

Diagrama de dispersión o nube de puntos

Hasta ahora los datos estadísticos estudiados estaban asociados a variables estadísticas unidimensionales, es decir se estudiaba un solo carácter de la población.

Sin embargo, en ocasiones es útil considerar a la vez varios caracteres de una misma población: la estatura, la edad y el peso de un grupo de 50 niños; el sueldo y la cualificación de un conjunto de asalariados; la extensión y el número de habitantes de los países europeos, la inversión en publicidad y la facturación de ciertas empresas, etc.

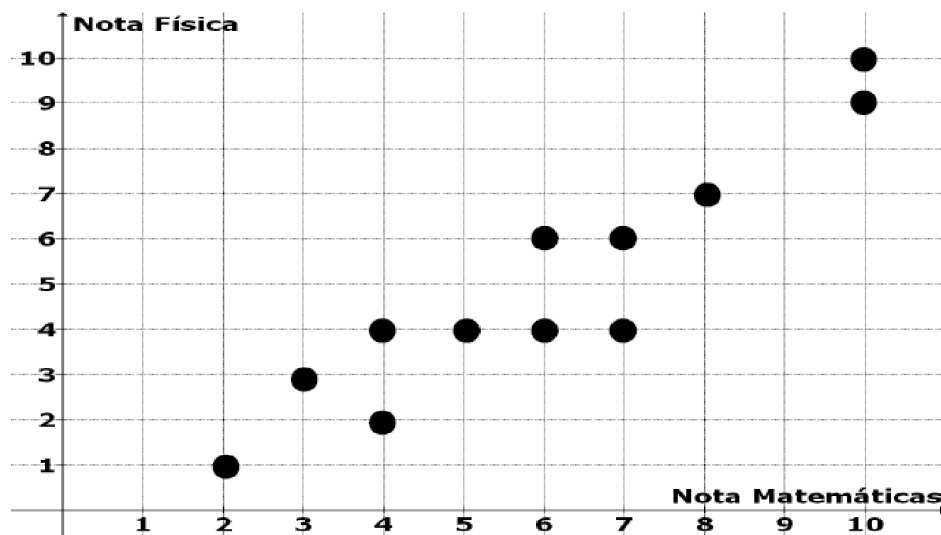
Nosotros reduciremos el estudio de una población a dos caracteres, teniendo así variables estadísticas bidimensionales.

Cuando se van a estudiar dos características X e Y de una misma población, los datos que se obtienen son parejas de valores  $(x_i, y_i)$ . El conjunto de datos  $(x_i, y_i)$  se llama distribución bidimensional y la representación gráfica de los puntos  $(x_i, y_i)$  se llama diagrama de dispersión o nube de puntos.

*Ejemplo:*

Notas de 12 alumnos en Matemáticas y Física

Alumno	a	b	c	d	e	f	g	h	i	j	k	l
Matemáticas	2	3	4	4	5	6	6	7	7	8	10	10
Física	1	3	2	4	4	4	6	4	6	7	9	10



Tablas de doble entrada

En las distribuciones bidimensionales, cuando hay pocos pares de valores, se procede como hemos hecho, es decir, enumerándolos.

Si algún par está repetido se pone dos veces, pero cuando el número de datos es grande, se recurre a las tablas de doble entrada.

En cada casilla se pone la frecuencia correspondiente al par de valores que definen esa casilla.

*Ejemplo*

x/y	0	1	2
0	2	1	0
1	3	4	1
2	0	5	3

Lo que indica el número de veces que está cada par. El par (0, 1) está 3 veces, (1, 2) está 5 veces. etc.

Correlación entre dos variables

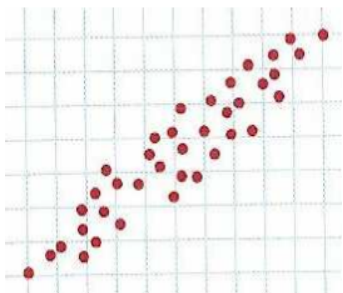
La correlación entre dos variables es el grado de concentración de la nube de puntos en torno a una línea imaginaria (creciente o decreciente) y nos indica si hay más o menos relación entre las X e Y. En el ejemplo anterior, la nube de puntos se concentra en torno a una línea. Podemos deducir que las notas de un determinado número de alumnos en física y matemáticas son parecidas.

La correlación será positiva o directa si la línea es creciente y negativa o inversa si es decreciente y será más fuerte cuanto mayor sea la concentración de los puntos entorno a esa línea.

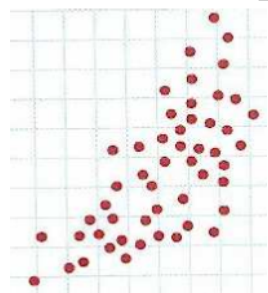
Si la nube se ajusta perfectamente a una línea se dice que la correlación es funcional. En este caso, los puntos forman parte de la gráfica de una función.

Se dice que hay correlación lineal si la nube de puntos pertenece a una recta.

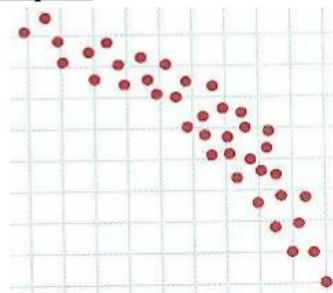
Ejemplos:



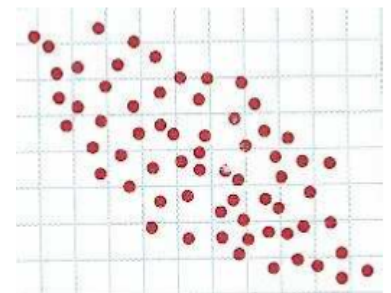
Correlación lineal directa fuerte



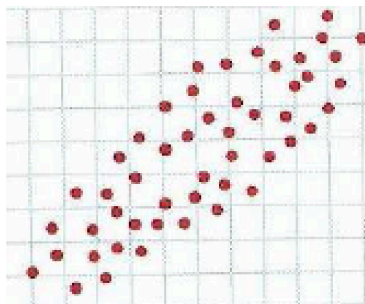
Correlación directa débil



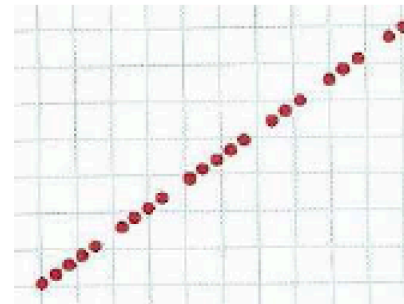
Correlación inversa fuerte



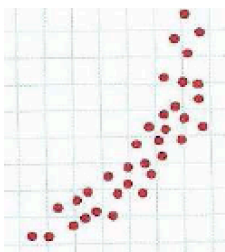
Correlación lineal inversa débil



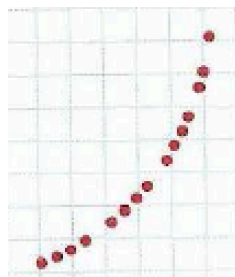
positiva débil



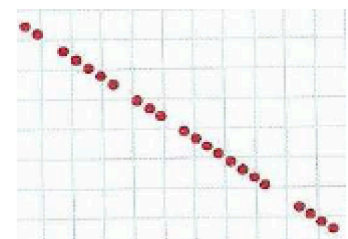
positiva funcional (lineal)



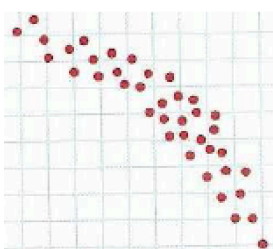
positiva curvilínea fuerte



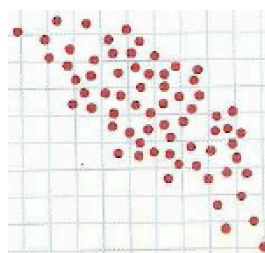
positiva funcional (curvilínea)



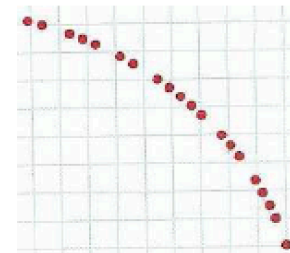
negativa funcional (lineal)



negativa curvilínea fuerte

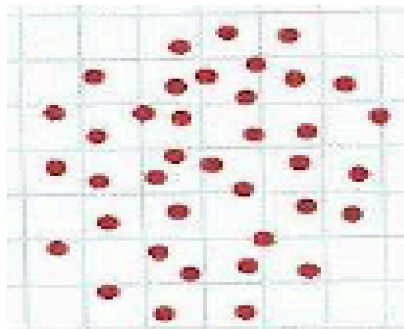


negativa curvilínea débil



negativa funcional (curvilínea)

Si los puntos están esparcidos sin concentrarse en torno a ninguna línea, se dice que no hay relación entre las variables o que la correlación es nula.

