

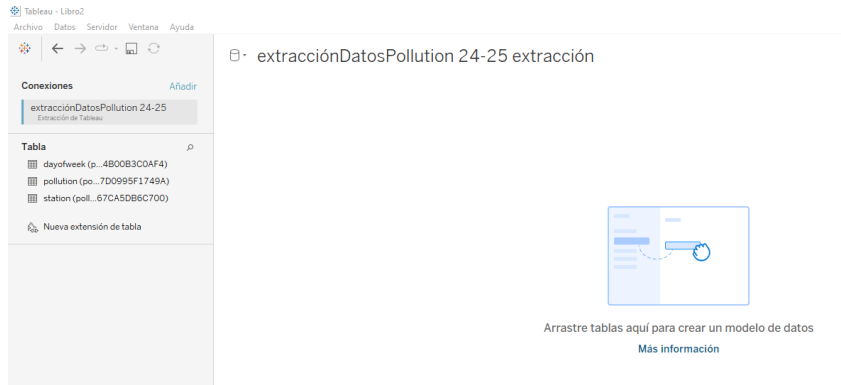
Sección #8 – Conexión y empezando con Tableau

Creación y conexión de Tableau

Crea un proyecto nuevo de Tableau y haz las siguientes ejercicios:

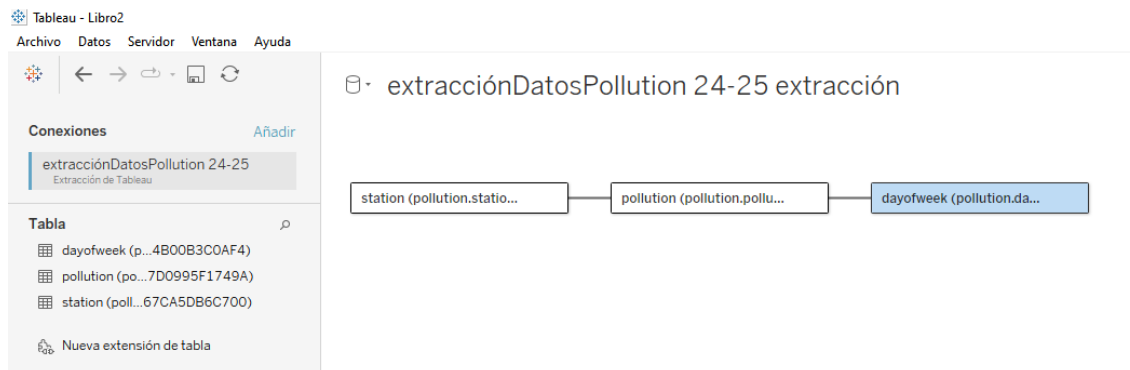
5. Carga la extracción de datos que encontrarás en Sallenet.

Se empieza cargando la extracción de datos de la base de datos *Pollution*.



6. Carga todas las tablas que encontrarás en la bbdd “Pollution”. Asegúrate que la carga es correcta así como las relaciones entre tablas.

Continuamos cargando todas las tablas del esquema de la BD relacionando correctamente las tablas para completar el modelo relacional.



7. Escoge un campo de contaminantes de la siguiente lista: NO, NO2, O3, SO2.

NO2.

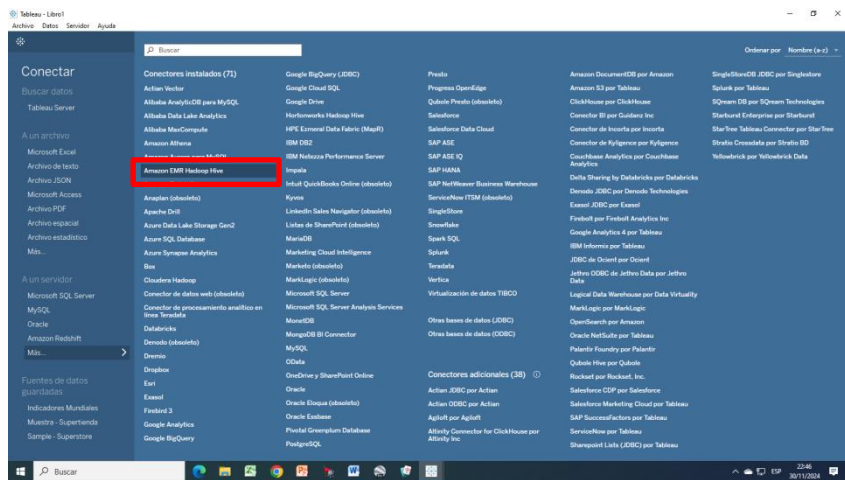
8. Indica aquí el elemento elegido, busca información por internet y descríbelo en 2 líneas.

El contaminante escogido en mi caso es el NO2. El dióxido de nitrógeno es un contaminante atmosférico, de origen principalmente antropogénico, cuyas fuentes fundamentales son el tráfico rodado, así como las emisiones de determinadas industrias y grandes instalaciones de combustión. Niveles altos de este pueden provocar numerosas patologías respiratorias.

