

91%

Văn bản đơn giản

	A	B	C
Hạng mục 1	15%	22%	42%
Hạng mục 2	40%	36%	20%
Hạng mục 3	35%	17%	34%
Hạng mục 4	30%	29%	26%
Hạng mục 5	55%	30%	58%
Hạng mục 6	11%	25%	49%

Bảng biểu

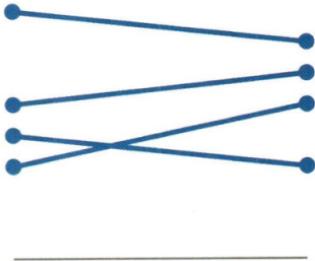
Biểu đồ phân tán



Biểu đồ đường

	A	B	C
Hạng mục 1	15%	22%	42%
Hạng mục 2	40%	36%	20%
Hạng mục 3	35%	17%	34%
Hạng mục 4	30%	29%	26%
Hạng mục 5	55%	30%	58%
Hạng mục 6	11%	25%	49%

Bản đồ nhiệt

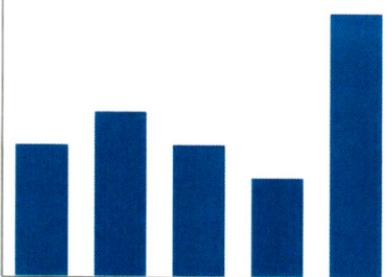


Biểu đồ đường xiên

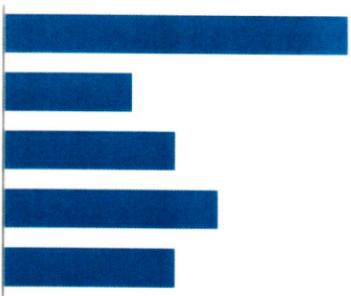
Hình 2.1 Hình ảnh trực quan tôi thường sử dụng nhất



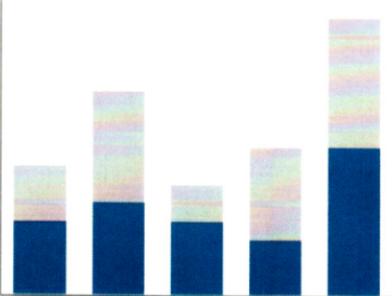
tinyurl.com/magnacappuccino



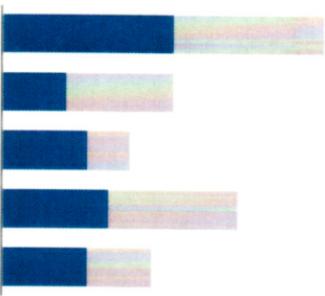
Biểu đồ cột đứng



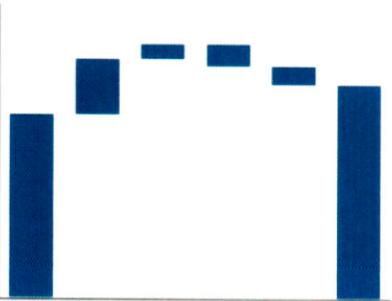
Biểu đồ cột ngang



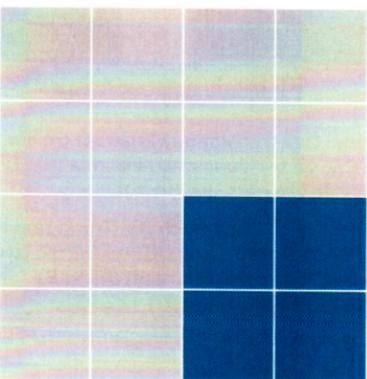
Biểu đồ cột chồng dạng đứng



Biểu đồ cột chồng nằm ngang



Biểu đồ thác nước



Biểu đồ vùng



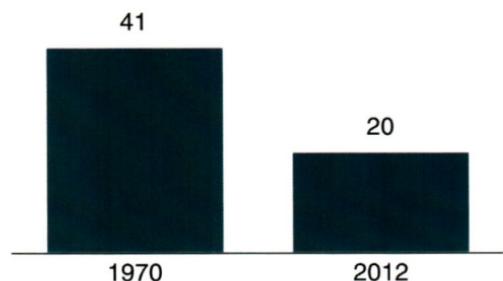
Văn bản đơn giản

Khi bạn chỉ có một hoặc hai thông số cần chia sẻ, một văn bản đơn giản thôi cũng quá đủ để truyền đạt rồi. Hãy nghĩ đến cảnh bạn chỉ dùng độc nhất một con số – phóng nó lên càng to càng tốt – rồi thêm vào vài câu từ bổ nghĩa nhằm nêu rõ ý của bạn. Ngoài nguy cơ gây hiểu lầm, bảng biểu hoặc đồ thị chỉ có một hoặc hai thông số sẽ khiến những con số đó với đi phần nào “tiếng nói” của mình. Khi bạn cần truyền đạt một hoặc hai thông số, hãy xem xét sử dụng trực tiếp thông số đó.

Để mô tả rõ hơn khái niệm này, chúng ta hãy cùng xem qua ví dụ. Một biểu đồ tương tự như Hình 2.2 phác họa một bản báo cáo của Trung tâm Nghiên cứu Pew về những người nội trợ vào tháng 4 năm 2014.

Trẻ em có mẹ là người nội trợ “truyền thống”

Tỷ lệ % trẻ em có mẹ làm nội trợ
và bố đi làm



Ghi chú: Dựa trên kết quả khảo sát trẻ em dưới 18 tuổi. Các bà mẹ được phân loại dựa trên tình trạng công việc vào năm 1970 và 2012.

Nguồn: Kết quả phân tích của Trung tâm Nghiên cứu Pew dựa trên Cơ sở Dữ liệu Dân số từ Cuộc khảo sát Toàn dân Hiện hành tháng Ba (IPUMS-CPS), vào năm 1971 và 2013

Thông tin từ TRUNG TÂM NGHIÊN CỨU PEW

Hình 2.2 Đồ thị nguyên mẫu về những bà mẹ nội trợ



tinyurl.com/magncaffuccino

Sự thật là bạn có một, hai thông số không đồng nghĩa với việc cần phải lập biểu đồ đâu! Trong Hình 2.2, hai thông số này chiếm quá nhiều không gian lắn câu từ. Thậm chí biểu đồ này còn chẳng giúp được gì nhiều trong việc lý giải các con số nữa (đó là chưa kể đến vị trí sắp xếp các nhãn dữ liệu nằm ra ngoài cột có thể làm đối tượng mục tiêu nhận thức sai về chiều cao tương đối vì 20 chưa bằng một nửa 41, gây phản trực quan).

