V JORNADA DE FORMACIÓN PARA ELABORAR PROYECTOS DE ESTADÍSTICA

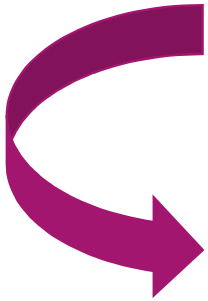
#SGaPEIO
Sociedade Galega para a Promoción da Estatística e da Investigación de Operacións

Presentación de datos estadísticos y gráficos

Facultade de Matemáticas, Campus Vida, s/n
15782 - Santiago de Compostela - A Coruña
http://www.sgapeio.es/

Cómo hacer un proyecto de estadística para presentar a la Incubadora?

Se trata de dar respuesta a una pregunta relevante usando técnicas estadísticas.



Recogida de datos

- Una encuesta
- Un plan experimental
- Explotación de bases de datos

Análisis estadístico e interpretación de resultados

Planificación de la incubadora

- Bases → Finales de noviembre
- Inscripción→ Primera quincena de marzo
- Envío de trabajos→ Segunda quincena de marzo
- Selección de finalistas→ Finales de abril-principios de mayo
- Presentación de finalistas y elección de los ganadores→ Segunda quincena de mayo

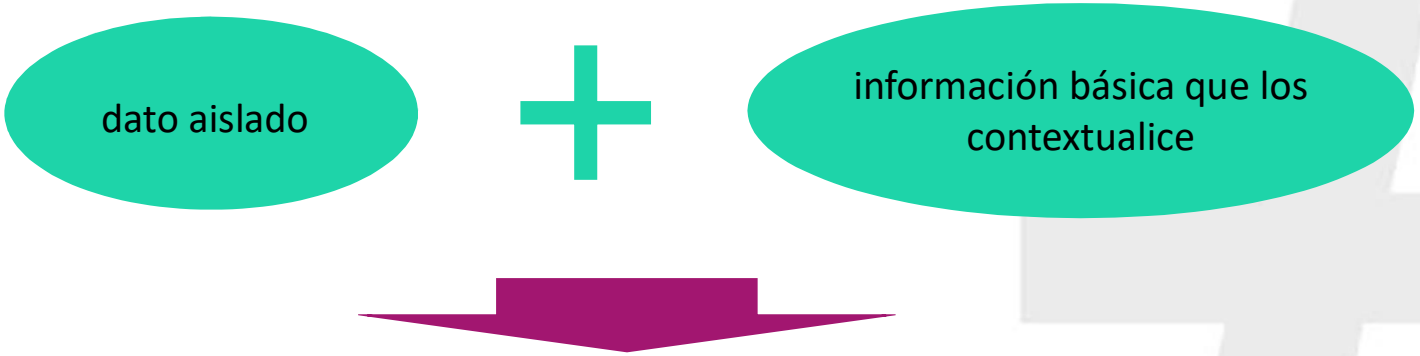
Presentación de la información estadística

Para que un dato estadístico tenga sentido debe ser correctamente presentado, dándole una orientación clara a aquello que se quiere trasmitir.

Una efectiva presentación de los datos requiere de una combinación de distintos elementos: texto, tablas, gráficos, mapas..

El texto

Un texto que trate sobre estadísticas no debe únicamente enumerar datos en forma de palabras, también debe contarnos algo sobre ellos.



qué ha sucedido
quién lo ha hecho
cuándo
dónde
por qué
cómo ha sucedido

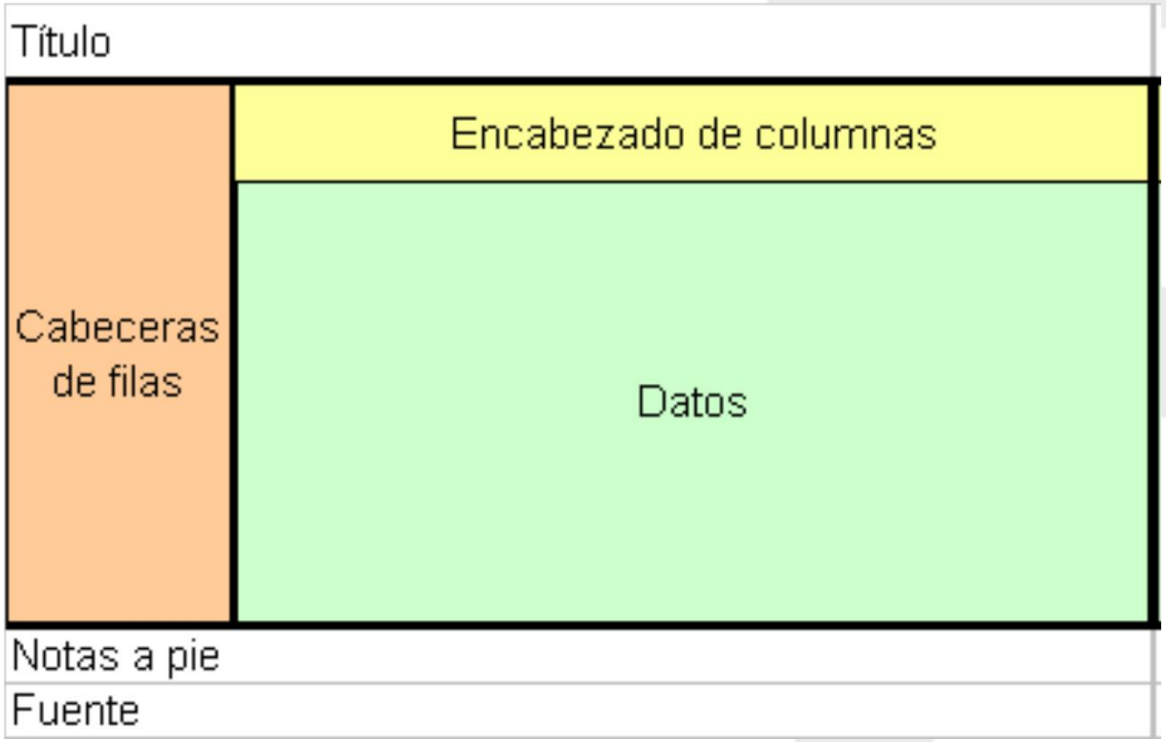
un conocimiento general sobre la información proporcionada
situándola en un contexto
visto desde una perspectiva determinada

