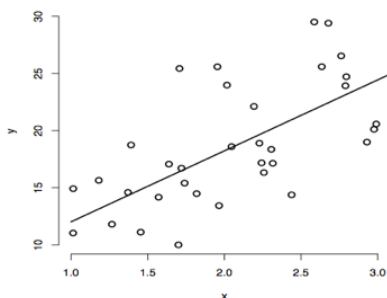


Asignatura: ESTADÍSTICA. 1º RELACIONES LABORALES. CURSO 2019-2020

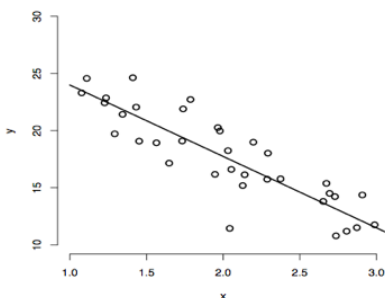
**BOLETÍN: TEMA 3**

1. Cada uno de los gráficos corresponde a una de las características que se describen a continuación. Relaciónalos y razona la respuesta.

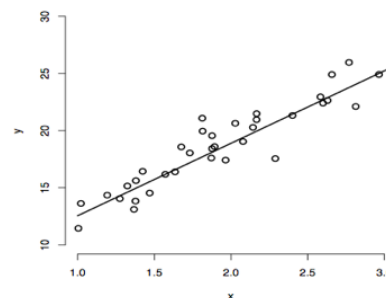
**Gráfico 1**



**Gráfico 2**



**Gráfico 3**



- a) La recta de regresión de  $Y$  sobre  $X$  es  $Y = 6,38X + 6,42$ , con  $R^2 = 0,87$   
 b)  $s_{xy}/s_x^2 = 6,21$  e  $R^2 = 0,44$   
 c)  $\bar{X} = 2,02$ ,  $\bar{Y} = 17,72$ ,  $\frac{1}{n} \sum_{i=1}^n X_i Y_i = 33,76$
2. Se desea estudiar si el precio del agua tiene un impacto en su consumo. De un estudio del instituto Nacional de Estadística se obtuvo la siguiente información resumida sobre 8 comunidades autónomas españolas para las variables  $Y$  = “consumo de agua”(en litros por habitante y día) y  $X$  = “precio del m3 de agua”(en céntimos de euro):

$$\sum_{i=1}^8 X_i = 680; \quad \sum_{i=1}^8 Y_i = 1275; \quad \sum_{i=1}^8 X_i^2 = 72796; \quad \sum_{i=1}^8 Y_i^2 = 209221; \quad \sum_{i=1}^8 X_i Y_i = 99828$$

En base a los datos anteriores:

- a) Estima, mediante el modelo de regresión lineal simple, los coeficientes de la recta de regresión que explica el consumo de agua en función de su precio.  
 b) Interpreta con detalle los coeficientes obtenidos en el apartado anterior. Existe relación entre el consumo de agua y el precio? ¿De qué tipo?  
 c) Calcula el coeficiente de determinación e interprétalo.  
 d) ¿Cuál será el consumo estimado de agua en una comunidad donde se establece un precio de 50 céntimo euro por cada m3?
3. Se desea estudiar la posible relación entre los gastos en alimentación ( $Y$ ) de una familia y su Ingresos globales ( $X$ ), en miles de euros, a través de una recta del tipo  $y = \beta_0 + \beta_1 x$ . Para esto se tomaron datos de  $n = 62$  familias y se obtuvieron los siguientes resultados:

$$\sum_{i=1}^{62} X_i = 91,11, \sum_{i=1}^{62} X_i^2 = 195,71, \sum_{i=1}^{62} Y_i = 24,47, \sum_{i=1}^{62} Y_i^2 = 12,25, s_{xy} = \frac{1}{62} \sum_{i=1}^{62} (X_i - \bar{X})(Y_i - \bar{Y}) = 0,19$$

- Calcula la recta de regresión de la variable gasto ( $Y$ ) frente a la variable ingreso ( $X$ ).
- Calcula el coeficiente de correlación e interprétalo.
- Calcula el coeficiente de determinación e interprétalo.
- ¿Qué gastos se espera para una familia que ingresa 1800 euros?