Trong trường hợp này, một câu đơn giản như “*20% trẻ em năm 2012 có mẹ là nội trợ so với con số 41% vào năm 1970*” là quá đủ.

Nhờ cách đó, trong một bản thuyết trình hay báo cáo, hình ảnh trực quan của bạn sẽ phần nào giống với Hình 2.3 sau đây.

20%

20% trẻ em năm 2012 có
mẹ là người nội trợ truyền thống
so với con số 41% vào năm 1970

Hình 2.3 Văn bản đơn giản thể hiện dữ liệu về những bà mẹ nội trợ

Ngoài ra, trong ví dụ cụ thể này, bạn có thể xem xét trình bày thông số dưới một dạng hoàn toàn khác. Ví dụ, bạn có thể chỉnh sửa lại văn bản theo kiểu tăng/giảm tỷ lệ: “Số lượng trẻ em có mẹ làm nội trợ truyền thống đã giảm hơn 50% giữa năm 1970 và 2012”. Tuy nhiên, tôi khuyên các bạn cũng nên cẩn thận vì việc giảm bớt thông số xuống chỉ còn một thông số có thể gây lạc mất bối cảnh. Trong trường hợp này, tôi nhận thấy rằng số lượng các thông số hiện có (20% và 41%) là quá hữu ích trong việc lý giải và thấu hiểu sự chuyển biến này rồi.



Khi bạn muốn truyền đạt một hoặc hai thông số: hãy sử dụng trực tiếp chính các thông số đó.

Tuy nhiên, khi cần phải truyền đạt nhiều dữ liệu hơn, thường thì bạn sẽ dùng đến bảng biểu hoặc đồ thị. Bạn cần phải nhớ rằng người ta sẽ tương tác khác nhau với hai mẫu hình ảnh trực quan này. Chúng ta sẽ cùng nghiên cứu riêng biệt từng mẫu, đồng thời xem qua một loạt các kiểu mẫu khác nhau cũng như trường hợp áp dụng.

Bảng biểu

Bảng biểu tương tác với hệ thống ngôn từ của chúng ta, nghĩa là ta phải đọc chúng. Thường khi gặp một bảng biểu, tôi sẽ giơ ngón trỏ ra: vừa chỉ, vừa đọc từng dòng, từng cột hay đang so sánh các giá trị với nhau. Bảng biểu tuyệt thế đấy! Chúng có chức năng truyền đạt thông tin đến từng đối tượng mục tiêu tùy theo mối quan tâm của họ. Nếu như cần truyền đạt nhiều đơn vị đo lường khác nhau, bạn có thể dễ dàng làm thế với bảng biểu thay vì đồ thị.

Bảng biểu khi thuyết trình trực tiếp

Sử dụng bảng biểu khi thuyết trình không phải lúc nào cũng là ý hay. Vì khi bạn đọc, đối tượng mục tiêu sẽ không lắng nghe hay để tâm đến những gì bạn nói. Khi bản thân đang sử dụng bảng biểu trong các buổi thuyết trình hay báo cáo, hãy tự vấn rằng: mình đang cố gắng trình bày điều gì đây? Nhiều khả năng bạn sẽ tìm ra một cách hay hơn để làm rõ và trực quan hóa một hay nhiều điểm đáng lưu ý. Lỡ như bạn cảm thấy mình bị thất thoát quá nhiều nội dung khi làm thế, hãy cân nhắc chèn một bảng biểu đầy đủ thông tin vào phần phụ lục kèm theo đường dẫn hay mục tham khảo để đáp ứng nhu cầu đối tượng mục tiêu.

Hãy nhớ thiết kế bảng biểu sao cho đừng quá nổi bật và hãy để phần dữ liệu là nhân vật chính. Đừng để những khung viền dày cộm hay bóng đốp quá đậm trở thành trọng tâm. Thay vào đó, hãy dùng những khung viền mỏng hoặc để trắng hoàn toàn nhằm phân biệt phần nội dung trong bảng biểu.



Hãy cùng xem qua những mẫu bảng biểu trong Hình 2.4. Các bạn nhớ vừa xem, vừa để ý xem độ nổi của dữ liệu so với khung viền ở hình thứ hai và ba (viền mỏng, viền tối giản).

Viền đậm

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Viền mỏng

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Viền tối giản

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Hình 2.4 Viền kẻ bảng

Nhiệm vụ của khung viền chỉ là làm nền cho phần dữ liệu. Các bạn hãy nghĩ cách làm chúng chìm đi bằng cách tô xám hoặc xóa hẳn. Dữ liệu là trọng tâm, không phải khung viền.

Gợi ý bạn đọc

 Để biết thêm thông tin về thiết kế bảng biểu, hãy nhớ tìm mua cuốn sách *Show Me the Numbers* (tạm dịch: *Thông số của tôi đâu*) của Stephen Few. Ông dành hẳn ra một chương chỉ để trình bày về thiết kế bảng biểu, cùng với những lời bàn luận về yếu tố xây dựng bảng biểu và những ví dụ thực tiễn tốt nhất về chủ đề này.

Còn bây giờ, chúng ta hãy xét đến một trường hợp vô cùng đặc thù của bảng biểu: Bản đồ nhiệt.



tinyurl.com/magnacappuccino

Bản đồ nhiệt⁹

Một cách tiếp cận thú vị cho phép bạn dung hòa những thông tin cần thiết trên bảng biểu nhưng vẫn tận dụng được tính trực quan chính là thông qua bản đồ nhiệt. Chúng ta có thể dùng bản đồ nhiệt để trực quan hóa dữ liệu dưới dạng bảng biểu. Điểm nổi bật ở đây là thay vì nhấn mạnh (hoặc thêm thắt vào) những con số, bạn sử dụng những ô màu nhãm truyền tải tầm mức của thông số đó.

Nhìn Hình 2.5, chúng ta sẽ thấy hai hình ảnh hiển thị cùng một nhóm các thông số, một bên ở dạng bảng biểu, một bên ở dạng bản đồ nhiệt.

Bảng biểu

	A	B	C
Hạng mục 1	15%	22%	42%
Hạng mục 2	40%	36%	20%
Hạng mục 3	35%	17%	34%
Hạng mục 4	30%	29%	26%
Hạng mục 5	55%	30%	58%
Hạng mục 6	11%	25%	49%

Bản đồ nhiệt

	A	B	C
Hạng mục 1	15%	22%	42%
Hạng mục 2	40%	36%	20%
Hạng mục 3	35%	17%	34%
Hạng mục 4	30%	29%	26%
Hạng mục 5	55%	30%	58%
Hạng mục 6	11%	25%	49%

Hình 2.5 Hai góc nhìn của cùng một dữ liệu

Ở bảng biểu Hình 2.5, bạn phải tự đọc toàn bộ dữ liệu. Tôi nhận thấy rằng mình phải đảo mắt liên tục khắp các hàng và cột để hiểu xem bản thân đang nhìn cái gì, ở đâu cho thấy thông số cao hay thấp hơn, đồng thời cũng phải tự sắp xếp thứ hạng của các đề mục trong bảng biểu.

