

TEMA 0. – INTRODUCCIÓN. LOS DATOS GEOGRÁFICOS.

Se van a estudiar tres conceptos fundamentales:

- Cartografía temática.
- Datos geográficos.
- Fuentes fundamentales.

CARTOGRAFÍA TEMÁTICA.

Se puede definir como un conjunto de mapas que contienen una información adicional propia, distinta de la puramente topográfica. Se basa en una base de referencia locacional e incluye una información adicional.

Las cualidades que definen la cartografía temática son:

- *Claridad y Legibilidad* del mapa, evitando, por ejemplo, incluir demasiada información en él.
- *Esquematismo*. Al pasar de la realidad a un mapa, tendremos que generalizar la información en función del fenómeno y la escala
- *Rigurosidad*. Definida como la precisión. Se trata de colocar el fenómeno, no sólo en el sitio exacto, sino no intentar ser más preciso que la propia fuente de información. La mayoría de los mapas incluyen datos estadísticos que han tenido que ser tratados previamente de una manera rigurosa.
- *Poder de evocación*, que seamos capaces de hacer entender la información. Por ejemplo, los colores oscuros son mejores, en este sentido, que los claros.

TEMA 1. – LOS DATOS GEOGRÁFICOS.

1.1. *Las variables zonales y sus especificaciones.*

1.1.1. Clases de variables geográficas.

Al hablar de los datos geográficos nos estamos refiriendo a dos aspectos fundamentales: la Entidad Geográfica, como base de referencia locacional, y la Componente Temática.

La **entidad geográfica** puede ser de dos tipos: natural o artificial. La primera significa como está en el propio territorio, por ejemplo, el uso del suelo; si estudio el terreno por zonas(p.e. cuadrículas de 200 Has), se tratará de una entidad geográfica artificial.

Problema de la Unidad Espacial Modificada(UEM).

Suponemos que hay una zona de conflictividad social(CS) en una ciudad. Si divido el cuadrante en 4 zonas, la parte de CS en cada una se diluye con la mayoría dominante del mismo, con lo que se enmascara la información.

En el segundo caso sí aparece reflejada la zona de conflictividad social, por lo que la información no queda enmascarada.

Los elementos que definen la entidad geográfica son:

- Geometría: convierte la realidad en tres elementos
- Topología: proceso matemático que estudia las relaciones espaciales entre los elementos de un mapa.

En el terreno nos vamos a encontrar con tres tipos de entidades geográficas: *puntos*, *líneas* y *superficies*. Algunos investigadores incluyen al Volumen.

- Punto: objeto con posición pero sin dimensión. Se definen por un par de coordenadas.
- Línea: objeto unidimensional que une dos o más puntos.
- Área: objeto bidimensional que ocupa una superficie delimitada por un perímetro formado por, al menos, tres líneas.
- Volumen: objeto tridimensional delimitado por, al menos, dos objetos superficiales.

En muchos casos, la propia escala indica la generalización de los datos, de manera que una superficie pueda aparecer como un punto. Por ejemplo, la representación de Jaén en un mapa de España.

La **componente temática** es variable; muy pocas veces es constante. Desde el punto de vista estadístico, hay que tener claros tres conceptos:

- VARIABLE. Es el concepto general, se suele representar por la letra mayúscula(X).
- VALOR. Cantidad que toma en un momento determinado. Se representa con la misma letra, en minúscula(xi).
- FRECUENCIA. Número de veces que se repite un valor en un periodo determinado(ni).

La componente temática se puede clasificar desde cuatro diferentes puntos de vista:

- *Dependiente*

Independiente. Su valor no depende de otro.

- *Continua*. Permite fraccionar sus unidades, presenta el valor en todos sus puntos (p.e. la Temperatura: 17°; 17, 03°;).

Discreta. No permite la fragmentación (p.e. la Población: 1 hab.; 2.000 hab.;).

- *Cualitativa*. Expresa una cualidad del fenómeno a representar (carretera, río).

Cuantitativa. Expresa una cantidad del fenómeno estudiado (Autopista, Autovía,).

- *Absoluta*. Cantidad en sí mismo (Número de hab.)

Relativa. Cantidad en relación con otro valor (tasa de natalidad). Proviene de alguna operación con datos absolutos.

Organización de los Datos.

Denominamos POBLACIÓN a todo el conjunto de variables geográficas que van a constituir la base para el estudio de cualquier dato. Una MUESTRA es un subconjunto de la población basándose en un criterio determinado.

Ejemplo 1: Población: provincias españolas.

Muestra: Provincias andaluzas.

Ejemplo 2: Población: movimientos migratorios en África.

Muestra: los movimientos de los países situados al Norte de una determinada latitud.

La muestra se puede establecer siguiendo, fundamentalmente, dos criterios:

- AZAR, sin criterio preestablecido.
- ESTRATIFICADA, que implica establecer algún tipo de criterio para una agrupación lógica.

Los datos se pueden organizar en tres tipos de agrupaciones:

- Números Absolutos.
- Frecuencias.
- Intervalos.

Números Absolutos: dan un valor determinado, en un momento determinado, de una variable.

Agrupación de frecuencias: marcamos el número de veces que se repite una variable en un conjunto. Por ejemplo, en las notas de un examen:

	n.	N	n. %	N%
A	25	25	53	53
B	15	40	32	85
C	5	45	11	96
D	2	47	4	100
	47		100	

Intervalos: implica incorporar cada valor dentro de un subconjunto.

