

Gramática de las gráficas

Pistas para mejorar
las representaciones de datos

JOAQUÍN SEVILLA MORÓDER

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Gramática de las gráficas
Pistas para mejorar las representaciones de datos

JOAQUÍN SEVILLA MORÓDER

Gramática de las gráficas

Pistas para mejorar
las representaciones de datos

Gramática de las gráficas

Pistas para mejorar las representaciones de datos

JOAQUÍN SEVILLA MORÓDER

Universidad Pública
de Navarra
Nafarroako
Unibertsitate Publikoa



Título: *Gramática de las gráficas: Pistas para mejorar las representaciones de datos*

Autor: Joaquín Sevilla Moróder

Edita: Universidad Pública de Navarra : Nafarroako Unibersitate Publikoa

Fotocomposición: Pretexto

Imprime: Ona Industria Gráfica

Depósito Legal: NA-2577/2005

ISBN: 84-9769-106-7

© Joaquín Sevilla Moróder

© Universidad Pública de Navarra

Impreso en papel ecológico

Esta publicación no puede ser reproducida, almacenada o transmitida total o parcialmente, sea cual fuere el medio y el procedimiento, incluidas las fotocopias, sin permiso previo concedido por escrito por los titulares del copyright.

Coordinación y distribución: Dirección de Publicaciones
Universidad Pública de Navarra
Campus de Arrosadía
31006 Pamplona
Fax: 948 169 300
Correo: publicaciones@unavarra.es

Índice

PRÓLOGO	9
INTRODUCCIÓN	11
Capítulo 1. COMUNICACIÓN GRÁFICA	25
Capítulo 2. QUÉ ES UNA GRÁFICA	39
Capítulo 3. EL TEXTO GRÁFICO	51
Capítulo 4. ORTOGRAFÍA	69
Capítulo 5. SINTAXIS	91
Capítulo 6. SEMÁNTICA	105
Capítulo 7. ESTILOS LITERARIOS. CONCLUSIONES FINALES	119
Anexo 1. BREVE HISTORIA DE LAS GRÁFICAS	125
Anexo 2. REVELAR LOS DATOS	133
Anexo 3. CUÁNDO NO HACER UNA GRÁFICA	141
REFERENCIAS	149
PROCEDENCIA DE LAS GRÁFICAS	155

Prólogo

Cuando uno comienza a preparar un texto (la tesis doctoral, el proyecto final de carrera, el primer reportaje en su nuevo puesto de periodista, etc.) la preocupación por el fondo de la cuestión que tiene entre manos es inmensa. El contenido técnico, los fundamentos, las fuentes, el detalle sobre el proceso, todas estas cosas ocupan tanto la mente que para las formas no suele quedar hueco: ni tiempo ni esfuerzo.

Tendemos a pensar que un buen contenido brilla por sí mismo. Si hemos hecho un buen experimento (o práctica de laboratorio), el artículo o informe en que lo exponamos será bueno. Si disponemos de unos excelentes datos estadísticos de una compleja encuesta el estudio que los detalle será también excelente. Pero esto dista mucho de ser cierto. A partir de unos buenos contenidos generar un buen resultado exige, cuando menos, una presentación correcta, un cierto esfuerzo en el proceso de darles su forma final.

En este sentido, hace casi una década comencé a impartir una asignatura de doctorado en la Universidad Pública de Navarra en la que nos preocupábamos por el formato de la actividad científica más que por su contenido. La propuse con un cierto reparo, temiendo que mis colegas considerasen una pérdida de tiempo para los estudiantes este esfuerzo dedicado a la forma en vez de al contenido. Afortunadamente el tiempo ha demostrado que ese es-

fuerzo está más que recompensado. Es fundamental dedicar una parte de la actividad científica a la comunicación efectiva de sus resultados, y un aprendizaje más o menos formalizado de esa comunicación ahorra mucho tiempo y sufrimiento a largo plazo.

Uno de los temas de ese curso está dedicado a las gráficas, dado que constituyen uno de los vehículos de expresión de contenidos científicos más útiles y habituales. A raíz de ese tema comencé a leer los periódicos fijándome especialmente en las gráficas desde un punto de vista puramente formal y recorté muchas de esas gráficas. En realidad tengo una gran tendencia a recortar de todo en los periódicos, aunque casi nunca sé qué hacer con los recortes al final. En este caso, al contrario, los recortes fueron colocándose en carpetas con títulos, y fue surgiendo así una estructuración de los errores en las gráficas.

Finalmente, el resultado de unas elecciones (negativo para mi candidatura) de las que salpican la vida política en la universidad me dejó desocupada (que no es exactamente lo mismo que libre) una cantidad de tiempo de la que no tenía prevista la capacidad de disponer. Decidí darle un empujón más a la estructuración de esos recortes de prensa como para poderlos publicar de alguna forma.

El resultado es el texto que sigue, y espero que resulte de utilidad a muchas de las personas que, disponiendo de unos datos excelentes, se propongan convertirlos en una gráfica que les haga justicia. He intentado centrarme en lo puramente formal, válido tanto para una gráfica de periódico deportivo como para la que se incluirá en una tesis doctoral.

El hilo argumental consiste en explotar la analogía entre las reglas para construir el lenguaje (la gramática) y las que existen para construir las gráficas. Sólo el capítulo 1 se sale de ese guión para ofrecer una cierta fundamentación teórica al lenguaje gráfico. En caso de resultar pesado, no hay ningún problema en saltarlo, dado que el resto conserva todo su sentido sin él.

Llegados a este punto sólo queda agradecer a los que han hecho posible que el texto vea la luz: a la derrota electoral por el tiempo que me desocupó, a la familia (esposa e hijos) por el tiempo no disfrutado y por el apoyo en lo esencial, a los estudiantes de todos estos años por lo que me han enseñado, a amigos y colegas por el apoyo, al editor por el cuidado en la publicación... A todos ellos el agradecimiento por lo que de valor haya.

Pamplona, 16 de junio de 2005

Introducción

Un punto de partida

Leyendo revistas y periódicos, en artículos científicos y libros técnicos o corrigiendo trabajos de los estudiantes, a menudo se encuentran gráficas. No es infrecuente que al observar algunas de esas gráficas se experimente una cierta sensación de que algo no va bien, de que no todo encaja. Pero casi nunca prestamos más atención; lo importante es el texto que se está leyendo, y la gráfica la apreciamos como una mera ilustración del mismo, de igual manera que contemplaríamos una fotografía o un dibujo.

La primera vez que tomé conciencia de que esa sensación de que las cosas no van bien se debía a flagrantes incorrecciones fue leyendo un periódico en plenas vacaciones. Recorté la página en cuestión y la guardé: en el periódico local de una ciudad costera se analizaba lo que había empeorado el tráfico en los últimos años, y para ilustrarlo se incluía la figura reproducida aquí (fig. 1).

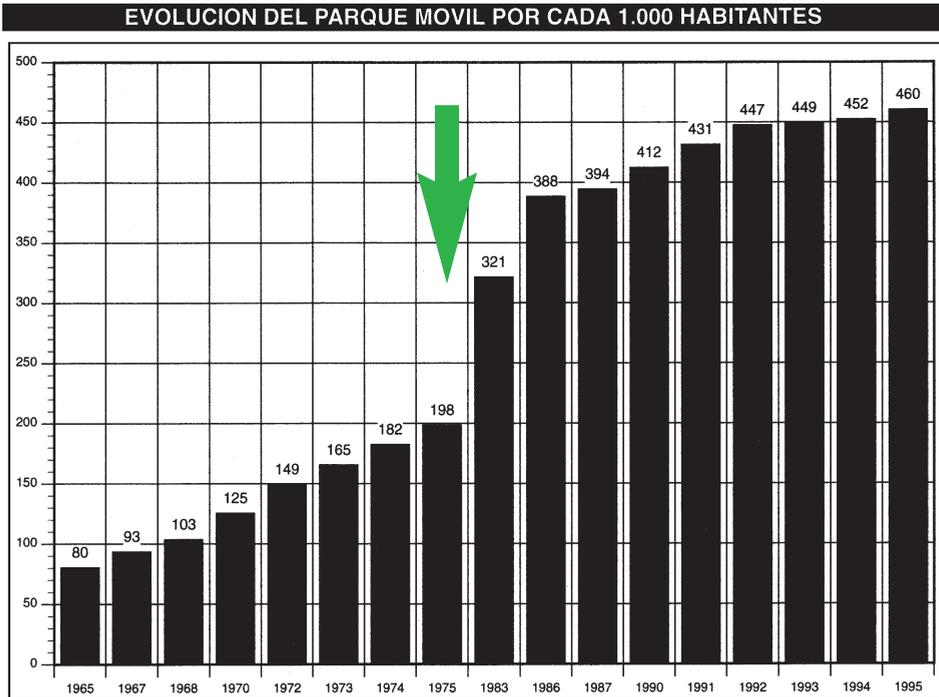


Figura 1. Evolución del parque móvil tomada de un periódico. El escalón que se aprecia en el centro (indicado por una flecha que no estaba en el original) se debe a un artificio de la representación, no a los datos representados.

En la figura se representa la evolución temporal del número de vehículos por cada 1000 habitantes en España. Del primer golpe de vista se aprecia un escalón brusco de subida, aproximadamente en el centro (indicado en la figura 1 con una flecha, que lógicamente no está en la figura del periódico). Adelantándome a la lectura, supuse que el redactor ponderaría que el crecimiento de las infraestructuras no puede ser tan brusco como el de los automóviles o algo así. Pero en realidad nada de eso aparecía en el artículo.

Al fijarme mejor en la gráfica para ver en qué momento se producía ese significativo crecimiento vi que era entre 1975 y 1983. Antes de empezar a elucubrar con un posible despegue económico en esos años o algo así me di cuenta de que las barras del gráfico no están colocadas de forma homogénea en el tiempo. Así entre dos barras consecutivas puede haber una separación temporal de un año, ¡o de 8!

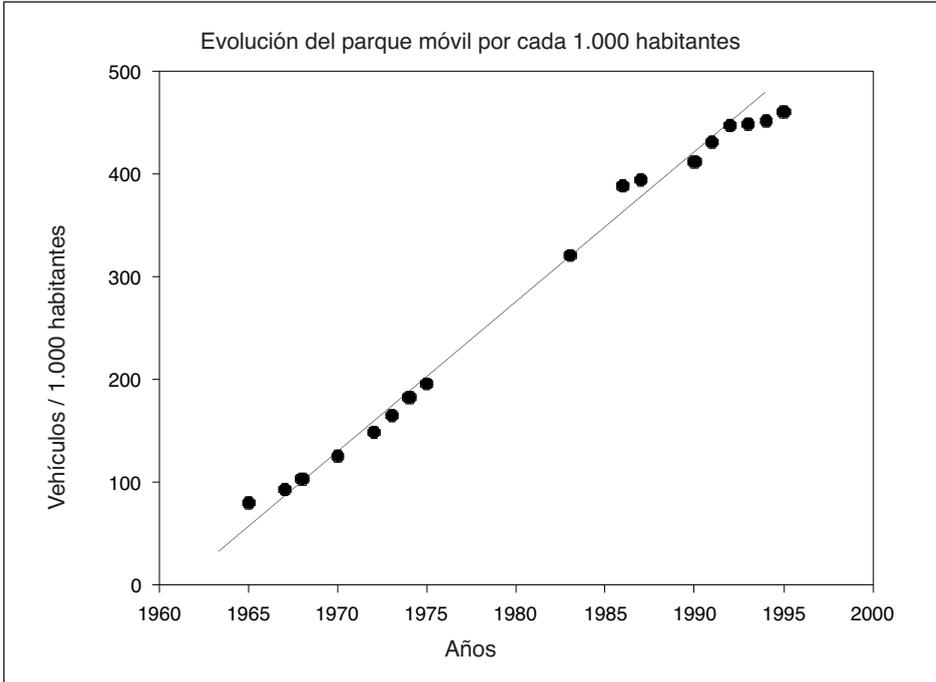


Figura 2. Los mismos datos de la figura 1 representados de forma que las distancias en el eje x correspondan con el paso del tiempo de una forma homogénea.

Si se vuelven a dibujar los mismos datos de una forma más respetuosa con el paso de los años (como la mostrada en la figura 2), ya no aparece ningún escalón. El crecimiento del parque móvil, a partir de los datos disponibles, es muy lineal entre 1965 y 1992. El intervalo que en la figura 1 da lugar al escalón, en la figura 2 se revela como un período del que se carece de información, no se dispone de los datos.

Una colección de datos incompleta más una elección incorrecta del tipo de gráfica dan lugar a un artificio gráfico que percibe el lector como un enunciado falso: no existe un aumento brusco puntual como el que claramente muestra la primera gráfica.

Bueno, se puede ser indulgente y pensar que en agosto los periódicos están llenos de becarios mientras los redactores avezados veranean, podría ser un despiste anecdótico.

Pero el incidente me creó un cierto recelo ante cualquier gráfica, las miraba todas ya con atención. No pasaron ni dos meses antes de que, al poco de

comenzar el curso, encontrara exactamente el mismo error formal en una gráfica incluida en uno de los guiones de prácticas que me entregó un grupo de estudiantes de último año de ingeniería (ver figura 3).

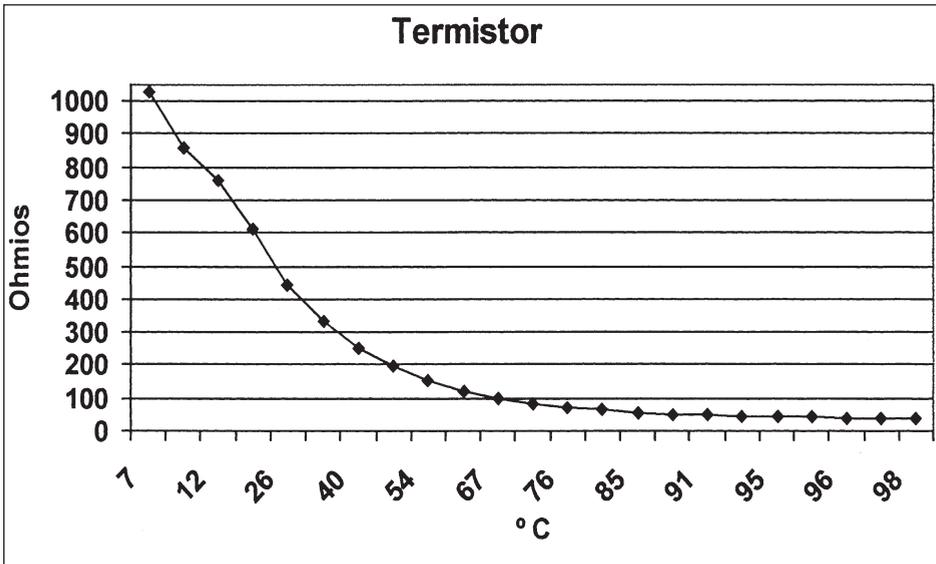


Figura 3. En una experiencia de laboratorio, unos estudiantes de ingeniería industrial representan la resistencia de un dispositivo en función de la temperatura. La representación adolece del mismo error que la de la figura 1: intervalos no equiespaciados en el eje x.

En el eje x de la figura cada intervalo lo mismo puede valer 1 que 14 °C. No es fácil saber qué aspecto tendría la gráfica si los intervalos de temperatura estuvieran correctamente espaciados. Me gustaría pensar que la incorrección estuvo causada por un esfuerzo voluntario del estudiante en distorsionar la figura para que se pareciera más a lo que el profesor espera encontrar, pero no es nada probable. Casi seguro que el error procede de una situación análoga a la del anterior redactor: los datos disponibles son los que son y no se le dedica un segundo a elegir un tipo de gráfico que se adecue a ellos.

En este caso ya no es tan fácil ser indulgente, un estudiante a escasos meses de obtener el título de Ingeniero no parece que deba cometer ese tipo de errores. Pero el caso mostrado en la figura 3 no deja de ser uno más: el error no es infrecuente en estudiantes. Pero tampoco lo es en los periódicos

(incluso cuando pasado agosto ya no se puede culpar a los becarios), en revistas, en informes técnicos o en casi cualquier texto que contenga gráficas¹.

Esta toma de conciencia de la ubicuidad del error gráfico lleva a plantearse muchas cuestiones. ¿Cómo se hace bien una gráfica? ¿Cómo se hace mal? ¿Por qué es tan frecuente encontrarlas incorrectas? ¿Alguien enseñó al estudiante de ingeniería cómo hacerlas bien? ¿Qué es, en definitiva, una gráfica? ¿Por qué utilizamos gráficas? ¿Para qué? ¿Desde cuándo? ¿Se ha teorizado mucho sobre gráficas?

La respuesta a la mayor parte de estas cuestiones no es trivial. Pero algunas sí se pueden responder categóricamente: nadie enseñó al estudiante de Ingeniería cómo hacer representaciones gráficas. A priori es fácil suponer que es algo suficientemente evidente como para requerir una enseñanza formal, o que tiene que traerse aprendido de antes de entrar en la Universidad. Lo que debe o no saberse antes de entrar a la universidad no es objeto de este texto, pero una reflexión sobre las gráficas sí. Una reflexión que, sin resultar demasiado pesada, pudiera haberle resultado de utilidad al redactor de la primera gráfica, al estudiante de la tercera, y a muchos más.

El objetivo, por tanto, es preparar un texto sencillo que pueda ayudar a no cometer el tipo de errores antes comentado a estudiantes, periodistas, ingenieros, economistas y quien sea menester. Visto dónde hay que llegar, ya sólo falta saber por dónde empezar.

Explorando posibles caminos

Tras la anécdota de la figura 1, estuve varios años recortando periódicos y otros documentos, y recopilando todo tipo de gráficas que me resultaran de algún modo incorrectas. También se lo encargué como trabajo a estudiantes de doctorado. Al cabo de algunos años me encontré con varios cientos de gráficos comentados, aunque mucho menos ordenados de lo que me hubiera gustado.

Agrupando e intentando sistematizar los recortes llegué a algunas conclusiones y entonces es cuando se me ocurrió la idea de preparar un texto. Pero la simple enumeración de gráficas erróneas comentadas no parece algo muy provechoso. En realidad sí tiene interés, al igual que lo tiene un libro de

1. Sólo las revistas científicas profesionales se salvan un poco de esta tendencia al error.

recetas de cocina; pero no es un procedimiento adecuado de aprendizaje. Aprender a cocinar desde cero con libros de recetas es bastante difícil. No, la “aproximación” enumerativa, bien de trucos para hacerlo, bien de errores a evitar, no parece la adecuada. Además, quizá por ser lo más evidente, es lo que resulta más fácil encontrar en la literatura sobre el tema, aunque sea en inglés².

Tras comprobar que la aproximación que surge de una forma más evidente no es la mejor, cabe preguntarse sobre de qué otras formas puede uno acercarse a la cuestión de aportar algo de ayuda en la realización de gráficas.

Un punto de partida que siempre resulta interesante es comenzar por la definición oficial. Así, el diccionario de la Real Academia³ dice respecto de “gráfica”:

gráfico, ca.

(Del lat. *graphicus*, y este del gr. *γραφικός*).

1. adj. Pertenciente o relativo a la escritura y a la imprenta.
2. adj. Dicho de una descripción, de una operación o de una demostración: Que se representa por medio de figuras o signos. U. t. c. s.
3. adj. Dicho de un modo de hablar: Que expone las cosas con la misma claridad que si estuvieran dibujadas.
4. m. Representación de datos numéricos por medio de una o varias líneas que hacen visible la relación que esos datos guardan entre sí.
5. f. **gráfico** (representación por medio de líneas).

Salvo la primera, las otras cuatro acepciones aportan ideas interesantes para lo que nos ocupa: representado por signos o figuras, expresado de una forma muy clara, representado por medio figuras o signos... Y ya la acepción 4ª es directamente una definición bastante precisa de las gráficas en las que aquí nos centramos.

El hecho de que los dibujos puedan aportar mucha claridad a una explicación, el que puedan hacer visibles relaciones o gradaciones, lleva a un posible análisis de las gráficas desde el punto de vista de la percepción, y ésta di-

2. Ver, en el capítulo sobre bibliografía, la sección tutoriales y galerías.

3. <http://www.rae.es/>

rectamente lleva a disciplinas como la psicología o las neurociencias⁴. Un camino sin duda interesante, pero que de nuevo aleja a nuestro redactor veraniego o a nuestro estudiante de ingeniería de mejorar en la tarea de realizar gráficas con un esfuerzo ajustado.

“Representación de datos numéricos por medio una o varias líneas...”. Otros diccionarios utilizan la palabra “coordenadas” en vez de líneas⁵. Esto recuerda al plano cartesiano: parejas ordenadas de números naturales que tienen asociado de una forma biunívoca un punto en una rejilla formada por dos ejes. Algo así estudiábamos en mi generación en octavo de EGB. Por un lado aquello queda muy lejos, por otro lado las construcciones matemáticas mínimamente formales resultan largas y engorrosas. Además, es un hecho aceptado entre los editores de libros de divulgación científica que cada fórmula matemática divide por dos el número de lectores de la obra. Esto me dejaría con algunos ingenieros y algunos redactores; probablemente muy motivados pero pocos.

Basta ojear un suplemento salmón de periódico para observar un buen número de tipos de gráficas diferentes. Lo del plano cartesiano podría abrir un camino a algunas de ellas, pero la gran variabilidad de tipos de gráficas haría el camino matemático enormemente largo⁶. Quizá entonces una enumeración de tipos de gráficas con sus características básicas... También siguiendo esta aproximación taxonómica hay textos disponibles. En el fondo esta aproximación lleva también a algo bastante parecido al libro de cocina: algo excelente como elemento de consulta para el que sabe, pero un mal comienzo para el aprendiz.

Otro punto de partida que siempre da sus frutos es el desarrollo histórico: ¿Desde cuándo se utilizan gráficas? ¿Son un invento reciente con autor conocido o es una construcción ancestral? A diferencia de lo que se podría pensar, las primeras gráficas se atribuyen al trabajo de Playfair a finales del siglo XVIII. Disciplinas relacionadas con la representación visual de entidades

4. Ver, por ejemplo, “percepción en visualización”: <http://www.csc.ncsu.edu/faculty/healey/PP/index.html>

5. http://diccionarios.elmundo.es/diccionarios/cgi/lee_diccionario.html. En este ponen “Representación de datos numéricos por medio de coordenadas o dibujos que hacen visible la relación o gradación que esos datos guardan entre sí: la gráfica de una ecuación”.

6. O bien muy abstracto, ya que sólo utilizando conceptos abstractos se pueden enunciar principios de aplicación a diversos tipos de gráficas. En todo caso difícil de seguir por los menos expertos.

físicas como la cartografía, o con el manejo de conceptos muy abstractos como las matemáticas contaban con milenios de antigüedad para cuando a alguien (Playfair) se le ocurrió utilizar la metáfora del reparto de una tarta redonda para presentar porcentajes. Apasionante, tanto que no renunció a incluir un breve resumen de esta historia en un anexo⁷. Pero de nuevo estamos en un punto que no conduce a un camino breve hacia la evitación de errores.

La lista de posibles puntos de partida, seguramente no exhaustiva, que hemos repasado incluiría⁸:

1. Listas de errores (o aciertos): recetas.
2. Definiciones, aproximación académica.
3. Historia de los gráficos, otra aproximación académica.
4. Aproximación psicológica, perceptiva.
5. Aproximación taxonómica: tipos de gráficos y sus características.
6. Aproximación matemática, constructiva.

Todos ellos son puntos de partida hacia la solución a preguntas como: ¿qué es una gráfica? Y ¿cómo hacer bien una gráfica? Parece pues que estas preguntas llevan a respuestas demasiado largas o complejas como para resultar útiles a personas que no son profesionales de las gráficas. Y a pesar de ello seguimos siendo muchos los que necesitamos realizar gráficas de vez en cuando: estudiantes, periodistas, vendedores, profesores, economistas, científicos... Y sería estupendo contar con algunas guías sencillas que ayudaran a hacerlo bien.

Quizá un camino fructífero se encuentre analizando esa necesidad amateur: ¿para qué necesitamos realizar gráficos? O, dicho en forma impersonal:

7. Una breve historia de las representaciones gráficas se incluye como anexo 1. El tema, como se verá, es muy interesante, pero en este punto distrae del hilo que se pretende.

8. El excelente libro de L. WILKINSON, *The Grammar of Graphics* dedica un tercio del capítulo introductorio a explicar en qué no consiste el sistema que propone. Algo parecido es lo que se presenta a continuación, salvando las distancias, que son muchas, ya que en este texto, a diferencia de Wilkinson, no se propone una gramática formal (y menos “la” gramática) para la construcción o análisis de gráficas.

Para qué sirve una gráfica

Buscando luz sobre esta cuestión se me ocurrió preguntarles a estudiantes universitarios que habían entregado trabajos⁹, bien como memoria de realización de prácticas, bien en otros tipos de trabajos. Un buen número de respuestas se podían resumir de la siguiente forma:

- Queda más bonito.
- Lo mandaba el profesor.
- Todo el mundo lo hace.
- Ocupa una porción de espacio con lo que acabo antes (o relleno más).
- No lo sé.

Desde un punto de vista subjetivo todas estas razones son legítimas de alguna manera (salvo la última, que muestra una involuntariedad excesiva), pero es bastante probable que estos estudiantes, con estas motivaciones, hayan realizado gráficas inadecuadas como las mostradas anteriormente.

Salvo en situaciones en las que el autor no sabe para qué hace la gráfica (o lo intuye de una forma tan vaga que es como si no lo supiera), el resultado suele dar satisfacción al objetivo buscado. El proceso de contemplar la tarea realizada para ver si cumple con las expectativas puestas en él requiere un nivel de competencia al alcance de cualquiera, y se lleva a cabo incluso de forma inconsciente; pero es imprescindible tener alguna expectativa para poderla realizar.

De las razones del listado anterior hay dos respuestas que no aportan ninguna expectativa sobre el aspecto final, que no suponen nada respecto al resultado obtenido con la representación gráfica. Incluir una gráfica porque se supone que lo ha mandado el profesor (o el redactor jefe, o lo espera el consejo de administración, etc.), o porque todos los colegas lo hacen habitualmente, no ofrece la menor referencia sobre el aspecto que debe tener dicha gráfica.

Si lo que se busca es aportarle un valor estético al documento o aumentar su volumen, entonces sí que se dispone una expectativa sobre el resultado final, quizá no la más adecuada, pero al menos existe una expectativa.

9. Sería interesante preguntar a otro tipo de autores como periodistas en sus periódicos, científicos en sus artículos, etc. Pero de momento no me he atrevido.

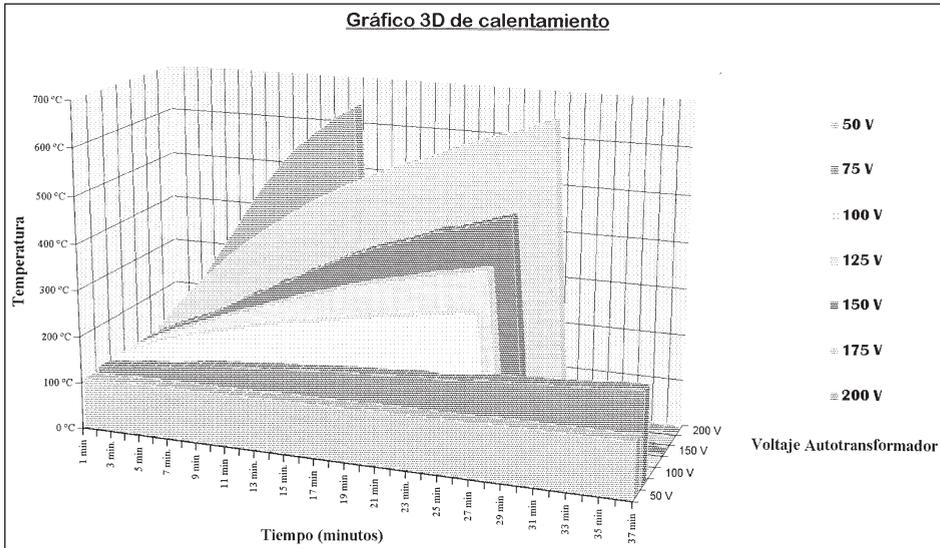


Figura 4. Ejemplo de gráfica realizada con una intención exclusivamente estética (a ojos del autor). Los datos representados ya habían sido analizados en el documento del que procede, un guión de prácticas de estudiantes, en otras gráficas más adecuadas.

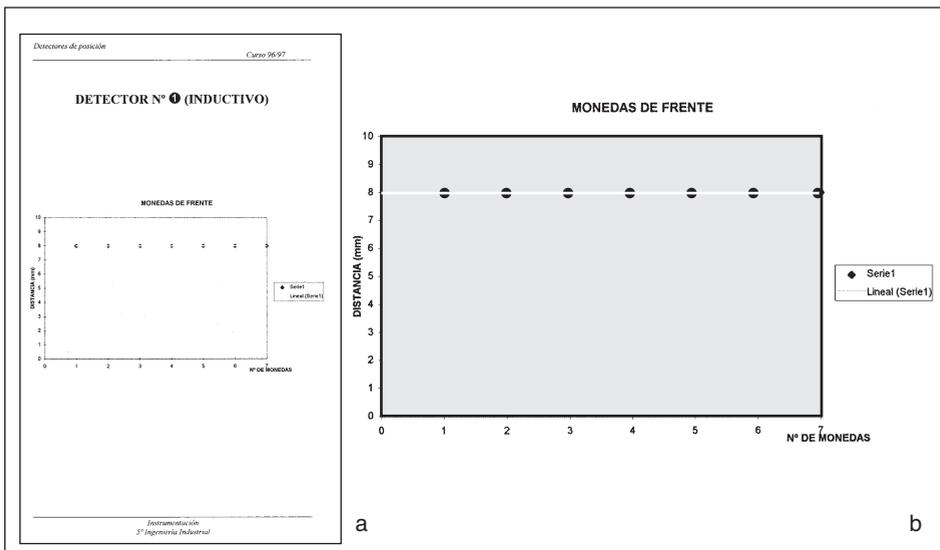


Figura 5. Ejemplo de una gráfica realizada con la sola intención de rellenar espacio. La figura a muestra la página completa, la b amplía la gráfica. Con siete puntos se rellena un folio completo.

Si lo que se busca es rellenar espacio pueden encontrarse gráficos, como el de la figura 5, en los que la cantidad de información ocupa un espacio muy pequeño en comparación con todo el arrebatado al texto. En la figura 5.a se muestra (reducida) la página completa del documento, y en la 5.b la gráfica. Para decir: “en siete medidas se obtuvo el mismo resultado” se gasta un folio entero, la expectativa se cumple extraordinariamente. Como el estudiante que dibujó esa gráfica hizo dos docenas de experimentos de ese tipo, consiguió entregar un impresionante documento de casi 30 páginas, tan costoso como absurdo.

Algunos estudiantes tuvieron respuestas algo más prometedoras a mi improvisada encuesta, y respondieron frases del tipo: “para mostrar los datos”. Esa es una la idea que se puede encontrar en la literatura del género (es la idea central de Tufte¹⁰). En una relación algo más detallada de la utilidad de las gráficas, Dürsteler¹¹ enumera:

El objetivo que persigue el gráfico. Los gráficos se pueden hacer por varias razones, entre las que destacan:

- **Comunicar un mensaje.** Las ventas han mejorado pero aún vamos por detrás del presupuesto...
- **Presentar grandes cantidades de información** de forma compacta y fácil de entender. Un mapa de carreteras es un ejemplo arquetípico de este tipo de objetivo.
- **Revelar los datos.** Descubrir relaciones de causa-efecto, conocer lo que pasa. Da la sensación de que en entornos de negocio se piensa más en los gráficos para presentar lo que ya se sabe que para descubrir lo que no se sabe.
- **Controlar la evolución** de parámetros de forma periódica. La evolución de la bolsa, las ventas, el presupuesto.

Analizando con un poco de profundidad todas las posibilidades (las antes citadas u otras que podamos plantear) se ve que en realidad siempre se están detallando casos particulares de comunicación.

Puede ser que el mensaje a transmitir no sea evidente y se haga necesario mostrar los datos en un formato que “revele” relaciones no evidentes a priori. O puede ser que los datos a representar sean dinámicos y que, por tan-

10. E. TUFTE, *The visual display of quantitative information*. Ver apartado de referencias.

11. www.Infovis.net (revista, mensaje 108).

to, el mensaje vaya cambiando con el tiempo¹². O puede ocurrir que el receptor del mensaje sea el mismo emisor que genera la gráfica, y que estemos asistiendo a una curiosa situación reflexiva de “autocomunicación”¹³. Sea cual sea la situación, de entre las anteriores, en todos los casos la podemos describir como un caso particular de comunicación: un emisor codifica un mensaje soportado por unos datos¹⁴ en un formato gráfico para que sea percibido por un receptor.

Contemplar las gráficas como elementos de comunicación es la idea fundamental. El único criterio general que se puede utilizar para decidir sobre la conveniencia o adecuación de una representación gráfica en un caso particular va a ser su efectividad como vehículo del mensaje que se pretende transmitir. Al igual que veíamos cómo una gráfica hecha para llenar espacio habría de ser grande, una gráfica hecha para comunicar debe hacerlo de forma efectiva. Y esa efectividad comunicativa la podemos juzgar tanto en una gráfica de tarta como en una de dispersión tridimensional, como en cualquier otra. Además de versátil en cuanto a tipos de gráficas, lo es en cuanto a la profundidad del análisis: la eficacia comunicativa se puede limitar a una percepción intuitiva instantánea, o ser objeto de un estudio de mucho detalle.

Conclusión final del libro

Hemos visto pues que las gráficas correctas tienen una finalidad comunicativa, mientras que las incorrectas son fruto de motivaciones espurias, como el rellenar espacio, la imitación, o la hipotética ganancia estética de un documento. Podría terminar aquí el libro, la conclusión final ya se ha obtenido: ***para hacer bien una gráfica no hay más que tomar conciencia de su utilidad comunicativa y hacer lo necesario para que esta resulte mínimamente efectiva.***

12. Este sería el caso por ejemplo de gráficas de cotizaciones de valores en bolsa, en las que sobre representaciones predeterminadas se van colocando los datos automáticamente a medida que se generan.

13. Esta situación resulta muy interesante, y además de en el texto se trata separadamente en el anexo 2.

14. La exigencia de unos “datos numéricos” en este tipo de comunicación es clara, una especificidad de este tipo de comunicación. Aparece ya en la definición del diccionario, 4ª acepción de “gráfica” en el DRAE.

En efecto, a partir de ahí todo lo demás será glosar los términos que aparecen en esa conclusión y dar algunas pistas que ayuden a concretar eso de “hacer lo necesario para”. Sólo espero que tener claro el punto de llegada desde el primer momento permita captar el hilo conductor y de unidad a una serie de cuestiones que corren el riesgo de resultar, a pesar de toda esta disertación, una colección de recetas de cocina.

Capítulo 1

La comunicación gráfica

Hemos finalizado la introducción con la intención de ir glosando la conclusión final: ***para hacer bien una gráfica no hay más que tomar conciencia de su utilidad comunicativa y hacer lo necesario para que ésta resulte mínimamente efectiva.***

Un punto adecuado para comenzar a profundizar puede ser la cuestión de la comunicación. También podría serlo el hecho de “tomar conciencia”, pero ahí nos encontramos con una de las cuestiones más profundas y peliagudas de la filosofía y la psicología. Supongamos pues que para la toma de conciencia nos conformamos con la idea naïf de la asimilación, de la capacidad de repetir la idea, y pasemos a la comunicación. Y más en concreto a la comunicación gráfica; no olvidemos que aún antes de la conciencia tenemos “para hacer bien una gráfica”, y ese es el objetivo final de verdad.

En cualquier caso, es posible que el sustrato sobre la comunicación gráfica que aporta este capítulo resulte demasiado teórico para algunos. No hay problema en saltarse este capítulo, el resto del texto resulta perfectamente comprensible sin él.

Comunicación gráfica

Por comunicación denominamos el proceso mediante el cual un emisor transmite un mensaje a un receptor. Al medio físico que soporta el mensaje se le denomina canal, y para completar la imagen clásica de la comunicación falta recordar que este proceso tiene lugar en un entorno concreto. Estas ideas proporcionan un modelo del fenómeno de la comunicación, ilustrado en la figura 6 y que, aunque en apariencia es muy simple¹, proporciona un marco muy adecuado para profundizar en algunas cuestiones de interés.

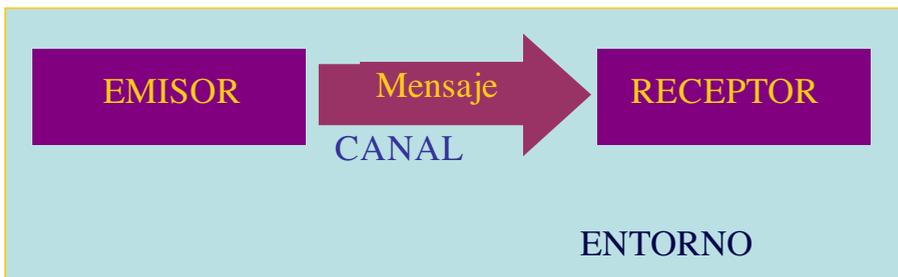


Figura 6. Esquema clásico del proceso de comunicación.

Cuando pensamos en la comunicación la primera idea que nos viene a la mente es la comunicación hablada: dos personas conversando bien frente a frente o quizá por teléfono. En este caso el mensaje está codificado en sonidos, mediante un lenguaje, y es emitido en forma de ondas sonoras, que constituirían el canal. Esa misma codificación en palabras se puede cambiar de canal si las escribimos. Tanto en el caso hablado como en el escrito es el lenguaje la herramienta que nos permite convertir el mensaje abstracto, que está inicialmente en la mente del emisor y debe pasar al receptor.

El lenguaje natural es un potentísimo sistema que permite expresar virtualmente cualquier idea que pueda existir. Probablemente los milenios de evolución del ser humano desde que adquirió esa capacidad han hecho que se convierta en una herramienta tan inextricable del pensamiento que es difícil imaginar mensajes inexpresables con el lenguaje y viceversa².

1. El modelo de comunicación podría incluir ruido en el canal, filtros culturales y muchos otros elementos, pero con los cinco indicados basta para la discusión que aquí nos ocupa.

2. Derek BICKERTON, *Lenguaje y especies*, Alianza Editorial, Madrid, 1990.

A pesar de su potencia y su indudable interés, lo que ahora nos interesa no es la comunicación basada en el lenguaje sino la basada en gráficos: es decir, aquellos procesos de transferencia de mensajes en los que éstos se codifican de forma gráfica. De una forma general podemos decir que hablamos de comunicación gráfica cuando el mensaje se materializa en forma de disposiciones geométricas de elementos que han de ser percibidos a través de la vista.