Để đỡ tốn chất xám, chúng ta có thể áp dụng khái niệm **độ bão hòa màu¹⁰** nhằm tăng thêm tính trực quan, giúp đôi mắt và não bộ nhanh chóng hướng đến điểm trọng yếu khả thi. Trong hình ảnh thứ hai bên phải có tên “Bản đồ nhiệt”, độ bão hòa của màu xanh dương càng đậm, thông số càng lớn. Điều này đã giúp cho quy trình xác định thông số trọng yếu trong nhóm này – thông số thấp nhất (11%) và thông số cao nhất (58%) – dễ dàng và nhanh hơn nhiều so với bảng biểu nguyên mẫu bên trái – vốn không mang tính trực quan để giúp điều hướng sự chú ý.

9. Bản đồ nhiệt: Dạng biểu đồ thể hiện cường độ của một số liệu, sự việc bằng màu sắc đậm, nhạt

10. Độ bão hòa màu liên quan đến cường độ màu trong ảnh. Về mặt kỹ thuật, nó là biểu thức của bảng thông ánh sáng từ một nguồn tinyurl.com/magncaffuccino

Ứng dụng lập biểu đồ (như Excel) thường sẽ tích hợp sẵn chức năng giúp bạn dễ dàng tạo nên những định dạng như trong Hình 2.5. Mỗi khi lập bảng biểu, hãy nhớ thêm vào những chú thích nhằm giúp đối tượng mục tiêu lý giải đúng các dữ liệu (trong trường hợp này là dòng THẤP – CAO phía trên bản đồ nhiệt với màu sắc tương ứng cùng định dạng phù hợp phục vụ mục đích trên).

Tiếp theo, hãy cùng bàn đến những hình ảnh trực quan mà chúng ta thường sẽ nghĩ đến trước nhất mỗi khi nhắc đến truyền tải thông tin qua dữ liệu: đồ thị.

Đồ thị

Nếu như bảng biểu tương tác với hệ thống ngôn từ của chúng ta, đồ thị sẽ tác động đến thị giác, cơ quan giúp xử lý thông tin nhanh hơn. Điều này đồng nghĩa với việc một đồ thị được thiết kế tỉ mỉ sẽ truyền đạt thông tin nhanh chóng hơn so với một bảng biểu tương tự. Như tôi đã trình bày ở phần mở đầu chương này, có hàng đống những loại biểu đồ tràn lan khắp nơi. Tuy vậy, điều đáng mừng là chỉ cần một vài mẫu thô sơ cũng đủ đáp ứng nhu cầu thường ngày của bạn rồi.

Những loại đồ thị tôi thường dùng rơi vào bốn định dạng sau: chấm, đường, cột và vùng. Chúng ta sẽ nghiên cứu những định dạng này sâu hơn, đồng thời bàn luận về các thể loại con của những dạng đồ thị mà tôi vẫn đang dùng hàng ngày. Ngoài ra, bạn cũng sẽ được tiếp cận với những áp dụng thực tế và ví dụ cho mỗi loại.

Biểu đồ hay đồ thị?

Một số người phân biệt rõ ràng giữa biểu đồ và đồ thị. Thông thường, “biểu đồ” mang tính phổ quát hơn và “đồ thị” là một dạng của biểu đồ (các loại biểu đồ khác gồm có bản đồ và giản đồ). Tôi thì không hay phân biệt như thế vì gần như toàn bộ những biểu đồ tôi dùng thường nhật đều là đồ thị cả. Xuyên suốt cuốn sách này, tôi sẽ sử dụng luôn phiên hai từ *biểu đồ* và *đồ thị*.

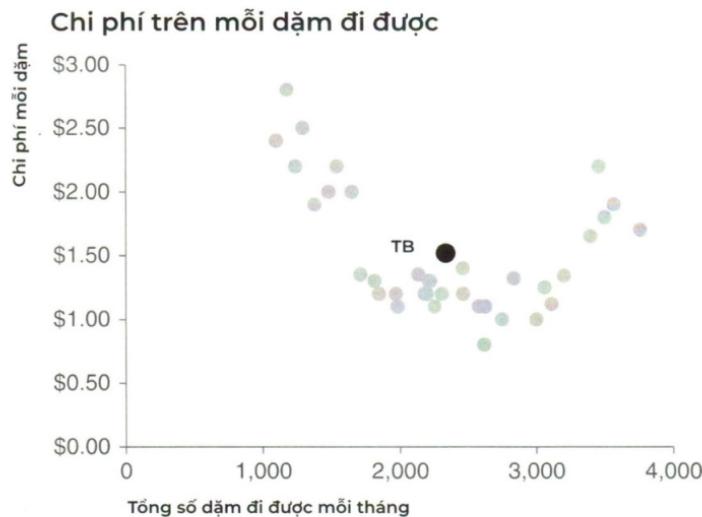


Dạng điểm

Biểu đồ phân tán

Biểu đồ phân tán thường hữu dụng trong việc thể hiện mối liên hệ giữa hai đối tượng bởi vì chúng cho phép bạn mã hóa đồng thời những dữ liệu trên trục x và trục y để xác định sự tồn tại cũng như định nghĩa một mối liên hệ nào đó. Biểu đồ này thường xuất hiện nhiều trong lĩnh vực khoa học (có lẽ vì lý do này mà những ai không quen thuộc với chúng sẽ cho rằng dạng biểu đồ này rất phức tạp). Dù ít người dùng, vẫn tồn tại những áp dụng thực tiễn của biểu đồ phân tán trong giới kinh doanh.

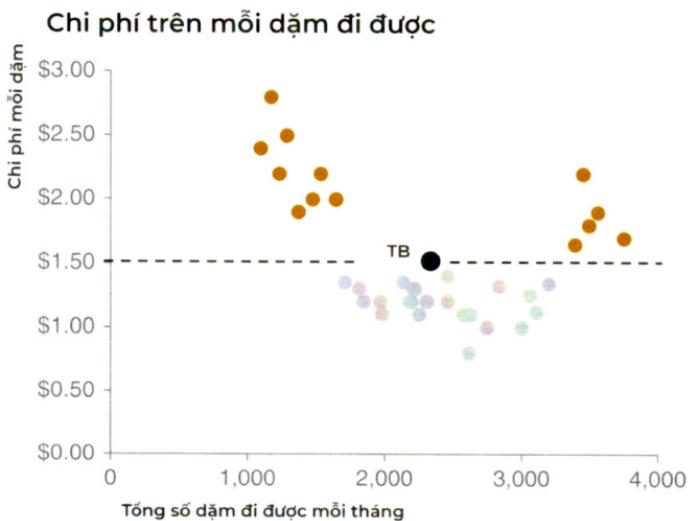
Lấy ví dụ như sau: Tưởng tượng rằng bạn là quản lý của một hãng xe buýt và đang mong muốn biết mối liên hệ giữa quãng đường di chuyển và chi phí mỗi dặm. Biểu đồ phân tán sẽ trông như Hình 2.6 sau.



