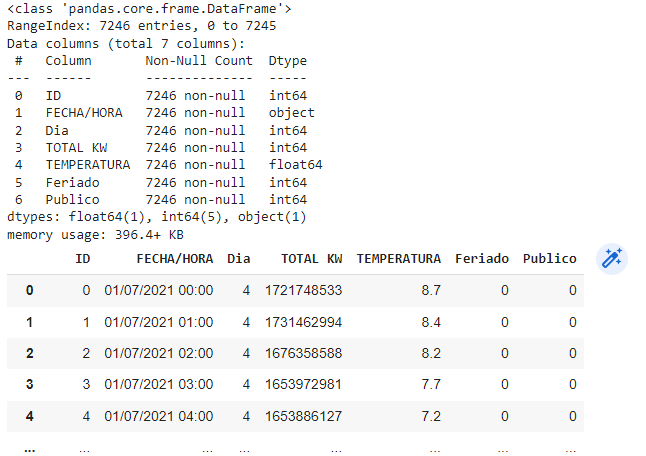
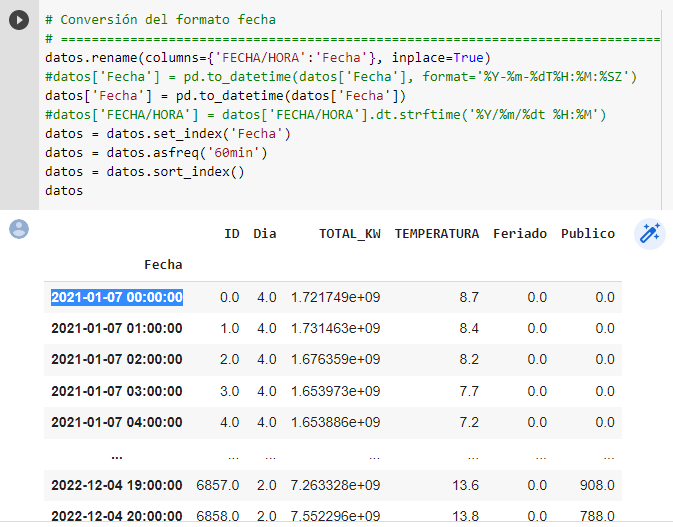
El estudio de las series de tiempo nos demandó el esfuerzo de entender como son tratados los datos referidos a esta variable

En el transcurso del desarrollo del trabajo nos encontramos que si dejamos que Python reclasifique los datos se producía un desfasaje de fechas, lo cual provocaba que el entrenamiento del modelo resultase en errores, ya que su interpretación de las fecha provocaba espacios en blanco del modelo, con lo cual los entrenamientos fallaban

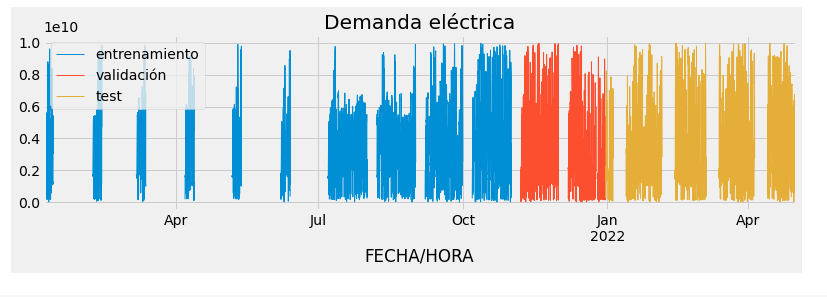
Datos en crudo:



Intentamos convertir las fechas sin tratamiento y nos encontramos con:



Lo que nos daba de resultado:



Estos espacios en blanco, provocaron un error del modelado de predicciones en la serie de tiempo



Esto se resuelve con un pequeño cambio en el modelo de datos, y un tratamiento correcto de las fechas en el set de datos

ID;Fecha;Dia;Demanda;Temperatura;Feriado;Publico;

0;2021/07/01 00:00;4;1721748533;8.70;0;0;

1;2021/07/01 01:00;4;1731462994;8.40;0;0;

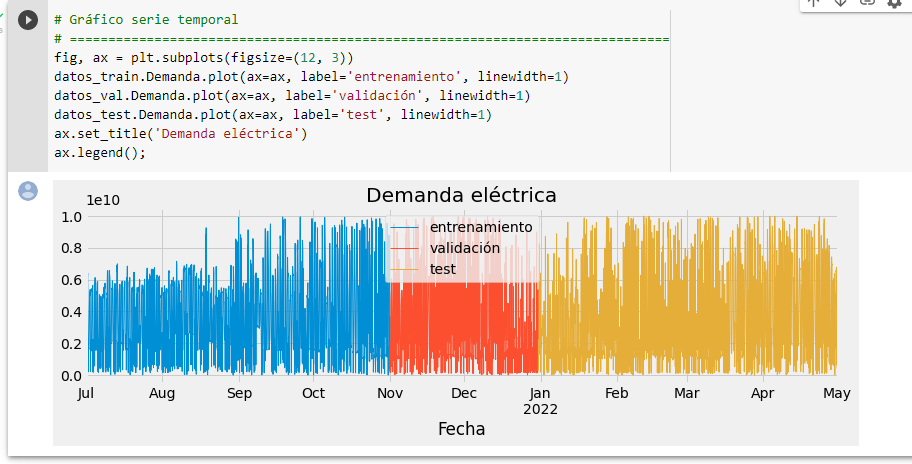
2;2021/07/01 02:00;4;1676358588;8.20;0;0;

