

O Fascínio pelo Círculo

Abstract.

O círculo é uma das formas com maior utilização na visualização de dados. Reconhecemos o gráfico circular ou de *pizza* como um dos gráficos circulares mais utilizados em todo o mundo. Este surgiu há mais de 200 anos por *William Playfair* e o seu uso foi ganhando força ao longo do tempo. No entanto muitos especialistas criticam o seu uso recorrente. O gráfico circular apresenta diversas limitações, pelo que não deve ser eleito para representações de dados complexos.

O gráfico de bolhas faz também uso do círculo para representar diversos dados. A área da forma representa a informação, exigindo a comparação entre diversas bolhas. No entanto, os gráficos de bolhas também apresentam diversas limitações, uma vez que a percepção da área do círculo é difícil para o sistema visual humano. A percepção de apenas uma unidade como o comprimento torna-se mais fácil e efetiva ao olho humano.

Neste sentido, o círculo apresenta diversas objeções no seu uso, seja qual for a sua representação. A questão é, se o círculo pode levar a erros de interpretação de dados, porquê o fascínio pelo seu uso no *design* de informação?

Palavras-chave: *design*, informação, círculo, forma circular, percepção, visualização.

1. Introdução

O *Design* de Informação é uma constante no nosso dia-a-dia, essencial para a compreensão do indivíduo de diversas matérias, desde as mais básicas às mais complexas. A visualização é o ponto de partida para o entendimento da informação, definida como o processo de formação de uma imagem mental, tão antigo quanto a própria história. Cairo (2012b) refere em "*The Functional Art*", a visão é o resultado do mapeamento do nosso ambiente baseado na informação que lhe está agregada e que os nossos olhos obtêm através de múltiplas fixações em pontos de interesses específicos (p. 83). Neste processo o olho é apenas o instrumento ótico, sendo o cérebro o motor da compreensão (Gregory, 1998).

Segundo a Teoria de *Gestalt*, o cérebro processa a imagem visual, a forma, tendo em consideração fatores como o equilíbrio, a clareza e a harmonia visual, como uma necessidade interna de organização. A percepção visual é realizada a partir da junção das formas e não do isolamento das mesmas. É um processo no qual se recorre à função cerebral para atribuir significado aos estímulos sensoriais (visão) recorrendo à memória.

A apreensão da informação visualizada depende, de forma significativa, da forma com que esta é apresen-

tada. O sistema visual é sensível a estímulos específicos como o movimento, as variações de cor, os contrastes formais, diferenças de orientação, entre outros, aos quais o cérebro dá uma prioridade específica. A forma é um dos primeiros aspectos formais a ter em conta na visualização de dados, e um dos quais mais influencia a percepção.

Na visualização de dados diversas formas são adotadas, no entanto, parece haver uma inclinação para o uso do círculo. A forma circular, mais especificamente o círculo, é uma das escolhas formais mais constantes, tendo vindo a ser utilizada muitas vezes sem consciência das suas capacidades e limitações.

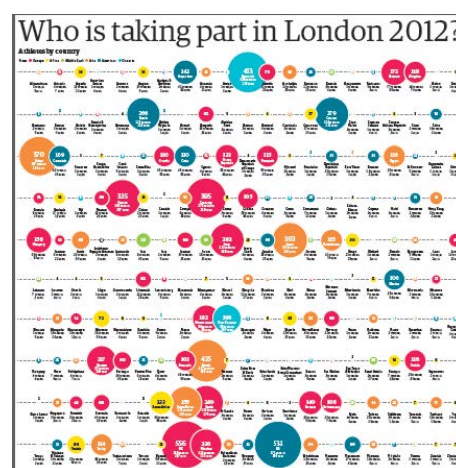


Fig 1. Visualização do número atletas de cada país a participar nos jogos Olímpicos de Londres em 2012.

Fonte: *The Guardian* (2012). Consultado a 22 de Dezembro de 2015 em <http://www.theguardian.com/sport/datablog/2012/jul/30/olympic-medals-visualised-guardian>

No exemplo acima, o círculo é usado numa timeline como elemento principal de representação de dados. A área do círculo é usada para representar o número de atletas a competir nos jogos Olímpicos de 2012 em Londres, por cada país. São identificadas seis categorias por seis diferentes cores: Europa a rosa escuro; África a amarelo; Médio Oriente a verde; Ásia a laranja; Américas a azul escuro e Oceânia a azul claro. Dentro das categorias existentes são representados diversos países. O que vemos no gráfico é uma organização de dados por ordem alfabética, não muito usual e, normalmente utilizada quando não é possível a organização por localização, tempo, categoria ou hierarquia. No entanto, neste caso, a organização podia ter sido realizada por qualquer das hipóteses, à exceção do tempo. A extensão dos círculos por diversas linhas torna o diagrama complexo e difícil de ler, assim como o largo número de círculos, que têm uma capacidade limitada de tamanho de dados. Há partida, neste caso, a representação dos dados poderia ser feita de forma simplificada e correta a partir de barras. No entanto, é explorada a forma circular com a mesma função das

barras - fazer comparações de dados - o que parece não ter sido uma escolha ponderada, uma vez que a capacidade de comparação do cérebro é mais efetiva quando relativa a comprimentos, ao invés de áreas.