Hình 2.6 Biểu đồ phân tán

Nếu như chỉ tập trung vào những trường hợp có chỉ số chi phí mỗi dặm cao hơn mức trung bình, chúng ta chỉ cần thêm thắt một chút vào biểu đồ phân tán là sẽ giúp đối tượng mục tiêu nhanh chóng nhận biết hơn. Kết quả sẽ phần nào giống như Hình 2.7.





Hình 2.7 Biểu đồ phân tán qua chỉnh sửa

Qua Hình 2.7, chúng ta có thể quan sát được chỉ số chi phí mỗi dặm sẽ cao hơn mức trung bình khi quãng đường di chuyển ít hơn 1.700 dặm và nhiều hơn 3.300 dặm như biểu đồ mẫu trên. Chúng ta sẽ bàn về cách và lý do lựa chọn định dạng trong những chương tiếp theo.

Dạng đường

Biểu đồ đường là định dạng thông dụng nhất để thể hiện những dữ liệu tiếp diễn. Từ những điểm được liên kết thông qua các đường nối, biểu đồ này ngụ ý cho thấy mối liên hệ giữa những điểm t่อ tùng chừng như chẳng liên quan gì đến nhau khi ở dạng dữ liệu theo mục (một cơ sở dữ liệu được phân loại ra thành từng hàng mục khác nhau). Thông thường, những dữ liệu tiếp diễn sẽ được đo lường bằng đơn vị thời gian: ngày, tháng, quý, hay năm.

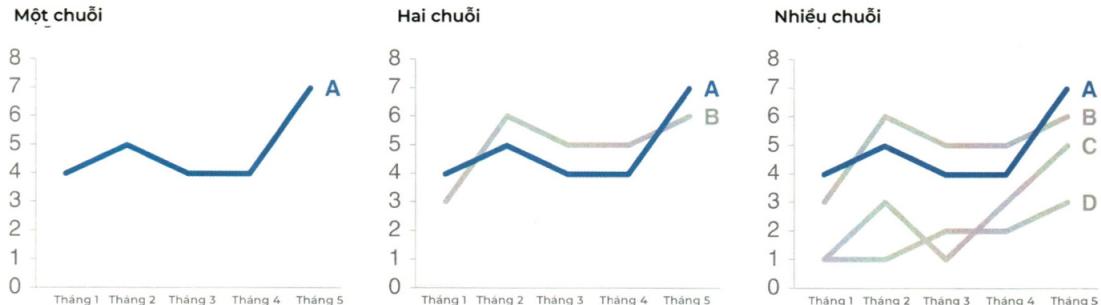
Ở định dạng này, tôi thường hay dùng đến hai loại biểu đồ: biểu đồ đường cắn bản và biểu đồ đường xiên.

Biểu đồ đường

Biểu đồ đường có thể thể hiện một, hai hay nhiều chuỗi dữ liệu khác nhau như Hình 2.8.



tinyurl.com/magnacappuccino



Hình 2.8 Biểu đồ đường

Các bạn lưu ý rằng khi vẽ đường biểu đồ theo các mốc thời gian trên trục x, các thông số thời gian đó bắt buộc phải nhất quán về mặt khoảng cách. Gần đây, tôi có xem qua một biểu đồ đường với thông số trên trục x biểu thị các mốc thời gian theo thập niên từ những năm 1900 trở về sau (1910, 1920, 1930...) rồi lại đột ngột chuyển sang hình thức thường niên sau năm 2010 (2011, 2012, 2013, 2014). Điều này đồng nghĩa với việc khoảng cách giữa các mốc thập niên và các mốc thường niên tương tự nhau. Thật là rối rắm khi trình bày dữ liệu như thế. Hãy nhớ rằng khoảng cách các mốc thời gian phải nhất quán như nhau.

Biểu thị thông số trung vị¹¹ trên khoảng giá trị¹² trong biểu đồ đường

Trong một số trường hợp, một đường trong biểu đồ của bạn có thể đại diện cho thông số đặc trưng nào đó, như thông số trung bình hoặc ước lượng¹³ theo như một dự đoán nào đó. Nếu như muốn thể hiện một khoảng giá trị (hoặc khoảng tin cậy¹⁴, tùy trường hợp), bạn có thể thêm trực tiếp lên biểu đồ bằng cách trực quan hóa khoảng giá trị đó. Ví dụ, Hình 2.9 cho thấy khoảng thời gian chờ đợi ngắn nhất, trung bình và dài nhất tại những điểm kiểm tra hộ chiếu thuộc các sân bay trong vòng 13 tháng.

11. Số trung vị: Trong lý thuyết xác suất và thống kê, số trung vị (tiếng Anh: median) là một số tách giữa nửa lớn hơn và nửa bé hơn của một mẫu, một quần thể, hay một phân bố xác suất.

12. Khoảng giá trị: Khoảng giá trị là sự chênh lệch giữa giá trị lớn nhất và giá trị nhỏ nhất trong một tập giá trị.

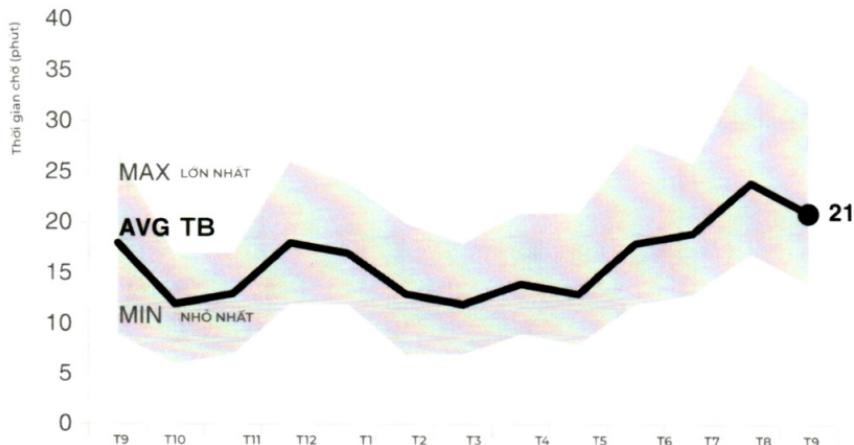
13. Ước lượng: Trong thống kê, một ước lượng là một giá trị được tính toán từ một mẫu thử và người ta hy vọng đó là giá trị tiêu biểu cho giá trị cần xác định trong dân số.

