

Contenido semántico y visualización de información: una propuesta de narrativa comunicacional basada en datos

Mario Pérez-Montoro

Universitat de Barcelona, España

<https://orcid.org/0000-0003-2426-8119>

Pérez-Montoro, M. (2025). Contenido semántico y visualización de información: una propuesta de narrativa comunicacional basada en datos. En J. Guallar, M. Váñez, & A. Ventura-Cisquella (Coords). *Comunicación digital. Tendencias y buenas prácticas* (pp. 89-109). Ediciones Profesionales de la Información. <https://doi.org/10.3145/cuvi-com.07.esp>

Resumen

Las gráficas son uno de los principales productos visuales que se utilizan para representar y transmitir la información. Una parte importante de nuestro trabajo diario consiste en la búsqueda, obtención y refinado de un conjunto de datos que nos permita narrar una historia. Pero los datos no hablan por sí solos. Necesitan ser visualizados para mostrar un análisis derivado y permitir que afloren los patrones que encierran. En este trabajo vamos a intentar caracterizar los diferentes y principales contextos y contenidos comunicacionales que podemos transmitir utilizando una gráfica y veremos cuál es la propuesta visual más adecuada para cubrir con eficiencia y eficacia esas intenciones comunicativas. O, dicho en otros términos, en este capítulo queremos desplegar de forma ordenada la metodología habitual que se utiliza para implementar de forma adecuada una correcta visualización de información: en primer lugar, se identifica el tipo de información a codificar, en segundo lugar, se elige el tipo de gráfica que mejor representa ese tipo de información y, por último, se diseñan los elementos visuales que conforman esa gráfica.

Palabras clave

Visualización de información; Visualización de datos; Comunicación visual de la información; Diseño de la información; Gráficas; Representación visual de la información.

1. Introducción

Las gráficas son uno de los principales productos visuales que se utilizan para representar y transmitir la información.

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

Para cubrir ese objetivo, y siguiendo ese esquema, vamos a tratar de desarrollar los siguientes puntos. En el primero de los apartados (apartado 2), abordaremos los principales tipos básicos de información o contenidos semánticos que podemos representar en una gráfica, ilustrándolos a partir de ejemplos que codificaremos en forma de tabla para que puedan ser entendidos de una forma más intuitiva. Y, en el apartado 3, examinaremos cuál es tipo de visualización que se adecua de forma más conveniente para comunicar cada uno de esos tipos de contenidos informativos.

2. Contenidos semánticos

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

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

Normalmente, el contenido semántico que intentamos transmitir con una gráfica es una proposición (una idea, en un sentido no técnico) en la que se afirma la existencia de una relación entre dos o más tipos de información. Concretamente, se afirma la relación entre dos (o más) ítems de información cuantitativa y/o cualitativa. O, dicho de otra manera, en esos contenidos se expresa algún tipo de relación entre valores que pertenecen a dos o más variables cuantitativas, entre las diferentes categorías o atributos que conforman dos o más variables cualitativas, o entre categorías o atributos y valores cuantitativos.

En este sentido, es posible establecer una clasificación de las principales relaciones expresadas en los contenidos semánticos representados por una gráfica (Zelazny, 2001; Few, 2012; Shneiderman, 1996; Pérez-Montoro, 2022). Entre estas, podríamos destacar, por ejemplo, la relación de comparación nominal, la de series temporales, la de ranking, la de parte-todo, la de desviación, la de distribución y la relación de correlación.

2.1. Relación de comparación nominal

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

En términos técnicos este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la comparación entre esos valores numéricos relacionados con esos ítems.

Figura 1

Cargo y sueldo de los empleados de la empresa ACME.

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

(Elaboración propia, datos simulados)

Un ejemplo de este tipo de contenido semántico lo podemos encontrar recogido en la figura 1. En esta figura, aparece una tabla en la que se representa la relación que mantienen los diferentes ítems, categorías o valores (gerente, analista, documentalista, administrativo, auxiliar y becario) de una variable cualitativa (los diferentes empleados (cargos o perfiles laborales) que podemos encontrar en la empresa ACME) con un valor numérico o cuantitativo (el sueldo bruto que reciben cada uno de esos empleados según el cargo que desempeñan en la empresa).

2.2. Relación de series temporales

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

De nuevo, desde una perspectiva conceptual, este contenido semántico se identifica como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable temporal para facilitar la comparación entre esos valores numéricos asociados a esas unidades temporales.

Figura 2

Ventas por trimestre en el ejercicio anterior.

Trimestre	Ventas
Primero	456.876
Segundo	189.387
Tercero	63.829
Cuarto	122.345

(Elaboración propia, datos simulados)

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

2.3. Relación de ranking

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

Figura 3

Ventas por provincias en Cataluña.

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

(Elaboración propia, datos simulados)

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

asignación de valores numéricos establece una ordenación (de mayor a menor o de menor a mayor) entre esas provincias catalanas.

2.4. Relación de parte-todo

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

Desde una perspectiva conceptual, este contenido semántico se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se señala que la agregación de la totalidad de esos ítems cualitativos conforma una entidad o un todo y, por tanto, cada una de esas categorías se identifica como una de las partes de ese todo. Es importante destacar también que la unidad común de medida más utilizada en este tipo de relaciones que expresan la contribución de las partes a un todo es el porcentaje. En ese contexto, el todo es el 100% y cada una de sus partes es una porción de ese 100%.

Figura 4

Actividades desarrolladas a lo largo de un día.

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

(Elaboración propia, datos simulados)

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

2.5. Relación de desviación

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

En un sentido más técnico, es un contenido semántico que se entiende como una proposición (una idea) en la que se afirma la existencia de una relación en la que se asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se comparan esos valores cuantitativos asociados respecto a un valor numérico concreto y de referencia. Normalmente, el valor numérico de referencia se encuentra también

asignado a un ítem cualitativo (de esa misma u otra variable cualitativa). En algunas circunstancias pueden existir dos o más valores de referencia respecto a los cuales se quiere realizar la comparación.

Figura 5

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

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

(Elaboración propia, datos simulados)

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

2.6. Relación de distribución

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

Figura 6

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

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

(Elaboración propia, datos simulados)

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

2.7. Relación de correlación

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

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

Figura 7

Inversión y beneficios por producto fabricado.

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

(Elaboración propia, datos simulados)

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

3. Visualización de contenidos

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

En un sentido práctico, una de las acciones críticas a la hora de generar una buena visualización es, aunque no sea la única, la elección correcta de la gráfica que mejor comunica (de forma eficaz y eficiente) la información que queremos representar. No se trata de una tarea sencilla. Depende de muchos factores, pero especialmente del tipo de información a representar y de nuestra intencionalidad comunicativa. Y es que cada tipo de contenido semántico reclama una representación gráfica determinada dependiendo también de nuestros objetivos comunicativos.

La mayoría de los programas de gestión de datos no nos asisten ni ayudan en la elección de la mejor gráfica para realizar esa visualización (Hugues y van Dam, 2013; Hearn, 2011; Shirley, 2009). Sólo algunas herramientas informáticas nos ofrecen algunas pequeñas orientaciones parciales (incompletas, en muchos casos) para hacerlo en función de la estructura de los datos a representar (Cherven, 2015; Jones, 2014; Khan, 2016).

En este apartado vamos a intentar afrontar ese problema. Vamos a intentar ofrecer una serie de recursos y consejos para poder elegir la gráfica comunicativamente más adecuada para representar cada uno de los contenidos semánticos que solemos codificar. Concretamente, introduciremos los mejores recursos para visualizar de forma adecuada la relación de comparación nominal, la de series temporales, la de ranking, la de parte-todo, la de desviación, la de distribución y la relación de correlación.

