

- 1- Introducción
- 2- Tipos de datos
- 3- Gráfico de barras
- 4- Pirámide de población
- 5- Gráfico de líneas
- 6- Gráfico de Pareto
- 7- Gráfico de sectores
- 8- Pictograma
- 9- Gráfico de dispersión
- 10- Cartograma
- 11- Bibliografía

Un gráfico estadístico es una representación visual de una serie de datos estadísticos. Es una herramienta muy eficaz, ya que un buen gráfico:

- capta la atención del lector;
- presenta la información de forma sencilla, clara y precisa;
- no induce a error;
- facilita la comparación de datos y destaca las tendencias y las diferencias;
- ilustra el mensaje, tema o trama del texto al que acompaña.

En primer lugar, vamos a empezar definiendo los tipos de datos que existen y a continuación veremos cómo se pueden representar a través de los gráficos.

Cualitativos: se refieren a cualidades o modalidades que no pueden expresarse numéricamente. Pueden ser:

- **ordinales:** si siguen un orden o secuencia (ej. el abecedario, los meses del año).
- **categoricos:** si no siguen ningún orden (ej. el estado civil de las personas: solteros, casados, viudos, divorciados y separados).

Cuantitativos: se refieren a cantidades o valores numéricos. Pueden ser:

- **discretos:** si toman valores enteros (0, 1, 2, 3...). Ej. el número de hijos, el número de alumnos de una clase...
- **continuos:** si pueden tomar cualquier valor dentro de un intervalo (ej. la estatura o el peso de las personas).

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

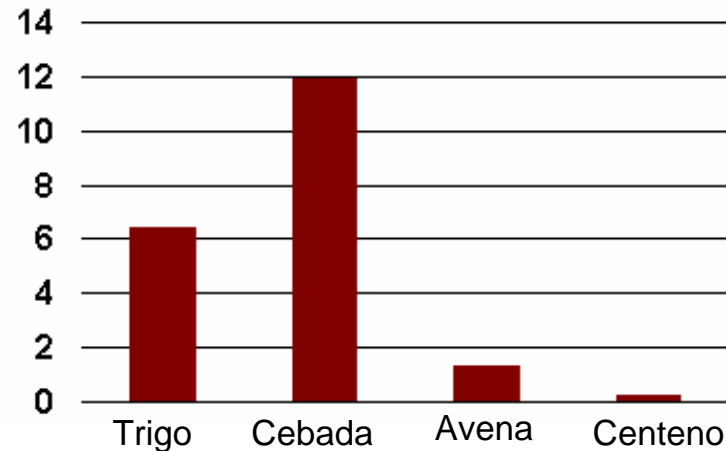
Producción Agrícola 2007

Cereales	Miles de toneladas
Cebada	11.945
Trigo	6.436
Avena	4.310
Centeno	261

Fuente: Ministerio de Agricultura,
Alimentación y Medio Ambiente

Producción de cereales en España. 2007

Millones de toneladas



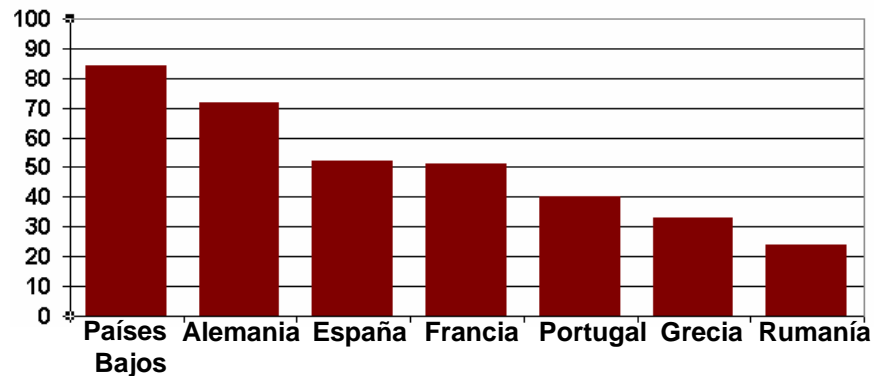
En uno de los ejes se posicionan las distintas categorías o modalidades de la variable cualitativa o discreta (en el ejemplo, el tipo de cereal) y en el otro el valor o frecuencia de cada categoría en una determinada escala (en el ejemplo, la producción en millones de toneladas de granos).

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. Suelen usarse cuando hay muchas categorías o sus nombres son demasiado largos.

Las categorías pueden **ordenarse alfabéticamente** facilitando su búsqueda o por sus **frecuencias** facilitando la comparación de los datos. Veamos el siguiente ejemplo del porcentaje habitantes usuarios de internet del año 2007 por países (Fuente: Unión Internacional de Telecomunicaciones).

Orientación vertical y orden por frecuencias



Orientación horizontal y orden alfabético

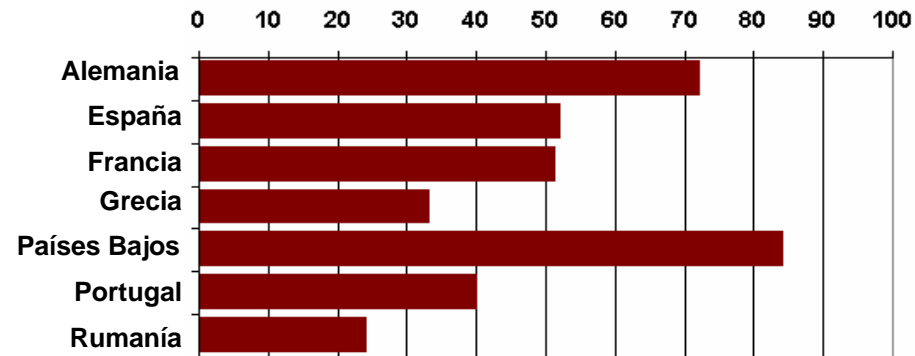


Gráfico de barras

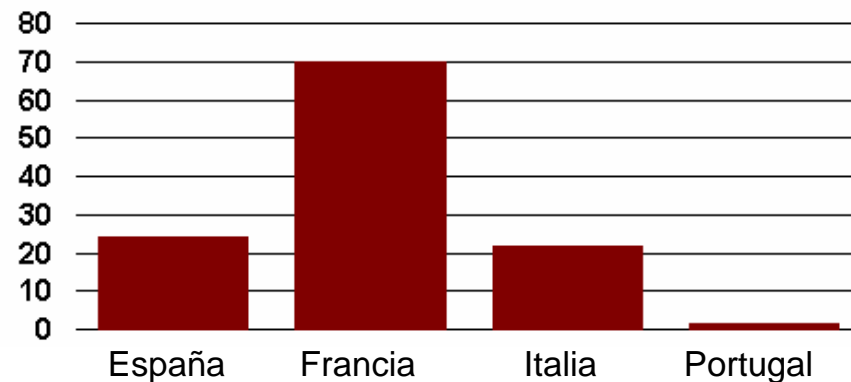
Se suelen usar para:

- **Comparar magnitudes** de varias categorías.

