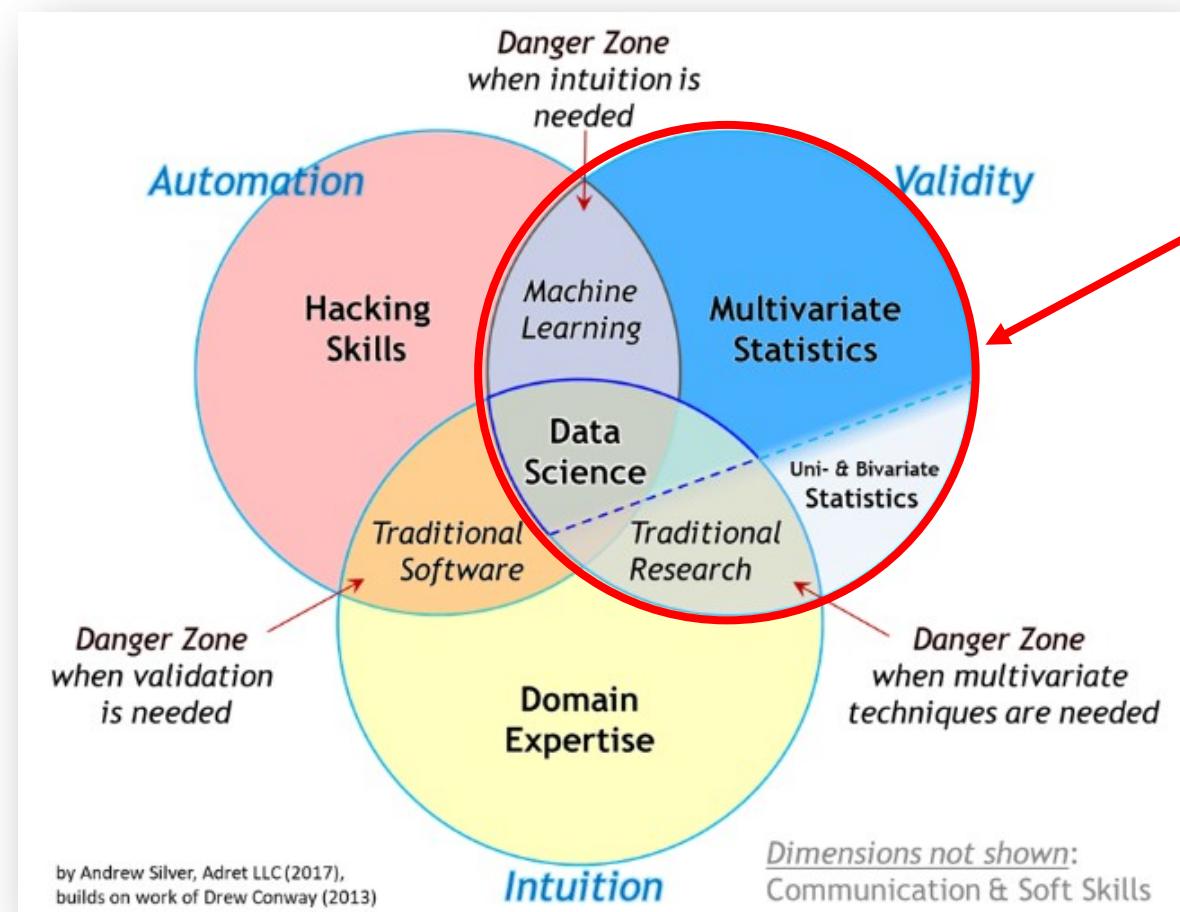


Estadística descriptiva y regresión lineal con Excel



Estadística Descriptiva y Regresión con Excel

OBJETIVO: Dominar el análisis de datos desde los fundamentos hasta técnicas avanzadas de regression

01

Fundamentos Estadísticos

Domina los conceptos esenciales: introducción a la estadística, identificación de fuentes de datos confiables, comprensión de conceptos fundamentales y clasificación de tipos de variables.

03

Visualización de Datos

Crea tablas de frecuencia profesionales y gráficos impactantes: barras, sectores e histogramas que comuniquen tus hallazgos con claridad.

05

Dispersión y Variabilidad

Evaluá la variabilidad mediante rango, varianza, desviación típica y coeficiente de variación para entender la distribución completa de tus datos.

07

Regresión Lineal

Calcula la recta de regresión, determina el coeficiente de correlación e interpreta resultados para hacer predicciones fundamentadas.

02

Escalas de Medición

Aprende a distinguir entre escalas nominal, ordinal, de intervalo y de razón. Conoce sus aplicaciones prácticas y ejemplos del mundo real.

04

Análisis Numérico

Calcula e interpreta medidas de tendencia central (media, mediana, moda) y medidas de posición (cuartiles, percentiles) para comprender tus datos.

06

Análisis Bivariante

Explora relaciones entre variables con tablas de contingencia y técnicas de análisis conjunto para descubrir patrones ocultos.