Algunos consejos para una escritura clara

- Utilizar frases cortas.
- Tratar una idea en cada frase.
- Dividir las frases largas.
- Iniciar cada párrafo con lo principal del mensaje.
- Realizar párrafos cortos (los párrafos no deben contener más de tres frases).
- Las frases que inician los párrafos no deben contener cifras.
- Listas o epígrafes para guiar la lectura y poder echar vistazos sobre aspectos concretos.
- En el caso de los verbos, evitar la voz pasiva y usar la voz activa. Los verbos en pasiva pueden ser confusos y hacer la escritura densa y menos directa.

Las tablas

Una tabla bien diseñada complementa al texto ayudando a minimizar el número de datos en el mismo y evitando tener que comentar datos poco significativos para el desarrollo de las ideas principales



Elementos de una tabla

1. El **título** de la tabla debe hacer una descripción clara y precisa de los datos. Debe responder las tres preguntas "qué", "dónde" y "cuándo". Ha de ser breve y conciso, evitándose el uso de verbos.
2. Los **encabezados de las columnas**, deben indicar qué datos hay presentes en cada columna y proporcionar las informaciones explicativas necesarias (unidad de medida, período de tiempo o área geográfica).
3. Las **cabeceras de las filas**, deben identificar los datos presentes en cada una de las filas de la tabla.
4. Las **notas a pie de tabla**, se puede utilizar para poner información adicional necesaria para comprender y utilizar correctamente los datos (definiciones, por ejemplo).
5. La **fuente de los datos**, es decir, el organismo que elaboró los datos y el método de recogida de datos (por ejemplo, censo de población o encuesta de población activa). Debe estar indicada en la parte inferior de la tabla.

Claridad en las tablas

Para asegurarse que una tabla es fácil de entender se debe tener en cuenta:

- El **título** de la tabla debe ser corto y describir el mensaje contenido en la tabla.
- Presentar los datos en **orden** o jerarquía, que haga su visualización más intuitiva, orden cronológico en el caso de series temporales.
- Utilizar **separadores de miles**.
- Redondear los datos, si éstos no requieren de un alto grado de precisión o usar mínimamente las **posiciones decimales**.
- **Alinear los números** hacia la coma decimal (o a la derecha en la ausencia de cifras decimales) hace claramente apreciable su valor relativo. No centres los números en una columna, a menos que todos ellos tengan la misma extensión.
- No dejar ninguna celda de datos vacía. Los valores que falten deben ser identificados como "no disponible“.

Los gráficos

Un gráfico es una presentación visual de los datos estadísticos. Es una herramienta muy eficaz ya que muestra datos de manera rápida y sencilla facilitando su comparación, lo que permite revelar tendencias y relaciones entre los datos.

Se debe considerar el uso de gráficos si se desea mostrar:

- **Comparación**: ¿Cuánto? ¿Qué elemento es más grande o más pequeño?
- **Cambios a lo largo del tiempo**: ¿Cómo evoluciona una variable?
- **Distribución de frecuencia**: ¿Cómo se distribuyen los elementos? ¿Cuáles son las diferencias?
- **Correlación**: ¿Están vinculadas dos variables?
- **Parte relativa de un todo**: ¿Cómo se encuentra un elemento en comparación con el total?

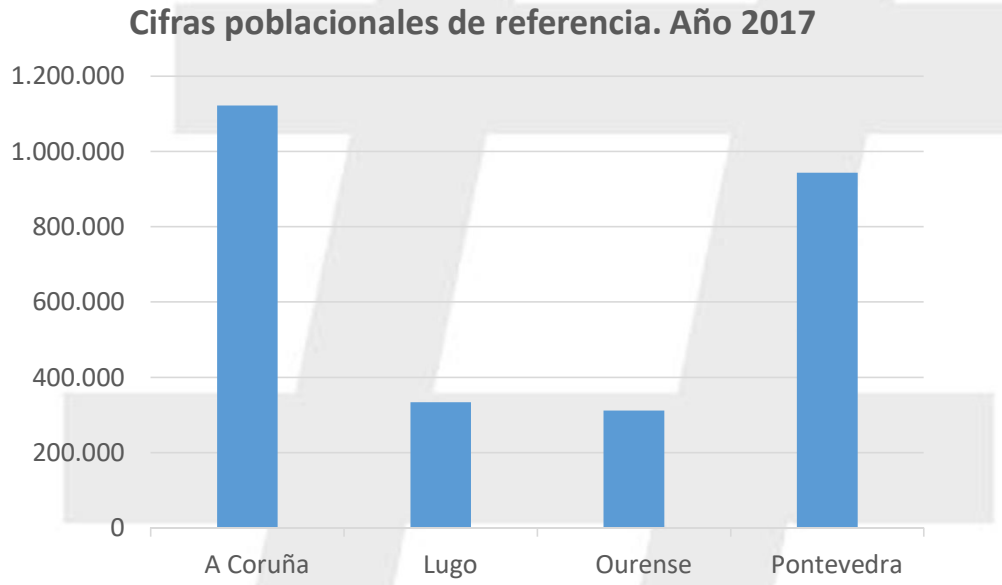
Gráfico de barras

Un gráfico de barras es una representación gráfica en un eje cartesiano de las **frecuencias de una variable cualitativa o discreta**.