Algumas correções podiam ser feitas através da interatividade, por exemplo ao clicar ou passar o rato sobre as bolhas serem exibidas as informações ocultas, ou com uma opção para reorganizar ou filtrar categorias agrupadas. No entanto, a organização dos dados podia ser feita de forma mais efetiva por localização e categoria, uma vez que cada categoria determina uma localização no globo. Assim, se a informação fosse disposta num mapa de bolhas, esta seria também agrupada por categorias.

Talvez assim fizesse mais sentido o uso do círculo, mas ainda assim estariam presentes as mesmas dificuldades de comparação de dados. Notamos, portanto, um fascínio pelo círculo, isto é, um uso não poderado e repetido da forma circular na representação de dados que não é entendido através de razões técnicas.

2. O círculo

O círculo é uma das formas básicas possíveis de encontrar na Natureza. A disposição equidistante dos pontos exteriores ao centro é impossível de desenhar com a mão humana. Caracteriza-se como uma linha curva fechada que divide o plano em duas partes: o interior e o exterior. As formas que têm linhas circulares geram uma resposta emocional positiva dos indivíduos. Os círculos não têm um fim nem um início definido e, como tal, transmitem unidade, infinidade e movimento livre (ex. roda de um carro). Representam o eterno e em todas as culturas são uma forma de representar o sol, a terra, a luz, o universo e os outros objetos celestiais. Pelas linhas curvas com que são compostos, são associados ao feminino, ao ventre, à gestação e conectados ao amor, energia e poder. São quentes, reconfortantes e transmitem harmonia. Os círculos protegem, suportam e restringem. Limitam o que está no interior e o que está no exterior. Oferecem segurança e conexão. Sugerem comunidade, integridade e perfeição. Representam o foco e derivam ideais e, como tal, atraem e criam ênfase. Representam o todo e o ciclo, como o intangível processo do tempo. Todo o fascínio pelo círculo poderá ser explicado pelo poder comunicativo e atrativo que ele revela?

3. A forma circular na Visualização de Dados

3.1. O gráfico circular

No *design* o ponto significa o todo, completo. Na infografia o círculo pode ser utilizado como uma metáfora para um ponto de começo em diversas formas de visualização (Klanten, Bourquin, Tissot & Ehmann (Eds.), 2008). Os mais simplistas e reconhecidos gráficos que fazem uso do círculo são os gráficos circulares, de fatias, *pizza* ou queijo, que oferecem uma forma fácil de sublinhar o efeito de largas diferenças. São um

dos tipos de gráficos com maior utilização em todo o mundo e evidenciam de forma muito particular o fanatismo pelo círculo. Ajudam a mostrar as proporções e percentagens entre categorias, no entanto, apresentam muitas limitações. Têm, assim, apesar da sua popularidade, sido criticados por vários especialistas que colocam em questão os limites da sua funcionalidade. No gráfico circular o círculo é dividido em segmentos proporcionais em que cada comprimento de arco representa uma proporção de cada categoria. No seu todo o círculo representa a soma total de todas as categorias, igual a 100%.

A sua origem remonta há mais de 200 anos. O primeiro gráfico do género apareceu em 1801, pelo engenheiro de naturalidade escocesa *William Playfair* (1759 – 1823), em *"The Statistical Breviary"* como elemento de dois grandes diagramas, considerado então o seu criador.

William Playfair foi conhecido primeiramente, em 1786, pela criação e popularização do gráfico de barras e do gráfico de linhas na representação de dados estatísticos temporais. Destes, o gráfico *pizza* foi a sua última maior invenção.

Em *"The Statistical Breviary"*, Playfair faz uma representação dos dados estatísticos do fim do século XIX dos países Europeus. No primeiro diagrama, *Playfair* retrata os países Europeus antes da Revolução Francesa de 1789. No segundo diagrama é feita uma representação desses países em 1801, o ano em que foi realizado o tratado de paz *Luneville* entre França e Áustria, constatando uma mudança.

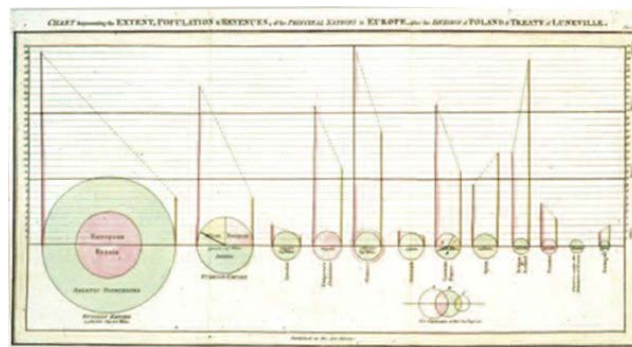


Fig 2. Segundo diagrama de *William Playfair* em *Statistical Breviary* (1801), que representa dados estatísticos dos países Europeus em 1801, o ano do tratado de paz *Luneville* entre França e Áustria.

Fonte: Spence, Ian (2005). *No Humble Pie: The Origins and Usage of a Statistical Chart*.

