①

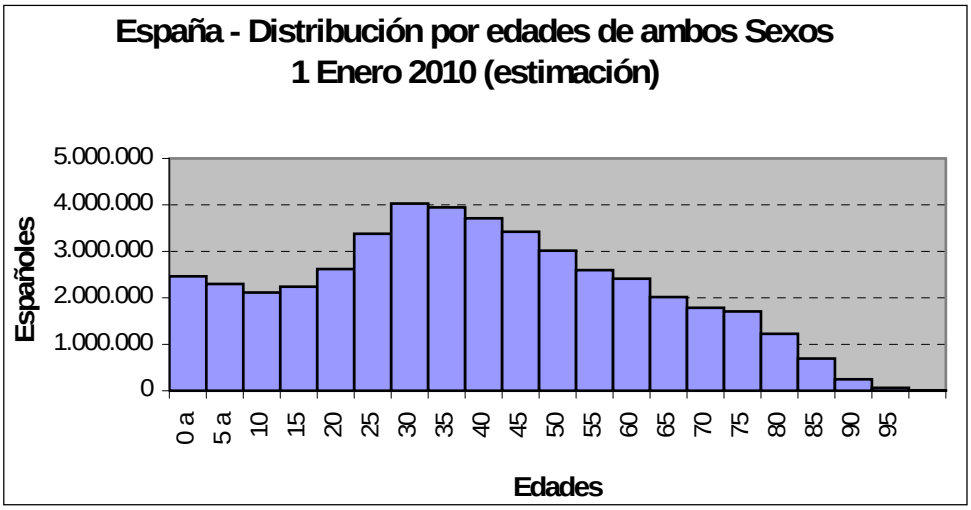
Variable objeto de estudio

Vamos a estudiar la distribución de población de ambos sexos por segmentos de edad, de cuatro en cuatro años estimada para el 1 de enero de 2010. Están sacados de la web del Instituto Nacional de Estadística español.
(http://www.ine.es/jaxiBD/tabla.do?per=01&type=db&divi=EPOB&idtab=3)

Tabla con los datos

Total Nacional	x	f	F	f·x	f·x²
0 a 4 años	2	2.461.227	2.461.227	4.922.454	9.844.908
5 a 9 años	7	2.297.895	4.759.122	16.085.265	112.596.855
10 a 14 años	12	2.113.106	6.872.228	25.357.272	304.287.264
15 a 19 años	17	2.240.229	9.112.457	38.083.893	647.426.181
20 a 24 años	22	2.619.905	11.732.362	57.637.910	1.268.034.020
25 a 29 años	27	3.379.985	15.112.347	91.259.595	2.464.009.065
30 a 34 años	32	4.030.195	19.142.542	128.966.240	4.126.919.680
35 a 39 años	37	3.945.775	23.088.317	145.993.675	5.401.765.975
40 a 44 años	42	3.711.892	26.800.209	155.899.464	6.547.777.488
45 a 49 años	47	3.423.294	30.223.503	160.894.818	7.562.056.446
50 a 54 años	52	3.013.424	33.236.927	156.698.048	8.148.298.496
55 a 59 años	57	2.595.058	35.831.985	147.918.306	8.431.343.442
60 a 64 años	62	2.411.466	38.243.451	149.510.892	9.269.675.304
65 a 69 años	67	2.015.458	40.258.909	135.035.686	9.047.390.962
70 a 74 años	72	1.786.020	42.044.929	128.593.440	9.258.727.680
75 a 79 años	77	1.706.879	43.751.808	131.429.683	10.120.085.591
80 a 84 años	82	1.226.794	44.978.602	100.597.108	8.248.962.856
85 a 89 años	87	691.867	45.670.469	60.192.429	5.236.741.323
90 a 94 años	92	248.325	45.918.794	22.845.900	2.101.822.800
95 a 99 años	97	63.071	45.981.865	6.117.887	593.435.039
100 y más años	102	7.152	45.989.017	729.504	74.409.408
		45.989.017		**1.864.769.469**	**98.975.610.783**

Representación gráfica



Medidas de centralización y dispersión

Aplicadas las fórmulas, hemos obtenido los siguientes resultados:
Vemos que la población está bastante envejecida, ya que el primer cuarto de la población ocupa aproximadamente 22 años, el segundo cuarto 15, el tercero 20 y el último 47 años. Se observa bien en el diagrama de cajas y bigotes.