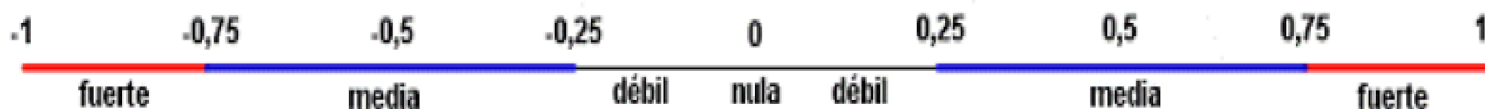


Correlación nula

El grado de concentración de la nube de puntos respecto a una línea se mide con un número llamado coeficiente de correlación,  $r$ , y tiene las siguientes propiedades:

- 1) No tiene dimensiones. Es decir, no depende de las unidades en las que se expresan los valores de las dos variables. Por tanto, si se realiza un cambio de unidades, el valor de  $r$  no varía.
- 2) El valor de  $r$  está comprendido entre  $-1$  y  $1$ .
- 3) Si la correlación es perfecta (puntos de la nube alineados), entonces  $|r| = 1$ , es decir,  $r = 1$  ó  $r = -1$ .
- 4) Si la correlación es fuerte,  $|r|$  es próximo a  $1$ .
- 5) Si la correlación es débil,  $|r|$  es próximo a  $0$ .
- 6) Si  $-1 < r < 0$  la correlación es negativa
- 7) Si  $0 < r < 1$  la correlación es positiva

La correlación es tanto más fuerte a medida que  $r$  se aproxima a  $-1$  ó  $1$  y es tanto más débil a medida que se aproxima a  $0$ .



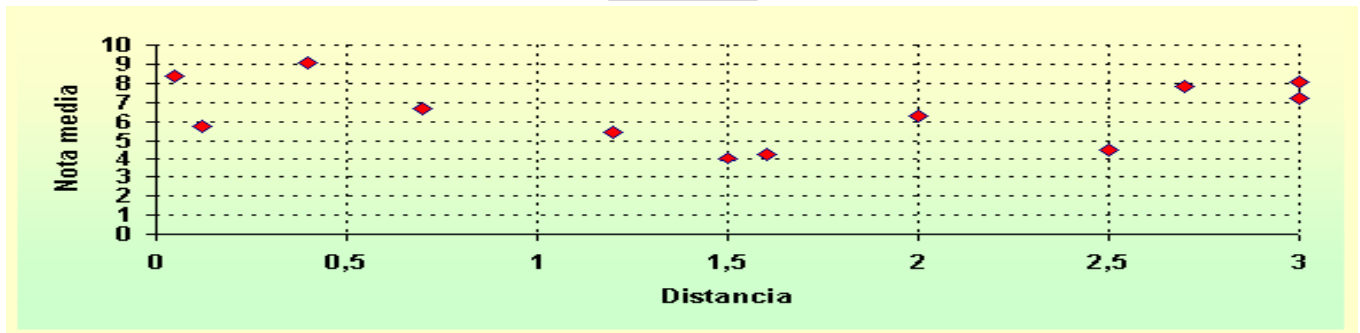
**Actividades resueltas**

1) A 12 alumnos de un centro se le preguntó a qué distancia estaba su residencia del Instituto, con el fin de estudiar si esta variable estaba relacionada con la nota media obtenida. Se obtuvieron los datos que figuran en la siguiente tabla:

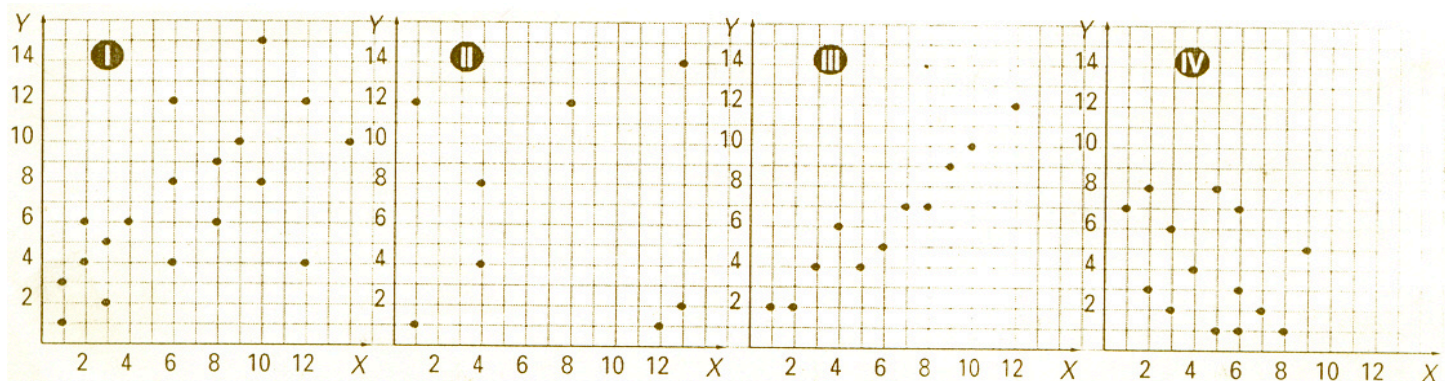
X = distancia (km)	0,0	0,1	0,	0,	1,	1,	1,	2	2,	2,	3	3
Y = nota media	8,4	5,7	9,	6,	5,	4	4,	6,	4,	7,	7,	8,
			1	7	4		1	2	5	9	2	1

Dibuja el diagrama de dispersión

**Resolución**



2) Los coeficientes de correlación de 4 distribuciones bidimensionales son (no necesariamente por ese orden):  $-0,04$ ;  $0,96$ ;  $0,65$  y  $-0,37$ . ¿Cuál corresponde a cada gráfico?



**Resolución**

Solución: I  $\rightarrow$   $0,65$     II  $\rightarrow$   $-0,04$     III  $\rightarrow$   $0,96$     IV  $\rightarrow$   $-0,37$

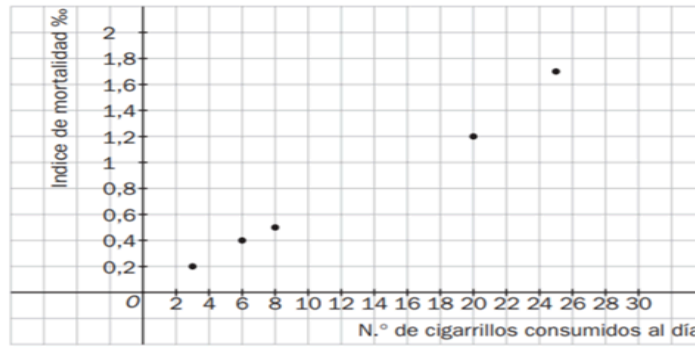
Observa la siguiente variable bidimensional.

N.º de cigarrillos consumidos al día	3	6	8	20	25
Índice de mortalidad	0,2	0,4	0,5	1,2	1,7

- 3) a) Representa la nube de puntos.  
b) Indica el tipo de correlación.

**Resolución**

a) La nube de puntos es la siguiente:



b) Como al aumentar el número de cigarrillos consumidos aumenta el índice de mortalidad, la correlación es positiva.

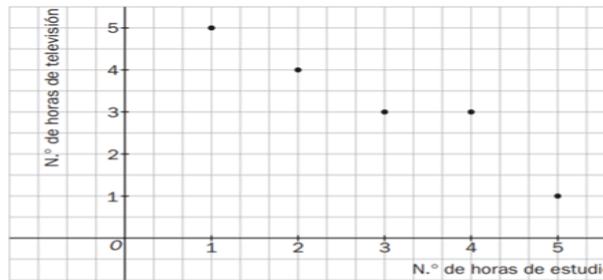
Una variable bidimensional viene dada por la siguiente tabla.

N.º de horas de estudio	1	2	3	4	5
N.º de horas de televisión	5	4	3	3	1

- 4) a) Representa el diagrama de dispersión.  
b) Indica el tipo de correlación.

**Resolución**

a) La nube de puntos es la siguiente:



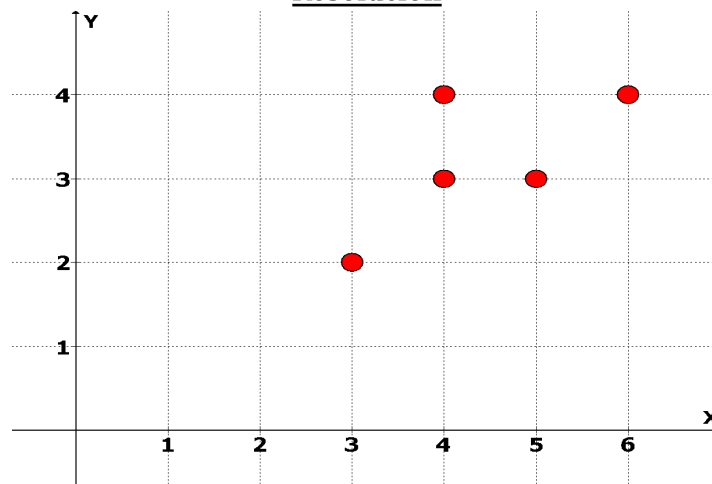
b) Como al aumentar el número de horas de estudio disminuye el número de horas de televisión, la correlación es negativa.

