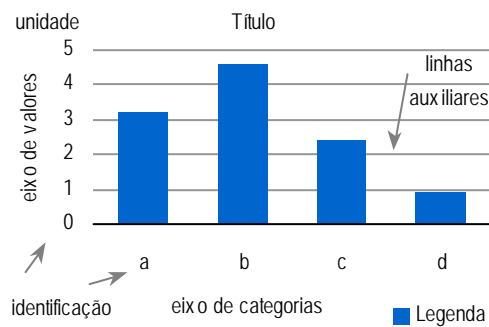


Figura 5 – Elementos de um gráfico

Estes elementos são constituídos por símbolos gráficos (pontos, linhas, números, letras, etc.) e sua variação (cor, valor, etc.).

A área do gráfico pode conter todos estes elementos, ou apenas alguns, sistematizados em duas áreas complementares: a 'área do desenho' (*plot area*) onde está contida a representação gráfica propriamente dita e a 'área exterior' (*chart area*) onde normalmente estão posicionadas as componentes de auxílio à leitura (título, legenda e identificações).

Área exterior:

Título

O título deve estar presente em qualquer tipo de representação gráfica e ser escrito com vista a orientar o leitor na sua interpretação. Para tal, deve ser redigido por forma a responder às perguntas: O Quê, Onde e Quando. Simultaneamente, deve ser conciso, relevante e claro, ou seja, conter apenas informação essencial para uma interpretação correcta do gráfico. Por exemplo, um gráfico

integrado numa publicação temática, relativa a uma dada região ou a um certo período temporal não necessita de incluir sistematicamente a mesma referência regional ou temporal. Sugere-se, igualmente, o posicionamento do título antes do gráfico funcionando como um cabeçalho, centrado horizontalmente (SCHMID, 1992) ou alinhado à esquerda (WALLGREN, 1996).

Identificações (ou rótulos)

Neste conceito genérico enquadra-se toda a informação escrita posicionada na área exterior: as designações dos eixos de valores e categorias, a referência às respectivas unidades e eventuais notas (fontes da informação, esclarecimentos, etc.).

A orientação de todas as palavras deve ser, preferencialmente, horizontal e estar de acordo com o sentido da leitura das palavras

escritas na língua, no nosso caso, da esquerda para a direita.

Na maior parte dos gráficos ou tabelas não se justifica uma grande precisão nos dados apresentados.

Um número excessivo de casas decimais (separadas das unidades por uma vírgula), ou mesmo uma casa decimal em valores elevados, envolve um rigor desnecessário e

prejudicial à leitura. Para ser mais legível, a formatação de valores acima dos milhares pode ser feita com um espaço em vez de com um ponto ou uma vírgula.

Os valores da escala devem ser expressos em valores arredondados múltiplos de 1, 2 e 5

(ex. 5, 10, 25, 50, 100, etc.). Aconselha-se a que não se apresentem números com mais de 5 dígitos, adaptando, caso seja preciso, a unidade para milhares ou milhões.

Legenda

Uma boa legenda deve fazer mais do que simplesmente etiquetar as componentes do gráfico. Deve dizer-nos o que é importante e qual é o objectivo do gráfico: informar o leitor e obrigar quem faz o gráfico a estruturar a informação (CLEVELAND, MCGILL, 1984a).

A legenda é constituída por símbolos e respectivas designações. O preenchimento dos símbolos (cor ou outros) deve ser realizado de modo a que não haja lugar para qualquer confusão visual entre eles e, conseqüentemente, para que exista uma ligação clara entre os símbolos e a componente representada. As designações, por seu lado, devem ser claras e concisas, deixando para notas adjacentes eventuais esclarecimentos.

Os símbolos devem aparecer na mesma ordem que as respectivas componentes: horizontalmente quando estão lado a lado (Figura 6) e verticalmente quando estão umas sobre as outras (WALLGREN, 1996).

Aconselha-se a manutenção da legenda para gráficos em que as componentes surjam mais do que uma vez (Figura 6).

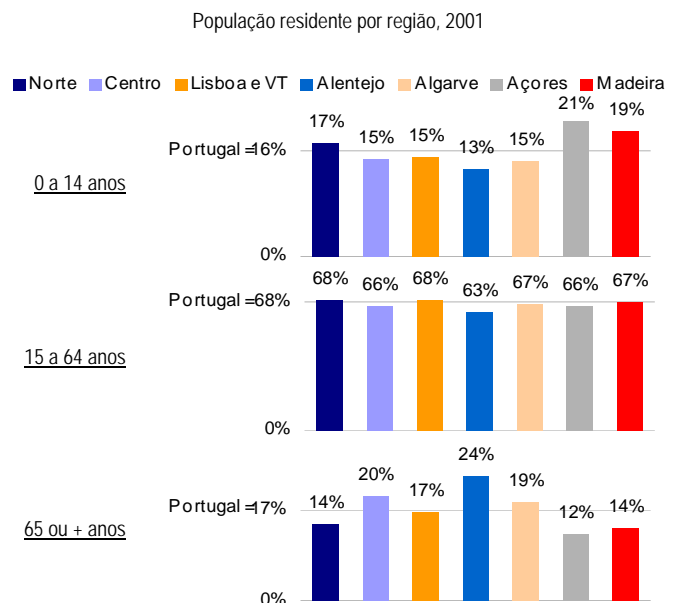


Figura 6 – Gráfico com uma legenda comum

Note-se que a localização da legenda na área exterior obriga o sistema visual a alternar a procura de informação entre a legenda e o gráfico, dificultando a sua interpretação imediata. Por este facto, é aconselhada sempre que possível a omissão da legenda e o posicionamento das designações junto das respectivas componentes, nomeadamente em gráficos de linhas (ver Figura 8) e circulares.

As designações da legenda podem ser deslocadas da 'área externa' para a 'área do desenho', permitindo não só que o próprio gráfico ocupe menos espaço, mas também diminuir a distância percorrida pelo sistema visual (ver Figura 10, onde essas designações surgem junto às linhas dos dados).

Área do desenho:

Eixo de categorias ou variáveis

Neste eixo estão posicionadas as variáveis ou categorias que se pretendem retratar. No caso de gráficos que representam séries que evoluem ao longo do tempo, a este eixo estão associados os períodos temporais, em que a cada mês, trimestre, ano ou outro, corresponderá apenas um ponto ou uma barra no gráfico. Esta relação é obviamente

unívoca, ou seja, não faz sentido representar numa mesma barra valores anuais e semestrais, ou no eixo anos e décadas, ou no mesmo espaço valores anuais e trimestrais (TUFTE, 1983).

O eixo das categorias deve ser visualmente mais ‘pesado’ do que as restantes linhas auxiliares (Figura 5) (SCHMID, 1992).

Linhas auxiliares (ou linhas de grelha)

Um dos elementos gráficos visualmente mais monótono são as linhas auxiliares. Devem, por isso ser suprimidas ou abafadas de tal forma que a sua presença se torne implícita. Ainda que possam auxiliar a leitura dos dados, a maioria das linhas auxiliares escuras tem um grande peso visual, encobrindo muitas vezes, o mais importante do gráfico: a informação. Quando forem realmente necessárias deve-se optar por usar uma cor neutra e, no caso particular de um fundo branco, a cor cinzenta (Figura 7).

Em certos casos, em particular nas séries temporais, pode ser considerado importante incluir linhas auxiliares verticais como auxílio à leitura de valores, por forma a complementar a leitura evolutiva da série com a leitura de valores em particular (Figura 8).

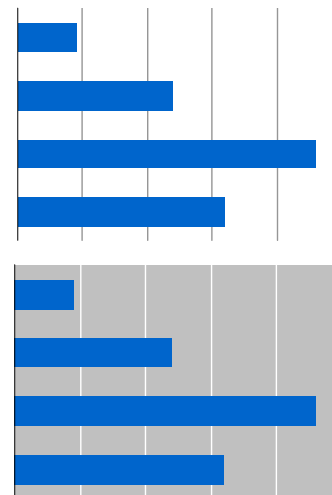


Figura 7 – Linhas auxiliares em fundo branco e de cor

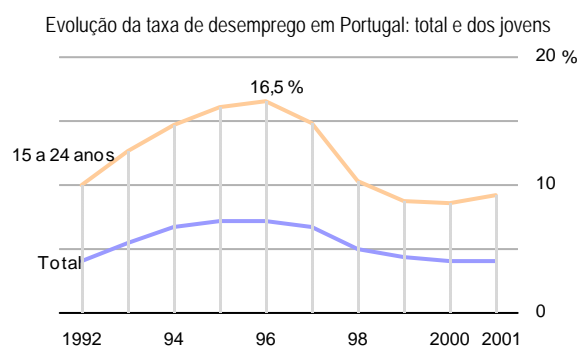


Figura 8 - Linhas auxiliares verticais num gráfico de linhas

Eixo de valores

Na maioria dos gráficos de séries temporais, os dados mais recentes estão situados à direita e longe das identificações do eixo dos valores, normalmente localizados à esquerda (Figura 9), fazendo com que o olho humano tenha que se movimentar alternadamente entre os dados e os valores ao longo das margens do gráfico.

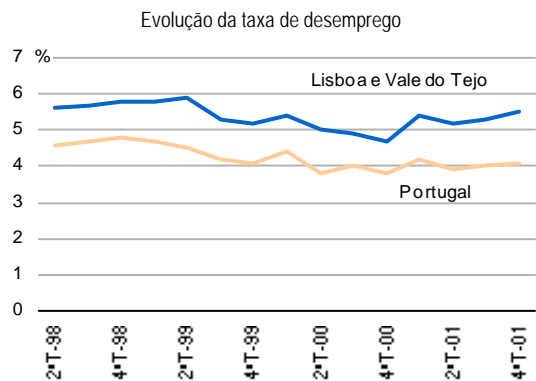


Figura 9 – Eixo de valores com identificações à esquerda

Esta imprecisão na leitura pode ser atenuada posicionando o eixo à direita junto dos dados mais recentes (ver Figura 8), duplicando o eixo (Figura 10), ou posicionando os valores junto das coordenadas respectivas (TUFTE, 1983).