moda:	32
media:	40,5481
Varianza:	508,0053
desviacion típica:	22,5390

Un porcentaje estimable de la población pertenece al intervalo: (18, 63)

Primer Cuartil:	22
Segundo Cuartil:	37
Tercer Cuartil:	57

Diagrama de Caja y Bigotes

② $\bar{x}=0 \Rightarrow 0=\frac{-3+3+a+b}{4} \Rightarrow a+b=0 \Rightarrow \boxed{a=-b}$

$\sigma=\sqrt{17} \Rightarrow \sqrt{17}=\sqrt{\frac{(-3)^2+3^2+a^2+b^2}{4}-0^2} \Rightarrow 17=\frac{9+9+b^2+b^2}{4}$; $50=2b^2$; $\boxed{b=5}$

$\boxed{a=-5}$

③ a) $\boxed{100 \text{ estudiantes}}$

b) $\boxed{\begin{matrix} a=54 \\ b=75 \end{matrix}}$

④ $630=\frac{\sum_{i=1}^{9} x_i}{9}$; $810=\frac{\sum_{i=10}^{12} x_i}{3}$

$\Downarrow$ $\Downarrow$

$5670=\sum_{i=1}^{9} x_i$ $\qquad 2430=\sum_{i=1}^{12} x_i$

$\sum_{i=1}^{12} x_i = 8100 \Rightarrow \bar{x}=\frac{8100}{12}=\boxed{675}$

⑤ $\boxed{18, 18, 19}, 19, \boxed{20, 22, 22}, 23, \boxed{27, 28, 28}, 31, \boxed{34, 34, 36}$

$\boxed{A=18,\ B=19,\ C=23,\ D=31,\ E=36}$

$Q_3-Q_1 = 31-19 = \boxed{12}$

⑥ a) mediana $= \boxed{24}$

$\boxed{154 \text{ taxis}}$

b) 40% de 200 $= 80 \rightarrow 22\$ \Rightarrow a = 22\$ \cdot \frac{1 \text{ km}}{0'55\$} = \boxed{40 \text{ km}}$

c) 90 km $\rightarrow$ 50 km $\cdot \frac{0'55\$}{1 \text{ km}} = 49'5\$ \rightarrow 186 \rightarrow 200 - 186 = 14$ taxis

$\frac{14}{200} \cdot 100 = \boxed{7\%}$

(7)

b) $\frac{100}{4} = 25 \Rightarrow Q_1 = 160$

$\frac{100}{4} \cdot 3 = 75 \Rightarrow Q_3 = 255$

Rango intercuartil $= 255 - 160 = \boxed{95}$

c)

P	f	F
$0 < P \leq 100$	12	12
$100 < P \leq 200$	46	58
$200 < P \leq 300$	29	87
$300 < P \leq 400$	7	94
$400 < P \leq 500$	6	100
	100	