14. Khoảng tin cậy: Là đại lượng tính bằng phần trăm cho biết độ tin cậy của một số liệu



Thời gian chờ xác thực hộ chiếu

Trong 13 tháng qua



Hình 2.9 Thể hiện thông số trung bình trong một phạm vi của biểu đồ đường

Biểu đồ đường xiên

Biểu đồ đường xiên vô cùng hữu ích khi bạn muốn so sánh hai khoảng hoặc mốc thời gian, đồng thời nhanh chóng thể hiện sự tăng/giảm giữa hai bên hoặc chỉ ra những điểm khác nhau thuộc nhiều phạm trù giữa hai điểm dữ liệu.

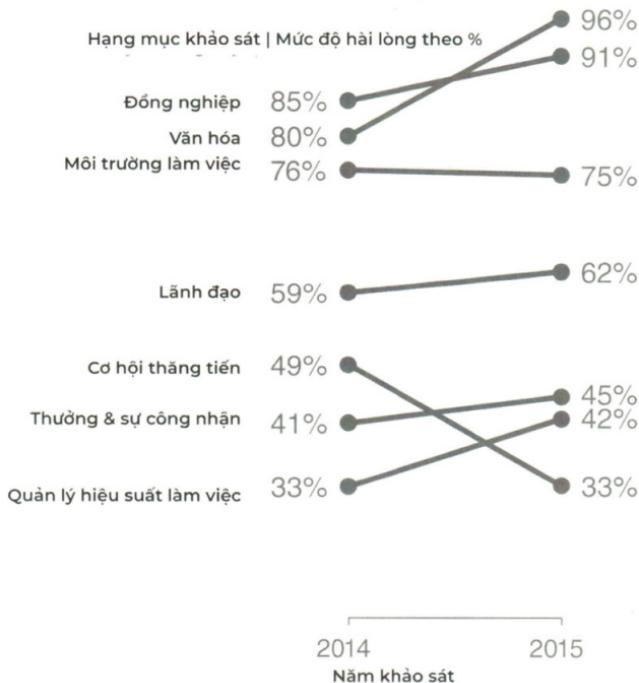
Phương thức giải thích rõ ràng nhất giá trị và cách ứng dụng biểu đồ đường xiên là bằng những ví dụ cụ thể. Hãy tưởng tượng bạn đang phân tích và truyền đạt dữ liệu từ kết quả cuộc khảo sát phản hồi của nhân viên gần đây. Để thể hiện sự thay đổi tương quan giữa kết quả hai cuộc khảo sát năm 2014 và 2015, chúng ta có thể tham khảo Hình 2.10.

Biểu đồ đường xiên chứa đựng vô vàn thông tin. Ngoài những thông số giá trị (hoặc điểm) tuyệt đối, biểu đồ này còn có những đường liên kết chúng lại giúp bạn thấy được mức độ biến thiên tăng/giảm của chúng (qua những đường xiên hoặc thẳng) mà không cần phải giải thích chức năng hay định nghĩa “mức độ biến thiên” là gì – chỉ cần nhìn thôi, đối tượng mục tiêu cũng đủ hiểu.



tinyurl.com/magnacappuccino

Phản hồi từ nhân viên theo thời gian



Hình 2.10 Biểu đồ đường xiên

Mẫu biểu đồ đường xiên

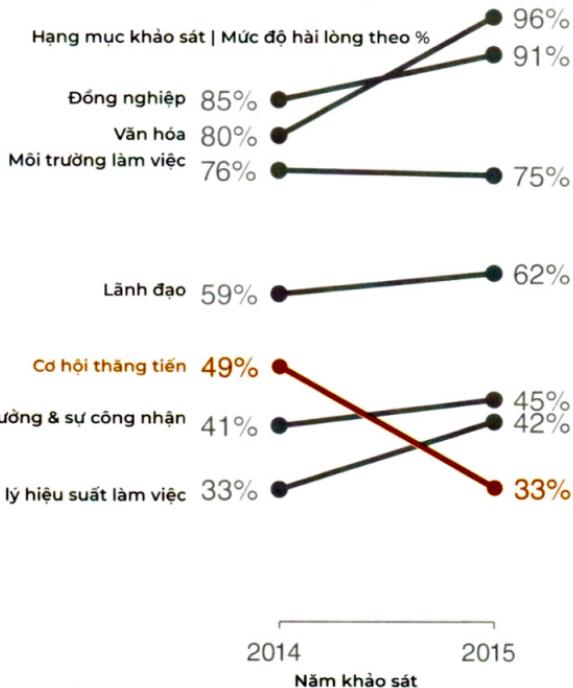
Có thể các bạn sẽ cần một chút kiên nhẫn khi tạo lập biểu đồ đường xiên vì thường ít có ứng dụng nào bao gồm chúng như một đồ thị tiêu chuẩn cả. Các bạn có thể tải bản mẫu trên Excel cùng với một ví dụ về biểu đồ đường xiên và hướng dẫn tùy chỉnh với mục đích sử dụng cá nhân qua địa chỉ: storytellingwithdata.com/slopegraph-template.

Biểu đồ đường xiên có hữu ích hay không cũng tùy vào trường hợp bạn sử dụng chúng. Nếu như có quá nhiều đường nằm chồng lên nhau, biểu đồ có thể kém hữu hiệu đi hẳn. Tuy nhiên, trong vài trường hợp, bạn vẫn có thể làm nổi bật từng lúc từng chuỗi dữ liệu một. Ví dụ, chúng ta có thể khiến đối tượng mục tiêu phải chú ý vào một phạm trù cụ thể nào đó đang giảm dần theo thời gian từ chính ví dụ trên.



tinyurl.com/magnacappuccino

Phản hồi từ nhân viên theo thời gian



Hình 2.11 Biểu đồ đường xiên qua chỉnh sửa

Ở Hình 2.11, chúng ta chú ý ngay đến sự suy giảm ở phạm trù “Phát triển sự nghiệp” trong khi những phạm trù còn lại vẫn được giữ nguyên, không cần cố làm nổi bật quá. Chúng ta sẽ đề cập chiến lược này sau ở phần những đặc tính nhận thức trong vô thức trong Chương 4.