08

Modelos Avanzados

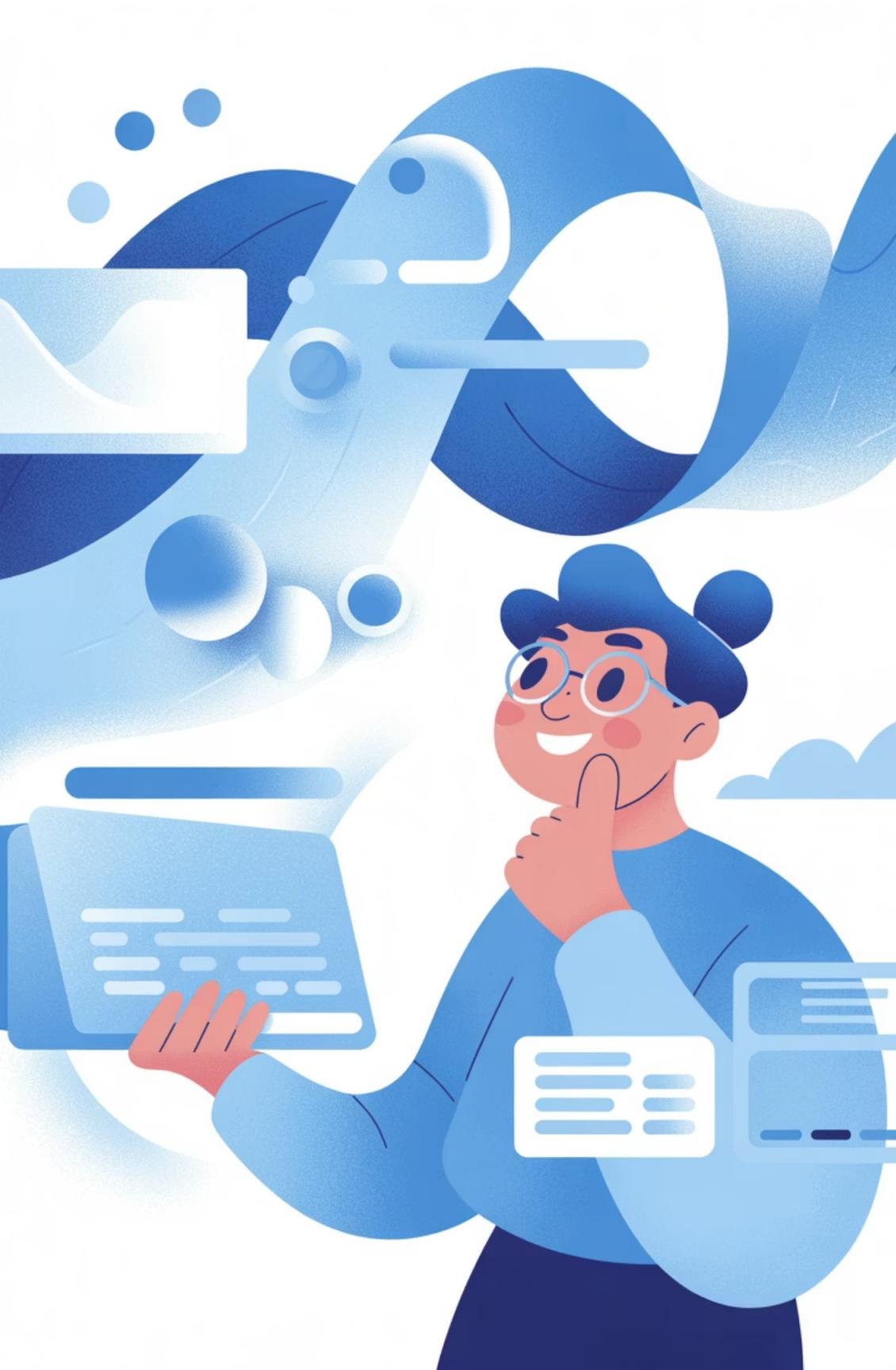
Ajusta curvas no lineales y aplica modelos exponenciales, logarítmicos y polinómicos para análisis sofisticados de datos complejos.

¿Por qué Excel para aprender los conceptos imprescindibles de Estadística?

- Funciones estadísticas integradas y fáciles de usar
- Visualización de datos instantánea y profesional
- Compatible con grandes volúmenes de información
- Herramientas de análisis de regresión incorporadas

Fundamentos Estadísticos

Presentación Módulo AADD (IES Trassierra) © 2025 por Elena Fernández Chirino tiene licencia CC BY-NC-ND 4.0.
Para ver una copia de esta licencia, visite <https://creativecommons.org/licenses/by-nc-nd/4.0/>



¿Qué es la Estadística?

Ciencia del sentido

Extrae conclusiones significativas de conjuntos de datos numéricos

Datos útiles

Observaciones y medidas: tasas de paro, ingresos, notas, resultados electorales

Conocimiento aplicado

Describe fenómenos, extrae conclusiones y facilita la toma de decisiones

La recopilación y el formato de los datos son el comienzo indispensable para utilizarlos en la solución de problemas

Fuentes de Datos Estadísticos

1

Fuentes primarias

Instituciones que recopilan datos directamente

- Oficinas gubernamentales
- Instituciones de prestigio
- Mayor fiabilidad y garantías



Recomendación: Utiliza siempre fuentes primarias cuando sea posible para garantizar la calidad de tus análisis

2

Fuentes secundarias

Publicaciones y trabajos derivados

- Elaborados por terceros
- Información reinterpretada
- Requieren verificación

3

Investigación propia

Cuando no existe información disponible

- Observación directa
- Experimentación controlada
- Encuestas diseñadas

Principales fuentes primarias en España y Europa



INE e IAEST

Instituto Nacional de Estadística
y servicios autonómicos



Banco de España

Datos económicos y
financieros nacionales



EUROSTAT y BCE

Estadísticas europeas y del
Banco Central



Organismos internacionales

Census Bureau, World Bank, OCDE

Conceptos Fundamentales

01

Población

Conjunto completo de elementos objeto de estudio, definidos sin ambigüedad

02

Individuo

Cada elemento de la población: personas, empresas, provincias, edificios...

03

Variables

Características o comportamientos que queremos estudiar de los individuos

04

Dato estadístico

Observación o medida de la intensidad con la que se verifica un carácter

De datos a conocimiento



Datos

Valores en bruto sin orientación

Ejemplo: 1,8

Información

Datos con estructura y unidades

Ejemplo: 1,8 m

Conocimiento

Interpretación en contexto específico

Ejemplo: 1,8 m de altura para una mujer

Censo

Recopilación de información de **todos** los individuos de la población

Muestra

Selección de una **parte significativa** de la población para el estudio

Fundamentos Estadísticos

Presentación Módulo AADD (IES Trassierra) © 2025 por Elena Fernández Chirino tiene licencia CC BY-NC-ND 4.0.
Para ver una copia de esta licencia, visite <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Tipos de Variables

Las técnicas estadísticas y la validez de las conclusiones dependen del tipo de variable que manejemos

Variables Cualitativas

Reflejan **cualidades** mediante categorías o modalidades

- No necesariamente numéricas
- Categorías bien definidas
- Mutuamente excluyentes
- Todos los individuos incluidos

Ejemplos: nivel de instrucción, estado civil, profesión

Variables Cuantitativas

Representan **cantidad** mediante números con significado aritmético

Discretas

- Valores finitos o numerables
- Número de estudiantes
- Miembros de familia
- Precios de acciones

Continuas

- Valores infinitos en intervalo
- Estatura, peso
- Temperatura
- Concentraciones



Nota práctica: Las variables continuas suelen medirse de forma discreta debido a limitaciones de los instrumentos de medición

Introducción a las Escalas de Medida

La cuantificación de una observación se puede realizar de acuerdo con diferentes reglas, lo que da lugar a distintas escalas de medida. La escala de medida es la referencia, numérica o no, que se utiliza para medir el valor u observar la categoría que corresponde a una característica de un individuo.