$\boxed{a = 7}$ $\boxed{b = 6}$

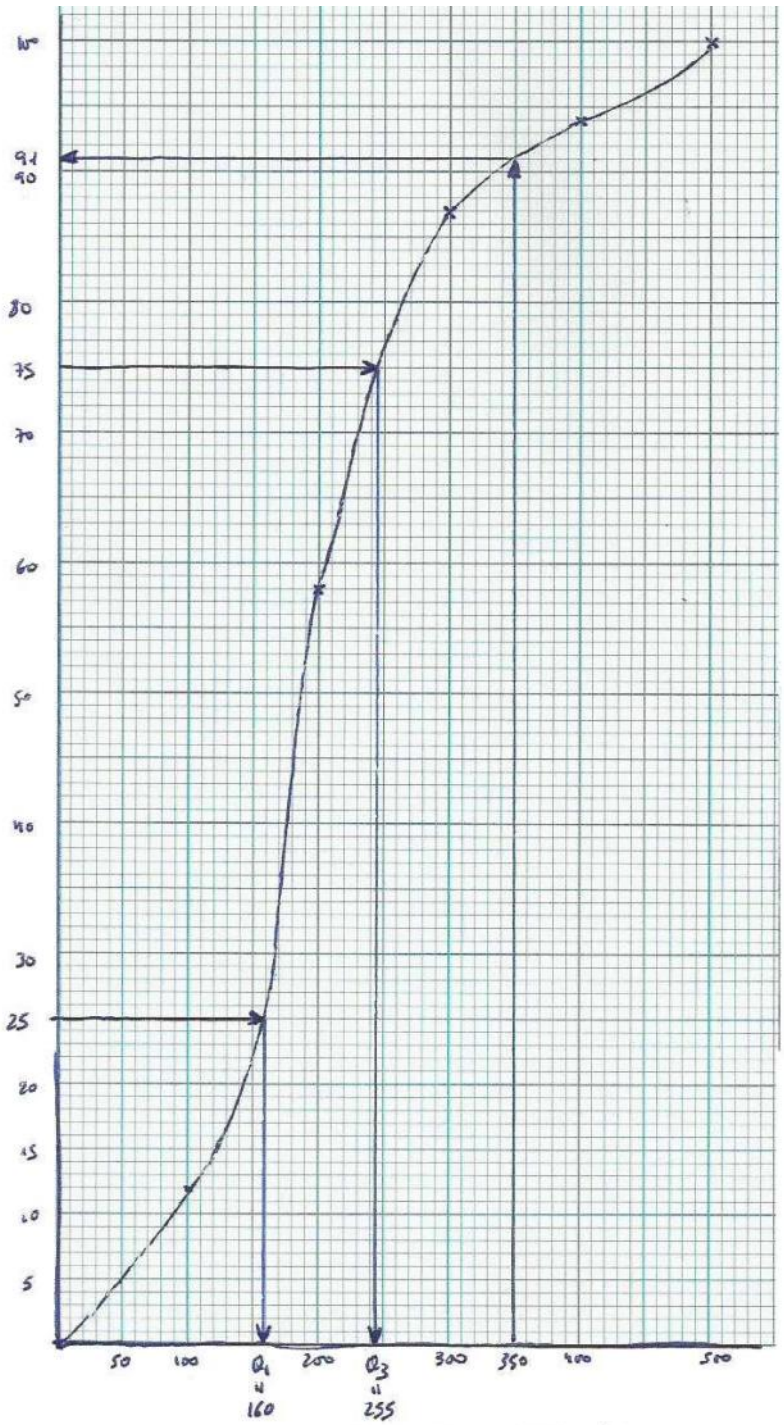
d)

P	f	f·P
50	12	600
150	46	6900
250	29	7250
350	7	2450
450	6	2700
	100	19900

$\overline{P} = \frac{19900}{100} = 199 \rightarrow \boxed{199.000\ \$}$

e) Para $x = 350 \rightarrow F = 91 \Rightarrow$

$100 - 91 = \boxed{9 \text{ casas 'de Luxe'}}$



(8)

X	f	F
1	5	5
2	10	15
3	P	15+P
4	6	21+P
5	2	23+P
	23+P	

Como la moda es 2, P es menor que 10

$P < 10 \Rightarrow 23 + P < 23 + 10$

Entonces el nº de datos estará entre 23 y 32

nº datos = 23 $\Rightarrow$ mediana ~~= 2~~

nº datos = 24 $\Rightarrow$ mediana = 2

nº datos = 30 $\Rightarrow$ mediana = ~~$\frac{2+3}{2}$~~ = 2'5

nº datos = 31 $\Rightarrow$ mediana = 3

nº datos = 32 $\Rightarrow$ mediana = 3

$\Rightarrow \boxed{P = 8 \text{ ó } P = 9}$

(9) A, B, C, D, E

mediana = 4 $\Rightarrow$ $\boxed{C=4}$

mode = 3 $\Rightarrow$ $\boxed{A=B=3}$

media = 5 $\Rightarrow$ $\frac{3+3+4+D+E}{5}=5$; $D+E=15$ $\Rightarrow$ $\boxed{\begin{matrix}D=5\\E=10\end{matrix}}$; $\boxed{\begin{matrix}D=6\\E=9\end{matrix}}$; $\boxed{\begin{matrix}D=7\\E=8\end{matrix}}$

