|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Coefficients: | | | | |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 7.721e+01 | 3.187e-01 | 242.25 | <2e-16 \*\*\* |
| datosarchivo$Afluencia | 3.003e-04 | 2.103e-05 | 14.28 | <2e-16 \*\*\* |

---

|  |
| --- |
| Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |

|  |  |
| --- | --- |
| Residual standard error: | 46.1 on 48424 degrees of freedom |
| Multiple R-squared: | 0.004193, Adjusted R-squared: 0.004173 |
| F-statistic: | 203.9 on 1 and 48424 DF, p-value: < 2.2e-16 |

Interpretación:

La ecuación que describe la relación entre las variables:

x=Estación del metro

y= Número de pasajeros que ingresa en esa estación (afluencia)

es:

y=0.0003003041x + 77.21026

El modelo encontrado muestra una relación positiva entre ambas variables, es decir, que la estación realmente influye en el número de pasajeros que ingresan al metro. Aun más allá de ello, a partir del análisis del coeficiente p, generado entre los “Coeficientes”, podemos ver que este toma un valor 0.00000000000000022. Dicho valor es menor a la hipótesis nula del test de normalidad según la cual el valor de p<0.5 para que exista relación entre las variables. En dicho caso, este valor es mucho menor al 0.5, por lo cual podemos decir que en función de la estación del metro que se esté analizando esta define en gran medida la cantidad de pasajeros que ingresa al metro.

**Gráficas de la línea de regresión lineal**

Escala de tiempo

Descripción generada automáticamente

Nuestra línea de regresión lineal es plana, por lo que no logra representar la recta que queremos.

Imagen que contiene Gráfico

Descripción generada automáticamente

Podemos observar que nuestra línea de regresión lineal aumenta conforme x y y lo hacen.

Pruebas de normalidad

Para analizar los datos de la afluencia, se optó por utilizar las pruebas de Kolmogorov para verificar si los datos son normales o no. Para ello, se utilizó la función shapiro.test() la cual nos arrojó los siguientes resultados:

Lilliefors (Kolmogorov-Smirnov) normality test

data: datosarchivo$Afluencia

D = 0.14475, p-value < 2.2e-16

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Interpretación: Tal como se había comprobado al obtener la ecuación de relación entre variables, el p-value es mucho menor a 0.5, lo cual demuestra que los datos de la afluencia de personas en el metro siguen una distribución normal.