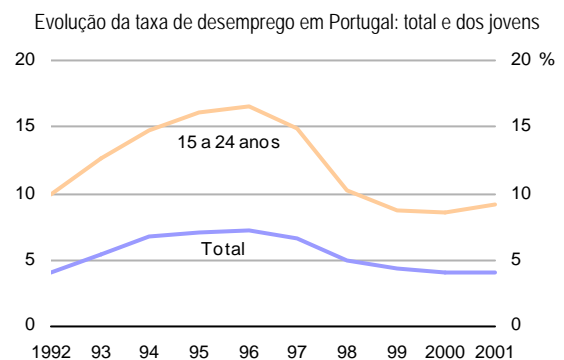


Figura 10 – Gráfico com duplicação do eixo

Os gráficos com dois eixos distintos são normalmente utilizados quando se têm diferentes unidades de medida (Figura 11) ou existem diferenças consideráveis de valores nas categorias de uma variável. Este tipo de gráficos deve ser evitado dado que é normalmente de difícil interpretação e, em muitos casos, bastante confuso (SCHMID, 1992).

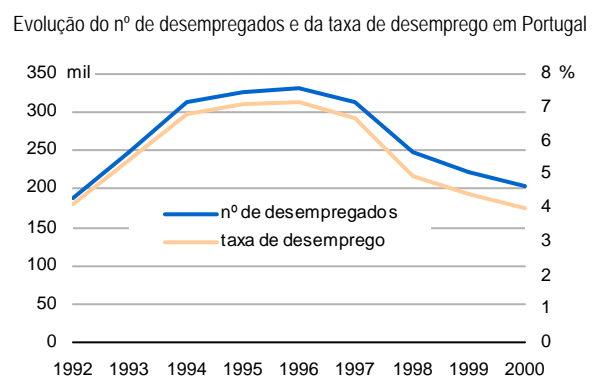


Figura 11 – Gráfico com dois eixos distintos

Quebra de escala

Por princípio, deve privilegiar-se a escala completa (com início em zero ou noutra valor de referência) em nome da honestidade na apresentação (Figura 12 - A). Contudo, essa quebra é admissível nos casos em que a informação apresenta pequenas variações, desde que acompanhada por uma simbologia perceptível ao leitor (Figura 12 - B).

Para melhor compreender os dados na fase da análise exploratória não existe qualquer problema em manipular as escalas e extrapolar eventuais variações, mas na fase da divulgação, deve existir algum cuidado para não evidenciar graficamente alterações nos dados que na verdade não ocorreram.

A quebra de escala é um exemplo de como se pode distorcer a mensagem transmitida. Quando o efeito nos dados é significativamente

diferente do efeito no gráfico, os valores aparecem visualmente sub ou sobre-avaliados (TUFTTE, 1983).

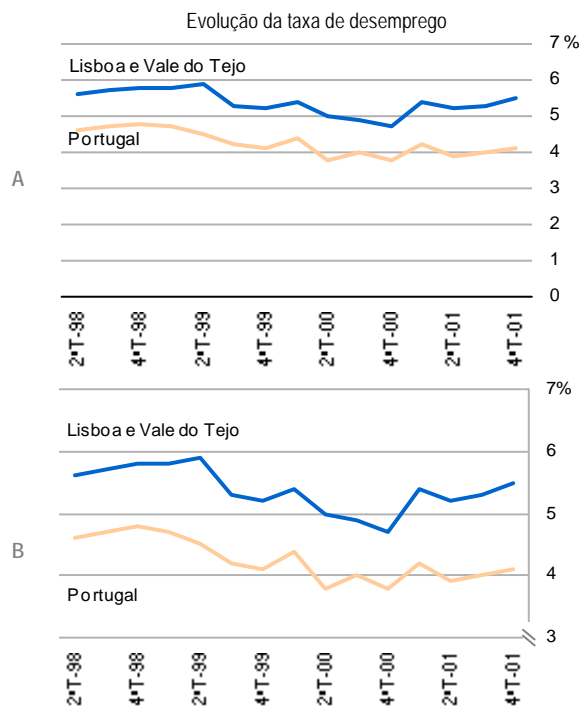


Figura 12 – Gráficos sem e com quebra de escala

Existem dois tipos de leitura possíveis num gráfico com mais de uma série temporal: a comparação vertical em que se confronta a dimensão relativa de uma série face a outra (ex: Portugal tem uma taxa de desemprego cerca de $\frac{3}{4}$ da de Lisboa e Vale do Tejo) e a comparação de declives em que é feita uma análise da evolução de ambas as séries.

No caso de se terem duas séries aparentemente constantes, a comparação entre elas apenas pode ser feita na vertical,

dado que dificilmente se detectam, visualmente, variações na sua evolução.

Neste caso, a utilização da quebra de escala permite detectar melhor as diferenças nos declives mas a comparação vertical entre as linhas deixa de fazer qualquer sentido (WALLGREN, 1996). É esta a razão pela qual não se devem fazer quebras de escala em gráficos de barras verticais, a comparação vertical entre as barras, após uma quebra de escala, não pode ser feita.

Variáveis visuais

Jacques BERTIN, em *Sémiologie graphique* (1973, 2ª ed.), foi o primeiro a sistematizar os conhecimentos sobre a aparência visual dos símbolos gráficos, criando uma tipologia com as seguintes variáveis visuais.

Localização – dada através das duas dimensões x,y do plano;

Tamanho – variação em comprimento, largura ou área, estando naturalmente ligado à importância numérica dos dados;

Valor – refere-se à variação (percebida) claro-escuro da cor ou à variação preto-branco;

Textura – tamanho e espaçamento dos elementos gráficos que constituem o símbolo (pontos, linhas ou outros), expresso pelo número desses elementos que se repetem por unidade de comprimento;

Cor – sensação pela qual se diferencia entre porções particulares do espectro electro-magnético, isto é, azul, verde, vermelho, etc.;

Orientação – também designada por direcção, corresponde ao ângulo com a linha de leitura;

Forma – pode ser geométrica (como quadrados ou círculos) ou então irregular.

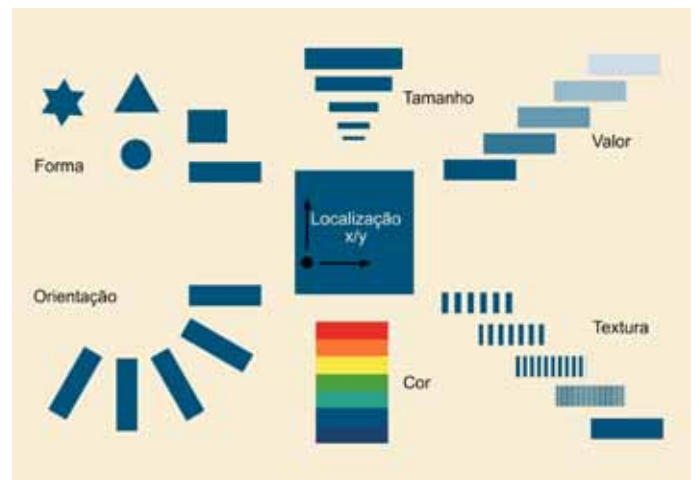


Figura 13 – As variáveis visuais segundo Bertin

1.2. Gráficos de barras

Os gráficos de barras são uma das formas mais populares de representar informação, em parte pela facilidade quer de execução, quer de leitura.

São para apresentar um conjunto de dados e também para comparar vários conjuntos de dados. Devem ser utilizados para representar variáveis discretas ou qualitativas, em termos absolutos ou relativos, ou para comparar categorias de variáveis quantitativas.

Podem, igualmente, representar a evolução de uma variável ao longo do tempo.

Neste tipo de gráficos, o leitor extrai os valores dos dados através da visualização da posição das barras relativamente a uma escala comum (CLEVELAND, MCGILL, 1984).

Normalmente, as barras começam no eixo das categorias, o que facilita a comparação das posições relativas.

Gráficos de barras simples (verticais ou horizontais)

Num gráfico de barras, as frequências podem ser indistintamente representadas no eixo das abcissas ou das ordenadas, ou seja, as barras podem ser horizontais ou verticais (Figura 14).

Apesar do gráfico de barras verticais ser o mais comum, existem situações em que é preferível optar pela outra disposição. O gráfico de barras horizontais é considerado de leitura mais fácil, quando é expressiva a diferença entre o valor mínimo e o valor máximo da variável. Num contexto de limitação do espaço disponível para posicionar o gráfico, é igualmente preferível optar pelo gráfico de barras horizontal, uma vez que

permite a inclusão de variadas categorias sem aumentar significativamente o espaço ocupado.

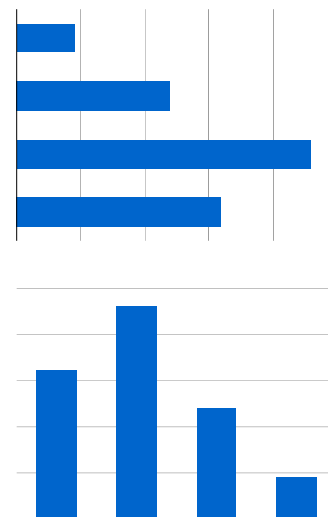


Figura 14 – Gráfico de barras horizontal e vertical

Aconselha-se o gráfico de barras horizontais para variáveis cujas categorias têm designações extensas, dado que nos gráficos de barras verticais o espaço para as designações é curto (Figura 15). Relembre-se que as designações não devem ser abreviadas, nem posicionar-se de forma a dificultar a leitura (verticalmente ou obliquamente) acabando, muitas vezes, por ocupar mais espaço do que o próprio gráfico.

Refira-se também que os gráficos de barras horizontais mostram, de forma mais clara, as diferenças entre os dados uma vez que possuem um eixo dos valores mais amplo. A Figura 15 é exemplo disso: apesar de ambos os gráficos ocuparem a mesma área, provocam efeitos visuais distintos quando se observam as categorias com maior frequência.