Figura 7. Un ejemplo de lenguaje gráfico.

De una forma más concreta podemos presentar ejemplos como los semáforos, las señales de tráfico, etc.³ En la figura 7 se presenta un ejemplo ilustrativo. Sobre una puerta hay colocado un símbolo que nos indica lo que nos encontraremos detrás de ella (unos aseos públicos) y otros dos símbolos que nos advierten de dos actividades que no se espera que realicemos en ese lugar.

Hay bastantes diferencias entre la comunicación basada en lenguajes y la comunicación basada en gráficos. El lenguaje permite una variedad mucho mayor de mensajes, así como un grado de matización que es difícil de imaginar utilizando gráficos. En la figura 8 se presenta una breve historieta (obviamente humorística), expresada tanto en palabras como de forma gráfica. Está claro que si viésemos la sucesión de símbolos sin ninguna explicación seríamos totalmente incapaces de reconstruir una imagen mental del mensaje transmitido con el texto. Esa es parte de la broma.

El mensaje de la figura 8 es demasiado largo y complejo para poderlo expresar de una forma gráfica efectiva. No es el caso de la figura 7, en la que los mensajes resultan bien claros, rotundamente claros. Recordemos la tercera acepción de la palabra gráfico del diccionario: “Se aplica a lo que expresa las cosas con la misma claridad que un dibujo: manifestó su negativa con un gesto muy gráfico”⁴. Cuando existe un dibujo ade-

3. En realidad los “lenguajes visuales” constituyen un ámbito de estudio que, si bien no es muy antiguo, ha desarrollado un corpus extenso y complejo. Ver por ejemplo: K. MARRIOT y B. MEYER (eds.), *Visual Language Theory*, ISBN 0-387-98367-8; Springer-Verlag 1998.

4. Del diccionario: http://diccionarios.elmundo.es/diccionarios/cgi/lee_diccionario.html

cuando a un mensaje, la potencia comunicativa del mismo es enorme; es en esos casos en los que una imagen vale más que mil palabras.

 One day, he was walking	 he saw a woman sleeping	 he felt desire burning inside him	 his adrenaline started pumping	 he took the plunge	 he invited her to have a coffee	 then to the restaurant	 they went on a trip
 they did different activities	 he took her to his house	 she told him she was on the pill	 and she laid down on the bed	 she spreaded one leg	 then the other	 then both	
 he reaction was immediate	 he penetrated her	 he went in and out	 he discovered that she wasn't a virgin	 he suggested some other positions	 she refused	 but she asked him to go faster	
 she made comments on his equipment	 When she saw all the colours of the rainbow,	 she shouted Stop!	 She hadn't told him the truth:	 she wasn't on the pill	 But he lost his self-control	 and reached the point of no return	
 she called him 9 months later	 from the hospital	 he had 2 children!	 his world crumbled	 he wanted to die	 The morale:	 for not making a woman pregnant	 wear protection

Figura 8. Un mensaje expresado tanto de forma gráfica como basada en un lenguaje natural (el idioma inglés).

Por otro lado, el lenguaje natural tiene una particularización cultural mucho mayor que el gráfico. Los símbolos que representan de forma abstracta en el primer caso son las palabras, mientras que en el segundo son iconos gráficos. Las palabras están estructuradas en idiomas (de los que existen alrededor de 4.000 diferentes en el mundo⁵) que requieren un proceso de aprendizaje largo, complejo y costoso. Los iconos gráficos son mucho más evidentes y sencillos. Algunos, como los semáforos u otras indicaciones de circulación vial, también requieren un aprendizaje que se desarrolla en un entorno cultural concreto. Aunque normalmente la amplitud del entorno cultural de validez significativa de estos iconos es mayor que la de los idiomas, ésta no será global.

5. http://www.fundacionflors.com/spip/article.php3?id_article=33

Casi con seguridad los nativos de las tribus amazónicas remotas no comprenderán el significado de un semáforo. Sin embargo, muchos de los iconos sí pueden ser de validez prácticamente universal. El primero de los símbolos que veíamos en la figura 7 (reproducido de forma aislada en la figura 9) sí que podría ser comprendido por los mismos indios amazónicos. En cualquier caso, aunque el icono sea reconocible, lo más probable es que nuestros indios del ejemplo no comprendieran el mensaje. Entenderían que la figurilla representa un niño orinando (más que una representación es una reproducción en miniatura de la escena real), pero seguramente no entenderían que el pedazo de madera rectangular al que está adherido es una puerta, ni esperarían encontrar tras ella un urinario público: puerta, bar y urinario son conceptos muy ajenos a su experiencia cotidiana.

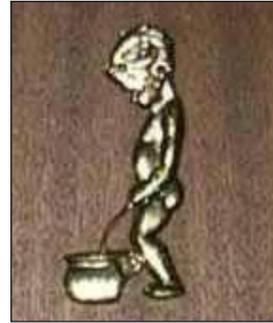


Figura 9. Ejemplo de icono gráfico de comprensión casi universal.

El ejemplo anterior nos trae a colación la importancia de algunos otros elementos que aparecían en nuestro modelo sencillo del proceso de comunicación (figura 6), en particular del entorno. El entorno en el que la imagen de la figura 2 adquiere su significado pleno es la puerta de los servicios de un local público. Por entorno entendemos toda la información disponible, perceptible por cualquier medio, que permite contextualizar el mensaje de forma que resulte significativo para el receptor.

Supongamos que un alcalde quiere que se construya en su municipio un campo de fútbol. Para ello contrata unos operarios de la localidad, les indica que deben buscar una extensión de hierba y les transmite unas instrucciones muy detalladas sobre las dimensiones que el campo debe tener y las líneas que lo delimitan. Si los operarios no tienen más información que las instrucciones recibidas (supongamos que son inmigrantes recién llegados de un remoto país) no sería extraño que el resultado de su trabajo fuera el mostrado en la figura 10. La información que no se transmitió resulta tan obvia (la extensión de hierba debía ser plana y despejada) que el alcalde la omitió, la supuso parte del entorno⁶.

6. El famoso “gag” de Tip y Coll sobre cómo se llena un vaso de agua es un ejemplo maravilloso de la inmensa cantidad de información tácita, que de tan obvia que resulta parece que no existe.



Figura 10. Resultado de la instrucción de construir un campo de fútbol sin suficiente información. Lo que faltaba era tan obvio (la superficie debía ser plana y despejada) que el emisor supuso que formaba parte del “entorno”.

El papel que el entorno juega en la comunicación gráfica es extremadamente importante. En realidad esto es válido para cualquier forma de comunicación. Si decimos “sube el gato al coche”, el mensaje es muy distinto si acabamos de cambiar una rueda que si emprendemos un viaje con la mascota en brazos. Pero en el caso de la información gráfica, precisamente por estar construida con iconos más difusos y reglas de construcción más débiles, la importancia de la contextualización, del entorno, es mayor si cabe.

Los desajustes entre el entorno en el que se percibe un mensaje gráfico en un momento dado y aquel para el que fue diseñado es causa de curiosos malentendidos que, en algunos casos, se convierten en chistes gráficos (figura 11).

Tomando en consideración ahora la construcción de los mensajes, nos encontramos con la existencia de reglas para la manipulación de los símbolos de forma que estos se constituyan de verdad en mensajes efectivos, inteligibles. En el caso de la comunicación basada en lenguajes naturales estas reglas no sólo son explícitas, sino que conforman un cuerpo de conocimiento exten-

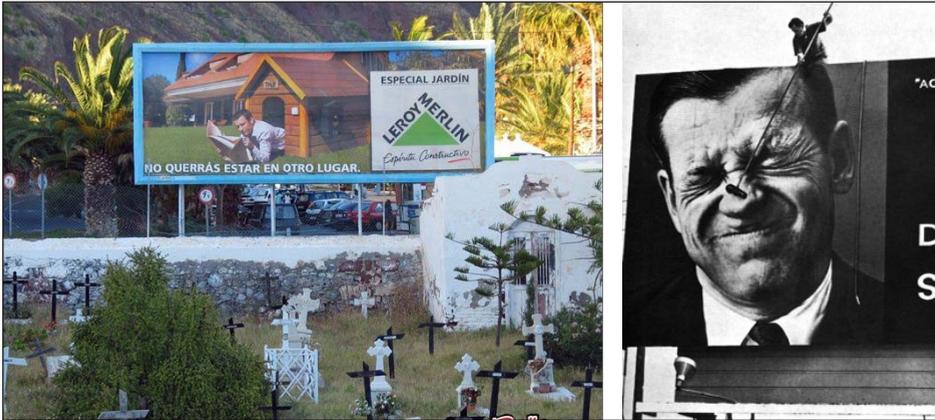


Figura 11. Dos chistes gráficos en los que la situación graciosa procede de la percepción del mensaje gráfico en un contexto para el que no fue diseñada.

so y profundo denominado gramática. A lo largo de todo el período de escolarización todos los niños reciben una extensa enseñanza formalizada de la gramática de su lengua materna. También se estudia ordenadamente la gramática de las lenguas extranjeras que uno aprende a lo largo de la vida⁷. Así, los conceptos de léxico, ortografía, sintaxis o semántica no nos resultan ajenos.

En el caso de la comunicación gráfica también existen reglas para la construcción de los mensajes a partir de los iconos que simbolizan las ideas primarias. Incluso la creación de estos iconos es también un proceso que está sometido a algunas reglas. Es relativamente sencillo diseñar nuevos iconos. También es posible crear nuevas palabras en los lenguajes naturales. De hecho ocurre a menudo, pero no es un proceso tan evidente, y es necesario un proceso de difusión cultural de las nuevas creaciones para que éstas sean aceptadas. En la figura 12 se presenta una señal de tráfico nueva, que no estudiamos al sacar el carné de conducir y sin embargo es claramente comprensible. Esto es así porque utiliza las reglas típicas de la construcción de las señales de tráfico (tamaño, forma, color, tipologías y textos...), pero añadiendo un icono nuevo (el de “abducciones extraterrestres”) que, aun no habiéndolo visto nun-

7. Este cuerpo de conocimiento es tan extenso que da lugar a los estudios universitarios de filología y a la profesión de lingüista.

ca antes, resulta de evidente comprensión. Aunque existe, esta “gramática gráfica” no está ni mucho menos tan estructurada y formalizada como la de los lenguajes naturales.



Figura 12. Un ejemplo de utilización de las reglas de construcción de mensajes gráficos.



Figura 13. Semáforos que, aunque en un contexto adecuado, obviamente no respetan las reglas que hacen que se construyan mensajes gráficos inteligibles.

Además de reglas de construcción de elementos comunicativos gráficos, existen reglas de utilización. Como se ha mencionado con anterioridad los semáforos son unos artefactos muy comunes que, mediante un código de colores iluminados, transmiten los mensajes gráficos necesarios para ordenar el tráfico. Esto es así siempre y cuando se respeten unas reglas para su utilización. Al igual que ocurría al analizar el entorno, las reglas resultan tan obvias que normalmente no reparamos en ellas, pero eso no quiere decir que no estén ahí: deben colocarse en las intersecciones de

las calles, deben ser visibles por los conductores cuya circulación se pretende regular... Es más fácil de hacer que enumerar un conjunto suficientemente extenso de reglas, como lo demuestra la evidente incorrección “gráfico-gramatical” mostrada en la figura 13.

De una forma un tanto intuitiva, apoyada por figuras humorísticas que dan vueltas por internet, hemos visto la existencia de una comunicación basada en elementos gráficos, que da lugar a lenguajes visuales en los que existen reglas para la construcción de los elementos comunicativos y para su utilización. Hemos visto también algunas similitudes y diferencias entre los lenguajes visuales y los léxicos. Pero además de ideas intuitivas, estas cuestiones se estudian seriamente. A continuación vamos a adentrarnos en

una formalización de los lenguajes gráficos que nos permitirá identificar de forma precisa las “gráficas” que nos interesan fundamentalmente en este libro dentro del proceloso océano de la comunicación gráfica en general.

Estructura básica de la comunicación gráfica

Como se ha mencionado ya la gramática de los lenguajes naturales está estructurada y bien establecida desde hace siglos⁸, mientras que las primeras conceptualizaciones sobre la gramática de los lenguajes gráficos son de las últimas décadas⁹. La incorporación social del canal de comunicación que ofrecen las nuevas tecnologías de la información y las comunicaciones, en particular la aparición de la red internet, ha despertado un interés renovado por el estudio del lenguaje gráfico, y así se pueden encontrar unas cuantas obras recientes que profundizan en estas cuestiones¹⁰.

Las raíces últimas del estudio de la información gráfica o información visual se encuentran en la psicología, dado que los procesos elementales que participan en ella son los relativos a la percepción (la captación inicial del mensaje visual) y a la cognición (la generación de una imagen mental operativa a partir de lo percibido).

Siguiendo a Yuri Engelhardt¹¹ podemos plantearnos un lenguaje visual particular como un “esquema de notación”¹², un nombre más técnico que nos permitirá precisiones. Los seres humanos utilizamos distintos tipos de esquemas mentales para estructurar cognitivamente nuestras experiencias, y los “esquemas de notación” son los esquemas mentales que tratan con el uso de representaciones.

El esquema de notación de una representación gráfica está compuesto de dos partes: un “esquema de componentes” y un “esquema de composición”. El esquema de componentes engloba la información sobre los componentes

8. La primera gramática de la lengua castellana es la de Antonio de Nebrija, de 1492.

9. La primera de la que tengo noticia es Jaques BERTIN, “Sémiologie Graphique”, 1967.

10. Una gran compilación de artículos y referencias bibliográficas se puede encontrar en www.infovis.net. Ver también el capítulo de referencias.

11. Yuri ENGELHARDT, <http://www.iaa.upf.es/activitats/semirec/LanguageofGraphics/> “The language of graphics: the lecture”.

12. La psicología cognitiva nos da una definición de esquema de notación como un esquema mental para tratar con representaciones visuales. En inglés “notational schema”.

visuales que aparecen en la representación gráfica (es el “vocabulario visual” de la representación), y sobre el significado de esos componentes visuales.

El esquema de composición, por otro lado, incluye la información sobre los principios de composición visual según los cuales se deben combinar los componentes visuales (sería la “gramática visual” de la representación), y sobre el significado de esos principios de composición.

Como veíamos anteriormente de una forma más intuitiva, existen algunas normas para generar iconos y otras para combinarlos. Técnicamente las primeras se conocen como “esquema de componentes” y las segundas como “esquema de composición”. Y el conjunto de ambas (el esquema de notación) es el análogo a la gramática del lenguaje gráfico particular considerado.

El diseñador de la señal de abducción de extraterrestres mostrada en la figura 12 tenía un esquema de notación en mente cuando la creó, y posteriormente quienes la observan tendrán (deseablemente) el mismo esquema de notación, o parte de él, cuando interpretan el significado de la señal. Este proceso se muestra esquemáticamente en la figura 14.

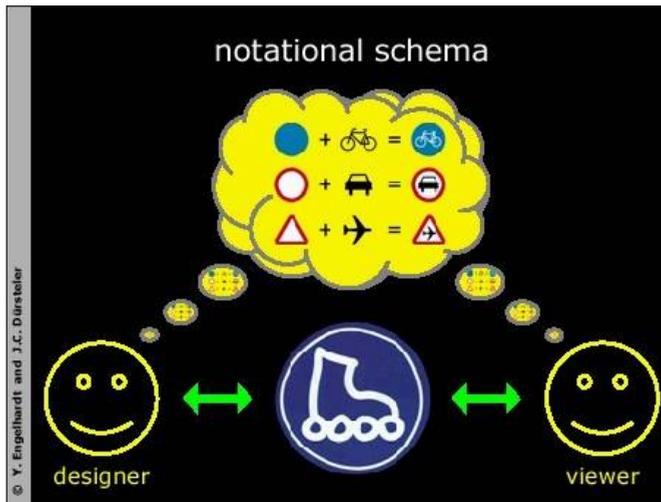


Figura 14. Representación gráfica de un “esquema de notación” (concepto subyacente a las representaciones gráficas).

Tras recopilar multitud de mensajes gráficos y analizarlos, intentando desmenuzar al máximo sus reglas de construcción, Engelhardt llega a la con-

clusión de que existen un número bastante reducido de esquemas de composición básicos o fundamentales, en concreto siete. Estos esquemas básicos o puros, se combinan a su vez mediante una serie de mecanismos (combinación simultánea, anidamiento, etc.) también limitados y conocidos. De este modo todas las relaciones entre los iconos aislados (componentes) que aparecen en un gráfico quedan descritas por un conjunto pequeño de procesos básicos. Cada uno de los siete esquemas de composición descritos se relaciona con una característica perceptiva geométrica sencilla:

1. Secuenciación lineal.
2. Enlace (mediante líneas o flechas)¹³.
3. Separación.
4. Agrupación por proximidad.
5. Contención (cerramiento en cajas).
6. Superposición.
7. Eje métrico.

Los distintos esquemas de composición se basan en disposiciones geométricas sencillas cuya interpretación surge de forma automática, instintiva, en base a una dilatada experiencia cotidiana. La secuenciación lineal la percibimos al hacer cola para subir al autobús, al escribir una palabra detrás de otra, incluso el devenir del tiempo como secuenciación de eventos lo visualizamos en una línea. La misma familiaridad por proximidad con las vivencias diarias hacen significativos de forma intuitiva a todos y cada uno de los esquemas de composición listados.

Este análisis es aplicable a los lenguajes visuales, a la comunicación gráfica en general, pero nuestro objetivo aquí es bastante más modesto. Por interesantes que resulten, no vamos a tratar ni de señales de tráfico ni de mapas topográficos; nos interesan las gráficas. Como veremos, las gráficas son un caso particular de comunicación que se enmarca en el séptimo caso de la lista: en el esquema de composición de eje métrico.

Merece pues una especial atención dicho esquema de eje métrico. El ejemplo típico es el reloj analógico, en el que la posición de una aguja sobre la escala graduada de la esfera indica la hora (figura 15). Vemos que, además de los “componentes de información” (los iconos significativos), es necesaria la

13. Casos particulares familiares del esquema de enlace serían los árboles (los organigramas de empresas) o las redes (mediante las que se representan los circuitos eléctricos).

existencia de un “componente de referencia”, en el ejemplo la escala grabada en la esfera del reloj.



Figura 15. El reloj como ejemplo de representación gráfica basada en el esquema de composición de eje métrico.

Además de la presencia de una escala graduada (de ahí el nombre del esquema: de eje métrico) se puede comprobar que el tipo de información que se expresa mediante esquemas de este tipo siempre incluye datos numéricos. Este esquema es especialmente adecuado para expresar información cuantitativa¹⁴.

Nos encontramos por tanto con que en el esquema de composición gráfica de eje métrico disponemos de algún dato cuantitativo y de un elemento de referencia (algún tipo de escala): lo que se está representando mediante este esquema gráfico es el proceso de medida. Aunque resulte un concepto más abstracto que la secuenciación lineal o el cerramiento en cajas, el proceso de medir es también muy natural (siquiera sea de forma rudimentaria). La medida de longitudes a palmos, pies o pasos seguro que es milenaria.

El diccionario de la Real Academia, en su primera acepción del verbo medir dice: Comparar una cantidad con su respectiva unidad, con el fin de averiguar cuántas veces la primera contiene la segunda. La definición encaja perfectamente con la necesidad de incluir un elemento de referencia (la respectiva unidad) junto con el elemento de información (la cantidad a comparar); y también con la existencia de información cuantitativa, como el resultado del proceso de comparación.

14. Esta relación no es biunívoca, ya que se pueden expresar algunos mensajes con contenido cuantitativo en formatos gráficos pertenecientes a otras categorías. Sin embargo la relación directa sí es cierta, ningún gráfico expresado en el esquema de eje métrico puede carecer de datos numéricos.

Las gráficas que nos interesan en este texto son las figuras estándar de representación de información cuantitativa estadística, científica, económica, etc. En la literatura se pueden encontrar denominaciones como “gráficos cuantitativos”, “gráficas estadísticas” y otros equivalentes para referirse a este tipo de construcciones. El paradigma sería la representación bidimensional clásica en el plano cartesiano (figura 16) o gráfica de dispersión, pero también lo son los gráficos de barras, las tartas, y muchos otros. Para evitar confusiones, a partir de ahora el término “gráfica” (utilizado como sustantivo y en femenino) hará siempre referencia a ese tipo de entidades, e intentaremos definir las de una forma algo más precisa en el siguiente capítulo.

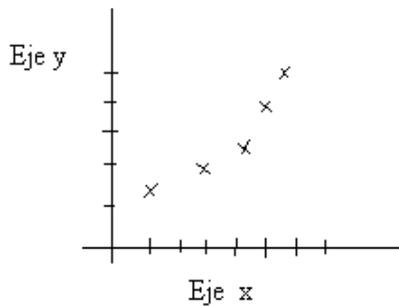


Figura 16. Gráfica cuantitativa en el plano cartesiano.

Sean de barras, tartas o de dispersión, las gráficas comparten el hecho de presentar datos numéricos y el de requerir escalas u otros elementos de referencia que permitan apreciar los valores numéricos representados. Esas características son precisamente las del esquema de composición de eje métrico dentro del esquema general que hemos visto anteriormente. Concluimos pues que las gráficas que nos interesan aquí son casos particulares de elementos de lenguajes visuales basados en el esquema de composición de eje métrico.

El lenguaje de las gráficas

Con lo que llevamos visto hasta ahora, podemos plantearnos la idea de que las gráficas constituyen un lenguaje, al menos en dos sentidos muy concretos. Por un lado su utilidad esencial, la motivación de su generación como

elementos de comunicación. Y por otra parte, la existencia de reglas más o menos definidas que rigen su construcción.

Así pues, de entre los lenguajes gráficos, aquí nos va a interesar el subconjunto que se ocupa de “las gráficas para la presentación de datos cuantitativos mediante el esquema de composición de eje métrico”, o de forma abreviada, el lenguaje de las gráficas. Y en particular, vamos a acercarnos a él intentando explotar al máximo la analogía entre el lenguaje natural, la lengua que hablamos a diario y el lenguaje de las gráficas. Como todas las analogías, vale en lo que vale y llevada al extremo dará lugar a absurdos. Sin embargo es un procedimiento muy potente que, con poco esfuerzo, nos permite reutilizar un conocimiento que ya tenemos (el de la lengua hablada) en otro contexto¹⁵.

Tanto el texto de Wilkinson como la mayoría de los existentes sobre esta cuestión abordan el estudio de las gráficas describiendo su proceso de síntesis, centrándose en los pasos a seguir para su construcción. Aquí le vamos a dar la vuelta al enfoque y vamos a analizar gráficos ya hechos para ir desgranando de ellos principios generales.

De las muchas fuentes posibles de gráficas, la mayoría de las que se utilizarán en esta obra están tomadas de periódicos. Por una parte es una fuente accesible, familiar a cualquier lector, y que presenta una casuística con toda la riqueza necesaria. Por otra parte, el lenguaje de los periódicos se utiliza en ocasiones en estudios lingüísticos¹⁶ como una línea de referencia de la complejidad del lenguaje. De la misma forma, podremos considerarlo aquí un nivel de referencia para el lenguaje de las gráficas, manteniendo la analogía entre el lenguaje textual y el gráfico desde el primer momento. La procedencia precisa de cada una de las figuras está listada al final del texto.

15. Será pues esta una obra necesariamente de introducción que no pretende ni mucho menos agotar el tema.

16. Un ejemplo se puede encontrar en el artículo: Donald P. HAYES, “The growing inaccessibility of science”, *Nature*, vol. 356, 30 April 1992, pp 739-40.

Capítulo 2

Qué es una gráfica

Tras tomar conciencia de que las gráficas son elementos de comunicación, y analizar las ideas básicas sobre la comunicación en general y sobre la comunicación gráfica en particular, parece que va siendo hora de abordar la cuestión de qué es una gráfica. Como casi siempre, las definiciones formales y rigurosas no son sencillas, y requieren de mucho esfuerzo para incorporar algunos casos marginales. En el extremo opuesto, una frase como “un dibujo basado en un conjunto de números” resulta tan vaga que tampoco es satisfactoria.

En el diccionario, como ya vimos, se encuentra: “f. Representación de datos numéricos por medio de coordenadas o dibujos que hacen visible la relación o gradación que esos datos guardan entre sí”¹. En este capítulo vamos a profundizar en algunas de las ideas presentadas en esta definición.

1. Como vimos también, el diccionario de la Real Academia en vez de “coordenadas o dibujos” dice “una o varias líneas”. La diferencia no es del todo trivial, pero ahora no nos interesa este camino.

En primer lugar, la forma en que las representaciones “hacen visible” es, como vimos con anterioridad, mediante elementos de lenguaje visual basados en el esquema de composición de eje métrico.

En segundo lugar, profundizaremos en los “datos numéricos”, y plantearemos algunas ideas sobre la estructuración de esa materia prima anterior a la gráfica y que es su sustrato inexcusable.

Por último trataremos la cuestión de la “representación de datos”. Detrás del término representación se esconde un proceso, en ocasiones muy complejo, de establecer una relación entre elementos del conjunto de los datos y elementos gráficos perceptibles visualmente en el dibujo final. Esa relación entre elementos de dos conjuntos (el de los datos y el de los elementos gráficos) se puede plantear matemáticamente como una aplicación, que se podría caracterizar con rigor². Aquí nos vamos a quedar sólo con una idea simple: la relación debe ser reversible (la aplicación biunívoca, como dirían los matemáticos).

Un elemento de comunicación gráfica con esquema de composición de eje métrico

Hemos visto ya cómo los elementos gráficos de comunicación se estructuran en lenguajes gráficos que se pueden contemplar como “esquemas de notación”, y que éstos a su vez incluían un esquema de componentes y un esquema de composición. Esto es una expresión técnica, en términos de la psicología cognitiva, del hecho de que la comunicación gráfica está basada en reglas, tanto para construir elementos como para utilizarlos.

De las ingentes cantidades de elementos gráficos de comunicación, a nosotros nos interesan aquí las gráficas, un tipo particular que, como vimos en el capítulo anterior, está basado en el esquema de composición de eje métrico. Esto supone que no cualquier dibujo que incluya números lo podemos denominar “gráfica”. Un buen ejemplo lo encontramos en la figura 17.

En esta figura se presentan datos sobre centros comerciales por comunidades autónomas, en concreto su número, la superficie que ocupan y el número de locales que contienen. La forma de mostrarlos es presentando una

2. Ver por ejemplo: <http://www.spss.com/research/wilkinson/Publications/gpl.pdf>

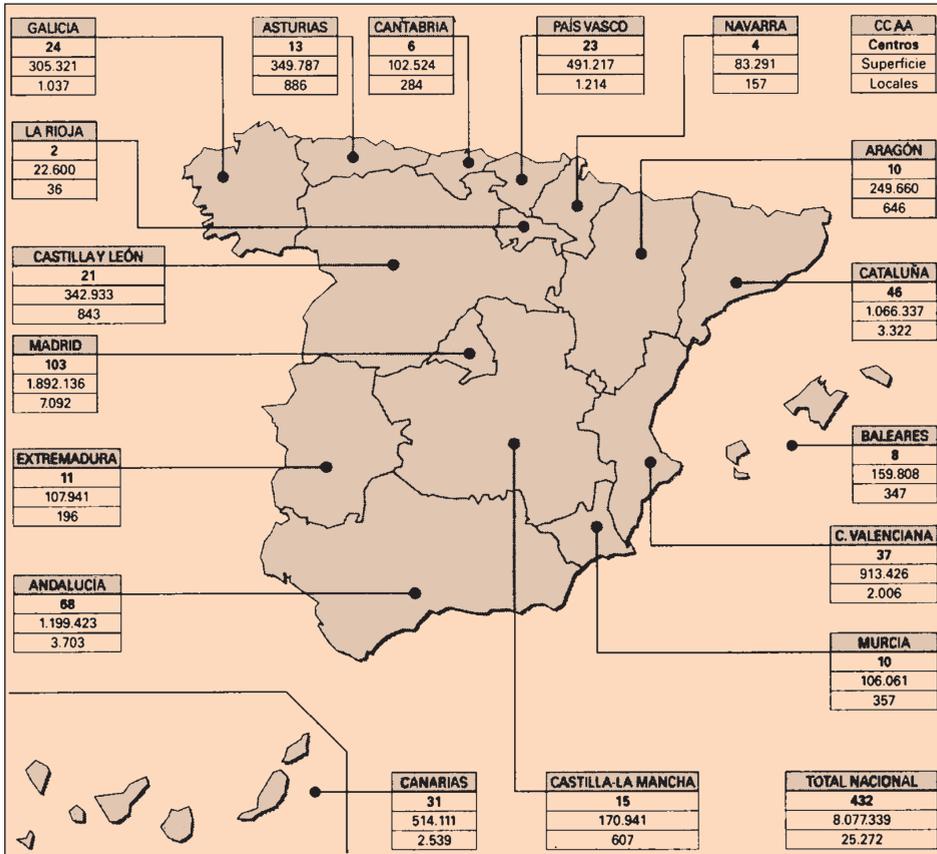


Figura 17 Infografía publicada en un periódico en la que se muestra información sobre centros comerciales por comunidades autónomas. Para nosotros esto no será una gráfica.

pequeña tabla para cada comunidad y asociarla mediante una línea a la comunidad en cuestión sobre un mapa de España.

Sin duda este es un elemento gráfico de comunicación, y la información que aporta va más allá de la que obtendríamos de una tabla simple con los mismos datos. Por ejemplo, a simple vista se aprecia sobre el mapa el tamaño relativo de las comunidades, que en relación con el número de centros da una idea de densidad que no se obtendría sin el mapa. El lenguaje visual utilizado en el elemento mostrado en la figura 1 se basa en el esquema de composición

“de enlace”³, la metáfora visual es la de asociación de elementos mediante una línea. En este caso uno de los elementos enlazados contiene datos cuantitativos⁴, pero éstos están más bien presentados que “representados”: se muestran tal cual, sin estar asociados a un elemento gráfico que los sustituya (que los represente).

Como ya hemos dicho, la particularidad que le vamos a pedir a un dibujo para que podamos considerarlo una gráfica es que el lenguaje visual presentado utilice el esquema de composición que denominamos “de eje métrico”: la metáfora visual esencial en la que se basan las gráficas es el proceso de medida.

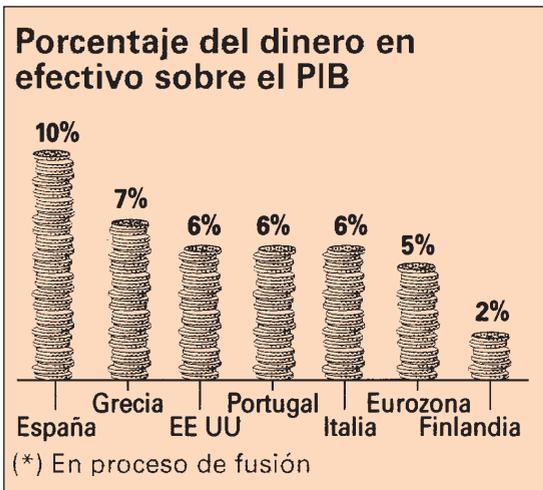


Figura 18 Gráfico de barras en el que se acentúa la metáfora visual de comparación de alturas representando las barras como pilas de monedas.

Esa metáfora visual puede referirse de una forma muy pura al proceso de medir, como en el caso de un gráfico de barras (ver figura 18), en el que las barras se comparan unas con otras igual que niños pequeños que miden su altura unos frente a otros para ver cuál ha crecido más⁵. Puede referirse

3. Los esquemas de composición los hemos discutido en el capítulo anterior siguiendo a Engelhardt: <http://www.iaa.upf.es/activitats/semirec/semi-engelhardt.htm>

4. También el mapa, obviamente, contiene información cuantitativa (posiciones, áreas, distancias...), pero los mapas no son gráficas en el sentido que nos interesa aquí.

5. Con o sin escala: dos niños se pueden comparar uno contra otro para ver cuál es más alto, o hacerlo marcando en la pared y midiendo cuánto más.

también a un proceso de medida que enfatiza el concepto de reparto, como ocurre, por ejemplo, en los gráficos de tarta en los que la analogía es evidente con el reparto de una tarta circular, hasta en el nombre que se les da. En cualquier caso, el proceso de medida está patente como analogía en cualquier elemento gráfico que asociamos al concepto de gráfica, de representación de datos.

Recordando una vez más la definición del diccionario, la gráfica supone una representación de datos con elementos “que hacen visible la relación o gradación que esos datos guardan entre sí”. Relación, gradación, reparto, comparación; todos son términos asociados al proceso de medida, y la metáfora visual del mismo es el esquema de composición de eje métrico.

Autores como Tufte hablan de representaciones cuantitativas (“quantitative representations”) o gráficos estadísticos, lo que incluye representaciones que aquí dejamos fuera; sería por ejemplo el caso de los mapas. Willkinson también considera en su Gramática Gráfica algunas representaciones que no pueden considerarse de eje métrico (aunque muy escasas). Entre el listado de algunos datos numéricos y una representación gráfica de los mismos en el sentido que aquí estamos considerando hay muchas otras posibilidades⁶. Sin duda son interesantes todas, pero el objetivo elegido aquí se centra en las representaciones gráficas tal como las venimos definiendo, basadas en esquemas de eje métrico⁷.

Incluye información cuantitativa: números, datos

Dado que, como hemos visto ya, las gráficas utilizan el proceso de medida como metáfora visual, la existencia de datos numéricos es inexcusable. Medir es comparar una cantidad con su respectiva unidad con el fin de averiguar cuántas veces la primera contiene la segunda. Ese número de veces es un dato cuantitativo: un número.

Entendemos por dato la cuantificación numérica de alguna magnitud. De momento no vamos a plantearnos ese proceso de cuantificación, sino que

6. Esto se discute en detalle en el anexo 3.

7. Además de que otros autores delimitan el ámbito de su estudio de otra forma, es interesante notar que en inglés existen más términos para referirse a representaciones gráficas (chart, plot, graph, graphic), lo que permite matizar más que en castellano (donde sólo disponemos de las dos últimas).

vamos a considerar un conjunto de datos como totalmente correcto e indiscutible. Se podría retomar esta cuestión dado que, al no ser perfecto ningún proceso de cuantificación, los resultados numéricos estarán afectados por alguna incertidumbre, algún error. El error con el que se generaron los datos también puede ser objeto de expresión geométrica en las gráficas, pero lo podemos considerar como un elemento aislable y obviarlo por ahora.

Para mostrar un solo dato es muy improbable, por no decir imposible, que una gráfica sea una buena opción como elemento de comunicación. Para decir que “en España había 40.847.371 habitantes según el censo de 2001” la frase en sí es la mejor forma de decirlo (ver anexo 2).

Ahora, si lo que queremos comunicar es la distribución de esa población por comunidades autónomas y provincias, comenzar con “De ellos, en Andalucía había 7.357.558, de los cuales en Almería habitaban 536.731, en Cádiz 1.116.491, ...” y seguir con un texto en el que ir intercalando todos los datos no parece que sea la mejor opción. Una forma más adecuada de mostrar esos números es de forma ordenada en una tabla (figura 19).

Así pues, a la hora de pensar en un conjunto de datos como sustrato de una gráfica, lo habitual es que esos datos los tengamos ordenados formando una tabla. Las distintas filas de la tabla son repeticiones de la misma estructura⁸ de datos. La estructura de la tabla recogida en la figura 1 tiene dos campos, dos columnas: cada dato individual está compuesto por dos entidades. El primero de los campos contiene un texto que hace referencia a la provincia o comunidad autónoma de la que se trata, mientras que el segundo contiene un número⁹ que indica la cantidad de habitantes de ese lugar (siempre según el censo del año 2001).

Los dos campos no contienen el mismo tipo de información, sólo el segundo contiene información cuantitativa, mientras que el primero contiene rótulos o etiquetas: categorías¹⁰ que sirven para saber a qué se refieren los datos cuantitativos. La estructura de datos mínima que puede contener una tabla susceptible de convertirse en gráfica es la del ejemplo: un campo categó-

8. Considerando los datos ordenados en columnas, como en la figura 3. También puede darse una tabla de la forma contraria, con la estructura de datos repetida por filas, pero es menos habitual.

9. Entero en este caso.

10. En la literatura sobre gráficas, y en los programas de ordenador diseñados para representarlas, es este término “categoría” el que se utiliza para los campos que contienen textos identificativos.

Cifras de población del Censo 2001			
ESPAÑA	40.847.371	Cataluña	6.343.110
Andalucía	7.357.558	Barcelona	4.805.927
Almería	536.731	Girona	565.304
Cádiz	1.116.491	Lleida	362.206
Córdoba	761.657	Tarragona	609.673
Granada	821.860	Comunidad Valenciana	4.162.776
Huelva	462.579	Alicante/Alacant	1.461.925
Jaén	643.820	Castellón/Castelló	484.566
Málaga	1.287.017	Valencia/València	2.216.285
Sevilla	1.727.803	Extremadura	1.058.503
Aragón	1.204.215	Badajoz	654.882
Huesca	206.502	Cáceres	403.621
Teruel	135.858	Galicia	2.695.880
Zaragoza	861.855	Coruña (A)	1.096.027
Asturias (Principado de)	1.062.998	Lugo	357.648
Baleares (Illes)	841.669	Ourense	338.446
Canarias	1.694.477	Pontevedra	903.759
Palmas (Las)	887.676	Madrid (Comunidad de)	5.423.384
Santa Cruz de Tenerife	806.801	Murcia (Región de)	1.197.646
Cantabria	535.131	Navarra (Cdad. Foral de)	555.829
Castilla y León	2.456.474	País Vasco	2.082.587
Ávila	163.442	Alava	286.387
Burgos	348.934	Guipúzcoa	673.563
León	488.751	Vizcaya	1.122.637
Palencia	174.143	Rioja (La)	276.702
Salamanca	345.609	Ceuta	71.505
Segovia	147.894	Melilla	66.411
Soria	90.717		
Valladolid	498.094		
Zamora	199.090		
Castilla-La Mancha	1.760.516		
Albacete	364.835		
Ciudad Real	478.957		
Cuenca	200.346		
Guadalajara	174.999		
Toledo	541.379		

Figura 19. Tabla de datos de población por provincias y comunidades autónomas según el censo de 2001. Fuente: INE, España en cifras (<http://www.ine.es/escif/escif/escif02.htm>).

rico y uno numérico¹¹. A partir de ahí, son posibles estructuras de datos más complejas.

11. Los estadísticos y los programadores de bases de datos informatizadas han desarrollado mucho la cuestión de los tipos de datos. Una referencia en la red, a título de ejemplo: <http://www.graphics.stanford.edu/courses/cs448b-02-spring/lectures/encoding/>

Podemos concluir hasta aquí que toda gráfica está asociada a un conjunto de datos numéricos, y que estos datos están organizados como la repetición de una estructura básica. Esta estructura contendrá campos, que pueden ser numéricos o categóricos, y vamos a definir (provisionalmente) la dimensión de una estructura de datos como el número de campos numéricos que contiene.