4. Se midieron simultáneamente la temperatura del aire ( $X$ ) y la temperatura del agua de mar ( $Y$ ), ambos en  $^{\circ}\text{C}$ , en una determinada zona costera en 12 épocas diferentes del año, obteniéndose las siguientes observaciones:

$X_i$	14	15	18	24	26	30	31	30	29	24	20	15
$Y_i$	11	11	13	19	23	26	28	27	25	22	16	10

- Calcula la recta de regresión de  $Y$  sobre  $X$  e interpreta el resultado.
  - Calcula el coeficiente de determinación. ¿Cuál es la temperatura esperada del agua si la temperatura del aire es de  $25^{\circ}\text{C}$ ?
5. El peso de la arcilla utilizada en ladrillos tiene implicaciones para el diseño estructural, acústico y térmico de los edificios. Los datos se obtuvieron de un estudio sobre el mortero utilizado en la fabricación de ladrillos. Las variables observadas son  $X$  = “contenido de aire”(en%) e  $Y$  = “densidad seca”(en g/cm<sup>3</sup>). Los datos son los siguientes:

$X$	5.7	6.8	9.6	10.7	12.6	14.4	15.0	15.3	16.2	17.8	18.7	19.7	20.6	25.0
$Y$	2.10	2.14	2.09	2.19	1.98	2.01	1.98	2.03	1.97	1.89	1.92	1.90	1.88	1.85

- Construye un diagrama de dispersión de estos datos para verificar la posible relación entre las variables. ¿Tu dirías que hay una fuerte relación lineal entre  $X$  e  $Y$ ? ¿Por qué? ¿Cuál es la característica principal de esto posible adicción? Calcule el coeficiente de correlación de la muestra entre  $X$  e  $Y$ .
  - Calcula la línea de regresión de  $Y$  sobre  $X$  y evalúe su bondad de ajuste a través del coeficiente de determinación.
  - Haz predicciones de los valores de la variable  $Y$  para los valores 5, 10, 15, 20 y 25 de la variable  $X$ .
6. Una hoja de datos de Excel contiene información sobre las variables  $X$  = “consumo de gasóleo”(en litros) e  $Y$  = “emisión de CO<sub>2</sub>”(en g) de 50 casas. A modo ilustrativo, en la tabla aparecen algunos datos.
- Si se pretende estudiar la posible relación de dependencia que puede existir entre las variables  $X$  e  $Y$ , ¿qué tipo de gráfico usarías? ¿Cómo se construye?
  - Utilizando algún programa estadístico, se han calculado las siguientes medidas resumen de las variables  $X$  e  $Y$ :

- número de observaciones: 50

- media muestral de  $X$ : 6.05

- media muestral de  $Y$ : 153.6

- varianza muestral de  $X$ : 0.7837

- varianza muestral  $Y$ : 480.7600

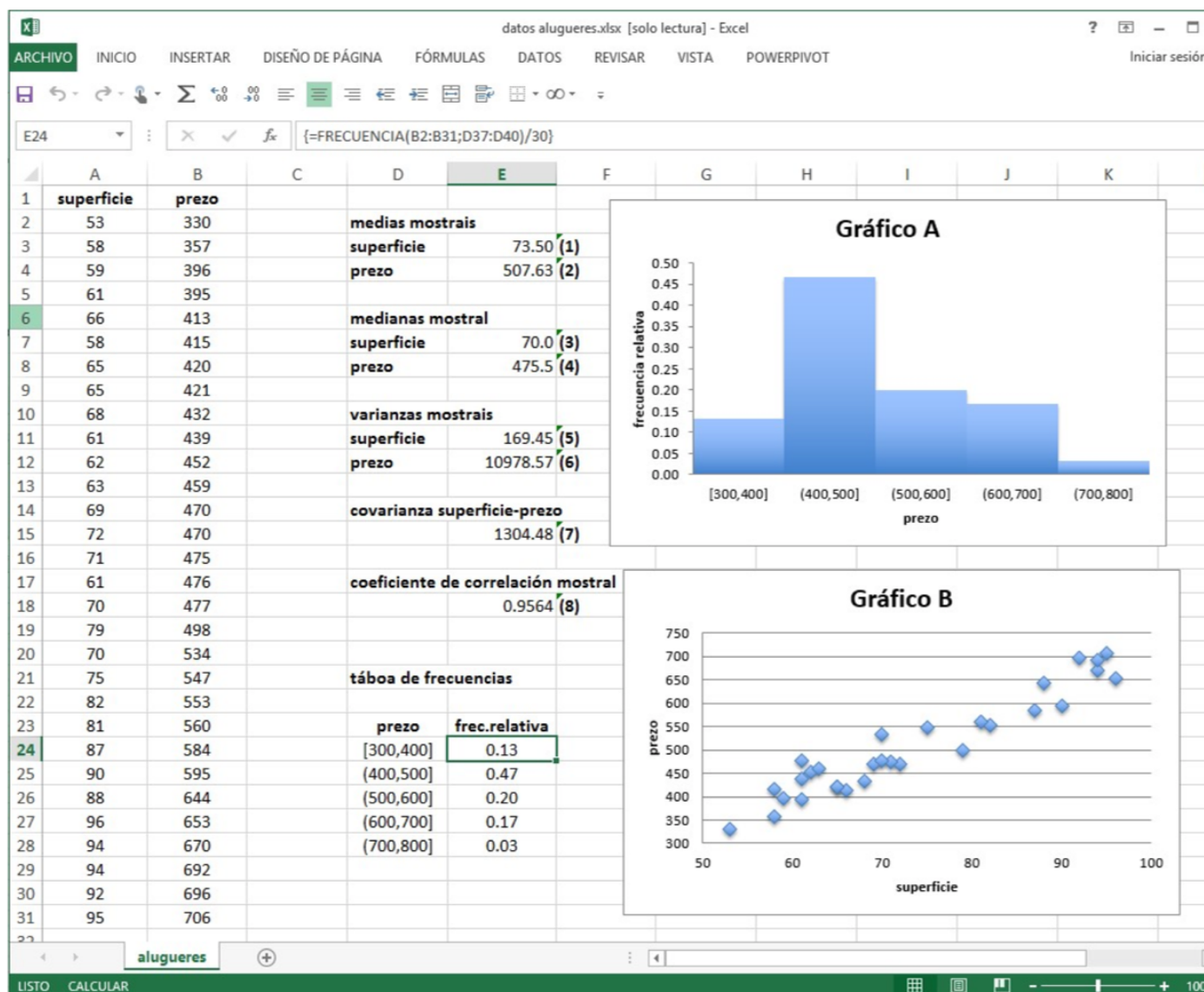
- covarianza muestral entre  $X$  e  $Y$ : 18.7040