Os círculos de diferentes dimensões representam, proporcionalmente, as áreas territoriais de cada país. A Rússia é a primeira a ser representada pelo círculo de maior dimensão, enquanto que Portugal está associado a um círculo muito pequeno. Quase todos os países têm a indicação do valor do tamanho do terreno em milhas quadradas logo abaixo do diâmetro horizontal do círculo. Para além do tamanho territorial, estão representados os dados relativos ao tamanho

da população e às receitas de cada país. A massa populacional é representada pela linha vertical vermelha ao lado esquerdo de cada círculo. A receita de cada país é representada pela linha vertical amarela à direita. A carga fiscal de cada país é representada pela linha tracejada que liga o topo das duas linhas verticais. No entanto, o declive desta linha varia, inevitavelmente, com o diâmetro dos círculos, como *Funkhouser* (1937) observou (*cit. in Spence, 2005*). Como tal, não é possível ter um índice preciso da carga fiscal. Podemos apenas observar se o declive é negativo ou positivo. Através das cores, também é feita a distinção de países Europeus com potências marítimas, a verde, de países com potências terrestres, a vermelho. (*Spence, 2005*).

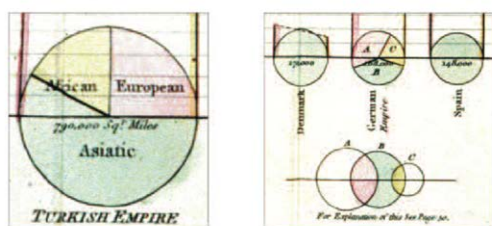


Fig 3. Detalhe do diagrama da figura anterior. Gráfico de fatias que representa o Império Turco à esquerda. À direita gráfico de fatias e diagrama de Venn que representam o Império Alemão.

Fonte: Spence, Ian (2005). *No Humble Pie: The Origins and Usage of a Statistical Chart*.

Para ilustrar as subdivisões políticas em alguns países, *Playfair* usou uma variedade de métodos. No caso do Império Russo, este é dividido em dois círculos concêntricos que representam o domínio Europeu e o domínio Asiático, ao que chamamos hoje gráficos de bolhas. As áreas separadas representam as partes do império em cada continente. O domínio Europeu é colorido de vermelho sendo representado como uma potência terrestre. Já o domínio Asiático a verde representa-se como uma potência marítima. No caso do Império Turco, o segundo maior, o uso do diagrama circular, mais especificamente de três círculos concêntricos, tornaria a comparação visual das áreas mais difícil que no caso do Império Russo. Como tal, *Playfair* dividiu o círculo em três fatias ou setores, cada um proporcional ao território Asiático, Europeu e Africano. O setor relativo ao território Asiático é colorido a verde como representação de uma potência marítima, o setor Europeu é colorido a vermelho como representação de uma potência terrestre, enquanto que o setor Africano é representado a amarelo, sem nos ser dada uma explicação racional para tal associação. Já no caso do Império Alemão, *Playfair* ilustra a subdivisão do Império em partes pertencentes à Áustria e Prússia e, ainda do restante território sob o controle dos príncipes alemães, através de dois modos de representação: o gráfico de pizza como no Império Russo, e um gráfico com três círculos sobrepostos lado a lado ao qual hoje chamamos diagrama de Venn, usado

para mostrar propriedades conjuntas. O círculo mais à esquerda (A) representa os interesses austríacos, o círculo mais à direita as autoridades prussianas (C), e o círculo central representa o Império Alemão (B). A interseção a vermelho representa a porção governada pela Áustria, a interseção a amarelo indica a área sobre o governo da Prússia, e o verde apresenta a província dos príncipes alemães. (*Spence, 2005*).

Notamos, portanto, que *Playfair*, em apenas dois diagramas, introduziu três variedades de gráficos circulares: o diagrama circular a que chamamos atualmente gráfico de bolhas ou bubble chart (com círculos concêntricos), o gráfico circular ou de fatias e o diagrama de Venn.

Playfair não foi rapidamente reconhecido pelas suas invenções gráficas. Na Grã-Bretanha foi ignorado devido à sua má reputação. Foi na Alemanha e França que as suas publicações atraíram as atenções, particularmente em Paris.

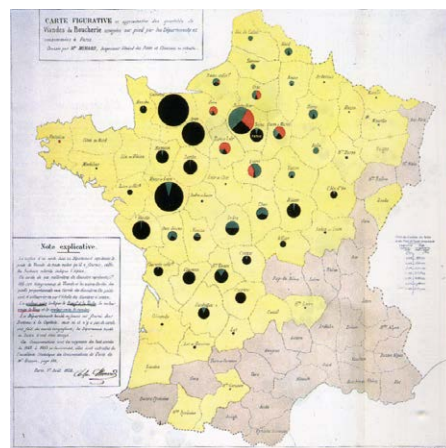


Fig 4. Mapa estatístico de Joseph Minard, 1858

Fonte: Spence, Ian (2005). *No Humble Pie: The Origins and Usage of a Statistical Chart*.

Apesar de nem toda a atenção ser favorável, em 1858, o engenheiro francês *Charles Joseph Minard* usou o gráfico circular para representar a percentagem de gado para consumo enviado de diferentes regiões de França para Paris. As áreas dos círculos são utilizadas para representar quantidades variáveis, embora existam razões fundamentadas que questionam a adequação desta forma geométrica particular.

Em meados do séc. XX, vários peritos em estatística defenderam fortes opiniões contra o gráfico pizza, tendo sido realizados vários estudos posteriores para analisar a efetividade do mesmo. Diversos estudos compararam o gráfico de pizza com o gráfico de barras, uma vez que as percepções de ângulos são geralmente menos precisas que as percepções de comprimento. *Cleveland e McGill* (1984) concluíram que o gráfico pizza é menos preciso que o gráfico de barras ao estimar a proporção ou ao comparar proporções. (*cit. in Spence, 2005*). No entanto, *Simkin e Hastie* conseguiram demonstrar que, quando a tarefa é estimar a proporção de uma barra ou segmento no seu todo,

o gráfico de *pizza* é superior ao gráfico de barras (*cit. in Spence, 2005*). Já *Spence* (1990) defendeu que as proporções em gráficos de fatias são mais facilmente percebidas que noutras formas básicas (*cit. in Spence, 2005*).