Nos Absolutos			Frecuencias			Intervalos					
				n	N		M. clase				
						10 !	1	1			
14	17	13				11 !	2	3			
17	14	10				12 !	2	5			
14	15	12				13 !	1	6	10-12	11	5
15	16	11				14 !	3	9	13-15	14	7
16	12	15				15 !	3	12	16-18	17	6
17	11	16				16 !	3	15			
						17 !	3	18			

Veamos otro ejemplo. Supongamos los datos de las temperaturas medias anuales . Si reducimos la tabla como en , estaremos organizando los datos por frecuencias. También se pueden organizar los datos por intervalos, como en el caso

En la organización de los datos hay que seguir un proceso de distribución de los mismos:

- **Recorrido:** diferencia entre los valores máximo y mínimo: $R = V_{\text{M}} - V_{\text{m}}$
- **Número de clases.** Para establecerlo, nos podemos valer de una referencia matemática, supeditada siempre al criterio del cartógrafo.

Podemos utilizar tres métodos:

- MORTICLIFFE: $n \text{ casos} = \sqrt{n}$
- BROOKS: $n \text{ casos} = 5 \log_{10} N$
- HUNTSBERGE: $n \text{ casos} = 1 + 3.3 \log_{10} N$

siempre con menos de 8 intervalos de cada clasificación.

- **Intervalos de clase.** Son los valores que actúan como límite en los subconjuntos. Más adelante se verá que la manera de hacer los intervalos es:
 - Iguales
 - Sistemáticamente desiguales: Prog. Aritmética: aumento índice constante.

descenso índice constante.

Prog. Geométrica aumento índice constante.

descenso índice constante.

- Irregulares.

Es muy importante la notación de los intervalos. Si se utiliza el corchete, significará que el valor está incluido en el intervalo; si lo que se usa es un paréntesis, significará que dicho valor no pertenece al intervalo en cuestión.

- **Marca de clase.** Es el punto central de cada intervalo que servirá para realizar los pertinentes cálculos estadísticos

1.1.2. Continuidad y Uniformidad.

Un dato es continuo cuando presenta valores en todos sus puntos, por ejemplo, la temperatura. La población es un dato discontinuo.

Un dato es uniforme cuando presenta un valor cuyos cambios, de un lugar a otro, son transitorios, no bruscos. Por ejemplo, la presión atmosférica. El PNB(Producto Nacional Bruto) es un ejemplo de un dato no uniforme.

1.1.3. Escalado de las variables geográficas.

Los datos geográficos se pueden escalar siguiendo cuatro métodos diferentes:

- Nominal.
- Ordinal.
- Intervalo.
- Índice.

El **escalado nominal** sólo implica una clasificación en función de una cualidad, sin diferencias cuantitativas.

El **escalado ordinal**, partiendo de la diferencia cualitativa, establece también una cuantitativa, porque incorpora lo que se llama un Rango (pueblo, ciudad, aldea,).

Las **escalas de intervalo** establecen las diferencias cuantitativas en función de unos límites para esos rangos; por ejemplo, ciudades: de 1.000 a 5.000 hab., de 5.000 a 25.000 hab. Utiliza unas clasificaciones arbitrarias, sin ninguna referencia al CERO.

Las **escalas de índices** utilizan un valor como establecimiento de partida, un valor de referencia: el CERO ABSOLUTO.

1.2. Representación gráfica de la información geográfica. Utilización de los gráficos.

El objetivo es mostrar relaciones entre magnitudes y señalar qué papel predominante juega una variable respecto al total.

Características de los gráficos en la representación cartográfica.

Nos permiten analizar, de forma visual, las relaciones entre distintas magnitudes, así como señalar, dentro de un conjunto, cual es el valor y la representatividad de cada dato con respecto al total.

Los tipos de gráficos que nos vamos a encontrar pueden ser:

- Gráficos lineales.
- Gráficos de barras
- Gráficos sectoriales.
- Diagramas triangulares.
- Gráficos de escala semilogarítmica.

Gráficos lineales.

Sirven para poner de manifiesto la relación de un elemento respecto a otro, la variación de un fenómeno a lo largo de un período de tiempo o para series temporales.

Su realización implica que la variable independiente se representa en abscisas y la dependiente, que suele variar irregularmente, en el eje de ordenadas.

Cuando se use más de una información, hay que mantener la misma escala. A veces es conveniente variar la escala para apreciar mejor las variaciones, pero sin falsear la realidad.

En cualquier caso, será preferible utilizar dos gráficos lineales en lugar de representar las dos variables en un único gráfico.

Gráficos de barras.

Se usan para representar una información donde los fenómenos tienen una naturaleza muy variable. También

si se quieren representar valores individualizados que no implican un crecimiento/decrecimiento continuo.

Gráficos sectoriales.

Tienen una gran ventaja: permiten representar dos informaciones al mismo tiempo. Por un lado el tamaño del círculo, gracias al cual sabremos el peso del fenómeno.

Por otro lado, dentro del círculo, según el área del sector, sabremos el peso de cada sector respecto al global del círculo.

Diagramas triangulares.

Se usan para representar una cantidad constante en tres elementos variables. Por ejemplo, la población en grandes grupos de edades: 0–15; 15–65; más de 65.

Se realiza un triángulo equilátero, donde cada lado se divide de 0 a 100, y cada punto del interior tendrá, por tanto, tres coordenadas. La división se hace de izquierda a derecha del eje.

Siempre se formarán ángulos de 60° entre los lados contiguos.

Por ejemplo, un diagrama de los tres sectores:

	SP	SS	ST	
Andalucía	22.6	25	52.4	= 100

El inconveniente que tienen es que son gráficos difíciles para medir y, al limitarse a tres ejes, en muchos casos hay que reagrupar las limitaciones.