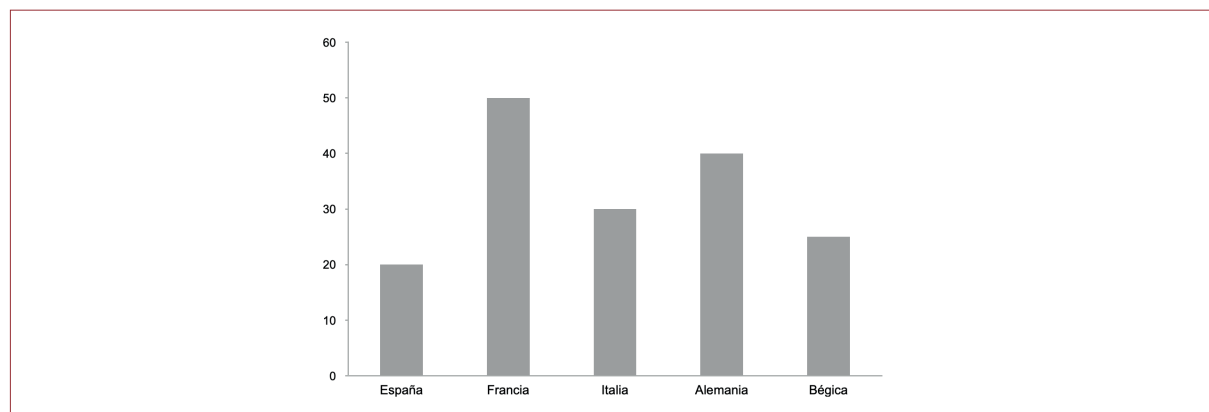
3.1. Visualización de la relación de comparación nominal

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

Sin perder de vista esta restricción semántica, los objetos visuales que mejor codifican esta relación son las barras (o columnas) (figura 8) y los puntos. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo) mostrando que visualmente son independientes entre ellas y que, por tanto, no existe una conexión intrínseca entre las mismas. Al hacerlo de ese modo, facilitan así la comparación visual entre esos valores numéricos relacionados con esos ítems.

Figura 8

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

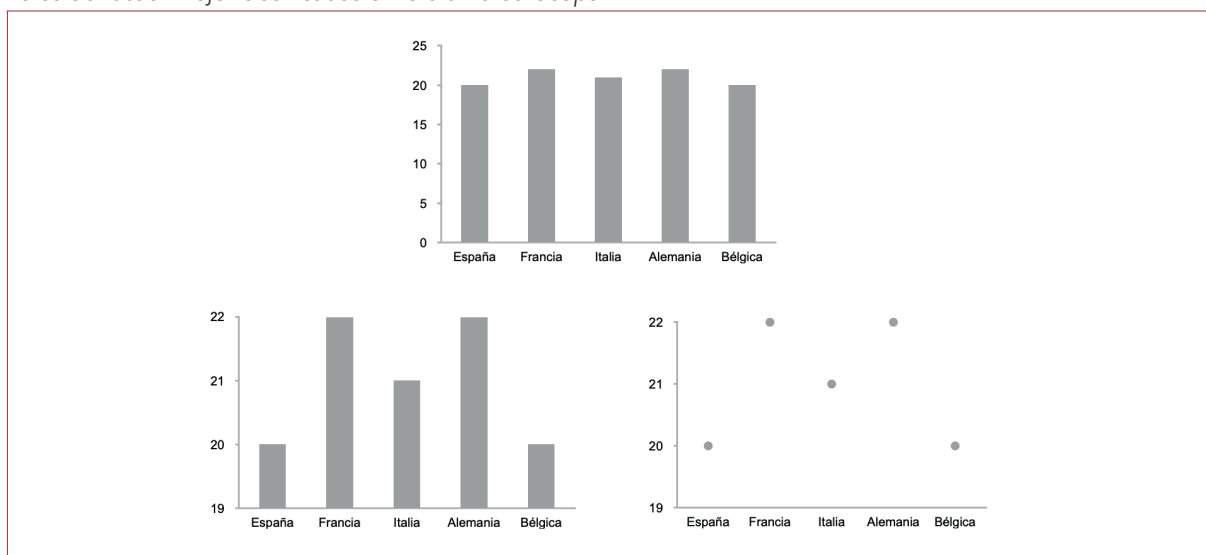


(Elaboración propia, datos simulados)

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

Figura 9

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



(Elaboración propia, datos simulados)

En esos casos, si representamos esos datos mediante una gráfica de columnas (o barras), puede costar apreciar la diferencia entre la longitud de esas figuras geométricas y, dificultar, de esa manera, la comparación visual entre las mismas y los valores numéricos que codifican. Para representarlos de una forma alternativa con columnas (o barras) y hacer evidente las diferencias entre los valores, podemos tener la tentación de comenzar el eje vertical Y (o X, en el caso de la gráfica de barras) en un valor superior a 0, en un valor poco menor que el valor numérico más pequeño representado y hacer visualmente más grandes la diferencia entre la longitud de esas figuras geométricas y de los valores que codifican. Pero esa práctica es incorrecta. Las columnas (o barras), al codificar el valor cuantitativo mediante su longitud, necesitan comenzar de 0 en el eje vertical (u horizontal, en el caso de la gráfica de barras). Si no lo hacen, no codifican correctamente el valor numérico.

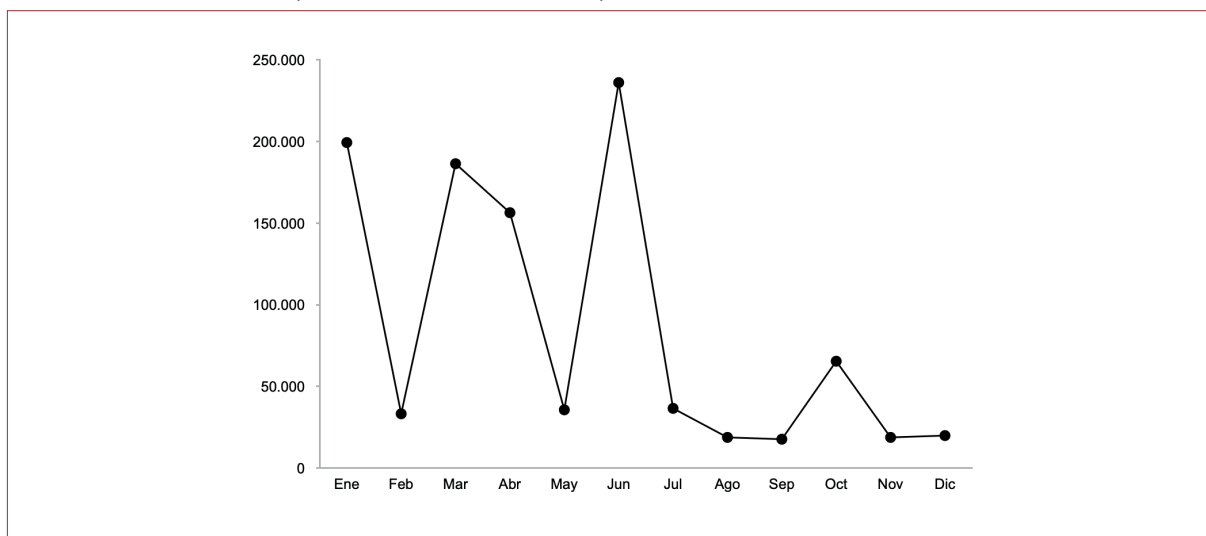
En cambio, si utilizamos una gráfica de puntos como alternativa, esa comprensión errónea no debería producirse ya que los puntos reciben su valor cuantitativo y cualitativo asociado a través de la posición que ocupan en el plano generado por los ejes X e Y (figura 9).

3.2. visualización de la relación de series temporales

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

Figura 10

Ventas mensuales de la empresa ACME durante el año pasado.