5) En los siguientes casos dibuja el diagrama de dispersión

a) Se ha realizado una encuesta preguntando por el número de personas que habitan en el hogar familiar y el número de dormitorios que tiene la casa. La tabla siguiente recoge la información obtenida:

X = número de personas	3	5	4	6	4
Y = número de dormitorios	2	3	4	4	3

**Resolución**

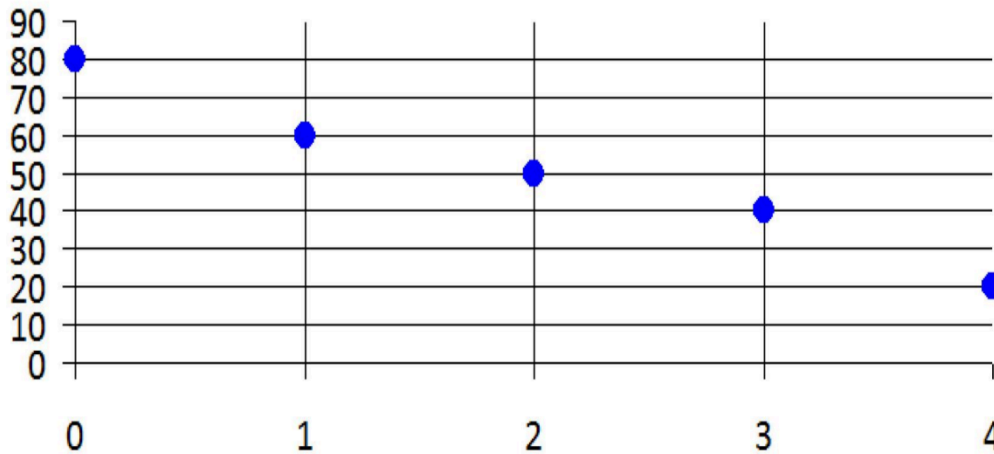


b) El número de gérmenes por  $\text{cm}^3$  en un enfermo que se está curando viene dado por la tabla

X = horas	0	1	2	3	4
-----------	---	---	---	---	---

Y = número de gérmenes	8	6	5	4	2
	0	0	0	0	0

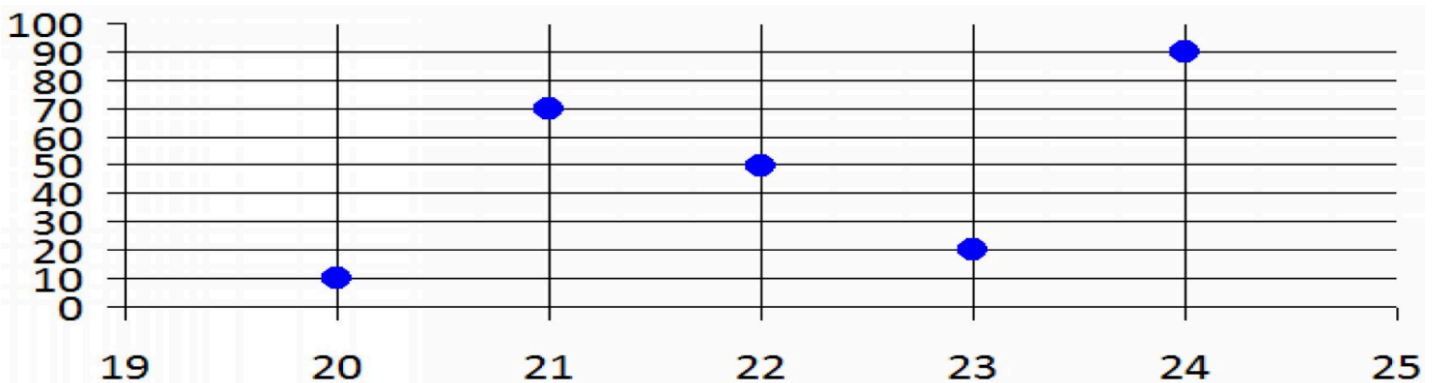
**Resolución**



c) El número de libros vendidos en una librería y la temperatura del día

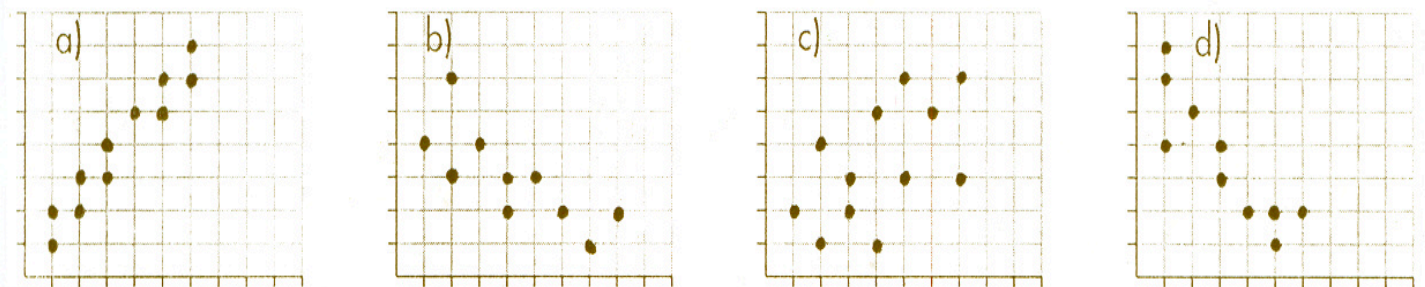
X = temperatura (°C)	2	2	2	2	2	2
	0	1	2	3	4	5
Y = número de libros	1	7	5	2	9	1
	0	0	0	0	0	0

**Resolución**



6) Resuelve las siguientes cuestiones:

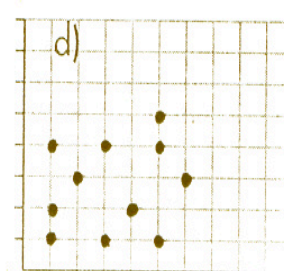
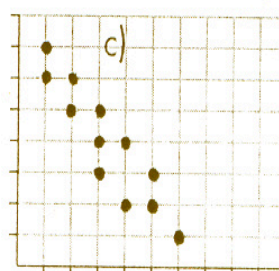
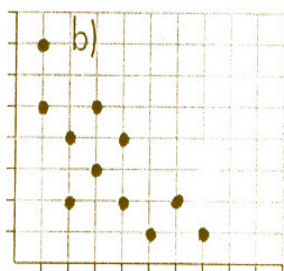
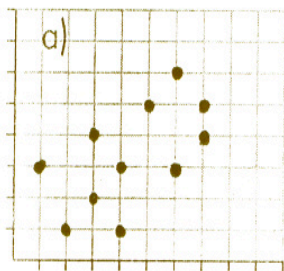
a) Los números 0,2; -0,9; -0,7 y 0,6 corresponden a los coeficientes de correlación de las siguientes distribuciones bidimensionales. Asigna a cada gráfica el suyo:



**Resolución**

Solución: a) → 0,6    b) → -0,7    c) → 0,2    d) → -0,9

b) Los coeficientes de correlación de estas distribuciones bidimensionales, son, en valor absoluto, 0,55; 0,75; 0,87 y 0,96. Asigna a cada uno el suyo, cambiando el signo cuando proceda.



**Resolución**

Solución: a) → 0,75    b) → -0,87    c) → -0,96    d) → 0,55

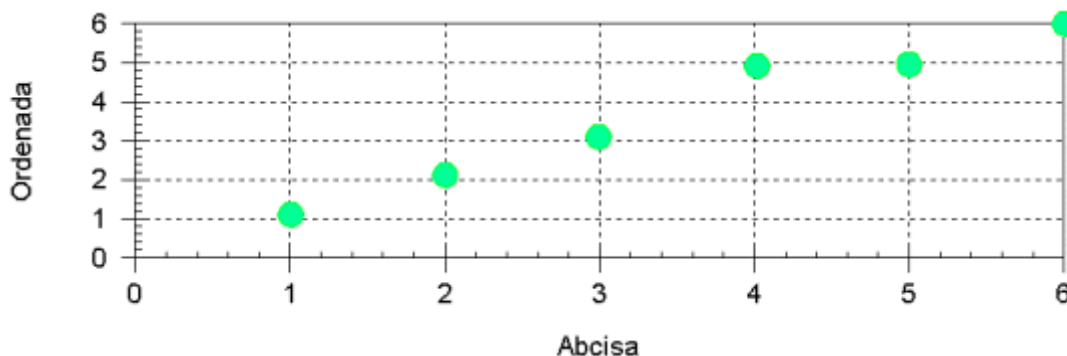
Más actividades

Una nube de puntos o diagrama de dispersión consiste en representar cada par de valores de las variables en un sistema de coordenadas cartesianas en el que los ejes X e Y representan las variables de la distribución bidimensional.

Por ejemplo, supongamos los siguientes datos:

x	1	2	3	4	5	6
y	1	2	3	5	5	6

La representación gráfica correspondiente sería:



De esta forma es sencillo verificar el tipo de relación que se establece entre las dos variables.

**Relación funcional**

Dos variables x e y están relacionadas funcionalmente cuando conocida la primera se puede saber con exactitud el valor de la segunda.

**Ejemplos**

Si se deja caer una piedra, existe una fórmula que nos permite calcular exactamente, la altura a la que se encuentra en función del tiempo transcurrido.

$$h = \frac{1}{2} g t^2$$

**Relación estadística**

Dos variables x e y están relacionadas estadísticamente cuando conocida la primera se puede estimar aproximadamente el valor de la segunda.

**Ejemplo:**

Ingresos y gastos de una familia.

Producción y ventas de una fábrica.

Gastos en publicidad y beneficios de una empresa.

**Variable estadística bidimensional**

Una variable bidimensional es una variable en la que cada individuo está definido por un par de caracteres, (X, Y).

Estos dos caracteres son a su vez variables estadísticas en las que sí existe relación entre ellas, una de las dos variables es la variable independiente y la otra variable dependiente.

Distribuciones bidimensionales

Son aquellas en las que a cada individuo le corresponden los valores de dos variables, las representamos por el par  $(x_i, y_i)$ .

Si representamos cada par de valores como las coordenadas de un punto, el conjunto de todos ellos se llama nube de puntos o diagrama de dispersión.

Sobre la nube de puntos puede trazarse una recta que se ajuste a ellos lo mejor posible, llamada recta de regresión.