(10) a) $\overline{w} = \frac{14\cdot7+23\cdot12+32\cdot13+41\cdot10+50\cdot8}{50} = 32$ ✓

b) $30 = \frac{14\cdot7+23\cdot12+32\cdot13+41\cdot10+50\cdot(8-x)}{50-x}$

$1500-30x = 1200+400-50x$

$20x = 100$ $\boxed{x=5}$

c) $\frac{14\cdot7+23\cdot12+32\cdot13+41\cdot(10+y)+50\cdot8}{50+y} \leq 33$

$1600+41y \leq 1650+33y$

$8y \leq 50$

$y \leq 6'25$ $\boxed{\text{Mayor nº de cajas} = 6}$

(11)

x	f
10	40
30	70
50	100
70	50
90	10
	270

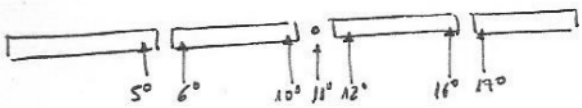
$\overline{x} = \frac{10\cdot40+30\cdot70+50\cdot100+70\cdot50+90\cdot10}{270} = \boxed{44'0741}$ años

(12)

x	f	F
4	2	2
5	5	7
6	4	11
7	3	14
8	4	18
10	2	20
12	1	21
	21	

a) Mediana: $\frac{21}{10}\ \frac{12}{}$ $\Rightarrow$ 11º $\Rightarrow$ $\boxed{\text{mediana} = 6}$

b) Cuartiles: $\frac{21}{5}\ \frac{4}{}$



$Q_1 = \frac{5^\circ+6^\circ}{2} = \frac{5+5}{2} = \boxed{5}$

$Q_3 = \frac{16^\circ+17^\circ}{2} = \frac{8+8}{2} = \boxed{8}$

c) $Q_3 - Q_1 = 8-5 = \boxed{3}$

①③

Variables objeto de estudio

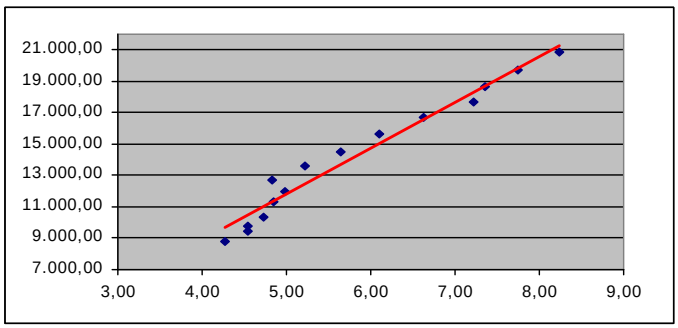
Vamos a estudiar la posible relación entre el crecimiento económico en España, concretamente el PIB medido en euros por habitante, y la natalidad medida con el número de nacimientos por 1.000 mujeres. Disponemos de datos de ambas variables desde el año 1991 hasta el 2005 y los hemos obtenido de la web del Instituto Nacional de Estadística español. (http://www.ine.es/daco/daco42/sociales06/sociales.htm)

Tabla con los datos

	nº nacimientos/1000 mujeres	PIB (€/hab)			
	x	y	x²	y²	x·y
1991	4,28	8.798,10	18,308	77.406.511,778	37.644,880
1992	4,54	9.444,55	20,571	89.199.582,147	42.836,376
1993	4,54	9.740,83	20,642	94.883.792,185	44.256,399
1994	4,74	10.332,14	22,422	106.753.090,237	48.924,406
1995	4,85	11.353,85	23,570	128.909.811,061	55.121,414
1996	4,99	12.002,66	24,905	144.063.894,161	59.899,300
1997	4,83	12.730,62	23,356	162.068.697,274	61.524,286
1998	5,23	13.581,69	27,337	184.462.365,617	71.012,176
1999	5,64	14.524,98	31,831	210.974.949,376	81.948,044
2000	6,11	15.653,20	37,327	245.022.685,310	95.634,092
2001	6,62	16.715,47	43,858	279.406.834,138	110.699,296
2002	7,22	17.650,33	52,077	311.534.090,172	127.372,078
2003	7,36	18.629,66	54,156	347.064.275,451	137.097,637
2004	7,74	19.678,42	59,908	387.240.023,131	152.310,933
2005	8,24	20.863,89	67,898	435.301.845,973	171.918,442
	86,93	**211.700,38**	**528,165**	**3.204.292.448,010**	**1.298.199,759**

