

COMUNICACIÓN VISUAL DE LA INFORMACIÓN

Qué y cómo podemos narrar con datos

MARIO PÉREZ-MONTORO

Comunicación visual de la información

qué y cómo podemos
narrar con datos

© 2022 Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

Esta publicação está disponível em acesso livre ao abrigo da licença Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). Ao utilizar o conteúdo da presente publicação, os usuários aceitam os termos de uso do Repositório UNESCO de acesso livre (www.unesco.org/open-access/terms-use-ccbysa-port).

Esta publicação tem a cooperação da UNESCO no âmbito do projeto “Ampliação e Modernização das Ações do IBICT relacionadas às Atividades de Coleta, Armazenamento, Sistematização, Análise, Disseminação e Preservação de Dados e Informações Relativos à Ciência, Tecnologia e Inovação” (Prodoc 914BRZ2005). As indicações de nomes e a apresentação do material ao longo deste livro não implicam a manifestação de qualquer opinião por parte da UNESCO a respeito da condição jurídica de qualquer país, território, cidade, região ou de suas autoridades, tampouco da delimitação de suas fronteiras ou limites. As ideias e opiniões expressas nesta publicação são as dos autores e não refletem obrigatoriamente as da UNESCO nem comprometem a Organização.



CONSELHO EXECUTIVO

- › Gustavo Saldanha (Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT; Universidade Federal do Estado do Rio de Janeiro – Unirio)
- › Paulo César Castro (Escola de Comunicação – ECO/UFRJ)

CONSELHO CIENTÍFICO DA COLEÇÃO

- › Cecília Leite (Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT)
- › Miguel Ángel Rendón Rojas (Universidade Nacional Autónoma de México - UNAM)
- › Muniz Sodré (Universidade Federal do Rio de Janeiro - UFRJ)
- › Ivana Bentes (Universidade Federal do Rio de Janeiro - UFRJ)
- › Naira Christofoleti Silveira (Universidade Federal do Estado do Rio de Janeiro – Unirio)
- › Rafael Capurro (Unesco)

Comunicación visual de la información

qué y cómo podemos
narrar con datos

Mario Pérez-Montoro



Rio de Janeiro
2022

Capa: Fernanda Estevam

Ilustração: GK Vector (br.freepik.com)

Projeto Gráfico: Paulo César Castro

Normalização e catalogação: Selo Nyota

Diagramação: Nathália Romeiro e Fernanda Estevam

P438m

Comunicación visual de la información: qué y cómo podemos narrar con datos /
Mario Pérez-Montoro. – Rio de Janeiro: IBICT, 2022. – (Coleção PPGCI 50 anos)

174p.

Inclui Bibliografia.

Disponível em: <https://ridi.ibict.br/>

ISBN 978-65-89167-63-1 (digital)

1. Comunicação audiovisual. 2. Dados. 3. Mídias digitais. I. Pérez-Montoro, Mario.
II. Título.

CDD 020



Projeto editorial em colaboração com o Programa de Educação Tutorial (PET) da Escola de Comunicação (ECO-UFRJ): Paulo César Castro (tutor) / aluno(a)s: Carolina Torres, Dandara Campello, João Maurício Maturana, Juliana Sorrenti, Kethury Santos, Lianne Henriques, Mariana da Paz, Ludmila Rancan, Moniqui Frazão, Robertha Braga, Sabrina Oliveira e Sara Maluf.



Programa de Pós-Graduação em Ciência da Informação (PPGCI), desenvolvido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia, do Ministério da Ciência e Tecnologia e Inovação (IBICT/MCTI) em convênio com a Escola de Comunicação da Universidade Federal do Rio de Janeiro (ECO/UFRJ).

Rua Lauro Muller, 455 - 4º andar
Botafogo - Rio de Janeiro - RJ
<http://www.ppgci.ufrj.br>

*A mis padres, José y Dolores, in memoriam
Y a todos los que en algún momento de sus vidas los amaron y quisieron*

Sumario

9 **PREFACIO**

13 **CONCEPTOS BÁSICOS**

- 13** 1. Introducción
- 17** 2. Visualización de la información
- 20** 3. Infografía y visualización de datos
- 26** 4. Percepción y visualización

37 **CARACTERIZACIÓN Y ANATOMÍA DE LAS GRÁFICAS**

- 37** 1. Introducción
- 38** 2. Definición de gráfica
- 45** 3. Gráficas versus tablas
- 55** 4. Tipología de gráficas

103 **COMUNICACIÓN EFICIENTE CON GRÁFICAS**

- 103** 1. Introducción
- 104** 2. Contenidos semánticos
- 113** 3. Visualización de contenidos
- 144** 4. Diseño visual de gráficas

169 **BIBLIOGRAFÍA**

Prefacio

Mario Pérez-Montoro¹

SEGURAMENTE, EN MUCHOS MOMENTOS DE NUESTRA VIDA ESCOLAR O LABORAL, con cierta asiduidad nos hemos tenido que enfrentar al problema de realizar una búsqueda, obtención y refinado de un conjunto de datos que nos permita apoyar, argumentar y narrar una historia que comunique cierta información a unos destinatarios concretos.

Pero, independientemente de la complejidad de esa búsqueda y refinado, los datos tampoco hablan por sí solos. No nos ofrecen de forma automática la información que albergan. Necesitan ser visualizados para mostrar un análisis derivado y permitir que afloren los patrones que encierran.

En esta misma línea, para realizar una buena visualización es imprescindible, también, conocer los principios lógicos y perceptivos que garantizan su eficiencia comunicativa. Es necesario, en definitiva, poseer un buen grado de alfabetización visual.

Con este trabajo, vamos a intentar contribuir, de forma modesta, a esa alfabetización visual necesaria, introduciendo de forma práctica esos principios que nos permitirán diseñar y utilizar este útil instrumento de representación de datos en nuestros productos comunicacionales.

Concretamente, en términos generales, vamos a presentar y revisar los conceptos y principios básicos que sustentan la visualización de la información y conocer sus fundamentos perceptivos elementales. También vamos a ofrecer una base metodológica para seleccionar la mejor visualización con gráficas para cada tipo de contenido semántico. Y, por último, vamos a aprender a realizar un correcto diseño visual de gráficas de acuerdo con los principios cognitivos y éticos que las regulan. Y para cubrir esos objetivos generales, hemos dividido este trabajo en tres capítulos.

¹ Doctor en Filosofía y Ciencias de la Educación. Profesor Catedrático de Universidad, Facultad de Información y Medios Audiovisuales, Universidad de Barcelona (España), perez-montoro@ub.edu

El primer capítulo, titulado *Conceptos básicos*, tiene como principal propósito abordar algunos de los principales conceptos elementales que conforman y permiten entender de forma adecuada la disciplina de la visualización de información.

Para alcanzar esa meta, en el primero de los apartados del capítulo, se explicará cómo podemos entender esa disciplina y cuáles son algunas de las funciones básicas que posee cualquier producto de visualización concreto y que podemos aprovechar cuando lo creamos y utilizamos en un contexto determinado (el almacenamiento de información, su comunicación y el soporte al razonamiento). En un segundo movimiento, en el siguiente apartado, introduciremos la distinción técnica entre los dos tipos principales de representaciones visuales que conforman la visualización de la información: la infografía y la visualización de datos. Por último, en un apartado posterior, acabaremos abordando la articulación que la visualización de información mantiene con el tema de la percepción, centrandó nuestra atención sobre tres aspectos primordiales: la dependencia que presenta la percepción respecto a contexto en el se produce, la existencia de límites en la discriminación perceptiva, y el principio de la ratio dato/tinta.

En el segundo capítulo, titulado *Caracterización y anatomía de las gráficas*, vamos a intentar especificar y describir uno de los principales productos visuales que se utilizan para representar y transmitir la información: las gráficas. Si dejamos a un lado el tema de las tablas, la gráfica se presenta como uno de los productos de visualización más utilizados y más presente en los contextos comunicativos que incluyen algún tipo representación de datos. Por tanto, para su correcta utilización, debemos, primero, tener claro en qué consiste este tipo especial de productos.

Para cubrir el objetivo de caracterizar las gráficas, vamos a tratar de desarrollar los siguientes puntos. En el primero de los apartados del capítulo, se explicará cómo podemos definir una gráfica, introduciendo las características básicas que estructuran este tipo de productos e ilustrando a partir de un par de ejemplos cómo se articulan esas características básicas en una gráfica. De la mano de esa definición analizaremos también los componentes estructurales que conforman las gráficas. Concretamente, introduciremos los elementos codificadores de información y los elementos de soporte que podemos encontrar en una gráfica, y los conectaremos con la distinción entre las diferentes clases de tinta que podemos utilizar en un producto visual de este tipo y que recogíamos en el principio de la ratio dato/tinta.

En el siguiente apartado, abordaremos las principales diferencias básicas existentes entre los dos principales productos de visualización: las gráficas y las tablas. Veremos cuáles son los contextos comunicacionales más propicios para utilizar cada uno de esos productos y, utilizando una serie de ejemplos ilustrativos, mostraremos las ventajas y limitaciones de cada uno de esos usos. Por último, acabaremos

el capítulo abordando la tipología de las principales gráficas que solemos utilizar con intención comunicativa. Veremos su anatomía, la ilustraremos con ejemplos, y adelantaremos algunos de los usos que podemos darle a cada uno de esos tipos.

En el último de los capítulos, titulado *Comunicación eficiente con gráficas*, teniendo en cuenta todo lo desarrollado hasta ese punto, intentaremos caracterizar los diferentes y principales contextos y contenidos comunicacionales que podemos transmitir utilizando una gráfica y veremos cuál es la propuesta visual más adecuada para cubrir con eficiencia y eficacia esas intenciones comunicativas. O, dicho en otros términos, en ese capítulo queremos desplegar de forma ordenada la metodología habitual que se utiliza para implementar de forma adecuada una correcta visualización de información.

Para alcanzar ese propósito, vamos a tratar de desarrollar los siguientes apartados. En el primero de los apartados del capítulo, abordaremos los principales tipos básicos de información o contenidos semánticos que podemos representar en una gráfica, ilustrándolos a partir de ejemplos que codificaremos en forma de tabla para que puedan ser entendidos de una forma más intuitiva. En el siguiente apartado, y recuperando la tipología de gráficas que desarrollamos en el segundo capítulo, examinaremos cuál es tipo de visualización que se adecua de forma más adecuada para comunicar cada uno de esos tipos de contenidos informativos. Por último, revisaremos algunos aspectos relacionados con el diseño eficiente de los principales elementos codificadores de información y componentes de soporte que conforman una gráfica y que detallamos también el capítulo anterior.

Conceptos básicos

1. Introducción

VIVIMOS INMERSOS EN UNA CULTURA VISUAL. EL DESARROLLO DE LAS TECNOLOGÍAS ha consolidado el uso generalizado de las imágenes como estrategia comunicativa. La prensa, el cine o la televisión, y ahora también internet, utilizan las imágenes para respaldar sus mensajes y hacerlos más persuasivos.

Este uso de la imagen se ha visto trasladado también al ámbito de la información. Desde los primeros trabajos de William Playfair (1786 y 1801) a finales el siglo XVIII, se ha ido introduciendo la imagen, la visualización, como estrategia complementaria al texto para codificar, representar y comunicar información de una forma más económica y eficiente (Dick, 2020).

Actualmente, existe un auge del uso intensivo de la visualización en todos los contextos. En el profesional, por ejemplo, las empresas utilizan cada día más visualizaciones en sus comunicaciones internas y externas. E incluso, en nuestra esfera más personal, gracias a la aparición de nuevos sencillos instrumentos, utilizamos visualizaciones para reforzar nuestra comunicación digital en internet.

Las causas del auge de este uso son diversas. Por un lado, el crecimiento exponencial de los datos y necesidad de gestionarlos. Si somos conscientes de que aproximadamente el 90% de los datos existentes se ha generado en los últimos 5 años, la visualización se convierte en una de las estrategias más efectivas para gestionar ese volumen ingente y hacerlo más asimilable. Por otro, han ido apareciendo en el mercado un número importante de herramientas tecnológicas baratas o gratuitas que facilitan y simplifican la creación de visualizaciones.

También ha sido importante la adopción de la visualización por parte de los medios de comunicación. Esos medios han visto en la visualización de datos una nueva manera de comunicar sus contenidos, superando estrategias narrativas más clásicas y apostando por su uso como una ventaja competitiva frente a sus competidores. Por último, una de las razones más importantes ha sido además el carácter altamente persuasivo de la visualización. Los mensajes que se refuerzan o se fun-

damentan sobre visualizaciones ofrecen a la audiencia un aire o pátina de cientificidad, de objetividad, de respetabilidad argumentativa, que cuesta mucho más de apreciar cuando estos recursos gráficos están ausentes.

Sin embargo, esta situación no es tan armoniosa como aparenta. Convive con el importante problema de la existencia de un desfase o desajuste acusado entre el uso intensivo que realizamos de la visualización y el escaso nivel de conocimientos que poseemos sobre lo visual. O, dicho en otros términos, nuestro grado de alfabetización visual no se corresponde ni avanza con el mismo ritmo que nuestro uso de la visualización.

A diferencia de lo que puedan pensar algunos usuarios, la creación y el uso correcto de una buena visualización de información está sujeto a reglas o principios que garantizan la efectividad comunicativa de la misma.

Muchos de esos principios provienen de la psicología y, más concretamente, de la teoría de la percepción. La correcta combinación de colores en las visualizaciones, la idoneidad de utilizar en tablas y gráficas tipografías *sans-serif* en detrimento de las *serif* o la imposibilidad cognitiva que tenemos los humanos de poder comparar visualmente ángulos con precisión y por tanto la recomendación de no utilizar de forma generalizada ciclogramas (gráficos de pastel), son algunos ejemplos de principios básicos derivados de la teoría de la percepción que rigen la construcción de buenas visualizaciones.

Como veremos más adelante, otros principios provienen de la propia estructura de la información y de la capacidad de representación visual que tienen las tablas y las gráficas. En ese contexto, por ejemplo, es posible enunciar un sencillo principio: cada tipo de información reclama un tipo especial de visualización. O, dicho en otras palabras, no cualquier información se puede representar utilizando cualquier tipo de visualización.

Este principio se sustenta sobre la idea de combinar y articular dos aspectos complementarios: la variación de estructuras en la información y las diferentes potencialidades representativas de cada tipo de tabla o gráfica. Por un lado, en un sentido técnico, no todo ítem de información posee una misma estructura proposicional. Así, por ejemplo, algunos ítems expresan la comparación entre dos variables, otros la evolución de una variable en el tiempo o, sin ir más lejos, la relación que mantiene un todo con cada una de sus partes. Y, por otro, no toda visualización posee la misma capacidad de representación visual. En este sentido, por ejemplo, una gráfica clásica de líneas representa con mucha eficiencia la evolución y la tendencia de una variable en el tiempo, pero le cuesta dar cuenta visualmente de la relación que mantiene un todo con cada una de sus partes. En cambio, un *TreeMap* representa de forma eficiente esa última relación, pero no sirve para mostrarnos cómo evoluciona en el tiempo una variable.

Sin embargo, al margen de estos que acabamos de enumerar, existen también una serie de principios conceptuales básicos derivados de la propia naturaleza de la información y la de sus representaciones visuales. Y entre estos, destaca uno que hay que tener siempre en cuenta cuando decidamos utilizar una visualización dentro de nuestras estrategias comunicativas: el principio de la objetividad.

Este principio nace de la disciplina de la filosofía. Concretamente de la filosofía analítica. En 1990, los filósofos estadounidenses David Israel y John Perry (profesores de la *Stanford University* (EEUU)) publican un trabajo donde recogen las principales ideas y argumentos que habían desarrollado con el lógico Jon Barwise en los años anteriores. Entre esas ideas formulan un principio en el que se expresa la naturaleza objetiva del flujo informativo.

- a) *Principio de la objetividad de la información*: el contenido informativo asociado a una señal es siempre una proposición verdadera (si una señal s transporta la información de que P entonces es verdad que P) (Israel y Perry, 1990).

El análisis de este principio nos permite extraer una serie de lecciones que debemos poner en práctica a la hora de crear una visualización de información como complemento de nuestras estrategias comunicativas.

En este contexto, respecto a este principio, la lección es rotunda: informar siempre significa contar la verdad. O, dicho de otra manera, no debemos mentir de forma intencionada utilizando visualizaciones. Una visualización que represente de manera parcial –o que, de forma interesada, no represente adecuadamente la información– no es una visualización correcta. En este sentido, debemos utilizar las visualizaciones de manera apropiada, sin diseñarlas de forma que disimulen los datos que no queremos presentar o acentuando de forma distorsionada los datos que queremos resaltar.

Ejemplos de estas manipulaciones los hemos podido encontrar en los debates televisivos correspondientes a las últimas campañas políticas para elegir al presidente del estado. En esos debates, los candidatos han blandido gráficas en las que se presentaba parcialmente la información que no apoyaba sus argumentos o se exageraba la que sí lo hacía. Steve Jobs, el que fue *CEO* de la compañía Apple, utilizaba en sus famosas presentaciones de productos visualizaciones en 3D para ofrecer, jugando con la orientación de las mismas, la sensación errónea de que los números de su sistema operativo móvil iOS eran mejores que los de la competencia Android de Google. O incluso, en muchos informes y también en artículos científicos podemos encontrar gráficas de columnas que, para apoyar los resultados de

sus proyectos, comienzan interesadamente el eje de ordenadas (eje Y) en un valor positivo distinto a cero para ofrecer de forma distorsionada, por exagerada, la diferencia entre ciertos valores concretos.

La conclusión que podemos extraer es que la visualización es un lenguaje universal para transmitir información a través de figuras visuales y que se rige por unos principios que garantizan su eficacia y eficiencia comunicativa (Horm, 1998; Cairo, 2019). Y que, lamentablemente, su uso intensivo por parte de la población no se ha visto acompañado del conocimiento de esos principios.

Para la consolidación del uso correcto de ese lenguaje narrativo es necesaria una alfabetización visual (Hattwig et al., 2011; Avgerinou y Pettersson, 2011; y Harris, 2006). Es imprescindible que los usuarios sean capaces de comunicarse correctamente con visualizaciones, no dejándose engañar por las mal diseñadas y construyendo buenas propuestas visuales basadas en esos principios (Drucker, 2014 y 2020; Balchin, 1976 y 1985; Boardman, 1990). Ese tipo de alfabetización visual (*graphicacy*, en inglés) se concretaría en la habilidad de entender y presentar correctamente información a partir de recursos visuales como diagramas, mapas, gráficas y otros formatos no textuales (Aldrich y Sheppard, 2000; Roth et al., 2005). Y, junto otros tipos de alfabetización como la lectora y la informacional (entre otras), formaría parte del arsenal y utillaje cognitivo mínimo para afrontar con garantías el desarrollo y el desenvolvimiento personal y social de todo ciudadano en la época en la que nos ha tocado vivir (Snavey y Cooper, 1997; Mutch 1997; Carbo 1997; Behrens, 1994; Doyle, 1994; Dess, 1991; Olsen y Coons, 1989; Kulthau, 1987; McClure, 1994; Bruce, 1997; McGarry, 1993 y 1994; o Bawden, 2001; entre otros muchos).

Con este primer capítulo, vamos a comenzar a contribuir, de forma modesta, a esta alfabetización visual necesaria. El capítulo que ahora iniciamos tiene como principal objetivo intentar abordar algunos de los principales conceptos básicos que conforman y permiten entender de forma adecuada la disciplina de la visualización de información. Y, para cubrir este objetivo, vamos a tratar de desarrollar los siguientes puntos.

En el primero de los apartados (apartado 2), se explicará cómo podemos entender esa disciplina y cuáles son algunas de las funciones básicas que posee cualquier producto de visualización concreto y que podemos aprovechar cuando lo creamos y utilizamos en un contexto determinado (el almacenamiento de información, su comunicación y el soporte al razonamiento). En un segundo movimiento, en el apartado 3, introduciremos la distinción técnica entre los dos tipos principales de representaciones visuales que conforman la visualización de la información: la infografía y la visualización de datos. Por último, en el apartado 4, acabaremos abordando la articulación que la visualización de información mantiene con el tema

de la percepción, centrando nuestra atención sobre tres aspectos primordiales: la dependencia que presenta la percepción respecto al contexto en el que se produce, la existencia de límites en la discriminación perceptiva, y el principio de la ratio dato/tinta.

2. Visualización de la información

Si abordamos de forma sistemática la literatura sobre la disciplina de la visualización de la información, una de las cosas que más rápidamente nos puede llamar la atención es que no existe una única, consensuada y compartida caracterización de la misma. Cada autor propone su propia definición, que se suma a un ecosistema teórico-conceptual densamente poblado de otras propuesta competidoras.

En ese escenario, nos encontramos frente a un término ambiguo que circula en algunos contextos profesionales y especializados y que seguramente aglutina conceptos pertenecientes a diversas disciplinas como la tecnología, la psicología, la comunicación, la información e, incluso, el ámbito artístico.

De todas formas, aunque no exista esa definición unánimemente compartida, es posible identificar una manera entender la disciplina de la visualización de la información y, también, cierto consenso en los mínimos conceptuales que comparten esas definiciones. Son unos mínimos que aglutinan una serie de conceptos (como los de representación, contenidos, diagramas, gráficas o comunicación, entre otros) y que se repiten en casi todas las definiciones.

Es justo a partir de la combinación y articulación de esos conceptos comunes de dónde vamos a extraer la definición de visualización de la información que vamos a utilizar a lo largo de todo este trabajo.

- b) *Definición de Visualización de la Información*: disciplina transversal que se encarga de la representación visual de contenidos proposicionales mediante el uso de diagramas, gráficas y esquemas para facilitar el almacenamiento, la aprehensión, la interpretación, la transformación y la comunicación de esos contenidos a través de esas representaciones visuales (Pérez-Montoro y Golkhosravi, 2014).

Teniendo en cuenta esta definición, la visualización de información persigue cubrir tres funciones básicas: el almacenamiento de información, su comunicación y el soporte al razonamiento.

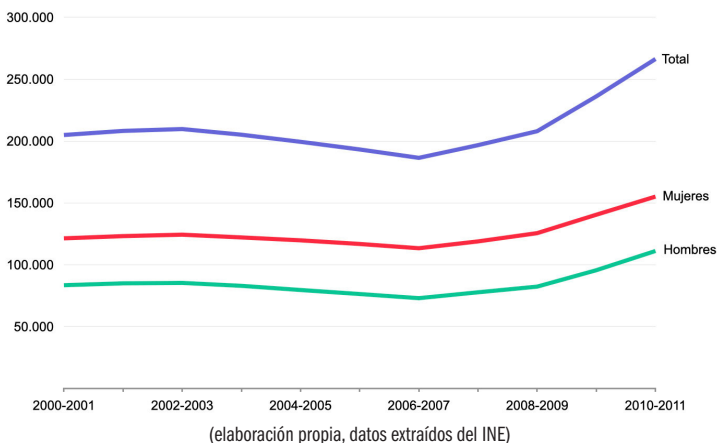
Respecto al almacenamiento, las representaciones visuales consiguen registrar y nos permiten recuperar más tarde la información que codifican. Así, por ejemplo, podemos generar una gráfica con los datos de la evolución de los graduados (antes

llamados “licenciados”) durante la primera década del siglo XXI en España (figura 1) y estaremos almacenando esos datos que la generan independientemente de la tabla que originariamente los albergaba. Pasado el tiempo, podemos recuperar esa gráfica y recuperar esos datos.

Pero la visualización no sólo permite el almacenamiento, sino también la comunicación de información. Con una gráfica como la de la figura 1 podemos comunicar a otros individuos muchas cosas. Entre otras, podemos compartir con otros y persuadirlos de que en esa década hubo más graduadas que graduados. O que en la época de bonanza económica disminuyó el número de graduados ya que el mercado laboral ofrecía un abanico de salidas profesionales sin pasar por la universidad a los jóvenes en edad de trabajar. Y, también, que, con la llegada de la crisis económica en 2006 y la retracción de ese mercado laboral, los alumnos optaron por regresar de nuevo a las aulas para retomar sus estudios universitarios.

Pero una de las funciones más interesantes que cubre la visualización es la de ofrecer un soporte al razonamiento. Cuando visualizamos información, activamos el proceso de materializar o cosificar (convertir en cosas materiales) entidades conceptuales (o proposicionales, si empleamos un lenguaje técnico) y abstractas que circulan por nuestras mentes (números, relaciones o conceptos, por ejemplo) y convertirlas en algo físico (manchas de colores y formas diferentes en un papel o en una pantalla) para facilitar de esta manera el análisis y la comprensión de esas entidades abstractas. Nos permite representar el universo de los conceptos en un plano físico y tangible para facilitarnos así el pensamiento y todas las operaciones cognitivas que utilicen como materia prima esos objetos de naturaleza proposicional y etérea.

Figura 1. Graduados universitarios en la primera década del siglo XXI en España



Pensemos, sólo a modo de ejemplo, lo difícil que sería multiplicar mentalmente un número de 6 cifras por otro de 8 cifras. Para los que no estamos entrenados, nos costaría unos minutos hacerlo y correríamos el riesgo de, seguramente, equivocarnos. En cambio, si cogiéramos un papel y lápiz e hiciésemos la multiplicación utilizando esos recursos (si visualizáramos el proceso, en definitiva), el tiempo se reduciría drásticamente y las posibilidades de equivocarnos en el cálculo disminuirían sensiblemente.

Ese es el (súper) poder de la visualización. En cierta manera, es una disciplina que se encarga de hacer visible lo invisible: se encarga de hacer aflorar, mediante representaciones visuales, aspectos y relaciones que no son evidentes o visibles en un conjunto de datos.

Hace un tiempo, con unos colegas de otra universidad, participamos en una investigación en la que se intentaba identificar en qué se fijan los usuarios de Google para elegir y visitar un resultado concreto cuando realizan una búsqueda en esa página: ¿en el título de ese contenido digital?, ¿en el resumen? o ¿en su URL? Para tratar de averiguarlo, utilizamos un *eye tracker*, una herramienta que registra los movimientos oculares de un usuario cuando visiona una pantalla e identifica en qué zonas de esa pantalla ha detenido su mirada y por cuánto tiempo lo ha hecho.

El fichero con los datos en bruto resultante del experimento pesaba varias *gigas* y tenía un aspecto como se recoge en la figura 2. Observando ese fichero, ¿alguien se atrevería a responder a las preguntas anteriores? Nadie podría barruntar en qué se fijan los usuarios cuando analizan la página de resultados tras una búsqueda con Google y deciden visitar uno de los resultados ofrecidos.

Figura 2. Aspecto del fichero con los datos en bruto resultantes del experimento de las búsquedas de Google

```

127489 URLEnd 1024 0 0 file:///C:/Documents%20and%20Settings/usuario/
Datos%20de%20programa/Mozilla/Firefox/Profiles/2mf7tb5z.default/ScrapBook/data/
20090428133334/watch_001.html
128128 ScreenRecStopped 16384 0 0
128902 InstructionStart 2048 0 0 Instruction Element
136782 LeftMouseClicked 1 32 64
136848 InstructionEnd 4096 0 0 Instruction Element
137127 InstructionStart 2048 0 0 Instruction Element
151254 LeftMouseClicked 1 32 64
151255 InstructionEnd 4096 0 0 Instruction Element
151681 ScreenRecStarted 8192 0 0
153662 URLStart 512 0 0 file:///C:/Documents%20and%20Settings/
usuario/Datos%20de%20programa/Mozilla/Firefox/Profiles/2mf7tb5z.default/ScrapBook/data/
20090428133417/index.html
169815 LeftMouseClicked 1 342 496
170078 URLEnd 1024 0 0 file:///C:/Documents%20and%20Settings/usuario/
Datos%20de%20programa/Mozilla/Firefox/Profiles/2mf7tb5z.default/ScrapBook/data/
20090428133417/index.html
170963 URLStart 512 0 0 file:///C:/Documents%20and%20Settings/
usuario/Datos%20de%20programa/Mozilla/Firefox/Profiles/2mf7tb5z.default/ScrapBook/data/
20090428133417/a001009.html
191674 URLEnd 1024 0 0 file:///C:/Documents%20and%20Settings/usuario/
Datos%20de%20programa/Mozilla/Firefox/Profiles/2mf7tb5z.default/ScrapBook/data/
20090428133417/a001009.html
192165 ScreenRecStopped 16384 0 0
192801 InstructionStart 2048 0 0 Instruction Element
200232 LeftMouseClicked 1 343 335
200762 InstructionEnd 4096 0 0 Instruction Element
201628 InstructionStart 2048 0 0 Instruction Element
207648 LeftMouseClicked 1 347 298
207656 InstructionEnd 4096 0 0 Instruction Element
208001 ScreenRecStarted 8192 0 0
209707 URLStart 512 0 0 file:///C:/Documents%20and%20Settings/
usuario/Datos%20de%20programa/Mozilla/Firefox/Profiles/2mf7tb5z.default/ScrapBook/data/

```

Sin embargo, si visualizamos esos datos utilizando un mapa de calor (donde la escala del amarillo al rojo representa gradualmente las zonas más visionadas y durante más tiempo), podemos identificar muchas cosas: que los usuarios se fijan más en los enlaces patrocinados y menos en los orgánicos, que los contenidos con mapas reclaman más su atención o que los resúmenes son visionados antes de clicar en un resultado concreto, entre otras muchas cuestiones (figura 3).

3. Infografía y visualización de datos

Una vez presentada la definición y repasadas sus principales funciones, es importante introducir la distinción técnica entre los dos tipos principales de representaciones visuales que conforman la visualización de la información: la infografía y la visualización de datos.

Comencemos señalando que, si revisamos la literatura científica o los trabajos de los profesionales dedicados al diseño de la información, comprobamos que, en muchas ocasiones, en sus propuestas los términos “infografía” y “visualización de datos” se utilizan como sinónimos para nombrar a cualquier producto de visualización utilizado en un contexto comunicativo.

Existen verdaderos ríos de tinta sobre si deben o no ser considerados como términos sinónimos. Muchos autores defienden que las dos son representaciones o visualizaciones de información, y que lo que las diferencia es que las visualizaciones son interactivas o dinámicas, mientras que las infografías no lo son (Cairo, 2012

Figura 3. Visualización de los datos resultantes del experimento de las búsquedas de Google mediante un mapa de calor



(elaboración propia, datos obtenidos del experimento)

y 2013). Pero esa distinción no se adecua totalmente a los productos que podemos encontrar, por ejemplo, navegando por internet: existen infografías y visualizaciones tanto estáticas como dinámicas.

Es necesario distinguirlos con otro tipo de criterios. En este sentido, una infografía puede ser considerada como una representación gráfica (estática o dinámica) que combina imagen y texto con objetivos narrativos. O dicho en otros términos: es una representación discursiva (que puede incluir o no alguna visualización de datos). Se trata de una representación figurativa, donde el objeto que representa posee una apariencia similar, una isomorfía (en sentido técnico), a la del objeto representado.

A lo largo de la historia del arte podemos encontrar muchos productos visuales que en cierta manera podrían ser considerados como infografías. Sin ir más lejos, en los libros de notas de Leonardo Da Vinci sobre el embarazo, podemos encontrar un ejemplo clásico de infografía (figura 4). Allí, el autor nos representa la evolución del feto en un embarazo humano. Utiliza representaciones (manchas en el papel) que mantienen una relación figurativa, de similitud, con lo que están representando para narrarnos la evolución del proceso del desarrollo del feto.

Pero, más allá de la historia del arte, cotidianamente también podemos encontrar infinidad de ejemplos de infografía en los periódicos o en sus versiones digitales. Se trata de recursos visuales con los que reforzar comunicacionalmente los textos en esos productos periodísticos o con lo que dar cuenta de aspectos que se escapan al código de representación textual. La narración sobre la ascensión al Eve-

Figura 4. Libros de notas de Leonardo Da Vinci (1519)



(upload.wikimedia.org/wikipedia/commons/thumb/c/c9/Da_Vinci_Studies_of_Embryos_Luc_Viatour.jpg/200px-Da_Vinci_Studies_of_Embryos_Luc_Viatour.jpg)

rest que presentamos en la figura 5 podría ser considerado como un típico ejemplo de infografía en este tipo de medios. En esa infografía (que incluye también algunas visualizaciones de datos (la gráfica de columna con las cumbres y las muertes que aparece en la parte inferior, por ejemplo)) nos explica cómo se estructura y se ejecuta el ascenso a la montaña en cuestión. Para ello utiliza representaciones figurativas (con todo tipo de detalles) de los alpinistas, los porteadores y de la propia montaña. E, incluso, utiliza ese tipo de figuras para mostrarnos los problemas respiratorios que puede provocar el ascenso entre los escaladores.

Figura 5. Infografía sobre la subida al Everest

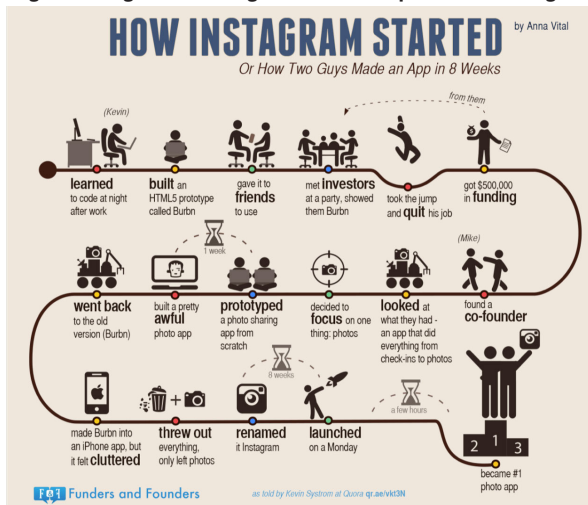


Por otro lado, en los últimos años, con el aumento del uso de las redes sociales, este tipo de productos visuales están teniendo mucha difusión. En el contexto de este tipo de comunicación social, compartir una infografía se presenta como una de las formas más eficaces y eficientes de hacer llegar a otras personas conceptos e ideas difíciles de transmitir mediante otros sistemas de codificación. En la figura 6 podemos encontrar un ejemplo de la típica infografía que circula por las redes sociales. En ésta, se describe, utilizando representaciones figurativas, todo el proceso de análisis, diseño e implantación de la conocida aplicación Instagram (figura 6). Nos narra como se fue gestando hasta colocarse en el número uno de las aplicaciones para compartir socialmente fotografías.

Es importante remarcar también que estas dos últimas infografías (la del ascenso al Everest y la de la generación de Instagram) son dos claros ejemplos de lo que se denomina también “*infopóster*”. Este tipo de infografías son productos visuales unitarios narrativamente, son unidades discursivas en si mismos, que no necesitan de otros elementos auxiliares (como un texto) para que puedan ser entendidos por sus potenciales receptores.

En cambio, una visualización de datos puede ser entendida como una representación gráfica (estática o dinámica) de un conjunto de datos para presentar aspectos y relaciones entre esos datos. O dicho de otra manera, frente a la infografía como representación discursiva, la visualización es una representación analítica. Se trata de una representación abstracta, donde el objeto que representa

Figura 6. Infografía sobre la generación de la aplicación de Instagram



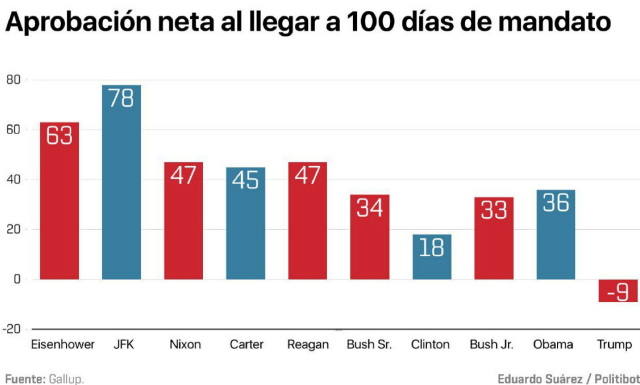
(<https://blog.adioma.com/how-instagram-started-infographic>)

no posee una apariencia similar, no es isomorfo (en sentido técnico), a la del objeto representado.

La gráfica donde se representa la aprobación de los últimos presidentes estadounidenses tras los primeros cien días de mandato (figura 7), puede ser considerada como un ejemplo típico de visualización de datos. En esa gráfica de columnas se codifica visualmente (mediante la longitud de la columna) cuál fue la cifra de aceptación de cada mandatario y para hacerlo no utiliza una representación figurativa, sino que cada columna, una figura abstracta, está codificando simultáneamente a cada uno de esos mandatarios y la variable cuantitativa asociada. Esa representación analítica nos permite identificar de forma cómoda que el presidente más aceptado fue John Fitzgerald Kennedy, y el que menos, Donal Trump; o que Bill Clinton tuvo la mitad de aceptación que Barack Obama en el mismo periodo de tiempo.

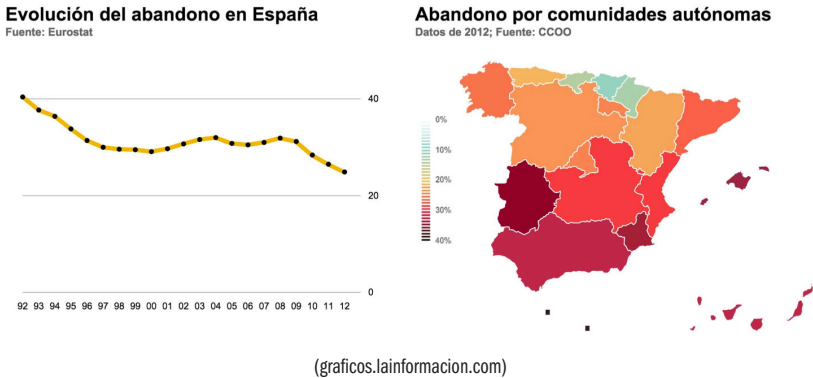
Otro ejemplo típico de visualización de datos lo encontramos en el mapa y la gráfica del abandono escolar en España en 2012 incluidos en la figura 8. En esta propuesta se nos representa mediante una gráfica de líneas la evolución en el tiempo (de 1992 a 2012) de la tasa de abandono escolar en el país peninsular. La línea es una figura abstracta que no mantiene ningún tipo de relación o similitud formal con lo representado (los que abandonan la escuela). Y nos facilita el análisis de esas cifras, permitiendo observar, entre otras muchas cosas, los repuntes de abandono en los años 2004 y 2008. Pero también nos representa, mediante la geolocalización de los datos a través de un mapa, la tasa de abandono por comunidades. Cada mancha en el mapa representa de forma abstracta esa tasa de abandono escolar en cada comunidad.

Figura 7. Aprobación de los últimos presidentes estadounidenses tras sus primeros cien días de mandato



(<https://www.politibot.io>)

Figura 8. Mapa y gráfica del abandono escolar en España en 2012



Por tanto, a modo de conclusión, a la luz de las definiciones ofrecidas y de todos los ejemplos de productos de visualización de información presentados, podemos decir que el factor discriminatorio entre una infografía y una visualización de datos no es la interactividad (su capacidad dinámica), sino el carácter discursivo y narrativo o analítico (su intencionalidad comunicativa) junto al tipo de representación (figurativa o abstracta) utilizada.

En este sentido, existen infografías no interactivas, como las que hemos introducido en las figuras 5 y 6. Pero también existen infografías interactivas, como la que nos ofrece el New York Times a modo de reportaje (los que se conoce técnicamente como “*webdoc*”, cercano al formato documental) sobre el ascenso a la montaña El Capitán (figura 9). En este producto visual se incluye un mapa interactivo que, mediante el uso de la navegación arriba-abajo (“*parallax effect*”, en un sentido técnico), nos va ofreciendo textos y material audiovisual de esa escalada a través del movimiento de nuestro ratón al hacer *scroll*. El *parallax* provoca que, al hacer ese *scroll*, el fondo parece moverse a una velocidad distinta a la del contenido ofreciendo así un efecto de linealidad narrativa, pero sobre todo de profundidad (con cierto grado de tridimensionalidad), que lo acerca a otros medios como el videojuego o el cine.

De la misma forma, también existen visualizaciones no interactivas, como las que se recoge en las figuras 7 y 8; e interactivas, como la de la figura 10, donde podemos interactuar y seleccionar nuestro perfil para calcular la tasa de paro asociada a ese perfil y luego comparar esa tasa con la del resto de España.

En este trabajo no vamos a abordar el tema de la infografía. Necesitaríamos un volumen de páginas similar para abordar también con cierta profundidad este tipo de interesantes productos visuales (Kosara y Mackinlay, 2013; Segel y Heer, 2010).

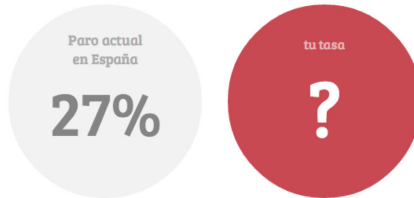
Figura 9. Ascenso a la montaña El Capitán, publicado por el New York Times



(www.nytimes.com/interactive/2015/01/09/sports/the-dawn-wall-el-capitan.html)

Figura 10. Calculador de la tasa de paro

Calcula tu tasa de paro



(tutasadeparo.es)

Nos vamos a centrar de forma exclusiva en la visualización de datos. Y, por tanto, a partir de ahora y en este contexto, vamos a convenir en utilizar como sinónimas las expresiones “visualización de datos” y “visualización de información”.

4. Percepción y visualización

Una vez introducida la definición y clarificados los términos principales que circulan erróneamente como sinónimos, para completar los conceptos básicos que sustentan la visualización, vamos a abordar ahora la articulación que ésta mantiene con el tema de la percepción.

Comencemos a partir de la siguiente idea justificadora: ¿por qué es necesario tener conocimiento sobre el tema de la percepción para poder visualizar información de una forma adecuada?

La respuesta es amplia y aborda diferentes razones de calado. La primera razón se fundamenta sobre una idea muy intuitiva: entendimiento y visión se encuentran estrechamente conectados (Cleveland, 1993). La visión domina nuestro espacio sensorial. Una parte importante de los receptores sensoriales de nuestro cuerpo están dedicados a la visión. Y es esa visión una de las fuentes de estímulos y materia prima de nuestro raciocinio (Ware, 2021).

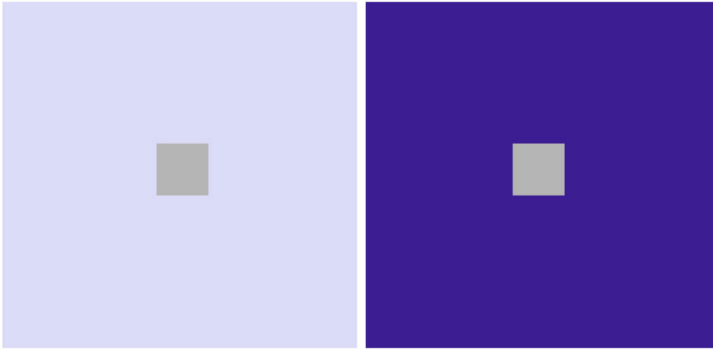
Dentro de este esquema intuitivo, para poder visualizar información de forma efectiva y poder también entender y extraer esa información de una visualización concreta de manera adecuada, es necesario tener algunos conocimientos sobre percepción visual. Es imprescindible saber qué es lo que sí funciona visualmente, qué es lo que no y por qué eso funciona o no funciona correctamente (Golombisky y Hagen, 2010).

Teniendo en cuenta todo esto, vamos a centrar nuestra atención sobre tres aspectos primordiales relacionados con la percepción: la dependencia que presenta la percepción respecto a contexto en el que se produce, la existencia de límites en la discriminación perceptiva y el principio de la ratio dato/tinta.

El primero es bien simple: la percepción depende del contexto (Harris, 1996; Healey, Christopher y Enns, 2012; Hockley y Bancroft, 2011). O, dicho en otros términos: la percepción no es neutra. Cuando percibimos, nada es grande o pequeño en términos absolutos, ni ancho estrecho, ni claro u oscuro, ni está totalmente arriba o abajo. Sino que esos (y otros) atributos perceptivos (longitud, intensidad o posición, por ejemplo) dependen del contexto visual en el que se encuentran integrados. En cierta manera, por así decirlo, todo depende de con qué lo comparamos.

Sin ir más lejos, una línea no es corta ni larga en sí misma; depende con qué otra línea se compare. En una gráfica, un círculo no es rojo oscuro o verde claro en sí mismo; depende del color que presente el círculo con el cual se está comparando. O un punto azul no está muy arriba o muy abajo en un plano determinado, sino que depende de los ejes que delimitan y estructuran la geografía de ese espacio bidimensional.

Esa dependencia de los atributos perceptivos respecto al contexto en el que están integrados la podemos ilustrar a partir de algunos ejemplos concretos. De esta manera, si nos fijamos en la figura 11 podemos observar que aparecen dos cuadrados pequeños insertados en otros dos cuadrados de mayor tamaño. La pregunta que nos asalta es ¿cuál de los dos cuadrados pequeños es más oscuro? Una primera respuesta sería que el cuadrado pequeño de la izquierda parece más oscuro que

Figura 11. Ejemplo de la dependencia del atributo perceptivo del color respecto al contexto

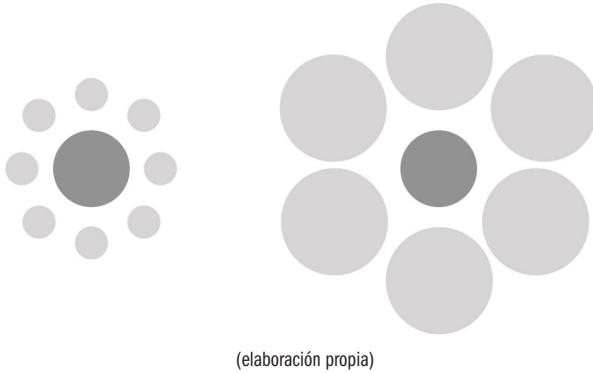
(elaboración propia)

el de la derecha. Pero eso es sólo una ilusión: los dos cuadrados pequeños son exactamente del mismo color. Que el izquierdo nos parezca más oscuro se justifica porque se encuentra dentro de un cuadrado mayor con un color más claro. El de la derecha parece más claro por lo contrario, por contraste, al encontrarse integrado en un cuadrado mayor mucho más oscuro.

Algo parecido ocurre en la figura 12, pero respecto al tamaño. Si observamos esta figura, podemos observar que aparecen dos círculos centrales (uno a la izquierda y otro a la derecha) rodeados por otros círculos a modo de satélites. Ahora, la pregunta es ¿cuál de los dos círculos es más grande? Una primera respuesta sería que el círculo de la izquierda parece más grande que el de la derecha. Pero eso es sólo una ilusión: los dos círculos son exactamente del mismo tamaño. Que el izquierdo nos parezca más grande se justifica porque se encuentra rodeado de círculos más pequeños. Y el de la derecha parece más pequeño por lo contrario, al encontrarse rodeado de círculos más grandes.

Esta dependencia de los atributos perceptivos respecto al contexto se traslada también a nuestras acciones más cotidianas. Por ejemplo, cuando preparamos una presentación y utilizamos diapositivas, debemos ser conscientes de que la visibilidad y legibilidad de lo que escribimos dependerá de la combinación del color del texto utilizado con el color del fondo. Así, por ejemplo, el texto negro sobre un fondo blanco (la combinación habitual en una presentación) o texto de color blanco sobre un fondo negro nos ofrece una buena visibilidad y legibilidad. En cambio, un texto en amarillo sobre un fondo blanco o un texto en azul sobre fondo negro, no son tan legibles para nuestra audiencia.

El segundo aspecto relacionado con la percepción que queremos abordar es el de la existencia de límites en la discriminación perceptiva que tenemos los huma-

Figura 12. Ejemplo de la dependencia del atributo perceptivo del tamaño respecto al contexto

nos (Tufte, 2006). Esa limitación se presenta, entre otros muchos contextos, en el de la percepción del color.

Aparentemente, los seres humanos acostumbramos a percibir de forma adecuada toda la escala de colores e intensidades. Pero eso no es del todo así. No todas las personas discriminamos de forma precisa entre los colores. Esa disfunción perceptiva se denomina técnicamente “protanomalia”.

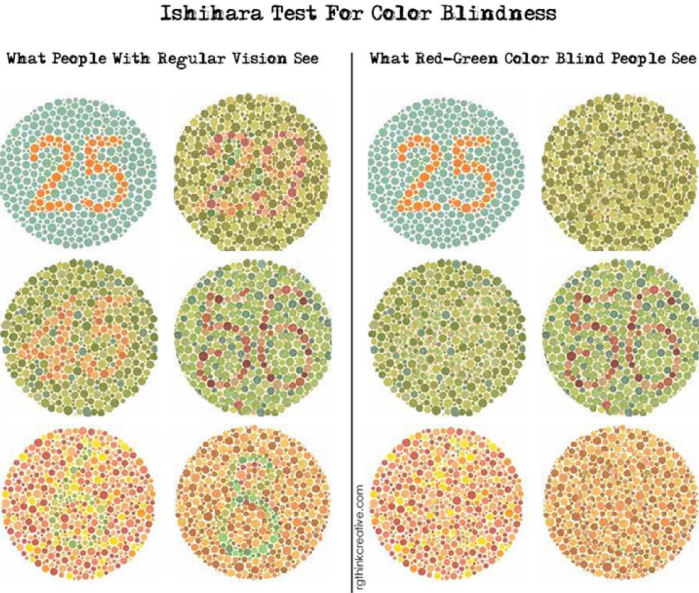
Esta disfunción visual normalmente se concreta en cierto grado de ceguera respecto al rojo y al verde y a sus combinaciones y está derivada de cierta anomalía presente en los conos oculares. Es mucho más común de lo que parece, sobre todo entre los hombres.