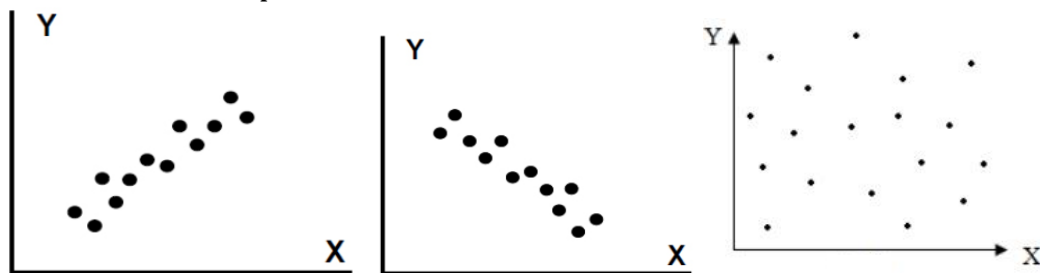
Ejemplo:

Las notas de 12 alumnos de una clase en Matemáticas y Física son las siguientes:

Matemáticas	2	3	4	4	5	6	6	7	7	8	10	10
Física	1	3	2	4	4	4	6	4	6	7	9	10

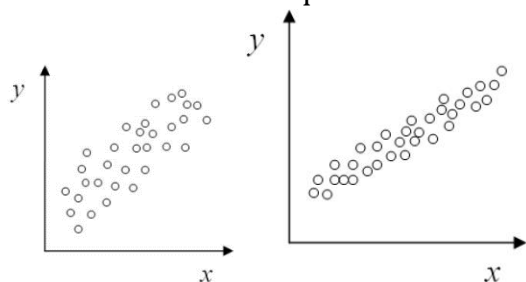
**TEST ESTADÍSTICA BIDIMENSIONAL**

Indica cómo es el tipo de correlación entre las variables X e Y



- positiva o directa
- negativa o inversa
- nula

Selecciona el caso en que la correlación entre las variables sea más débil



¿Qué tipo de correlación hay entre las variables X e Y si el coeficiente de correlación  $r = 0$   $r$  está entre  $-1$  y  $-0,75$   $r$  está entre  $-0,75$  y  $-0,25$   $r$  está entre  $-0,25$  y  $0$   $r$  está entre  $0$  y  $0,25$   $r$  está entre  $0,25$  y  $0,75$   $r$  está entre  $0,75$  y  $1$ ?

- nula
- directa fuerte
- directa media
- directa débil

inversa fuerte  
 inversa media  
 inversa débil

Si la varianza de X es 4, la de Y es 9 y el coeficiente de correlación es  $-0,4$  entonces la covarianza es ...

2,4  
 9  
 $-0,1$   
 $v-2,4$   
 $-0,4$   
 $-1$   
 $-15$   
 2  
 3

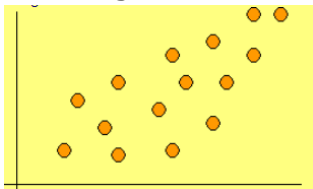
Una de las siguientes afirmaciones sobre el coeficiente de correlación  $r$  es falsa. Indícala.

Si  $r = 0$ , la correlación es nula  
 Si la covarianza es positiva,  $r$  también lo es.  
 $v$  Si  $r = -0,8$ , la correlación es muy débil.  
 Si  $r = 0,9$ , la correlación es directa y fuerte.

Si de las variables X e Y se sabe que  $S_x = 2$ ,  $S_y = 3$   $S_{xy} = -5,4$  entonces el coeficiente de correlación es igual a ...

0,9  
 $v-0,9$   
 0,81  
 $-0,81$

De los siguientes coeficientes de correlación di cuál corresponde al diagrama de dispersión de la figura



$-0,7$   
 $v0,6$   
 $-1$   
 0

La recta que mejor se ajusta a una nube de puntos se llama ...

recta del regreso  
 recta de la covarianza  
 $v$ recta de regresión  
 recta de correlación

La recta de regresión de y sobre x ...

$v$ permite estimar los valores de y dados los de x  
 permite estimar los valores de x dados los de y  
 no permite realizar estimaciones

La recta de regresión de  $x$  sobre  $y$ ,  
permite estimar los valores de  $y$  dados los de  $x$   
permite estimar los valores de  $x$  dados los de  $y$   
no permite realizar estimaciones

En qué caso es más fuerte la correlación entre las variables  
cuando el ángulo que forman las rectas de regresión es muy pequeño  
cuando el ángulo que forman las rectas de regresión es muy grande

La correlación trata de establecer la relación o dependencia que existe entre las dos variables que intervienen en una distribución bidimensional.

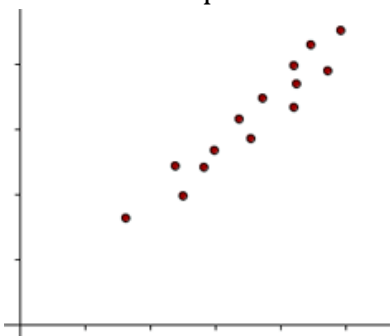
Es decir, determinar si los cambios en una de las variables influyen en los cambios de la otra. En caso de que suceda, diremos que las variables están correlacionadas o que hay correlación entre ellas.

Tipos de correlación

### 1º Correlación directa

La correlación directa se da cuando al aumentar una de las variables la otra aumenta.

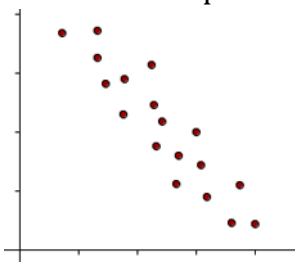
La recta correspondiente a la nube de puntos de la distribución es una recta creciente.



### 2º Correlación inversa

La correlación inversa se da cuando al aumentar una de las variables la otra disminuye.

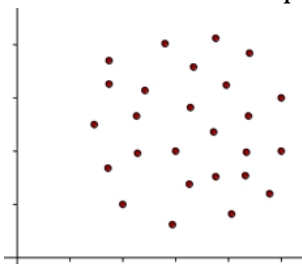
La recta correspondiente a la nube de puntos de la distribución es una recta decreciente.



### 3º Correlación nula

La correlación nula se da cuando no hay dependencia de ningún tipo entre las variables.

En este caso se dice que las variables están incorreladas y la nube de puntos tiene una forma redondeada.

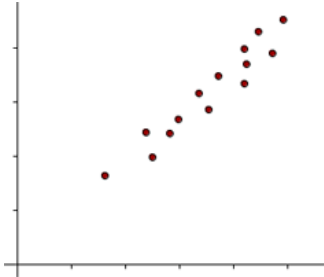


Grado de correlación

El grado de correlación indica la proximidad que hay entre los puntos de la nube de puntos. Se pueden dar tres tipos:

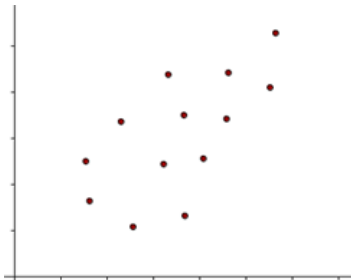
1. Correlación fuerte

La correlación será fuerte cuanto más cerca estén los puntos de la recta.

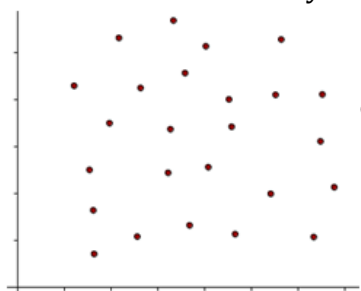


2. Correlación débil

La correlación será débil cuanto más separados estén los puntos de la recta.

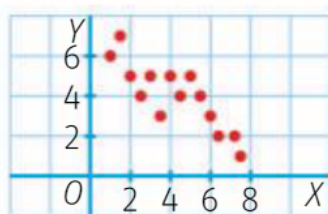
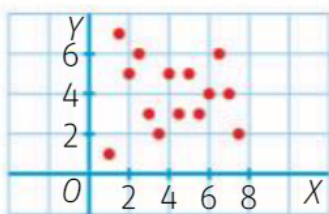


3. Correlación muy débil o prácticamente nula



\*\*\*\*\*

Observa los siguientes diagramas de dispersión e indica el tipo de relación entre las variables.



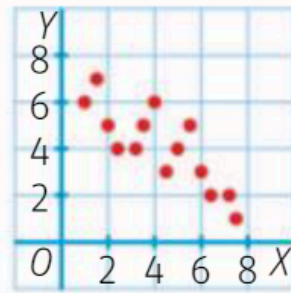
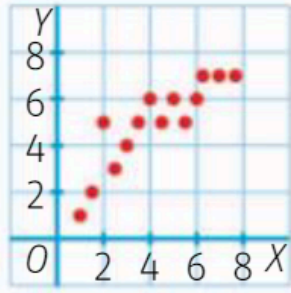
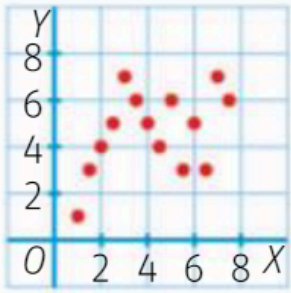
Solución

a) Correlación débil.

b) Correlación fuerte inversa.