Desde un punto de vista constructivista nos podemos plantear ahora cuándo es conveniente utilizar una gráfica para presentar un conjunto de datos numéricos. Dürsteler¹² (a partir de Tufte) plantea que de los varios formatos posibles (frase, tabla y gráfico) la elección, sin ser sencilla, puede basarse en el número de datos a presentar. Así, la frase sería lo adecuado hasta 2 ó 3 datos, la tabla entre 3 y 20 y el gráfico entre 2 e infinito, resultando más útil cuantos más datos haya. En ese artículo Dürsteler presenta esa información (la del número de datos para cada medio de presentación) de las tres formas, y resulta muy interesante apreciar los matices en cada caso. Tufte¹³ plantea más categorías, como el texto tabular y el semigráfico, a la hora de elegir la forma de mostrar los datos.

Nosotros aquí hemos considerado el hecho comunicativo como hilo conductor, y en ese sentido debemos concluir que la gráfica¹⁴ será un elección adecuada cuando:

1. Exista un mensaje que transmitir.
2. Ese mensaje esté claramente soportado por los datos.
3. La gráfica lo transmita adecuadamente.

Una relación reversible entre los datos y su representación

Una gráfica supone una “representación de los datos”: un proceso, en ocasiones complejo, que permite establecer una relación entre elementos del conjunto de los datos y elementos gráficos perceptibles visualmente en el dibujo final. Esa relación entre elementos de dos conjuntos (el de los datos y el de los elementos gráficos) se puede plantear matemáticamente como una

12. J.C. DÜRSTELER, Infovis.net. mensaje nº 108 de 2 dic 2002.

13. E. TUFTE, *The visual Display of Quantitative Information*, cap 9.

14. Tanto Dürsteler como Tufte se refieren en este contexto con el término “gráfico” a modalidades de representación más amplias que las “gráficas” que consideramos aquí, como se ha mencionado antes.

aplicación. Sin entrar en más detalles sobre la caracterización matemática de aplicaciones, sí hay una característica relevante: la reversibilidad. El autor genera una representación a partir de los datos numéricos; el lector debe tener la posibilidad de revertir el proceso y obtener los datos a partir de la representación.

En todas las gráficas que se han presentado hasta ahora esta reversibilidad existe. Si tomamos la gráfica 20, por ejemplo¹⁵, podemos ver que la densidad de población de la Comunidad de Madrid es mayor que 6 y menor que 7 habitantes por hectárea. Incluso se ve claramente que es mayor que 6,5, situándose alrededor de 6,75 (de hecho el valor representado era 6,7556).

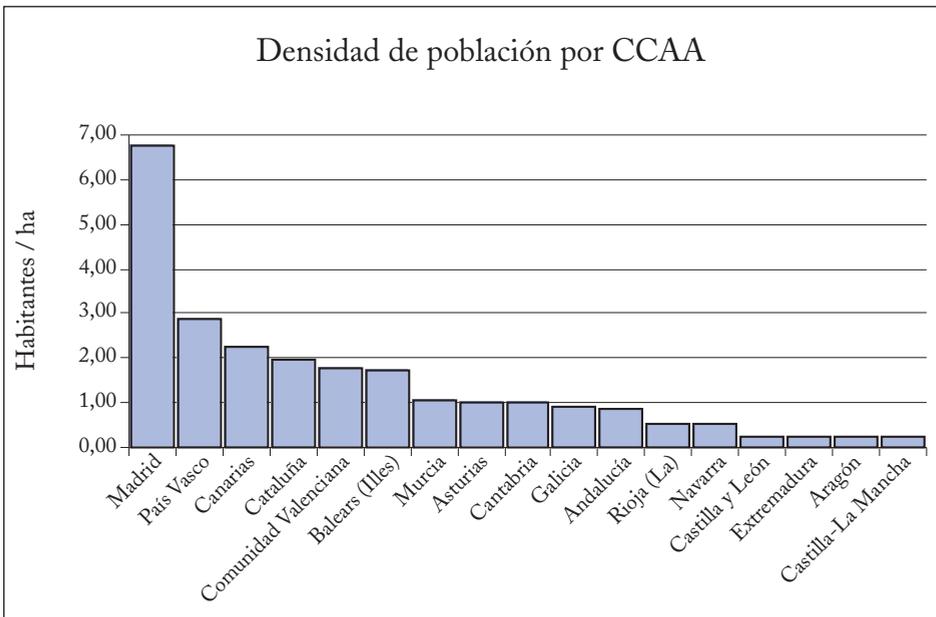


Figura 20. Gráfico de barras de la densidad de población de las comunidades autónomas españolas.

Esa reversibilidad lógicamente no es perfecta, una parte de la precisión que contienen los números que constituyen los datos originales se pierde a fa-

15. Esta gráfica aparece en el anexo 2, "Revelar los datos".

vor de la visualización gráfica de los mismos. Pero admitiendo un pequeño margen de imprecisión, vemos que en las gráficas es posible recuperar la información cuantitativa original.

Esta recuperación de la información original es esencial para que el receptor del mensaje, el lector, pueda profundizar en el análisis generando enunciados cuantitativos o incluso realizando sus propias representaciones, en las que busque enfatizar aspectos no muy evidentes en la gráfica original. Este tipo de estudio no es infrecuente en el caso de gráficas incluidas en artículos científicos.

Sin embargo, hay algunas representaciones muy extendidas en las cuales la reversibilidad entre los datos y su representación se ha perdido. El caso más paradigmático es el de las representaciones de tartas en perspectiva, y en general todas las representaciones en perspectiva, mal llamadas tridimensionales.

Siguiendo con el ejemplo de la sección anterior, nos puede interesar representar la distribución de la población española entre comunidades autónomas (o al menos entre las más pobladas). Una representación típica de este tipo de información es la que se muestra en la figura 21.

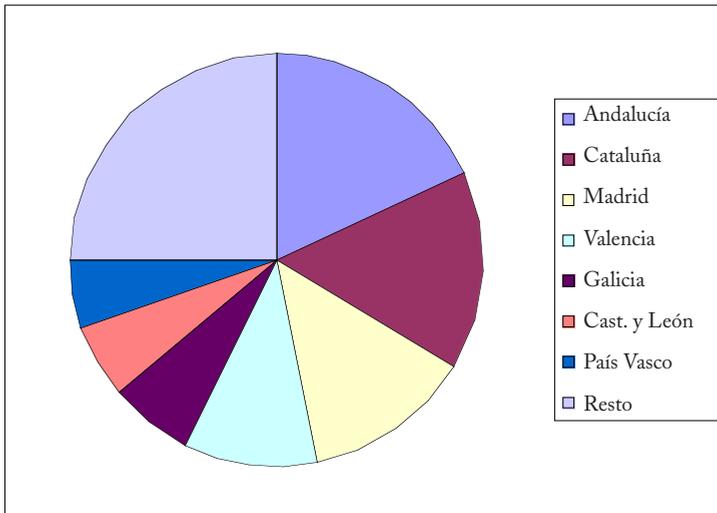


Figura 21. Gráfico de tarta en el que se presenta la distribución de la población española entre comunidades autónomas. Se muestran las 7 más pobladas, agrupando la población de las 10 restantes bajo el rótulo "resto".

Esta representación es una gráfica en el sentido de que cumple el criterio que estamos introduciendo en esta sección. La recuperación del porcentaje de población de cada comunidad no puede realizarse de una forma tan inmediata como en el caso de la gráfica 20, dado que no se incluye ningún elemento de referencia: no hay escala. Sin embargo basta tomar un transportador de ángulos, y de una forma sencilla se puede establecer la proporción entre 360° , que corresponden al 100% de la circunferencia, y los ángulos medidos, que nos llevará al valor original buscado.

No es infrecuente, por otra parte, que la representación que se realice no sea como la de la figura anterior sino como la mostrada en la figura 22, o en la figura 23. Las representaciones mostradas en estas dos figuras no podemos considerarlas formalmente como gráficas, dado que al presentar una cierta perspectiva de los sectores originales (los de la figura 21) que es desconocida para el lector, se imposibilita relacionar la medida angular de un sector con el valor del porcentaje que lo generó.

Por otra parte, este tipo de representaciones tampoco son adecuadas desde un punto de vista informativo dado que presentan una disfunción perceptiva importante. En realidad la característica que da lugar a la percepción del tamaño es el área observada, y el área de cada sector se ve modificada por la posición que ocupa respecto del ángulo de visión. En los ejemplos de las figuras 22 y 23, las representaciones son idénticas salvo el ángulo de comienzo de la serie de sectores. Esto hace que, por ejemplo, el sector azul oscuro (correspondiente al País Vasco) dé la impresión de ser bastante mayor en el caso de la figura 23 que en el de la figura 22. Se obtiene la impresión contraria para los sectores representativos de Valencia o Madrid, que parecen mayores en la figura 22.

Hemos descrito dos razones importantes, una formal y otra perceptiva, para que no resulte conveniente denominar “gráficas” a este tipo de representaciones. En realidad lo mejor es no utilizarlas en absoluto.

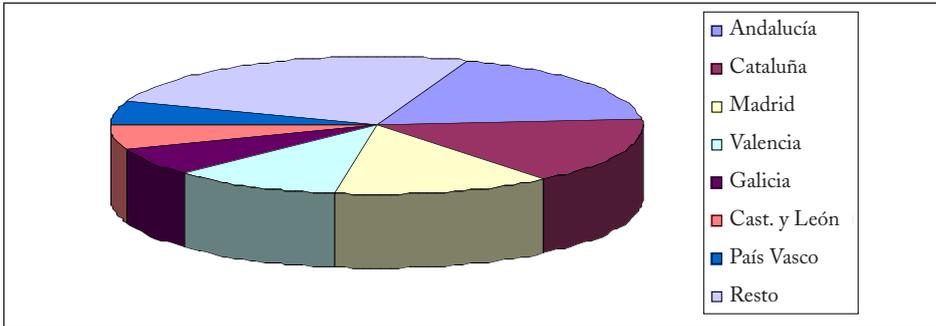


Figura 22. Representación análoga a la de la figura 10, que no podemos considerar formalmente una gráfica dado que se ha perdido la posibilidad de recuperar los datos originales.

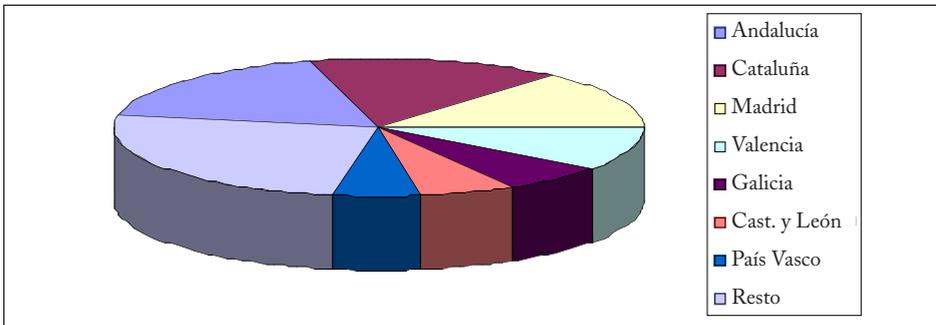


Figura 23. Representación análoga a la de la figura 11 en la que sólo cambia el ángulo de comienzo de la presentación secuencial de los sectores. Se puede apreciar como esa modificación altera sustancialmente la percepción relativa de los mismos.

Resumen del capítulo

En este capítulo hemos analizado tres características que le vamos a exigir a una representación para que podamos llamarla “gráfica” formalmente. Estas características son:

- Que se base en el esquema de composición de eje métrico.
- Que incluya datos: información cuantitativa.
- Que los datos se puedan recuperar a partir de la representación.

Las dos primeras características están muy relacionadas entre sí, son aspectos del hecho de que las gráficas presentan metáforas del proceso de medida. La tercera implica excluir toda representación en perspectiva (“tridimensional”) y en especial las tartas, dado que resultan equívocas y de datos irrecuperables.

Capítulo 3

El texto gráfico

El texto gráfico: Párrafos, oraciones simples y compuestas, oraciones subordinadas

Hemos visto ya con cierto detalle qué es lo que podemos denominar una gráfica, y se ha repetido de forma insistente su característica esencial de elemento de comunicación. En muchas ocasiones encontramos elementos más complejos, con los que se pretende transmitir un conjunto de ideas que excede la capacidad expresiva de una gráfica aislada. Lo que nos encontramos en estos casos es lo que en lenguaje periodístico se denomina una “infografía”: un conjunto de elementos gráficos (siendo gráficas muchos de ellos, incluso todos) diseñados como un todo con una intención comunicativa conjunta.

Con el fin de explotar al máximo la analogía del lenguaje de las gráficas con el lenguaje natural que utilizamos, podemos considerar que estas infografías son “textos gráficos”, o más concretamente “párrafos gráficos”, recordando las agrupaciones de oraciones consecutivas, separadas por puntos, en un texto convencional.

En las figuras 24, 25 y 26 se presentan ejemplos de párrafos gráficos compuestos exclusivamente de gráficas. Se puede ver cómo el elemento grá-

fico en su conjunto, el párrafo, se diferencia del texto léxico del periódico mediante un recuadro oscuro. Al envolver a todas las gráficas, este recuadro le aporta también unidad al elemento, percibiéndose las gráficas como partes de un conjunto. En los tres casos el recuadro presenta un título de conjunto al margen de los títulos que también están presentes en las gráficas individuales. Este título muestra con bastante claridad la intención comunicativa del párrafo como un elemento integrado, mientras que el título de cada una de las gráficas ilustra lo que en ella se representa.

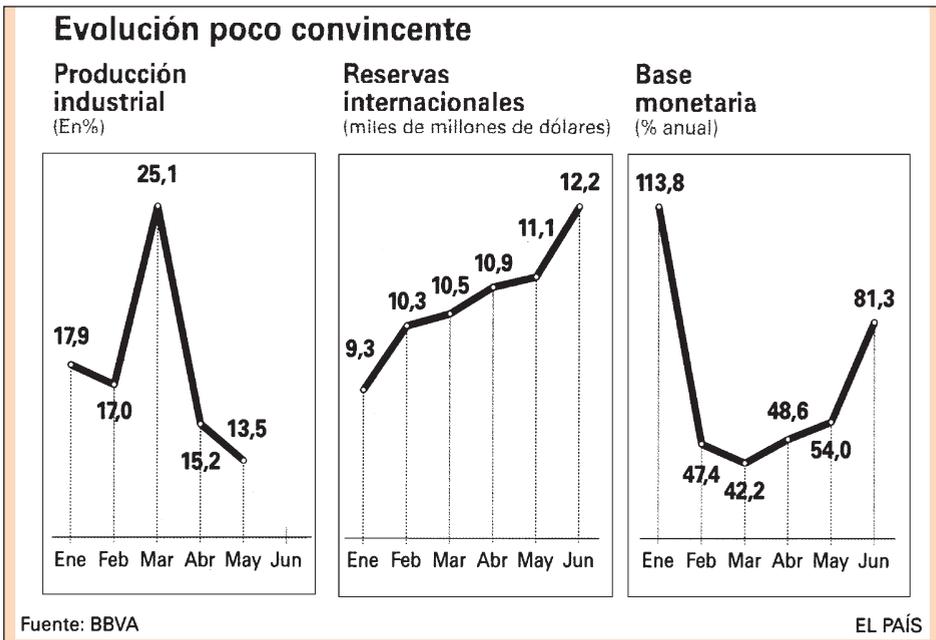


Figura 24. Párrafo gráfico compuesto por tres gráficas sencillas que pretende ilustrar la evolución económica reciente en Argentina.

Completando la analogía tendremos que considerar que cada una de las gráficas que conforman un párrafo gráfico es una “oración gráfica”. Así, la figura 24 muestra un párrafo compuesto por tres oraciones, al igual que el de la figura 25, mientras que el tercero está compuesto por cinco oraciones.

Las oraciones que componen la figura 24 son formalmente iguales (muestran la evolución de una variable a lo largo del mismo semestre con el mismo

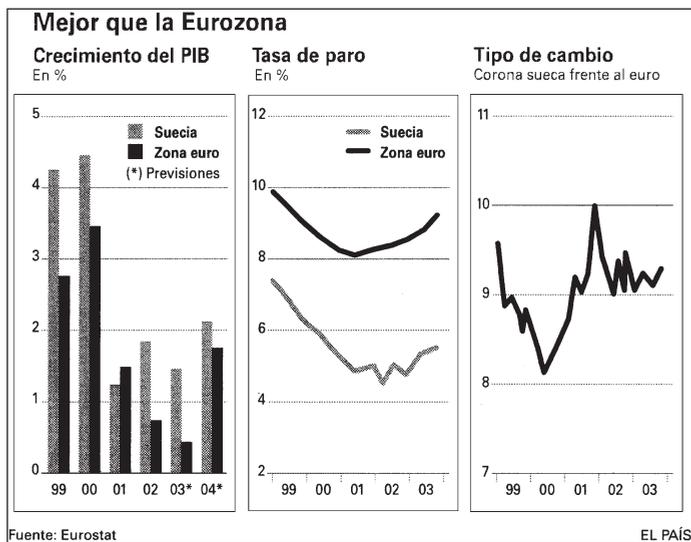


Figura 25. Párrafo gráfico, que compara algunos parámetros económicos entre Suecia y la eurozona, compuesto por tres gráficas.

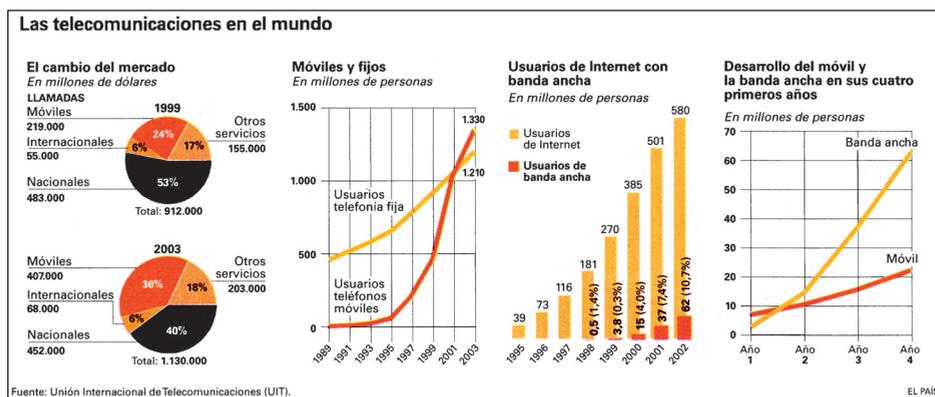


Figura 26. Párrafo gráfico relativamente complejo que muestra la situación y evolución de las telecomunicaciones a nivel mundial.

tipo de representación) y en cada una de ellas hay un solo trazo, una única serie de datos.

En el párrafo mostrado en la figura 25 son también tres las magnitudes que le interesa mostrar al autor, pero con una intención centrada más en com-

parar la situación de dos sujetos que en evidenciar la evolución de uno de ellos (como ocurría en la figura anterior). Ello lleva a que las oraciones incluyan dos trazos, dos series de datos¹.

La diferencia entre las oraciones que componen los párrafos de las figuras 24 y 25 la vamos a recoger denominando a las primeras “oraciones simples” y a las segundas “oraciones compuestas”. Consideraremos como “simples” las oraciones (gráficas) que muestren una única serie de datos², lo que da lugar a un único trazo. Del mismo modo, las gráficas en que se represente más de una serie de datos serán las que consideraremos como oraciones compuestas.

La representación de unos datos como una oración compuesta es una opción que siempre tiene como alternativa la representación de cada una de las series de datos como una oración simple. Es decir, que las oraciones compuestas se pueden descomponer en oraciones simples manteniendo básicamente el mismo mensaje³.

Las gráficas que estamos denominando oraciones compuestas requieren un elemento nuevo, que habitualmente se denomina “leyenda”, que permita identificar cuál es el sujeto de cada una de las trazas de la gráfica. Un elemento de referencia que muestre la categoría que diferencia cada una de las series de datos representadas en la gráfica.

Con las ideas que hemos introducido aquí vemos que la figura 26 presenta un párrafo con 5 oraciones. De ellas, las dos gráficas de tarta del primer recuadro serían oraciones simples mientras que las tres restantes serían compuestas, en los tres casos de dos series de datos.

En los tres casos analizados hasta ahora, las oraciones componentes de los párrafos gráficos son totalmente independientes. El autor podría haber prescindido de alguna de ellas o añadido otras sin alterar la estructura. La magnitud representada en cada una de las gráficas no guarda relación formal⁴ con ninguna de sus compañeras de párrafo. Esto, sin embargo, no siempre es

1. Bueno, sólo las dos primeras lo hacen, ya que la tercera presenta una magnitud que ya es en sí misma una comparación entre los dos sujetos (el tipo de cambio).

2. Una única serie de datos acorde con el tipo de gráfica que se presenta: con concordancia en la dimensión. Este tema se expondrá con detalle en el segundo epígrafe del capítulo 4.

3. Si bien la percepción de la comparación, que es lo que se busca en estas representaciones, se dificulta cuando se descompone una oración gráfica compuesta en dos (o más) simples.

4. Obviamente si guardan relación semántica, su significado tiene que estar relacionado como para que el autor las considere partes de un todo mayor como es el párrafo, pero no hay relación desde un punto de vista formal.

así. En las figuras 27 y 28 se presentan ejemplos de párrafos en los que algunas de las gráficas guardan una relación formal clara con otras.

En la figura 27 se ve en primer lugar el “resumen del programa inversor” de la empresa que se analiza. Este programa se divide en cuatro grandes bloques, lo que se representa mediante un gráfico de tarta que evidencia las proporciones relativas de cada una de esas partes. Pero interesa mostrar la subdivisión de dos de ellos en subprogramas menores, y para ello se representan gráficos de barras que se relacionan con la tarta, enmarcando estas gráficas en un recuadro con una flecha que señala la porción de tarta que desglosan, como un globo de texto en una viñeta de cómic. Estas gráficas de barras relacionadas directamente con porciones de la gráfica de tarta las podemos considerar como “oraciones subordinadas” de la oración principal (que sería la tarta).

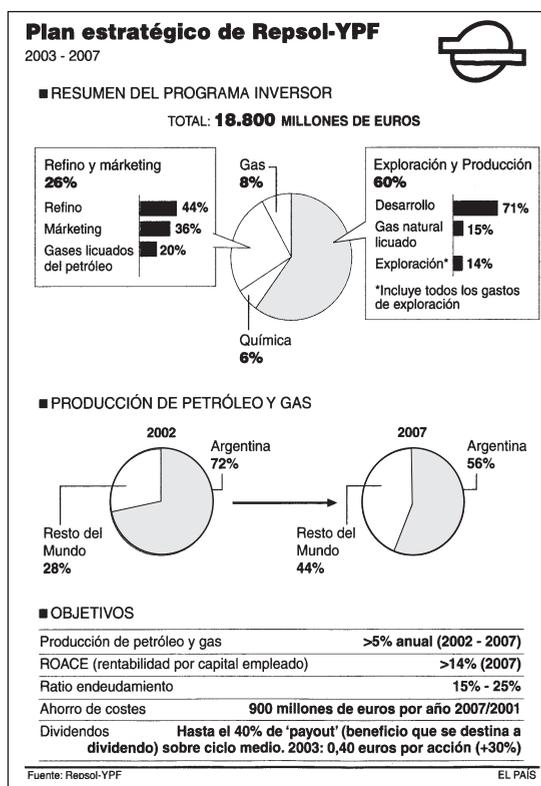


Figura 27. Párrafo gráfico formado por cinco gráficas que no se pueden considerar como totalmente independientes. Podemos considerar unas “subordinadas” de otras.

La composición de un párrafo gráfico a partir de las oraciones que lo componen utiliza esquemas de composición distintos del de eje métrico que es el que caracteriza lo que hemos denominado como “gráfica”⁵. Así, esa relación de subordinación se está indicando con un dibujo que mezcla los esquemas de “contención” (dado que las gráficas de barras están contenidas en un recuadro), “agrupación por proximidad” (al colocar los recuadros cerca de las porciones de la tarta con las que se relacionan) y “enlace” (por la flecha del globo que asocia los elementos).

Del mismo modo, en la segunda parte de la figura 4, donde se presenta la producción de petróleo y gas, se muestran dos oraciones simples (dos tartas con la proporción entre Argentina y el resto del mundo) asociadas mediante una flecha. Cada una de las gráficas se refiere a un año distinto, y la flecha representa la evolución temporal entre los dos años correspondientes a cada gráfica (una mezcla de esquemas de composición de “secuenciación lineal” y “enlace”).

En la figura 28 se muestra un párrafo gráfico que se compone de tres bloques de gráficas y una fotografía. Cada bloque tiene una gráfica de tarta como oración principal y una gráfica de barras subordinada de ésta (dos en el primer bloque). Aquí la subordinación se indica por contención y proximidad⁶. La inclusión de la fotografía clarifica de un primer vistazo el sujeto de toda la información (el colectivo de médicos) y además completa la composición del párrafo dando lugar a un rectángulo completamente lleno. Es curioso notar que para que la composición resultara más adecuada se ha tenido que invertir la fotografía original. Esto se aprecia por el rótulo “EXIT” de la esquina superior izquierda, que está invertido.

Como resumen de esta sección, se ha presentado el concepto de párrafo gráfico y el de sus componentes, las oraciones gráficas. De estas, a su vez, se ha visto que pueden ser, por una parte simples o compuestas, y por otra independientes o subordinadas. La forma de relación de los elementos que componen un párrafo gráfico se basa en esquemas de composición distintos del de eje métrico, que es la base de las gráficas.

5. En general, aunque alguna vez pudiera ocurrir que apareciesen gráficas dentro de gráficas.

6. En este caso no hay flechas para indicar la relación.

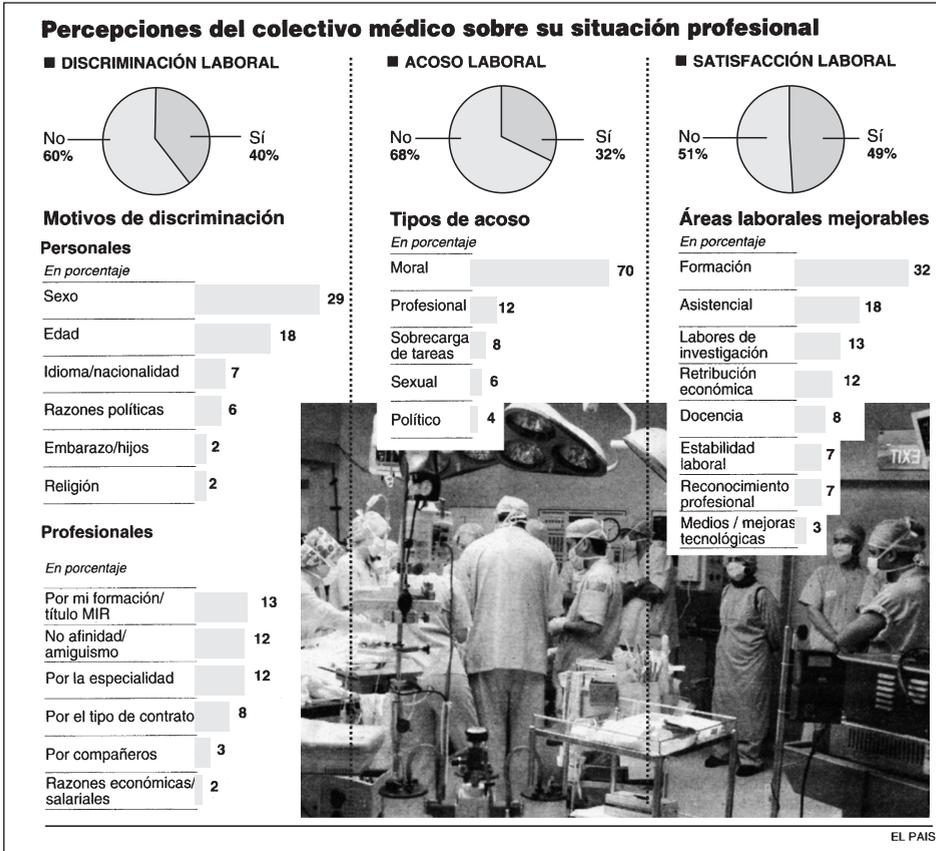


Figura 28. Párrafo gráfico que muestra siete gráficas agrupadas en forma de tres series en las que se puede apreciar, en cada caso, una oración principal (las tartas) y otras subordinadas (los gráficos de barras).

Sujeto y predicados

Hemos visto ya que las gráficas son representaciones de datos numéricos en las que se utiliza el esquema de composición de eje métrico: la metáfora visual del proceso de medida. Sabemos también que medir es comparar una cantidad con su respectiva unidad. Tanto la cantidad a medir como la unidad son porciones particulares de una misma magnitud. De este modo la gráfica, al representar unas cuantificaciones concretas de la magnitud evidencia alguna característica de la misma, o dicho de otro modo una gráfica expresa un

predicado sobre la magnitud que en ella se representa. Podemos pues extender la analogía entre el lenguaje léxico y el gráfico considerando que el sujeto de las oraciones gráficas es la magnitud representada. Del mismo modo, el predicado de la oración gráfica es la percepción visual llamativa que se obtiene a partir de la traza de la representación.



Figura 29 Evolución temporal de la superficie construida de centros comerciales.

En la figura 29 se presenta la superficie ocupada por la construcción de nuevos centros comerciales, año a año desde 1988 hasta 2002. El sujeto de la oración es “la superficie de nuevos centros comerciales”, y de él lo primero que percibimos, el predicado, es que “crece con el paso del tiempo”. Así, en una primera aproximación, la gráfica se puede leer como: “la superficie de nuevos centros comerciales crece con el paso del tiempo”. Es claro que esa era la intención del periodista que incluyó esa gráfica, dado que la coloca bajo el titular de “Un desarrollo imparable”⁷.

La gráfica dice muchas más cosas: “se ha multiplicado por más de 6 en 15 años”, “entre 1993 y 2000 se mantuvo prácticamente constante”, “tuvo un

7. Se entiende que de los centros comerciales, y no tanto por la gráfica como por el titular de la noticia y el resto del “entorno”.

año especialmente pobre en el 1999”, “se aproxima ya al millón de metros cuadrados”, etc. Pero todas esas cosas se predicán del mismo sujeto: “la superficie de nuevos centros comerciales”, que es el sujeto de la oración gráfica.

Las gráficas, como todos los sistemas de presentación visual de información cuantitativa, manifiestan muchos niveles de lectura. La misma representación transporta desde el mensaje más inmediato hasta otros mucho más sutiles que se van apreciando con una observación pausada. En todo caso, es interesante darse cuenta de que todos esos mensajes son predicados de un mismo sujeto. Una gráfica sólo dice cosas de la magnitud en ella representada.

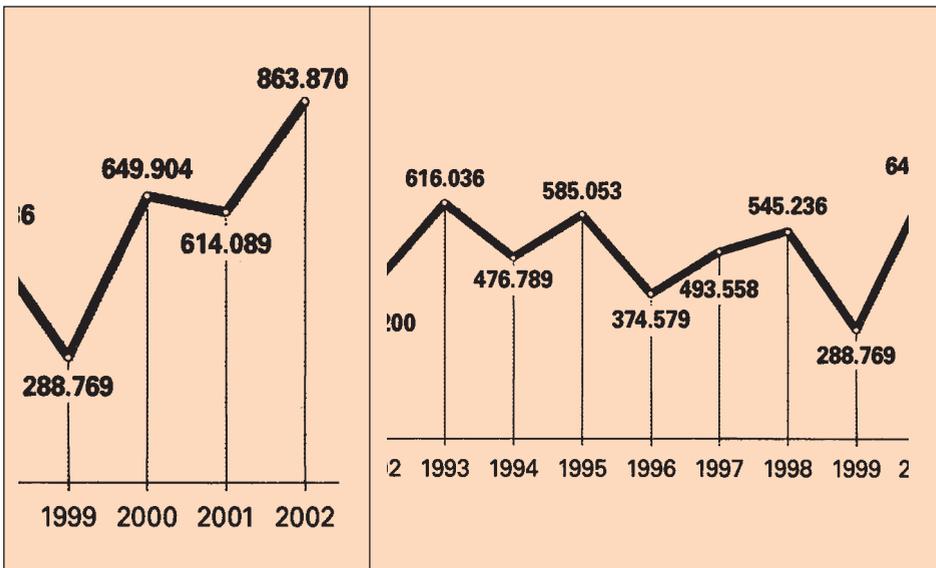


Figura 30. Dos pedazos de la figura 6, recortados y ligeramente deformados (estirándolos en una de las direcciones).

Otra cuestión diferente es cuánto de soportados están los predicados que se expresan por los datos reales. En vez de la representación de la figura 29 podríamos haber visto las de la 30 (no sólo son los mismos datos sino incluso la misma representación, sólo que recortada y deformada). En el caso de la figura 30.a se aprecia un crecimiento rapidísimo (se triplica en tres años), mientras que en el caso de la figura 30.b lo que se ve es un lento declinar que acaba disminuyendo a la mitad en 6 años. Obviamente los mensajes son ciertos, pero los dos son parciales. Del grado de “veracidad” de los predicados ha-

brá que ocuparse más adelante, bajo el epígrafe de *Semántica*, con el que se denota la relación entre la representación y el significado. Pero en todo caso, los dos predicados hacen referencia al mismo sujeto: la magnitud representada.

Como se muestra en el ejemplo anterior, los predicados que se desprenden de una representación son múltiples y de distintos niveles de profundidad de análisis. Esta diversidad en el mensaje, y la condensación de una gran cantidad de información en un espacio reducido, son características que hacen especialmente interesantes las representaciones gráficas. Pero a pesar de esta multiplicidad siempre hay una primera impresión, un mensaje principal, que queda después matizado por el resto del análisis. Este mensaje principal, que se capta “al primer golpe de vista”, está muy relacionado con las características perceptivas del sentido de la visión.

Algunos de estos “predicados a primera vista” serían:

- Tendencia evolutiva: Crecimiento, decrecimiento, mantenimiento, etc.
- Comparación de valor (cualitativa o cuantitativa).
- Reparto.
- Correlación.
- Sucesos singulares.

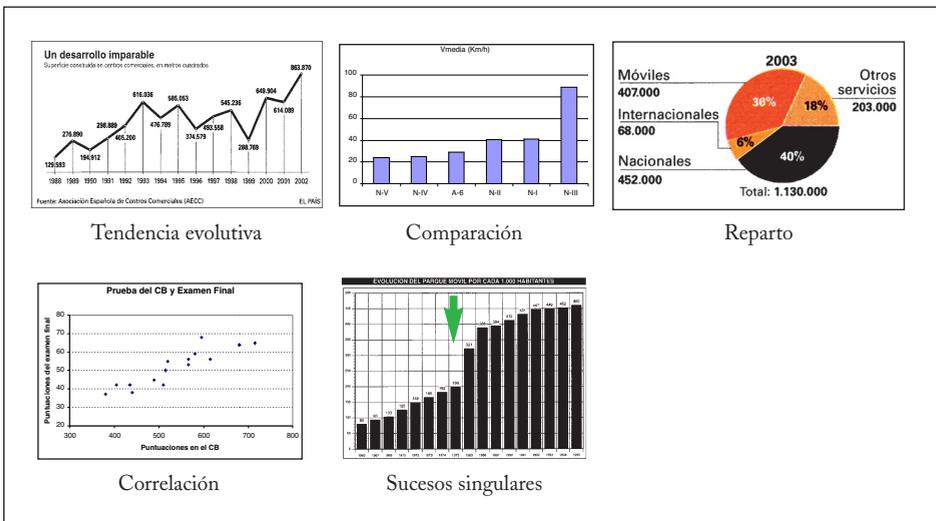


Figura 31. Cinco tipos de predicados básicos de las oraciones gráficas junto con un ejemplo de cada uno.

Por una parte hay unas características gráficas que se perciben mejor que otras⁸, y por otra algunas disposiciones geométricas resultan más significativas que otras para la mente. De la conjunción de ambas cuestiones se deriva la existencia de unos tipos de gráfica especialmente adecuados para cada uno de estos predicados básicos.

Los estudios sobre percepción de características gráficas⁹ muestran (ver figura 32) que hay tres que el ojo percibe mucho mejor que cualquier otra: posición, longitud e inclinación (ángulo). Así las comparaciones se manifiestan como diferencias de longitud de barras, los repartos como ángulos de porciones de tarta y los sucesos singulares como alteraciones de inclinación.

Por otra parte, la cuestión cognitiva se analizó ya en el capítulo sobre la comunicación gráfica, siguiendo a Engelhardt y su clasificación de los “esquemas de notación”, como la parte de la representación mental que describe las relaciones entre elementos gráficos¹⁰. Se puede ver cómo también los esquemas de notación básicos descritos por Engelhardt se pueden reconocer en los tipos básicos de predicado.

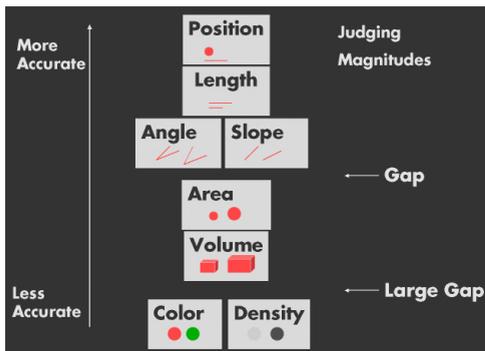


Figura 32. Características gráficas perceptivas (canales perceptivos), ordenados en función de la precisión que ofrecen¹¹.

8. Ver comunicaciones de Munzner, por ejemplo.

9. Codificación visual. Por ejemplo: Tamara MUNZNER, <http://www.cs.ubc.ca/~tmm/talks/sig04/sig04final.1/> O de las citas que ella incluye:

<http://www.csc.ncsu.edu/faculty/healey/PP/index.html> entra en muchos detalles sobre percepción para la visualización. También en el capítulo sobre “estética” de Wilkinson. También se puede ver en CLEVELAND, W. S. (1994). *The Elements of Graphing Data* (Rev. Ed.). Summit, NJ: Hobart Press.

10. Estos a su vez tenían unas reglas de construcción denominadas “esquemas de componentes”.

11. <http://www.graphics.stanford.edu/courses/cs448b-02-spring/lectures/encoding/> La figura está tomada de esa dirección, y está inserta en una presentación muy interesante sobre el tema de la codificación gráfica de datos cuantitativos.