Para entender esta disfunción podemos observar la figura 13. Como vemos existen dos conjuntos formados por seis círculos cada uno separado por una barra vertical. Centrémonos en el conjunto de círculos de la izquierda. ¿Son capaces de leer los números que albergan esos seis círculos? La mayoría de ustedes lo harán sin ningún tipo de esfuerzo y responderán rápidamente: 25, 29, 45, 56, 6 y 8. Pero no todos serán capaces de leerlos.

Centrémonos ahora en el conjunto de círculos de la derecha. ¿Ven ustedes los mismos números? Seguramente no. Sólo pueden leer cómodamente el 25 del primer círculo y con alguna dificultad el 56 del cuarto círculo. No pueden leer los números que albergan los otros cuatro círculos restantes. Pues las personas que sufren esa disfunción visual ven exactamente igual los seis círculos de la izquierda como los de la derecha. O dicho de otra manera, debido a sus problemas perceptivos, en los dos conjuntos sólo son capaces de leer el número 25 y, con dificultades, el 56.

Por tanto, esta disfunción visual debe ser tenida en cuenta a la hora de diseñar visualizaciones. Así, por ejemplo, es mucho más aconsejable hacer una gráfica de

Figura 13. Test para detectar problemas en la discriminación visual entre colores



(www.fastcodesign.com/3027162/a-teenager-redesigns-the-web-for-the-color-blind?utm_source=Facebook)

líneas con variaciones de intensidad de un mismo color que combinar líneas de color rojo y verde que se solapan visualmente en algunas zonas de la gráfica.

El último aspecto relacionado con la percepción que queremos tratar es el relacionado con lo que se conoce como el principio de la ratio dato/tinta (*data-ink ratio*, en inglés). Se trata de un principio propuesto en 1983 por Edward Tufte, un importante teórico de la disciplina de la visualización, y que introduce una de las condiciones básicas que debemos cumplir si perseguimos una visualización eficiente.

Si, como vimos anteriormente en la introducción, el principio de la objetividad de la información defiende la verdad frente a la manipulación o el engaño a la hora de construir una buena visualización, el principio de la ratio dato/tinta nos recomienda centrarnos en el criterio de la eficacia comunicativa por encima del efecto estético.

Para abordar el alcance de este principio de la ratio dato/tinta, hemos de introducir primero algunas ideas básicas. Comencemos señalando que en toda visualización podemos distinguir dos tipos diferente de tinta. Por un lado, tendríamos la tinta-dato, tinta utilizada en la visualización para representar datos. Una especie de tinta semántica, podríamos decir en cierto sentido. Y, por otro, la tinta-no dato, tinta utilizada en la visualización pero que no representa datos. Una especie de tinta sintáctica, en cierta manera y un sentido más técnico.

En las figuras 14 y 15 se pueden encontrar ilustrados estos dos tipos de tinta en los típicos ejemplos de productos de visualización: en una tabla y en una gráfica. En la primera figura (14), la tabla de arriba incluye tanto la tinta-dato como la tinta-no dato. En la tabla de abajo a la izquierda, se destaca en negro sólo la tinta-dato, la que se utiliza para representar datos (toda la tinta menos la utilizada para dibujar las líneas y el sombreado). Así, tenemos la tinta para escribir los nombres de las filas y las columnas, los valores numéricos y el título de la tabla. Y, en la tabla de abajo a la derecha, se destaca en negro y en un gris más intenso sólo la tinta-no dato, la tinta que aparece en la tabla pero que no ha sido utilizada para representar datos, como las líneas y el sombreado.

En la segunda figura (15), la gráfica de arriba incluye tanto la tinta-dato como la tinta-no dato. En la gráfica de abajo a la izquierda, se destaca en negro y en grises más intensos sólo la tinta-dato (toda la tinta menos la utilizada para dibujar las líneas de los ejes). Así, tenemos la tinta utilizada para escribir el nombre de las categorías que aparecen en el eje X, las categorías y las marcas de color en la leyenda, el título de la gráfica, las etiquetas numéricas del eje Y y las columnas que representan los fumadores en Italia y España. Y, en la gráfica de abajo a la derecha, se destaca en negro sólo la tinta-no dato, la tinta que aparece en la gráfica pero que no ha sido utilizada para representar datos, como las líneas de los ejes X e Y.

Con esa distinción en la mano, podemos introducir ahora la noción de ratio dato/tinta. La ratio dato/tinta es la proporción entre la cantidad de tinta utilizada para representar los datos (tinta-dato) y la cantidad de tinta total utilizada en la visualización (tinta-dato + tinta-no dato). O dicho de otra manera: tinta-dato / (tinta-dato + tinta-no dato).

Figura 14. Tinta-dato y tinta-no dato en una tabla

Gastos	Departamento	
	Marketing	Tecnología
Viajes	4.567	7.895
Dietas	29.876	26.456
Material	15.695	12.456
Total	50.138	46.807

Tabla 1. Gastos por departamento

Gastos	Departamento	
	Marketing	Tecnología
Viajes	4.567	7.895
Dietas	29.876	26.456
Material	15.695	12.456
Total	50.138	46.807

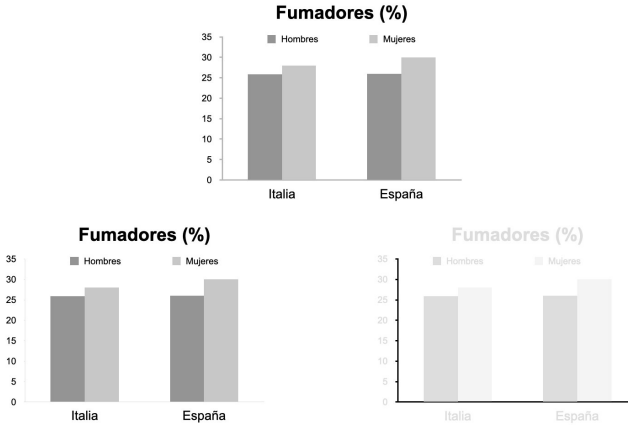
Tabla 1. Gastos por departamento

Gastos	Departamento	
	Marketing	Tecnología
Viajes	4.567	7.895
Dietas	29.876	26.456
Material	15.695	12.456
Total	50.138	46.807

Tabla 1. Gastos por departamento

(elaboración propia, datos simulados)

Figura 15. Tinta-dato y tinta-no dato en una gráfica



Aclarados estos conceptos, ya podemos introducir el principio ratio dato/tinta: toda buena visualización debe obtener un valor de esa ratio cercano (o igual) a 1. O, dicho en otros términos, la mayoría de la tinta que aparezca en una visualización debe estar dedicada a representar datos y no a otro tipo de funciones, como la de llamar la atención del receptor de la visualización.

- c) *Principio de la Ratio dato/tinta*: en una gráfica, la cantidad de tinta utilizada para representar datos dividida por la cantidad de tinta total utilizada en una visualización debe ser igual o cercana a 1.

Es importante destacar un par de cosas relacionadas con este principio. La primera es ver qué significan los casos extremos en los que se cumple o no se cumple el principio enunciado. Por un lado, cuanto mayor es esa ratio (cercana a 1), más tinta en la visualización se utiliza para representar datos. El caso extremo acontece cuando esa ratio es igual 1. En ese caso, toda la tinta utilizada (el 100%) representa información. Por otro lado, cuanto menor es esa ratio (cercana a 0), menos tinta en la visualización se utiliza para representar información. El caso extremo se produce cuando el valor de esa ratio es igual a 0. En ese caso, ninguna cantidad de tinta utilizada (el 0%) está representando información.

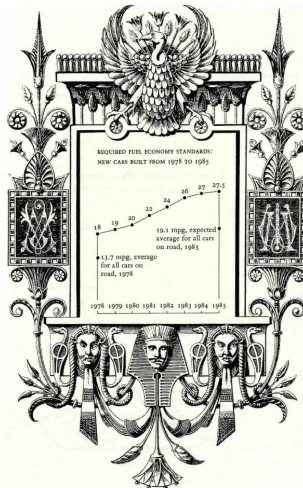
La segunda cosa que queremos destacar es que, en cierta manera, en el principio no se exige un valor de la ratio exacto a 1 (y sí, en cambio, nos recomienda un valor cercano a 1) porque podemos distinguir dos tipos de tinta-no dato dentro de una visualización.

Por un lado, tendríamos la tinta-no dato que, aunque estrictamente no esté representando datos en esa visualización, nos ayuda a estructurar y entender mejor la información representada visualmente. Esta tendría una función básicamente sintáctica dentro de la visualización. Un ejemplo de este tipo de tinta-no dato sería la utilizada en las tablas para dibujar las líneas que facilitan una lectura por filas o por columna o el fondo de color utilizado para destacar un conjunto de datos concretos dentro de la tabla. O, en el caso de las gráficas, un ejemplo de este tipo de tinta-no dato sería la utilizada para dibujar los ejes X e Y.

Y, por otro, tendríamos la tinta-no dato que, aunque estrictamente no esté representando datos en esa visualización, nos ayuda a generar otros efectos más indirectos como la mejora de la estética de la visualización o provocar la sorpresa y el interés en el receptor a partir de ese producto visual. En cierta manera, este segundo tipo de tinta-no dato podría ser considerado como una especie de ruido visual. En un sentido más técnico, este tipo de tinta se denomina “*chartjunk*” (en inglés), y puede ser entendido como un elemento gráfico superfluo o innecesario para comprender la información representada en una visualización.

El propio Tufte (1983) nos suministra en la figura 16 un ejemplo clásico de visualización poco eficiente porque no cumple el principio formulado. Allí aparece una gráfica rodeada de un marco barroco innecesario y cargada de tinta superflua

Figura 16. Visualización con una escasa ratio dato/tinta



Tufte, Edward (1983). *The Visual Display of Quantitative Information*

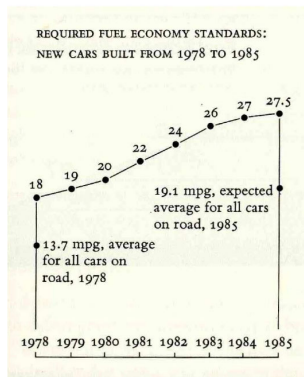
(<https://strajk.me/notes/2012/the-visual-display-of-quantitative-information-by-edward-r-tufte>)

comunicativamente como la utilizada para representar todos los adornos de máscaras y otros motivos florales y estéticos. En la figura 17, en cambio, podemos encontrar una versión de la gráfica anterior con una mejor ratio, donde sólo aparece tinta dedicada a la representación de datos, más una pequeña proporción de tinta (sintáctica, si utilizamos la distinción anterior), como la que dibuja el eje X, que, aunque no representa estrictamente datos, nos permite articular y entender mejor la información representada en esa visualización.

Otro ejemplo clásico que ilustra el principio introducido lo podemos encontrar en la figura 18. La figura, del diseñador gráfico Nigel Holmes y recogida por Tufte (1990), nos introduce una gráfica en el que se intenta representar la evolución del precio de los diamantes desde 1978 hasta 1982. Pero, lamentablemente, no toda la tinta que aparece en este producto visual está dedicada a representar la evolución de esos datos en el tiempo. Presenta una cantidad ingente de tinta (más del 70%, podríamos decir) que representan otras cosas y que no contribuyen a la satisfacción de las necesidades informativas de sus potenciales receptores. Motivado por algún dudoso y poco políticamente correcto gusto estético, el autor invierte tinta en el dibujo de una mujer tocada de un sombrero de copa, con un corsé, cargada de diamantes y vestida con una mallas que aprovecha para hacerla coincidir con la cuadrícula de la gráfica. La línea de la evolución temporal se aprovecha para contornear la pierna derecha de la mujer.

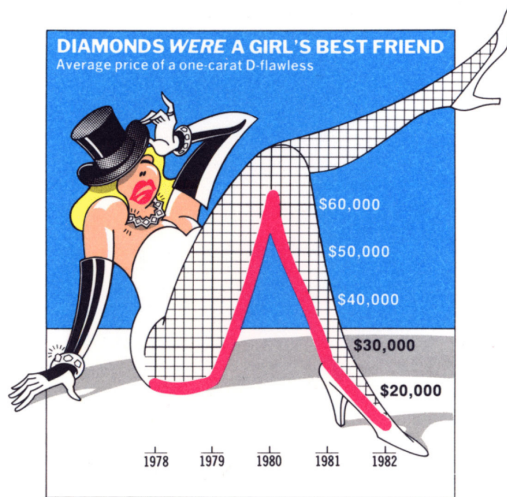
Con estas ideas y estos ejemplos presentados, ya estamos en disposición de poder introducir las consecuencias que acarrea este principio a la hora de crear una visualización de información. O, dicho de otra manera: cómo podemos utilizar este

Figura 17. Visualización con una mayor ratio dato/tinta



Tufte, Edward (1983). *The Visual Display of Quantitative Information*

(<https://strajk.me/notes/2012/the-visual-display-of-quantitative-information-by-edward-r-tufte>)

Figura 18. Visualización con una ratio dato/tinta muy baja

(www.danielpradilla.info/blog/wp-content/uploads/2012/11/Tufte-Chartjunk.png)

principio para generar visualizaciones de información más comunicacionalmente eficientes (Tufte, 1983).

En primer lugar, y en términos generales, cumplir el principio asegura que en la visualización se persiga por encima de todo comunicar, no sólo crear un efecto estético o de sorpresa en el receptor. Está claro que el efecto estético de las visualizaciones refuerza las estrategias comunicativas que las incluyen. Pero apostar sólo por su aspecto estético y descuidar su eficiencia representativa se vuelve en contra de nuestra propia estrategia comunicativa. En este sentido, esa tendencia actual en las redes sociales y en los cibermedios de utilizar lo que en la jerga de la disciplina se conoce como visualizaciones “*eye candy*” (“visualizaciones bombón”, podríamos llamarlas; visualizaciones cuyo objetivo no es tanto representar información como causar un efecto estético) dejan de considerarse como buenas y eficientes visualizaciones.

Por otro lado, para garantizar esa eficiencia, hay que maximizar la ratio dato/tinta dentro de lo razonable. Hay que introducir la tinta-no dato de tipo sintáctico en su justa medida. La que necesitamos para garantizar la comprensión de lo que queremos comunicar por parte de sus receptores. Y evitar de forma exhaustiva la incorporación de ruido visual. Eso significa no añadir tinta, de cualquier tipo (dato o no dato), superflua o redundante.

También debemos justificar la aparición de cualquier superficie o porción de tinta en la visualización que estamos creando. La justificación principal es la

transmisión efectiva y eficiente de información. Si la aparición de cierta porción de tinta no encuentra una justificación en el aseguramiento de la representación adecuada de información o su correcta aprehensión, no debemos incluirla en la visualización.

Además, si añadimos más tinta a una visualización en la que estamos trabajando, es porque añadimos nueva información a nuestro producto visual. No porque queramos mejorar su aspecto estético o busquemos llamar la atención de sus potenciales destinatarios.

Y, por último, la recomendación de cumplir el principio en el diseño de toda visualización (ya sea una gráfica o una tabla) se implementa sobre dos acciones distintas pero a la vez articuladas: reduciendo la tinta-no dato (tinta que no representa información) al mínimo y aumentando y enfatizando al máximo la tinta-dato (tinta que representa información).

La reducción al mínimo de la tinta-no dato se consigue de dos maneras distintas. Por un lado, eliminando al máximo la tinta-no dato innecesaria. Y, por otro, atenuando y regularizando la tinta-no dato que se mantiene en la representación; que esté presente para que podamos utilizarla como receptores, pero sin destacar en exceso.

En cambio, el aumento y enfatizado máximo de la tinta-dato se consigue por dos vías. Primero, eliminando al máximo la tinta-dato innecesaria, para que aflore la que sí está representando la información que queremos transmitir. Y, segundo, enfatizando con recursos gráficos la tinta-dato más importante que se mantiene en esa visualización (ya sea en una tabla o en una gráfica).

En definitiva, para generar buenos productos visuales, lo que tenemos que hacer es cumplir este principio de la ratio dato/tinta, pero sin olvidar también la máxima que circula normalmente entre los buenos diseñadores de la información: sin información, no hay visualización. La transmisión información debe ubicarse en el centro y debe convertirse en el motor de la comunicación a través del uso de una visualización. O como dicen coloquialmente los diseñadores cuya lengua materna es el inglés: *no info, no infographics*.

Caracterización y anatomía de las gráficas

1. Introducción

En el capítulo anterior hemos podido introducir algunos de los conceptos básicos que nos permiten comenzar a entender la disciplina de la visualización de información. En ese capítulo, revisamos en qué consiste esa disciplina, introdujimos la distinción entre los conceptos de infografía y visualización de datos y abordamos algunas de las nociones básicas relacionadas con la percepción que debemos tener en cuenta a la hora de diseñar y crear buenos productos visuales para la comunicación de información.

En este segundo capítulo, vamos a intentar caracterizar uno de los principales productos visuales que se utilizan para representar y transmitir la información: las gráficas. Si dejamos a un lado el tema de las tablas, la gráfica se presenta como uno de los productos de visualización más utilizados y más presente en los contextos comunicativos que incluyen algún tipo representación de datos. Por tanto, para su correcta utilización, debemos, primero, tener claro en qué consiste este tipo especial de productos.

Para cubrir el objetivo de caracterizar las gráficas, vamos a tratar de desarrollar los siguientes puntos. En el primero de los apartados (apartado 2), se explicará cómo podemos definir una gráfica, introduciendo las características básicas que estructuran este tipo de productos e ilustrando a partir de un par de ejemplos cómo se articulan esas características básicas en una gráfica. De la mano de esa definición analizaremos también los componentes estructurales que conforman las gráficas. Concretamente, introduciremos los elementos codificadores de información y los elementos de soporte que podemos encontrar en una gráfica, y los conectaremos con la distinción entre las diferentes clases de tinta que podemos utilizar en un producto visual de este tipo y que recogíamos en el principio de la ratio dato/tinta.

En el siguiente apartado (apartado 3), abordaremos las principales diferencias básicas existentes entre los dos principales productos de visualización: las gráficas y las tablas. Veremos cuáles son los contextos comunicacionales más propicios para

utilizar cada uno de esos productos y, utilizando una serie de ejemplos ilustrativos, mostraremos las ventajas y limitaciones de cada uno de esos usos. Por último, en el apartado 4, acabaremos el capítulo abordando la tipología de las principales gráficas que solemos utilizar con intención comunicativa. Veremos su anatomía, la ilustraremos con ejemplos y adelantaremos algunos de los usos que podemos darle a cada uno de esos tipos.

2. Definición de gráfica

En el capítulo anterior pudimos introducir la definición de la disciplina de la visualización de la información. De forma intuitiva, defendimos que se trataba de una disciplina que persigue la representación de contenidos proposicionales (de ideas, en definitiva) mediante el uso de algunos productos visuales.

Entre esos productos visuales se incluyen, entre otros, los diagramas, las tablas, los mapas o los esquemas (Tukey, 1977). Pero, sin duda, el producto más utilizado para la representación y comunicación de datos es la gráfica (Spence, 2000; Yau, 2011 y 2013; Zelazny, 2001).

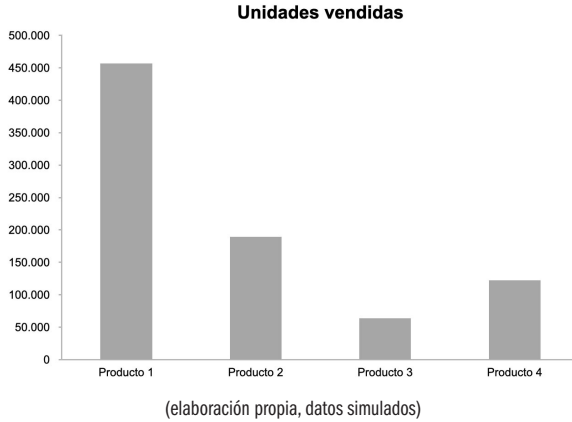
Una gráfica incluye elementos visuales que se relacionan entre ellos para representar información, principalmente de tipo cuantitativo. Si atendemos a esos elementos visuales y a sus posibles relaciones, podemos introducir una definición de gráfica (Few, 2012).

- a) *Definición de gráfica*: instrumento para visualizar información cuantitativa que presenta una serie de características. Por un lado, los valores cuantitativos se codifican como elementos u objetos gráficos. Por otro, esos valores se representan dentro de un área delimitada por uno o más ejes que nos ayudan a ubicar localizaciones espaciales en un plano bidimensional. Y, por último, los ejes proporcionan escalas (cuantitativas y cualitativas) que se utilizan para asignar valores y etiquetas a esos elementos gráficos.

Si fijamos nuestra atención en la figura 1, podemos ver ilustrados todos esos componentes y las relaciones que mantienen entre ellos y cómo se cumple la definición de gráfica.

En la figura se están representando las unidades de los cuatros productos fabricados por una misma empresa y que han sido vendidos durante el pasado año. Como vemos, se cumplen las características atribuibles a una gráfica.

Por un lado, los valores se codifican como objetos o elementos gráficos. En este caso, las unidades vendidas se representan mediante columnas de color gris

Figura 1. Unidades vendidas de los cuatro productos fabricadas por una misma empresa durante el pasado año

(objetos o elementos gráficos) que se ubican de una determinada manera respecto a esos dos ejes para que adquieran un significado concreto.

Por otro, esos valores cuantitativos se representan dentro de un área delimitada por uno o más ejes que nos ayudan a ubicar localizaciones espaciales en un plano bidimensional. En este sentido, la cantidad de unidades vendidas de cada producto fabricado se representan dentro del espacio generado por dos ejes. Esos ejes reciben el nombre técnico de “coordenadas cartesianas” o “ejes cartesianos”. El eje X recibe el nombre de “abscisas” y el eje Y el de “ordenadas”. La intersección de los dos ejes se denomina técnicamente “origen de coordenadas”.

Y, por último, los ejes nos proporcionan las escalas (cuantitativas y cualitativas) que se utilizan para asignar valores y etiquetas a esos elementos gráficos. En el eje X se representa la variable cualitativa (cada uno de los cuatro productos vendidos); y en eje Y se representa la variable cuantitativa (la cantidad de unidades vendidas). Esos dos ejes dotan de sentido a cada uno de los elementos gráficos (las columnas) que representan las unidades vendidas. El eje X nos permite saber a qué producto está representando una columna gris determinada. La posición extrema superior de esa figura geométrica gris respecto al eje Y determina el valor numérico, el número de unidades vendidas del producto representado por la columna.

Esta caracterización podría ser considerada como una definición funcional de gráfica. Pero en una gráfica no sólo conviven ejes y elementos gráficos. Es posible acercarnos a una gráfica y poder discriminar todos los elementos que la conforman (Cleveland, 1994).

Para identificar esos componentes estructurales de una gráfica, podemos recuperar la distinción introducida por Tufte (1983) dentro de su principio de la ratio

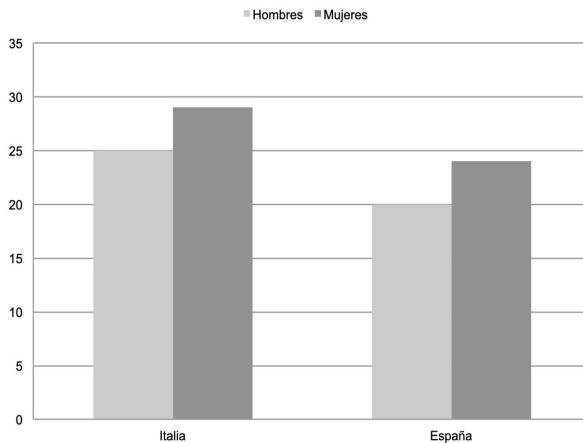
dato/tinta. Según este autor, podemos distinguir dos tipos de tinta dentro de una gráfica: la tinta-dato (tinta dedicada a la representación de datos en una gráfica) y la tinta-no dato (tinta utilizada en la visualización pero que no representa datos). Y de la mano de esta distinción, podemos identificar dos tipos principales de componentes estructurales dentro de una gráfica: los elementos codificadores de información y los elementos de soporte (Few 2012).

Los elementos codificadores de información podrían ser identificados como aquellos objetos que aparecen en una gráfica, que han sido dibujados utilizando tinta-dato y que están representando información dentro de esa gráfica. Serían algo así como los *componentes semánticos* (como las palabras en el lenguaje natural) de las gráficas, dotados de significado semántico, al estar representando información.

Entre esos componentes semánticos podemos incluir, principalmente, los elementos codificadores de valores cuantitativos, los elementos codificadores de ítems cualitativos, las escalas, la leyenda y el texto complementario. Para ilustrar todos elementos podemos utilizar la figura 2.

Los elementos codificadores de valores cuantitativos son aquellos objetos visuales (normalmente se utilizan puntos, líneas, barras, columnas o áreas, entre otros) que aparecen en una gráfica y que están representando la información numérica en ésta (Bertin,1983). En nuestro ejemplo de la figura 2, cada una de las cuatro columnas grises que están representado los porcentajes de fumadores (hombres y mujeres) en España e Italia sería un componente de este tipo. Cada una de esas

Figura 2. Fumadores en Italia y España durante el pasado año (%)



Gráfica 1. Fumadores en Italia y España durante el pasado año (%)

(elaboración propia, datos simulados)

columnas codifica esa información numérica confrontando su longitud respecto el eje Y y a sus marcas y etiquetas.

Esos mismos objetos visuales (puntos, líneas, barras, columnas o áreas, entre otros) también pueden ser considerados simultáneamente como elementos codificadores de ítems cualitativos. Son elementos bifrontes, por así decirlo, dotados de la capacidad de codificar o representar al mismo tiempo tanto información cuantitativa como ítems cualitativos. Se trata de objetos que aparecen en una gráfica y que están representando cada una de las categorías que forman las variables cualitativas incluidas en esa gráfica. En la figura 2, cada una de las cuatro columnas grises que están representado las mujeres y los hombres italianos y españoles que fuman sería un componente de este tipo. Cada una de esas columnas codifica ese ítem cualitativo confrontando su posición respecto el eje X y a sus marcas y etiquetas.

Otro de los componentes semánticos de una gráfica son las escalas que forman parte de los ejes. En nuestro ejemplo, la escala cuantitativa (que va del valor 0% al 35%) recogida en el eje Y; y la escala categorial plasmada en el eje X (que incluye las categorías “Italia” y “España”). La leyenda es otro de esos elementos codificadores de información. Una leyenda es una especie de micrográfica dentro de la gráfica dónde se nos da la clave para, junto a la información suministrada por lo ejes, poder dotar de significado a los objetos visuales que representan simultáneamente los valores cuantitativos y los ítems cualitativos. En nuestro ejemplo, la leyenda está formada por dos cuadrados grises y las etiquetas textuales “Hombres” y “Mujeres” y nos permite entender qué está representando cada una de las cuatro columnas que aparecen en la gráfica. Por último, entre los elementos semánticos hay que incluir también el texto complementario que aparece en la gráfica y que nos ayuda a completar y entender la información que se pretende transmitir utilizando ese recurso visual; especialmente, y entre otros, el título de esa gráfica. En la figura 2, el título de la gráfica (“Fumadores en Italia y España durante el pasado año (%)”) es parte de ese texto complementario al que aludimos y nos permite contextualizar en el espacio y en el tiempo los porcentajes de fumadores representados en la gráfica.

Por otro lado, y frente a los componentes semánticos, los elementos de soporte podrían ser identificados como aquellos objetos que aparecen en una gráfica, que han sido dibujados utilizando tinta-no dato y que están realzando u organizando la información dentro de esa gráfica. Serían algo así como los *componentes sintácticos* (como los signos de puntuación o el uso de negritas en el lenguaje natural escrito, por ejemplo) de las gráficas, no dotados de un significado semántico concreto y sin la función de representar información. Su misión principal se centra en la estructuración sintáctica de las gráficas para facilitar su lectura e interpretación.

Entre esos componentes sintácticos podemos incluir, principalmente, los ejes, la región de datos y las líneas de cuadrícula. Para ilustrar todos elementos podemos volver a utilizar la figura 2.

Los ejes son dos líneas perpendiculares que se cruzan en un punto concreto. Cómo ya adelantamos, el eje horizontal se denomina “eje X” o “eje de abscisas” y el eje vertical viene nombrado como “eje Y” o “eje de ordenadas”. En nuestro ejemplo, en el eje horizontal (eje X) se ubican los ítems cualitativos (hombres y mujeres italianos y españoles que fuman) que se van a representar en la gráfica. Y en el eje vertical (eje Y) se ubica la escala que nos va a permitir averiguar el valor numérico (en porcentaje) de esos fumadores.

Técnicamente, la región de datos coincide con el área generada por el cruce de los dos ejes y en la que se visualiza o se representa la información. En la figura 2, la región de datos está generada por los ejes X e Y y es la zona o área donde representamos, mediante el uso de objetos visuales, los valores cuantitativos (porcentaje de fumadores) asociados a los ítems cualitativos (los diferentes tipos (separados por género) de fumadores españoles e italianos).

Por último, las líneas de cuadrícula (denominadas también como “retícula” o “parrilla”) son combinaciones de líneas verticales y/o horizontales que ocupan toda la región de datos y que se cruzan entre ellas formando rectángulos para facilitar la identificación, lectura y comprensión de la información representada en una gráfica por parte de sus potenciales usuarios. En nuestro ejemplo sólo se han dibujado las líneas horizontales y se han obviado las verticales. Estas líneas se corresponden con cada una de las marcas de la escala representada en el eje Y y nos permiten y facilitan la asignación de un valor numérico a cada una de las cuatro columnas grises que están representando a los diferentes tipos de fumadores (hombres y mujeres) de España e Italia.

Nos gustaría acabar este apartado dedicado a la definición de gráfica describiendo también otros elementemos más secundarios (y ofreciendo además los términos técnicos que se utilizan para identificarlos) que, aunque no los hayamos incluido entre los componentes codificadores de información y los elementos de soporte, pueden aparecer en una gráfica y pueden permitirnos entender mejor su estructura y la información que se pretende transmitir con ésta. Abordar esta terminología es una forma de suministrar, también, un utillaje léxico adecuado para facilitarnos la lectura posterior de fuentes bibliográficas sobre la disciplina de la visualización de información.

Entre esos componentes más secundarios podemos incluir, principalmente, el subtítulo, las fuentes y la autoría, el borde de la gráfica, el título del eje Y, el título del eje X, la etiqueta del eje Y, la etiqueta del eje X, las marcas de los ejes, la línea de

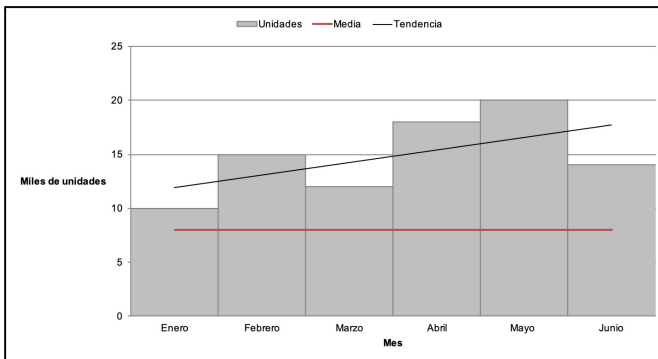
tendencia y la línea de referencia. Para ilustrar todos estos componentes podemos utilizar, en este caso, la figura 3.

El subtítulo es una composición textual que viene a completar y contextualizar la información representada en el título de la gráfica. En nuestro ejemplo, el subtítulo es la composición textual “Primer semestre del año anterior” y nos permite aclarar que la información representada por el título (las unidades que han sido vendidas) se corresponden con las del primer semestre de ese año concreto.

La declaración de las fuentes y el origen de los datos utilizados para realizar la gráfica y la autoría de la misma son dos elementos importantes que se recomienda incluir en toda gráfica bien construida. Son elementos que ayudan a contextualizar la calidad y el alcance del mensaje representado en ese producto visual. En la gráfica 3 hemos colocado esos elementos entre paréntesis (“Fuente: CompletelyFakeData Inc.; Autoría: ACME”) a continuación del subtítulo, para que el potencial usuario de la gráfica compruebe de dónde se han extraído los datos para construirla y quién es el autor de esa visualización.

En cambio, el borde de la gráfica es uno de los elementos más opcionales que podemos incluir en nuestro producto de visualización. Es una línea que normalmente conforma un rectángulo y que suele rodear a la gráfica. En nuestro ejemplo, es el rectángulo que encapsula a la totalidad de la gráfica, incluida la leyenda. Su incorporación es opcional. Solo se suele utilizar si nuestra gráfica debe convivir con otras gráficas en un área muy pequeña y corremos el riesgo de que la cercanía de las gráficas provoque un solapamiento visual de las mismas. En esos casos, está justificada la utilización del borde para diferenciar visualmente dónde empieza y

Figura 3. Unidades vendidas. Primer semestre del año anterior



Gráfica 1. Unidades vendidas. Primer semestre del año anterior
(Fuente: CompletelyFakeData Inc.; Autoría: ACME)

(elaboración propia, datos simulados)

termina cada una de las gráficas cercanas. Si no es ese el caso, es mejor no incluir ese borde. Si lo incorporáramos de forma innecesaria estaríamos incluyendo tinta-no dato en la gráfica de forma gratuita. Y, de esta manera, haríamos disminuir de forma arbitraria e injustificada la ratio dato-tinta de esa gráfica y su propia eficiencia comunicativa (utilizaremos más tinta para comunicar exactamente lo mismo).

Los títulos de los ejes X e Y son otros de los componentes que debemos incluir en una gráfica. Esos títulos nos ofrecen las claves para poder interpretar correctamente los objetos visuales (puntos, líneas, barras, columnas o áreas, entre otros) que representan la información en una gráfica. En la gráfica 3, el título del eje X (“Mes”) nos permite saber que cada una de las columnas grises están representando las unidades vendidas en un mes de ese semestre. En cambio, el título del eje Y (“Miles de unidades”) nos permite interpretar el valor cuantitativo asignado a la longitud de cada una de las seis columnas grises. Es importante recordar que ambos títulos deben estar escritos formando un renglón horizontal (y no en vertical o en diagonal, como nos proponen algunos softwares de visualización) para facilitar una lectura cómoda y natural por parte del posible usuario de la gráfica.

Podemos completar el título de los ejes añadiendo también etiquetas en esos mismos ejes. Esas etiquetas nos ayudarán a identificar con mayor finura y de forma individualizada la información representada por los objetos visuales a los que hacíamos referencia. En nuestro ejemplo, las etiquetas del eje X (“Enero”, “Febrero”, “Marzo”, “Abril”, “Mayo” y “Junio”) nos permiten inferir qué mes se corresponde con cada una de las seis columnas grises. En el mismo ejemplo, las etiquetas del eje Y (“0”, “5”, “10”, “15”, “20” y “25”) nos permiten identificar con exactitud el valor numérico de las unidades vendidas que se corresponden con la longitud de cada una de las seis columnas grises.

En una gráfica, para saber el lugar exacto donde se aplican esas etiquetas en los ejes X e Y, utilizamos las marcas de eje (*tick marks*, en inglés). Se trata de pequeñas secciones de línea que se adhieren al eje (a veces justo a la izquierda, otras a la derecha o traspasándolo) de forma perpendicular para mostrarnos en qué punto exacto de ese eje se aplica una etiqueta concreta. En la figura 3, podemos ver que el eje Y está dividido por pequeñas marcas que nos permiten saber el punto exacto del eje que se corresponde con cada una de las etiquetas (“0”, “5”, “10”, “15”, “20” y “25”) del mismo. De esta manera, confrontado frente a esas marcas y etiquetas, podemos hacer una lectura muy fina y certera de las unidades vendidas en un mes concreto a partir de la longitud de la columna gris que está representando ese mes en la gráfica. En este caso, en el eje X no es necesario incluir esas marcas ya que la ubicación de las propias columnas nos indica con qué mes se corresponden.

Los últimos elementos que queremos destacar en una gráfica son las líneas de tendencia y de referencia. En algunas ocasiones, nos interesa resaltar en una misma gráfica, no sólo los elementos que están representando los valores cuantitativos asociados a ítems cualitativos, sino también otros datos más globales e indirectos asociados a esa representación más primaria.

Por ejemplo, para mostrar la evolución global de un conjunto de datos, podemos utilizar la línea de tendencia. Se trata de una línea que suelen generar la mayoría de las herramientas informáticas para la creación de gráficas y que representan ese patrón que ofrece un conjunto de datos concretos. En nuestro ejemplo, la línea de tendencia aparece inclinada, mostrándonos que la tendencia de las unidades vendidas a lo largo del semestre ha ido aumentando desde el mes de enero al de junio.

En cambio, a veces podemos querer representar en una gráfica cómo un conjunto de datos varía frente a un dato concreto. En esos casos podemos utilizar una línea de referencia: representamos el dato respecto al cual comparar el conjunto como una línea y podemos contrastar visualmente el comportamiento de los datos que forman parte de ese conjunto. En la figura 3, con la línea de referencia (línea roja horizontal) hemos representado la media de unidades vendidas a lo largo del año anterior y de esta manera podemos comprobar cómo ha variado, respecto a esa media, la cantidad de unidades vendidas en cada uno de los meses del primer semestre de ese año.

3. Gráficas versus tablas

Como ya adelantamos en el capítulo anterior, dentro del paraguas de la visualización de información se cobijan y conviven una diversidad de tipos de propuestas visuales que nos permiten representar contenidos proposicionales con una intención comunicativa y/o de análisis. Así, para conseguir ese tipo de representación podemos generar diagramas, esquemas o mapas, entre otras muchas alternativas.

Pero, sin duda, los dos productos más utilizados para la representación visual de la información cuantitativa son las tablas y las gráficas. Ambas propuestas de representación aparecen en la mayoría de los trabajos donde se intentan analizar y comunicar aspectos relacionados con un conjunto de datos, independientemente de su tamaño y del contexto en el que éstos han sido recogidos.

Pero esos dos tipos de representaciones (las gráficas y las tablas) no son libremente intercambiables a la hora de codificar la información. Cada uno de ellos presenta unas características y un poder de representación propio, derivado de su estructura sintáctico-visual, que los habilita para representar de forma eficiente ciertos contenidos semánticos y que, a su vez, impiden que puedan codificar otros

contenidos alternativos de forma adecuada. O dicho en otros términos: no todos los contenidos se pueden representar de forma indiferente con una tabla o con una gráfica. Las características proposicionales de ese contenido y las características básicas estructurales (la estructura sintáctica, por así decirlo) de las tablas y las gráficas favorecen o no la eficiencia de esa representación. En función de la intención comunicativa que persigamos, sólo ciertos contenidos pueden ser representados adecuadamente con una gráfica, mientras que otros nos reclaman el uso de una tabla.

Para abordar esa adecuación en el uso de las gráficas o las tablas, vamos a comenzar refrescando las características básicas estructurales de cada uno de estos dos productos visuales.

Como ya hemos podido ver en el apartado anterior, una gráfica es un instrumento para visualizar información que codifica valores cuantitativos asociados a ítems cualitativos mediante el uso de elementos gráficos referenciados respecto a dos ejes perpendiculares.

En cambio, una tabla es una propuesta de visualización que fundamenta su poder representativo en la presentación y organización de los datos codificados (Few, 2012).

- b) *Definición de tabla*: instrumento para visualizar y organizar información cuantitativa que presenta una serie de características. Por un lado, los datos se encuentran codificados de forma alfanumérica, a partir de la utilización de números y texto. Por otro, esos datos están organizados en forma de cuadrícula mediante una estructura de filas y columnas. Y, por último, esas filas y columnas se suelen reforzar visualmente a través del uso de líneas que ayudan a generar una parrilla formada por celdas y a facilitar la interpretación de esa estructura por parte de sus potenciales usuarios.

Si, como hicimos en el caso de las gráficas, fijamos ahora nuestra atención en la figura 4, podemos ver ilustrados todos esos componentes y las relaciones que mantienen entre ellos y cómo se cumple la definición de tabla.

Como podemos comprobar, en la figura, los datos se encuentran codificados de forma alfanumérica, a partir de la utilización de números y texto. Así, para codificar a los empleados y sus cargos utilizamos unidades textuales (como “Empleado A” o “Gerente”, por ejemplo). Y para codificar los sueldos de cada uno de esos empleados utilizamos combinaciones numéricas (como “150.000” o “40.000”, por ejemplo). Por otro lado, esos datos están organizados en forma de cuadrícula mediante una estructura de filas y columnas. La tabla incluida en la figura está estructurada en tres columnas (cuyas cabeceras son “Empleado”, “Cargo” y “Sueldo”) y cada una de las filas está representado un empleado concreto con su cargo y sueldo.

Figura 4. Cargo y sueldo de los empleados de la empresa ACME

Empleado	Cargo	Sueldo
Empleado A	Gerente	150.000
Empleado B	Analista	40.000
Empleado C	Documentalista	38.000
Empleado D	Administrativo	32.000
Empleado E	Auxiliar	25.000
Empleado F	Becario	8.000
Sueldo medio		48.833

(elaboración propia, datos simulados)

Y, por último, esas filas y columnas se suelen reforzar visualmente a través del uso de líneas que ayudan a generar una parrilla formada por celdas y a facilitar la interpretación de esa estructura por parte de sus potenciales usuarios. En la figura hemos utilizado dos líneas horizontales (denominadas técnicamente como “filetes”). Una para separar los nombres de las columnas (“Empleado”, “Cargo” y “Sueldo”) de sus posibles valores. Y otra para separar los datos de la tercera columna de un valor promedio (en este caso la media aritmética) que representa el conjunto de los datos recogidos en esa columna (la de los sueldos).

Es importante señalar un par de aspectos que se derivan de la definición que acabamos de introducir y de ilustrar.

Por un lado, que la estructura de las tablas respeta lo que se conoce como el modelo entidad/relación, concepto central en el contexto de las bases de datos y de la disciplina de la recuperación de la información. En ese contexto, las filas de una tabla pueden ser consideradas como los registros, como las entidades representadas en la tabla, en definitiva. En nuestro ejemplo, los registros serían los empleados representados en la tabla. Y, en ese mismo contexto, las columnas de una tabla podrían ser identificadas como los campos de esos registros, como atributos de las entidades representadas. En la figura 4, las columnas (código del empleado (A, B, etc.), cargo y sueldo) serían los campos de esos registros (o los atributos de esas entidades, de esos empleados que estamos representando).

Y, por otro, se puede defender también de forma complementaria que las tablas presentan una estructura de variables y observaciones (Wickham, 2017). En este sentido, cada columna debe considerarse como una variable que puede adquirir una serie de valores. En nuestro ejemplo, existen tres variables: empleado, cargo y sueldo. Y cada fila debe ser considerada como una observación de esa variable, como uno de los posibles estados o valores que puede adoptar esa variable. Así,

en la figura, la variable *Empleado* recibe en cada una de las celdas que forman la columna uno de los posibles seis empleados (A, B, C, D, E y F), la variable *Cargo* puede adquirir uno de los puestos de trabajo desempeñados en la empresa ACME de los siguientes valores (gerente, analista, documentalista, administrativo, auxiliar y becario), y la variable *Sueldo* se instancia en cada uno de los salarios brutos anuales que reciben esos cargos. Cada una de las filas será una observación donde se representa el empleado, su cargo y su sueldo. En esta misma línea, la tabla en su conjunto puede ser entendida como una unidad de observación; y, en nuestro ejemplo, como una descripción de la empresa ACME en un momento de tiempo determinado.

Una vez que tenemos caracterizadas tanto las gráficas como las tablas, ya podemos analizar el poder de representación propio, derivado de su estructura sintáctico-visual, que las habilita para representar de forma eficiente ciertos contenidos semánticos y que, a su vez, impide que puedan codificar otros contenidos alternativos de forma adecuada.

Comencemos señalando que las tablas codifican los datos representados de forma alfanumérica. Estructurando esos datos en filas y columnas, nos permiten, a través de nuestra competencia lectora, adquirir la información que allí aparece representada. En cambio, las gráficas codifican esos datos de forma visual. Utilizando objetos gráficos (líneas, barras, puntos o formas, entre otros), nos permiten, a través de nuestra percepción visual, aprehender la información que nos tratan de comunicar.

Y estas diferencias en la codificación nos van a permitir analizar los poderes de representación de las tablas y las gráficas y limitar de esa manera los contextos comunicacionales en los que cada uno de esos dos tipos de propuestas visuales funcionan mejor y son más adecuados si queremos garantizar con su uso la transmisión eficiente de información (Few 2012).

Centrémonos primero en las tablas. Veamos cómo éstas presentan una serie de ventajas frente a las gráficas a la hora de visualizar algunos tipos concretos de información.

Si analizamos la estructura de las tablas, podemos comprobar que éstas ofrecen una primera ventaja competitiva frente a las gráficas: facilitan y hacen cómoda la búsqueda e identificación de los valores representados. La estructura matricial de una tabla genera un espacio cartesiano dividido por las coordenadas que nos ofrecen las filas y las columnas. Y, utilizando como criterio esas columnas y filas, podemos identificar fácilmente los datos que estemos buscando. Así, por ejemplo, en la figura 4 podemos identificar fácilmente el cargo y el sueldo de cualquiera de los seis empleados representados si utilizamos las coordenadas de las columnas *Empleado*, *Cargo* y *Sueldo*, combinándolas con el criterio de las filas (*Empleado A*, *Empleado B*, *Empleado C*, *Empleado D*, *Empleado E* o *Empleado F*).

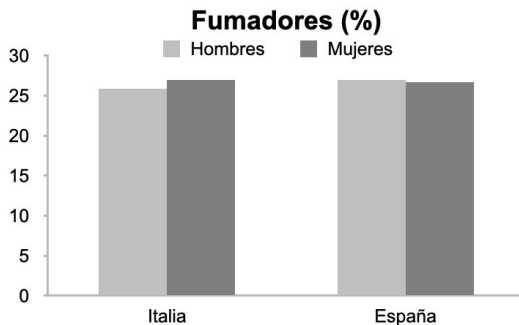
Esta estructura matricial nos permite utilizar las tablas para visualizar relaciones simples entre valores cuantitativos e ítems cualitativos con los que éstos están relacionados, ofreciendo así a los usuarios una manera sencilla de poder localizarlos, identificarlos y compararlos entre sí. En nuestro ejemplo de la figura 4, la tabla nos permite localizar los sueldos asociados a empleados y cargos concretos y comparar estos salarios con los que reciben otros empleados y cargos alternativos.

La segunda ventaja competitiva de las tablas frente a las gráficas radica en el tipo de codificación de datos que implementa. Las tablas codifican los datos de forma alfanumérica. Es decir, codifican los datos utilizando caracteres alfabéticos y numéricos. Y la aprehensión de estos datos a partir de esta codificación es relativamente cómoda. Nuestra competencia lectora nos habilita esta aprehensión con extremada naturalidad. Si dirigimos nuestra mirada a una tabla, podemos hacernos de una forma cómoda con los datos allí codificados simplemente a partir de su lectura. No es necesario ningún esfuerzo cognitivo extra.

Con la tabla que recogemos en la figura 5 podemos ilustrar esta característica. Si queremos hacernos con los datos allí codificados, sólo tenemos que leer los caracteres alfanuméricos que aparecen. De esta manera, podemos leer con comodidad que el porcentaje de los hombres fumadores en España en ese año es del 27,0065%, mientras que el porcentaje de las mujeres fumadoras en Italia en ese año es del 26,9876%.

Figura 5. Tabla y gráfica sobre los fumadores en Italia y España del año en curso (%)

País	Hombres (%)	Mujeres (%)
Italia	25,8906	26,9876
España	27,0065	26,6784



(elaboración propia, datos simulados)

Esa misma comodidad en la aprehensión no la encontramos en una gráfica. En una gráfica los datos no se encuentran codificados exclusivamente de forma alfanumérica. Necesitamos un esfuerzo cognitivo extra. Esos datos se codifican mediante el uso de elementos gráficos posicionados en relación con unos ejes. Por tanto, si queremos hacernos con los datos que aparecen representados en una gráfica no podemos recurrir exclusivamente a nuestra capacidad lectora. Debemos observar esos elementos, enfrentarlos a los ejes, identificar las etiquetas que estructuran esos ejes y extraer la información combinando y articulando todas estas acciones cognitivas. Si nos centramos en nuestra gráfica recogida en la figura 5, para poder saber cuál es el porcentaje de las mujeres fumadoras en Italia según esa visualización no podemos leerlo directamente, necesitamos interpretarlo, nos requiere un esfuerzo cognitivo suplementario. Es necesario que entendamos que la segunda columna (la primera más oscura que encontramos desde la izquierda) representa ese porcentaje gracias a las etiquetas del eje X y a la leyenda y que, según la escala que nos ofrece el eje Y, el valor numérico de ese porcentaje está alrededor del 27%.