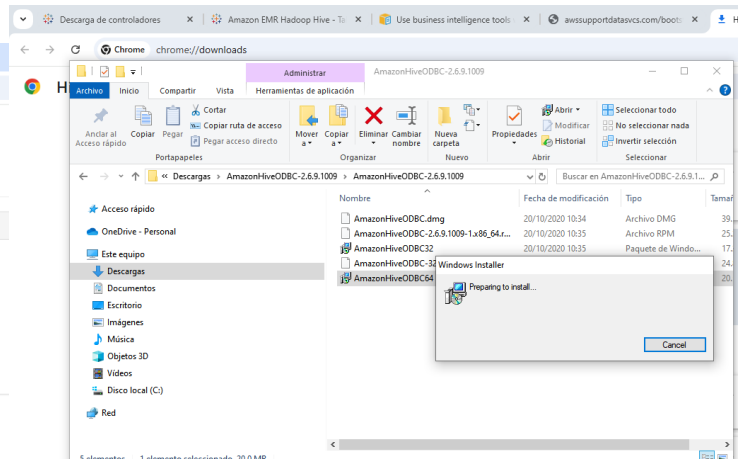
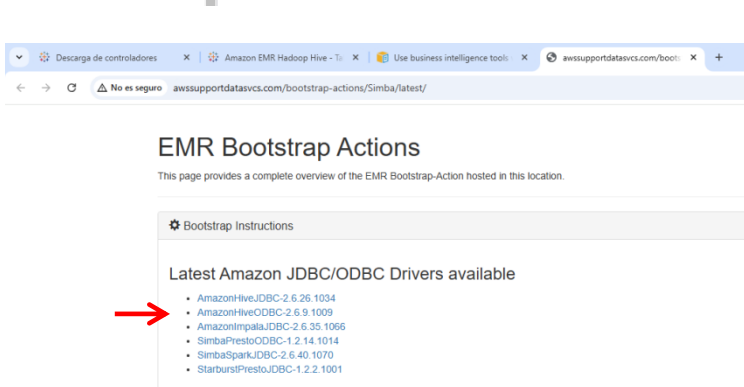
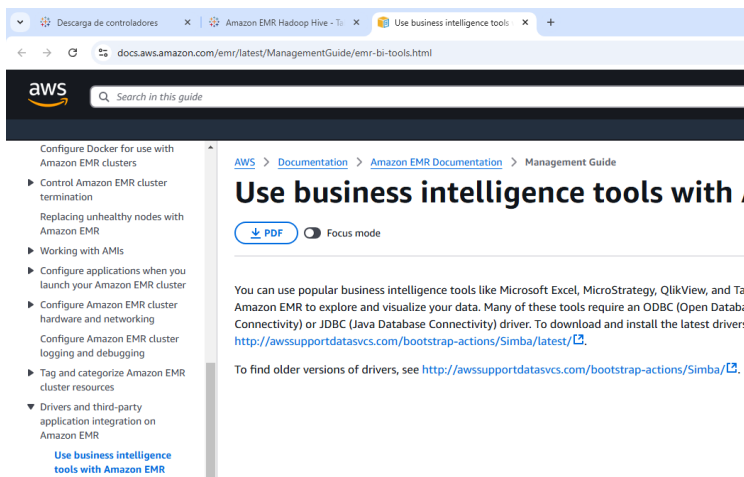
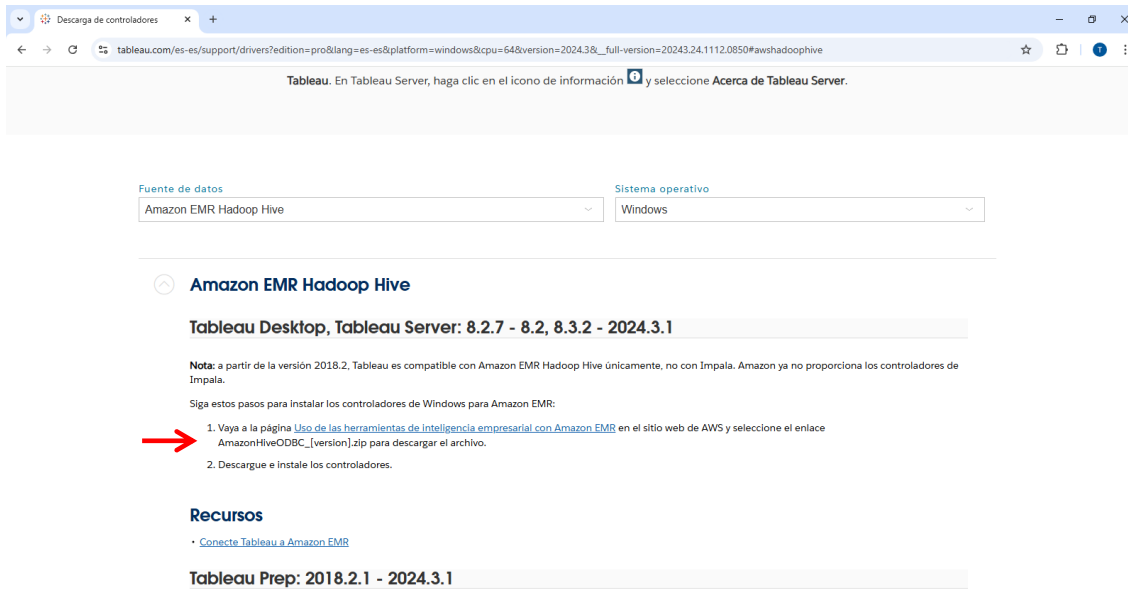
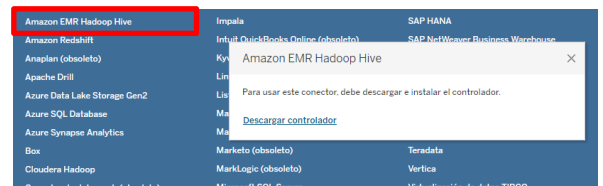
9. Hablad con el grupo para que cada miembro seleccione un elemento distinto. Es importante que no se repita porque os servirá para la entrega grupal.