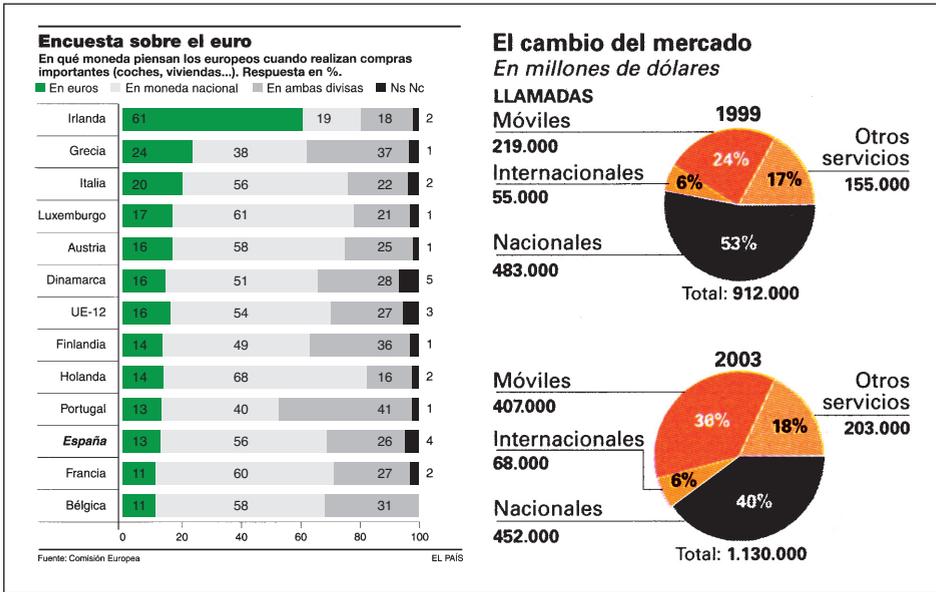


Figura 33. En el primer caso se muestra un predicado de “reparto más comparación” en una sola oración, y en el segundo “reparto más evolución temporal” mediante dos oraciones gráficas agrupadas.

Por último, conviene resaltar que el “predicado a primera vista” que transmite una gráfica no tiene por qué ser uno de los tipos básicos: puede ser una mezcla de ellos, como en los casos presentados en la figura 33.

En todo caso lo que comenzaba simplemente como una analogía entre las infografías y los párrafos de texto ha ido derivando hacia la psicología de la percepción y de la cognición, y no es mala cosa atisbar estas disciplinas, pero debemos seguir con ideas sencillas para mantener el objetivo del libro.

Apuntes de gramática de los textos gráficos

Tras los capítulos anteriores ya sabemos discernir entre lo que consideramos una gráfica y otros elementos de comunicación visual. También hemos introducido las ideas de párrafo gráfico y oración; y en éstas distinguimos entre simples y compuestas por una parte e independientes y subordinadas por otra. Hemos considerado, por último, cómo el sujeto de las oraciones gráficas

es la magnitud representada y la multiplicidad de predicados que se transmiten, así como la clasificación de los mismos.

El progreso natural del estudio del lenguaje incluiría las tres etapas clásicas: ortografía, sintaxis y semántica. Es decir la introducción de las palabras como elementos constituyentes de las oraciones construidas de unas formas particulares y no de otras (el vocabulario y la ortografía), las reglas fundamentales de interrelación entre palabras para construir oraciones correctas (la sintaxis) y la relación entre la expresión léxica y el significado que se pretende transmitir (la semántica). La trasposición de esos conceptos al lenguaje de las gráficas no es tan inmediata, y puede merecer la pena considerar las tres cuestiones de una forma interrelacionada, más o menos juntas.

Una gráfica tiene un conjunto de elementos constituyentes, que siguiendo la analogía serían las palabras. Componentes de una gráfica serían por ejemplo: título principal, título secundario o subtítulo, descripción del gráfico, región de datos y símbolos, eje horizontal y escala, eje vertical y escala, apuntadores, descriptores de señales y marcas (ver figura 34). Y para todos esos elementos hay unas formas de representación más comunes que otras. Sin embargo no es posible establecer reglas precisas sobre la corrección o incorrección de cada uno de estos elementos de forma aislada.

En el caso léxico no hay duda, es ortográficamente correcta la palabra si escribo “queso” e incorrecta si “keso”, sea cual sea la oración en que las coloque, mientras que en el caso de las gráficas el mismo elemento lo podremos considerar correcto o incorrecto en muchos casos dependiendo del resto de los elementos¹². No sería pues tanto un error ortográfico como sintáctico.

Del mismo modo, hay representaciones que, según una consideración puramente formal (en cuanto a ortografía y sintáctica) son correctas, pero que están construidas de una forma en que se exageran o incluso tergiversan los mensajes que realmente soportan los datos. Estas situaciones habríamos de considerarlas errores semánticos, que en muchas ocasiones se basan en artificios formales que, por sí mismos, no podemos considerar erróneos.

La palabra “queso” se escribe sólo de esa forma, y aunque podríamos entenderla si la escribimos “keso” “quesso” o “quéso”, es claro que estamos cometiendo errores en la ortografía de la palabra. La ausencia de una formalización

12. Un ejemplo podría ser la inclusión de una leyenda, lo que es correcto cuando hay varias series de datos y absurdo en caso contrario. Por cierto, MS Excell siempre la pone por defecto, y si no se tiene cuidado de borrarla cuando resulta innecesaria queda fatal.

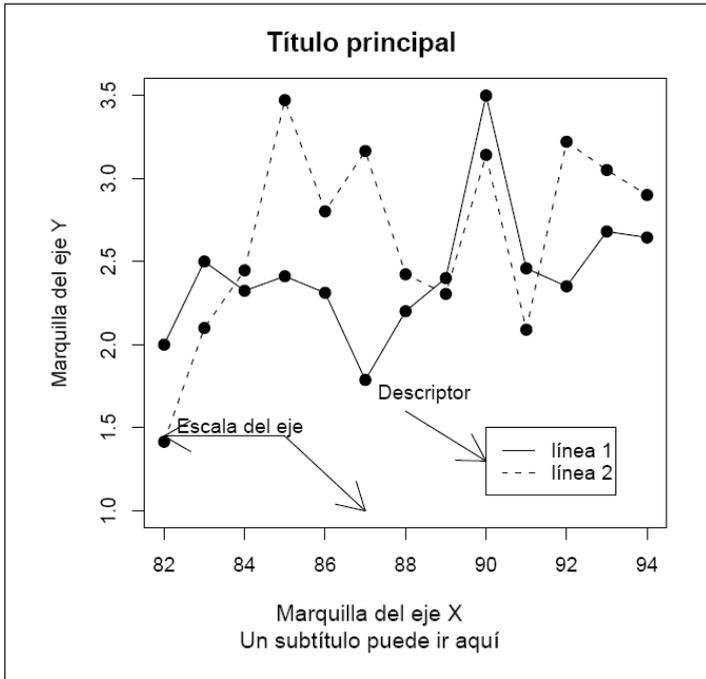


Figura 34. Componentes principales de una gráfica de dispersión (tomado de <http://cran.r-project.org/doc/contrib/grafi3.pdf>). Una figura análoga para gráficas de barras se puede encontrar en: <http://iilt.ilstu.edu/gmclass/pos138/datadisplay/goodcharts.htm>

precisa en el lenguaje gráfico deja a la inteligibilidad, a la capacidad comunicativa, como único juez sobre la corrección ortográfica de cada uno de los elementos que la componen.

Estas relaciones entre la estructura y el significado, por un lado, y entre unos elementos formales y otros, por otro, hacen que en la mayoría de los ejemplos reales no se puedan tipificar los errores de una forma inequívoca.

No parece pues que en el lenguaje gráfico sea tan conveniente la separación entre ortografía, sintaxis y semántica. Aun así, y por llevar la analogía entre lenguajes a sus últimas consecuencias¹³, vamos a conservar las tres palabras del análisis léxico como epígrafes para el gráfico, aunque sólo sea como ayuda nemotécnica.

13. Aunque como veremos eso de lugar a ciertas incoherencias.

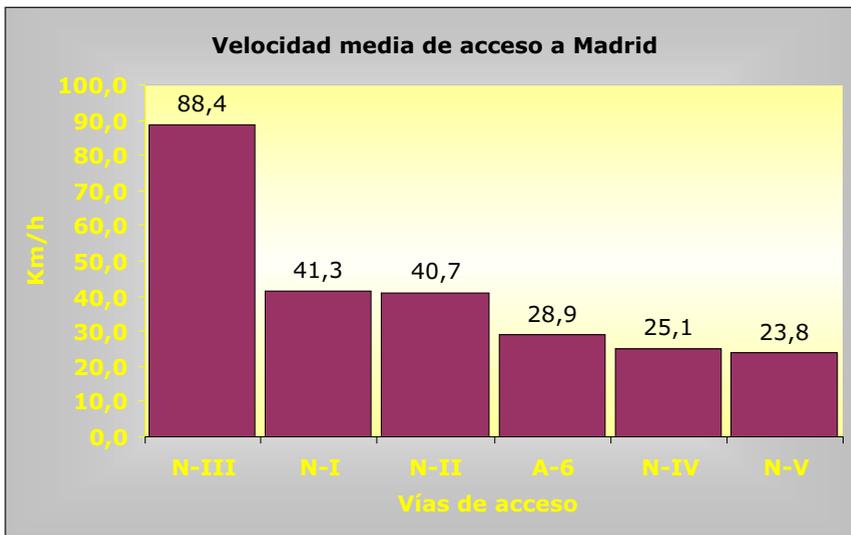
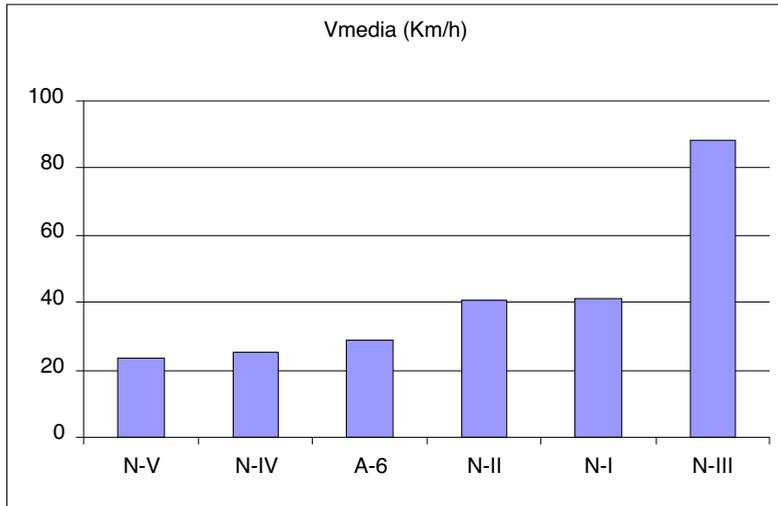


Figura 35. Quésos y Kesos. Los mismos datos¹⁴ en el mismo tipo de gráfica, sólo que el aspecto formal (tamaños, colores, tipos, etc.) de casi todos los elementos es distinto en una y otra. ¿Se aprecian errores “ortográficos”? ¿Cuál es más efectiva como elemento de comunicación?

14. La fuente de estos datos y una discusión sobre su representación original se encuentra en el capítulo 4 (figuras de la 40 a la 43).

Así, nos referiremos con “ortografía” a los elementos mínimos necesarios para que una gráfica sea inteligible, para que exista un mensaje transmitido. En concreto:

- Definición de las magnitudes representadas.
- Adecuación del tipo de gráfico a la estructura de los datos disponibles.
- Existencia de ejes y escalas (u otras referencias) que permitan la relación de los elementos gráficos con los datos originales.
- Leyendas u otros elementos identificativos de las series (en caso de haber más de una).

Bajo el epígrafe “sintaxis” veremos las relaciones entre los elementos que hacen que la gráfica sea un buen elemento de comunicación, que se entienda bien, sin dificultades añadidas:

- Tamaños de los ejes y elección de las escalas.
- Tamaño total de la gráfica.
- Proporción entre el área gráfica y el total ocupado por la representación.
- Trazos, puntos y colores legibles, identificables...
- Etc.

Por último, una vez que tenemos algún significado (ortográficamente correcto) que se percibe adecuadamente (sintácticamente correcto) falta asegurarse de que ese significado es coherente con los datos disponibles: la “semántica”.

Estas tres cuestiones las trataremos a continuación en capítulos separados.

Resumen del capítulo

En este capítulo se han sentado las bases fundamentales de la analogía entre los lenguajes gráfico y léxico. En primer lugar se asocia “párrafo gráfico” con “infografía”, y “gráfica”, tal y como se definió en el capítulo anterior, con “oración”.

Las oraciones gráficas se ha visto que pueden ser simples o compuestas por un lado, e independientes o subordinadas por otro. Se ha introducido el “sujeto” como la magnitud representada en una oración simple (si se representan varias magnitudes la oración será compuesta).

A la impresión visual que sobre el sujeto transmite la gráfica le hemos denominado “predicado”, y hemos comprobado que una gráfica puede transmitir múltiples predicados a medida que se avanza en la profundidad de su análisis. Los predicados fundamentales están relacionados con características de la percepción visual.

Finalmente se ha presentado el programa que se abordará en los siguientes capítulos para profundizar en la analogía que se ha presentado aquí.

Capítulo 4 Ortografía

Cuando decimos que dos cosas son análogas no pretendemos que sean idénticas, las analogías tienen similitudes pero también diferencias. Las similitudes resultan muy fructíferas para la imaginación, aunque las diferencias también pueden ser interesantes. A lo largo de este texto exploramos la analogía entre el formalismo del lenguaje natural y el del lenguaje gráfico (en particular de las gráficas)¹. Y las definiciones que vamos haciendo al hilo de esta analogía nos pueden llevar, como de hecho ocurre ahora, a contradicciones internas. En todo caso, mientras entendamos la analogía como una forma de inspiración y/o un sistema mnemotécnico, estas contradicciones no deben suponer mayor problema. Desde luego, no pretendo mantener que la similitud entre las estructuras de los lenguajes léxico y gráfico tenga un trasfondo teórico más esencial.

El sujeto es un elemento de una oración, y consideraciones sobre la necesidad de explicitarlo o no, etc. pertenecen al estudio de las estructuras de las

1. Aunque sea de un formalismo un tanto naïf y obsoleto como el tema del sujeto y el predicado...

oraciones: la sintaxis. Sin embargo nosotros hemos definido aquí, al final del capítulo anterior, como cuestiones “ortográficas” aquellos elementos mínimos que hacen de un dibujo que sea efectivamente una gráfica. Como vamos a ver a continuación, la definición de la magnitud representada es vital para que una gráfica pueda ser comprendida. Será pues una regla básica (ortográfica) la necesidad de definir la magnitud representada. Pero en atención a otros aspectos de la analogía hemos llamado “sujeto” a dicha magnitud, por lo que entramos en contradicción con lo dicho al principio... El tema del sujeto es ortografía en un cierto sentido y sintaxis en otro. Avisados de la ambigüedad, no debe suponer mayor problema

El sujeto debe estar definido

Uno de los peores problemas que se pueden presentar a la hora de contemplar una gráfica es la indefinición del sujeto de la misma. Si no sabemos qué es lo que se representa, de quién se está hablando, todo lo que se diga resulta inútil, lo mismo da que se diga mejor o peor.

Un ejemplo lo encontramos en la figura 36. Vemos que en ella se representan los “minutos de publicidad emitidos sobrepasando el límite de 12 min./hora” para cinco semanas bien definidas y medidos en una escala de minutos correcta². El predicado parece también claro: hay una cadena sistemáticamente peor que las demás.

Pero si queremos analizar cuánto más exagerada es la publicidad de esa cadena y nos fijamos en los datos, nos encontramos con valores superiores a 60 minutos en todas las semanas. Dado que el límite legal está expresado en minutos por hora, uno esperaría *a priori* que así estuvieran representados los datos, pero valores superiores a 60 minutos dejan claro que no es así, obviamente no se pueden emitir más de 60 minutos por hora de publicidad.

Establecido que no son medias horarias podemos especular con que la representación sea de medias diarias o quizá semanales. A favor de la hipótesis de las medias semanales está el hecho de que las barras del gráfico corresponden a semanas (como se muestra en el eje horizontal). Pero un exceso de

2. Esa representación se repite para cuatro cadenas de televisión, haciendo de la gráfica una oración compuesta con cuatro trazas, cuatro barras. Esto no afecta al argumento que se presenta aquí.



Figura 36. Un ejemplo de gráfica en la que la magnitud representada está insuficientemente definida.

60 minutos en una semana respecto de los 12 minutos por hora supone que, en vez de 12 se han emitido 12,36 minutos en promedio cada hora de la semana³, una cantidad que un espectador normal, sin estar pendiente del cronómetro, no puede apreciar.

Pero todo el anterior argumento no deja de ser una elucubración ya que no podemos saber si es cierta la hipótesis de que lo que se representa son va-

3. Dado que una semana hay 168 horas, basta dividir 60 entre esta cantidad, suponiendo que se emite durante 24 horas al día. Si fueran 12 horas diarias de emisión el exceso sería de 0,71 min/hora.

lores integrados para toda una semana. Tampoco sabemos las horas semanales de emisión de cada cadena. En resumen, la magnitud representada no se ha definido suficientemente, y ello impide que la oración tenga verdadero sentido⁴.

Más exagerada aún resulta la indefinición del sujeto en la figura 37. En ella se presenta un párrafo gráfico con dos oraciones (la primera simple y la segunda compuesta), referidas a una magnitud denominada “Burnout”. En la primera se muestra su incidencia por profesiones y en la segunda su “prevalencia” en médicos de atención primaria españoles. Pero ¿qué es el “Burnout”? El titular del artículo en que se encuadra esta infografía ya nos indica que se trata de la situación laboral de “estar quemado”, en el texto del artículo se describen los síntomas de esta afección, pero en ningún lugar se encuentra información sobre su cuantificación: cómo se mide, en qué unidades, etc.

Al igual que en el ejemplo anterior podemos comenzar a hacer suposiciones con las pistas que ofrecen los gráficos. Parece que, en la primera gráfica, la medida es en porcentaje de un grupo de población, lo que lleva a pensar que el quemado laboral no es gradual sino es algo que se tiene o no de manera completa. Así, en la gráfica se representaría el porcentaje de un grupo que sí posee la afección. Nos vamos con esa hipótesis a la segunda gráfica y vemos que las magnitudes representadas son: “cansancio emocional”, “despersonalización” y “realización personal”, sin información sobre las unidades en que se miden (ni si la “realización personal” contribuye a aumentar o a disminuir el quemado). La falta de información es tan grande que no es posible establecer una hipótesis razonable sobre lo que puedan significar las gráficas de esta figura.

En los dos ejemplos que acabamos de ver se comprueba que la indefinición del sujeto hace que una gráfica carezca por completo de significado. Puede ocurrir que a esta situación se llegue por descuido del autor (igual que veíamos en la introducción), con la intención de llenar espacio o con la intención de ofrecer un mensaje tergiversado, no soportado por los datos. Pero también puede ocurrir que se deba simplemente a que el receptor de la gráfica carezca del contexto adecuado.

La indefinición del sujeto en una oración gráfica no es una cuestión que se pueda juzgar de manera absoluta, es necesario tener información sobre el

4. Otra cosa es la existencia de un significado aparente, a primera vista, que no está realmente soportado por la información cuantitativa disponible.

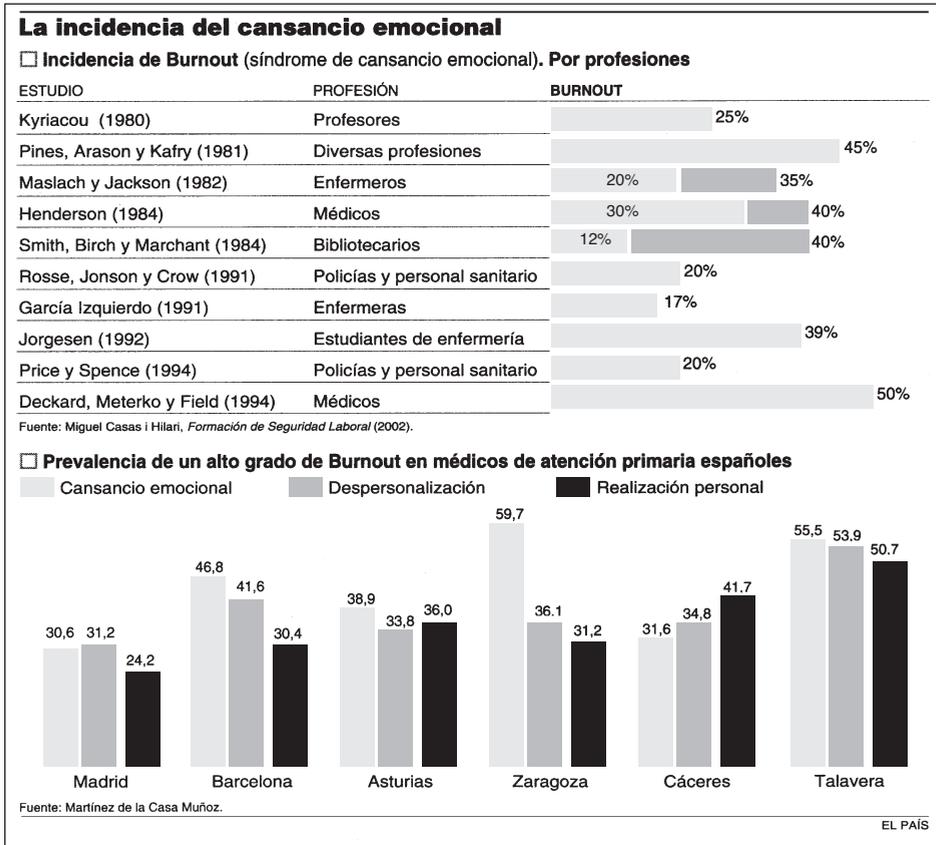


Figura 37. (El País 28 oct 2003- La incidencia del cansancio emocional)
 Párrafo gráfico con dos oraciones en las que los sujetos no están definidos en absoluto.

grado de conocimientos del receptor de la gráfica⁵ acerca del tema de que trata. Las gráficas de la figura 2, mostradas en un congreso de especialistas⁶ sobre el “Burnout”, probablemente no necesiten más definiciones para ser entendidas: todos allí sabrán qué es, cómo se mide, qué rasgos contribuyen a él y todo lo demás. También puede ser que en algún ambiente profesional las gráficas

5. Volvemos a la gráfica como codificación de un mensaje en un proceso de comunicación, esquematizado al comienzo del libro.

6. Según el texto del artículo las gráficas provienen, en efecto, de un congreso de especialistas en el tema.

como la de la figura 36 sean una forma habitual de controlar la publicidad emitida en las cadenas y que allí resulte más que evidente su significado.

A la vista de lo escrito en el artículo, y de las fuentes de los datos de la figura 37, parece que el periodista ha transcrito unas gráficas de las ponencias de un congreso a su artículo, y eso ha dado lugar a una descontextualización que las vacía de contenido. Los elementos gráficos, al igual que los léxicos, encuentran el significado que les dio su autor original en el contexto en el que los colocó.

Como resumen de esta sección, hemos visto cómo la indefinición del sujeto de una oración gráfica la hace incomprensible, y cómo esa definición depende por una parte del receptor del mensaje y del contexto de la gráfica por otra.

Concordancia entre el tipo de gráfica y el tipo de datos

En el capítulo segundo vimos que las gráficas representan datos, y que éstos pueden ser de distintos tipos. Aunque esta cuestión de la tipología de los datos puede complicarse bastante, para lo que aquí nos interesa nos podemos quedar con dos tipos: números y categorías⁷. La elección de un tipo de gráfico adecuado a una estructura de datos supone una idea muy simple: para las variables categóricas hay que utilizar ejes categóricos, y para las variables numéricas, ejes numéricos⁸. Obviedad que a veces no lo es tanto...

En la figura 38 vemos un ejemplo de una elección incorrecta del tipo de gráfico. En ella se representa la variación del precio de la vivienda respecto de la media europea para diferentes países. Se presentan tres series de la misma magnitud (correspondientes a los años 1999 y 2002 y a la media del periodo). Es por tanto una oración compuesta, aunque para lo que aquí interesa resaltar basta con considerar una sola de las series. La elección de una recta para unir los datos correspondientes a cada país es de lo más desafortunada, dado

7. La cosa puede ser un poco más compleja que datos numéricos o categóricos. Desde la estadística y la programación de bases de datos ha sido necesario afinar un poco más. Por ejemplo ver:

<http://www.graphics.stanford.edu/courses/cs448b-02-spring/lectures/encoding/> y http://rrpac.upr.clu.edu:9090/~amenend/conf96390.htm#_Toc5857201

8. Esta concordancia es lo mínimo para tener una gráfica correcta. Para que sea efectiva como elemento de comunicación habrá que preocuparse de más cosas.

que esa recta representa una interpolación entre los valores de los que se tiene información. Pero en el caso de países esto no tiene ningún sentido. Entre el precio de la vivienda en Alemania y el precio en Portugal no hay nada, mientras que la línea creciente sí que sugiere una tendencia. Lo adecuado en este caso habría sido una representación como gráfico de barras en la que cada barra correspondiese a un país, y entre una y otra no hay nada. Al utilizar líneas uniendo los puntos se convierte al eje horizontal en un eje métrico, cuando en realidad los datos representados son categorías.

Al profundizar en el análisis la gráfica presenta más sorpresas: ¿cuál es el dato correspondiente a cada país? He intentado recuperarlo a partir de la gráfica para intentar rehacerla en otro formato. Sería razonable esperar un par de cosas: (i) que el dato de cada país corresponda al centro del intervalo que se le asigna y (ii) que las inflexiones se produzcan en los datos (suponiendo que las líneas lo que hacen es unir datos reales). Para Alemania vale, pero para Rumanía... las inflexiones no sólo no se producen en el centro del intervalo, sino cada una de las trazas aparece en un sitio distinto. La traza correspondiente al 2002 entre el mínimo de Holanda y el dato de España parece curvada, con lo que no se encuentran inflexiones. Al final no me atrevo a hacer suposiciones y, por tanto, renuncio a redibujar esta gráfica. ¿Qué se está representando? Con lo fácil que hubiera sido hacer un gráfico de barras...

Volviendo a la idea que nos interesa resaltar aquí, se ha utilizado un eje numérico para representar datos categóricos. Así, unos datos que tienen sólo

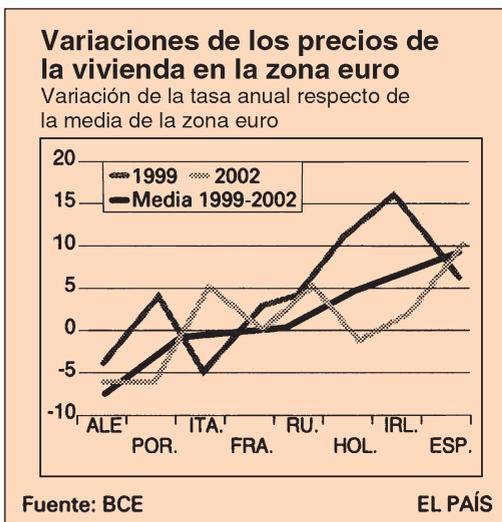


Figura 38. Representación de datos unidimensionales en un formato bidimensional. La interpolación entre dos países es algo totalmente carente de sentido.

una serie numérica⁹ se representan con dos ejes numéricos, mientras que la corrección “ortográfica” exige coherencia entre estas dos magnitudes. Podemos llamar “dimensión” a cada una de estas magnitudes: dimensión de una estructura de datos a la cantidad de datos numéricos que contiene cada dato para cada sujeto, y dimensión de una gráfica al número de ejes métricos que contiene. Con esas definiciones, la obviedad del principio se puede reescribir de una forma que suena mucho más científica: debe coincidir la dimensión de una gráfica con la de la estructura de datos que en ella se representa.

Más habitual es encontrar el error de la falta de concordancia en sentido contrario: representando datos bidimensionales en una gráfica unidimensional. Las dos primeras gráficas de este texto (en la introducción), la del parque móvil en España y la de la resistencia de un termistor, eran ejemplos de esta situación. Aquí añadimos, como figura 39, un ejemplo más.

En la figura 40 se presenta el porcentaje de viviendas en alquiler en España en función del paso del tiempo. O al menos lo que sugiere el tipo de representación es que el eje horizontal representa el paso del tiempo. El tiempo es una variable continua, numérica, sin embargo se representa como si de una categoría se tratase. Pasar de un punto al siguiente no supone la misma cantidad de tiempo en distintos puntos del eje¹⁰. Esto hace que las pendientes, saltos o inflexiones que se aprecian en la traza no correspondan a ninguna característica real de los datos sino que sean puros artificios¹¹.

El resultado es muy parecido al que veremos en el capítulo siguiente al hablar de la homogeneidad del sujeto: las características visuales más impactantes, las que se perciben a primera vista, no se corresponden con el mensaje que soportan los datos, mientras que éste queda oculto, y si se quiere reconstruir, hay que entresacarlo de los datos. En este sentido la presentación de los datos en una tabla resultaría más eficiente como medio de comunicación ya que, aunque es necesario que el lector haga un esfuerzo en encontrar el mensaje visual, al menos no ha de destruir antes uno falso.

Hay otra característica formal que diferencia unas series de datos de otras a la hora de relacionarlas con los tipos de gráficas en las que se pueden

9. Para cada una de las magnitudes representadas, cada uno de los sujetos.

10. Probablemente el redactor sólo disponía de esos datos, pero la excusa no exime del error: los podía haber buscado, o presentado de una forma más adecuada.

11. De hecho el brusco salto que aparece desde el 60 al 21% es una tendencia lineal suave cuando se representa correctamente, igual que ocurría con el parque de automóviles, al comienzo de la introducción.

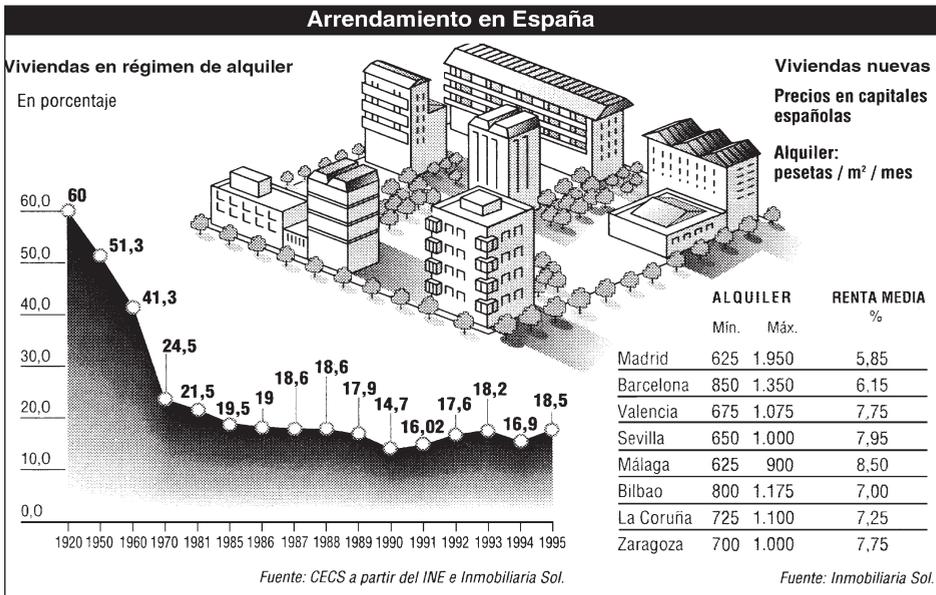


Figura 39. Al igual que en el caso anterior, el descenso brusco que se aprecia al comienzo de esta gráfica no es más que un artificio debido a la selección no homogénea de los datos a representar.

representar. Se trata de la existencia o no de una “relación de clausura” entre todos los datos de una serie numérica; nombre un tanto pretencioso para referirse al hecho de que la suma total de las cifras de una columna deba dar una cantidad determinada. El ejemplo típico lo encontramos cuando una columna expresa porcentajes: obviamente la suma de la columna debe ser 100, pero hay otros muchos casos, como el reparto de escaños de un parlamento, de acciones de una sociedad, etc.

Para expresar series de datos con relación de clausura se dispone de tipos de gráficos específicos. Para los datos numéricos en general, explota el proceso de medir para crear las metáforas gráficas que dan lugar a los diversos formatos. Cuando existe una condición de cierre es el proceso de reparto de una totalidad el que da pie a estos tipos de gráficos. El ejemplo paradigmático es el “gráfico de tarta”¹².

12. Aunque no es el único, el de “hemiciclo” es también típico, también se utiliza el reparto de una porción lineal, especialmente para comparar repartos (ver figura 33, en el capítulo 3).

En muchos casos cuando se busca una representación que acentúa el concepto de reparto se pierde información sobre el valor absoluto a repartir (y lógicamente más sobre el valor real de los pedazos). Por el contrario, resaltar el valor absoluto tiende a oscurecer la apreciación de las proporciones del reparto.

Unas referencias suficientes: Presencia, integridad y referencia de los ejes

Las gráficas son sistemas de representación visual de datos numéricos haciendo uso del esquema de composición de eje métrico, la metáfora visual del proceso de medida. Y dado que medir es comparar una cantidad con su respectiva unidad, las gráficas necesitarán una referencia de las unidades en que se efectúa la representación: necesitan una escala. El eje numérico¹³, aquel contra el que se representan los datos numéricos, debe explicitar su escala de alguna forma que resulte inteligible, y mejor todavía, que sirva además para comunicar eficazmente el mensaje de la gráfica.

Ausencia total de ejes

A lo largo de este capítulo veníamos analizando otros errores. Si nos fijamos ahora en los ejes vemos que las figuras 36, 37 y 39 poseen su correspondiente escala en el eje de ordenadas. No ocurre lo mismo con la figura 38, la del “burnout”, en la que en vez de una escala para comparar se da ya el valor de la comparación como dato numérico encima de cada una de las columnas del gráfico de barras. Pero mejor comencemos el análisis con casos exagerados.

En la figura 40 vemos una gráfica en la que el eje de ordenadas se ha omitido completamente. De esta forma, la idea cualitativa indicada en la gráfica con los rótulos de más o menos velocidad media es totalmente imposible de cuantificar. Quizá haya sido un poco tramposo sacar esta gráfica de su contexto. En realidad era parte del párrafo gráfico (ver figura 41). Vemos que en otra parte del párrafo los datos numéricos se muestran claramente, incluso que hay otra representación, como gráfico de barras, que cuantifica la velocidad media para cada vía.

13. El que veíamos antes que da lugar a la “dimensión” del gráfico. Una gráfica puede tener varios ejes métricos, pero nunca menos de uno.

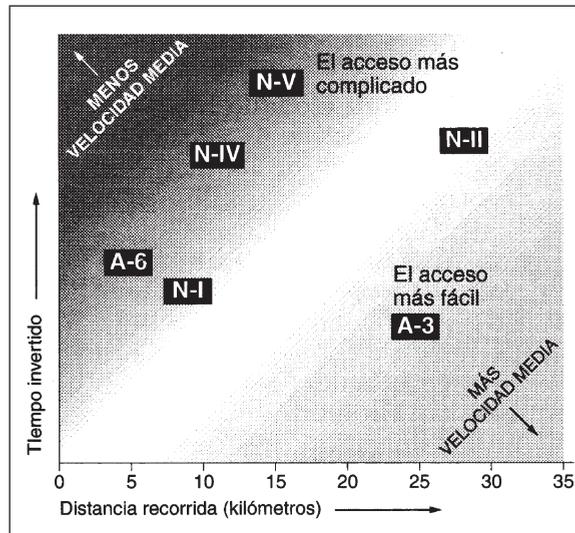


Figura 40. Para dar una idea de la rapidez de los accesos a Madrid, se presentan datos del tiempo que se ha tardado en recorrer una cierta distancia. Cada punto corresponde al trayecto por una vía, y la escala de tiempos no existe.

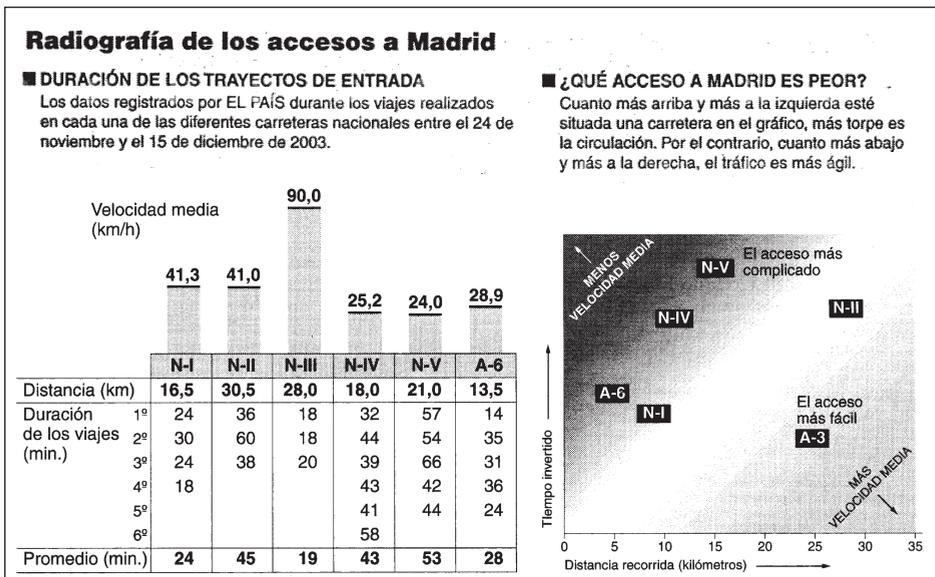


Figura 41. Párrafo gráfico del que se extrajo la gráfica de la figura 5.

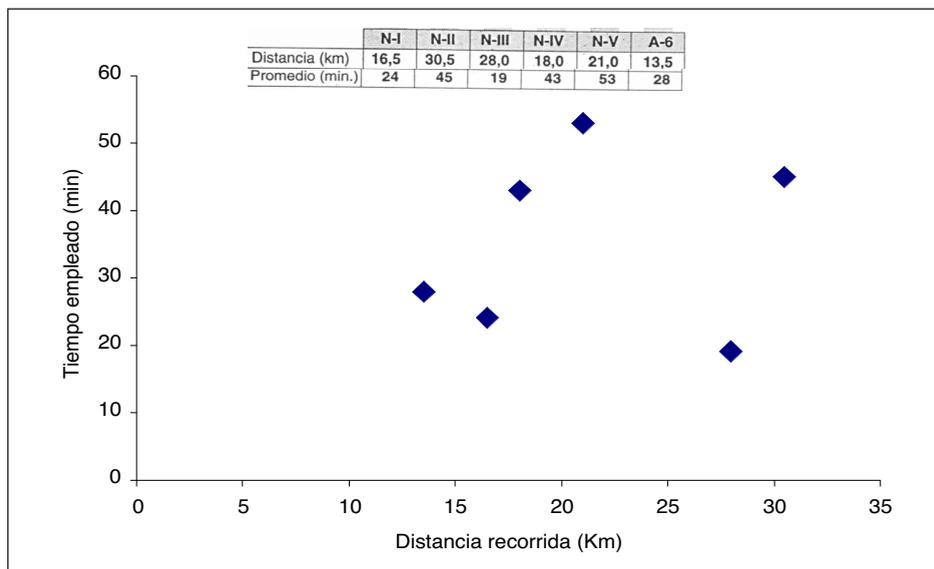


Figura 42. Tomando de la figura 6 los datos que supuestamente aparecían en la figura 5, se representan aquí respecto de unos ejes bien definidos.