Os gráficos circulares são ideais para dar ao leitor uma ideia rápida da distribuição proporcional dos dados. A representação de dados classificados em categorias nominais ou ordinais pode ser feita de forma efetiva através do gráfico circular. Os dados nominais são categorizados de acordo com informações descritivas ou qualitativas (ex. terra natal ou tipo de animal de estimação que temos) Os dados ordinais são semelhantes mas as diferentes categorias também podem ser classificados, por exemplo, num formulário as respostas podem ser classificada como péssimo, mau, regular, bom ou ótimo.

No entanto, críticos defendem que a existência de mais de 5 ou 6 categorias dificulta a percepção das diferenças dos tamanhos relativos e, por conseguinte, a interpretação da informação. *Hollands* e *Spence* (1992, 1998, 2001) conduziram diversas experiências que demonstraram que, desde que utilizados para representar um diminuto número de proporções, os gráficos *pizza* conseguem tornar-se tão efetivos quanto outros gráficos vulgarmente utilizados (*cit. in Spence, 2005*).

Ainda assim as suas limitações são variadas. Para além da inadequação da representação de grandes quantidades de dados, os gráficos de fatias dificultam comparações minuciosas e precisas entre proporções de ângulos. Desta forma, o impacto dos dados é perdido ou mal interpretado. Neste caso, um gráfico de barras pode ser uma solução, uma vez que a comparação de dados é facilitada em comprimento.

Ainda assim, *Spence* (2005), nascido a 1944, psicólogo escocês-canadense e professor emérito do Departamento de Psicologia da Universidade de Toronto, defende que o criticismo à volta do gráfico de fatias deve-se muito ao facto de se fazer uma utilização deste além das suas possibilidades. Justifica dizendo que o gráfico de fatias é um simples gráfico informativo cujo propósito é mostrar a relação da parte com o todo, não sendo uma escolha correta para representar situações complexas (p. 363). Ainda assim, o seu uso foi popularizado em diversas situações, que nem sempre permitem unicamente mostrar a relação da parte com o todo.

No exemplo apresentado abaixo, é feita uma representação dos mercados de um seletivo número de países através de diversos gráficos de fatias. O diagrama, realizado pelo *Harvard Business Review* em 2010, pretende retratar a saúde dos mercados e as dificuldades dos mesmos durante a crise económica, de modo a identificar quais os países que terão recursos para recuperar e quais não. Os círculos de diferentes dimensões identificam o PIB (Produto Interno Bruto) de cada país em escala, e a sua posição referencia a posição

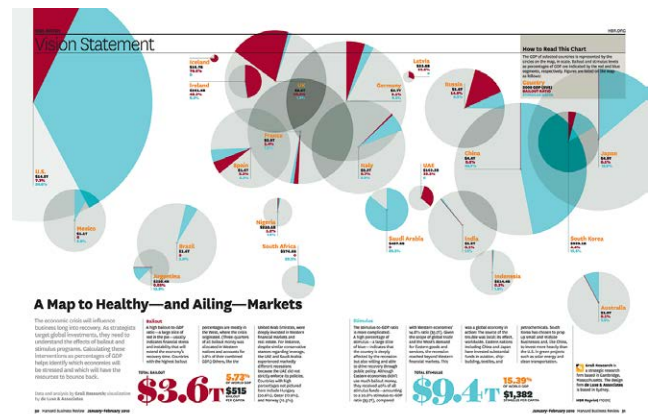


Fig 5. Vision Statement: A Map to Healthy—and Ailing—Markets por Grail Research e de Luxe & Associates. Fonte: *Harvard Business Review* (2010). Consultado a 22 de Dezembro de 2015 em <https://hbr.org/2010/01/vision-statement-a-map-to-healthy-and-ailing-markets>

de cada país no globo. Cada círculo é dividido então em três setores: a cinza o PIB total do país; vermelho o dinheiro dispensado a programas *"Bailout"*, isto é, a injeção de liquidez dada a entidades falidas ou próximas de falência, a fim de poderem honrar seus compromissos a curto prazo; a azul o dinheiro entregue a programas estímulo ou de incentivo.

O diminuto número de setores permite-nos identificar rapidamente quais os países com maior percentagem de programas *"Bailout"* e quais os países com uma maior percentagem de programas de incentivo. Como conclusão, é referido no diagrama que uma maior percentagem de programas *"Bailout"* normalmente indica stress e instabilidade financeira que irá aumentar o tempo de recuperação do País; por outro lado, uma maior percentagem de programas de incentivo indica que o País é severamente afetado pela recessão mas também que é capaz de recuperar da crise económica através de políticas públicas.

No entanto, podemos notar que os diversos círculos são sobrepostos uns aos outros o que dificulta a sua leitura em alguns casos. Como resultado disso também, os setores azul e vermelho são colocados de forma aleatória dentro do círculo não tendo como começo as 12 horas, criando maiores dificuldades na comparação. Outro aspeto que podemos referenciar é o facto de que países com baixo PIB não são facilmente analisados uma vez que os círculos são demasiado pequenos.