01

Escala Nominal

Clasificación en categorías sin orden establecido

Ejemplos: Comunidad Autónoma de residencia, tipo de grupo sanguíneo, estado civil, raza, religión, nacionalidad, sexo.

02

Escala Ordinal

Categorías con orden pero sin distancia cuantificable

Ejemplos: Dureza de un mineral, estatus socioeconómico, ranking de vinos, grado de satisfacción con un producto, nivel de estudios, clasificación de la liga de fútbol.

03

Escala de Intervalo

Unidad de medida con distancias cuantificables

Ejemplos: Temperatura en grados centígrados, puntuación en test de inteligencia, fechas de calendario.

04

Escala de Razón

Punto cero absoluto con significado de ausencia

Ejemplos: altura, peso, total de ventas, número de hijos en una familia, distancia geográfica, etc.



Tipos de Datos y Estadística

Clasificación según Temporalidad

1

Datos Transversales o de Sección Cruzada

Se refieren a distintos individuos en un determinado período o instante de tiempo.

Ejemplo: número de empleados de pequeñas empresas a 31 de diciembre de 2006.

2

Series Temporales

Están referidos a un mismo individuo en distintos períodos o instantes de tiempo.

Ejemplo: número de empleados de una empresa desde 1970 hasta 2006.

3

Panel de Datos

Conjuga la transversalidad y la temporalidad, siendo una combinación de series temporales y datos de sección cruzada.

Ejemplo: número de empleados de cada empresa en 36 años.

Estadística Univariante

Tratamiento de una única característica de los individuos de una población.

Estadística Multivariante

Tratamiento de dos o más características simultáneamente.

Tabulación y Representación Gráfica de Datos Univariantes

Presentación Módulo AADD (IES Trassierra) © 2025 por Elena Fernández Chirino tiene licencia CC BY-NC-ND 4.0.
Para ver una copia de esta licencia, visite <https://creativecommons.org/licenses/by-nc-nd/4.0/>

¿Qué es la Tabulación?

Denominamos **tabulación** al proceso de ordenación y agrupamiento de un conjunto de datos. Como resultado del mismo se obtienen las **distribuciones de frecuencias** de las observaciones.

En lo que sigue, emplearemos la siguiente **notación básica**:

- Tamaño de la población o muestra: N
- Número de modalidades: k
- Valores de la variable: $\{x_1, x_2, \dots, x_k\}$
- Supuesto (cuando tenga sentido): $x_1 < x_2 < \dots < x_k$

Frecuencias

Sea X una variable o característica medida en una población de tamaño N que toma los valores $\{x_i; i = 1 \dots k\}$.

La **Frecuencia Absoluta** se define como el número de veces que se observa o repite el valor x_i , y la denotamos por n_i . Se cumple que:

$$N = \sum_{i=1}^k n_i$$

La **Frecuencia Relativa** se define como el porcentaje de veces que se observa el valor x_i , y la denotamos por f_i : $f_i = n_i/N$. Se cumple que:

$$\sum_{i=1}^k f_i = 1$$

Frecuencias acumuladas

Las frecuencias absolutas y relativas se pueden obtener siempre, independientemente del tipo de variable que estemos manejando. Para aquellos datos que presentan alguna ordenación, es decir, todos excepto los de tipo nominal, podemos obtener las **frecuencias acumuladas**.

La **Frecuencia Absoluta Acumulada** se define como el número total de observaciones menores o iguales que x_i , y la denotamos por N_i :

$$N_i = \sum_{j=1}^i n_j$$

Se cumple que $N_k = N$

La **Frecuencia Relativa Acumulada** se define como el porcentaje de observaciones que son menores o iguales que x_i , y la denotamos por F_i :

$$F_i = \sum_{j=1}^i f_j \quad \text{o} \quad F_i = \frac{N_i}{N}$$



Tipos de Frecuencias

Frecuencia Absoluta (n_i)

Número de veces que se observa el valor x_i

$$N = \sum_{i=1}^k n_i$$

Frecuencia Relativa (f_i)

Porcentaje de veces que aparece x_i : $f_i = n_i/N$

$$\sum_{i=1}^k f_i = 1$$

Frecuencias Acumuladas

Absoluta (N_i): observaciones $\leq x$

Relativa (F_i): porcentaje $\leq x_i$

Distribuciones de Frecuencias según Tipo de Variable

01

Variables Cualitativas Nominales

Solo frecuencias absolutas (n_i) y relativas (f_i). No tienen orden, por lo que no se calculan frecuencias acumuladas.

02

Variables Ordinales y Discretas

Se calculan todas las frecuencias: absolutas, relativas y sus correspondientes acumuladas (N_i, F_i).

03

Variables Continuas Agrupadas

Datos agrupados en intervalos ($L_{i-1}, L_i]$ con marca de clase x_i , amplitud a_i y densidad $d_i = n_i/a_i$.

Distribuciones de Frecuencias

01

Variables Cualitativas Nominales

Solo frecuencias absolutas (n_i) y relativas (f_i). No tienen orden, por lo que no se calculan frecuencias acumuladas.

X	n_i	f_i
x_1	n_1	$f_1 = \frac{n_1}{N}$
x_2	n_2	$f_2 = \frac{n_2}{N}$
:	:	:
x_k	n_k	$f_k = \frac{n_k}{N}$
\sum	N	1

Distribuciones de Frecuencias

02

Variables Ordinales y Discretas

Se calculan todas las frecuencias:
absolutas, relativas y sus
correspondientes acumuladas (N_i , F_i).