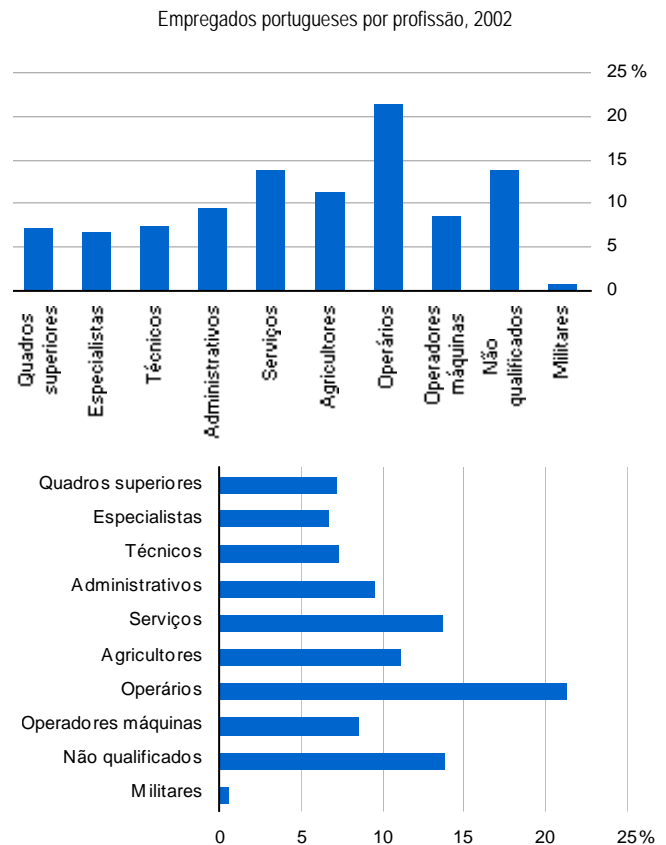


Figura 15 – Designações num gráfico de barras vertical e horizontal

Representação de valores negativos

A representação de valores negativos é desaconselhada em gráficos de barras horizontais, dado que, convencionalmente, aos valores negativos está associada uma barra numa posição descendente (Fig. 16).

De facto, a associação visual entre esquerda e direita e valores negativos e positivos, respectivamente, pode não ser directa para um leitor menos experiente. Por essa razão, devem ser utilizados gráficos de barras verticais quando existem valores negativos.

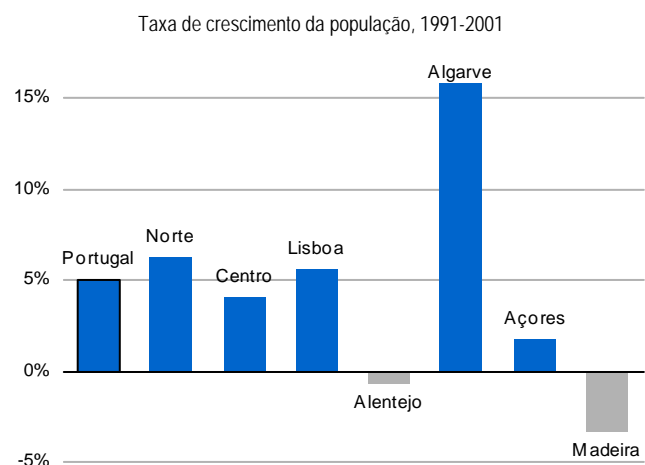


Figura 16 – Representação de valores negativos

Algumas regras relacionadas com a construção dos gráficos de barras

Escala no eixo dos valores

Nos gráficos de barras não é admissível a quebra de escala por deixar de ser possível efectuar comparações verticais entre categorias.

Uma quebra de escala é enganadora, porque mostra visualmente a existência de grandes variações nos dados que, de facto, não existem (Figura 17 A e B).

Olhando para a Figura 17 b, um leitor menos atento poderia dizer que em 1991 existiam cerca de um terço das pessoas de 2001, o que é falso.

No entanto, quando uma das barras assume um valor anormal e ocupa muito espaço na imagem, é admissível truncá-la. Tal terá que ser feito de forma clara e compreensível para o leitor, apresentando, por exemplo, o valor respectivo e também uma simbologia que permita compreender que a barra foi interrompida (Figura 18).

Pode ser indicado, em certos casos, fazer variar a escala entre 0 e 100 % (Figura 19) para que o leitor possa perceber quanto é que falta em cada barra para atingir os 100%.

Sempre que for possível, é aconselhável comparar as categorias com o total - neste caso Portugal – enriquecendo, desta forma, a leitura do gráfico (Figura 19).

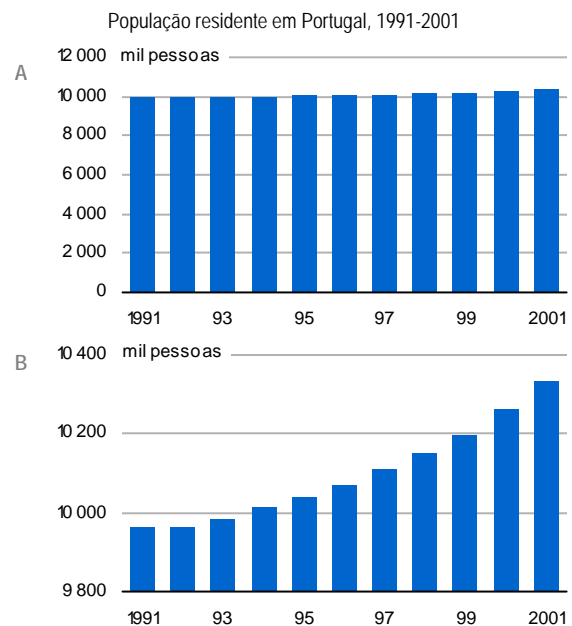


Figura 17 – Gráfico sem quebra de escala e erradamente com quebra de escala

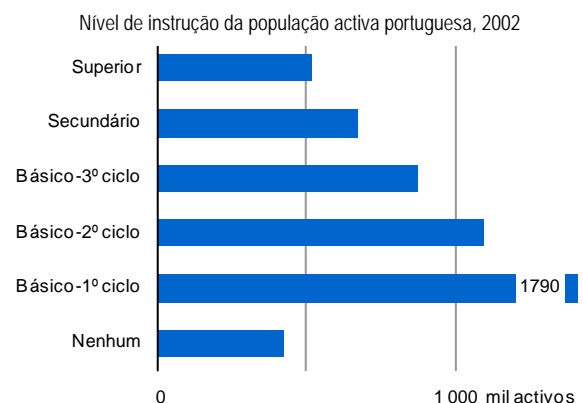


Figura 18 – Gráfico com barra truncada

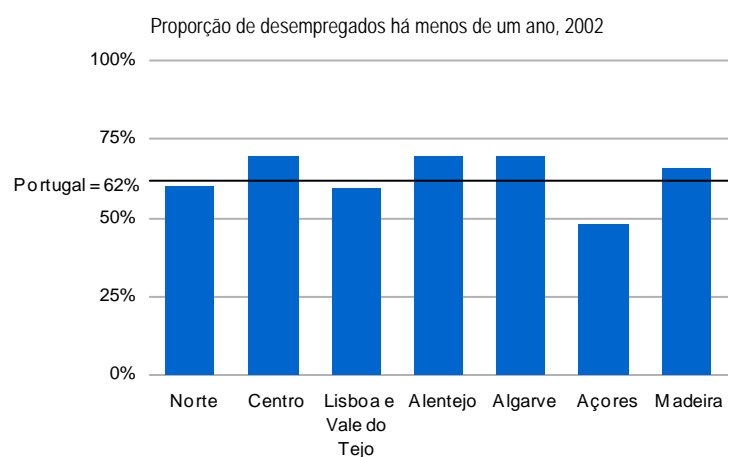


Figura 19 - Gráfico com escala entre 0 e 100%

Equilíbrio visual: espaços entre as barras e linhas auxiliares

Os espaços entre as barras devem estar construídos de forma a que não se dificulte a comparação (Figura 20 - B) nem se assemelhe a um histograma (C), sugerindo uma continuidade quando, afinal, a variável representada é discreta. É aconselhado um espaço entre as barras aproximadamente igual ao tamanho das mesmas (A).

As linhas auxiliares existem para ajudar o sistema visual a fazer comparações e ler valores aproximados. Um gráfico com demasiadas linhas auxiliares (B) dá mais peso visual do que deve a estes elementos secundários, sem que daí advenham vantagens significativas ao nível da leitura de valores aproximados. Por outro lado, um gráfico com poucas linhas auxiliares não traz grande valor acrescentado à leitura (C) (WALLGREN, 1996).

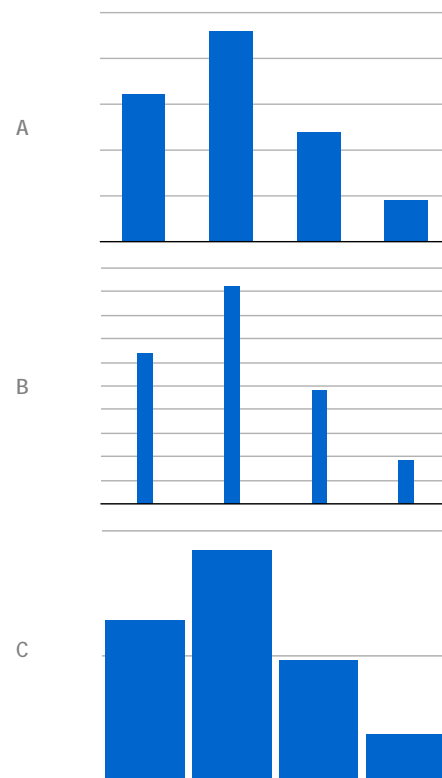


Figura 20 – Espaçamento de barras e linhas auxiliares

Ordenação

Na representação da informação, por vezes, é importante organizar as categorias por ordem crescente ou decrescente (Figura 21) para melhor compreender certos fenómenos implícitos.

É igualmente comum ordenar alfabeticamente (ou geograficamente) as designações das categorias, nomeadamente nos casos em que se representam países ou outro tipo de unidades administrativas, mas tal nem sempre é a melhor opção.

Se o mesmo conjunto de categorias é apresentado em mais do que um gráfico, então a posição relativa de cada categoria deve manter-se, ou seja, as categorias devem

aparecer na mesma ordem em todos os gráficos. Da mesma forma, o tamanho e a escala dos gráficos deve ser o mesmo, se o objectivo for a comparação entre eles.

