Que número é este?

Um guia sobre estatísticas para jornalistas

PORDATA

FUNDAÇÃO FRANCISCO MANUEL DOS SANTOS

Os números têm de ser interrogados. Sem isso, são apenas uma sucessão inerte de algarismos e uma potencial fonte de equívocos. Partindo de tal premissa, este livro apresenta, em 22 tópicos, questões essenciais para o bom uso das estatísticas nas notícias, auxiliando na sua interpretação e comunicação. Um guia prático, de fácil leitura e que se espera útil para o dia-a-dia dos jornalistas.

Que número é este?

FUNDAÇÃO
FRANCISCO MANUEL DOS SANTOS

Largo Monterroio Mascarenhas, n.º 1, 7.º piso,
1099-081, Lisboa · Telf: 21 001 58 00
ffms@ffms.pt

Título: *Que número é este? Um guia sobre estatísticas para jornalistas*
Autores: Ricardo Garcia, Maria João Valente Rosa e Luísa Barbosa
Revisão do texto: Joana Vicente Pinto
Design e paginação: Guidesign
Impressão e acabamentos: Guide – Artes Gráficas, Lda.
ISBN: 978-98-8838-88-9
Depósito legal: 419 886/16

Os dados estatísticos apresentados foram consultados
a 2 de Dezembro de 2016.

Que número é este?

Um guia sobre estatísticas para jornalistas

Ricardo Garcia
Maria João Valente Rosa
Luísa Barbosa

ACADEMIA PORDATA

FUNDAÇÃO FRANCISCO MANUEL DOS SANTOS

PORDATA

ÍNDICE

A sua notícia baseia-se em dados estatísticos?

Então leia este livro.

Prólogo

Estatística e estatísticas não são apenas a mesma palavra no singular e no plural. A primeira é uma disciplina científica. A segunda é o resultado prático da sua aplicação.

Respeitando esta fronteira, o presente livro não se propõe ser um guia sobre Estatística, mas sim um guia de estatísticas. O seu objectivo é simplesmente oferecer aos jornalistas instrumentos que facilitem o seu trabalho de comunicação das estatísticas e, como corolário, da realidade que as mesmas representam.

Este projecto segue-se a outras iniciativas da Fundação Francisco Manuel dos Santos para o reforço da qualidade das notícias baseadas em dados estatísticos. Entre elas estão, no âmbito da Academia Pordata, os cursos de literacia em estatísticas para jornalistas, cujos conteúdos ajudaram a construir a espinha dorsal deste livro.

Este é um guia pensado para estar em cima da secretária ou no ambiente de trabalho do computador de qualquer jornalista. É um manual de consulta rápida, para quando surge uma dúvida, mas também se presta a uma leitura mais alongada. O livro está dividido de forma a proporcionar ambas as abordagens.

A sua introdução é para ser lida antes de qualquer outro capítulo. Em poucas páginas, ali está o essencial do pensamento que conforma o guia. É um texto que procura estimular a aplicação, aos números, do espírito crítico próprio do jornalismo no tratamento da informação. As estatísticas devem ser interrogadas, entrevistadas. É essa a mensagem deste capítulo.

Segue-se aquilo a que se poderia chamar "bagagem obrigatória" para quem vai fazer notícias a partir de dados estatísticos. Nessa mala imaginária estão cinco itens, na verdade, cinco competências sem as quais a viagem pelo mundo da comunicação dos números se torna mais difícil: calcular percentagens, fazer arredondamentos, utilizar regras de três simples, construir gráficos e escrever de forma acessível.

Os capítulos subsequentes inspiram-se na tradicional pergunta quíntupla a que uma notícia deve por norma responder: o quê, quando, como, onde e porquê? Com uma ligeira alteração – substituindo-se o "porquê" por "quanto" –, é ao longo destas questões que este guia caminha.

Cada um dos capítulos é independente do resto do livro, de tal modo que a sua leitura não necessita seguir uma ordem determinada. Neles, encontram-se definições, exemplos, alertas e explicações de como se fazem os cálculos.

Um índice remissivo no final permite, ainda, ao leitor ir directo ao assunto sempre que pretenda esclarecer uma dúvida.

Não se encontram aqui, obviamente, todos os conceitos que constituem a ampla galáxia das estatísticas. Mas apenas uma selecção entre os que estão mais directamente associados à informação estatística *per se*.

Com este guia prático, de fácil acesso e que se espera útil para o dia-a-dia dos jornalistas, a Fundação Francisco Manuel dos Santos, através da Pordata, junta-se, assim, a outras entidades europeias responsáveis pela divulgação ampla de dados estatísticos que já lançaram guias ou iniciativas semelhantes, tendo em vista a promoção de uma sociedade melhor informada, mais esclarecida e, por isso, mais livre.

Boa leitura e boas estatísticas!

Introdução

Lidar com estatísticas requer dos jornalistas aquilo que a própria profissão lhes exige todos os dias: que façam perguntas. Sem ser questionado, um número é apenas um número, uma sucessão inerte de algarismos.

E números é o que não falta. Só o Instituto Nacional de Estatística (INE) divulga anualmente milhares de actualizações de dados sobre os mais diversos temas, da população ao emprego, da economia à saúde, das condições de vida ao comércio, entre tantos outros. Somem-se ao INE outras entidades com responsabilidade na produção de estatísticas oficiais, governos, institutos públicos, universidades, empresas, organizações não-governamentais. E teremos uma longa cordilheira daquilo a que uma publicação da agência de estatísticas da Noruega qualifica como “vulcões numéricos”[1].

O poder mediático dos números é incomensurável. Se mostram que algo mudou, são notícia. Se causam surpresa, são notícia. Se são inéditos, são notícia. Não será exagero dizer que se é um número, é notícia.

Dominados por este inescapável magnetismo, nem sempre paramos para interrogar a facilidade com que se transformam

1 Estatísticas da Noruega, Paris 21 (2009). *Apresentação Amigável de Estatísticas*, p. 6

dados estatísticos em títulos noticiosos. A variação que revelam é mesmo importante? Por que razão surpreendem?

Um exemplo: se a esperança de vida à nascença, em Portugal, passou de 80,2 anos para 80,4 anos entre 2013 e 2014, esta alteração merecerá mesmo um grande destaque nas notícias? O aumento, que representa cerca de dois meses e meio, é rigorosamente o mesmo que se observou todos os anos desde 2010. Qual é a novidade então?

Questionar um dado estatístico vai muito além de indagar se vale a pena ou não fazer dele uma notícia. Os números que chegam às redacções em comunicados de imprensa, anúncios, relatórios ou entrevistas não revelam, logo à partida, tudo o que importa saber sobre eles. Sós, são números frios que dificilmente interessarão ao cidadão comum. As estatísticas precisam de explicação, de contexto, de vida.

Estar consciente desta necessidade é um pré-requisito para quem faz notícias a partir de dados estatísticos. O pequeno guia *Making Sense of Statistics*, publicado em 2010 no Reino Unido, deixa a seguinte questão: "Se uma estatística é a resposta, então qual era a pergunta?"[2]. Um relatório da BBC, publicado seis anos depois, e que curiosamente adoptou o mesmo título, recomenda que os seus jornalistas e editores façam mais para dar enquadramento às estatísticas, "de modo a que a audiência compreenda o seu significado"[3].

Antes da audiência das notícias, no entanto, quem deve em primeiro lugar compreender o significado dos dados estatísticos é o próprio jornalista. E como? Fazendo perguntas.

Cada um, com a sua própria sensibilidade, saberá o melhor caminho a seguir. Mas aqui sugere-se um roteiro simples de sete interrogações essenciais:

---

2 Hawkes, N., Sierra, L. et al. (2010). *Making Sense of Statistics*, p. 5

3 BBC Trust (2016). *BBC Trust Impartiality Review: Making Sense of Statistics*, p. 11

1. Como é possível?

Não, ninguém está aqui a incutir um espírito de desconfiança profunda em relação às estatísticas em geral, como se qualquer número fosse uma mentira. O que aqui se reforça é a necessidade de se aplicar aos dados estatísticos o mesmo espírito crítico que qualquer jornalista destina a todo o tipo de informação com que trabalha.

Algumas situações, é claro, mais facilmente despertarão este saudável cepticismo. Suponhamos que é divulgado um comunicado de imprensa a referir que o número de casos de uma determinada doença subiu 500% num ano. Uma percentagem de 500% é forte candidata a manchete em qualquer parte do mundo. Mas como será possível que tenha havido tamanha subida?

Mesmo em relação a dados que pareçam inquestionáveis – como as estimativas da população, por exemplo –, inquiri-los funciona como alimento à curiosidade, estimulando a busca do que está por detrás ou além de um simples número.

Este é o ponto de partida para se trabalhar com estatísticas. É a questão-chave que abre a porta a outras. É como dizer: "Vamos lá ver o que isto significa".

2. Que números são estes?

Pode parecer uma banalidade, mas sem compreender os números que temos à frente, é muito mais fácil deixarmo-nos apanhar pelo erro. Tal como a um entrevistado, este é o momento de se questionar os números de modo a conhecer o que são e o que os caracteriza. Por exemplo:

- Qual é a sua ordem de grandeza e escala – se estamos a falar em milhões, milhares de milhões, biliões, etc.;
- Qual é a sua unidade de medida, ou seja, o que está a ser contado – por exemplo, indivíduos, euros, empresas ou registos de actos administrativos;

- Que valor é aquele – se são números absolutos ou relativos; se são valores exactos ou arredondados; se se referem a uma proporção, a um rácio, a uma taxa, a um índice, se são apresentados em números-índice, se são médias, taxas de variação.

3. De onde vêm?

Trata-se aqui de identificar quem é o responsável pelos dados estatísticos e por que estão a ser divulgados naquele momento. Se a pergunta fosse dirigida aos números em si, seria algo do género: "Quem vos mandou aqui e porquê?".

A entidade responsável pelos dados é uma autoridade estatística? É uma empresa? Um organismo público? Uma organização não-governamental? Tem um interesse directo nas notícias que possam ser produzidas com aqueles números? Alguma vez divulgou estatísticas inconsistentes ou enganadoras?

Se for um estudo académico, os resultados foram publicados numa revista com arbitragem científica? Os investigadores têm trabalho reconhecido nas suas áreas? Quem financiou o estudo?

Por outro lado, convém perguntar por que razão uma determinada informação estatística surge numa determinada altura. Em inúmeras situações, não haverá nenhuma razão especial a não ser o calendário estatístico. É o caso das meras actualizações periódicas de séries de estatísticas oficiais, que confirmam ou não a continuidade de evoluções passadas.

Por vezes, porém, há dados que surgem por outras razões – a publicação de um estudo científico, por exemplo – ou sem uma motivação que se conheça mas que é preciso investigar.

A razão do momento em que os dados surgem não tem de ser, necessariamente, óbvia. Mas é preciso estar atento. Há casos-limite que, embora não sejam tão comuns, não são inéditos. Dados mediáticos podem muito bem ser apenas um veículo de propaganda de um produto, de uma empresa, de um sector de actividade ou de uma campanha. Um inquérito sobre o consumo de bens de higiene pessoal, por exemplo, pode coincidir, não por acaso, com o lançamento de um novo champô.

4. Como foram produzidos?

Não se deve negligenciar a forma como um dado estatístico foi recolhido ou produzido, porque a metodologia tem influência no resultado, bem como na sua credibilidade.

Muitas estatísticas são produzidas a partir de registos administrativos – como o número de nascimentos ou de óbitos, de casamentos ou de divórcios, de processos judiciais ou de constituição de novas empresas.

Outras estatísticas resultam de inquéritos. E os inquéritos dependem de amostras. E as amostras podem ou não ser representativas da população. E mesmo que se diga que são representativas, de facto serão? E que perguntas estão na origem dos resultados?

Há ainda dados que não são observados ou medidos de forma directa, mas antes provêm de cálculos, de agregações e de ponderações, como o PIB.

Naturalmente, há curvas e cantos obscuros no labirinto de alternativas de como se recolhem dados estatísticos. Não é razoável pensar que um jornalista perca horas a estudar a fundo os métodos e procedimentos que estão na base da produção de cada dado estatístico. Mas isto não deve servir de desculpa para não se indagar como se chegou àqueles resultados.

5. O que mostram e o que não mostram?

Quando se pergunta o que um dado ou uma série de dados estatísticos está a mostrar, há pelo menos duas dimensões de resposta.

A primeira é a que permite saber do que um número está a falar em concreto. Se os dados falam do desemprego, como é que a entidade que os produziu ou divulgou define a situação de desempregado? Se falam do número de casamentos em diferentes municípios, o que se considera na contabilização: onde o casal vai morar ou onde a união foi registada? Se são dados anuais, estamos a falar de uma média dos doze meses, da soma dos valores mensais ou de uma contagem num dia específico?

A segunda dimensão é a que se refere ao valor em si e ao seu significado. É uma simples contagem em valores absolutos ou uma variação? Se se tratar de uma variação, qual é a sua relevância? Um aumento de 500% no número de casos de uma doença pode significar tanto uma variação de uma para seis ocorrências como de 10.000 para 60.000 – duas situações obviamente muito distintas.

Picos em séries de dados estatísticos podem apenas constituir situações episódicas – cujas razões merecem ser explicadas –, e não tendências de fundo. E há casos de variações que são justificadas apenas por alterações sazonais ao longo do ano, como as que ocorrem com a taxa de inflação ou de desemprego. Ainda, nalgumas circunstâncias, bruscas diferenças no valor de um indicador de um ano para o outro podem não significar mais do que uma quebra de série, devido, por exemplo, a uma alteração da metodologia de recolha.

Estes são apenas exemplos que ilustram a necessidade de compreender melhor o significado do dado estatístico.

Procurar saber o que os dados mostram acaba por fornecer algumas respostas também sobre o que não mostram, conduzindo a outras interpretações. A iniciativa cabe ao jornalista. Antes de um dever, é uma oportunidade.

Imaginem, por exemplo, uma notícia a dizer que a população de Portugal reduziu 0,4% em 2015. Como informação, é pouco. Qual a razão desta queda? Houve mais mortes do que nascimentos? Ou houve mais pessoas a emigrar do que a imigrar? Ou ambos? Terá havido alguma alteração substancial nestes indicadores naquele ano? Por quê? Porventura, encontrar-se-á um ângulo, um título, uma história diferente, mais rica e mais interessante do que a simples apresentação de um dado estatístico geral.

Vale a pena ter em mente que as estatísticas seleccionadas para os títulos dos comunicados de imprensa resultam de escolhas que não esgotam todas as leituras que o mesmo conjunto de dados é capaz de permitir.

6. O que fazer com eles?

A abordagem jornalística mais simples às estatísticas é repetir esses dados tal como chegam à redacção. É a mais simples, mas a mais aborrecida. O resultado é, muitas vezes, uma notícia acrítica, descritiva, carregada de números e inspiradora de bocejos.

O seu reverso são peças analíticas, se possível enriquecidas com gráficos ou outras formas de visualização, em que os dados são avaliados em maior detalhe, aprofundados, complementados por outros e colocados em contexto. Por esta via, contam-se melhores histórias. Produzir notícias a partir de dados estatísticos é, sobretudo, isso.

Para atingir este objectivo, mais uma vez, é preciso ir além dos números. A palavra-chave aqui é investigar. Olhar para outros conjuntos de dados, procurar outras explicações, pedir a opinião de outros especialistas, não se contentar com interpretações prontas, evitar o caminho mais fácil.

Os números contidos nas estatísticas são apenas um ponto de partida. O ponto de chegada é definido por quem os interpreta.

7. Que erros evitar?

Num *ranking* de erros a evitar quando se lida com estatísticas, o primeiro lugar está reservado para a crença cega nos números. Não se deve dispensar nenhum dado estatístico, seja qual for a sua origem, de ser interrogado.

No topo da lista também está a produção de notícias sem pensar na sua audiência. Esta norma básica do jornalismo em geral é ainda mais pertinente no trato com estatísticas, sempre recheadas de números de difícil leitura, conceitos técnicos e jargão incompreensível. A simplicidade é o melhor caminho.

No caldo de complexidade do oceano estatístico, não é difícil cair-se em erros corriqueiros que facilmente comprometem uma notícia – um arredondamento equivocado, um cálculo inapropriado, um gráfico mal construído ou a clássica confusão entre percentagem e pontos percentuais, por exemplo.

E os erros ocorrem com relativa frequência. Um estudo que analisou cerca de 1100 notícias com informação matemática, publicadas em cinco jornais portugueses nos primeiros três meses de 2013, concluiu que aproximadamente uma em cada três continha erros[4]. Os mais frequentes referiam-se não a algo que tenha sido mal feito, mas a algo que deveria ter sido feito e não o foi. Por exemplo, não informar qual é a amostra em causa num inquérito, não referir a margem de erro, esquecer-se das escalas e unidades em gráficos e tabelas ou utilizar números sem dar pistas sobre a sua dimensão e contexto.

Conclusões apressadas também merecem um lugar no *ranking* dos erros a evitar. Uma brecha por onde elas aparecem está na facilidade com que se confunde correlação com causalidade.

4 Pereira. S. (2015). *A Matemática na Imprensa Portuguesa*, pp. 100-124

Um exemplo muito citado é o das cegonhas e dos bebés. Algumas análises identificaram uma coincidência entre o número de cegonhas por quilómetro quadrado e o número de bebés por quilómetro quadrado. Diz-se, assim, que há uma correlação entre cegonhas e bebés, pois o número de ambos varia mais ou menos da mesma forma. Mas isso não permite concluir que a causa para os bebés são as cegonhas ou vice-versa. Um terceiro factor é responsável por ambas variarem em sintonia: a urbanização. Este exemplo é risível, mas serve apenas para mostrar que atribuir de imediato uma relação de causalidade ao resultado de uma correlação forte é um erro.

Em síntese

Quase todos os jornalistas lidam com estatísticas no dia-a-dia. Mas muitos se sentirão desconfortáveis para trabalhar com mais detalhe a imensa variedade de dados divulgados periodicamente por inúmeras entidades, limitando-se a reproduzi-los.

Trabalhar com estatísticas não é algo reservado aos profissionais de comunicação social que tenham tido uma boa formação em matemática. Escrever sobre estatísticas, aliás, não é escrever sobre matemática. É contar histórias que falam sobre a sociedade em que vivemos.

Compreender o essencial das estatísticas é uma arma poderosa em vários sentidos. Alarga o leque de abordagens, permite novas leituras e ajuda a evitar erros, omissões e manipulações na divulgação de números.

Com perguntas essenciais e algum conhecimento de base, qualquer um navega com facilidade e maior segurança nesse universo.

ANTES DE MAIS

QUANTOS POR CENTO?

Percentagens

As percentagens estão de tal forma impregnadas no nosso dia-a-dia que já nem damos pela sua enorme utilidade. Servem para expressar diferentes tipos de valores relativos, tais como: a parte de um todo (proporções), uma relação entre valores de natureza distinta (rácios), a diferença relativa entre dois números (taxas de variação).

Não admira que sejam as percentagens a principal estrela nas notícias baseadas em dados estatísticos.

COMO SE CALCULA

Se tiver um número, resultado de uma divisão, e quiser convertê-lo numa percentagem:

1. Multiplicar o número por 100 (avançar duas casas decimais);
2. Acrescentar o símbolo "%".

$1{,}44 = 144\%$

$0{,}53 = 53\%$

Se tiver dois números quaisquer e quiser calcular a percentagem de um em relação ao outro:

1. Dividir um valor pelo outro;
2. Multiplicar por 100;
3. Acrescentar o símbolo "%".

Exemplo: *Qual é a proporção de mulheres na população residente em Portugal?*

$$\frac{\text{Mulheres}}{\text{População}} = \frac{5.445.489}{10.358.076} \approx 0{,}526 \xrightarrow[\times 100 \text{ e "\%"}]{} 52{,}6\%$$

Fontes: INE/Pordata (dados de 2015)

A partir de uma percentagem e do valor total a que ela se refere, é possível chegar-se a uma aproximação ao valor absoluto que ela representa.

Exemplo: *Se a proporção de mulheres na população é de 52,6%, quantas mulheres há aproximadamente?*

$$\frac{\text{Percentagem} \times \text{População}}{100} = \frac{52{,}6 \times 10.358.076}{100} \approx 5.448.348$$

O valor é aproximado porque a percentagem que representa a proporção de mulheres é um número já arredondado (ver capítulo sobre *arredondamentos*).

Ao se fazerem multiplicações com percentagens, é importante ter em mente que o número que tem o símbolo "%" foi multiplicado por 100. Por isso, é preciso dividi-lo por 100 antes de se fazerem os cálculos.

$$10\% = \frac{10}{100} = 0{,}10$$

Assim: $10\% \times 10\% \neq 100\%$

E sim: $10\% \times 10\% = 0{,}10 \times 0{,}10 = 0{,}01 = 1\%$

COMO SE LÊ

Quando lemos "tantos por cento de", o que isto significa, no fundo, é que há "tantos em cada 100" ou "tantos por cada 100".