	A	B
1	$X$	$Y$
2	5.3	141
3	4.9	128
4	6.8	179
5	5.9	152
6	7.2	179
...	...	...
47	5.0	142
48	5.5	134
49	5.1	127
50	6.1	149
51	7.2	190

Teniendo en cuenta esta información, ¿cuánto vale el coeficiente de correlación muestral entre  $X$  e  $Y$ ? ¿Indica esto que existe una fuerte relación entre estas dos variables?

- Calcula la recta de regresión que explica el comportamiento de la variable  $Y$  en función de la variable  $X$ . ¿Qué cantidad de CO<sub>2</sub> se preveé para una casa con un consumo de 6 litros?

7. Una hoja de datos de Excel contiene información sobre 30 pisos de Vigo. Las variables consideradas son: superficie = “superficie de piso”(en m<sup>2</sup>) y precio = “precio de alquiler mensual”(en euros). El siguiente figura muestra una hoja de Excel con datos y algunas medidas descriptivas y gráficas.



- Si se deseara estudiar la posible asimetría de la distribución de la variable “precio”, qué gráfico sería apropiado, gráfico A o gráfico B? ¿Cómo se llama este gráfico?
- ¿Qué porcentaje de apartamentos tiene un alquiler de 600 euros o menos?
- ¿Cuál es la línea de regresión que explica el comportamiento de la variable “precio” de acuerdo con la variable “Superficie”? ¿Cómo se interpreta la pendiente de la línea de regresión en este caso?
- Calcula el coeficiente de determinación. Cómo se interpreta su valor?
- Según el modelo de regresión anterior, ¿cuál es el precio promedio de alquiler de un piso de 65 m<sup>2</sup>?

8. En varios modelos de automóviles, se observan los datos correspondientes a la potencia del motor en CV (variable  $X$ ) y a la aceleración medida como el número de segundos necesarios para acelerar de 0 a 100 km/h (variable  $Y$ ). La siguiente tabla enumera los valores obtenidos:

X	50	75	90	100	120	100
Y	15	12	10.5	10	9	8

- Calcula las siguientes medidas de resumen de datos: media muestral de  $X$ , media muestral de  $Y$ , varianza y desviación típica muestral de  $X$ , varianza y desviación típica muestral de  $Y$ , covarianza muestral entre  $X$  e  $Y$  y coeficiente de correlación muestral entre  $X$  e  $Y$ . Interprete el valor del coeficiente de correlación.
- Haz el gráfico de dispersión de  $X$  e  $Y$ .
- Estima, mediante el modelo de regresión lineal simple, los coeficientes de la recta de regresión que explica la aceleración en función de la potencia del motor. Interpreta los coeficientes obtenidos en detalle.
- Calcula el coeficiente de determinación como medida de bondad de ajuste del modelo de regresión lineal obtenido en la sección anterior. Interpretalo.
- ¿Cuál será la aceleración estimada para un automóvil de 130 hp?