\*\*\*\*\*

Asocia los coeficientes de correlación  $r$ , con su nube de puntos:



A.  $r = -0,8$

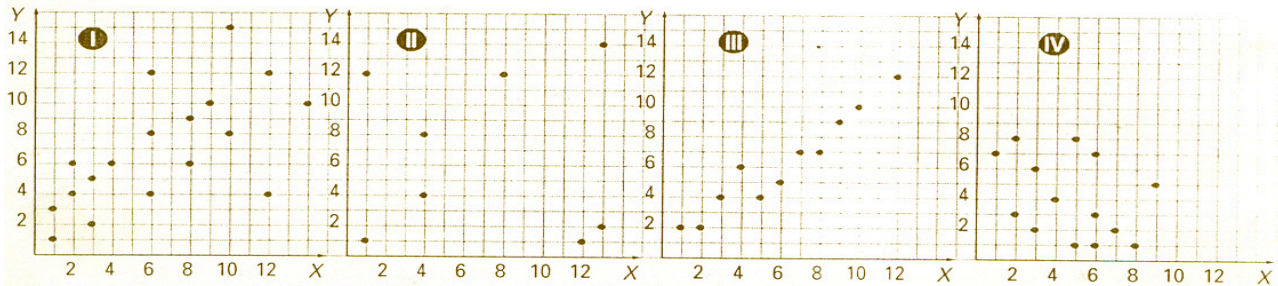
B.  $r = 0,4$

C.  $r = 0,9$

Solución

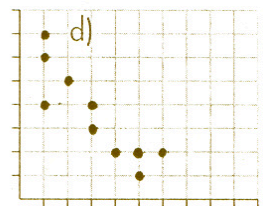
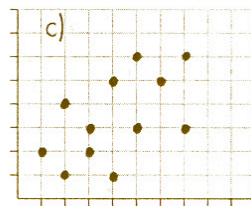
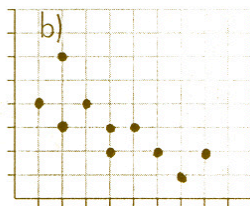
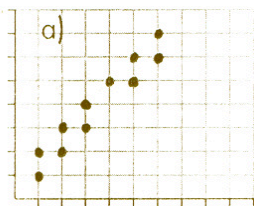
- a) La primera nube de puntos no se ajusta claramente a ninguna recta, por lo que su coeficiente de correlación lineal debe ser  $r = 0,4$ .
- b) La segunda nube de puntos se ajusta a una recta de pendiente positiva, por lo que su coeficiente de correlación debe ser  $r = 0,9$ .
- c) La tercera nube de puntos se ajusta a una recta de pendiente negativa, por lo que su coeficiente de correlación debe ser  $r = -0,8$ .

Los coeficientes de correlación de 4 distribuciones bidimensionales son (no necesariamente por ese orden):  $-0,04$ ;  $0,96$ ;  $0,65$  y  $-0,37$ . ¿Cuál corresponde a cada gráfico?



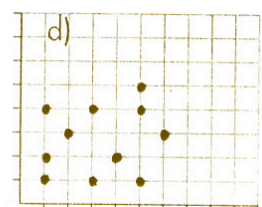
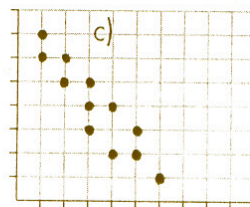
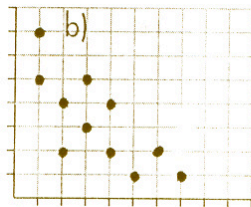
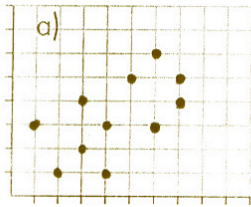
Solución: I  $\rightarrow 0,65$     II  $\rightarrow -0,04$     III  $\rightarrow 0,96$     IV  $\rightarrow -0,37$

Los números  $0,2$ ;  $-0,9$ ;  $-0,7$  y  $0,6$  corresponden a los coeficientes de correlación de las siguientes distribuciones bidimensionales. Asigna a cada gráfica el suyo:



Solución: a)  $\rightarrow 0,6$     b)  $\rightarrow -0,7$     c)  $\rightarrow 0,2$     d)  $\rightarrow -0,9$

Los coeficientes de correlación de estas distribuciones bidimensionales, son, en valor absoluto,  $0,55$ ;  $0,75$ ;  $0,87$  y  $0,96$ . Asigna a cada uno el suyo, cambiando el signo cuando proceda.



Solución : a)  $\rightarrow 0,75$     b)  $\rightarrow -0,87$     c)  $\rightarrow -0,96$     d)  $\rightarrow 0,55$

¿Qué tipo de correlación existe en cada caso?

a)  $r = 1$                       b)  $r = 21$                       c)  $r = -0,1$                       d)  $r = 0,5$

\*\*\*\*\*

Obtén la nube de puntos y estudia la correlación de los siguientes datos:

<b>x</b>	2	3	4	4	5	5	6	7	8	8
<b>y</b>	7	6	5	6	3	4	4	2	2	1

\*\*\*\*\*

Construye la tabla de doble entrada correspondiente a los siguientes datos:

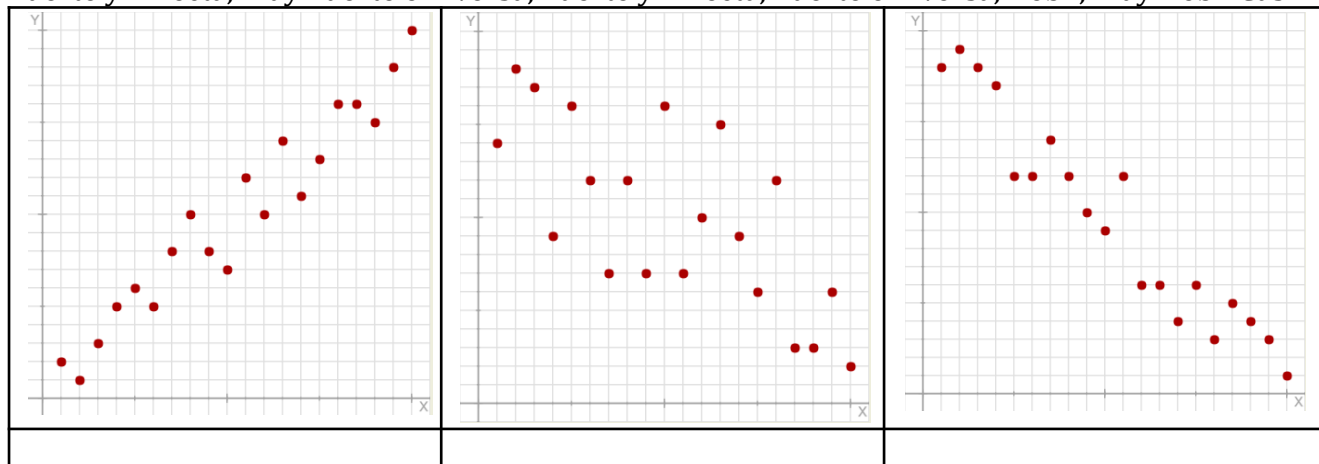
(8,0), (8,1), (9,2), (9,4), (10,2), (10,2), (13,2), (15,3)

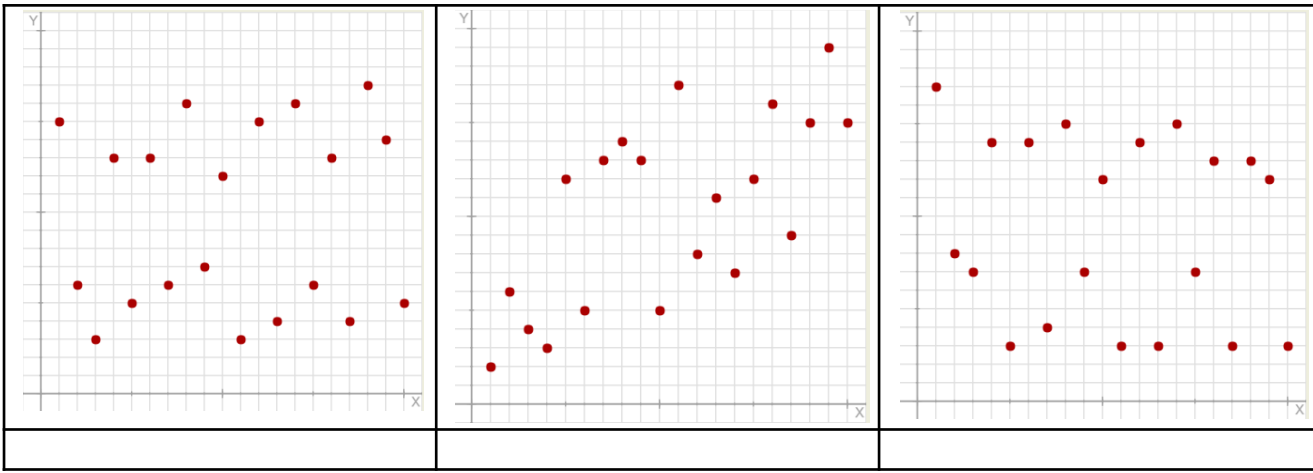
\*\*\*\*\*

X	4	5	5	5	6	6	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	9	9	9	9	9	9	1
Y	3	5	5	6	7	7	7	7	7	7	8	8	8	8	7	7	8	8	8	8	8	8	8	8	8	8	1	1	1
																										0	0	0	

\*\*\*\*\*

Indica de qué tipo de correlación es cada uno de las siguientes nubes de puntos, entre las opciones: Muy Fuerte y Directa, Muy Fuerte e Inversa, Fuerte y Directa, Fuerte e Inversa, Débil, Muy Débil Casi Nula.