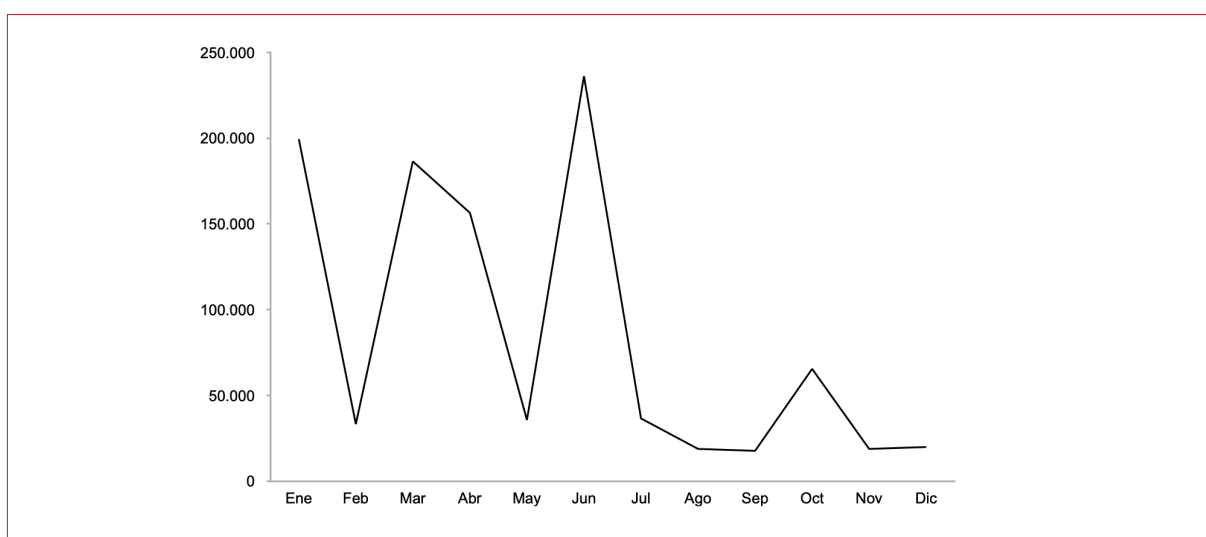


(Elaboración propia, datos simulados)

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

Figura 11

Ventas mensuales de la empresa ACME durante el año pasado.



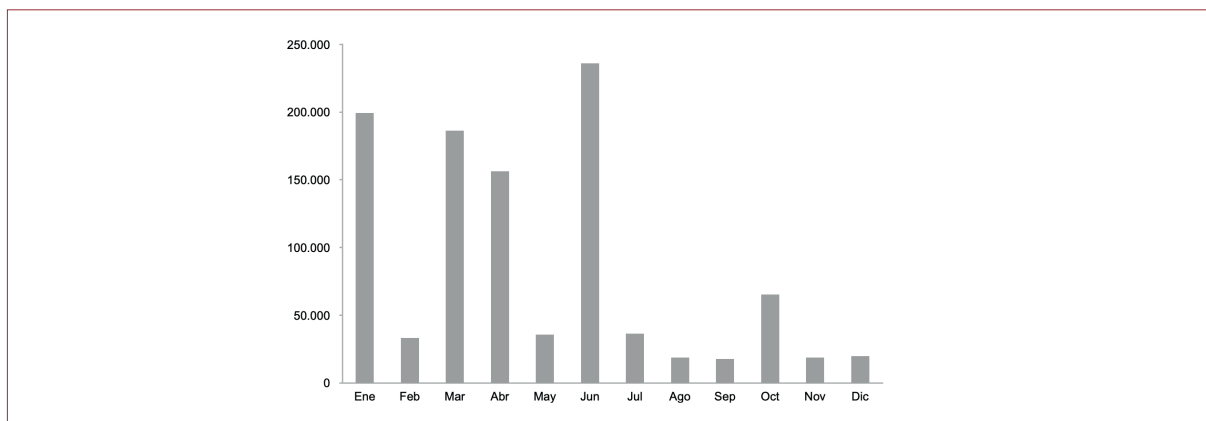
(Elaboración propia, datos simulados)

Teniendo en cuenta todo esto, los objetos visuales que mejor codifican esta relación son la combinación de puntos y líneas (figura 10), las líneas utilizadas de forma no combinada (figura 11) y las columnas. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo)

mostrando que visualmente están conectadas con la siguiente, que son el eslabón o estadio dentro de una cadena temporal. Al hacerlo de ese modo, representan esa continuidad temporal entre esas duplas.

Figura 12

Ventas mensuales de la empresa ACME durante el año pasado.



(Elaboración propia, datos simulados)

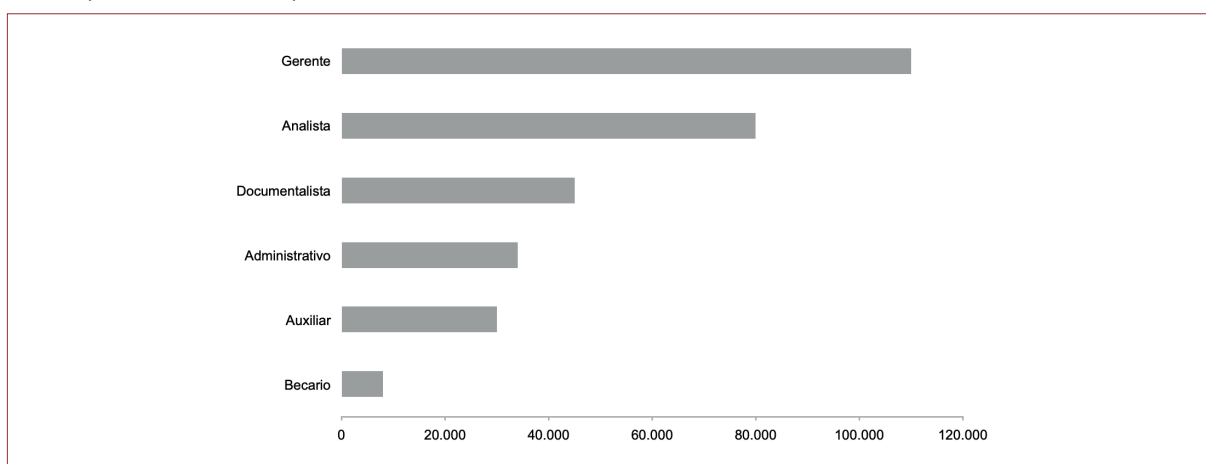
La gráfica de columnas, en cambio, la utilizaremos cuando queramos representar una relación de series temporales, pero queramos poner nuestro foco comunicativo en la transmisión de los valores cuantitativos individuales asociados a los momentos de tiempo de una variable temporal (cuando busquemos facilitar la comparación entre esos valores) y no en la evolución global de la serie (figura 12).

3.3. Visualización de la relación de *ranking*

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

Figura 13

Sueldo por cargo de la empresa ACME.