Representación gráfica

La nube de puntos y la recta de regresión ***y sobre x*** están representadas en el siguiente diagrama:



Medidas de centralización, dispersión y regresión

Aplicadas las fórmulas, hemos obtenido los resultados de la tabla adjunta

media_x :	5,795
des_tip_x :	1,275
media_y :	14.113,359
des_tip_y :	3.799,027
covar :	4.755,921
coef_corr :	0,982

El coeficiente de correlación lineal tiene un valor muy cercano a 1, por lo que se puede calificar de **positiva y muy fuerte** la relación existente entre el PIB y la natalidad. Es decir, que al ir aumentando la riqueza en España también va aumentando la natalidad.

Rectas de regresión

Si escribimos la ecuación de la recta de regresión *y sobre x* en la forma: $y = m \cdot x + n$ los coeficientes que hemos hallado son:

Recta de regresión y sobre x	
m:	2925,1674
n:	10383,4973

Si escribimos la ecuación de la recta de regresión *x sobre y* en la forma: $x = m \cdot y + n$ los coeficientes que hemos hallado son:

Recta de regresión x sobre y	
m:	0,0003
n:	1,1446

Estimación de resultados

Vamos a estimar el nº de nacidos por 1.000 mujeres para un PIB de 25.000€ por habitante.
Sustituimos $y = 25.000$ en la recta de regresión *x sobre y* obteniendo $x = 9,38$
Es decir, que para dicho PIB habría 9,38 nacidos por 1.000 mujeres en un año.

(14) Hay una 1ª nube con correlación ligeramente positiva ⇒ $\boxed{0'7}$

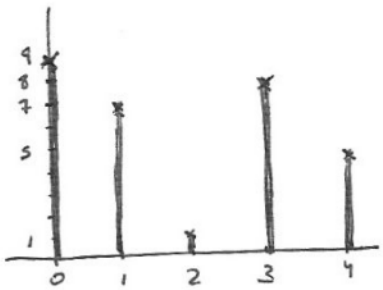
La 2ª nube tiene correlación ligeramente negativa ⇒ $\boxed{-0'8}$

La 3ª nube es informe ⇒ $\boxed{0'1}$

La 4ª nube es una relación lineal decreciente ⇒ $\boxed{-1}$

(15) a)

x	f	f_r
0	9	$9/30 = 30\%$
1	7	$7/30 = 23\%$
2	1	$1/30 = 3\%$
3	8	$8/30 = 27\%$
4	5	$5/30 = 17\%$
	30	100%



b)

y	f	f·y	$f \cdot y^2$
15	10	150	2250
16	10	160	2560
17	10	170	2890
	30	480	7700

$\bar{y} = \frac{480}{30} = \boxed{16}$ $\quad \sigma_y^2 = \frac{7700}{30} - 16^2 = \boxed{0'6}$

(16) a) $\bar{y} = \frac{8 \cdot 2 + 6 \cdot 8 + 5 \cdot 6 + 4 \cdot 3 + 2 \cdot 1}{20} = 5'4$

$\sigma_y = \sqrt{\frac{8^2 \cdot 2 + 6^2 \cdot 8 + 5^2 \cdot 6 + 4^2 \cdot 3 + 2^2 \cdot 1}{20} - 5'4^2} = 1'3191$

Un porcentaje estimable de los resultados pertenece al intervalo $(\bar{y} - \sigma_y, \bar{y} + \sigma_y)$:

$(5'4 - 1'3191, 5'4 + 1'3191) = \boxed{(4'0809, 6'7191)}$

b) Son 20 datos, al ser un nº par, buscamos el 10º y el 11º y hacemos su media aritmética:

$\text{Mediana} = \frac{6 + 5}{2} = \boxed{5'5}$

c) $\bar{x} = 43$ $\qquad \bar{y} = 5'4$

$\sigma_x = 20'2731$ $\qquad \sigma_y = 1'3191$

$\sigma_{xy} = -26'2$

$\boxed{r = -0'9797}$

El tiempo de ordenador y la calificación están fuertemente relacionados de manera inversa.