La tercera ventaja competitiva de las tablas frente a las gráficas se fundamenta sobre la escalabilidad en su poder de codificación. Las tablas, al codificar los datos de forma alfanumérica, permiten expresar los valores numéricos con el grado de precisión que deseemos. Sin ningún tipo de limitación. Podemos codificar cualquier valor cuantitativo con todo el grado de finura que necesitemos. En la tabla de nuestro ejemplo de la figura 5, podemos expresar los porcentajes de los fumadores con toda la precisión que deseemos, incluyendo el número de decimales que necesitemos, sin límite.

Expresar esos valores cuantitativos con tanta finura en una gráfica es una tarea muy difícil, por no decir imposible. Al codificar los datos mediante el uso de elementos gráficos posicionados en relación con unos ejes y requerir una interpretación adicional (más allá de la lectura simple) no podemos alcanzar esa precisión. De nuevo, si nos centramos en la gráfica recogida en la figura 5, es imposible poder saber cuál es el porcentaje exacto de las mujeres fumadoras en Italia según esa visualización. Al no poder leerlo directamente y tener que interpretarlo confrontando visualmente la columna correspondiente con el eje Y, lo único que podemos concluir es que el valor numérico de ese porcentaje está alrededor del 27%, pero en ningún caso que es del 26,9876% (un número entero con cuatro decimales).

Es importante remarcar que la única opción para alcanzar el mismo nivel de precisión que en las tablas es incluir en la gráfica (o completarla con) una etiqueta alfanumérica que exprese ese valor con la finura deseada. Este tipo de estrategias visuales se suelen implementar en las gráficas interactivas que podemos encontrar

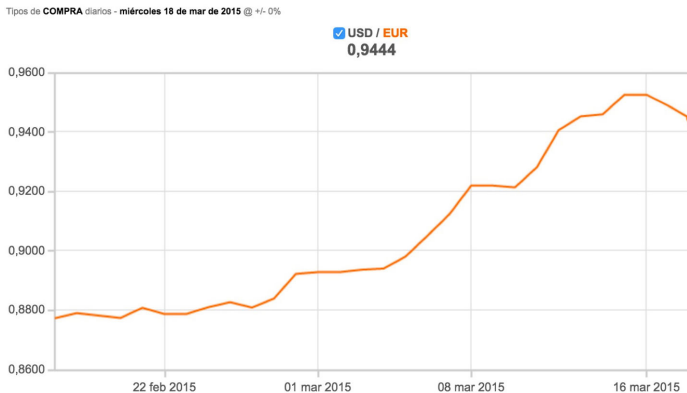
en internet. En esas gráficas, y utilizando la estrategia del *mouseover* (pasada del ratón), podemos descubrir el valor cuantitativo exacto (con todo tipo de precisión) asociado al elemento gráfico que lo está codificando. Un ejemplo de este tipo de estrategias visuales lo podemos observar en la figura 6. En la gráfica que allí aparece, podemos conocer el cambio exacto del Dólar/Euro el día 18 de marzo de 2015 (0,944) si colocamos el ratón sobre el último punto de la línea que conforma la gráfica y que se corresponde con ese día.

La última ventaja competitiva de las tablas frente a las gráficas que queremos destacar se centra sobre un tipo especial de información representada. Concretamente, las tablas son muy eficientes cuando lo que se persigue es comunicar información en la que aparecen distintas unidades de medida. En esas situaciones, las tablas nos permiten codificar de forma cómoda y sencilla múltiples conjuntos de valores cuantitativos expresándolos en diferentes unidades de medida.

Si centramos ahora nuestra atención en la tabla que recogemos en la figura 7, observamos que ésta es capaz de representar diferentes valores cuantitativos expresados en diferentes escalas de medida. Así, podemos representar simultáneamente el número de ítems vendidos de los cuatro productos que fabrica la empresa ACME, los diferentes precios en euros de cada uno de esos ítems y el porcentaje del aumento de ventas de esos ítems respecto al ejercicio anterior. O, dicho de otra manera: ítems, euros y porcentajes conviven sin tensión en una tabla que se muestra eficiente en la comunicación simultánea de estos datos tan dispares.

Frente a este poder comunicacional de las tablas, las gráficas (al menos las clásicas) no puede competir. En conjuntos de datos que incorporan valores que pre-

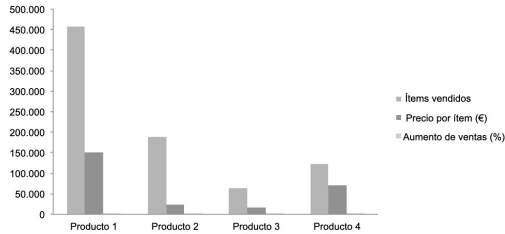
Figura 6. Evolución del cambio del Dólar/Euro



(<https://www1.oanda.com/lang/es/currency/convert/>)

Figura 7. Tabla y gráfica sobre los ítems vendidos, precio por ítem y aumento de las ventas de los cuatro productos fabricados por la empresa ACME

Producto	Ítems vendidos	Precio por ítem (€)	Aumento de ventas (%)
Producto 1	456.876	150.678	40,3
Producto 2	189.387	22.864	90,2
Producto 3	63.829	15.547	20,5
Producto 4	122.345	70.876	15,6



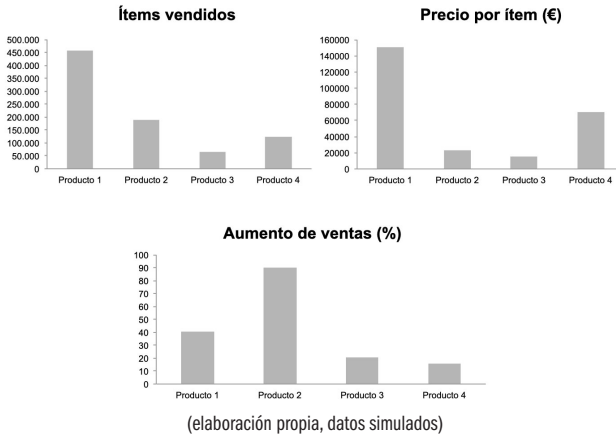
(elaboración propia, datos simulados)

sentan un rango muy amplio por incluir unidades diferentes, las gráficas presentan la dificultad de representar con finura frente a un único eje valores tan dispares. Esta limitación se puede ilustrar si revisamos la gráfica recogida en la figura 7. La gráfica debe representar frente al eje Y valores muy bajos (los que se corresponden con el porcentaje del aumento de las ventas, que van de 15,6 % a 90,2%) junto a otros muy elevados (como el de los ítems vendidos, con un valor máximo de 456.876 respecto al producto 1). Y esto provoca que escalemos ese eje para poder representar con columnas los valores de la zona alta del rango (los ítems vendidos). Pero al hacer esto, las columnas que representan el porcentaje del aumento de las ventas, al incluir valores tan bajos, no se pueden apreciar visualmente, fallando en la comunicación de estos porcentajes.

Para solventar este problema comunicativo sin renunciar a la utilización de gráficas, nos vemos obligados, entre otras alternativas, a plantear tres (y no una única, como en el caso de la tabla) visualizaciones distintas. Como recogemos en la figura 8, tendríamos que desagregar el conjunto original con ese rango tan amplio y construir tres gráficas independientes para representar respectivamente cada uno de los tres conjuntos de valores (ítems vendidos, precio por ítem y porcentaje del aumento de las ventas) de manera que sus respectivos ejes Y se escalen en función de los rangos de esos conjuntos de datos. Con esta alternativa más costosa en tinta y esfuerzo conseguimos que los ítems vendidos, el precio por ítem y el porcentaje del aumento de las ventas de los cuatro productos pudieran comunicarse de forma efectiva.

Pero no todo son desventajas o limitaciones cuando utilizamos gráficas (y no tablas) para visualizar información. Está claro, como hemos visto, que no son una

Figura 8. Ítems vendidos, precio por ítem y aumento de las ventas de los cuatro productos fabricados por la empresa ACME



buena opción porque no podemos recurrir a la estructura de filas/columnas para la búsqueda de valores concretos, porque no podemos utilizar sólo nuestra competencia lectora para aprehender la información que visualizan, porque no permiten expresar los valores numéricos con el grado de precisión que deseemos o porque no podemos representar un conjunto de datos que incluya valores expresados en distintas unidades de medida y que presenta un rango muy amplio.

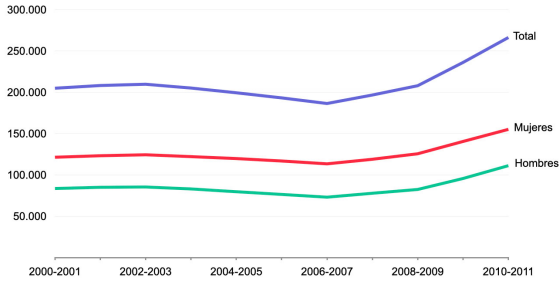
Sin embargo, en ciertas circunstancias comunicativas, una gráfica puede ser la mejor opción para visualizar una información determinada. Gracias a que las gráficas codifican valores cuantitativos de forma visual (y no alfanumérica, como ocurre en el caso de una tabla), éstas nos representan mucho más que una colección de datos o de valores. Nos ofrecen una figura o imagen global de los datos representados que nos permiten identificar patrones, anomalías y tendencias dentro de esos datos, y, simultáneamente, observar los valores atípicos y la relación que éstos mantienen con el resto de los datos de esa colección. Las tablas, al estar basadas en otro tipo de codificación más textual, se encuentran conceptualmente limitadas para ofrecernos esa capacidad representativa global que nos muestran las gráficas.

Esta ventaja representacional de las gráficas frente a las tablas la podemos ilustrar a partir de las dos visualizaciones incluidas en la figura 9. Estas dos visualizaciones (una tabla y una gráfica) se generan a partir de una misma colección de datos que incluye el número total de graduados (desagregado también por sexos) en el sistema universitario español durante la década de los años 2000.

Si nos centramos en la gráfica de líneas que aparece en la figura, y gracias a la codificación visual de los datos, podemos obtener de forma cómoda una visión

Figura 9. Tabla y gráfica que representan el número de graduados por el sistema universitario español durante la primera década del siglo XXI en España

	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011
Hombres	83.515	85.017	85.368	83.008	79.633	76.383	73.036	77.764	82.374	95.645	111.195
Mujeres	121.389	123.199	124.353	122.134	119.768	116.838	113.395	118.920	125.593	140.493	155.136
Total	204.904	208.216	209.721	205.142	199.401	193.221	186.431	196.704	207.967	236.138	266.331



(elaboración propia, datos extraídos de INE)

global de los datos donde destacar algunos patrones como, entre otros, que el número de mujeres graduadas es siempre superior al de hombres a lo largo de toda esa década o que el número total de graduados fue bajando del curso 2002-2003 al 2006-2007, pero que a partir de ese año volvió a crecer. También podemos observar con claridad la tendencia global de crecimiento positivo (en forma de palo de jockey) que nos muestran esos datos en conjunto, aunque en un periodo intermedio el número de graduados descienda parcialmente. Y, sólo por señalar un ejemplo más, nos permite la comparación entre los conjuntos de datos de los hombres y las mujeres identificando las coincidencias, diferencias y excepciones entre esos dos grupos de población.

Frente a este poder comunicacional de las gráficas, las tablas no pueden competir. Al ser interpretados de forma unitaria (y no en su conjunto) cada uno de sus valores codificados en la tabla, sin una memoria prodigiosa que retuviese en nuestra cabeza de forma vívida y simultánea todos los datos (numéricos y textuales) albergados en sus 48 celdas, no podríamos identificar mentalmente todos esos patrones y contenidos informativos descritos en el párrafo anterior y que de forma tan cómoda podemos extraer al visionar simplemente la gráfica.

En cierta manera, a modo de metáfora y como conclusión, podemos ver esta ventaja de las gráficas frente a las tablas de la siguiente manera. Imaginemos que un señor acostumbra a sacar su perro a pasear cada día a las once de la noche. Para ello utiliza una de esas cadenas extensibles que, aunque pueden ser bloqueadas en cualquier momento por el dueño, le ofrecen al animal cierto margen de maniobra a lo largo del paseo. El paseo dura 45 minutos. El perro, en su collar, lleva implantada

una luz amarilla que brilla sin cesar a lo largo de todo el paseo. El señor lleva también una luz, en este caso roja, encima de la gorra que se pone cada vez que saca el perro a esas horas de noche cerrada. Supongamos que esta noche salen a pasear y grabamos el trayecto de forma cenital utilizando un dron que vuela a 20 metros de altura. De forma simultánea, recogemos también las coordenadas (latitudes y longitudes) de cada uno los dos transeúntes (humano y canino) mediante sendos GPS.

En este contexto imaginario, la tabla sería algo así como la recolección sistemática de todas esas coordenadas geográficas generadas por el señor y su perro (recogidas por el GPS) y presentadas de forma estructurada mediante una cuadrícula con filas y columnas. Con esta tabla, podríamos conocer de forma muy precisa la posición exacta del señor y su mascota en un momento de tiempo determinado. Y también podríamos comparar esas coordenadas con las que presentan en otro momento de tiempo distinto. Pero, difícilmente, a no ser que seamos unos superdotados y seamos capaces de retener simultáneamente en nuestras cabezas cientos de esas coordenadas generadas, podremos intuir cómo se ha desarrollado el paseo de nuestros dos protagonistas.

En ese mismo contexto, la gráfica sería como la grabación cenital del dron. Esa gráfica estaría formada por dos líneas (una roja y otra amarilla) ubicadas en dos ejes que le dan significación geográfica y temporal a cada uno de los puntos que las conforman. Por un lado, esa gráfica nos permitiría ver los movimientos en zigzag del perro (de la luz amarilla), sus fluctuaciones momentáneas. Pero, por otro, si detenemos nuestra atención en la línea que recoge la luz roja (la del dueño del animal) podríamos observar el itinerario completo del trayecto, la tendencia de ese paseo.

4. Tipología de gráficas

Hasta el momento, en este capítulo, hemos abordado la caracterización de las gráficas, incluyendo su definición, sus características básicas, sus componentes estructurales (los elementos codificadores de información y los elementos de soporte) y las principales diferencias básicas respecto a su potencial de codificación existentes entre los dos principales productos de visualización (las gráficas y las tablas).

En este último apartado, finalizaremos el capítulo abordando la tipología de las principales gráficas que solemos utilizar con intención comunicativa. Existe un conjunto extenso de propuestas visuales para la representación de la información (Shneiderman, 1996), pero de forma concreta, en este trabajo abordaremos la gráfica de barras o columnas, la gráfica de barras apiladas, la gráfica de líneas, la gráfica de combinación de barras (o columnas) y líneas, la gráfica de áreas, la gráfica circular y el *treemap* (o mapa de árbol).

Para cada una de esas gráficas, introduciremos una definición, analizaremos su anatomía, la ilustraremos con ejemplos y adelantaremos algunos de los usos que podemos darle a cada uno de esos tipos.

4.1. Gráfica de barras o de columnas

Comencemos por una de las propuestas visuales más utilizadas cuando tratamos de representar información: la gráfica de barras (o, su variante, la gráfica de columnas).

Desde un punto de vista visual, este tipo de gráficas se caracteriza por presentar una serie de formas o superficies rectangulares dentro de un espacio generado por dos ejes perpendiculares. Cuando esas formas rectangulares se presentan de manera horizontal (el lado más largo del rectángulo coincide con la anchura de la superficie) en ese espacio, la propuesta visual resultante recibe el nombre de “gráfica de barras”. En cambio, cuando esas superficies rectangulares se presentan de forma vertical (el lado más largo del rectángulo coincide con la altura de la superficie) en ese espacio, la propuesta visual resultante recibe el nombre de “gráfica de columnas”.

Se acostumbra a utilizar para mostrar y comparar información cuantitativa asociada a ítems cualitativos (o a valores de variables discretas). Dentro de ese contexto, son justo esas formas rectangulares las que representan simultáneamente (de forma bifronte) los ítems cualitativos y los valores cuantitativos o numéricos con los cuales se asocian. Desde un punto de vista anatómico, la posición de la forma rectangular en el plano generado por los dos ejes determina el valor o ítem cualitativo que representa. En cambio, la longitud de esa forma rectangular es la que representa el valor numérico asociado.

Respecto a la precisión, cada uno de los ejes que generan el espacio donde se ubican las formas rectangulares otorgan la clave para interpretar el significado cualitativo y numérico asociado a cada una de las formas rectangulares. En el caso de la gráfica de barras, es el eje Y el que suministra los ítems cualitativos asociados a esas formas rectangulares y es el eje X el que nos ofrece valor numérico asociado a éstas. Sin embargo, en el caso de la gráfica de columnas la ecuación es la contraria: es el eje X el que suministra los ítems cualitativos asociados a esas formas rectangulares y es el eje Y el que nos ofrece valor numérico asociado a éstas.

Si articulamos todas estas características que acabamos de destacar de una forma resumida, podemos ofrecer una versión condensada de la definición de gráfica de barras (o de columnas).

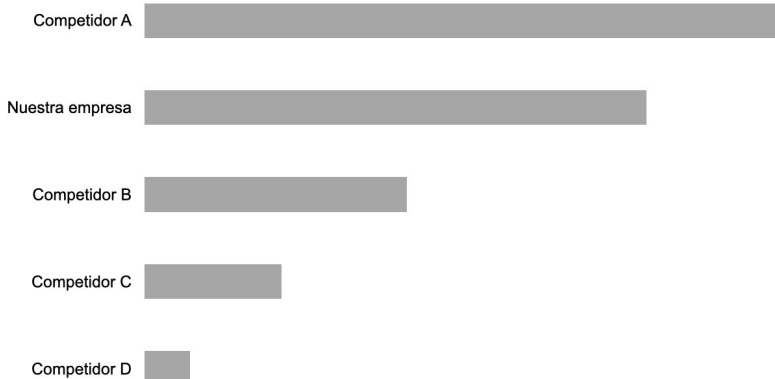
- c) *Definición de gráfica de barras (o de columnas):* instrumento para visualizar información cuantitativa que presenta una serie de características. Por un

lado, los valores cuantitativos asociados a ítems cualitativos se codifican simultáneamente como rectángulos dispuestos de forma horizontal (gráfica de barras) o vertical (gráfica de columnas). Por otro, esos valores se representan dentro de un área delimitada por dos ejes que nos ayudan a ubicar localizaciones espaciales en un plano bidimensional. Y, por último, los ejes proporcionan escalas (cuantitativas y cualitativas) que se utilizan para asignar valores y etiquetas a esas formas rectangulares.

Pasemos ahora a ilustrar estas características a través de un par de sencillos ejemplos. Si centramos nuestra atención en la figura 10, allí podemos identificar una típica gráfica de barras. Ésta presenta cinco formas o superficies rectangulares dentro de un espacio generado por dos ejes perpendiculares. Se trata de una gráfica de barras (y no de columnas) porque esas formas rectangulares se presentan de manera horizontal en ese espacio.

Como su título indica, la estamos utilizando para representar las ventas anuales del ejercicio anterior (información cuantitativa) que han tenido las firmas que forman parte de la competencia sectorial de nuestra empresa (ítems cualitativos). Esa representación nos permite mostrar las ventas de cada una de esas empresas y compararlas con las del resto. Esos rectángulos grises codifican simultáneamente los ítems cualitativos (las empresas que forman parte del sector) y los valores cuantitativos o numéricos con los cuales se asocian (las ventas anuales del ejercicio anterior). La posición de cada rectángulo en el plano nos indica cuál es la empresa representada (nuestra empresa, el competidor A, el competidor B, el competidor C

Figura 10. Ventas anuales en el ejercicio anterior de la competencia sectorial de nuestra empresa



(elaboración propia, datos simulados)

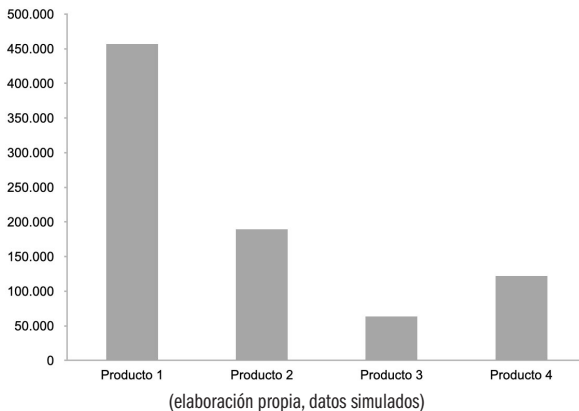
o el competidor D). La longitud de cada rectángulo representa las ventas de cada una de las empresas. El eje Y, con sus etiquetas, nos permite averiguar con qué empresa se corresponde cada uno de los rectángulos grises. El eje X, en cambio, nos ofrece el valor de las ventas anuales del año anterior de esas empresas.

Por otro lado, si centramos ahora nuestra atención en la figura 11, allí podemos observar una típica gráfica de columnas. Ésta presenta cinco formas o superficies rectangulares dentro de un espacio generado por dos ejes perpendiculares. Se trata de una gráfica de columnas (y no de barras) porque esas formas rectangulares se presentan de manera vertical en ese espacio.

Como su título indica, la estamos utilizando para representar las unidades vendidas en el ejercicio anterior (información cuantitativa) de los cuatro productos que fabrica la empresa ACME (ítems cualitativos). Esa representación nos permite mostrar las unidades vendidas de cada producto y compararlas con las de los otros tres productos. Esos rectángulos grises codifican simultáneamente los ítems cualitativos (los cuatro productos) y los valores cuantitativos o numéricos con los cuales se asocian (las unidades vendidas). La posición de cada rectángulo en el plano nos indica cuál es el producto representado (producto 1, producto 2, producto 3 y producto 4). La longitud de cada rectángulo representa las unidades vendidas de cada uno de esos productos. El eje X, con sus etiquetas, nos permite averiguar con qué producto se corresponde cada uno de los rectángulos grises. El eje Y, en cambio, nos ofrece el valor de las unidades vendidas de cada uno de los productos.

Respecto a este tipo de gráficas, es importante señalar un par de aspectos importantes: los criterios para elegir una gráfica de barras o una gráfica de columnas y el tipo de variables discretas que estas gráficas suelen representar.

Figura 11. Unidades vendidas en el ejercicio anterior de los cuatro productos que fabrica la empresa ACME



Comencemos por los criterios para elegir una gráfica de barras o una gráfica de columnas. Como hemos visto, estas dos variantes sirven para representar valores cuantitativos asociados a ítems cualitativos. El único rasgo que las diferencia es que una, la de barras, utiliza rectángulos en posición horizontal y la otra, la de columnas, los coloca en posición vertical. Frente a esta distinción nos asalta una duda: si son casi iguales ¿cuándo utilizaremos cada una de esas variantes? Más adelante abordaremos otros criterios complementarios para dar respuesta a esta pregunta. Pero, ahora, de momento, podemos identificar un criterio claro que nos va a permitir elegir cuál de las dos variantes es más adecuado utilizar.

Concretamente, cuando estamos utilizando una gráfica de columnas y las etiquetas de los ítems o categorías representadas son tan extensas (en el número de caracteres) que no caben en el eje de X de forma que el usuario pueda leerlas de una forma natural y nos veamos obligados a ubicarlas en ese eje con algún tipo de inclinación, utilizaremos mejor una gráfica de barras (y no de columnas) para garantizar la comunicación eficiente de los datos.

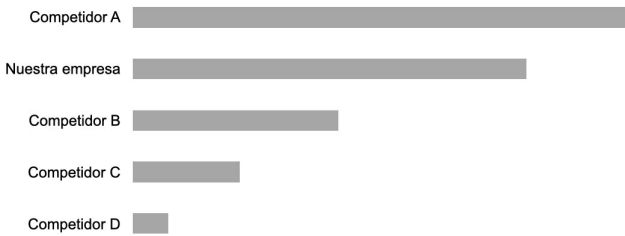
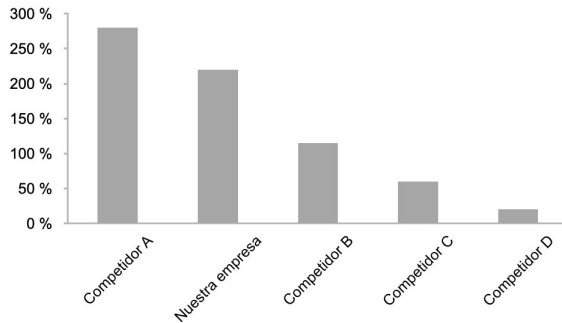
Veámoslo a través de un ejemplo. Imaginemos que queremos representar visualmente las ventas anuales del ejercicio anterior que han tenido las firmas que forman parte de la competencia sectorial de nuestra empresa. Y para ello decidimos utilizar una gráfica de columnas. Pero al hacerlo, comprobamos que las etiquetas de las empresas (ítems cualitativos) representadas son tan extensas que no caben en el eje de X y nos vemos obligados a ubicarlas en ese eje con una inclinación u orientación en diagonal (figura 12). Pero al ubicarlas de esta manera, provocamos que el usuario de esa gráfica no pueda leerlas de una forma natural y tenga que girar su cabeza para poder interpretar esas etiquetas. En esas situaciones, es mejor apostar por una gráfica de barras (como la que mostramos, abajo, en la misma figura 12). Como podemos comprobar, al colocar en este tipo de gráficas las etiquetas de los ítems categoriales en el eje Y (y no en el X como en la de columnas), podemos reducir la región de los datos y ubicar esas etiquetas con una orientación horizontal de manera que los usuarios de la gráfica puedan leerlos de una forma natural y cómoda.

Terminemos este apartado abordando el tipo de variables discretas que estas gráficas de barras o columnas suelen representar.

Como hemos ya señalado, este tipo de gráficas se utiliza principalmente para mostrar y comparar información cuantitativa asociada a ítems cualitativos (o a valores de variables discretas).

Abundemos un poco más sobre la tipología de las variables. Tradicionalmente, a groso modo, las variables se distinguen en dos grandes grupos: las variables cuantitativas y las variables cualitativas. Las cuantitativas son variables que miden un evento determinado y que pueden tomar un valor numérico. Como, por ejem-

Figura 12. Gráfica de columnas y de barras para representar las ventas anuales del ejercicio anterior de la competencia sectorial de nuestra empresa



(elaboración propia, datos simulados)

plo, la edad o el número de hijos de una persona. Las cualitativas, son variables que mencionan cualidades de un evento concreto. Como, por ejemplo, el sexo o la nacionalidad de una persona.

A su vez, las cuantitativas se dividen en continuas (cuando pueden tomar cualquier valor dentro de un intervalo) y discretas (cuando no pueden tomar ningún valor entre dos consecutivos). La estatura o el peso de una persona son ejemplos de variables cuantitativas continuas. El número de árboles de un parque o el de hijos de una persona son ejemplos de variables cuantitativas discretas.

En la misma línea, las cualitativas se dividen en nominales (formadas por ítems que conforman un espacio temático común sin un orden o jerarquía intrínseca) y ordinales (formadas por ítems que conforman un espacio temático común y que presentan un orden intrínseco). Los continentes que conforman la superficie terrestre o la nacionalidad de una persona son ejemplos de variables cualitativas nominales. La clasificación de la liga de fútbol española del año pasado o los diferentes niveles educativos de un país concreto (primaria, secundaria, bachillerato y universidad, por ejemplo) son dos muestras de variables cualitativas ordinales.

Aclarado este punto, nosotros no vamos a seguir esta distinción clásica cuantitativo-cualitativo y vamos a convenir en clasificar las variables sólo en dos grandes

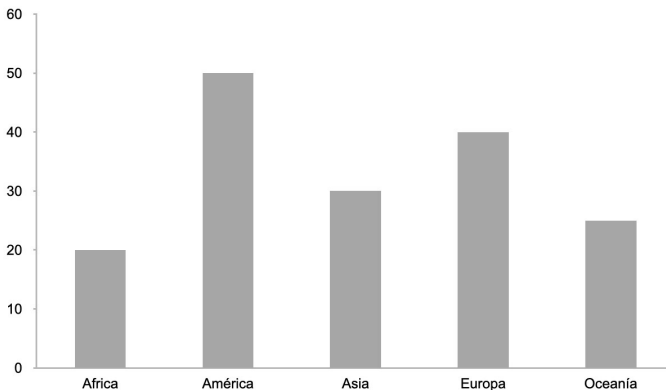
grupos: las variables continuas y las variables discretas. En esta convención, una variable es discreta cuando no puede tomar ningún valor intermedio entre dos valores consecutivos y es continua cuando puede tomar cualquier valor dentro de un intervalo (para cada dos valores cualesquiera, siempre puede existir otro valor intermedio).

Teniendo en cuenta esta distinción propuesta, las gráficas de barras o columnas nos permiten representar de forma adecuada información cuantitativa asociada a tres tipos concretos de variables discretas: nominales, ordinales y de intervalo (Few, 2006). Como veremos más adelante, esa distinción nos permitirá comparar el poder de codificación de este tipo de gráficas frente al de otras alternativas visuales.

Una variable nominal se encuentra formada por ítems discretos que conforman un espacio temático común y que no encierran un orden prefijado que debe ser respetado cuando pasan a ser representados gráficamente. Los diferentes continentes que conforman la superficie terrestre podrían ser considerados como los posibles valores de un ejemplo de este tipo de variables. Las gráficas de barras o columnas funcionan especialmente bien cuando codifican visualmente los valores cuantitativos asociados a este tipo de variables. En la figura 13 podemos ver cómo, con este tipo de gráficos, es posible representar adecuadamente el porcentaje de población que posee un vehículo de transporte propio asociado a cada uno de los continentes terrestres.

Por otro lado, una variable ordinal se encuentra formada por ítems discretos que conforman un espacio temático común, pero que, a diferencia de la nominal, encierran también un orden prefijado que debe ser respetado cuando pasan a ser representados gráficamente. Los equipos ordenados por su clasificación en la liga

Figura 13. Porcentaje de población que posee un vehículo de transporte propio asociado a cada uno de los continentes terrestres



(elaboración propia, datos simulados)

de fútbol española del año pasado podrían ser considerados como los posibles valores de un ejemplo de este tipo de variables. Las gráficas de barras o columnas también funcionan especialmente bien cuando codifican visualmente los valores cuantitativos asociados a este tipo de variables. En la figura 14 es posible ver cómo, con este tipo de gráficos, podemos representar adecuadamente el número de goles recibidos asociado a cada uno de los primeros cinco clasificados de esa competición liguera que se jugó en la temporada anterior.

Por último, en cierta manera, una variable de intervalo puede ser considerada como variable cuantitativa continua reconvertida en variable discreta. Esta reconversión se obtiene al dividir el rango de valores de la escala numérica en categorías que representan series secuenciales de rangos más pequeños (de igual medida o medidas diferentes). Por tanto, se trata de una variable formada por ítems discretos que, como la escala ordinal, encierran un orden prefijado que debe ser respetado cuando vayan a ser codificados; pero que, en este caso, representan además valores cuantitativos. Cada una de las categorías resultantes de dividir la población española por edades (de 0 a 15 años, de 16 a 30, de 31 a 45, de 46 a 60 y de 61 a 75 años, por ejemplo) podrían ser considerados como los posibles valores de un ejemplo de este tipo de variables. Las gráficas de barras o columnas también funcionan especialmente bien cuando codifican visualmente los valores cuantitativos asociados a este tipo de variables. En la figura 15 podemos ver cómo, con una versión especial de este tipo de gráficos (un histograma), es posible representar adecuadamente el número de pacientes por franja de edad en una población de 200 enfermos de epilepsia.

Figura 14. Goles recibidos por cada uno de los primeros cinco clasificados en la liga española de la temporada anterior

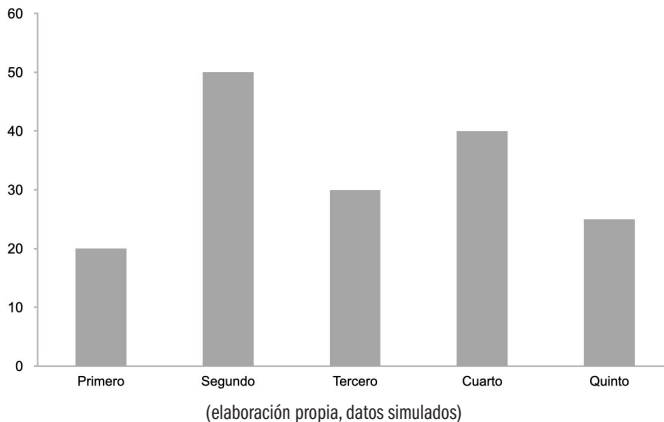
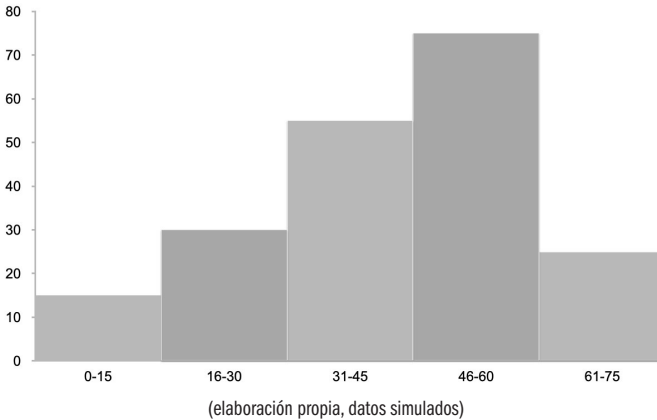


Figura 15. Pacientes por franja de edad en una población de enfermos de epilepsia

Es importante destacar que para estos casos hemos utilizado una variante especial de las gráficas de columnas: el histograma. Anatómicamente hablando, un histograma es una gráfica de columnas en la que esas columnas aparecen una a continuación de la otra. O, dicho en otros términos, entre esas figuras rectangulares no existe un espacio en blanco. La ausencia de ese espacio permite representar la idea de que existe una continuidad entre los intervalos representados por cada una de las columnas.

4.2. Gráfica de barras (o de columnas) apiladas

Pasemos ahora a abordar el segundo tipo de propuestas visuales orientadas a la representación y codificación de información: la gráfica de barras apiladas (o, su variante, la gráfica de columnas apiladas).

Desde un punto de vista visual, este tipo de gráficas (*stacked bar graph*, en inglés) puede ser considerado como una variación de las gráficas de barras o columnas. Se caracterizan por presentar diferentes series de formas o superficies rectangulares apiladas (una encima de otra) dentro de un espacio generado por dos ejes perpendiculares. Cuando esas series constituidas por formas rectangulares se presentan de manera horizontal (el lado más largo del rectángulo formado por la serie coincide con la anchura de esa superficie) en ese espacio la propuesta visual resultante recibe el nombre de “gráfica de barras apiladas”. En cambio, cuando las series constituidas por superficies rectangulares se presentan de forma vertical (el lado más largo del rectángulo formado por la serie coincide con la altura de esa superficie) en ese espacio la propuesta visual resultante recibe el nombre de “gráfica de columnas apiladas”.

Estas gráficas representan mediante una barra o columna (una forma rectangular) un valor (un todo) que se obtiene a partir de la agregación de la representación, en forma de cuadrilátero, de los valores de sus partes. Por tanto, se acostumbra a utilizar para mostrar cómo diferentes categorías más grandes se dividen en categorías más pequeñas y cuál es la relación cuantitativa de cada parte sobre la cantidad total. También se utilizan para comparar numéricamente esas categorías grandes, y también sus partes, entre sí. O, dicho en otros términos, se utilizan para representar y comparar diferentes instancias de un todo y sus partes enfatizando el todo (y no sus partes).

Dentro de ese contexto, son justo esas formas rectangulares (las totales y las partes) las que representan simultáneamente (de forma bifronte) los ítems cualitativos y los valores cuantitativos o numéricos con los cuales se asocian. Desde un punto de vista anatómico, la posición de las formas rectangulares en el plano generado por los dos ejes determina el ítem cualitativo (qué todo y a qué todo pertenece cada parte) que representan. Complementariamente, en el de las partes, el color de las superficies apiladas especifica de forma precisa el ítem cualitativo (qué parte del todo) que representan. En cambio, la longitud de esas formas rectangulares es la que representa el valor numérico asociado.

Igual que en el caso de las gráficas abordadas en el apartado anterior, respecto a la precisión, cada uno de los ejes que generan el espacio donde se ubican esas formas rectangulares otorgan la clave para interpretar el significado cualitativo (y también el color, en el caso de las partes) y numérico asociado a cada una de las formas o superficies. En el caso de la gráfica de barras apiladas, es el eje Y (junto al color, en el caso de las partes) el que suministra los ítems cualitativos asociados a esas formas rectangulares y es el eje X el que nos ofrece valor numérico asociado a éstas. Sin embargo, en el caso de la gráfica de columnas apiladas la ecuación es la contraria: es el eje X (junto al color, en el caso de las partes) el que suministra los ítems cualitativos asociados a esas formas rectangulares y es el eje Y el que nos ofrece valor numérico asociado a éstas.

Si articulamos todas estas características que acabamos de destacar de una forma resumida, podemos ofrecer una versión condensada de la definición de gráfica de barras (o de columnas) apiladas.

- d) *Definición de gráfica de barras (o de columnas) apiladas*: instrumento para visualizar información cuantitativa que presenta una serie de características. Por un lado, los valores cuantitativos asociados a ítems cualitativos (a unidades totales) se codifican simultáneamente como rectángulos dispuestos de forma horizontal (gráfica de barras apiladas) o vertical (gráfica de

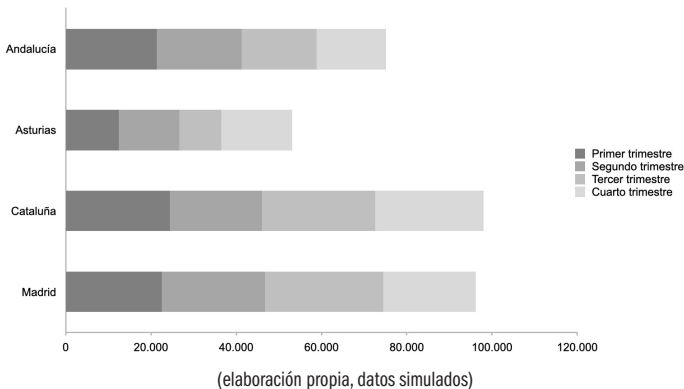
columnas apiladas). Por otro, esas figuras geométricas están formadas a la vez por unidades rectangulares menores apiladas que codifican los valores cuantitativos asociados a las partes de esas unidades totales. Todos esos valores (unidades totales y sus partes) se representan dentro de un área delimitada por dos ejes que nos ayudan a ubicar localizaciones espaciales en un plano bidimensional. Y, por último, los ejes y el color proporcionan escalas (cuantitativas y cualitativas) que se utilizan para asignar valores y etiquetas a todas esas formas rectangulares.

Ilustremos estas características a través de un par de sencillos ejemplos. En la figura 16 podemos identificar una típica gráfica de barras apiladas. Ésta presenta cuatro formas o superficies rectangulares dentro de un espacio generado por dos ejes perpendiculares. Se trata de una gráfica de barras (y no de columnas) porque esas formas rectangulares se presentan de manera horizontal en ese espacio.

Como su título indica, la estamos utilizando para representar las ventas (información cuantitativa) por trimestre (parte o sección de los ítems cualitativos) en las cuatro regiones (ítems cualitativos) donde ACME distribuye sus productos. Esa representación nos permite mostrar las ventas por trimestres en cada una de las regiones y compararlas con las del resto de regiones y trimestres.

Esos rectángulos formados por secciones que presentan diferentes intensidades de color gris codifican simultáneamente los ítems cualitativos (las regiones (rectángulos) y los trimestres (secciones del rectángulo)) y los valores cuantitativos o numéricos (las ventas) con los cuales se asocian respectivamente. La posición de cada rectángulo y sus secciones en el plano nos indica cuál es la región (Asturias,

Figura 16. Ventas por trimestre en las cuatro regiones donde ACME distribuye sus productos



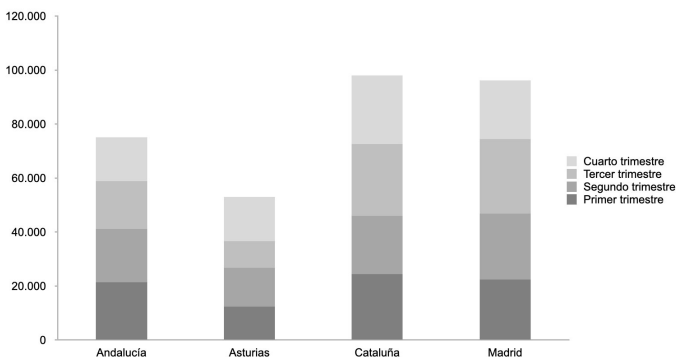
Andalucía, Cataluña o Madrid) y el trimestre representado. La longitud de cada rectángulo agregado y de las secciones que lo conforman representa las ventas en cada una de las regiones y en cada uso de los trimestres, respectivamente. El eje Y, con sus etiquetas, nos permite averiguar con qué región se corresponde cada uno de los rectángulos formado por secciones grises. Como muestra la leyenda, la intensidad del color gris nos indica el trimestre (primer trimestre, segundo trimestre, tercer trimestre o cuarto trimestre). El eje X, en cambio, nos ofrece el valor de las ventas en esas regiones por trimestre.

Por otro lado, en la figura 17, podemos observar la alternativa (generada con los mismos datos) en forma de gráfica de columnas apiladas. Se trata de una gráfica de columnas (y no de barras) porque esas formas rectangulares formadas por secciones de color gris con diferentes intensidades se presentan de manera vertical en ese espacio. Su funcionamiento es idéntico al descrito respecto a la figura 16. Lo único que cambia es que, en este caso, es el eje X, con sus etiquetas, el que nos permite averiguar con qué región se corresponde cada uno de los rectángulos formado por secciones grises. La intensidad del color gris, como muestra la leyenda, también nos indica el trimestre. Y es el eje Y, en cambio, el que nos ofrece el valor de las ventas en esas regiones por trimestre.

Respecto a este tipo de gráficas (de barras o columnas apiladas), es importante señalar un par de aspectos importantes. Por un lado, su tipología de variantes visuales. Y, por otro, sus limitaciones a la hora de representar ciertos tipos de información.

Comencemos abordando la tipología de variantes visuales. Desde una perspectiva estructural es posible discriminar entre dos tipos de gráficos de barras (o columnas) apiladas diferentes: las simples y las porcentuales.

Figura 17. Ventas por trimestre en las cuatro regiones donde ACME distribuye sus productos



(elaboración propia, datos simulados)

Las gráficas de barras (o columnas) apiladas simples colocan cada parte o segmento al lado o a continuación del anterior. El valor total de la barra (o columna) coincide con la agregación de los valores de sus partes en términos absolutos. Las figuras 16 y 17 son buenos ejemplos, respectivamente, de gráfica de barras y columnas apiladas simples.

Los gráficos de barras (o columnas) apiladas porcentuales muestran, en cambio, el valor de la totalidad de cada grupo a partir de la agregación del porcentaje que supone cada una de las partes respecto a la cantidad total del grupo al que pertenecen. En este sentido, el eje respecto al que se referencian los valores cuantitativos (X en la de barras e Y en el de la de columnas) se presenta con una escala porcentual y la longitud de cada barra o columna total se corresponde con el valor del 100%. Teniendo en cuenta esta caracterización, las figuras 18 y 19 son buenos ejemplos, respectivamente, de gráfica de barras y columnas apiladas porcentuales.

Figura 18. Porcentaje de ventas por trimestre en las cuatro regiones donde ACME distribuye sus productos

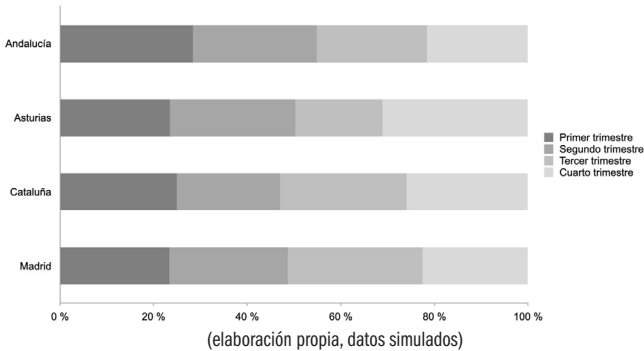
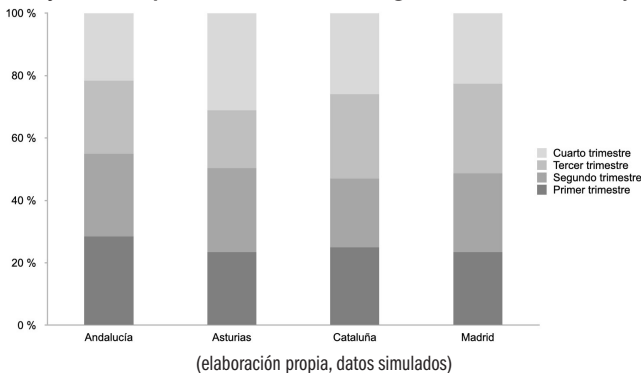


Figura 19. Porcentaje de ventas por trimestre en las cuatro regiones donde ACME distribuye sus productos



Por otro lado, es importante destacar el uso extendido de este tipo de gráficas, tanto las simples como las porcentuales. Ese uso puede venir justificado por la facilidad de integrar paletas de colores en esta clase de visualizaciones. Al ser humano le llaman mucho la atención los colores. Y la posibilidad de integrarlos en los textos a través de las gráficas la identificamos como una buena oportunidad de mejorar estéticamente nuestros aburridos (desde una perspectiva cromática) trabajos en blanco y negro.

Sin embargo, independientemente de su uso intensivo, hay que señalar también que este tipo de propuestas visuales presenta una serie de limitaciones a la hora de representar ciertos tipos de información.

La principal limitación representacional se concentra en el hecho de que éstas son más difíciles de entender y aprehender la información que codifican a medida que aumenta el número de segmentos o partes que conforman esas barras o columnas apiladas. La dificultad aflora al no permitir comparar visualmente de forma cómoda y exacta las partes de un todo entre ellas ni con otras partes de otro todo, excepto el valor de la primera parte o segmento que se encuentra tocando al eje donde se ofrecen las etiquetas que nos permiten identificar los ítems cualitativos.

Veamos esta dificultad a través de un ejemplo. Recuperemos las figuras 16 y 17. En cualquiera de las dos propuestas, podemos comparar cómodamente las ventas anuales en las cuatro regiones. La longitud de las barras o las columnas formada por la agregación de las secciones grises representa adecuadamente esas ventas. Y la alineación de esas figuras agregadas respecto al eje que ofrece los ítems cualitativos con los que se corresponden, permite ver sin problemas cuáles son las diferencias en sus longitudes.

El problema surge cuando intentamos comparar esos sectores entre sí, ya sea con el resto que forman parte de la misma barra o columna o con las secciones o porciones del resto de esas figuras agregadas. Si comparamos entre sí los sectores (sus longitudes) que representan las ventas del primer trimestre de las cuatro regiones, tampoco tenemos grandes problemas. Porque sus bases están perfectamente alineadas respecto a un mismo eje (Y en la gráfica de barras y X en el de las columnas apiladas) y es fácil la comparación visual de sus longitudes. Podemos ver que la región donde más se vende ese trimestre es Cataluña, seguida de Madrid, Andalucía y Asturias.

Pero ¿qué ocurre con el resto de los sectores? ¿Qué ocurre si queremos incluir en nuestra comparación también otros sectores que no codifiquen las ventas del primer semestre? Aquí la cosa ya no es tan sencilla. Por ejemplo, nos cuesta mucho comparar visualmente las ventas en el segundo semestre en Asturias con las del tercer trimestre en Andalucía. ¿Dónde y cuándo se vendió más? Lo mismo ocurre

al comparar las ventas del primer y el tercer trimestre en Cataluña. Nos cuesta mucho comparar visual y mentalmente la longitud de esos sectores. Esa dificultad se justifica por el hecho de que cuando comparamos dos o más sectores que no están alineados sobre un mismo eje (sobre una misma línea) no podemos evaluar adecuadamente sus diferencias de longitud y, por tanto, no podemos extraer conclusiones claras sobre el valor cuantitativo que están representando. En esos casos, en definitiva, cada sector comienza en un punto diferente (el del eje donde se representan los ítems cualitativos (para los del primer trimestre) o el generado por la agregación de longitud de los sectores que le preceden (para el resto de los trimestres)), lo que dificulta la comparación exacta de sus longitudes.