Exemplos:

- *53% da população são mulheres:* em cada 100 habitantes, 53 são mulheres;
- *Índice de envelhecimento atinge 144%:* por cada 100 jovens, há 144 idosos;
- *Salários aumentaram 5%:* por cada 100 euros de salário, houve um aumento de cinco euros.

Convém não esquecer que uma percentagem não faz sentido sozinha e depende umbilicalmente do valor que está na base (denominador). Afinal, dizemos sempre "tantos por cento *de*" alguma coisa.

Voltando aos exemplos anteriores:

- Mulheres são 53% *da* população;
- Idosos equivalem a 144% *dos* jovens;
- Aumento foi de 5% *do* salário anterior.

ALERTA

Permilagem é *por mil*

Permilagem (‰) é uma medida semelhante à percentagem, mas significando “tantos em cada 1000” ou “tantos por cada 1000”. É calculada da mesma forma: divide-se um número pelo outro e o resultado é multiplicado, neste caso, por 1000.

MAIS OU MENOS QUANTO?

Arredondamentos

Arredondar um número significa descomplicá-lo, cortando-lhe algumas gorduras. O resultado é um valor que não corresponde exactamente ao número original, mas anda lá perto. Fazemos isso muitas vezes, por exemplo com os preços, quando dizemos que algo custa “mais ou menos” tantos euros.

Os arredondamentos seguem regras próprias que praticamente se resumem ao seguinte: de 0 a 4, arredondar para baixo; de 5 a 9, arredondar para cima. Mas vale a pena revê-las.

COMO SE CALCULA

Para arredondar qualquer número, basta identificar o primeiro algarismo da parte a dispensar e:

- Se for menor do que 5 (0, 1, 2, 3 ou 4), o algarismo anterior mantém-se inalterado;
- Se for igual ou maior do que 5 (5, 6, 7, 8 ou 9), aumenta-se o algarismo anterior em 1.

Exemplos:

Arredondar números decimais

	Zero casas	**Uma casa**	**Duas casas**
100,994	101	101,0	100,99
23,452	23	23,5	23,45

Arredondar números inteiros

<table>
<tr><th></th><th></th><th>Milhares</th><th>Milhões</th></tr>
<tr><td rowspan="2">859.600</td><td>Zero casas decimais</td><td>860 mil</td><td>1 milhão</td></tr>
<tr><td>Uma casa decimal</td><td>859,6 mil</td><td>0,9 milhões</td></tr>
<tr><td rowspan="2">10.450.200</td><td>Zero casas decimais</td><td>10.450 mil</td><td>10 milhões</td></tr>
<tr><td>Uma casa decimal</td><td>10.450,2 mil</td><td>10,5 milhões</td></tr>
</table>

Arredondar não significa, portanto, apenas deitar fora a parte que não se deseja e deixar todo o resto sem alterações. Se assim fosse, 14,9 seria arredondado para 14 e não para 15 e, com isso, a distância em relação ao valor real seria maior (0,9 em vez de 0,1).

ALERTA

Não arredondar arredondados

Não se deve arredondar um número sucessivamente. O arredondamento tem sempre de levar em conta o número original.

Incorrecto: 10,445 $\longrightarrow$ 10,45 $\longrightarrow$ 10,5 $\longrightarrow$ 11

Correcto: 10,445 $\longrightarrow$ 10

As regras do arredondamento significam a distribuição equilibrada do erro, com uma parte dos valores a ser ajustada para cima e outra parte, para baixo.

Ainda assim, um número arredondado não é um número preciso e, por isso, merece alguma cautela no momento de se fazerem cálculos.

ALERTA

Nunca é 100% exacto

É preciso ter muita atenção ao fazer cálculos com números arredondados. Os resultados nunca serão exactos, mas apenas uma aproximação.

Este alerta é particularmente importante, tendo em vista que muitos dados estatísticos são apresentados com valores já arredondados.

Por exemplo, em 2014 os idosos em Portugal representavam 20,1% da população. Mas se multiplicarmos aquela percentagem pelo número de residentes no país, a conta não bate certo, isto é, não obtemos o valor exacto da população idosa.

$$20{,}1\% \times 10.401.062 = \frac{20{,}1 \times 10.401.062}{100} = 2.090.613{,}462$$

O resultado obtido não só tem casas decimais – algo incompatível com a noção de indivíduo – como acrescenta cerca de 3000 idosos aos que realmente existiam naquele ano (2.087.505). E porquê? Porque a percentagem real é de 20,0701139941287%. Os 20,1% apresentados são um valor arredondado.

QUANTO VALE x?

Regra de três simples

Se por cada 100 casamentos há 70 divórcios, quantos serão os divórcios se houver 200 casamentos? A resposta é intuitiva: desde que o rácio entre casamentos e divórcios seja o mesmo, se as uniões duplicam, as separações também duplicam e, portanto, serão 140.

É esta a regra de três simples – um antiquíssimo método para se resolver problemas envolvendo grandezas proporcionais, onde temos três números e é preciso descobrir um quarto.

COMO SE CALCULA

Em 2014, havia 81 dentistas por cada 100.000 habitantes em Portugal, e a população era de 10.401.062 habitantes. Qual era o número total de dentistas?

Assumindo que os valores estão **directamente relacionados**, ou seja, que se há mais população há um aumento proporcional de dentistas, efectua-se o seguinte cálculo:

1. Atribuir x ao número que se quer descobrir (total de dentistas);
2. Alinhar os pares de dados, da seguinte forma:

$$\begin{array}{rcl} 100.000 & \text{———} & 81 \\ 10.401.062 & \text{———} & x \end{array}$$

81(*100.000 está para 81…*)

x (*…assim como* 10.401.062 *está para* x)

3. Fazer uma multiplicação cruzada e resolver:

$$\begin{array}{rcl} 100.000 & \times & 81 \\ 10.401.062 & & x \end{array} \qquad 100.000\,x = 10.401.062 \times 81$$

$$x = \frac{10.401.062 \times 81}{100.000} \qquad x \approx 8425$$

Ou seja, havia aproximadamente 8400 dentistas em Portugal em 2014.

Se uma distância é percorrida em duas horas a 75 km/h, quanto dura a viagem se a velocidade for de 100 km/h?

Neste caso, é uma **relação inversa**, isto é, um aumento num valor provoca uma diminuição proporcional no outro. Nestas situações, a regra de três simples é feita de forma semelhante, apenas invertendo-se a ordem de um dos pares.

$$\begin{array}{rcl} 75 & — & 2 \\ 100 & — & x \end{array} \longrightarrow \begin{array}{rcl} 75 & — & x \\ 100 & — & 2 \end{array} \longrightarrow \begin{array}{rcl} 75 & \times & x \\ 100 & & 2 \end{array}$$

$$x = \frac{75 \times 2}{100} = 1{,}5 \text{ horas}$$

Resultado: a viagem demoraria uma hora e meia.

ALERTA

Por que "aproximadamente"?
Dados estatísticos apresentados sob a forma de rácios, taxas ou proporções normalmente já vêm em números arredondados. É o caso dos 81 dentistas – um valor arredondado – por cada 100.000 habitantes. Por isso, quaisquer cálculos feito com eles nunca nos conduzirão aos valores exactos que estiveram na sua origem. (Ver capítulo sobre *arredondamentos.*)

No primeiro exemplo, o que a regra de três simples está a dizer é que se por cada 100.000 habitantes havia 81 dentistas em Portugal em 2014, então, sabendo que a população do país era de 10.401.062 pessoas, deveria haver aproximadamente 8400 dentistas.

Este exemplo mostra como a regra de três simples é útil para estimar, de forma aproximada, dados "escondidos". Permite partir de um rácio, taxa ou proporção e chegar a um valor absoluto (ver capítulos sobre *rácios*, *taxas* e *proporções*).

Regras de três simples são essenciais também para se calcular números-índice, que são uma forma simplificada de medir a distância entre dois valores (ver capítulo sobre *números-índice*). Por exemplo, como o PIB *per capita* de Portugal se comparava com o da União Europeia (UE) em 2015?

Os valores, em PPS (paridade do poder de compra padrão)[5], eram:

	UE	Portugal
PIB *per capita*	28.743	22.243

Fontes: Eurostat/Pordata (dados provisórios)

Se assumirmos que os 28.743 PPS por habitante da UE equivalem a 100, então os 22.243 PPS por habitante de Portugal equivalerão a x. Ou seja:

28.743 ——— 100 (*28.743 está para 100…*)
22.243 ——— x (*…assim como 22.243 está para x*)

$$28.743\,x = 22.243 \times 100$$

$$x = \frac{22.243 \times 100}{28.743} \approx 77$$

Resultado: se o PIB per capita *da UE equivale a 100, o de Portugal equivale a 77.*

5 PPS é uma unidade monetária artificial que elimina as diferenças no poder de compra ao igualar, teoricamente, os níveis de preços entre países.

COMO POUPAR MIL PALAVRAS?

Gráficos

Embora haja actualmente infinitas alternativas para a visualização de dados, os gráficos tradicionais continuam a ser um aliado imprescindível para a comunicação das estatísticas. Vemo-los em várias notícias, mas muitas outras mereciam a sua utilização.

Desde que bem feitos, cumprem com eficácia o ditado de que “uma imagem vale mais que mil palavras”.

Auto-explicativo, adequado, simples e claro são as palavras-chave para um gráfico de sucesso.

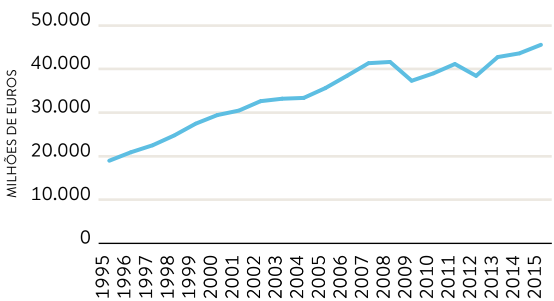
Destacam-se três aspectos a ter em conta:

- O tipo de gráfico a utilizar;
- Os elementos que deve conter;
- Os erros a evitar.

Gráficos de linhas, de barras (horizontais), de colunas (verticais) e os circulares são os mais comuns no dia-a-dia.

Gráfico de linhas: particularmente adequado para mostrar a evolução temporal de um indicador. Por exemplo, as receitas das administrações públicas em impostos, ao longo dos anos:

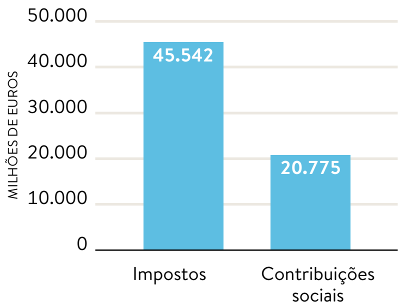
Receitas dos impostos das administrações públicas em Portugal



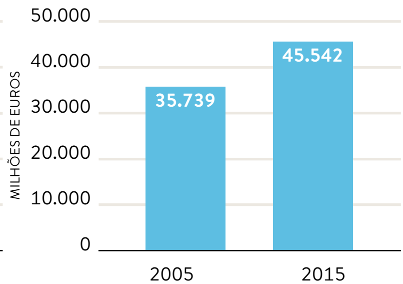
Fontes: INE/MF/Pordata

Gráfico de barras/colunas: é especialmente útil para pôr lado a lado valores de categorias distintas num mesmo momento ou da mesma categoria em momentos diferentes do tempo. Por exemplo:

Receitas dos impostos e das contribuições sociais das administrações públicas em 2015



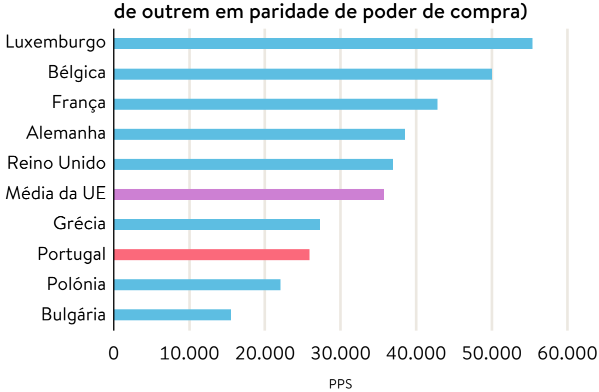
Receitas dos impostos das administrações públicas em 2005 e 2015



Fontes: INE/MF/Pordata

O mesmo tipo de valores também pode ser apresentado na horizontal (gráfico de barras), comparando diferentes categorias, territórios, países ou regiões num determinado ano.

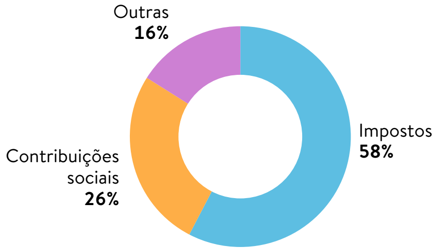
Salário médio em 2015
(Remuneração anual dos trabalhadores por conta de outrem em paridade de poder de compra)



Fontes: Eurostat/Pordata

Gráfico circular: utilizado para ilustrar o modo como um "bolo" está repartido num determinado momento, mostrando a proporcionalidade das várias partes que compõem o total. Por exemplo, as receitas das administrações públicas, segundo as origens:

Receitas das administrações públicas em 2015



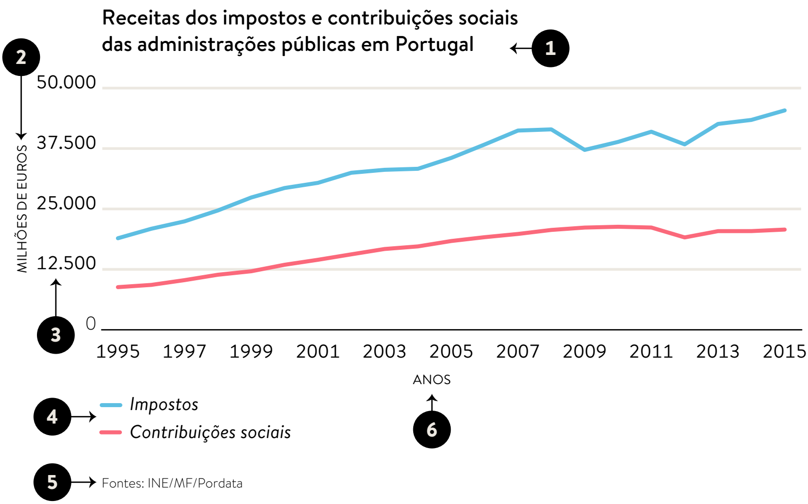
Fontes: INE/MF/Pordata

ALERTA

Gráficos circulares = 100%

Se a soma das partes de um gráfico circular for diferente de 100%, é provável que exista um erro. É uma questão básica de matemática: a soma de todas as partes de um todo, em percentagem, é sempre 100%. Quando se usam dados arredondados, no entanto, o resultado da adição pode ter uma diferença muito pequena em relação a 100%.

Os tipos de gráficos aqui mostrados são apenas alguns dos habitualmente utilizados. Seja qual for o seu estilo, há elementos essenciais sem os quais não é possível compreendê-los. Estes são os que, em princípio, devem acompanhar um gráfico com dois eixos (horizontal e vertical):



1. Título
Deve dizer claramente do que trata o gráfico.

2. Legenda do eixo *y*
Unidade de medida da escala. Em alternativa, pode estar no título.

3. Escala de valores do eixo *y*
Por norma, os valores começam em zero.

4. Legenda
Necessária se estiverem representados mais de um indicador.

5. Fonte
De onde vêm os dados.

6. Legenda do eixo *x*.

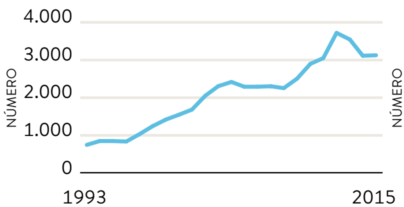
Há amplas oportunidades para se cometerem erros num gráfico, ou mesmo manipulá-los de modo a dar uma interpretação ora mais, ora menos dramática da realidade. Alguns *sites* na Internet, aliás, dedicam-se a encontrar alguns desses exemplos enganadores nas notícias.

Algumas situações comuns a evitar:

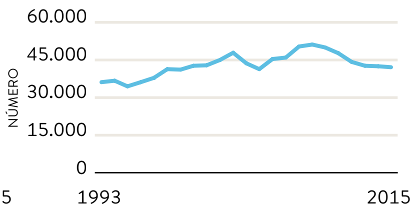
Comparar gráficos com escalas diferentes

Se colocarmos lado a lado dois gráficos sobre os crimes contra o Estado e contra as pessoas em Portugal, em escalas diferentes, facilmente ficaremos mais preocupados com o primeiro e, possivelmente, seria com base nele que se faria uma notícia:

Crimes contra o Estado registados pela PSP



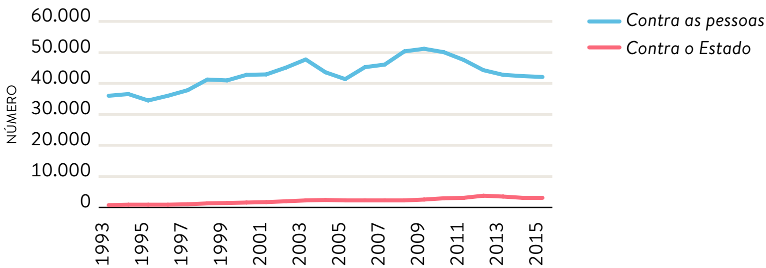
Crimes contra as pessoas registados pela PSP



Fontes: DGPJ/Pordata

Mas juntando ambas as linhas num só gráfico com a mesma escala, a realidade tem outra leitura: embora o número de crimes contra o Estado esteja a subir, ainda são relativamente poucos quando comparados com os crimes contra as pessoas.

Crimes contra o Estado e contra as pessoas registados pela PSP

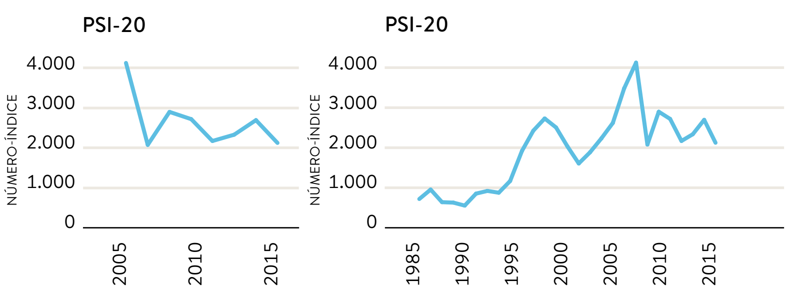


Fontes: DGPJ/Pordata

Na verdade, aqui não há soluções únicas e cada tipo de gráfico tem os seus prós e contras. Os primeiros dois não são comparáveis entre si, mas mostram a evolução de cada tipo de crime individualmente. O terceiro mostra quais os mais frequentes, mas esmaga a linha dos crimes contra o Estado quase numa recta. Neste caso, a apresentação dos três gráficos seria a solução ideal. Mas muitas vezes, por limitações de espaço, isso não é possível numa notícia.

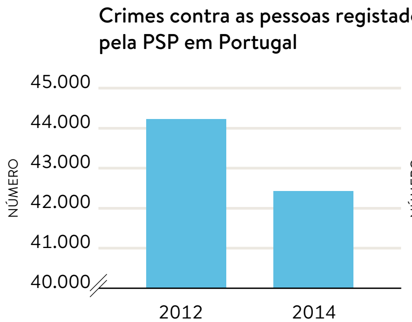
Mostrar apenas parte de uma série temporal

Nem sempre utilizamos todos os dados de que dispomos para construir um gráfico. E isto faz sentido quando o que nos interessa é, por exemplo, realçar apenas um breve período temporal. É importante ter em conta, no entanto, que, ao fazê-lo, se está a eliminar parte da história. Veja-se o caso do gráfico abaixo, à esquerda, que mostra que o índice bolsista PSI-20 sofreu uma queda abrupta depois de 2007. Já o da direita ajuda a relativizar a queda, revelando que o valor de 2008 é, afinal, superior ao valor de 2002 e a grande parte dos observados nos anos 90.