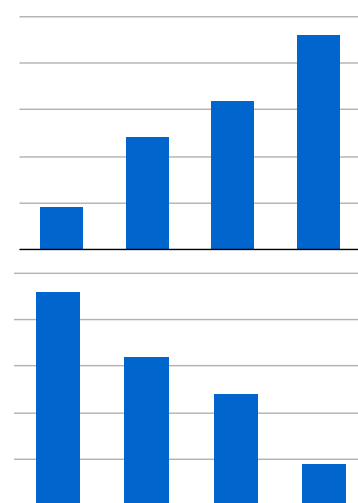


Figura 21 – Gráfico de barras por ordem crescente ou decrescente

Quando as categorias não são todas discriminadas, existindo, por exemplo, uma que reúne as restantes categorias sob a designação de 'Outros', é aconselhável não a incluir na ordenação e reservar-lhe o último lugar (WALLGREN, 1996; SCHMID, 1992) (Figura 22). Caso se utilizem cores para diferenciar as categorias, a categoria 'Outros', por ser a menos importante, deve ter uma cor que não se destaque (ex: cinzento).

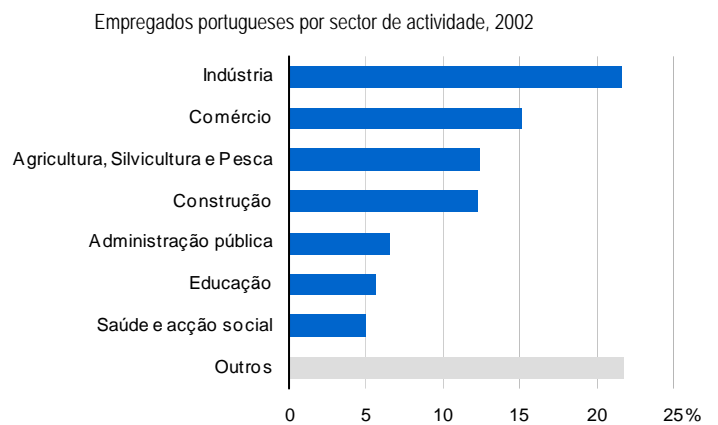


Figura 22 – Ordenação das categorias

Gráficos de barras agrupadas

Os gráficos de barras agrupadas são utilizados para descrever, simultaneamente, duas ou mais categorias, para uma dada variável discreta, ou quando se pretende realçar o valor das categorias em detrimento do valor total das variáveis (WALLGREN, 1996).

As diferentes categorias são representadas por barras sendo a distinção entre elas feita recorrendo às variáveis visuais (cor ou valor). Os grupos de entidades devem estar separados por um espaço em branco, mas não deve existir qualquer espaço entre as categorias de cada grupo.

Dado que a comparação entre barras adjacentes ao nível da estimação de valores é mais eficaz, em termos perceptivos, do que entre barras mais afastadas, o agrupamento escolhido deve estar de acordo com as categorias a que se pretende dar ênfase. Assim, em termos visuais são comparadas primeiro as categorias que constam da legenda e só depois são relacionadas as desagregações da variável (Figura 23 - A e B).

As barras podem apresentar indiferentemente valores relativos ou absolutos, consoante o tipo de análise, sendo por vezes de extremo interesse projectar ambos quando existem diferenças significativas (Figura 23 – B e C).

Este processo é tanto menos legível quanto maior for o número de categorias

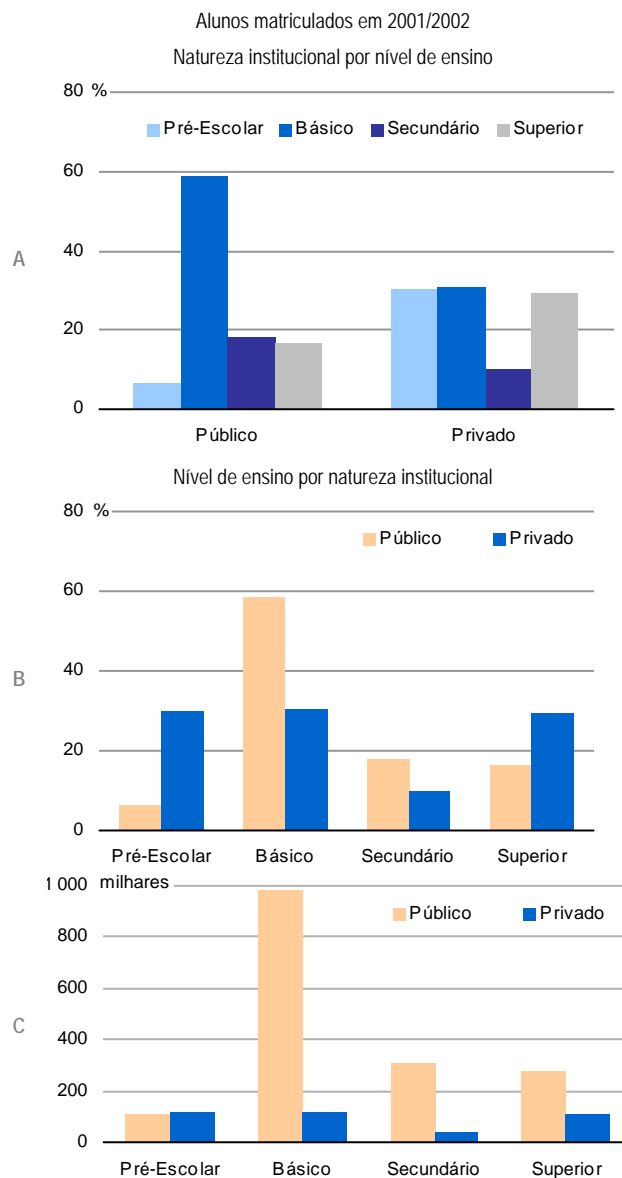


Figura 23 – Gráfico de barras agrupadas em quatro e duas categorias, em valores relativos e absolutos

representadas, sendo aconselhável não incluir mais do que três/quatro categorias, por variável, num gráfico.

Nos casos em que existem diversos grupos compostos por variadas categorias, é preferível construir-se diferentes gráficos em vez de acumular a informação num só.

Sobreposição em gráficos de barras agrupadas

Nos gráficos agrupados, as barras, que representam as categorias de cada grupo, podem tocar-se ou mesmo sobrepor-se (SCHMID, 1992). A sobreposição permite ordenar as categorias para além de poupar espaço e incluir mais informação, Note-se que as barras que se localizam num plano mais distante (e com uma cor menos forte) são percebidas como sendo menos importantes (Figura 24).

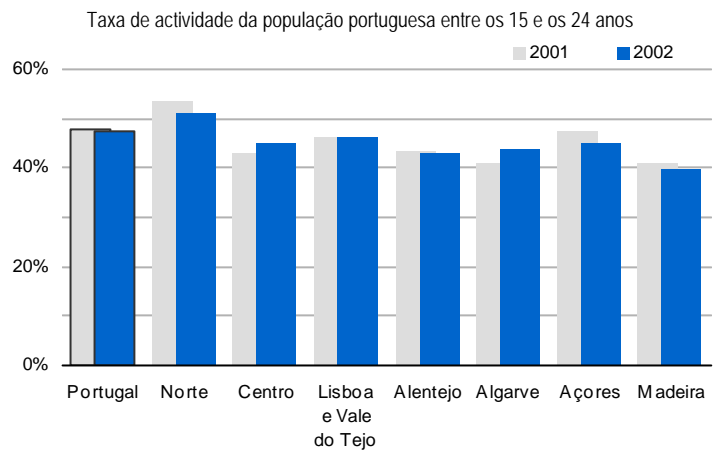


Figura 24 - Gráfico de barras agrupadas parcialmente sobrepostas

É, igualmente, proposta a sobreposição de barras nos casos em que os valores são sistematicamente menores numa categoria do que na outra (Figura 25). Realçar valores ou acontecimentos é também uma forma de análise dos dados. Por vezes, é importante dar ênfase visual a um determinado valor ou a uma determinada categoria.

Neste caso, e a título de exemplo, tornou-se mais grossa a linha auxiliar referente a 50% dos empregados - a única que tem um valor numérico associado - e deixou-se a leitura dos restantes valores para as linhas auxiliares não numeradas (Figura 25).

Para realçar a categoria referente a Portugal, pode-se utilizar uma moldura (A) ou uma cor mais escura (B). Apenas se apresentam os valores das categorias que se considerem dignas de análise (A - os Açores apresentam a maior diferença entre sexos) em vez de carregar demasiado o gráfico (B).

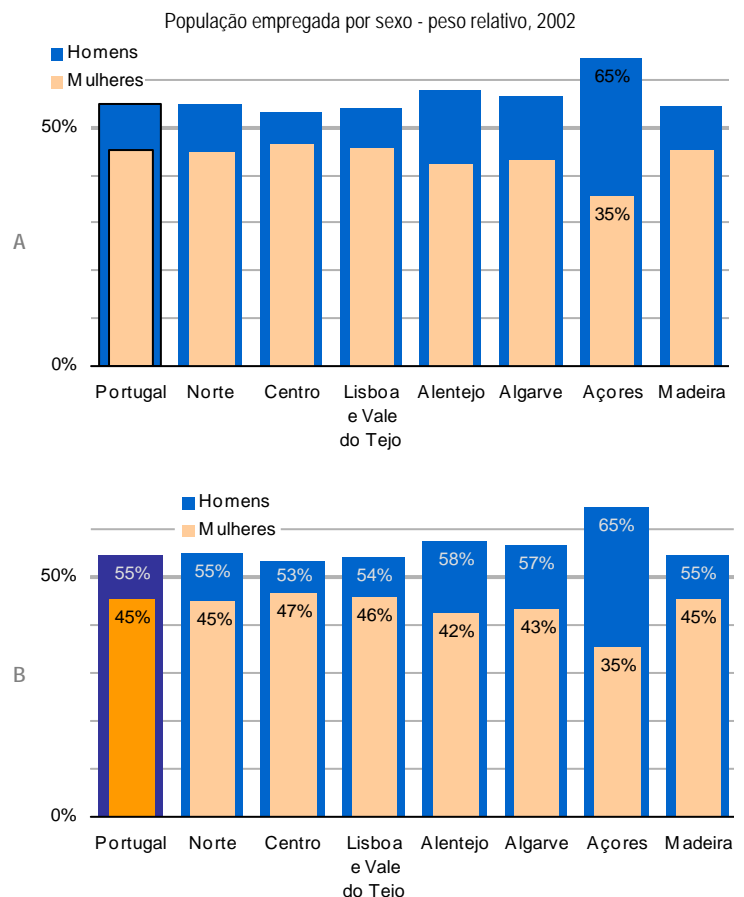


Figura 25 – Gráficos de barras agrupadas totalmente sobrepostas

Gráficos de barras empilhadas

Recorre-se aos gráficos de barras empilhadas (Figura 26) em situações análogas aos gráficos de barras agrupadas, ou seja, quando o conjunto de dados contém duas ou mais categorias.