X	n_i	f_i	N_i	F_i
x_1	n_1	$f_1 = \frac{n_1}{N}$	$N_1 = n_1$	$F_1 = f_1 \text{ ó } F_1 = \frac{N_1}{N}$
x_2	n_2	$f_2 = \frac{n_2}{N}$	$N_2 = n_1 + n_2$	$F_2 = f_1 + f_2 \text{ ó } F_2 = \frac{N_2}{N}$
\vdots	\vdots	\vdots	\vdots	\vdots
x_k	n_k	$f_k = \frac{n_k}{N}$	$N_k = n_1 + \dots + n_k = N$	$F_k = f_1 + \dots + f_k = 1 \text{ ó }$ $F_k = \frac{N_k}{N} = 1$
\sum	N	1		

Distribuciones de Frecuencias

Variables Continuas Agrupadas

Datos agrupados en intervalos $(L_{i-1}, L_i]$

marca de clase:

$$x_i = \frac{L_{i-1} + L_i}{2}$$

En el caso de que los intervalos no sean todos de la misma amplitud, se debe proceder a calcular una nueva característica de la población, su **densidad de frecuencia**, que puede ser diferente para cada intervalo, y se define como el cociente entre la frecuencia absoluta de cada intervalo y su correspondiente **amplitud**: $a_i = L_i - L_{i-1}$.

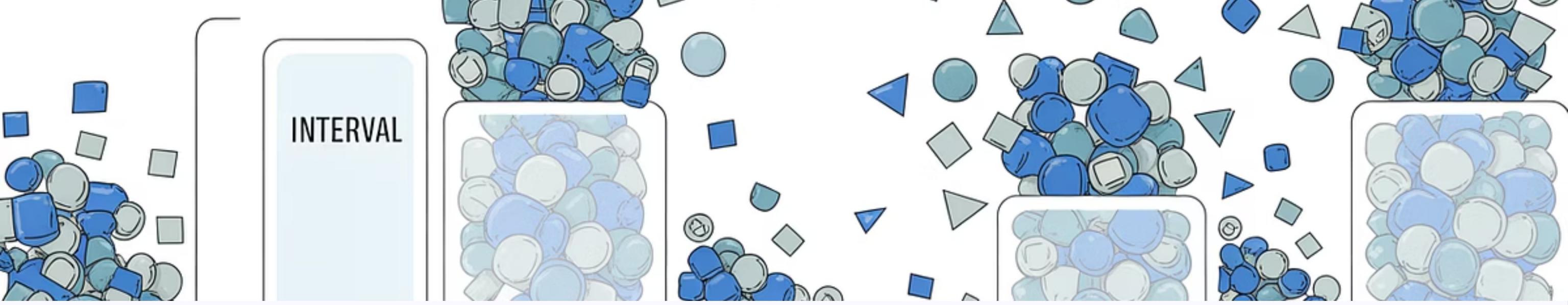
$$d_i = \frac{n_i}{a_i}$$

También es posible calcular la densidad empleando la frecuencia relativa:

$$d_i = \frac{f_i}{a_i}$$

Con todo ello podemos obtener la **distribución de frecuencias agrupadas**:

Intervalo	Marca clase	Amplitud	Densidad	Frecuencias			
				n_i	f_i	N_i	F_i
$(L_{i-1}, L_i]$	x_i	a_i	d_i				
$(L_0, L_1]$	x_1	a_1	d_1	n_1	f_1	N_1	F_1
$(L_1, L_2]$	x_2	a_2	d_2	n_2	f_2	N_2	F_2
...
$(L_{k-1}, L_k]$	x_k	a_k	d_k	n_k	f_k	N_k	F_k
<i>Totales</i>				N		1	



Representación Gráfica de Datos

Variables Cualitativas

Cartogramas, pictogramas, diagramas de sectores y diagramas de barras.

Variables Discretas

Diagrama de barras para datos no agrupados.

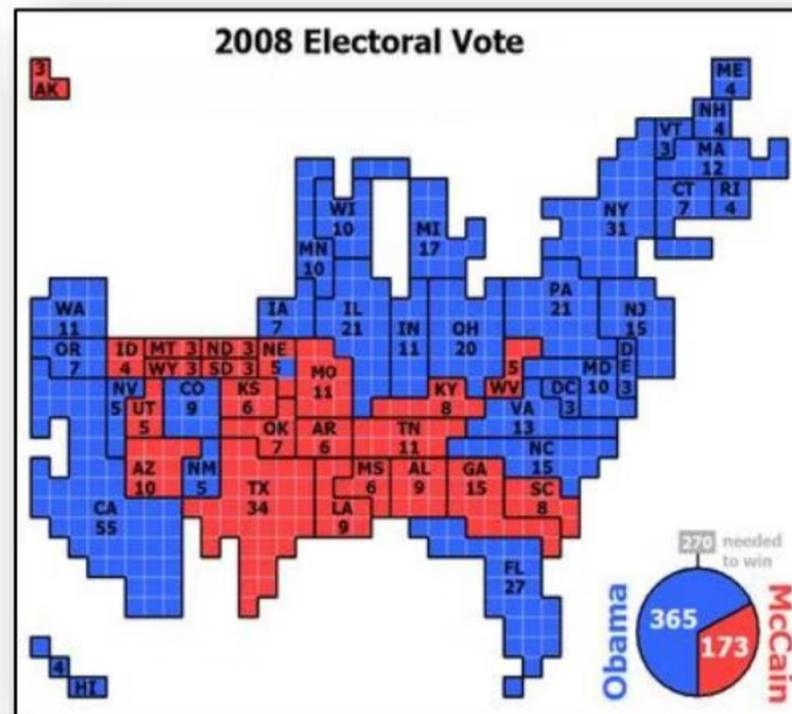
Variables Continuas

Histogramas y polígonos de frecuencias para datos agrupados.