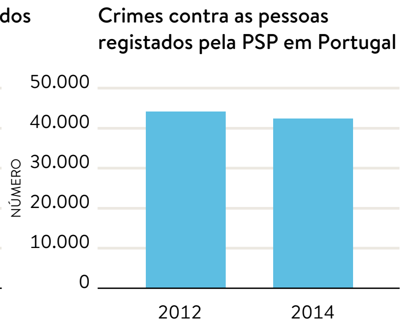


Fontes: BdP/Pordata

Gráficos em que a escala de valores do eixo *y* não começa em zero
Normalmente a escala do eixo *y* dos valores deve começar em zero. Caso contrário teremos uma imagem distorcida da realidade que se está a representar. Por exemplo:



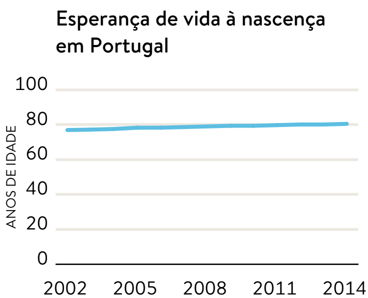
Fontes: DGPJ/Pordata

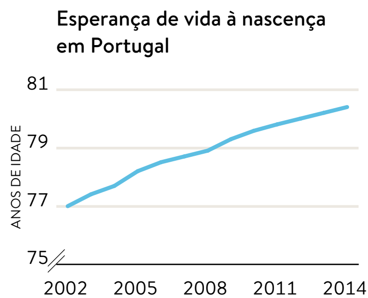


Errado: escala de valores do eixo *y* começa em 40.000 e o número de crimes em 2014 parece ser metade do de 2012.

Certo: escala de valores do eixo *y* começa em zero e dá a dimensão certa da diferença no número de crimes entre os dois anos.

Embora os gráficos de linhas também devam ter a escala no eixo *y* a começar em zero, por vezes pode fazer sentido iniciá-la noutro ponto para salientar pequenas variações importantes que de outra forma passariam despercebidas. Por exemplo, ligeiros aumentos ou reduções numa população com milhões de habitantes são quase imperceptíveis num gráfico em que a escala do eixo *y* começa em zero. O mesmo acontece com a evolução da esperança de vida à nascença.





Fontes: INE/Pordata

Escala a começar em zero: não se vê a evolução.

Escala a começar em 75 anos: percebe-se melhor a evolução.

ALERTA

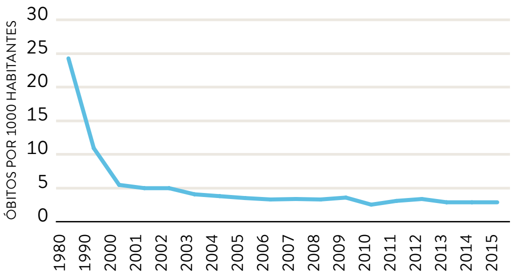
Indicar a quebra de escala

Quando a escala do eixo *y* de um gráfico não começa em zero deve-se inserir o símbolo // no seu lugar.

Distâncias iguais para diferentes amplitudes temporais

Imaginemos, por hipótese, que as estatísticas sobre taxa de mortalidade infantil só tivessem dados para 1980, 1990, 2000 e, a partir daí, para todos os anos. Fazer um gráfico com espaçamentos idênticos entre todas as datas para as quais existem dados conhecidos resultaria numa imagem errada, pois assumiria que o intervalo entre dois anos seguidos seria o mesmo do que entre uma década e outra. A distância entre os anos no eixo x (horizontal) do gráfico deve ser sempre igual ou, no caso de falharem anos, ser proporcional à amplitude temporal. Além disso, a linha que une os anos de 1980, 1990 e 2000 deve ser representada a tracejado, significando que não há dados para os demais anos naquele intervalo.

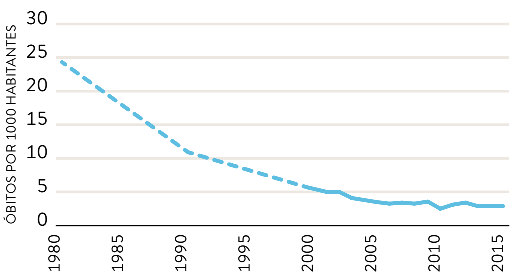
Taxa de mortalidade infantil em Portugal



Fontes: INE/Pordata

Errado: décadas e anos com o mesmo espaçamento.

Taxa de mortalidade infantil em Portugal



Fontes: INE/Pordata

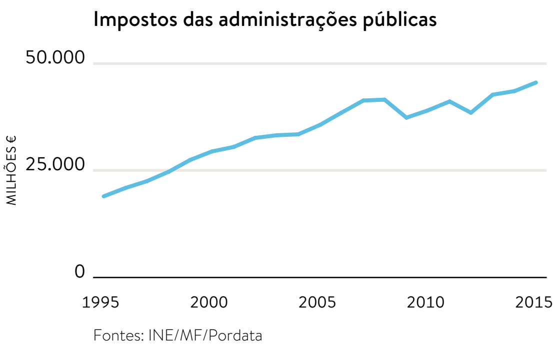
Certo: anos igualmente espaçados.

ALERTA

Opte pela simplicidade

O público-alvo das notícias é muito diverso, por isso, a simplicidade é o melhor farol para que um gráfico seja compreendido pelo maior número possível de pessoas. Evite congestioná-lo com muita informação.

A necessidade de simplificação é tanto maior quanto menor for o espaço disponível para se introduzir um gráfico numa notícia. Pode ser preciso utilizar títulos mais curtos, escalas com poucos valores ou unidades de medida abreviadas. O importante, no entanto, é não se esquecer dos elementos fundamentais para que a imagem seja compreensível, independentemente do corpo da notícia. Por exemplo:



Este é o mesmo gráfico que se apresentou anteriormente neste capítulo, mas simplificado para caber numa coluna de jornal. O título foi reduzido e, para a imagem não ficar muito carregada, as escalas têm menos valores, a legenda do eixo *y* está abreviada e a do eixo *x* foi eliminada.

TRADUZINDO?

Escrita simples

A formação bruta de capital fixo em Portugal aumentou 3,47% em 2014, chegando a 25.993,1 milhões de euros a preços correntes, o que representa uma variação absoluta de 871,1 milhões de euros face a 2013, ano em que este indicador sofreu uma contracção de 5,81% em relação a 2012.

Difícil de entender? Não admira. O texto acima, felizmente hipotético, é a antítese da regra número um da comunicação de estatísticas: ser simples, curta e directa.

Tudo no exemplo da página anterior conspira a favor da complicação: a frase é longa, tem muitos números, alguns deles de difícil leitura, não apresenta uma síntese clara, inclui termos incompreensíveis e todo o texto é intrinsecamente enfadonho.

Embora o jornalismo tenha as suas próprias normas para evitar notícias que ponham qualquer um a dormir, eis algumas sugestões específicas para lidar com números:

Contar uma história

Pode parecer um conselho desnecessário, pois este é o dia-a-dia do jornalismo. Mas, na verdade, muitas notícias com dados estatísticos são apenas peças descritivas, apresentando uma sucessão interminável de números despejados directamente dos comunicados que chegam às redacções. Facilmente derivam para aquilo que um guia da ONU baptizou de “estatísticas de elevador”: este indicador subiu, aquele desceu; neste ano subiu, naquele desceu; neste caso subiu, naquele desceu…[6]

A necessidade de contar uma história tem de estar na cabeça do jornalista antes de escrever a notícia. E contar uma história com números não é aliar a matemática à ficção. É somente interpretar as estatísticas, dar-lhes vida e extrair-lhes o que tiverem de maior interesse para o cidadão.

Começar com palavras

Ser curto e directo não implica trazer logo para a primeira frase da notícia o número porventura mais significativo de uma informação estatística. Pelo contrário, uma sentença inicial só com palavras, destacando antes a interpretação dos números, desperta mais interesse do que os números em si – que podem não ter apelo de

6 UNECE (2009). *Making Data Meaningful, Part 1*, p. 6

espécie alguma. Esta não é uma regra universal, é claro, pois há casos em que é inevitável trazer um número para a primeira frase. No entanto, na maioria das situações, será melhor começar de outra forma. Convém lembrar que a primeira frase de uma notícia é aquela que agarra ou afasta o público.

EXEMPLOS

Em vez de: *Portugal recebeu 29.896 imigrantes em 2015, mais 53,2% face ao ano anterior, mas 40.377 pessoas emigraram, menos 18,5% do que em 2014.*

Talvez: *A emigração diminuiu e a imigração aumentou, mas, ainda assim, houve mais gente a sair do que a entrar no país em 2015.*

Em vez de: *A população de Portugal caiu 0,32% em 2015, pelo sexto ano consecutivo, fruto de um saldo migratório e de um saldo natural negativos, somando 10.341.330 pessoas, menos 33.492 do que em 2014.*

Talvez: *A população de Portugal voltou a cair em 2015, com mais mortes do que nascimentos e mais emigrantes do que imigrantes.*

Evitar o jargão técnico

As estatísticas estão recheadas de jargão e termos técnicos indigestos. É imperativo vertê-los para uma linguagem próxima e que todos entendam.

EXEMPLOS

Em vez de: *A formação bruta de capital fixo em Portugal aumentou 3,5% em 2014.*

Talvez: *O investimento em Portugal aumentou 3,5% em 2014.*

Em vez de: *O índice sintético de fecundidade em Portugal subiu para 1,30 em 2015.*

Talvez: *O número médio de filhos por mulher em idade fértil em Portugal subiu para 1,30 em 2015.*

Mesmo os itens mais populares do jargão estatístico merecem ser traduzidos. Por exemplo, variação “homóloga” é um termo que prolifera nas notícias como pipocas, possivelmente porque imaginamos que todos sabem o que significa. Mas por que não dizer antes “em relação ao mesmo período do ano anterior”?

Evitar tecnicismos não implica necessariamente ignorar os termos técnicos. Pode ser relevante mencioná-los, mas, no plano ideal, devem aparecer apenas depois da sua versão mais compreensível.

Seja como for, não se pode esquecer que a prioridade é simplificar, sem que isso prejudique o rigor da notícia.

Citar poucos números

Polvilhar uma notícia com muitos números é meio caminho para o aborrecimento total. Seleccionar os mais importantes e reparti-los por frases curtas, com apenas um ou dois em cada uma, torna a compreensão muito mais fácil.

EXEMPLO

Em vez de: *Cerca de 2,1 milhões de hóspedes foram recebidos nos estabelecimentos hoteleiros em Portugal em Setembro de 2016, representando um crescimento homólogo de 7,5%, uma variação mais elevada do que a registada no mês anterior (3,6%). O número de dormidas chegou a 5,9 milhões, com um aumento homólogo de 6,5%, contra 4,2% em agosto.*

Talvez: *O movimento nos hotéis em Portugal aumentou em Setembro em relação ao mesmo mês do ano passado. O número de hóspedes subiu 7,5%, e o de dormidas, 6,5%. Pouco mais de dois milhões de pessoas foram recebidas nos estabelecimentos hoteleiros nesse mês.*

Encurtar e arredondar

Números muito grandes são difíceis de ler, mas estão em todo o lado nas estatísticas. Um número como 25.993,1 milhões de euros torna-se mais acessível se arredondado a 26 mil milhões de euros. Da mesma forma, é muito mais fácil compreender que a população de Portugal era de cerca de 10,3 milhões de habitantes no final de 2015, do que se dissermos o valor completo: 10.341.330.

EXEMPLOS

Em vez de: *No próximo domingo, 9.741.377 eleitores vão escolher o novo Presidente da República.*

Talvez: *No próximo domingo, cerca de 9,7 milhões de eleitores vão escolher o novo Presidente da República.*

Em vez de: *Portugal perdeu 33.492 habitantes em 2015.*

Talvez: *Portugal perdeu cerca de 33 mil habitantes em 2015.*

Evitar muitas casas decimais

Se as casas decimais significarem mais ruído do que informação, o melhor é deixar os números inteiros.

EXEMPLOS

Em vez de: *Segundo as últimas estimativas, 2,64% dos habitantes têm 85 ou mais anos.*

Talvez: *Segundo as últimas estimativas, cerca de 3% dos habitantes têm 85 ou mais anos.*

Em vez de: *O Estado arrecadou 12.695,7 milhões de euros em IRS em 2015.*

Talvez: *O Estado arrecadou 12.696 milhões de euros em IRS em 2015.*

Casas decimais devem ser mantidas nas situações em que façam sentido. É o caso de indicadores nos quais pequenas variações têm um impacto significativo. Por exemplo: a taxa de inflação ou o crescimento do PIB. Nos preços também, quando os cêntimos importam – por exemplo, nos combustíveis.

Em números menores do que um, naturalmente, também se devem manter alguns algarismos após a vírgula, conforme a necessidade.

Por norma, no entanto, deve utilizar-se o menor número possível de casas decimais – uma, no máximo duas, raramente mais.

Traduzir percentagens

É sempre melhor traduzir as percentagens para uma razão que se compreenda com maior facilidade. Por exemplo, 25% é o mesmo que "um em cada quatro". Atenção que a correspondência nem sempre é um valor certo: 33% equivalem, *aproximadamente*, a um em cada três. Aqui fica uma cábula (valores certos a negrito, os demais são aproximados):

%	Equivale a…	%	Equivale a…
5%	**Um em cada 20**	17%	Um em cada seis
6%	Um em cada 17	20%	**Um em cada cinco**
7%	Um em cada 14	25%	**Um em cada quatro**
8%	Um em cada 13 (ou 12)	33%	Um em cada três
9%	Um em cada 11	34%	Um em cada três
11%	Um em cada nove	40%	**Dois em cada cinco**
12%	Um em cada oito	60%	**Três em cada cinco**
13%	Um em cada oito	66%	Dois em cada três
14%	Um em cada sete	67%	Dois em cada três
15%	Um em cada sete	75%	**Três em cada quatro**
16%	Um em cada seis	80%	**Quatro em cada cinco**

EXEMPLO

Em vez de: *Dados do Ministério da Justiça indicam que 58,3% dos crimes registados pela PSP são contra o património.*

Talvez: *Dados do Ministério da Justiça indicam que cerca de três em cada cinco crimes registados pela PSP são contra o património.*

Usar analogias e comparações

Números muito grandes ou muito pequenos necessitam de algo que lhes dê um sentido de proporção. Comparações são uma alternativa.

EXEMPLO

Em vez de: *A área ardida nos incêndios florestais em Portugal Continental em 2003 foi de 426 mil hectares.*

Talvez: *A área ardida nos incêndios florestais em Portugal Continental em 2003 foi de 426 mil hectares, ou seja, 5% da superfície de Portugal Continental.*

Melhor ainda: *Um em cada 20 hectares da superfície de Portugal Continental foi varrido pelas chamas em 2003.*

Analogias, com os devidos cuidados para não serem absurdas ou inverosímeis, também são uma opção. Há algumas que são clássicas: um hectare equivale a um campo de futebol, 2500 metros cúbicos de água são suficientes para encher uma piscina olímpica; 40.000 quilómetros é o mesmo que uma volta ao mundo. Mas a imaginação, aliada ao bom senso, é o limite.

EXEMPLO

Em vez de: *O investimento em Portugal aumentou quase 900 milhões de euros no ano passado.*

Talvez: *O investimento em Portugal aumentou quase 900 milhões de euros no ano passado, como se uma nova Ponte Vasco da Gama tivesse sido construída.*

Transformar um valor anual em mensal, diário ou horário (dividindo-o por 12, 365 ou 8760) pode ser outro caminho, desde que faça sentido e fique claro tratar-se não de um valor observado mas de uma média estimada.

EXEMPLO

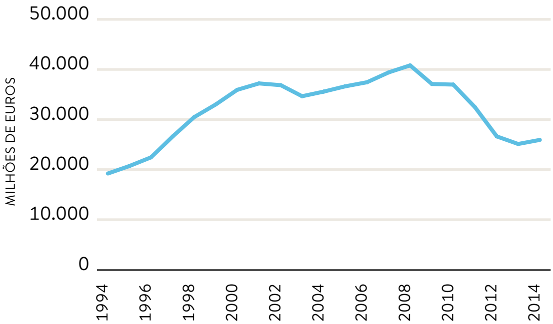
Em vez de: *Hospitais e centros de saúde registaram um total de 8,7 milhões de urgências no ano passado.*

Talvez: *Hospitais e centros de saúde registaram um total de mil urgências por hora, em média, no ano passado.*

Usar gráficos

Um gráfico, por mais simples ou pequeno que seja, ajuda sempre na comunicação dos números. Uma notícia sobre a subida de 3,5% no investimento em Portugal em 2014, por exemplo, beneficiaria muito com uma imagem da evolução desse indicador ao longo dos anos.

Investimento em Portugal
(Formação bruta de capital fixo)



Fontes: INE/BdP/Pordata

Gráficos são fáceis de fazer, mas é preciso estar atento às suas regras e aos erros que facilmente se podem cometer (ver capítulo sobre *gráficos*).

Adicionar contexto

Os números são frios. Sem um contexto que os enquadre, dizem pouco. Mencionar apenas que o investimento em Portugal aumentou 3,5% não fornece grandes pistas sobre o que esta subida significa. O gráfico anterior já é um grande auxílio, pois situa o último dado disponível em relação à evolução temporal do investimento, mostrando que:

- Os 3,5% a mais representam uma inversão da tendência de queda desde 2008;

- Apesar disso, o valor do investimento em 2014 ainda era muito mais baixo do que em anos anteriores.

Só com estes dois elementos é possível construir uma notícia com um arranque mais analítico e contextualizado:

> O investimento em Portugal voltou a crescer em 2014, depois de cinco anos de quedas sucessivas. A formação bruta de capital fixo – um indicador do investimento – aumentou 3,5% em relação a 2013. O valor chegou a 26 mil milhões de euros, mas, ainda assim, é o segundo mais baixo desde 1997.

Adicionar contexto significa, sobretudo, ir à procura de mais. Implica olhar para mais informação e não se limitar a reproduzir os números muito gerais que chegam à redacção.

Conferir tudo

Os números têm uma desvantagem em relação às palavras: não há um corrector automático – como aqueles dos processadores de texto – que permita identificar os erros. Por isso, cabe ao jornalista confirmar se cada número que utilizou está certo, se não há gralhas, más transcrições, contas erradas ou outros deslizes. É muito fácil incorrer num lapso qualquer.

Fazer uma *checklist*

Ao finalizar uma notícia com números, vale a pena revê-la e perguntar se uma pessoa que nada sabe sobre o assunto será capaz de compreender o que ali está. Para isso, uma *checklist* pode ser útil:

1. Os números são claros?
2. Não há informação a mais?
3. Há algum termo técnico difícil de entender?

4. O contexto é suficiente?
5. Há um gráfico a acompanhar?
6. E sobretudo: alguém se interessará por isso?

Se a resposta for “não” às primeiras cinco perguntas, certamente também será negativa para a última. Se for “sim” para as cinco primeiras e “não” para a última, é porque o problema está no pouco interesse que os dados podem ter. Nesse caso, vale a pena avaliar se a notícia merece mesmo ser publicada.


O QUÊ

QUEM TEM MAIS?

Números absolutos e relativos

Que país da União Europeia tem mais presença de estrangeiros? A resposta é a Alemanha, com 7,5 milhões de residentes de nacionalidade estrangeira em 2015. Ou então o Luxemburgo, com apenas 260 mil.

Estranho, não? Sim, é verdade. Ambos os países são de facto candidatos ao pódio: a Alemanha por ter o maior número de estrangeiros e o Luxemburgo por ter a maior proporção de estrangeiros na população. No primeiro caso, estamos a falar de um valor absoluto, e no segundo, de um valor relativo.

O QUE É

Número absoluto é o que representa quantidades reais, como 12 empresas, 100 euros ou 2500 automóveis. Por vezes a realidade que o número representa não é tangível ou não resulta de uma única contagem, como é o caso do PIB.

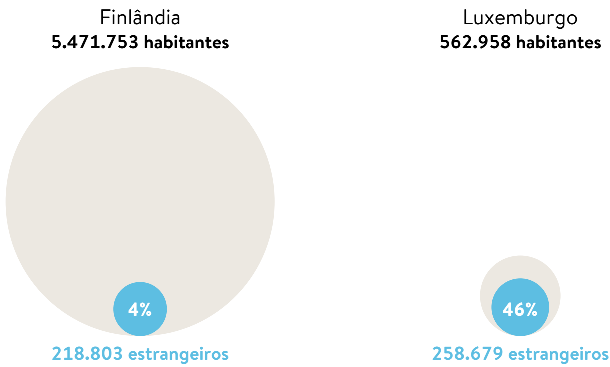
Número relativo é uma medida que expressa a relação entre dois valores, como a população estrangeira no total da população residente, o número de médicos por 100.000 habitantes ou a quantidade de processos judiciais resolvidos em relação aos novos.

Ambos os números são centrais nas estatísticas. Os absolutos dão-nos a dimensão concreta de uma realidade. Os relativos permitem-nos comparar realidades de dimensão diferente. E pensar nos dois em simultâneo ajuda-nos a evitar ratoeiras numéricas.

À primeira vista, por exemplo, o Luxemburgo estava relativamente próximo da Finlândia em número de habitantes com nacionalidade estrangeira em 2015:

Fontes: Eurostat/Pordata

O panorama muda totalmente de figura, porém, quando levamos em conta a população total de cada país:



Fontes: Eurostat/Pordata

Ou seja, considerar apenas os números absolutos esconde, neste caso, uma disparidade que só os números relativos são capazes de mostrar. O total de estrangeiros é muito próximo, mas no Luxemburgo eles representam quase metade da população (46%), enquanto na Finlândia são apenas 4%, isto é, menos de um em cada 20.