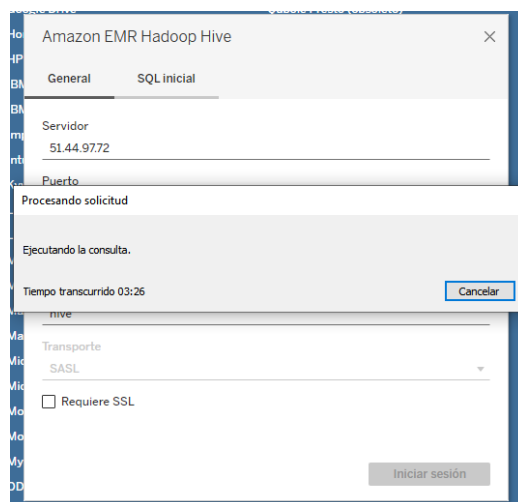
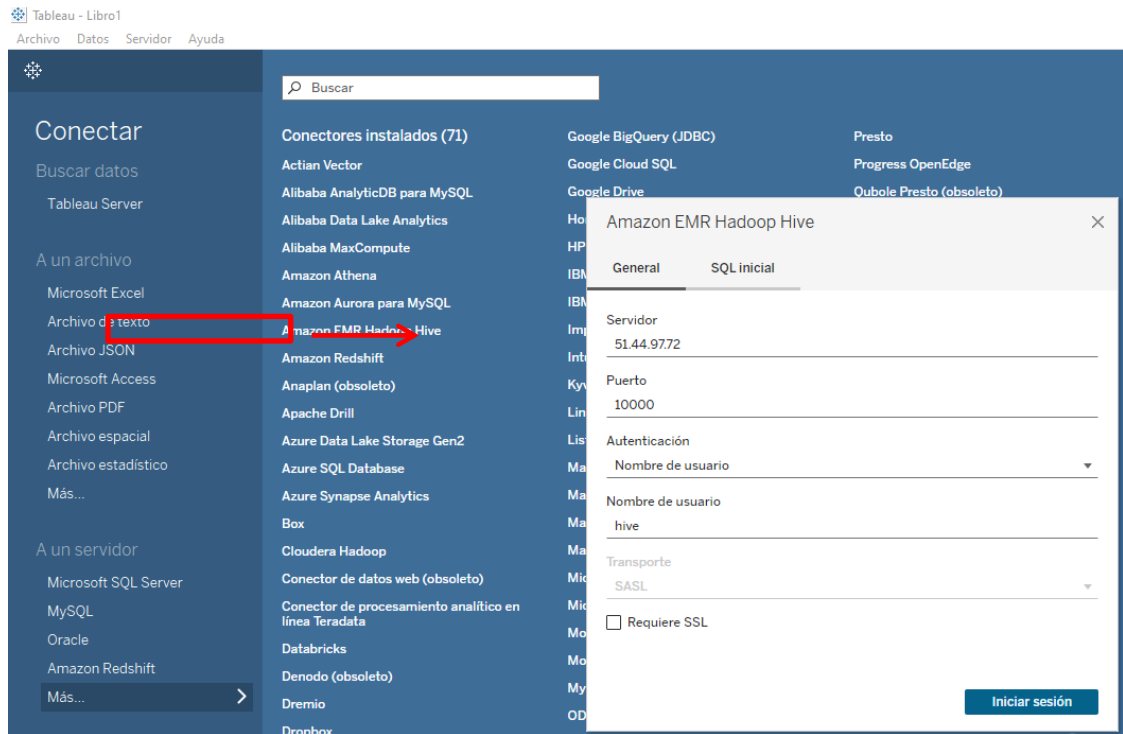
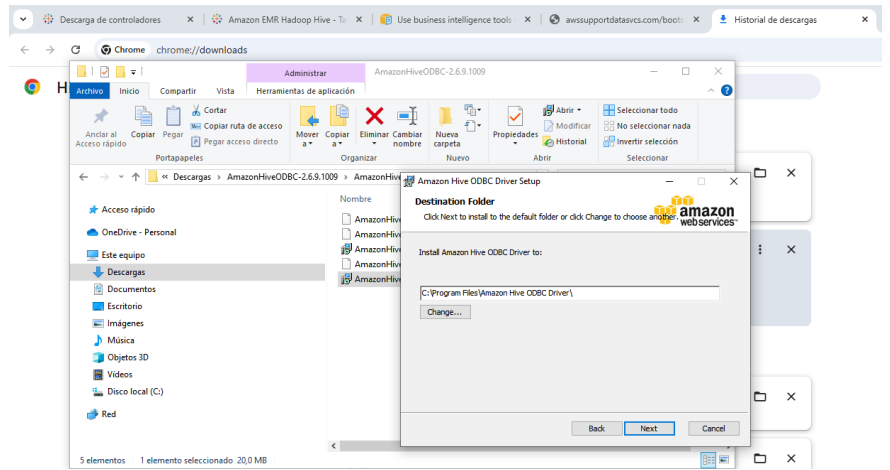
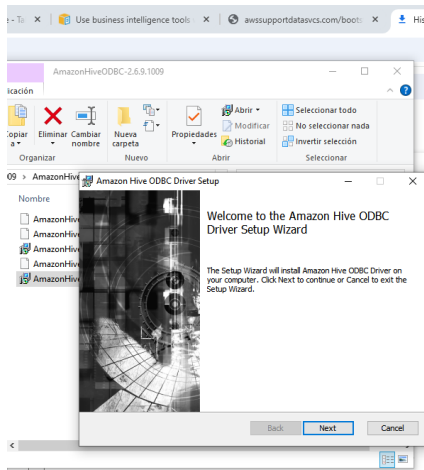
Se han repartido los contaminantes entre los miembros del grupo.