Gráficos para Variables Cualitativas

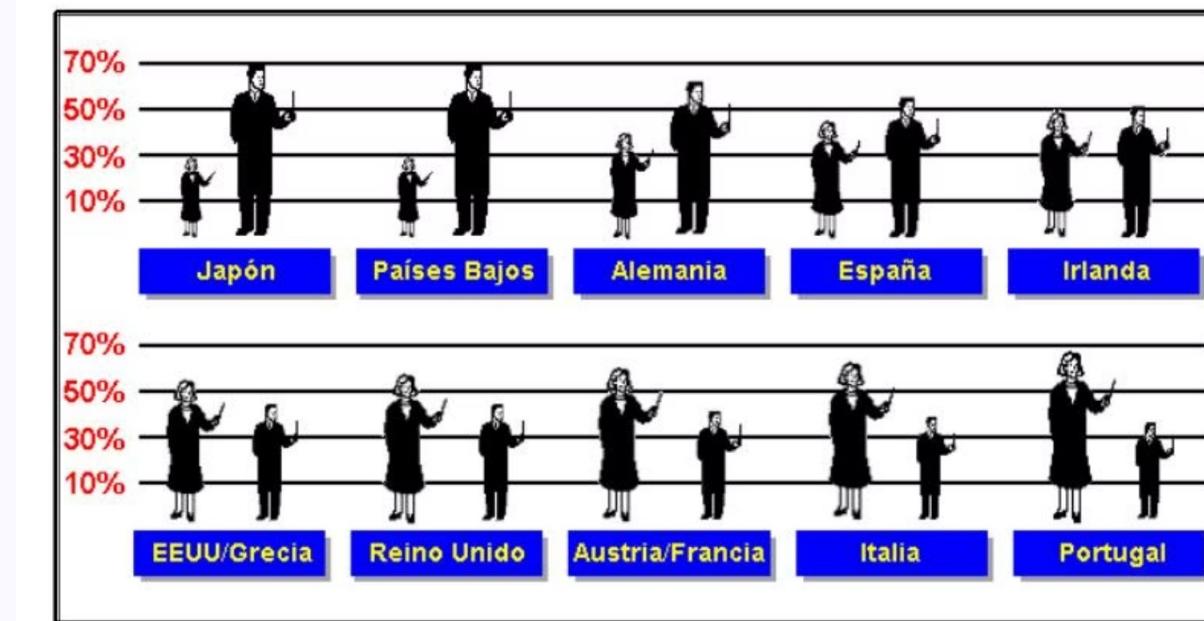
Cartogramas

Representación geográfica de datos sobre mapas.



Pictogramas

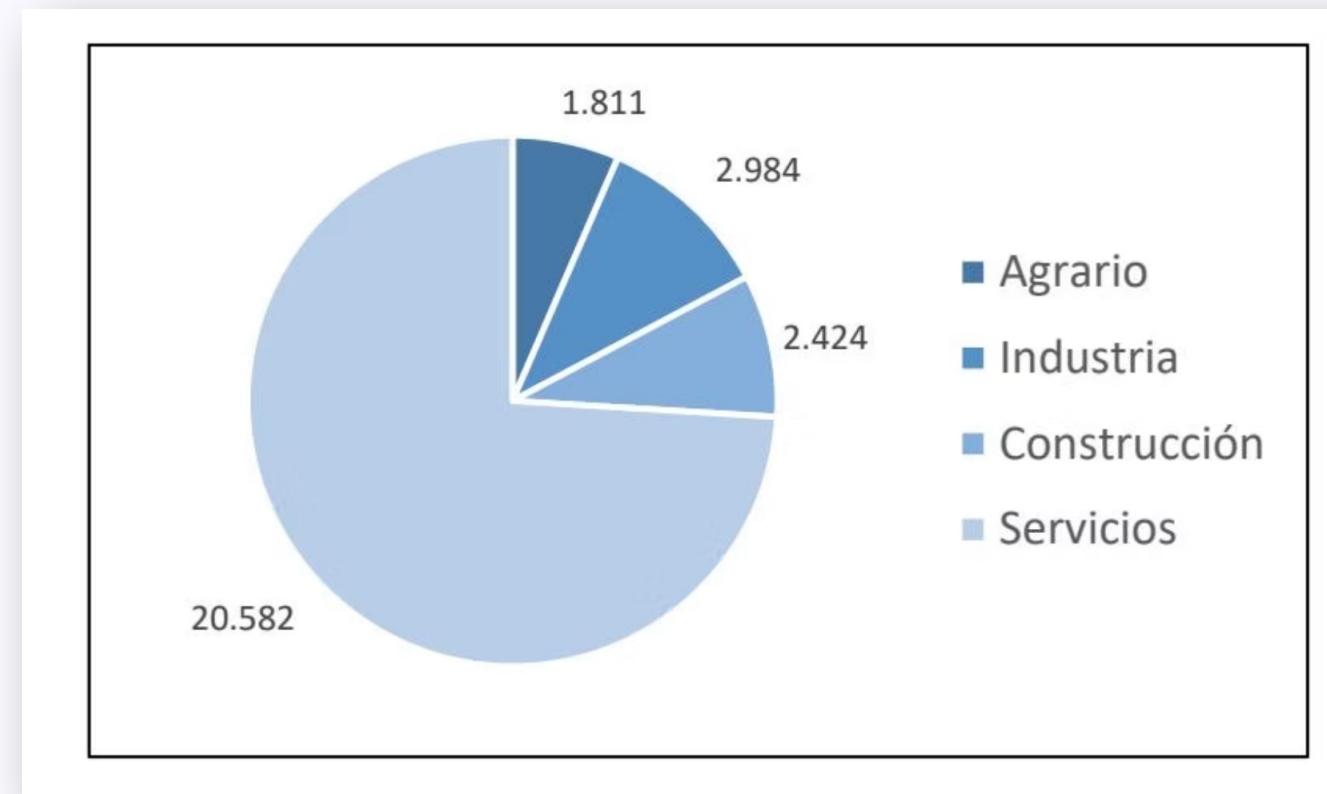
Uso de símbolos o iconos para representar cantidades.



Diagramas de Sectores

Diagrama de Sectores

Especialmente indicado para variables cualitativas nominales. Cada sector representa una proporción del total.