COMO SE LÊ

Os **números absolutos** lêem-se como vieram ao mundo, sempre acompanhados da sua unidade de medida:

- 258.679 estrangeiros

Os **números relativos** podem ser expressos de várias formas:

- *Cerca de 46% dos habitantes são estrangeiros;*
- *Cerca de metade dos habitantes é estrangeira;*
- *Aproximadamente um em cada dois habitantes é estrangeiro;*
- *Há 46 estrangeiros em cada 100 habitantes.*

O mesmo raciocínio vale no sentido inverso, ou seja, os números relativos por vezes também não contam toda a história. Em 2015, tanto em Espanha como na Noruega, 10% da população eram cidadãos com nacionalidade estrangeira. No primeiro caso, no entanto, estão em causa cerca de 4,5 milhões de pessoas, e no segundo, 500 mil. Nesta situação, a igualdade nos números relativos não revela a grande diferença que há nos valores absolutos.

Assim, quando se pergunta "quem tem mais", normalmente haverá duas ou mais respostas, conforme se utilizem valores absolutos ou relativos.

ALERTA

Qual é mais importante?

Os números absolutos não são mais importantes do que os relativos, nem vice-versa. São tipos de valores diferentes, aplicados para finalidades distintas e que se podem complementar.

Onde há mais acidentes de trânsito? Tanto poderá ser nos pontos onde o número absoluto de acidentes é maior, como nos pontos onde a razão entre o número de acidentes e o número de carros a circular é maior. O primeiro indicador é importante para os serviços de assistência pois localiza os pontos com maior probabilidade de haver muitos acidentes – algo que pode ser apenas o corolário natural de uma via com intenso tráfego. O segundo indica onde há maior probabilidade de qualquer carro ter um acidente e é relevante para quem é responsável pela condição das estradas, pois naqueles sítios, de certeza, alguma coisa vai mal.

ALERTA

Para valores pequenos, números absolutos

Sempre que estamos perante valores muito pequenos, os números absolutos devem ter primazia sobre os relativos. Por exemplo, se o número de acidentes de viação passou de um para dois, houve um aumento de 100% (número relativo) mas, na verdade, estamos a falar de um único acidente adicional.

Por permitir interpretar melhor a realidade, a combinação entre números absolutos e números relativos constitui a espinha dorsal das estatísticas. Para um mesmo tema, pode haver indicadores diferentes que contemplem estes dois tipos de valores. Exemplos:

Números absolutos	Números relativos
Dívida pública	Dívida pública em % do PIB
Médicos	Médicos por habitante
População inactiva	Taxa de inactividade
Pequenas e médias empresas	% de pequenas e médias empresas

É importante ter em conta que os números relativos surgem nas estatísticas já de modo simplificado e, via de regra, não sabemos à partida quais os valores absolutos que lhes deram origem. Dizer que há 4% de estrangeiros na população é uma forma quase gráfica de comunicar uma realidade, mas não nos esclarece sobre quantos são os estrangeiros, nem sobre a dimensão da população.

QUANTO EM RELAÇÃO AO TOTAL?

Proporções

A Internet está a chegar à idade sénior. E a prova disso está nas estatísticas: em 2015, a disseminação da *web* entre os mais velhos em Portugal já era de 27%, contra apenas 2% dez anos antes.

Mas 27% de quê? Muitos jornalistas já terão feito esta pergunta ao olharem para um quadro com dados em percentagem. E têm razão para isso, pois há diferentes tipos de valores que podem ser apresentados assim. Os mais usuais são as proporções.

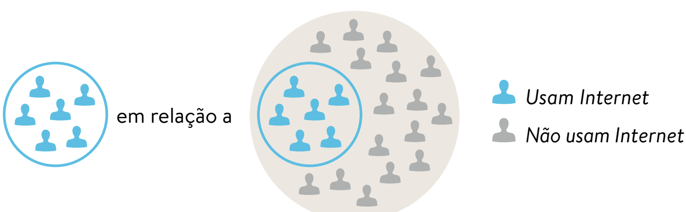
O QUE É

Uma **proporção** é uma medida que representa a parte de um todo, como a fatia de um bolo. Resulta da divisão de um conjunto por outro conjunto que o contém. Por exemplo, o número de votos num partido dividido pelo número total de votos.

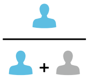
Dois atributos caracterizam uma proporção:

- Primeiro, relaciona sempre conjuntos de observações do mesmo tipo e que, por isso, tenham a mesma unidade de medida – pessoas com pessoas, euros com euros, votos com votos, crimes com crimes.
- Segundo, como já referido, relaciona uma parte de um conjunto com o todo (universo) que o inclui. Por exemplo, três maçãs verdes num cesto com 12 maçãs. A proporção de maçãs que são verdes equivale a três em cada doze 12 maçãs, ou uma em cada quatro.

Os 27% citados no início deste capítulo são uma proporção apresentada em percentagem, isto é, cujo resultado da relação é multiplicado por 100. Representam a quota de idosos que utilizam a Internet em relação a todos os idosos considerados. Ou seja:



Assim sendo:

Proporção = 

numerador: a parte

denominador: o todo

Nas proporções, portanto, o valor que está na parte de cima da divisão (o numerador) faz parte do valor que está em baixo (o denominador).

COMO SE LÊ

As proporções podem ser apresentadas de várias formas:

- *Cerca de 27% dos idosos utilizam a Internet;*
- *Por cada 100 idosos, cerca de 27 utilizam a Internet;*
- *Cerca de um quarto dos idosos utiliza a Internet;*
- *Aproximadamente um em cada quatro idosos utiliza a Internet;*
- *A proporção de idosos que utiliza a Internet é de 0,27.*

ALERTA

Nunca mais de 100%

Uma proporção nunca pode ser superior a 100% (ou maior do que um, caso o resultado não esteja em percentagem), pois a parte de um todo não pode ser maior do que o todo em si. Se encontrar um valor superior a 100% numa tabela com percentagens, é porque se trata de outro tipo de valor, como um rácio ou uma taxa de variação.

COMO SE CALCULA

Para o cálculo de uma proporção:

- Dividir a parte (numerador) pelo todo (denominador).

Para exprimir o valor em percentagem:

- Multiplicar o resultado por 100 e acrescentar o símbolo “%”.

Se as casas decimais não forem importantes, arredondar para um valor facilmente compreensível (ver capítulo sobre *arredondamentos*).

Exemplo: qual era a proporção de idosos em Portugal em 2014?

Idosos	2.087.505
População	10.401.062

Fontes: INE/Pordata

1. Dividir a parte pelo todo:

$$\frac{2.087.505}{10.401.062} \approx 0{,}2007$$

2. Multiplicar por 100 e adicionar o símbolo “%”:

$$0{,}2007 \xrightarrow[\times 100 \text{ e “\%”}]{} 20{,}07\%$$

3. Arredondar (opcional): 20%

Por vezes, precisamos de fazer o cálculo ao contrário: partir de uma proporção para se chegar a uma quantidade absoluta. Por exemplo, sabemos que 20% dos habitantes do país são idosos, mas queremos obter o número concreto de pessoas nessa faixa etária. Matematicamente falando, a solução seria muito fácil: multiplicar a percentagem pelo universo total considerado (toda a população) e dividir por 100 (ver capítulo sobre *regra de três simples*). Na prática, no entanto, este exercício só permite chegar a um valor aproximado do número total de idosos, pois as proporções aparecem nas estatísticas já como números arredondados, mesmo que contenham uma ou mais casas decimais. E efetuar cálculos com números arredondados é uma boa receita para equívocos (ver capítulo sobre *arredondamentos*).

ALERTA

Quando evitar proporções

Se o universo de que estamos a falar for muito reduzido, as proporções não são medidas úteis. Que sentido faz dizer que, em Marvão, 67% dos casamentos em 2015 foram católicos e 33% apenas civis, sabendo que só foram celebrados três matrimónios? Um único casamento civil adicional empataria as contas em 50-50% e mais um outro viraria o jogo para 40-60%.

É preciso ter em atenção que algumas proporções aparecem com outros nomes, como "taxas" ou "índices". Exemplos:

$$\text{Taxa de abstenção} = \frac{\text{Eleitores que não votaram}}{\text{Todos os eleitores}}$$

$$\text{Taxa de desemprego} = \frac{\text{Desempregados}}{\text{Desempregados + Empregados}}$$

$$\text{Índice de longevidade} = \frac{\text{Pessoas com 75 ou mais anos}}{\text{Pessoas com 65 ou mais anos}}$$

$$\text{Taxa de analfabetismo} = \frac{\text{Analfabetos}}{\text{População}}$$

ALHOS COM BUGALHOS?

Rácios

Alhos não se *comparam* com bugalhos, bem o sabemos. Mas não há mal algum em *relacionar* uns com os outros, desde que tal relação faça sentido. É para isso que serve um tipo de valor próprio das estatísticas: os rácios.

O QUE É

Um **rácio** é uma divisão entre dois números que representam conjuntos distintos: camas e habitantes, quilómetros e horas, homens e mulheres, jovens e idosos, divórcios e casamentos. O resultado é um número relativo, só comparável a outros que sejam obtidos da mesma forma, isto é, que relaciona os mesmos conjuntos de forma igual. Por exemplo, o rácio jovens/idosos de Portugal é comparável com o rácio jovens/idosos de Espanha, mas não com o rácio idosos/jovens.

São os rácios que nos permitem desconfiar que há mais filas nas farmácias em Mafra do que em Portalegre. Como? Através da relação entre o número de habitantes e o número de farmácias de cada um destes concelhos. São dois elementos diferentes – de um lado pessoas, do outro, estabelecimentos. Mas podemos dividir um pelo outro, sem cerimónia, para se chegar a um valor indicativo que possibilite uma comparação:

	Mafra	**Portalegre**
Habitantes	80.723	23.720
Farmácias	16	14
Habitantes por farmácia	5045	1694

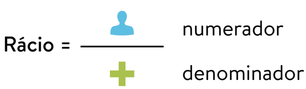
Fontes: INE/Pordata (dados de 2014)

Sem combinar alhos com bugalhos num valor relativo, chamado rácio, não seria possível concluir que uma farmácia em Mafra serve em média quase três vezes mais habitantes do que uma em Portalegre.

Num rácio, portanto, o que está do lado de cima da fracção (o numerador) é diferente do que está do lado de baixo (o denominador). Nalguns casos, pode haver alguma sobreposição dos universos considerados, com um conjunto parcialmente contido no outro. Por exemplo, uma vez que parte das despesas do Estado entra para o cálculo do PIB, no rácio entre estes dois indicadores, o numerador está parcialmente contido no denominador.

O normal, porém, é uma relação entre elementos distintos, como no caso das farmácias e da população:

Ou:



COMO SE LÊ

Genericamente, os rácios lêem-se da seguinte forma: *por cada tantas unidades de uma coisa há tantas unidades de outra coisa*.

Por exemplo, se a divisão do número de homens pelo número de mulheres numa população é de 0,91, diz-se:

Por cada mulher há 0,91 homens.

Se multiplicarmos este resultado por 100, temos um rácio de 91%, ou seja:

Por cada 100 mulheres, há 91 homens.

Uma boa parte dos rácios que figuram nas estatísticas aparecem sob a forma de "tantos casos" por cada "tantos habitantes". É uma relação intuitiva e eficaz para comparações, pois normaliza os dados de diferentes regiões, como se todas tivessem a mesma base populacional.

COMO SE CALCULA

Para calcular um rácio, basta dividir um valor pelo outro:

$$\frac{23.720 \text{ habitantes}}{14 \text{ farmácias}} \approx 1694 \text{ habitantes por farmácia}$$

Se invertêssemos os conjuntos, dividindo as farmácias pelos habitantes:

$$\frac{14 \text{ farmácias}}{23.720 \text{ habitantes}} \approx 0{,}00059 \text{ farmácias por habitante}$$

Como o resultado é inferior a um, com mais casas decimais do que a paciência algébrica do cidadão, multiplica-se o resultado por uma potência de 10 (10, 100, 1000, 10.000 ou mais), para tornar a leitura compreensível:

0,00059 x 10 = 0,0059 farmácias por cada 10 habitantes
0,00059 x 100 = 0,059 farmácias por cada 100 habitantes
0,00059 x 1000 = 0,59 farmácias por cada 1000 habitantes
0,00059 x 10.000 = 5,9 farmácias por cada 10.000 habitantes

ALERTA

Não confundir com proporções
Muito cuidado com os rácios apresentados em percentagem, pois facilmente os confundimos com proporções (ver capítulo sobre *proporções*). Um dos sintomas que os distingue é o facto de os rácios, ao contrário das proporções, poderem ser superiores a 100% (ou maiores do que um, no caso de o resultado não estar em percentagem).

Muitos rácios prestam-se a essa confusão. Quando dizemos que houve 70 divórcios por cada 100 casamentos em 2013, isto não significa que 70% dos matrimónios celebrados naquele ano colapsaram logo a seguir – o que seria uma proporção assustadora. Na verdade, o que este rácio quer dizer é que por cada 100 casamentos realizados em 2013, houve 70 divórcios também em 2013, certamente de uniões celebradas noutros anos.

Outro exemplo para salientar a diferença entre rácios e proporções, em que tanto o cálculo quanto a forma como o resultado é lido se afastam: se em cada quatro pessoas numa multidão houver um homem e três mulheres, o rácio homens/mulheres será de um para três. Ou seja:

$$\frac{1}{3} \approx 0{,}33 = 33\%$$

Certo: <u>*Por cada*</u> *100 mulheres, há 33 homens.*
Errado: *A proporção de homens é de 33%.*

Quanto à proporção de homens na multidão, ela é de um para cada quatro pessoas:

$$\frac{1}{4} = 0{,}25 = 25\%$$

Certo: *<u>Em cada</u> 100 pessoas há 25 homens.*
Certo: *A proporção de homens é de 25%.*

Faz toda a diferença dizer "x <u>por cada</u> y" (rácio) ou "x <u>em cada</u> y" (proporção).

Não há limites para os rácios, podemos dividir o que quisermos, sempre que haja uma lógica na relação que se está a estabelecer. As estatísticas estão cheias desses exemplos, como:

Indicador estatístico em rácio	**Numerador**	**Denominador**
PIB *per capita*	Euros	População
Densidade populacional	População	Área
Relação de masculinidade	Homens	Mulheres
Índice de envelhecimento	Idosos	Jovens
Divórcios por 100 casamentos	Divórcios	Casamentos
Consultas médicas por 1000 habitantes	Consultas	População

QUAL É O POTENCIAL?

Taxas

Pode parecer absurdo, mas embora ninguém se case em Odivelas (ver capítulo sobre *âmbito geográfico*), é neste concelho onde nascem em Portugal mais bebés *per capita*.

"Bebés *per capita*" não soa nada bem, por isso falemos em "nascimentos por cada 1000 habitantes". Pronto, eis uma taxa. Neste caso, a taxa bruta de natalidade.

O QUE É

Uma **taxa** é uma medida que relaciona um certo número de acontecimentos ocorridos com o universo que pode dar origem ao mesmo tipo de eventos. Na prática, é a divisão de um valor (por exemplo, número de nascimentos, mortes, empregados, falências, imigrantes, doenças) por uma espécie de universo *potencial*.

A ideia de "potencialidade" é o fio condutor do conceito estatístico de taxa. No caso da taxa bruta de natalidade, o que se faz é dividir o número de bebés nascidos num determinado local – um concelho, uma região, um país – pelo número total de habitantes desse território. Com isso, é possível avaliar o potencial procriador de uma população.

COMO SE LÊ

As taxas interpretam-se de forma semelhante aos rácios: *há tantas unidades de uma coisa por tantas unidades de outra coisa.*

Taxa de emprego = 51,2%
Por cada 100 habitantes com 15 ou mais anos, há 51 pessoas empregadas.

Taxa bruta de natalidade = 8,3‰ (‰ = por mil)
Por cada 1000 habitantes, há cerca de oito nascimentos num ano.

Uma taxa é, portanto, um caso particular de rácio (ver capítulo sobre *rácios)* em que se dividem dois conjuntos distintos, mas no qual o segundo (o denominador) está de certo modo relacionado com os acontecimentos do primeiro (o numerador), pois é o seu universo potencial. Ou seja, o numerador está potencialmente contido no denominador.

Sob o chapéu desta filosofia genérica, cabem escolhas diferentes para o indicador desse "valor potencial", umas melhores do que outras. A população geral, para retomar o caso da taxa bruta de natalidade, não é um bom indicador do universo possível de nascimentos, porque só as mulheres é que podem ter filhos, e numa determinada faixa de idade. A demografia criou, por isso, uma outra taxa, mais perfeita, que divide os nascimentos verificados num ano pelo número de mulheres em idade fértil. O denominador está, assim, muito próximo do número máximo de nascimentos que poderia ter ocorrido.

As taxas estão, por isso, associadas, de algum modo, a uma noção de probabilidade. Dizem-nos o que é maximamente expectável que ocorra: quantos nascimentos na população, quantos acidentes entre os empregados, quantos estudantes entre os jovens em idade escolar.

Uma das áreas onde as taxas são muito utilizadas é a da saúde, para medir, com base nas ocorrências observadas, a incidência de determinada doença.

Em regra, o universo potencial utilizado como denominador da taxa é muito maior do que os casos que de facto ocorrem. Daí que as taxas resultem em valores normalmente muito reduzidos e que, por isso, tenham de ser multiplicadas por 1000, 10.000 ou 100.000.

COMO SE CALCULA

Para calcular uma taxa, basta dividir um valor pelo outro.

Por exemplo, a incidência da tuberculose em Portugal em 2013:

$$\text{Taxa} = \frac{\text{Casos de tuberculose}}{\text{População}} = \frac{2393}{10.457.295} \approx 0{,}000229 \text{ casos por habitante}$$

Como o valor é muito pequeno, multiplicamo-lo por 100.000:

Taxa = 22,9 casos por 100.000 habitantes.

A fronteira que distingue o que devemos chamar rácios, taxas ou proporções pode ser difusa. Veja-se o exemplo dos licenciados. Em 2015, em Portugal, por cada 100 alunos matriculados pela primeira vez no ensino superior, houve 74 outros estudantes que o concluíram. É um rácio, pois os que se licenciam não pertencem ao grupo dos que acabam de entrar para a universidade, são conjuntos distintos.

Nesse mesmo ano, em cada 100 alunos que estavam matriculados no ensino superior, 22 concluíram-no. É uma proporção, pois os que se licenciam estão incluídos no conjunto de todos os alunos matriculados.

E ainda para 2015, podemos também dizer que por cada 1000 habitantes, sete pessoas obtiveram um diploma de ensino superior. É uma taxa, pois os residentes no país são aqueles que se poderiam hipoteticamente ter licenciado – ou seja, são o seu universo potencial.

Por isso, é preciso manter os olhos bem abertos de modo a compreender o que está a ser dividido pelo quê.

ALERTA

Taxas que não são taxas

As estatísticas estão repletas de indicadores com a designação de "taxa" mas que não o são. Por exemplo, a taxa de analfabetismo é, na verdade, uma proporção, dado que divide a população que não sabe ler e escrever pela população total – ou seja uma parte pelo todo no qual se insere.

Há outros exemplos de taxas que não são taxas. Entre eles, o da taxa de cobertura das importações pelas exportações. É um rácio, pois além de dividir dois conjuntos diferentes (o valor dos bens e serviços que entram no país pelo valor dos bens e serviços que saem), o denominador não constitui um potencial do numerador.

Já a taxa de inflação é um caso diferente: é uma taxa de variação (ver capítulo sobre *variações e taxas de variação*).

Há, porém, muitos indicadores que são taxas no verdadeiro sentido estatístico.

Exemplos:

- *Taxas de imigração e emigração*
- *Taxas de natalidade e mortalidade*
- *Taxa de escolarização*
- *Casos de tuberculose por 100 mil habitantes*
- *Doutoramentos por 100 mil habitantes*

MUDOU MUITO?

Variações e taxas de variação

Tudo sobe e desce nos noticiários. Os preços, as vendas, os lucros, a Bolsa, os turistas, os crimes, o desemprego, a temperatura, os votos. Há duas razões para isso: é natural que indicadores como estes estejam constantemente a mexer, e tudo o que mexe é notícia. O jornalismo gosta de variações.

Estas são intrínsecas às estatísticas, pois qualquer contagem ou medição periódica tem um só objetivo, que é o de perceber se algo está a mudar e como.

O QUE É

Variação é a medida da diferença entre dois valores.

Variação absoluta é a diferença que separa um valor de outro, obtida através de uma simples subtracção.

Variação relativa ou taxa de variação é quanto a variação absoluta representa em relação a um dos valores. Normalmente, é expressa em percentagem.