Gráficos de escala semilogarítmica.

Permiten relacionar un elemento constante a lo largo del eje horizontal con un elemento variable que cambia proporcionalmente en el eje vertical.

Se utiliza cuando queremos representar características muy dispares.

Por ejemplo, la población en Jaén(100000 Hab), Baeza(16000 Hab) y Jamilena(1000 Hab)

En cuanto a las escalas, en abscisas la elección es libre y en ordenadas, escala semilogarítmica. Si en los dos ejes se usan escalas logarítmicas, el gráfico se llama Logarítmico.

1.3. Síntesis de la información geográfica: el empleo de índices.

Su fundamento radica en relacionar dos valores, obteniendo una cifra que permite establecer comparaciones entre diferentes unidades territoriales.

Pueden ser de dos tipos: proporción y densidad.

- **PROPORCIÓN:** cuando hacemos el cociente entre dos fenómenos. Si la proporción es muy pequeña, se multiplica por 10, 100, 1000, , con lo que se pasa a hablar de TASA.
- **DENSIDAD:** cuando relacionamos el fenómeno con una superficie.

1.3.1. Cocientes y tasas.

Índice de Demangeon. Es el índice básico. Se mide por k ,

$$k = \frac{E \cdot N}{T}$$

donde E: Población fuera de la capital; N: número de entidades de la capital; T: población total.

Por ejemplo, con Martos y Santiago–Pontones:

MARTOS: SANTIAGO–PONTONES:

T: 20.630 T: 4.687

N: 8 N: 52

NP: 18.172 NP: 1.483

E: 2.458 E: 3.204

$$k_M = \frac{2.458 \cdot 8}{20.630} = 0.95$$

$$k_{SP} = \frac{3.204 \cdot 52}{4.687} = 35.54$$

Demangeon oscila entre la máxima concentración y la máxima dispersión.

Concentración máxima < 0.01
Concentración media 0.01 – 0.1
Concentración mínima 0.1 – 1
Dispersión mínima 1 – 10
Dispersión media 10 – 100
Dispersión máxima > 100

En el ejemplo,

	T	N	NP	E
MARTOS	17.078	6	12.844	4.234
S–PONTONES	5.579	34	1.702	3.877

Índice de dependencia. Es la relación entre la población potencial activa y la población en edad no activa.

$$ID = \frac{(< 15) + (> 65)}{15 - 64} \cdot 1.000$$

En el ejemplo:

	MARTOS	S-PONTON.
< 15	4.733	763
15-64	12.885	3.045
> 65	3.014	880

MARTOS: 601,24

SANTIAGO-PONTONES: 539,57. Cada 1.000 habitantes en edad activa tienen que mantener 601/539 habitantes en edad no activa.

Densidad de población.

$$D = \frac{n \text{ habitantes}}{\text{Sup}(\text{Km}^2)}$$

En el ejemplo: (en 1.991)

	Población	Sup(Km2)	DENSIDAD
MARTOS	20.630	265,9	77,59 Hab/Km2
S-PONTONES	4.687	630,9	7,42 Hab/Km2

Coefficiente de localización industrial de Sargent-Florence. Estudia la estructura industrial de un país.

$$CL = \frac{\frac{\text{Empleo provincial de la rama que se estudia}}{\text{Empleo provincial total}}}{\frac{\text{Empleo nacional de la rama que se estudia}}{\text{Empleo nacional total}}}$$

Por ejemplo,

En la provincia de Álava: En España:

Empleo del COMERCIO: 9.388 Hab Empleo del COMERCIO: 1.515.574 Hab.

Empleo TOTAL: 97.554 Hab Empleo TOTAL: 11.883.143 Hab.

$$CL = \frac{\frac{9.388}{97.554}}{\frac{1.515.574}{11.883.143}}$$

$$CL_{LAVA} = 0,76$$

Permite zonificar en un mapa las áreas donde la especialización de un determinado aspecto prevalece sobre los demás.

1.3.2. Medidas de Concentración.

Índice de Semejanza. Sirve para comparar una característica geográfica a escala nacional con una distribución al nivel de unidades territoriales menores.

$$IS = 1 - \frac{\sum d}{100}$$

Por ejemplo, tenemos 5 barrios (A, B, C, D, E):

	A	B	C	D	E
%sobre pobl.	7	15	25	13	40
%obreros especialistas	21	19	4	50	6
d (b-a)	14	4	-21	37	-34
(-)= (+)= 55					

$$IS = 1 + 55/100 = 0,45$$

- Si tiende a CERO, la población se concentra en una unidad y la característica lo hace en otra diferente.
- Si tiende a UNO, habrá identidad entre las dos características

Coeficiente de GINI. Es la medida del grado de concentración. Si se acerca a CERO, hay una mejor equidistribución, y si se acerca a UNO, es que la concentración es mayor.

$$Ig = 1 - \frac{\sum_{i=1}^{k-1} q_i}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{\sum \frac{x_i n_i}{U}}{\sum \frac{n_i}{N}} \cdot 100$$

OJO!!! Los datos se ordenan de menor a mayor.

N Frecuencia acumulada

.

Ejemplo: Dos Comunidades Autónomas con 4 y 5 provincias, respectivamente, reparten sus recursos de la forma que se indica en la siguiente tabla. ¿En cuál de ellos se produce un reparto más equitativo? (¿Ig 0?).