Al disponer de los datos numéricos que supuestamente se representan en la gráfica que estamos analizando (figura 40), podemos comprobar claramente que la situación de los puntos sobre el eje x es incorrecta. Comparando con una representación fidedigna, como la mostrada en la figura 42, parece que el intervalo que en la gráfica original se rotulaba de cero a 35, en realidad debería empezar en 10 en vez de en cero.

En todo caso, lo que es claro es que la gráfica publicada no respeta la ortografía mínima de la representación: en el eje x los puntos no están situados en la posición indicada por el eje, y en el eje y no existe escala.

Continuando la crítica a la gráfica de la figura 40, falta considerar la extraña representación escogida por el autor. La información que parece querer resaltar es la velocidad media de las distintas vías, y para ello la gráfica de barras es una muy buena opción¹⁴. Parece que el diseño experimental utilizado

14. Tal y como aparece en la infografía, o con algunas mejoras quizá como la ordenación los datos. Las gráficas presentadas en la figura 35 del capítulo 3 son representaciones en forma de gráfico de barras de estos mismos datos.

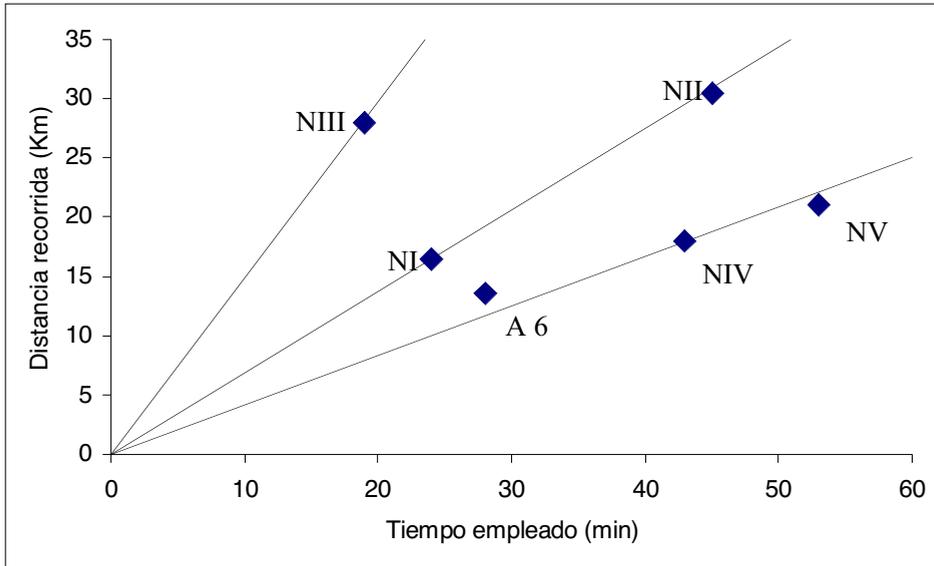


Figura 43. Una nueva representación de los datos de la figura 7. En este caso al intercambiar los ejes, las rectas que parten del origen unen puntos con la misma velocidad.

para obtener esos valores medios también le interesa al autor. Lo que se ha hecho es elegir un trayecto de unos determinados kilómetros (distintos para cada carretera) y cronometrar su recorrido repitiéndolo unas cuantas veces (también distinto número para cada carretera). Al menos eso es lo que parece deducirse de la tabla de datos incluida en la infografía. Desde un punto de vista científico podríamos discutir sobre la diferencia en el número de repeticiones para cada vía, la aleatoriedad en los horarios escogidos o la longitud de los trayectos, pero desde un punto de vista de una aproximación periódica no hay nada que objetar.

Si los datos se quieren presentar como gráfica de dispersión, es mucho más habitual, y fácil de interpretar, una versión que tiene los ejes cambiados respecto de la elegida en la figura original. Si en vez de representar tiempo frente a espacio se representa espacio frente a tiempo (ver figura 43), las líneas que parten del origen unen puntos con la misma velocidad; y cuanto mayor es la pendiente mayor es la velocidad. Así es fácil ver en la figura 43 cómo la NIII es la más rápida, seguida de las N-II y N-I, que tienen aproximadamente la misma, y como más lento está el grupo restante.

Ausencia de eje pero no de referencias (eje “hágaselo usted mismo”)

Con el ejemplo anterior hemos visto los errores ortográficos más exagerados de ausencia total de un eje o, peor aún, la situación en que los datos no se correspondan a su posición en un eje que sí se incluye. Ahora vamos a considerar el caso, menos grave quizá aunque más habitual, de ausencia de un eje explícito con inclusión, a cambio, de algunos valores numéricos que ayuden a reconstruirlo mentalmente¹⁵.

En la figura 44 vemos otro caso de una gráfica en la que no se da el eje vertical de forma explícita, pero a cambio se dan unas referencias que ayudan a situar los puntos. Lo que se da es el valor numérico correspondiente a cada dato. Esta colección de números es algo así como un “eje Ikea”: constrúyalo usted mismo. En efecto, podemos coger una escuadra y un cartabón, tirar líneas horizontales que pasen por los puntos y luego medir los espacios entre estos. Mediante sencillas reglas de tres a partir de estas medidas podemos obtener cualquier otro valor. En la figura 45 se presenta un breve resumen del proceso y se incluye el eje reconstruido para la gráfica de la figura anterior.

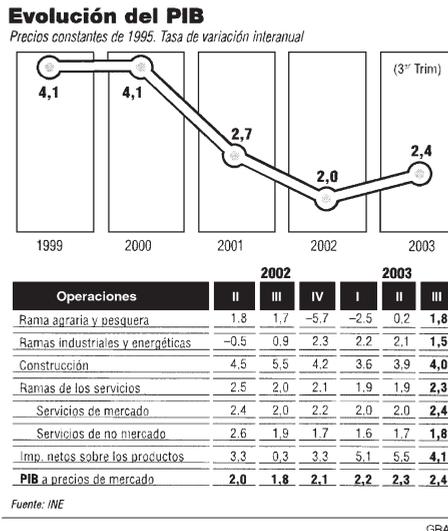


Figura 44. Una representación de la evolución del producto interior bruto en la que no hay eje y, pero hay datos numéricos que sirven de referencia.

15. Un ejemplo de esta situación lo veíamos ya en la figura 37, en este capítulo.

La reconstrucción de un eje a partir de los datos incluidos en la gráfica no es una tarea excesivamente difícil, pero tampoco es una trivialidad. El concepto de “hágaselo usted mismo” en el caso de los muebles es obviamente ventajoso: se transporta más fácilmente y el coste de montaje se transfiere al cliente. Pero en el caso de los ejes ¿alguien saca alguna ventaja de su elusión? Sólo en el caso de que el mensaje que se desea transmitir no sea exactamente el que soportan los datos. En el capítulo sobre semántica volveremos sobre esta cuestión.

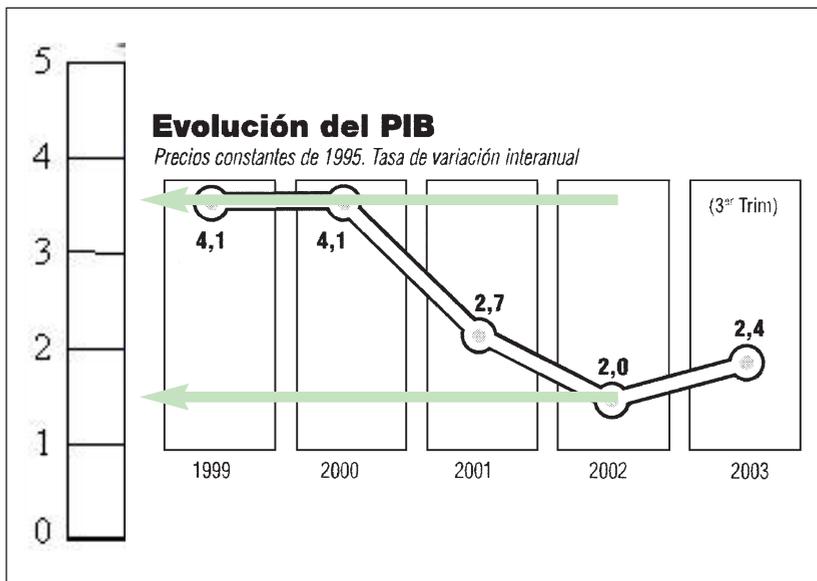


Figura 45. Reconstrucción del eje que no se explicitaba en la figura anterior.

Ausencia de la referencia de cero

En la figura 45 podemos apreciar claramente cuál es una de las consecuencias más evidentes de los ejes “hágaselo usted mismo”: El origen de la escala (el valor cero) puede estar en cualquier sitio. Para saber dónde está es estrictamente necesario realizar las operaciones que conducen a recuperar el eje.

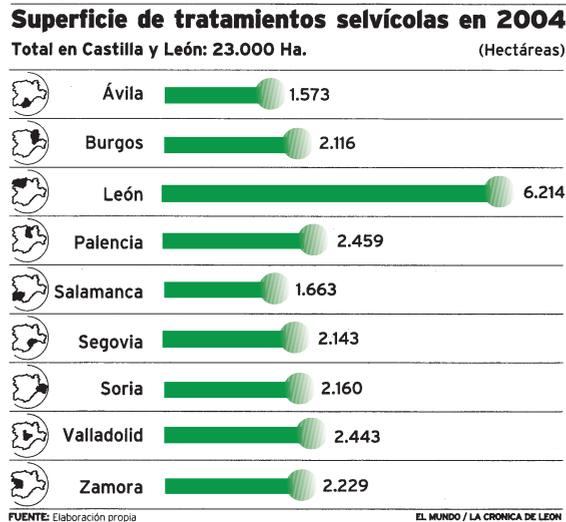


Figura 46. Una representación en forma de barras de comienzo difuso y corona indefinida, cuya longitud no se puede apreciar.

En la figura 45 (o mejor en la original, en la 44) al menos hay una cosa que no es difícil apreciar: el cero está fuera de la zona dibujada, ya que ésta queda casi acotada por los valores máximo y mínimo representados (2 y 4,1, que es casi 4). Pero hay situaciones en las que ni siquiera se puede saber a primera vista dónde queda el cero. ¿Dónde queda en la gráfica de la figura 46?

La cuestión de la posición del origen de la escala en la que se representan los datos no es en absoluto trivial. Los mismos datos representados con diferentes amplitudes de escala transmiten una impresión muy diferente. Al recortar una parte común a los datos representados se acrecientan las diferencias, y se aprecian mejor los detalles de la evolución, pero se pierde la referencia absoluta. En las gráficas de la figura 47, se presentan los mismos datos (que son los mismos de la figura 44) de las dos formas. En el caso en que se incluye el cero se ve directamente que la columna de 2002 es la mitad que la de 2000, como realmente ocurre. En cambio en la representación de la derecha la misma columna es 9 veces mayor que la de 2002. La duplicación del tamaño de una columna ha perdido sentido como predicado visual, ahora hay que fijarse en la escala y realizar la operación mental correspondiente para obtener una idea de las proporciones de los valores absolutos. A cambio, las diferencias entre los valores de los tres últimos años se pueden percibir mejor.

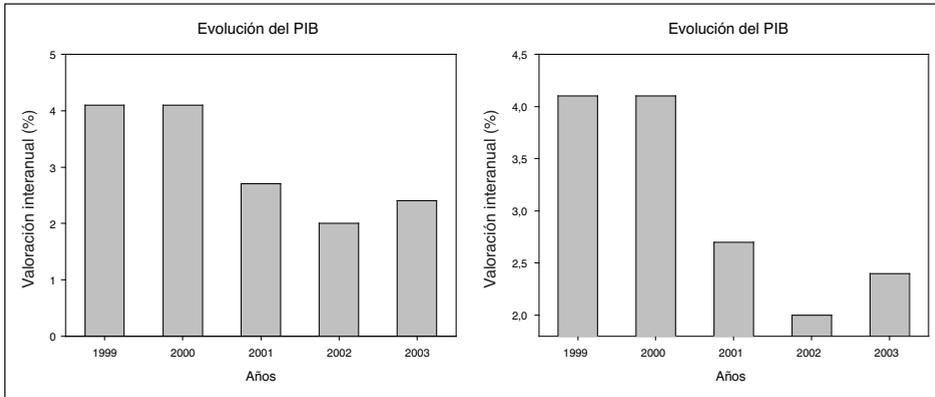


Figura 47. Representaciones de los mismos datos de la figura 44 incluyendo el cero en el primer caso y maximizando las diferencias en el segundo.

Volviendo a la figura 46, la respuesta a la pregunta antes plantada no es fácil. Las barras de ese gráfico tienen la extraordinaria propiedad de estar indefinidas tanto en su comienzo como en su final. El comienzo degradado hace que no se pueda saber el punto exacto en que arrancan. En el extremo opuesto, la columna está rematada por una bola (un arbusto, una mancha o lo que sea)¹⁶ que no se sabe en qué medida pertenece a la columna. Estas indefiniciones hacen que sea especialmente complejo el bricolaje de reconstrucción del eje, y con él del origen. Creo que no merece la pena entrar en detalles sobre dicho proceso, pero su resultado sí se muestra en la figura 48, superpuesto a la gráfica original. Resulta que una parte del degradado no pertenece a la longitud coherente de la barra, y el “adorno” de las columnas tampoco contribuye a la longitud coherente con el valor representado.

En la figura 49 vemos un ejemplo más de la diferente impresión que produce la inclusión o no del origen en el eje. En ambas series de datos el valor máximo es aproximadamente doble del mínimo, sin embargo no transmiten las dos esa idea. ¿Cuál es la correcta?

En cualquier caso, la ausencia de un eje explícito, por mucho que se pueda reconstruir de bricolaje, unida a la colocación arbitraria del origen del mismo lo debemos considerar un error ortográfico importante, dado que genera un mensaje incierto.

16. Este tipo de elementos que no aportan nada a la representación, salvo quizá una suelta “estética” les dedica Tufte mucha atención en su libro.



Figura 48. Barras de longitud proporcional a los valores indicados superpuestas a la gráfica original presentada en la figura 11.



Figura 49. Dos gráficas de series de datos muy parecidas que dan una impresión muy distinta por la inclusión o no del origen (¿Cuál es la que lo incluye?).

Ejes deformados

Hemos visto como errores la inexistencia total de ejes en las gráficas, o la inclusión de referencias que, aun permitiendo la reconstrucción de los ejes despistan la posición del origen. El caso que nos falta por considerar no es muy habitual, se trata de ejes en los que las escalas van variando a lo largo de su longitud: ejes de plastilina.

Como no es un error muy habitual, debo decir que el ejemplo de que dispongo (figura 51) es un tanto exótico, o mejor dicho especializado, ya que procede de una revista científica. Este origen hace especialmente difícil su comprensión detallada (qué se representa, qué hay en el eje de abscisas, las unidades del eje de ordenadas, etc.) pero afortunadamente nada de eso no es importante aquí. Lo llamativo es la escala del eje vertical en las dos gráficas (reproducidos separados del resto de sus gráficas en la figura 50). Según en qué parte del eje nos encontremos, la misma distancia corresponde a distinto número de unidades. Se ha comprobado¹⁷ que no es una errata tipográfica a la hora de imprimir el artículo: los datos efectivamente responden a la escala indicada.

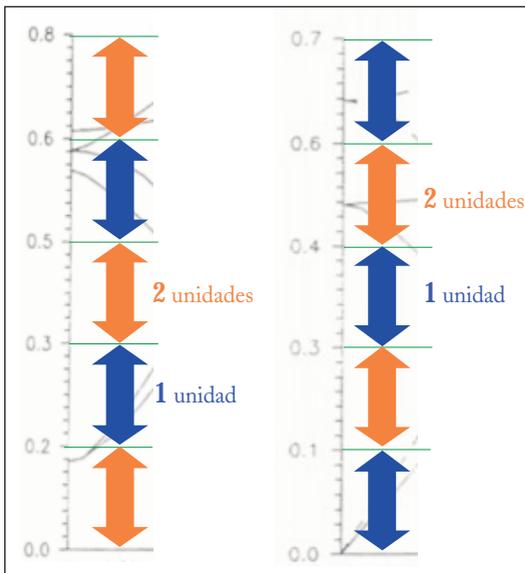
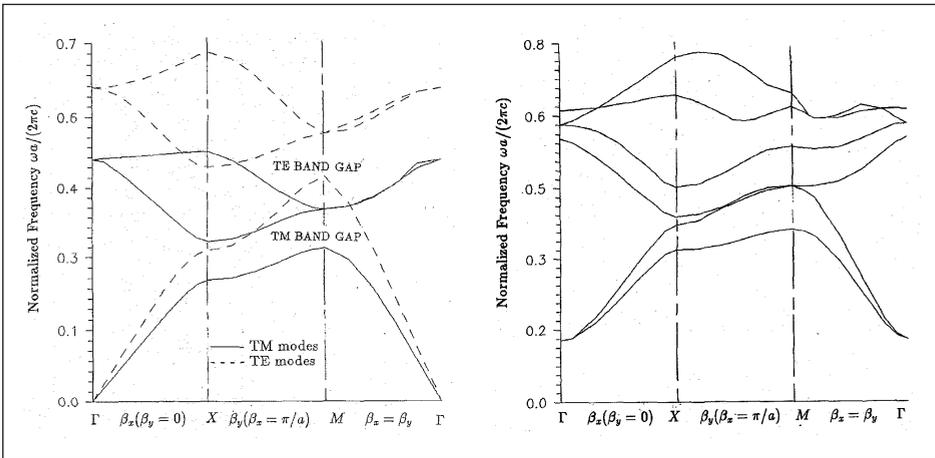


Figura 50. Los ejes de abscisas de las dos gráficas mostradas en la figura 51 se presentan aquí separados del resto de la gráfica para apreciar mejor el argumento. Espacios iguales (representados por las flechas) corresponden a diferentes unidades, una para las flechas azules y dos para las anaranjadas.

17. O al menos eso me dijo el estudiante de doctorado que me llamó la atención sobre esta gráfica y que es un auténtico experto en el tema de fondo. Gracias Mikel Gomez Laso.

La razón que llevó al autor de estas gráficas a tomarse la molestia de deformar los ejes de una manera tan extraña parece que tiene que ver con el deseo de acentuar alguna característica del aspecto de las trazas referidas a dichos ejes (en concreto los picos). El deseo es perfectamente legítimo, los picos están ahí, los datos presentan realmente mínimos y máximos, y esa es una característica importante para el tema que se está estudiando. Pero habrá que esforzarse en comunicar este hecho, en transmitirlo de una forma adecuada, mientras que la forma escogida no se puede considerar correcta.

Hay que darse cuenta de que el distinto valor de longitud atribuido a la misma magnitud en distintas partes del eje no se produce de ninguna forma ordenada, no se trata de un eje logarítmico, ni fraccionado ni nada parecido.



Figuras 51 a y b. Tomadas de una revista científica. Tanto las magnitudes representadas como el eje x o las unidades del y son muy específicas de dicho campo científico. El escalado del eje vertical, comprobado que no corresponde a una errata, es, cuando menos, llamativo.

En el peor de los casos el autor debería haber llamado la atención, tanto en el texto como en el pie de figura, sobre el artificio utilizado; y mejor aún buscar otros procedimientos expresivos para comunicar el mensaje de una manera correcta. Hay recursos de uso común que se pueden explotar a la búsqueda de una representación que haga justicia al mensaje que se pretende dar. Una posibilidad es introducir rupturas explícitas del eje (como se muestra en

la figura 16 para otros datos). Quizá también un “*insert*”, una gráfica de menor tamaño dentro de la principal en la que se enfatiza algunas característica.

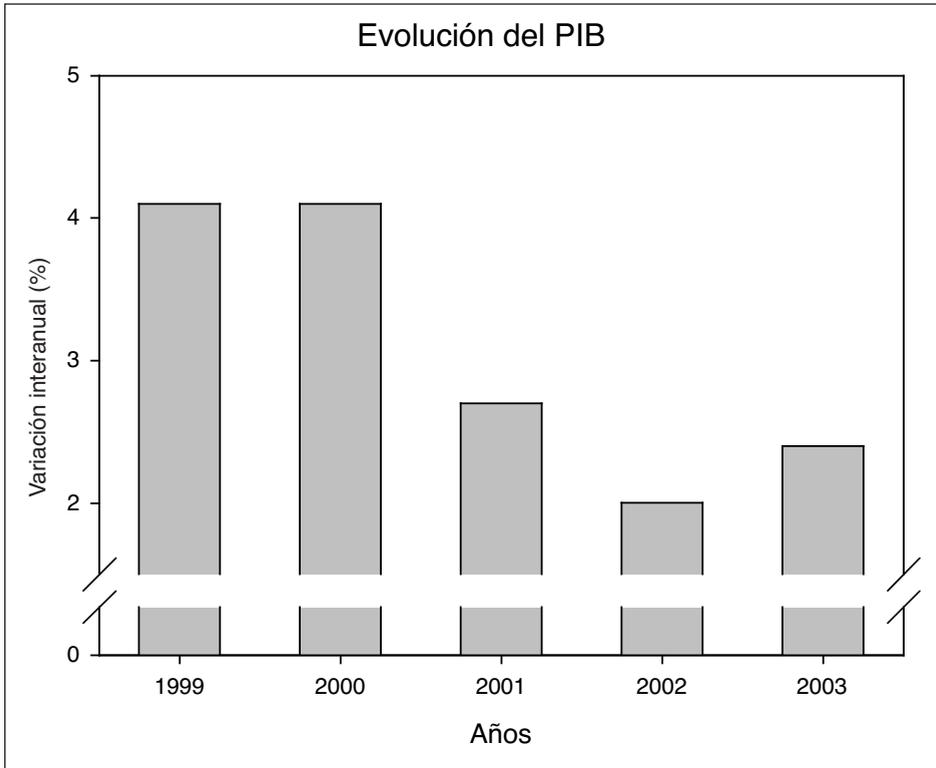


Figura 52. Una nueva versión de los datos de la figura 12, en la que se incluye la fractura del eje de ordenadas.

Es más fácil detectar una expresión incorrecta que sugerir una correcta y que además incluya toda la potencia expresiva que buscaba el autor. Aquí estamos asentando los rudimentos de una gramática de las gráficas, algo muy lejano de un curso sobre creación literaria en ese lenguaje. Lo que es claro es que los ejes deben mantenerse íntegros, sin deformaciones, dado que lo contrario conduce a mensajes que no se entienden adecuadamente.

Resumen del capítulo

Hemos visto en primer lugar como la indefinición del sujeto de una oración gráfica la hace incomprensible, y como esa definición depende por una parte del receptor del mensaje y del contexto de la gráfica por otra.

En segundo lugar comprobamos que la elección de un tipo de gráfico debe adecuarse a la estructura de los datos disponibles: para las variables categóricas hay que utilizar ejes categóricos, y para las variables numéricas, ejes numéricos.

Finalmente analizamos la necesidad de elementos de referencia en las gráficas: debemos disponer de ejes, explícitos, con el origen claro y que no estén deformados.

Capítulo 5

Sintaxis

Consideramos ortográficos los errores que hacían incomprendible la gráfica, los que daban lugar a una gráfica con un mensaje inexistente, imposible de recuperar o muy alejado del que realmente soportan los datos. Aquellos se derivaban de la utilización inadecuada de uno de los elementos constituyentes de la gráfica. Vamos a adentrarnos ahora en el estudio de un conjunto de situaciones que tienen más que ver con combinaciones inadecuadas de elementos, y que dan lugar generalmente a situaciones menos graves. No es el caso de la coherencia del sujeto, lo primero que veremos, que da lugar al mismo tipo de problemas, serios, que la discordancia entre tipo de gráfica y estructura de datos.

Coherencia del sujeto: unicidad en la magnitud representada

En el capítulo anterior vimos que el sujeto de las gráficas tiene que estar definido para que se pueda comprender algo. Pero además tiene que ser coherente, homogéneo. Una oración puede contener más de un sujeto. Cuando ocurre eso decimos que tenemos una oración compuesta, y la podemos des-

componer en oraciones simples, una para cada sujeto. Pero una vez que estamos considerando una oración simple, todos los datos han de ser cuantificaciones de la misma magnitud. Si el sujeto es la producción de peras, no podré mezclar datos de producción de manzanas. Estos comentarios resultan demasiado obvios o demasiado abstractos. Un ejemplo ayudará a ponerlos en su justo término.

En la figura 53 se presenta un párrafo gráfico ilustrativo del mercado hipotecario, compuesto por cuatro gráficas de dispersión y dos de barras, todas ellas simples. Me encontré con esta gráfica en un avión, y casi salto de alegría al ver que el precio de la vivienda estaba a punto de estabilizarse tras la brutal subida de los últimos años. Eso, al menos, es lo que se percibe al primer golpe de vista de la última de las gráficas de la infografía.

Pero al mirarla un poco más en detalle se comprueba que esa impresión procede de un artificio (¡el precio de la vivienda sigue subiendo!). La gráfica está mezclando datos anuales con datos trimestrales, y la suavización de la pendiente se debe al cambio de un tipo de datos a otro. La representación de los datos trimestrales como equivalentes a los anuales (o peor aún dado que el intervalo es algo mayor) hace que la traza no se corresponda con la realidad. La gráfica del ejemplo está representando datos no homogéneos, el sujeto no es coherente. Y lo mismo ocurre en las dos primeras gráficas de la misma figura.

Esta falta de homogeneidad la hemos podido detectar porque queda explicitada en el eje x de las tres gráficas incorrectas. No es que exista información insuficiente o tergiversada para entender lo que se presenta, lo que ocurre es que la impresión visual no se corresponde con la realidad. Para saber lo que pasa (si realmente está dejando de subir el precio de la vivienda o no) hay que leer los datos y reconstruir mentalmente la tendencia. Para eso es mejor dar los datos en forma de tabla. Así al complejo proceso de construir un mensaje a partir de unos números al menos le eliminamos la fase de destruir un mensaje incorrecto.

Otro ejemplo, quizá más claro, lo encontramos en la figura 54. En ella vemos un párrafo gráfico con dos oraciones relacionadas en las que la subordinada es una gráfica de tarta. La gráfica principal es una gráfica de dispersión en la que se representa el número de inmigrantes atendidos, en abscisas, frente al año en que se atendieron, en ordenadas. De los tres puntos que contiene uno no es conmensurable con los otros dos. Los puntos correspondientes a 2001 y 2002 representan el total de atenciones del año, mientras que el de 2003 muestra datos hasta agosto (según se indica en el propio gráfico).



Figura 53. Varias de las gráficas de este párrafo mezclan datos de años con datos de trimestres en la misma representación. Además de modificar el eje x, induce una gran confusión en la percepción de las pendientes.

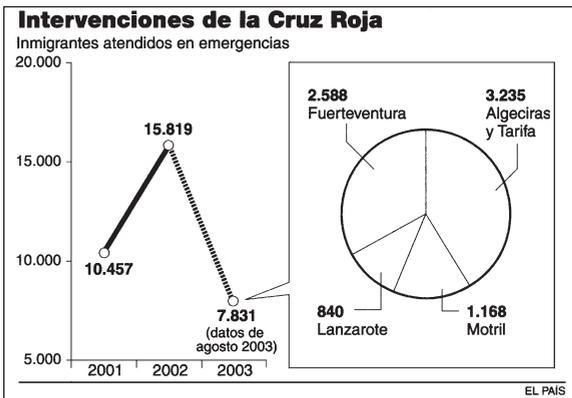


Figura 54. La primera gráfica de esta infografía muestra sólo tres puntos, y el segundo es inconmensurable con los otros dos.

La comparación entre datos relativos a un período y datos relativos a una fracción (7/12) del mismo no es posible. No sabemos si las atenciones se producen homogéneamente a lo largo de todo el año o si hay meses de mucha más actividad (como casi seguro ocurre). Esos datos imposibles de comparar, cuando se plasman en una representación gráfica que enfatiza la percepción visual de las comparaciones no pueden llevar más que a confusión. De hecho, a primera vista (antes de reparar en la diferencia del último punto) lo que parece es que la curva desciende abruptamente, cuando en realidad la tendencia probablemente es justo la contraria.

En esta gráfica no se ha respetado la coherencia del sujeto: se han mezclado peras y manzanas. “Inmigrantes atendidos a lo largo de un año” e “inmigrantes atendidos a lo largo de 7 meses” son magnitudes distintas, y una oración simple como la que estamos considerando no puede construirse correctamente sin unicidad en la magnitud representada, sin coherencia en el sujeto¹.

La heterogeneidad en la definición de la magnitud representada altera las pendientes de forma que se pierde la impresión visual que transmiten las gráficas. El hecho de explicitar la heterogeneidad, bien en el rótulo del dato, como en la figura 54, bien en la escala, como en la figura 53, sirve para llamar a la reflexión y observar la gráfica con más detenimiento, pero en cualquier caso hace que se pierda la misión fundamental de la representación: transmitir un mensaje claro con un golpe de vista.

Proporción en ejes y escalas

Todos los elementos gráficos se caracterizan por ser percibidos visualmente. Las gráficas producen una impresión visual que es la que genera la comprensión del mensaje: fisiología de la visión y psicología cognitiva. Ya nos hemos asomado a estas disciplinas y hemos visto las ideas principales. Por ejemplo el hecho de que las pendientes (los ángulos) son una de las características que se percibe de una forma más clara.

1. Si para preservar la unidad se elimina el punto sobrante queda una gráfica con dos puntos, el mínimo posible, y una muy mala opción para expresar la información que contienen.

En la figura 55 se muestra una gráfica que representa la evolución temporal del número de trabajadores² de la telefónica vasca. La forma de representar esos datos presenta una exagerada desproporción en el tamaño de sus dos ejes, muy largo el x y muy corto el y. La traza resultante es muy alargada, y apenas transmite impresión ninguna, parece casi horizontal. Para apreciar mejor el significado de estos datos podemos redibujarlos, pero también podemos ser más drásticos y, utilizando las ventajas del mundo digital deformarla hasta hacer sus dos ejes casi iguales (figura 56). El resultado queda tan comprimido que no se pueden leer los rótulos, pero el mensaje gráfico se hace mucho más patente: “sin dejar de crecer, a partir del 2000 esto ocurre a un ritmo mucho menor”. Un predicado que no se corresponde exactamente con el titular de la infografía: “Expansión de la telefonía vasca”.

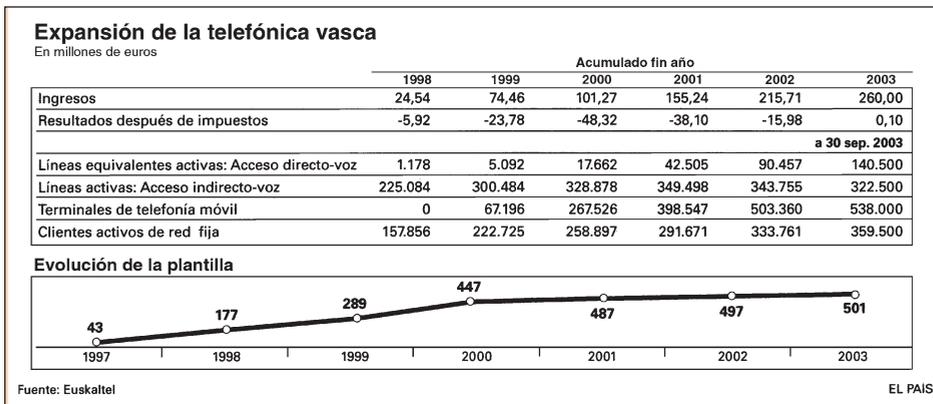


Figura 55. Una gráfica con una exagerada desproporción en el tamaño de sus dos ejes, resultando por ello poco efectiva.

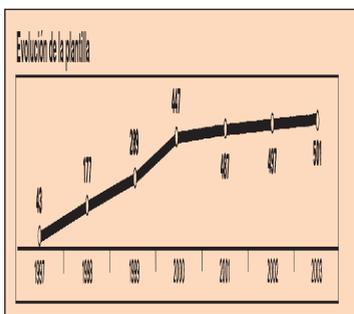


Figura 56. Versión deformada de la grafica de la figura anterior. Se aprecia claramente ahora un cambio de tendencia en el centro.

2. La ausencia de un eje explícito no deja claro este hecho, pero es de suponer que cuando dice “evolución de la plantilla” esta viene cuantificada en número de trabajadores.

Quizá exista una motivación detrás de la desproporcionada representación de los datos, el intento de que los datos no estropeen un buen titular³, pero quizá fuera una simple cuestión de composición: se trataría de llenar el mismo espacio que con la tabla de datos que va en el mismo párrafo gráfico. Sea como fuere, sintácticamente los ejes deben guardar una mínima proporción.

Proporción en la densidad de información

No sólo los ejes deben estar proporcionados entre sí, también esperamos una proporción adecuada entre la cantidad de información y el espacio ocupado por la gráfica para que resulte un buen elemento de comunicación.

Un ejemplo de desproporción lo vemos en la figura 57, donde se reproduce una figura que vimos ya en la introducción, a propósito de intereses espurios de los autores a la hora de incluir una gráfica. En este caso con 7 datos se rellena una gráfica que ocupaba en su versión original 153 cm² (9 x 17). Como además esta gráfica se presentaba sola en una página, el área dedicada a los siete puntos asciende al folio⁴ completo: 630 cm².

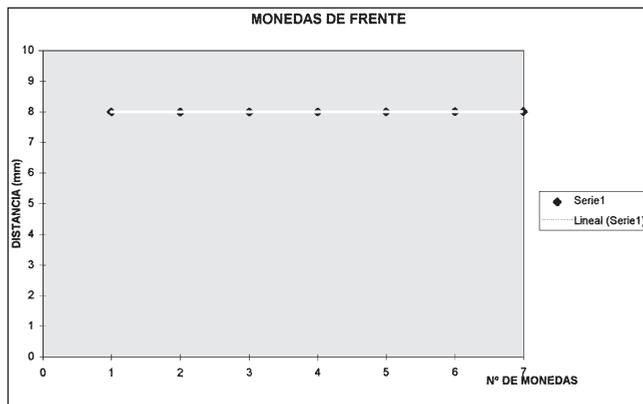


Figura 57. Siete datos de un experimento dan el mismo resultado.

La representación gráfica de los mismos no aporta más valor que una frase de texto.

3. Que es la adaptación a las gráficas de la famosa máxima de los malos periodistas: “no dejes que la realidad te estropee un buen titular”. Más de 30 referencias a esta frase en google.

4. En realidad se trataba de un DIN A4 (210 x 297 mm). Entiéndase “folio” para referirse a una hoja entera.

Estos datos nos permitirían calcular un índice numérico de densidad de información en forma de datos por centímetro cuadrado⁵, obteniendo unos valores de 45 milésimas para la gráfica, o de 11 si consideramos el folio completo. Es un indicador interesante, sobre todo porque está bien definido, y se puede calcular para cualquier gráfica. Lo que ocurre es que no sirve para decidir si una gráfica es mejor que otra.

En la figura 58 se presenta otra gráfica del mismo estudiante que la anterior, procedente de un experimento análogo. La gráfica original ocupaba aproximadamente el mismo espacio, pero el número de datos es menor, ahora son sólo 5. El índice de densidad de datos es de 33 milésimas frente a las 45 de la figura 57. Sin embargo, la inutilidad del gráfico no es ni mucho menos tan grande. Estos datos sí muestran una variación, hay un predicado no trivial: “a medida que aumentamos el área de la pieza de madera que acercamos a un detector más lejos lo detecta, y la relación entre ambos es aproximadamente lineal, aunque para áreas muy pequeñas parece que detecta peor...”.

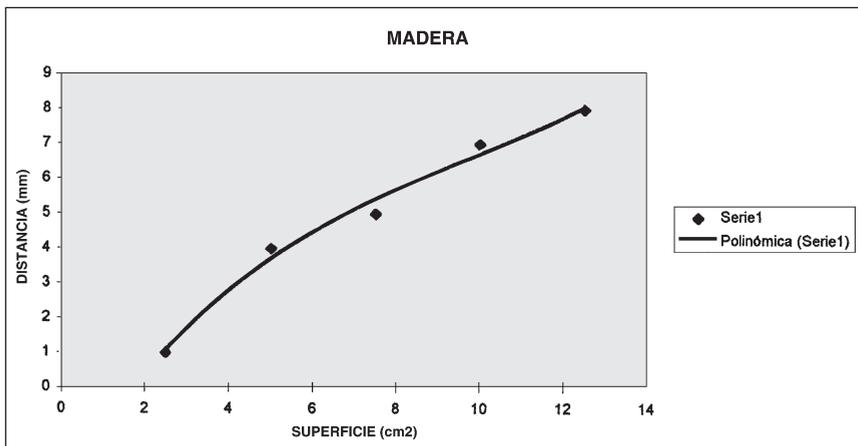


Figura 58. Una figura con menor densidad de datos por unidad de área pero con mayor información.

El indicador que nos ayudaría a decidir cuándo una gráfica es mejor que otra sería el de “información” por unidad de área, pero esa “información” es

5. Así define Tufte un índice en su libro *The visual display of quantitative information*, en la página 162, sólo que en pulgadas, como buen americano que es.

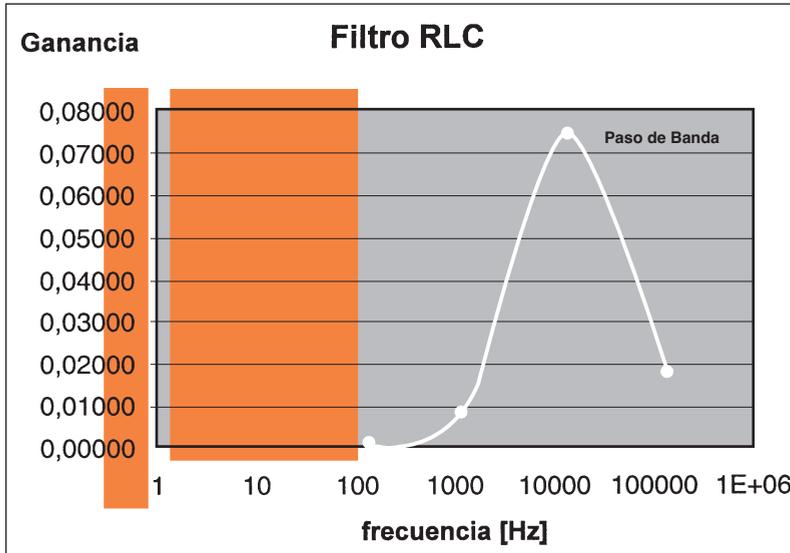


Figura 59. Sin necesidad de indicadores cuantitativos es claro que el área de las zonas anaranjadas se debería eliminar, dado que no aporta nada a la representación

una magnitud mucho más difícil de definir. Si no pretendemos desarrollar una ciencia sobre todo esto sino ofrecer unas guías sencillas, todo intento de cuantificación resulta como matar moscas a cañonazos. Nos quedamos pues con el análisis cualitativo de algunos casos que llaman la atención.

