

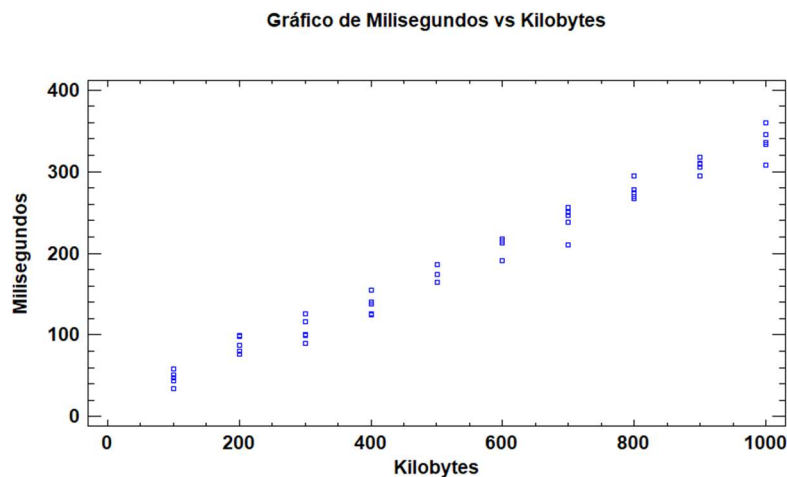
La recta de regresión

Práctica 2

Rafael Rus Rus Grupo 82

Ejercicio 1

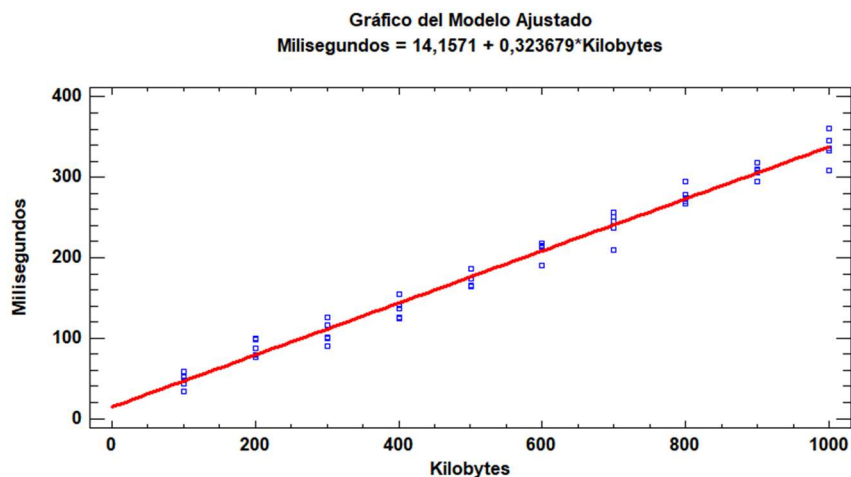
- a) Análisis gráfico de la relación entre el tamaño del fichero que se lee (variable explicativa X expresada como Kilobytes) y el tiempo empleado en su lectura (variable respuesta Y expresada como Milisegundos).



En la muestra se puede observar el gráfico de dispersión donde se aprecia que la relación entre ambas variables es bastante lineal y muy fuerte. Parece entonces sensato utilizar una recta de regresión para predecir Y en función de X. Destacar también que la nube de puntos alrededor de la recta de regresión que se traza es estrecha (a pesar de que los puntos están muy dispersos con respecto al eje de la variable X).

- b) Recta de regresión:

Cálculo de la recta de regresión que sirve para predecir el tiempo empleado en leer un fichero en función de sus tamaños en kilobytes: $y_i = a + bx_i$, donde y_i son milisegundos y x_i son kilobytes.



$$b = \text{Pendiente} = \frac{\text{Cov}(x,y)}{S_x^2} = 0.32$$

$$a = \text{Intercepto} = \bar{y} - b\bar{x} = 14.16$$

No es necesario realizar transformaciones sobre la recta de regresión pues el gráfico es lineal en todo el estudio.

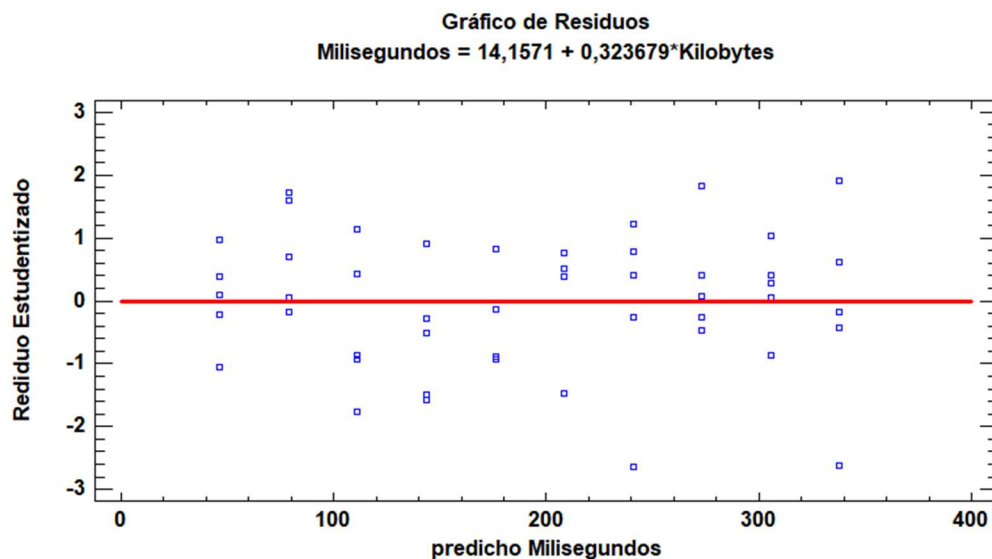
- c) Se tardará aproximadamente 127 milisegundos en leer un archivo de 350 kilobytes.

$$\text{Milisegundos} = 14.16 + 0.32 * 350 \text{ Kilobytes}$$

- d) Calculo que en 200 milisegundos podremos leer 574 kilobytes.

$$574 \text{ Milisegundos} = 14.16 + 0.32 * \text{Kilobytes}$$

- e) Evaluación de la regresión:



En primer lugar, podemos observar que el estadístico R-Cuadrada del programa indica que el modelo ajustado explica el 98.3012 % de la variabilidad en Milisegundos (por lo que la relación es bastante lineal). Además, en la gráfica, los puntos están distribuidos al azar en torno a la recta.

Otro dato que destacar es que el coeficiente de correlación es igual a 0.99147, lo cual indica una relación relativamente fuerte entre las variables. En cuanto al error estándar del estimado, indica que la desviación estándar de los residuos es 12.4736. Por su parte, el estadístico de Durbin-Watson (DW) examina los residuos para determinar si hay alguna correlación significativa basada en el orden en el que se presentan en el archivo los datos. Puesto que el valor-P es mayor que 0.05, no hay indicación de una autocorrelación serial en los residuos con un nivel de confianza del 95.0%.