Neste tipo de gráficos, cada barra subdivide-se em pelo menos duas categorias, com distintas cores ou padrões, permitindo mostrar a relação entre cada categoria (Homens/Mulheres) e o respectivo subtotal (ex: Comércio e Administração). As categorias surgem assim posicionadas umas sobre as outras, se for um gráfico de barras vertical (ou lado a lado, se o gráfico for horizontal), sendo que a altura (ou a largura) de cada componente corresponde ao valor absoluto ou relativo da categoria.

O gráfico em valor absoluto (A) adequa-se aos casos em que se pretende evidenciar mais o valor total das variáveis do que das respectivas categorias (WALLGREN, 1996), dado que o todo é apreendido com maior precisão do que as partes. Tal precisão advém de, para o total, ser comparada a posição relativa numa mesma escala, enquanto que na estimação dos valores das categorias são confrontados e ordenados os tamanhos respectivos.

Se o maior objectivo destes gráficos é indicar graficamente a soma total, mais do que estimar visualmente as respectivas categorias, valerá então a pena questionar porque não se opta por representar apenas o total ou então substituir esta por outra forma de representação.

No gráfico em valor relativo (B) apenas se pode estimar o valor das categorias observando o tamanho das barras que lhes correspondem.

Alunos matriculados no ensino superior por área de estudo segundo o sexo, 2001/02

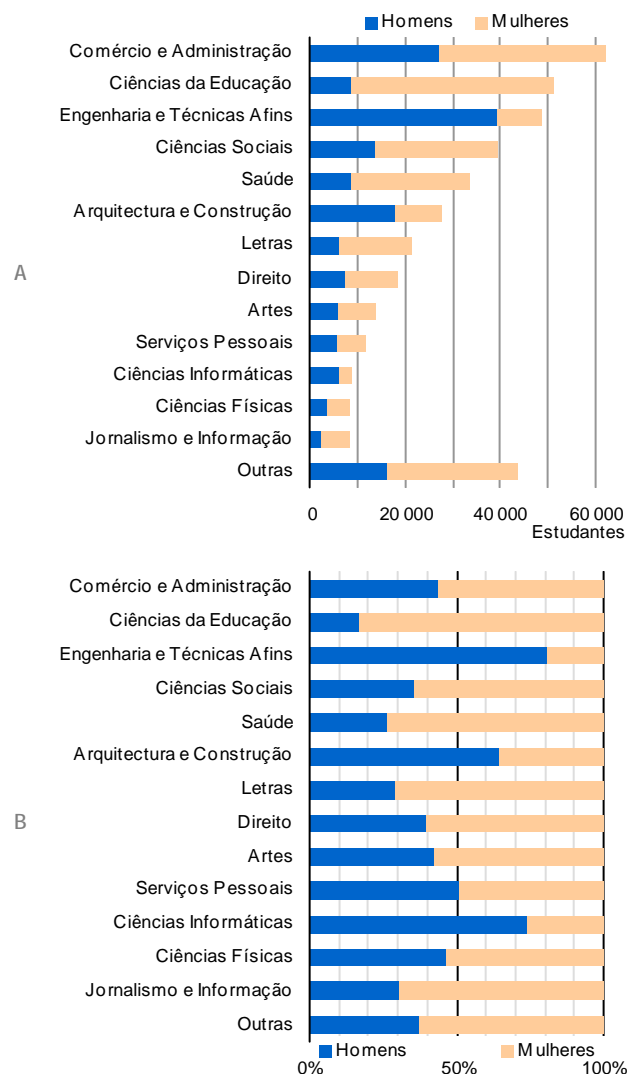


Figura 26 – Gráfico de barras empilhadas horizontalmente em valores absolutos e relativos

Com duas categorias torna-se mais fácil estimar os valores, dado que a base e o topo da escala servem de ponto de referência, mas com mais de duas categorias a leitura é consideravelmente mais difícil.

Desvantagem dos gráficos de barras empilhadas

De facto, as primeiras componentes são facilmente comparáveis por começarem junto ao eixo, mas nas seguintes apenas se consegue inferir aproximadamente os valores, sendo tanto mais difícil quanto maior for a variação da primeira categoria (Figura 27).

Por conseguinte, as flutuações e o peso excessivo da primeira categoria podem comprometer a leitura das restantes variáveis representadas.

Se a comparação entre categorias com base no tamanho pode envolver erros, não negligenciáveis, entre os verdadeiros valores e os estimados visualmente, a ordenação entre as categorias de uma mesma barra pode até ser incorrectamente realizada, pondo em causa a validade desta forma de apresentação de informação (CLEVELAND, MCGILL, 1984a).

É por esta razão que os gráficos de barras empilhadas devem ser limitados a um conjunto restrito de variáveis e categorias. Em certos casos é preferível substituir por um gráfico de barras agrupadas, porque melhora a estimação dos valores individuais, apesar de não facilitar a comparação entre categorias.

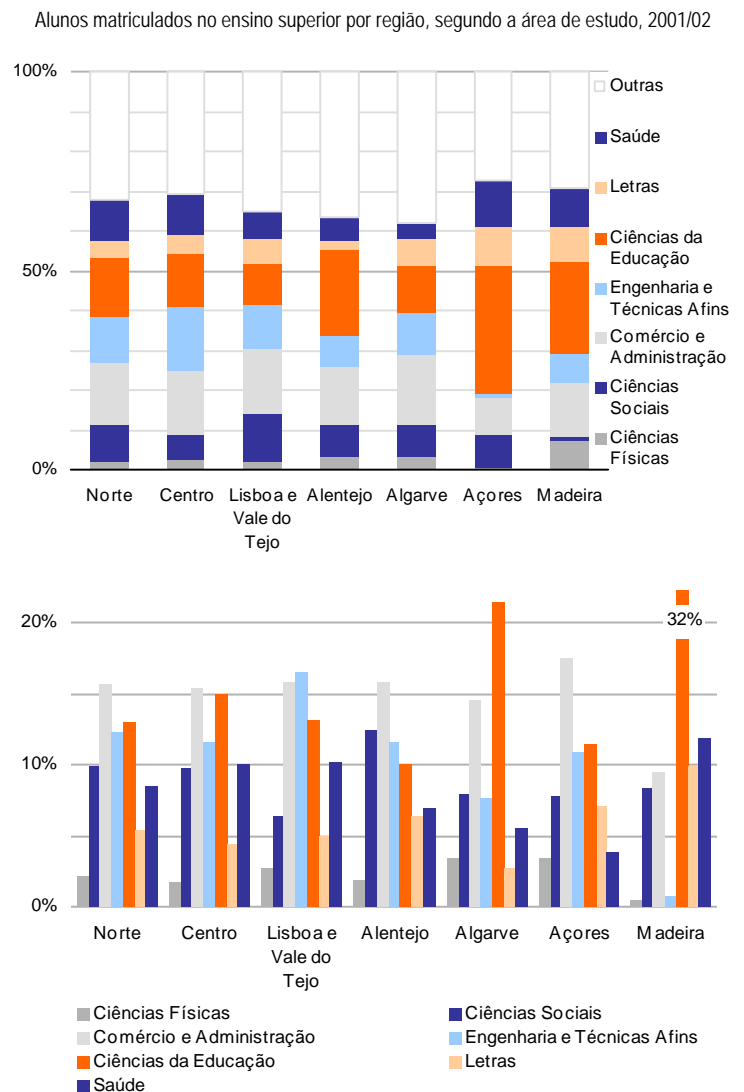


Figura 27 – Gráfico de barras empilhadas verticalmente e gráfico de barras agrupadas

Histograma

Um histograma mostra a distribuição de valores de uma variável contínua através de um gráfico de barras unidas. Contudo, se uma variável discreta apresentar muitos valores distintos, também pode ser representada por um histograma. Normalmente, os histogramas são representados por barras com bases

iguais em que a altura (ou o comprimento) varia em função da frequência relativa ou absoluta. De facto, no caso em que os intervalos têm a mesma amplitude, a área depende apenas da altura. Mas, quando as classes têm diferentes dimensões, a área de cada barra já não é proporcional à altura,

devendo ser calculada a altura por forma que a área de cada rectângulo seja proporcional à frequência relativa de cada classe. Enquanto no primeiro caso o eixo dos valores transmite a informação alusiva à frequência relativa de cada classe, no segundo caso este eixo não tem qualquer significado sendo o leitor obrigado a comparar áreas para interpretar a informação, o que se revela bastante mais difícil.

Esta forma gráfica permite indicar valores extremos e enviesamentos, demonstrando visualmente se a variável segue uma distribuição normal.

A representação das percentagens permite também comparar conjuntos de dados de diferentes dimensões.

Proporção da população feminina no total da população portuguesa, 2001

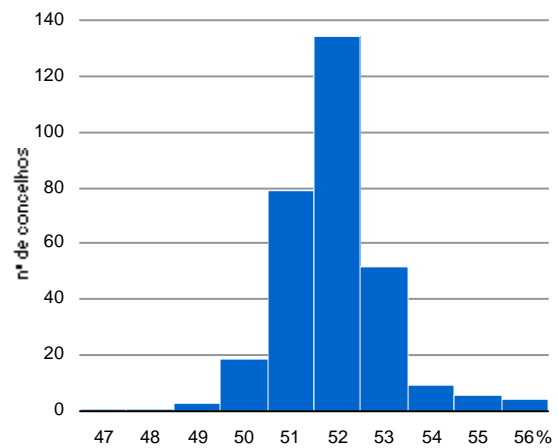


Figura 28 – Histograma

Pirâmide Etária

A pirâmide etária é também um histograma e é muito utilizada em análises demográficas por permitir visualizar numa única imagem a distribuição da população por idades e simultaneamente compará-la entre os dois sexos. A sua representação é feita em dois eixos horizontais (um para os efectivos masculinos e outro para os femininos) podendo esta ser em valores absolutos ou relativos.