PROVINCIAS	COMUNIDAD A	COMUNIDAD B
1	300.000	1.500.000
2	500.000	1.000.000

3	200.000	2.000.000
4	150.000	1.000.000
5		2.000.000

COMUNIDAD A.

xi	ni	N	xini	Ui	pi	qi	
150.000	1	1	150.000	150.000	25	13,04	Ig=1-100/150 Ig = 0,3333
200.000	1	2	200.000	350.000	50	30,44	
300.000	1	3	300.000	650.000	75	56,52	
500.000	1	4	500.000	1.150.000	100	100	
			1.150.000	$\sum_{i=1}^{k-1} p_i = 150$ $\sum_{i=1}^{k-1} q_i = 100$			

COMUNIDAD B.

xi	ni	N	xini	Ui	pi	qi	
1.000.000	2	2	2.000.000	2.000.000	40	26,667	Ig = 1-73,34/100
1.500.000	1	3	1.500.000	3.500.000	60	46,667	
2.000.000	2	5	4.000.000	7.500.000	100	100	Ig = 0,2666
			7.500.000	$\sum_{i=1}^{k-1} p_i = 100$ $\sum_{i=1}^{k-1} q_i = 73,34$			

Por tanto, hay un reparto más equitativo en la Comunidad B.

Curva de Lorentz. Se trata de una curva de frecuencias acumulativas que mide la desigualdad entre una distribución teórica y otra real.

La manera de trabajar es la misma que para calcular el índice de GINI. Veámoslo con un ejemplo:

SALARIO/HORA	TRABAJADORES
500-1500	3
1500-2500	7
2500-3500	8
3500-4500	4
4500-5500	2
5500-6500	1

xi	ni	N	xini	Ui	pi	qi	
1.000	3	3	3.000	3.000	12	4,11	Ig = 1-253,41/308 Ig = 0,1772
2.000	7	10	14.000	17.000	40	23,29	
3.000	8	18	24.000	41.000	72	56,16	

Ig = 0,1772

4.000	4	22	16.000	57.000	88	78,08
5.000	2	24	10.000	67.000	96	91,78
6.000	1	25	6.000	73.000	100	100,00
			73.000		EJE X	EJE Y
$\sum_{i=1}^{k-1} p_i = 308$ $\sum_{i=1}^{k-1} q_i = 253,41$						

1.3.2. Medidas de Concentración.

Índice de Combinación de . Se usa, genéricamente, para la delimitación de zonas agrícolas.

$$W = \frac{\sum d^2}{N}$$

donde: d " diferencia entre la superficie ocupada por un cultivo en la realidad y en una combinación ideal

N " Número de elementos que intervienen.

Ejemplo. Combinación de cultivos en Bailén.

Herbáceos(H): 650 6,32%

Olivar(O): 8.192 79,65%

Viñedo(V): 275 2,67%

Prados(P): 717 6,97%

Espacios arbóreos/frutales(V): 450 4,37%

TOTAL: 10.284

NOTA: Todo lo que represente menos de un 1%, no se utiliza.

	O	O-P		O-P-H			O-P-H-F				O-P-H-F-V				
%I	1C	2C		3C			4C				5C				
%R	100	50	50	33,3	33,3	33,3	25	25	25	25	20	20	20	20	20
	79,65	79,65	6,97	79,65	6,97	6,32	79,65	6,97	6,32	4,37	79,65	6,97	6,32	4,37	2,67
d	20,35	29,65	43,03	46,35	26,33	26,98	54,65	18,03	18,68	20,63	59,65	13,03	13,68	15,63	17,38
d2	414,12	2742,6		3569,5			4086,23				4459,62				
N	1	2		3			4				5				
W	414,2	1371,3		1189,8			1021,5				891,9				

Se trata de zonificar espacios que diversifican sus cultivos. Indica el predominio de cultivos en una determinada zona.

Índice neto de diversificación industrial.

$$IN = \frac{\text{ndice bruto de diversificación provincial} - \text{ndice bruto de diversificación nacional}}{\text{ndice bruto de diversificación mínima} - \text{ndice bruto de diversificación nacional}} \cdot 1.000$$

$$IN = \frac{IBDP - IBDN}{IBDM - IBDN} \cdot 1.000$$

Ejemplo. Tenemos 4 provincias con 5 sectores de actividad. Vamos a ver el % de empleo industrial en cada caso.

	% empleo industrial						A		B		C		D			%	%ac	%	%ac	%	%ac	%	%ac
	1	2	3	4	5																		
PROVINCIAS	1	2	3	4	5		40	40	60	60	30	30	20	20									
A	40	30	20	5	5		30	70	20	80	30	60	20	40									
B	60	20	10	5	5		20	90	10	90	20	80	20	60									
C	30	30	20	10	10		5	95	5	95	10	90	20	80									
D	20	20	20	20	20		5	100	5	100	10	100	20	100									
							IBDP	395		425		360		300									

(%ac = % acumulado)

Con 5 sectores de actividad IBDM 100 100

IBDM=500 0 200 Acumulando IBDM en cada uno de los 5 sectores.

- 300
- 400
- **500**

Calculamos el promedio de los porcentajes de los sectores en las provincias:

1> 37,5 37,5

2> 25 62,5

3> 17,5 80

4> 10 90

5> 10 100

IBDN = 370

Solución:

$$W_A = \frac{395 - 370}{500 - 370} \cdot 1.000 = 192,3$$

$$W_B = \frac{425 - 370}{500 - 370} \cdot 1.000 = 423,1$$

$$W_C = \frac{360 - 370}{500 - 370} \cdot 1.000 = -76,9$$

$$W_D = \frac{300 - 370}{500 - 370} \cdot 1.000 = -538,5$$

Conclusión:

Cuanto más negativo sea W, mayor diversificación industrial

1.3.4 Indicadores socioeconómicos.

Normalmente se utilizan en lo que se conoce como variables sintéticas, capaces de sintetizar variables heterogéneas para dar un identificador común que caracterice a distintos territorios. Son difíciles de calibrar.