O título *Preço do gasóleo caiu 24 cêntimos entre 2012 e 2015* refere-se a uma variação absoluta. A mesma diferença expressa em termos relativos resultaria em: *Preço do gasóleo baixou 17% de 2012 para 2015.*

As variações frequentemente referem-se à evolução de um indicador em dois momentos no tempo, como no exemplo acima. No entanto, também podem ser calculadas entre dois valores observados no mesmo momento, mas em territórios diferentes – como países, regiões ou municípios – ou em universos e categorias distintos. Por exemplo: *Salários no Luxemburgo estavam 55% acima da média da UE em 2015*. Ou: *Em 2014, as mulheres ganhavam menos 20% do que os homens*.

COMO SE LÊ

Variações absolutas exprimem-se em:

- Unidades: *O preço do leite aumentou cinco cêntimos;*
- Pontos percentuais: *A taxa de juro subiu 1,8 pontos percentuais.*

Variações relativas exprimem-se em:

- Percentagens: *Crimes violentos diminuíram 30%;*
- Dobro, triplo, etc.: *População de Lisboa é o triplo da de Almada;*
- Fracções: *Valor das acções caiu um* terço.

Qualquer variação relativa depende umbilicalmente do valor de base que se escolhe para a comparação. Óbvio, não? Pois é, mas se o salário médio num país subiu de 500 para 1000 euros em 20 anos, para calcular a variação relativa nesse período podemos escolher como base tanto o valor mais recente como o mais antigo. Dessa forma, podemos dizer que agora ganha-se 100% mais do que há duas décadas ou que antes ganhava-se 50% menos do que agora. Ou seja, aqui o dobro equivale à metade – um intrigante mistério matemático. E no momento de se escolher uma manchete, 100% ou 50% fazem toda a diferença.

Por convenção e bom senso, quando estamos perante valores observados em dois momentos do tempo, o melhor para o cálculo da variação relativa é adoptar como base o valor inicial.

COMO SE CALCULA

Variação absoluta

Basta subtrair um valor (B) ao outro (A):

Variação absoluta = Valor B - Valor A

Variação relativa ou taxa de variação (em %):

1. Subtrair um valor (B) ao outro (A);
2. Dividir o resultado por um dos valores (normalmente pelo valor A);
3. Multiplicar por 100 e acrescentar o símbolo "%";
4. Arredondar (opcional).

$$\text{Variação relativa} = \frac{(\text{Valor B - Valor A})}{\text{Valor A}} \times 100$$

Exemplo: quanto variou o preço médio do gasóleo entre 2011 e 2012?

2011	1,37 €/litro
2012	1,45 €/litro

Fontes: DGEG/Pordata

Variação absoluta	Variação relativa (taxa de variação)
Subtrair o valor recente ao inicial: 1,45 − 1,37 = 0,08 *O preço aumentou oito cêntimos.*	1. Subtrair o valor recente ao inicial: 1,45 − 1,37 = 0,08 2. Dividir pelo valor inicial: $\frac{0,08}{1,37} \approx 0,058$ 3. Multiplicar por 100 e acrescentar "%": $0,058 \xrightarrow[\times 100 \text{ e } \text{“\%”}]{} 5,8\%$ 4. Arredondar (opcional): *6%* *O preço aumentou 6%.*

ALERTA

Atenção aos sinais

- Positivo = aumento
- Negativo = diminuição

Muitas vezes, temos de calcular a variação de valores que já estão expressos em percentagens, como a taxa de desemprego ou a proporção de votos num partido. Embora as contas sejam semelhantes, há uma regra essencial a não esquecer:

ALERTA

Percentagem *vs.* pontos percentuais

A variação absoluta entre duas percentagens é dada em **pontos percentuais** e não em percentagem. Se um partido obteve 20% de votos numa eleição e 30% noutra, a variação absoluta foi de **10 pontos percentuais** e não de 10%.

É sempre possível calcular a variação relativa entre duas percentagens, da mesma forma que se faz com qualquer outro par de valores. No exemplo acima, a subida de 20% para 30% representa uma variação relativa de 50% na proporção de votos naquele partido.

De igual modo, se a taxa de inflação passou de 1% para 2%, houve um aumento absoluto de um ponto percentual, mas um aumento relativo de 100%. Que número representa melhor a variação? Os 100% podem ser mais bombásticos, mas escondem o facto de estarmos a falar de apenas um ponto percentual. Se a inflação fosse de 30%, um aumento relativo daquela magnitude teria outro impacto e atiraria a taxa de inflação para 60%.

Variações relativas e absolutas enfrentam-se, assim, nessa luta pelo protagonismo, de modo que qualquer opção entre uma ou outra tem de ser aferida pelo bom senso.

Cada situação diferente ditará, é claro, a escolha do que deve ser salientado em primeiro plano. Quem vive do salário mínimo terá mais interesse em saber que pode contar com mais 25 euros no bolso ao final do mês (variação absoluta), do que ser informado de que houve uma subida de 5% (taxa de variação). Por outro lado, de pouco vale dizer que o PIB nacional engordou 6,5 mil milhões de euros – um valor impalpável para qualquer um –, mas sim que o país ficou 3,7% mais rico.

ALERTA

Contexto obrigatório

Uma variação relativa deve ser contextualizada com:

- O valor inicial a partir do qual foi calculada;
- A diferença absoluta dos valores em questão;
- O período em causa.

Exemplo: *O gasóleo subiu 6% em 2012. O preço do litro aumentou oito cêntimos, de 1,37 para 1,45 euros por litro.*

Ter em conta a diferença absoluta entre os valores é particularmente importante para afastar a sedução mediática de grandes taxas de variação que, na realidade, têm muito menor significado do que parecem.

ALERTA

Se os números forem pequenos...

As taxas de variação são muito sensíveis a números pequenos. Poucas unidades a mais podem resultar em taxas elevadas, mas apenas porque a base de partida é pequena.

A produção de energia solar em Portugal, por exemplo, aumentou cerca de 2500% entre 2007 e 2014. Não admira, pois na primeira data quase não havia painéis fotovoltaicos no país.

De igual forma, numa aldeia com um nascimento por ano, basta um bebé adicional para haver um aumento de 100% na natalidade.

O resultado de uma variação em percentagem traduz-se, também, em "número de vezes". Mas esse terreno, onde o português se encontra com a matemática, é pantanoso. Um aumento de 200% não resulta no dobro, mas no triplo do valor original – porque onde havia um, passou a haver mais dois, totalizando três.

ALERTA

De % ao "número de vezes"

- 50% de aumento: valor final é **uma vez e meia** o valor inicial
- 100% de aumento: valor final é **o dobro** do valor inicial
- 150% de aumento: valor final é **duas vezes e meia** o valor inicial
- 200% de aumento: valor final é **o triplo** do valor inicial
- 300% de aumento: valor final é **quatro vezes** o valor inicial

Passe a redundância, as taxas de variação podem variar ao longo do ano devido apenas a efeitos sazonais. No Algarve, o desemprego cai no Verão, quando o turismo está em alta, e volta a aumentar no Inverno. Portanto, o que faz sentido é comparar Verão com Verão, ou seja, dois períodos *homólogos*.

O QUE É

Taxa de variação homóloga é aquela que compara um valor com outro referente a um mesmo período, normalmente do ano anterior.

Variações, como referido no princípio deste capítulo, são o que há de mais sedutor nas estatísticas. São destacadas em comunicados, relatórios e estudos, referidas em depoimentos e entrevistas, ou calculadas directamente nas redacções.

Exemplos:

- *Taxa de inflação*
- *Taxa de crescimento do PIB*
- *Taxa de crescimento da população*

ABAIXO OU ACIMA?

Números-índice

Quem sabe com quanto Portugal tem contribuído para o problema das alterações climáticas? Cá vai: com o equivalente a 60.656.000 toneladas de dióxido de carbono atiradas para o ar em 1990, com 62.531.000 em 1991, com 66.928.000 em 1992, com 65.582.000 em... e assim por diante até aos dados mais recentes.

São números que não dizem absolutamente nada à maior parte dos mortais, a começar porque não permitem perceber com facilidade em que proporção as emissões de CO_2 estão a subir ou a descer. Para ajudar a sua digestão, as estatísticas dispõem de um instrumento muito prático: os números-índice.

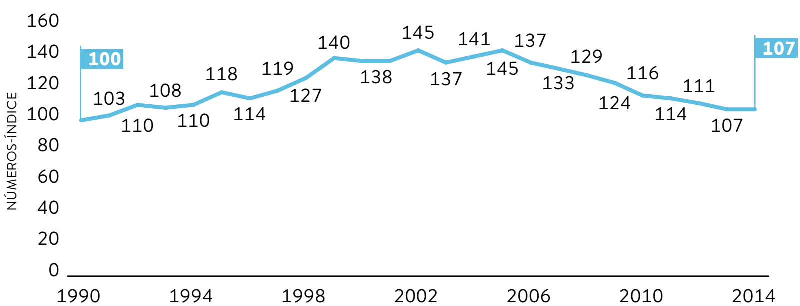
O QUE É

Número-índice é uma medida simplificada para se mostrar a variação entre um valor e um ponto de referência. Em vez de se utilizarem os números originais, assume-se normalmente que o valor de referência equivale a 100. Por exemplo, 10.000 equivalem a 100. Depois, recalcula-se o segundo valor em função daquela mesma base. Assim, se 10.000 equivalem a 100, então 15.000 equivalem a 150.

O ponto de referência é escolhido *a priori*. Pode, por exemplo, ser o valor num determinado momento ou num determinado território.

Vamos atribuir às cerca de 61 milhões de toneladas de 1990 o índice 100. Será o ponto de partida da comparação. Se 61 milhões equivalem a 100, então, utilizando uma regra de três simples, os quase 63 milhões de 1991 equivalem a 103 (ver capítulo sobre *regra de três simples*). Procedendo da mesma forma com toda a série de dados, chegamos a um gráfico simpático, que nos diz logo que as emissões subiram, depois desceram e quanto variaram proporcionalmente em cada ano em relação ao valor inicial.

Emissões de gases com efeito de estufa (GEE) em Portugal (1990=100)



Fontes: Eurostat/Pordata

COMO SE LÊ

O valor de base dos números-índice está sempre identificado nas estatísticas. Por exemplo:

- 1990=100: ou seja, 100 é o valor observado em 1990;
- UE=100: ou seja, 100 é o valor observado na União Europeia.

O valor do número-índice num dado momento indica o sentido da variação:

- Se for superior a 100, o valor é maior do que a base / houve um aumento;
- Se for inferior a 100, o valor é menor do que a base / houve uma diminuição.

Subtraindo-se 100 ao número-índice, obtém-se a taxa de variação percentual:

- Emissões de CO_2 em 2014 = 107 (1990=100)
  107 – 100 = 7. Ou seja: *houve um aumento de 7% face a 1990*.
- Emissões de óxidos de enxofre em 2014 = 15 (1995=100)
  15 – 100 = – 85. Ou seja: *houve uma redução de 85% face a 1995*.

ALERTA

Valores iguais, números-índice diferentes

Um mesmo valor num determinado momento pode resultar em números-índice diferentes, conforme a escolha da base.

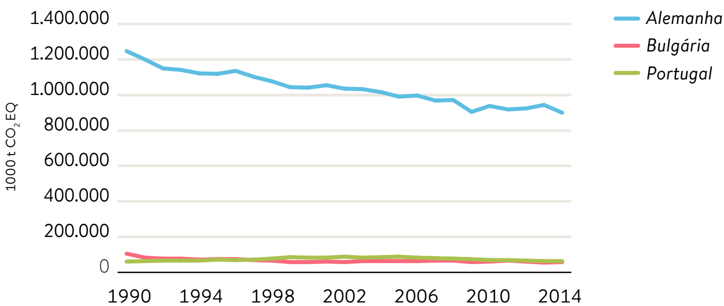
Exemplo: salário mínimo em Portugal.

Salário mínimo		Número-índice (2014 = 100)		Número-índice (2015 = 100)	
2014	485 €	2014	100	2014	96
2015	505 €	2015	104	2015	100
2016	530 €	2016	109	2016	105

Fontes: DGERT/Pordata

Números-índice são também muito úteis para comparar territórios diferentes. Se analisarmos em simultâneo as emissões de CO_2 de Portugal, Alemanha e Bulgária, o gráfico da sua evolução no tempo, em valores absolutos, dá-nos apenas uma leitura: a Alemanha tem muito mais emissões. Mas mal se percebe o que está a acontecer em Portugal e na Bulgária, achatados lá em baixo quase numa recta.

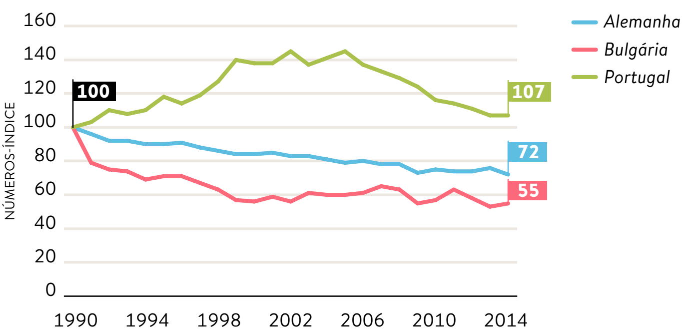
Emissões de gases com efeito de estufa



Fontes: Eurostat/Pordata

Mas se transformarmos as três séries em números-índice, todas a começarem em 100 no mesmo ano, o resultado é outro:

Emissões de gases com efeito de estufa (1990=100)



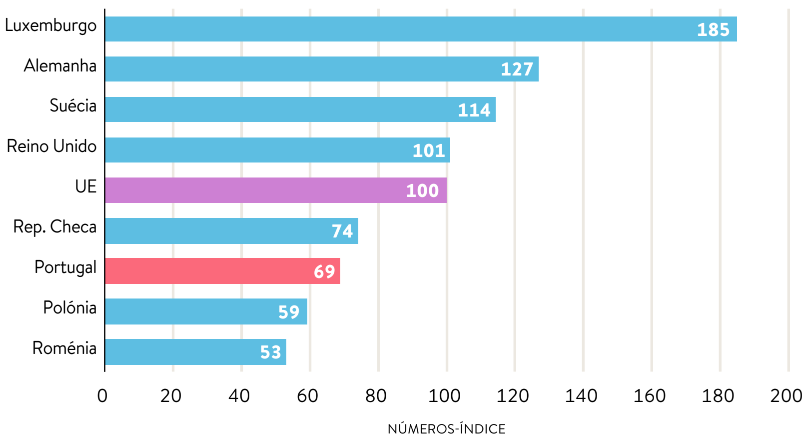
Fontes: Eurostat/Pordata

Agora temos a noção clara de como cada país se comportou individualmente e ficamos a saber que até 2014, em relação a 1990, as emissões tinham subido 7% em Portugal e caído na Alemanha e na Bulgária 28% e 45%, respectivamente.

Números-índice são muito utilizados para medir variações ao longo do tempo, utilizando um ano como ponto de partida (ver gráfico anterior). No entanto, também podem utilizar outra base, por exemplo, uma média ou um valor absoluto referente a um território – como acontece com frequência nas estatísticas europeias. Nestes casos, a média da União Europeia vale 100 e o desempenho dos países é avaliado em relação àquela baliza comum.

Eis uma comparação assim feita:

Produtividade por hora de trabalho em 2015 (UE=100)



Fontes: Eurostat/Pordata

Do gráfico acima, constata-se que a produtividade em Portugal em 2015 tinha o número-índice 69, portanto era 31% inferior à da União Europeia (UE). Já a da Alemanha, com o valor 127, era 27% superior à da UE.

ALERTA

Distância só à base

Os números-índice apenas dão uma noção imediata da distância de um valor em relação à base. Não se pode usar o mesmo tipo de raciocínio entre dois outros valores. Por exemplo, a produtividade do trabalho na Alemanha não é 58% (127 menos 69) maior do que em Portugal.

Qualquer pessoa que não se tenha esquecido da regra de três simples (ver capítulo sobre *regra de três simples*) pode converter um conjunto de dados de uma série em números-índice.

COMO SE CALCULA

Os números-índice podem calcular-se através de uma regra de três simples. Na prática:

1. Escolher o valor de referência (base);
2. Multiplicar por 100 o valor que se quer comparar;
3. Dividir o resultado pelo valor de referência;
4. Arredondar (opcional).

Por exemplo, o PIB *per capita* em Portugal, tomando como base o valor de 2010:

	Euros		Números-índice
2010	**17.018**	**17.018 equivale a 100**	**100**
2011	16.686	(16.686 x 100) / 17.018	98
2012	16.015	(16.015 x 100) / 17.018	94
2013	16.282	(16.282 x 100) / 17.018	96
2014	16.641	(16.641 x 100) / 17.018	98

Fontes: INE/BdP/Pordat*a*

ALERTA

Atenção aos nomes

Não confundir a designação de “índice” (ver capítulo sobre *índices*) com “números-índice”.

Muitos indicadores estatísticos são vulgarmente apresentados na forma final de números-índice. Acresce ainda que, embora o valor habitual de base seja 100, ele pode ser um outro múltiplo de 10, devendo essa base ser devidamente referenciada.

Exemplos

- *Índice de bem-estar (2004=100)*
- *Poder de compra* per capita *(Portugal=100)*
- *PSI-20 (05/01/1988=1000)*

EM QUE PÉ ESTAMOS?

Índices

Apesar dos altos e baixos, algumas medidas da condição anímica de Portugal revelam sinais de melhoria neste princípio de século. O índice de desenvolvimento humano do país subiu de 0,782 em 2000 para 0,830 em 2014. O índice global de bem-estar também aumentou de 100 para 118, entre 2004 e 2015, embora o índice das condições materiais de vida tenha caído de 100 para 88, sendo contrabalançado, porém, pelo índice de qualidade de vida, que chegou a 132.

Confuso? Bem-vindo ao mundo dos índices.

O QUE É

Um **índice** estatístico é uma medida que visa aferir o estado de conceitos complexos e difíceis de medir por um único indicador – como o bem-estar, o desenvolvimento, a sustentabilidade ou até a felicidade.

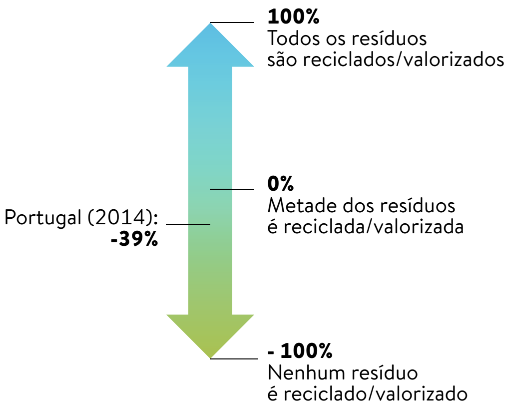
É um valor que agrega vários indicadores, que podem pesar de modo diferente para o resultado final, transformando-os numa medida única.

Índices compostos são os que incluem, no seu cálculo, apenas indicadores do mesmo tipo. Exemplo: o índice de preços no consumidor (IPC), que combina preços de vários produtos.

Índices compósitos: são os que agregam indicadores com unidades de medida diferente. Exemplo: o índice de bem-estar, que leva em conta indicadores tão díspares como taxa de desemprego, desigualdade na distribuição de rendimento, taxa de mortalidade infantil, índice de consumos culturais, índice de participação eleitoral, proporção de praias com bandeira azul e total de emissão de gases com efeito de estufa.

Os índices podem ter um valor mínimo e um valor máximo pré-definidos. E esses valores têm sempre um significado muito preciso. É o caso do índice de hierarquia na gestão de resíduos urbanos, que mede a que distância o país se encontra do ideal da economia circular, em que todo o lixo é reaproveitado. O seu valor máximo é 100% e significa que todos os resíduos urbanos são reciclados ou valorizados organicamente, isto é, transformados em fertilizante para a agricultura ou biogás. O valor mínimo é -100%, representando uma situação em que todo o lixo é depositado em aterros ou incinerado, e nada é reciclado. Simplificando:

Índice de hierarquia na gestão de resíduos



Fontes: INE/APA/Pordata

Este é um índice composto, pois o que está em causa é a quantidade de resíduos por cada tipo de destino, ou seja, indicadores do mesmo género.

Índices compósitos, por sua vez, são mais como uma receita de bolo em que entram ingredientes de diferentes tipos. Por vezes, estes índices desmultiplicam-se em várias dimensões. Por exemplo, o índice de bem-estar desdobra-se em dois grandes grupos: qualidade de vida e condições materiais de vida, cada um deles com vários subníveis.

Já o índice de desenvolvimento humano, divulgado anualmente pelo Programa das Nações Unidas para o Desenvolvimento, combina esperança de vida, educação e rendimento. São três dimensões expressas por indicadores precisos. O seu valor situa-se entre 0 (desenvolvimento mínimo) e 1 (desenvolvimento máximo). Em 2014, a Noruega estava no topo do *ranking* mundial, com 0,944, e o Níger na cauda, com 0,348.