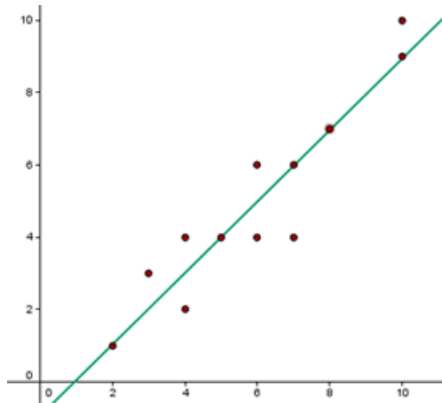
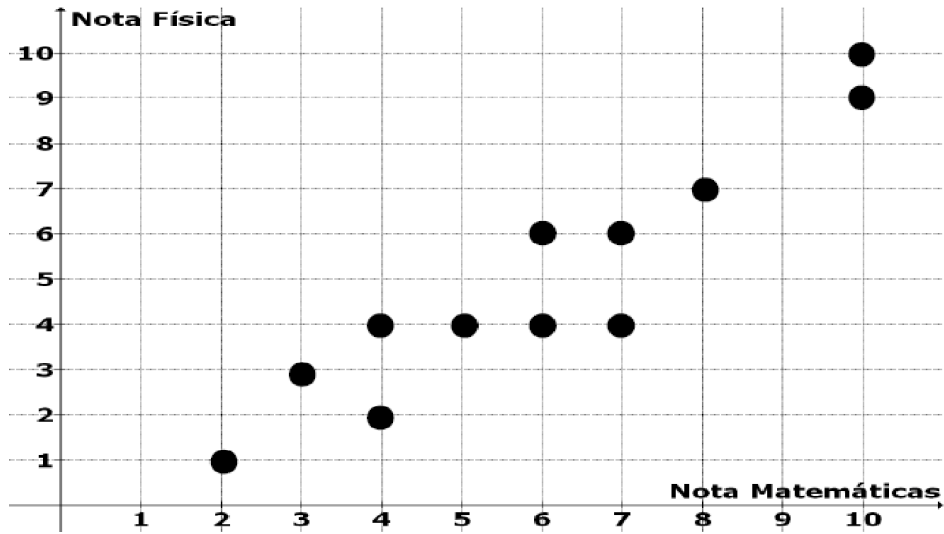


Completa la siguiente tabla con los tipos de correlación en función de los valores de "r":

Si $r = 1$ ó $r = -1$	
Si $0,5 < r < 1$	
Si $-1 < r < 0,5$	
Si $r = 0$ ó próximo	

Notas de 12 alumnos en Matemáticas y Física

Alumno	a	b	c	d	e	f	g	h	i	j	k	l
Matemáticas	2	3	4	4	5	6	6	7	7	8	10	10
Física	1	3	2	4	4	4	6	4	6	7	9	10

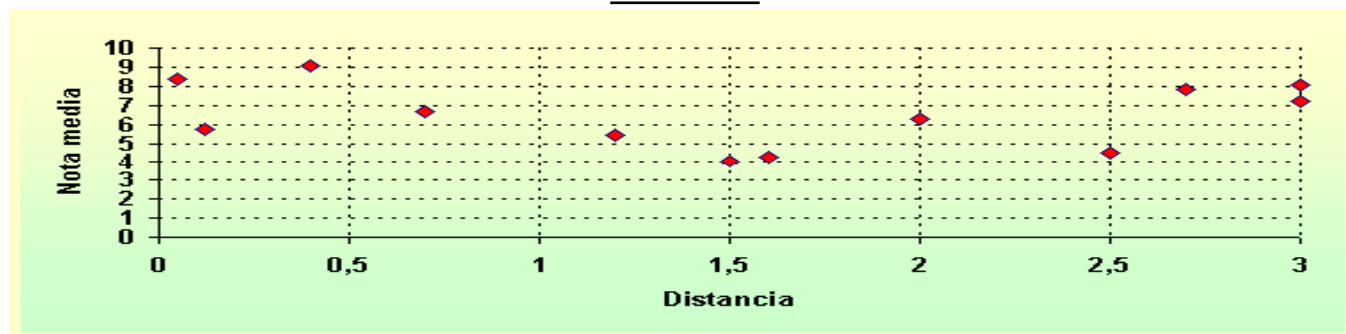


1) A 12 alumnos de un centro se le preguntó a qué distancia estaba su residencia del Instituto, con el fin de estudiar si esta variable estaba relacionada con la nota media obtenida. Se obtuvieron los datos que figuran en la siguiente tabla:

X = distancia (km)	0,0	0,1	0,	0,	1,	1,	1,	2	2,	2,	3	3
Y = nota media	8,4	5,7	9,	6,	5,	4	4,	6,	4,	7,	7,	8,
			1	7	4		1	2	5	9	2	1

Dibuja el diagrama de dispersión

**Resolución**



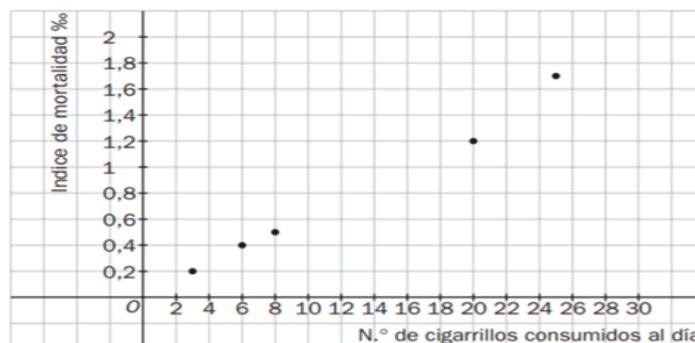
Observa la siguiente variable bidimensional.

N.º de cigarrillos consumidos al día	3	6	8	20	25
Índice de mortalidad	0,2	0,4	0,5	1,2	1,7

- 3) a) Representa la nube de puntos.  
b) Indica el tipo de correlación.

**Resolución**

a) La nube de puntos es la siguiente:



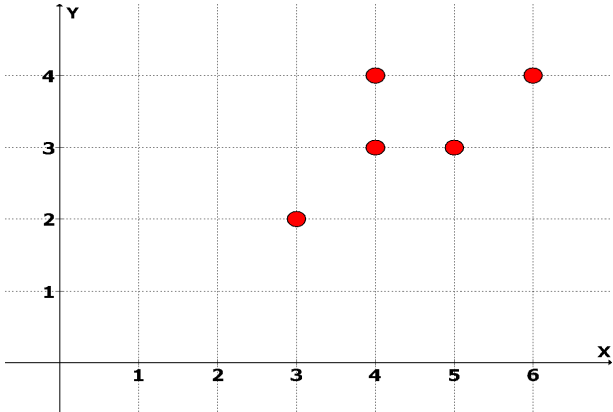
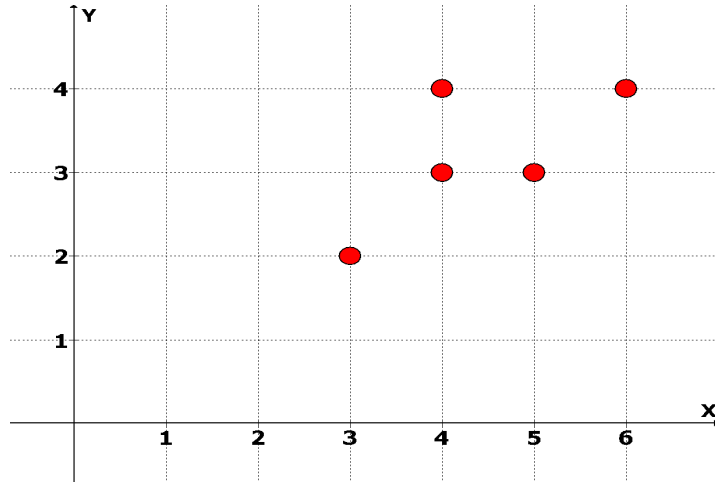
b) Como al aumentar el número de cigarrillos consumidos aumenta el índice de mortalidad, la correlación es positiva.

5) En los siguientes casos dibuja el diagrama de dispersión

a) Se ha realizado una encuesta preguntando por el número de personas que habitan en el hogar familiar y el número de dormitorios que tiene la casa. La tabla siguiente recoge la información obtenida:

X = número de personas	3	5	4	6	4
Y = número de dormitorios	2	3	4	4	3

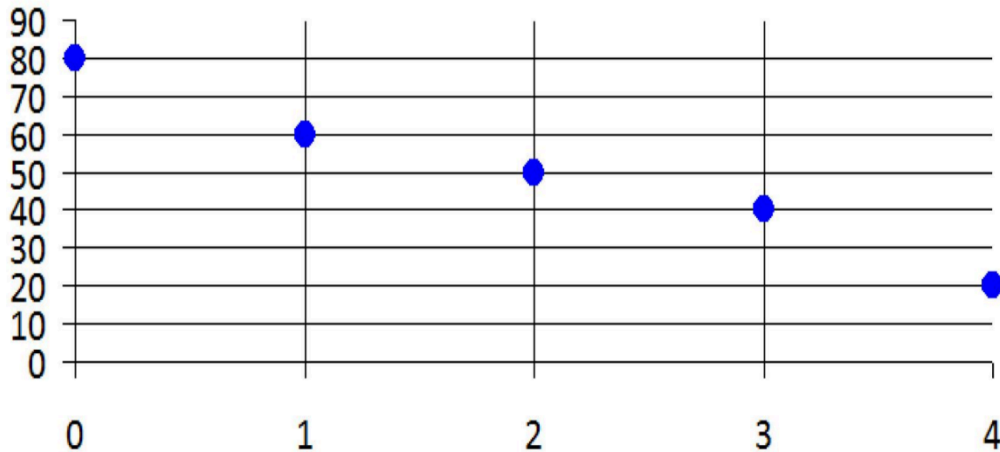
**Resolución**



b) El número de gérmenes por  $\text{cm}^3$  en un enfermo que se está curando viene dado por la tabla