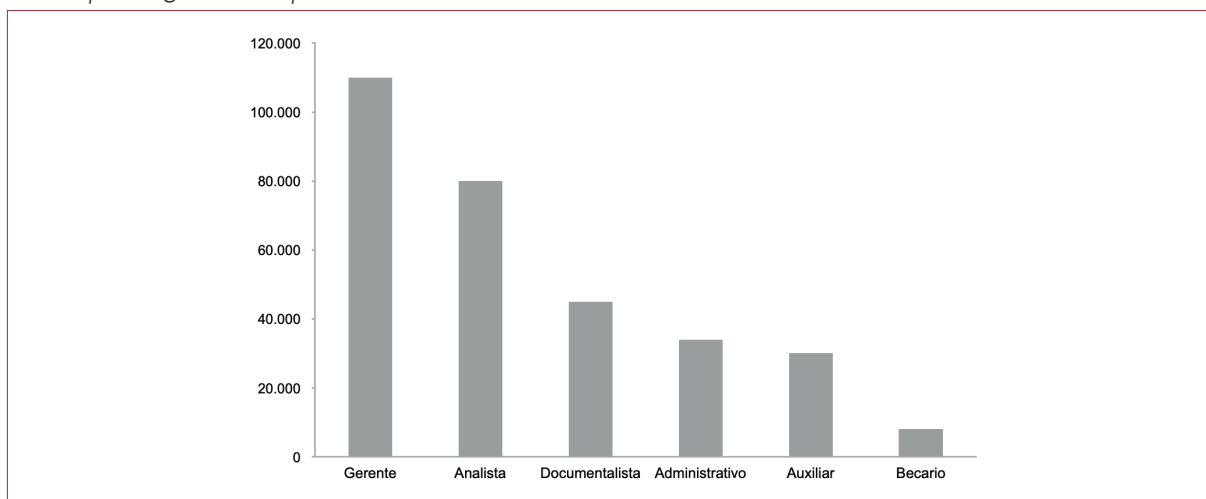


(Elaboración propia, datos simulados)

Teniendo en cuenta esto, una gráfica que represente adecuadamente ese tipo de contenido semántico debe utilizar unos objetos visuales que codifique simultáneamente el valor numérico y la categoría con la que se encuentra asociado y que muestren ese orden intrínseco entre las categorías. Para que estos objetos puedan codificar estas duplas semánticas de forma adecuada, en uno de los ejes debe ubicarse esa variable cualitativa respetando su ordenación interna y, en el otro, la escala cuantitativa.

Figura 14

Sueldo por cargo en la empresa ACME.

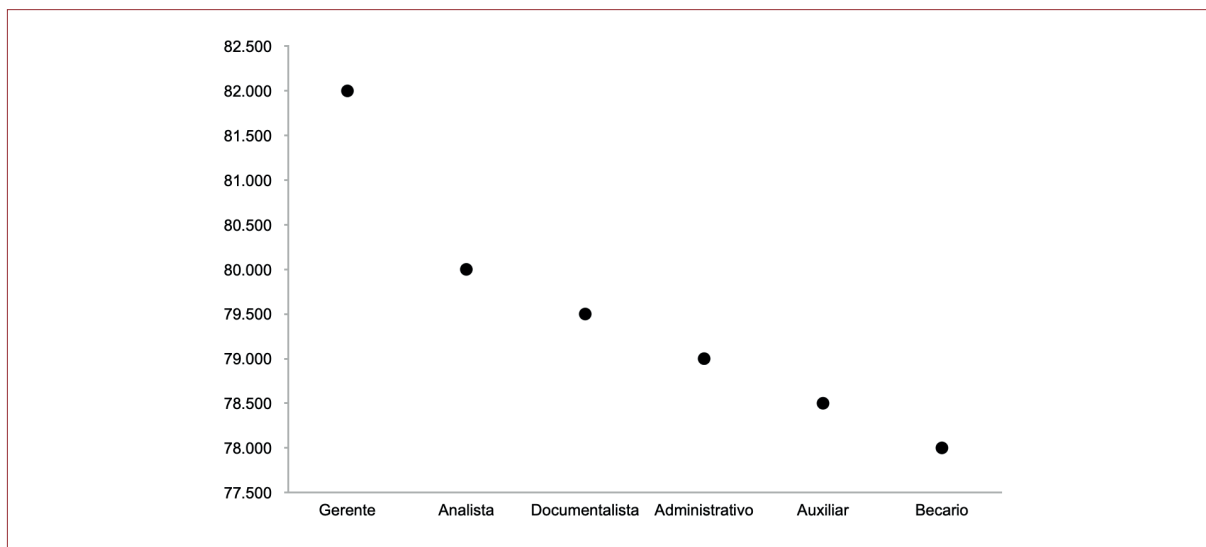


(Elaboración propia, datos simulados)

A la luz de estas condiciones, los objetos visuales que mejor codifican esta relación son las barras (figura 13), las columnas (figura 14) y los puntos. Y lo son porque codifican esas duplas (valor cuantitativo e ítem cualitativo) mostrando visualmente el orden interno establecido entre los ítems de la variable cualitativa.

Figura 15

Sueldo por cargo en la empresa ACME.



(Elaboración propia, datos simulados)

Frente a la de barras o a la de columnas, la gráfica de puntos la utilizaremos en aquellos casos en los que la diferencia entre los valores cuantitativos que se representan sea muy escasa y queramos visualizar de forma cómoda la relación de ranking (figura 15).

3.4. Visualización de la relación de parte-todo

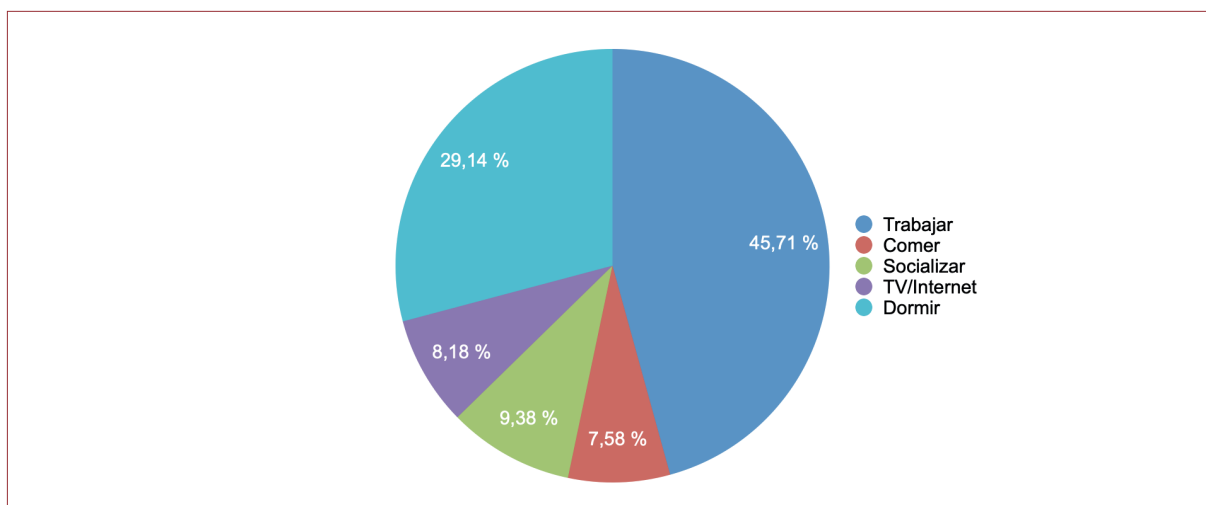
Como ya señalamos, la relación de parte-todo asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa que, a su vez, conforman una entidad o un todo.

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

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

Figura 16

Actividades desarrolladas a lo largo de un día.

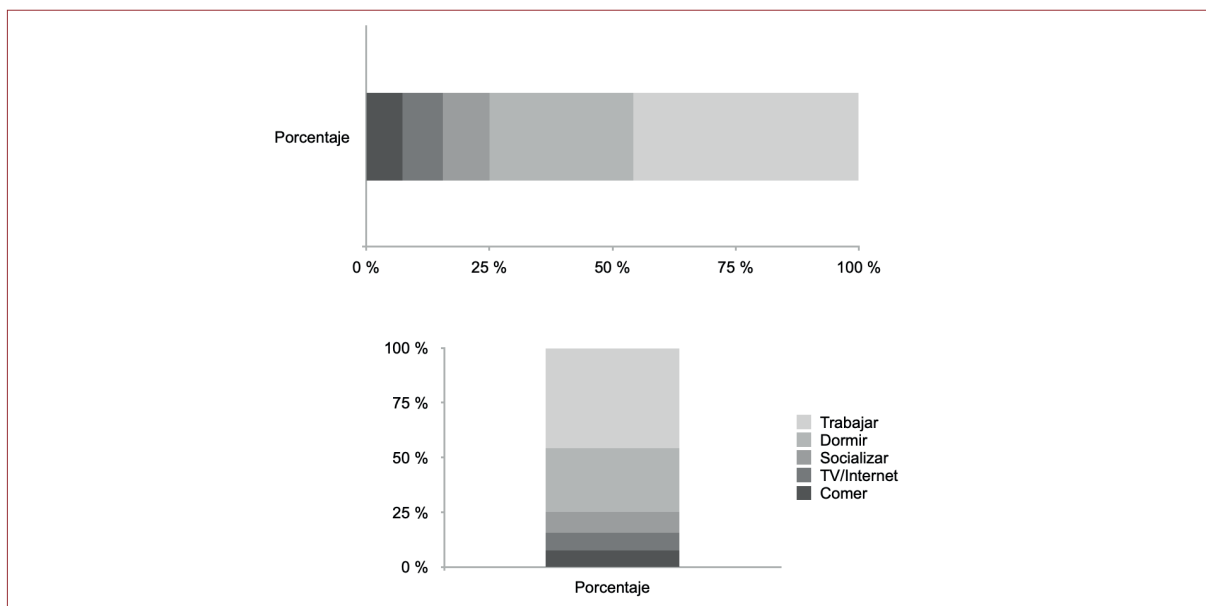


(Elaboración propia, datos simulados)