- Ver la **evolución en el tiempo** de una magnitud concreta.

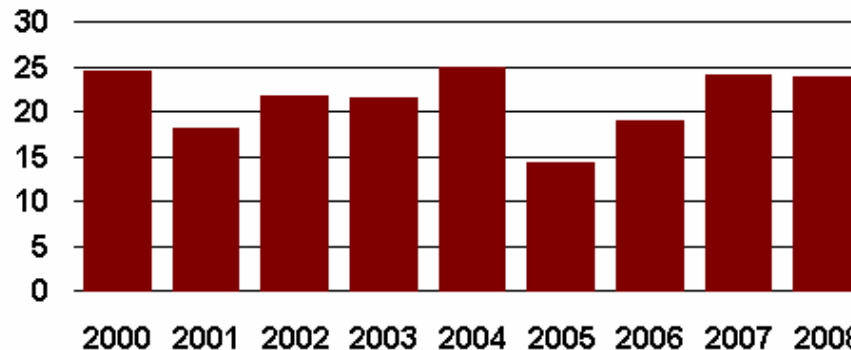
Producción de cereales. 2008

Millones de toneladas



Producción de cereales en España

Millones de toneladas



Fuente: Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO)

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.

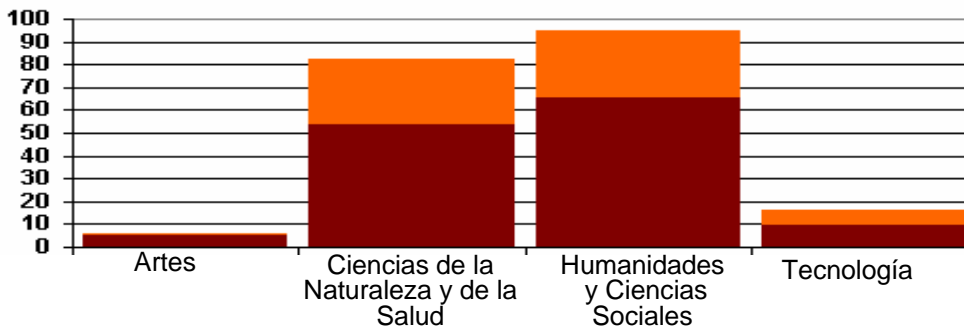
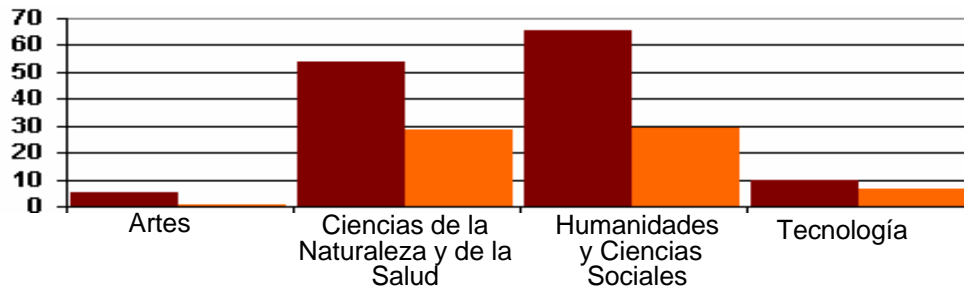
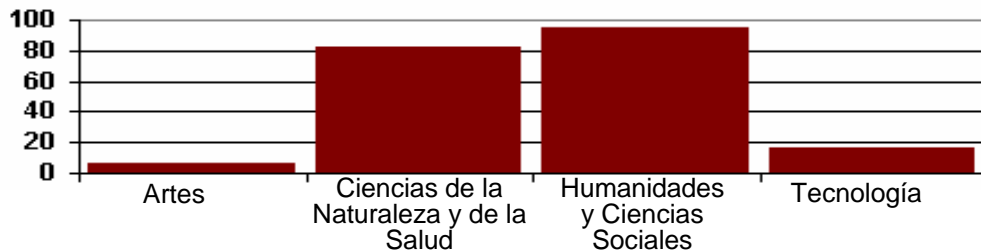
- **Apilado:**

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

Alumnado que terminó Bachillerato por su opción académica

Miles de alumnos

■ Centros Públicos ■ Centros Privados



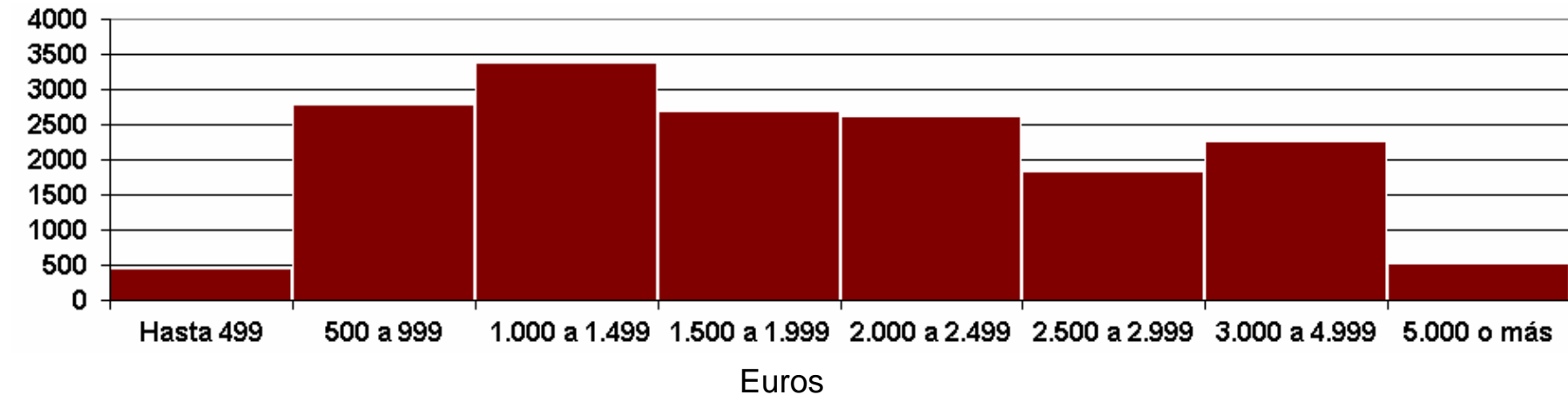
•Histograma:

Se usa para representar las frecuencias de una **variable cuantitativa continua**.

En uno de los ejes se posicionan las clases de la variable continua (los intervalos o las marcas de clase que son los puntos medios de cada intervalo) y en el otro eje las frecuencias. No existe separación entre las barras.

