BIG DATA

VISUALIZACION DE DATOS QUICKSIGHT

EDUARD LARA

1. INTRODUCCION HERRAMIENTAS AWS

Las distintas herramientas para la visualización de datos son:.

- QuickSight
- □ Zepellin
- ☐ Jupiter Notebook
- □ D3.js
- Microstrategy

2. QUICKSIGHT

☐ Es un servicio de análisis de datos basado en la nube. ☐ Permite construir visualizaciones y hacer análisis personalizado de tus datos. ☐ Casos de uso ☐ Análisis de datos de marketing, de ventas, financieros, operaciones obtenidos de ficheros y de bases de datos. ☐ Las fuentes de datos pueden ser ☐ Bases de datos: Redshift, Aurora, MariaDB, SQL Server MySQL y PostgreSQL. ☐ Ficheros tanto de S3 como a servidores locales y serían en formato Excel, CSV, TSV, formato de log, etc. ☐ Las bases de datos pueden ser en el cloud o también podrían ser en servidores locales o discos locales.

3. ZEPPELIN

☐ Es un block de notas basado en web ☐ Permite análisis de datos interactivos, ☐ Permite utilizar Scala, Python, Spark SQL y Hive SQL ☐ Para manipular datos y visualizar resultados. ☐ Además, permite compartir documentos, ☐ Se integra fácilmente con S3 ☐ Casos de uso ■ Motores de recomendación ☐ análisis de datos en tiempo real.

4. JUPITER NOTEBOOK

Tupiter Notebook, que es muy similar.
s un block de notas basado en web.
Permite utilizar Python, R, Julia, Scala y otros mucho: enguajes.
Permite la colaboración de otros con otros usuarios compartiendo documentos
También es utilizado para DataScience
Ambos blocks de notas son muy similares.
Zeppelín se suele utilizar más para la parte de Spark, nientras Júpiter se utiliza mucho con Python

5. D3.JS

- □ D3.js que es una librería JavaScript que produce visualizaciones de datos dinámicas interactivas en navegadores web.
- ☐ Permite leer ficheros CSV, TSV, JSON
- ☐ Permite generar tablas HTML, gráficos de barras, dashboard en tiempo real, mapas y gráficos interactivos que funciona en un navegador web.

6. MICROSTRATEGY

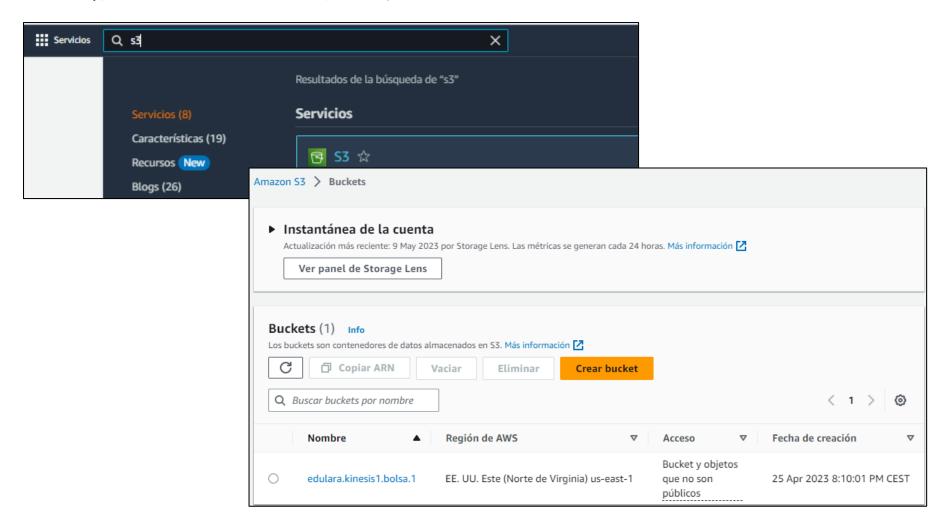
Es un software comercial de análisis de datos y BI (Bussiness Intelligent) que ayuda a los negocios a tomar mejores decisiones.
Permite crear también cuadros de mando, informes, consultas, etc
Se puede encontrar en el MarketPlace de AWS
Se puede ejecutar en una instancia de EC2
Tiene integraciones con Athena, Aurora, EMR, y Redshift
Tiene competidores, como pueden ser: Tableau, Cognos, Looker