As idades encontram-se representadas no eixo vertical, servindo de legenda a ambos os gráficos e são normalmente apresentadas em grupos etários de cinco anos, mas também podem ser representadas ano a ano.

A representação em valores absolutos fornece a dimensão dos dados mas impede qualquer tipo de comparação no espaço ou no tempo, que apenas é possível se os dados forem apresentados em termos relativos

(NAZARETH, 1996; INE, DRLVT, 2001). No entanto, esta forma de apresentação pode ser aplicada a outro tipo de informação demográfica (como, por exemplo, o nível de instrução) ou até para representar variáveis contínuas com uma legenda comum (WALLGREN, 1996).

População portuguesa por sexo e grupo etário, 2001

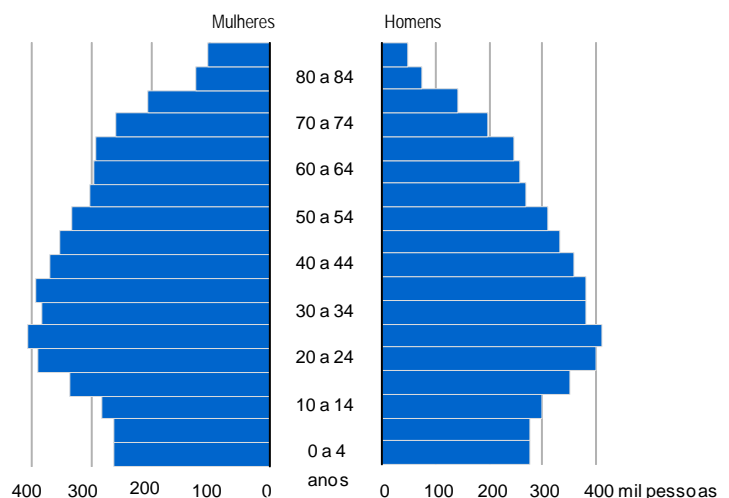


Figura 29 – Pirâmide etária

Séries temporais em Gráficos de barras

Um gráfico de barras verticais pode ter datas no eixo das categorias, possibilitando a representação de evoluções ao longo do tempo.

Os gráficos de barras podem substituir os gráficos de séries temporais nos casos em que a série de dados é muito curta. São igualmente indicados quando se pretendem fazer comparações verticais de determinadas variáveis num período específico, ou seja, quando se dá importância ao valor da variável em cada período e se pretende sobretudo relacionar quantidades individuais.

Para uma única série de dados, ambas as possibilidades (barras e linhas) são adequadas para mostrar tendências, mas para mais de uma série de dados, os gráficos de linhas são claramente preferíveis (JACOBS, 1997). Por isso, não é aconselhável utilizar os gráficos de barras para representar várias séries de dados. Quando as variáveis assumem valores sistematicamente inferiores ainda é possível acompanhar a sua evolução (Figura 30) mas quando as variáveis se entrecruzam o gráfico torna-se ilegível (Figura 31).

Nos casos em que a informação contida no gráfico é tal que impede uma correcta visualização, deve ser considerada a sua substituição por uma tabela de dados, ou então, a partição em vários gráficos.

Evolução dos alunos matriculados no ensino superior português, por tipo de ensino

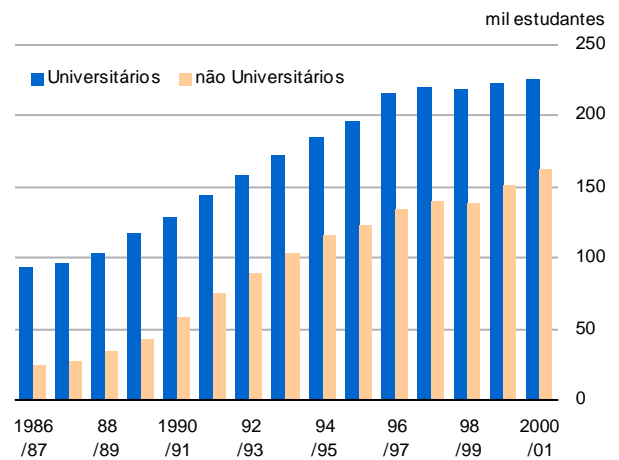


Figura 30– Gráfico de barras com duas séries temporais

Evolução da população portuguesa por grupos etários, 1991-2001

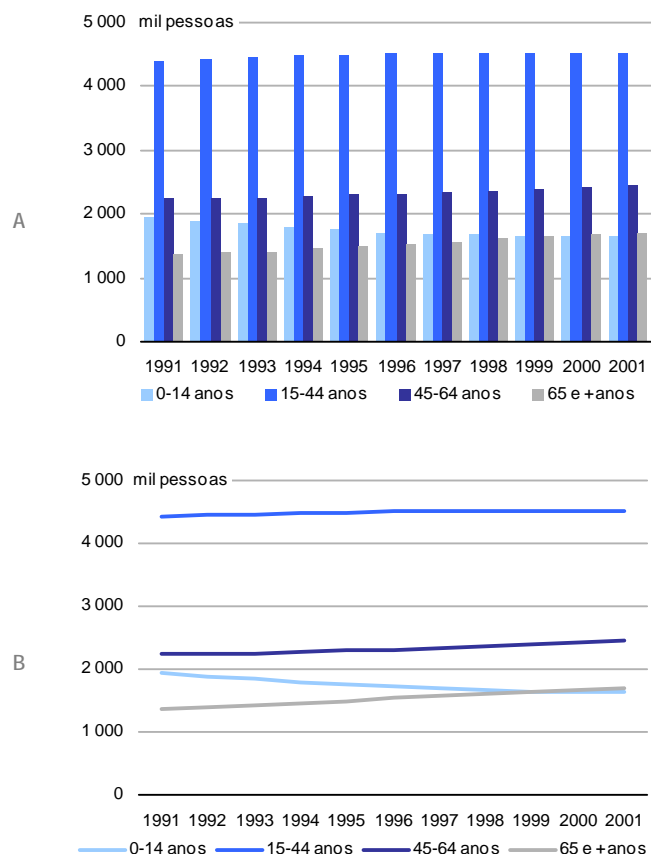


Figura 31 – Gráfico de séries temporais: barras e linhas

1.3. Gráficos de linhas

O gráfico de linhas é indicado para mostrar tendências e evoluções de uma variável contínua por outra variável contínua.

O mais comum é aquele que representa séries temporais (ou cronológicas), em que uma determinada variável contínua é analisada ao longo do tempo. O eixo do y mede a(s) variável(eis) em estudo, enquanto o eixo do x apresenta as unidades temporais dispostas cronologicamente em intervalos iguais de tempo, começando à esquerda com a data mais antiga (Figura 32).

Num gráfico de linhas, ao contrário dos gráficos de barras, as séries podem ser longas. O objectivo nestes gráficos é comparar os declives das curvas por forma a responder a perguntas do tipo: Em que períodos a variação foi significativa? Quando foram os pontos de inflexão? (WALLGREN, 1996).

Visualmente, para um determinado conjunto de dados, a união dos pontos (pares de coordenadas: x,y), é feita através de uma linha que sugere a continuidade.

Não devem ser incluídas mais do que três linhas por gráfico, caso contrário tornam o gráfico de difícil leitura (SCHMID, 1992; TUFTE, 1983). Quando muitas linhas se sobrepõem (Figura 33), é preferível substituir o gráfico de linhas por vários gráficos.

Evolução dos alunos matriculados no ensino superior português, por sexo

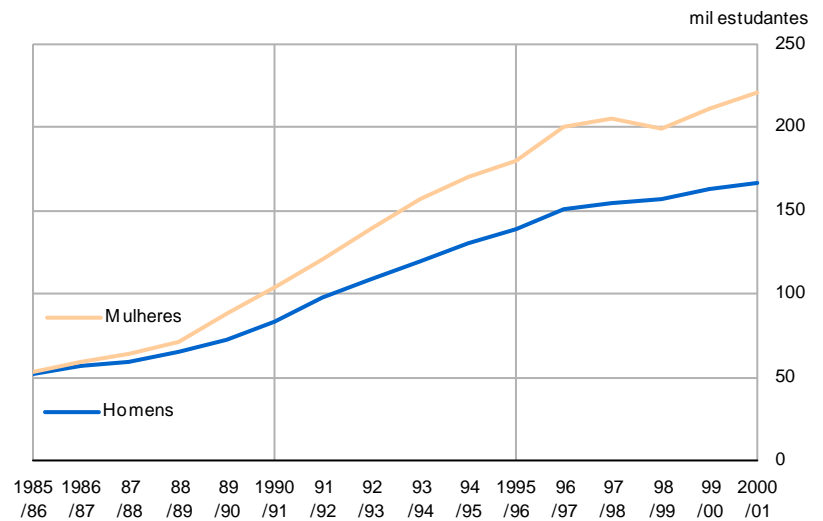


Figura 32 – Gráfico de séries temporais

Deve ser usado um estilo de linha diferente para cada gráfico, recorrendo à cor, forma, tamanho ou valor. Mesmo se as linhas se diferenciarem pela cor, pode ser necessário distinguir as linhas de outra forma, para facilitar a interpretação nos casos de impressão a preto e branco ou de reprodução através de fotocópias. Porém, tal opção pode dar uma ordem visual às linhas, não coincidente com a realidade, dado que, por exemplo, uma linha a tracejado é visualmente menos importante que uma linha a cheio.