Número de hogares según ingresos. 2008

Miles de hogares



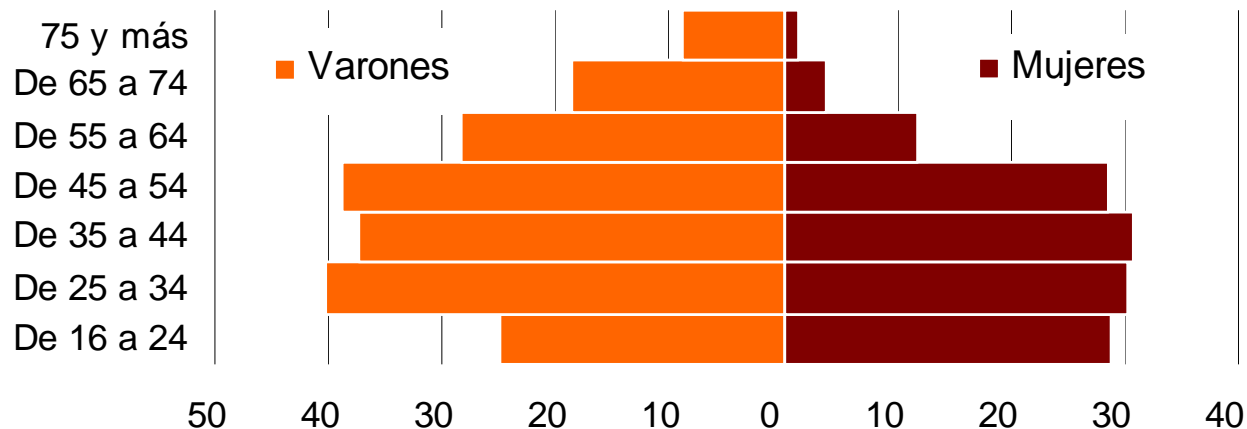
Fuente: Encuesta de Presupuestos Familiares. INE

•Bi-direccional:

Tiene orientación **horizontal** y contiene **dos series** de datos cuyas barras de frecuencias crecen en sentidos opuestos. Las más comunes son las pirámides de población de las que hablaremos más adelante.

Consumo de tabaco según sexo y grupos de edad

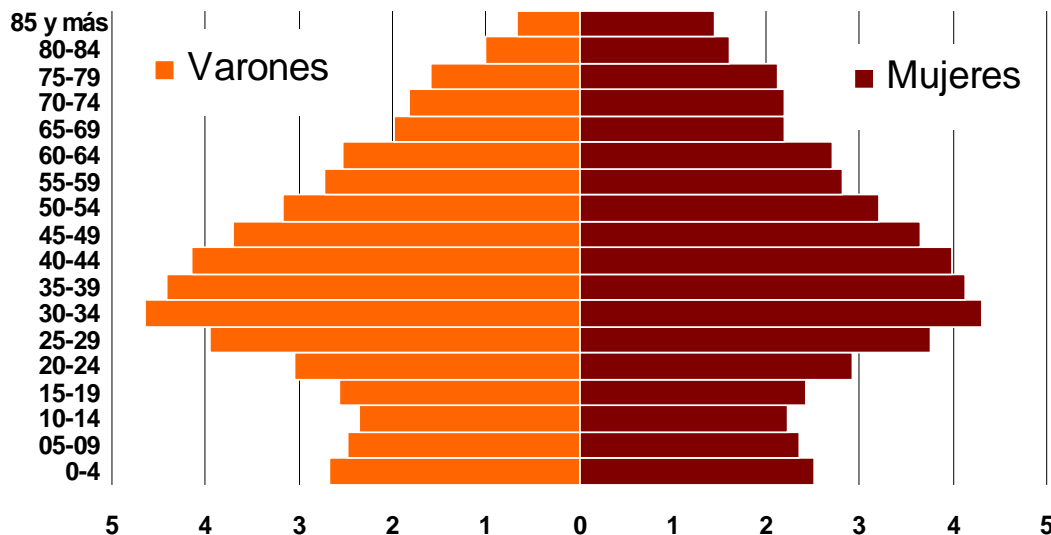
Fumadores diarios (porcentajes)



Fuente: Encuesta Nacional de Salud 2006. INE

Una pirámide de población es un **histograma bi-direccional** que muestra la **estructura demográfica de una población**, por sexo y edad, en un momento determinado.

Pirámide de la población española. 2009



Fuente: Explotación estadística del Padrón. INE

En el eje vertical se posicionan los **rangos de edades** y en el horizontal los **porcentajes de población**. En una de las direcciones se colocan las barras que representan la distribución (% sobre el total de población) por edad de los varones y en la otra la distribución por edad de las mujeres.

Estos gráficos nos dan una visión de la juventud, madurez o vejez de una población y, por tanto, su **grado de desarrollo**.

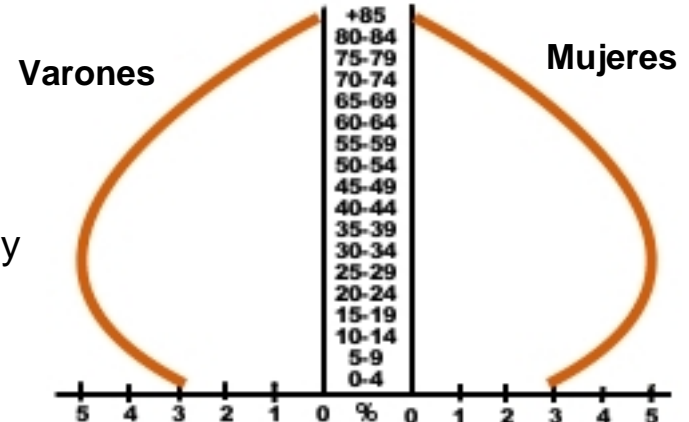
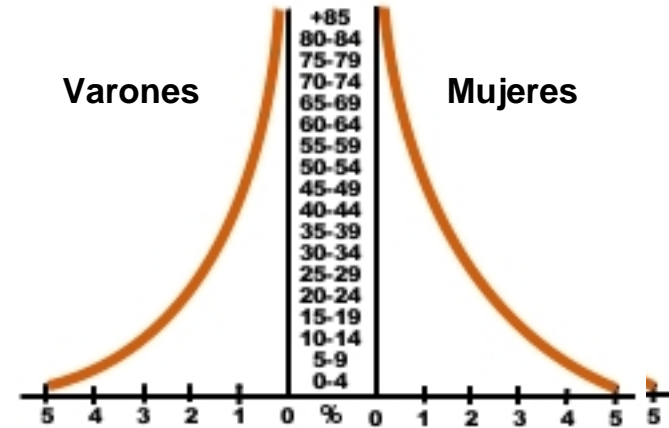
Según su forma puede haber distintos **tipos de pirámides**:

- **Progresiva:**

Existe un alto porcentaje de población joven que va desapareciendo según avanzan las edades. Son típicas de países subdesarrollados cuya esperanza de vida es baja y la tasa de natalidad alta.