Porém, conseguimos perceber que o diagrama é efetivo no sentido que faz uma comparação das proporções com o todo, como *Spence* refere. Neste aspeto o círculo é a única forma básica que consegue transmitir a ideia de um todo sem a necessidade de leitura de dados. Já a comparação entre os dois setores, vermelho e azul, é de difícil leitura quando os seus ângulos são semelhantes. Isto deve-se ao facto, já referenciado, de que a percepção de ângulos é geralmente menos precisa que a percepção de comprimentos.

Podemos, desta forma, concluir que as limitações do gráfico de fatias são largas, e a sua utilização deve ser feita em representações de dados simples. Assim que a complexidade dos dados aumenta, a leitura dos mesmos é dificultada, diminuindo a sua eficácia. Ainda assim, a sua utilização é vasta, essencialmente na imprensa popular, ao contrário das áreas científicas e empresariais, como Ian Spence conclui pelo seu estudo analítico referenciado no artigo "No Humble Pie: The Origins and Usage of a Statistical Chart". Porém, a razão da popularização do seu uso não é clara.

3.2 Gráfico "donut"

Uma das variações criadas a partir do gráfico de pizza é o gráfico *donut*, em que a área do centro é eliminada, criando uma espécie de anel. O gráfico *donut* tem uma ligeira vantagem em relação ao gráfico de pizza. Uma vez que é o comprimento do arco que é percebido e não a área da fatia, o foco recai mais nas mudanças de valores globais, ainda assim, a percepção do comprimento do arco não é tão efetiva como a percepção de comprimentos em linha reta.

O exemplo a seguir mostra o orçamento de estado dos governos da Grã-Bretanha de 2011-12 e a mudança de percentagem desde o último orçamento. A azul mostra o montante de gastos dividido por áreas principais; a amarelo, as receitas divididas nas principais fontes de rendimento; e a vermelho o déficit do país, isto é, o dinheiro pedido por empréstimo, que representa a diferença de valor entre os gastos e as receitas. Os dados apresentam-se em milhões de libras em todas as categorias e sub-categorias, apesar dessa referência não ser dada nas subcategorias. As percentagens existentes mostram a diferença de orçamento do ano anterior para o ano em análise.

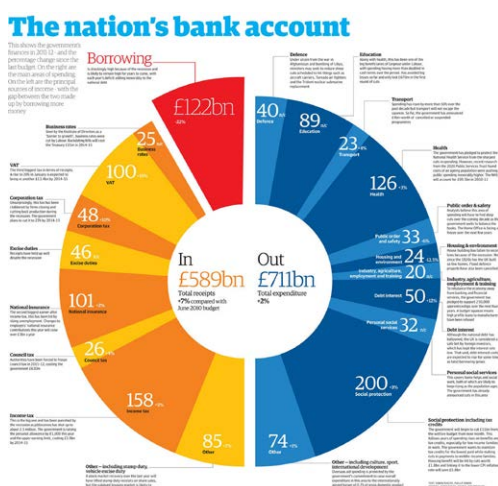


Fig 6. UK Orçamento 2011: rendimentos, empréstimos e gastos. Ilustração: Jenny Ridley para *The Guardian*.
 Fonte: *The Guardian* (2011). Consultado a 22 de Dezembro de 2015 em <http://www.theguardian.com/uk/datablog/2011/mar/23/budget-2011-measures-list#data>

Neste caso não é feita uma comparação do todo, re-

ferente a 100%, com as partes - função principal do gráfico circular - mas sim a representação de dois "todos" e da diferença de valor entre esses. Podemos sim comparar categorias e sub-categorias através do comprimento de arco, tornando-se ligeiramente superior ao gráfico de pizza, onde a comparação é feita pela percepção da área. No entanto, num gráfico de barras a comparação entre as diferentes sub-categorias seria mais efetiva. Porém, para realizar um gráfico de barras com toda a informação, perderíamos a comparação das categorias, a menos que se formasse um outro gráfico sem subcategorias. Neste sentido, o círculo efetiva-se como a melhor hipótese, uma vez que conseguimos fazer a comparação de categorias e subcategorias no mesmo diagrama.

No entanto, uma vez que pode ser bastante difícil comparar diferentes seções do gráfico *donut*, assim como num gráfico de pizza, a comparação das diversas subcategorias é mais difícil de realizar, embora não pareça ser a principal leitura do gráfico. Os gráficos de *donut* são mais utilizados para comparar uma fatia com o todo, em vez de comparar seções individuais entre si, o que não é o caso do exemplo dado. Contudo, notamos que o objetivo principal do gráfico é a comparação das principais categorias entre si, isto é, do orçamento reservado em gastos versus o orçamento obtido de receitas que se percebe na terceira porção, o que garante a sua efetividade.

Em comparação com os gráficos pizza, os gráficos *donut* dispõem o espaço central para colocar mais informação, o que permite também que outros erros sejam feitos como a inserção de outro gráfico *donut* no interior, o que tende a dificultar a leitura dos dados. Todas as advertências referidas sobre os gráficos pizza também se aplicam aos gráficos *donut*: não devem existir mais de 5 ou 6 categorias; a soma das partes do gráfico deve totalizar os 100%; As seções devem ser legendadas porque os ângulos das partes são difíceis de ler.