X = horas	0	1	2	3	4
Y = número de gérmenes	8	6	5	4	2
	0	0	0	0	0

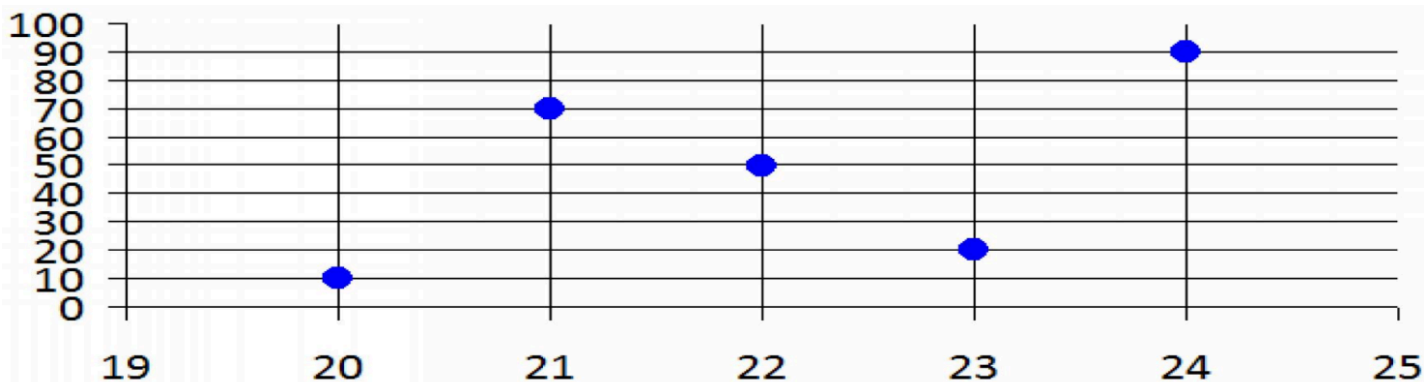
**Resolución**



c) El número de libros vendidos en una librería y la temperatura del día

X = temperatura ( $^{\circ}\text{C}$ )	2	2	2	2	2	2
Y = número de libros	0	1	2	3	4	5
	1	7	5	2	9	1
	0	0	0	0	0	0

**Resolución**



El precio de las fotocopias de una copistería

$x_i = n^{\circ}$ de copias	1	2	3	4	5
$y_i =$ precio (ctmos)	3	6	9	12	15

Se han medido los pesos y las alturas de 6 personas, obteniéndose los datos de la tabla adjunta.

Pesos (kg)	Alturas (m)
65	1'7
60	1'5
63	1'7
63	1'7
68	1'75
68	1'8

En las bibliotecas de 6 poblaciones se han analizado la afluencia de lectores ( $x$  expresadas en miles de personas) y el número de libros prestados ( $Y$ ), obteniéndose los datos de la tabla:

$X$	0,50	1	1,3	1,7	2	2,5
$Y$	180	240	250	300	340	400

\*\*\*\*\*

En un barrio de cierta ciudad se ha dado la siguiente distribución.

Año	Jóvenes mayores de 15 años	Estudiantes de Bachillerato
2001	40	12
2002	37	15
2003	30	13
2004	39	14
2005	42	17
2006	45	16
2007	50	20
2008	55	18

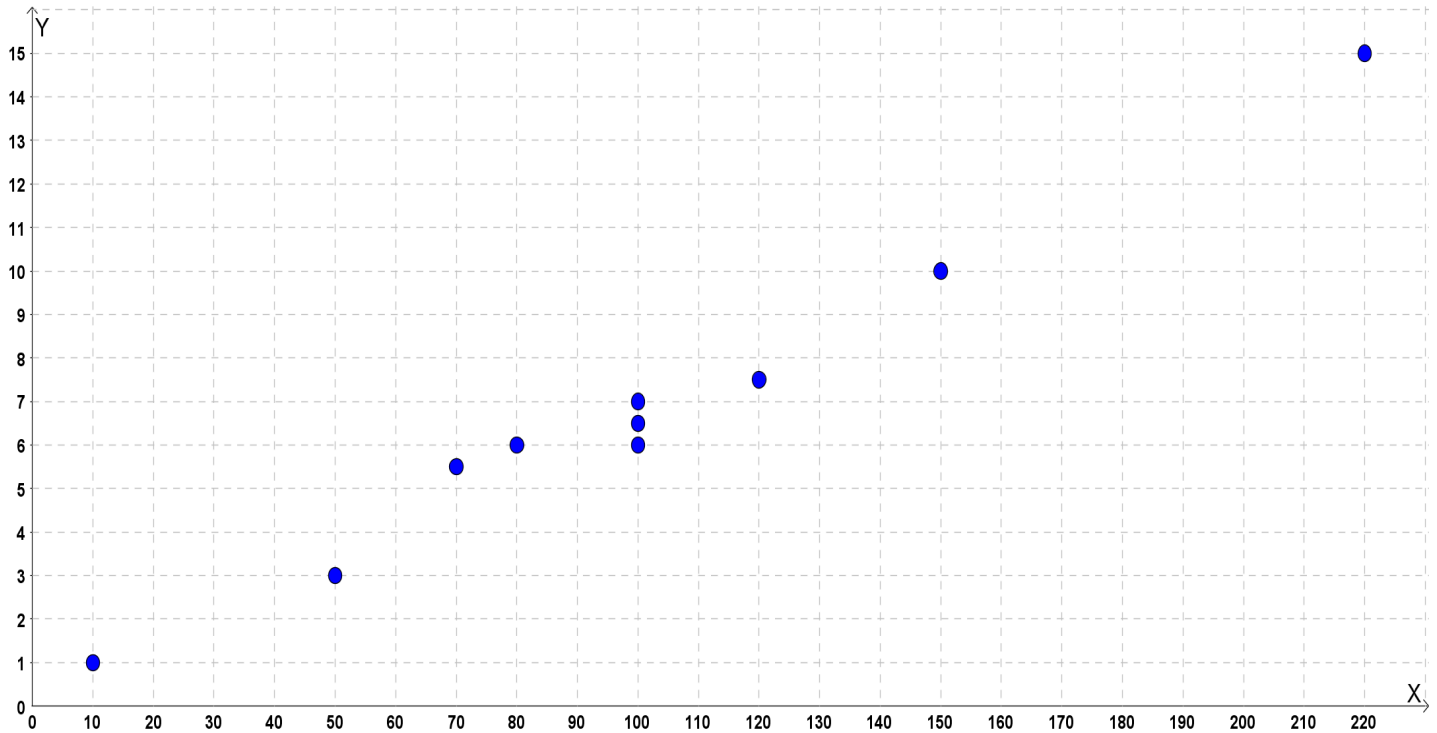
\*\*\*\*\*

Durante 10 días hemos realizado mediciones sobre el consumo de un coche

(litros consumidos y kilómetros recorridos). Los datos obtenidos han sido los siguientes:

X: km	10	8	5	10	1	10	70	12	15	22
	0	0	0	0	0	0		0	0	0
Y: litros	6,5	6	3	6	1	7	5,5	7,5	10	15

Dibuja la nube de puntos y estudia gráficamente la correlación entre los litros consumidos y los kilómetros recorridos.



La correlación es directa y fuerte

Los resultados de una encuesta realizada a un grupo de alumnos sobre el número de horas que se estudia al día y el número de suspensos obtenidos en una evaluación, se reflejan en la siguiente tabla:

Nº horas (x)	2	0	1	0,5	1	2	3	1,5	2,5	2	1	1,5	0,5	2	1,5	0	0,5	1,5	1	3
Nº suspensos (y)	1	6	3	4	2	0	0	1	0	3	1	0	3	2	2	5	5	3	2	0

Representar mediante una nube de puntos estos resultados e indicar el tipo de correlación.

\*\*\*\*\*

A un enfermo que tiene mucha fiebre se le administra un medicamento para que ésta descienda. Se observa al paciente controlándole la temperatura cada 30 minutos, durante 6 horas. Apuntamos los datos en una variable bidimensional en la cual:

X = Tiempo transcurrido desde la administración del medicamento

Y = Temperatura del paciente

Si el medicamento ha sido efectivo, ¿qué tipo de correlación existe entre estas dos variables, positiva o negativa?

\*\*\*\*\*

Se desea investigar el ganado caprino y el ganado ovino de un país. En la tabla de doble entrada adjunta se presentan los resultados de un estudio de explotaciones ganaderas, seleccionadas aleatoriamente del censo agropecuario. Se proporcionan las frecuencias conjuntas del número de cabezas (en miles) de cabras (X) y ovejas (Y) que poseen las explotaciones.