Bất chấp biểu đồ dạng đường rất hữu ích trong việc trình bày dữ liệu theo thời gian, tôi lại thường hay dùng đến biểu đồ dạng cột để minh họa dữ liệu tùy theo hạng mục. Chúng giúp sắp xếp các thông tin thành từng nhóm khác nhau.

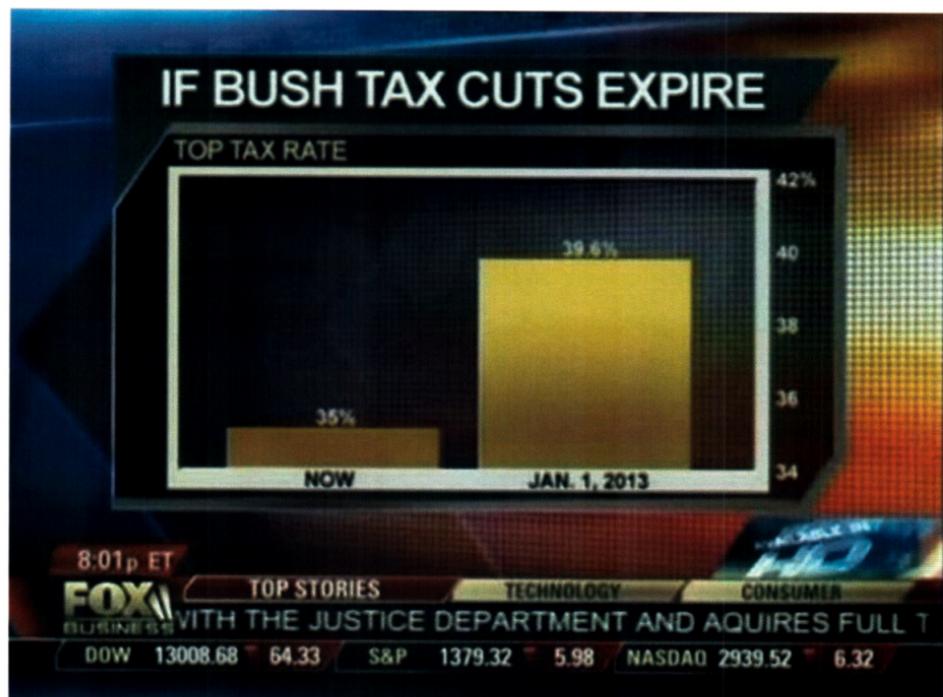


Dạng cột

Thông thường người ta ít dùng đến biểu đồ dạng cột vì chúng quá phổ biến. Thật sai lầm. Lẽ ra chúng ta nên tận dụng biểu đồ dạng cột vì *chính lý do đó*, bởi như vậy sẽ giúp đối tượng mục tiêu đỡ phải suy nghĩ quá nhiều. Thay vì loay hoay tìm cách đọc biểu đồ, đối tượng mục tiêu sẽ động não xét xem đâu là thông tin đáng lưu tâm từ chính biểu đồ đó.

Biểu đồ dạng cột cũng thân thiện với thị giác nữa. Cặp mắt có xu hướng so sánh điểm cuối của các cột cho nên chúng ta có thể dễ dàng nhận thấy hạng mục nào có quy mô lớn nhất, nhỏ nhất cũng như mức độ khác biệt giữa các hạng mục. Nhưng hãy nhớ rằng chính vì xu hướng so sánh điểm cuối giữa các cột với nhau, giá trị của đường cơ sở (điểm giao nhau giữa hai trục x và y) phải luôn bằng 0 nhằm tránh sai lệch khi đối chiếu.

Chúng ta hãy xem qua Hình 2.12 trên kênh Fox News sau đây.



Hình 2.12 Biểu đồ cột trên kênh Fox News



tinyurl.com/magnacappuccino

Như ví dụ trên, hãy tưởng tượng bản thân quay về mùa thu năm 2012. Bây giờ, chúng ta tự ngẫm xem mọi chuyện sẽ ra sao nếu chính sách cắt giảm thuế của tổng thống Bush hết hiệu lực. Ở cột bên trái, ta nhận thấy tỷ lệ thuế ở mức cao nhất hiện là 35%, còn ở bên khác, con số sẽ là 39.6% vào ngày 1 tháng Giêng.

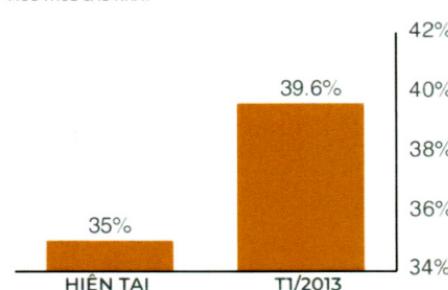
Khi nhìn vào biểu đồ này, bạn sẽ cảm thấy thế nào khi chính sách cắt giảm thuế có khả năng hết hiệu lực trong tương lai? Có lẽ sẽ lo sốt vó lên vì mức độ gia tăng quá khủng khiếp chăng? Giờ thì nhìn kỹ vào nào.

Các bạn chú ý vào con số phía dưới cùng bên phải của trục ngang là 34 thay vì 0. Vậy nên, theo lý thuyết, biểu đồ này ắt hẳn phải còn kéo dài xuống dưới cùng nữa. Thực chất, nếu vẽ biểu đồ theo cách này, mức độ gia tăng phải đến 460% (chiều cao của cột bằng $35 - 34 = 1$ và $39.6 - 34 = 5.6$, nên $(5.6 - 1) / 1 = 460\%$). Còn nếu chúng ta vẽ biểu đồ với giá trị đường cơ sở bằng 0 thì chiều cao các cột sẽ được thể hiện chính xác (35 và 39.6), đồng thời chúng ta sẽ thấy được mức độ gia tăng thật sự là 13% ($(39.6 - 35) / 35$). Hãy cùng so sánh hai biểu đồ cột nằm cạnh nhau trong Hình 2.13.