Cifras poblacionales de referencia. Año 2017

	Total	Hombres	Mujeres
Galicia	2.710.216	1.309.481	1.400.736
La Coruña	1.121.770	539.064	582.706
Lugo	333.610	162.077	171.533
Orense	311.578	150.333	161.246
Pontevedra	943.257	458.006	485.251

Fuente: IGE. Cifras poblacionales de referencia
Nota: Los totales pueden no coincidir con las sumas correspondientes debido a redondeos en las estimaciones. Estas cifras están baseadas Cfiras de población del INE.



Fuente: IGE. Cifras poblacionales de referencia

En uno de los ejes se posicionan las distintas categorías o modalidades de la variable y en el otro el valor o frecuencia de cada categoría en una determinada escala

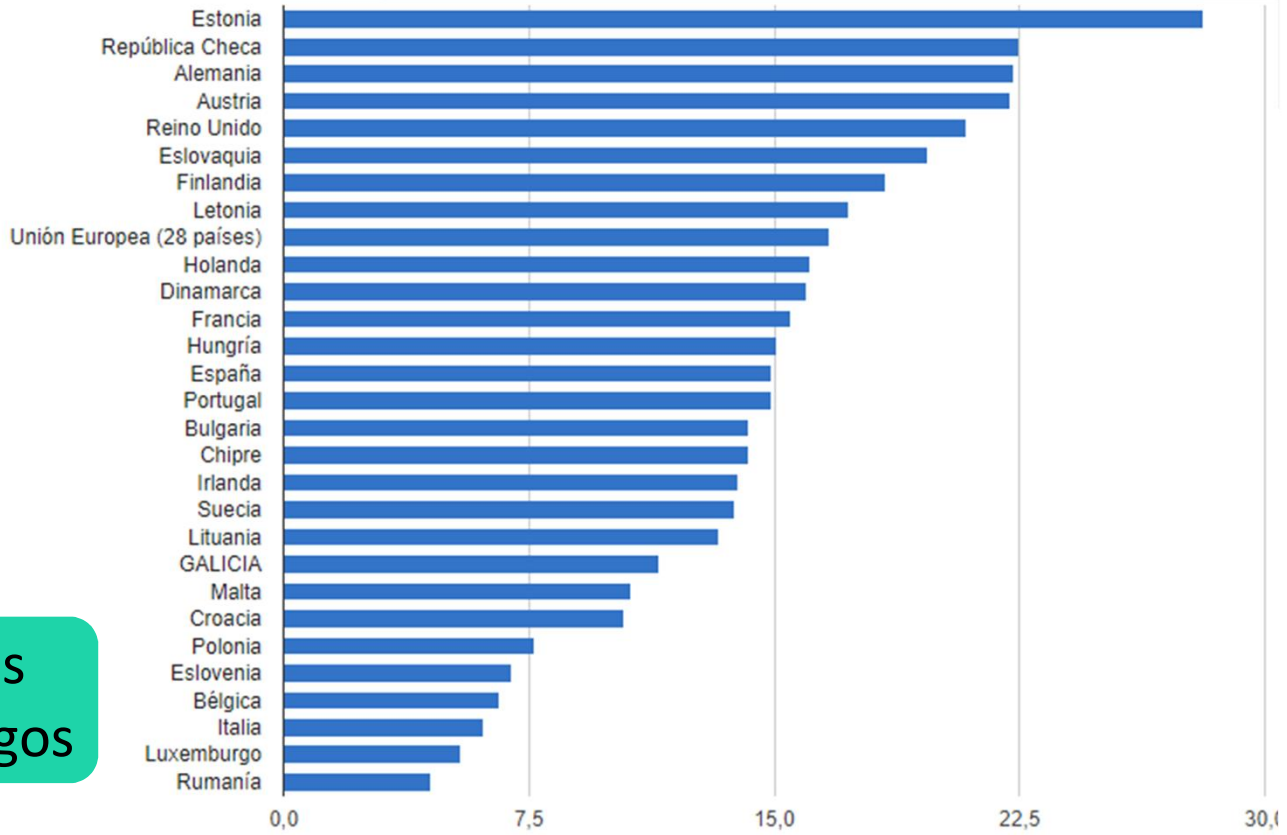
Gráfico de barras

La **orientación** del gráfico puede ser:

- **Vertical:** las distintas categorías están situadas en el eje horizontal y las barras de frecuencias crecen verticalmente
- **Horizontal**: las categorías se sitúan en el eje vertical y las barras crecen horizontalmente.

Muchas categorías
Nombres muy largos

Brecha salarial en los países de la UE 27. Año 2014



Fuente: IGE. Estudio sobre salarios

Gráfico de barras

Las categorías pueden **ordenarse alfabéticamente** facilitando su búsqueda o por sus **frecuencias** facilitando la comparación de los datos.

Se suelen usar para:

- **Comparar magnitudes** de varias categorías.
- Ver la **evolución en el tiempo** de una magnitud concreta.