Paso O. Vamos a hacer ahora un ejercicio de visualización y análisis de datos

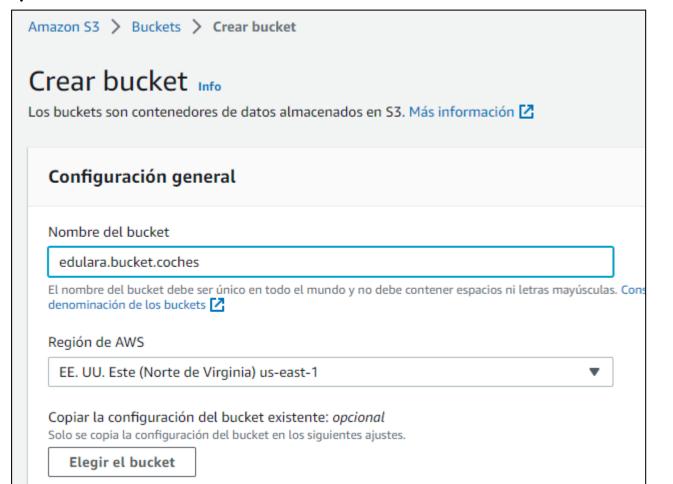
Cargaremos un fichero de datos en un bucket de 53 y mediante Quicksight crearemos gráficos a partir de los datos cargados en ese bucket.



Paso 1. Empezamos creando nuestro bucket en S3, pulsamos S3 y vamos a crear este bucket.

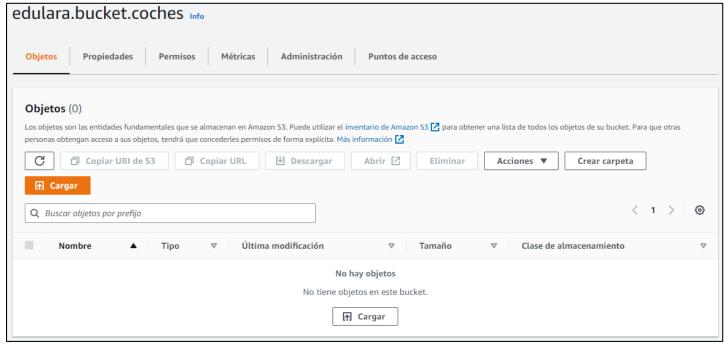


Paso 2. Le damos al botón de crear Bucket. Le ponemos un nombre único, por ejemplo, nuestro_nombre.bucket.coches. El resto lo vamos a dejar por defecto .Y le damos al botón Crear bucket

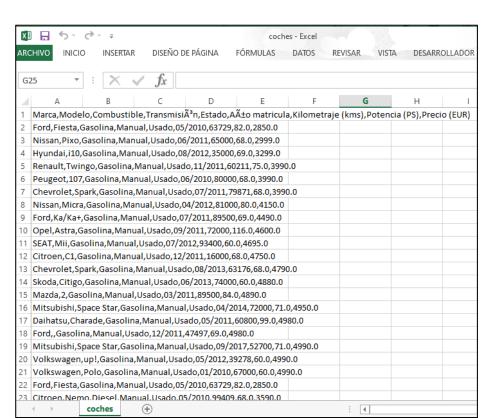


Paso 3. El bucket se ha creado correctamente. Si clickamos encima de él, podemos ver sus detalles

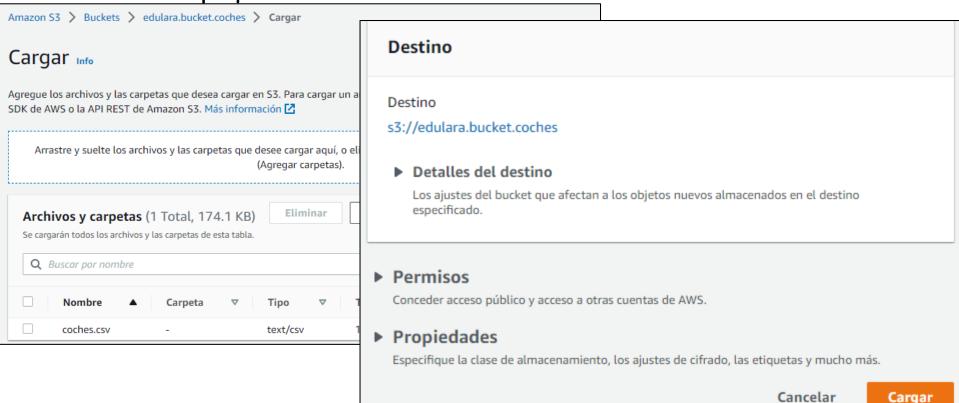
Los bud	ckets (2) Info ckets son contenedores de datos alma	acenados en S3. Más información 🔼	C 🗇 Copiar ARN Vac	iar	Eliminar	Crear	bucke >	et ©	
	Nombre 🔺	Región de AWS	▽	Acceso	▽	Fecha de creac	ión		∇
0	edulara.bucket.coches	EE. UU. Este (Norte de Virginia) us-east-1		Bucket y objetos que no son públicos		10 May 2023 1	1:01:20 P	M CES	Т
0	edulara.kinesis1.bolsa.1	EE. UU. Este (Norte de Virginia) us-east-1		Bucket y objetos que no son públicos		25 Apr 2023 8:	10:01 PM	CEST	