Không có đường cơ sở: như hình nguyên bản

NẾU CHÍNH SÁCH GIẢM THUẾ CỦA TỔNG THỐNG BUSH ĐÁO HẠN

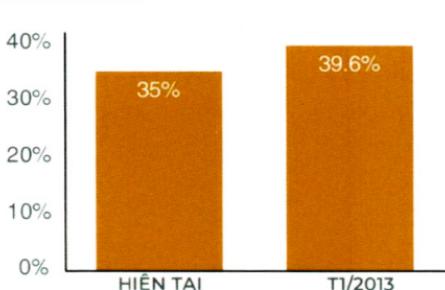
MỨC THUẾ CAO NHẤT



Có đường cơ sở: đáng lẽ phải thế này

NẾU CHÍNH SÁCH GIẢM THUẾ CỦA TỔNG THỐNG BUSH ĐÁO HẠN

MỨC THUẾ CAO NHẤT



Hình 2.13 Biểu đồ cột buộc phải có đường cơ sở

Trong Hình 2.13, mức độ gia tăng khủng khiếp ở hình bên trái đã được giảm đi đáng kể ở hình còn lại sau khi đã được "tái thiết" phù hợp hơn. Có lẽ mức thuế tăng lên không quá ngại, hoặc ít nhất không khủng khiếp, như hình gốc minh họa. Một lần nữa, vì mắt có xu hướng so sánh điểm cuối giữa các cột với nhau, biểu đồ cần phải thể hiện toàn bộ bối cảnh liên quan nhằm giúp đối tượng mục tiêu có thể đối chiếu chính xác.



tinyurl.com/magncaffuccino

Các bạn có thể để ý thấy cách cấu trúc của biểu đồ cũng có chút thay đổi sau khi được làm lại. Nhãn dữ liệu trên cột y vốn nằm bên phải cột trong hình gốc nay đã được chuyển sang bên trái (để thấy cách lý giải dữ liệu trước khi tiến hành phân tích). Nhãn dữ liệu của các cột, từ vị trí bên ngoài, đã được chuyển vào trong nhằm tránh gây rối mắt. Nếu được tự do minh họa theo ý mình, có lẽ tôi đã bỏ luôn cột y và chỉ để lại nhãn dữ liệu bên trong cột nhằm giảm thiểu những thông tin thừa. Tuy nhiên, trong trường hợp này, tôi vẫn giữ nguyên cột y để các bạn thấy rõ được thông số xuất phát điểm của nó bằng 0.

Trục biểu đồ và nhãn dữ liệu

Khi trực quan hóa dữ liệu qua biểu đồ, thông thường người ta sẽ phải quyết định xem giữ lại nhãn dữ liệu các cột hoặc loại bỏ hoàn toàn và đính trực tiếp thông số dữ liệu lên biểu đồ. Mỗi khi lâm vào tình thế này, các bạn hãy xem xét mức độ cụ thể cần thiết. Nếu bạn muốn đối tượng mục tiêu tập trung vào xu thế tổng thể, hãy cân nhắc giữ lại các cột, nhưng tô xám chúng nhằm giảm độ nổi bật. Nếu như một giá trị thông số cụ thể nào đó rất quan trọng, có lẽ cách tốt hơn hết là đính trực tiếp thông số dữ liệu lên biểu đồ. Ở trường hợp sau, loại bỏ hoàn toàn các trục là cách tối ưu nhất tránh thông tin thừa. Luôn luôn cân nhắc xem bạn muốn đối tượng mục tiêu tiếp cận hình ảnh trực quan như thế nào và xây dựng chúng sao cho phù hợp.

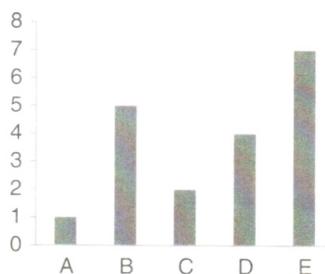
Quy tắc mà tôi muốn nói đến ở đây là *giá trị đường cơ sở của biểu đồ dạng cột buộc phải bằng 0*. Chúng ta không áp dụng quy tắc này cho biểu đồ đường. Vì đối với biểu đồ đường, điểm đáng lưu ý là vị trí tương quan giữa các đường trong khoảng giá trị (thay vì chiều dài từ đường cơ sở hoặc cột), cho nên đường cơ sở có giá trị khác 0 cũng chẳng sao cả. Dẫu sao chúng ta cũng đừng chủ quan – hãy nói rõ cho đối tượng mục tiêu biết rằng bạn đang sử dụng đường cơ sở có giá trị khác 0 cũng như bối cảnh cũng thay đổi tương ứng. Nhờ đó, bạn sẽ không quá tập trung cũng như cường điệu quá mức một thay đổi, hoặc một điểm khác biệt, không đáng kể.



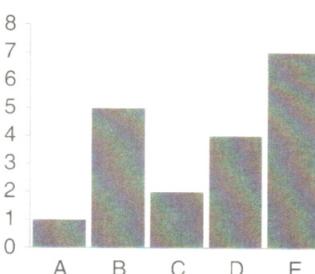
Nhưng sẽ ra sao nếu bạn thay đổi tỷ lệ trên biểu đồ cột hoặc tự ý chỉnh sửa các dữ liệu nhằm nhấn mạnh hơn ý mình? Lối suy nghĩ sai lệch này, cùng với hành động trực quan hóa dữ liệu lèo lạc, thật không ổn chút nào. Ngoài phạm trù quy chuẩn, đây cũng là một việc vô cùng rủi ro. Chỉ cần một đối tượng tinh ý phát hiện ra vấn đề (ví dụ như cột y của biểu đồ bắt đầu bằng một con số khác 0) là toàn bộ công sức, thậm chí cả uy tín của bạn, đều đổ sông đổ biển.

Tiện thể trong lúc xem xét độ dài của các cột, chúng ta cũng hãy bàn luận về *bề ngang* của chúng nữa. Chẳng tồn tại quy tắc nào về vấn đề này cả, nhưng nhìn chung, bề rộng các cột nên to hơn so với diện tích khoảng trắng giữa các cột với nhau. Tuy nhiên, các bạn cũng đừng vẽ cột quá to, kẻo đối tượng mục tiêu lại đối chiếu diện tích cột thay vì chiều cao. Chúng ta hãy cùng xem hình ảnh minh họa “nguyên lý Goldilocks¹⁵” của biểu đồ cột: quá mỏng, quá dày và vừa phải.

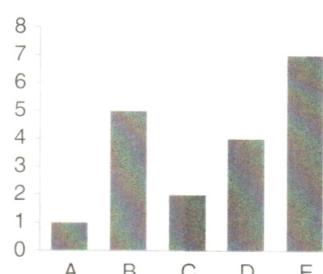
Quá mỏng



Quá dày



Vừa phải



Hình 2.14 Độ rộng của cột

Chúng ta đã đi qua một số ví dụ thực tiễn nhất về biểu đồ dạng cột nói chung. Giờ đây, ta sẽ bàn đến một số biến thể của nó. Có sẵn trong tay một số định dạng biểu đồ cột sẽ giúp bạn linh hoạt hơn khi đứng trước những thử thách trực quan hóa dữ liệu khác nhau. Chúng ta sẽ đề cập đến những định dạng mà tôi cho rằng chẳng lạ lẫm gì với các bạn đâu.