① $\left.\begin{array}{l} c-a=10 \\ \boxed{b=11} \\ \frac{a+b+c}{3}=9 \end{array}\right\} \left.\begin{array}{l} c-a=10 \\ a+c=16 \end{array}\right\}$

$2c=26 \;;\; \boxed{c=13} \Rightarrow \boxed{a=3}$

② a)

x	$0 \le x < 20$	$20 \le x < 40$	$40 \le x < 60$	$60 \le x < 80$	$80 \le x < 100$
f	22	52	64	42	20
F	22	74	138	180	200

b) 40% de 200 = 80 → **42 puntos**

③ a) $\overline{x} = \frac{2230}{45} = \boxed{49'5}$

b) $\sum_{i=1}^{47} x_i = \sum_{i=1}^{45} x_i + 37 + 30 = 2230 + 37 + 30 = 2297 \longrightarrow \overline{x} = \frac{2297}{47} = \boxed{48'8723}$

④ a)

b)

⑤ (c)(i) $\frac{80}{2} = 40 \longrightarrow$ mediana $\simeq \boxed{91'7\ g} \longrightarrow \boxed{92\ g}$

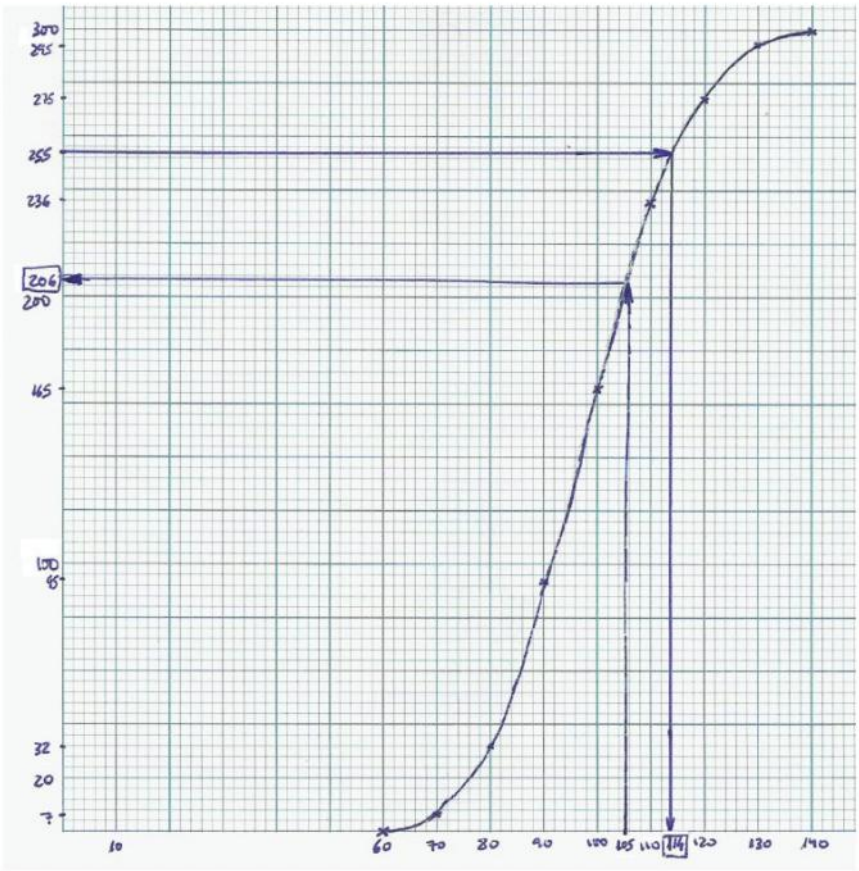
(ii) $\frac{80 \cdot 3}{4} = 60 \longrightarrow Q_3 = \boxed{101'25\ g} \longrightarrow \boxed{101\ g}$

(d) $(w_1 - \overline{w}) + (w_2 - \overline{w}) + \cdots + (w_{80} - \overline{w}) = (w_1 + w_2 + \cdots + w_{80}) - (\overline{w} + \overline{w} + \cdots + \overline{w}) =$