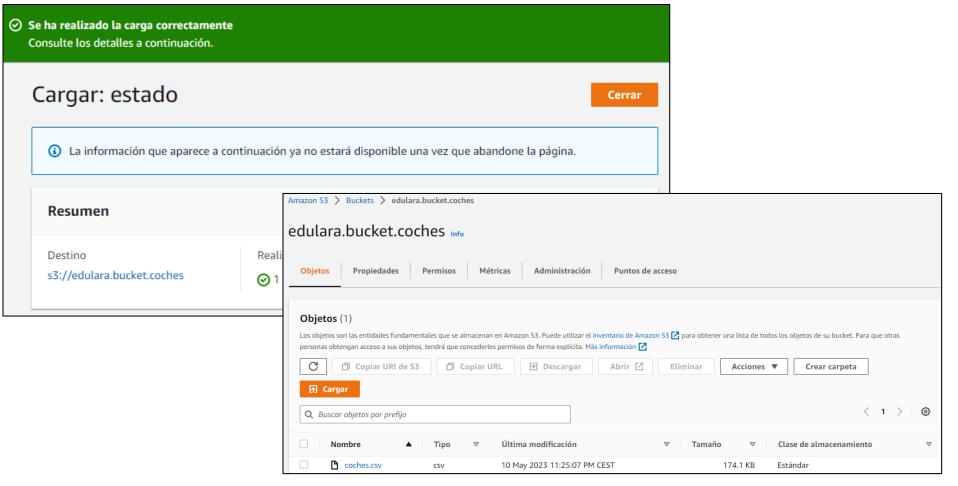
Paso 4. Cargaremos en este bucket el fichero de datos coches.csv, donde tenemos información de coches. Contiene una serie de columnas que son la marca, el modelo, combustible, transmisión, etc. y con los valores correspondiente en las filas de abajo. Cada fila es un coche que tiene valores para cada una de las columnas que están arriba.



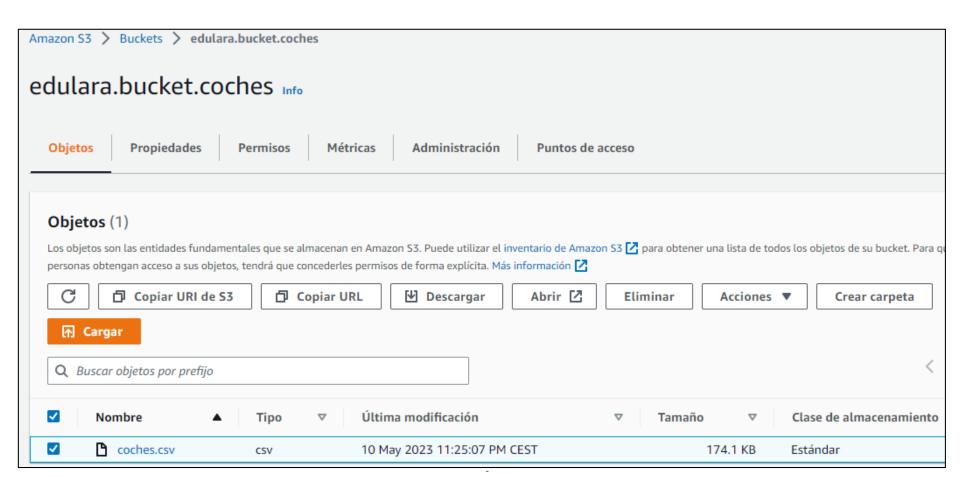
Paso 5. Cargamos este fichero csv en el bucket que acabamos de crear. Le damos al botón de cargar, después al botón Agregar archivos, lo seleccionamos, y el damos Abrir. Si bajamos a la parte de abajo podemos darle al botón de cargar. Lo que hace esto es subir el fichero al paquete.



Paso 6. Nos indica que la carga se ha realizado correctamente. Le damos al botón de cerrar y entonces vemos nuestro bucket con el fichero subido coches. CSV.