Una posible alternativa para superar esta dificultad sería generar una gráfica de columnas múltiples representando cada una de las partes y la entidad completa (el todo) en forma de barras o columnas. En esa figura alternativa, todos esos elementos visuales codificadores de información estarían alineados en sus bases sobre un mismo eje y, por tanto, podríamos comparar cómodamente la longitud cada una de las secciones y de los todos sin los problemas que nos presentaba la gráfica apilada.

Las figuras 20 y 21 podrían ser consideradas como una propuesta alternativa en esta línea. Respectivamente, se trata de una gráfica de barras (figura 20) y de columnas múltiples (figura 21) donde las ventas de cada trimestre y el total de ventas en una región se representan mediante unos rectángulos alineados por la base sobre el eje que nos codifica con qué región se corresponden cada uno de esos rectángulos.

Esta propuesta alternativa nos ofrece ventajas interesantes. Por un lado, nos permite representar claramente la relación parte-todo, la relación que mantienen cada una de las partes con el todo, con la entidad completa. Siguiendo con los mismos ejemplos (figuras 20 y 21), éstos nos proporcionan la posibilidad de entender

Figura 20. Ventas totales y por trimestre en las cuatro regiones donde ACME distribuye sus productos

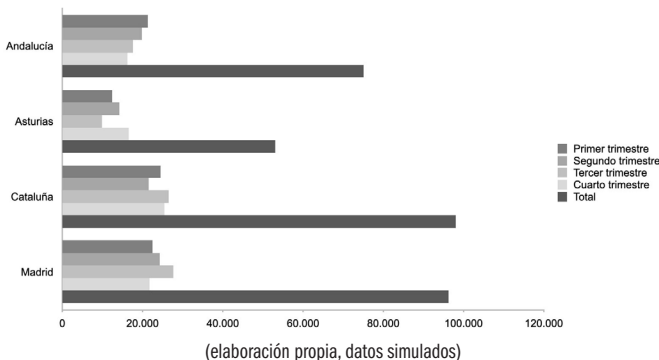
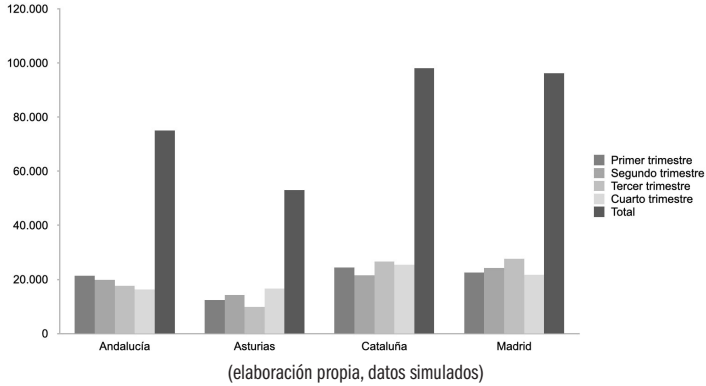


Figura 21. Ventas totales y por trimestre en las cuatro regiones donde ACME distribuye sus productos

cómo las ventas de cada trimestre contribuyen a las ventas totales en cada una de las regiones representadas. Tan sólo tenemos que comparar la longitud de las superficies rectangulares que representan las ventas de cada trimestre con la de la que representa las ventas totales en la región. Esta comparación es más complicada, como ya hemos visto, en el caso de las gráficas de barras o columnas apiladas alternativas.

Por otro, nuestra propuesta alternativa, de la misma manera que las apiladas, nos permite también comparar perfectamente los todos entre ellos y las entidades completas entre ellas. En los dos ejemplos, sólo tenemos que comparar la longitud de las barras o columnas que representan las ventas totales en cada región.

Y, por último, pero a diferencia de las gráficas apiladas, nos facilitan la comparación de todas las secciones o partes (pertenecientes a la misma entidad o a otra diferente) entre sí al estar alineadas por la base sobre el mismo eje. En los ejemplos, podemos comparar las ventas por trimestre entre sí confrontando la diferencia respecto a la longitud de las barras (o columnas) que las representan.

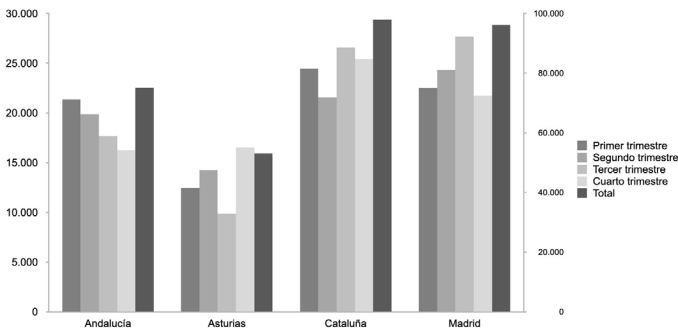
Pero no todo son ventajas en estas propuestas visuales alternativas. También presentan una limitación importante. Cuando las partes tiene un valor muy alto, la diferencia visual entre la longitud de las barras (o columnas) que representan esas partes y las que representan la entidad completa (el todo) es muy grande y, por tanto, las partes se representan con barras (o columnas) con una longitud muy pequeña y esto puede dificultar la comparación entre esas partes. Como ocurre en las figuras 20 y 21, se observa una diferencia importante entre las longitudes de las barras (o columnas) que representan las ventas por trimestre y las que representan las ventas totales. Eso provoca que, al encajar las barras (o columnas) de las ventas totales en el eje que nos suministra las marcas que nos permiten identificar el valor

cuantitativo que representan, la escala resultante no nos permite identificar de forma cómoda los valores cuantitativos asociados a las que representan las ventas por trimestres. Así, podemos identificar de forma sencilla las ventas totales por región, pero no podemos precisar con finura los valores de las barras (o columnas) que representan las ventas por trimestres en esas regiones.

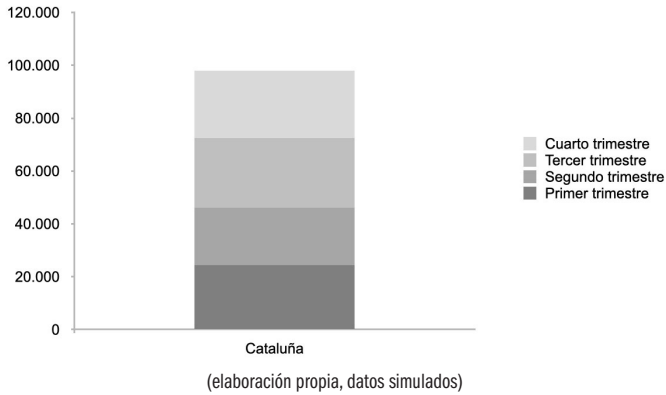
Para solucionar esta limitación sin abandonar la propuesta alternativa, podemos duplicar el eje que nos ofrece la escala cuantitativa (el eje X en el caso de las gráficas de barras múltiples, y el eje Y en el caso de las columnas múltiples). Y al duplicarlo, podemos encajar la barra (o columna) que representan la entidad completa respecto a uno de esos ejes y las barras que representan las partes o secciones respecto al eje duplicado. Al hacer esto reducimos la longitud de las barras (o columnas) que representan las entidades completas (con valores cuantitativos altos) y habilitamos que puedan convivir en una misma distribución visual con las barras (o columnas) que representan las partes de esas entidades (con valores cuantitativos más bajos).

Apliquemos esta solución a la gráfica que recogimos en la figura 21 generando una versión que recogemos en la figura 22. En esta nueva gráfica hemos duplicado el eje Y, de forma que hemos podido encajar respectivamente en esos ejes dos escalas cuantitativas alternativas. Una de esas escalas, la del eje Y (el de la izquierda) nos permite identificar con un alto grado de precisión la longitud de las columnas que representan cada una de las partes de cada todo. Cosa que no ocurría en la figura 21. O, dicho de otra manera, nos permiten identificar las ventas por trimestre realizadas en cada región. La otra escala, la del eje Y' (el de la derecha) nos permite identificar también con un alto grado de precisión la longitud de las columnas que representan cada una de las entidades completas. O, dicho de otra manera, nos permiten identificar las ventas totales realizadas en cada región.

Figura 22. Ventas totales y por trimestre en las cuatro regiones donde ACME distribuye sus productos



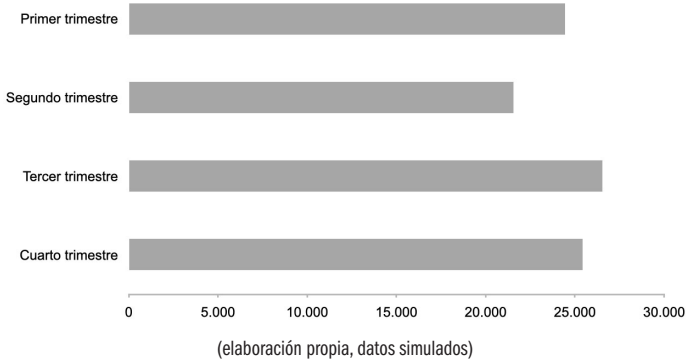
(elaboración propia, datos simulados)

Figura 23. Ventas totales y por trimestre de la empresa ACME en Cataluña

Acabemos este apartado señalando también que en el caso de tener que representar una única instancia de una entidad completa y sus partes, es recomendable no utilizar una gráfica de barras (o columnas) apiladas. Volveríamos a tener el mismo problema al comparar e identificar visualmente con precisión los valores cuantitativos asociados a las secciones que representan las partes.

Esta dificultad la podemos apreciar en la gráfica recogida en la figura 23. Con esta gráfica de columnas apiladas se pretende representar las ventas totales y por trimestre de la empresa ACME en Cataluña. Identificar las ventas totales es sencillo: sólo tenemos que identificar la longitud de la columna formada por la agregación de los sectores rectangulares. En la misma línea, también podríamos identificar fácilmente las ventas del primer trimestre de esa empresa, confrontando contra el eje Y la longitud del primer sector rectangular, el que se encuentra alineado con el eje X. Pero los problemas aparecen cuando abordamos el resto de los trimestres. Por un lado, nos cuesta mucho más identificar con precisión la longitud y, por tanto, los valores de los trimestres (segundo, tercer y cuarto) que representan los sectores correspondientes. Tenemos que identificar dónde comienzan y acaban respecto al eje Y y hacer el cálculo mentalmente para calcular esas longitudes. Y, por otro, y por la misma razón, también es muy incómodo comparar visualmente los valores de los trimestres una vez identificados, ya que tenemos que realizar operaciones cognitivas sin referencias visuales.

Para superar esas limitaciones, como alternativa podemos generar una gráfica de barras simple, como la recogida en la figura 24. En esta gráfica, a diferencia de lo que ocurría en la de columnas apiladas (figura 23), podemos, por un lado, identificar de forma cómoda la longitud de cada una de las barras que representa cada trimestre y, por tanto, concretar con precisión los valores de las ventas en esos

Figura 24. Ventas totales y por trimestre de la empresa ACME en Cataluña

periodos de tiempo. Y, por otro, a partir de esta cómoda identificación, podemos comparar esos valores sin apenas inversión o carga cognitiva.

4.3. Gráfica de líneas

Pasemos ahora a abordar el tercer tipo de propuestas visuales orientadas a representación y codificación de información: la gráfica de líneas. Se trata, junto a la gráfica de barras o columnas, de una de las visualizaciones más utilizadas en ese tipo de representación.

De forma ortodoxa, y desde un punto de vista visual, este tipo de gráficas se caracteriza por presentar una serie de puntos distribuidos en el espacio generado por dos ejes perpendiculares y conectados secuencialmente (de izquierda a derecha) mediante líneas. El resultado es una línea quebrada distribuida en ese espacio y en la que en cada uno de esos vértices de la línea quebrada se ubica un punto. Si bien, ese es su aspecto clásico, en muchas de las versiones de este tipo de visualizaciones se suprimen, por redundantes visualmente, los puntos ubicados en los vértices.

Estas gráficas se acostumbran a utilizar para mostrar cómo una variable cuantitativa evoluciona en un período o intervalo de tiempo continuo concreto. En este sentido, normalmente se generan para mostrar tendencias en la evolución de esa variable cuantitativa o para manifestar las relaciones respecto a la evolución temporal de diferentes variables cuantitativas (representando cada una de esas variables con una línea dentro de la misma gráfica). En definitiva, estas gráficas ayudan a ofrecer una visión global o general de esa evolución durante ese período o intervalo.

Dentro de ese contexto, es justo la línea la que representa simultáneamente los valores cuantitativos o numéricos de la variable y cada uno de los momentos de

tiempo de ese intervalo. Desde un punto de vista anatómico, la posición de cada punto que coincide con los vértices de la línea en el plano generado por los dos ejes determina el valor cuantitativo y el momento de tiempo que representan.

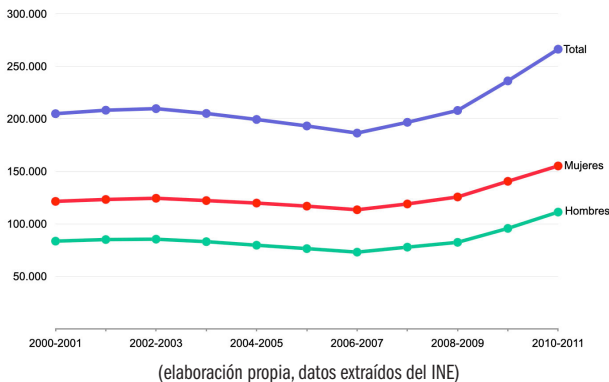
Igual que en el caso de las gráficas abordadas anteriormente, respecto a la precisión, cada uno de los ejes que generan el espacio donde se ubican esos puntos o vértices otorgan la clave para interpretar el significado temporal y numérico asociado a cada una de esas formas. El eje Y suministra los valores cuantitativos asociados a esos puntos o vértices y es el eje X el que nos ofrece el valor temporal asociado a éstos. Es importante señalar que los valores negativos se pueden mostrar por debajo del eje X.

Si articulamos todas estas características que acabamos de destacar de una forma resumida, podemos ofrecer una versión de condensada de la definición de gráfica de líneas.

- e) *Definición de gráfica de líneas:* instrumento para visualizar información cuantitativa que presenta una serie de características. Por un lado, los valores cuantitativos asociados a unidades temporales se codifican simultáneamente como puntos o vértices que forman parte de una línea. Por otro, todos esos valores se representan dentro de un área delimitada por dos ejes que nos ayudan a ubicar localizaciones espaciales en un plano bidimensional. Y, por último, los ejes proporcionan escalas (cuantitativas y temporales) que se utilizan para asignar valores y etiquetas a todos esos puntos o vértices.

Ilustremos estas características a través de un par de ejemplos. En la figura 25 podemos identificar una típica gráfica de líneas, en su modo más ortodoxo, que

Figura 25. Graduados universitarios en la primera década del siglo XXI en España



incorpora puntos en los vértices. Ésta presenta tres líneas dentro de un espacio generado por dos ejes perpendiculares. La gráfica recogida en la figura 26 sería la versión más utilizada, donde desaparecen los puntos de los ejes, pero manteniendo su poder de representación.

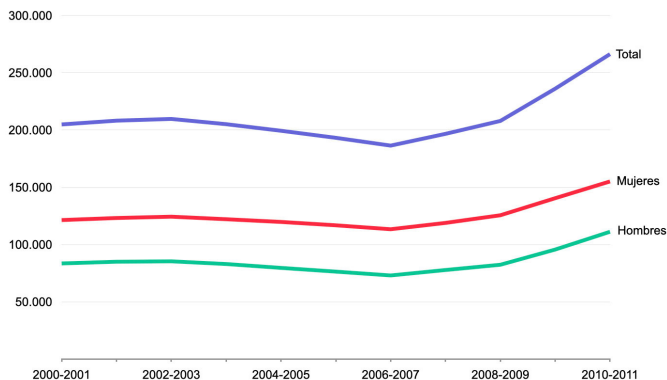
Como sus títulos indican, estamos utilizando ambas gráficas para representar la evolución del número total (y desagregado por sexo) de graduados universitarios (información cuantitativa) durante la primera década del siglo XXI (momentos temporales dentro de un intervalo continuo) en España. Esa representación nos permite mostrar la evolución de esos graduados (totales, hombres y mujeres) a lo largo de esa década y poder comparar esas evoluciones entre sí.

Esos puntos o vértices unidos por líneas codifican simultáneamente los valores cuantitativos o numéricos de la variable (número de graduados (totales, mujeres y hombres)) y cada uno de los momentos de tiempo de ese intervalo (los cursos académicos).

La posición de cada punto que coincide con los vértices de la línea en el plano generado por los dos ejes nos indica el número de graduados universitarios (totales o desagregados por sexo) en ese curso académico representado. El eje Y, con sus etiquetas, nos permite averiguar número de graduados que se corresponde con cada uno de los puntos o vértices. El eje X, en cambio, nos ofrece el curso académico con el que se corresponde ese punto o vértice.

Respecto a este tipo de gráfica, es importante resaltar un par de aspectos: su poder de representación del continuo temporal y su comportamiento frente a otros tipos de gráficas respecto a la codificación de variables discretas.

Figura 26. Graduados universitarios en la primera década del siglo XXI en España



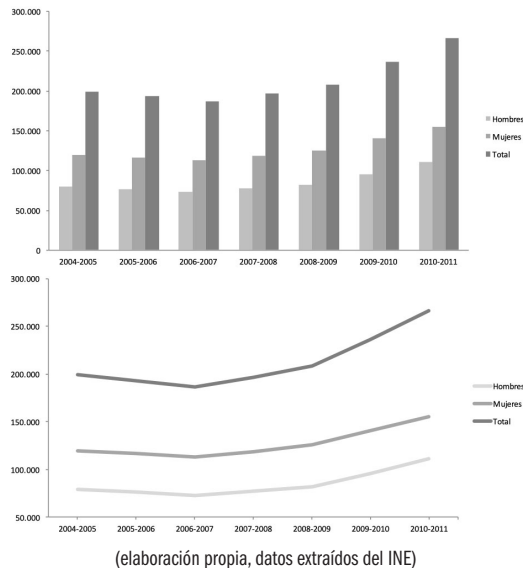
(elaboración propia, datos extraídos del INE)

Comencemos abordando su poder de representación del continuo temporal. En términos generales, es importante señalar que el uso de líneas en visualización lleva implícito la idea de que existen valores intermedios entre los valores conectados (los puntos o vértices) por las líneas. O, dicho de otra manera, representan la existencia de un continuo de valores entre aquellos puntos o vértices conectados por la línea.

Veamos este aspecto a través de un ejemplo. En la figura 27 hemos representado a partir de los mismos datos la evolución de los graduados universitarios españoles durante la primera década de este siglo utilizando dos gráficas distintas (una de columnas y otra de líneas). La gráfica de líneas permite visualizar mejor la evolución en el tiempo de los graduados (mujeres, hombres y totales). La de columnas, en cambio, permite percibir mejor los valores individuales de los graduados en un año concreto. Hay que señalar también que, como se ve en el ejemplo, a diferencia de la de barras, la gráfica de líneas no necesita comenzar su eje de coordenadas en el valor cero o ya que el énfasis no se pone en los valores individuales sino en la tendencia en la evolución.

Veamos, por último, el comportamiento de las gráficas de líneas frente a otros tipos de gráficas respecto a la codificación de variables discretas. Para ello, recordemos primero, como ya vimos anteriormente los tres tipos principales de variables discretas que se suelen representar utilizando gráficas: la variable nominal, la ordinal y la de intervalo.

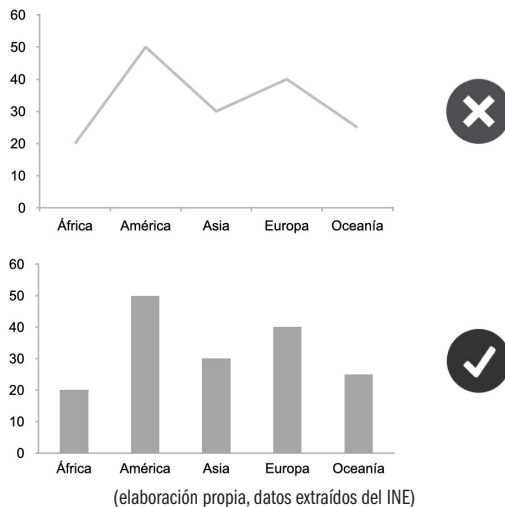
Figura 27. Gráfica de líneas y de columnas para representar el número de graduados universitarios en la primera década del siglo XXI en España



Frente a esta clasificación, la gráfica de barras (o columnas) es la mejor estrategia (frente a la de líneas) para representar valores cuantitativos vinculados a variables nominales u ordinales. El peso visual de las barras permite enfatizar el valor individual de cada categoría o ítem de la variable discreta y facilita la comparación cómoda entre los valores de las categorías confrontando la altura de las barras.

Ilustremos esta situación a través de un ejemplo. En la figura 28 recogemos la representación de unos valores cuantitativos (porcentaje de personas propietarias de un vehículo de motor) asociados a los ítems de una variable nominal (los continentes terrestres) a través de una gráfica de líneas y otra de columnas. Como podemos apreciar, la representación correcta se consigue con la gráfica de columnas. La longitud de las columnas nos permite representar correctamente esos porcentajes de propietarios de vehículos en cada continente. La gráfica de líneas, en cambio, no consigue representar esa información de forma adecuada. La utilización de la línea para codificar la información nos da entender de forma errónea que existe una conexión o continuidad entre el porcentaje de propietarios de África y América, entre los de América y Asia, entre los de Asia y Europa y entre los de Europa y Oceanía. Pero esto no es cierto. El porcentaje de propietarios de vehículos en un continente es independiente de ese mismo porcentaje en otro continente diferente. De hecho, si cambiáramos el orden de los ítems cualitativos dispuestos en el eje X (poniendo el orden justo al revés, comenzando por Oceanía y terminado por África, por ejemplo), estaríamos representando los mismos datos, pero la línea resultante sería totalmente distinta a la presentada en la gráfica de la figura 28.

Figura 28. Gráfica de líneas y de columnas para representar el porcentaje de propietarios de vehículos en cada continente



En la figura 29 ocurre lo mismo, pero con una variable ordinal. En esta figura recogemos la representación de unos valores cuantitativos (goles en contra) asociados a los ítems de una variable ordinal (los cinco equipos ordenados por su clasificación en la liga de fútbol española del año pasado) a través de una gráfica de líneas y otra de columnas. Como podemos apreciar, la representación correcta se consigue de nuevo con la gráfica de columnas. La longitud de las columnas nos permite representar ordenada y correctamente esos goles en contra de cada uno de los cinco equipos. La gráfica de líneas, en cambio, no consigue representar esa información de forma adecuada. La utilización de la línea para codificar la información nos da entender de forma errónea que existe una conexión o continuidad entre los goles en contra del primero y segundo clasificados, entre el segundo y el tercero, entre el tercero y el cuarto y entre el cuarto y el quinto. Pero esto no es cierto. La cantidad de goles recibidos por cada equipo es independiente de la que recibe otro.

En cambio, la gráfica de líneas es la mejor estrategia (frente a la de barras o columnas) para representar valores cuantitativos vinculados a variables de intervalos. En esos casos, la línea permite obtener una visión conjunta y continua de los valores representados y de la serie completa. O, dicho en otros términos, esa línea codifica adecuadamente la continuidad entre esos valores del intervalo.

Para ilustrar este funcionamiento, en la figura 30 recogemos la representación de unos valores cuantitativos (número pacientes) asociados a los ítems de una va-

Figura 29. Gráfica de líneas y de columnas para representar el número de goles recibidos por cada uno de los primeros cinco clasificados la liga española del año pasado

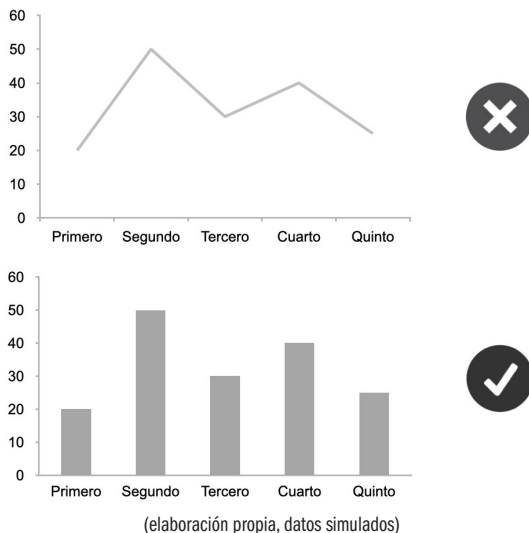
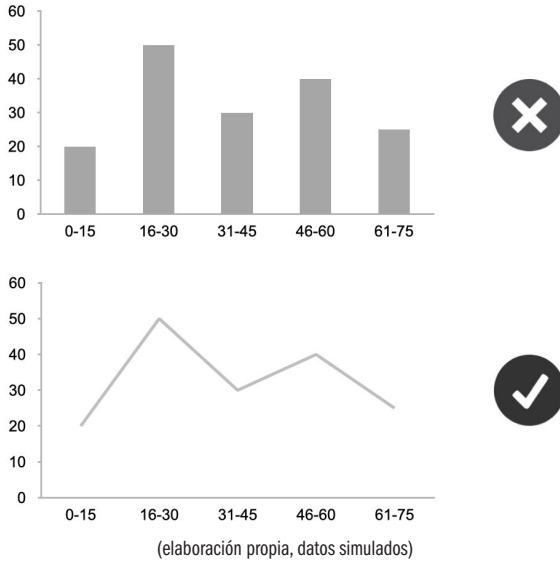


Figura 30. Gráfica de líneas y de columnas para representar el número de pacientes por franja de edad en una población de enfermos de epilepsia

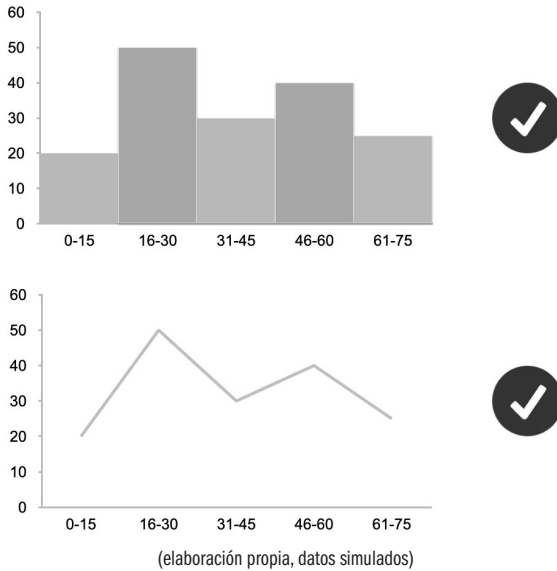


riable de intervalos (las franjas de edad de una población de pacientes que sufren episodios de epilepsia) a través de una gráfica de líneas y otra de columnas. Como podemos apreciar, la representación correcta se consigue con la gráfica de líneas. La posición en el plano de los puntos que conforman la línea y la continuidad visual que imprime esa figura nos permite representar correctamente esa evolución del número de pacientes en función de la franja de edad en la que se encuentre el enfermo. La gráfica de columnas, en cambio, no consigue representar esa información de forma adecuada. La utilización de columnas independientes para codificar la información nos da a entender de forma errónea que no existe conexión o continuidad en la evolución del número de pacientes de una franja de edad a la siguiente.

Terminemos este apartado destacando también que, en algunas situaciones especiales, se pueden utilizar las gráficas de barras o columnas en forma de histograma (junto a las de líneas) para representar variables de intervalo. Concretamente, cuando se quiere enfatizar los valores individuales por encima de la evolución de todos los valores representados.

Podemos ilustrar esta alternativa a través de un ejemplo. En la figura 31 recogemos la representación de los mismos valores cuantitativos (número pacientes) asociados a los ítems de una variable de intervalos (las franjas de edad de la población de unos pacientes que sufren episodios de epilepsia) a través de un histograma y

Figura 31. Histograma y gráfica de líneas para representar el número de pacientes por franja de edad en una población de enfermos de epilepsia



una gráfica de líneas. Como podemos apreciar, la representación correcta se consigue tanto con el histograma como con la gráfica de líneas. La justificación del uso de la gráfica de líneas ya la hemos abordado. La utilización del uso del histograma se justifica porque en este caso las columnas, al no tener espacio que las separe, no son visualmente independientes. De hecho, estas conforman un área continua que nos permite entender de forma correcta que existe una conexión o continuidad en la evolución del número de pacientes de cada franja de edad a la siguiente y nos permite además representar adecuadamente esa evolución en función de la franja de edad en la que se encuentren los pacientes.

4.4. Gráfica de combinación de barras (o columnas) y líneas

Desde un punto de vista visual, las gráficas de combinación de barras (o columnas) y líneas se caracterizan por presentar una serie de formas o superficies rectangulares combinadas con alguna línea (una o más) encajadas dentro de un espacio generado por dos ejes perpendiculares. Cuando esas formas rectangulares se presentan de manera horizontal (el lado más largo del rectángulo coincide con la anchura de la superficie) en ese espacio la propuesta visual resultante recibe el nombre de “gráfica de combinación de barras y líneas”. En cambio, cuando esas superficies rectangulares se presentan de forma vertical (el lado más largo del rec-

tángulo coincide con la altura de la superficie) en ese espacio la propuesta visual resultante recibe el nombre de “gráfica de combinación de columnas y líneas”.

Se acostumbran a utilizar para mostrar y comparar información cuantitativa asociada a ítems cualitativos. Dentro de ese contexto, son justo los vértices (o puntos) de las líneas y esas formas rectangulares las que representan simultáneamente los ítems cualitativos y los valores cuantitativos o numéricos con los cuales se asocian. Desde un punto de vista anatómico, la posición del vértice o punto de la línea y de la forma rectangular en el plano generado por los dos ejes determina el valor o ítem cualitativo que representa. En cambio, la longitud de esa forma rectangular y la posición en el plano del vértice o punto en el caso de la línea son los que representan el valor numérico asociado.

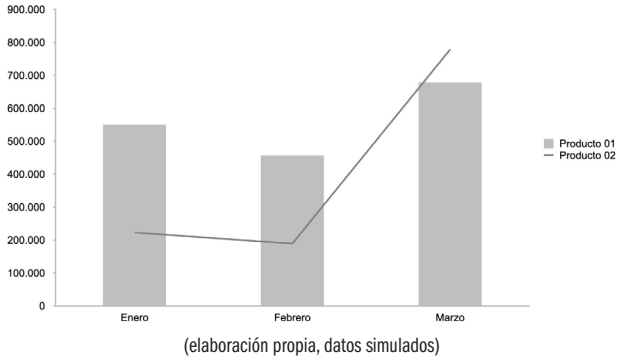
De la misma forma que en el caso de la gráfica de barras (o columnas), respecto a la precisión, cada uno de los ejes que generan el espacio donde se ubican las formas rectangulares y los vértices otorgan la clave para interpretar el significado cualitativo y numérico asociado a cada una de las formas rectangulares. En el caso de la gráfica de combinación de barras y líneas, es el eje Y el que suministra los ítems cualitativos asociados a esos vértices y formas rectangulares y es el eje X el que nos ofrece valor numérico asociado a éstas. Sin embargo, en el caso de la gráfica de combinación de columnas y líneas la distribución es la contraria: es el eje X el que suministra los ítems cualitativos asociados a esas formas rectangulares y vértices y es el eje Y el que nos ofrece valor numérico asociado a éstas.

Si articulamos todas estas características que acabamos de destacar de una forma resumida, podemos ofrecer una versión de condensada de la definición de gráfica de combinación de barras (o de columnas) y líneas.

- f) *Definición de gráfica de combinación de barras (o de columnas) y líneas:* instrumento para visualizar información cuantitativa que presenta una serie de características. Por un lado, los valores cuantitativos asociados a ítems cualitativos se codifican simultáneamente como rectángulos dispuestos de forma horizontal (gráfica de barras) o vertical (gráfica de columnas) y como vértices de una o varias líneas. Por otro, esos valores se representan dentro de un área delimitada por dos ejes que nos ayudan a ubicar localizaciones espaciales en un plano bidimensional. Y, por último, los ejes proporcionan escalas (cuantitativas y cualitativas) que se utilizan para asignar valores y etiquetas a esos vértices y formas rectangulares.

Pasemos ahora a ilustrar estas características a través de un ejemplo. En la figura 32 podemos identificar una típica gráfica de combinación de columnas y líneas.

Figura 32. Ventas mensuales de los dos productos fabricados por ACME en el primer trimestre del año pasado



Ésta presenta tres formas o superficies rectangulares y una línea con dos puntos extremos y un vértice dentro de un espacio generado por dos ejes perpendiculares. Se trata de una gráfica de combinación de líneas y columnas (y no de barras) porque esas formas rectangulares se presentan de manera vertical en ese espacio.

Como su título indica, la estamos utilizando para representar las ventas de los dos productos fabricados por ACME (información cuantitativa) en cada uno de los meses del primer trimestre del año (ítems cualitativos). Esa representación nos permite mostrar las ventas mensuales de cada uno de los productos y compararlas entre sí.

Por un lado, esos tres rectángulos grises codifican simultáneamente los ítems cualitativos (los meses del trimestre) y los valores cuantitativos o numéricos con los cuales se asocian (las ventas mensuales del producto 1). La posición de cada rectángulo en el plano nos indica cuál es mes representado (Enero, Febrero o Marzo). La longitud de cada rectángulo representa las ventas del producto 1 en cada uno de los meses.

Por otro lado, los tres puntos (extremos y vértice) grises de la línea codifican simultáneamente los ítems cualitativos (los meses del trimestre) y los valores cuantitativos o numéricos con los cuales se asocian (las ventas mensuales del producto 2). La posición de cada punto en el plano nos indica cuál es el mes representado (Enero, Febrero o Marzo) y las ventas del producto 2 en cada uno de los meses.

Por último, el eje X, con sus etiquetas, nos permite averiguar con qué mes se corresponde cada uno de los rectángulos y los puntos de la línea gris oscura. El eje Y, en cambio, nos ofrece el valor de las ventas mensuales de cada uno de los productos 1 y 2.

Normalmente, este tipo de gráficas se suelen utilizar para representar tres tipos de información diferente: el resaltado en una comparación nominal, el acumulado de las partes de un todo y un tipo especial de relación parte-todo (Few 2006).

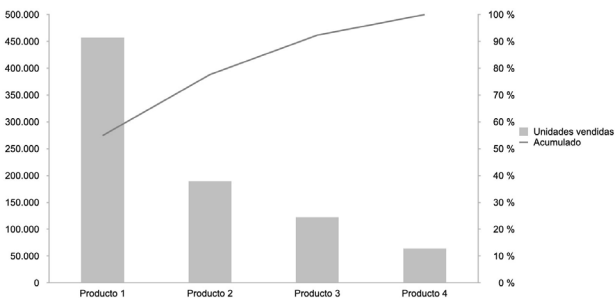
El uso más habitual de este tipo de gráficas se produce cuando queremos resaltar algunos de los elementos que forman parte de una comparación de valores cuantitativos asociados a ítems cualitativos. Retomemos como ejemplo la visualización recogida en la figura 32. En esta gráfica hemos utilizado unas columnas para representar las ventas mensuales del producto 1 en el primer trimestre del año pasado, pero hemos representado la evolución de esas ventas respecto al producto 2 mediante una línea para resaltarlas y confrontarlas con lo representado por las columnas. Utilizamos la combinación columnas y líneas para resaltar cada tipo de información representada por esas figuras encajadas en el plano.

Otro de los usos que se le suele dar a este tipo de gráficas es el de la representación del acumulado de las partes de un todo. Concretamente, se utiliza para codificar y clasificar (de mayor a menor peso) las partes de un todo y para visualizar la contribución del acumulado parcial y total de esas partes respecto al todo.

Veámoslo a través de un ejemplo. Imaginemos que queremos ver cuál es la importancia de cada producto que fabrica ACME respecto al total de las ventas de la empresa y que queremos representar también cómo se acumula, de mayor a menor, esa importancia en esas ventas. Para conseguir ese objetivo, podemos generar una gráfica de Pareto. La gráfica (o diagrama) de Pareto no es nada más que una visualización formada por una gráfica de columnas (donde se representan de mayor a menor las partes de un todo) completada con una línea que codifica la importancia acumulada de esas partes respecto al todo.

En la figura 33 hemos generado una gráfica de Pareto con esa información. Hemos representado mediante columnas las ventas de cada uno de los cuatro productos que fabrica (el eje Y (primario) nos ofrece ese valor cuantitativo absoluto). Y hemos completado esa visualización mediante una línea que va codificando con cada punto la contribución porcentual de las ventas de todos los productos que se

Figura 33. Contribución de los cuatro productos fabricados por ACME a las ventas totales de la empresa



(elaboración propia, datos simulados)

encuentran a la izquierda de ese punto respecto al total de las ventas de la empresa (el eje Y' (secundario) nos ofrece ese valor porcentual). Así, por ejemplo, según el primer punto de la línea, el producto 1 contribuye al 55% de las ventas de ACME. El segundo punto o vértice de esa línea nos indica que el agregado de los productos 1 y 2 supone un 77% de las ventas totales. El tercer punto o vértice nos representa que el agregado de los productos 1, 2 y 3 supone un 92% de las ventas totales. Y, el último punto o vértice de esa línea nos codifica que el agregado de los productos 1, 2, 3 y 4 supone un 100% de las ventas totales de ACME.

Como alternativa a la figura 33, podemos también generar una gráfica donde se obvien los valores absolutos de las partes (las ventas de cada producto) y pongamos exclusivamente el peso o importancia en el agregado porcentual de las ventas respecto al total, escamoteando para ello el eje Y (primario) y manteniendo exclusivamente el eje Y' (gráfica de Pareto recogida en la figura 34).

El último uso que queremos destacar de este tipo de gráficas es el de la representación de un tipo especial de la relación parte-todo. Concretamente, podemos representar las partes de un todo o una entidad en diferentes momentos temporales mediante columnas múltiples y, simultáneamente, mediante una línea podemos codificar la evolución de ese todo o entidad a través del período generado por esos momentos de tiempo.

En la figura 35 recogemos esa posibilidad. Imaginemos que en ACME sólo se fabrican tres productos. Por un lado, podemos ver las ventas mensuales de cada producto presentándolos mediante columnas múltiples asociadas a cada mes del trimestre (Enero, Febrero y Marzo). Y, por otro, mediante una línea queremos representar las ventas totales de la empresa a lo largo de ese trimestre.

Es importante remarcar que, para que las ventas mensuales puedan visualizarse de forma adecuada (y aparezcan representadas todas las cantidades, por peque-

Figura 34. Contribución de los cuatro productos fabricados por ACME a las ventas totales de la empresa

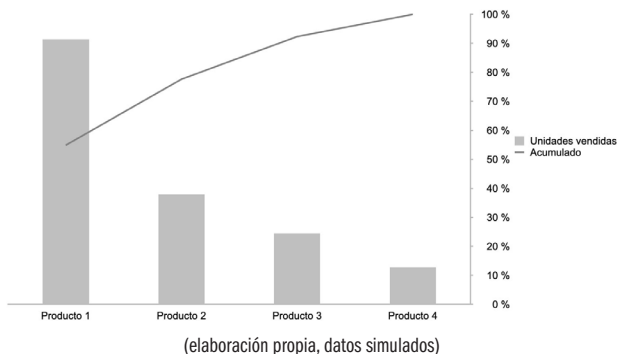


Figura 35. Contribución de los tres productos fabricados por ACME a las ventas totales de la empresa durante el primer trimestre del año

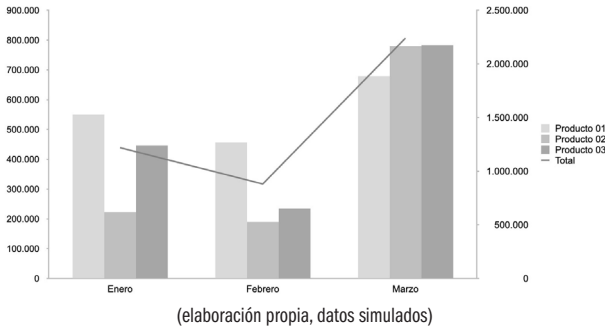
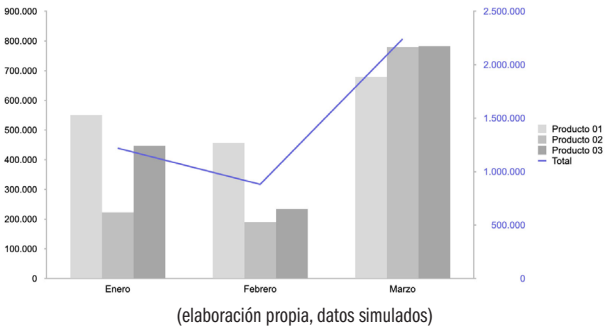


Figura 36. Contribución de los tres productos fabricados por ACME a las ventas totales de la empresa durante el primer trimestre del año



ñas que sea) hemos decidido desdoblar el eje Y en dos. En el eje Y primario hemos ofrecido la escala de las ventas de los tres productos. El eje Y' (o secundario) nos sirve para representar la evolución de las ventas totales. Si no queremos informar al usuario del desdoblamiento de los ejes Y en la leyenda o como nota asociada a la gráfica, podemos generar una representación alternativa donde tanto las etiquetas del eje Y' como la propia línea se hayan dibujado utilizando un mismo color (azul, en la figura 36), pero diferente al resto del registro cromático de la gráfica. De esta manera, el usuario entenderá sin mayores instrucciones que la escala de la línea se corresponde con el eje dibujado utilizando el mismo color (el eje Y').

4.5. Gráfica de áreas

Las gráficas de áreas son propuestas de visualización que se concretan en la disposición de superficies o áreas 2-D encajadas en un doble eje de coordenadas.

Desde un punto de vista visual, este tipo de gráficas se caracteriza por presentar una serie de puntos distribuidos en el espacio generado por dos ejes perpendicu-

lares y conectados secuencialmente (de izquierda a derecha) mediante líneas y por mostrar coloreado el espacio o área generada entre esa línea resultante (que une los puntos) y esos ejes de coordenadas. En cierta manera, en un sentido coloquial, son una versión coloreada de las gráficas de líneas. En este sentido, el resultado es una línea quebrada distribuida en ese espacio (en la que en cada uno de esos vértices de la línea quebrada se ubica un punto) y en la que el área generada por esa línea y los ejes aparece coloreada.

Al igual que las de líneas, estas gráficas se acostumbran a utilizar para mostrar cómo una variable cuantitativa evoluciona en un período o intervalo de tiempo continuo concreto. En este sentido, normalmente se generan para mostrar tendencias en la evolución de esa variable cuantitativa o para manifestar las relaciones respecto a la evolución temporal de diferentes variables cuantitativas (representando cada una de esas variables con una línea que genera un área coloreada determinada dentro de la misma gráfica). En definitiva, estas gráficas ayudan a ofrecer una visión global o general de esa evolución dentro de ese período o intervalo.

De la misma forma que en la gráfica de líneas, es justo la línea la que representa simultáneamente los valores cuantitativos o numéricos de la variable y cada uno de los momentos de tiempo de ese intervalo. Desde un punto de vista anatómico, la posición de cada punto que coincide con los vértices de la línea en el plano generado por los dos ejes determina el valor cuantitativo y el momento de tiempo que representan.

Siguiendo el mismo esquema expositivo, respecto a la precisión, cada uno de los ejes que generan el espacio donde se ubican esos puntos o vértices otorgan la clave para interpretar el significado temporal y numérico asociado a cada una de esas formas. El eje Y suministra los valores cuantitativos asociados a esos puntos o vértices, y es el eje X el que nos ofrece valor temporal asociado a éstos. Es importante señalar que los valores negativos se pueden mostrar por debajo del eje X.

Si articulamos todas estas características que acabamos de destacar de una forma resumida, podemos ofrecer una versión de condensada de la definición de gráfica de áreas.

- g) *Definición de gráfica de áreas:* instrumento para visualizar información cuantitativa que presenta una serie de características. Por un lado, los valores cuantitativos asociados a unidades temporales se codifican simultáneamente como puntos o vértices que forman parte de una línea. Por otro, todos esos valores se representan dentro de un área delimitada por dos ejes que nos ayudan a ubicar localizaciones espaciales en un plano bidimensional. Además, el área o espacio generado por la línea y los ejes aparece coloreado. Y,

por último, los ejes proporcionan escalas (cuantitativas y temporales) que se utilizan para asignar valores y etiquetas a todos esos puntos o vértices.

Ilustremos estas características a través de un ejemplo. En la figura 37 podemos identificar una típica gráfica de áreas. Como su título indica, estamos utilizando la gráfica para representar la evolución del número total (y desagregado por sexo) de graduados universitarios (información cuantitativa) durante la primera década del siglo XXI (momentos temporales dentro de un intervalo continuo) en España. Esa representación nos permite mostrar la evolución de esos graduados (totales, hombres y mujeres) a lo largo de esa década y poder comparar esas evoluciones entre sí.

Esos puntos o vértices unidos por líneas codifican simultáneamente los valores cuantitativos o numéricos de la variable (número de graduados (totales, mujeres y hombre)) y cada uno de los momentos de tiempo de ese intervalo (los cursos académicos).

La posición de cada punto que coincide con los vértices de la línea en el plano generado por los dos ejes nos indica el número de graduados universitarios (totales o desagregados por sexo) en ese curso académico representado. El eje Y, con sus etiquetas, nos permite averiguar el número de graduados que se corresponde con cada uno de los puntos o vértices. El eje X, en cambio, nos ofrece el curso académico con el que se corresponde ese punto o vértice.

Es importante resaltar que existen dos tipos diferentes de gráficas de áreas: las simples y las apiladas. Las gráficas de áreas simples se caracterizan por el hecho de que todos los espacios coloreados generados por sus líneas y ejes comienzan en el valor 0 del eje Y. Las gráficas incluidas en las figuras 37 y 38 son dos ejemplos de este tipo. Como podemos observar, en ambas gráficas las áreas generadas por sus líneas y ejes comienzan en el valor 0 del eje Y.

En cambio, en el caso de las gráficas de área apiladas, las áreas generadas por sus líneas y ejes comienzan en los puntos donde acaba el área generada por la serie de datos anterior. La gráfica incluida en la figura 39 es un ejemplo de este segundo tipo. Como podemos observar, en la gráfica las áreas generadas por sus líneas (las líneas que se corresponden con cada producto) comienzan justo donde acaba el área generada por la línea de la serie de datos anterior. La primera área que codifica el producto 1, comienza en el valor 0 del eje Y. En cambio, el área que representa el producto 2 comienza justo donde terminaba el área que se corresponde con el producto 1. De la misma manera, el área que codifica el producto 3 comienza justo donde terminaba el área que se corresponde con el producto 2.

Dentro de esta distinción, es importante señalar que las gráficas de áreas simples presentan una dificultad que no aparece cuando utilizamos las apiladas.

Figura 37. Graduados universitarios en la primera década del siglo XXI en España

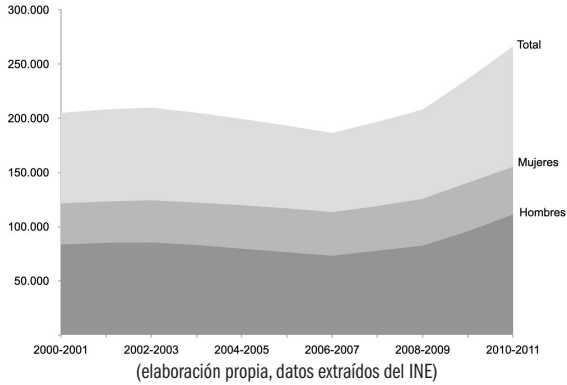
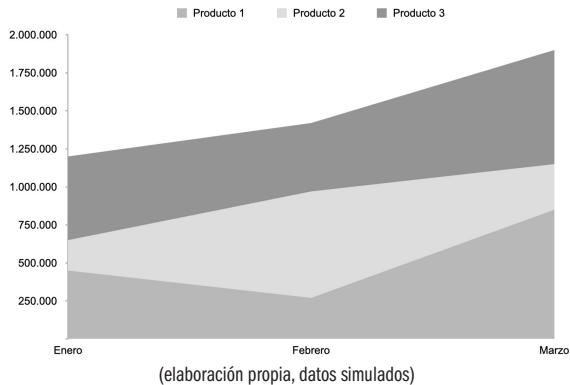


Figura 38. Ventas de los tres productos fabricados por ACME durante el primer trimestre del año



Figura 39. Ventas de los tres productos fabricados por ACME durante el primer trimestre del año



Concretamente, se materializa cuando en esa gráfica se representa más de una serie de valores en el tiempo, algunos de los puntos que generan las líneas que delimitan las áreas pueden quedar escondidos tras otras áreas y no podemos identificar visualmente esos valores ocultos. O, dicho de otra manera, puede producirse un eventual solapamiento visual que impide identificar algunos de los puntos que generan esas áreas no apiladas. Con el ejemplo de la figura 38, podemos ilustrar este problema. En esa gráfica podemos identificar todos los valores asociados al producto 1 a lo largo del trimestre representado. En cambio, respecto al producto 2, sólo podemos leer de forma cómoda el valor correspondiente con el mes de febrero, pero hemos de inferir visualmente (por el solapamiento con el área del producto 1) los valores que se asignan a los meses de enero y marzo. Algo parecido ocurre con el producto 3: podemos leer de forma cómoda el valor correspondiente con el mes de enero, pero hemos de inferir visualmente (por el solapamiento con el área del producto 1 y 2) los valores que se asignan a los meses de febrero y marzo.