COMO SE LÊ

Na leitura de um índice, o valor em si não é suficiente. É fundamental referir o conceito que o valor representa e os seus limites, caso existam.

Exemplo: *Portugal tinha um índice de desenvolvimento humano de 0,830 e a Noruega de 0,944, numa escala de 0 (mínimo) a 1 (máximo).*

Os índices podem ser apresentados em:

- Percentagem: *proporção de poder de compra da região Norte = 32%;*
- Número absoluto: *índice de desenvolvimento humano = 0,944;*
- Número-índice: *índice de bem-estar (2004=100) = 118.*

ALERTA

Índices que não são índices
Há inúmeros indicadores chamados de "índices" nas estatísticas, mas que são apenas meras divisões entre dois valores. Por exemplo, os índices de longevidade e de envelhecimento: o primeiro é uma simples proporção, e o segundo, um rácio (ver capítulos sobre *proporções* e *rácios*).

O QUE SERIA O NORMAL?

Médias

Em Nisa, as noites de cinema não passam despercebidas. Em 2015, foi o concelho do país com mais espectadores na sala por sessão: 173, em média. É um valor que representa quase oito vezes a média nacional, que foi de 23.

Parabéns a Nisa, mas, na realidade, aquele número pode não representar uma sessão típica no cine-teatro do concelho. As médias são uma espécie de varinha mágica capaz de sintetizar dezenas, centenas, milhares, milhões de dados num único valor. Por isso, são extremamente populares no dia-a-dia e, como corolário, nas notícias. Mas têm das suas e podem pregar partidas.

O QUE É

Média é uma medida sobre o ponto de equilíbrio numa distribuição de valores, indicando uma tendência central, potencialmente representativa.

Existem vários tipos de média. A mais frequente é a média aritmética, obtida através da soma de todos os valores e dividida pelo número total de casos observados.

O advérbio "potencialmente" é aqui importante, pois as médias nem sempre são suficientes para retratar os dados que as originam.

ALERTA

Médias iguais, realidades diferentes

Duas médias iguais podem representar distribuições diferentes de valores, ou seja, realidades muito distintas.

Imaginem que em quatro sessões de cinema, a média de espectadores tenha sido 20. As duas realidades abaixo são igualmente possíveis para esse mesmo resultado:

Sessões	1	2	3	4	Média
Espectadores	20	20	20	20	$\frac{20 + 20 + 20 + 20}{4} = 20$
	2	55	9	14	$\frac{2 + 55 + 9 + 14}{4} = 20$

No segundo caso, nenhum valor está sequer próximo de 20. Por isso, a média não é, ali, uma boa síntese da forma como os 80 espectadores que frequentaram aquela sala estiveram distribuídos pelas quatro sessões.

No jargão estatístico, diz-se que aqueles dados têm uma *dispersão* muito elevada. Em quatro casos observados, é fácil ver isso. No entanto, se fossem centenas ou milhares de observações, seria preciso calcular o desvio-padrão da distribuição – quanto menor, mais próximos da média estão os valores.

No cine-teatro de Nisa, houve 11 sessões em 2015. Mentalmente imaginamos que há 173 pessoas a assistirem ao filme de cada vez. Mas a distribuição poderá ter sido outra. Por exemplo:

Sessões	1	2	3	4	5	6	7	8	9	10	11
Espectadores	23	50	50	50	50	50	180	300	350	400	400

A média aritmética também é 173. Mas:

- Pelo menos metade das sessões teve 50 ou menos espectadores;
- A lotação mais comum foi de 50 espectadores.

Estas conclusões adicionais exemplificam dois conceitos que andam de mãos dadas com a média: a mediana e a moda.

O QUE É

Mediana é o valor que está exatamente no meio de uma distribuição de dados ordenados. Isto é, metade dos valores observados é igual ou superior à mediana e a outra metade é igual ou inferior. No exemplo acima, a mediana seria 50.

Moda é o valor mais frequente da distribuição, ou seja, aquele que aparece mais vezes. Ainda no exemplo acima, a moda também seria 50.

Entre os 102 concelhos do país que tinham cinema em 2015, a média de ecrãs é de cinco por município. No entanto, a mediana é dois – ou seja, pelo menos metade dos concelhos tem duas ou menos salas. E a moda é um, isto é, o mais comum é haver uma só sala.

A moda e a mediana são complementos da média. A mediana, em especial, merece sempre ser vista em conjunto com a média. Se estão próximas, é porque a média é um farol suficiente da tendência central. Ou seja, há uma distribuição equilibrada dos valores acima e abaixo da média. Se estão distantes, vale a pena investigar os dados, pois há alguma coisa que está a puxar a média para um lado ou para o outro – como um "peso pesado" ou um "peso leve" num dos extremos.

Por exemplo: num determinado bairro vendem-se quatro apartamentos por 70.000, 100.000, 110.000 e 120.000 euros. O mercado está pronto para fechar o ano com um preço médio de 100.000 euros por imóvel. Mas em Dezembro, um palacete é comercializado por 2.000.000 euros. A média sobe para 480.000 euros, um valor que está longe de representar a maior parte dos negócios que ali se efectuaram.

ALERTA

Valores que distorcem

Valores extremos numa série de dados podem distorcer uma média:

- Se muito altos em relação aos demais, aumentam a média;
- Se muito baixos em relação aos demais, diminuem a média.

Este risco é tanto maior quanto menor for a série de observações que está a ser considerada. Se forem poucos os elementos, um valor extremo torna a média sem sentido. Entre milhares de dados, pode passar praticamente despercebido.

COMO SE CALCULA

Média: somar todos os valores e dividir pelo número de casos observados:

2, 31, 24, 8, 15 (cinco casos observados)

$$\text{Média} = \frac{2 + 31 + 24 + 8 + 15}{5} = 16$$

Muitas vezes, os valores a dividir já vêm prontos (números sobre o total de espectadores e sessões de cinema, por exemplo) e basta dividir um pelo outro.

Mediana: ordenar os valores e encontrar o que está no meio da distribuição.

- Se o número de elementos for **ímpar**, é o valor que está de facto no meio:

  2, 8, **15**, 24, 31

  Mediana = 15.
- Se o número de elementos for **par**, é a média dos dois centrais:

  2, 8, **15**, **19**, 24, 31

$$\text{Mediana} = \frac{15 + 19}{2} = 17$$

ALERTA

Médias de médias?

Não se devem fazer médias de médias relativas a universos de diferentes dimensões, pois cada uma tem um "peso" distinto.

É muito fácil incorrer neste erro. Por exemplo: qual terá sido a média de espectadores por sessão de cinema em Lisboa e Coimbra em 2015, considerando-se os dois concelhos ao mesmo tempo?

	Sessões	Espectadores	Média
Lisboa	109.672	2.820.559	26
Coimbra	23.500	472.475	20
Total	**133.172**	**3.293.034**	**???**

Fontes: ICA/Pordata

Se cairmos na tentação de fazer a média das médias de cada concelho, temos:

$$\frac{26 + 20}{2} = 23$$

Certo? Errado. A média, por definição, deverá ser a soma de todos os espectadores, dividida pelo número total de sessões, ou seja:

$$\frac{3.293.034}{133.172} \approx 25$$

Vale a pena notar que a média dos dois concelhos juntos está mais próxima da de Lisboa, do que da de Coimbra. E faz todo o sentido, pois há muito mais sessões de cinema na capital.

Há centenas de indicadores estatísticos apresentados em médias.

Exemplos:

- *Remuneração média mensal*
- *Temperatura média do ar*
- *Dimensão média das famílias*
- *Esperança média de vida à nascença*
- *Preço médio de venda da gasolina*


QUANTO

QUANTO?

Números e escalas

59.032.120.694. Aí está um número assustador. Pronunciá-lo é um desafio à musculatura bucal: cinquenta e nove milhares de milhão, trinta e dois milhões, cento e vinte mil, seiscentos e noventa e quatro.

Ninguém diz isso. E, na realidade, não é preciso. Grandezas numéricas como esta – neste caso o valor das importações de bens em Portugal em 2014 – são candidatas inevitáveis à simplificação. É muito mais fácil dizer que as importações de bens somaram cerca de 59 mil milhões de euros. O valor é praticamente o mesmo – salvo os arredondamentos – apenas expresso numa *escala* diferente.

O QUE É

A **escala** de um valor é a dimensão em que este é apresentado. Na prática, indica quantas unidades estão a ser contadas de cada vez – se são dezenas, centenas, milhares, milhões, etc.

Eis o mesmo valor das importações de bens em várias escalas:

- 59.032.120.694 euros (escala = unidades)
- 59.032.121 mil euros (escala = milhares)
- 59.032 milhões de euros (escala = milhões)
- 59 mil milhões de euros (escala = milhares de milhão)
- 0,059 biliões de euros (escala = biliões)

O número, a escala e a unidade de medida (ver capítulo *unidades de medida*) formam uma *troika* inseparável. É essencial identificar cada um dos três elementos quando se está perante um dado estatístico. Ou seja:

- Que número temos à frente?
- Em que escala está representado?
- Qual é a sua unidade de medida?

Vejam-se as respostas para duas formas de representação do valor das importações:

	Número	**Escala**	**Unidade de medida**
59.032.120.694	Cinquenta e nove milhares de milhão, trinta e dois milhões, etc.	Unidades	Euros
59 mil milhões	Cinquenta e nove	Mil milhões	Euros

Compreender um número significa começar por saber lê-lo. Qualquer jornalista – em especial da área económica – já terá tropeçado em trambolhos numéricos como 59.032.120.694, ficando a contar o número de algarismos para tentar perceber que grandeza é aquela.

COMO SE LÊ

Para verificar se são milhares, milhões, milhares de milhão, biliões, etc., **contar os algarismos após o primeiro ponto a partir da esquerda** (se o número tiver uma vírgula, não considerar os algarismos da parte decimal).

N.º de algarismos	Número	Lê-se
3	1.000	Mil
6	1.000.000	Um milhão
9	1.000.000.000	Mil milhões
12	1.000.000.000.000	Um bilião
15	1.000.000.000.000.000	Mil biliões
18	1.000.000.000.000.000.000	Um trilião

Como 59.032.120.694 tem nove algarismos após o primeiro ponto, então estamos a falar de um número na ordem dos milhares de milhão.

No caso dos números inferiores a um, para saber se são décimas, centésimas, milésimas, milionésimas, etc., **contar as casas decimais**:

Casas decimais	Número	Lê-se
1	0,1	Uma décima
2	0,01	Uma centésima
3	0,001	Uma milésima
4	0,0001	Uma décima de milésima
5	0,00001	Uma centésima de milésima
6	0,000001	Uma milionésima

ALERTA

Biliões ou mil milhões?

Nalguns países, incluindo Portugal, um bilião equivale a um milhão de milhões. Noutros, como os EUA e o Brasil, um bilião equivale a mil milhões.

Número	Portugal	EUA
1.000.000.000	Mil milhões	Um bilião
1.000.000.000.000	Um bilião	Um trilião

Vírgula ou ponto?

Nos países anglo-saxónicos e outros, como Índia, China e Japão, a vírgula é utilizada para separar o milhar; e o ponto, para separar as décimas. Em Portugal, tal como em muitos outros países, a vírgula separa as décimas e o ponto os milhares.

Portugal	EUA
10,2	10.2
10.200	10,200

Na apresentação de dados estatísticos, a mudança de escalas é essencial por uma razão bastante pragmática: há números tão extensos que é preciso simplificá-los para caberem nas tabelas e serem mais facilmente apreendidos.

COMO SE CALCULA

Para **simplificar valores grandes**, reduzindo o número de algarismos a apresentar, caminha-se com a vírgula **para a esquerda** e aumenta--se a escala (1000 vezes por cada três algarismos). Habitualmente arredonda-se o resultado (ver capítulo sobre *arredondamentos*).

Por exemplo, a população de Portugal em 2015:

vírgula retrocede três casas — *vírgula retrocede três casas*

10.358.076 (escala = unidades) → 10.358,076 mil (escala = milhares) → 10,358076 milhões (escala = milhões)

escala aumenta — *escala aumenta*

A população era de 10,4 milhões de habitantes.

Para **"desdobrar" valores**, o caminho é o inverso: avança-se com a vírgula **para a direita**, acrescentam-se zeros (se necessário), e **reduz-se** a escala (1000 vezes por cada três algarismos):

vírgula avança três casas ou acrescentam-se zeros — *vírgula avança três casas ou acrescentam-se zeros*

10,4 milhões (escala = milhões) → 10.400 mil (escala = milhares) → 10.400.000 (escala = unidades)

escala diminui — *escala diminui*

ALERTA

Desdobramentos são aproximações

É preciso ter em atenção que quando desdobramos um número nunca chegamos ao seu valor original. Por exemplo, se vemos escrito algures que a população somava 10,4 milhões de pessoas em 2015, não sabemos que, na realidade, eram 10.358.076. E ao desdobrar aquele número, o que obtemos são 10.400.000 pessoas.

Por vezes, também somos confrontados com números muito pequenos, da ordem das centésimas, milésimas, décimas de milésimas, de difícil leitura ou compreensão. Por exemplo, em 2014 foram lançadas para a atmosfera, em Portugal, 11.424 toneladas de óxido nitroso – um dos gases que estão a aquecer o planeta. Se quisermos calcular as emissões deste gás *per capita*, temos de dividir o valor pelo da população em 2014, que correspondia a 10.401.062 pessoas.

O resultado não diz nada a ninguém – aproximadamente 0,001098 toneladas por habitante – e merece ser simplificado.

COMO SE CALCULA

Para **simplificar números muito pequenos**, menores do que um, anda-se com a vírgula para a direita, reduz-se a escala (1000 vezes por cada três algarismos) e arredonda-se o resultado:

vírgula avança três casas

0,001098 toneladas (escala = toneladas) → 1,098 milésimas de tonelada (escala = milésimas de tonelada)

escala diminui

Arredondando-se o resultado para uma casa decimal, temos 1,1 milésimas de tonelada, que mais não são do que 1,1 quilos (1 tonelada = 1000 quilos).

A escala de uma série estatística é normalmente apresentada com o dado ou referida na tabela onde este figura.

Escala

Euro - Milhares

Anos	PIB
2008	178.872.582
2009	175.448.190
2010	179.929.812
2011	176.166.578
2012	138.397.969
2013	170.269.327

Fontes: INE/BdP/Pordata

Nalguns casos, as escalas são apresentadas através da sua notação científica, em potências de 10:

Escala	Notação científica
Milhares	10^3 ou 10^3
Milhões	10^6 ou 10^6
Mil milhões	10^9 ou 10^9
Biliões	10^{12} ou 10^12

A referência à escala também é uma componente da metainformação associada aos quadros estatísticos (ver capítulo sobre *metainformação*).

DE QUÊ?

Unidades de medida

Quando se olha para as estatísticas sobre o uso de telemóveis em Portugal, constata-se que o número de unidades ultrapassou o número de habitantes do país em 2004. E em 2011, já havia quase duas unidades por habitante.

Mas unidades de quê? Bem, até 2009, o que se contava era o número de assinaturas de serviço móvel. Desde 2010, o que se observa é o número de equipamentos. Assinaturas e telemóveis são as *unidades de medida* desta série estatística.

O QUE É

Unidade de medida é a referência que indica o que está a ser contado. Em 20 milhões de equipamentos de telemóvel, a unidade de medida é equipamentos de telemóvel, assim como em 200 hectares, a unidade de medida é hectares, em 179 milhões de euros, são euros, ou em 500 toneladas, são toneladas.

A unidade de medida pode ser expressa através de um símbolo (€, °C), uma abreviatura (km, kWh), uma palavra (freguesias, edifícios), ou uma combinação entre unidades (€/pessoa, habitantes/km^2).

Muitas unidades de medida são óbvias. A população, os estudantes, os médicos, os juízes, os desportistas contam-se em indivíduos. Os registos notariais contam-se em registos, as patentes de invenções em patentes, os internamentos hospitalares em internamentos, os livros das bibliotecas em livros, as empresas em empresas.

Mas, às vezes, o que parece óbvio afinal não é. Uma leitura apressada dos dados oficiais levaria qualquer um a concluir que em 2014 havia cerca de 38 mil investigadores em Portugal. Uma análise mais atenta mostra, no entanto, que a contagem não se refere aos investigadores em si, mas ao equivalente a investigadores a tempo integral (ETI). Por exemplo, dois profissionais que dedicam 50% do seu tempo a actividades de investigação equivalem a um investigador a 100% e, por isso, são contabilizados como um ETI.

A unidade de medida, neste caso, não é indivíduos mas sim equivalente a tempo integral – algo que não é compreensível de imediato. Muitas outras unidades de medida, como toneladas de CO_2 equivalente, quilowatts-hora ou gigajoules, soam igualmente estranhas aos ouvidos menos treinados, bem como à vista, dadas as suas siglas (tCO_2eq, kWh, GJ).

ALERTA

Atenção às quebras de série
As unidades de medida de uma série estatística podem ser alteradas num determinado momento. Quando isto ocorre, há uma *quebra de série* e os valores antes e depois desse momento não são imediatamente comparáveis (ver capítulo sobre *quebras de série*).

No caso das estatísticas sobre os telemóveis, conforme referido no exemplo anterior, o que se contabilizava até 2009 era o número de assinaturas. A partir de 2010, a unidade de medida passou a ser os equipamentos activos. É fácil compreender por que o antes e o depois são coisas diferentes:

- Em 2009, havia 16,1 milhões de assinaturas de telemóvel;
- Em 2010, havia 19,7 milhões de telemóveis activos.

Ambos os números podem estar na mesma série de dados, em anos consecutivos, mas não são directamente comparáveis. E, por isso, o ano em que a alteração ocorreu deve ser assinalado com uma quebra de série.

ALERTA

Uma unidade, diferentes abordagens
Há unidades de medida que podem ser apresentadas com diferentes abordagens. Por exemplo: euros a *preços correntes* ou *a preços constantes* (ver capítulo sobre *valores nominais e reais, preços correntes e constantes*).

A unidade de medida nem sempre diz tudo. Em casos menos óbvios, pode haver um conceito associado que esclarece o seu significado (ver capítulo sobre *conceitos*). Na evolução do número de hospitais em Portugal, a unidade de medida é sempre “estabelecimentos de saúde”. Mas, até 2009, um centro hospitalar, mesmo tendo vários hospitais, era por vezes contado apenas como um estabelecimento. A partir do ano seguinte, todas as unidades locais passaram a fazer parte do cálculo. E subitamente, em 2010, as estatísticas registaram o aparecimento de 43 novos hospitais no país – uma realidade quase impossível de acontecer.

Não se deve negligenciar a necessidade de olhar sempre para a unidade de medida de um valor. Regra geral, elas vêm associadas aos dados e também na metainformação respectiva (ver capítulo sobre *metainformação*).

Unidade de medida

Euro - Milhares

Anos	PIB
2008	178.872.582
2009	175.448.190
2010	179.929.812
2011	176.166.578
2012	138.397.969
2013	170.269.327

Fontes: INE/BdP/Pordata

Exemplos:

- *PIB: euros (€)*
- *População: indivíduos*
- *Densidade populacional: indivíduos por* km^2
- *Água: metros cúbicos* (m^3)
- *Chuva: milímetros (mm)*
- *Área ardida: hectares (ha)*
- *Resíduos: toneladas (t)*
- *Preço da eletricidade: euros por quilowatt-hora (€/kWh)*
- *Gases com efeito de estufa: toneladas de* CO_2 *equivalente* (tCO_2eq)

QUANTO VALE HOJE?

Valores nominais e reais, preços correntes e constantes

É inacreditável pensar que o salário mínimo nacional em Portugal subiu 2400% em 40 anos. Mas é verdade. O seu valor, em notas contadinhas uma a uma, aumentou de 20 euros (4000 escudos) em 1975 para 505 euros em 2015.

É claro que não podemos soltar foguetes com estas contas. Afinal, a inflação comeu o valor do dinheiro. Com um euro, hoje, compra-se incomparavelmente menos do que com o mesmo valor em escudos há 40 anos.

São duas formas de olhar para o dinheiro: através do seu *valor nominal* e do seu *valor real*.

O QUE É

Valor nominal do dinheiro é o montante tal como o conhecemos num momento. Diz-se que é o dinheiro a **preços correntes**.

Valor real do dinheiro é o valor nominal ajustado de modo a anular o efeito da inflação. Diz-se que é o dinheiro a **preços constantes**.