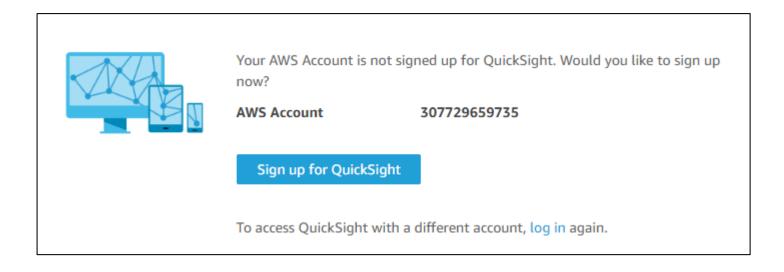


Paso 7. Si queremos verlo tendremos que darle a abrir o a descargar.

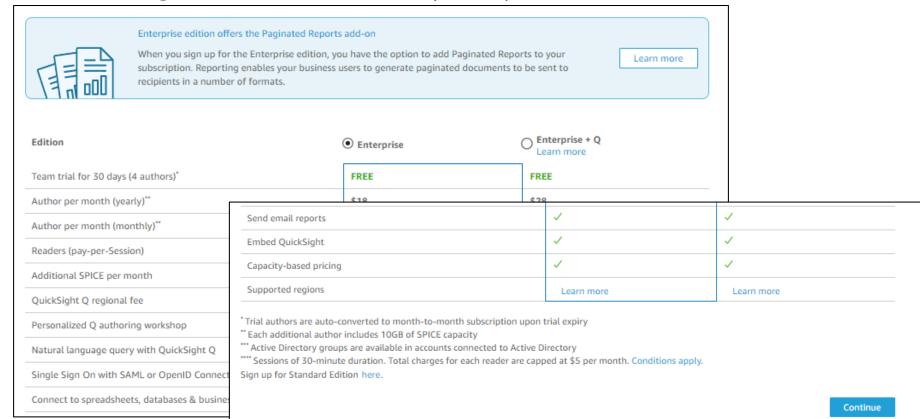


Paso 8. Buscamos el servicio QuickSight - Análisis empresarial rápido y fácil de usar. Como es la primera vez que entramos, tenemos que configurarlo. Le damos a Sign up for QuickSight.

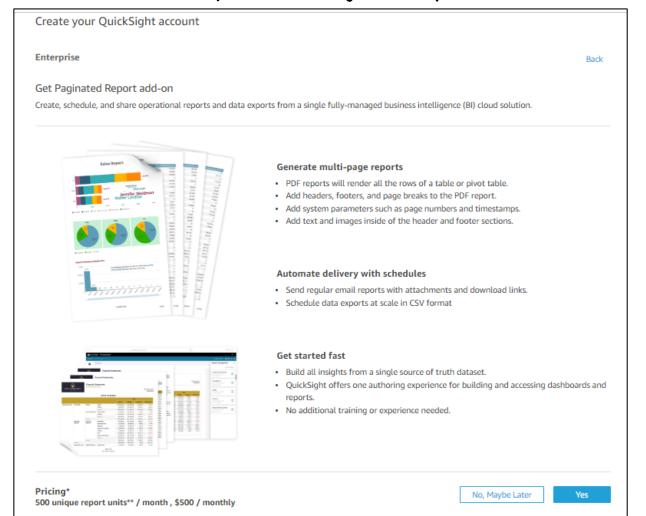




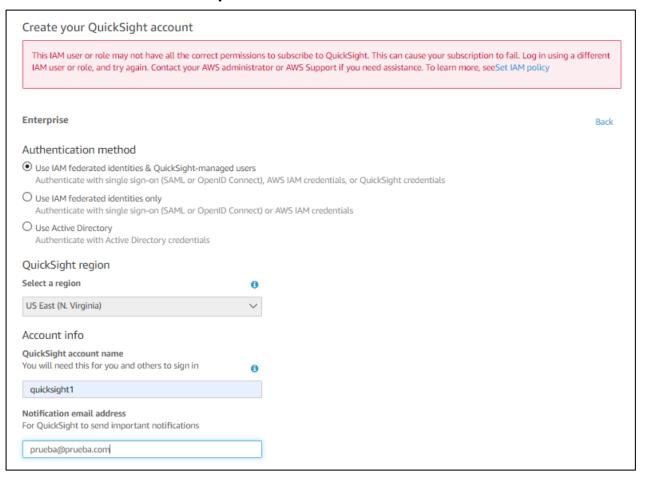
Paso 9. Elegimos la opción Enterprise, que tiene una versión gratuita durante 30 días y luego esos serian los costes. Ahora lo vamos a configurar y utilizar, y luego al final del laboratorio veremos como se desconfigura la instancia que vamos a crear para que no tenga ningún coste. Elegimos la versión Enterprise y le vamos a Continuar.