- **Regresiva:**

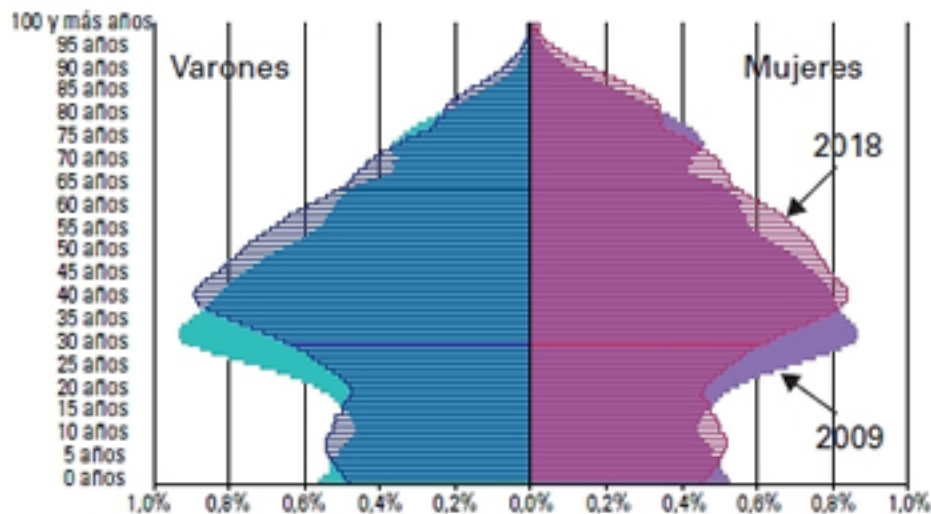
En la base existe menos población que en el medio y la población envejecida es considerable. Son típicas de países desarrollados cuya natalidad está descendiendo y la esperanza de vida es alta.



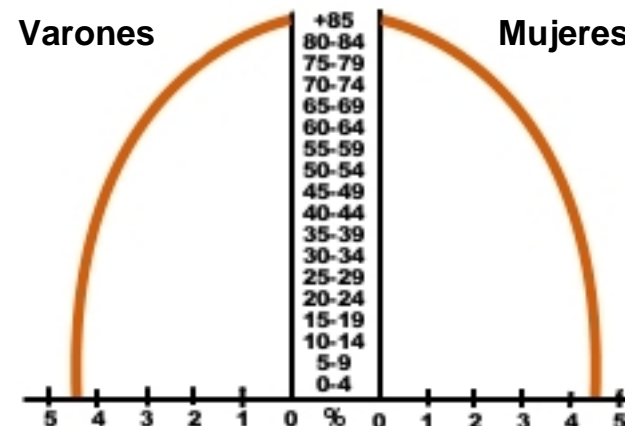
•Estancada:

Los tramos intermedios de edades tienen la misma población que la base. Son típicas de países en vías de desarrollo donde se ha controlado la mortalidad y se empieza a controlar la natalidad.

Pirámide de Población 2009 y 2018



Fuente: Explotación estadística del Padrón, estimaciones y proyecciones. INE



Nos podemos encontrar **varias series de datos** en un mismo gráfico, como por ejemplo, la población en distintos años o de distintos países.

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.

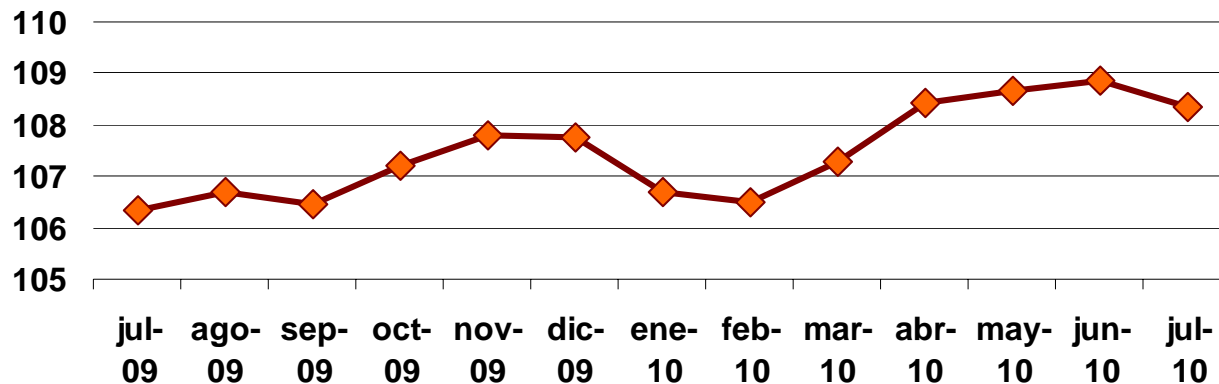
IPC 2009 - 2010

Base 2006

	<u>IPC</u>
jul-09	106,3
ago-09	106,7
sep-09	106,4
oct-09	107,2
nov-09	107,8
dic-09	107,8
ene-10	106,7
feb-10	106,5
mar-10	107,3
abr-10	108,4
may-10	108,7
jun-10	108,9
jul-10	108,4

Índice de precios al consumo

Base 2006



Fuente: IPC. INE

En cada eje se representa cada una de las variables cuya relación se quiere observar (en el ejemplo, en el eje horizontal los meses y en el eje vertical la media mensual del IPC correspondiente a esos meses).

Gráfico de líneas

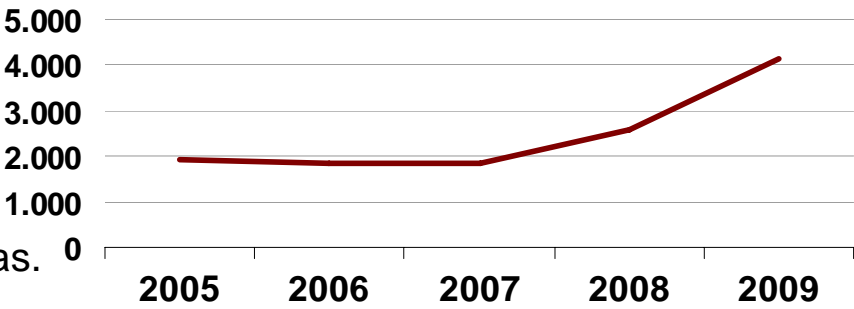
Se suelen usar para presentar **tendencias temporales**.

En el eje horizontal se ha de posicionar la variable que indica las unidades de tiempo y en el vertical se introduce la escala de la variable cuya variación en el tiempo queremos ver.

Pueden aparecer **varias variables** para compararlas.