4.6. Gráfica circular

Pasemos ahora a abordar uno de los tipos de propuestas visuales orientadas a la representación y codificación de información más utilizados: la gráfica circular.

Desde un punto de vista visual, este tipo de gráficas (ciclograma, gráfica de sectores, gráfica de pastel, gráfica de torta, gráfica de quesitos o *pie chart* (en inglés)) se caracteriza por presentar un área o superficie circular dividida en diferentes sectores por una o varias rectas en forma de radios.

Estas gráficas representan mediante el área circular un valor (un todo) que se obtiene a partir de la agregación de la representación, en forma de sectores, de los valores de sus partes. Por tanto, se acostumbra a utilizar para mostrar cómo una categoría más grande se divide en categorías más pequeñas y cuál es la relación cuantitativa (en términos porcentuales) de cada parte sobre la cantidad total. O, dicho en otros términos, se utilizan para representar y comparar un todo con sus partes.

Dentro de ese contexto, son justo esas formas o superficies que conforman los sectores del círculo (generadas cada una por dos radios y un arco, como si porciones de un pastel se trataran) las que representan simultáneamente los ítems cualitativos y los valores cuantitativos o numéricos con los cuales se asocian. Desde un punto de vista anatómico, el color de esa superficie que conforma cada sector dentro del círculo determina el ítem cualitativo (parte del todo) que representan. En cambio, el ángulo que generan los radios que delimitan junto al arco cada uno de esos sectores es el que representa el valor numérico asociado.

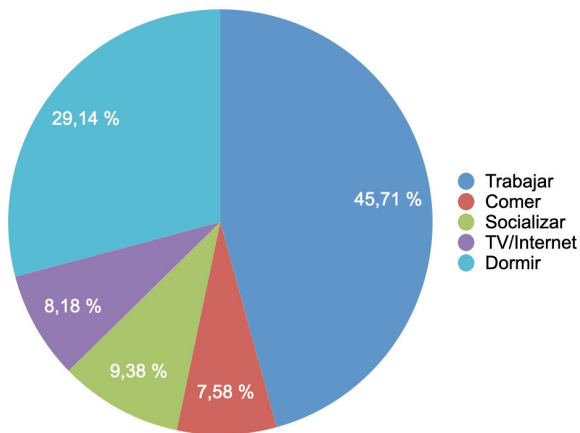
Si articulamos todas estas características que acabamos de destacar de una forma resumida, podemos ofrecer una versión de condensada de la definición de gráfica circular.

- h) *Definición de gráfica circular*: instrumento para visualizar información cuantitativa que presenta una serie de características. Por un lado, el todo (o unidad) se codifica como un área o superficie circular. Por otro, cada una de las partes de ese todo se representa como un sector de esa superficie circular generado por dos radios y un arco. El color de esa superficie que conforma cada sector dentro del círculo determina el ítem cualitativo (parte del todo) que representa. Por último, en cambio, el ángulo que generan los radios que delimitan junto al arco cada uno de esos sectores es el que representa el valor numérico asociado.

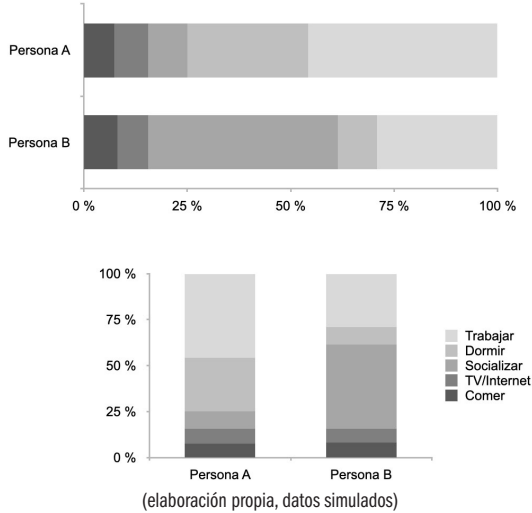
Ilustremos estas características a través de un sencillo ejemplo. En la figura 40 podemos identificar una típica gráfica circular. Ésta presenta un área o superficie circular dividida o formada por cinco sectores generados a su vez por dos radios y un arco. Como su título indica, estamos utilizando la gráfica para representar el peso en tiempo de cada una de las tareas que una persona concreta desarrolla a lo largo de una jornada.

Cada uno de esos sectores o porciones codifica simultáneamente una de las actividades realizadas (trabajar, comer, socializar, ver TV/internet y dormir) a lo largo

Figura 40. Actividades desarrolladas a lo largo de un día



(elaboración propia, datos simulados)

Figura 41. Actividades desarrolladas a lo largo de un día por dos personas diferentes

del día y el valor cuantitativo o numérico asociado a esa actividad (el porcentaje de tiempo del día que le dedicamos a esa actividad). El color de cada sector dentro del círculo nos indica, a partir de la leyenda, la actividad que representa. En cambio, el ángulo que generan los radios que delimitan junto al arco cada uno de esos sectores es el que representa ese porcentaje de tiempo que invertimos en esa actividad.

Es importante resaltar que las gráficas circulares son muy populares y suelen ser incluidas en muchos de los textos donde se aborde un análisis de información cuantitativa y se quiera representar la relación parte-todo. Su popularidad le viene por esa capacidad de representación de la información. Sin embargo, de la misma manera que en el caso de las gráficas de barras o columnas apiladas, ese uso intensivo puede también venir justificado por la facilidad de integrar paletas de colores en esta clase de visualizaciones. Como ya apuntamos, al ser humano nos llaman la atención los colores y la posibilidad de integrarlos en los textos a través de las gráficas la vemos como una buena oportunidad de mejorar estéticamente nuestros aburridos trabajos en blanco (papel) y negro (texto).

De todas formas, independientemente de ese uso destacado, hay que señalar también que este tipo de propuestas visuales no siempre pueden ser utilizadas para representar esa relación parte-todo. Se deben producir una serie de condiciones para que esa función representacional se desarrolle de forma eficiente.

La primera de las condiciones para su uso que queremos señalar es que sólo exista una serie de datos a representar. O, dicho de otra manera, en una misma grá-

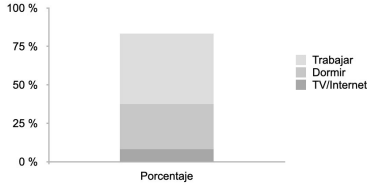
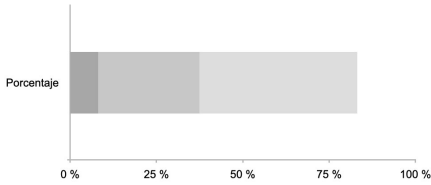
fica circular no podemos representar más de un todo junto a sus partes. Como vemos en la figura 40, podemos representar adecuadamente el porcentaje de tiempo asociado a las tareas que realiza diariamente una persona concreta. Sin embargo, no podemos representar simultáneamente, en la misma gráfica, la distribución del porcentaje de tiempo de las tareas diarias que ha realizado también una segunda persona. En una gráfica de barras o columnas apiladas, por ejemplo, sí que podríamos representar las actividades diarias de esas dos personas A y B (esos dos todos, esas dos series) mediante dos columnas y cada uno de los porcentajes de esas actividades (sus partes) como las superficies rectangulares que forman esas dos columnas (figura 41).

La segunda condición que debe respetarse es que tengamos los datos de todas las partes que conforman el todo. Si nos falta uno o más de esos valores, la gráfica no puede construirse. Nos quedaría un hueco (o varios) en la superficie circular que representa el todo. Si no supiéramos el porcentaje de tiempo dedicado a comer o a socializar, por ejemplo, no podríamos representar mediante el área circular las actividades diarias de una persona. En cambio, en una gráfica de barras o columnas apiladas, por ejemplo, sí que podríamos representar las actividades diarias de una persona mediante una columna (o barra), aunque nos faltase alguno de los porcentajes de esas actividades (comer y socializar, por ejemplo) (figura 42).

Por otro lado, la tercera de las condiciones es que ninguno de los valores de las partes (de la serie) sea negativo. Hablando de forma estricta y si nos centramos en la relación parte-todo, conceptualmente es un sinsentido, ya que ninguna parte de un todo puede tener un valor negativo. Pero lo que queremos poner sobre la mesa es que la gráfica circular no puede representar una serie de datos que incluya algún valor negativo. Es imposible representar un valor negativo a partir del ángulo (siempre es un valor positivo) entre dos radios. Convengamos, por ejemplo, aunque no tenga mucho sentido, que el porcentaje de tiempo dedicado a dormir es negativo. En ese caso, sería imposible representarlo utilizando el ángulo que forman dos líneas (radios). En cambio, por ejemplo, en una gráfica de barras o columnas apiladas, por ejemplo, sí que podríamos representar ese valor, aunque sea negativo (figura 43).

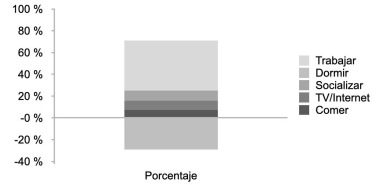
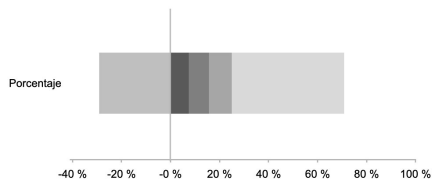
En la misma línea, la cuarta condición que debe respetarse es que ninguno de los datos de las partes sea igual a cero. Es imposible representar el valor o a partir del ángulo (siempre es un valor positivo) entre dos radios. Convengamos, por ejemplo, que el porcentaje de tiempo dedicado a dormir es igual a cero. En ese caso, sería imposible representarlo utilizando el ángulo que forman dos líneas (radios). En cambio, por ejemplo, en una gráfica de barras o columnas apiladas, por ejemplo, sí que podríamos representar ese valor, aunque sea igual a cero. En ese caso, la parte

Figura 42. Actividades desarrolladas a lo largo de un día



(elaboración propia, datos simulados)

Figura 43. Actividades desarrolladas a lo largo de un día

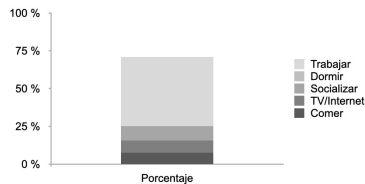
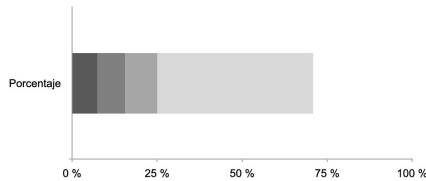


(elaboración propia, datos simulados)

(Dormir) aparece en la leyenda, pero no se le asigna superficie rectangular en la barra o columna (figura 44).

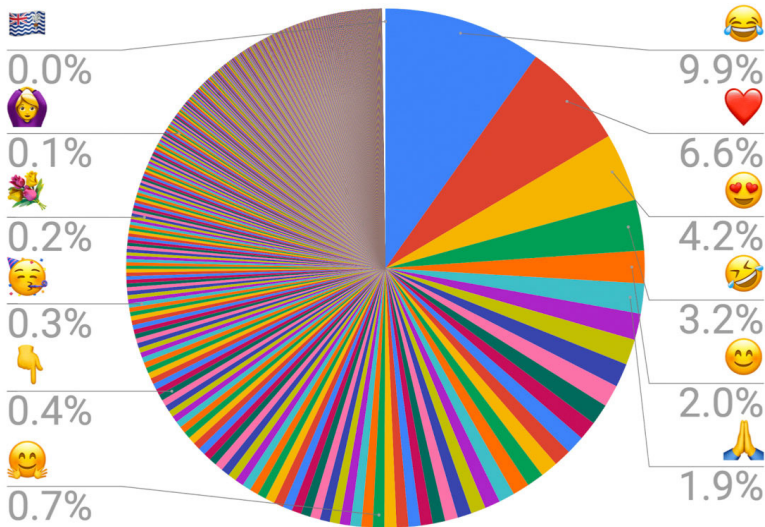
La quinta condición que se exige es que la serie no incluya más de siete valores. En los casos que se supera esa cantidad de valores, las superficies formadas por los dos radios y el arco que representa cada valor se hacen muy pequeñas, dificultando de esta manera la identificación de las mismas y el análisis visual por parte de un usuario. Un ejemplo de esta situación lo podemos encontrar en la figura 45. En esa figura aparece una gráfica circular con la que el Consorcio Unicode ha querido

Figura 44. Actividades desarrolladas a lo largo de un día



(elaboración propia, datos simulados)

Figura 45. Porcentaje de uso de los diferentes emoticonos en las redes sociales

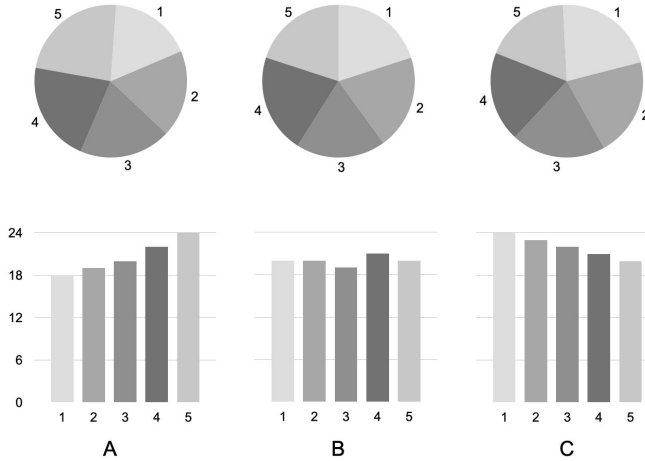


(<https://home.unicode.org/emoji/emoji-frequency/ados>)

representar el porcentaje o frecuencia de uso de los diferentes emoticonos en las redes sociales. Como vemos, el porcentaje de uso de cada uno de esos recursos expresivos se representa por una sección en la circunferencia generada por dos radios y un arco. Al intentar representar tantos valores dentro de una misma circunferencia, la mayoría, debido a su escaso tamaño, no se pueden identificar ni analizar visualmente de forma adecuada.

Por otro lado, para utilizar este tipo de gráfica de manera eficiente es importante también que los valores de la serie (las partes) sean lo suficientemente diferentes, que exista un contraste entre esos valores, para que se aprecien esas diferencias visualmente. Un ejemplo en el que no se cumple esa condición lo podemos observar en la figura 46. En esa figura, en la parte superior se representan 3 todos (A, B y C) con sus cinco secciones (secciones 1, 2, 3, 4 y 5) utilizando para ello tres gráficas circulares respectivamente. Como vemos, todas esas secciones son muy parecidas en tamaño, no existe mucho contraste entre los valores que están representando. Si observamos esas gráficas, nos costaría mucho comparar visualmente esas secciones y obtener unas conclusiones claras. Frente a las preguntas ¿la sección 4 de A es mayor o menor que la sección 2 de B? y ¿y la 5 de B respecto a la 3 de C?, las repuestas no son fáciles de obtener. Nos cuesta mucho esa comparación visual.

En cambio, si generamos 3 gráficas de columnas utilizando los mismos datos representados en las gráficas circulares, la cosa cambia radicalmente. Aunque esos

Figura 46. Gráficas circulares frente a sus gráficas de columnas equivalentes

(elaboración propia, datos simulados)

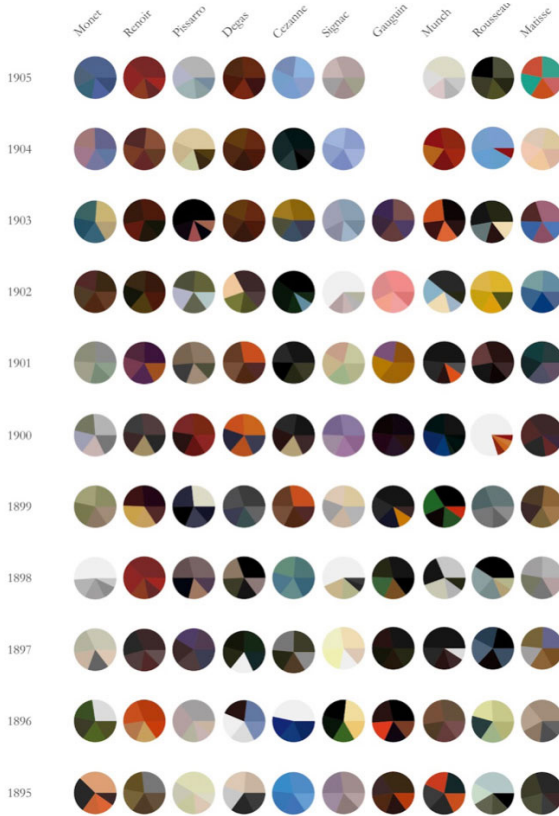
valores siguen siendo muy similares, podemos compararlos visualmente sin ningún tipo de problemas y contestar de forma rápida y sin temor a equivocarnos a las preguntas anteriores: 4 de A es mayor que 2 de B y 5 de B es menor que 3 de C.

Hay que destacar también, que en algunas situaciones, no es importante comunicar los valores cuantitativos exactos de las variables, sino que nuestra intención comunicativa se focaliza más en informar sólo sobre las proporciones y porcentajes que existen entre estos. En esos casos, podemos utilizar de forma cómoda las gráficas circulares y no otras propuestas visuales.

Ilustremos esta recomendación a través del ejemplo que recogemos en la figura 47. En esa figura, en 2012, Buxton y Ruths decidieron analizar los cuadros pintados por Monet, Renoir, Pissarro, Degas, Cezanne, Signac, Gauguin, Munch, Rousseau y Matisse entre 1985 y 1905. E identificaron los 5 colores más utilizados por cada uno de esos pintores en esos años elegidos. Después hicieron una gráfica circular por cada pintor y año para representar el porcentaje de uso de cada uno de los colores utilizados respecto a la suma de esos cinco colores. El resultado no nos permite identificar con precisión la proporción de cada uno de los colores en un pintor y en un año concreto. Pero, en cambio, sí que nos habilita a descubrir de forma cómoda, entre otras muchas cosas, la evolución cromática de esos artistas, la etapas de colorido por las que pasaron o, incluso, las coincidencias en el uso del color que compartieron. Y eso lo conseguimos con la utilización de las gráficas circulares.

Por último, y aunque no sea una condición de uso en un sentido estricto, para facilitar su lectura es recomendable que, si decidimos utilizar una gráfica de pastel,

Figura 47. Evolución cromática de diez artistas



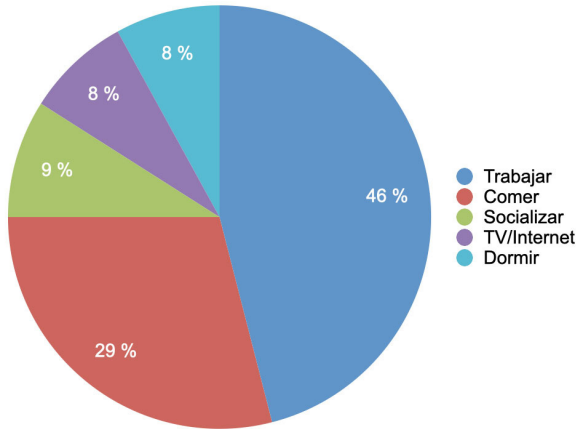
Ten Artists, Ten Years.
 Each chart represents an individual painting with the five most prominent colours shown proportionally.
 Arthur Buxton & Derek Ruths, 2012

(<https://flowingdata.com/2012/08/10/paintings-as-pie-charts>)

mejor coloquemos las secciones de mayor tamaño del gráfico circular en la parte superior (comenzando a las 12 en punto, por expresarlo de forma coloquial), que es donde primero se posa la mirada y, en orden decreciente, las sucesivas secciones más pequeñas a continuación de ésta. Así, por ejemplo, en la gráfica incluida en la figura 48, colocamos la tarea mayor (Trabajar) en la parte superior por ser la sección mayor y a continuación en orden decreciente las sucesivamente más pequeñas (Dormir, Socializar, TV/Internet y Comer).

Hasta aquí llegarían las condiciones que deben cumplirse para la utilización adecuada de una gráfica circular para la comunicación efectiva de la información. De todas formas, queremos concluir este apartado dedicado a este tipo de propuestas visuales, señalando que éstas comunican de una forma menos eficiente lo

Figura 48. Actividades desarrolladas a lo largo de un día



(elaboración propia, datos simulados)

mismo que podemos comunicar de forma más adecuada con otros gráficos alternativos (Few 2007).

Una de esas alternativas es, por ejemplo, la gráfica de barras. Este tipo de gráfica representa de una manera más clara y efectiva que la circular la relación que mantienen las partes con el todo al que pertenecen. La razón que justifica esa ventaja en la representación radica en los atributos que codifican los valores cuantitativos en cada uno de esos dos tipos de gráficas. Mientras que en la de barras es la longitud de esas figuras rectangulares la que codifica ese valor, en el caso de las circulares es el ángulo formado por los dos radios que delimitan la sección junto al arco el que lo hace.

Y el ser humano, a lo largo de su evolución se ha visto obligado a aprender a discriminar entre longitudes. Ha tenido que decidir si una zanja era lo suficientemente larga como para saltarla o, por el contrario, rodearla. O ha tenido que decidir si un peligro amenazante (otro hombre con actitud agresiva, por ejemplo) se encontraba demasiado cerca como para verse forzado a luchar y no poder contemplar la huida o una estrategia evasiva. Igual nos ha pasado con la discriminación entre los colores. El ser humano ha tenido que discriminar entre los colores porque su vida ha dependido de ello: había que discriminar entre el color del depredador (amenaza) y el color de la presa (comida) o entre las bayas moradas (venenosas) y las rojas (comestibles).

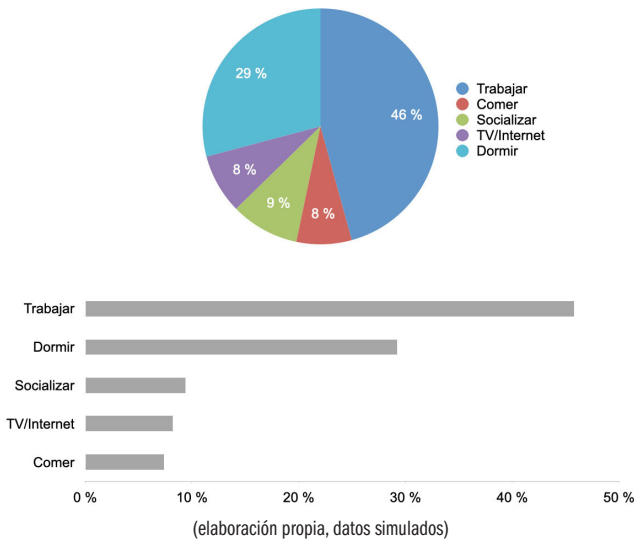
Sin embargo, el ser humano no se ha visto obligado a discriminar entre ángulos. Nuestra vida y el futuro de nuestra especie nunca ha dependido de esa habilidad cognitiva. Por eso, aunque nos gusten las gráficas circulares por el juego

cromático que nos ofrecen, éstas no nos permiten comunicar con eficacia la información cuantitativa encerrada en los datos. Al codificar a partir de los ángulos, nos entorpecen la comparación y el análisis visual de la información que comunican.

Este contraste entre los atributos utilizados en la codificación y sus limitaciones lo podemos ilustrar con la figura 49. En esta figura, la gráfica de barras nos permite comparar visualmente de forma más cómoda los valores de los porcentajes de las actividades diarias. Esa comodidad proviene de la utilización del atributo visual de la longitud de la barra para codificar el valor cuantitativo. Ese atributo es fácilmente aprehensible y cognitivamente es cómoda una comparación visual entre longitudes.

Por su lado, en cambio, la gráfica circular representa este valor mediante un área limitada por un semicírculo y dos rectas formando un ángulo. El atributo del ángulo (el verdadero codificador) no es fácilmente aprehensible; de la misma manera que tampoco es cognitivamente cómoda una comparación visual entre diferentes ángulos. Por esa razón, en muchos casos, para asegurarnos la comunicación de esos valores que representan las secciones, nos vemos obligados a utilizar etiquetas numéricas (como las que hemos introducido en la gráfica circular de la figura 49). Y, si para asegurar la comunicación tenemos que ofrecer esas etiquetas, nos surge una reflexión: ¿por qué utilizamos una gráfica circular para esa representación? Igual, sería mucho más eficiente generar como alternativa una tabla bien diseñada para comunicar esos valores a los usuarios.

Figura 49. Actividades desarrolladas a lo largo de un día



4.7. Treemap

La última gráfica que queremos abordar en este trabajo es el *treemap* (o gráfico de árbol). Este tipo de gráficas son propuestas de visualización consistentes en la disposición de superficies rectangulares anidadas jerárquicamente en un rectángulo mayor.

Desde un punto de vista visual, este tipo de gráficas se caracteriza por presentar una gran superficie rectangular que incluye en su interior (ocupando toda esa superficie) otras diferentes superficies rectangulares que pueden albergar, a su vez, de forma anidada, rectángulos más pequeños. Conceptualmente, todo rectángulo incluido en la figura es susceptible de estar formado simultáneamente por rectángulos de menor tamaño anidados en su interior.

Propuesto por Ben Shneiderman (*Human-Computer Interaction Lab, University of Maryland, EEUU*) a principios de la década de los noventa del pasado siglo para visualizar la estructura en carpetas del disco duro de un ordenador, este tipo de gráficas se acostumbra a utilizar para representar, en el mínimo espacio posible, grandes conjuntos de datos que se encuentran estructurados de una forma categorial o jerárquica; codificando además, no sólo esa estructura jerárquica, sino también el valor cuantitativo asociado a cada una de las categorías que forman parte de esa estructura.

La estructura jerárquica se representa mediante la disposición anidada de las superficies rectangulares. Cada categoría que se encuentra en un nivel superior en la jerarquía (las “categorías padre”, en un sentido coloquial) se representa mediante un rectángulo que alberga las superficies rectangulares que codifican las categorías que dependen de ésta en la estructura jerárquica (las “categorías hijo”, en un sentido lato). El color de esas superficies rectangulares ayuda a identificar visualmente esa jerarquía entre categorías. El valor cuantitativo asociado a cada una de las categorías (padres e hijos, por así decirlo) que forman parte de la gráfica se codifica mediante el tamaño de la superficie o área rectangular que la representa. El tamaño del área de la categoría padre es el total de sus subcategorías.

Si articulamos todas estas características que acabamos de destacar de una forma resumida, podemos ofrecer una versión de condensada de la definición de gráfica de áreas.

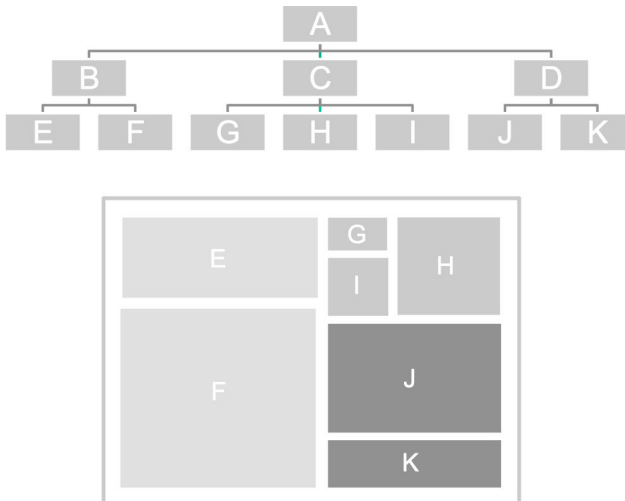
- i) *Definición de TreeMap*: instrumento para visualizar información cuantitativa que presenta una serie de características. Por un lado, se constituye como una gran superficie rectangular que incluye en su interior otras diferentes superficies rectangulares que pueden albergar, a su vez, de forma anidada, rectángulos más pequeños. Por otro, todo rectángulo incluido en la figura es susceptible de estar formado simultáneamente por rectángulos de menor tamaño anidados en su interior. Se utiliza para representar conjuntos de

datos que se encuentran estructurados de forma jerárquica. La estructura jerárquica se representa mediante la disposición anidada de las superficies rectangulares y el color de las mismas. El valor cuantitativo asociado a cada una de las categorías se codifica mediante el tamaño de la superficie o área rectangular que la representa.

Ilustremos estas características a través de un ejemplo. En la figura 50 podemos identificar una estructura jerárquica (en la parte superior) junto a su representación a través de un *treemap*. Como vemos, la estructura jerárquica está formada por una categoría raíz (o categoría padre, por seguir la misma nomenclatura) A de la que dependen tres categorías hijo (B, C y D). A su vez, esas tres categorías hijo están formadas por otras subcategorías (o categorías nieto): B está constituida por E y F; C por G, H e I; y D por J y K.

Esa estructura jerárquica aparece representada mediante el *treemap* que se ubica debajo. La categoría raíz A está representada por el rectángulo mayor, el que engloba el resto de los rectángulos más pequeños. Las tres categorías en el siguiente nivel de la jerarquía (B, C y D) se encuentran codificadas por tres rectángulos coloreados en tres tonalidades de gris. B está codificado mediante el color gris más claro, D por el más oscuro y C por el gris intermedio. Las subcategorías que forman el siguiente nivel de la jerarquía (y que dependen de B, C y D) también están

Figura 50. Estructura jerárquica representada a través de un *treemap*



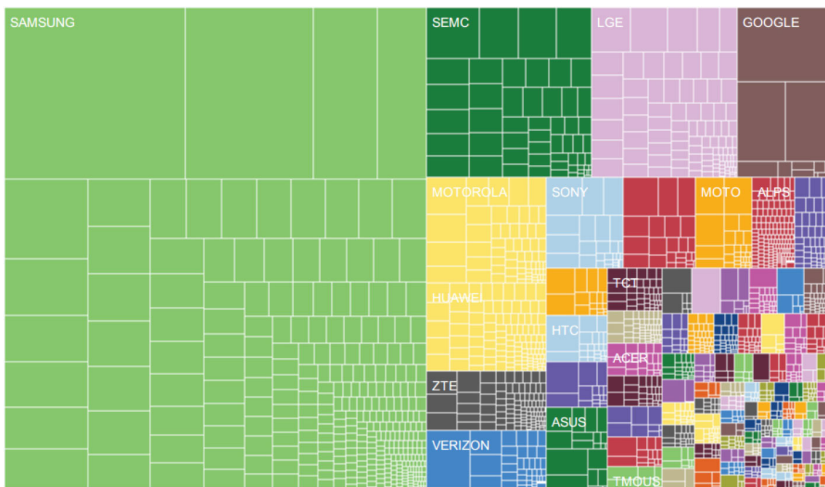
(elaboración propia, datos simulados)

representadas con el color de su categoría padre. Por tanto, E y F están codificados por dos rectángulos del mismo color que B; G, H e I por figuras del mismo color que C; y J y K por áreas del mismo color que D. El valor cuantitativo asociado a cada una de éstas últimas siete categorías se representa a partir del tamaño del área que la codifica. El valor cuantitativo de B, C y D se representa por la agregación del tamaño de las superficies de las categorías que se incluyen respectivamente en el nivel más bajo de la jerarquía. El valor cuantitativo asociado a la categoría A está codificado por el tamaño total de la figura rectangular que forma la gráfica, la suma de las áreas que representan las categorías B, C y D.

Este tipo de gráficas se utilizan mucho, sobre todo en el ámbito de la comunicación digital, ya sea en el contexto de la web o en el de las redes sociales. Un ejemplo de este uso lo encontramos en la figura 51. En la figura encontramos un *treemap* que nos muestra la fragmentación del ecosistema de fabricantes de teléfonos móviles con el sistema operativo Android. En la propuesta visual, el color resalta la agrupación categorial de forma que todos los dispositivos representados con el mismo color pertenecen a la misma marca; el tamaño de cada rectángulo codifica, en cambio, el valor numérico de las ventas del modelo representado.

Para finalizar, es importante destacar también que, para facilitar su lectura, en los contextos digitales este tipo de gráficas normalmente suelen ser interactivas de forma que al pasar el ratón nos aparezca el valor cuantitativo asociado a la superficie rectangular que estamos explorando visualmente (Murray, 2013).

Figura 51. Fragmentación del ecosistema de los teléfonos móviles con el sistema operativo Android



(<https://www.siliconweek.com/wp-content/uploads/2013/07/fragmentacionAndroid.png>)

Comunicación eficiente con gráficas

1. Introducción

EN EL CAPÍTULO ANTERIOR HEMOS PODIDO CARACTERIZAR UNO DE LOS PRINCIPALES productos visuales que se utilizan para representar y transmitir la información: las gráficas. En ese capítulo, abordamos cómo podemos definir una gráfica, analizamos los componentes estructurales que la conforman, revisamos las principales diferencias existentes respecto a su uso comunicacional frente a otros productos de visualización como las tablas y, por último, examinamos la tipología de las principales gráficas que solemos utilizar para representar y analizar información cuantitativa.

En este tercer capítulo, teniendo en cuenta todo lo desarrollado hasta este punto, vamos a intentar caracterizar los diferentes y principales contextos y contenidos comunicacionales que podemos transmitir utilizando una gráfica y veremos cuál es la propuesta visual más adecuada para cubrir con eficiencia y eficacia esas intenciones comunicativas. O, dicho en otros términos, en este capítulo queremos desplegar de forma ordenada la metodología habitual que se utiliza para implementar de forma adecuada una correcta visualización de información: en primer lugar se identifica el tipo de información a codificar, en segundo lugar, se elige el tipo de gráfica que mejor representa ese tipo de información y, por último, se diseñan los elementos visuales que conforman esa gráfica.

Para cubrir ese objetivo, y siguiendo ese esquema, vamos a tratar de desarrollar los siguientes puntos. En el primero de los apartados (apartado 2), abordaremos los principales tipos básicos de información o contenidos semánticos que podemos representar en una gráfica, ilustrándolos a partir de ejemplos que codificaremos en forma de tabla para que puedan ser entendidos de una forma más intuitiva. En el apartado 3, y recuperando la tipología de gráficas que desarrollamos en el capítulo anterior, examinaremos cuál es tipo de visualización que se adecua de forma más conveniente para comunicar cada uno de esos tipos de contenidos informativos. Por último, en el apartado 4, revisaremos algunos aspectos relacionados con el di-

seño eficiente de los principales elementos codificadores de información y componentes de soporte que conforman una gráfica y que detallamos también el capítulo anterior.

2. Contenidos semánticos

De forma intuitiva podemos afirmar que cualquier tipo de estrategia de comunicación se fundamenta sobre tres elementos básicos: el mensaje o contenido semántico que queremos transmitir, el receptor al cuál va dirigido y el contexto que rodea todo el acto comunicativo.

El caso de la comunicación a través del uso de gráficas no es diferente. Esa comunicación depende del mensaje o contenido informativo que pretendemos hacer llegar a nuestra audiencia, el público al cuál queremos que llegue ese mensaje y el contexto (la unidad documental en la que se inserta la gráfica, el texto que acompaña a la gráfica, el canal de transmisión, nuestra intención comunicativa o el momento de su consumo, entre otros muchos factores) que arropa todo ese intento de comunicar con gráficas. Centrémonos ahora exclusivamente en esos contenidos que queremos transmitir.

Normalmente, el contenido semántico que intentamos transmitir con una gráfica es una proposición (una idea, en un sentido no técnico) en la que se afirma la existencia de una relación entre dos o más tipos de información. Concretamente, se afirma la relación entre dos (o más) ítems de información cuantitativa y/o cualitativa. O, dicho de otra manera, en esos contenidos se expresa algún tipo de relación entre valores que pertenecen a dos o más variables cuantitativas, entre las diferentes categorías o atributos que conforman dos o más variables cualitativas, o entre categorías o atributos y valores cuantitativos. Así, por ejemplo, podemos utilizar una gráfica o una propuesta visual para comunicar que existe (o que no existe) una relación entre la inversión que hacemos en la reforma de unas viviendas (variable cuantitativa) y el precio final por el cual se venden esas viviendas (variable cuantitativa). También podemos transmitir con una propuesta visual la idea de que existe una relación entre el género de una persona (hombre o mujer; una variable cualitativa) y su relación con el tabaco (fumador o no fumador; otra variable cualitativa). Pero también podemos representar con el uso de una gráfica la relación existente entre el número de unidades vendidas de un producto concreto (variable cuantitativa) y la nacionalidad (variable cualitativa) de los diferentes vendedores que conforman el departamento comercial de la empresa que fabrica ese producto.

Más allá de ejemplos concretos, es posible establecer una clasificación de las principales relaciones expresadas en los contenidos semánticos representados por una gráfica (Zelazny, 2001; Few, 2012; Shneiderman, 1996). Entre estas, podríamos

destacar, por ejemplo, la relación de comparación nominal, la de series temporales, la de ranking, la de parte-todo, la de desviación, la de distribución y la relación de correlación.

A estos tipos de contenidos semánticos vamos a dedicar los siguientes apartados. Para cada uno de ellos, introduciremos una definición y, para completarla, trataremos de ilustrarlo mediante el uso de una tabla que recoja ese contenido conceptual.

2.1. Relación de comparación nominal

La relación de comparación nominal podría ser considerada como el contenido semántico más básico y habitual que solemos representar utilizando una gráfica.

En términos técnicos este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la comparación entre esos valores numéricos relacionados con esos ítems. En este sentido, y entre otras muchas cosas, este contenido semántico nos permite identificar y expresar, por ejemplo, cuál es el ítem cualitativo que lleva asociado el valor cuantitativo más alto, cuál es la categoría cualitativa que lleva asociada el valor cuantitativo más bajo, o si a una de las categorías en concreto se le asigna un valor numérico mayor o menor que a otro de esos valores cualitativos.

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 1. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (gerente, analista, documentalista, administrativo, auxiliar y becario) de una variable cualitativa (los diferentes empleados (cargos o perfiles laborales) que podemos encontrar en la

Figura 1. Cargo y sueldo de los empleados de la empresa ACME

Cargo	Sueldo
Gerente	150.000
Analista	40.000
Documentalista	40.000
Administrativo	32.000
Auxiliar	25.000
Becario	8.000

(elaboración propia, datos simulados)

empresa ACME) con un valor numérico o cuantitativo (el sueldo bruto que reciben cada uno de esos empleados según el cargo que desempeñan en la empresa).

En este sentido, este contenido semántico nos permite, además de otras cosas, identificar y expresar, por ejemplo, que el sueldo más alto de la empresa lo cobra el gerente, que el más bajo lo recibe el becario, que el analista y el documentalista perciben el mismo sueldo o que el salario del gerente es seis veces más alto que el del auxiliar.

2.2. Relación de series temporales

La relación de series temporales es otro de los contenidos semánticos habituales que acostumbramos a comunicar mediante el uso de gráficas.

De nuevo, desde una perspectiva conceptual, este contenido semántico se identifica como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable temporal para facilitar la comparación entre esos valores numéricos asociados a esas unidades temporales. Es uno de los contenidos semánticos más representado en los contextos económicos o de negocio y persigue expresar la evolución de una variable cuantitativa a través del tiempo.

En esta línea, este contenido semántico nos permite identificar y expresar, por ejemplo, cuál ha sido la evolución en el tiempo de una variable cuantitativa, cuál es la unidad temporal que tiene asignado el valor cuantitativo más alto, cuál es la que tiene el más bajo, si existe algún tipo de patrón en esa evolución o si existen algunos valores atípicos que se escapan del patrón normal de esa evolución en el tiempo.

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 2. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (primero, segundo, tercero y cuarto) de una variable temporal (los diferentes trimestres del año pasado) con un valor numérico o cuantitativo (las ventas que ha generado la empresa en esas unidades temporales).

Figura 2. Ventas por trimestre en el ejercicio anterior

Trimestre	Ventas
Primero	456.876
Segundo	189.387
Tercero	63.829
Cuarto	122.345

(elaboración propia, datos simulados)

En este sentido, este contenido semántico nos permite identificar y expresar, por ejemplo, que el primer trimestre tuvo las ventas más altas, que en el tercero fue cuando menos se vendió y que la evolución de los ingresos comenzó con unas ventas altas en el primer trimestre, que luego bajaron durante los trimestres segundo y tercero y que repuntaron en el último trimestre del ejercicio anterior.

2.3. Relación de ranking

Pasemos ahora a abordar la relación de ranking. Este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se señala que esta asignación de valores numéricos establece una ordenación (de mayor a menor o de menor a mayor) entre los ítems de la variable cualitativa.

El objetivo es expresar ese orden entre las categorías cualitativas derivado de la asignación y poder facilitar también la comparación entre esos valores numéricos relacionados con esos ítems. En este sentido, este contenido semántico nos permite identificar y expresar, por ejemplo, cuál es el ítem cualitativo que lleva asociado el valor cuantitativo más alto, cuál es la categoría cualitativa que lleva asociada el valor cuantitativo más bajo, si a una de las categorías en concreto se le asigna un valor numérico mayor o menor que a otro de esos valores cualitativos o el orden (ascendente o descendente) de todos los ítems cualitativos utilizando como criterio los valores numéricos asociados.

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 3. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (Barcelona, Tarragona, Gerona y Lérida) de una variable cualitativa (las diferentes provincias que conforman la región española de Cataluña) con un valor numérico o cuantitativo (las ventas que una empresa ha conseguido en esa provincia). Esta asignación de valores numéricos establece una ordenación (de mayor a menor o de menor a mayor) entre esas provincias catalanas.

Figura 3. Ventas por provincias en Cataluña

Ranking	Provincia	Ventas
1	Barcelona	456.876
2	Tarragona	189.387
3	Gerona	122.345
4	Lérida	63.829

(elaboración propia, datos simulados)

En este contexto, además de otras cosas, este contenido semántico nos permite identificar y expresar, por ejemplo, que la provincia donde más se vendió fue Barcelona, que Lérida fue la provincia con las ventas más bajas o que, si utilizamos como criterio las ventas producidas en esas localidades, podemos establecer un orden (ranking) descendente entre esas provincias (en primera posición Barcelona, luego Tarragona, en tercera posición Gerona y, por último, Lérida).

2.4. Relación de parte-todo

La relación de parte-todo podría ser considerada como otro de los contenidos más habituales que solemos representar utilizando una gráfica.

Desde una perspectiva conceptual, este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se señala que la agregación de la totalidad de esos ítems cualitativos conforma una entidad o un todo y, por tanto, cada una de esas categorías se identifica como una de las partes de ese todo.

El objetivo en este caso es expresar esa relación cuantitativa que mantienen cada una de las partes (su peso, por así decirlo) respecto al todo y poder facilitar también la comparación entre esos valores numéricos relacionados con esos ítems. En este sentido, este contenido semántico nos permite identificar y expresar, por ejemplo, cuál es el ítem cualitativo (la parte del todo) que lleva asociado el valor cuantitativo más alto (respecto al todo), cuál es la categoría cualitativa (la parte del todo) que lleva asociada el valor cuantitativo más bajo (respecto al todo) o si a una de las categorías en concreto (a una parte del todo) se le asigna un valor numérico mayor o menor que a otro de esos valores cualitativos (a otra parte del todo).

Es importante destacar también que la unidad común de medida más utilizada en este tipo de relaciones que expresan la contribución de las partes a un todo es el porcentaje. En ese contexto, el todo es el 100% y cada una de sus partes es una porción de ese 100%.

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 4. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (trabajar, comer, socializar, ver TV/internet y dormir) de una variable cualitativa (las diferentes actividades realizadas a lo largo del día) con un valor numérico o cuantitativo (el porcentaje de tiempo del día que le dedicamos a esa actividad). Esta asignación de valores numéricos expresa la relación cuantitativa que mantienen cada una de las partes (su peso, por así decirlo) respecto al todo y facilita también la comparación entre esos valores numéricos relacionados con esos ítems.

Figura 4. Actividades desarrolladas a lo largo de un día

Actividad	Porcentaje
Comer	7,40%
TV/Internet	8,20%
Socializar	9,40%
Dormir	29,20%
Trabajar	45,80%
Total	100,00%

(elaboración propia, datos simulados)

En esta línea, y entre otras muchas cosas, este contenido semántico nos permite identificar y expresar, por ejemplo, que trabajar es la actividad a la cual dedicamos más tiempo a lo largo del día, que comer es la actividad diaria en la que gastamos menos tiempo o que invertimos en dormir tres veces más tiempo que en navegar por internet y/o ver la televisión.

2.5. Relación de desviación

De forma intuitiva, la relación de desviación expresa como varía un conjunto de datos cuantitativos respecto a un valor numérico concreto.

En un sentido más técnico, es un contenido semántico que se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se comparan esos valores cuantitativos asociados respecto a un valor numérico concreto y de referencia. Normalmente, el valor numérico de referencia se encuentra también asignado a un ítem cualitativo (de esa misma u otra variable cualitativa). En algunas circunstancias pueden existir dos o más valores de referencia respecto a los cuales se quiere realizar la comparación.

El objetivo en este caso es expresar esa relación cuantitativa que mantienen cada uno de esos valores asignados a esos ítems respecto al valor de referencia. En este sentido, este contenido semántico nos permite identificar y expresar, por ejemplo, cuál es el ítem cualitativo que lleva asociado el valor cuantitativo más alejado del valor de referencia, cuál es la categoría cualitativa que lleva asociada el valor cuantitativo más cercano al valor de referencia o cómo es cuantitativamente hablando esa distancia entre esos valores numéricos comparados.

Es importante destacar también que esta comparación cuantitativa respecto al valor de referencia suele expresarse en términos porcentuales. Y que puede imple-

Figura 5. Comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa

Empresa	Ventas
Competidor A	280%
Competidor B	220%
Competidor C	115%
Competidor D	60%
Competidor E	20%

(elaboración propia, datos simulados)

mentarse de dos formas distintas. Por un lado, el valor de referencia se identifica con el 100% y el resto de los valores son expresados como una proporción o porcentaje respecto a ese valor primario o de referencia. Y, por otro, el valor de referencia se identifica con el 0% y el resto de los valores son expresados como una proporción o porcentaje positivo o negativo respecto a ese valor primario o de referencia.

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 5. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (competidor A, competidor B, competidor C, competidor D y competidor E) de una variable cualitativa (las diferentes empresas que compiten sectorialmente con la nuestra) con un valor numérico o cuantitativo (el porcentaje de ventas de esa firma respecto a las ventas de nuestra empresa). Esta asignación de valores numéricos expresa la relación cuantitativa que mantienen las ventas de cada empresa respecto a las de la nuestra y facilita también la comparación entre esos valores numéricos relacionados con esos ítems. En este caso, el valor de referencia (las ventas de nuestra empresa) se identifica con el 100% y el resto de los valores (las ventas de la competencia) son expresados como una proporción o porcentaje respecto a ese valor primario o de referencia.

En esta línea, este contenido semántico nos permite identificar y expresar, por ejemplo, que el competidor A es la firma que más vende en el sector (casi el triple de lo que vende nuestra empresa, un 280%), que el competidor B vende un poco más del doble de lo que vende nuestra organización (un 220%), que el competidor C vende un poco más que nuestra empresa (un 115%) o que los competidores D y E venden menos que nosotros (un 60% y 20% de las ventas de nuestra empresa, respectivamente).

2.6. Relación de distribución

La relación de distribución podría ser considerada como un contenido semántico que se corresponde con una proposición (una idea) en la que se afirma la

existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable de intervalo para facilitar la comparación entre esos valores numéricos asociados a esas unidades de intervalo. Ese contenido expresa cómo los valores cuantitativos de un conjunto se encuentran repartidos o diseminados a lo largo de un rango (intervalo entre el valor máximo y el valor mínimo) de ese conjunto, desde el valor más bajo hasta el más alto, a través de las categorías que conforman esa variable de intervalo.

En este sentido, este contenido semántico nos permite identificar y expresar, por ejemplo, cuál es el valor cuantitativo más alto dentro de ese conjunto de datos (el límite superior del rango), cuál es el valor cuantitativo más bajo dentro de ese conjunto de datos (el límite inferior del rango), si la mayoría de los valores se agolpan en la parte baja del rango, si la mayoría de los valores se concentran en la parte alta del rango, si la mayoría de los valores se aglutinan en la parte central del rango o si los valores se reparten de manera uniforme a lo largo de ese rango del conjunto de datos o valores cuantitativos.

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 6. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (0 a 15, 16 a 30, 31 a 45, 46 a 60 y 61 a 70) de una variable de intervalo (las franjas de edad de una población de pacientes que sufren episodios de epilepsia) con un valor numérico o cuantitativo (el número de pacientes). Esta asignación de valores numéricos expresa la distribución del número de pacientes (cómo están repartidos o diseminados) por franja de edad de una población de 200 enfermos de epilepsia.