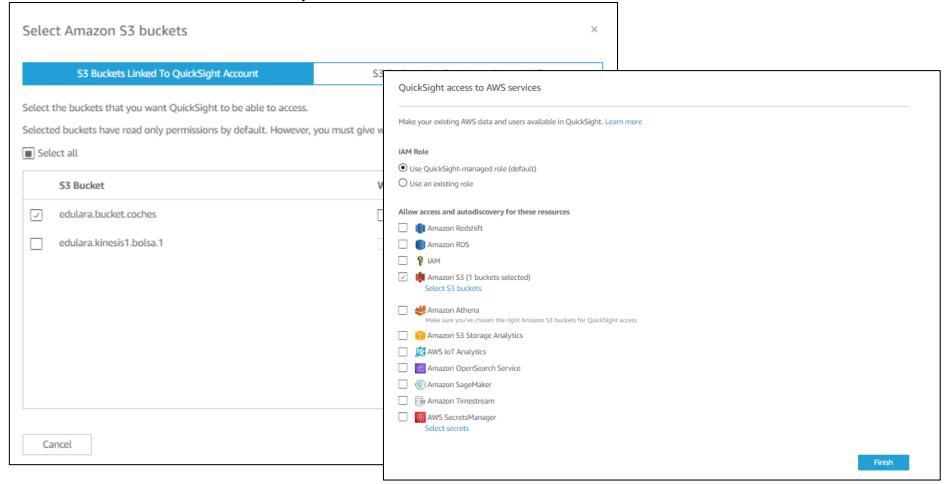
Paso 10. Nos preguntan si queremos el complemento de informes paginados. Indicamos que a lo mejor despues



Paso 11. Aquí tendremos que meter nuestros datos. Dejamos por defecto todas las opciones de Metodo de Autentificacion y Región. Ponemos nuestro nombre y nuestra dirección de correo



Paso 12. Deseleccionamos todos los recursos y seleccionamos sólo 53. Vamos a la pantalla de los buckets de 53 y seleccionamos el bucket de coches, que acabamos de crear anteriormente



Paso 13. Le damos al botón Finish. Esto nos configurará la cuenta para poder QuickSight. Esperamos 4 o 5 minutos y ya podremos utilizar la herramienta.

This IAM user or role may not have all the correct permissions to subscribe to QuickSight. This can cause your subscription to fail. Log in using a different IAM user or role, and try again. Contact your AWS administrator or AWS Support if you need assistance. To learn more, seeSet IAM policy

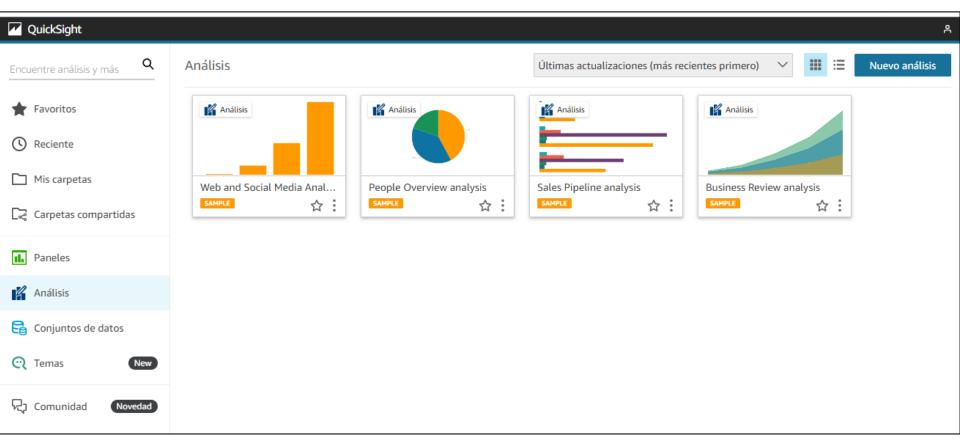
Congratulations! You are signed up for Amazon QuickSight!

Access QuickSight with the following information
Account name: quicksight11

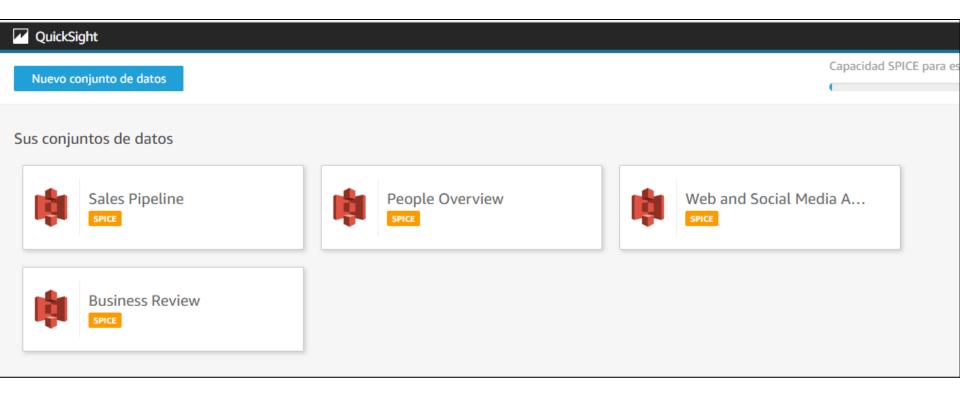
Go to Amazon QuickSight

NOTA: Al final de este laboratorio la quitaremos para darle ningún coste.