$= 80 \cdot \overline{w} - 80 \cdot \overline{w} = \boxed{0}$

(6)

v	marca clase	f	$v \cdot f$
$v \leq 60$	–	0	
$60 < v \leq 70$	65	7	455
$70 < v \leq 80$	75	25	1875
$80 < v \leq 90$	85	63	5355
$90 < v \leq 100$	95	70	6650
$100 < v \leq 110$	105	71	7455
$110 < v \leq 120$	115	39	4485
$120 < v \leq 130$	125	20	2500
$130 < v \leq 140$	135	5	675
$v > 140$	–	0	
		300	29450



a)

$$\bar{x} = \frac{29450}{300} = \boxed{98'16}$$

b)

f	F	
7	7	
25	32	
63	95	
70	165	← a
71	236	
39	275	← b
20	295	
5	300	

$\boxed{a = 165}$

$\boxed{b = 275}$

(i) $v = 105 \longrightarrow F = 206$

$v > 105 \longrightarrow N = 300 - 206 = 94$ $\qquad \frac{94}{300} \cdot 100 = \boxed{31'3\,\%}$

(ii) 15% de $300 = 45$ $\qquad 300 - 45 = 255 \longrightarrow \boxed{v = 114 \text{ km/h}}$

(7) a) Son 12 datos, la mediana es la media del 6° y el 7° datos:

$$\frac{9+r}{2} = 9'5 \Rightarrow \boxed{r = 10}$$

El tercer cuartil será la media del 9° y 10° datos:

$$\frac{s+13}{2} = 13 \Rightarrow \boxed{s = 13}$$

b)
$$\frac{5+6+7+7+9+9+10+10+13+13+13+t}{12} = 10 \Rightarrow \boxed{t = 18}$$

(8) $a < b < c = d$

moda $= 11 \Rightarrow \boxed{c = d = 11}$

rango $= 8 \Rightarrow d - a = 8 \Rightarrow \boxed{a = 3}$

media $= 8 \Rightarrow \frac{a+b+c+d}{4} = 8 \Rightarrow \frac{b+25}{4} = 8 \Rightarrow \boxed{b = 7}$

(9) $p = 32 + 33 = \boxed{65}$

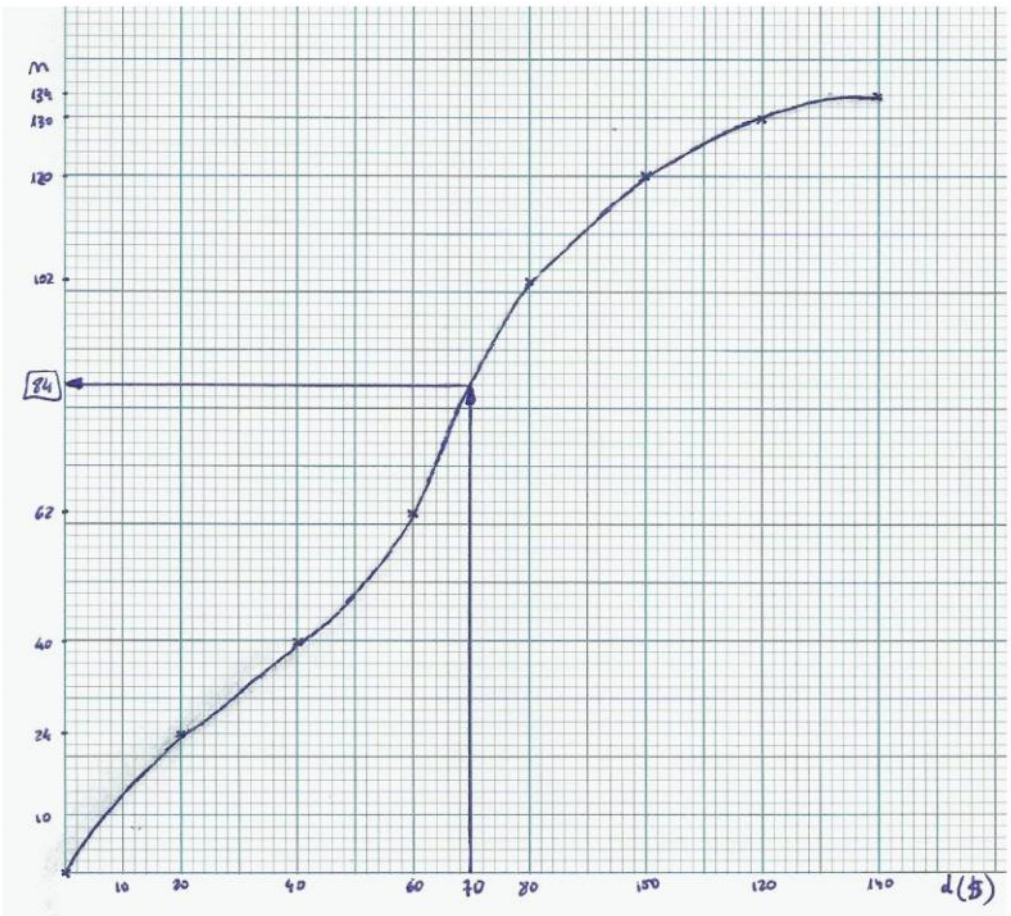
$65 + q = 99 \Rightarrow \boxed{q = 34}$

