9.- En una empresa los empleados se encuentran clasificados en tres categorías: A, B y C, de tal forma que, en diciembre de 2010, se tenía:

	Número de empleados	Salario medio por mes (10^3€)	Desviación típica (10^3€)
A	130	145	22.5
B	50	200	42
C	20	300	70

a) Calcular el salario medio para el conjunto de la empresa. Dar un indicador de su representatividad.

b) Para fijar los salarios en 2012 se propusieron las tres alternativas siguientes:
   1. Elevación de todos los salarios en un 5%.
   2. Incremento lineal de 100€.
   3. Aumento de un 10% en la categoría A, un 8% en la categoría B y un 4% en la C.

   Determinar los salarios medios totales con cada alternativa, estudiando cual de ellas reduce más la dispersión relativa inicial de los salarios para el conjunto de la empresa.

$$\bar{x}_A = \frac{\sum_A n_i x_i}{n_A} \Rightarrow \sum_A n_i x_i = n_A \bar{x}_A = 130.145 = 18850$$

$$\sum_B n_i x_i = n_B \bar{x}_B = 50.200 = 10000 \quad \sum_C n_i x_i = n_C \bar{x}_C = 20.300 = 6000$$

$$\sum_T n_i x_i = 18850 + 10000 + 6000 = 34850 \quad \bar{x}_T = \frac{\sum_T n_i x_i}{n_T} = \frac{34850}{200} = 174{,}25$$

$$s_A^2 = \frac{\sum_A n_i x_i^2}{n_A} - \bar{x}_A^2 \Rightarrow \sum_A n_i x_i^2 = n_A\left(s_A^2 + \bar{x}_A^2\right) = 130\left(22{,}5^2 + 145^2\right) =$$

$$= 130(506{,}25 + 21025) = 2799062{,}5$$

$$\sum_B n_i x_i^2 = 50\left(42^2 + 200^2\right) = 2088200 \quad \sum_C n_i x_i^2 = 20\left(70^2 + 300^2\right) = 1898000$$

$$\sum_T n_i x_i^2 = 2799062{,}5 + 2088200 + 1898000 = 6785262{,}5$$

$$s_T^2 = \frac{\sum_T n_i x_i^2}{n_T} - \bar{x}_T^2 = \frac{6785262{,}5}{200} - 174{,}25^2 = 33926{,}3125 - 30363{,}0625 =$$

x_i	n_i	N_i	p_i	x_i*n_i	(xi*ni)↑	q_i
10	10	10	0,08	100	100	0,021
20	42	52	0,43	840	940	0,198
40	35	87	0,73	1400	2340	0,493
65	20	107	0,89	1300	3640	0,767
85	13	120		1105	4745	
Total	120		2,13	4745	4745	1,479

$$I_G = \frac{\sum_{i=1}^{k-1}(p_i - q_i)}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{\sum_{i=1}^{k-1} q_i}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{1,479}{2,13} = 0,306$$

$= 3563{,}25 \quad s_T = \sqrt{3563{,}25} = 59{,}69 \quad CV_T = \frac{s_T}{\bar{x}_T} = \frac{59{,}69}{174{,}25} = 0{,}3425$

B) $\bar{x}_{T1} = 1{,}05.174{,}25 = 182{,}98 \quad \bar{x}_{T2} = 174{,}25 + 0{,}1 = 174{,}35$

$\sum_T n_i x_i = 1{,}1.18850 + 1{,}08.10000 + 1{,}04.6000 =$

$= 20735 + 10800 + 6240 = 37775$

$$\bar{x}_{T3} = \frac{\sum_{T3} n_i x_i}{n_T} = \frac{37775}{200} = 188{,}875$$

En la alternativa 1 $CV_{T1} = CV_T$ puesto que hemos efectuado sólo un cambio de escala, y el cociente de variación de Pearson, permanece invariante frente a los cambios de escala.

En la alternativa 2 hemos efectuado sólo un cambio de origen de trabajos y la varianza permanece igual luego $CV_{T2} = \frac{s_{T2}}{\bar{x}_{T2}} = \frac{59{,}69}{174{,}35} = 0{,}34235$ y la dispersión relativa experimenta una ligera disminución

En la alternativa 3 hemos efectuado sólo un cambio de escala en cada grupo, que no altera las varianzas iniciales, luego la varianza total sigue siendo la misma $CV_{T3} = \frac{s_{T3}}{\bar{x}_{T3}} = \frac{59{,}69}{188{,}875} = 0{,}317$ luego esta alternativa es la que más reduce la dispersión relativa.