Paso 14. Una vez ha finalizado, le damos al botón de Go to Amazon QuickSight y entonces ya nos abriría la herramienta de servicio de QuickSight de Análisis de datos para Amazon AWS. Vamos a crear un nuevo análisis, En la pantalla inicial, le damos al botón Nuevo Analisis.

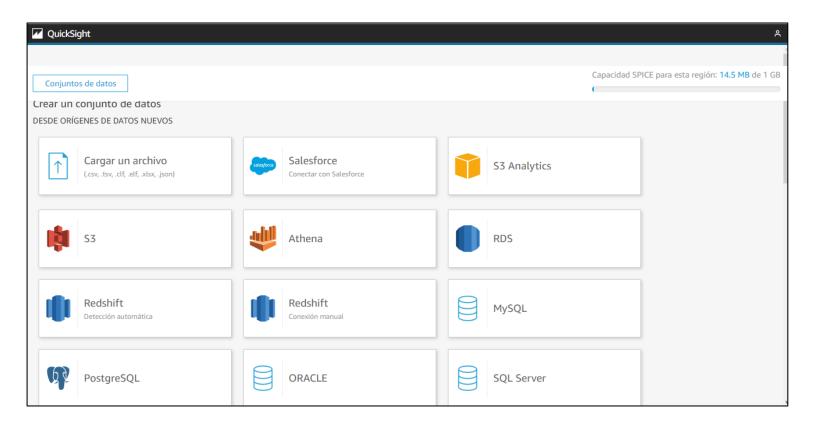


Paso 15. Aquí vamos a crear un nuevo conjunto de datos. Le damos al botón Nuevo conjunto de datos



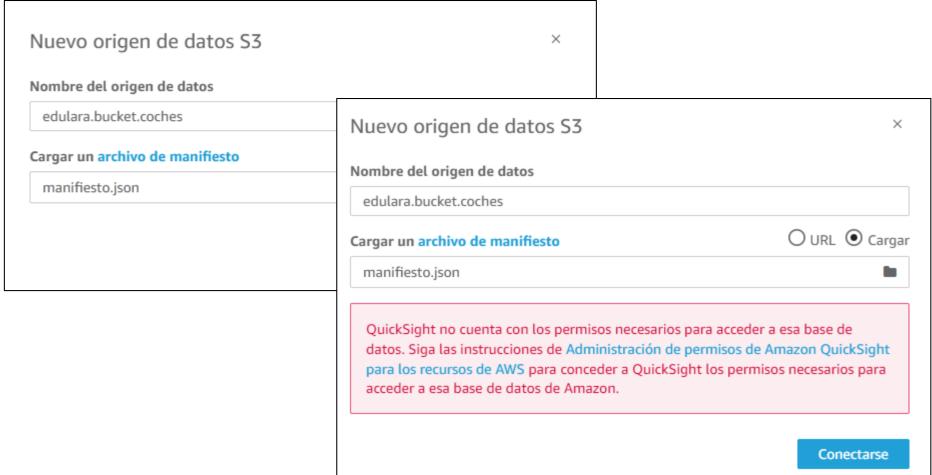
23

Paso 16. Tenemos diferentes formas de añadir un nuevo conjunto de datos, como puede ser crear un archivo directamente, Salesforce, S3 Analytics, bases de datos, etc. Nosotros vamos a acceder a S3, ya que hemos subido el fichero de coches.csv al bucket de S3



Paso 17. Le tenemos que decir el nombre de nuestro bucket y tenemos que subir un manifiesto. El manifiesto es un fichero JSON que tiene una serie de propiedades. Descargamos el fichero y sólo tenemos que cambiar el nombre de nuestro bucket. Lo demás deja por defecto: tenemos un fichero CSV, delimitador entre campos es una coma, etc Salvamos el archivo.