Evolución de la tasa bruta de alumnado que se gradúa en ESO. Galicia

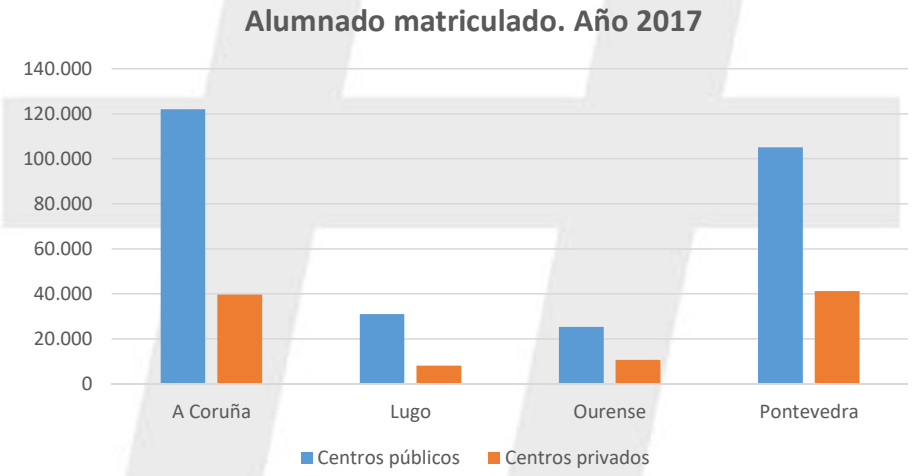
Fuente: IGE, MECD. Estadística de las Enseñanzas no Universitarias, INE. Estimaciones intercensales de la población y Cifras de población

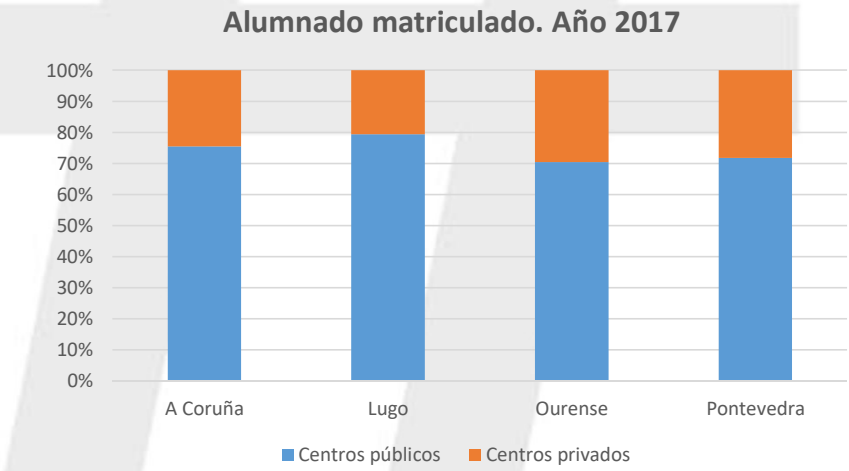
Tipos de gráficos de barras

Sencillo: Contiene una **única serie** de datos

Agrupado: Contiene **varias series** de datos y cada una se representa por un tipo de barra de un mismo color o textura.

Apilado: Contiene **varias series** de datos. La barra se divide en segmentos de diferentes colores o texturas y cada uno de ellos representa una serie.





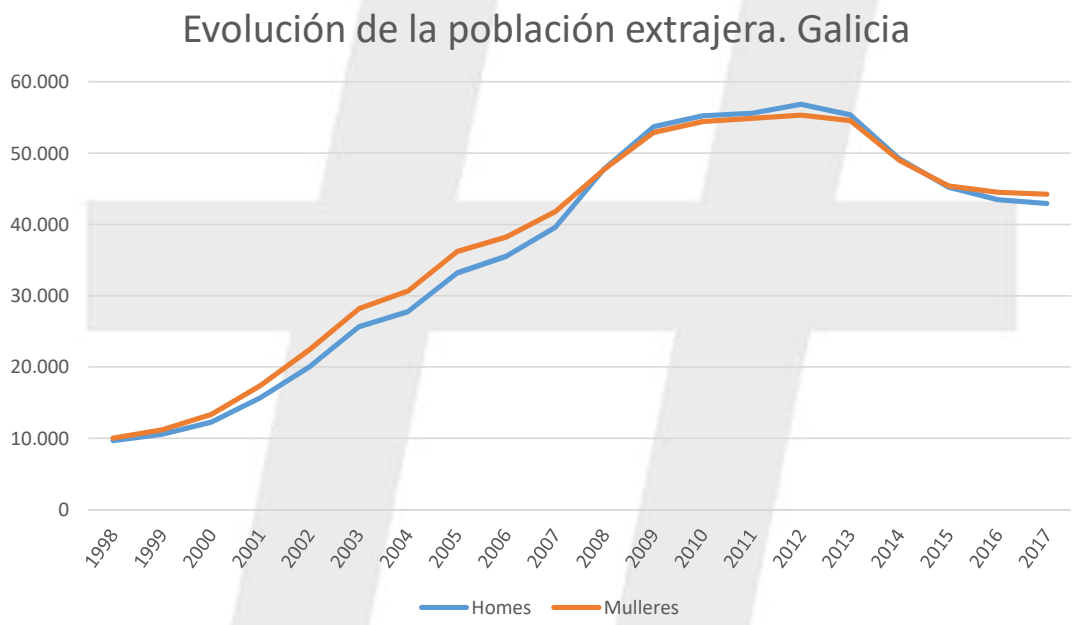
Fuente: IGE

Gráfico de líneas

Un gráfico de líneas es una representación gráfica en un eje cartesiano de la **relación** que existe entre dos variables reflejando con claridad los **cambios** producidos.

- Se suelen usar para presentar **tendencias temporales**.