3.3. Gráfico de bolhas

O círculo é objeto de representação noutro género de gráficos chamados os gráficos de bolhas, onde os dados são representados pelo tamanho do círculo. Estas visualizações de dados têm também sido largamente utilizadas. O original gráfico de bolhas, chamado "bubble chart" é apresentado sobre dois eixos (x, y), como no gráfico de dispersão, ou seja, usa um sistema de coordenadas cartesianas para situar pontos sobre uma grelha onde os eixos X e Y são variáveis distintas, porém, ao contrário de um gráfico de dispersão, a cada ponto é atribuído um rótulo ou categoria (seja exibido ao lado ou numa legenda). Cada ponto traçado representa uma terceira variável pela área do círculo. As cores também podem ser objeto de distinção entre categorias ou na representação de uma variável de dados adicional, assim como o brilho ou a opacidade. O tempo pode ser uma das variáveis num dos eixos

ou ser representado através de uma animação onde os dados vão sendo alterados.

Os gráficos de bolha são normalmente utilizados para comparar e mostrar as relações entre os círculos categorizados, pelo uso do posicionamento e das proporções. O quadro geral do diagrama pode servir também para analisar padrões.

No exemplo seguinte, é feita uma representação de dados sobre o estado da bolsa, em que o eixo horizontal (x) representa a evolução do valor das ações na última semana e o eixo vertical (y) representa a evolução do valor das ações no período de um ano (últimos 12 meses). A terceira variável - área dos círculos - representa o volume das transações realizadas numa semana por cada empresa.

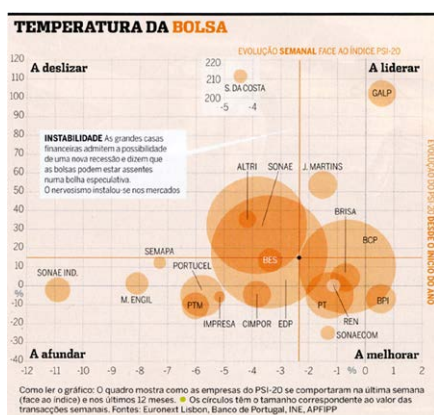


Fig 7. Temperatura da Bolsa, Revista Visão.

Através da análise do gráfico conseguimos ter a percepção dos seguintes dados:

- As empresas com maior volume de transações realizadas numa semana são a EDP, seguida da SONAE e BCP, não sendo clara qual das duas últimas está na frente.
- A empresa em liderança parece ser a GALP, uma vez que, apesar de não ter grande volume de transações, evoluiu positivamente numa semana e anualmente. Ainda assim S. DA COSTA evoluiu cerca de 210% nos últimos 12 meses, apesar da evolução negativa na última semana.
- SONAE IND. é a empresa em pior situação na última semana, com evolução negativa também no último ano.
- SONAE COM é a empresa em pior situação considerando o último ano, apesar de ter evolução positiva no período de uma semana, o que constata que está a melhorar.

As linhas a laranja, horizontal e vertical, representam ainda a média de evolução das empresas, na última semana e nos últimos 12 anos, respetivamente.

O gráfico apresentado permite assim condensar informação complexa, que pode então ser facilmente entendida através da visualização. No entanto, como qualquer matéria complexa, ele requer atenção para se perceber o modo de representação dos dados e perceber a informação. Constata-se que o gráfico

representa de forma evidente os dados em causa e permite retirar grandes relações, inclusive por comparação dos círculos. No entanto não é possível retirar valores precisos do mesmo, uma vez que, na representação do círculo, o mesmo é colocado no sistema cartesiano a partir do seu ponto central que não é facilmente entendido e a recorrência aos valores dos eixos não permite saber exatamente a evolução percentual das empresas, para além de que a área de cada círculo não permite retirar valores concisos do volume de transações e não é facilmente comparada, limitando a comparação desses volumes entre empresas.

3.4. Mapa de bolhas

O mapa de bolhas é similar ao gráfico de bolhas, mas o sistema cartesiano é eliminado e substituído por uma referência geográfica, daí o seu nome. Num mapa de bolhas os círculos são dispostos sobre uma região geográfica onde a área do círculo é proporcional ao seu valor no conjunto de dados. Um exemplo disso é o mapa apresentado abaixo, onde é feita uma representação das emissões de dióxido de carbono por país. À semelhança do exemplo da figura 1, são apresentadas 7 categorias distintas que identificam as seguintes áreas: América do Norte (vermelho); América do Sul (laranja); Europa (roxo); África (castanho); Euroásia (verde); Médio Oriente (amarelo); Ásia-Oceânia (azul).

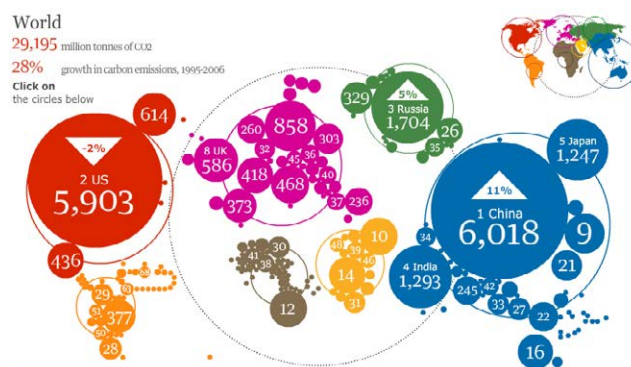


Fig 8. Dados informativos sobre a emissão de CO2 por cada país - Climate change: The Carbon atlas.