En la figura 59 vemos una vez más cómo un estudiante presenta pocos datos en mucho espacio. De ese espacio hay una parte que sobra de forma evidente. Casi la mitad de la escala del eje x no contiene información, ese trozo sobra. Además, el formato en que están dados los datos del eje y resulta muy inadecuada y, cuando menos, todos los ceros del final de los números deberían no estar (estos elementos sobrantes se indican en naranja en la figura).

Además de la densidad de datos por unidad de superficie utilizada, Tufte⁶ introduce la razón datos-tinta, como la proporción entre la tinta empleada para representar los datos y la tinta total en el gráfico. Sostiene a continuación que los gráficos son mejores cuanto mayor sea ese cociente. Aceptando ese principio, por supuesto los fondos de las gráficas de las figuras 57 a 59 ha-

6. E. TUFTE, *The visual display of quantitative information*, p. 93.

bría que eliminarlos, pero además habría que continuar con las marcas de los ejes: no hace falta indicar unidad a unidad, con la del comienzo o la del final basta. Continuando el procedimiento se llega a unas gráficas ciertamente minimalistas que, en muchos casos resultan mucho más difíciles de comprender que versiones ligeramente más “recargadas”.

Ya desde la introducción sabemos que lo que necesitamos para hacer gráficas adecuadas es “tomar conciencia de su utilidad comunicativa y hacer lo necesario para que ésta resulte mínimamente efectiva”. La cuestión ahora es si hay algún índice numérico que proporcione una receta infalible para la efectividad comunicativa. Y la respuesta, según hemos ido viendo, es que no. Hay que buscar un equilibrio entre el espacio ocupado, la tinta empleada y la información transmitida de forma que la comprensión se optimice. No deberemos incluir espacios inútiles ni añadir escalas, leyendas, fondos y otros elementos (tinta) que no contribuyan, pero no hay que exagerar la poda. El criterio habrá de estar en el sentido común del autor.

Una mínima composición (efectos visuales)

Describiendo la historia reciente de los gráficos estadísticos, Tufte incluye una elegía de John W. Tukey, una de cuyas obras alaba porque “huye de intentos torturados de construcción de *estándares gráficos* en un desesperado intento por acabar con las distorsiones”⁷. Leer una frase tan cruel le hace a uno recapacitar. Espero que estas notas no resulten ni torturadoras ni desesperanzadoras, porque la verdad es que sí que intentan generar, si no unos estándares, al menos unas guías para una mejor realización de gráficas que, entre otras cosas, eviten las distorsiones. Ahora bien, las mismas obras de Tufte parecen tener el mismo propósito...

En todo caso, a medida que nos alejamos de las cuestiones más fundamentales que son inexcusables para que una gráfica tenga sentido, las recomendaciones van siendo más débiles. Ya veíamos en el apartado anterior que sólo podemos sugerir equilibrio. Lo mismo nos va a ocurrir con las restantes cuestiones de composición que vamos a considerar a continuación.

En la figura 60, se presenta una nueva gráfica de esa fuente inagotable de sorpresas que constituyen los estudiantes. En ella se presentan datos de calen-

7. De la misma obra, p. 53.

tamiento de un horno, correspondiendo los datos a la temperatura alcanzada al cabo de un tiempo dado de su conexión (hay dos trazas para dos potencias de calentamiento). La primera cosa sorprendente es la elección de los ejes; lo habitual, y lo que hicieron todos los demás compañeros es representar la temperatura alcanzada (eje y) para un tiempo dado (eje x). Aquí se presenta el tiempo necesario (eje y) para llegar a una temperatura dada (eje x), que es lo mismo pero de una forma inusual. Por otro lado se le dedica al rótulo identificativo de las trazas (la leyenda) la misma superficie que al conjunto completo de los datos. Además el título del eje y está escrito en horizontal, con lo que se desperdicia otra buena porción de superficie, y en cambio las marcas del eje de temperaturas están en vertical, y resulta antinatural su lectura.

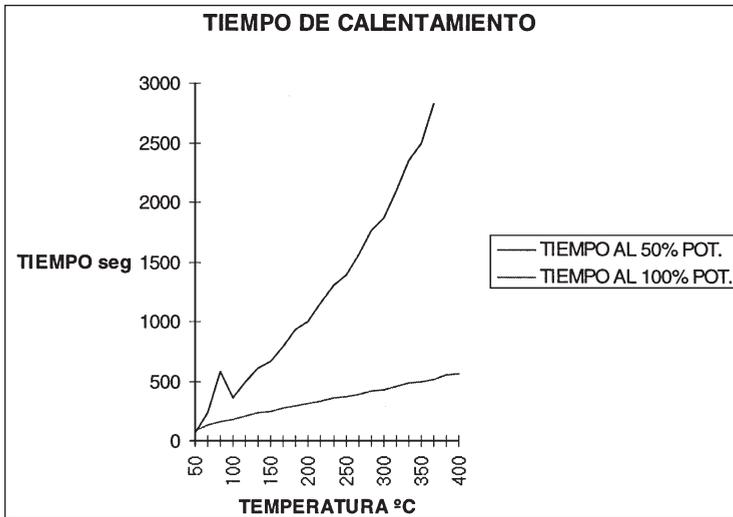


Figura 60. Un ejemplo de elección incorrecta de elementos de la composición: rótulos verticales en el eje x. Título horizontal del eje y, tamaño de la leyenda, etc.

Ninguna de las objeciones antes indicadas hace que la gráfica resulte ininteligible, pero entre todas la hacen fea e incómoda. ¿Se puede dar una lista para que esto no pase? Difícilmente. Ya veíamos cómo Tufte considera desesperados y torturados los intentos de hacerlo. Baste pues llamar la atención sobre este tipo de cuestiones básicas de composición del gráfico y esperar de las capacidades del autor su adecuada consideración.

Un cuestión de composición que merece una atención especial es la formación de efectos visuales. Determinadas configuraciones gráficas son perci-

bidas y procesadas por el cerebro de tal forma que se toma conciencia de cosas que no están en la figura realmente. En la figura 61, se presenta un caso muy exagerado, en el que unas combinaciones de franjas circulares dibujadas (y por tanto estáticas) se perciben como dotadas de un movimiento giratorio. Se haga lo que se haga para intentar aislarse de esa percepción errónea, es imposible (al menos yo no lo consigo).

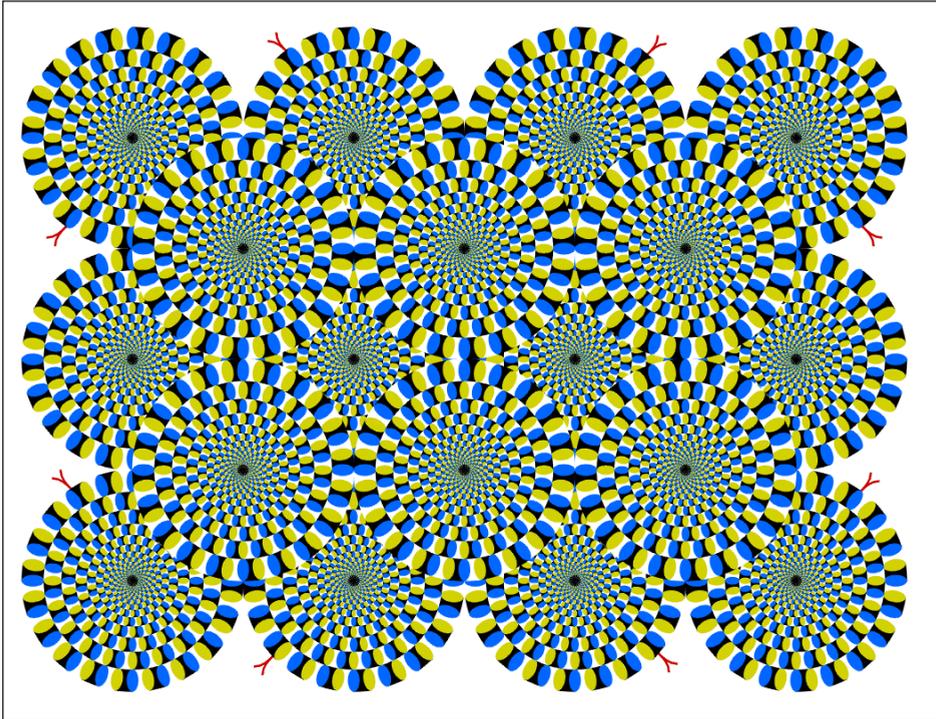


Figura 61. Ejemplo de ilusión óptica exagerada, aunque sea difícil de creer, nada se mueve.

Sin necesidad de que el efecto sea tan exagerado, hay una multitud de efectos visuales, franjas de Moiré, ilusiones ópticas, etc. que generan percepciones incorrectas o distorsionadas de lo representado. Tal como hemos analizado con anterioridad, la comprensión de una gráfica (entre otros gráficos) se basa en la fisiología de la percepción y en la psicología cognitiva. Así pues, todos los trucos que se conocen bien en estas disciplinas, si aparecen en una gráfica, darán lugar a problemas en su interpretación. Un caso mucho menos

espectacular que el de la figura 61, pero procedente de un estudiante, se presenta en la figura 62 ⁸.

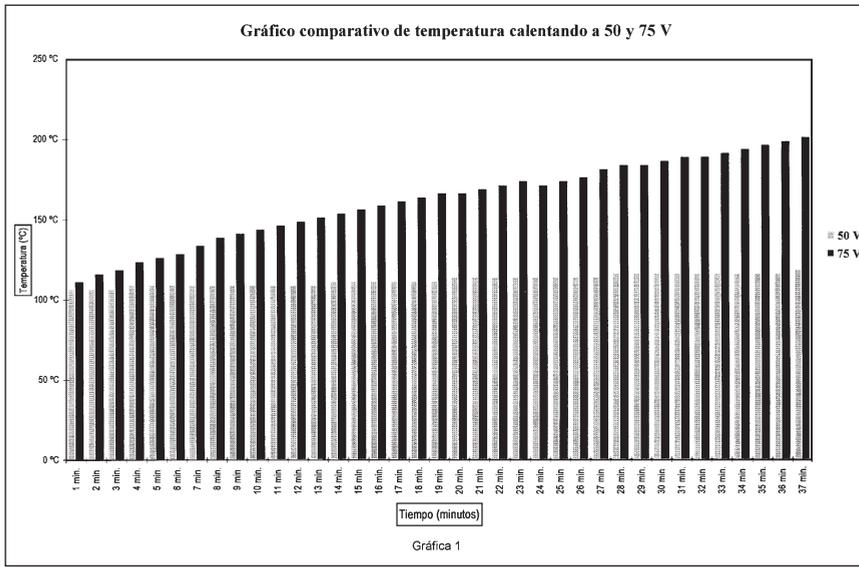


Figura 62. Una gráfica en la que las dos series de barras (y la parte blanca entre ellas como una tercera) producen una cierta impresión de vibración.

Resumen del capítulo

En este capítulo dedicado a las relaciones entre los elementos que componen un gráfico hemos analizado cinco ideas fundamentales:

8. Aunque para nuestros rudimentos de gramática gráfica estas consideraciones son más que suficientes, el tema de las ilusiones ópticas es apasionante. Se puede encontrar más información en:

- Aplicada a gráficas, en Tufte, capítulo 5.
- Un libro: Cassin, C. *Visual Illusions in Motion with Moiré Screens: 60 Designs and 3 Plastic Screens*. New York: Dover, 1997.
- En internet: <http://www.sandlotscience.com/>
- Otro libro: I. Amidror, *The Theory of the Moiré Phenomenon*, Kluwer Academic Publishers, 2000.
- Más internet: <http://www.exploratorium.edu/snacks/iconperception.html>
- Y un poco más: <http://www.bu.edu/smec/lite/moire/>

1. En la misma serie de datos (el sujeto de una oración simple) no se pueden mezclar números que no corresponden a la misma magnitud.
2. Los ejes deben estar proporcionados
3. El espacio ocupado (y la tinta consumida) deben ser acordes a la cantidad de información transmitida.
4. Hay que tener cuidado de que la composición global resulte correcta (las orientaciones de las letras en los rótulos, las posiciones de los ejes, leyendas, etc.)
5. Conviene evitar efectos visuales indeseados (franjas de Moiré, ilusiones ópticas, etc.)

Salvo la primera (y un poco la última) que es más normativa, las demás cuestiones apelan a la proporción, algo que admite bastante variación subjetiva.

Capítulo 6

Semántica

Tras estudiar las características formales que hacen que una gráfica resulte inteligible (ortografía), y las relaciones entre elementos que hacen que ese significado se perciba sin dificultades añadidas (sintaxis), nos falta adentrarnos en la relación entre la forma de la representación y el significado, la semántica.

Desde el comienzo del texto estamos haciendo consideraciones respecto de cuál podía ser la intención del autor al realizar la representación de una forma concreta al hilo de otras cuestiones. Este es el momento de centrarse en estas cuestiones en sí mismas.

“Para muchas personas la primera palabra que viene a la cabeza cuando piensan sobre gráficas estadísticas es *mentira*”¹. Con una afirmación tan rotunda comienza Tufte un capítulo sobre la “integridad gráfica”. Si pensamos en mentira “dolosa”, con voluntad de mentir, la frase es seguramente muy

1. “For many people the first word that comes to mind when they think about statistical charts is ‘lie’”. Tufte, p. 52. No es fácil encontrar la relación exacta entre los términos en inglés “chart”, “graph”, “plot” y “graphic” y los castellanos “gráfico” y “gráfica”.

exagerada, pero si consideramos las distorsiones debidas a descuidos y errores probablemente no lo sea tanto.

Vamos a ver a continuación la conveniencia de explicitar el mensaje a la hora de enfrentar una representación gráfica: para que en la gráfica no se introduzca una mentira por descuido nada mejor que conocer la verdad. A continuación veremos algunos ejemplos de mensajes gráficos que no se corresponden directamente con los datos que allí se representan².

Adecuación de la representación al mensaje buscado

A poco que los datos disponibles no sean demasiado triviales del sujeto, de la magnitud representada, se pueden decir muchas cosas distintas, y cada uno de estos predicados resultará mejor transmitido con una representación diferente. Un ejemplo lo podemos ver en las gráficas de la figura 63. Los datos allí representados proceden de una práctica de laboratorio en la que los estudiantes debían diseñar un sistema automático de control de temperatura de un horno.

Se desea mostrar la estabilidad del dispositivo de control de la temperatura. Se requería en la práctica mantener la temperatura a 300 °C durante varias horas, no superar los 400 °C en ningún momento y alcanzar el valor estable deseado (300 °C) en un tiempo corto. Habrá que buscar una representación en la que se pueda comprobar si estos requisitos se cumplen, como la mostrada en la figura 63.

Por el contrario, si se desea mostrar la inestabilidad del dispositivo, habrá que buscar otra representación distinta. A una variación temporal la podemos calificar de estable o inestable dependiendo de la situación, del mismo modo que 37cc de whisky se pueden percibir como botella medio llena o medio vacía. El concepto de inestabilidad depende del problema, y en la práctica se solicitaba que el dispositivo mantuviese la temperatura en 300 °C con una desviación máxima de $\pm 5\%$. Necesitamos una escala que, a diferencia de la de la figura 63, permita apreciar este tipo de variaciones. La gráfica de la figura 64 es una representación de los mismos datos que en la figura 63, diseñada para apreciar el nivel de inestabilidad, que por cierto no cumple las especificaciones requeridas.

2. Dado que son ejemplos reales de periódicos, da un poco de respeto hablar directamente de “mentira dolosa”, aunque dan toda la pinta.

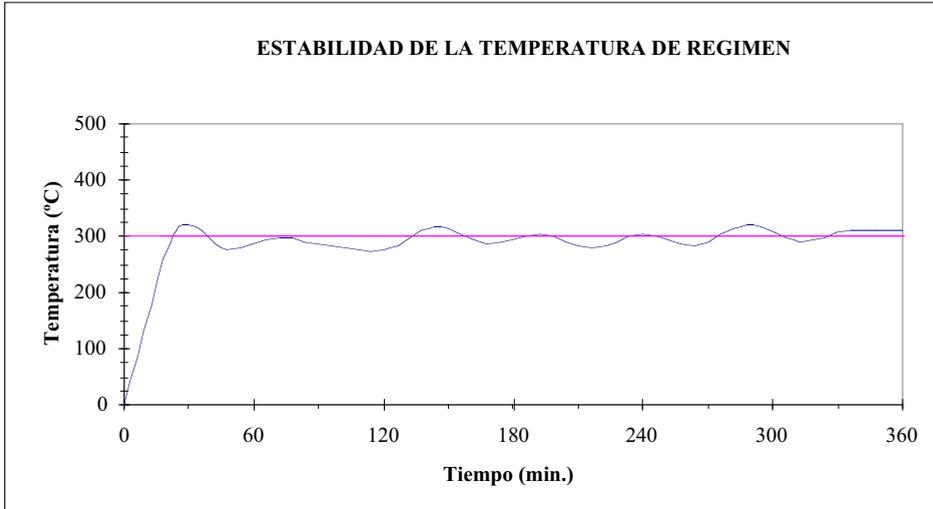


Figura 63. Evolución de la temperatura de un horno controlada por un dispositivo diseñado en prácticas por unos estudiantes. La representación está pensada para apreciar una serie de requisitos de estabilidad que se hacían en la práctica.

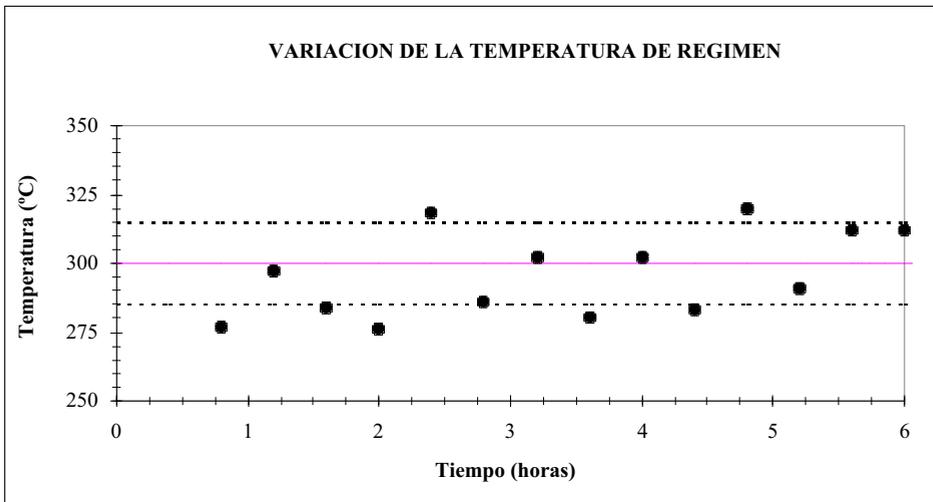


Figura 64. Los mismos datos de la figura anterior en esta representación evidencian el no cumplimiento de un criterio más.

Ninguna de las dos gráficas es incorrecta, y una no es mejor que la otra. El sujeto es el mismo en ambos casos, no sólo la magnitud, los datos son los mismos, pero la representación está especialmente adaptada en cada caso para transmitir el mensaje que de ella se pretende. En el caso de la figura 63, velocidad de subida, estabilización alrededor de 300° y no superación de 400°. En el caso de la figura 64, el predicado es que la estabilidad se sale, aunque ligeramente, de la banda de más menos 5% de los 300 °C.

Los estudiantes que incluyeron estas dos gráficas en su guión acertaron a encontrar una forma gráfica sencilla de comprobar el grado de cumplimiento de cada uno de los requisitos solicitados por el profesor: con dos gráficas diferentes. Aunque sean representaciones de los mismos datos la información que se transmite no es redundante, cada gráfica está adaptada a un mensaje distinto y, si se quieren transmitir los dos, ambas son necesarias.

Se puede uno plantear si hay formas de transmitir todos los mensajes que se buscan en este ejemplo con una sola gráfica. Casi seguro que es posible, y probablemente de muchas formas distintas, pero esas cuestiones corresponden más a un curso de literatura que a un manual de iniciación a la gramática. Lo que aquí nos interesa es remarcar que la representación gráfica debe adaptarse al mensaje que se pretende transmitir. En general, eso se podrá conseguir de muchas formas diferentes, y unas serán más eficaces, otras más bonitas, otras más compactas, otras más vistosas, otras más rigurosas... Hay diferentes estilos literarios gráficos.

Demagogia en el párrafo (¿De verdad se pueden comparar esas gráficas?)

Los estudiantes que presentaron las anteriores gráficas fueron especialmente honestos, dado que el mensaje de la segunda gráfica era que no se cumplía uno de los requisitos que buscaba el profesor. Desgraciadamente esto no ocurre así siempre. Muy al contrario, es frecuente encontrar gráficas que transmiten un mensaje que, no sólo no está soportado por los datos, sino que en muchas ocasiones es opuesto. Hay múltiples formas de conseguir representaciones que no transmiten lo que los datos sugieren³, vamos a ver a con-

3. El capítulo de Tufte al que se hacía referencia al comienzo del capítulo es un buen ejemplo, en el que se define incluso un “factor de mentira” que permite cuantificar el nivel de

tinuación las desviaciones del significado de los datos que pueden aparecer por la agregación de gráficas, en el nivel del párrafo.

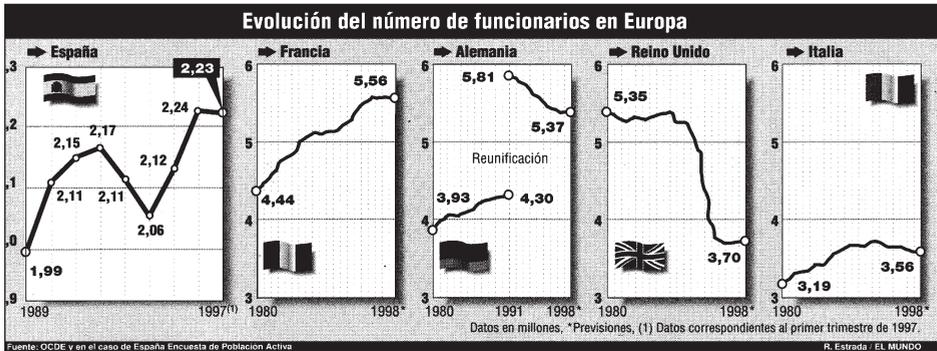


Figura 65. Demagogia en el párrafo. Cada una de las gráficas no plantea problemas relevantes, es la comparación entre la primera y las demás la que conduce a una impresión visual no soportada por los datos (ver figura 66).

La primera situación a considerar es aquella en la que se transmite un mensaje torcido sin incluir defectos formales⁴. Un ejemplo se muestra en la figura 65, en la que se reproduce una infografía de un periódico insertada para ilustrar la evolución del número de funcionarios en España. Esa ilustración se basa en la comparación con distintos países. El formato elegido es el de un párrafo compuesto por oraciones simples: los datos de cada país representados en una gráfica separada, y todas las gráficas se presentan juntas. La primera impresión visual es que España, junto con Francia quizás, presenta una evolución creciente muy contraria al signo de los tiempos marcado por los otros países⁵.

Para proceder a una segunda lectura más profunda se fija uno en las escalas, y ahí comienzan las sorpresas. Las gráficas de referencia, las de los otros cuatro países, están representadas de una forma idéntica: del 80 al 98 en años y de 3 a 6 en millones de funcionarios; tamaños, tipografías, todo idéntico, imposibilitando interpretaciones no apoyadas en los datos. Pero cuando mi-

fraude en algunos de los casos. Otra lista de formas de mentir en gráficas la encontramos en la dirección de internet: <http://www.math.yorku.ca/SCS/Gallery/> (Darts).

4. O siendo éstos muy nimios.

5. Esta idea es la que sostiene el titular de la noticia en la que se insertaba la infografía.

ramos la gráfica de España, la que se pretende comparar con esa línea de referencia trazada por los otros cuatro países, lo que encontramos es que las escalas son totalmente distintas.

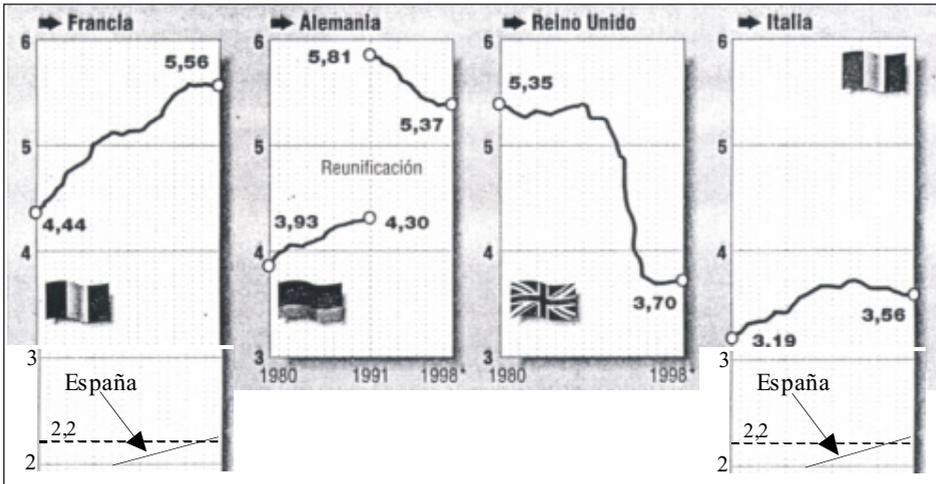


Figura 66. Los mismos datos de la figura 3, en los que la gráfica correspondiente a España se ha redibujado en las mismas escalas que las otras, y se presenta junto con la de Francia e Italia.

Con el fin de comprobar el efecto visual que produce una representación en escalas análogas se ha incluido la figura 66. En ella los datos de funcionarios españoles se han redibujado en la misma escala que los de referencia, y el dibujo resultante se ha añadido a dos de las cuatro gráficas de referencia. Lo primero que llama la atención es que en las escalas originales estos datos quedaban fuera del dibujo; tan fuera y tan pequeños que los detalles de la evolución temporal se tornan irrelevantes⁶. Lo que surge ahora a primera vista es que el número de funcionarios en España está muy fuera de la tendencia europea, y que el crecimiento que se observa es tan suave que se tardará décadas en llegar a esos valores. Una idea totalmente contrapuesta a la que plantea la primera representación.

6. De hecho se ha sustituido por una línea recta en la figura 66 la evolución que, vista en la figura 65 muestra un máximo en el año 92 y un mínimo en el 94. En la nueva escala esos detalles resultan totalmente irrelevantes.

Este ejemplo muestra cómo se puede transmitir una idea muy alejada de la realidad con gráficas que resultan impecables. La desviación en el mensaje aparece en el nivel del párrafo, no en el de las oraciones. Es la comparación entre distintas gráficas la que genera la impresión mental, y esa comparación se presenta con escalas incomparables. Por su parte, cada una de las gráficas del de forma aislada es perfecta⁷.

Esta situación se puede empeorar si las gráficas que conforman el párrafo adolecen, aun de forma aislada, de errores ortográficos o sintácticos. Este es el caso de la infografía presentada en la figura 67. En ella las gráficas que conforman el párrafo carecen de ejes para la variable representada (la inflación), sólo disponemos de unas indicaciones en algunos datos que nos pueden servir de guía para reconstruir el eje por nuestra cuenta⁸. Los valores presentados y la ausencia de líneas de referencia no facilitan esta tarea.

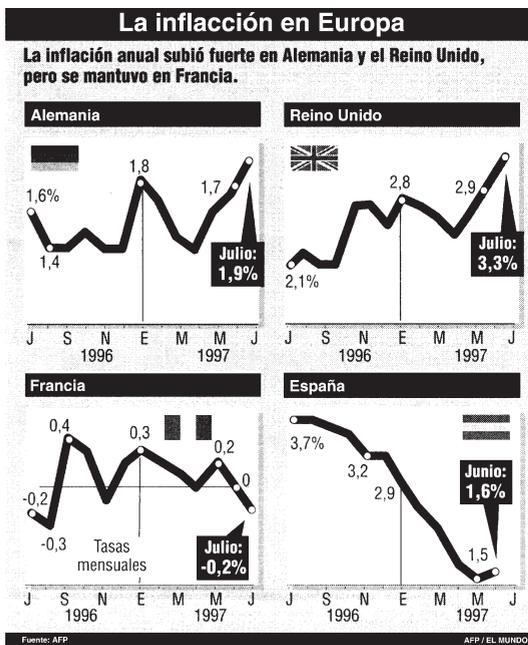


Figura 67. Demagogia en el párrafo, como en el caso de la figura 65, solo que aquí la ausencia de ejes explícitos de ordenadas (un error sintáctico), hacen más difícil desentrañar el engaño y recuperar la impresión original.

7. Suficientemente adecuada para lo que se discute aquí. Si se entra en detalles, la escala de años tiene 20 divisiones cuando debe corresponder a 18 años, lo que dificulta mucho extraer los datos que no vienen rotulados. Se podría entrar en otras objeciones, pero todo ello es muy menor.

8. Estos terribles ejes “hágaselo usted mismo”.

Pero hay una cosa que sí se puede observar rápidamente: los datos de Alemania y Francia son tan bajos que en la escala de España quedarían muy fuera, por debajo, del dibujo presentado; más incluso que en el ejemplo de la figura 66. Al elegir una representación basada en oraciones simples, y no incluir elementos de referencia adecuados, se hace casi imposible la comparación, que es justamente el objetivo comunicativo principal del párrafo. Sin el esfuerzo previo de análisis, la percepción a primera vista asume proporciones y referencias equivalentes y deriva mensajes equivocados.

Demagogia basada en el sujeto (¿Qué se representa exactamente?)

El ejemplo de los funcionarios que veíamos en las figuras 65 y 66, todavía nos ofrece una posibilidad de análisis más. Como veíamos, la magnitud representada en todas las gráficas era el número de funcionarios de cada país (en millones). Una magnitud bien definida, en unas unidades adecuadas... pero que no necesariamente es útil para comparar. Seguro que a nadie se le ocurriría comparar el número de funcionarios de Andorra con el de China, y no sólo por la diferencia de regímenes económicos, sino principalmente por la evidente diferencia de tamaño. Quizá entonces también fuera necesario considerar el tamaño para proceder a una comparación más adecuada entre países.

El ejercicio es, en efecto, muy revelador (figuras 68 y 69). Si consideramos proporciones en vez de números absolutos, España está en el rango de la mayoría, mientras que la anomalía resulta ser Francia. Como se ve en la figura 69a, de cada 100 habitantes aproximadamente 6 son funcionarios (o lo eran en 1998, año al que corresponden los datos), un poquito menos en España y un poquito más en Alemania, pero la cifra está cerca de 6 salvo en Francia que está en 9,6. Si la comparación se hace respecto de la población activa, en vez de la total, lo que se ve es que de cada 100 trabajadores 13,5 son funcionarios (13,2 en el Reino Unido y 14,4 en España son los valores extremos) salvo en Francia que hay 22,7 (figura 69b).

El ejemplo nos sirve para comprobar que el sujeto elegido para las gráficas de la figura 66 no era el más conveniente. Se puede argumentar que para ver la evolución temporal no es tan importante tener magnitudes estrictamente comparables, pero en realidad la evolución temporal puede resultar irrelevante si los valores cuya evolución comparamos no son equivalentes.

Para transmitir un mensaje sobre un tema es importante elegir un sujeto que efectivamente resulte representativo del tema en cuestión. Si queremos

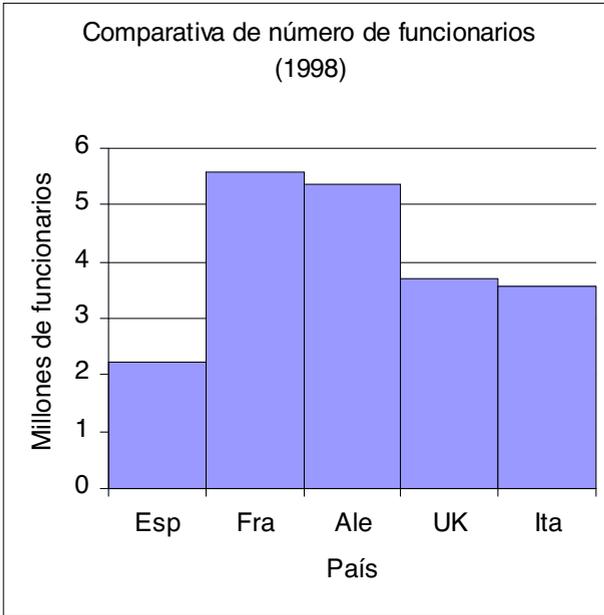


Figura 68. Comparación del número de funcionarios en el año 1998 entre distintos países de la Unión Europea. Datos de número total de funcionarios.

comentar si tenemos muchos o pocos funcionarios, y su evolución es exagerada en su crecimiento, habrá que ser riguroso en la elección de los indicadores adecuados⁹. En esta línea de elegir un sujeto adecuado para una argumentación nos podemos encontrar con casos verdaderamente exóticos, como el mostrado en la figura 70.

Todos sabemos que el precio de la vivienda en España es alto y sobre todo que está subiendo desmesuradamente en los últimos años. Lo que no tenemos tan claro es si en esa subida nos acompaña el resto de países vecinos o si somos una excepción. Para aclarar esta cuestión se puede ir a las fuentes de datos estadísticos y mostrar algo así como la tasa de crecimiento del precio de una vivienda media (el porcentaje de subida) para distintos lugares, o alguna otra variable fácil de entender. Eso sí, a lo mejor resulta que las subidas en España no son tan exageradas comparativamente, igual que pasaba con los funcionarios.

9. Quizá se le haya dedicado demasiado detalle a este ejemplo, pero el carácter de funcionario del autor impelía a dejar en este caso el contenido tan claro como las conclusiones respecto de las gráficas utilizadas.

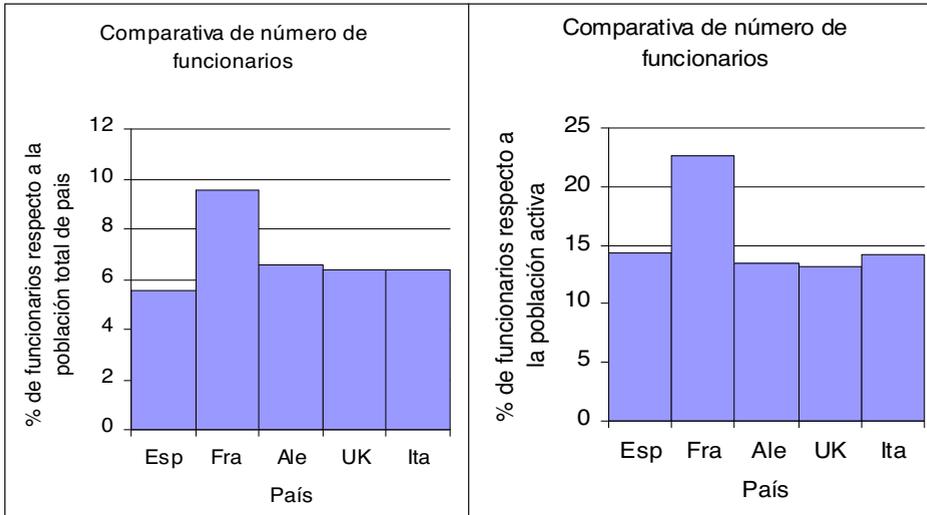


Figura 69. Comparación del número de funcionarios en el año 1998 entre distintos países de la Unión Europea. Datos de porcentaje de funcionarios respecto de la población total del país (a la izquierda) y respecto de la población activa de cada país (a la derecha).

Por alguna razón desconocida, quizá buscando un indicador en el que España quedase como campeón europeo, en la figura 70 se representó “la media entre los años 1999 y 2002 de la variación de la tasa anual de aumento de precio de la vivienda respecto de la tasa media de la zona euro”. Sencillo ¿verdad? Bueno, y esto suponiendo haber interpretado bien la información que da la gráfica sobre la variable representada.

Además de escoger una variable ciertamente exótica para la representación, los valores correspondientes a cada país se unen con una línea. Ya vimos en el capítulo sobre ortografía un caso análogo. Esto supone tratar una variable categórica como si fuese numérica, un error importante de dimensión del gráfico. Además la mezcla de esa representación lineal con otros dos gráficos de barras (datos de las variaciones de la tasa anual respecto de la media europea para dos años concretos) resulta confusa. Resumiendo, podemos encontrar en esta gráfica errores de todos los tipos. Como no puede ser de otra forma el resultado es nulo desde el punto de vista de la comunicación: no se entiende nada. O quizá peor que nulo: a primera vista deja un mensaje confuso e indocumentado: “el precio de la vivienda en España es el que más crece de Europa”.

Nuevas previsiones del BCE para la zona euro

Crecimiento económico

Dic. 2003	2003	2004	2005
BCE	0,2-0,6	1,1-2,1	1,9-2,9

Inflación

Dic. 2003	2003	2004	2005
BCE	2,0-2,2	1,3-2,3	1,0-2,2

Diferencias en las variaciones de los precios de la vivienda entre los países de la zona euro

Variación de la tasa anual respecto de la media de la zona euro

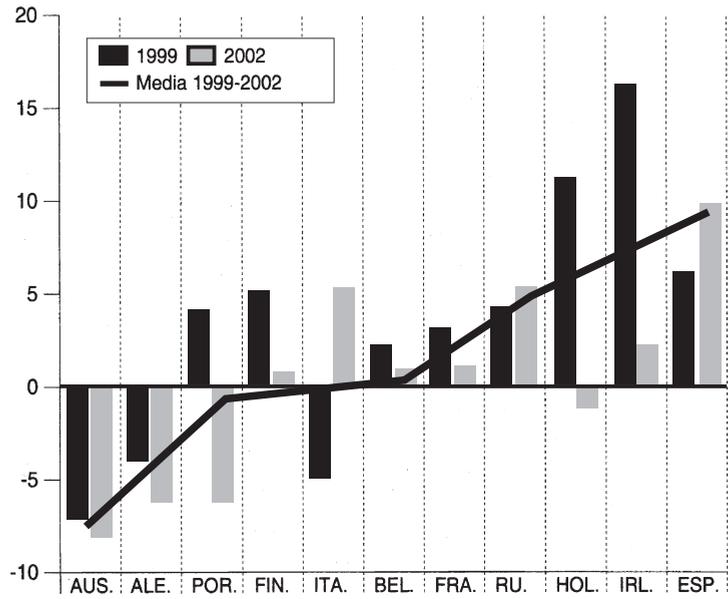


Figura 70. Representación de “la media entre los años 1999 y 2002 de la variación de la tasa anual de aumento de precio de la vivienda respecto de la tasa media de la zona euro”. Exótica variable en la que España resulta vencedora (al menos de los países representados).