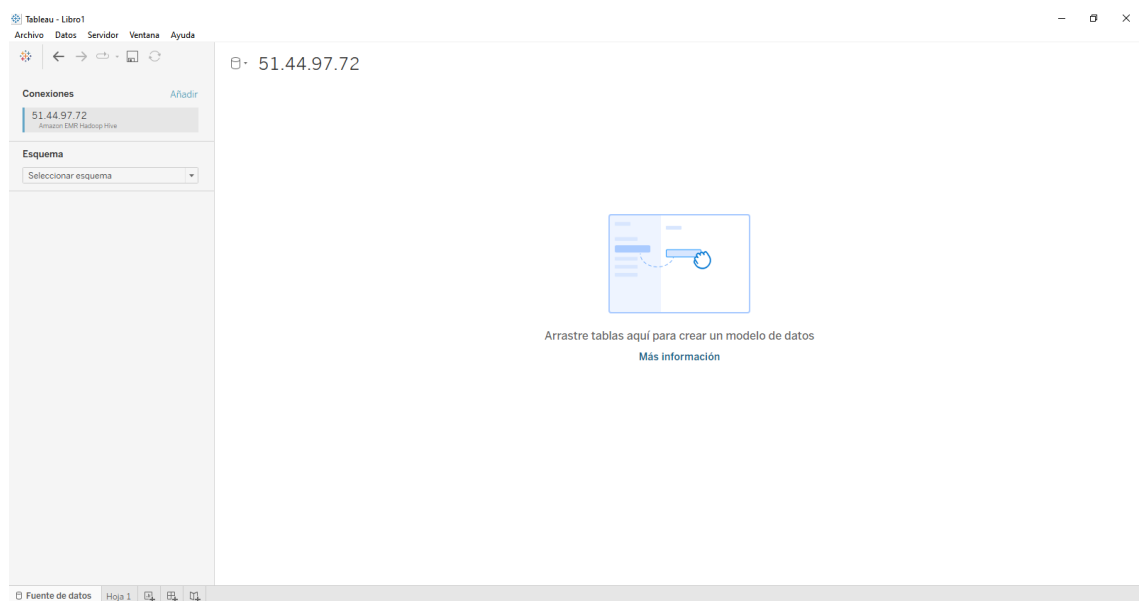
10. EXTRA: (Este apartado es optativo, su realización beneficiará a la nota de forma aditiva). La base de datos de Hadoop estará abierta durante un breve período de tiempo. Si logras conectarte exitosamente aporta pruebas de ello y explica cómo lo has logrado.



Para usar este conector, debe descargar e instalar el controlador.







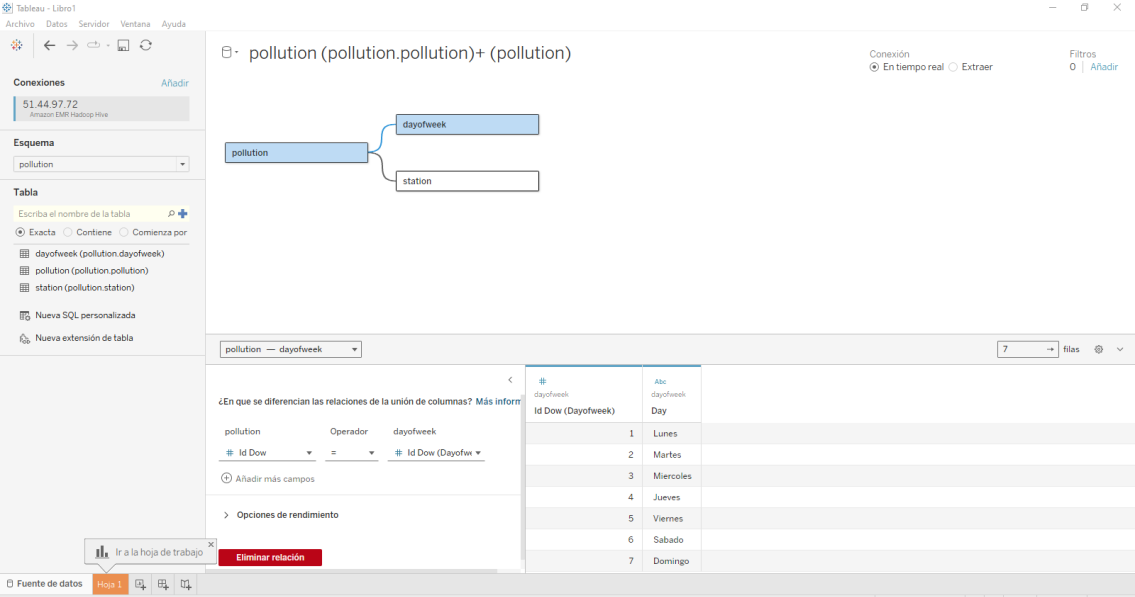
Selecciono el esquema de la base de datos → pollution. Arrastro y suelto sobre el dashboard la tabla que quiero visualizar y analizar.

#	pollution Id	pollution Fecha	#	pollution Id Dow	#	pollution Dom	Abc	pollution Hora	#	pollution Id Estacion	#	pollution Pm1	#	pollution Pm25	#	pollution Pm10	#	pollution No
1	01/01/2020	3	1	0:00:00	1	NULL	770000	83.0000	16.000									
2	01/01/2020	3	1	1:00:00	1	NULL	65.0000	70.0000	20.000									
3	01/01/2020	3	1	2:00:00	1	NULL	47.0000	51.0000	16.000									
4	01/01/2020	3	1	3:00:00	1	NULL	43.0000	46.0000	16.000									
5	01/01/2020	3	1	4:00:00	1	NULL	36.0000	39.0000	10.000									
6	01/01/2020	3	1	5:00:00	1	NULL	28.0000	30.0000	2.000									
7	01/01/2020	3	1	6:00:00	1	NULL	25.0000	27.0000	2.000									

Información de la relación entre las dos tablas.