- En el eje horizontal →variable que indica las unidades de tiempo
- En el vertical → escala de la variable cuya variación en el tiempo queremos ver.
- Pueden aparecer **varias variables** para compararlas



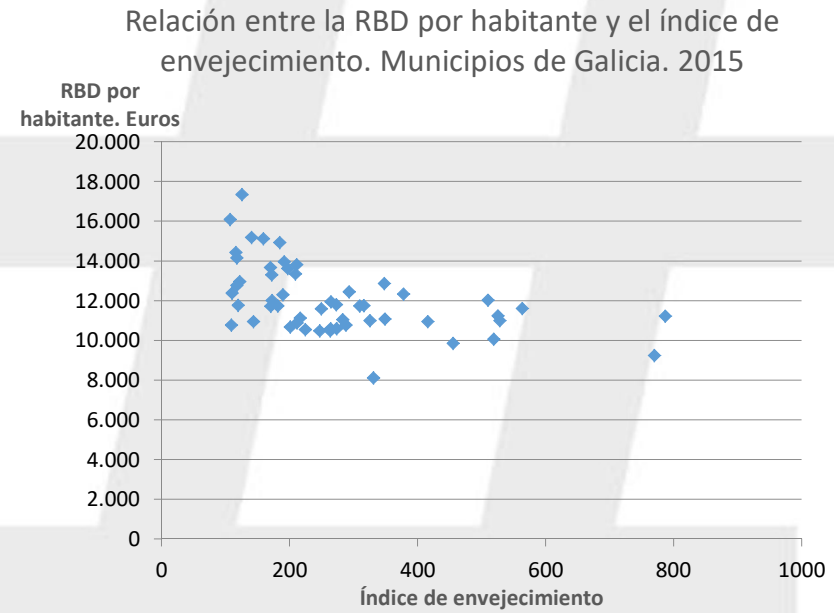
Fuente: INE Padrón municipal de habitantes

Gráfico de dispersión

Este gráfico nos informa del **grado de correlación** entre dos variables

muestra si el incremento o disminución de los valores de una de las variables altera de alguna manera los valores de la otra

variable independiente → eje horizontal
variable dependiente → el eje vertical.



El **tipo de correlación** se puede deducir según la forma de la nube de puntos

Gráfico de sectores

Un gráfico de sectores es una representación circular de las **frecuencias relativas de una variable cualitativa o discreta**

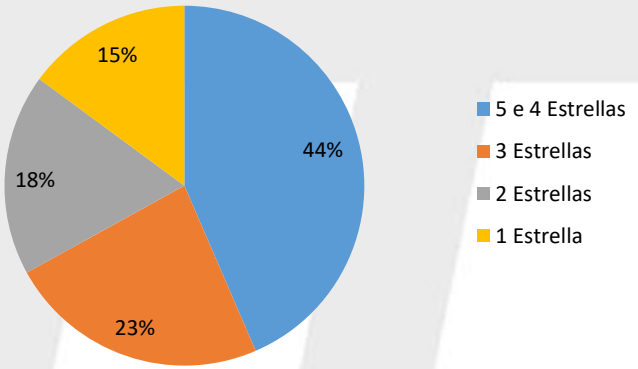
Comparación

El **círculo** representa la **totalidad** que se quiere observar y cada porción, llamadas **sectores**, representan la **proporción de cada categoría** de la variable respecto el total.

Suele expresarse en **porcentajes**

Presentar los segmentos en **orden** facilita la comparación

Viajeros alojados en hoteles. Galicia.
Año 2017

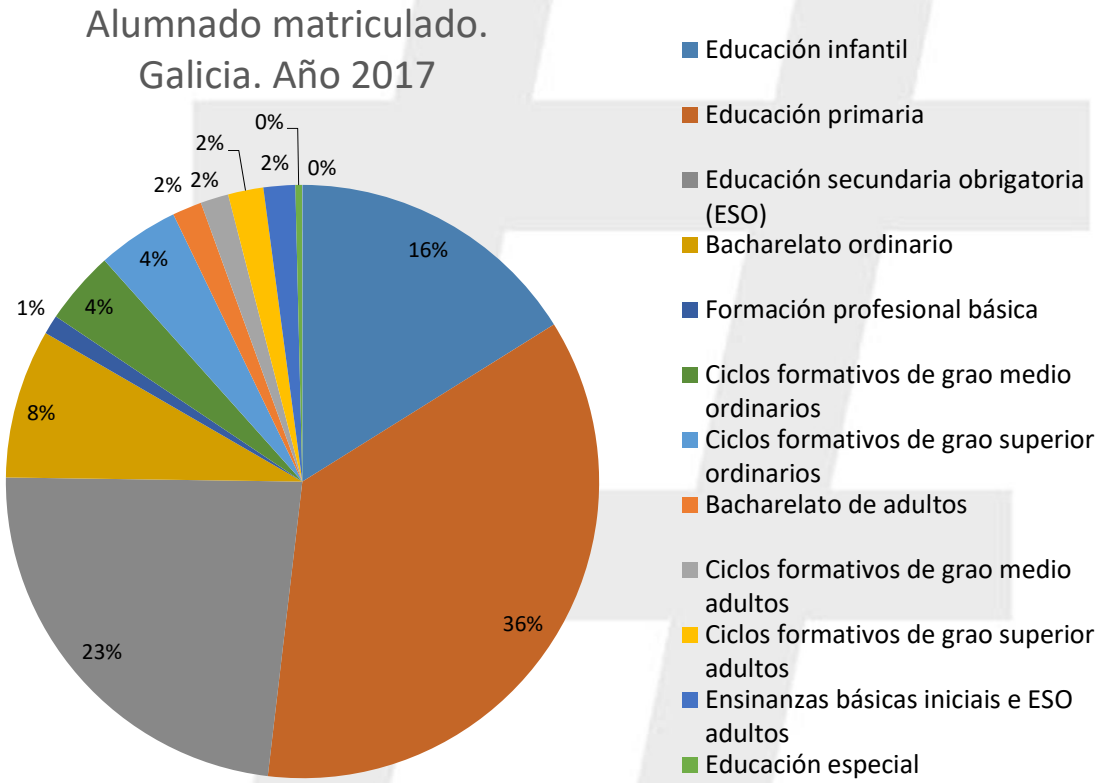


Fuente: IGE, INE. Encuesta de Ocupación Hotelera

Gráfico de sectores

Se deben mostrar un pequeño número de categorías, por lo general no superior a seis.