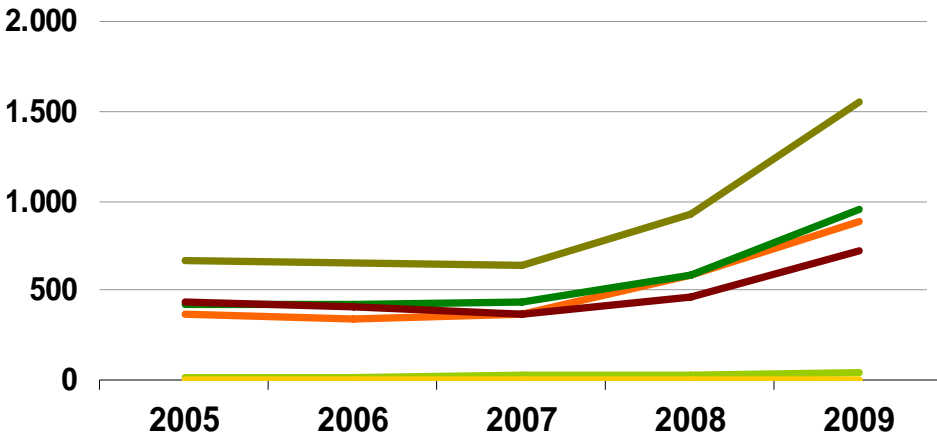
Parados en España

Miles de personas



Parados en España por nivel de formación alcanzada

Miles de personas



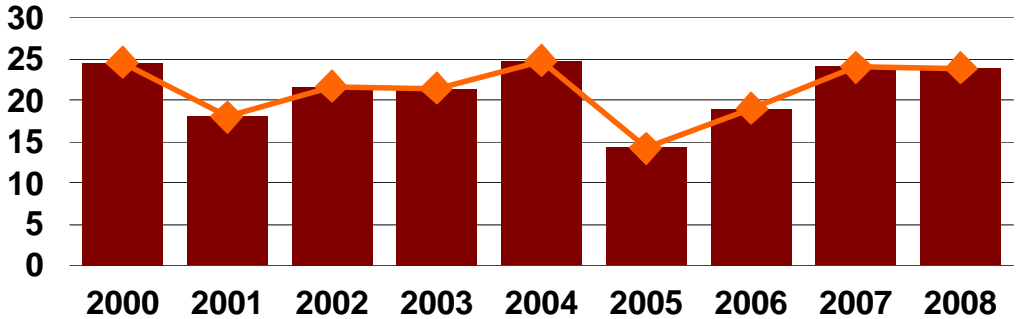
- Analfabetos
- Educación Primaria
- Educación Secundaria 1ª etapa y formación e inserción laboral correspondiente
- Educación Secundaria 2ª etapa y formación e inserción laboral correspondiente
- Educación superior, excepto doctorado
- Doctorado

Gráfico de líneas

Casos particulares:

- Si se unen los puntos medios de las bases superiores de las barras en los gráficos de barra se obtiene el **polígono de frecuencias**.

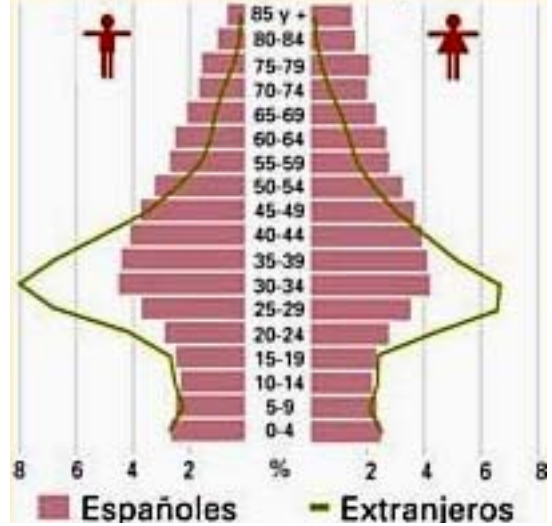
Producción de cereales en España
Millones de toneladas métricas



Fuente: Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO)

- Las **pirámides de población** se pueden presentar como un gráfico de líneas. Es útil cuando se quieren presentar varias series de datos, por ejemplo, si se quieren comparar pirámides de varios países, ver la población extranjera o comparar la estructura de población de varios años.

Pirámide de población de España 2010



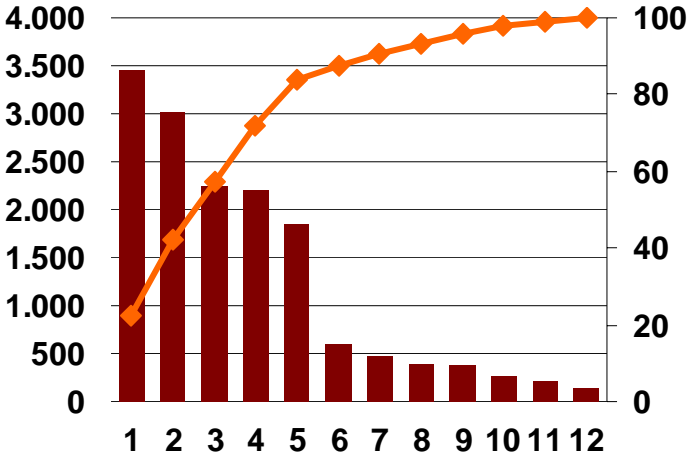
Fuente: Explotación estadística del Padrón. INE

Gráfico de Pareto

Un gráfico de Pareto es un tipo de gráfico de barras vertical ordenado por frecuencias de forma descendente que **identifica y da un orden de prioridad a los datos.**

<u>Causas externas de mortalidad</u>	<u>Defunciones</u>	<u>%</u>	<u>% acumulados</u>
1 Suicidio y lesiones autoinfligidas	3.457	23	23
2 Accidentes de tráfico de vehículos de motor	3.030	20	43
3 Ahogamiento, sumersión y sofocación accidentales	2.258	15	58
4 Otros accidentes	2.212	14	72
5 Caídas accidentales	1.845	12	84
6 Envenenamiento accidental por psicofármacos y drogas de abuso	599	4	88
7 Otros accidentes de transporte	476	3	91
8 Complicaciones de la atención médica y quirúrgica	404	3	94
9 Agresiones (Homicidios)	378	2	96
10 Otros envenenamientos accidentales	281	2	98
11 Accidentes por fuego, humo y sustancias calientes	202	1	99
12 Otras causas externas y sus efectos tardíos	147	1	100
Total	15.289	100	

Fuente: INE



En el eje horizontal se representan las categorías de la variable que queremos estudiar (diferentes causas externas de mortalidad). En el eje vertical derecho se muestra la escala de porcentajes y en el eje vertical izquierdo la escala de frecuencias (número de defunciones).

Las barras muestran las frecuencias de las categorías de la variable y la línea representa el porcentaje acumulado de dichas frecuencias respecto el total.

Este gráfico muestra el **Principio de Pareto**, formulado por el estadístico italiano Vilfredo Pareto (1848-1923) :

pocos vitales, muchos triviales

Es decir, hay muchos problemas sin importancia frente a unos pocos graves ya que, por lo general, el 20% de las causas totales hacen que sean originados el 80% de los efectos.

Pareto formuló este Principio tras un estudio sobre la distribución de la riqueza con el que estableció que la desigualdad económica es inevitable en cualquier sociedad.