#	station Id Estacion (Station)	Abc	station Name
1	01/01/2020	3	1

Modelo relacional con las tres tablas que aparecen en el esquema de la BD.



[CS01] Capturas de pantalla de la conexión: (0,2 puntos)

Elección y explicación del contaminante: (0,2 puntos)

[CS02 EXTRA] Conexión Hadoop: (0,5 puntos)

Creación del Dashboard

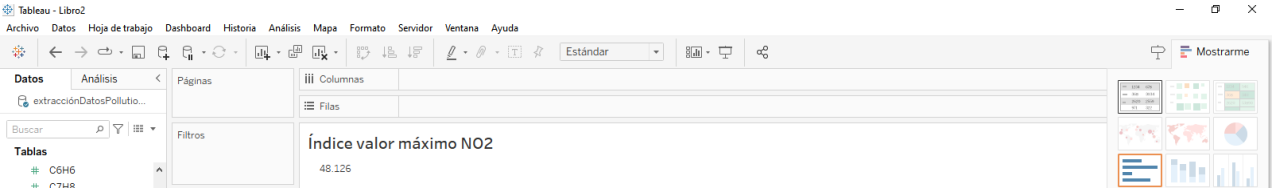
Realiza las siguientes tareas sobre el contaminante elegido:

11. Crea un índice que muestre el valor máximo detectado. Ponle un título.

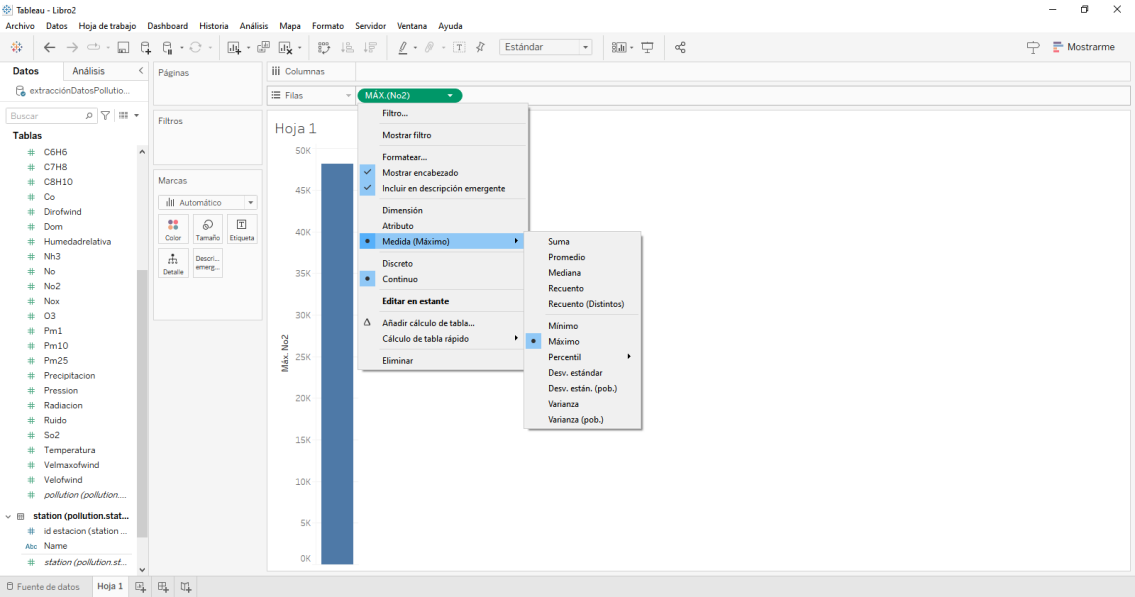
En la hoja de trabajo (Hoja 1) creo un índice que muestre el valor máximo del contaminante NO2. Para ello, vamos a mostrar tablas de texto y seleccionamos como medida el valor máximo.



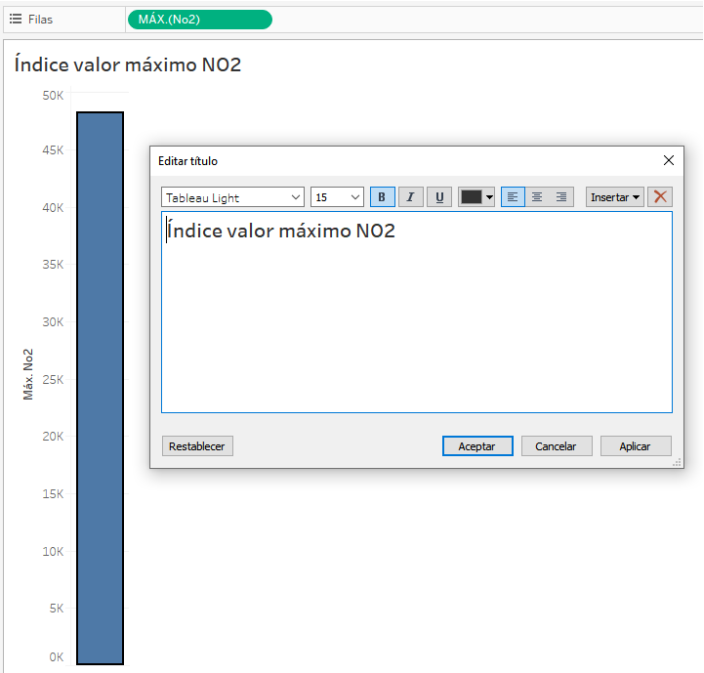
Le ponemos un título.