Diagramas de Barras

Sobre un par de ejes se representan en el eje de las abscisas las diferentes modalidades de la variable y se levanta, para cada una de ellas, una barra con una altura proporcional a la frecuencia (absoluta o relativa) de dicha modalidad.

Diagrama de Barras variable cualitativa

En el caso de variables nominales, el eje de abscisas representa cada una de las modalidades, sin importar el orden seguido.

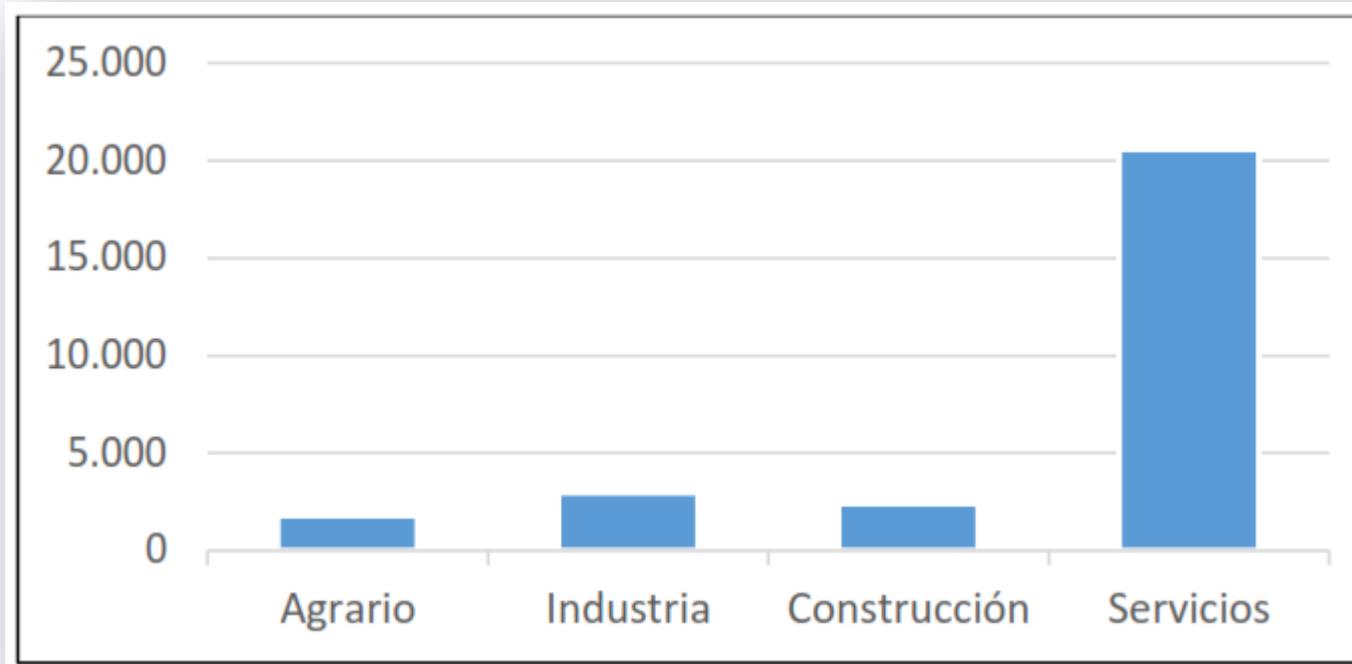
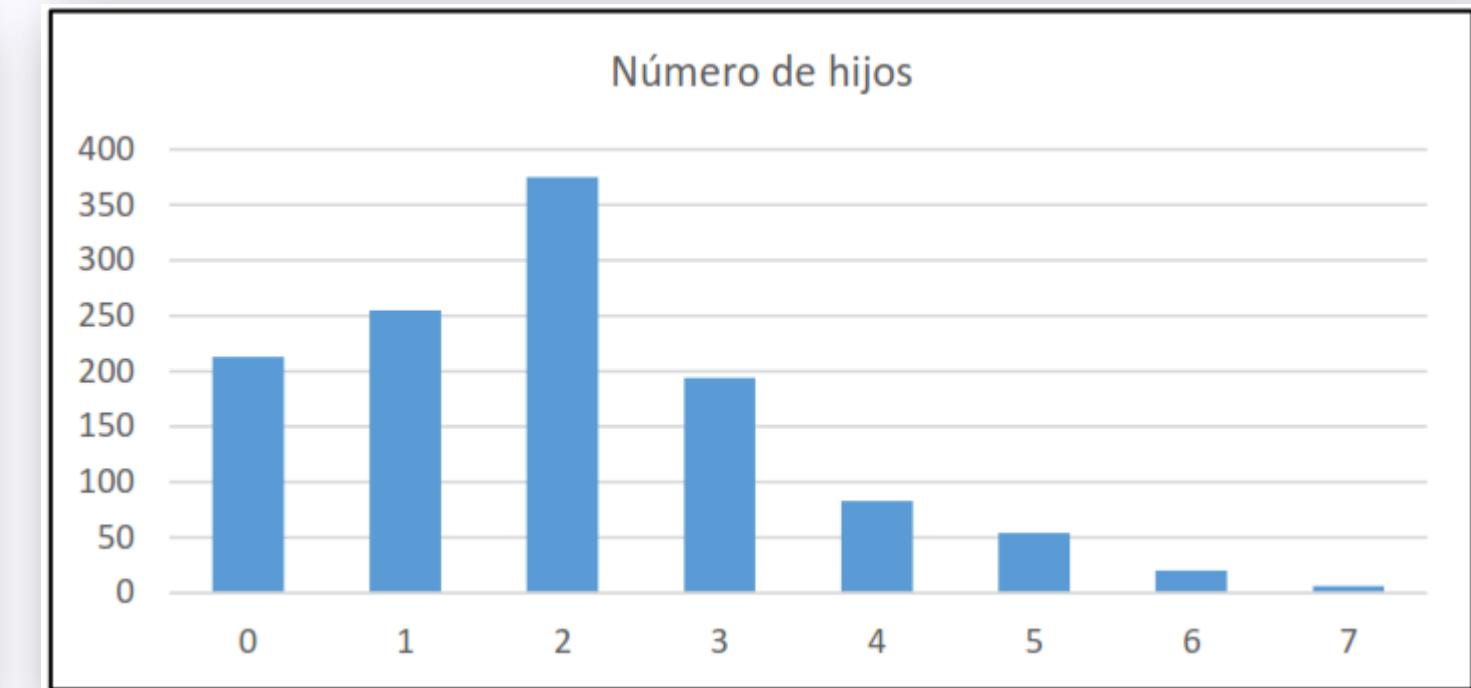


Diagrama de Barras variable cuantitativa discreta

Barras con altura proporcional a la frecuencia.