En una gráfica circular, cada sector nos codifica un valor cuantitativo (en función del ángulo de los radios que los limitan) asociado a un ítem cualitativo (determinado por el color que le asigna la leyenda) y nos muestra las relaciones cuantitativas que mantienen cada una de esas partes o sectores entre sí y con el todo (figura 16).

Figura 17

Actividades desarrolladas a lo largo de un día.

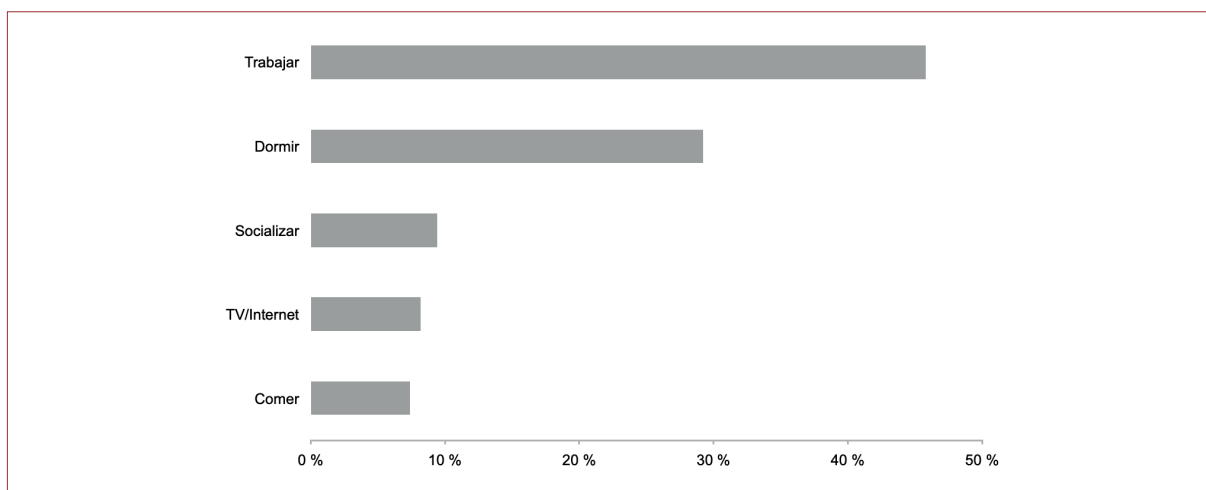


(Elaboración propia, datos simulados)

La otra propuesta habitual que se utiliza para representar la relación de parte-todo es la gráfica de barras (o columnas) apiladas. Estas gráficas representan mediante una barra o columna (una forma rectangular) un valor (un todo) que se obtiene a partir de la agregación de la representación, en forma de cuadrilátero, de los valores de sus partes. Igual que en el caso de la gráfica circular, nos muestra las relaciones cuantitativas que mantienen cada una de esas partes entre sí y con el todo (figura 17).

Figura 18

Actividades desarrolladas a lo largo de un día.

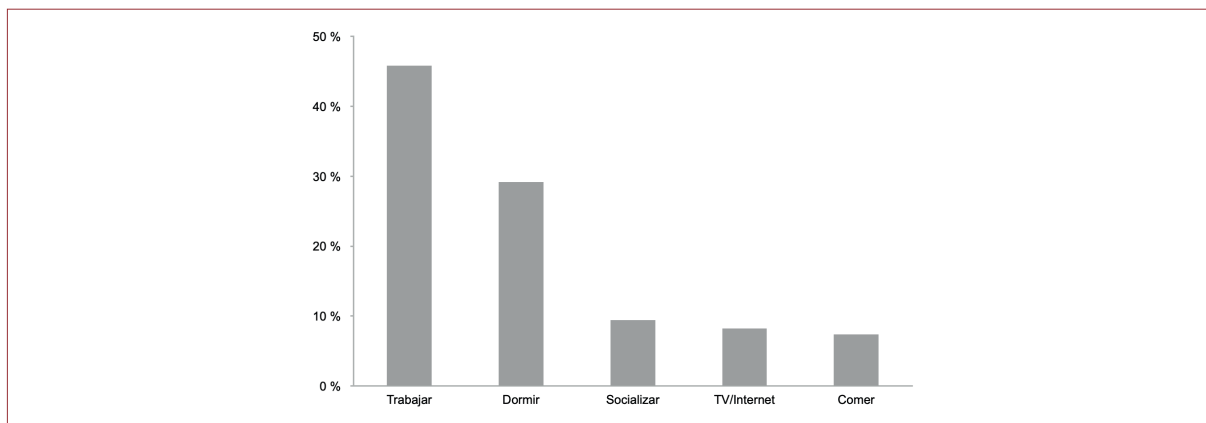


(Elaboración propia, datos simulados)

Frente a estos dos tipos de gráfica que normalmente se utilizan para representar la relación de parte-todo, los objetos visuales que mejor codifican esta relación son las barras (figura 18) y las columnas (figura 19) no apiladas.

Figura 19

Actividades desarrolladas a lo largo de un día.



(Elaboración propia, datos simulados)

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

3.5. Visualización de la relación de desviación

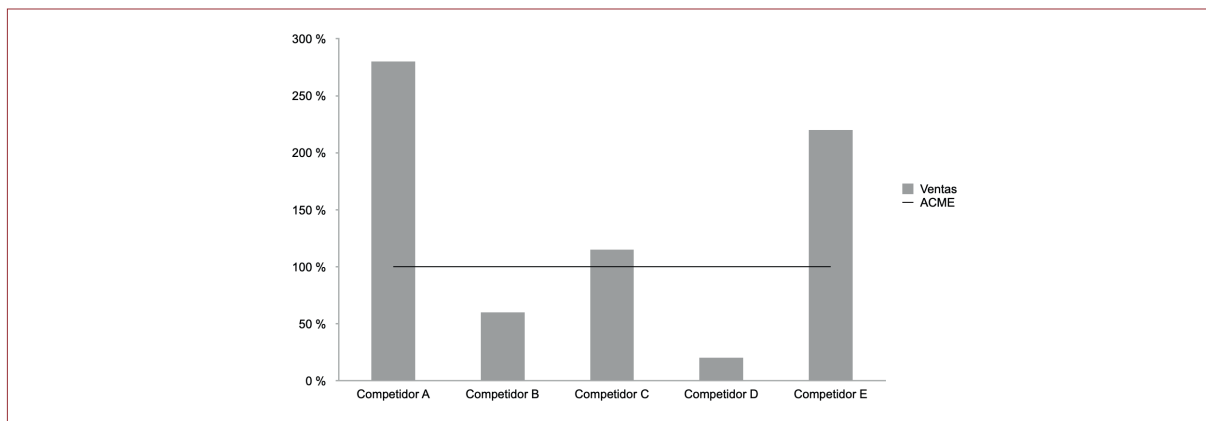
Como ya planteamos, la relación de desviación asigna o conecta un valor cuantitativo a cada uno de los ítems (valores o categorías) de una variable cualitativa y en la que se comparan esos valores cuantitativos asociados respecto a un valor numérico concreto y de referencia.