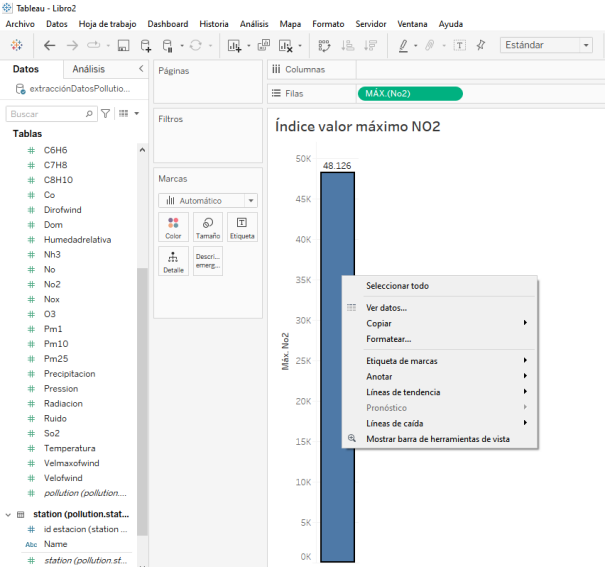
Otra opción es mostrar la barra.



Le ponemos un título.



Se añade la etiqueta con el máximo a la barra.



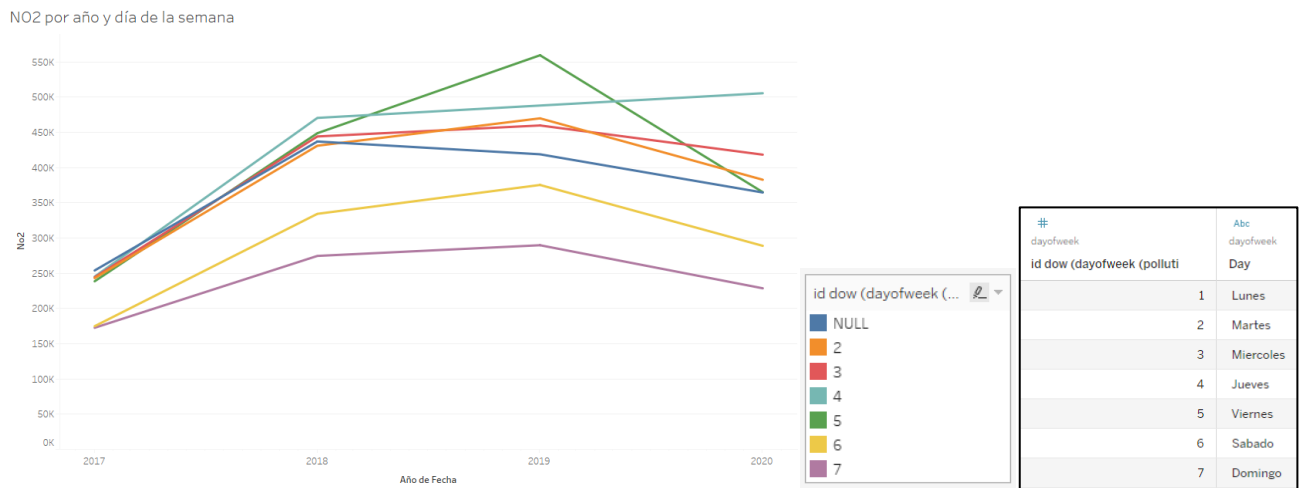
[CS03] Capturas de pantalla del dashboard - índice: (0,25 puntos)

12. Realiza 2 gráficos. Puedes usar de sectores, de barras, un mapa de calor... Los que tu elijas pero que sean 2 distintos. Selecciona los datos que consideres interesantes:

- o Para cada uno de ellos aplica filtros como (año, estación...)
- o Deberán de incluir tu contaminante escogido, además de otros campos que elijas
- o Explica qué muestra cada gráfico

Gráfico 1. Diagrama de líneas (continuas).

Muestra la suma del contaminante NO2 que se ha registrado en los años del 2017 al 2020. Se toman las fechas en años aunque también existe la opción de utilizar: trimestre, mes, número de semana y día. Las siete líneas que aparecen coloreadas representan los días de la semana (dayofweek).



Como puede verse en la imagen anterior, en el año 2019 se registra el pico más alto de este contaminante, en concreto, los viernes (línea verde). Por el contrario, los domingos es cuando los valores de NO2 son más bajos (línea morada). También se observa que los valores en 2020 son superiores a 2017 para todos los días de la semana.