En esta línea, este contenido semántico nos permite identificar y expresar, por ejemplo, que la franja de edad que presenta mayor número de pacientes (75) va de los 46 a los 60 años, que el menor número de pacientes (15) se encuentra en la franja de los 0 a los 15 años, que de los 16 a los 30 años aumenta respecto a la franja anterior el número de pacientes (subiendo hasta los 30), que en la franja siguiente (de

Figura 6. Pacientes por franja de edad en una población de enfermos de epilepsia

Franja edad	Pacientes
0-15	15
16-30	30
31-45	55
46-60	75
61-75	25

(elaboración propia, datos simulados)

31 a 45 años) se vuelve a disparar el número de pacientes (llegando a las 55) o que en la última franja de edad, la que va de los 61 a los 75 años, el número de pacientes vuelve a disminuir (25).

2.7. Relación de correlación

En un sentido intuitivo, la relación de correlación intenta expresar si dos conjuntos de valores cuantitativos se encuentran relacionados entre sí y, si es el caso, cómo varía uno respecto al otro.

Desde una perspectiva más conceptual y en su formulación más básica, este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asignan o conectan dos valores cuantitativos (cada uno perteneciente a una variable cuantitativa distinta) a cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la identificación de algún patrón en la variación entre esos dos tipos de valores numéricos relacionados con esos ítems. Intenta expresar si existe o no una situación en los que los valores de una de las dos variables cuantitativas se modifican o cambian de manera sistemática con respecto a los valores de la otra.

En este sentido, y entre otras muchas cosas, este contenido semántico nos permite identificar y expresar, por ejemplo, si existe o no ese patrón entre la variación de dos variables cuantitativas asociadas a una variable cualitativa, en el caso de que exista el patrón nos indica el signo de ese patrón o esa relación (positiva o negativa) o, también el caso de que exista el patrón, nos muestra el grado de esa relación (alta o baja).

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 7. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (producto 1, producto 2 y producto 3) de una variable cualitativa (los diferentes productos que fabrica la empresa ACME) con dos valores numéricos o cuantitativos que pertenecen respectivamente a dos variables cuantitativas (la inversión realizada en la fabricación y distribución de un producto, por un lado; y los beneficios totales obtenidos en la vida comercial de cada producto, por otro).

Figura 7. Inversión y beneficios por producto fabricado

Producto	Inversión	Beneficios
Producto 01	22.864	678.543
Producto 02	45.789	778.765
Producto 03	50.678	783.213
Total	119.331	2.240.521

(elaboración propia, datos simulados)

En este sentido, además de otras cosas, este contenido semántico nos permite identificar y expresar, por ejemplo, que existe cierto patrón o relación clara entre la inversión que se realiza en un producto y los beneficios que se obtienen del mismo. Vemos que a medida que varía esa inversión, varía también el beneficio que se obtienen de cada producto. Y podemos afirmar también que esa variación es positiva, ya que cuando aumenta la inversión aumentan también los beneficios.

3. Visualización de contenidos

Una vez revisados los tipos de gráficas que se suelen utilizar habitualmente para codificar información y los principales contenidos semánticos que queremos comunicar utilizando gráficas, nos queda ahora abordar el proceso de visualización de esos contenidos o unidades proposicionales a través de la representación gráfica. O, dicho en otros términos, queremos analizar cuáles son las mejores soluciones gráficas para codificar adecuadamente cada uno de esos tipos de contenidos semánticos.

En un sentido práctico, normalmente, cuando disponemos de un conjunto de datos en un sistema de gestión de datos (Excel o Numbers (Apple), por ejemplo) y nuestro objetivo es visualizarlo, nos asalta siempre la misma duda: ¿qué gráfica es la que tenemos que utilizar? Una de las acciones críticas a la hora de generar una buena visualización es, aunque no sea la única, la elección correcta de la gráfica que mejor comunica (de forma eficaz y eficiente) la información que queremos representar. No se trata de una tarea sencilla. Depende de muchos factores, pero especialmente del tipo de información a representar y de nuestra intencionalidad comunicativa. Y es que cada tipo de contenido semántico reclama una representación gráfica determinada dependiendo también de nuestros objetivos comunicativos.

Si un procesador de textos nos ayuda a escribir mejor (nos detecta errores orto-tipográficos y sintácticos en nuestros textos), los programas que solemos utilizar para codificar información no nos asisten, en cambio, en esta línea. La situación ideal sería que ese programa informático con el que gestionamos los datos nos generara de forma adecuada la mejor de las visualizaciones de acuerdo con nuestras intenciones comunicativas.

Pero ese software no existe. La mayoría de los programas de gestión de datos no nos asisten ni ayudan en la elección de la mejor gráfica para realizar esa visualización (Hugues y van Dam, 2013; Hearn, 2011; Shirley, 2009). Los softwares más populares que podemos utilizar para crear visualizaciones (Excel, Numbers o la hoja de cálculo de Drive, por ejemplo) no nos ayudan mucho en esa tarea de selección de esa gráfica. En muchos casos, seleccionamos los datos que tenemos en el fichero y el programa nos ofrece un catálogo fijo de visualizaciones entre las cuáles

tenemos que elegir la que consideremos como más adecuada, pero no qué gráfica debemos utilizar en función del tipo de datos que tenemos y de nuestras intenciones comunicativas. Sólo algunas herramientas informáticas (como Tableau, por ejemplo) nos ofrecen algunas pequeñas orientaciones parciales (incompletas, en muchos casos) para hacerlo en función de la estructura de los datos a representar (Cherven, 2015; Jones, 2014; Khan, 2016).

En este apartado vamos a intentar afrontar ese problema. Vamos a intentar ofrecer una serie de recursos y consejos para poder elegir la gráfica comunicativamente más adecuada para representar cada uno de los contenidos semánticos que solemos codificar. Y lo haremos utilizando los principales elementos visuales disponibles para representar datos en gráficas: los puntos, las líneas, las barras y las columnas.

Concretamente, introduciremos los mejores recursos para visualizar de forma adecuada la relación de comparación nominal, la de series temporales, la de ranking, la de parte-todo, la de desviación, la de distribución y la relación de correlación.

3.1. Visualización de la relación de comparación nominal

Como ya vimos anteriormente, la relación de comparación nominal asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la comparación entre esos valores numéricos relacionados con esos ítems.

Teniendo esto en cuenta, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifique simultáneamente el valor numérico y el ítem cualitativo con el que se encuentra asociado. Y deben codificar esta relación mostrando que cada una de las duplas resultantes formadas por un valor cuantitativo y una categoría cualitativa son independientes entre sí, que no existe una conexión entre estas duplas.

Sin perder de vista esta restricción semántica, los objetos visuales que mejor codifican esta relación son las barras (o columnas) y los puntos. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo) mostrando que visualmente son independientes entre ellas y que, por tanto, no existe una conexión intrínseca entre las mismas. Al hacerlo de ese modo, facilitan así la comparación visual entre esos valores numéricos relacionados con esos ítems.

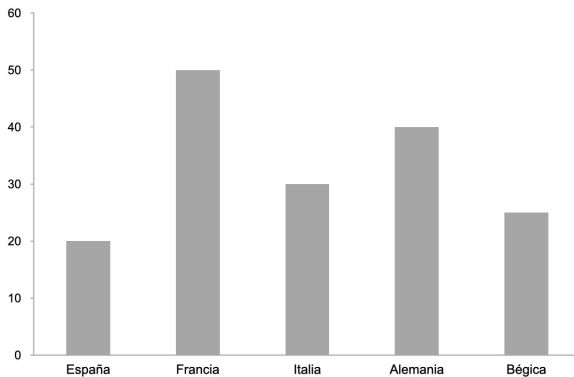
En una gráfica de barras o columnas, y respecto a su ubicación frente a los ejes X e Y, cada barra (o columna) nos codifica una pareja de valores cuantitativo y cualitativo de forma aislada, independiente, sin representar o mostrarnos una relación visual entre esas barras (o columnas) que conforman la gráfica. Esta representación independiente habilita la comparación visual entre esos valores numéricos relacionados con esos ítems.

Lo mismo ocurre con los puntos. En una gráfica de puntos, y, de nuevo, respecto a su ubicación frente a los ejes, cada punto nos codifica una pareja de valores cuantitativo y cualitativo de forma aislada, independiente, sin representar o mostrarnos una relación visual entre esos puntos que forman parte la gráfica. Con esa codificación independiente se habilita la comparación visual entre esas duplas formadas por valores numéricos asociadas con esos ítems cualitativos.

Ilustremos ese comportamiento representacional a través de un par de ejemplos. En la figura 8, aparece una gráfica de columnas en la que se representa la relación de comparación nominal. Concretamente, se intenta representar los goles en contra recibidos por las cinco selecciones nacionales de fútbol mejor clasificadas en la última Eurocopa de fútbol para facilitar la comparación entre esos datos. Gracias a su ubicación respecto a los ejes X e Y, cada columna representa de forma simultánea una de esas selecciones y los goles en contra que ésta recibió en ese campeonato. Codifica esta información de forma aislada, independiente, sin mostrarnos una relación visual con el resto de las columnas que representan a las otras selecciones. Esta representación independiente habilita la comparación visual entre el número de goles en contra recibidos por esas selecciones.

En ciertos contextos, la gráfica de puntos puede considerarse como una buena alternativa a la gráfica de columnas (o barras) en la representación de este tipo de contenido semántico. Concretamente, cuando la diferencia entre los valores cuantitativos que se representan es muy escasa. O, dicho en otros términos, cuando el rango en el que fluctúan esos valores es pequeño, cuando existe un grado de dispersión baja entre ese conjunto de valores.

Figura 8. Goles en contra recibidos por las cinco selecciones nacionales de fútbol mejor clasificadas en la última Eurocopa de fútbol



(elaboración propia, datos simulados)

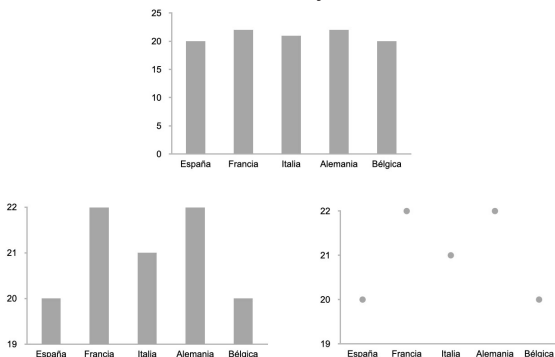
En esos casos, si representamos esos datos mediante una gráfica de columnas (o barras), puede costar apreciar la diferencia entre la longitud de esas figuras geométricas y, dificultar, de esa manera, la comparación visual entre las mismas y los valores numéricos que codifican. Para representarlos de una forma alternativa con columnas (o barras) y hacer evidente las diferencias entre los valores, podemos tener la tentación de comenzar el eje vertical Y (o X, en el caso de la gráfica de barras) en un valor superior a cero, en un valor poco menor que el valor numérico más pequeño representado y hacer visualmente más grandes la diferencia entre la longitud de esas figuras geométricas y de los valores que codifican.

Pero esa práctica es incorrecta. Las columnas (o barras), al codificar el valor cuantitativo mediante su longitud, necesitan comenzar en el punto cero del eje vertical (u horizontal, en el caso de la gráfica de barras). Si no lo hacen, no codifican correctamente el valor numérico.

Si detenemos nuestra atención en la figura 9, podemos observar esta limitación. En esa figura, en la parte superior aparece una gráfica de columnas donde se representan los goles en contra recibidos por las cinco selecciones nacionales de fútbol mejor clasificadas en la última Eurocopa de fútbol. Como la diferencia entre esos goles en contra recibidos por las selecciones es escasa, la comparación visual entre las columnas se hace difícil.

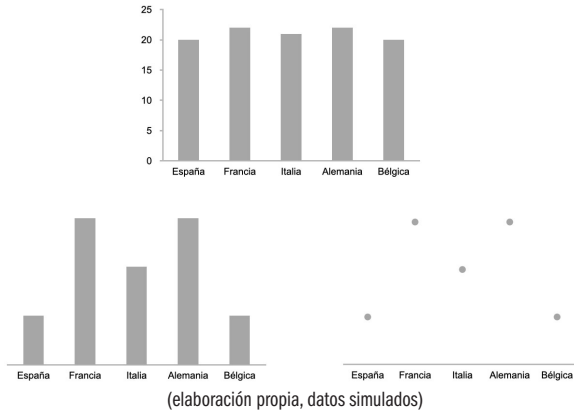
Para hacer evidentes las diferencias entre los goles en contra, podemos tener la tentación de comenzar el eje vertical Y de la gráfica de columnas en un valor superior a 0 (en 19, por ejemplo) y hacer visualmente más grande la diferencia entre las columnas que representan esos goles en contra. Como hemos hecho en la gráfica de columnas situada en la parte inferior e izquierda de la figura 9. Pero esta

Figura 9. Gráficas de columnas y de puntos para representar los goles en contra recibidos por las cinco selecciones nacionales de fútbol mejor clasificadas en la última Eurocopa



(elaboración propia, datos simulados)

Figura 10. Gráficas de columnas y de puntos para representar los goles en contra recibidos por las cinco selecciones nacionales de fútbol mejor clasificadas en la última Eurocopa de fútbol



representación no es correcta. Si un usuario se fija en esa gráfica puede entender erróneamente que Francia o Alemania recibieron el doble de goles en contra que España. El error se justifica por que, al comenzar el eje Y en un valor distinto cero, estamos enmascarando la longitud real de las barras que representan los goles en contra.

En cambio, si utilizamos una gráfica de puntos como alternativa, esa comprensión errónea no debería producirse ya que los puntos reciben su valor cuantitativo y cualitativo asociado a través de la posición que ocupan en el plano generado por los ejes X e Y. En la gráfica de puntos que incluimos en la figura 9 (abajo, a la derecha), el eje también comienza en un valor distinto a cero, en 19. Por tanto, como los puntos representan el valor numérico por su posición respecto al eje Y (y no su longitud, como en el caso de la columna o las barras), sí que codifican, de forma adecuada y de manera visualmente más cómoda, la diferencia entre los goles recibidos por esas selecciones nacionales.

Esto se hace más evidente si borramos el eje Y en las gráficas en las que éste no comenzaba en cero, como hacemos en la figura 10. En esta figura, la gráfica de columnas inferior continúa representando erróneamente que Francia o Alemania recibieron el doble de goles en contra que España. En cambio, la gráfica de puntos nos impide concluir, cuando es falso, que Francia o Alemania recibieron el doble de goles en contra que España, ya que deberíamos saber cómo es la escala del eje Y para poder observar eso. En cambio, esa misma gráfica, nos permite ver, de forma correcta y entre otras cosas, que la diferencia de goles encajados entre España y Alemania (2 goles) es el doble que la diferencia de goles encajados entre España e Italia (1 gol).

3.2. Visualización de la relación de series temporales

Como ya destacamos, la relación de series temporales asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable temporal para facilitarnos la comparación entre esos valores numéricos asociados a esas unidades temporales.

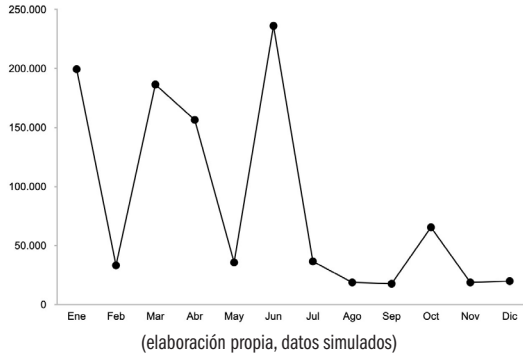
A la luz de esa descripción, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente el valor numérico y la categoría temporal con la que se encuentra asociado. Para que estos objetos puedan codificar estas duplas semánticas de forma adecuada, en uno de los ejes debe ubicarse esa variable temporal con cada una de las etiquetas que representan un momento de tiempo (año, mes, etc.). Y, como los valores temporales poseen un orden natural, ese orden debe ser también respetado en la visualización. Por convención, en nuestra cultura occidental este tipo de orden se suele visualizar horizontalmente, de izquierda a derecha, a través del eje X.

Teniendo en cuenta todo esto, los objetos visuales que mejor codifican esta relación son la combinación de puntos y líneas, las líneas utilizadas de forma no combinada y las columnas. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo) mostrando que cada una de esas duplas está visualmente conectada con la siguiente, que son el eslabón o estadio dentro de una cadena temporal. Al hacerlo de ese modo, representan esa continuidad temporal entre esas duplas.

En una gráfica de puntos y líneas combinados, y respecto a su ubicación frente a los ejes X e Y, cada punto nos codifica un valor cuantitativo asociado a un momento de tiempo, y nos conecta mediante una línea esa figura circular con el punto anterior y posterior para que entendamos que, a partir de esa conexión visual mediante la línea, se desarrolla un proceso temporal donde el punto es sólo uno de sus momentos.

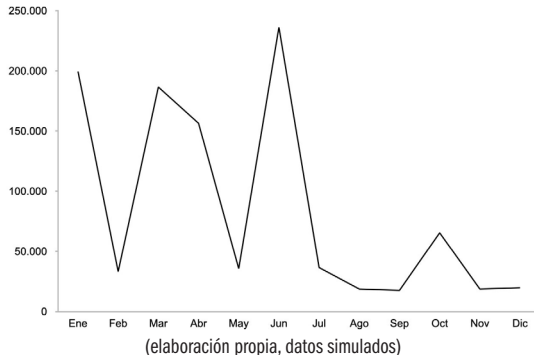
Con la figura 11 podemos ilustrar este efecto. En esa gráfica de puntos y líneas combinados, cada punto nos codifica el valor de las ventas de la empresa ACME (respecto al eje Y) en un mes concreto del año pasado (respecto al eje X) y nos conecta esas ventas mediante una línea con las de los meses anteriores y posteriores para que entendamos que, a partir de esa conexión visual mediante la línea, se ha producido una evolución de las ventas a lo largo de ese año.

Lo mismo ocurre en una gráfica de líneas. En relación con su ubicación respecto a los ejes X e Y, cada vértice de la línea nos codifica un valor cuantitativo asociado a un momento de tiempo y nos conecta mediante una línea ese punto concreto en el plano con el vértice anterior y posterior para que entendamos que, a partir de esa conexión visual mediante la línea, se desarrolla un proceso temporal donde cada uno de los vértices es sólo uno de sus momentos en esa cadena temporal.

Figura 11. Ventas mensuales de la empresa ACME durante el año pasado

En la figura 12 podemos ver un ejemplo de este tipo de gráficas. En este caso, en esta gráfica de líneas, cada vértice es el que nos codifica el valor de las ventas de la empresa ACME (respecto al eje Y) en un mes concreto del año pasado (respecto al eje X). Y, de la misma forma que en el caso anterior, nos conecta esas ventas mediante una línea con la del mes anterior y posterior para que entendamos que, a partir de esa conexión visual mediante la línea, se ha producido una evolución de las ventas a lo largo de ese año.

La gráfica de columnas, en cambio, la utilizaremos cuando queramos representar una relación de series temporales, pero queramos poner nuestro foco comunicativo en la transmisión de los valores cuantitativos individuales asociados a los momentos de tiempo de una variable temporal (cuando busquemos facilitar la comparación entre esos valores) y no en la evolución global de la serie. En una gráfica de columnas, y respecto a su ubicación frente a los ejes X e Y, cada columna nos codifica un valor cuantitativo asociado a un momento de tiempo, pero al estar

Figura 12. Ventas mensuales de la empresa ACME durante el año pasado

separadas por un espacio y no estar conectadas mediante una línea, no podemos poner el acento comunicativo en la evolución global de esos valores en la cadena temporal.

Con la figura 13 podemos ilustrar este uso. En esa gráfica de columnas, cada rectángulo nos codifica el valor de las ventas de la empresa ACME (respecto al eje Y) en un mes concreto del año pasado (respecto al eje X). Esta representación nos permite poner nuestro acento y enfatizar comunicativamente la comparación entre las ventas que se han producido en cada uno de esos meses y no la evolución global de esas ventas durante el año representado.

Es importante destacar respecto al uso de la gráfica de columnas, que la podemos utilizar con esa intención siempre y cuando no convivan en un mismo eje X demasiadas unidades temporales. Si las utilizamos cuando existe una serie formada por muchos ítems temporales, el resultado es una gráfica con un número elevado de columnas que entorpecen la comparación de los valores numéricos asociados a éstas. En esos casos, es mejor utilizar una gráfica de líneas.

En la figura 14 utilizamos una gráfica de líneas para representar la inversión anual en educación secundaria (línea azul) y en educación primaria (línea roja) durante el periodo que va de 1992 a 2011. Con las líneas podemos ver la evolución de esa inversión. En cambio, si hubiéramos elegido una gráfica de columnas para representar esos 20 años de inversión, el resultado hubiese sido un plano generado por los dos ejes poblado por una inflación visual de 40 columnas (20 para cada tipo de educación) que no nos hubiera facilitado la comparación visual entre las inversiones anuales realizadas en cada uno de esos dos tipos de educación.

Frente a la combinación de puntos y líneas, líneas y columnas, los objetos visuales que no utilizaremos para representar las series temporales son las barras y los puntos (no combinados con líneas).

Figura 13. Ventas mensuales de la empresa ACME durante el año pasado

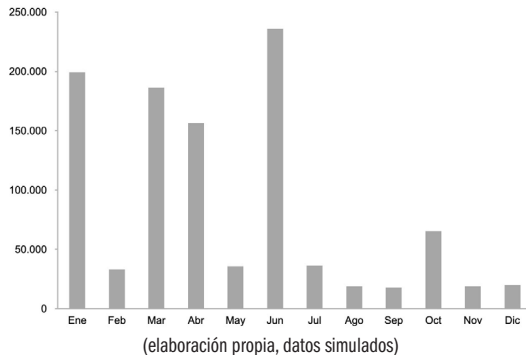
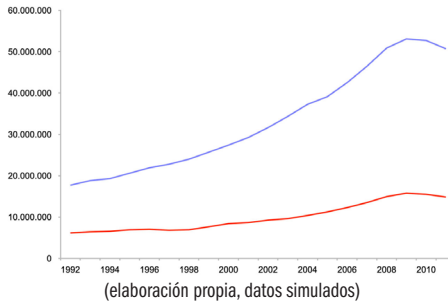


Figura 14. Inversión anual en educación secundaria y en educación primaria durante el periodo que va de 1992 a 2011

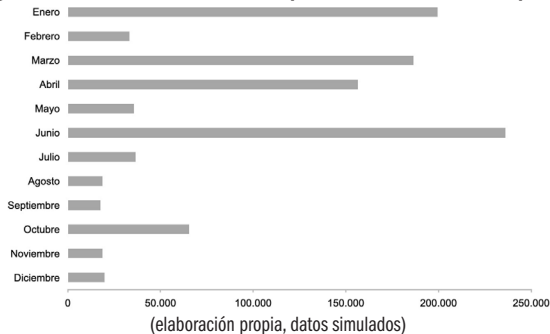


Las gráficas de barras no funcionan bien a la hora de representar este contenido semántico ya que utilizan el eje X para las etiquetas cuantitativas y el eje Y para las subdivisiones temporales. Y al hacerlo, no ofrecen una representación de izquierda a derecha de la evolución de unos valores en el tiempo, sino una codificación de la evolución temporal de arriba a abajo. Y esto va en contra de las convenciones en nuestra cultura occidental. En nuestra cultura, cuando imaginamos el desarrollo de un período en el tiempo, lo hacemos como algo que va de izquierda a derecha y no de arriba a abajo.

En la figura 15 ilustramos esa limitación. Como podemos comprobar, la gráfica de barras que allí aparece nos representa la evolución de las ventas mensuales de ACME como un proceso que se desarrolla de arriba a bajo y no de izquierda a derecha como estamos acostumbrados en nuestro contexto cultural.

En cambio, las gráficas de puntos (sin combinación con líneas) no funcionan bien a la hora de representar este contenido semántico por otro motivo diferente.

Figura 15. Ventas mensuales de la empresa ACME durante el año pasado



Lo hacen porque al no estar los puntos conectados mediante líneas, el resultado son unos puntos aislados que, aunque codifican de forma correcta cada valor cuantitativo asociado a un momento de tiempo, no nos permiten entender que se esté representando un proceso temporal donde el punto es sólo uno de sus momentos. Sin las líneas, se nos escamotea esa conexión que une en el tiempo a todos los puntos y que nos ofrece la imagen de continuidad temporal.

En la figura 16 podemos observar esta dificultad. En esa gráfica de puntos, cada figura circular nos codifica el valor de las ventas de la empresa ACME (respecto al eje Y) en un mes concreto del año pasado (respecto al eje X). Sin embargo, al no conectarnos entre sí esas ventas mediante una línea, nos cuesta mucho percibir con sólo los puntos (sin líneas) esa evolución de las ventas a lo largo de ese año.

En cierta manera, tras esta limitación se oculta un efecto muy similar al que nos producían aquellos pasatiempos que antiguamente incorporaban lo periódicos y en los que se tenía que descubrir qué imagen se encontraba dibujada en una zona de la página uniendo con un lápiz una serie de puntos numerados. Solo con los puntos, no podías saber cuál era la imagen. Hasta que no uníamos esos puntos de forma ordenada mediante una línea, el dibujo que escondía no se nos hacía visible.

3.3. Visualización de la relación de ranking

Anteriormente pudimos ver que la relación de ranking asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se señala que esta asignación de valores numéricos establece una ordenación (de mayor a menor o de menor a mayor) entre los ítems de esa variable cualitativa.

Teniendo en cuenta esto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifique simul-

Figura 16. Ventas mensuales de la empresa ACME durante el año pasado

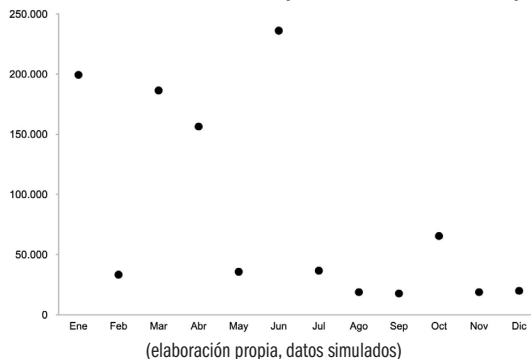
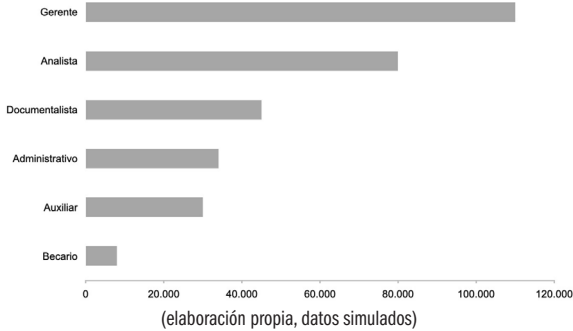


Figura 17. Sueldo por cargo en la empresa ACME

táneamente el valor numérico y la categoría con la que se encuentra asociado y que muestren ese orden intrínseco entre las categorías. Para que estos objetos puedan codificar estas duplas semánticas de forma adecuada, en uno de los ejes debe ubicarse esa variable cualitativa respetando su ordenación interna y, en el otro, la escala cuantitativa.

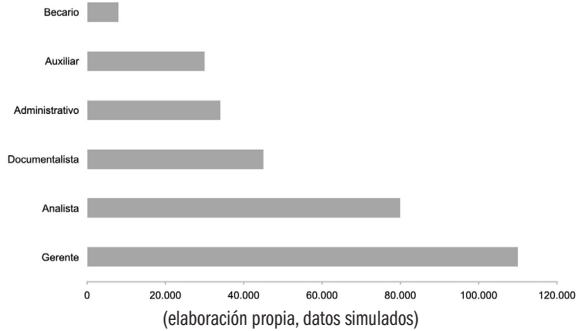
A la luz de estas condiciones, los objetos visuales que mejor codifican esta relación son las barras, las columnas y los puntos. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo) mostrando visualmente el orden interno establecido entre los ítems de la variable cualitativa.

En una gráfica de barras, cada figura rectangular nos codifica un valor cuantitativo (respecto al eje X) asociado a ítem cualitativo (respecto al eje Y) y nos muestra también la ordenación intrínseca que subyace entre esos ítems por la posición que ocupan las barras a lo largo del eje Y.

El orden o ranking que puede representar una gráfica de barras puede ser descendente o ascendente. En el descendente, la barra que representa el valor cuantitativo más alto se ubica en la parte superior del eje Y y el resto se ordenan de mayor a menor en ese eje utilizando como criterio el valor de la parte cuantitativa de la dupla representada. Como se muestra en la figura 17, donde aparece una gráfica de barras que representa el ranking descendente de los sueldos por cargo que se cobran en una empresa. El gerente es la persona que más cobra, el becario el que menos y el resto de los cargos se ordenan entre esos dos valores en función del salario.

En el ranking ascendente, la barra que representa el valor cuantitativo más bajo se ubica en la parte superior del eje Y y el resto se ordenan de menor a mayor en ese eje utilizando como criterio el valor de la parte cuantitativa de la dupla representada. Como se muestra en la figura 18, donde aparece una gráfica de barras que representa el ranking ascendente de los sueldos por cargo que se cobran en una

Figura 18. Sueldo por cargo en la empresa ACME

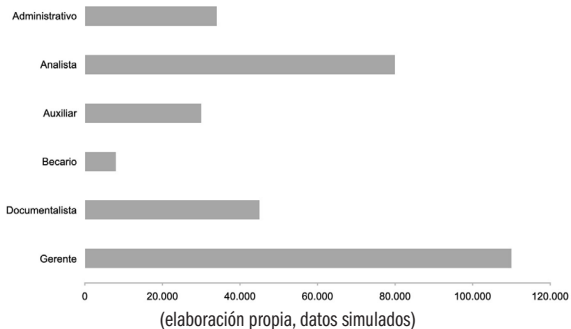


empresa. El becario es la persona que menos cobra, el gerente el que más y el resto de los cargos se ordenan entre esos dos valores en función del salario.

Es importante señalar que si no respetamos el orden (ascendente o descendente) entre las categorías (derivado de su valor cuantitativo asociado) y las ubicamos en el eje Y siguiendo otro criterio distinto, el resultado no es una gráfica que nos represente una relación de ranking, sino una visualización que nos codifica una mera comparación nominal. En la gráfica recogida en la figura 19 hemos construido con los mismos datos una gráfica de barras, pero para ordenar las barras no hemos utilizado como criterio el valor cuantitativo asociado, sino el orden alfabético de las categorías cualitativas. El resultado es una gráfica que codifica correctamente la relación de comparación nominal, pero que no representa la relación de ranking.

En una gráfica de columnas ocurre algo muy parecido. Cada figura rectangular nos codifica un valor cuantitativo (respecto al eje Y) asociado a ítem cualitativo (respecto al eje X) y nos muestra también la ordenación intrínseca que subyace entre esos ítems por la posición que ocupan las barras a lo largo del eje X.

Figura 19. Sueldo por cargo en la empresa ACME



Como ocurría con las gráficas de barras, el orden o ranking que puede representar una gráfica de columnas puede ser también descendente o ascendente. En el descendente, la columna que representa el valor cuantitativo más alto se ubica en la parte izquierda del eje X y el resto se ordenan de mayor a menor en ese eje utilizando como criterio el valor de la parte cuantitativa de la dupla representada. Como se muestra en la figura 20, donde aparece una gráfica de columnas que representa el ranking descendente de los sueldos por cargo que se cobran en una empresa. El gerente es la persona que más cobra, el becario el que menos y el resto de los cargos se ordenan entre esos dos valores en función del salario.

En el ranking ascendente, la columna que representa el valor cuantitativo más bajo se ubica en la parte izquierda del eje X y el resto se ordenan de menor a mayor en ese eje utilizando como criterio el valor de la parte cuantitativa de la dupla representada. Como se muestra en la figura 21, donde aparece una gráfica de columnas que representa el ranking ascendente de los sueldos por cargo que se cobran en una empresa. El becario es la persona que menos cobra, el gerente el que más y el resto de los cargos se ordenan entre esos dos valores en función del salario.

Figura 20. Sueldo por cargo en la empresa ACME

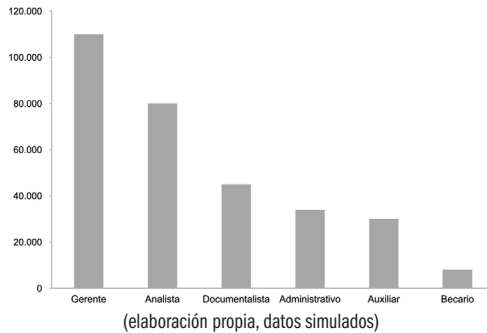
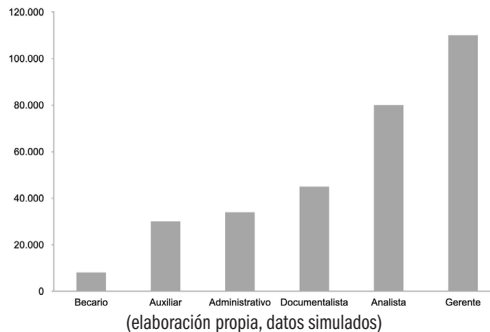


Figura 21. Sueldo por cargo en la empresa ACME



Igual que en el caso de las barras, es importante señalar que si no respetamos el orden (ascendente o descendente) entre las categorías (derivado de su valor cuantitativo asociado) y las ubicamos en el eje X siguiendo otro criterio distinto, el resultado no es una gráfica que nos represente una relación de ranking, sino una visualización que nos codifica una mera comparación nominal. En la gráfica recogida en la figura 22 hemos construido con los mismos datos una gráfica de columnas, pero para ordenar las columnas no hemos utilizado como criterio el valor cuantitativo asociado, sino el orden alfabético de las categorías cualitativas. El resultado es una gráfica que codifica correctamente la relación de comparación nominal, pero que no representa la relación de ranking.

Frente a la de barras o a la de columnas, la gráfica de puntos la utilizaremos en aquellos casos en los que la diferencia entre los valores cuantitativos que se representan sea muy escasa y queramos visualizar de forma cómoda la relación de ranking. O, dicho de otra manera: cuando queremos representar esa relación y el rango en el que fluctúan esos valores es pequeño, cuando existe un grado de dispersión baja entre ese conjunto de valores.

Como ya vimos cuando abordamos la visualización de la relación de comparación nominal, en esos casos, si representamos esos datos mediante una gráfica de columnas (o barras), puede costar apreciar la diferencia entre la longitud de esas figuras geométricas y dificultar, de esa manera, la comparación visual entre las mismas y los valores numéricos que codifican.

En esos contextos comunicativos, utilizar una gráfica de columnas (o barras) cuyo eje Y (o X, en el caso de la gráfica de barras) comience en un valor superior a cero para hacer evidentes las diferencias entre los valores, no es una alternativa. Ya que, como señalamos, las columnas (o barras), al codificar el valor cuantitativo mediante su longitud, necesitan comenzar en el punto que se corresponde con el

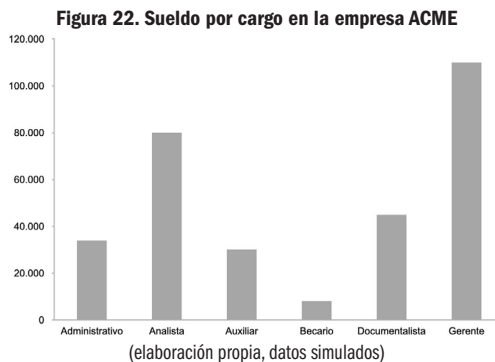
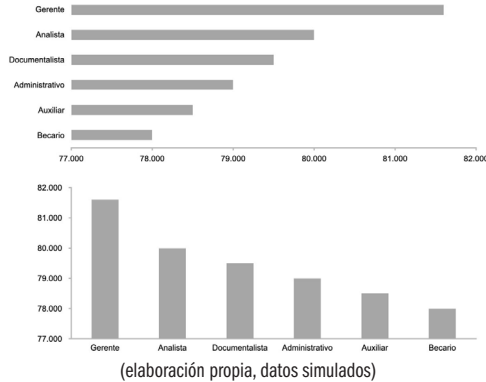


Figura 23. Sueldo por cargo en la empresa ACME



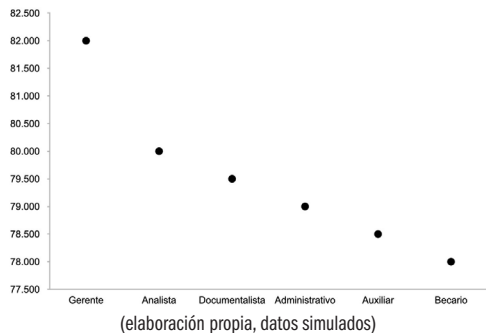
cero del eje vertical (u horizontal, en el caso de la gráfica de barras). Si no lo hacen, no codifican correctamente el valor numérico (como se aprecia en la figura 23).

Como en el caso de la visualización de la relación de comparación nominal, la alternativa correcta es utilizar una gráfica de puntos, ya que estas figuras circulares reciben su valor cuantitativo y cualitativo asociado a través de la posición que ocupan en el plano generado por los ejes X e Y. Y esta asignación de contenidos semánticos no se ve alterada si el eje Y comienza en un valor superior a cero (como ilustramos en la figura 24).

3.4. Visualización de la relación de parte-todo

Como ya señalamos, la relación de parte-todo asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa que, a su vez, conforman una entidad o un todo.

Figura 24. Sueldo por cargo en la empresa ACME



Teniendo en cuenta esto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente el valor numérico y la categoría con la que se encuentra asociado y que expresen también esa relación cuantitativa que mantienen cada una de las partes (su peso, por así decirlo) respecto al todo y poder facilitar así la comparación entre esos valores numéricos relacionados con esos ítems.

En muchos contextos comunicacionales, las propuestas visuales que habitualmente se utilizan para codificar esta relación son las gráficas circulares y las gráficas de barras (o columnas) apiladas.

En una gráfica circular, cada sector nos codifica un valor cuantitativo (en función del ángulo de los radios que lo limitan) asociado a un ítem cualitativo (determinado por el color que le asigna la leyenda) y nos muestra las relaciones cuantitativas que mantienen cada una de esas partes o sectores entre sí y con el todo.

En la figura 25 podemos identificar una gráfica circular que presenta un área o superficie circular dividida o formada por cinco sectores generados a su vez por dos radios y un arco. Como su título indica, estamos utilizando la gráfica para representar el peso en tiempo de cada una de las tareas que una persona concreta desarrolla a lo largo de una jornada.

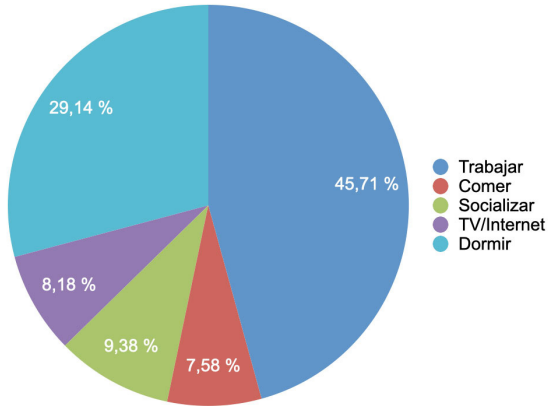
Pero como ya apuntamos, aunque su uso sea muy habitual, este tipo de gráficas suelen generar problemas cuando tratamos de comunicar con ellas la relación de parte-todo. Al representar el valor cualitativo mediante el ángulo que forman los radios que limitan el sector, estas propuestas visuales nos permiten codificar de una manera adecuada si una de las partes es más grande o pequeña que otra, pero los usuarios pueden encontrar dificultades para identificar el valor exacto cuantitativo que está representando cada uno de los sectores que conforman el círculo. Esto puede obligarnos a incluir los valores numéricos en cada uno de esos sectores para garantizar la correcta comunicación (como hacemos en la figura 25).

La otra propuesta habitual que se utiliza para representar la relación de parte-todo es la gráfica de barras (o columnas) apiladas. Estas gráficas representan mediante una barra o columna (una forma rectangular) un valor (un todo) que se obtiene a partir de la agregación de la representación, en forma de cuadrilátero, de los valores de sus partes. Igual que en el caso de la gráfica circular, nos muestra las relaciones cuantitativas que mantienen cada una de esas partes entre sí y con el todo.

En la figura 26 podemos identificar una gráfica de barras apiladas y otra de columnas también apiladas que representan el peso en tiempo (porcentaje) de cada una de las tareas que una persona concreta desarrolla a lo largo de una jornada.

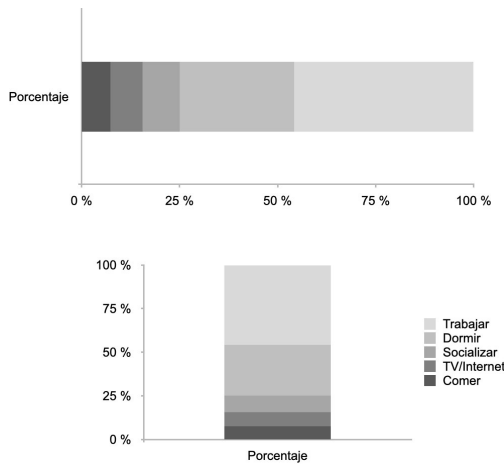
Sin embargo, aunque su uso sea muy común, igual que ocurría con las gráficas circulares, este tipo de gráficas también presentan problemas. Como ya apuntamos,

Figura 25. Actividades desarrolladas a lo largo de un día



(elaboración propia, datos simulados)

Figura 26. Actividades desarrolladas a lo largo de un día

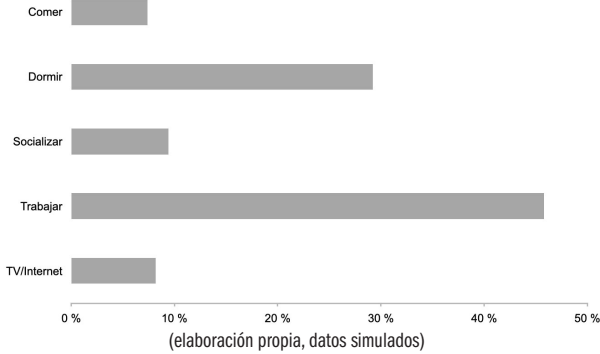


(elaboración propia, datos simulados)

el problema principal es que, si bien permiten comunicar visualmente de forma razonable cuándo el valor de una de las partes es mayor o menor que otro, el usuario puede encontrarse con dificultades a la hora de asignar el valor exacto cuantitativo que está representando el cuadrilátero que codifica cada parte. Como vemos en la figura 26, cada una de esas gráficas nos permite inferir de forma cómoda el valor exacto de la primera actividad (comer), pero no el del resto de las actividades.

Frente a estos dos tipos de gráfica que normalmente se utilizan para representar la relación de parte-todo pero que presentan las dificultades señaladas, los objetos visuales que mejor codifican esta relación son las barras y las colum-

Figura 27. Actividades desarrolladas a lo largo de un día

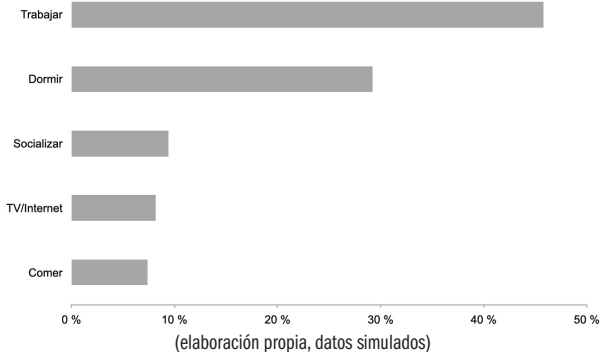


nas (no apiladas). Y lo hacen superando esos problemas en la representación de información.

Las gráficas de barras, por ejemplo, asociando a sus figuras rectangulares el valor cuantitativo que se corresponde con cada una de las partes del todo a partir de la propia longitud de esas figuras, nos permiten una forma cómoda de comparar numéricamente las partes entre sí y respecto al todo (como mostramos en la figura 27). Y nos facilitan esto sin tener que pensar en los grados que tiene un ángulo (como en el caso de la gráfica circular) o vernos obligados a calcular la longitud de un sector de una barra sin referencias visuales suficientes (como en el caso de la gráfica de barras apiladas).

Para mejorar su eficacia comunicativa, cuando representemos la relación parte-todo con una gráfica de barras, es interesante ordenar esas figuras que representan las partes en forma de ranking (descendente o ascendente) a partir del criterio de su valor numérico asociado (como hacemos en la figura 28).

Figura 28. Actividades desarrolladas a lo largo de un día



En el caso de las gráficas de columnas, las ventajas son las mismas. Asociando a sus figuras rectangulares el valor cuantitativo que se corresponde con cada una de las partes del todo a partir de la propia longitud de esas figuras, nos permiten una forma cómoda de comparar numéricamente (sin cálculos de ángulos o faltas de referencias visuales) las partes entre sí y respecto al todo (como mostramos en la figura 29).

De la misma forma, para mejorar su eficacia comunicativa, cuando representemos la relación parte todo con una gráfica de columnas, es mejor ordenar esos objetos visuales que representan las partes en forma de ranking (descendente o ascendente) a partir del criterio de su valor numérico asociado (como hacemos en la figura 30).

Para acabar este apartado es importante señalar que, frente a las barras y columnas, los elementos visuales que no utilizaremos para representar la relación de parte-todo son las líneas y los puntos (o la combinación de líneas y puntos).

Figura 29. Actividades desarrolladas a lo largo de un día

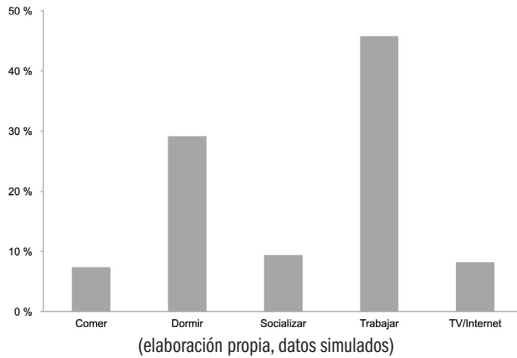
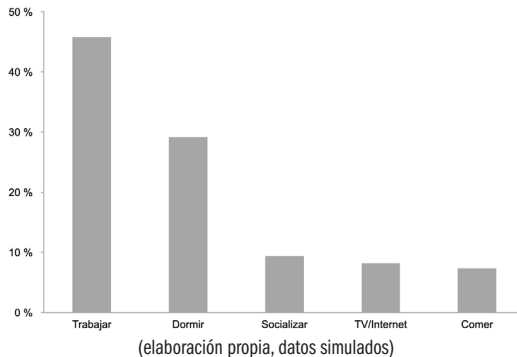


Figura 30. Actividades desarrolladas a lo largo de un día



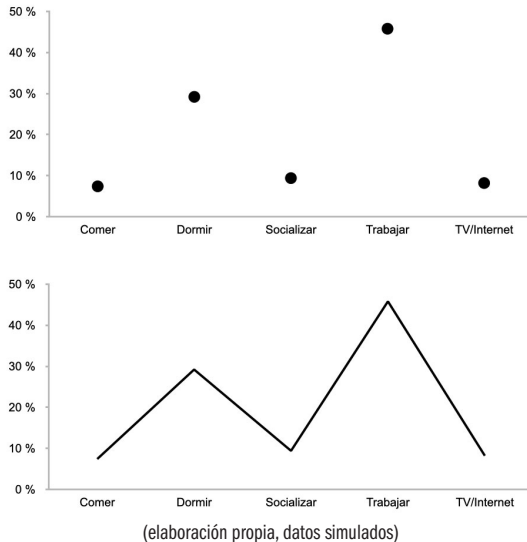
Las gráficas de puntos no funcionan bien a la hora de representar este contenido semántico ya que no poseen las dimensiones visuales suficientes para poder representar este tipo de relaciones. Como podemos ver en la figura 31, utilizando una gráfica de puntos no podemos intuir que ahí se está representado un todo formado por sus partes. Y algo parecido ocurre con la gráfica de líneas (o la gráfica de combinación de líneas y puntos). Como recogemos en la figura 31, ésta no nos permite representar adecuadamente esa relación al conectar mediante líneas cada una de las partes y hacernos entender, de forma errónea, que hay algún tipo de conexión u orden temporal entre las mismas.

3.5. Visualización de la relación de desviación

Como ya planteamos, la relación de desviación asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se comparan esos valores cuantitativos asociados respecto a un valor numérico concreto y de referencia.

Teniendo en cuenta esto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente el valor numérico y la categoría con la que se encuentra asociado y que permitan la comparación cuantitativa de esos datos numéricos respecto al valor de referencia.

Figura 31. Actividades desarrolladas a lo largo de un día



Contemplando esa restricción, los objetos visuales que mejor codifican esta relación son las barras y las columnas. Este tipo de elementos representan adecuadamente esas duplas (valor cuantitativo e ítem cualitativo) mostrando visualmente qué relación cuantitativa mantienen con el valor de referencia.