Evolução dos alunos matriculados em Portugal, por tipo de ensino

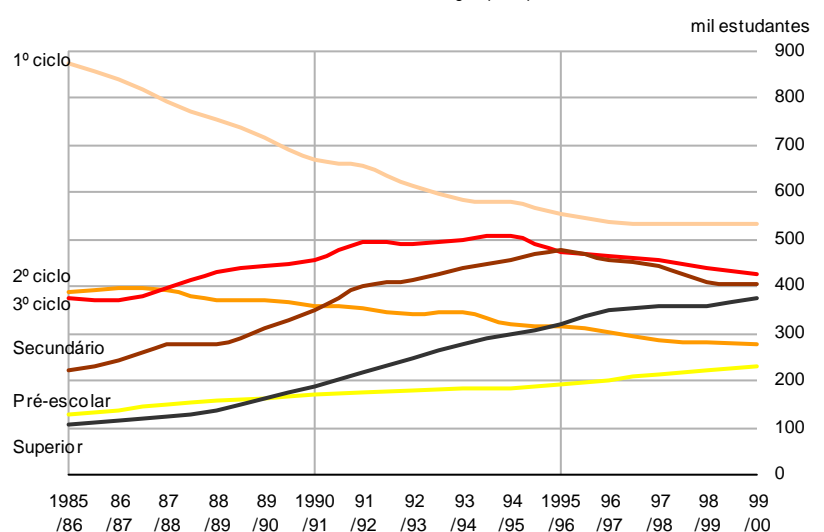


Figura 33 – Gráfico com demasiadas linhas

A variável medida no eixo das categorias nos gráficos de linhas não pode ser qualitativa (Figura 34). De facto, neste caso, a evolução da série não têm qualquer significado, ou seja, entre o Algarve e a Madeira não se pode afirmar que existe uma quebra na série de dados, mas apenas que os Açores têm um valor inferior. Também não é possível estimar os valores intermédios entre as categorias da variável, neste caso, não se pode dizer que existem x% de desempregados no Oceano Atlântico (gráfico correcto: Figura 19, pag. 20)

Os períodos devem estar igualmente espaçados se forem consecutivos e proporcionalmente espaçados se forem descontínuos, ou seja, quando ocorrem intervalos irregulares de tempo é indicado um ajustamento no espaçamento das colunas. Por exemplo, o espaço entre dados de 1998 e 2000 deve ser o dobro do que entre 2000 e 2001 (Figura 35).

Quando se pretendem comparar duas curvas que apresentam comportamentos muito semelhantes (Figura 36 - A), é preferível projectar a diferença entre elas, neste caso entre homens e mulheres (Figura 36 - B) em vez das curvas propriamente ditas.

Uma modificação repentina nos dados pode ser encoberta se o gráfico começar depois dessa modificação, mostrando uma estabilidade incorrecta (WAINER, 1984). Pelo contrário, uma alteração pode tornar-se brusca se o gráfico apenas representar aquele período e não o contextualizar, como, por exemplo, em séries com uma sazonalidade forte.

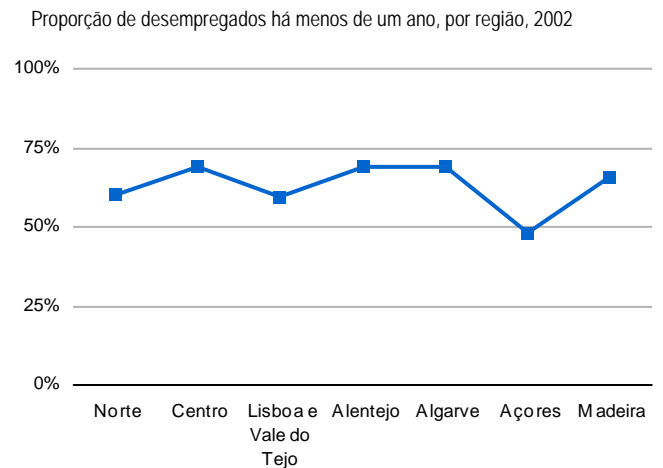


Figura 34 – Gráfico de linhas incorrecto

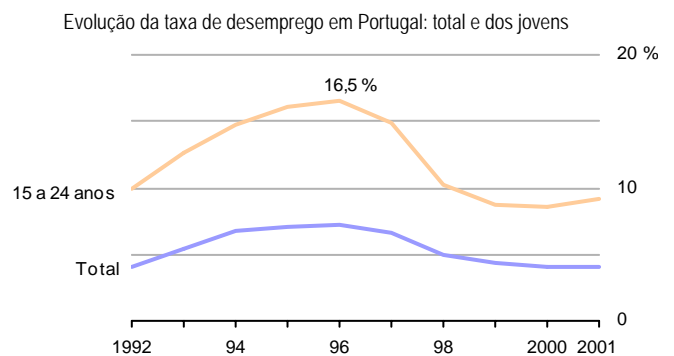


Figura 35 – Espaço entre os valores no eixo das categorias

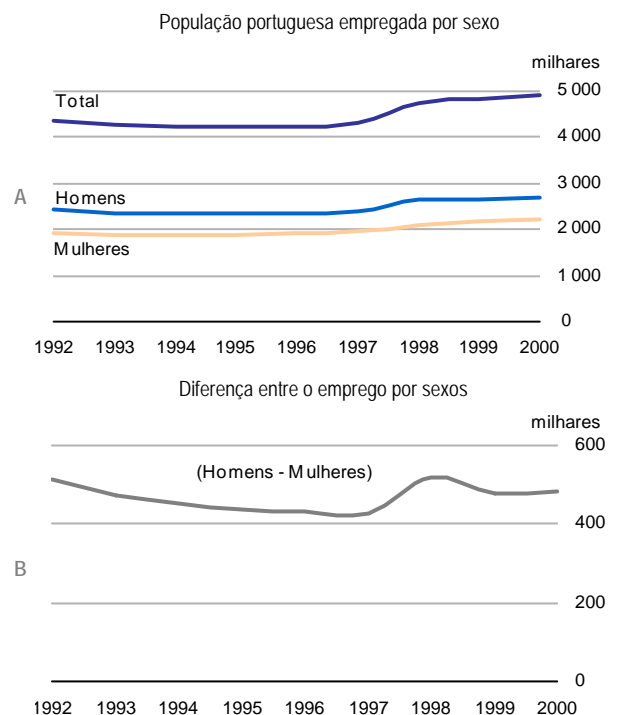


Figura 36 – Comparação de séries paralelas

Gráficos de área

Recorre-se aos gráficos de área quando se pretende visualizar simultaneamente a evolução do total e das respectivas componentes. Tal como nos gráficos de barras empilhados, existem poucas vantagens nesta forma de apresentação dado não ser possível responder de forma imediata a perguntas sobre o crescimento ou decréscimo ao longo do tempo, sobretudo quando a primeira das componentes apresenta oscilações significativas.

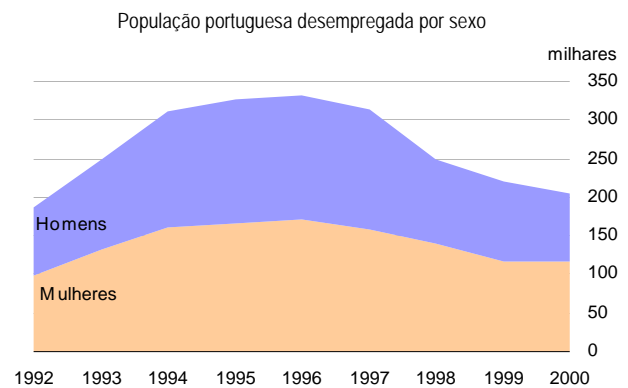


Figura 37 – Gráfico de área empilhada

Os gráficos de área são utilizados como alternativa aos gráficos de linhas. No entanto, trazem dificuldades acrescidas quando as áreas se intersectam porque deixa de ser possível seguir a evolução das componentes.

1.4. Gráficos circulares

O gráfico circular tornou-se muito comum em publicações direccionadas para um público alargado, mas tem vindo a ser amplamente contestada pela sua falta de capacidade informativa (WAINER, 1990; TUFTE, 1983; BERTIN, 1977, etc.).

Os gráficos circulares exibem as partes do todo como se fatias de um bolo se tratassem; a isso se deve a denominação inglesa 'pie chart' traduzida em português para queijo ou tarte. Para um determinado período temporal, a

variável em análise é projectada num círculo correspondendo a cada componente um ângulo, por forma a que as componentes no seu conjunto perfaçam os 360° (Figura 38).

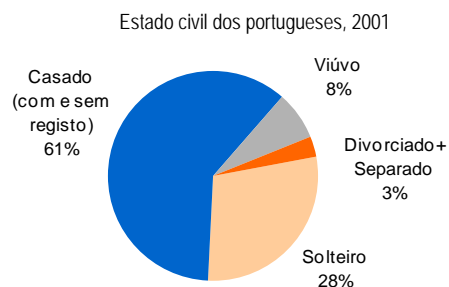


Figura 38 – Gráfico circular

A sua utilização é desaconselhada quando se pretende comparar mais do que um período temporal, para variáveis que contenham mais de cinco componentes ou quando as componentes têm aproximadamente o mesmo peso, sendo neste caso, preferível substituir o gráfico circular por um gráfico de barras (SCHMID, 1992). Muitas fatias ou fatias demasiadamente estreitas são dificilmente interpretáveis, sendo por isso necessário complementar o gráfico com os valores respectivos (Figura 38) ou associar um subconjunto de valores a outro gráfico circular de tamanho proporcional à quantidade que representa (Figura 39).

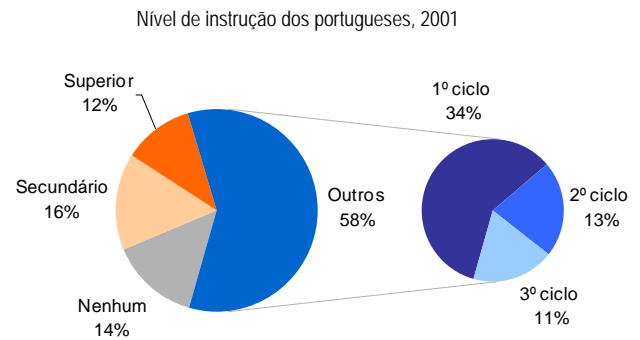


Figura 39 – Gráfico circular subdividido

Assim, a utilização dos gráficos circulares é apenas referida positivamente nos casos em que uma ou duas componentes dominam o total para dar uma ideia genérica dos dados, mas poder-se-á questionar se não será melhor recorrer a uma tabela.