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

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

Figura 20

Comparación de las ventas en el ejercicio anterior de la competencia sectorial respecto a nuestra empresa.

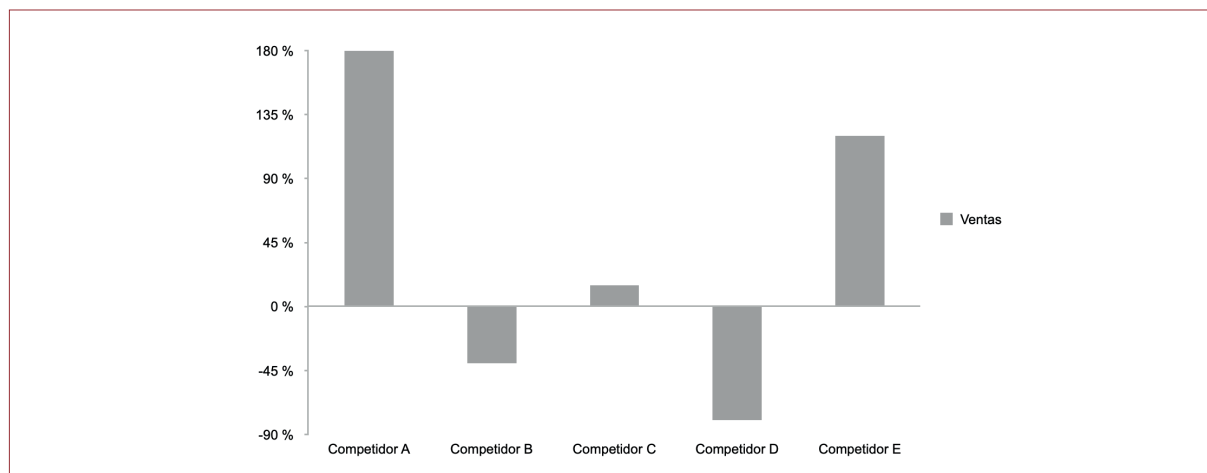


(Elaboración propia, datos simulados)

En una gráfica de barras (o columnas), cada figura rectangular nos codifica un valor cuantitativo (respecto al eje X en el caso de las barras, y respecto al eje Y en el caso de la columnas) asociado a ítem cualitativo (respecto al eje Y en el caso de las barras y respecto al eje X en el caso de la columnas), y, si representamos el valor de referencia mediante una línea o haciéndolo coincidir con uno de los ejes (con el eje Y en el caso de las barras y con el eje X en el caso de la columnas), nos expresa también la relación cuantitativa que mantienen esos valores numéricos respecto al de referencia (figura 20).

Figura 21

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



(Own elaboration, simulated data)

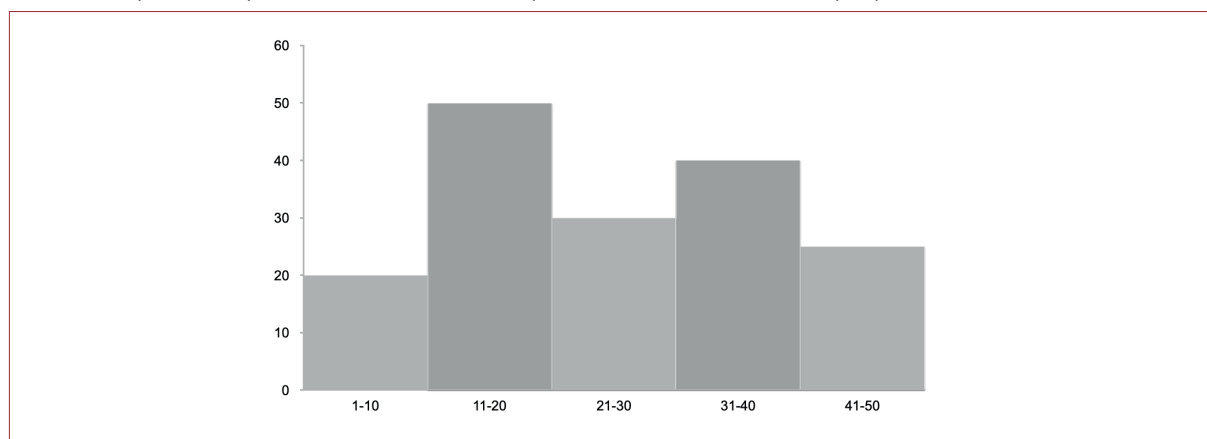
Frente a esta gráfica mixta, podemos ofrecer una versión alternativa sin línea, que incluya las columnas como únicos objetos visuales. En este caso, para no tener que utilizar la línea, haremos coincidir el valor de referencia (las ventas de ACME) con el eje X (figura 21).

3.6. Visualización de la relación de distribución

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

Figura 22

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

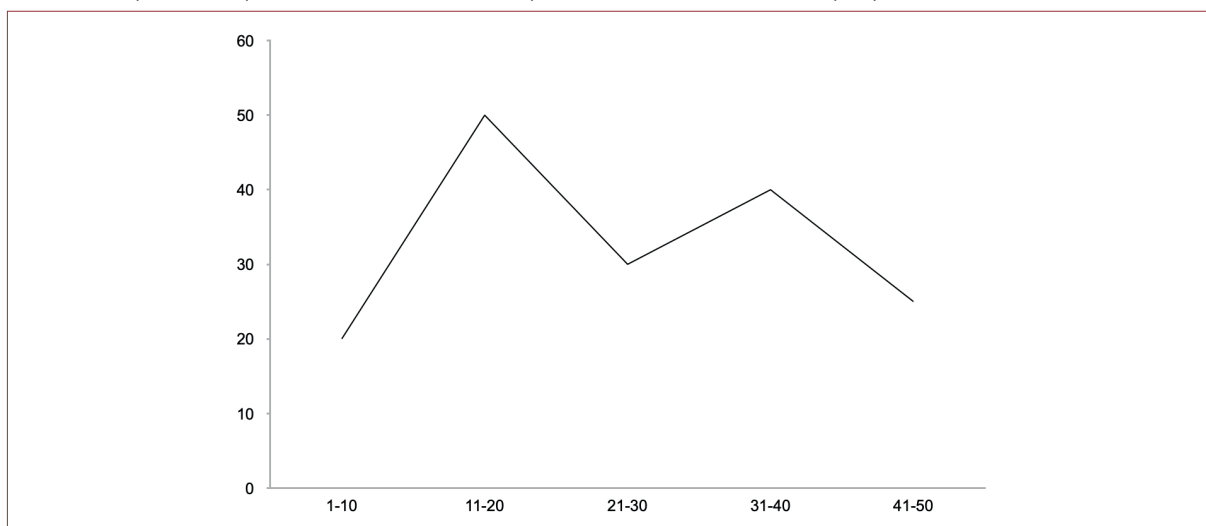


(Elaboración propia, datos simulados)