Es eficaz para visualizar la importancia relativa de una categoría sobre un total.



Gráficos animados

La animación y el vídeo son dos importantes técnicas emergentes de **visualización de datos**.

El desarrollo de nuevas TIC nos permite recibir mensajes a través de imágenes en movimiento.

Este formato hace más fácil contar la historia, mediante la combinación de descripciones en audio o texto con ilustraciones gráficas que explican el significado que se esconde tras los números.

Hans Rosling, cofundador de Gapminder, ha tenido un gran éxito con el uso de la animación para ilustrar datos, como medio para comunicar las estadísticas.

Gapminder es una fundación que pone a disposición de los usuarios tanto datos como herramientas para la visualización http://www.gapminder.org/

https://www.gapminder.org/answers/how-does-income-relate-to-life-expectancy/

Gapminder Offline

Es una programa descargable que no requiere de internet

https://www.gapminder.org/tools-offline/

Podemos visualizar datos precargados o utilizar nuestros propios datos

Posee gran variedad de gráficos animados:

- Burbujas
- Ránkings
- Líneas
- Áreas (sólo para datos precargados)

Ejemplos:

Datos precargados: Burbujas, ránkings

Datos propios desde google docs:

https://drive.google.com/open?id=1Ea7mFMqLm1J0smqQJ4qmGoiXDs5PolySYGGETEwp-0s

Componentes de apoyo de un gráfico

El **título** del gráfico debe dar una idea clara acerca de lo trata el gráfico. Tiene que ser corto y conciso. Se pueden utilizar dos tipos de títulos:

- Un **título informativo** proporciona toda la información necesaria para entender los datos

"qué", "dónde" y "cuándo"

- Un **título descriptivo** pone de relieve las principales pautas o tendencias representadas en el gráfico y plantea en pocas palabras aquello que ilustra el gráfico.

Las **etiquetas** de los ejes deben identificar los valores mostrados en el gráfico. Las distintas categorías se muestran horizontalmente en ambos ejes.

Los **títulos de los ejes** deben identificar la unidad de medida de los datos (*miles, %, edad*). No es necesario cuando la unidad de medida es evidente ("años" para series temporales).

Componentes de apoyo de un gráfico

Las **líneas de división** se pueden añadir en los gráficos de barras y de líneas para ayudar a los usuarios a leer y comparar los valores de los datos.

Las **leyendas** y **etiquetas** de datos deben identificar los símbolos, patrones o colores utilizados para representar los datos en el gráfico.

En la parte inferior del gráfico aparecerán:

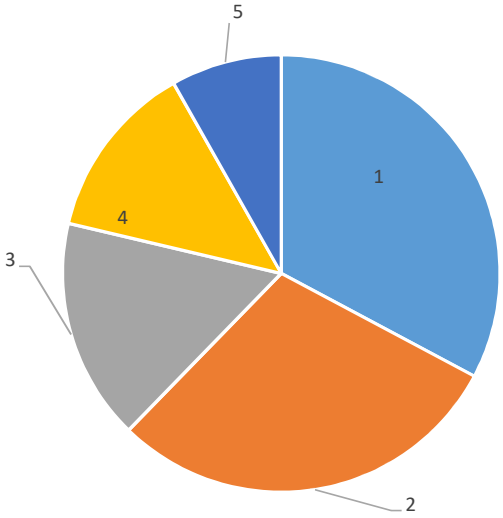
Nota/s que pueden ser utilizadas para proporcionar definiciones.

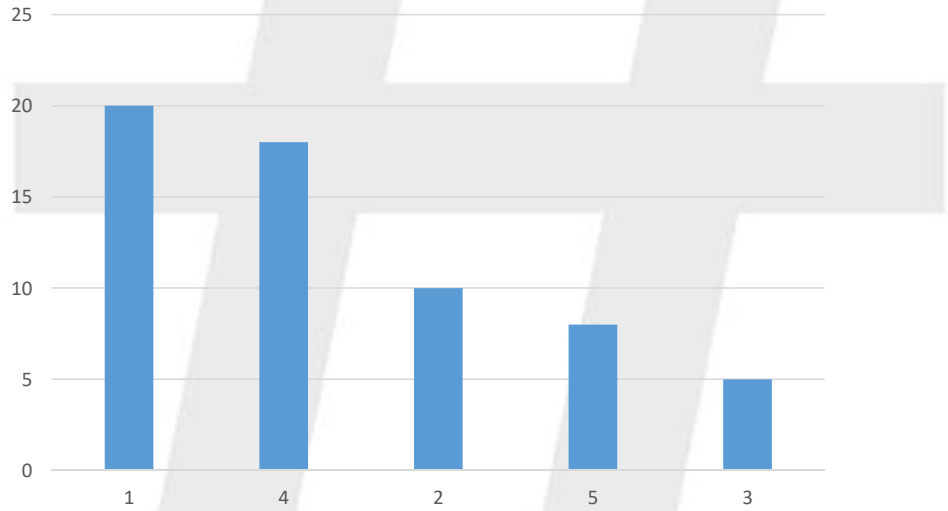
Fuente de los datos

¿Qué tipo de gráfico se debe usar?

Se debe tomar una **decisión** al respecto. Una buena práctica es experimentar con diferentes tipos de gráficos para seleccionar la herramienta más apropiada para comunicar el mensaje.

¿Cuál resulta más claro?