Un caso más de extraña variable para representar es el mostrado en la figura 71. En este caso la variable propiamente dicha no es extraña, se trata del porcentaje de embarazos que acaban en aborto, en función de la edad de la madre y del año en que se produce el embarazo. El año se diferencia en la representación por el nivel de gris de la barra, se realizan distintas gráficas para distintas situaciones de edad de la madre. Y aquí es donde aparece la extrañeza. Las tres gráficas representadas cubren: de 15 a 19 años la primera, menores de 18 la segunda y de 18 a 19 la tercera.

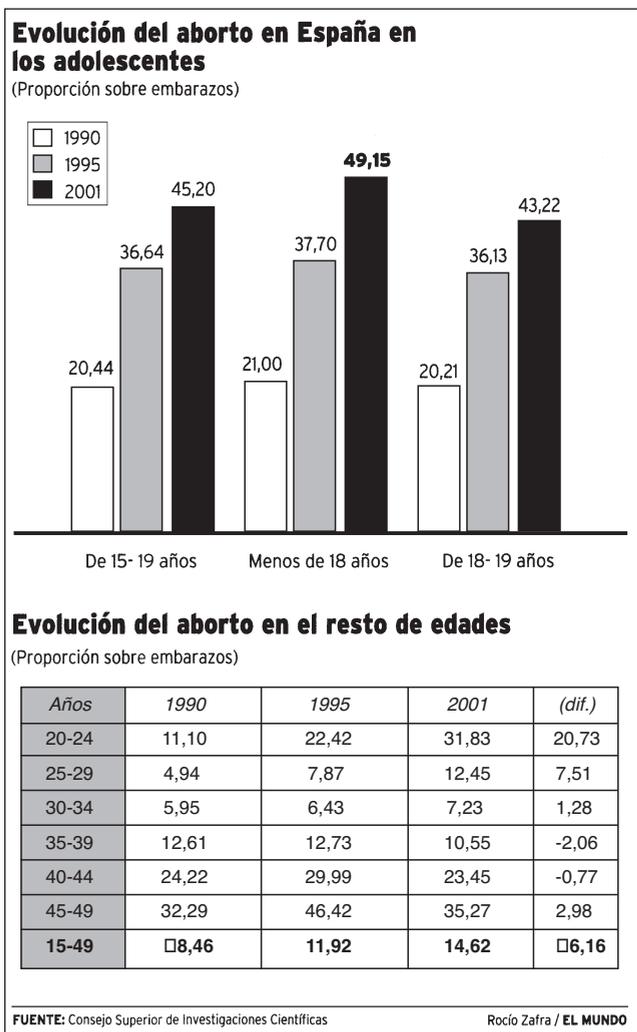


Figura 71. La variable representada es extraña, y lo es por las agrupaciones de edades seleccionadas: de 15 a 19, menor de 18 y de 18 a 19. ¿Qué pasa con otras agrupaciones de años? ¿Por qué no se da de año en año?

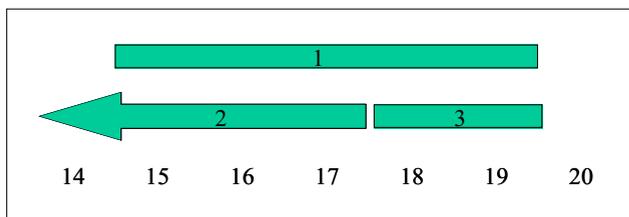


Figura 72. Cobertura en años de la madre de cada una de las tres gráficas de la figura anterior.

Esa distribución de agrupaciones de edades de la madre no se observa que responda a ningún patrón. Como se puede ver más claramente en la figura 72, no hay relación posible entre los tres grupos de edades (a pesar de que los datos originales son necesariamente redundantes). Como además son datos de porcentaje, no se pueden hacer sumas o restas con los resultados... ¿Qué se pretendía al elegir esta representación? ¿Cuál es el mensaje? Si se trata de mostrar que las madres adolescentes tienden a abortar más con los años, parece claro. Pero esa tendencia también se observa hasta los 30 años¹⁰ de la madre. Por otro lado una sola de las tres gráficas deja la idea rotundamente clara: la proporción de embarazos en madres menores de edad que acaban en aborto ha pasado de uno de cada cinco a uno de cada dos en una década (esto es lo que nos dice la gráfica del centro).

De los ejemplos anteriores concluimos que es muy importante elegir un sujeto, una magnitud a representar, que efectivamente sea representativo del mensaje que se pretende transmitir con la gráfica. Del mismo modo, una elección rebuscada puede conducir mensajes no soportados por los datos: demagogia gráfica.

Resumen del capítulo

En este capítulo hemos estudiado la relación entre la forma de representar los datos y el mensaje que con ello se busca. En primer lugar vimos como distintas representaciones de los mismos datos hacen patentes predicados muy distintos.

A continuación entramos en la demagogia gráfica. Bajo las preguntas ¿de verdad son comparables esas gráficas? y ¿qué es lo que se representa exactamente? vimos dos formas comunes de transmitir mensajes no soportados por los datos, bien a través del párrafo (la infografía completa) bien simplemente con una oración, con una gráfica.

10. Aunque en forma de tabla, se presentan los mismos datos de porcentaje de embarazos que terminan en aborto para las siguientes edades de la madre, agregados en paquetes de 5 años, y para los mismos años en la infografía.

Capítulo 7

Estilos literarios y síntesis final

Estilos y literatura

En los lenguajes escritos podemos identificar estilos muy diferentes. Una de las características que se puede analizar es la longitud de las oraciones. Hay escritores como William Faulkner que, antes de poner un punto, encadenan una oración subordinada de otra hasta más de diez veces. También Alejo Carpentier es un ejemplo de prosa muy intrincada en la que encontrar un punto y aparte es una aventura de varias páginas. Por el contrario hay prosas sincopadas, con oraciones simples y breves. Bert Easton Ellis en *Menos que cero*, o *Historias del Kronen* de José Ángel Mañas, son ejemplos claros. En ambos casos los escritores acompañan con esta característica de su estilo el tipo de mensaje que transmiten. El estilo enmarañado acompaña perfectamente las tormentosas historias que nos narra Faulker, del mismo modo que la sencillez esquemática de Mañas casi dibuja la simpleza de sus personajes.

A la hora de representar gráficas, es frecuente que nos encontremos en situaciones en las que el mensaje a transmitir exceda la capacidad de una gráfica simple y no se trate de una, dos o tres series de números sino de muchas

más. Ante esta situación también se puede optar por diferentes “estilos narrativos” a la hora de plantear el conjunto completo de información gráfica necesaria (el párrafo gráfico, la infografía).

La opción “historias del Kronen” optaría por representar cada pedazo de información (cada serie) en una gráfica separada, ajustando las escalas de unas y otras y su distribución geométrica en el papel de forma que se evidencien las comparaciones entre ellas que sustancian el mensaje a transmitir. El estilo Faulkner tendería a buscar un tipo de gráfico más sofisticado que soporte bien la inclusión en un mismo espacio de todas las series de datos que se quieren analizar, y además colocadas de una forma que facilite dicha comparación.

Al igual que cada escritor tiene su estilo y no se puede decir que uno sea mejor que otro, del mismo modo ambos enfoques para abordar mensajes gráficos complejos en textos gráficos son perfectamente posibles, y a priori no hay ninguna preferencia entre un estilo y otro. Puede que la dificultad técnica de unos estilos y otros no sea la misma, pero en todo caso la clave para llevar a cabo la tarea con éxito es la misma de siempre: explicitar el mensaje que se quiere transmitir con el párrafo y asegurarse de que el resultado final efectivamente transporta ese mensaje.

Cuando un gráfico es subyugado por formas decorativas o basura de ordenador, cuando las medidas y estructuras de los datos se convierten en Elementos de Diseño, cuando el diseño completo busca más el Estilo Gráfico que la información cuantitativa, entonces al gráfico resultante se le puede llamar “pato” en honor a la tienda en forma de pato “Big Duck”¹. Con esta frase introduce Tufte un “estilo literario” en el que la estética es el único objetivo, no sólo por encima de la finalidad comunicativa sino hasta por encima de los datos cuantitativos que necesariamente subyacen a las representaciones gráficas. La arquitectura moderna, al eliminar los ornamentos, generó edificios que eran íntegramente ornamento². La propia tienda “Big Duck” de Flandes a que se refiere Tufte, o el Guggenheim de Bilbao serían ejemplos paradigmáticos de esta situación. Obviamente no parece que este estilo literario gráfico pueda ser aceptable.

1. TUFTE, página 116 y siguientes. En su libro incluye una fotografía de la tienda, un edificio con forma de pato sentado en la hierba.

2. Robert VENTURI, citado en Tufte, página 117.

En el extremo contrario, Tufte propone maximizar los índices de densidad de datos (tanto en área como en tinta), lo que lleva a un estilo minimalista que no necesariamente es el mejor para una comunicación efectiva.

Entramos en el terreno de la literatura. La poesía tiene un cierto interés comunicativo, pero no necesariamente para un público amplio ni de forma exclusiva el interés estético es el fundamental. Desde luego no es un método de diseminación de mensajes sencillos a públicos masivos. ¿Existe la literatura gráfica? En la medida en que podamos decir de una creación que es bella o no, en cuanto se hace sujeto de juicio estético es susceptible de ser arte. Y las gráficas de forma muy análoga a la literatura. La función principal de las palabras es la comunicación entre personas, pero cuando la dimensión estética es buscada por el autor y apreciada por los lectores se entra en el terreno artístico.

Quizá fuera más adecuado considerar las gráficas sujeto de “artesanía” que de arte. El arte tiene esa potencial inutilidad esencial de la creación *per se* que se le niega a la artesanía. Un collar ha de poderse llevar al cuello, y esa restricción impone suficientes límites como para que la orfebrería sea artesanía y no arte: hay collares más bonitos y más feos, pero han de seguir siendo collares. Del mismo modo las gráficas tienen demasiadas restricciones como para cruzar la frontera de la artesanía.

Literatura, novela y poesía, arte o artesanía. Desde luego estamos ya muy lejos del objetivo de proporcionar guías sencillas para la realización de gráficas adecuadas por parte de no profesionales. Queden estas cuestiones como un punto de partida para la reflexión futura, pero fuera ya de esta publicación.

Síntesis final

Este parece un buen punto para cerrar un camino comenzado unas decenas de páginas atrás. La idea central del libro, enunciada al final de la introducción era: ***para hacer bien una gráfica no hay más que tomar conciencia de su utilidad comunicativa y hacer lo necesario para que esta resulte mínimamente efectiva.***

Hemos dedicado unas cuantas páginas a profundizar más en qué tipo de cosas son “las necesarias” para conseguir el objetivo. Tras ellas un hilo conductor: ese elemento de comunicación que son las gráficas está estructurado como un lenguaje. Y de ese lenguaje presentamos los rudimentos de una gramática. Como ya se ha mencionado antes, esa estructuración lingüística se

utiliza en este texto como una “analogía”, como una asociación de ideas que nos haga más fácil estructurar y recordar las ideas. No se trata, ni mucho menos, de una formalización precisa³.

En primer lugar nos detuvimos en la consideración de la comunicación gráfica, con especial atención a la tipología de formas de construir mensajes gráficos a partir de sus elementos (los esquemas de composición) que nos ofrece la psicología cognitiva. Ahí encontramos el lugar que ocupan las gráficas para representar datos: en el esquema de composición de eje métrico.

A continuación intentamos acotar más el tipo de representaciones que denominamos “gráficas” y que son las que nos interesan en este texto. Tres son las ideas fundamentales. En resumen concluimos que las gráficas (i) son elementos de comunicación gráfica basados en el esquema de composición de eje métrico, (ii) que incorporan información cuantitativa (datos numéricos) y (iii) en los que la relación entre datos y representación es reversible.

En el siguiente capítulo presentamos propiamente la analogía entre los textos gráficos y los léxicos. El resumen de los conceptos relacionados de ambos mundos lo vemos en la tabla 1.

Tabla 1
Analogía entre elementos de las construcciones gráficas y conceptos lingüísticos

<i>Texto gráfico</i>	<i>Texto</i>
Gráfica	Oración
Infografía (grupo de gráficas)	Párrafo
Magnitud representada	Sujeto de la oración
Impresión visual que produce la gráfica	Predicado de la oración
Gráfica con una sola serie de datos	Oración simple
Gráfica con más de una serie de datos	Oración compuesta
Gráfica relacionada con otras en una infografía	Oración subordinada

A partir de ese punto, y basándonos en ejemplos de gráficas con errores, fuimos enumerando una serie de recomendaciones, que agrupamos en tres

3. Para lectores interesados en esta cuestión, y con habilidades matemáticas, una obra de referencia es *The Grammar of Graphics* de Leland WILLKINSON.

capítulos (ortografía, sintaxis y semántica), intentando seguir explotando la analogía con la lingüística. En realidad las ideas no son muchas, como se puede ver en la tabla 2 en la que se han resumido. En total en 12 líneas se recogen todas las ideas, que además no son tan complejas... En realidad, como hemos visto al ir analizándolas de una en una, son cuestiones menos triviales de lo que parecen listadas así, en su esquema más esquelético.

Tabla 2
Resumen de las recomendaciones analizadas a lo largo del libro

Ortografía	El sujeto debe estar definido Debe haber concordancia entre el tipo de gráfica y el tipo de datos Deben incluir ejes explícitos, íntegros y bien referenciados
Sintaxis	La magnitud representada ha de ser la misma para todos los datos Ejes y escalas deben guardar una cierta proporción Adecuación de la densidad de información (respecto a espacio o tinta) Debe incorporarse una mínima composición y evitarse los efectos visuales
Semántica	Con los mismos datos hay distintas representaciones según el mensaje Se puede hacer demagogia con gráficas impecables Pequeños errores formales ayudan a los efectos «demagógicos» La elección de la magnitud a presentar debe ser acorde con el mensaje
Estilos literarios	Y en todo lo anterior queda mucho espacio para el gusto personal

Unas últimas reflexiones casi filosóficas apuran la analogía, si consideramos la capacidad literaria y artística del lenguaje de las gráficas. En este punto la analogía ya ha quedado agotada, al menos como medio para proporcionar pistas sencillas que ayuden a representar datos a todos los que tienen que hacerlo sin ser estadísticos profesionales.

Anexo 1

Breve historia de las gráficas

No es infrecuente hacer una analogía entre lo costoso que ha resultado un descubrimiento para la humanidad, y lo tardío y lo difícil que nos pueda resultar a cada uno de nosotros como individuos comprenderlo.

En esta idea, la lógica de Aristóteles o la geometría de Euclides resultan naturales, de comprensión prácticamente obvia. El acervo cultural de antes de Jesucristo es algo sencillo, que se aprende en la escuela. Ya Newton y Descartes no resultan tan evidentes en todo. Eso de que la fuerza que hace caer la manzana sea la misma que mantiene a la luna en su sitio ya no es tan fácil de roer. Y la geometría no euclídea, o la relatividad son directamente incomprensibles.

Del mismo modo que el desarrollo embrionario (la ontogénesis) repite el desarrollo evolutivo (la filogénesis)¹, el desarrollo cognitivo individual rememora de alguna manera el desarrollo histórico de la humanidad. Si esto es

1. “La ley de la repetición de la filogénesis en la ontogénesis no presenta dudas” dice el Dr. C. Enrique González Suárez en http://bvs.sld.cu/revistas/aci/vol12_2_04/aci0904.htm aunque la cosa no está tan clara cuando se entra en detalles.

así, sería interesante saber de cuando datan las representaciones gráficas que nos han ocupado en este texto. ¿Hacia Pitágoras gráficas de barras?

Si Pitágoras o alguno de sus coetáneos representaron datos gráficamente no ha quedado ningún rastro de ello. La historia que concluirá en las gráficas comienza con los mapas, representaciones a escala del mundo físico². Los agrimensores egipcios utilizaban coordenadas físicas tres mil años antes de Jesucristo para localizar puntos en la tierra³. Latitud y longitud como abscisas y ordenadas, y en las intersecciones accidentes geográficos. Un procedimiento para confeccionar mapas: representaciones bidimensionales a escala de realidades casi bidimensionales (con alturas mucho menores que latitudes y longitudes). La relación entre el objeto representado y su representación es muy próxima, el proceso mental no es excesivamente complicado.

La evolución hacia las gráficas requiere un proceso de abstracción a partir de los mapas. Estos se fueron sofisticando y mejorando, pero sin salir del esquema de representación sencillo (latitud y longitud para indicar accidentes geográficos). Está documentada⁴ una primera gráfica en la que se representan las inclinaciones de las órbitas planetarias en función del tiempo en algún momento alrededor del siglo X. A pesar de los 5.000 años transcurridos desde los primeros mapas, esta representación es un caso aislado que demuestra que, aunque una mente sí concibió la abstracción que supone sustituir la longitud por el tiempo, el proceso estaba muy lejos de poderse considerar generalizado.

Uno de los primeros mapas de datos es la carta de Edmond Halley, de 1686, en la que representa datos meteorológicos en vez de geográficos⁵ (direcciones de vientos y monzones). Las representaciones comienzan a incor-

2. Hay tres fuentes principales como documentación de este anexo citadas a continuación. La mayor parte de las referencias que se incluyen proceden a su vez de alguna de estas tres:

- “Presentation graphics” L. Wilkinson, *International Encyclopedia of the Social & Behavioral Sciences*, consultable en: <http://www.spss.com/research/wilkinson/Publications/iesbs.pdf>
- “Visual Revelations” H. Wainer, *Chance* VOL. 17, NO. 2, 2004, consultable en <http://www.math.yorku.ca/SCS/Papers/NobodysPerfect.pdf>
- Edward R. Tufte, *The Visual Display of Quantitative Information*, Graphics Press, 1983.
- 3. BENIGER, J.R. and ROBYN, D.L. (1978). “Quantitative graphics in statistics: A brief history”. *The American Statistician*, 32, 1-11, citado por Willkinson.
- 4. FUNKHOUSER, H.G. & WALKER, H.M. (1935). “Playfair and his charts”, *Economic History*, 3, 103-109, citado por Wilkinson y por Tufte.
- 5. Tufte, p. 23, cita a N.J.W. THROWER, “Edmond Halley as a Thematic Geo-Cartographer”, *Annals of the Association of American Geographers*, 59 (December, 1969), 625-676

porar abstracciones a la representación fidedigna del espacio físico que se consigue con los mapas milenarios ya.

Las coordenadas cartesianas toman su nombre de su inventor, Descartes (1596-1650), quien las introdujo para localizar puntos en espacios abstractos, concretamente para representar las ecuaciones que aparecían en su geometría analítica. Pero a pesar de esta creación explícita, su uso no se extiende. Son muy escasas las representaciones de datos empleadas a lo largo de los siglos XVII y XVIII⁶. Los científicos de la Ilustración tendían a considerar los gráficos como diletantismo, y cuando los utilizaban era para representar diagramas geométricos o para dibujar funciones algebraicas, nunca para representar datos⁷. Por su parte, los científicos sociales de este periodo se centraron en recoger tablas estadísticas del Estado que se pudieran leer detalladamente sin el recurso de la aproximación gráfica⁸.

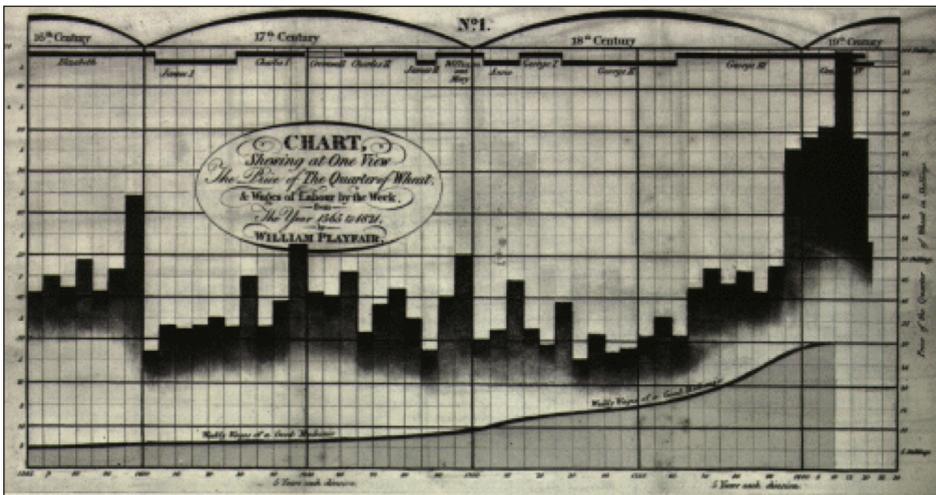


Figura 73. Una de las primeras gráficas de Playfair, en la que se representan tres series paralelas: precios, salarios, y los reinados de reyes y de reinas británicos (tomada de <http://www.wmich.edu/ssc/about.html>)

6. Tufte incluye un ejemplo de una serie temporal de J.H. Lambert publicada 1779 como una de las primeras.

7. TILLING, L. (1975). Early experimental graphics. *The British Journal for the History of Science*, 8, 193-213 y Funkhouser 1937 (nota 4) según Wilkinson

8. BENIGER, J.R. and ROBYN, D.L. (1978). Quantitative graphics in statistics: A brief history. *The American Statistician*, 32, 1-11

Se han documentado algunas excepciones a esta tendencia⁹, sin embargo hay un acuerdo general en que la figura clave, auténticamente revolucionaria es William Playfair (1759-1823). Economista político escocés, Playfair publicó en 1786 el “Atlas político y comercial” conteniendo 44 gráficas: 43 series temporales y un diagrama de barras. Esta obra se considera como el auténtico arranque de las gráficas estadísticas.

Representando los datos del estado¹⁰ Playfair, junto con J.H. Lambert (1728-1777) inventaron las gráficas de tartas, los gráficos de círculos y los diagramas de barras categóricos. En resumen todas las representaciones básicas de datos se deben a estos dos autores.

El astrónomo británico John Frederick William Herschel (1792-1871) es el candidato más probable a inventor del diagrama de dispersión¹¹ (*scatterplot*). El uso de Herschel del diagrama de la dispersión tiene lugar en su «investigación de las órbitas de las estrellas dobles giratorias» que leyó en la Real Sociedad Astronómica en enero el 13 de 1832 y publicó un año más tarde.

Las gráficas de líneas o barras, series temporales y tartas, parecen ser las formas más frecuentes de presentación de información cuantitativa en los medios populares, mientras que los diagramas de dispersión ocupan un lugar preeminente en las páginas de las revistas científicas y técnicas. A partir de esta observación considera Wainer¹² que no es extraño que los inventores de unas y otra fueran diferentes, y lógico además que el primero fuera un estadístico y el segundo un astrónomo.

Las gráficas de dispersión se han calificado como “el invento más versátil y polimórfico de toda la historia de los gráficos estadísticos”¹³. También Tufte dice que: “...El gráfico relacional – en su forma más simple la gráfica de dispersión y sus variantes- es el más grande de todos los diseño gráficos. Relaciona al menos dos variables animando, incluso implorando al observador a que verifique la posible relación causal entre las variables representadas”¹⁴. En

9. Gráficas de Wren, Halley, Huygens, Plot, Lister o Priestley, documentadas en Beniger and Robyn (1978) (nota 8) así como en: WAINER, H. y VELLEMAN, P.F. (2001). Statistical graphs: Mapping the pathways of science. *The Annual Review of Psychology*, 52

10. Origen de la palabra “estadística”.

11. FRIENDLY, M. y DENIS, D. (2004). Who invented the scatterplot? Manuscript submitted for publication. citado de esta forma por Wainer

12. WAINER 2004 (nota 2).

13. Wainer 2004 (nota 2), citando allí el manuscrito sujeto a revisión de Friendly & Denis, 2004.

14. Tufte página 47.

efecto, el predicado fundamental que transmiten este tipo de gráficas es la correlación y en su caso la causalidad, conceptos ambos esenciales en la actividad científica. No es de extrañar, por tanto, que sea en las revistas científicas donde este tipo de representación resulte tan común. En línea con la admiración que le produce, Tufte define un nuevo índice, en este caso para medir el grado de “sofisticación” de una publicación, como la relación entre el número de gráficas relacionales respecto del total de gráficas encontradas en una muestra aleatoria de la publicación. Los resultados de la aplicación a grandes rasgos de este índice no son muy sorprendentes: en las publicaciones científicas es muy alto, y en los periódicos muy bajo. Más que una medida de la sofisticación podemos considerarlo como una medida del tipo de predicados que transmita cada publicación, muy centrados en la búsqueda de la causalidad en las revistas científicas y mucho más descriptivos en la prensa cotidiana.

Volviendo al desarrollo histórico, a partir de los seminales trabajos de Playfair y Herschel comenzaron a utilizarse progresivamente las representaciones de datos, si bien no puede considerarse esto una técnica habitual hasta un siglo después de su invención. Paulatinamente se fueron añadiendo nuevos tipos de representaciones, hoy comunes, a lo largo de los siglos XIX y XX. Un resumen de algunos de ellos¹⁵ incluiría la utilización por Guerry en 1833 de barras de densidad para mostrar datos de crímenes agrupados por grupos de edad. Pearson adoptó esta forma de representación para representar una distribución de frecuencias y lo denominó “histograma”. Minard desarrolló un conjunto de mapas esquemáticos superponiendo datos estadísticos a los datos geográficos. La pirámide de población (un doble histograma de población por edad y sexo) lo introdujo F.A. Walker, superintendente del censo de los Estados Unidos en 1874.

Una parte importante del pensamiento sobre gráficos estadísticos del siglo XX ha estado dedicado a la cuestión de cómo unos gráficos poco profesionales pueden engañar a observadores poco entrenados, dejando de lado cuestiones importantes como el uso de gráficos en el análisis serio de datos. No muy lejos de la idea de la ilustración del diletantismo, se asume que las gráficas sirven para mostrar lo obvio al ignorante¹⁶. Fue a finales de los años

15. Siguiendo a BENIGER and ROBYN (1978) (nota 8) y STIGLER, S. (1983). *The History of Statistics*. Cambridge, MA: Harvard University Press, citados por Wilkinson.

16. TUFTE, p. 53.

60 cuando las gráficas estadísticas adquirieron un estatus de respetabilidad científica, gracias al trabajo de John Tukey¹⁷



Figura 74. John W. Tukey (1915-2000) tomado de: <http://stat.bell-labs.com/who/tukey/>

Fecundo matemático y estadístico, Tukey deja tras de sí una extensa obra sobre un conjunto de temas, probablemente es más conocido sea la Transformada Rápida de Fourier (FFT). En la historia que nos interesa aquí destaca el Análisis Exploratorio de Datos¹⁸ (Exploratory Data Analysis, EDA) desarrollado en su libro del mismo título de 1977. EDA es una filosofía básicamente gráfica de exploración de datos estadísticos. Por ello muchas veces se la confunde con la estadística gráfica, aunque EDA va más allá. Lo interesante, a nuestros efectos, de EDA, es la potencia que añaden los gráficos, la representación visual, a las herramientas estadísticas. Los gráficos proporcionan a los resultados estadísticos una gran ayuda para la comprensión de su significado. Se convierten en una herramienta potente y formalizada para “revelar los datos”¹⁹. Un buen final para la historia que comenzaron los agrimensores egipcios de hace milenios.

17. Ver, por ejemplo, <http://stat.bell-labs.com/who/tukey/memo/techttools.html>
<http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Tukey.html>
<http://www.infovis.net> número 6 de la revista

18. Sobre tema hay una sección específica en el apartado del texto sobre referencias.

19. Ver anexo 2 de esta publicación.

Podemos pues concluir que Pitágoras no hacía gráficos de barras. Como hemos visto, esa representación no fue inventada hasta el último cuarto del siglo XVIII. Por tanto, si aceptamos que el desarrollo cognitivo individual rememora de alguna manera el desarrollo histórico de la humanidad (como la ontogénesis repite la filogénesis), no es raro que haya tantas equivocaciones gráficas. Las representaciones gráficas no son desarrollos históricamente triviales.

Ya en el siglo XXI el estudio de las gráficas ha tomado un renovado auge como una de las partes de un concepto más amplio: la visualización de la información. Este campo multidisciplinar analiza toda forma de transmitir información de forma eficiente explotando el canal visual. Desde las instrucciones de montaje de los muebles de Ikea, a los horarios de trenes, pasando por la “usabilidad” de páginas web o las gráficas cuantitativas, todo es “Visualización”²⁰.

20. También sobre visualización hay una sección específica en el apartado del texto sobre referencias.

Anexo 2

Revelar los datos (Autocomunicación)

Estamos enfatizando la consideración de las gráficas como elementos de comunicación. Como tales serán necesariamente portadoras de un mensaje, y si este mensaje es explícito, mejor. Un mensaje es algo más que un conjunto de datos. Es necesario observar el conjunto de datos a representar hasta ser capaz de explicitar la intención de realizar con ellos una gráfica.

Hay muchas situaciones en las que la intención es evidente desde el primer momento, y en las que ese proceso resultará trivial: evolución de las ventas del último trimestre, proporción de escaños tras unas elecciones...

En otros casos, en cambio, más que una observación se requiere una auténtica investigación que suponga incluso la representación de muchas gráficas hasta dar con la clave que se busca. Estas gráficas que podríamos denominar “**gráficas de trabajo**” tienen por misión descubrir relaciones no evidentes en los datos de partida. Esta capacidad de manifestar relaciones que se desconocían a priori es una de las características más potentes de las representaciones gráficas.

A esa capacidad de ayudar a encontrar relaciones no evidentes en los datos de partida se la denomina a menudo en la literatura “revelar los da-

tos”¹. Este proceso de “revelado” es análogo al método científico clásico: se parte de la observación directa para establecer una hipótesis, y se trabaja para intentar corroborar o refutar dicha hipótesis.

Desde el punto de vista de la teoría de la comunicación, en la elaboración de las gráficas de trabajo nos encontramos con un proceso de comunicación atípico, en el que el emisor y el receptor son la misma persona, mientras que el mensaje no es explícito. Este proceso de investigación basado en gráficos de trabajo suele producirse de forma iterativa: partiendo de una representación estándar, se realiza una primera observación, que lleva a modificar la representación intentando enfatizar lo observado y así sucesivamente hasta que se dispone de una representación en la que se hace patente un hecho que no se conocía al comienzo del proceso. Al finalizar, el hecho descubierto es el mensaje explícito a transmitir y la gráfica que lo hacía evidente es una representación adecuada para el mismo. Ya se está en condiciones de utilizar la gráfica para transmitir el mensaje a receptores diferentes del propio autor (i.e. publicar). Es muy posible que a lo largo de este proceso de investigación no sólo haya sido necesario cambiar la forma de representar gráficamente los datos, sino que se hayan tenido que modificar los propios datos (promedios, resta de línea base, cálculo de diferencias, derivadas, transformada de Fourier, etc.).

En el camino hacia un mensaje explícito y una representación adecuada al mismo, puede ocurrir que el conjunto de datos de partida no sea el más adecuado y que no sea fácil, quizá hasta imposible, sostener con ellos el mensaje. En estos casos hay que rehacer el conjunto de datos original. No se trata ahora simplemente de transformar los datos de que partíamos sino de buscar otros, actividad que queda fuera del objetivo que nos proponemos aquí. Vamos a ver en este anexo un ejemplo concreto de este proceso de autocomunicación.

Ejemplo de gráficas de trabajo (población en España)

Partimos en este ejemplo de una tabla de datos que, con una parte de la información de la figura 19 en el capítulo 2, se completa con el conjunto de datos de superficie dando lugar al mostrado en la figura 75.

1. Dürsteler en Infovis.net y Tufte “show the data”...

Figura 75
Tabla de datos en la que a la población,
se añade la superficie de las comunidades autónomas²

<i>Comunidad Autónoma</i>	<i>Pobl. 2001</i>	<i>Superficie geográfica total(ha)</i>
Andalucía	7.357.558	8.735.919
Aragón	1.204.215	4.766.838
Asturias	1.062.998	1.056.428
Balears (Illes)	841.669	494.202
Canarias	1.694.477	748.023
Cantabria	535.131	532.134
Castilla y León	2.456.474	9.401.032
Castilla-La Mancha	1.760.516	7.922.531
Cataluña	6.343.110	3.194.728
Comunidad Valenciana	4.162.776	2.325.912
Extremadura	1.058.503	4.160.166
Galicia	2.695.880	2.947.667
Madrid	5.423.384	802.792
Murcia	1.197.646	1.131.738
Navarra	555.829	1.039.135
País Vasco	2.082.587	725.072
Rioja (La)	276.702	503.388

A partir de esta tabla de datos se pretende realizar una representación gráfica, pero supongamos que, al comienzo, no es evidente lo que se pretende comunicar con ella. Una representación estándar de estructuras de datos bi-dimensionales es el gráfico cartesiano de dispersión. Así, una primera gráfica a partir de estos datos sería la mostrada en la siguiente figura.

2. La población está tomada de la anterior figura 19. Los datos de superficie están tomados del Ministerio de Agricultura, Pesca y Alimentación, anuario del año 1999. http://www.mapya.es/info/pags/anuar_99/cap03_distribucion/distribucion_suelo_5.htm

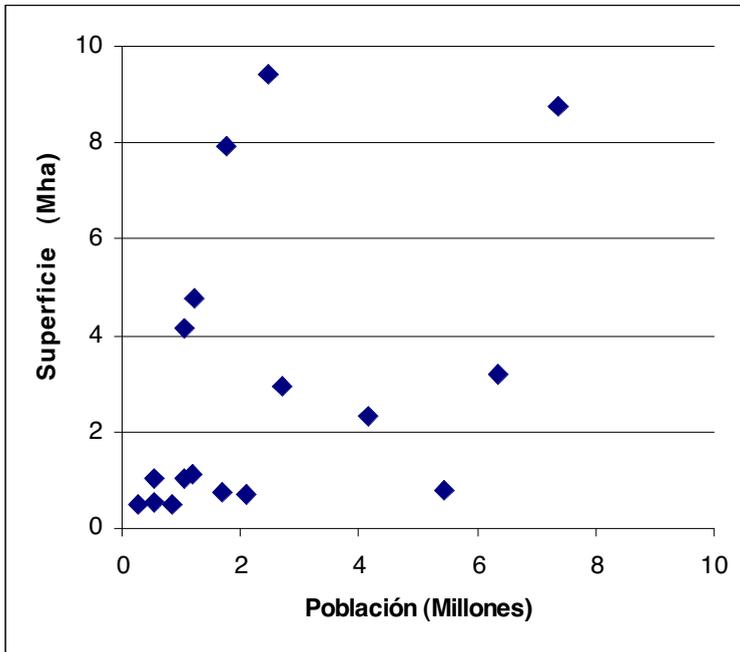


Figura 76. Una primera representación gráfica de los datos de la figura 1.

En la gráfica de la figura 76 la bisectriz del cuadrante une puntos con la misma densidad de población, cuanto más lejos del origen con más superficie y más población, pero en la misma proporción. De hecho eso ocurre para todas las líneas rectas que pasan por el origen de coordenadas. Así, el proceso de trabajo puede continuar añadiéndose unas líneas de referencia que indiquen distintas densidades de población (figura 77).

En esta nueva representación, que en realidad es la misma que la anterior añadiendo unos elementos de referencia, se aprecia que los puntos se agrupan con bastante nitidez siguiendo las líneas de igual densidad de población, sugiriendo que hay unos conjuntos de comunidades autónomas que poseen una densidad de población muy parecida.

Para aclarar más esta idea, y para saber qué comunidades son las que pertenecen a cada uno de los grupos, una nueva representación consistiría en añadir rótulos a los puntos en los que se explicitara el nombre de la comuni-

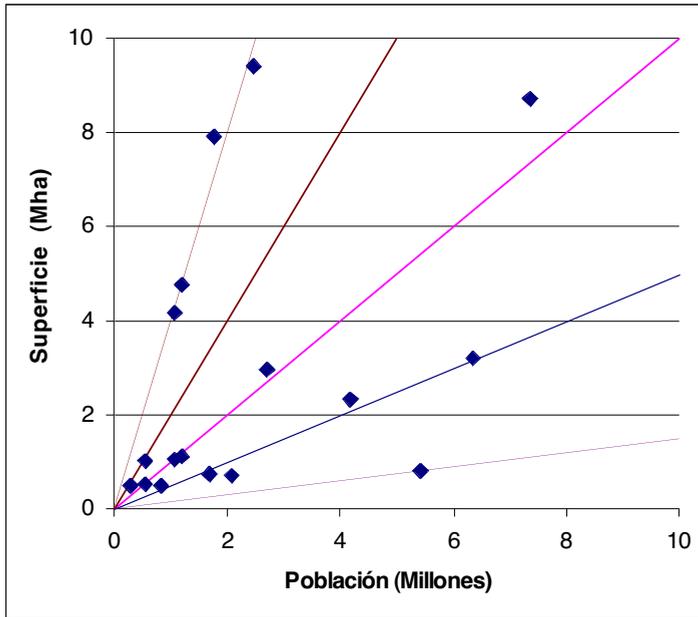


Figura 77. Representación de la superficie frente a la población de las comunidades autónomas españolas. Cada punto corresponde a una comunidad. Las líneas indican las posiciones con la misma densidad de población, en concreto y de menor a mayor (de más vertical a más horizontal): 0,25 0,5 1 2 y 6,66 habitantes por hectárea.

dad autónoma que representan. Dado que cuando se hace eso en una sola gráfica con todos los puntos el resultado queda muy emborronado, se puede dividir la gráfica en dos. El resultado de este proceso se muestra en la figura 78. Esta figura aporta mucha información ya que permite apreciar para cada comunidad tanto el valor absoluto de su población como el de su superficie, la posición relativa de ambas entre las demás y la densidad de población (tanto en valor absoluto como relativa al resto de comunidades).

El proceso de trabajo se puede dar por concluido. Esta gráfica manifiesta (“revela”) todas las relaciones entre los datos originales. A partir de esa información ya puesta de manifiesto es necesario seleccionar el mensaje a transmitir y con él la mejor representación gráfica.