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

Figura 23

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

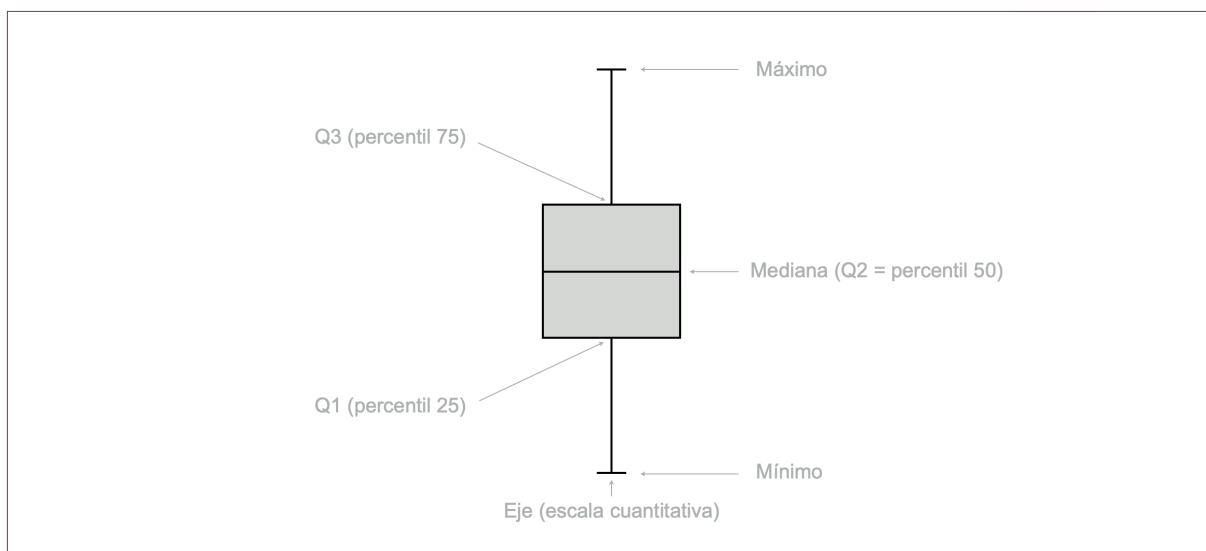


(Elaboración propia, datos simulados)

Otra de las restricciones que debemos tener en cuenta es si queremos representar la distribución de un único conjunto de datos o, por el contrario, queremos visualizar simultáneamente, en una misma propuesta, la distribución de dos o más conjuntos de datos.

Figura 24

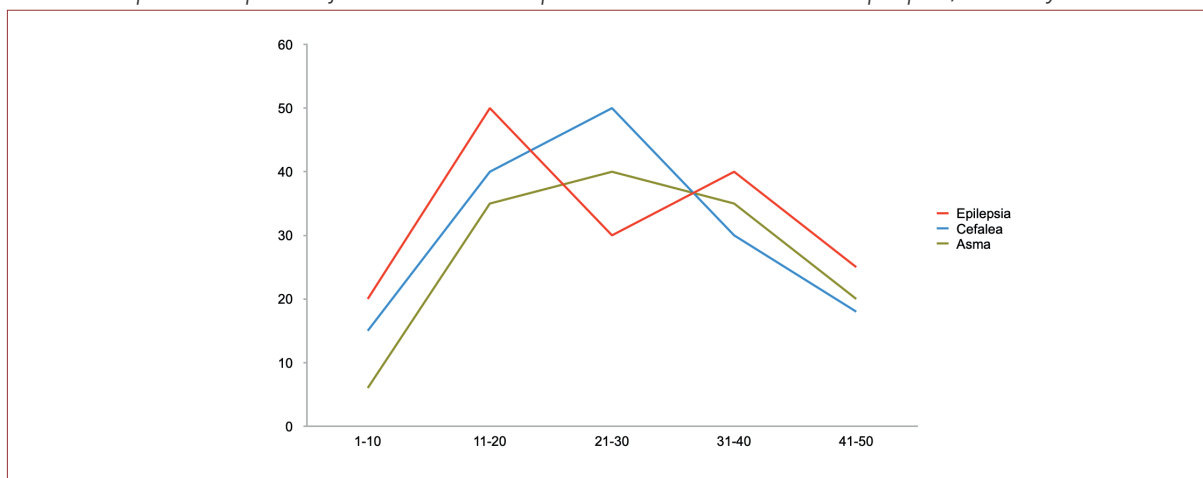
Diagrama de caja



(Elaboración propia, datos simulados)

Figura 25

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

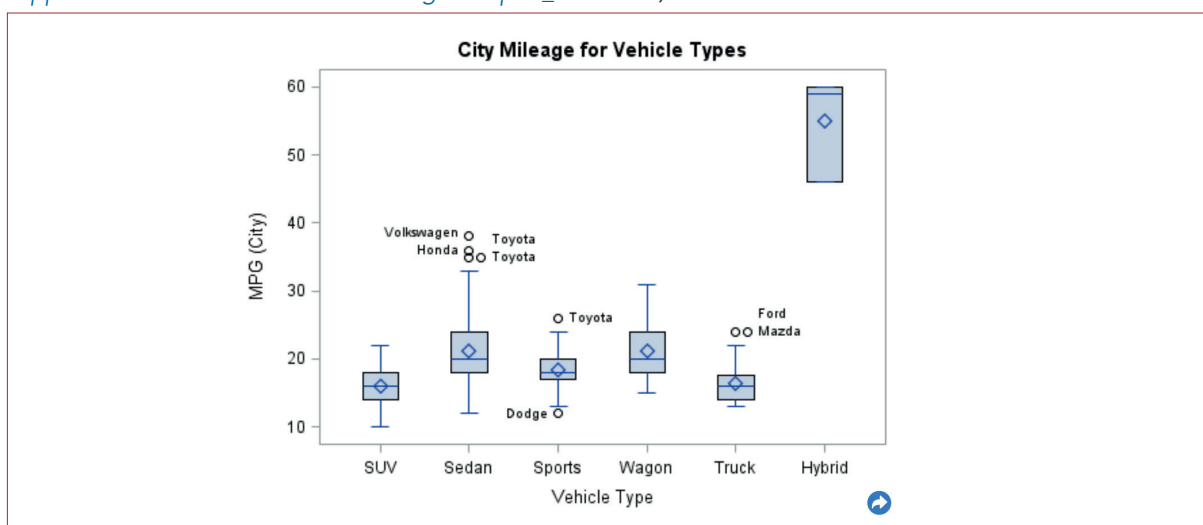


(Elaboración propia, datos simulados)

Centrémonos en el primero de los casos. Cuando nuestro objetivo es representar la distribución de un único conjunto de valores, las gráficas de distribución simple que mejor codifican esta relación son el histograma (figura 22), el polígono de frecuencia (figura 23) y el diagrama de caja (figura 24).

Figura 26

Distribución de vehículos según el modelo y el número de millas que recorren por galón de combustibles (https://support.sas.com/rnd/datavisualization/gtl/boxplot_sect2.htm)



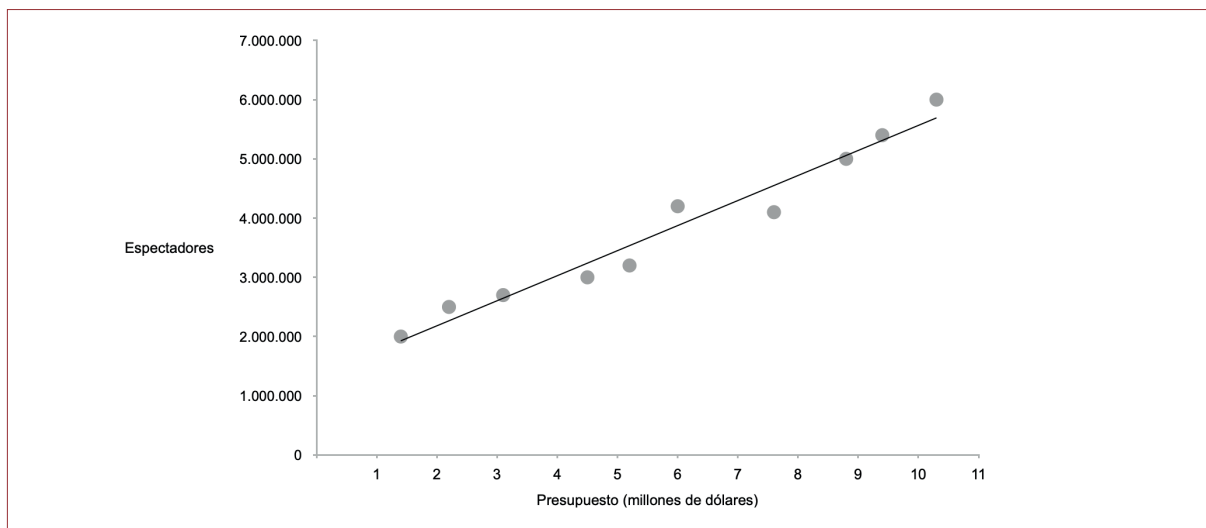
Abordemos ahora el segundo de los casos. Cuando nuestro objetivo es representar simultáneamente la distribución de dos o más conjuntos de valores, las gráficas de distribución múltiple que mejor codifican esta relación son la articulación de diferentes polígonos de frecuencia en una misma gráfica (figura 25) y la combinación de distintos diagramas de caja en la misma propuesta visual (figura 26).