Índice de Bienestar Social.

$$I_{BS} = \frac{\sum z_i}{N}$$

$$I = \frac{z_i - \bar{z}}{\delta_i}$$

Ejemplo:

En las tablas de las páginas -4- y -5-, se encuentra un ejemplo para el cálculo del índice de bienestar social teniendo en cuenta los siguientes factores: Médicos por cada 10.000 habitantes, Tasa de escolaridad por cada 1.000 habitantes, Ingreso medio por persona, la superficie útil de la vivienda y el consumo de energía eléctrica en alumbrado y uso doméstico.

FUENTES ESTADÍSTICAS PARA LA CARTOGRAFÍA TEMÁTICA.

Pueden ser públicas o privadas. Las primeras son más o menos periódicas, mientras que las segundas son más coyunturales. En general, las fuentes adolecen de una gran pobreza informativa, ya que algunas son muy difíciles de interpretar. También se puede dudar de su fiabilidad, ya que tienen una construcción propia (por ejemplo, el Censo Agrario de 1.972).

Unas veces carecen de actualización; otras están muy atomizadas o desagregadas, sin término medio. No existe homogeneidad.

(sigue en las páginas -12- y siguientes)

TEMA 2. – LA COMPONENTE TEMÁTICA. TRATAMIENTO ESTADÍSTICO.

Hay que tener en cuenta unos ciertos conocimientos de estadística para cotejar la información que tenemos.

Manejaremos tres aspectos:

- Medidas de tendencia central y dispersión;
- Correlación y regresión; y
- Estadística espacial.

La representación cartográfica emana de los datos estadísticos. Tendremos en cuenta tres aspectos fundamentales para usar las fuentes estadísticas: *fiabilidad*, *adecuación* y *familiaridad*.

- **Fiabilidad:** En algunas ocasiones, los datos no son del todo fiables; lo primero que hay que hacer es cotejar la fuente estadística que ha realizado el trabajo.
- **Adecuación:** Todas las fuentes cartográficas no trabajan sobre la misma unidad de análisis. Por ello, no pueden compararse todas directamente; deben estar adecuadas o trabajar sobre la misma unidad o, si no es así, fundir las informaciones de una manera coherente.
- **Familiaridad:** Ésta la da nada más que el uso frecuente. Hay que saber qué se va buscando, dónde se va a buscar, cómo se va a encontrar,

2.1. Medidas de tendencia central y dispersión.

Son las más utilizadas: Media, Mediana y Moda.

2.1.2. Medidas de tendencia central.

2.1.1.a. MEDIA.

Se divide en:

Media Aritmética. Cuando tienen una realidad espacial homogénea. Los datos no se desajustan mucho de la media.

$$\bar{x} = \frac{\sum x}{N} \quad \text{SIMPLE}$$

$$\bar{x} = \frac{\sum x_i n_i}{N} \quad \text{FRECUENCIAS}$$

$$\bar{x} = \frac{\sum x_i n_i}{N} \quad \text{INTERVALOS}$$

x_i = marca de clase.

Ejemplo: Queremos saber las Has medias de un municipio.

Límite Expl.	Frec(ni)	xi	xini
10-19	1	14,5	14,5
20-29	6	24,5	147
30-39	11	34,5	379,5
40-49	31	44,5	1379,5
50-59	42	54,5	2289
60-69	32	65,5	2064
70-79	15	74,5	1117,5
80-89	10	84,5	845
90-99	2	94,5	189
	150		8425

Media Ponderada. Se usa cuando se desea incluir una modificación a la media. En el ejemplo, la ponderación será la superficie. En estos casos, la media ponderada se llama media geográfica.

$$\bar{x} = \frac{\sum x_i a_i}{A}$$

Media armónica. Cuando se desea medir tasas, cambios, velocidades,

$$Ma = \frac{N}{\sum \frac{1}{x_i}}$$

donde N=nº casos.

$$\text{Ejemplo: } Ma = \frac{4}{\frac{1}{10} + \frac{1}{20} + \frac{1}{30} + \frac{1}{40}} = 19,2 \text{ Km / h}$$

Tramos : 1 2 3 4

Velocidad(Km/h): 10 20 30 40.

Media Geométrica. Cuando queremos descubrir promedios de crecimiento o cualquier magnitud cuyos cambios tiendan a ser proporcionales con el tamaño de esa magnitud.

$$Mg = \sqrt[n]{x_1 x_2 x_3 \dots x_n}$$

$$Mg = \text{anti log } \frac{\sum \log x_i}{N}$$

Ejemplo:

Población en Jaén en 1.960 64.553 Hab.

Población en Jaén en 1.970 77.317 Hab. ¿Cuál es la población en 1.965?

$$Mg = \text{anti log } \frac{\log 64.553 + \log 77.317}{2}$$

. La población en 1.965 es 70.648 Habitantes.

2.1.1.b. MEDIANA.

La mediana es el valor central, una vez ordenada la serie. Cuando la serie es par, se hace el promedio de los dos valores centrales.

Cuando tenemos frecuencias, hacemos la frecuencia acumulada:

Si es por intervalos:

$$Md = 65 + \frac{8 - 5}{5} \cdot 5 = 68$$

2.1.1.c. MODA.

Es el valor que se presenta con mayor frecuencia, el más representativo de una serie. Es de gran ayuda la realización de un histograma de frecuencias.