En una gráfica de barras (o columnas), cada figura rectangular nos codifica un valor cuantitativo (respecto al eje X en el caso de las barras, y respecto al eje Y en el caso de las columnas) asociado a ítem cualitativo (respecto al eje Y en el caso de las barras y respecto al eje X en el caso de las columnas), y, si representamos el valor de referencia mediante una línea o haciéndolo coincidir con uno de los ejes (con el eje Y en el caso de las barras y con el eje X en el caso de las columnas), nos expresa también la relación cuantitativa que mantienen esos valores numéricos respecto al de referencia.

Ilustremos todo esto a través de un ejemplo. En la figura 32 aparece una gráfica de columnas en la que se representan las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa ACME. Esta gráfica incluye también una línea negra que codifica el valor de referencia (las ventas de nuestra empresa).

Esta visualización combinada de columnas y línea expresa la relación cuantitativa que mantienen las ventas de cada empresa respecto a las de la nuestra y nos permite facilitar también la comparación entre esos valores numéricos relacionados con esos ítems. En este caso, el valor de referencia (las ventas de nuestra empresa) se identifica con el 100% y el resto de los valores (las ventas de la competencia) son expresados como una proporción o porcentaje respecto a ese valor primario o de referencia.

Frente a esta gráfica mixta, podemos ofrecer una versión alternativa sin línea, que incluya las columnas como únicos objetos visuales. En este caso, para no tener

Figura 32. Comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa

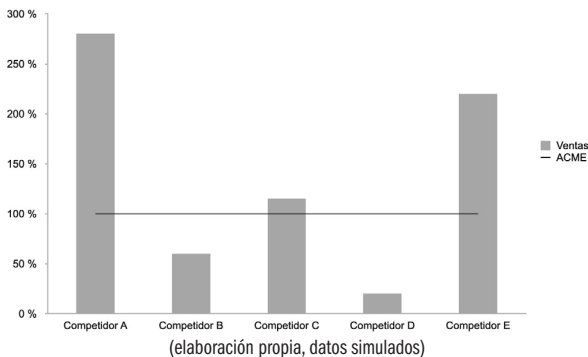
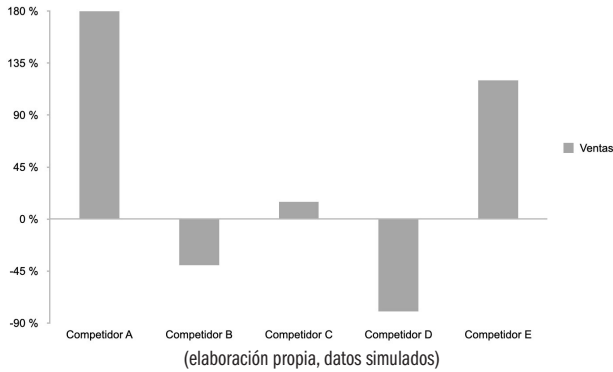


Figura 33. Comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa



que utilizar la línea, haremos coincidir el valor de referencia (las ventas de ACME) con el eje X. Como mostramos en la figura 33, si restamos 100 a cada uno de los valores porcentuales anteriores de las ventas de la competencia, podemos construir una gráfica de columnas que nos represente, también en clave porcentual, el porcentaje neto de ventas de la competencia respecto a nuestra empresa ACME (cuyo valor coincidirá visualmente con el eje X).

Es importante señalar que, aunque en ciertos contextos comunicacionales relacionados con la economía se utilicen de una forma bastante habitual, es recomendable no recurrir a las gráficas de líneas o de líneas y puntos para representar este contenido semántico. Si utilizamos este tipo de gráficas, corremos el riesgo de que el usuario pueda pensar que estamos visualizando la evolución de una variable cuantitativa y no una relación de desviación.

Y esto ocurriría tanto en la versión que incluye el valor de referencia como línea (figura 34) como la que lo hace coincidir con el eje X (figura 35). Al estar conectados por una línea, el usuario podría entender erróneamente que hay algún tipo de evolución o conexión entre el valor de las ventas del competidor A y el del competidor B, el del competidor B y el del competidor C, el del competidor C y el del competidor D y entre el valor de las ventas del competidor D y el del competidor E.

3.6. Visualización de la relación de distribución

Anteriormente definimos que la relación de distribución asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable de intervalo para facilitar la comparación entre esos valores numéricos asociados a esas unidades de intervalo.

Figura 34. Comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa

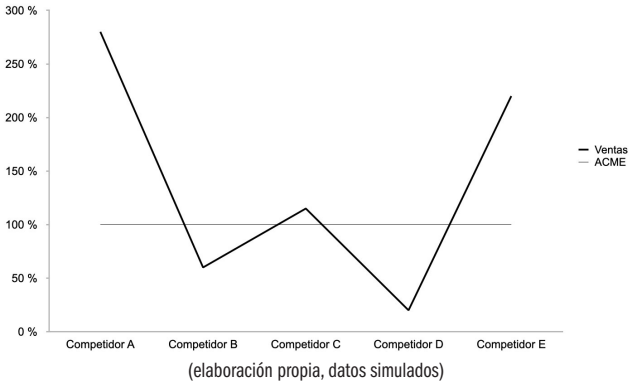
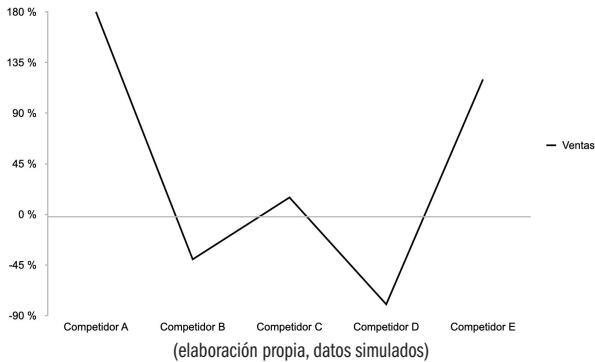


Figura 35. Comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa



Por tanto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente el valor numérico y la categoría con la que se encuentra asociado y que permitan expresar cómo esos valores cuantitativos se encuentran repartidos o diseminados a lo largo del rango del conjunto al cual pertenecen, a través de las categorías que conforman esa variable de intervalo.

Otra de las restricciones que debemos tener en cuenta es si queremos representar la distribución de un único conjunto de datos o, por el contrario, queremos visualizar simultáneamente, en una misma propuesta, la distribución de dos o más conjuntos de datos. O, dicho de otra manera, si queremos proponer una gráfica de distribución simple (muestra la distribución de un único conjunto de valores) o una gráfica de distribución múltiple (codifica la distribución de diferentes conjuntos de valores para poder compararlas).

Centrémonos en el primero de los casos. Cuando nuestro objetivo es representar la distribución de un único conjunto de valores, las gráficas de distribución simple que mejor codifican esta relación son el histograma, el polígono de frecuencia y el diagrama de caja.

En un histograma, cada columna nos codifica (respecto al eje Y) el número de ocurrencias asociado a uno de los ítems que conforman una variable de intervalos (respecto al eje X). Y nos permite poner el acento comunicativo en esos valores numéricos, por encima de una visión global de los valores distribuidos.

En la figura 36 recogemos la representación del número pacientes asociados a los ítems de una variable de intervalos (las franjas de edad de la población de unos pacientes que sufren episodios de epilepsia) a través de un histograma. Como podemos comprobar, las columnas, al no tener espacio que las separe, no son visualmente independientes. De hecho, estas conforman un área continua que nos permite entender de forma correcta que existe una conexión o continuidad en la evolución del número de pacientes de cada franja de edad a la siguiente y nos permite además representar adecuadamente esa evolución o distribución en función de la franja de edad en la que se encuentren los pacientes.

El polígono de frecuencia es otra alternativa para representar el mismo contenido semántico. Estrictamente hablando, aunque reciba este nombre, es una variación de la gráfica de líneas. En esta gráfica, cada vértice de la línea nos codifica (respecto al eje Y) el número de ocurrencias asociado a uno de los ítems que conforman una variable de intervalos (respecto al eje X). En esos casos, la línea permite obtener una visión conjunta y continua de los valores representados y de la serie completa. O, dicho en otros términos, esa línea codifica la continuidad entre esos valores del intervalo. Y nos permite poner el acento comunicativo en una visión global de los valores distribuidos, por encima de los valores numéricos individuales.

Figura 36. Número de pacientes por franja de edad en una población de enfermos de epilepsia

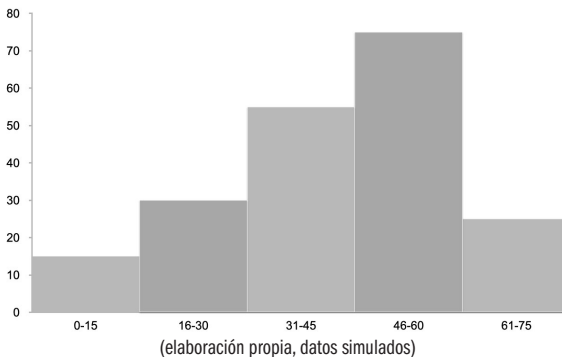
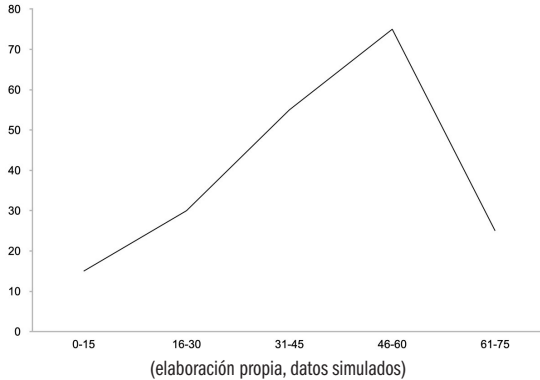


Figura 37. Número de pacientes por franja de edad en una población de enfermos de epilepsia

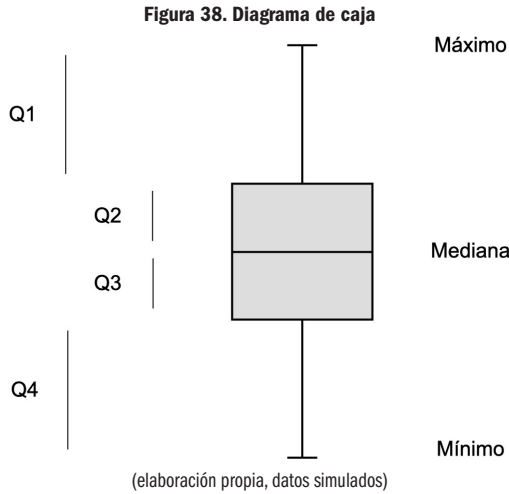
En la figura 37 recogemos la representación de unos valores cuantitativos (número de pacientes) asociados a los ítems de una variable de intervalos (las franjas de edad de una población de pacientes que sufren episodios de epilepsia) a través del polígono de frecuencia. La posición en el plano de los puntos que conforman la línea y la continuidad visual que imprime esa figura nos permite representar correctamente esa evolución del número de pacientes en función de la franja de edad en la que se encuentre el enfermo.

El diagrama de caja (*box plot* o *box-and-whisker plot*, en inglés) es otra buena alternativa para representar el mismo contenido semántico. Se trata de un gráfico, basado en cuartiles, mediante el cual se visualiza un conjunto de datos. En este contexto, los valores que determinan los cuartiles de un conjunto ordenado de datos deben entenderse como los tres puntos de corte que dividen ese conjunto de datos en cuatro grupos de la misma medida.

Como podemos comprobar en la figura 38, anatómicamente hablando, un diagrama de caja está compuesto por un rectángulo (la caja) y dos brazos (lo que se conoce como “bigotes”). Es un gráfico que suministra información sobre los valores mínimo y máximo del conjunto representado (los extremos de los bigotes, respectivamente).

También nos representa los límites superiores de cada uno de los cuartiles. El límite superior del cuartil Q_4 sería la mediana del intervalo que va del valor mínimo al valor de la mediana del intervalo total o rango. El límite superior del cuartil Q_3 se identificaría como la mediana del intervalo total del conjunto de valores. El límite superior del cuartil Q_2 coincidiría con la mediana del intervalo que va del valor de la mediana del intervalo total o rango al valor máximo.

Por último, este tipo de propuesta visual nos permite codificar también la posible existencia de valores atípicos (que destacan por no ajustarse a la distribución

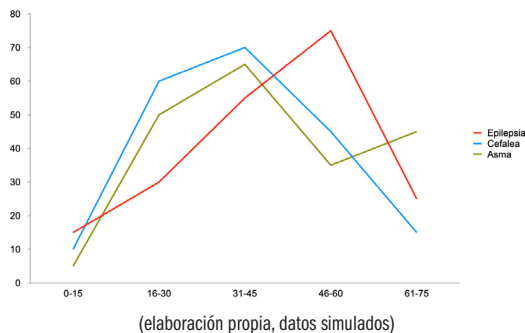


del resto de los valores representados y que se encuentran fuera de los límites inferior y superior) y la simetría o no de la distribución.

Abordemos ahora el segundo de los casos. Cuando nuestro objetivo es representar simultáneamente la distribución de dos o más conjuntos de valores, las gráficas de distribución múltiple que mejor codifican esta relación son la articulación de diferentes polígonos de frecuencia en una misma gráfica y la combinación de distintos diagramas de caja en la misma propuesta visual.

En la figura 39 representamos, mediante el uso de diferentes polígonos de frecuencia en una misma gráfica, el número de pacientes por franja de edad de unas poblaciones de enfermos de epilepsia, cefalea y asma. Cada línea codifica la distribución por franja de edad de cada una de esas poblaciones. Se acompaña de

Figura 39. Número de pacientes por franja de edad en unas poblaciones de enfermos de epilepsia, cefalea y asma



la leyenda para que el usuario pueda identificar con qué conjuntos de datos se corresponde cada una de las líneas. Para garantizar su correcta lectura y evitar el solapamiento visual es recomendable no representar simultáneamente más de cinco conjuntos de datos.

En la figura 40 representamos, mediante el uso combinado de distintos diagramas de caja en la misma propuesta visual, la distribución del aprovechamiento de carburante (MPG: millas por galón en ciudad) por tipo de coche (SUV (del inglés “Sport Utility Vehicle”), sedán, deportivo, familiar, furgoneta e híbrido).

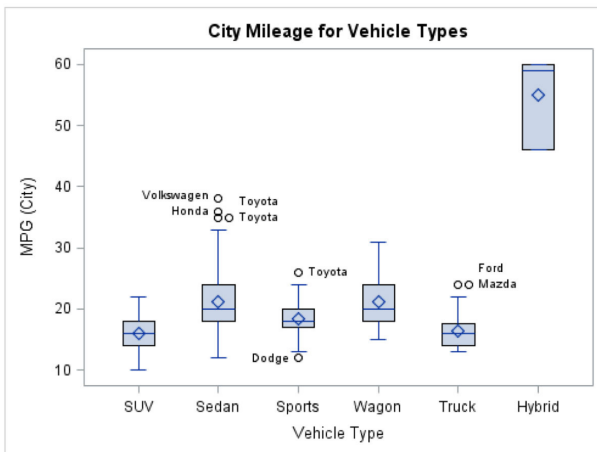
Cada gráfica de cajas nos codifica la distribución de los modelos de una clase especial de vehículo (deportivo, por ejemplo) que las marcas incluyen en su catálogo utilizando como criterio para ordenar esa distribución la distancia (en millas en un contexto urbano) que pueden recorrer esos coches a partir de la misma unidad de combustible (un galón, que equivale a 3,785 litros). Como vemos, los vehículos híbridos son los que más distancia recorren con la misma cantidad de carburante.

En este caso, y frente a los polígonos de frecuencia, como la posibilidad de que se produzca solapamiento entre los diagramas es muy escasa, se recomienda el uso de esta propuesta visual para la codificación simultánea de más de 5 distribuciones.

3.7. Visualización de la relación correlación

Abordemos ahora la visualización del último de los contenidos semánticos. Como ya vimos anteriormente, una relación de correlación asigna o conecta dos valores cuantitativos (cada uno perteneciente a una variable cuantitativa distinta) a

Figura 40. Millas por galón recorridas en circuito urbano por cada uno de los tipos de vehículo de los diferentes fabricantes



(https://support.sas.com/md/datavisualization/gt/boxplot_sect2.htm)

cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la identificación de algún patrón (fuerte o débil) en la variación entre esos dos tipos de valores numéricos relacionados con esos ítems.

Por tanto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifiquen simultáneamente los dos valores numéricos y la categoría con la que se encuentran asociados y que permitan expresar si existe o no una situación en los que los valores de una de las dos variables cuantitativas se modifican o cambian de manera sistemática con respecto a los valores de la otra.

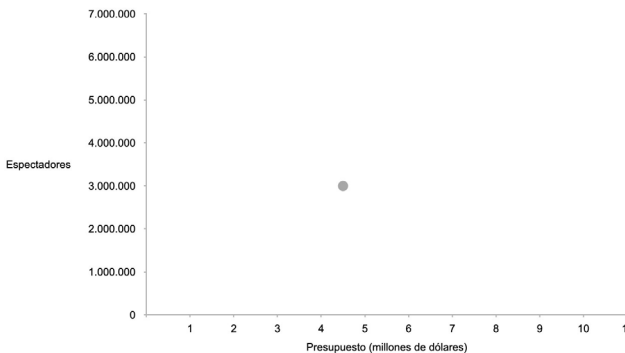
La propuesta visual que habitualmente se utiliza en la mayoría de los contextos comunicativos para representar este tipo de contenido semántico es el gráfico de dispersión (*scatter plot*, en inglés), que puede ser entendido como una variante particular de la gráfica de puntos.

Como el objetivo comunicativo del uso de esta propuesta visual es mostrar la relación entre los valores cuantitativos de ambos conjuntos y no entre las categorías a las cuales se asignan, en este tipo de gráficas los ejes X e Y presentan escalas cuantitativas mostrando cada punto de la gráfica como la intersección de un par de valores numéricos (suministrados por esos ejes X e Y).

En la figura 41 recogemos la génesis de un gráfico de dispersión. Imaginemos que queremos visualizar si existe o no (y si existe, cómo es de fuerte o débil), una posible correlación entre el presupuesto de las películas estrenadas la última semana y el número de espectadores que terminan consumiendo esas películas.

Para representar esa correlación, primero se codifica con un punto una película concreta, la película A. Ese punto que codifica la película A se ubica en el plano generado por los ejes X e Y a partir de los dos valores cuantitativos (presupuesto y

Figura 41. Presupuesto y espectadores de una de las películas estrenadas la última semana



(elaboración propia, datos simulados)

espectadores) asociados a la misma: sus 4,5 millones de dólares de presupuesto y los 3.000.000 espectadores que han disfrutado de esa obra.

El gráfico de dispersión se completa representando con puntos el resto de películas a partir de sus presupuestos y espectadores (figura 42). El resultado es una gráfica de puntos que codifica el presupuesto y los espectadores de las películas estrenadas la última semana.

La gráfica de dispersión suele culminarse incorporando la línea de tendencia. Se trata de una línea que nos calcula el propio software de visualización a partir de los datos cuantitativos representados (figura 43), y que nos marca si existe o no algún tipo de correlación (si existe una línea que va de abajo a arriba o de arriba abajo), si es positiva o negativa (es positiva si la línea va de abajo a arriba; es negativa si la línea va de arriba abajo) y si es fuerte o débil (dependiendo del ángulo que forma esa línea al cruzarse con el eje Y).

Algunos autores defienden que, principalmente en el contexto empresarial, el gráfico de dispersión presenta dificultades de interpretación entre los usuarios que no están familiarizados con este tipo de propuestas visuales (Few, 2012). Como

Figura 42. Presupuesto y espectadores de las películas estrenadas la última semana

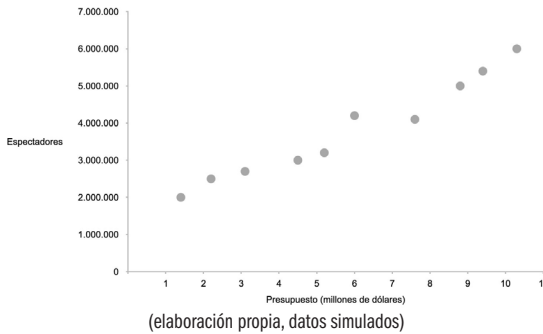
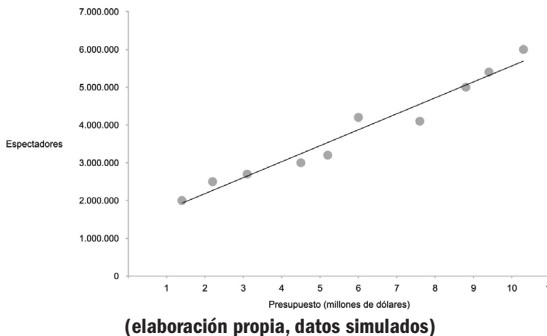


Figura 43. Presupuesto y espectadores de las películas estrenadas la última semana



alternativas más intuitivas, cuando estamos tratando de visualizar conjuntos de datos no muy extensos, es posible utilizar una gráfica de columnas de correlación o una gráfica de barras emparejadas.

Una gráfica de columnas de correlación (*correlation bar graph*, en inglés) es una propuesta visual en la que, dentro del espacio generado por un eje X y dos ejes Y (Y e Y'), se ubican dos conjuntos de columnas dispuestas en la misma dirección, unas embebidas dentro de las otras, y que representan, respectivamente, los valores de las dos variables cuantitativas cuya correlación queremos analizar (figura 44).

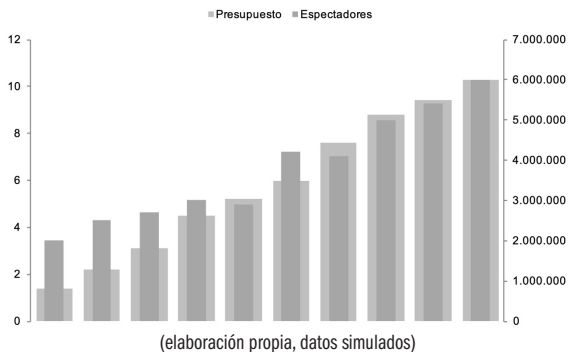
Como vemos en el ejemplo, se decide que las columnas que representan una de las variables cuantitativas (número de espectadores de la película) aparezcan dibujadas o embebidas dentro de las columnas que representan la otra variable cuantitativa (presupuesto de la película). En este sentido, cada película está representada por una pareja de columnas, una dentro de la otra.

Cada conjunto de columnas recibe respectivamente su valor cuantitativo respecto a los ejes Y e Y': las columnas que representan el presupuesto respecto al eje Y y las que codifican el número de espectadores respecto al eje Y'. Y ordenamos de izquierda a derecha cada par de columnas que representan una película utilizando como criterio el valor ascendente de sus presupuestos.

Incluyendo una leyenda para que se interprete bien cada columna, la gráfica nos permite identificar y comparar visualmente los valores de las dos variables cuantitativas y apreciar que existe algún grado de correlación positiva entre el presupuesto y los espectadores de las películas estrenadas la semana pasada. Y se aprecia esa correlación ya que cuando crece el presupuesto, crece también el número de espectadores.

Esta visualización se puede completar añadiendo, si el software lo permite, las líneas de tendencia de cada uno de los conjuntos de valores representados (figura 45).

Figura 44. Presupuesto y espectadores de las películas estrenadas la última semana



Estas líneas, al evolucionar con direcciones similares, nos muestran también que existe cierta tendencia entre el incremento de presupuesto y el incremento de espectadores

Por último, la gráfica de barras emparejadas es también una alternativa al diagrama de dispersión cuando analizamos visualmente conjuntos de datos no muy extensos. Esta gráfica utiliza dos escalas y dos conjuntos de barras dispuestas en direcciones opuestas para codificar respectivamente cada conjunto de valores cuantitativos (figura 46).

Como vemos en el ejemplo, cada película se representa a partir de dos barras enfrentadas que codifican respectivamente su presupuesto y su número de espectadores. Esas parejas se ordenan de forma descendente utilizando como criterio el valor de su presupuesto. La gráfica nos permite identificar y comparar visualmente los valores de las dos variables cuantitativas y apreciar que existe algún grado de correlación positiva entre el presupuesto y los espectadores de las películas estrenadas la semana pasada. Y se aprecia esa correlación ya que cuando decrece el presupuesto, disminuye también el número de espectadores.

Esta visualización se puede completar añadiendo las líneas de tendencia de cada uno de los conjuntos de valores representados (figura 47). Estas líneas, al evo-

Figura 45. Presupuesto y espectadores de las películas estrenadas la última semana

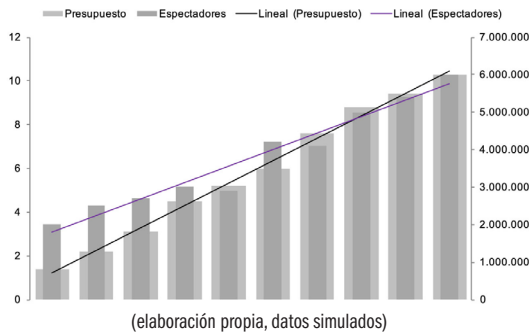


Figura 46. Presupuesto y espectadores de las películas estrenadas la última semana

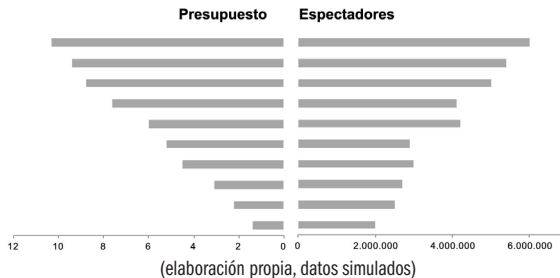
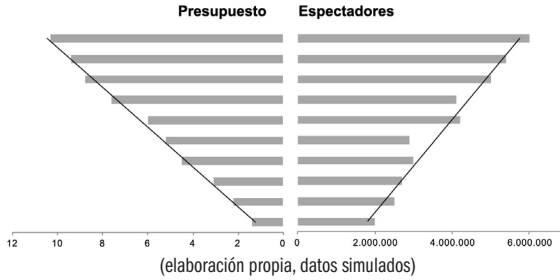


Figura 47. Presupuesto y espectadores de las películas estrenadas la última semana



lucionar con direcciones convergentes, nos muestran también que existe cierta tendencia entre el incremento de presupuesto y el incremento de espectadores.

4. Diseño visual de gráficas

A lo largo de este capítulo hemos intentado cubrir varios objetivos. Primero, en el apartado 2, hemos abordado los principales tipos básicos de información o contenidos semánticos que podemos representar en una gráfica, ilustrándolos a partir de ejemplos que codificamos en forma de tabla para que pudiesen ser entendidos de una forma más intuitiva. A continuación, en el apartado 3, hemos examinado cuál es el tipo de visualización que se adecua de forma más idónea para comunicar cada uno de esos tipos de contenidos semánticos. Ahora, por último, en este apartado, vamos a revisar una serie de aspectos relacionados con el diseño eficiente de los principales elementos que conforman una gráfica (Robbins, 201; Chen, 1999; Tondreau, 2009; Wong, 2010; Card, 2003; Fekete et al., 2008; North, 2005; van Wijk, 2006).

En definitiva, vamos a intentar ofrecer una serie de heurísticos, de consejos o reglas estructurales, que nos van a ayudar a diseñar unas propuestas visuales que funcionen de forma más adecuada a la hora de comunicar y analizar la información que alberga un conjunto de datos (Kirk, 2012; Spence, 2014; Nussbaumer, 2015; Yau, 2011)

Y vamos a presentar esos heurísticos, consejos o reglas relacionándolos con los elementos de una gráfica que regulan. Concretamente, los vamos a introducir de la mano de algunos de los principales elementos codificadores de información y componentes de soporte que conforman una gráfica y que detallamos también el capítulo anterior: los puntos, las barras (o columnas), las líneas, las escalas, la leyenda, los ejes y la cuadrícula (Few, 2012; Miller, 1956; Meirelles, 2013; Cotgreave et al. 2017).

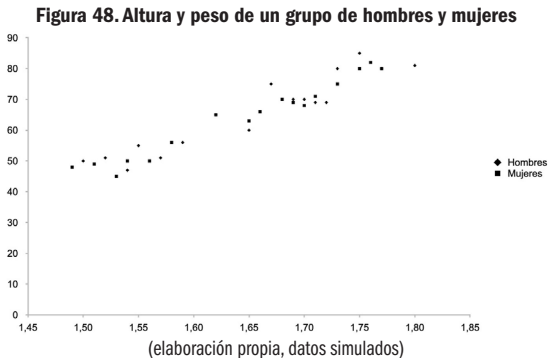
4.1. Puntos

Un punto es un objeto visual constituido por una forma (círculo, cuadrado, rectángulo, circunferencia o cruz, entre otras muchas) que, normalmente, dependiendo de su ubicación en el plano generado por los dos ejes X e Y, suele asociar dos valores cuantitativos entre sí o vincular un valor cuantitativo con uno de los ítems (valores o categorías) de una variable cualitativa.

Cuando intentamos representar ese tipo de información utilizando una gráfica de puntos, normalmente nos enfrentamos a dos dificultades importantes. Por un lado, los propios puntos pueden solaparse entre ellos de manera que sea muy difícil discriminarlos visualmente. Y, por otro, los puntos, cuando se combinan con líneas, pueden aparecer escondidos tras éstas de manera que no podamos identificarlos con claridad. Abordemos qué soluciones podemos aplicar para resolver estas dificultades.

Veamos la primera de las dificultades a través de un ejemplo. En la figura 48 recogemos una gráfica de puntos que se solapan entre ellos y, al hacerlo, dificultan su correcta visualización.

Concretamente, con cada punto estamos representando la altura y el peso de una persona concreta. Dependiendo de su sexo, si la persona es un hombre se representa con un punto en forma de rombo; si es una mujer, se visualiza mediante un cuadrado. En esa gráfica hay dos parejas de puntos que se solapan totalmente entre ellos al encontrarse (cada pareja) situados en las mismas coordenadas respecto a los ejes X e Y; y, que, por tanto, es imposible discriminarlos entre sí. En una de las parejas que se solapan, en el mismo espacio del plano, coinciden de forma exacta dos puntos que representan respectivamente a un hombre y a una mujer. Esas dos personas, al tener la misma altura (1,62 metros) y el mismo peso (65 kilogramos), están representadas por unos puntos que ocupan el mismo lugar en la gráfica y no permiten su discriminación visual. Lo mismo ocurre con la otra



pareja que se solapan: representan dos personas de la misma altura (1,77 metros) y el mismo peso (80 kilogramos) y, por tanto, esos puntos que los representan coinciden exactamente en el mismo lugar en la gráfica y no permiten su correcta identificación visual.

Para resolver este problema se pueden proponer diversas soluciones. Por ejemplo, podemos agrandar (en una o dos dimensiones) la gráfica para ver más separados los puntos. O, también, con el mismo fin, podemos disminuir el tamaño de los puntos o los colores de estos. Pero, desgraciadamente, el problema de los puntos que se solapan de forma exacta continua.

Para solucionarlo, lo mejor es utilizar puntos sin relleno (no sólidos, sólo dibujados con las líneas de contorno) y, de forma simultánea, proponer formas y colores diferentes para cada tipo de puntos. Esta solución la podemos ver en la figura 49. Allí hemos utilizados puntos sin relleno para representar a las personas. Y para discriminar entre sexos hemos propuesto círculos azules para representar a los hombres y cruces rojas para codificar a las mujeres. Como podemos comprobar, con esta solución podemos discriminar de forma cómoda entre todos los puntos, incluso en el caso del solapamiento completo de los mismos.

Abordemos ahora la segunda de las dificultades. En muchas ocasiones, cuando los puntos aparecen combinados con líneas en una misma gráfica, éstos pueden aparecer escondidos tras esas líneas de manera que no podamos identificarlos con claridad. Esta dificultad la podemos ilustrar con la figura 50.

Como vemos, en la gráfica que se incluye, se intenta representar las ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos. Y para hacerlo se codifican las ventas de cada región mediante una línea que se discrimina del resto por la forma de los puntos que une. Pero en esta ocasión, al ser el tamaño de los puntos inferior a la anchura de la línea que los conecta, es imposible identificar cuál es la línea que representa a cada una de las ventas regionales.

Figura 49. Altura y peso de un grupo de hombres y mujeres

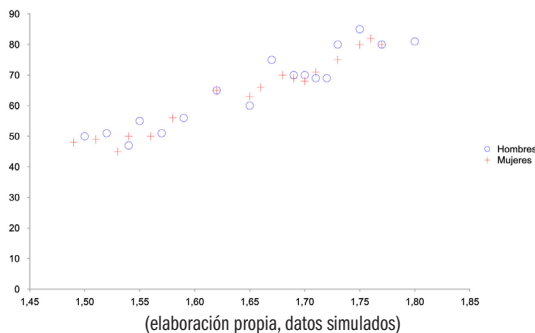
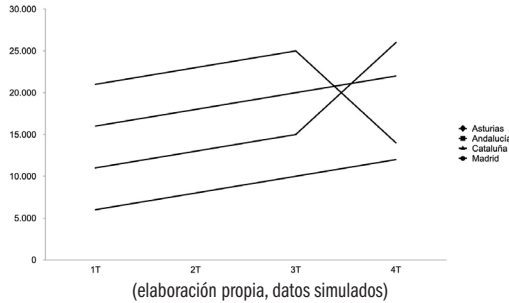


Figura 50. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos



Para solucionar esta dificultad, la solución visual más directa es aumentar el tamaño de los puntos y disminuir el tamaño de las líneas (como proponemos en la figura 51). De esta manera podemos distinguir de forma cómoda cada una de las líneas a partir de la forma distinta que presentan los puntos que éstas conectan.

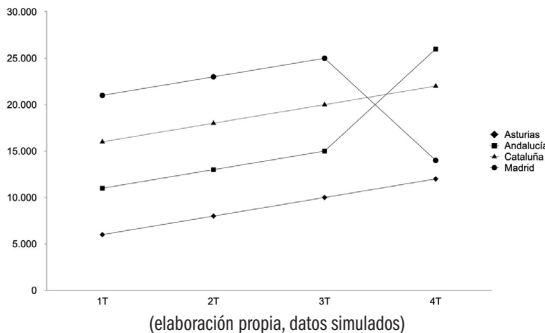
4.2. Barras (o columnas)

Cómo ya abordamos anteriormente, una barra (o columna) es un objeto visual constituido por un rectángulo que, normalmente, dependiendo de su ubicación en el plano generado por los dos ejes X e Y, suele asociar un valor cuantitativo con uno de los ítems (valores o categorías) de una variable cualitativa.

Respecto a estos objetos visuales, las principales recomendaciones respecto a su diseño visual las encontramos relacionadas con su orientación y con el punto de ubicación de su base.

Ya vimos que, teniendo en cuenta su orientación, estas figuras geométricas pueden estar dispuestas de dos maneras diferentes. En vertical formando una gráfica de columnas; o en horizontal, como parte de una gráfica de barras.

Figura 51. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos



En los contextos donde no existan otras restricciones, la disposición horizontal en forma de barras la utilizaremos cuando las etiquetas de los ítems de la variable cualitativa que está representando sean excesivamente extensos y no puedan aparecer en la gráfica de forma que faciliten una lectura natural de las mismas por parte del usuario.

En la figura 52 recogemos una gráfica de barras en la que se representan los sueldos por cargo en el seno de una empresa.

Como vemos, hemos dispuesto los objetos visuales rectangulares de forma horizontal, como barras, porque las etiquetas de los ítems de la variable cualitativa (los cargos) son excesivamente extensos. Y al colocarlos de esta manera, esas etiquetas se disponen también de forma horizontal habilitando una lectura cómoda y natural de las mismas por parte del usuario.

Si no lo hubiésemos hecho de esta manera y hubiéramos orientado esos objetos visuales rectangulares de forma vertical, como columnas, esas etiquetas hubieran aparecido dispuestas de tal manera que el usuario de la misma se hubiera visto obligado a invertir mayor esfuerzo cognitivo en su lectura.

Así, por ejemplo, si hubiéramos apostado por una gráfica de columnas, para que las etiquetas no se solaparan visualmente bajo el eje X, hubiéramos tenido que incluirlas de forma vertical provocando una lectura incómoda de abajo a arriba (figura 53), de arriba abajo (figura 54), en diagonal (figura 55) o de forma apilada (figura 56).

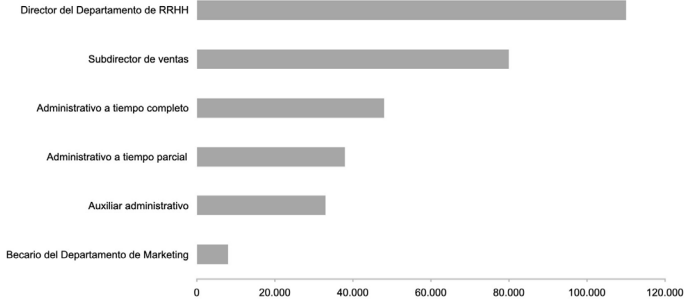
Pero ninguna de esas lecturas sería tan cómoda como la que podemos ofrecer si construimos esa gráfica utilizando barras y no columnas.

Otro de los aspectos importantes de diseño que debemos tener en cuenta respecto a estas figuras rectangulares, es el punto de ubicación de su base. Esas figuras rectangulares presentan dos extremos diferenciados: la cabeza (que forma el extremo o punta derecha en el caso de la barra o superior en el de la columna) y la base (que forma el extremo o punta izquierda en la barra o inferior en la columna). El punto de la cabeza es el que representa el valor cuantitativo asociado al ítem cualitativo.

Teniendo en cuenta esa distinción, es importante ubicar siempre la base de la barra o de la columna en el valor 0, junto a la línea del eje x (en el caso de las columnas) o junto a la línea del eje Y (en el caso de las barras). Porque, como vimos, la longitud de la barra o de la columna es el atributo visual que representa el valor cuantitativo que lleva asociada esa figura.

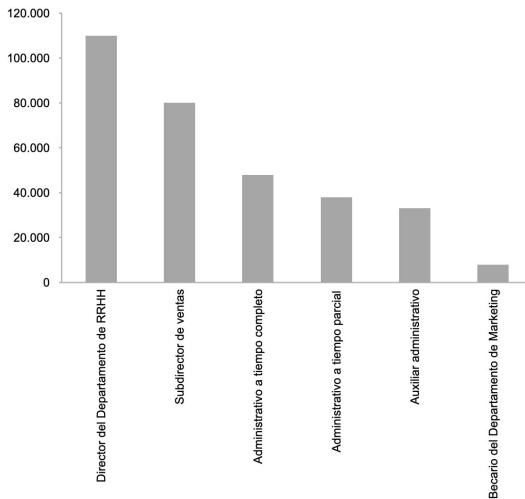
Y esto debe cumplirse incluso en el caso en el que se representen valores positivos y negativos, como en la gráfica recogida en la figura 57, donde representamos mediante una gráfica de columnas una comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa.

Figura 52. Sueldo por cargo en la empresa ACME



(elaboración propia, datos simulados)

Figura 53. Sueldo por cargo en la empresa ACME

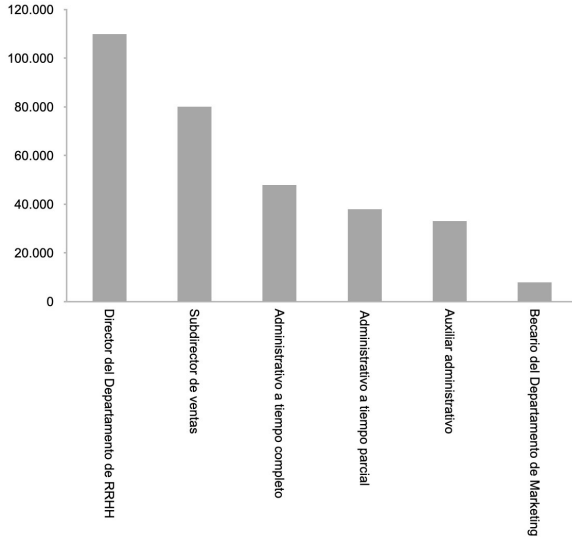


(elaboración propia, datos simulados)

Esa recomendación puede no seguirse en algunos casos. Como cuando queremos visualizar un rango de valores y no un único valor determinado mediante una gráfica de columnas. Como en la figura 58, donde representamos mediante gráfica de barras flotantes (o barras de rango) el rango de salarios por cargo en una empresa.

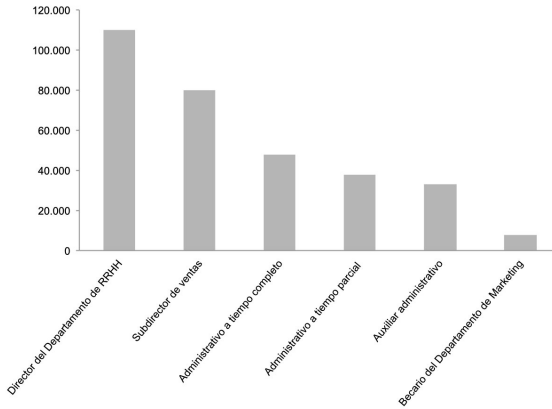
Este tipo de gráficas suelen completarse representando, además del valor mínimo y máximo del rango, algún dato de referencia, como, por ejemplo, la media. En la figura 59 se recoge una gráfica de columnas flotantes (o columnas de rango) que representa el rango de salarios por cargo en una empresa, completándose cada una de esas columnas mediante una marca que representa el sueldo medio dentro de cada uno de esos perfiles laborales.

Figura 54. Sueldo por cargo en la empresa ACME



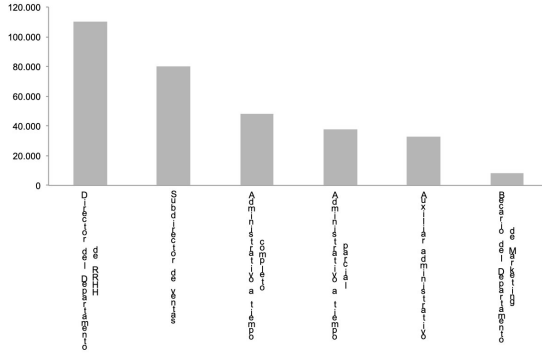
(elaboración propia, datos simulados)

Figura 55. Sueldo por cargo en la empresa ACME



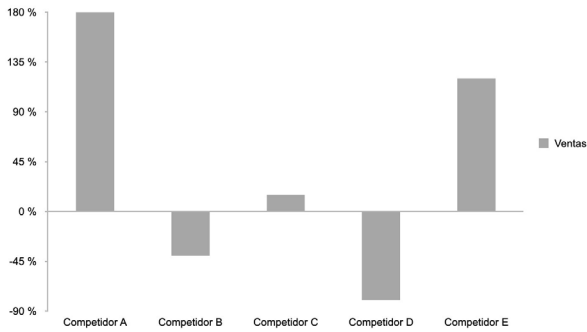
(elaboración propia, datos simulados)

Figura 56. Sueldo por cargo en la empresa ACME



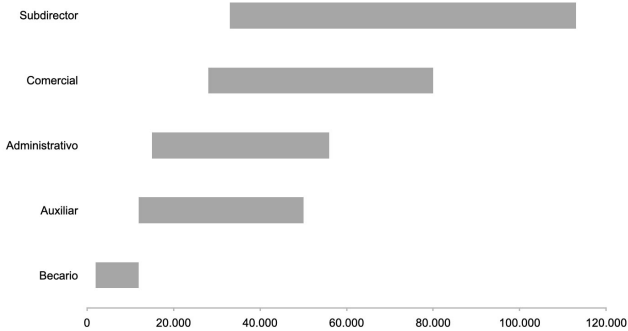
(elaboración propia, datos simulados)

Figura 57. Comparativa de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa



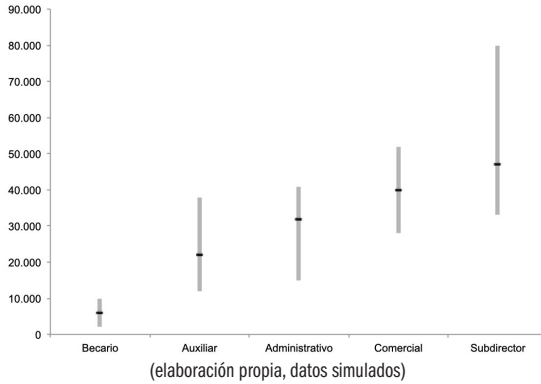
(elaboración propia, datos simulados)

Figura 58. Sueldo por cargo en la empresa ACME



(elaboración propia, datos simulados)

Figura 59. Sueldo por cargo en la empresa ACME



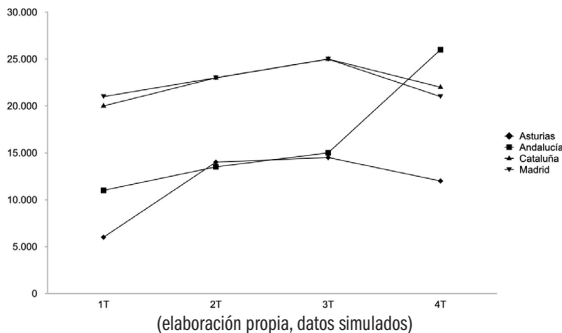
4.3. Líneas

En el contexto de la visualización de información, una línea estándar (no de referencia o de tendencia) debe ser considerada como un objeto formado por una sucesión continua de puntos que acostumbra a representar la evolución de los valores de una variable cuantitativa a lo largo de un intervalo temporal. Se suelen utilizar para enfatizar la continuidad y el flujo o evolución de un valor cuantitativo individual a otro.

El principal problema de diseño asociado a este tipo de objetos visuales aparece cuando tratamos de representar de forma simultánea y en una misma gráfica la evolución temporal de diversas variables cuantitativas y corremos el riesgo de que las líneas no se distingan bien al producirse algún grado de solapamiento entre ellas.

En la figura 60 recogemos una gráfica de líneas en la que se presenta ese problema. Allí se está representando la evolución de las ventas trimestrales en las cua-

Figura 60. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos



tro regiones donde ACME distribuye sus productos. Las líneas que codifican esas ventas se solapan totalmente en algunos tramos y en otros dificultan su discriminación visual.

Para solucionar este problema tenemos varias alternativas. En primer lugar, en contextos comunicativos en los que está habilitada la opción de publicar las gráficas en color, podemos optar por utilizar diferentes colores para cada línea, como recogemos en la figura 61.

En contextos en los que esté más limitada la publicación en color, podemos optar por utilizar diferentes intensidades de color para cada línea. Se trata de un método un poco menos efectivo, como recogemos en la figura 62.

En aquellos contextos donde la publicación está restringida al blanco y negro, podemos utilizar también la variación en la intensidad del color de las líneas o generar la gráfica con diferentes estilos o patrones de línea.

En la figura 63 incluimos esta última opción. En la gráfica representamos respectivamente las ventas trimestrales de cada región utilizando para ello una línea

Figura 61. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos

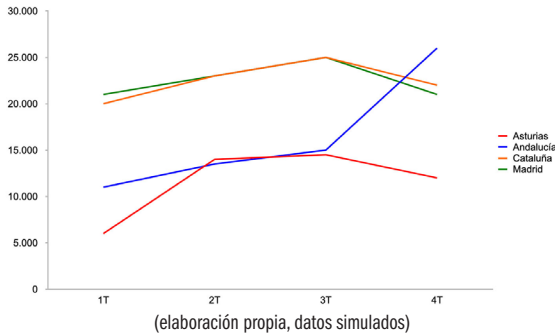


Figura 62. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos

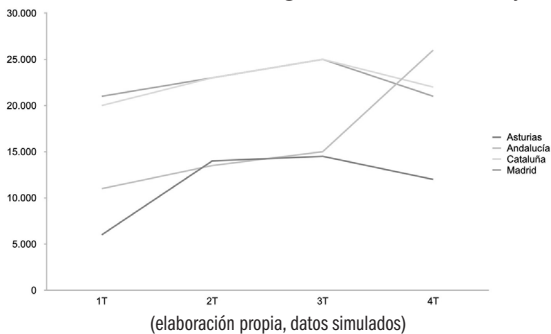
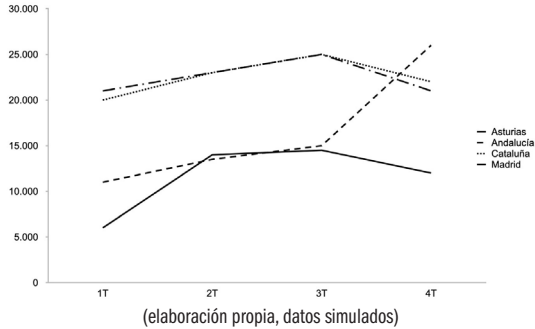


Figura 63. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos

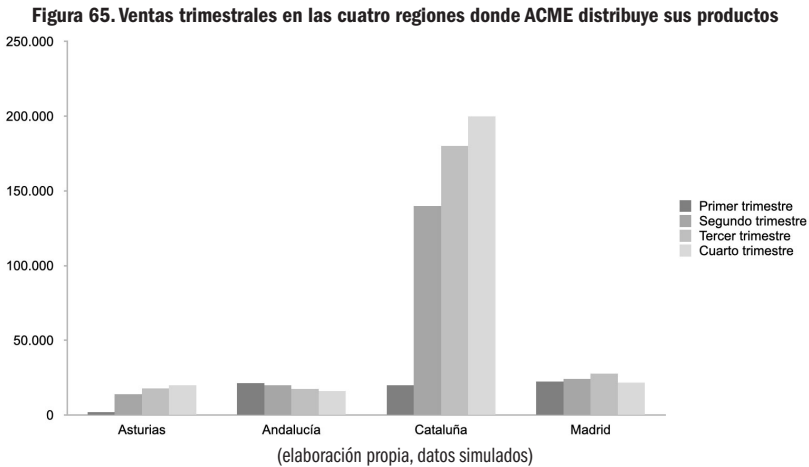
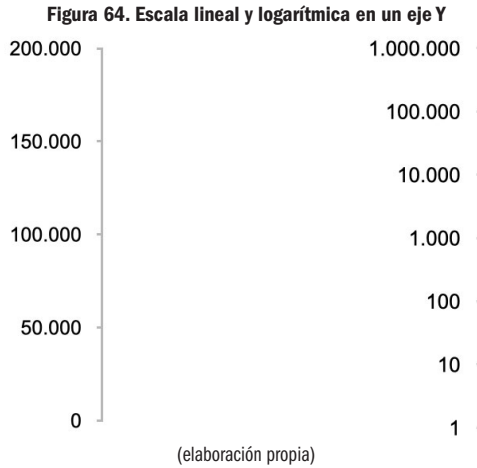
que presenta un patrón visual (línea sólida, punto y raya, guiones o puntos) distinto al de las demás. Se trata de la opción menos efectiva (y aconsejable) de las tres, ya que, en ciertos contextos, puede introducir confusión y generar problemas de lectura en sus potenciales usuarios.