$\frac{125}{4} = 31'25 \Rightarrow 32^{\circ} \longrightarrow \boxed{\text{mediana} = 16}$

$\overline{x} = \frac{15 \cdot 11 + 16 \cdot 21 + 17 \cdot 33 + 18 \cdot 34 + 19 \cdot 18 + 20 \cdot 8}{125} = \frac{2176}{125} = \boxed{17'408}$

(10) $\overline{x} = \frac{10 \cdot 24 + 30 \cdot 16 + 50 \cdot 22 + 70 \cdot 40 + 90 \cdot 18 + 110 \cdot 10 + 130 \cdot 4}{134} = \frac{7860}{134} = \boxed{58'6567}$

d	<20	<40	<60	<80	<100	<110	<140
m	24	40	62	102	120	130	134



c) i) $\overline{t} \simeq 2 \cdot \overline{d}^{2/3} + 3 = \boxed{33'1932}$

ii) $t = 37 \Rightarrow 37 = 2 \cdot d^{2/3} + 3 \Rightarrow d = \left[\frac{37-3}{2}\right]^{3/2} = 70'1$

$d \geq 70'1 \Rightarrow m = 134 - 84 = \boxed{50}$

(11) a) $\frac{80}{2} = 40 \longrightarrow$ **mediana = 32**

b) $\frac{80}{4} = 20 \longrightarrow Q_1 = 22$

$\frac{80 \cdot 3}{4} = 60 \longrightarrow Q_3 = 40$ $\Rightarrow Q_3 - Q_1 =$ **22**

c)

	F	f
$0 \le t < 10$	5	5
$10 \le t < 20$	16	11
$20 \le t < 30$	36	20
$30 \le t < 40$	60	24
$40 \le t < 50$	74	14
$50 \le t < 60$	80	6

(12) a)

y	f	f_r
1'5	8	8/30
4'5	15	15/30
7'5	7	7/30
	30	

b)

x	f	f·x
18	10	180
44	17	748
60	3	180
	30	1108

$\bar{x} = \frac{1108}{30} = 36'93$ años

(13) a) $\bar{x} = \frac{189'3536}{15} =$ **12'6236**

$\bar{y} = \frac{344'0143}{15} =$ **22'9343**

$\sigma_x = \sqrt{\frac{2669'3347}{15} - 12'6236^2} =$ **4'2138**

$\sigma_y = \sqrt{\frac{8498'1153}{15} - 22'9343^2} =$ **6'3686**

$\sigma_{xy} = \frac{4742,0056}{15} - 12'6236 \cdot 22'9343 =$ **26'6203**

b) $r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = \frac{26'6203}{4'2138 \cdot 6'3686} =$ **0'9920**

c) $y - 22'9343 = \frac{26'6203}{4'2138^2} \cdot (x - 12'6236)$

$y = 1'4992x + 4'0088$

$x = 12\% \Rightarrow y = 21'9992 \simeq$ **22%**

Es muy fiable la estimación, porque r = 0'9920 es muy cercano a 1

(14) a) $\bar{x} = 35'3$ $\quad \bar{y} = 4233'3$

$\sigma_x = 8'6923$ $\quad \sigma_y = 708'6764$

$\sigma_{xy} = 6122'2$

$r = 0'9939$

$y = 1370'2941 + 81'0294x$

b) Si vendemos x sillas, ingresamos $120x$ €

los costes estimados de fabricación serían:

$y = 1370'2941 + 81'0294x$

Beneficios $\Rightarrow 120x \ge 1370'2941 + 81'0294x$

$x \ge 35'1623$

Debe vender, al menos, **36 sillas**.