Si los datos vienen dados por intervalos:

$$Mo = l_i + \frac{n_{i+1}}{n_{i+1} + n_{i-1}} a_i$$

donde: l_i = límite inferior del intervalo;

n_{i+1} = frecuencia de la clase superior a la media;

n_{i-1} = frecuencia de la clase inferior a la media;

a_i = amplitud del intervalo.

2.1.3. Medidas de posición no central.

2.1.2.a. Cuantiles.

Son medidas que dividen la muestra con igual número de observaciones. Cuando la división se hace en cuatro partes, hablamos de *cuartiles*, que son tres: $N/4$, $N/2$ y $3N/4$.

Quintiles: $Q/5$, $2Q/5$, $3Q/5$ y $4Q/5$.

Deciles: $D/10$, $D/5$,

Percentiles: $P/100$, $P/50$,

Ejemplo 1:

- calcular al número máximo de cigarrillos que fuma uno de los fumadores del 25% que menos fuma.
- ¿Qué porcentaje de individuos fuma entre 12–22 cigarrillos al día?

	n	N
2.5–5.5	10	10
5.5–10.5	20	30
10.5–15.5	30	60
15.5–20.5	25	85
20.5–25.5	15	100
	100	

a) ¿Dónde está el 25%? 5.5–10.5

$$P_{25} = 5.5 + \frac{25 - 1}{20} \cdot 5 \quad P_{25} = 9.25 \text{ cigarrillos}$$

b)

$$12 = 10.5 + \frac{\frac{100 - \alpha_1}{1} - 30}{30} \cdot 5 \quad \alpha_1 = P = 39\%$$

$$22 = 20.5 + \frac{\frac{100 - \alpha_2}{2} - 85}{30} \cdot 5 \quad \alpha_2 = P = 87.5\%$$

$$- = 87.5 - 39. \text{ SOLUCIÓN: } 48.5\%$$

Ejemplo 2:

En la tabla adjunta se muestra la distribución de población en Hab./Km2 de 120 municipios. Si se consideran muy densas las zonas sobre el P79, ¿qué densidad mínima habría de tener?. Calcular el % de municipios con densidad de población < 103 Hab/Km2.

	n	N
60–70	2	2
70–80	3	5
80–90	25	30
90–100	46	76
100–110	35	111
110–120	5	116

120-130	3	119
130-140	1	120

a) $79 \cdot 120 / 100 = 94.8 \text{ Hab/Km}^2$; $P_{79} = 100 + \frac{94.8 - 76}{35} \cdot 10 \quad P_{79} = 105.37 \text{ Hab / Km}^2$

b) $\frac{120}{x} - 76$
 $103 = 100 + \frac{100}{35} \cdot 10 \quad x = 72.08 \text{ Hab / Km}^2$

Vamos a trabajar con series de datos y series agrupadas en intervalos y frecuencias, para ver algunos ejemplos. Imaginamos que partimos de 2 series de datos:

Serie A: 1, 4, 4.5, 5, 5.5, 6, 9.

Serie B: 1.5, 2, 3, 4.5, 7, 8, 9. En ambos casos $\bar{x} = 5$. Parece que el 5 caracteriza mejor a la primera serie.

Veamos la desviación media de los valores respecto al valor central, utilizando la desviación típica:

$(x_i - \bar{x})^2$
 SerieA: $(16+1+0.25+0+0.25+1+16) = 34.5$
 $\delta_1 = \sqrt{\frac{34.5}{7}} = 2.22$

$(x_i - \bar{x})^2$
 SerieB: $(12.25+9+4+1.25+4+9+16) = 54.5$
 $\delta_2 = \sqrt{\frac{54.5}{7}} = 2.8$

A la vista de los resultados de δ_1 y δ_2 , vemos que los valores de la serie B se alejan más de la media.

Supongamos que vamos a trabajar con frecuencias. Tenemos una tabla de distribución del número de hijos en España en un periodo de tiempo. Calcular:

– Media, – Mediana; – Desviación típica; – Desviación media;

– Desviación semintercuartílica.

Nº hijos:	0	1	2	3	4	5	6	7
ni :	8	13	10	5	2	1	0	1
N :	8	21	31	36	38	39	39	40

Media: $x_{ini} = 0+13+20+15+8+5+0+7 = 68$
 $\bar{x} = \frac{\sum x_i n_i}{N} = \frac{68}{40} \quad \bar{x} = 1.7$

Mediana: Posición 20 1. Mediana=1

Desviación semintercuartílica: 1er cuartil. – $Q1N / 4 = 10$ (Posición 10) $Q1 = 1$

3er cuartil. - $Q3N / 4 = 30$ (Posición 30) $Q3 = 2$. $Ds = (2-1)/2 = 0.5$

Desviación Media:

$$D_m = \frac{\sum |x - \bar{x}| n_i}{N} = \frac{45.4}{40}$$

$$\sum |x - \bar{x}| n_i = 45.4$$

$$D_m = 1.135$$

Desviación tipo:

$$\delta = \sqrt{\frac{\sum (x - \bar{x})^2 n_i}{N}} = \sqrt{\frac{88.4}{40}};$$

$$= 1.8$$

$$\sum (x - \bar{x})^2 n_i = 23.12 + 6.37 + 0.9 + 8.45 + 10.5 + 10.8 + 0 + 28.09 = 88.4$$

Ejemplo de distribución por intervalos para calcular la desviación típica:

Nº municipios	0-50	50-100	100-150	150-200	200-250	250-300	300-350	350-400	
Marca de clase	25	75	125	175	225	275	325	375	
Nº provincias(n)	3	15	6	7	12	4	1	2	N=50
N	3	18	24	31	43	47	48	50	

$$\text{Media} = 161; \quad \bar{x} = \frac{\sum x_i n_i}{N} = \frac{8050}{50} = 161$$

$$\delta = \sqrt{\frac{\sum (x - \bar{x})^2 n_i}{N}} = \sqrt{\frac{395200}{50}};$$

$$= 88.9$$

$$\sum (x - \bar{x})^2 n_i = 55448 + 110940 + 7776 + 1372 + 49152 + 51984 + 26896 + 91592 = 395200$$

En el siguiente ejemplo tengo la producción de trigo y maíz, sus medias y desviaciones:

	Producción	$\bar{x}$	
Zona A (Trigo)	16 Qm/Ha	12	3.5
Zona B (Maíz)	30 Qm/Ha	23	8

¿Cuál es más representativa respecto a su media?.

Para ello vamos a **normalizar las medias**: $\frac{x - \bar{x}}{\delta}$

; Trigo = $(16-12)/3.5 = 1.4$

Maíz = $(30-23)/8.2 = 0.85$. Por lo que es más representativo el cultivo de trigo que el del maíz.

Lo deseable en una distribución sería que coincidieran la media, la mediana y la moda en un mismo punto; es decir, que la serie estuviera normalizada. Lo normal es que las series sean asimétricas.

A partir de la campana de Gauss, podemos decir si la asimetría de la serie es positiva o negativa.

Pearson intentó medir el grado de asimetría de unas variables, utilizando el **coeficiente de asimetría**:

$$C_{asim} = \frac{3 (media - mediana)}{desviación}$$

Con la campana de Gauss sabemos la probabilidad de encontrar un dato entre la media y una, dos, tres, veces la desviación típica:

En una curva de distribución normal se cumple que:

$(\bar{x} - \bar{x})$ - 68.26 %

$(\bar{x} - 2\bar{x})$ - 95.44%

$(\bar{x} - 3\bar{x})$ - 97.74%

2.1.4. Medidas de dispersión.

2.1.4.a. Parámetros de dispersión relativos.

2.2. Correlación y Regresión.

TEMA 3. – CARTOGRAFÍA TEMÁTICA. SOLUCIONES GRÁFICAS Y CALSIFICACIÓN.

3.1. Distribución espacial de las variables cartográficas.

Al representar los elementos de un mapa, desde el punto de vista espacial, sólo vamos a poder cartografiar **puntos, líneas y superficies**, con independencia de que, debido a diversos factores, puedan quedar falseados.

3.2. Soluciones gráficas.

Cuando se quiere realizar un mapa temático hay que tener en cuenta:

- Símbolos: representación gráfica de un objeto con rigurosidad.

- Variables retinianas.

El simbolismo cartográfico consiste en la disposición convencional de los símbolos, teniendo en cuenta las variables retinianas, y que tengan implantación puntual, lineal o superficial.

3.2.1. VARIABLES RETINIANAS. Son: *forma, tamaño, color, orientación, grano y tono*.

- Forma. Para diferenciar cualitativamente los elementos.
- Tamaño. Es, junto con el color, la que permite establecer diferencias cuantitativas.
- Color.
- Orientación.
- Grano.
- Tono.

3.2.2. CLASES DE SÍMBOLOS. Son: signos, puntos, cartodiagramas, líneas y figuras geométricas.

SIGNOS. Cuando nos referimos a ellos, lo hacemos a dibujos más o menos esquemáticos, a los que se les asigna un significado.

- Símbolos convencionales. Son los que aparecen en un mapa, necesitados de una leyenda.
- Símbolos pictóricos. A través de un dibujo, representamos el elemento.

Los primeros suelen ser esquemas, en posición real, de un objeto no representable en el mapa por sus dimensiones.

PUNTOS. El concepto geométrico de punto excluye cualquier concepto de dimensión, aunque en el dibujo la tenga. Se pueden considerar como puntos los círculos cuyo " **2 mm**."

El punto, acompañado de color, nos permite establecer dos variables en el mapa: el punto, como tal, y el color que identifica a cada punto.

También puede tener un valor numérico: asociamos el tamaño del punto al volumen de lo que vamos a representar; o bien, como una unidad en una localización concreta.

Un mapa formado por puntos genera zonas de distinta concentración del fenómeno representado.

CARTODIAGRAMAS. Son mapas que localizan, en los puntos de control, gráficas que representan los valores de un fenómeno en un periodo de tiempo. El resultado es una estadística representada en un mapa. No tienen nada que ver con los cartogramas.

LÍNEAS. Son curvas irregulares que unen puntos con el mismo valor. En términos generales se les llama ISOLÍNEAS. Entre éstas, podemos establecer diferencias.

Líneas Isométricas. Éstas son líneas que expresan cantidades reales o derivadas (Temperatura, altitud, etc.) que se localizan en puntos determinados.

Isopletas. Un ejemplo de isopletas puede ser la línea que une puntos de igual Rendimiento por hectárea (Rto/Ha). Estos puntos no tienen valor real; normalmente, estos datos se localizan, por ejemplo, en las cabeceras de comarca.

Isocoropletas. Son las Líneas Isométricas acompañadas de una gradación de color.

FIGURAS GEOMÉTRICAS. Se diferencian planas y 3D. Se utilizan para representar.

No deben usarse distintas figuras geométricas para valores dentro de una misma serie. Por ejemplo:
Temperatura:

Tampoco más de dos o tres figuras geométricas en un mapa, porque puede resultar confuso.