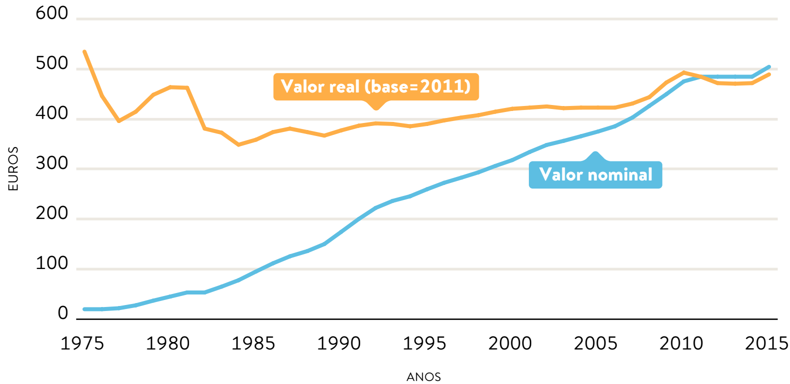
O conceito de valor nominal é muito fácil de entender. Uma nota de 20 euros terá sempre o valor nominal de 20 euros. É o que está lá, ponto final.

Já o valor real é algo um pouco mais complexo. Imaginemos, por hipótese, que em 2000 um frango assado custava cinco euros. Quem ganhasse 20 euros poderia comprar quatro. Em 2010, porém, as churrasqueiras já estavam a cobrar 10 euros por unidade. Para adquirir os mesmos quatro frangos, seriam então necessários 40 euros. Ou seja, se o custo do frango fosse representativo dos preços do consumo, 20 euros, em 2000, teriam um valor real de 40 euros *a preços de 2010*.

Os valores reais pressupõem a escolha de um ano como base, cujos preços servem de baliza para a comparação do poder de compra do dinheiro em diferentes momentos. No ano base, o valor real é igual ao valor nominal. Para os outros anos, todos os preços são recalculados.

Considerando o salário mínimo nacional, juntando num gráfico os seus valores nominais e reais desde 1975, e tomando-se como referência os preços de 2011 (base=2011), vê-se logo do que estamos a falar.

Salário mínimo nacional em valores nominais e reais



Fontes: DGERT/Pordata

A linha azul do gráfico mostra o valor que aparecia nos recibos de vencimento de quem ganhava o salário mínimo. A linha laranja representa o quanto aquelas quantias valeriam em 2011, ou seja, a preços de 2011.

Neste caso, os valores nominais estão quase sempre a subir. São números assim que costumamos ver nas notícias. Por exemplo, entre 2014 e 2015 houve um acréscimo de 20 euros (de 485 para 505 euros), representando 4% de incremento.

Já os valores reais sobem e descem, conforme a inflação tenha um efeito mais ou menos corrosivo sobre o aumento nominal. Resumo da história: 1975 foi o ano de ouro do salário mínimo nacional. Eram apenas 20 euros (ou seja, cerca de 4000 escudos), mas que, aos preços de 2011, valeriam 535 euros. Já o salário mínimo em 2015, cujo valor nominal era de 505 euros, equivaleria a 489 euros a preços de 2011. Ou seja, em termos reais, o salário mínimo de 2015 vale menos do que o de 1975.

COMO SE LÊ

A leitura do valor real inclui sempre a indicação do ano cujos preços servem de referência. Exemplos:

- *O salário mínimo de 1975 valeria 535 euros a preços de 2011.*
- *Para ter o mesmo poder de compra que tinha em 1975, o salário mínimo deveria ser de 535 euros em 2011.*

Para se transformarem os valores a preços correntes (nominais) em valores a preços constantes (reais) utilizam-se *deflatores*. São séries de valores, um para cada ano, que permitem ajustar as estatísticas monetárias aos efeitos da inflação. Existem vários deflatores, pois os níveis de inflação ou de variação de preços diferem consoante o tipo de produto que estejamos a considerar. Contudo, os mais usados são o índice de preços no consumidor (IPC) e o deflator do PIB.

Os deflatores aparecem num formato semelhante a este:

2009	0,951409
2010	0,964749
2011	1,000000
2012	1,027728
2013	1,030544
2014	1,027687
2015	1,032692

Neste caso, são deflatores do IPC entre 2009 e 2015, calculados a partir de dados do Instituto Nacional de Estatística. A base aqui é 2011, por isso o valor nesse ano é um (o valor real é igual ao valor nominal). Com estes deflatores, os cálculos serão feitos tendo como referência os preços de 2011.

COMO SE CALCULA

Para transformar valores a **preços correntes (nominais)** em valores a **preços constantes (reais)**, basta:

1. Escolher um deflator;
2. Dividir o valor nominal pelo valor do deflator no mesmo ano;
3. Mencionar o ano base no resultado (valores *a preços constantes de...*).

Por exemplo: quanto vale realmente o salário mínimo de 2009, de 2011 e de 2015, a preços de 2011?

	2009	**2011**	**2015**
Valor a preços correntes	450,00 €	485,00 €	505,00 €
Deflator do IPC (base=2011)	0,951409	1	1,032692
Valor a preços constantes de 2011	$\frac{450,00}{0,951409} \approx 473€$	$\frac{485,00}{1} = 485€$	$\frac{505,00}{1,032692} \approx 489€$

A resposta à pergunta clássica "Quanto vale isso hoje?" esbarra num problema: normalmente, os deflatores publicados não se referem ao ano mais recente. Por exemplo, se os deflatores do IPC têm a base 2011, só podemos calcular o valor real do salário mínimo de 2009 e 2015 em relação aos preços daquele ano – como fizemos acima. Para saber o valor do dinheiro em relação aos preços de uma data mais próxima, é preciso fazer alguns cálculos adicionais.

Assim, para dizer quanto o salário mínimo de 2009 valeria em 2015, temos de ajustar os deflatores para ambos os anos e calcular o valor real, da seguinte forma:

1. Atribuir ao deflator do salário mínimo de 2015 o valor um (valor real = valor nominal);
2. Recalcular o valor do deflator do salário mínimo de 2009 (ver capítulo sobre *regra de três simples*):

Ano	**Deflator (base=2011)**	**Deflator (base=2015)**
2009	0,951409	x
2015	1,032692	1

$$\frac{0{,}951409}{1{,}032692} = \frac{x}{1}$$

$$1{,}032692\, x = 1 \times 0{,}951409$$

$$x = \frac{0{,}951409}{1{,}032692} \approx 0{,}921290$$

3. Dividir o valor do salário mínimo de 2009 pelo novo deflator:

$$\frac{450{,}00}{0{,}921290} \approx 488 \text{ euros}$$

O salário mínimo de 2009 valeria 488 euros em 2015.

ALERTA

Que tipo de valor escolher?

Valores a preços correntes e a preços constantes são complementares, e há muitas estatísticas em euros para as quais interessa conhecer ambos. Exemplos: o PIB, salários, despesas, investimento.

COMO, ONDE E QUANDO

DADOS SOBRE OS DADOS?

Metainformação

Nem sempre um quadro estatístico diz tudo. Olhamos para os números e uma voz inaudível diz-nos que falta ali alguma coisa, algo que explique que tipo de valor é aquele, quem o recolheu e como, o que de facto está a ser contabilizado, a que período específico se referem os dados, qual a data da última actualização, etc.

É nessa altura que se deve recorrer a uma parte integrante de qualquer série estatística, mas muitas vezes negligenciada: a *metainformação*.

O QUE É

Metainformação significa, literalmente, informação sobre a informação. Nas estatísticas, a metainformação é uma espécie de ficha técnica que descreve os dados e o modo como foram produzidos.

Não há uma regra uniforme para a apresentação da metainformação. Mas, tal como as notícias, ela procura responder a perguntas essenciais sobre um conjunto de dados estatísticos: o quê, quando, onde, quanto, como e quem.

O QUÊ?
- Designação da série;
- Tipo de valor utilizado.

QUANTO?
- Unidade de medida;
- Escala.

QUEM?
- Entidades responsáveis pelos dados.

Mi
Metainformação

ONDE?
- Âmbito geográfico dos dados.

COMO?
- Tipo de operação estatística;
- Conceitos;
- Fórmulas de cálculo;
- Quebras de série e notas.

QUANDO?
- Periodicidade de recolha dos dados;
- Período/momento de referência dos dados;
- Data da última actualização.

ALERTA

Olhar sempre

É boa prática olhar sempre para a metainformação dos quadros estatísticos, mesmo para os que não inspiram quaisquer dúvidas. Por exemplo, uma estimativa anual da população parece apenas isso: a população num determinado ano. Mas há resultados diferentes para um mesmo ano e nem sempre o que os distingue está claramente expresso nos quadros onde se apresentam.

Há alguns casos de estatísticas muito comuns nas notícias que estão talhadas para equívocos de interpretação. As taxas de emprego e de desemprego estão nesta categoria (ver capítulo sobre *conceitos*). E é na metainformação que iremos encontrar o que interessa para as compreender.

ALERTA

Se não houver, desconfie

Um dado estatístico sem metainformação ou qualquer descrição do que é, bem como sobre os seus processos de produção, tem de ser investigado antes de utilizado numa notícia. Nem todos os números inspiram o mesmo nível de confiança (ver capítulo sobre *fontes e operações de recolha*).

A metainformação está por vezes identificada nos *sites* de estatísticas pelo símbolo *Mi*. Pode ser encontrada:

- Associada aos quadros estatísticos publicados *online;*
- Em notas técnicas das publicações estatísticas.

QUEM DISSE?

Fontes e operações de recolha

***Em 2014, as famílias em Portugal poupavam cerca de 6% do seu rendimento disponível.* Muito bem, mas quem disse isso? E como sabemos que esta informação é fiável?**

Com as redacções bombardeadas com números todos os dias, há duas perguntas elementares que têm de ser feitas quando se trabalha com dados estatísticos:

- Quem os produziu?
- Como foram obtidos?

O QUE É

A **fonte** de um conjunto de dados estatísticos é quem os recolhe, produz e/ou divulga. É a entidade responsável pelos números que ali estão.

- **Fontes primárias** são as que divulgam informação estatística que elas próprias produziram. Por exemplo, o INE com os dados do abandono escolar precoce ou o Instituto Português do Mar e da Atmosfera (IPMA) com os dados do clima.
- **Fontes secundárias** são as que divulgam informação produzida por outros. Por exemplo, a Pordata, ao publicar estatísticas produzidas por fontes oficiais com competências nas respectivas áreas.

ALERTA

Mencionar as duas

Quando há dados divulgados por uma fonte secundária, deve-se sempre mencionar a primária também, ou seja, a que produziu a informação. Por exemplo, se a Pordata publica dados do abandono escolar precoce produzidos pelo INE, a menção deve ser:

Fontes: INE/Pordata

Ninguém mais do que os jornalistas sabe da relevância de se identificar a fonte de uma informação e de se avaliar até que ponto ela é credível. Nas estatísticas, algumas pistas ajudam nesta tarefa. Uma delas é o adjectivo *oficial*, termo um pouco banalizado. O conceito de *estatísticas oficiais* encontra-se definido em Portugal e remete para as fontes que integram o Sistema Estatístico Nacional (SEN). São autoridades estatísticas, isto é, entidades com competência própria para desenvolver, produzir e divulgar estatísticas oficiais – como o Instituto Nacional de Estatística (INE), o Banco de Portugal e os serviços regionais de estatística das Regiões Autónomas dos Açores e da Madeira. O INE também delega competências noutras entidades para a realização de operações estatísticas determinadas.

A nível internacional, outras entidades, como o Eurostat, são igualmente fontes de estatísticas oficiais.

Tudo o que daí vem é oficial e ficamos com a garantia de que as estatísticas divulgadas seguem metodologias e procedimentos harmonizados e reconhecidos nacional e internacionalmente.

Existem, ainda, outras entidades oficiais responsáveis pela produção de estatísticas em áreas particulares, como o IPMA, a Caixa Geral de Aposentações e outras, mas que não integram o SEN.

Há inúmeras estatísticas não oficiais que provêm de outras entidades, como empresas, centros de investigação ou organizações não-governamentais. A sua qualidade é muito variável e nem sempre é fácil ajuizar até que ponto os dados são consistentes e fiáveis. Por vezes, estão validados com uma metodologia clara e reconhecida. Noutros casos, chegam às redacções em comunicados que são parcos em explicações sobre como foram produzidos, quais os conceitos envolvidos, o universo considerado e outros elementos que permitam enquadrá-los melhor.

ALERTA

Verificar antes de usar

Antes de se utilizar uma estatística não oficial numa notícia, é necessário um trabalho prévio de verificação da sua qualidade e significado.

Três sugestões para esta avaliação:

1. Verificar se os dados vêm acompanhados de uma ficha metodológica ou de outra forma de referência que descreva os elementos fundamentais que estão na base daquelas estatísticas, tais como a escala, unidade de medida, conceito, âmbito geográfico, operação de recolha, etc.
2. Contactar a entidade que os produziu ou divulgou, para esclarecer dúvidas sobre a metodologia e conceitos adoptados.
3. Recorrer a um especialista da área para uma opinião sobre a fiabilidade daqueles dados.

Mesmo no caso dos dados com o selo oficial, por vezes é preciso ir mais longe para os compreender. Duas entidades oficiais podem adoptar conceitos e operações de recolha distintos, resultando em números diferentes para o que parece ser uma mesma realidade.

Por exemplo, no caso do desemprego:

- O INE apresenta os dados dos desempregados com base num inquérito amostral à população;
- O Instituto de Emprego e Formação Profissional (IEFP) apresenta dados dos desempregados que se encontram inscritos nos centros de emprego.

Resultado: em 2015, o INE registou 647 mil desempregados, enquanto o IEFP contabilizou 561 mil.

Estamos, neste caso, perante não só duas fontes primárias diferentes, como duas operações de recolha distintas, com base em conceitos diferentes do que é um desempregado. E, naturalmente, os valores não coincidem.

O QUE É

Operação de recolha é a forma como os dados estatísticos são recolhidos. São distintas, conforme:

- Envolvam toda a população (recenseamento) ou uma parte (amostra);
- Obtenham os dados por via directa (inquéritos) ou administrativa (registos).

Diferentes combinações de operações de recolha são encontradas nas estatísticas. Por exemplo:

	Recenseamento	**Amostra**
Via directa	Censos da população	Inquérito ao emprego
Via administrativa	Estatísticas de casamentos	–

Há ainda várias situações que não resultam de uma recolha específica de informação, e sim de uma combinação de múltiplos dados pré-existentes, por vezes obtidos por diferentes fontes e/ou operações estatísticas. São, por isso, *estudos estatísticos*. Indicadores macroeconómicos, como o PIB ou o índice de preços no consumidor, estão nesta categoria.

As fontes e operações de recolha são uma componente da metainformação dos quadros estatísticos (ver capítulo sobre *metainformação*).

O QUE SIGNIFICA?

Conceitos

Provavelmente não há estatísticas mais confusas em Portugal do que as do emprego e do desemprego. É um universo tão labiríntico que até mete medo tentar explicá-lo. Tudo o que parece óbvio, afinal não é. "População activa" não é o mesmo que "população em idade activa". "Inactivos" não são todos os que não estão a trabalhar. A taxa de emprego não é o reverso da taxa de desemprego.

A lista é longa e resulta de algo que muitas vezes negligenciamos nas estatísticas: os *conceitos* que estão por detrás de cada dado ou indicador estatístico.

O QUE É

Conceito é a definição explícita do que representa um dado estatístico. Na prática, delimita o que está a ser contabilizado.

Desempregado, por exemplo, não é apenas aquele que não tem trabalho. O conceito formal, assumido no Inquérito ao Emprego, é definido por um texto com quase 200 palavras. Resumidamente, um desempregado é aquele que:

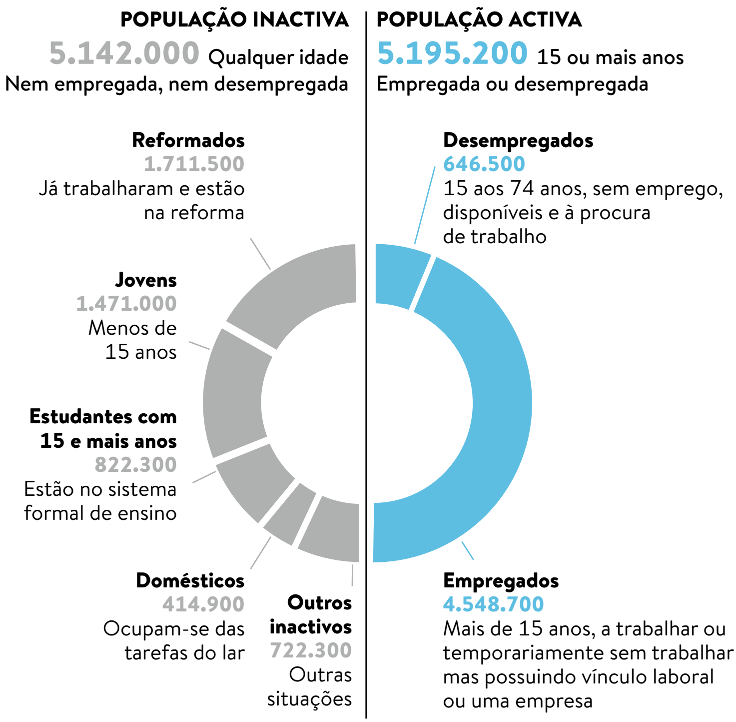
- Tem 15 a 74 anos;
- Não tem trabalho;
- Está disponível para trabalhar;
- Está a procurar trabalho activamente.

Todas estas condições têm de ser cumpridas para se entrar para as contas do desemprego.

Já os empregados são os que:

- Têm 15 anos ou mais;
- Estão numa destas situações:
  - Trabalham e recebem uma remuneração;
  - Possuem vínculo laboral ou têm uma empresa, mas temporariamente não estão a trabalhar;
  - Estão em pré-reforma, mas a trabalhar.

Como se vê, não é assim tão simples quanto parece. E querem mais confusão? Então eis como a população portuguesa estava dividida em 2015, em termos de ocupação, e os respectivos conceitos, aqui simplificados para evitar sonolências:



Fontes: INE/Pordata

Os conceitos, como tudo em estatística, são resultado de escolhas deliberadas. Não são, é claro, opções aleatórias, mas sim fruto de decisões baseadas em normas e metodologias sólidas. Uma vez acordados, os conceitos são adoptados transversalmente por diferentes entidades, garantindo, dessa forma, que os dados são comparáveis, seja qual for a fonte. Mas nem sempre é assim.

ALERTA

Coisas parecidas, conceitos diferentes
Dois dados estatísticos com designações similares podem estar ligados a conceitos diferentes, não sendo por isso comparáveis.

É por isso que as estatísticas sobre as mortes nas estradas em Portugal Continental têm duas versões. Uma delas leva em conta apenas as mortes no local do acidente ou a caminho do hospital. A outra inclui os óbitos que ocorrem, na sequência do sinistro, até 30 dias depois. Assim, o número de mortos nos acidentes rodoviários em 2015 pode ser 473 ou 593, conforme a definição que se adopte. É uma questão de conceitos.

Dessa forma, dois indicadores que parecem ter o mesmo significado podem ter resultados diferentes. Não se pode dizer que um esteja certo e o outro errado. Os conceitos que estão na origem dos dados é que são distintos.

Também, nalguns casos, termos que parecem complementares, como as duas metades de uma moeda, podem não o ser. As taxas de desemprego e de emprego ilustram esta ironia. Se a primeira for de 20%, então a segunda será de 80%, correcto? Não, incorrecto. E, mais uma vez, a culpa é dos conceitos:

- Taxa de desemprego: número de desempregados em relação à **população activa** (empregados + desempregados). Foi de 12,4% em 2015.
- Taxa de emprego: número de empregados em relação à **população com 15 ou mais anos.** Foi de 51,2% em 2015.

Na galeria dos conceitos com grande potencial baralhador, figura também o par "população activa" e "população em idade activa". A população activa, como visto no exemplo anterior, é simplesmente a soma dos empregados e desempregados – o que de si já baralha um pouco, dada a tentação de se considerar quem está no desemprego como inactivo. Já a população em idade activa é algo diferente: são todos os que estão entre os 15 e os 64 anos, seja qual for a sua situação laboral (trabalhadores, desempregados, estudantes, reformados, domésticos ou outros).

Por estas e muitas outras razões, trabalhar com dados estatísticos sem olhar para os seus conceitos é meio caminho andado para equívocos. Facilmente se confunde inactivo com desempregado, entre tantos outros possíveis desencontros.

ALERTA

Conceitos podem mudar

Por vezes, os conceitos são alterados ao longo do tempo. Por exemplo, em 2011, algumas modificações importantes foram introduzidas nos conceitos de empregado e desempregado. Quando isto acontece, a série estatística sofre uma quebra, pois os dados anteriores à modificação não são totalmente comparáveis com os posteriores (ver capítulo sobre *quebras de série*).

Os conceitos associados aos dados estatísticos podem ser encontrados:

- Na metainformação;
- Nos glossários das bases de dados.

ONDE?

Âmbito geográfico

Está lá nos números: em 2015, não há registo de casamentos em Odivelas. Vivem no concelho cerca de 150.000 pessoas, a sua população é mais jovem do que a média nacional, estão lá todos os ingredientes para que centenas de casais por ano dêem o nó. Porém, nada.