En el ejemplo anterior, casi el 80% de las defunciones por causas externas de mortalidad (el 72%), son debidas a 4 de las 12 causas detalladas (el 33%).

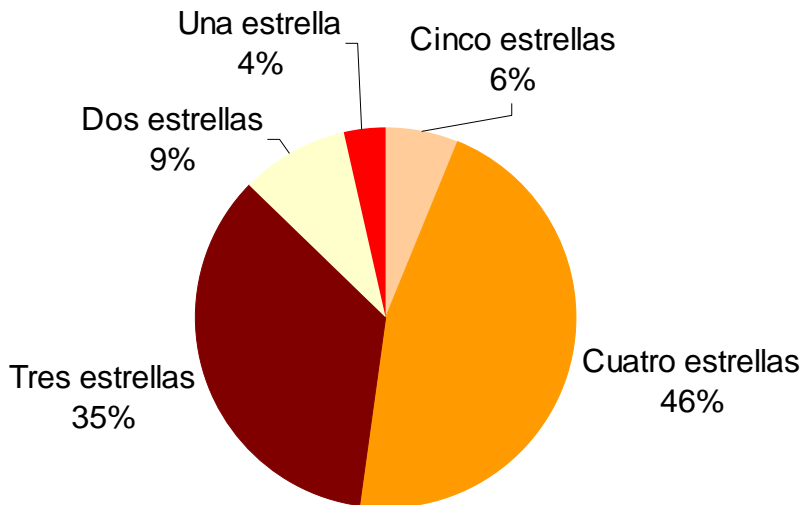
Un gráfico de sectores es una representación circular de las **frecuencias relativas de una variable cualitativa o discreta** que permite, de una manera sencilla y rápida, su comparación.

Alojamientos Turísticos. 2009

<u>Categoría</u>	<u>Número de viajeros</u>
Total	69.152.754
Cinco estrellas	4.216.253
Cuatro estrellas	31.960.442
Tres estrellas	24.079.125
Dos estrellas	6.331.715
Una estrella	2.565.219

Fuente: Encuesta de Ocupación en Alojamientos Turísticos

Viajeros hospedados en hoteles españoles por categoría del establecimiento. 2009



El **círculo** representa la **totalidad** que se quiere observar (en el ejemplo, total de viajeros hospedados en hoteles) y cada porción, llamadas **sectores**, representan la **proporción de cada categoría** de la variable (en el ejemplo, tipo de hotel) respecto el total. Suele expresarse en **porcentajes**.

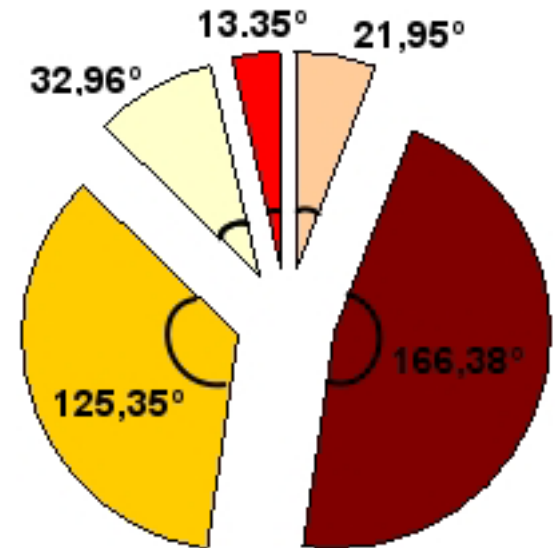
Obtención de los **ángulos** de cada sector:

<u>Categoría</u>	<u>Frecuencia absoluta</u>	<u>Frecuencia relativa</u>	<u>Porcentaje (%)</u>
Cinco estrellas	4.216.253	0,06	6
Cuatro estrellas	31.960.442	0,46	46
Tres estrellas	24.079.125	0,35	35
Dos estrellas	6.331.715	0,09	9
Una estrella	2.565.219	0,04	4
<u>Total</u>	<u>69.152.754</u>	<u>1,00</u>	<u>100</u>

$$\text{Frecuencia relativa} = \frac{\text{Frecuencia absoluta}}{\text{Suma de todas las frecuencias absolutas}}$$

$$\text{Porcentaje} = \text{Frecuencia relativa} \times 100$$

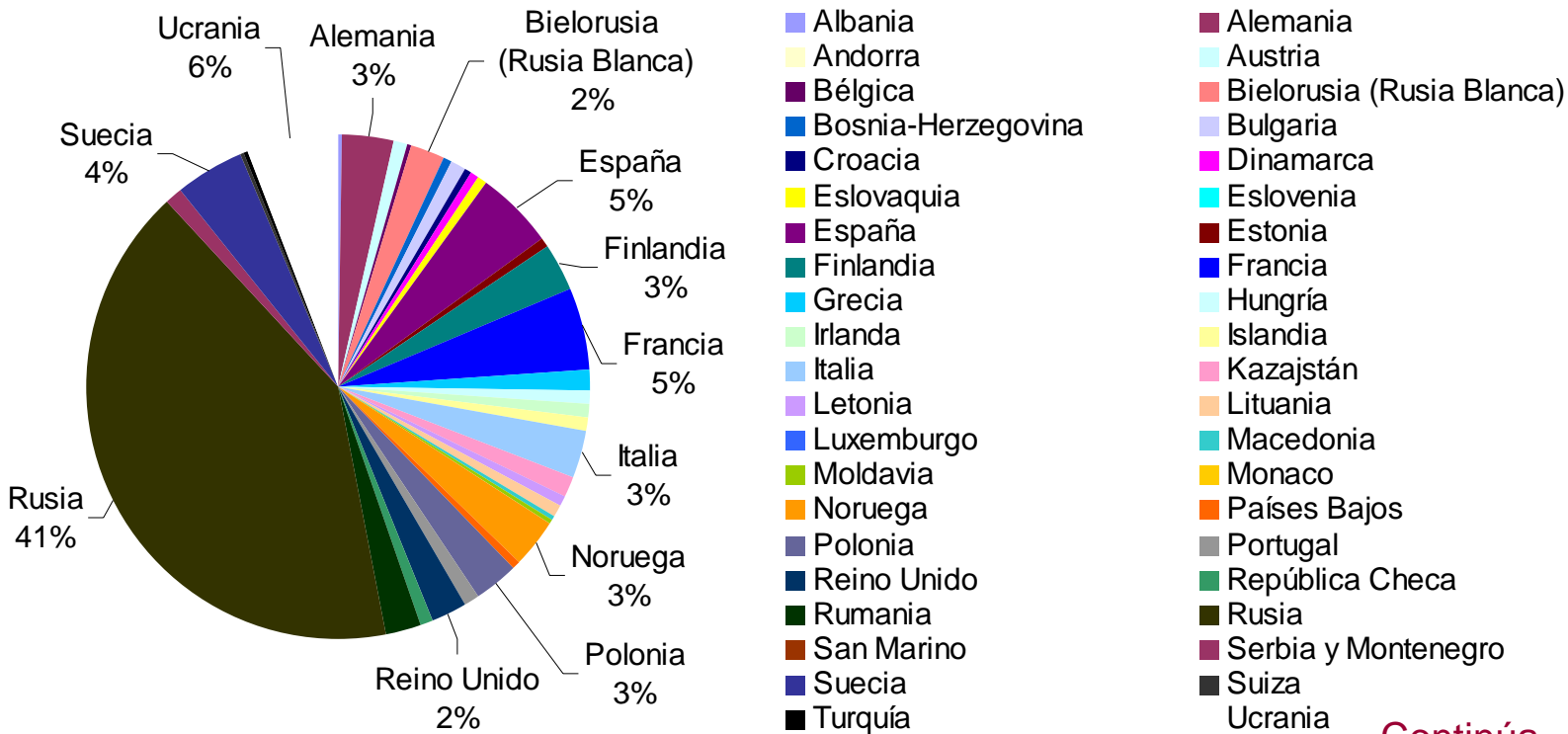
$$\text{Ángulo} = \text{Frecuencia relativa} \times 360 \text{ (grados que tiene una circunferencia)}$$