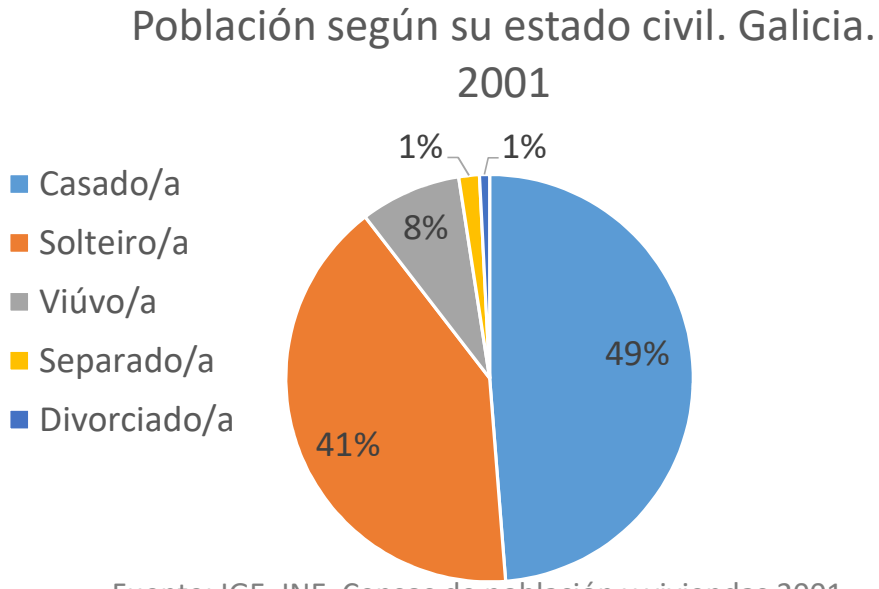
Lograr claridad en los gráficos

- Preferentemente, utiliza línea continua en lugar de otros estilos; y en los rellenos o fondos de gráfico un solo color y sin trama.
- Evita marcadores de datos en los gráficos de línea.
- Utiliza los valores de los datos sólo si estos no interfieren en la visualización del gráfico principal.
- Comienza en cero la escala del eje de ordenadas Y.
- Utiliza una sola unidad de medida por gráfico.
- Utiliza diseños bidimensionales. El gráfico tridimensional induce a error.
- En el texto del gráfico:
  - No utilices abreviaturas.
  - Evita los acrónimos.
  - Escribe las etiquetas de izquierda a derecha

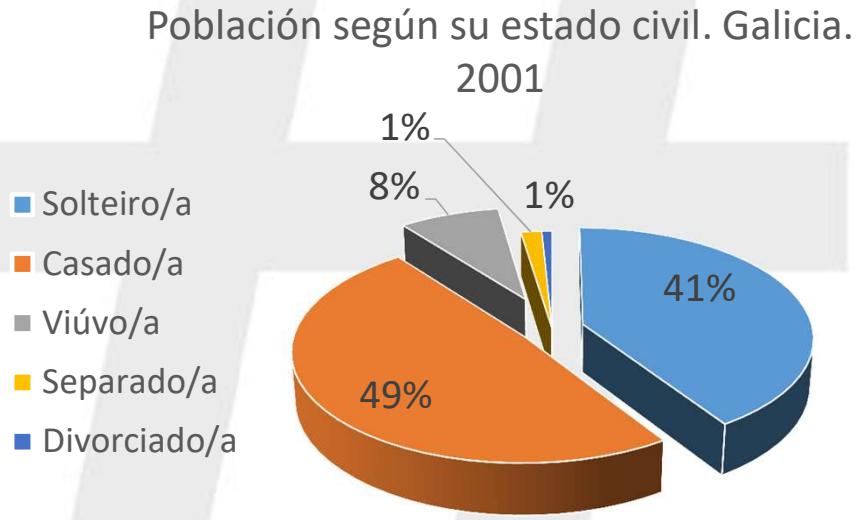
Lograr claridad en los gráficos

Los gráficos en tres dimensiones rara vez añaden valor y a menudo confunden a los lectores

Sin embargo, cuando se representan datos que tienen más de dos atributos, usar la profundidad de hecho puede permitir a los lectores comprender la idea con mayor claridad



