Pasos: después de integrar las dos bases de datos tanto la de calidad de aguas como la estación meteorológica empezamos con el análisis.

Primer paso: subir la base de datos y graficar una serie de tiempo con respecto a la variable Tiempo con cada una de las variables.

Paso 2: Para hacer un análisis de datos faltante e imputación de estos se opta por hacerlo de forma univariada. A continuación revisaremos variable a variable con respecto al tiempo.

Salinidad:

Primera vista:

Gráfico

Descripción generada automáticamente

En una reunión con el experto se concluyo esta variable carece de sentido que tenga valores menores al valor 25. Se opto por colocar los datos que estén por dabajo de este valor en datos faltantes para poder imputarlos. Después de realizar este proceso la grafica se mira de esta manera:

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Podemos ver que hay unos datos muy dispersos, entonces se procede a graficar un boxplot para tener una idea si hay valores atípicos.

Tabla

Descripción generada automáticamente con confianza baja

Según el grafico podemos decir que si hay valores atípicos, pero para corroborar esto vamos a hacer algunas pruebas.

Hecho el box plot, vamos a ver por medio de lof valores atípicos – luego por mas de 5 consecutivos