

## Proyecto de investigación de la segunda evaluación.

El proyecto de investigación, consiste en hacer un **estudio estadístico UNIDIMENSIONAL y otro BIDIMENSIONAL**. Al final tenéis un ejemplo con los dos estudios que pido.

Los datos deben de ser reales, y debéis especificar de dónde los habéis sacado, es decir debe estar la dirección de la página web de donde los sacasteis en el proyecto, para que yo pueda comprobarlos. Por ejemplo, yo los saqué de la página del instituto nacional de estadística.:

<http://www.ine.es/jaxiBD/tabla.do?per=01&type=db&divi=EPOB&idtab=3>

y copiáis el enlace para que yo lo vea.

Los dos estudios deben contener:

- Variables objeto del estudio y dirección de la página web de los datos
- Tabla con los datos
- Medidas de centralización
- Etc:
- Exactamente los mismos apartados que los dos ejemplos que encontrareis más adelante.

La página con los datos por supuesto, no debe ser la misma bajo ningún concepto para dos o más alumnos, de tal manera que antes de empezar el estudio, me diréis de donde sacaréis los datos, para que nadie más pueda sacar los mismos. Si hay alguien que tiene los mismos datos, consideraré que han copiado uno del otro y los dos tendrán 0 en el proyecto.

Los datos pueden provenir del periódico o de otra fuente, no es necesario que sean de una página web.

Mejor si los datos provienen de algún tema en que tengas interés: deporte, cine, música, etc.

Cualquier pregunta se me puede hacer en clase y los 5-10 primeros minutos de la clase los dedicaremos a los proyectos si es necesario.

Las gráficas no es necesario hacerlas al ordenador, se pueden hacer a mano.

**Ejemplos de los proyectos: ➔**

## EJEMPLO DE ESTUDIO ESTADÍSTICO DE UNA VARIABLE UNIDIMENSIONAL

### Variable objeto de estudio

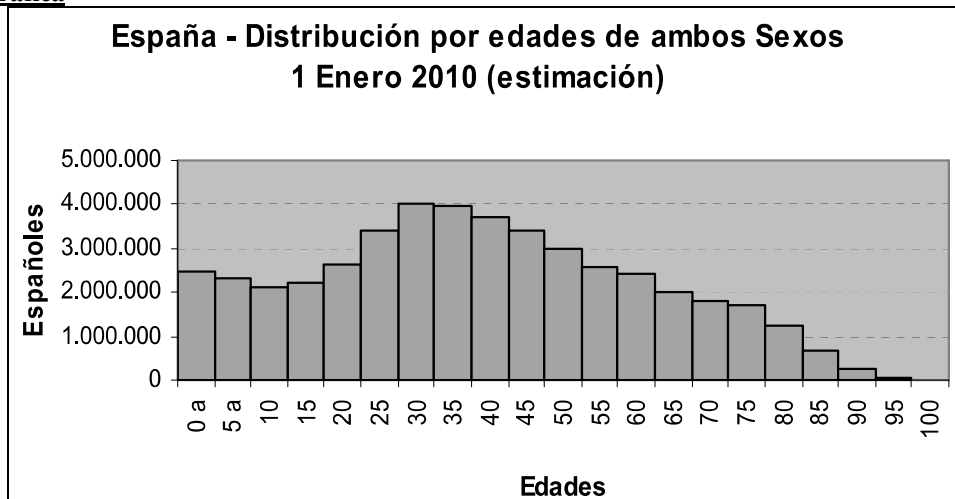
Vamos a estudiar la distribución de población de ambos sexos por segmentos de edad, de cuatro en cuatro años estimada para el 1 de enero de 2010. Están sacados de la web del Instituto Nacional de Estadística español.

(<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·x <sup>2</sup>
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
		<b>45.989.017</b>		<b>1.864.769.469</b>	<b>98.975.610.783</b>

### Representación gráfica



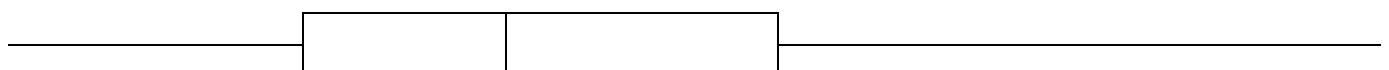
### Medidas de centralización y dispersión

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.

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

### Diagrama de Cajas y Bigotes



<b>moda:</b>	32
<b>media:</b>	40,5481
<b>Varianza:</b>	508,0053
<b>desviación típica:</b>	22,5390

<b>Primer Cuartil:</b>	22
<b>Segundo Cuartil:</b>	37
<b>Tercer Cuartil:</b>	57

## EJEMPLO DE ESTUDIO ESTADÍSTICO DE UNA VARIABLE BIDIMENSIONAL

### Variables objeto de estudio

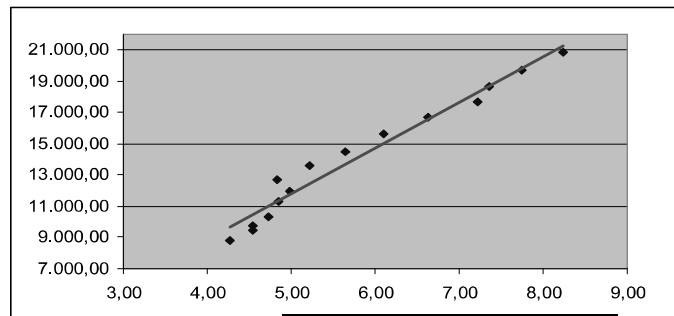
Vamos a estudiar la posible relación entre el crecimiento económico en España, concretamente el PIB medido en euros por habitante, y la natalidad medida con el número de nacimientos por 1.000 mujeres. Disponemos de datos de ambas variables desde el año 1991 hasta el 2005 y los hemos obtenido de la web 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	x <sup>2</sup>	y <sup>2</sup>	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
	<b>86,93</b>	<b>211.700,38</b>	<b>528,165</b>	<b>3.204.292.448,010</b>	<b>1.298.199,759</b>

### 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

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.

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

### 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:

<b>Recta de regresión y sobre x</b>
<b>m:</b> 2925,1674
<b>n:</b> 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:

<b>Recta de regresión x sobre y</b>
<b>m:</b> 0,0003
<b>n:</b> 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.