10.- Dada la siguiente distribución de salarios mensuales en euros de los trabajadores de una empresa:

Salario	n_i
400 - 600	5
600 - 900	40
900 - 1400	50
1400 - 1800	15
1800 - 2500	7
2500 - 4000	3

Si se supone que los salarios se distribuyen uniformemente en los distintos intervalos, se pide:

a) Determinar el sueldo más frecuente y representar gráficamente la distribución.

b) Calcular el coeficiente de variación de Pearson. Si este coeficiente suele estar en torno al 30% en empresas similares del entorno. ¿Se puede decir que en esta empresa los sueldos son más homogéneos que en las empresas similares del entorno?

c) Obtener el índice de concentración de Gini y representar la curva de Lorenz. ¿Se puede decir que los salarios de esta empresa están muy concentrados en unos pocos trabajadores?

d) Si dividimos a los trabajadores en dos grupos, los que más cobran y los que menos cobran, calcular el sueldo que delimitaría ambos grupos. Si la empresa solo dispone este mes de la mitad de la masa salarial y empieza pagando a los que menos cobran, calcular cuánto sería el último salario pagado.

e) Hacer un estudio de la forma de la distribución, obtener los coeficientes de asimetría y curtosis e interpretar los resultados.

a)

L.inf.	L.sup	n_i	h_i
4	6	5	2,5
6	9	40	13,33
9	14	50	10
14	18	15	3,75
18	25	7	1
25	40	3	0,2
Total		120	

$\hat{x} \in [6.9) \quad \hat{x} = 7,5$

b)

L.inf.	L.sup	xi	ni	Ni	xi-xo	x'i=(xi-xo)/a	ni*x'i	x'i^2	ni*x'i^2
4	6	5	5	5	-6,5	-1,3	-6,5	1,69	8,45
6	9	7,5	40	45	-4,0	-0,8	-32	0,64	25,6
9	14	11,5	50	95	0,0	0	0	0	0
14	18	16	15	110	4,5	0,9	13,5	0,81	12,15
18	25	21,5	7	117	10,0	2	14	4	28
25	40	32,5	3	120	21,0	4,2	12,6	17,64	52,92
Total			120				1,6		127,12

$$\bar{x}' = \frac{\sum n_i x_i'}{N} = \frac{1,6}{120} = 0,013 \quad \bar{x} = a\bar{x}' + x_0 = 5.0,013 + 11,5 = 11,567$$

$$s^2 = \frac{\sum n_i x_i'^2}{N} - \bar{x}'^2 = \frac{127,12}{120} - 0,013^2 = 1,0593 - 0,0003 = 1,0591$$

$$s^2 = a^2 s'^2 = 5^2.1,0591 = 26,275 \quad s = \sqrt{26,275} = 5,13$$

$$CV = \frac{s}{\bar{x}} = \frac{5,13}{11,567} = 0,44$$

Como $CV > 0,30$ no se puede decir que en esta empresa los sueldos son más homogéneos que en las empresas similares del entorno.

c)

xi	ni	Ni	pi	xi*ni	(xi*ni)↑	qi
5	5	5	0,04	25	25	0,018
7,5	40	45	0,38	300	325	0,2341
11,5	50	95	0,79	575	900	0,6484
16	15	110	0,92	240	1140	0,8213
21,5	7	117	0,98	150,5	1290,5	0,9298
32,5	3	120		97,5	1388	
Total	120		3,10	1388	1388	2,6517

$$I_G = \frac{\sum_{i=1}^{k-1}(p_i - q_i)}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{\sum_{i=1}^{k-1} q_i}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{2,6517}{3,10} = 0,145$$

No se puede decir que los salarios de esta empresa están muy concentrados en unos pocos trabajadores.

d) $\frac{N}{2} = 60 \quad x_M \in [9,14) \quad \left.\begin{matrix} 50 \_\_\_\_\_\_ 5 \\ 15 \_\_\_\_\_\_ x \end{matrix}\right\} x = \frac{5.15}{50} = 1,5 \quad x_M = 9 + 1,5 = 10,5$