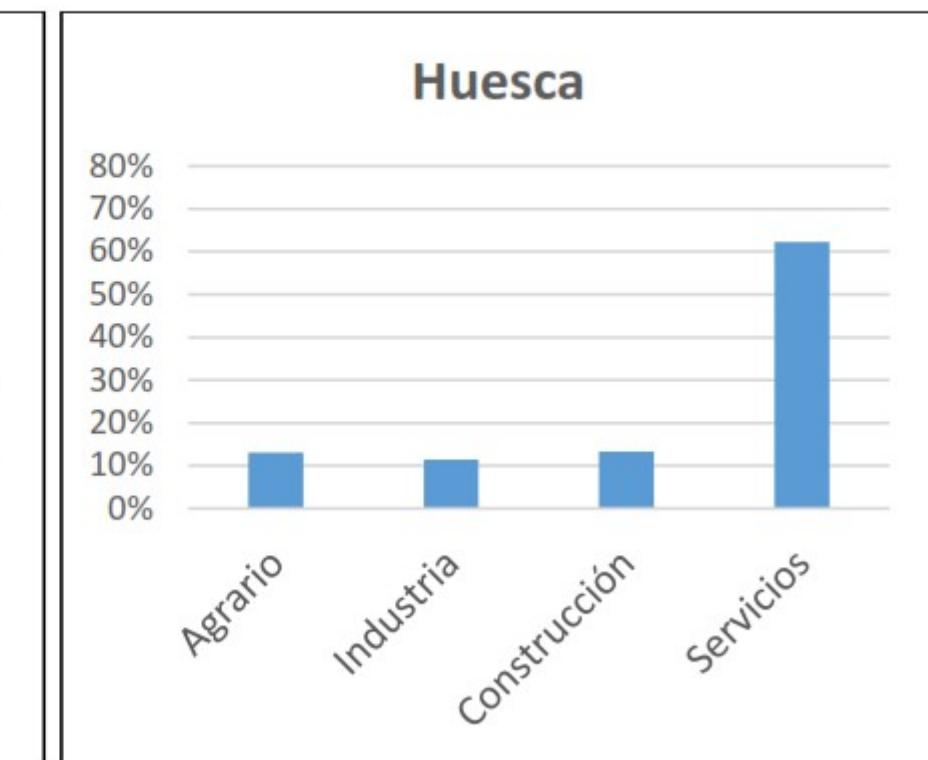
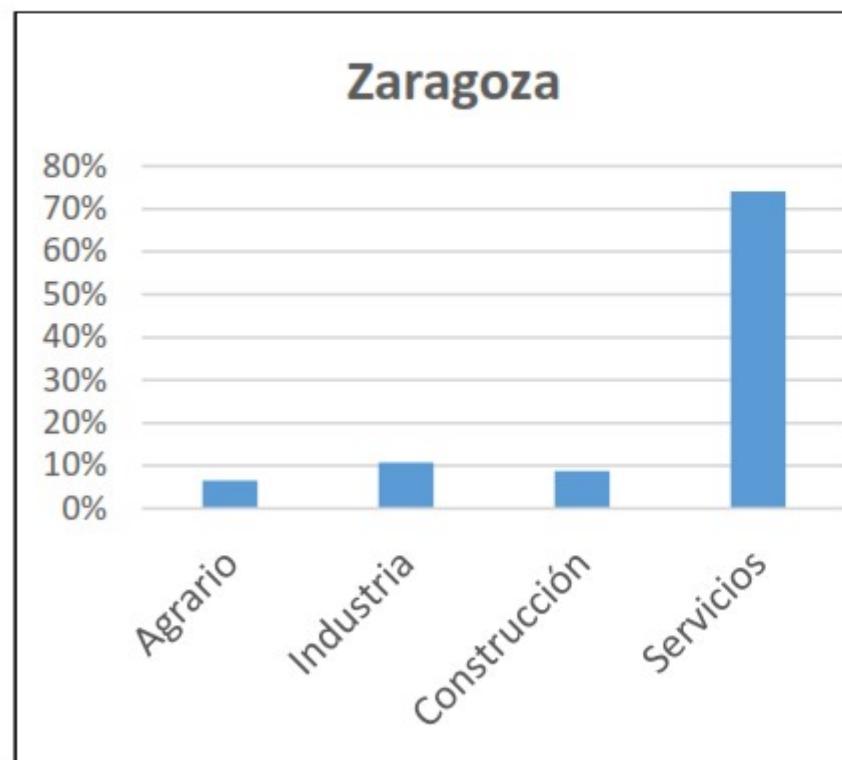