4.4. Escalas

En el contexto de una gráfica, las escalas que estructuran los ejes nos permiten clasificar los objetos visuales que están representando la información. Concretamente, las escalas suelen dividir los ejes en secciones idénticas y asignar medidas cuantitativas o etiquetas categoriales a esas secciones. Dependiendo del lugar que ocupan en la gráfica respecto a esas escalas, a los objetos se les asigna una información cualitativa y cuantitativa. Sin esas escalas, las gráficas carecerían de ese significado semántico y no podrían cubrir sus funciones comunicacionales y de análisis de la información.

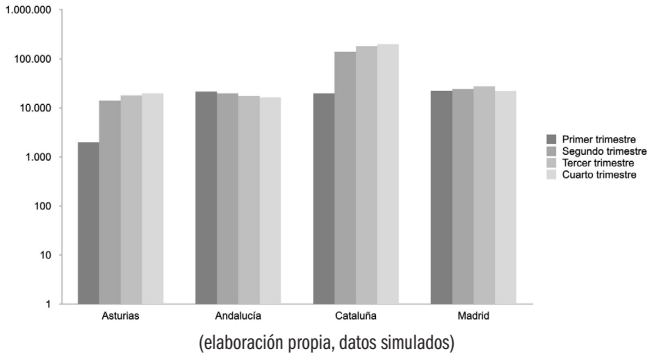
Si nos centramos en la dimensión cuantitativa, es posible distinguir dos tipos de escalas: la escala lineal y la escala logarítmica. En una escala lineal, el intervalo cuantitativo entre una marca de la escala y la siguiente es siempre el mismo. En cambio, en una escala logarítmica, el intervalo cuantitativo entre una marca de la escala y la siguiente es igual al valor del intervalo previo multiplicado por la base del logaritmo (normalmente base 10). El contraste entre esos dos tipos de escalas y sus marcas en el eje Y lo podemos ilustrar a través de la figura 64 (escala lineal a la izquierda y escala logarítmica (en base 10) a la derecha en la figura).

Es importante recomendar el uso de las escalas logarítmicas en gráficas que incluyen valores muy dispares, donde se represente un conjunto de datos con un rango muy extenso. En este tipo de gráficas, al convivir valores bajos y muy altos, se hace difícil representarlos con una escala lineal ya que, si proponemos marcas para poder representar los valores altos, perdemos las referencias visuales de esas marcas para poder interpretar los valores más bajos.



Un ejemplo de esa dificultad lo recogemos en la figura 65. Allí se incluye una gráfica de columnas donde se representan las ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos. En esta gráfica, al utilizarse una escala lineal en el eje Y e incluir los datos altos (las ventas de los trimestres segundo, tercero y cuarto en Cataluña), perdemos las referencias visuales para poder identificar los valores cuantitativos exactos del primero de los trimestres de zona geográfica y del resto de los trimestres de las otras regiones.

La solución es volver a crear la misma gráfica, pero sustituyendo la escala lineal en el eje Y por una logarítmica. Como vemos en la figura 66, con esa apuesta se consigue representar simultáneamente en un mismo espacio valores muy dispares (muy

Figura 66. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos

grandes y muy pequeños) sin perder las referencias visuales (a partir de las marcas del eje Y) que nos permitan identificar todos los valores cuantitativos asociados.

Las otras recomendaciones que queremos introducir respecto a las escalas se concentran sobre el número de marcas y sobre las etiquetas numéricas que éstas incorporan.

Sobre las marcas que debe incluir el eje que contiene la escala cuantitativa no podemos fijar o proponer un número exacto para que una gráfica funcione visualmente. El único consejo que podemos recomendar es que debemos incluir un número de marcas intermedio que cumpla la siguiente doble condición: que sea suficiente para facilitar la interpretación de la gráfica pero, a su vez, que no sobrecargue visualmente el eje y pueda abrumar al usuario en sus tareas de descodificación de la gráfica.

En las figuras 67 y 68 recogemos respectivamente un ejemplo en el que no se cumple una de las dos partes de la doble condición. En la gráfica incluida en la figura 67, aparece un número muy reducido de marcas (sólo dos) en la escala del eje Y que hacen muy difícil identificar, al faltarnos referencias visuales, el valor cuantitativo exacto asignado a cada una de las columnas. Y, en la gráfica incluida en la figura 68, aparece un número excesivo de marcas en la escala del eje Y que hacen difícil identificar, por sobrecarga visual, cuál es la marca concreta que nos puede ayudar a descodificar el valor cuantitativo exacto asignado a cada una de las columnas.

Respecto a las etiquetas que se incluyen en las escalas numéricas de una gráfica hay que procurar que incluyan de forma exclusiva números redondos (que terminen en 0 o en 5). Los utilizaremos en detrimento de otras etiquetas numéricas alternativas, porque el ser humano está acostumbrado a manejarse con este tipo de números (son más amigables, rápidos de leer y comprender) y a tomar decisiones en su vida cotidiana a partir de los mimos.

Figura 67. Unidades vendidas de los cuatro productos fabricadas por una misma empresa durante el pasado año

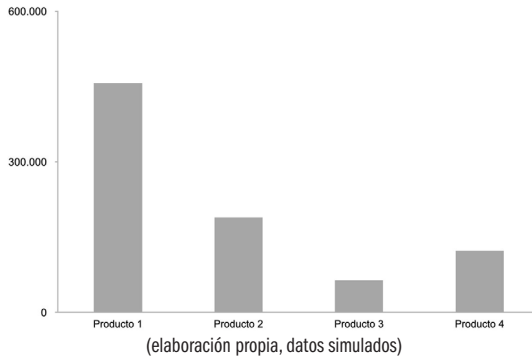
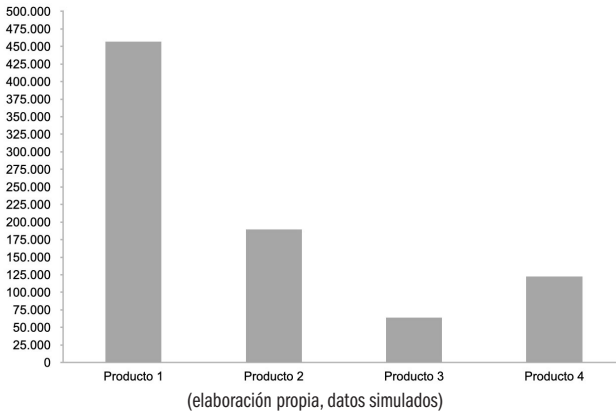


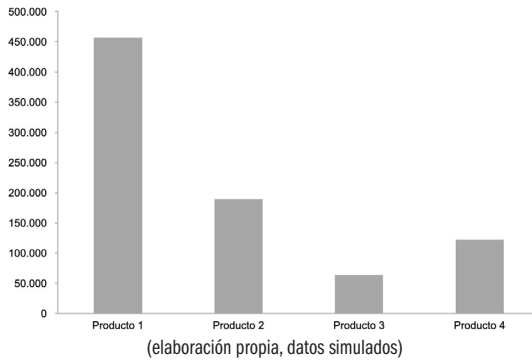
Figura 68. Unidades vendidas de los cuatro productos fabricadas por una misma empresa durante el pasado año



Como ejemplo para ilustrar esta relación especial que mantenemos con este tipo de números podemos recurrir al contexto de las redes sociales. Nadie dedicaría un post para decir que tiene o que ha alcanzado los 999 o los 1.001 seguidores en *Instagram*. En cambio, muchos de nosotros dedicaríamos unos minutos para generar un contenido en el que se anunciase que ya hemos conseguido alcanzar los 1.000 seguidores en esa red social.

Lo mismo ocurre en el contexto de una gráfica. Si, como recogemos en la figura 69, construimos una visualización en la que las etiquetas numéricas del eje que incorpora la escala cuantitativa acaban en 0 o 5 facilitaremos la lectura de esas etiquetas y habilitaremos también, de esa manera, una decodificación cómoda de los valores numéricos asignados a las columnas.

Figura 69. Unidades vendidas de los cuatro productos fabricadas por una misma empresa durante el pasado año

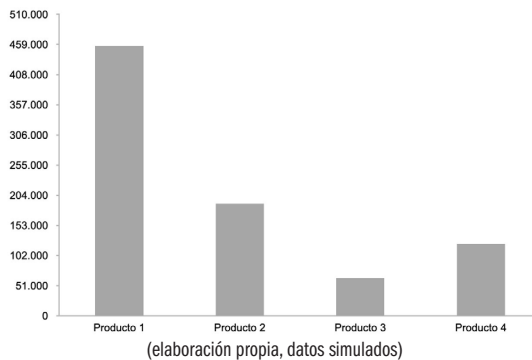


En cambio, si, como recogemos en la figura 70, proponemos una visualización en la que las etiquetas numéricas del eje que incorpora la escala cuantitativa no acaban en 0 ni en 5, al sernos más ajenas, dificultaremos la lectura de esas etiquetas y provocaremos, de esa manera, que la descodificación de los valores numéricos asignados a las columnas nos reclame un mayor esfuerzo cognitivo.

4.5. Leyenda

Como ya vimos en el capítulo anterior, la leyenda en una visualización es una especie de micrográfica (dotada también de cierta estructura arquitectónica similar a la de las tablas) dónde se nos ofrece la clave para, junto a la información suministrada por lo ejes, poder dotar de significado a los objetos visuales que representan simultáneamente los valores cuantitativos y los ítems cualitativos.

Figura 70. Unidades vendidas de los cuatro productos fabricadas por una misma empresa durante el pasado año



Respecto a ese elemento visual, podemos ofrecer unas recomendaciones que afectarían a su presencia o ausencia en la gráfica y al espacio o lugar de su emplazamiento o ubicación en la misma.

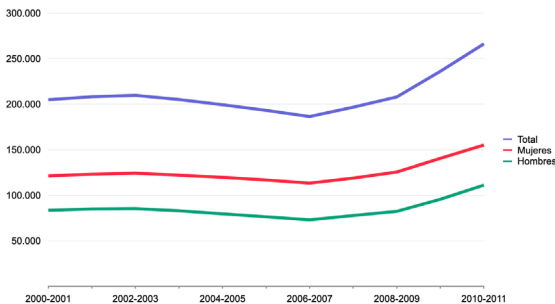
Comencemos abordando el tema de la presencia o la ausencia de la leyenda en una gráfica. El principal problema que nos podemos encontrar con la leyenda de una gráfica es que ésta suele ocupar un espacio importante dentro de la visualización. Y eso puede provocar que aumente el tamaño de la gráfica hasta el punto de que no encaje de forma adecuada en el documento en el cual está insertada. En esos caos concretos, podemos optar por eliminar directamente la leyenda y pasar a etiquetar directamente los atributos visuales.

Veamos este problema y esta solución a partir de un par de ejemplos. En la figura 71 recogemos una gráfica que representa el número de graduados universitarios en la primera década del siglo XXI en España. Esa gráfica incluye una leyenda (que nos ayuda a identificar el valor semántico de cada una de las líneas) pero que provoca que tengamos que ensanchar, de forma quizá excesiva, el espacio dedicado a la visualización.

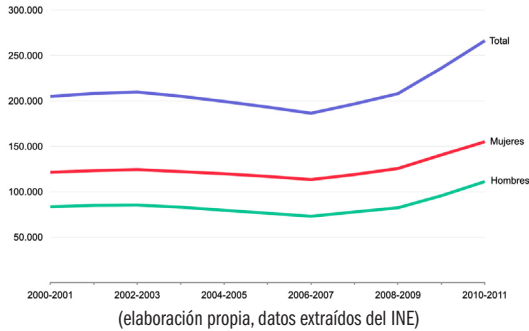
En esos casos, para reducir ese espacio, podemos eliminar directamente la leyenda de la gráfica y, para no perder la guía de codificación que ésta nos ofrecía, podemos incorporar unas etiquetas (“Total”, “Hombres” y “Mujeres”) al final de cada línea para poder identificar qué categoría o tipo de graduado está representando cada una de esas líneas (como recogemos en la figura 72).

Pero no en todas las gráficas podemos implementar esta solución. En algunos contextos, optar por eliminar la leyenda y renombrar directamente los objetos codificadores de información puede provocar una saturación y un solapamiento visual entre las etiquetas aportadas y esos propios objetos. Por ejemplo, como mostramos en la figura 73, si trasladamos las etiquetas de la leyenda a los extremos de las columnas que están representando las ventas trimestrales en las cuatro regiones

Figura 71. Graduados universitarios en la primera década del siglo XXI en España



(elaboración propia, datos extraídos del INE)

Figura 72. Graduados universitarios en la primera década del siglo XXI en España

donde ACME distribuye sus productos, obtenemos un mal resultado. Esas etiquetas se solapan visualmente entre sí y con las propias columnas, dificultando su lectura.

Frente a esta alternativa, como mostramos en la figura 74, es mucho más rentable y claro mantener la leyenda para obtener la clave de descodificación que nos permita identificar los valores de las ventas en esas cuatro regiones a lo largo de cada uno de los trimestres.

El lugar donde emplazar la leyenda es también importante. Como recomendación general, si decidimos que aparezca en la gráfica, debemos ubicar la leyenda fuera de la región de los datos, donde no se solape con los objetos visuales codificadores de información. Aunque, en algunos contextos, es posible insertarla dentro de ese espacio o región siempre y cuando no entorpezca la lectura de la gráfica.

En esta línea, en la gráfica de líneas que hemos recogido en la figura 75, hemos optado por esa última recomendación. Para no aumentar la anchura de la visualización, hemos sacado la leyenda de la zona de la derecha de la región de los datos y, como alternativa a la versión donde introducíamos las etiquetas que nos ayudan a descodificar las líneas (figura 72), la hemos insertado en ese espacio procurando que no se solapen con los objetos visuales (las líneas) que allí aparecen.

4.6 Ejes

Los ejes de coordenadas son uno de los elementos o componentes de soporte críticos dentro de una gráfica. No añaden o codifican directamente la información, pero suministran las referencias para poder descodificar la información representada por los objetos visuales (como los puntos, las líneas, las barras o las columnas) que aparecen en una visualización.

Normalmente, en una gráfica se suelen incluir dos ejes: el eje X y el eje Y. Pero en algunas circunstancias especiales podemos incluir algún eje extra o incluso borrar la línea de algunos de los ejes que allí aparecen.

Figura 73. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos

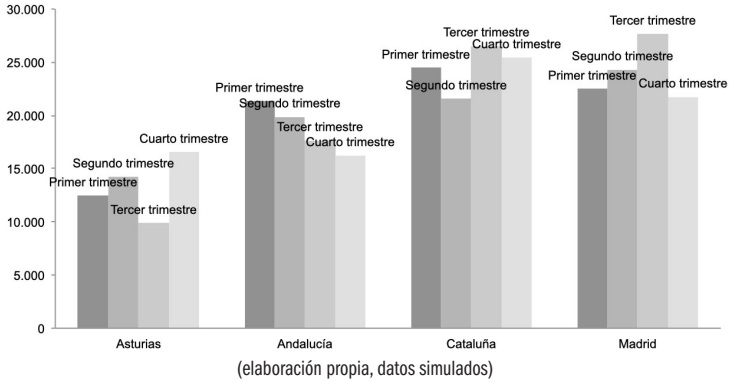


Figura 74. Ventas trimestrales en las cuatro regiones donde ACME distribuye sus productos

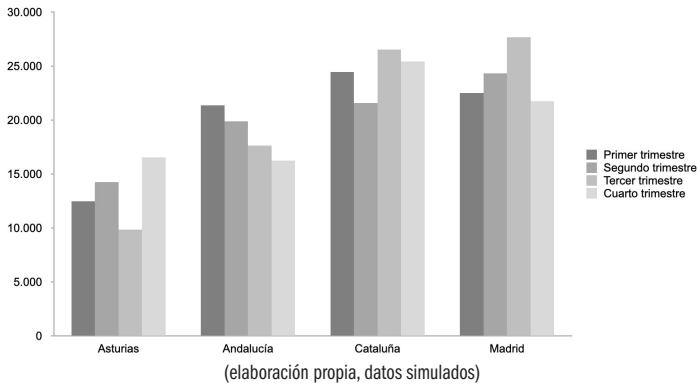


Figura 75. Graduados universitarios en la primera década del siglo XXI en España

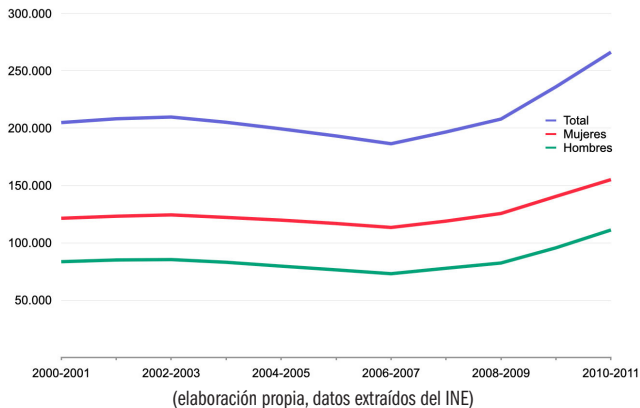
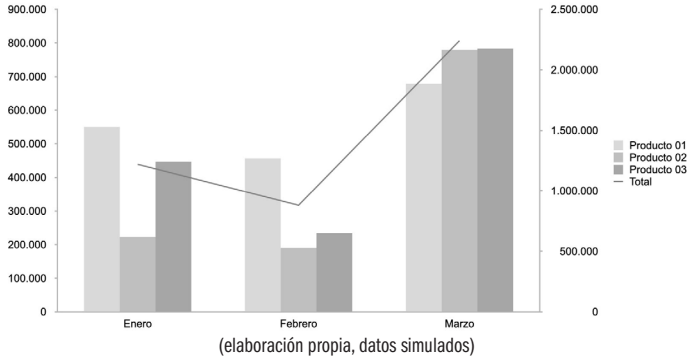


Figura 76. Ventas mensuales del primer trimestre del año pasado de los tres productos que fabrica ACME

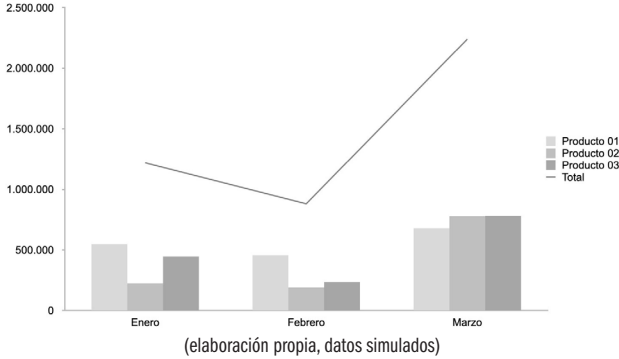
Comencemos por la posibilidad de incluir algún eje extra. En aquellas circunstancias en la que estemos representando un conjunto de datos con un rango muy extenso, incluir un eje extra puede ayudar a comunicar la información que queremos representar. Cuando en una misma gráfica debemos representar una serie de datos bajos junto a otra en la que son muy altos, podemos referenciar la serie de los bajos respecto a un de los ejes e incluir un tercer eje para referenciar la serie de los altos.

Veámoslo a través de un ejemplo. En la figura 76 hemos recogido una gráfica de columnas en la que se representan mediante columnas las ventas mensuales del primer trimestre del año pasado de los tres productos que fabrica ACME. Y para ofrecer el valor numérico asociado a cada columna hemos dispuesto el eje Y. En cambio, como los valores totales de las ventas mensuales de los tres productos es muy superior a las ventas mensuales de uno solo de esos productos, hemos decidido representar esos totales mediante una línea que se referencia respecto a un segundo eje Y.

De esta manera, conseguimos que convivan simultáneamente las ventas totales y las individuales y que se ofrezcan todas las referencias visuales para poder identificar los valores numéricos asociados a esos objetos (columnas y línea). En cierto sentido, al hacer esto estamos estandarizando visualmente los datos para poder comparar visualmente los valores de las dos series e identificar los valores concretos en las dos series.

Si no lo hubiéramos hecho de esa forma y hubiésemos referenciado las columnas y las líneas respecto a un único eje Y, la distancia entre esos valores nos hubiera impedido obtener en la gráfica las referencias visuales necesarias para poder leer con comodidad los valores cuantitativos asociados a las columnas (como ocurre en la gráfica de la figura 77).

Figura 77. Ventas mensuales del primer trimestre del año pasado de los tres productos que fabrica ACME



Centrémonos ahora en el tema de las líneas de los ejes. La percepción humana se rige por el principio de continuidad. Según este principio, tendemos a percibir como parte de un único todo aquellos objetos que están alineados con otros o aparecen a continuación de otros. O, dicho de tora forma: tendemos a percibir elementos como continuos, aunque estén interrumpidos entre sí.

Teniendo en cuenta este principio regulador de la percepción, en algunas circunstancias concretas, podemos decidir también borrar la línea de uno de los dos ejes de una gráfica. No es del todo recomendable, pero en una gráfica podemos mantener la línea del eje X y no dibujar en cambio la del eje Y cuando en este último eje sólo aparecen los términos o etiquetas de una variable cualitativa (y no una escala de valores cuantitativos). Por ejemplo, como ocurre en la gráfica recogida en la figura 78. En esa gráfica aparece dibujada la línea del eje X (la escala de valores

Figura 78. Sueldo por cargo en la empresa ACME

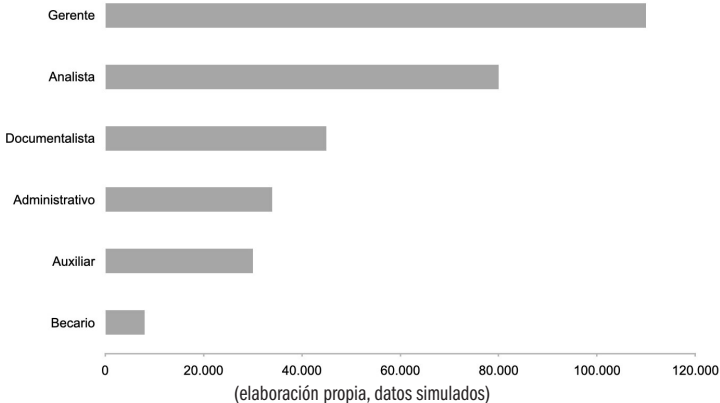
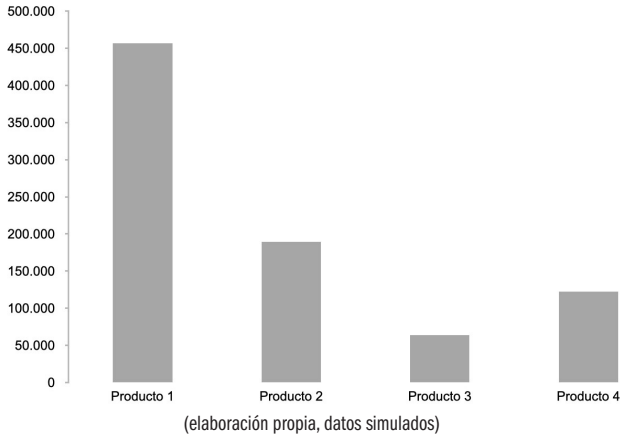


Figura 79. Unidades vendidas de los cuatro productos fabricadas por una misma empresa durante el pasado año
Unidades vendidas



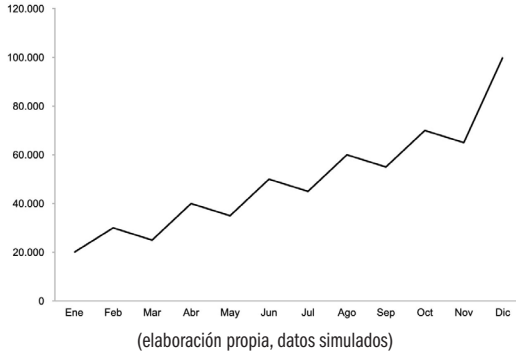
cuantitativos) pero no la del eje Y ya que la alineación de las barras por su base en este último nos hace percibir, en cierta manera y por así decirlo, esa línea ausente.

En la misma línea, en algunas ocasiones determinadas, en una gráfica también podemos mantener la línea del eje Y y no dibujar en cambio la del eje X cuando en este último eje sólo aparecen los términos o etiquetas de una variable cualitativa. Por ejemplo, como ocurre en la gráfica recogida en la figura 79. En esa gráfica aparece dibujada la línea del eje Y (la escala de valores cuantitativos) pero no la del eje X ya que la alineación de las columnas por su base en este último nos hace identificar, en cierta manera y por así decirlo, esa línea no presente.

Por otro lado, nos gustaría introducir una serie de recomendaciones sobre la ratio o proporción que debe mantener la longitud de uno de los ejes respecto a la del otro en el contexto de una gráfica. O dicho de otra manera, qué proporción entre la anchura y la altura de la región de datos de una gráfica es la más adecuada.

Ésta no es una cuestión trivial. Esa proporción tiene una gran influencia en la percepción de la gráfica por parte del usuario. Jugar intencionalmente con esa proporción puede dar lugar a una comunicación no del todo veraz de algún aspecto relacionado con un conjunto de datos.

Veámoslo a través de un ejemplo. Imaginemos que queremos comunicar cómo ha ido la venta de los productos fabricados por la empresa ACME durante el ejercicio del año pasado. Y para ello construimos la gráfica recogida en la figura 80. Con esa gráfica mostramos que las ventas han ido creciendo en forma de dientes

Figura 80. Productos vendidos por la empresa ACME durante el ejercicio anterior

de sierra a lo largo del año y que el balance es positivo, pasándose de las 20.000 unidades en enero a las 100.000 del mes de diciembre.

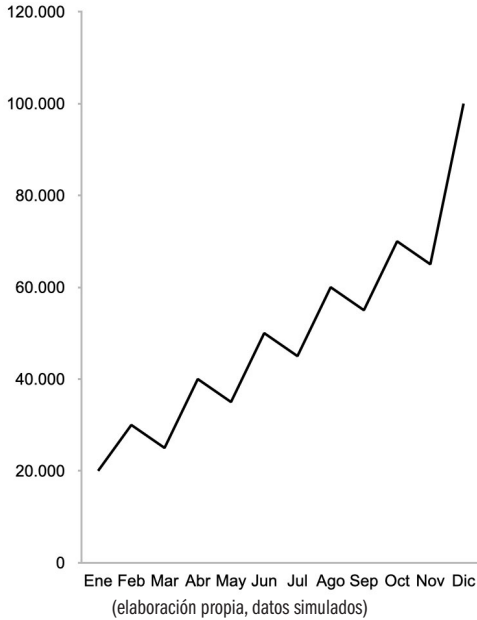
En cambio, manteniendo exactamente los mismos datos, podemos variar la proporción de la anchura de la gráfica respecto a su altura y transmitir una sensación totalmente diferente. Podemos, por ejemplo, hacer mucho más estrecha y alta la gráfica y dar a entender así que los resultados de las ventas han sido mucho más exagerados y positivos de lo que en un principio podría parecer (como mostramos en la gráfica incluida en la figura 81).

Frente a esta posibilidad de exageración, no existe una única ratio o proporción entre la anchura y altura de una gráfica que podamos recomendar. Siempre dependerá de los datos y de nuestra intención comunicativa a la hora de crear la visualización. Lo único que podemos exigir en ese diseño es intentar no manipular la ratio intencionadamente para distorsionar la comunicación y procurar hacer las gráficas más anchas que altas. De esta manera facilitaremos la lectura de la gráfica y respetaremos, además, el sentido normal de la lectura textual que podría aplicar un usuario al interactuar con la gráfica.

4.7. Cuadrícula

La cuadrícula es el último elemento sobre el que vamos a introducir alguna recomendación acerca de su diseño para que contribuya a mejorar la eficiencia en la comunicación que perseguimos con el uso de la gráfica en la que se incluye.

Como ya vimos en el capítulo anterior, las líneas de cuadrícula (denominadas también como “retícula” o “parrilla”) son combinaciones de líneas verticales y/o horizontales que ocupan toda la región de datos y que habitualmente (aunque no siempre) se cruzan entre ellas formando rectángulos para facilitar la identificación,

Figura 81. Productos vendidos por la empresa ACME durante el ejercicio anterior

lectura y comprensión de la información representada en una gráfica por parte de sus potenciales usuarios.

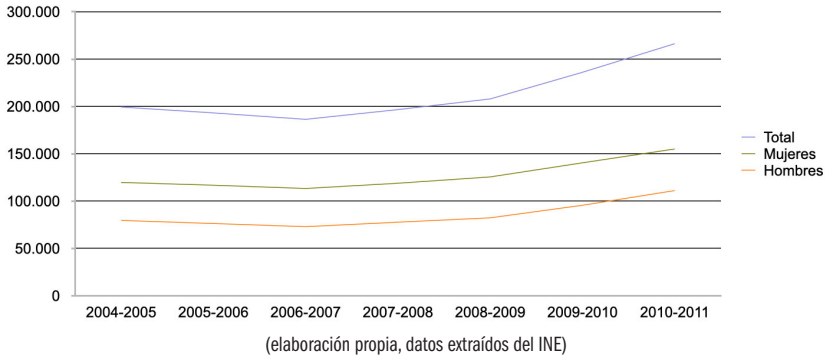
Respecto a este elemento de soporte, vamos a introducir unas recomendaciones sobre su aspecto y sobre su estructura en aquellas situaciones dónde nos interesa resaltar algún patrón entre los datos representados.

En relación con su aspecto, la recomendación principal es no utilizar un estilo de línea que tenga un aspecto similar a otros elementos codificadores de información en la misma gráfica. Si no respetamos este heurístico, el resultado visual puede provocar una lectura errónea y dificultosa de la gráfica por parte del usuario.

Veamos cómo funciona esta sugerencia a través de un par de ejemplos. En las figuras 82 y 83 recogemos respectivamente unas gráficas en las que no seguimos la recomendación y, por tanto, provocamos una situación en la que el usuario puede tener dificultades para entender ambas propuestas comunicativas.

En la gráfica de la figura 82, se han introducido las líneas de cuadrícula haciéndolas coincidir con valores primarios de la escala del eje Y. Además, estos elementos de soporten se han dibujado con el mismo grosor que el de las líneas que representan el número de graduados universitarios (totales, hombres y mujeres) en la primera década del siglo XXI en España. Al generarlos con el mismo grosor, esta

Figura 82. Graduados universitarios en la primera década del siglo XXI en España



decisión provoca que las líneas de cuadrículas se puedan confundir visualmente con los otros elementos lineales que están representando la evolución de esos tres grupos de graduados.

En el mismo sentido, en la gráfica de la figura 83, también se han introducido las líneas de cuadrícula, pero, en este caso, haciéndolas coincidir con valores primarios (los que llevan una etiqueta numérica asignada) y secundarios (valores intermedios que no llevan una etiqueta numérica asignada) de la escala del eje Y. De nuevo, todos estos elementos de soporte se han dibujado con el mismo grosor que el de las líneas que representan el número de graduados universitarios (totales, hombre y mujeres) en la primera década del siglo XXI en España. De la misma forma que en el caso anterior, al generarlos con el mismo grosor, esta decisión provoca que las líneas de cuadrículas puedan solaparse visualmente con los otros elementos lineales que están codificando la evolución de los tres grupos de graduados.

Figura 83. Graduados universitarios en la primera década del siglo XXI en España

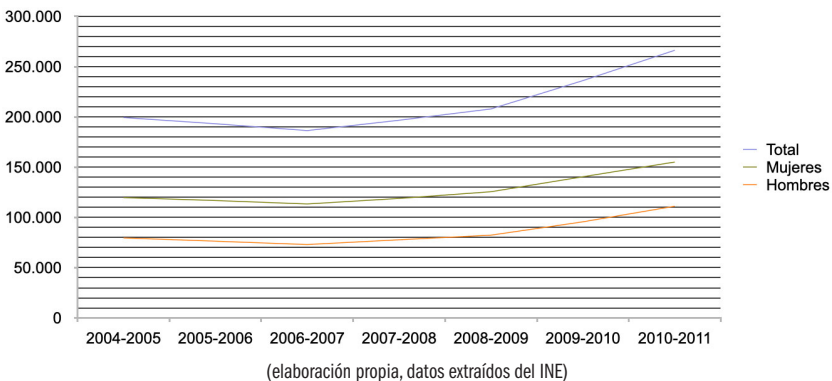
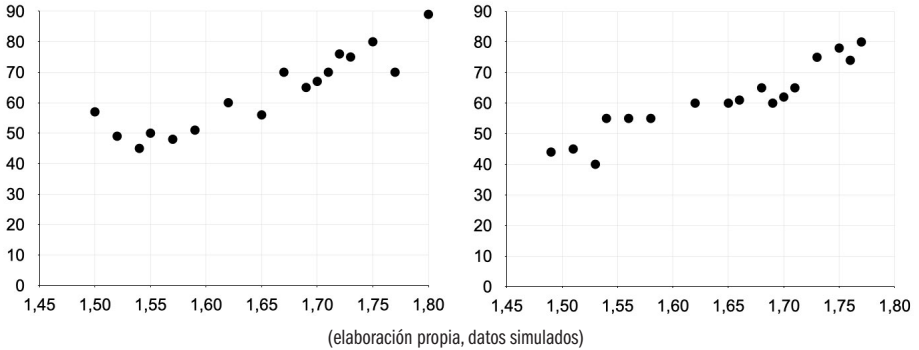


Figura 84. Altura y peso de un grupo de hombres y mujeres



La última de las recomendaciones afecta a aquellas situaciones comunicativas en las que nos interesa resaltar la percepción y comparación de algún patrón entre datos representados a través de diferentes gráficas. En esos casos, para facilitar esa identificación y comparación, es altamente recomendable repetir la misma estructura de cuadrículas y de etiquetas en los ejes de las gráficas implicadas.

Podemos ilustrar esta sugerencia con la última de las figuras (84). En esta figura recogemos dos gráficas que intentan representar si existe o no algún tipo de correlación entre el peso y la altura de una serie de personas, teniendo en cuenta también el sexo de los miembros de esa comunidad. La gráfica de la izquierda es un diagrama de dispersión que analiza esa correlación entre los hombres de esa población. En la de derecha, se codifica esa misma correlación, pero entre las mujeres. En este contexto, y para facilitar la identificación y la comparación de los patrones que presentan esos dos conjuntos de datos, nuestra propuesta visual se fundamentaría en dos diagramas de dispersión con las mismas proporciones y tamaños y donde sus respectivas parrillas formadas por las líneas de cuadrículas son idénticas en forma, tamaño, distribución y etiquetas numéricas.

Bibliografía

- ALDRICH, F.; SHEPPARD, L. Graphicacy; The fourth 'R'? **Primary Science Review**, vol. 64, p. 8-11. 2000.
- ANNING, A. Pathways to the graphicacy club: The crossroad of home and pre-school. *Journal of Early Childhood Literacy*, vol. 3, n. 1, p. 5-35. 2003.
- AVGERINOY, M. y PETERSSON, R. Toward a Cohesive Theory of Visual Literacy. *Journal of Visual Literacy*, vil. 30, n. 2, p. 1-19. 2011.
- BALCHIN, W. G. Graphicacy. *American Cartographer*, vol. 3, n. 1, p. 33-38. 1976.
- BALCHIN, W. G. Graphicacy comes of age. *Teaching Geography*, 1vol. 1, n. 1, p. 8-9. 1985.
- BAWDEN, David. Information and Digital Literacy: A review of concepts. *Journal of Documentation*, vol. 57, n. 2, p. 218-259. 2001.
- BEHRENS, S. J. A Conceptual Analysis and Historical Overview of Information Literacy. *College and Research Libraries*, vol. 55, n. 4, p. 309-322. 1994.
- BERTIN, Jacques. *Semiology of Graphics: Diagrams, Networks, Maps*. New York, NY: Esri Press. 1983.
- BOARDMAN, D. Graphicacy revisited: mapping abilities and gender differences. *Educational Review*, vol. 42, n. 1, p. 57-64. 1990.
- BRUCE, C. The Relational Approach: a New Model for Information Literacy. *The New Review of Information and Library Research*, vol. 3, p. 1-22. 1997.
- CAIRO, Alberto. *The Truthful Art*. San Francisco: New Riders. 2012.
- CAIRO, Alberto. *The Functional Art: An Introduction to Information Graphics and Visualization*. Berkeley, CA: New Riders. 2013.
- CAIRO, Alberto. *How Charts Lie. Getting Smarter about Visual Information*. New York: W. W. Norton & Co. 2019.
- CARBO, T. Mediacy: Knowledge and Skills to Navigate the Information Superhighway. *Proceedings of the Infoethics Conference, Monte Carlo, 10-12 Marzo 1997, UNESCO, Paris*. 1997.
- CARD, S. K. Information Visualization. *In.: JACKO, J. A. y SEARS, A. (eds.). The Human-Computer Interaction Handbook*. Mahwah, NJ: Lawrence Erlbaum Associates. 2003.
- CHEN, Chaomei. *Information Visualization and Virtual Environments*. New York: Springer Verlag. 1999.
- CHERVEN, K. *Mastering Gephi Network Visualization*. Birmingham: Ed. Packt Publishing Ltd. 2015.
- CLEVELAND, William S. *Visualizing Data*. Summit, NJ: Hobart Press. 1993.

- CLEVELAND, William S. *The Elements of Graphing Data*. Summit: Hobart Press. 1994.
- COTGREAVE, Andy; Shaffer, Jeffrey y Wexler, Steve. *The Big Book of Dashboards: Visualizing Your Data Using Real-World Business*. Hoboken, New Jersey: John Wiley & Sons. 2017.
- DESS, H. M. "Information Literacy: a Subject Source Survey and Annotated Bibliography". In.: VARLEJS, J. (ed.) (1991). *Information Literacy: Learning How to Learn*. Jefferson: Mcfarland. 1991.
- DICK, Murray. *The Infographic. A History of Data Graphics in News and Communications*. Cambridge: The MIT Press. 2020.
- DOYLE, C. S. *Information Literacy in An Information Society: A Concept for the Information Age*. Syracuse: ERIC Clearinghouse. 1994.
- DRUCKER, Johanna. *Graphesis: Visual Forms of Knowledge Production*. Cambridge: Harvard University Press. 2014.
- DRUCKER, Johanna. *Visualization and Interpretation. Humanistic Approaches to Display*. Boston: The MIT Press. 2020.
- FEKETE, J. D.; VAN WIJK, J. STASKO, J. NORTH, C. *The Value of Information Visualization*. In.: KERREN, A.; STASKO, J.; FEKETE, J. D. y NORTH, C. *Information Visualization: Human-Centered Issues and Perspectives*. New York: Springer Verlag. 2008.
- FEW, Stephen. *Information Dashboard Design*. Sebastopol (CA): O'Reilly Media Inc. 2006.
- FEW, Stephen. *Save the Pie for Dessert. Perceptual Edge. Visual Business Intelligence Newsletter*. 2007.
- FEW, Stephen. *Show Me the Numbers*. Oakland: Analytics Press. 2012.
- GOLOMBISKY, K. y Hagen, R. *White Space is Not Your Enemy: A Beginner's Guide to Communicating Visually through Graphic, Web & Multimedia Design*. Burlington, MA: Elsevier, Inc. 2010.
- HARRIS, Robert L. *Information Graphics: A Comprehensive Illustrated Reference*. New York, NY: Oxford University Press. 1996.
- Harris, Benjamin. *Visual Information Literacy via Visual Means: Three Heuristics*. *Reference Services Review*, vol. 34, n. 2, p. 213 -221. 2006.
- HATTWIG, D., BURGESS, J., BUSSERT, K. y MEDAILLE, A. *Visual Literacy Competency Standards for Higher Education*. Association of College & Research Libraries. 2011. <<https://www.ala.org/acrl/standards/visualliteracy>>. [Consulta: septiembre 2021].
- HEALEY, Christopher G. y ENNS, James. *Attention and Visual Memory in Visualization and Computer Graphics*. *IEEE Transactions on Visualization and*

- Computer Graphics, vol. 18, n. 7, p. 1170-1188. 2012.
- HEARN, Donald. Computer graphics with OpenGL. Boston: Pearson. 2011.
- HOCKLEY, William E. y BANCROFT, Tyler. Extensions of the Picture Superiority Effect in Associative Recognition. *Canadian Journal of Experimental Psychology*, vol. 65, n. 4, p. 236-56. 2011.
- HORM, Robert E. Visual Language. Bainmdrig: MacroVU. 1998.
- HUGUES, J. F. y VAN DAM, A. Computer Graphics: Principles and Practice. Boston: Addison-Wesley. 2013.
- ISRAEL, D. y PERRY, J. What is Information? *In.*: Hanson, Philip P. (ed.). Information, Language and Cognition. Vancouver: University of British Columbia Press. 1990.
- JONES, B. Communicating Data with Tableau: Designing, Developing, and Delivering Data Visualizations. London: Ed. Safari Books Online. 2014.
- KHAN, A. Jumpstart Tableau: A Step-By-Step Guide to Better Data Visualization. London: Ed. Apress. 2016.
- KIRK, Andy. Data Visualization: A Successful Design Process. Birmingham: Packt Publishing. 2012.
- KOSARA, Robert y MACKINLAY, Jock. Storytelling: The Next Step for Visualization. *Computer*, vol. 46, n. 5, p. 44-50. 2013.
- KULTHAU, C. Information Skills for an Information Society. Syracuse: ERIC Clearinghouse on Educational Resources. 1987.
- MCCLURE, C. Network Literacy: A Role for Libraries. *Information Technology and Libraries*, vol. 13, p. 115-125. 1994.
- MCGARRY, Kevin. Definitions and Meanings of Literacy. *In.*: Barker, K. y Lonsdale, R. (eds.). Skills for Life: The Value and Meaning of Literacy. London: Taylor Graham. 1994.
- MCGARRY, Kevin. The Changing Context of Information. London: Library Association Publishing. 1993.
- MEIRELLES, Isabel. Design for Information. Beverly, MA: Rockport Publishers. 2013.
- MILLER, George Armitage. The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review*, vol. 63, n. 2, p. 81-97. 1956.
- MURRAY, Scott. Interactive Data Visualization for the Web: An Introduction to Designing with D3. Sebastopol, CA: O'Reilly Media, Inc. 2013.
- MUTCH, M. Information Literacy: An Exploration?. *International Journal of Information Management*, vol. 17, n. 5, p. 377-386. 1997.
- NORTH, C. Information Visualization. *In.*: SALVENDY, G. (ed.). Handbook of

- Human Factors and Ergonomics. Hoboken, New Jersey: John Wiley & Sons. 2005.
- NUSSBAUMER, Cole. *Storytelling With Data: A Data Visualization Guide for Business Professionals*. Hoboken, New Jersey: John Wiley & Sons. 2015.
- OLSEN, J. K. y COONS, B. Cornell University's Information Literacy Program. *In.*: MENSCHING, G. y MENSCHING, T. (eds.). *Coping with Information Illiteracy: Bibliographic Instruction for The Information Age*. Ann Arbor: Pieran Press. 1989.
- PÉREZ-MONTORO, Mario y GOLKHOSRAVI, Mehrad. *Information Visualization*. *In.*: DÍAZ, José María; PÉREZ-MONTORO, Mario y SALTO, Francisco (coords.). *Interdisciplinary Elucidation of Concepts, Metaphors, Theories and Problems Concerning Information*. Santa Elena (Ecuador): Universidad Estatal Península de Santa Elena. 2014.
- PLAYFAIR, William. *Commercial and Political Atlas: Representing, by Copper-Plate Charts, the Progress of the Commerce, Revenues, Expenditure, and Debts of England, during the Whole of the Eighteenth Century*. London: Corry. 1786.
- PLAYFAIR, William. *Statistical Breviary; Shewing, on a Principle Entirely New, the Resources of Every State and Kingdom in Europe*. London: Wallis. 1801.
- ROBBINS, Naomi. *Creating More Effective Graphs*. Columbia: Chart House. 2013.
- ROTH, W.-M., POZZER-ARDENGHI, L., y HAN, J. Y. *Critical Graphicacy: Understanding Visual Representation Practices in School Science*. New York: Springer Verlag. 2005.
- SEGEL, Edward y HEER, Jeffrey. *Narrative Visualization: Telling Stories with Data*. *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, n. 6, p. 1139-1148. 2010.
- SHIRLEY, Peter. *Fundamentals of computer graphics*. Natick, MA: AK Peters. 2009.
- SHNEIDERMAN, Ben. *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations*. *Proceedings of the IEEE Symposium on Visual Languages*, IEEE Computer Society Press, p. 336-43. 1996.
- SNAVELY, L. y COOPER, N. *The Information Literacy Debate*. *Journal of Academic Librarianship*, vol. 23, n. 1, p. 9-20. 1997.
- SPENCE, Robert. *Information Visualization*. New York: Pearson Education. 2000.
- SPENCE, Robert. *Information Visualization: An Introduction*. New York: Springer Verlag. 2014.
- TONDREAU, Beth. *Layout Essentials: 100 Design Principles for Using Grids*.

- Beverly, MA: Rockport Publishers. 2009.
- TUFTE, Edward. *The Visual Display of Quantitative Information*. Cheshire: Graphic Press. 1983.
- TUFTE, Edward. *Envisioning Information*. Cheshire: Graphic Press. 1990.
- TUFTE, Edward. *Beautiful Evidence*. Cheshire: Graphics Press. 2006.
- TUKEY, John W. *Exploratory Data Analysis*. Reading, MA: Addison Wesley. 1977.
- VAN WIJK, J. Views on Visualization. *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, n. 4, p. 421-433. 2006.
- WARE, Collin. *Information Visualization: Perception for Design*. Cambridge, MA: Morgan Kaufmann. 2021.
- WICKHAN, Hadley. Tidy Data. *Journal of Statistical Software*, vol. 55, n. 2, p.1-24. 2017.
- WONG, Dona M. *The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures*. New York: W. W. Norton and Company. 2010.
- YAU, Nathan. *Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics*. Indianapolis, IN: Wiley Publishing, Inc. 2011.
- YAU, Nathan. *Data Points: Visualization that Means Something*. Indianapolis, IN: Wiley Publishing, Inc. 2011.
- ZELAZNY, Gene. *Say It with Charts*. New York: MacGraw-Hill. 2001.

50

Realização



Cooperação



Cooperação
Representação
no Brasil



ESTA OBRA É PARTE DA COLEÇÃO PPGCI 50 ANOS E FOI
COMPOSTA EM MINION PELO PROGRAMA DE EDUCAÇÃO
TUTORIAL DA ESCOLA DE COMUNICAÇÃO DA UFRJ EM
FEVEREIRO DE 2022.

“ Con este trabajo, vamos a presentar y revisar los conceptos y principios básicos que sustentan la visualización de la información y conocer sus fundamentos perceptivos elementales. También vamos a ofrecer una base metodológica para seleccionar la mejor visualización con gráficas para cada tipo de contenido semántico. Y, por último, vamos a aprender a realizar un correcto diseño visual de gráficas de acuerdo con los principios cognitivos y éticos que las regulan. ”

EM COOPERAÇÃO



United Nations
Educational, Scientific and
Cultural Organization