La mitad de la masa salarial es $\frac{1388}{2} = 694$ y es superada en el intervalo $[9,14)$.

$$\left.\begin{array}{l} 575 \underline{\qquad\qquad} 5 \\ 369 \underline{\qquad\qquad} x \end{array}\right\} x = \frac{5.369}{575} = 3,21$$ El último salario pagado sería 900+321=1221€

e)

$$m_3 = a_3 - 3a_1a_2 + 2a_1^3$$

$$m_4 = a_4 - 4a_1a_3 + 6a_1^2a_2 - 3a_1^4$$

$$a_3 = \sum_{i=1}^{k} f_i x_i^3 = \frac{257,734}{120} = 2,15$$

$$m_3 = 2,15 - 3.0,013.0,053 + 2.0,013^3 = 2,15 - 0,002 + 0,000 = 2,13$$

$$A_F = \frac{m_3}{s_x^3} = \frac{2,13}{1,06^3} = 1,79$$ Asimetría positiva

$$a_4 = \sum_{i=1}^{k} f_i x_i^4 = \frac{1086,01}{120} = 9,05$$

$$m_4 = 9,05 - 4.0,013.2,13 + 60,013^2.0,053 - 30,013^4 =$$
$$= 9,05 - 0,111 + 0,000 - 0,0000 = 8,94$$

$$C_F = \frac{m_4}{s_x^4} = \frac{8,94}{1,06^4} = 7,08$$

Salarios													
Valores			frecuencia		11,5	5							
L.inf.	L.sup	x_i	n_i	N_i	x_i-x_0	x'_i=(x_i-x_0)/a	n_i*x'_i	x'_i^2	n_i*x'_i^2	x'_i^3	n_i*x'_i^3	x'_i^4	n_i*x'_i^4
4	6	5	5	5	-6,5	-1,3	-6,5	1,69	8,45	-2,197	-10,985	2,8561	14,2805
6	9	7,5	40	45	-4,0	-0,8	-32	0,64	25,6	-0,512	-20,48	0,4096	16,384
9	14	11,5	50	95	0,0	0	0	0	0	0	0	0	0
14	18	16	15	110	4,5	0,9	13,5	0,81	12,15	0,729	10,935	0,6561	9,8415
18	25	21,5	7	117	10,0	2	14	4	28	8	56	16	112
25	40	32,5	3	120	21,0	4,2	12,6	17,64	52,92	74,088	222,264	311,1696	933,5088
Total			120				1,6		127,12		257,734		1086,0148

Estaturas													
Datos cuantitativos continuos													
Valores			Frecuencias		1,725	0,05							
L.inf.	L.sup	x_i	n_i	N_i	x_i-x_o	x_i=(x_i-x_o)/a	n_i*x_i	x_i^2	n_i*x'_i^2	x_i^3	n_i*x'_i^3	x'_i^4	n_i*x'_i^4
1,55	1,60	1,575	5	5	-0,15	-3	-15	9	45	-27	-135	81	405
1,60	1,65	1,625	11	16	-0,10	-2	-22	4	44	-8	-88	16	176
1,65	1,70	1,675	13	29	-0,05	-1	-13	1	13	-1	-13	1	13
1,70	1,75	1,725	8	37	0,00	0	0	0	0	0	0	0	0
1,75	1,80	1,775	5	42	0,05	1	5	1	5	1	5	1	5
1,80	1,85	1,825	6	48	0,10	2	12	4	24	8	48	16	96
1,85	1,90	1,875	1	49	0,15	3	3	9	9	27	27	81	81
1,90	1,95	1,925	1		0,20	4	4	16	16	64	64	256	256
Total			50				-26		156		-92		1032

Parámetros de centralización	
Intervalo modal	[1,65-1,70)
Moda	1,664
Intervalo mediano	[1,65-1,70)
Mediana	1,68462
Media	1,699
Parámetros de posición	
Primer cuartil	1,634
Tercer cuartil	1,755
Parámetros de dispersión	
Varianza	0,007124
Desviación típica	0,084
Cociente de variación	0,050
Parámetros de forma	
Coeficiente de asimetría de Bowley	0,164
Coeficiente de asimetría de Fisher	0,571
Coeficiente de curtosis de Fisher	2,667