- ❑ **Importante:** Para comparar gráficas se usan las frecuencias relativas

Diagramas de Barras para comparar variables

- ❑ **Importante:** Para comparar gráficas se usan las frecuencias relativas

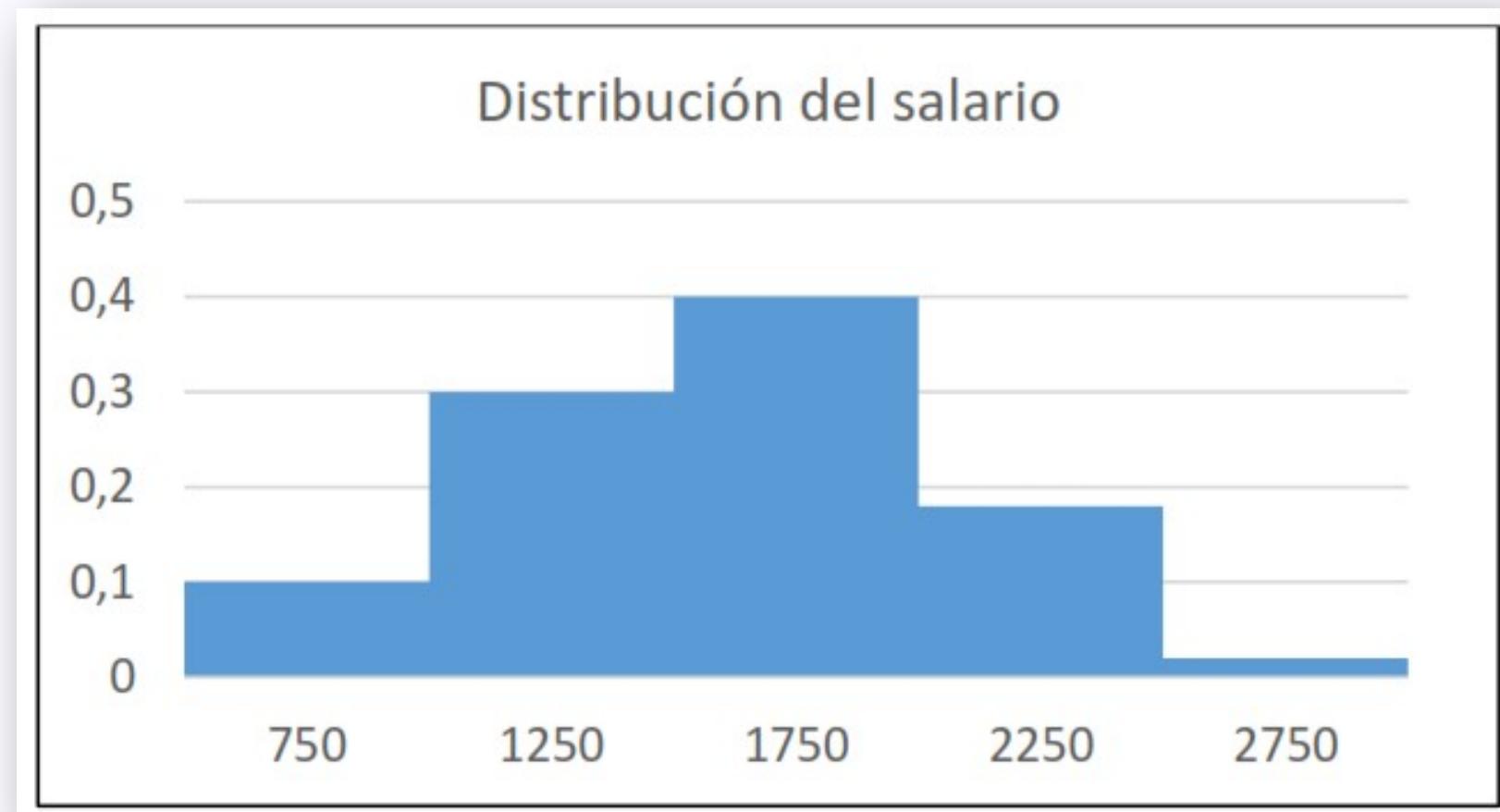
En el eje vertical podemos representar las frecuencias absolutas o relativas, pero si vamos a comparar gráficos debemos utilizar las frecuencias relativas:



Histogramas

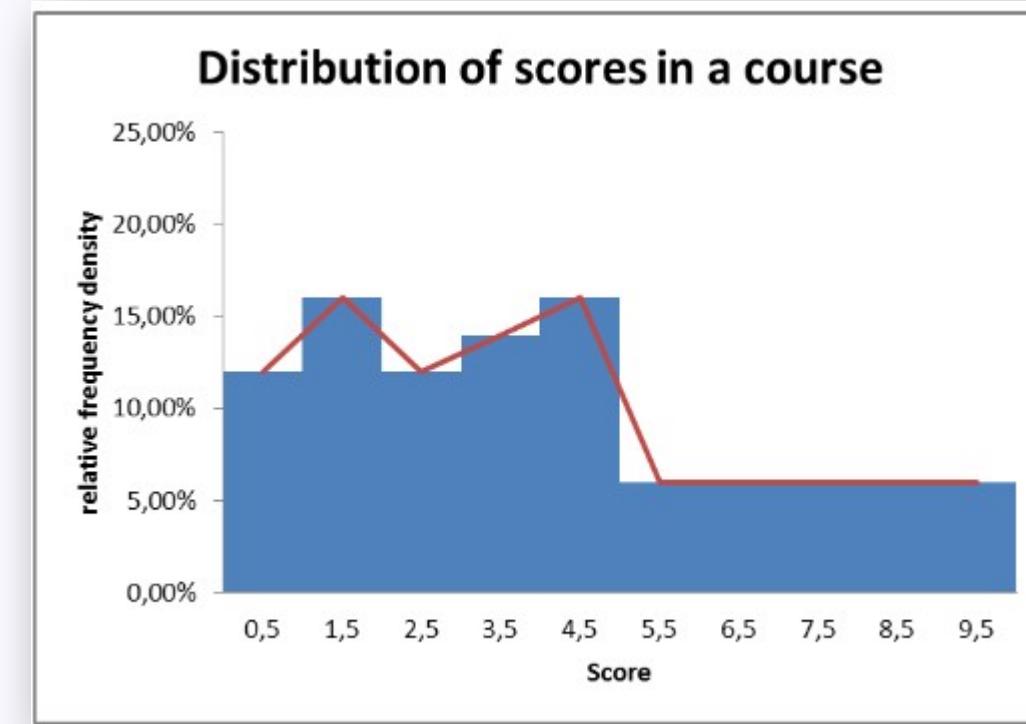
Representación gráfica mediante barras contiguas donde la **superficie es proporcional a la frecuencia**. Ideal para distribuciones agrupadas.

- Eje vertical: densidades de frecuencia (d_i)
- Eje horizontal: intervalos con marcas de clase
- Barras contiguas sin espacios



Polígonos de Frecuencias

Representación equivalente al histograma que se obtiene [uniendo los centros de las bases superiores](#) de los rectángulos del histograma.



Los polígonos permiten visualizar la forma de la distribución y facilitan la comparación entre diferentes conjuntos de datos.