3.7. Visualización de la relación correlación

Abordemos ahora la visualización del último de los contenidos semánticos. Como ya vimos anteriormente, una relación de correlación asigna o conecta dos valores cuantitativos (cada uno perteneciente a una variable cuantitativa distinta) a cada uno de los ítems (valores o categorías) de una variable cualitativa para facilitar la identificación de algún patrón (fuerte o débil) en la variación entre esos dos tipos de valores numéricos relacionados con esos ítems.

Figura 27

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

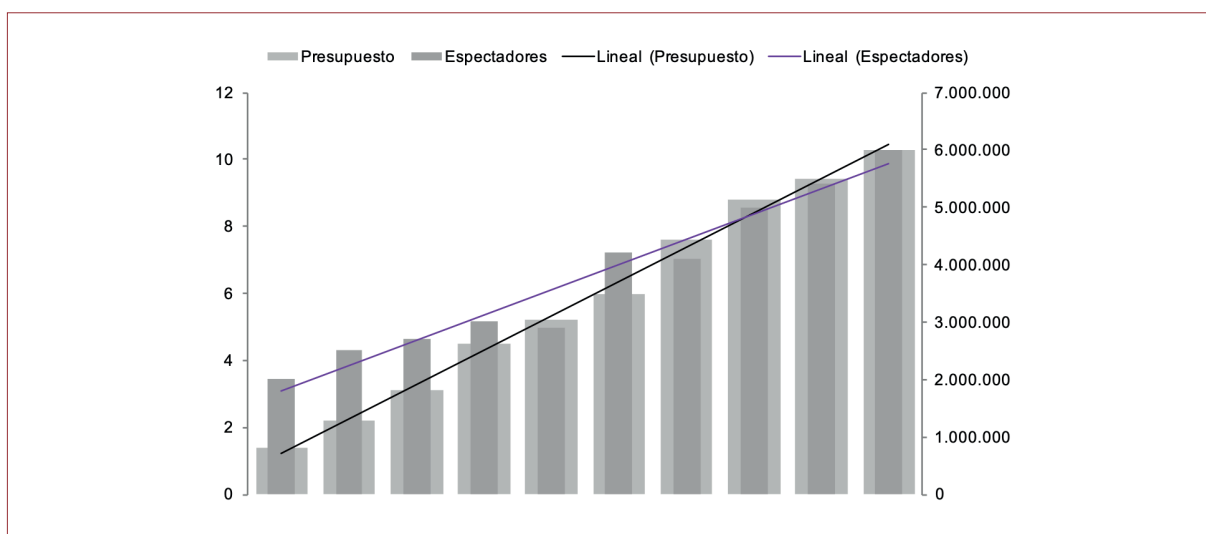


(Elaboración propia, datos simulados)

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

Figura 28

Presupuesto y espectadores de las películas estrenadas la última semana.

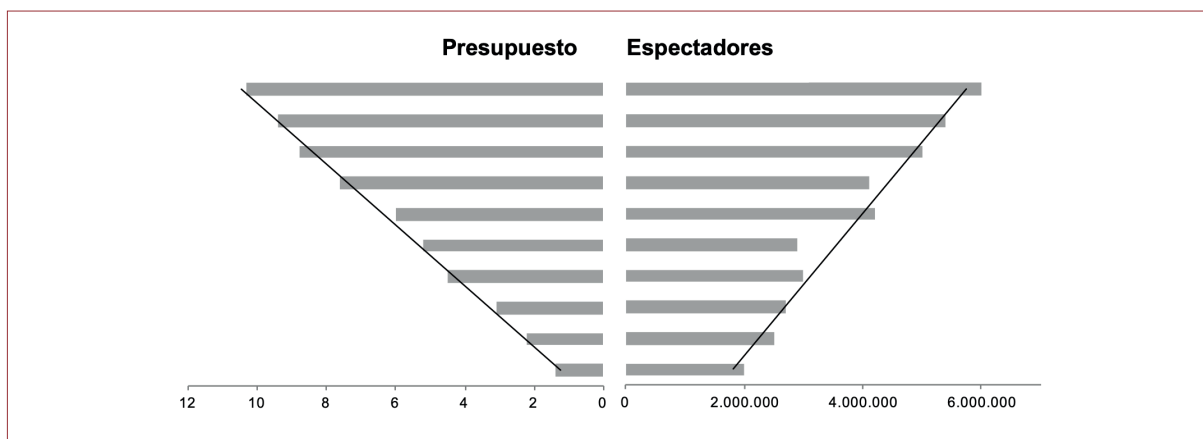


(Elaboración propia, datos simulados)

La propuesta visual que habitualmente se utiliza en la mayoría de los contextos comunicativos para representar este tipo de contenido semántico es el gráfico de dispersión (scatter plot, en inglés), que puede ser entendido como una variante particular de la gráfica de puntos (figura 27).

Figure 29

Presupuesto y espectadores de las películas estrenadas la última semana.



(Elaboración propia, datos simulados)

Algunos autores defienden que, principalmente en el contexto empresarial, el gráfico de dispersión presenta dificultades de interpretación entre los usuarios que no están familiarizados con este tipo de propuestas visuales (Few, 2012). Como alternativas más intuitivas, cuando estamos tratando de visualizar conjuntos de datos no muy extensos, es posible utilizar una gráfica de columnas de correlación (figura 28) o una gráfica de barras emparejadas (figura 29).

4. Nota

Este trabajo es fruto de la revisión, actualización y edición de una parte de un trabajo anterior (Pérez-Montoro, 2022).

5. Financiación

Este trabajo forma parte del proyecto "Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUCOM)". Ayuda PID2021-123579OB-I00 financiada por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.

6. Referencias

- Cherven, K. (2015). *Mastering Gephi Network Visualization*. Birmingham: Ed. Packt Publishing Ltd.
- Few, S. (2012). *Show me the numbers*. Oakland: Analytics Press.
- Hearn, D. (2011). *Computer graphics with OpenGL*. Boston: Pearson.
- Hugues, J. F., & Van Dam, A. (2013). *Computer Graphics: Principles and Practice*. Boston: Addison-Wesley.

- Jones, B. (2014). *Communicating data with Tableau: Designing, developing, and delivering data visualizations*. London: Ed. Safari Books Online.
- Khan, A. (2016). *Jumpstart Tableau: A step-by-step guide to better data visualization*. London: Ed. Apress.
- Pérez-Montoro, M. (2022). *Comunicación visual de la información. Qué y cómo podemos narrar con datos*. Río de Janeiro: IBICT-UNESCO.
- Shirley, P. (2009). *Fundamentals of computer graphics*. Natick, MA: AK Peters.
- Shneiderman, B. (1996). "The eyes have it: A task by data type taxonomy for information visualizations". *Proceedings of the IEEE Symposium on Visual Languages*, IEEE Computer Society Press, pp. 336-343.
- Zelazny, G. (2001). *Say it with charts*. New York: MacGraw-Hill.