Precisamente por la riqueza informativa que aportan este tipo de gráficas (figura 78), no transmiten un mensaje sencillo, y requieren de bastante atención del lector para su comprensión. Puede ocurrir que lo que se

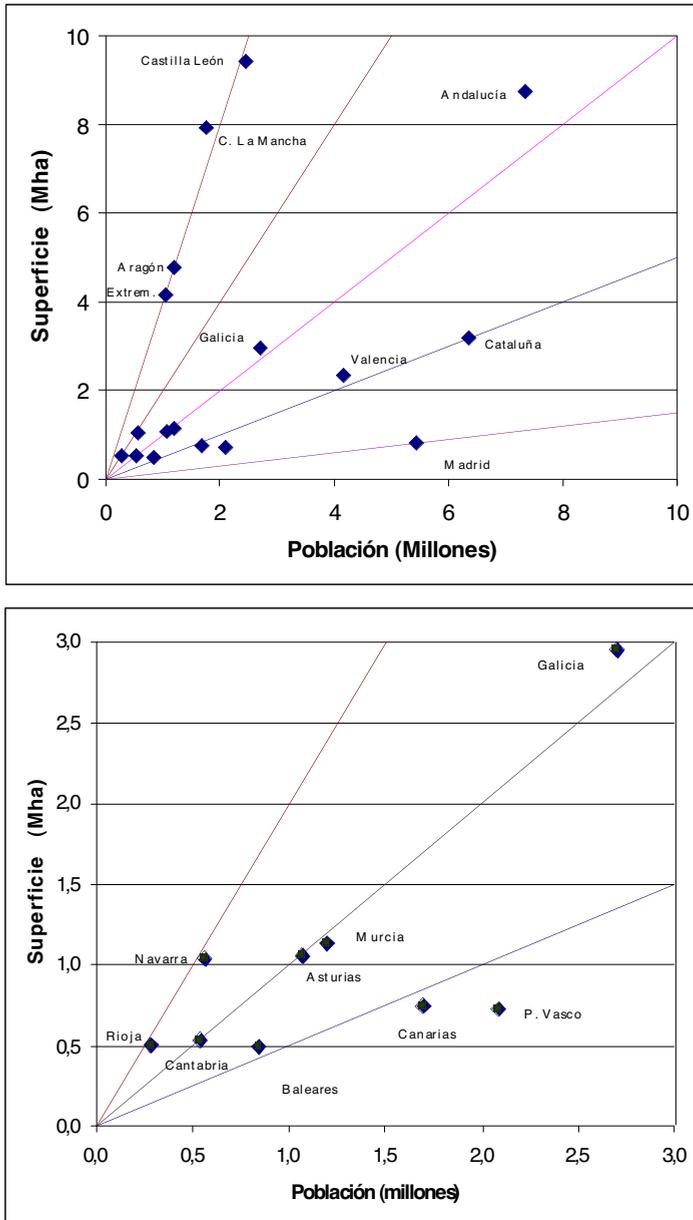


Figura 78. La misma representación que en la figura 3 añadiendo rótulos a los puntos que los identifiquen. Se incluye además una segunda representación de detalle para identificar las comunidades de menor tamaño (más próximas al origen).

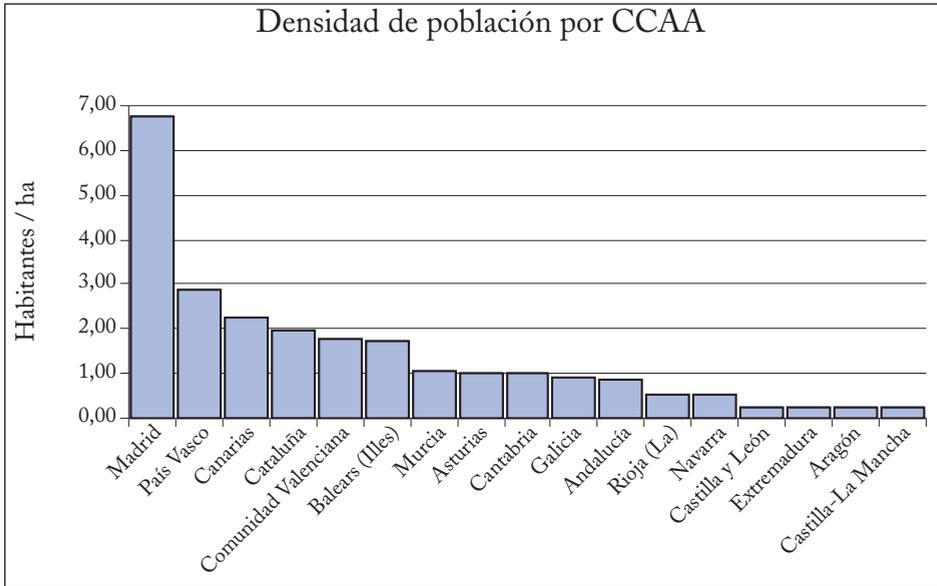


Figura 79. Gráfico de barras de la densidad de población de las comunidades autónomas españolas.

pretenda en un momento dado sea algo mucho más simple, como por ejemplo mostrar la densidad de población para cada caso. Entonces será mucho más adecuado calcular la densidad y representarla como gráfico de barras. Ese cálculo de la densidad supone una transformación (muy sencilla en este caso) de los datos de partida para obtener otros nuevos. La representación como gráfica de barras de los datos transformados se muestra como figura 79.

Al ser los mismos datos, son los mismos predicados los que se pueden establecer, sólo que enfatizando más algunos u obviando otros. En esta representación se ve claro que las comunidades se agrupan en valores de densidad muy próximos. Aquí aparecen como barras de igual altura, y lo mismo se aprecia en la gráfica de la figura 78 como agrupamientos sobre las líneas de referencia. Allí se identificaban los tamaños absolutos de las comunidades y aquí se pierde esa información; a cambio aquí se aprecia numéricamente la densidad de cada comunidad, lo que en la figura 78 no se podía apreciar fácilmente de un vistazo. Estas observaciones dejan claro que ninguna de las dos representaciones es mejor que la otra de una forma ab-

solita, sino que cada una será más adecuada en un proceso de comunicación concreto³.

El ejemplo ilustra el proceso de realización de gráficas de trabajo como elementos en el proceso de explicitar un mensaje a partir de unos datos dados. También se muestra cómo es posible, y en ocasiones conveniente, realizar transformaciones de los datos originales para buscar las representaciones más adecuadas al mensaje que se quiere transmitir.

3. Este mismo ejemplo se utiliza en el capítulo 2, al hilo de la precisión sobre lo que es una gráfica. En la figura 21 de aquel capítulo, se presenta una gráfica de tarta para mostrar la distribución de población entre comunidades; un predicado distinto que encuentra su mejor vehículo en otra representación.

Anexo 3

Cuándo no hacer una gráfica

A lo largo del texto nos hemos centrado fundamentalmente en cómo hacer bien una gráfica, dando por supuesto que tenemos claro el paso previo: queremos hacer una gráfica. Pero este primer paso también requiere un cierto análisis. ¿En que situaciones o bajo qué condiciones una gráfica es lo más adecuado?

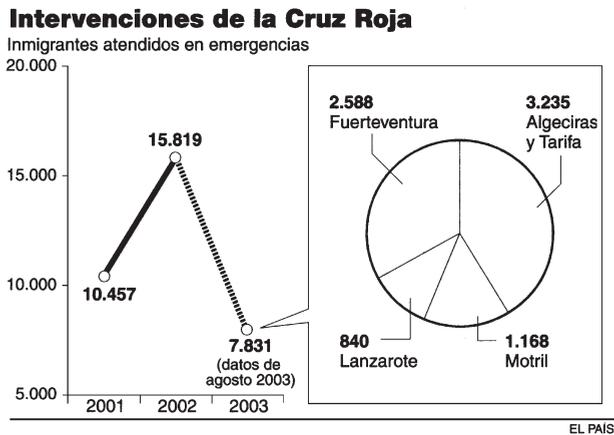


Figura 80. Serie temporal que incluye dos puntos, y un tercero que en realidad no es comparable con los otros dos.

La gráfica que se presenta aquí como figura 80 ya la habíamos analizado al referirnos a la sintaxis¹. Entonces analizamos el hecho de que el tercer punto de la gráfica no es comparable a los otros dos, ahora lo que nos interesa es la oportunidad de presentar estos datos en forma de gráfica. En rarísimas ocasiones una gráfica con un solo punto tiene sentido, y con dos prácticamente igual. No hay ningún predicado gráfico que transmitir. Es por tanto un esfuerzo vano intentar mostrar 10.497 y 15.819 en una gráfica. Este es un caso claro en el que la gráfica no es la mejor opción.

Consideremos otro ejemplo menos evidente: “En un hospital se está realizando un estudio sobre la influencia de la herencia en las enfermedades cardiacas. Para ello se toma un grupo de enfermos y se investiga qué porcentaje tenía antecedentes familiares de ese tipo de enfermedades, resultando el 72% frente a un 28% de enfermos que no los tenían. Para completar el estudio se repite un recuento análogo en un grupo de personas sanas, resultando que el 66% no tenían antecedentes y el 34% sí”. Los mismos resultados del párrafo anterior se pueden presentar en forma de tabla (de dos formas distintas), como se muestra en las figuras 81 y 82.

<i>% de personas que</i>	<i>Tienen la enfermedad</i>	
	SÍ	NO
Tienen Antecedentes	72	28
SÍ	72	34
NO	28	66

Figura 81. Tabla en la que se recoge la misma información expuesta en forma de texto en el párrafo anterior

<i>% con antecedentes en dos grupos: sanos y enfermos</i>		
<i>Enfermedad</i>	<i>Antecedentes</i>	
SÍ	SÍ	72
SÍ	NO	28
NO	SÍ	34
NO	NO	66

Figura 82. Otra tabla diferente en la que se vuelve a recoger la misma información que en la figura anterior y en el párrafo correspondiente.

1. Es la misma que la figura 54 del capítulo 5 (Sintaxis).

Hay una información que no resulta difícil perder según cómo se presenten los datos: los cuatro números que resultan del experimento suman 100 (como no puede ser de otra manera dado que son porcentajes), tomando las parejas de una determinada forma. En la disposición matricial de la figura 81 se puede apreciar cómo la suma de cada columna es la que da 100, pero no la suma de cada fila (que dan 106 y 94 respectivamente). Esto es coherente con el enunciado de texto, en el que se decía que se había analizado el tanto por cien de personas con antecedentes para un grupo de personas enfermas y otro grupo de personas sanas. Este hecho es el que se pretende enfatizar en la segunda versión de la tabla (figura 82), lo que a cambio lleva a una pérdida de estructura de los datos. En la estructura matricial se ve cómo los números de la “diagonal principal” (los sí, sí y no, no) son mucho mayores que los de los términos cruzados. Esa estructura hace patente la correlación entre las dos variables que se analizan (enfermedad y antecedentes), y es familiar para las personas que trabajan con este tipo de cuestiones.

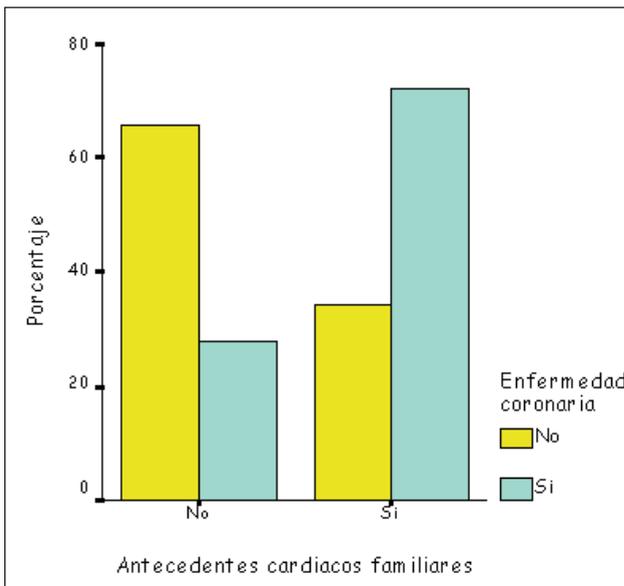


Figura 83. Representación gráfica de los datos que se han presentado en las tablas anteriores tal y como se encontró en la literatura². Las columnas que suman 100 no son las adyacentes, sino las del mismo color.

2. <http://www.fisterra.com/material/investiga/graficos/images/Image119.gif>

La información que se pretende transmitir la hemos presentado de tres formas, como texto (6 líneas, 77 palabras), como tabla de dos entradas en formato matricial y como tabla de dos entradas en formato lineal. ¿Cómo podríamos representar esos datos de forma gráfica? La representación que encontré en la literatura (y fuente original de los datos reelaborados en las tablas anteriores) es la que se muestra en la figura 83.

El resultado de esta representación no parece muy bueno. Es verdad que los datos son pocos (sólo 4) y además resultan liosos con todas las combinaciones de dos variables que pueden ser sí o no; pero precisamente porque los datos de partida pueden inducir a confusión hay que buscar una representación que la dificulte.

Formalmente no hay nada que objetar a la gráfica de la figura 83, y semánticamente tampoco: no transmite ningún mensaje que no esté soportado por los datos... En realidad no transmite ningún mensaje en absoluto³, y ese es el problema. La contemplación de los datos de este ejemplo sugiere una fuerte correlación entre el padecimiento de la enfermedad y la existencia de antecedentes familiares, pero ese predicado, esa fuerte correlación, no se desprende mejor de la gráfica que de la tabla (en cualquiera de sus dos formatos), ni en la tabla mejor que en el párrafo.

Una nueva representación de los datos del ejemplo se muestra en la figura 84, como dos gráficas de tarta en esta ocasión. Esta representación deja mucho más claro cuál es la categoría en la que se han calculado los porcentajes (y por tanto cuadra la suma), pero el predicado principal, la correlación entre las dos categorías no se desprende mejor que en los casos anteriores.

En este ejemplo hemos visto cinco formas distintas de presentar los mismos datos: el texto, dos tablas y dos gráficas. Se comprueba cómo el predicado esencial que nos gustaría extraer de ellos (la correlación entre enfermedad y antecedentes familiares) no se hace más claro en unas que en otras. Así pues nos encontramos con 5 formas de expresión, equivalentes en potencia comunicativa, pero no en otras características como la dificultad técnica de realización o el espacio de papel ocupado (ver tabla de la figura 85)

3. Al menos al lector no especialista en epidemiología. Ignoro si es una representación muy habitual en un campo y el hábito de sus especialistas hace que en él adquiera un significado muy obvio.

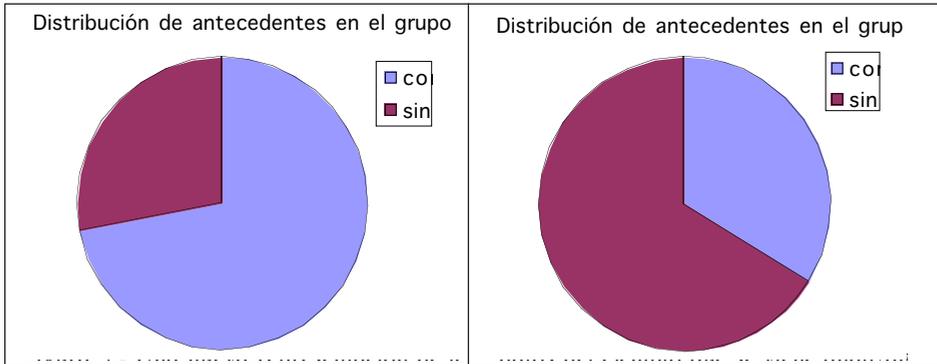


Figura 84. Una nueva representación de los datos del ejemplo que se viene analizando. ¿Se aprecia mejor aquí la idea de correlación que se desprende de los datos?

<i>Item</i>	<i>Espacio (ua)</i>	<i>Dificultad(ua)</i>
Texto	3	1
Tabla matricial	3,5	10
Tabla lineal	4	10
Gráfica barras	11	15
Gráficas tarta	8	20

Figura 85. Tabla en la que se resumen el espacio ocupado y la dificultad de realización de cada uno de los 5 medios comunicativos de los 4 datos del ejemplo. Las unidades arbitrarias en que se miden ambas magnitudes se describen en el texto.

Para estimar el espacio ocupado se ha medido, en centímetros, la longitud de folio (considerando que siempre se ocupa la anchura completa) empleada en cada caso. Las unidades arbitrarias en las que se estima la dificultad son los minutos que le ha costado al autor confeccionar cada representación, referidos al caso del texto, que fue el más rápido.

Esta misma cuestión la trata Tufte con detalle en el capítulo 9 de su libro. Del estudio de los ejemplos que va desgranando se deduce que a medida que aumenta el número de datos a presentar el texto va siendo menos conveniente y la gráfica más, quedando la tabla para tamaños intermedios. En todo caso las fronteras no son muy definidas y a cada conjunto de datos habrá que

buscarle la mejor solución. No se encuentra un guía operativa más concreta allí⁴.

Por su parte Dürsteler, en su magnífica página web sobre visualización⁵, a partir de la idea de Tufte sí establece unas fronteras numéricas en cuanto a la cantidad de datos, aunque sean unas fronteras no perfectamente definidas. El enunciado de esas fronteras lo plantea además de las tres formas analizadas: como texto (figura 86), como tabla (figura 87) o como gráfica (figura 88).

“Conviene utilizar frases para relaciones entre 2 ó 3 datos, tablas entre 3 y alrededor de 20 datos y gráficos a partir de 3 relaciones, y tanto más cuanto mayor sea el número de datos y relaciones entre ellos”

Figura 86. Texto en el que se establecen las fronteras en cuanto a número de datos para cada tipo de presentación de información numérica. Aunque no es habitual considerar un párrafo como figura, se hace aquí por analogía con las figuras 8 y 9. Tomado de Infovis.net.

	OK
Frases	Relaciones entre 2 ó 3 datos.
Tablas	Relaciones entre 3 y 20 datos (aprox)
Gráficos	Desde 2 en adelante. Especialmente útiles para grandes cantidades de datos.

Figura 87. Tabla que presenta la misma información que la figura 7 sólo que estructurada como tabla en vez de como texto. Tomado de Infovis.net.

4. Bueno, salvo que la densidad de información en los gráficos de tarta es tan baja que “should never be used”, idea apoyada además por Bertin (*Graphics and Graphic Information Processing*, 1981), que dice que las gráficas de múltiples tartas son completamente inútiles. A pesar de eso me he atrevido a dibujar la figura 84. El maximalismo teórico de Bertin o el estético de Tufte se relajan cuando se centra la atención en el carácter comunicativo de las gráficas. Yo sólo “prohíbo” las tartas en perspectiva (ver cap 3).

5. www.infovis.net, artículo número 108.

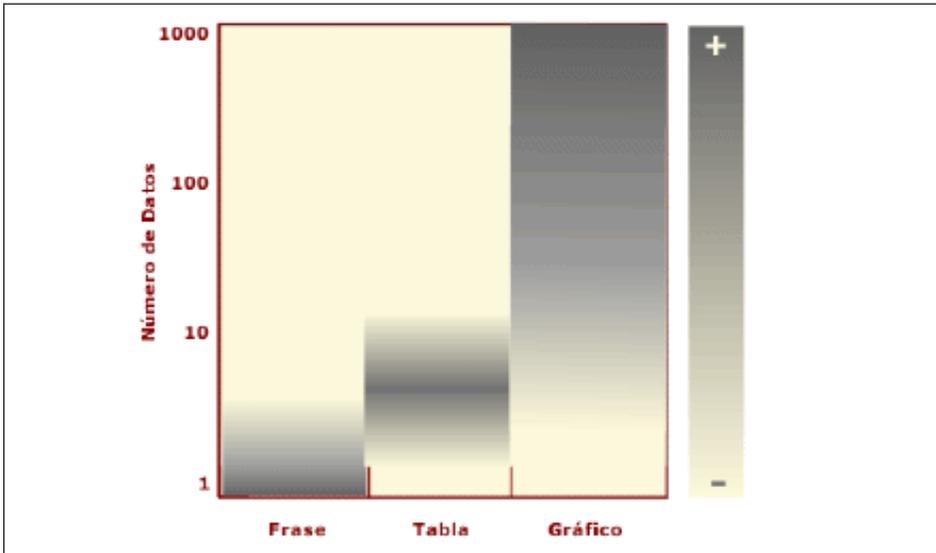


Figura 88. Representación gráfica de la misma información que en las figuras 86 y 87. “Las gradaciones de color y los solapamientos nos indican que ni 3 ni 20 son números definitivos, depende de la aplicación. El gráfico añade en este caso una expresividad extra a los números exactos de la tabla y reduce la verborrea necesaria para explicar este mensaje. Juzguen por sí mismos cuál de las tres es más apropiada para este caso.” Tomado de Infovis.net , al igual que la gráfica.

De los ejemplos analizados podemos extraer algunas conclusiones. En primer lugar que la gráfica no es siempre la mejor opción para comunicar información cuantitativa, se dispone también de texto y tablas, y todas son útiles en ocasiones. En segundo lugar la tendencia cuantitativa: para muy pocos datos el texto será la opción y para muchos las gráficas.

La idea principal que presentamos en este texto es que las gráficas son un medio de comunicación, y es en esta idea en la que podemos fundamentar la decisión final sobre la adecuación de una gráfica en una situación dada. Por una parte deberá existir un mensaje a transmitir, un predicado para la oración que es la gráfica. En ausencia de un predicado claro, una tabla mostrará los datos de una forma más directa, más adecuada. Obviamente ese mensaje a transmitir debe ser pertinente, debe estar apoyado por los datos de los que se dispone. En estas condiciones, si es posible encontrar una representación que de verdad muestre el mensaje, si se puede construir una oración gráfica con ese predicado, el resultado valdrá más que mil palabras o que cien tablas.

Referencias

Se puede comprar un buen montón de libros sobre el tema de las representaciones gráficas. Algunos son accesibles a lectores sin una formación específica, aunque la mayoría están orientados a matemáticos o estadísticos profesionales. Pero además, hoy día la cantidad de información disponible en internet es impresionante, tanto que muchas veces lo difícil es saber por dónde empezar. A continuación se incluyen algunos libros, pero sobre todo muchas direcciones de internet con las que poder profundizar.

Libros:

Algunos libros, por seleccionar alguno:

The Visual Display of Quantitative Information, Edward R. Tufte, Graphics Press, 1983 (ISBN 0-9613921-0-X) 197 pp.

Visualización de Información, Juan C. Dürstelner, Gestion 2000, 1ª edición, 2002 (ISBN 84-8088-836-9), 152 pp.

Visual Language Theory, K. Marriot y B. Meyer (eds.), Springer-Verlag, 1998 (ISBN 0-387-98367-8), 381 pp.

The Grammar of Graphics, Leland Wilkinson, Springer-Verlag, 1999 (ISBN 0-387987-74-6), 408 pp.

Y muchos más detallados y en ocasiones comentados en las siguientes listas:

<http://www.infovis.net/Biblio/Biblioteca.htm> Biblioteca de INFOVIS.NET De ahí se puede ir a la lista de libros sobre visualización de la información, entre otros.

<http://www.bell-labs.com/topic/societies/asagraphics/resources.html> Dentro de la página de la American Statistical Association, la parte de referencias, dentro de la sección de gráficos estadísticos.

http://www.chrp.org/pdf/HSR012502_handout.pdf

Tutoriales y galerías de gráficos comentadas (en línea)

<http://www.math.yorku.ca/SCS/Gallery/> por Michael Friendly, University of York. (Reino Unido). Se trata de un sitio estupendo en el que se puede encontrar: una lista de aciertos (laureles) y otra de errores (dardos) a la hora de hacer gráficos, una dedicatoria a Tukey, una lista de hitos históricos en el mundo de las gráficas, y algunas cosas más, como la lista de enlaces que incluye.

<http://www.math.sfu.ca/~cschwarz/Stat-301/Handouts/node8.html>.

Una colección de malos gráficos comentados. Dentro de la extensa documentación de una asignatura (statistics for life sciences) de la Universidad British Columbia de Canadá.

<http://www.alzado.org/think/graficos.html> Dentro de una empresa (alzado.org) dedicada a la mejora de páginas web, aparece ese artículo sobre las gráficas en el que se puede encontrar listas de lo que hay y no hay que hacer.

<http://rrpac.upr.clu.edu:9090/~amenend/conf96390.htm> Un curso de estadística aplicada a la educación, en realidad es casi todo estadística. Cuenta de una forma muy clara el tipo de datos, de escalas, correlación y causalidad, etc.

http://www.sapdesignguild.org/resources/diagram_guidelines/index.html. Un tutorial de SAP sobre las representaciones gráficas. Esto está dentro

de la página de “resources” dentro de la principal de SAP (<http://www.sapdesignguild.org/resources/resources.asp>). Allí, además, hay otras cosas de interés como un tutorial sobre ilusiones ópticas, sobre el color, y algunos otros temas que, relacionados con el diseño, también tienen que ver con las gráficas.

<http://my.execpc.com/~helberg/pitfalls/> Un interesante artículo sobre errores estadísticos, que incluye los errores gráficos más comunes en una de sus secciones. Está en la página personal de su autor, Clay Helberg, un estadístico de SPSS.

<http://lilt.ilstu.edu/gmclass/pos138/datadisplay/goodcharts.htm> Tutorial sobre tablas y gráficos de Gary Klass de la Illinois State University (EEUU).

Cursos y textos completos (en línea)

<http://www.iua.upf.es/activitats/semirec/LanguageofGraphics/> Resumen de una conferencia de Engelhardt en la UPF. El eje central del capítulo de la comunicación gráfica.

http://www.sciences-po.fr/cartographie/cartographie_html/5_page5theorie/graphique_bertin2001/e00_indice/e00_indice.html Semiología gráfica de Bertin, las ideas fundamentales del seminal libro de este autor en la red en castellano (entre otras lenguas)

<http://www.graphics.stanford.edu/courses/cs448b-04-winter/> Referencias de una asignatura sobre visualización en la Universidad de Stanford que da paso a algunos contenidos de la misma.

<http://www.csc.ncsu.edu/faculty/healey/PP/index.html> Percepción en visualización. Una página de C. H. Healey, profesor de la universidad de Carolina del Norte, que trata los aspectos de la visualización relacionados con la fisiología y psicología de la percepción. Hay unos interesantes simuladores en línea para hacer tus propios test de percepción.

<http://www.graphics.stanford.edu/courses/cs448b-02-spring/lectures/encoding/> Un curso de Stanford, del profesor Pat Hanrahan, sobre codificación de datos para visualización.

<http://www.fisterra.com/material/investiga/indice.htm> Es el índice de un curso completo que se refiere fundamentalmente a investigación médica: diseño de experimentos y análisis estadístico de los resultados. Un inmenso curso de estadística en línea, y dentro de él una lista de enlaces a programas

estadísticos para análisis de datos en internet.: http://www.fisterra.com/material/investiga/program_internet/program_internet.htm

<http://members.aol.com/johnp71/javasta2.html> Software gratuito para análisis estadístico. Eso está dentro de la pagina personal de John C. Pezzullo, profesor retirado de la universidad de Georgetown (Washington D.C.) en Estados Unidos .

Estadística en general y EDA en particular (en línea)

<http://lib.stat.cmu.edu/> Datos, software y noticias para la comunidad de estadísticos.

<http://www.r-project.org/> El proyecto R, dentro de la filosofía del software libre, un lenguaje y un entorno para la computación estadística y los gráficos.

<http://www.itl.nist.gov/div898/handbook/eda/eda.htm> Estupenda página sobre el EDA (Exploratory Data Análisis) de John Tukey alojada en el NIST.

Páginas personales de personas interesantes

<http://www.infovis.net> Proyecto personal de Juan Carlos Dürsteler, y una página llena de artículos breves en los que se condensa lo mejor sobre visualización.

<http://www.edwardtufte.com/tufte/index> La página principal de Tufte. Aunque en apariencia sea muy comercial, orientada a la venta de sus libros, poster y otros artículos, esconde una buena cantidad de información interesante... *a la* Tufte.

<http://www.research.ibm.com/people/l/lloyd/> Página personal de Lloyd Treinish, un matemático dedicado a la visualización.

<http://www.cs.ubc.ca/~tmm/> Tamara Munzner. Otra página personal de una especialista en visualización. Sus charlas sobre introducción a la visualización resultan muy aclaratorias.

<http://graphics.stanford.edu/~munzner/> Tamara Munzner en otro emplazamiento WEB, complementaria de la anterior.

<http://www.yuriweb.com/> La página de Yuri Engelhardt. No es que impresione por el diseño, pero apunta a mucha información de interés.

<http://www.spss.com/research/wilkinson/> Página personal de Leland Wilkinson.

<http://cm.bell-labs.com/cm/ms/departments/sia/wsc/index.html> Página personal de Cleveland.

<http://www.chrp.org/love/> La página personal de Love. Allí se encuentra una interesante charla sobre representaciones gráficas:

http://www.chrp.org/pdf/HSR012502_abst.pdf

Centros de investigación sobre Visualización de la Información

<http://www.hcrc.ed.ac.uk/Site/GAL.html>

<http://www.ticam.utexas.edu/CCV/index.html>

<http://www.graphics.stanford.edu/>

<http://www-2.cs.cmu.edu/~sage/>

<http://www.cc.gatech.edu/gvu/tccg/>

<http://www.palgrave-journals.com/ivs/journal/v3/n2/index.html>

<http://www.leepotts.com/tehi/>

Procedencia de las figuras

<i>Pág.</i>	<i>Capítulo</i>	<i>Número de figura</i>	<i>Procedencia</i>	<i>Fecha o comentarios</i>
12	0	1	Levante	26/8/96
13	0	2	Elabor propia (splot)	con datos tomados de la figura 1
14	0	3	Guiones de practicas de estudiantes	
20	0	4	Guiones de practicas de estudiantes	
20	0	5-a	Guiones de practicas de estudiantes	
20	0	5-b	Guiones de practicas de estudiantes	
26	1	6	Elabor propia (pwp)	
27	1	7	Internet	Recibida por correo electrónico
28	1	8	Internet	Recibida por correo electrónico
29	1	9	Internet (detalle de 8)	Recibida por correo electrónico
30	1	10	Internet	Recibida por correo electrónico
31	1	11-a	Internet	Recibida por correo electrónico
31	1	11-b	Internet	Recibida por correo electrónico
32	1	12	Internet	Recibida por correo electrónico
32	1	13	Internet	Recibida por correo electrónico
34	1	14	Internet	Recibida por correo electrónico
36	1	15	Internet	Recibida por correo electrónico
37	1	16	Elaboración propia	Con el Paint
41	2	17	El País	7/9/03
42	2	18	El País	7/12/03
45	2	19	Tabla de internet	
47	2	20	Elaboración propia	MS- Excell
48	2	21	Elaboración propia	MS- Excell

GRAMÁTICA DE LAS GRÁFICAS: PISTAS PARA MEJORAR LAS REPRESENTACIONES DE DATOS

<i>Pág.</i>	<i>Capítulo</i>	<i>Número de figura</i>	<i>Procedencia</i>	<i>Fecha o comentarios</i>
50	2	22	Elaboración propia	MS- Excell
50	2	23	Elaboración propia	MS- Excell
52	3	24	El País	7/9/03
53	3	25	El País	7/9/03
53	3	26	El País	23/10/03
55	3	27	El País	11/11/03
57	3	28	El País	2/12/03
58	3	29	El País	7/9/03
59	3	30	Elaboración propia	Modificaciones de la figura 29
60	3	31	Elaboración propia	Adición de figuras en MS. Power Point
61	3	32	Internet	Pg personal de T. Munzner
62	3	33-a	El País	16/12/03
62	3	33-b	El País	Fracción de la 26
64	3	34	Internet	En el pie de figura está la URL precisa
65	3	35-a	Elaboración propia	MS- Excell
65	3	35-b	Elaboración propia	MS- Excell
71	4	36	El Mundo	26/7/96
73	4	37	El País	28/10/03
75	4	38	El País	14-dici-03
77	4	39	El País	10/11/96
79	4	40	El País	Detalle de la figura 41
79	4	41	El País	30-dici-03
80	4	42	Elaboración propia	MS- Excell
81	4	43	Elaboración propia	MS- Excell
82	4	44	Diario de Navarra	25-novi-03
83	4	45	Elaboración propia	Añadidos a la figura 44
84	4	46	El Mundo	7-juli-04
85	4	47	Elaboración propia	MS- Excell
86	4	48	Elaboración propia	Añadidos a la figura 46
86	4	49	El País	2-novi-03
87	4	50	Elaboración propia	Añadidos a la figura 51
88	4	51	IEEE Trnas on Microw Theor and Tech	Vol 44, NO 12, Dec 1996 (2688)
89	4	52	Elaboración propia	MS- Excell
93	5	53	El País	6-novi-03
93	5	54	El País	11/11/2003 (es la 80)
95	5	55	El País	7-dici-03
95	5	56	Elaboración propia	Deformación de la figura 55

PROCEDENCIA DE LAS FIGURAS

<i>Pág.</i>	<i>Capítulo</i>	<i>Número de figura</i>	<i>Procedencia</i>	<i>Fecha o comentarios</i>
96	5	57	Guiones de practicas de estudiantes	Está en la introducc.
97	5	58	Guiones de practicas de estudiantes	
98	5	59	Guiones de practicas de estudiantes	
100	5	60	Guiones de practicas de estudiantes	
101	5	61	Internet	Recibida por correo electrónico
102	5	62	Guiones de practicas de estudiantes	
107	6	63	Elaboración propia	MS- Excell
107	6	64	Elaboración propia	MS- Excell
109	6	65	El Mundo	22-agos-03
110	6	66	Elaboración propia	Sobre la figura 65
111	6	67	El Mundo	13-agos-97
113	6	68	Elaboración propia	MS- Excell
114	6	69a	Elaboración propia	MS- Excell
114	6	69b	Elaboración propia	MS- Excell
115	6	70	El País	12-dici-03
116	6	71	El Mundo	7-juli-04
116	6	72	Elaboración propia	MS- Power Point
122	7	Tabla 1	Tabla	MS- Word
123	7	Tabla 2	Tabla	MS- Word
127	A1	73	Internet	URL detallada en el texto
130	A1	74	Internet	URL detallada en el texto
135	A2	75	Tabla	Datos obtenidos en internet
136	A2	76	Elaboración propia	MS- Excell
137	A2	77	Elaboración propia	MS- Excell
138	A2	78a	Elaboración propia	MS- Excell
138	A2	78b	Elaboración propia	MS- Excell
139	A2	79	Elaboración propia	MS- Excell
141	A3	80	El País	Es la misma figura vista como 54
142	A3	81	Tabla	MS- Word
142	A3	82	Tabla	MS- Word
143	A3	83	Internet	URL detallada en el texto
145	A3	84a	Elaboración propia	MS- Excell
145	A3	84b	Elaboración propia	MS- Excell
145	A3	85	Tabla	MS- Word
146	A3	86	Internet	URL detallada en el texto
146	A3	87	Internet	URL detallada en el texto
147	A3	88	Internet	URL detallada en el texto

Títulos de la colección:

1. Diseño de estructuras lineales.
Un modelo sistemático
de Análisis Estructural
Pedro Gonzaga Vélez
Faustino N. Gimena Ramos
2. Manual de Ingeniería Hidráulica
Armando Coutinho de Lencastre
Camilo Robles García (versión
española)
3. Cálculo de Estructuras
Francisco de Asís Ramírez Chasco
4. Guiones de prácticas de Metrología
Dimensional
Incluye también:
Cuaderno del alumno. Parte 1.
Medida de longitudes y ángulos
Cuaderno del alumno. Parte 2.
Medida de formas y microgeometría
Carmelo Javier Luis Pérez
Miguel José Ugalde Barbería
Ignacio Puertas Arbizu
Lucas Álvarez Vega
Francisco Javier Ripodas Agudo
5. Procesos de conformado por
fundición. Moldeo en arena
Carmelo Javier Luis Pérez
Miguel José Ugalde Barbería
Ignacio Puertas Arbizu
Lucas Álvarez Vega
- 6[4]. Cálculo mecánico de líneas eléctricas
aéreas de alta tensión
Daniel Narro Bañares
Isaac Cenoz Echeverría
- 7[5]. Prácticas para una introducción a la
Hidráulica
Javier Casali Sarasibar
José Javier López Rodríguez
8. Telecomunicaciones en la
construcción
Ignacio R. Matías Maestro
Carlos Fernández Valdivielso
9. Elasticidad y Resistencia de
materiales. Prácticas de Laboratorio.
José Javier Lumbreyas Azanza
Arturo Resano Lázaro
Jesús Zurita Gabasa
Daniel Narro Bañares
Isaac Cenoz Echeverría
Gerardo Fernández Andueza
10. Gramática de las gráficas. Pistas
para mejorar las representaciones
de datos
Joaquín Sevilla Moróder

Colección Ingeniería, 10

Periodistas, científicos, sociólogos, ingenieros, economistas... En multitud de profesiones nos encontramos con la necesidad de interpretar o de realizar gráficas. Las gráficas son elementos habituales en multitud de textos: periódicos, revistas, informes, publicaciones científicas, etc. Y no es de extrañar dado que constituyen un excelente método de comunicación de información cuantitativa. No es improbable, sin embargo, encontrar gráficas que no resultan adecuadas, que no se entienden o que transmitan impresiones equívocas.

Del mismo modo que los textos están sujetos a una gramática que aprendemos desde pequeños en la escuela, las gráficas también disponen de una estructura, sólo que mucho menos formalizada y, que desde luego, no se enseña en la escuela.

El objetivo de este texto es ofrecer unas guías muy simples, basadas en la analogía con la gramática del lenguaje, que ayuden a comprender y especialmente a realizar representaciones gráficas correctas. Está dirigido a todos los públicos, buscándose la máxima sencillez en las ideas, sin ninguna apoyatura matemática o lingüística. Del mismo modo, los ejemplos están tomados casi exclusivamente de

periódicos, de forma que el contenido de las gráficas tampoco despierte de su forma, que es lo que se pretende analizar aquí.

JOAQUÍN SEVILLA MORÓDER (Valencia, 1963). Es Doctor en Ciencias Físicas por la Universidad Autónoma de Madrid. Ha trabajado como profesor ayudante en la Escuela Universitaria de Fisioterapia de la ONCE, donde participó en un proyecto de diseño, construcción, evaluación e implantación de material didáctico para la enseñanza de la física a alumnos ciegos y deficientes visuales. También trabajó como Ingeniero de Diseño en Westinghouse España, participando en varios proyectos para las centrales nucleares españolas. En 1994 se incorporó a la Universidad Pública de Navarra, en la que es profesor Titular de Universidad desde 1996.

Su actividad docente e investigadora en la universidad se centra en temas de instrumentación, bien en el entorno industrial o en el biomédico. También viene impartiendo habitualmente el curso de doctorado "Aspectos básicos de la actividad investigadora".



Universidad Pública
de Navarra
Nafarroako
Unibertsitate Publikoa

ISBN 84-9769-106-7



9 788497 691062