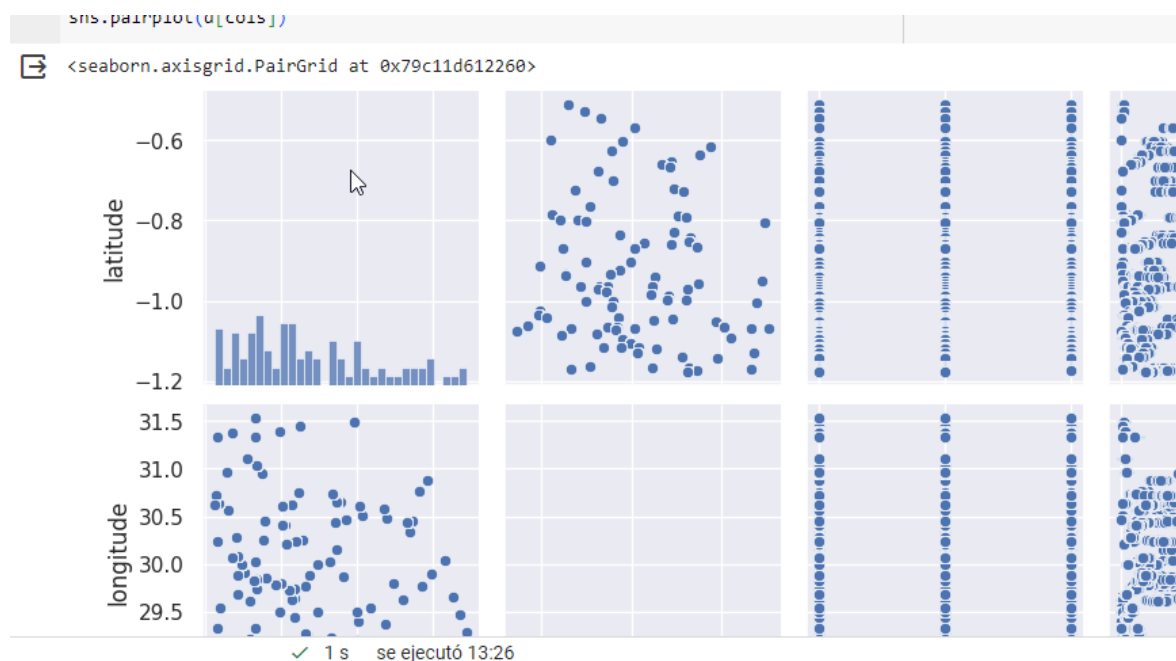


Preprocesamos el dataset, nos disponemos a hacer la limpieza de datos para poder realizar los primeros modelos.

Inicialmente verifica que el dataset no tiene variables categóricas, por lo tanto lo deberemos hacer convirtiendo algunas variables continuas.

Realizamos una exploración de las columnas del dataset para encontrar las posibles variables que tengan correlación con la variable objetivo, la cual es la variable emisión



Igualmente se verifica que hay un gran numero de datos faltantes pero no es aleatorio sino que hay dos grupos de columnas que tienen unas un valor de 2107 datos faltantes y otro grupo con solo un dato faltante.

Igualmente se visualiza algunas variables discretas cuyo valor parece no estar relacionado en absoluto con la variable objetivo.