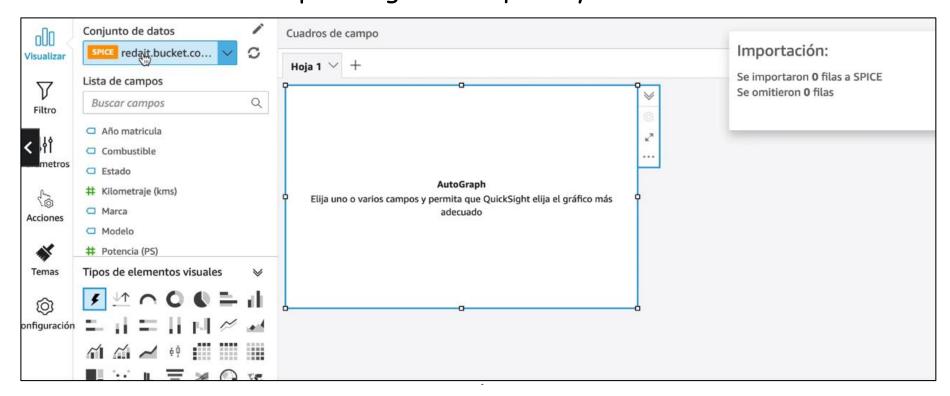
Paso 18. Una vez hemos puesto en Nombre de origen de datos nuestro bucket, buscamos y cargamos el archivo manifiesto JSON ya modificado y le vamos al botón Conectarse



Paso 19. Nos indica que va a guardar la información en SPICE y simplemente va a leer la información y tenerla disponible para agilizar los análisis y los gráficos.



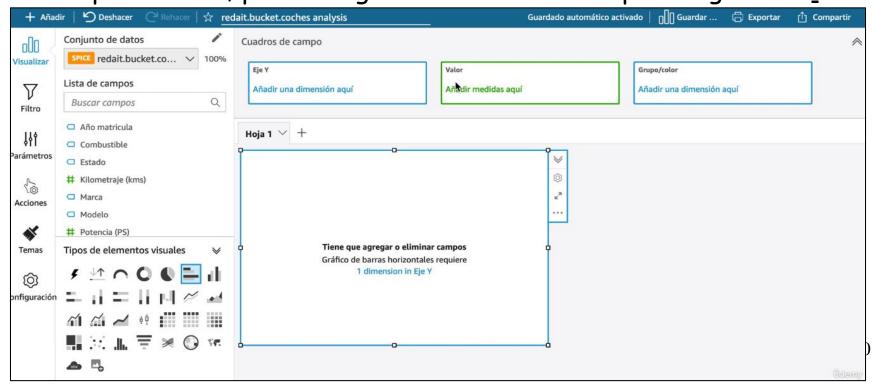
Paso 20. Le vamos a visualizar. Puede tardar 1 o 2 minutos en cargar el fichero, depende del tamaño. A partir de aquí ya podemos empezar a analizar la información que acabamos de subir del fichero de datos, usando todos los tipos de gráficos que hay



Paso 21. Una vez completada la carga 100% del fichero, nos indica el número de filas que se han cargado. Ya podemos empezar a hacer nuestros gráficos.

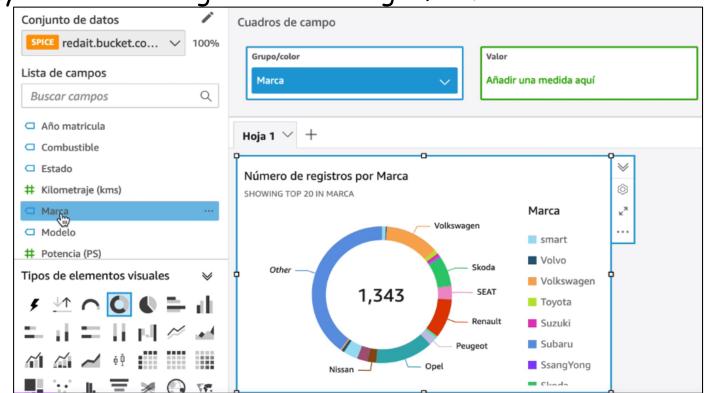


Paso 22. Según el tipo de gráfico seleccionado tenemos que poner valores en el eje X y en el eje Y. En la lista de campos estarían las columnas de la tabla, que son las columnas de nuestro CSV (la primera fila del fichero csv, el resto de filas son los valores). Simplemente poniendo el gráfico que queremos hacer y seleccionando las columnas correspondientes, podemos generar diferentes tipos de gráficos.