Son útiles cuando las **categorías son pocas**.

Si el gráfico tuviera muchas variables, no aportaría casi información y sería prácticamente incomprensible, como ocurre en el siguiente ejemplo.

Proporción de superficie de los países europeos



Pictograma

Un pictograma es un gráfico que representa mediante **figuras o símbolos** las **frecuencias de una variable cualitativa o discreta**.

Al igual que los gráficos de barras suelen usarse para **comparar magnitudes** o ver la **evolución en el tiempo** de una categoría concreta.

Tipos de pictogramas:

- **Gráficas de barras** cuyas barras están constituidas por **símbolos o figuras distorsionadas** que se adaptan a la longitud de la barra.



Fuente: Encuesta de Hogar y Medio ambiente

Icono obtenido de la web del ayuntamiento de Madrid área Medio ambiente

Porcentaje de viviendas cuyos residentes depositan residuos en su punto especificado de recogida. 2008

- **Gráficas de barras** cuyas barras están constituidas por **símbolos o figuras del mismo tamaño** que representan una cantidad específica (a mayor frecuencia, más acumulación de figuras).



- Símbolo o figuras cuyas **áreas** representan las frecuencias de cada categoría de la variable.

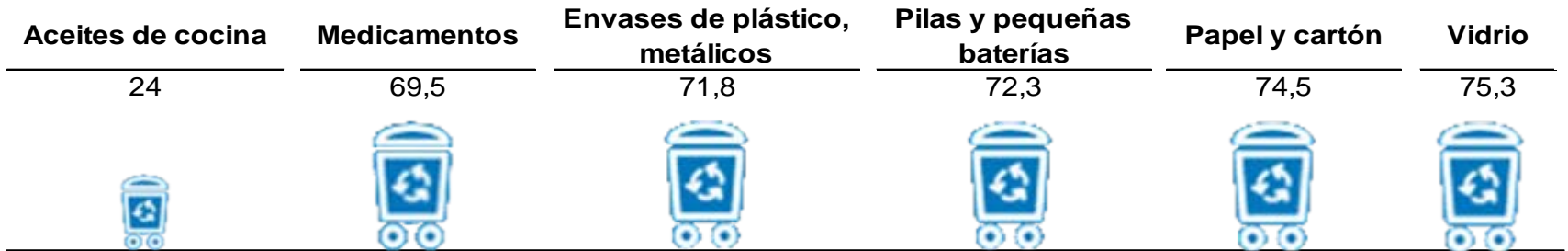


Gráfico de dispersión

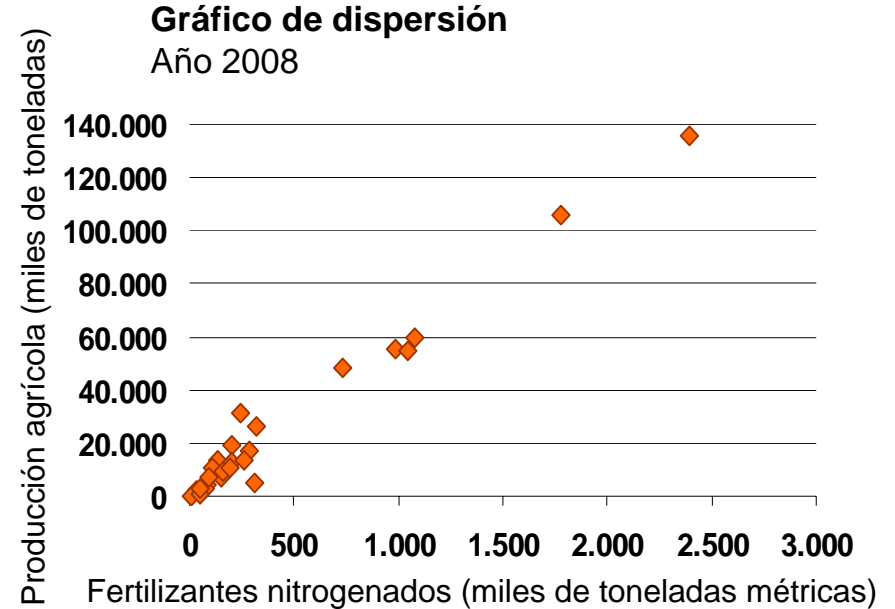
Un gráfico de dispersión muestra en un eje cartesiano la **relación** que existe entre dos variables.

País	Fertilizantes Nitrogenados*	Produccion agrícola**
Alemania	1.779.455	106.041
Austria	110.512	10.407
Bulgaria	138.464	13.294
Chipre	7.982	239
Dinamarca	205.829	19.361
Eslovaquia	91.673	7.350
Eslovenia	53.907	1.051
España	983.409	55.739
Estonia	33.500	1.830
Finlandia	161.159	9.186
Francia	2.395.024	135.819
Grecia	201.549	12.507
Hungría	316.520	26.130
Irlanda	312.655	5.133
Italia	729.420	48.634
Letonia	83.301	4.045
Lituania	153.797	7.309
Países Bajos	261.493	13.255
Polonia	1.080.261	59.860
Portugal	76.250	2.694
Reino Unido	1.048.100	54.754
República Checa	289.063	16.840
Rumanía	240.880	31.273
Suecia	190.838	11.012
Suiza	50.844	3.111

* Tonelada métrica de elemento nutritivo consumido

** Miles de toneladas

Fuente: Eurostat; datos referidos a 2008

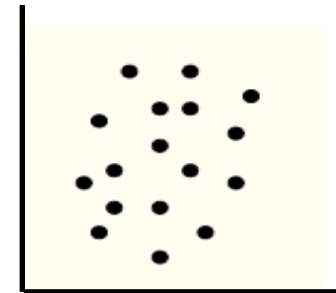


Este gráfico nos informa del **grado de correlación** entre las dos variables, es decir, nos muestra si el incremento o disminución de los valores de una de las variables, denominada variable independiente y que se suele representar en el eje horizontal, altera de alguna manera los valores de la otra, denominada variable dependiente y que representa generalmente en el eje vertical. En el ejemplo, el consumo de fertilizantes nitrogenados de cada país está representado en el eje horizontal y la producción agrícola en el vertical.