3;2021/07/01 03:00;4;1653972981;7.70;0;0;

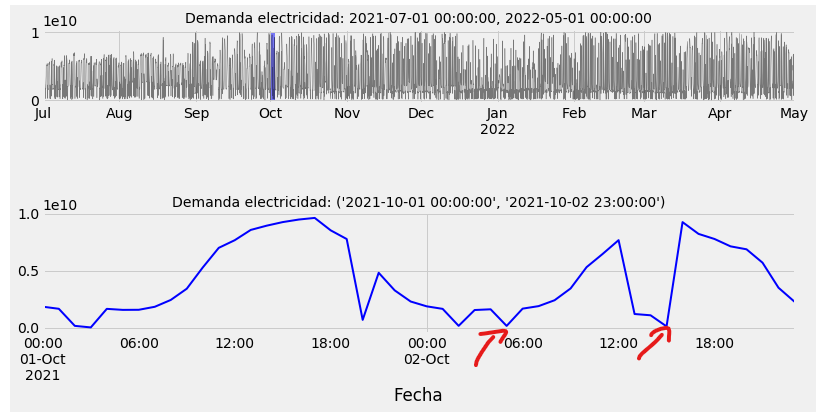
4;2021/07/01 04:00;4;1653886127;7.20;0;0;

5;2021/07/01 05:00;4;1697885847;6.70;0;0;

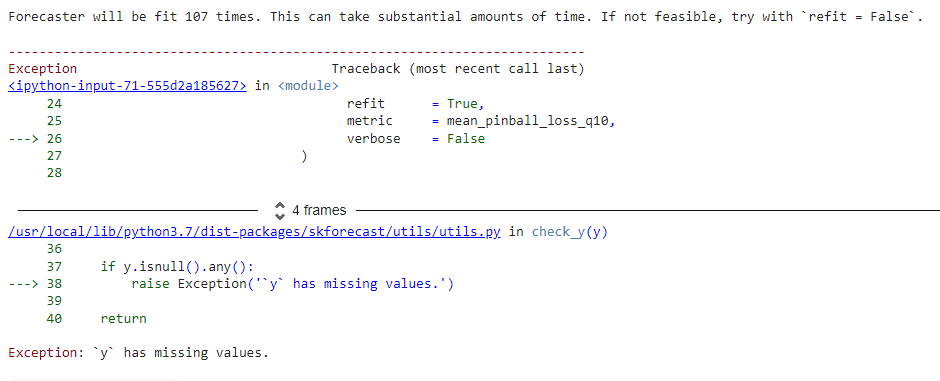
Lo que nos limpió los datos:



Al momento de iniciar los modelos predictivos, nos encontramos con problemas relacionados a la continuidad de los datos, ya que se aprecian en el set de datos, algunos valores que caen abruptamente.



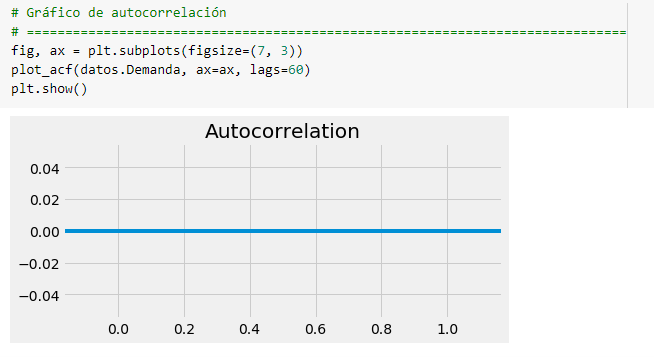
Estos valles nos producen problemas en los modelos de Predicción (backtest) al darnos valores faltantes.



Investigando el tema, se llaga a la conclusión de que la escases de datos, produce este tipo de errores, ya que la característica de los modelados forescat es que llena los espacios con null, sobretodos cuando llega a los extremos de la serie de datos.

Como conclusión, podemos decir que si bien el modelo teórico funciona, se debe tener un gran trabajo en la preparación de los datos y en el volumen del mismo, ya que gran parte del éxito de la predicción radica en obtener la mayor cantidad y calidad de datos y su preparación.

En nuestro caso particular, no llegamos a un resultado positivo, por diversos problemas, pero el mayor de todos se dio en que se disponía SOLO de 10 meses de datos y eso es un volumen muy escaso para realizar predicciones del tipo foward – back forecasting, mas con nuestra distribución de datos que en ese volumen no alcanzó a generar correlaciones apreciables, lo que fue un primer indicio de que el volumen de datos no alcanzaba a impactar en el modelo de predicción



Con los datos actuales y dado que el periodo de entrenamiento es muy pequeño se ha logrado un factor de cobertura aproximadamente 60%