Fonte: *The Guardian* (2008). Consultado a 17 de Dezembro de 2015 em <http://www.theguardian.com/global/interactive/2008/dec/09/climatechange-carbonemissions>

Um dos principais problemas de um gráfico de bolhas é a quantidade limitada de dados, o que se comprova também no mapa de bolhas, uma vez que as bolhas excessivamente grandes podem sobrepor-se a outras bolhas e regiões no mapa. No gráfico apresentado apenas temos leitura das bolhas maiores, o que parece ser a intenção final do diagrama - identificar que a China supera os EUA na emissão de CO2 e qual o top 5 dos maiores emissores (1. China; 2. EUA; 3. Rússia; 4. Índia; 5. Japão). Contudo, devido à interatividade do gráfico, ao clicarmos numa das categorias conseguimos visualizar esta individualmente e ter acesso aos dados de todos os círculos, mesmo os mais pequenos

(ao clicar sobre estes). Esta visualização não era possível sem esta face interativa disponível.

Ainda assim a comparação entre a China e os EUA só é possível devido à leitura dos números expostos, caso contrário, a área dos dois círculos, a olho humano, é muito similar tornando mesmo difícil a percepção de que uma das áreas é maior que a outra.

Como Cairo (2012b) explica, o cérebro humano não é efetivo a calcular tamanhos de áreas, é muito melhor a comparar apenas uma dimensão como a largura ou a altura. O cérebro é uma máquina irremediavelmente preguiçosa o que faz com que, quando confrontado com diversas situações como situações de perigo, não perca energia nem tempo a analisar comprimentos, larguras e áreas. Apenas o comprimento é suficiente, por exemplo.

Na percepção do tamanho de um círculo o sistema visual humano tende a comparar alturas e não as áreas. Isto é, quando os valores dos círculos não são apresentados o cérebro fica confuso e não consegue fazer comparações efetivas, como no exemplo acima no caso da China vs EUA. Como Alberto Cairo refere, as bolhas enganam, no sentido que fazem com que se subestimem diferenças.

Cairo (2012b) usa um teste muito simples para evidenciar este facto: coloca dois círculos de diferentes áreas concêntricos (como na representação do império Russo por Playfair) e pede que, com o dado informativo do círculo maior, se estime a informação do círculo menor. As respostas são em média cerca de 50% do tamanho do círculo maior, no entanto, ao colocar os mesmos dados sobre um gráfico de barras demonstra que o anterior círculo menor é apenas cerca de um terço do círculo maior. Desta forma conseguimos perceber a distorção dos dados a que a representação circular pode levar.



Fig 9. "A primeira bolha representa \$80 bilhões. A segunda bolha representa qual percentagem, em relação à primeira? Metade, talvez?" - Alberto Cairo

Fonte: Cairo (2012b). *The Functional Art: An introduction to information graphics and visualization*.

A área de um círculo, ao contrário do seu diâmetro ou circunferência não é proporcional ao seu raio, mas ao quadrado do raio. Dessa forma, se o tamanho dos círculos for escalado segundo o diâmetro, as diferenças de tamanho aparentes entre as bolhas serão não-lineares e falaciosas. Para obter uma escala adequada ponderada, é necessário dimensionar o raio de cada círculo para a raiz quadrada do valor dos dados correspondentes ao tamanho do mesmo. Este problema

de escala pode levar a más interpretações extremas, especialmente quando o intervalo de dados tem uma grande propagação.

Market Capitalization of Société Générale

Billions of dollars

Source: Bloomberg



Figure 2.10. Our friend, the bar chart, comes to the rescue.

Fig 10. "O nosso amigo, gráfico de barras, vem para salvar" - Alberto Cairo

Fonte: Cairo (2012b). *The Functional Art: An introduction to information graphics and visualization*.

O uso "desenfreado" do círculo como objeto de representação de números faz com que estas situações aconteçam diversas vezes e, dessa forma, se transmita informação de forma errada, para além de que a escolha do círculo seja muitas vezes feita sem considerar a sua função e as suas limitações. No entanto, não parece haver uma razão ponderada para o seu uso excessivo.

4. Conclusão

O uso do círculo em representações gráficas de visualização de dados parece ser uma constante. Porém o círculo é uma forma que apresenta muitas limitações como podemos verificar. Não é clara a razão pelo seu fascínio, mas o seu uso é claramente sobrestimado. A razão mais óbvia parece ser um apelo visual tendencial para a forma circular. O círculo, pelas suas linhas circulares sem início nem fim, forma um foco que parece aproximar e atrair. A verdade é que na percepção da forma percebemos que formas de linhas retas se tornam mais rígidas como o quadrado e o retângulo que garantem estabilidade, poder e racionalidade, enquanto que as linhas curvas lidam com o lado emocional uma vez que são formas ligadas à natureza que tranquilizam, oferecem segurança e conexão e sugerem comunidade, integridade e perfeição. São formas mais quentes, ao contrário das formas de linhas retas, mais frias e que parecem distanciar. Talvez seja esta a justificação pelo fascínio à volta do círculo.

Porém, a escolha dos elementos visuais para comunicar dados informativos não deve ser completamente livre. Deve ter em consideração a forma como o cérebro humano processa informações. A percepção tem capacidades limitadas e segue determinadas regras, truques e atalhos. Tendo esses dados em consideração, é possível antecipar e criar os gráficos mais convenientes para cada caso. Deve, por isso, ser feita uma utilização do círculo com consciência das suas limitações e funções (Cairo, 2012a).