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

- **Correlación nula:**

No existe ninguna relación entre las variables. Se dice que ambas son **independientes**.



- **Correlación lineal:**

Existe una relación lineal **negativa** si al aumentar los valores de la variable independiente disminuyen los valores de la variable dependiente y relación lineal **positiva** si al aumentar los valores de la variable independiente aumentan los valores de la variable dependiente.

- **Correlación no lineal:**

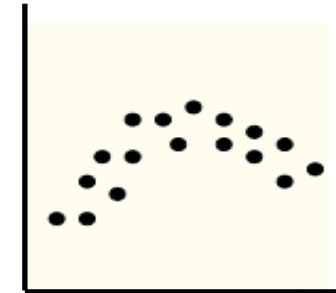
Existe una relación entre las variables pero no es lineal.



negativa

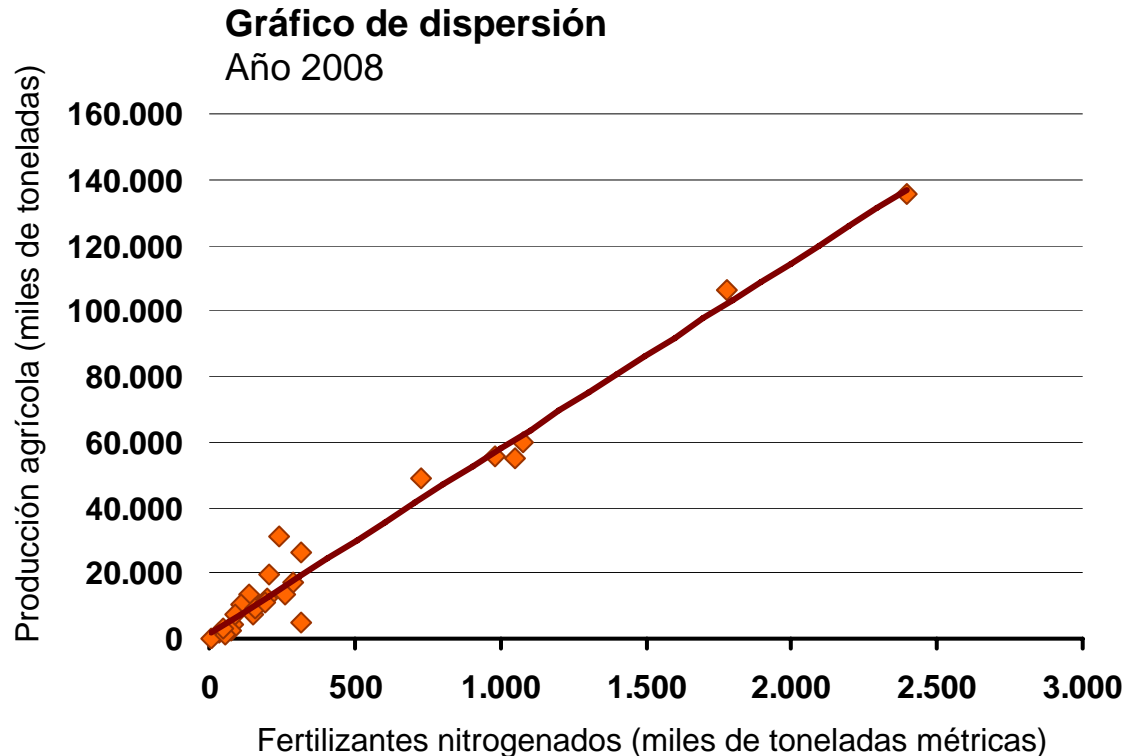


positiva



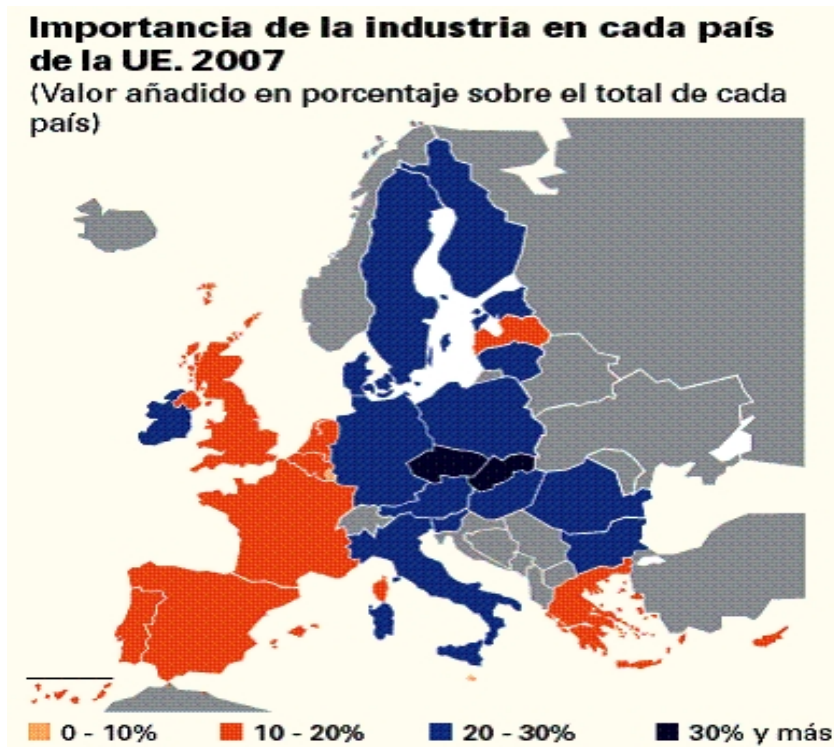
Dentro de la correlación lineal, en algunos gráficos se dibuja la **recta de regresión** obtenida gracias a un método matemático: el ajuste lineal.

Esta recta nos muestra más clara y rápidamente la tendencia de los puntos y, por tanto, la relación.



En el ejemplo se puede apreciar que los países con más producción agrícola consumen más fertilizantes nitrogenados. Se observa en principio una relación lineal positiva.

Un cartograma es un mapa en el que se presentan datos estadísticos por regiones bien poniendo el número o coloreando las distintas zonas en función del dato que representan.



Fuente: Panorámica de la industria. INE

Tasa de paro. 2009



Fuente: España en cifras. INE

- Instituto de Tecnologías Educativas. Ministerio de Educación
- Instituto Nacional de Estadística. Portugal
- Instituto Nacional de Estadística. España