•	X	•	•	1	•	•
\Y						

•	0	•	•	€	•	•
•	1	•	•	1	•	•
			0			
•	2	•	•	€	•	•
•	3	•	•	5	•	•
•	4	•	•	3	•	•

Haz el diagrama de dispersión indicando en cada punto su frecuencia absoluta  
Indica qué tipo de correlación existe entre las variables

\*\*\*\*\*

Sueldo medio en España de varones y mujeres en la industria y los servicios, según los datos de 1999 ofrecidos por el INE:

CC.AA.	Sueldo medio	
	Varones	Mujeres
Andalucía	248.389	158.901
Aragón	282.054	156.485
Asturias (Principado de)	275.406	177.203
Baleares (Islas)	253.681	176.835
Canarias	217.843	167.953
Cantabria	270.570	163.153
Castilla y León	260.336	171.002
Castilla-La Mancha	226.887	146.525
Cataluña	281.496	195.771
Comunidad Valenciana	244.350	159.117
Extremadura	220.644	133.952
Galicia	229.395	163.609
Madrid (Comunidad de)	308.122	235.456
Murcia (Región de)	218.924	144.544
Navarra (C. Foral de)	289.006	195.560
País Vasco	322.222	232.367
Rioja (La)	255.193	166.257

\*\*\*\*\*

Días de la semana	L	M	X	J	V	S	D
Temperatura máxima	32	31	32	33	32	31	33
Temperatura mínima	22	21	24	25	20	21	23

\*\*\*\*\*

Sea la siguiente distribución bidimensional de las variables X (medio de transporte) e Y (tiempo en minutos del recorrido)

Recuento		tiempo (mn)					Total
		0-10	10-20	20-30	30-40	Más de 40	
transp	Metro	15	12	17	8	8	60
	Bus	6	8	14	11	1	40
Total		21	20	31	19	9	100

Selecciona las frases que sean verdaderas:  
 vEl 25% utilizan el autobús y tardan entre 20 y 40 minutos.  
 Utilizan el autobús el 8% de los que tardan entre 10 y 20 minutos

- vEl 27,5% de los usuarios de autobús tardan entre 30 y 40 minutos.
- El 50% de los que tardan entre 20 y 40 minutos utilizan el metro.
- vDe los que usan el metro, la quinta parte tarda entre 10 y 20 minutos
- v Los 7/18 de los que tardan de 0 a 30 minutos usan el autobús

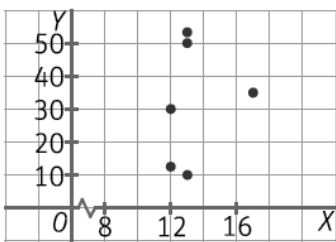
Las previsiones de la Agencia Estatal de Meteorología AEMET para una semana en Zaragoza han sido:

M 15				X 16		J 17		V 18	S 19
00-06 h	06-12 h	12-18 h	18-24 h	00-12 h	12-24 h	00-12 h	12-24 h		
6°C	9°C	10°C	6°C						
Probabilidad de precipitación									
75%	80%	5%	0%	5%	15%	5%	20%	50%	35%
Cota de nieve (m)									
900	900				1000	900	1300	1700	
Temperatura mínima y máxima (°C)									
6 / 13				5 / 13		2 / 12		2 / 13 6 / 17	

- a) Dibuja la nube de puntos que relaciona la temperatura máxima con la probabilidad media de precipitación cada día.
- c) ¿Existe alguna correlación entre la temperatura máxima y la probabilidad de precipitaciones? ¿De qué tipo?

Solución

a)



- c) Existe correlación directa entre ambas variables, pero muy débil

\*\*\*\*\*

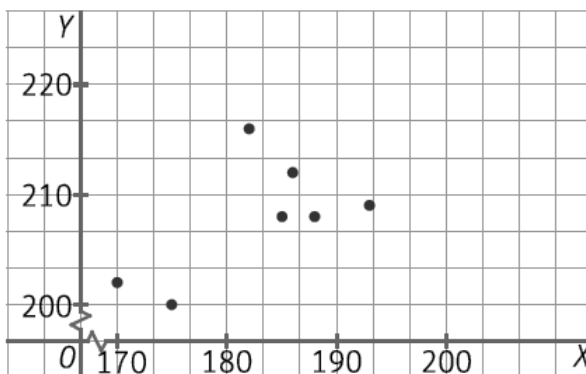
Luis quiere comprarse una moto de 1000 c. c. Compara precios, potencia y peso de varios modelos.

Marca	€	CV	kg
Tuzuki	14 300	185	208
Conda	15 700	175	199
MMM	16 500	170	202
ZVK	17 500	193	209
Tamiha	16 000	182	216
Asawaki	14 600	188	208
MV Felicia	18 500	186	212

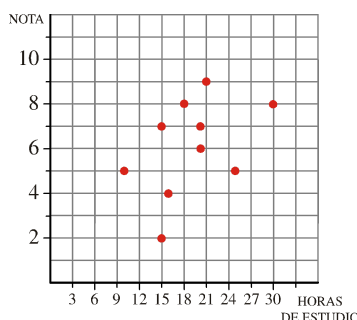
- b) Representa la nube de puntos en la que X sea la potencia, e Y, el peso. ¿Qué puedes indicar de la correlación entre ambas?

Solución

b) La nube de puntos es la siguiente:



Se ha representado el diagrama de dispersión correspondiente a la nota en un examen y las horas de estudio de un grupo de jóvenes



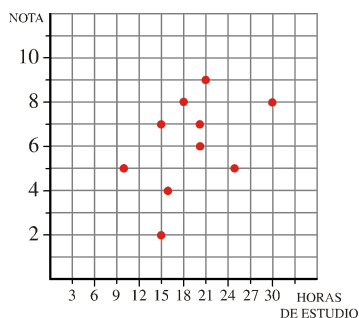
Indica cuál de estos valores es más apropiado para el coeficiente de correlación: 0,92; -0,44; -0,92; 0,44

Un grupo de 10 amigos se ha presentado a una prueba de oposición. Anotaron el número de horas que dedicaron a estudiar la semana antes del examen y la nota obtenida en la prueba. La información se recoge en la siguiente tabla:

<b>Horas de estudio</b>	<b>21</b>	<b>15</b>	<b>10</b>	<b>15</b>	<b>20</b>	<b>30</b>	<b>18</b>	<b>20</b>	<b>25</b>	<b>16</b>
<b>Nota</b>	<b>9</b>	<b>7</b>	<b>5</b>	<b>2</b>	<b>7</b>	<b>8</b>	<b>8</b>	<b>6</b>	<b>5</b>	<b>4</b>

Representa los datos mediante una nube de puntos e indica cuál de estos valores te parece más apropiado para el coeficiente de correlación: 0,92; -0,44; -0,92; 0,44.

Solución:



Observando la representación, vemos que el coeficiente de correlación es positivo y bajo. Por tanto,  $r = 0,44$ .

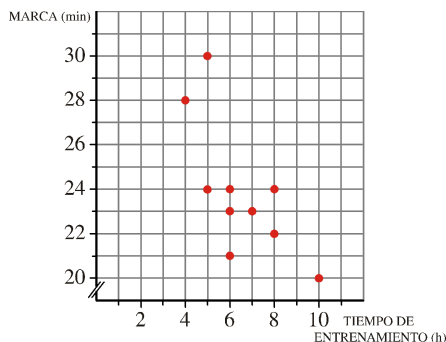
\*\*\*\*\*

Se ha medido el número medio de horas de entrenamiento a la semana de un grupo de 10 atletas y el tiempo, en minutos, que han hecho en una carrera, obteniendo los siguientes resultados:

<b>Horas de entrenamiento</b>	5	6	6	5	8	6	8	10	7	4
<b>Tiempo carrera</b>	30	23	24	24	22	21	24	20	23	28

Representa los datos mediante una nube de puntos y di cuál de estos valores te parece más apropiado para el coeficiente de correlación: 0,71; -0,71; 0,45; -0,32.

Solución:



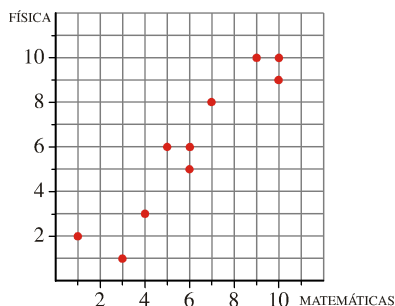
A la vista de la representación, observamos que el coeficiente de correlación,  $r$ , es negativo y relativamente alto. Por tanto,  $r \bullet \approx 0,71$ .

Las notas de 10 alumnos y alumnas de una clase en Matemáticas y en Física han sido las siguientes:

<b>Matemáticas</b>	7	6	4	5	9	10	3	1	10	6
<b>Física</b>	8	6	3	6	10	9	1	2	10	5

Representa los datos mediante una nube de puntos y di cuál de estos valores te parece más apropiado para el coeficiente de correlación: 0,23; 0,94; -0,37; -0,94.

Solución:



Viendo la representación, observamos que el coeficiente de correlación es positivo y alto. Por tanto,  $r = 0,94$ .

Para realizar un estudio sobre la utilización de una impresora en un determinado departamento se midió en un día los minutos transcurridos entre las sucesivas utilizaciones (X) y el número de páginas impresas (Y), obteniéndose los siguientes resultados:

<b>X</b>	9	9	4	6	8	9	7	6	9	9	9	8	8	9	8	9
<b>Y</b>	3	8	3	8	3	8	8	8	3	8	12	12	8	8	8	12
<b>X</b>	9	9	10	9	15	10	12	12	10	10	12	10	10	12	12	10
<b>Y</b>	12	20	8	20	8	8	20	8	8	12	8	20	20	3	3	20

a) Escribir la distribución de frecuencias conjunta. ¿Cuál es el porcentaje de veces que transcurre más de nueve minutos desde la anterior utilización y se imprimen menos de doce páginas? ¿Cuántas veces se

-----  
imprimen menos de doce páginas y transcurren nueve minutos desde la anterior utilización?

- b) Frecuencias marginales. ¿Cuántas veces se imprimen como mucho doce páginas? ¿Cuántas páginas como máximo se imprimen en el 80% de las ocasiones?
- c) Hallar la distribución de frecuencias del número de páginas impresas condicionada a que han transcurrido nueve minutos entre sucesivas utilizaciones.
- d) Dibujar el diagrama de dispersión.