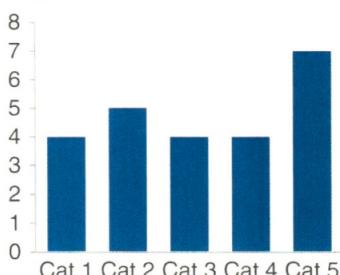
15. Nguyên lý Goldilocks: Cái gì “vừa đủ” là tốt nhất



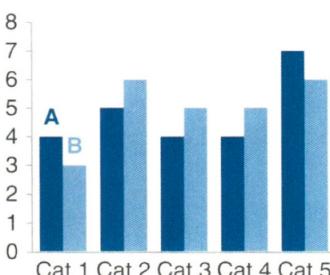
Biểu đồ cột đứng

Có thể nói biểu đồ cột đứng chính là định dạng cơ bản nhất, hay còn gọi là biểu đồ thanh. Giống như biểu đồ đường, biểu đồ cột đứng có thể bao gồm một, hai hoặc nhiều chuỗi dữ liệu. Nhưng cũng lưu ý rằng càng nhiều chuỗi dữ liệu, bạn sẽ càng khó tập trung và rút ra được thông tin từ tất cả cùng một lúc. Vậy nên hãy cẩn trọng khi thêm chuỗi dữ liệu vào biểu đồ cột. Bạn cũng hãy chú ý đến khái niệm visual grouping¹⁶ trong những biểu đồ cột có nhiều hơn một chuỗi dữ liệu vì đây là điều tất yếu mỗi khi giãn cách từng hạng mục. Điều này càng nhấn mạnh thêm tầm quan trọng của việc sắp xếp vị trí các hạng mục. Hãy cân nhắc xem bạn muốn đổi tượng mục tiêu đổi chiểu những gì, rồi sau đó sắp xếp phù hợp vị trí từng hạng mục nhằm đơn giản hóa tối đa điều đó.

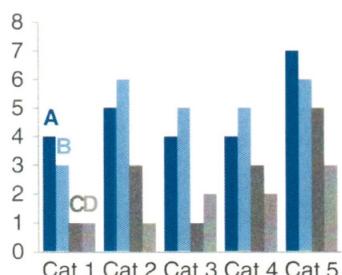
Một chuỗi



Hai chuỗi



Nhiều chuỗi



Hình 2.15 Biểu đồ cột

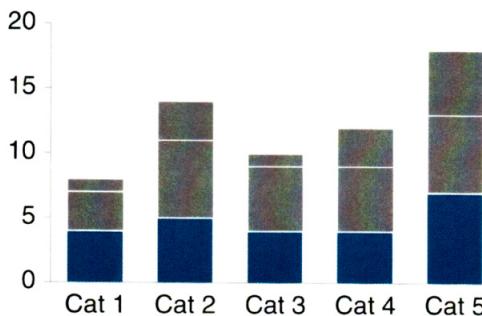
Biểu đồ cột chồng dạng đứng

Những áp dụng thực tiễn của dạng biểu đồ này hiếm hơn nhiều. Nó cho phép bạn đổi chiểu những thông số tổng thể giữa các hạng mục, đồng thời cũng cho thấy những thông số phụ trong cùng hạng mục đó. Tuy nhiên, điều này sẽ dễ khiến đổi tượng mục tiêu choáng ngợp trước những hình ảnh trực quan đó, đặc biệt với những định dạng nguyên bản nhiều màu trong đa số ứng dụng tạo đồ thị (chúng ta sẽ bàn sâu hơn sau). Bạn sẽ thấy rằng thật khó để có thể đổi chiểu những thông số phụ giữa các hạng mục ngoại trừ thông số phụ dưới cùng (thông số nằm tiếp giáp trực x) vì lúc này bạn dường như không còn đường cơ sở thăng lắp để so sánh nữa. Điều này khiến việc đổi chiểu bằng thi giác gặp nhiều trở ngại như Hình 2.16 sau.

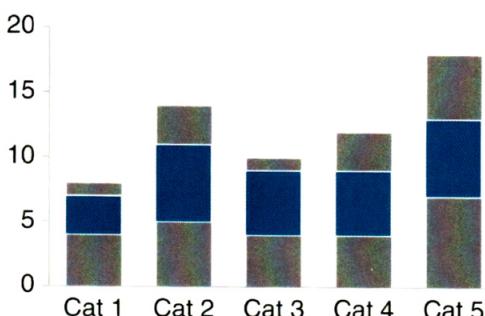
16. Visual grouping: Hình thức gộp chung nhóm những hình ảnh trực quan có cùng một hay nhiều điểm chung cụ thể



So sánh phần dữ liệu này thì dễ



So sánh phần dữ liệu này thì khó



Hình 2.16 So sánh dữ liệu bằng biểu đồ cột chồng

Chúng ta có thể xây dựng và thêm trực tiếp giá trị tuyệt đối vào biểu đồ cột chồng dạng đứng (ghi thông số hiện có như Hình 2.16) hoặc định tỷ lệ tất cả các cột ở mức 100% (bạn sẽ minh họa tỷ lệ % cho mỗi thông số phụ; chúng ta sẽ đi vào chi tiết ở Chương 9). Cách bạn chọn sẽ tùy thuộc những gì bạn muốn truyền đạt đến đối tượng mục tiêu. Khi sử dụng biểu đồ với tất cả các cột đạt tỷ lệ 100%, bạn hãy suy nghĩ xem biểu đồ vẫn truyền tải được ý nghĩa hay không nếu bạn thêm những giá trị tuyệt đối vào từng hạng mục cụ thể trên cột (dù thêm trực tiếp lên biểu đồ hay chỉ là ghi chú) nhằm hỗ trợ bạn lý giải dữ liệu.

Biểu đồ thác nước

Ta có thể sử dụng biểu đồ thác nước với mục đích phân tách từng hạng mục của biểu đồ cột để tập trung phân tích mỗi lúc một hạng mục riêng lẻ, hoặc nhằm biểu thị một điểm khởi đầu, sự tăng/giảm tỷ lệ và cuối cùng là một điểm kết thúc.

Cách tốt nhất để mô tả ví dụ thực tiễn biểu đồ thác nước là bằng ví dụ. Hãy tưởng tượng bạn làm trong mảng Đối tác Nhân sự¹⁷ và đang mong muốn thấu hiểu, cũng như truyền đạt thực trạng những thay đổi trong nguồn nhân lực năm vừa qua cho nhóm khách hàng mà bạn đang hỗ trợ.

Một biểu đồ thác nước biểu thị kết quả phân tích của bạn sẽ khá tương đồng với Hình 2.17 sau.

17. Đối tác Nhân sự: Một nghiệp vụ/bộ phận có trách nhiệm tư vấn và cung cấp dịch vụ cho các phòng ban nhằm đảm bảo mỗi phòng ban có được nguồn nhân lực tốt nhất