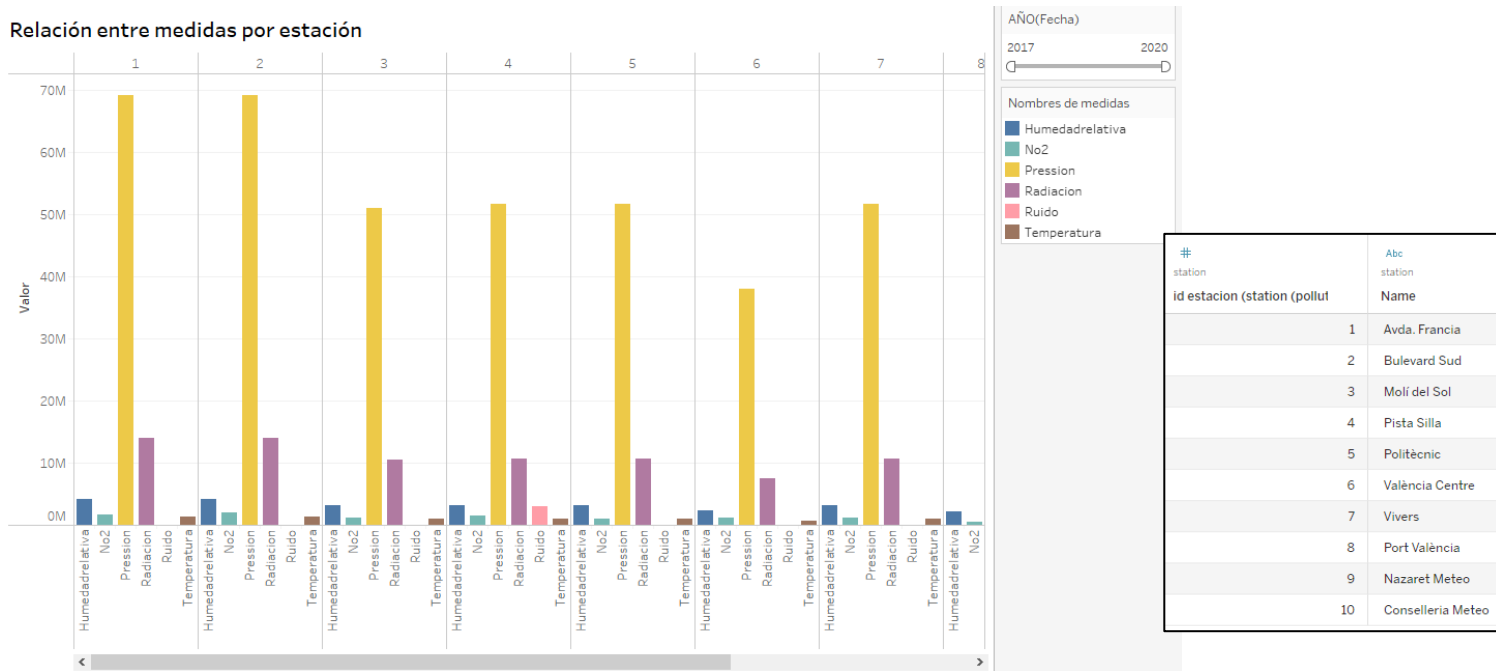
Ahora filtro únicamente las líneas para el último año.



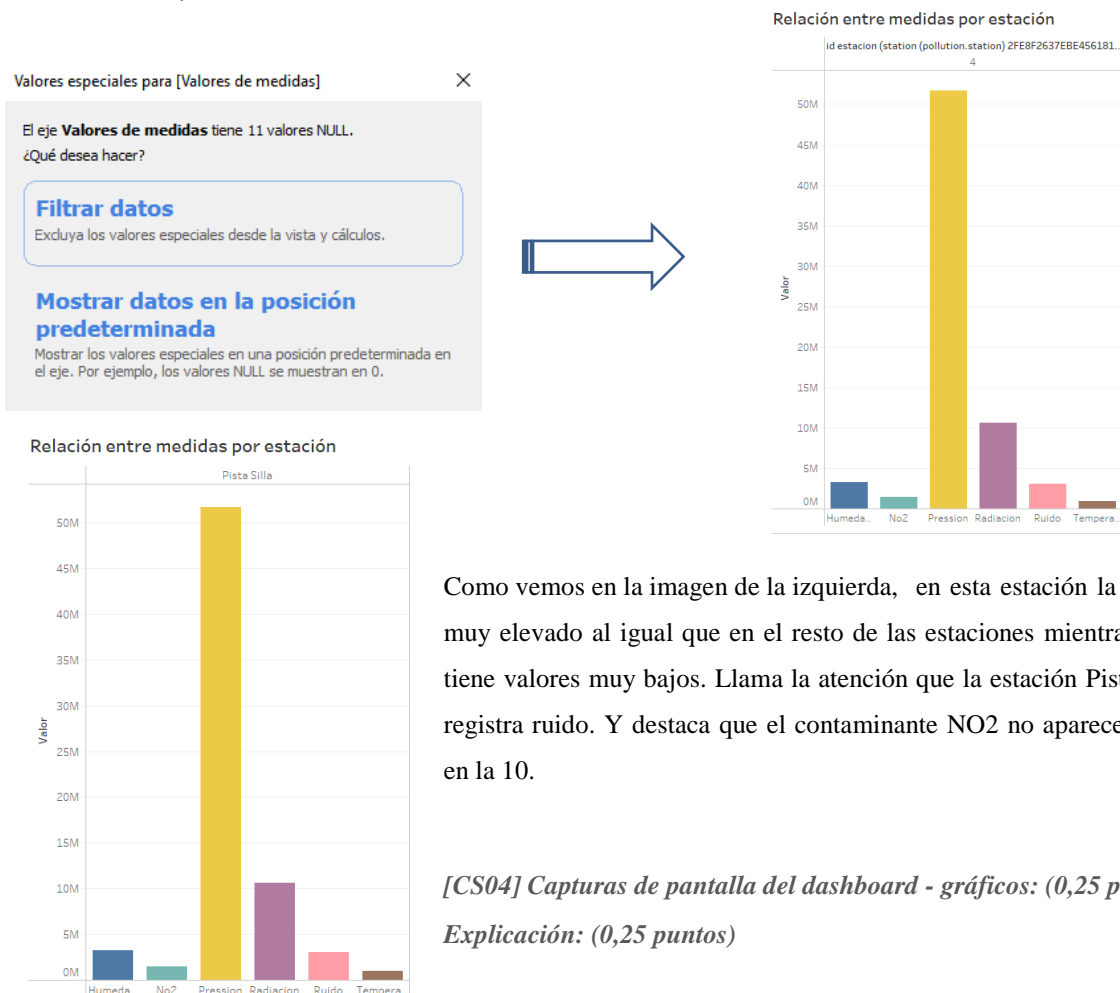
Se observa que este último año, los niveles de NO2 descendieron los viernes. En realidad, todos los días de la semana excepto los jueves (línea azul claro) sus valores decrecieron con respecto a 2019.

Gráfico 2. Diagrama de barras paralelas.