É um paradoxo perturbador mas que se explica por uma subtileza das estatísticas sobre os casamentos: uma das formas de os contabilizar não leva em conta onde o casal vive, mas onde o matrimónio é registado. E como em Odivelas não há conservatória de registo civil, os noivos têm de o fazer na vizinha Loures ou noutro concelho.

Se um jornalista não souber deste detalhe do *âmbito geográfico* daquelas estatísticas, certamente pensará que tem uma boa manchete: *Ninguém se casou em Odivelas em 2015.*

O QUE É

O **âmbito geográfico** é a delimitação territorial do dado estatístico. Na prática, indica em que território o valor está a ser observado.

Há uma primeira linha do âmbito geográfico, que é a forma como os dados estão desagregados – por exemplo, por municípios, por regiões, por países ou por grupo de países. Mas já lá vamos.

A segunda linha é aquela onde o diabo se esconde. É o chamado *âmbito geográfico de referência*, que especifica aquilo que é levado em conta na altura da recolha da informação. Se o âmbito geográfico de referência dos casamentos é o local de registo, o matrimónio de uma pessoa que vive em Lisboa mas se casa no Porto engrossará as estatísticas do segundo município e não do primeiro.

Ainda no domínio conjugal, aqui está mais uma ironia de Odivelas. Embora não haja, oficialmente, casamentos no concelho, cerca de 500 pessoas ficam ali viúvas todos os anos. E por quê? Porque o âmbito geográfico de referência para a dissolução de casamentos por morte de um dos cônjuges é o local de residência do falecido.

ALERTA

Verificar sempre

Certificar-se de qual é o âmbito geográfico de referência dos dados é essencial para a interpretação de estatísticas, especialmente as regionais.

Não se deve negligenciar este alerta, pois, em especial nas estatísticas regionais, há incontáveis particularidades no que toca a definir a quem pertence um número. Por exemplo:

Venda de combustíveis: a referência é a localização geográfica da venda. Se estivermos a falar de gasóleo e gasolina, será onde estão os postos de abastecimento e não onde vivem os consumidores. Por isso, se os moradores de uma cidade estiverem a atestar o depósito no município vizinho porque é mais barato, o seu consumo de combustível aparecerá nas estatísticas desse concelho.

Diplomados: as estatísticas sobre os alunos que concluíram um curso superior (diplomados) em cada concelho num determinado ano tomam por base o local do estabelecimento de ensino e não o de residência desses diplomados. Em 2015, Lisboa, Porto, Coimbra, Braga e Aveiro estão à frente, pois é lá que se localiza grande parte dos estabelecimentos de ensino superior. Já os dados dos Censos de 2011 sobre o número de pessoas que têm o ensino superior completo tomam por base o local de residência. Neste caso, Lisboa, Porto, Oeiras, Cascais e Sintra é que lideram a lista de municípios.

Médicos: as estatísticas de médicos por concelho em Portugal também podem exigir algum cuidado interpretativo. Um profissional pode ser contabilizado no Porto mas exercer a sua actividade em Aveiro. Isto porque certos dados estatísticos se baseiam nas inscrições desses profissionais na Ordem dos Médicos, onde o que conta é o local de residência e não o concelho de exercício da actividade profissional.

Pessoal nas empresas: imagine que trabalha na filial de uma empresa em Lisboa, mas com sede em Vila Nova de Gaia. A sua ligação à empresa pode entrar nas estatísticas para o segundo concelho e não para o primeiro. O âmbito geográfico de algumas estatísticas sobre o pessoal ao serviço das empresas corresponde ao local da sede da empresa.

A referência ao âmbito geográfico é essencial. Faz parte da história de um número, ajuda a compreender o que ele representa e evita interpretações incorrectas sobre os valores.

A outra linha do âmbito geográfico, tal como referido, é simplesmente a desagregação dos dados a nível territorial. É fácil compreender que desagregação é esta quando um quadro estatístico se refere, por exemplo, a freguesias, a concelhos, ao país como um todo ou a um grupo de países, como os Estados-membros da União Europeia. O cenário torna-se um pouco menos óbvio, porém, quando a referência territorial são as NUTS, que em inglês significa frutos de casca rija, mas que nas estatísticas nada têm a ver com comida (ver caixa *O que são NUTS?*).

Os títulos dos quadros estatísticos explicitam por norma o modo de desagregação dos dados.

Quanto ao âmbito geográfico de referência – que, como alertado mais acima, é essencial conferir – esta indicação encontra-se normalmente descrita na metainformação dos dados estatísticos (ver capítulo sobre *metainformação*).

O QUE SÃO NUTS?

NUTS são uma classificação territorial comum adoptada em todos os países da União Europeia para fins estatísticos. A sigla, traduzida para português, diz isso mesmo: Nomenclatura das Unidades Territoriais para Fins Estatísticos.

Foram criadas pelo Eurostat com o objectivo de garantir a comparabilidade das estatísticas entre os Estados-membros. Assim, todo o território da UE está subdividido em NUTS, em três níveis diferentes.

Portugal está assim repartido:

3 NUTS I
Portugal Continental, Região Autónoma dos Açores e Região Autónoma da Madeira

7 NUTS II
Norte, Centro, Alentejo, Algarve, Área Metropolitana de Lisboa, Região Autónoma dos Açores e Região Autónoma da Madeira

25 NUTS III
21 Comunidades intermunicipais, Área Metropolitana de Lisboa, Área Metropolitana do Porto, Região Autónoma dos Açores e Região Autónoma da Madeira

Falar em NUTS numa notícia é receita para dormir. Mas, frequentemente, é importante termos uma leitura supramunicipal do país. A delimitação geográfica das NUTS serve, ainda, de padrão para a aplicação das políticas europeias, como a dos apoios financeiros da UE, daí a sua importância em temas mediáticos.

QUANDO?

Período ou momento de referência

Quantos habitantes tinha Portugal em 2011?
a) 10.542.398
b) 10.557.560
c) 10.562.178

A resposta, segundo as estatísticas oficiais, é a), b) e c). Todas estão certas. Simplesmente, cada um dos valores diz respeito a um momento diferente. A primeira constitui a estimativa do número de habitantes a 31 de Dezembro. A segunda é uma estimativa da população a meio do ano. E a terceira é a contagem feita pelos Censos da População, referente ao dia 21 de Março.

Sem saber o *momento* ou *período de referência* de um dado estatístico, ficamos um pouco descalços para a sua interpretação.

O QUE É

Período ou momento de referência é o “quando”, o período de tempo ou a data específica a que se refere um dado estatístico.

Nas estatísticas, encontram-se diferentes períodos de referência:

- Ano civil: é o mais comum. É como se conta o número de espectáculos, de consultas médicas, de nascimentos, de óbitos e inúmeros outros indicadores.
- Um mês específico: é o caso da média da temperatura máxima do ar no mês mais quente do ano.
- Uma data específica: na maior parte das vezes 31 de Dezembro ou 1 de Janeiro. Por exemplo, os processos pendentes nos tribunais judiciais. Outras datas: a dos Censos e a das eleições.
- Outros casos: por exemplo, valores médios que consideram vários anos para o cálculo, como acontece com os dados sobre a esperança de vida, quando se referem a triénios.

ALERTA

Não confundir com periodicidade

Período de referência não é o mesmo que periodicidade – que é a frequência com que um dado estatístico é apurado. Por exemplo, os Censos contam a população num único dia – que é o seu momento de referência. Mas ocorrem, por norma, uma vez a cada dez anos, ou seja, a sua periodicidade é decenal.

Alguns indicadores próximos podem, por vezes, estar relacionados com períodos de referência diferentes. Por exemplo, os processos judiciais que entraram ou foram concluídos nos tribunais contam-se ao longo de um ano civil, mas os pendentes são aqueles que ainda se encontram por resolver a 31 de Dezembro. Logo, se relacionarmos o número de processos pendentes com o de concluídos, estamos a considerar dois períodos de referência distintos. E isto deve ser referido na notícia.

ALERTA

Há períodos que não se somam

Nem sempre se podem somar dados infra-anuais para se obter o resultado do ano. É o que acontece com os dados mensais do desemprego do IEFP, por exemplo. Os valores indicam quantas pessoas estão inscritas nos centros de emprego. Suponhamos que são 100.000 todos os meses. Ao final do ano, não haverá 1.200.000 desempregados, mas sim 100.000 desempregados ao longo de 12 meses.

Há indicadores que estão sujeitos a variações sazonais, ou seja, aquelas que se acentuam previsivelmente num determinado período do ano. É o que acontece em Portugal, por exemplo, com o desemprego, a ocupação hoteleira ou a inflação. Nestes casos, é útil ter em conta os períodos homólogos dos anos anteriores (ver capítulo sobre *variações e taxas de variação*).

Por norma, o momento ou período de referência de um quadro estatístico é uma componente da metainformação associada aos dados (ver capítulo sobre *metainformação*).

SEMPRE SE CONTOU ASSIM?

Quebras de série

O ano de 2004 foi fantástico para o mundo empresarial do país. As estatísticas parecem revelar que, em apenas doze meses, surgiram quase 450 mil novas empresas em Portugal. Eram pouco menos de 670 mil em 2003 e saltaram para 1,1 milhões no ano seguinte, um vertiginoso aumento de 67%.

Esperem lá... Nesse mesmo período, a taxa de desemprego não diminuiu, pelo contrário, aumentou. Há aqui algo que não está a bater certo.

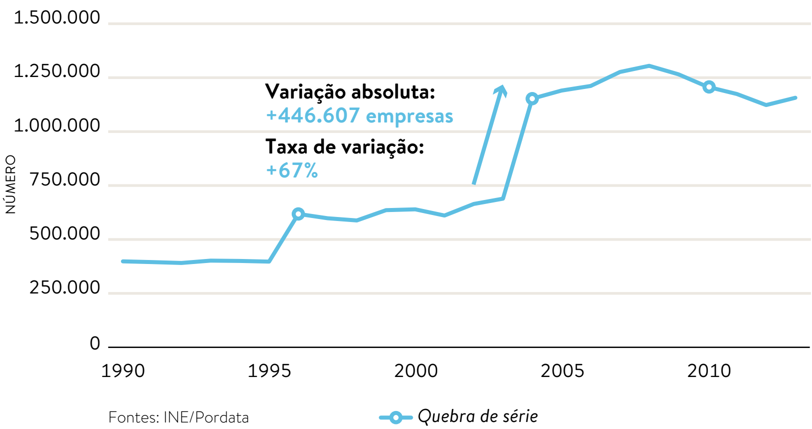
Pois há: uma *quebra de série* nos dados sobre as empresas.

O QUE É

Uma **quebra de série** ocorre quando se altera a forma como um dado estatístico é definido ou contado, afectando a sua comparabilidade com dados de momentos anteriores. Mudanças no tipo de recolha de dados, na metodologia aplicada, nos conceitos, no período de referência, na abrangência do universo, são exemplos de alterações que provocam quebras de série.

O que aconteceu às empresas foi muito simples: a partir de 2004, os trabalhadores independentes também passaram a ser considerados empresas individuais – tal como já o eram os empresários em nome individual. Ou seja, modificou-se o universo de referência do conceito de empresa. E o resultado foi este:

Número de empresas em Portugal



Fontes: INE/Pordata

ALERTA

Não se compara

Quando há uma quebra de série, os valores antes e depois daquele momento deixam de ser totalmente comparáveis entre si.

Sabendo disso, é claro que já não se podem fazer manchetes com aqueles 67% de aumento no número de empresas, pois é incorrecto afirmar que houve tal acréscimo entre 2003 e 2004. Os dados a partir de 2004 não são comparáveis com os anteriores àquela data.

ALERTA

Atenção aos símbolos

Quebras de séries são muito comuns e, por princípio, estão identificadas nos quadros estatísticos através de símbolos, como por exemplo:

⊥

b

Nos gráficos, os momentos onde há quebra de série também podem estar assinalados.

Olhar para a evolução de um dado estatístico, descobrir um valor que seria notícia em qualquer parte do mundo e depois concluir que apenas se deve a uma alteração metodológica ou conceptual é um verdadeiro anti-clímax jornalístico.

No entanto, embora as quebras de série retirem a comparabilidade directa entre o antes e o depois, o seu impacto real deve ser analisado caso a caso. Para isso, é preciso compreender as razões da quebra de série, de modo a ajuizar o seu efeito potencial sobre os dados.

ALERTA

Cuidados a ter

Sempre que uma notícia inclua estatísticas com quebra de série, deve-se:

- Rejeitar comparações directas entre o antes e o depois;
- Investigar as razões da quebra de série;
- Explicar, com esta informação, como os dados foram afectados.

Uma quebra de série nada mais é do que uma adaptação necessária das estatísticas a uma nova realidade, em função de evoluções metodológicas, tecnológicas, legais ou outras. Por exemplo:

- Alterações metodológicas: até 2003, as estimativas do número de empresas eram realizadas a partir de um inquérito amostral. Em 2004, os dados passaram a ser recolhidos por recenseamento administrativo.
- Alterações tecnológicas: em 2011, os dados sobre os recintos culturais sofrem uma quebra de série porque a informação deixa de ser recolhida em papel para passar a sê-lo por via electrónica.
- Reconfigurações territoriais: a criação da freguesia do Parque das Nações, em 2013, abrangendo uma área que pertencia ao concelho de Loures e passou para Lisboa, alterou as estatísticas dos dois concelhos.
- Alterações legislativas: a legalização do registo civil de casamentos de pessoas do mesmo sexo, em 2010, originou uma quebra de série nas estatísticas sobre os matrimónios. O universo dos dados deixou de considerar unicamente os casamentos entre pessoas de sexo diferente. O impacto da alteração foi, no entanto, reduzido, pois os novos registos

representaram, naquele ano e nos seguintes, apenas cerca de 1% do total.

- Alterações de conceitos: a proliferação do alojamento local levou a que as estatísticas passassem a considerar também esta tipologia entre os empreendimentos turísticos. Resultado: quebra de série em 2013.

Além de estarem indicadas nos quadros estatísticos através de símbolos, as quebras de série e as razões que as motivaram também são uma componente da metainformação associada aos dados (ver capítulo sobre *metainformação*).

Referências bibliográficas

BBC Trust (2016). *BBC Trust Impartiality Review: Making Sense of Statistics.* BBC Trust.

Blalock, H. (1979). *Social Statistics* (2.ª ed.). McGraw-Hill Book Company.

Bolton, P. (2007). *Statistical literacy guide. Index numbers.* House of Commons Library.

Cohen, S. (2001). *Numbers in the Classroom: Using Math and Statistics in the News.* Investigative Reporters and Editors.

Cohn, V., Cope, L. (2012). *News and numbers: A Writer's Guide to Statistics* (3.ª ed.). Wiley-Blackwell.

Deviant, S. (2010). *The Practically Cheating Statistics Handbook* (2.ª ed.). StatisticsHowTo.com.

Estatísticas da Noruega, Paris 21 (2009). *Apresentação Amigável de Estatísticas, Guia para a elaboração de uma estratégia e directrizes de disseminação para os países em desenvolvimento e em transição.* Paris 21 e Estatísticas da Noruega.

Harrison, S. (2009). How to: get to grips with numbers as a journalist. Journalism.co.uk. URL: https://www.journalism.co.uk/skills/how-to-get-to-grips-with-numbers-as-a-journalist/s7/a534975/. Acesso em Setembro de 2016.

Harrison, S. (2009). How to: get to grips with numbers as a journalist. Part 2. Journalism.co.uk.

Hawkes, N., Sierra, L. *et al*. (2010). *Making Sense of Statistics*. Sense about Science, Straight Statistics e Royal Statistical Society.

Hofer, T., Przyrembel, H., Verleger, S. (2004). New evidence for the Theory of the Stork. *Pediatric and Perinatal Epidemiology, 2004*, *18*, pp. 88-92.

Huff, D. (1954). *How to Lie with Statistics*. W. W. Norton & Company.

IBGE (s.d.). *Guia do Censo 2010 para Jornalistas.* Instituto Brasileiro de Geografia e Estatística.

INE (s.d.). *O que é o Índice de Bem-estar (IBE).* Instituto Nacional de Estatística.

INE (s.d.). Sistema de Metainformação. Instituto Nacional de Estatística. URL: http://smi.ine.pt/. Acesso entre Setembro e Novembro de 2016.

Kille, L.W. (2014). Math basics for journalists: Working with averages and percentages. Journalist's Resource. Harvard Kennedy School. URL: http://journalistsresource.org/tip-sheets/foundations/math-for-journalists. Acesso em Setembro de 2016.

Livingston. C., Voakes, P. (2005). *Working with Numbers and Statistics. A Handbook for Journalists.* Lawrence Erlbaum Associates.

Matthews, R. (2000). Storks deliver babies (p=0.008). *Teaching Statistics. Volume 22. Number 2. Summer 2000.* pp. 36-38.

McConway, K. (2012). Score and Ignore. A radio listener's guide to ignoring health stories. *Significance.* Outubro 2012. pp. 45-48.

Meyer, P. (2002). *Precision Journalism. A Reporter's Introduction to Social Science Methods* (4.ª ed.). Rowman and Littlefield.

Mosqueda, P. (2011). Entendendo estatísticas: um guia para jornalistas. Jornalismo nas Américas. Universidade do Texas em Austin. URL: https://knightcenter.utexas.edu/pt-br/blog/entendendo-estatisticas-um-guia-para-jornalistas. Acesso em Setembro 2016.

OECD (2008). *Handbook on Constructing Composite Indicators. Methodology and User Guide.* OECD, JRC.

Paulos, J. A. (1995). *A Mathematician Reads the Newspaper.* Basic Books.

Pereira, S., Machiavelo, A., Azevedo, J. (2016). Maths in the news: uses and errors in Portuguese newspapers. *JCOM (Journal of Science Communication).* 15(04), A03.

Pereira. S. (2015). *A Matemática na Imprensa Portuguesa.* Tese de doutoramento. Faculdade de Letras da Universidade do Porto.

Pordata (s.d.). Glossário. URL: http://www.pordata.pt/Glossario. Acesso entre Setembro e Dezembro de 2016.

Reis, E. (2008). *Estatística Descritiva* (7.ª ed.). Edições Sílabo.

Rosa, L.V. (2013). *Inquéritos e Sondagens. Dicionário.* UnYLeYa.

Royal Statistical Society (s.d.). *A dozen rules of thumb for journalists.* Royal Statistical Society.

Statistics How To (s.d.). Misleading Graphs: Real Life Examples. Statistics How To: Statistics for the rest of us!. URL: http://www.statisticshowto.com/misleading-graphs/. Acesso em Outubro de 2016.

UNDP (2015). *Human Development Report 2015. Work for Human Development.* United Nations Development Programme.

UNECE (2009). *Making Data Meaningful. Part 1: A guide to writing stories about numbers.* United Nations Economic Commission for Europe/Nações Unidas.

UNECE (2009). *Making Data Meaningful. Part 2: A guide to presenting statistics.* United Nations Economic Commission for Europe/Nações Unidas.

Vigen, T. (2015). *Spurious Correlations.* Hachette Books.

Wickham, K.W. (2003). *Math Tools for Journalists* (2.ª ed.). Marion Street Press.

Zeisel, H. (2010). *Fale com Números*. Assírio & Alvim, (Textos de Ciências Sociais e Humanas).

Fontes de dados

Eurostat: ec.europa.eu/eurostat/

Instituto Nacional de Estatística: www.ine.pt

Pordata: www.pordata.pt

Índice remissivo

U

V

Ricardo Garcia Jornalista desde 1988 nas áreas do ambiente, ciência e jornalismo de dados. Autor de *Sobre a Terra*, um guia ambiental, e *Nós no Mundo*, sobre a sustentabilidade.

Maria João Valente Rosa Doutorada em Sociologia. Professora universitária da FCSH/UNL. Dirige, desde 2009, a Pordata. É membro do Conselho Superior de Estatística e do Comité Consultivo Estatístico Europeu (ESAC).

Luísa Barbosa Socióloga, colaboradora da Pordata desde a sua criação, em 2009, e coordenadora de projecto de investigação sobre exibição alternativa de cinema.

QUANTO EM RELAÇÃO AO TOTAL?

PROPORÇÕES

QUAL É O POTENCIAL?

TAXAS

QUANTO VALE HOJE?

VALORES NOMINAIS E REAIS, PREÇOS CORRENTES E CONSTANTES

MUDOU MUITO?

VARIAÇÕES E TAXAS DE VARIAÇÃO

QUANTO VALE x?

REGRA DE TRÊS SIMPLES

COMO POUPAR MIL PALAVRAS?

GRÁFICOS

MAIS OU MENOS QUANTO?

ARREDONDAMENTOS

TRADUZINDO?

ESCRITA SIMPLES

ALHOS COM BUGALHOS?

RÁCIOS

DADOS SOBRE OS DADOS?

METAINFORMAÇÃO

9 789898 838889
ISBN 978-989-8838-88-9