Fuente: IGE. INE. Censos de población y viviendas 2001

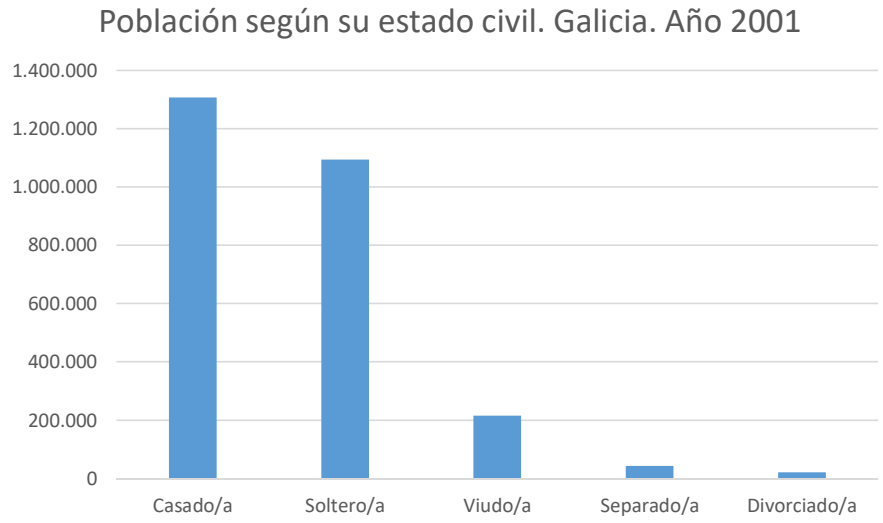


Fuente: IGE. INE. Censos de población y viviendas 2001

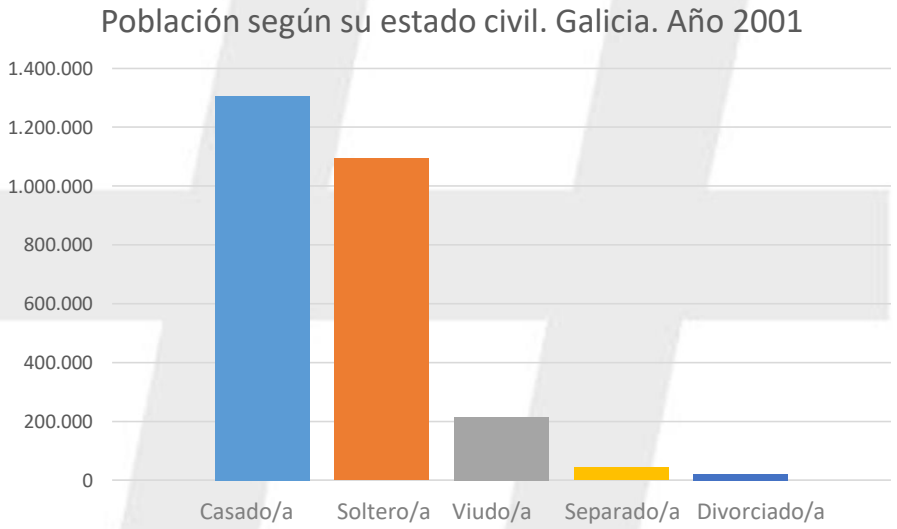
Lograr claridad en los gráficos

En general, se debe evitar añadir cualquier característica que no tenga relación con algún dato

Usar un color diferente para cada valor en un gráfico de barras dificulta la lectura



Fuente: IGE. INE. Censos de población y viviendas 2001



Fuente: IGE. INE. Censos de población y viviendas 2001

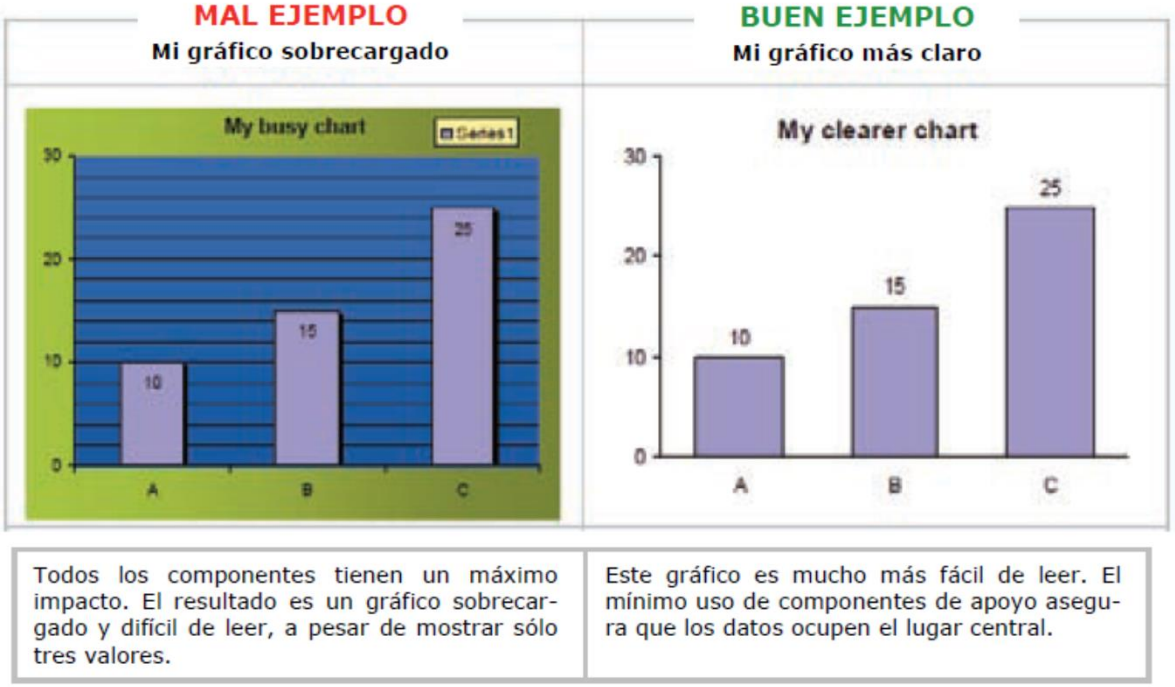
Lograr claridad en los gráficos

Para maximizar la eficiencia de un gráfico, los datos deben ocupar un lugar central. Los componentes de apoyo deben:

- **Ser expuestos sólo si procede**. Títulos de ejes, leyendas y etiquetas de datos pueden ser esenciales para la correcta comprensión del gráfico o no ser en absoluto necesarios, dependiendo de la naturaleza de los datos.

- **Ser sutiles**. Utiliza líneas más claras para los ejes y líneas de división

Los elementos decorativos no deben distraer la atención del lector



Enlaces de interés:

- Portal educativo. IGE: http://www.ige.eu/estatico/educacion/index.htm
- Portal educativo. IGE. ¿Cómo realizar una encuesta?: http://www.ige.eu/estatico/educacion/Matematicas/Enquisa/Teoria.htm
- Pirámides de población. IGE: http://www.ige.eu/igebdt/atlas_comarcal/atlas.jsp?id=1
- Portal divulgativo del INE: Explica https://www.ine.es/explica/explica.htm
- Manuales Naciones Unidas: Cómo hacer comprensibles los datos
  - Parte 1 Una guía para escribir sobre números http://www.unece.org/fileadmin/DAM/stats/documents/writing/MDM_Part1_Spanish.pdf
  - Parte 2: Una guía para presentar estadísticas http://www.unece.org/fileadmin/DAM/stats/documents/writing/MDM_Part2_Spanish.pdf

Muchas gracias por vuestra atención y animaros a participar en la Incubadora de Sondaxes e Experimentos!!

#SGaPEIO Sociedade Galega para a Promoción da Estatística e da Investigación de Operacións