Muestra diferentes valores de medidas (humedad relativa, NO2, presión, radiación, ruido y temperatura) en las diez estaciones de Valencia (ver leyenda) desde el año 2017 hasta el año 2020. En la imagen se observa la barra de desplazamiento sobre el eje horizontal que permite recorrer los valores de todas las estaciones.



Como vemos que indica 11 NULL, vamos a filtrar los datos. El resultado es que se queda con la estación 4 porque no tiene valores nulos. Quitamos la etiqueta del campo “id estación” y sustituimos el id 4 por el nombre de esta estación “Pista Silla” con la opción de editar alias.



Como vemos en la imagen de la izquierda, en esta estación la presión tiene un valor muy elevado al igual que en el resto de las estaciones mientras que el contaminante tiene valores muy bajos. Llama la atención que la estación Pista Silla es la única que registra ruido. Y destaca que el contaminante NO2 no aparece ni en la estación 9 ni en la 10.

[CS04] Capturas de pantalla del dashboard - gráficos: (0,25 puntos)

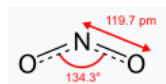
Explicación: (0,25 puntos)

13. Crea un dashboard con alguna imagen, texto y los indicadores/gráficos calculados en los apartados anteriores.

o Haz que los filtros sean compartidos entre los elementos del dashboard

o Guarda el dashboard en un fichero y compártelo con el grupo para crear una Story en la entrega grupal. (Ver siguiente sección)

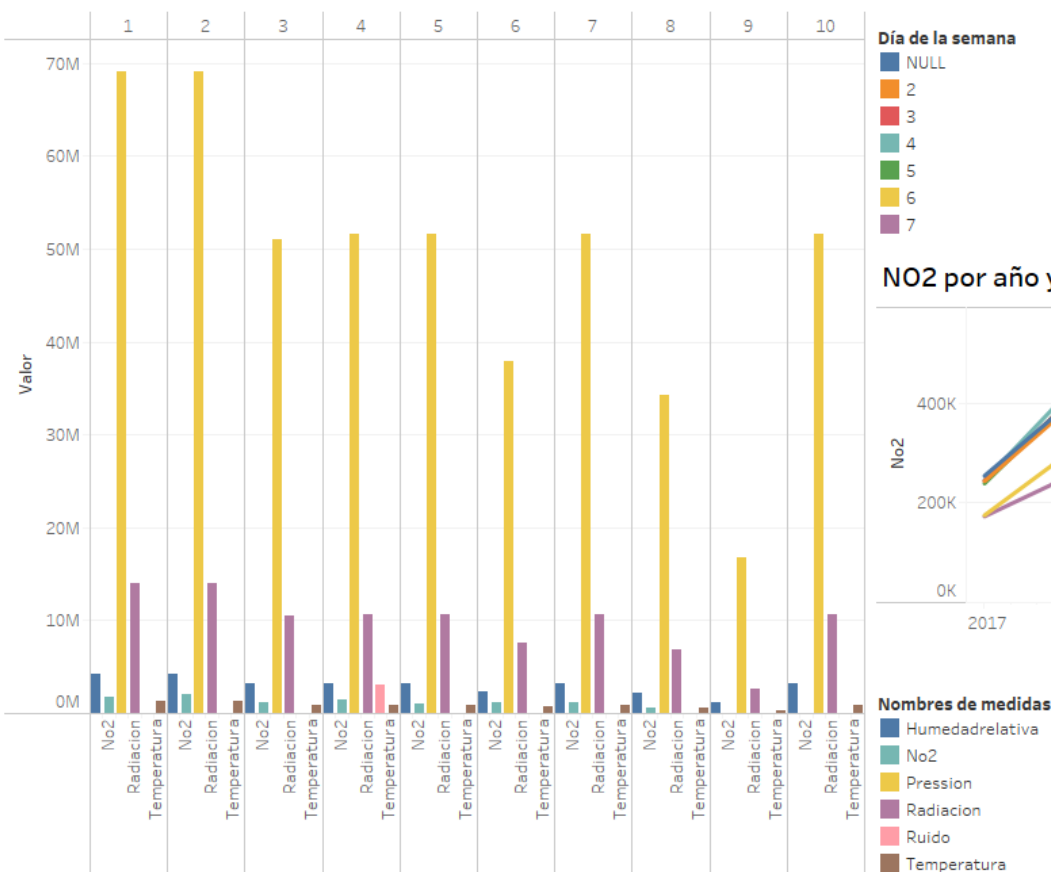
CONTAMINANTE NO2



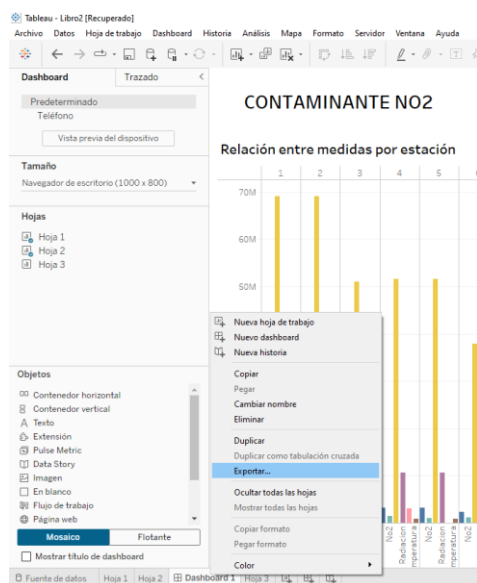
Para realizar este dashboard se ha cargado la extracción de datos de la base de datos *Pollution*.

Relación entre medidas por estación

Año de Fecha
2017 a 2020



Guardo el dashboard en un fichero para que creamos una Story todo el grupo en la siguiente sección.



Dashboard_Teresa