

1. Hemos recibido el encargo de desarrollar una aplicación para móvil que convierta archivos en formato .jpg a .pdf. Uno de los requisitos solicitados por el cliente es que, una vez elegido el archivo .jpg a convertir, se muestre en pantalla el tiempo estimado que se tardará en convertir el archivo. Para implementar tal funcionalidad, decidimos usar un modelo de regresión para predecir el tiempo en función del tamaño del archivo .jpg, tomando como base los datos recolectados en la fase de pruebas de la aplicación, reflejados en la siguiente tabla:

Tamaño archivo (MB)	20	10	50	17	5
Tiempo de conversión (segs.)	1.8	1	5.3	1.9	0.6

- a) Construye las dos rectas de regresión correspondientes a los datos en la tabla anterior.
- b) ¿Podemos concluir que la regresión es adecuada para resolver el problema planteado, en base a los datos de la tabla?
- c) Usando la recta de regresión adecuada, calcula el tiempo de conversión estimado que mostraremos en pantalla, para un archivo de 30MB.
2. Con objeto de determinar la relación entre el tiempo de respuesta (en segundos) de una determinada base de datos de consulta y el número de usuarios se han tomado 10 datos correspondientes a 2 semanas consecutivas, obteniéndose los siguientes resultados (tiempo | número de usuarios):

	Lun.	Mar.	Mié.	Jue.	Vie.
Semana 1	4.32 15	7.14 18	9.21 20	9.71 20	15.39 26
Semana 2	5.2 16	8.37 19	9.34 20	10.46 21	18.9 29

- a) ¿Cuál de las dos variables es más homogénea?
- b) Usando la recta de regresión adecuada, determina el número de usuarios activos, si el tiempo de respuesta es de 12 segundos.
- c) ¿Podemos concluir que la regresión es adecuada para resolver el problema planteado, en base a los datos de la tabla?
3. La siguiente tabla muestra los resultados de medir el tiempo que se tarda en transferir 5 archivos de distinto tamaño a través de una red:

Tamaño del archivo (KBytes)	300	500	120	600	400
Tiempo de transmisión (segundos)	1.1	1.9	0.3	2.1	1.5

- a) Usando regresión lineal, determina el tiempo de transmisión de un archivo de 200 KBytes.
- b) Usando regresión lineal, determina el tamaño de un archivo que tarda 0.5 segundos en ser transferido.
- c) Determina si el ajuste es fiable o no.
4. Se han obtenido los siguientes datos sobre número de usuarios en un aula de acceso libre y número de fallos en los PCs del aula:

Número de usuarios	47	41	54	50	42
Número de fallos	5	4	6	5	3

- a) Obtén la recta de regresión necesaria para predecir el número de usuarios en función del número de fallos.
- b) Determina cómo de fiable es el ajuste.
5. La siguiente tabla muestra los resultados de medir, en 5 instantes durante un día, el número de usuarios conectados a un servidor Linux y la cantidad de RAM (en GB) disponible en el sistema:

Número de usuarios	5	3	2	6	1
RAM disponible	1.1	2.6	3	1	3.5

- a) Usando regresión lineal, determina la cantidad de RAM disponible si hay 4 usuarios conectados.
- b) Usando regresión lineal, determina el número de usuarios conectados si la RAM disponible es de 1.5 GB.
- c) Determina cómo de fiable es el ajuste.
6. La siguiente tabla muestra los resultados de medir la cantidad de datos (X , en KB) tecleada por cinco operadores en un día de trabajo y el tamaño de su monitor (Y , en pulgadas).

X	150	175	210	230	276
Y	15	17	19	21	26

- a) Usando regresión lineal, determina la cantidad de datos que se puede predecir para un operador cuyo monitor sea de 24 pulgadas.
- b) Determina cómo de fiable es el ajuste.
7. Los siguientes datos representan el número de accesos (X) a un servidor remoto y el número de cortes en la comunicación (Y) sufridos durante diez días consecutivos:

X	170	160	210	140	180	240	160	140	210	230
Y	8	7	11	5	9	12	8	6	10	13

- a) Comprueba si las variables X e Y son independientes.
- b) Calcula, usando una recta de regresión, el número de cortes en la comunicación durante un día en el que se producen 150 accesos. ¿Es bueno el ajuste?
8. Los siguientes datos representan el número total de mensajes de correo electrónico (X) manejados por un servidor y el número de mensajes tipo SPAM (Y) correspondientes a diez días consecutivos:

X	170	160	210	140	180	240	160	140	210	230
Y	8	7	11	5	9	12	8	6	10	15

- a) ¿Cuál de las dos variables es más dispersa?
- b) Calcula, usando una recta de regresión, el número total de mensajes manejados durante un día en el que se recibieron 13 mensajes tipo SPAM. ¿Es bueno el ajuste realizado mediante la recta de regresión?
9. En una empresa los empleados se clasifican en tres categorías: técnicos, especialistas y administrativos. El salario medio mensual y la varianza de los salarios de cada categoría en el mes de Diciembre de 2015 son los que aparecen en el siguiente cuadro:

Categoría	Salario medio mensual (euros)	Varianza de los salarios (miles de <i>euros</i> ²)
Técnicos	2500	10
Especialistas	2000	25
Administrativos	1500	15

I.- ¿En qué grupo de empleados los salarios son más homogéneos?

II.- En la discusión para fijar los salarios de 2016 han sido propuestas dos alternativas:

A: La elevación de todos los salarios en un 5 %.

B: La elevación de todos los salarios en 50 euros mensuales.

Calcula los salarios medios que resultan de aplicar las dos alternativas y la dispersión relativa en cada caso. ¿Qué alternativa provoca menos dispersión en los salarios de los grupos?

10. Se han estudiado las calificaciones de 100 alumnos en dos asignaturas: Matemáticas y Estadística, obteniéndose los siguientes datos:

$$\bar{x} = 110, \bar{y} = 2,5, S_x = 10, S_y = 0,5, r = 0,85.$$

- a) ¿Qué nota se puede predecir para un alumno, que ha obtenido 125 puntos en Matemáticas, en la asignatura de Estadística?
- b) ¿Se puede decir que aquellos alumnos que obtienen mayor calificación en Matemáticas sean los mismos que obtienen mayor calificación en Estadística?
- c) ¿Cuál es la ecuación de la recta de regresión de X sobre Y ?