La proporcionalidad de la figura geométrica se hace por el método de la Raíz Cuadrada, donde se establece la raíz del dato al radio(círculo)/lado(cuadrado).

Método del Escalado Sicológico.

Fue utilizado por Flannery. A partir de un estudio sicológico, se dio cuenta de la capacidad de diferenciar los distintos tamaños y símbolos. Se hizo el Escalado Sicológico.

A partir de aquí, le aplicó los símbolos graduados.

Vectores.

Son una serie de líneas que nos marcan la dirección o el sentido del movimiento de un determinado fenómeno.

En muchos casos simbolizan un punto de origen y otro de destino.

Se utilizan para representar fenómenos que implican movimiento.

Además, normalmente se representan arcos para dar más impresión de movimiento natural; también se utiliza el grosor para la magnitud, bien como símbolo graduado, o bien que cada uno represente un grado de magnitud.

Ejemplo:

Los vectores fueron utilizados por primera vez en 1938, en Borneo, para representar el movimiento de viajeros por ferrocarril en Irlanda.

Los vectores constituyen un mapa de flujos.

COLOR.

El color debemos entenderlo en dos opciones:

- Factor delimitador; y
- Elemento adicional.
- Factor delimitador.

El color, además de tener un gran factor de atracción sobre el usuario, resulta muy rápido establecer relación entre ellos y su significado. Por ejemplo, verde!vegetación.

Juega un papel muy importante, además de tener un componente natural importante.

- Elemento adicional.

Además de la componente de magnitud, podemos encontrar una relación.

Cuando representamos cualidades, debemos tener cuidado con los tonos de color, porque podemos representar una cantidad cuantitativa.

Si, por ejemplo, represento los cultivos de una región, y le doy tonos del color de la impresión, habrá zonas de más valor y zonas de menos:

El color, cuando utilizamos mapas de puntos, no debemos utilizar colores muy próximos entre porque es muy difícil diferenciarlos.

El color también implica dificultades cuando utilizamos figuras geométricas planas.

3.3. Clasificación de los mapas temáticos.

3.3.1. Atendiendo a su contenido.

- Mapas analíticos:

3.3.2. Según la naturaleza de los datos.

CARTOGRAFÍA II

– 15 –

TEMA 2 La Componente Temática

– 27 –

n = número de casos.

n – frecuencia

N – frecuencia acumulada

n % – frecuencia porcentual

N % – frecuencia porcentual acumulada

25

1900 1910 1920 1930 1940 1950 . . . 1.991

1900 1910 1920 1930 1940 1950 . . . 1.991

100.000

10.000

1.000

100

10

0

100.000

50.000

25.000

JAÉN

BAEZA

JAMILENA

JAÉN

BAEZA

JAMILENA

Figura 1. Jamilena y Baeza apenas han variado, mientras que la población de Jaén ha crecido mucho en el último siglo

Figura 2. Utilizando un gráfico de escala semilogarítmica, tenemos una idea distinta de crecimiento

12 40 72 88 96 100

100

91,78

78,08

56,16

23,29

4,11

CURVA DE LORENTZ

52,4

SE MA CO HU GR CA JA AL

CONSUMO

22,6

PROVINCIAS

MA

SE

CA

AL

CO

GR

JA

HU

CASTILLA LA MANCHA

ANDALUCÍA

$$\bar{x} = \frac{x_i \cdot n_i}{N} = \frac{8.425}{150}; \bar{x} = 56,16 \text{ Ha}$$

xi	ai	xiai
170	300	170*300
21	19	21*19
5	12	5*12
	A= ai	

170

10

x = 155,46

5

	n	N	
1	10	10	
2	12	22	Posición 303Mediana
3	23	45	
4	11	56	
5	4	60	
		N=60El valor central es la posición 30	

	n	N	
55-60	1	1	
			(N+1)/2=8;

Mediana " (65-70)

60-65	4	5	
65-70	5	10	
70-75	3	13	
75-80	1	14	
80-85	1	15	
		N=15	

Podemos calcular la mediana dentro del intervalo:

$$Md = \text{Limite inferior} + \frac{Md - N_{\text{de la clase anterior}}}{n_{\text{clase}}} \times \text{amplitud del intervalo}$$

Cuartiles

N/4 N/2 3N/4

2.5 P12 P22 29.5

%

$$12 = 10.5 + \frac{\frac{100 - \alpha_1}{1} - 30}{30} \times 5 \quad \alpha_1 = P = 39\%$$

< 15°

15°-18°

>18°

Patrón Standard(no es fijo):

- Tamaño círculo máximo " Sup. max.
- Relación valores máximo y mínimo de radios.

$$x = \frac{\sqrt{V_{\min}} - a}{\sqrt{V_{\max}}}$$

Serie	
x1	$\sqrt{x_1} \quad r_1$
x2	$\sqrt{x_2} \quad r_2$
x3	$\sqrt{x_3} \quad r_3$

x_4	$\sqrt{x_4} \quad r_4$
-------	------------------------

! Lo asoció a una unidad base

Xi	log Xi	log Xi · 0,57	Antilogaritmo
X1	log X1	log X1 · 0,57	
X2	log X2	log X2 · 0,57	
X3	log X3	log X3 · 0,57	

Los vectores indican movimiento de mercancías.

Como vemos, el grosor nos indica la magnitud:

1 cm 750–1000 Tm

0,5 cm 500–750 Tm

0,3 cm 0–500 Tm

Aquí lo hago por rangos

Podemos añadir colores para diferenciar las mercancías

1000 Tm

150 Tm 400 Tm