É comum encontrar gráficos circulares distorcidos, ou seja, assumindo formas não circulares, para poupar espaço ou então por razões que a razão desconhece. Tornar uma figura circular numa elipse é altamente enganador, particularmente para os segmentos mais estreitos e deve ser evitado por desvirtuar completamente o gráfico original.

Outra prática corrente é a separação das fatias movendo-as radialmente para fora, provocando afastamentos desiguais entre fatias díspares. Como para manter as separações iguais é necessário posicionar as fatias de forma não circular, pelo que nenhuma das opções é formalmente correcta (BOUNFORD, 2000).

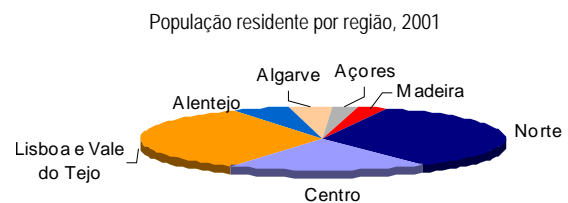


Figura 40 – Gráfico circular distorcido

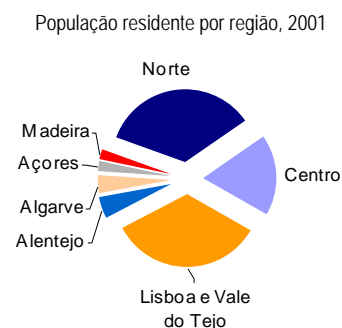


Figura 41 – Gráfico circular com fatias separadas

1.5. Pictogramas

Os pictogramas são gráficos comuns, mas com características decorativas. A sua utilização é indicada numa apresentação superficial em que o contacto com a imagem é breve, nomeadamente, em jornais ou revistas de âmbito alargado ou quando o público-alvo tem um nível educacional médio ou baixo.

Os pictogramas mais usuais são os baseados no critério do tamanho: em que a variação em área do tamanho das formas utilizadas é proporcional à variação da variável representada (Figura 42 - A).

No entanto, é vulgar encontrar imagens, particularmente nos média em que foi aumentada a altura e a largura simultaneamente, e não a área, tornando o desenho desproporcionado e transmitindo uma ideia completamente errada.

Os pictogramas constituídos por formas unitárias são também bastante utilizados. Neste caso, a cada elemento é atribuído um valor existindo, assim, tantos elementos quanto a dimensão da variável.

A pirâmide etária cujas barras são formadas por elementos que representam pessoas, é um dos mais difundidos. Um dos problemas surge com o tratamento dado às casas decimais. Modley (1952, in SCHMID, 1992) diz que as fracções de símbolos devem ser minimizadas, devendo-se, preferencialmente, arredondar os valores. De facto, é comum

Senão veja-se: na Figura 42 – B, Portugal tem 3 vezes mais estudantes do que Lisboa e Vale do Tejo, para ambos os sexos. Assim, a área do boneco referente a Portugal deve ser 3 vezes maior. Por isso, este tipo de apresentação é considerado como um dos mais enganadores (SCHMID, 1992; TUFTE, 1983).

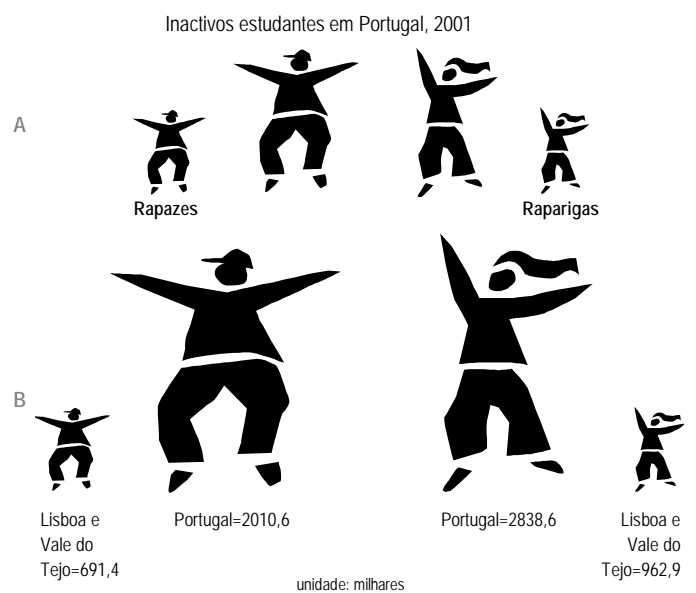


Figura 42 – Pictograma baseado no critério do tamanho

encontrar nas pirâmides etárias acima referidas, barras em que o último símbolo é fraccionado, ou seja, que terminam em braços, pernas ou cabeças (Figura 43).

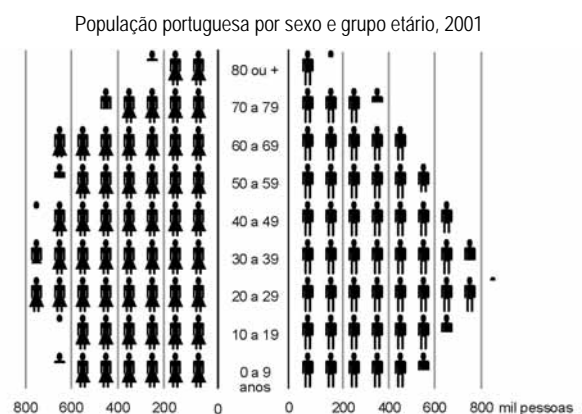


Figura 43 – Pictograma: pirâmide etária

1.6. Ver também ...

Neste dossiê são referidas, sucintamente, algumas das questões mais importantes associadas à representação gráfica, nomeadamente, as que se relacionam com a construção dos gráficos mais conhecidos e utilizados.

A informação utilizada para os gráficos aqui incluídos é bastante actual e pode ser encontrada em www.ine.pt. Todas as figuras, à excepção da última, foram construídas através do software *Excel*.

Este texto baseia-se na minha dissertação de mestrado intitulada: *Representação gráfica e cartográfica da informação estatística* e defendida, em Junho de 2003, no ISEGI/Universidade Nova de Lisboa.

Sobre os gráficos e a estatística existem diversos livros, artigos, *web sites*, dos quais se destacam os seguintes:

Publicações: 📖 : livros e 📄 : artigos em revistas

-
- 📄 BENIGER, James R.; ROBYN, Dorothy L. (1978), “Quantitative graphics in statistics: A brief history”, *The American Statistician*, 32 (1), p. 1-11.
 - 📖 BERTIN, Jacques (1973) 2.^a ed. (1^a ed. 1967) - *Sémiologie graphique*. Paris: Gauthier-Villars.
 - 📖 CHAMBERS, John C.; CLEVELAND, William S.; KLEINER, Beat; TUKEY, Paul A. (1998) 2^a ed. (1^a ed. 1983) - *Graphical methods for data analysis*. USA: Chapman & Hall.
 - 📄 CLEVELAND, William S.; MCGILL, Robert (1987a), “Graphical perception: The visual decoding of quantitative information on graphical displays of data”, *Journal of the Royal Statistical Society*, 150, p. 192-229.
 - 📄 CLEVELAND, William S.; MCGILL, Robert (1984a), “Graphical perception: Theory, Experimentation, and application to the development of graphical methods”, *Journal of the American Statistical Association*, 82, p. 419-423.
 - 📖 GRAPHICS GUIDELINES: *The theory and practice of presenting statistical data graphically, together with proposals for education of statisticians in appropriate use of graphics for presentation* (1994). COMMISSION OF THE EUROPEAN COMMUNITIES - EUROSTAT. Kent: White Waghorn Limited.
 - 📖 HUFF, Darrell (1991) 3^a ed. (1^a ed. 1954) - *How to lie with statistics*. England: Pinguin Books.
 - 📄 INE, DRLVT (2001), “As pirâmides de idades”, *Revista de Estudos Regionais* nº 2 (Conceitos e metodologias), *Instituto Nacional de Estatística*, p. 75-78.

- @ JACOBS, Bernhard (1997), "Experimental analysis of the graphical presentation of data in line graphs and bar charts in superposition and juxtaposition", <http://www.uni-saarland.de/philfak/MZ/graph/gesamtue.html>.
- 📖 NAZARETH, J. Manuel (1996) - *Introdução à demografia - Teoria e prática*. Lisboa: Editorial Presença.
- 📖 SCHMID, Calvin F. (1992) 2ª ed.; (1983, 1ª ed.) - *Statistical graphics - Design principles and practices*. Krieger.
- 📖 SILVA, Ana A. (2003) - *Representação gráfica e cartográfica da informação estatística*. Dissertação de mestrado defendida no Instituto Superior de Estatística e Gestão de Informação da Universidade Nova de Lisboa.
- 📖 TUFTE, Edward R. (1983) - *The visual display of quantitative information*. Cheshire-Connecticut: Graphic Press.
- 📖 TUKEY John W. (1977) - *Exploratory data analysis*. USA: Addison-Wesley.
- 📖 WAINER, Howard (1990), "Graphical Visions from William PLAYFAIR to John TUKEY", *Statistical Science*, 5 (3), p. 340-346.
- 📖 WAINER, Howard (1984), "How to display data badly", *The American Statistician*, 38 (2), p. 137-147.
- 📖 WALLGREN, Anders; WALLGREN, Britt; PERSSON, Rolf; JORNER, Ulf; HAALAND, Jan-Aage (1996) (English translation from Swedish "Statistikens Bilder - Att Skapa Diagram" Statistics Sweden 1995) - *Graphing statistics & data: Creating better charts*. California: SAGE Publications.

Páginas na Internet @

American statistical association - Section on Statistical Graphics:

<http://www.amstat-online.org/sections/graphics/>

Journal of computational and graphical statistics:

<http://www.amstat.org/publications/jcgs/>

Outros:

<http://www.edwardtufte.com/tufte/> (um dos melhores autores sobre esta temática – ver livros)

<http://www.mhhe.com/business/opsci/bstat/vistat.mhtml> (*visual statistics*)

<http://www.nas.nasa.gov/Groups/VisTech/visWeblets.html> (links sobre visualização científica)

<http://www.bell-labs.com/topic/societies/asagraphics/resources.html> (software, livros, revistas, etc.)

Gostaria de agradecer à Dra. Sandra Servinho e à Dra. Carolina Macedo, pelos seus comentários pertinentes.