William S. Cleveland e Robert McGill, dois investigadores dos laboratórios AT&T Bell Labs, publicaram um artigo, no ano 1984, no *Journal of the American Statistical Association* intitulado "Graphical perception: theory, experimentation, and application to the deve-

lopment of graphical methods", no qual propuseram um guia para eleger a forma correta de representação de dados segundo a função que cada gráfico tem. No guia referem que a posição sobre um eixo comum, a posição sobre eixos não alinhados e a longitude, direção e ângulo, por ordem de importância, permitem comparações mais precisas; enquanto que a área, volume, curvatura, intensidade e saturação de cor são mais precisos na percepção de padrões e tendências gerais (ver figura 11) (cit. in Cairo, 2012a).

Cleveland e McGill dizem que a escolha pela forma gráfica que envolve tarefas perceptivas elementares que levam a julgamentos mais precisos do que outras formas gráficas (com a mesma informação quantitativa) irá resultar numa melhor organização e aumenta as possibilidades de uma percepção correta de padrões e comportamentos (cit. in Cairo, 2012a).

Quanto mais detalhadas e precisas são as operações que os leitores devem completar num gráfico, mais alta na hierarquia deve estar a forma utilizada para representar os dados. Em outras palavras, e segundo Cairo (2012a) um gráfico de barras é sempre melhor que um gráfico de bolhas se o objetivo central do designer é facilitar comparações precisas entre números. A escala de Cleveland e McGill também permite verificar a razão pela qual os gráficos de pizza quase nunca são úteis (Cairo, 2012a).

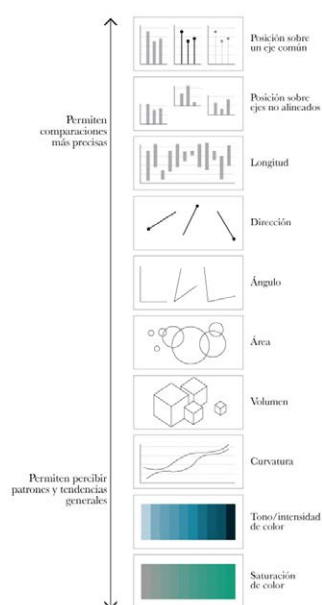


Fig 11. Guia para eleger forma correta de representação de dados por Cleveland e McGill.

Fonte: Cairo (2012a). *Blog El País: Periodismo com Futuro*.

Quanto mais alta está uma representação na escala de Cleveland e McGill, mais exatas serão as comparações que os leitores são capazes de fazer. No entanto, nem sempre o objetivo de um gráfico é a exatidão mas sim a percepção dos padrões gerais, as concentrações, tendências, etc. Nestes casos, quanto mais baixa a representação está na mesma escala mais efetiva poderá ser (Cairo, 2012a).

Não é errado representar informação através de círculos. A infografia deve ser uma ferramenta na qual as formas devem ser escolhidas segundo os objetivos ou intenções definidas. E o círculo é uma forma capaz de fazer representações efetivas, dependendo da intenção e forma como que é utilizado. Existem muitas representações de círculos na infografia que comprovam que não é dispensado o tempo suficiente para

pensar na melhor forma de representar a informação. É necessária a reflexão sobre qual o melhor tipo de gráfico para representar certo tipo de informação de forma a retirar o maior partido de cada forma. Os formatos padrão de visualização como o gráfico de barras ou o gráfico de linhas tornaram-se muito familiares, e arriscam ser ignorados muito rapidamente pelo leitor. No entanto, o uso pela forma circular, como o gráfico de fatias ou de bolhas (também formas padrão) que parecem atrair maior atenção, podem distorcer os dados.

O uso por formas padrão de visualização, quer circulares ou não, parece ser uma solução elegante e fácil de seguir, mas os profissionais de visualização de dados não devem confiar demasiado em formas estabelecidas no final de 1700. Pelo contrário, deve haver uma procura por novas formas de representação de dados que desafiem os limites da forma num esforço de envolver mais efetivamente o leitor.

Bibliografia.

Cairo, Alberto (2012a). Cómo eleger el tipo de infográfico más adecuado para cada historia. *Blog El País: Periodismo com Futuro*. Disponível em: <http://blogs.elpais.com/periodismo-con-futuro/2012/02/elegirgrafico.html>

Cairo, Alberto (2012b). *The Functional Art: An introduction to information graphics and visualization*. Disponível em: <https://books.google.pt/books?id=xwjhh6Wu-VUC&pg=PT53&lpg=PT53&dq=The+Bubble+Plague+Alberto+Cairo&source=bl&ots=nLwGEqKQ5e&sig=e7C22ygP2Sz1KOL-t5-YyXS6qc-E&hl=pt-PT&sa=X&ved=0ahUKewiEsPuinOH-JAhVCRxoKHWbiAnwQ6AEIjJAB#v=onepage&q&f=true>

Gregory, Richard (1997). *Eye and Brain: The psychology of seeing*. OUP Oxford; 5 edition: England.

Klanten, R., Bourquin, N., Tissot, T. & Ehmann, S. (Eds.) (2008). *Data Flow: Visualising Information in Graphic Design*. Gestalten: Berlin.

Spence, Ian (2005). No Humble Pie: The Origins and Usage of a Statistical Chart. *University of Toronto*. Disponível em: <http://www.psych.utoronto.ca/users/spence/Spence%202005.pdf>