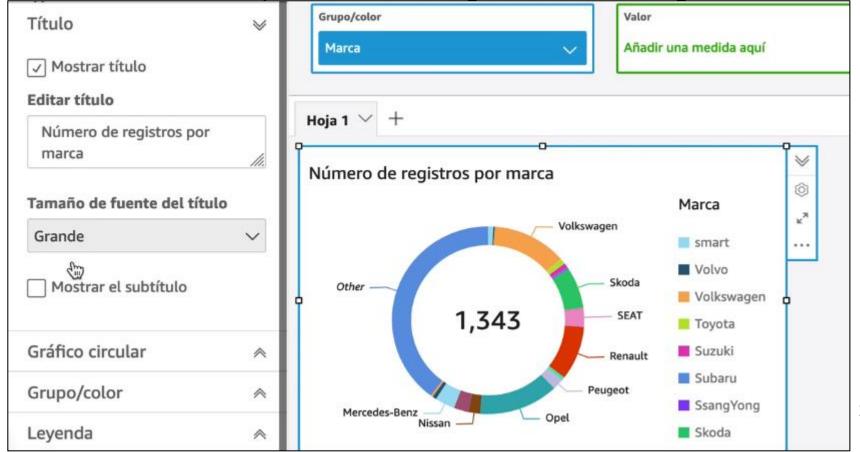
Paso 23. Vamos a crear nuestro primer gráfico. Usaremos un gráfico circular. Añadiremos en el Grupo/Color, la columna marca.

Automáticamente con los datos que tiene guardados en bucket de 53 construye un gráfico circular con las marcas. Volkswagen tiene el 13% de artículos, Renault el 9%, Opel tiene el 13%, etc. De una manera muy sencilla se ha generado este gráfico.



₩∅^x...

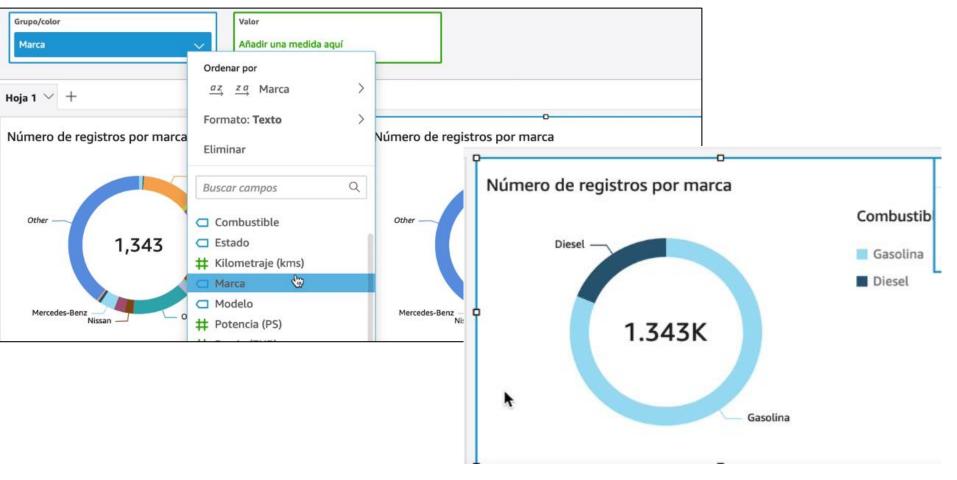
Paso 24. Se puede personalizar, cambiándole el título. Le damos a este_botón de edición y podemos editar el título. Podemos poner un tamaño grande, pequeño, etc podemos mostrar o no el subtítulo. De esta forma sencilla podemos ir configurando nuestro gráfico



Paso 25. Si queremos generar otro grafico parecido (poner en Grupo/Color en lugar de marca, modelo o tipo de motor, etc) podemos copiarlo haciendo click en el botón derecho Duplicar elemento visual. Nos genera uno igual.



Paso 26. Si seleccionamos el nuevo grafico (queda marcado en azul), podemos cambiar en lugar de marca, por ejemplo combustible. Nos genera un grafico donde diésel es el 9% y gasolina el 81%



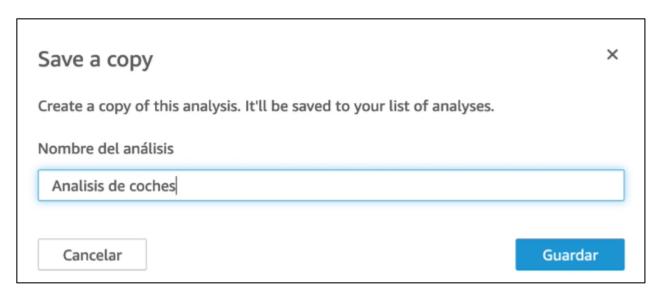
Paso 27. Si queremos ver el estado de los coches, porcentaje de coches de segunda mano o nuevos, vemos que un 9% son nuevos y un 91% son de segunda mano



Paso 28. Podemos guardar el nombre de este análisis/grafico.



Podemos poner de nombre Análisis de coches, y QuickSight nos guarda esta información que podemos visualizar más adelante, sin tener que volver a crearla nuevamente.

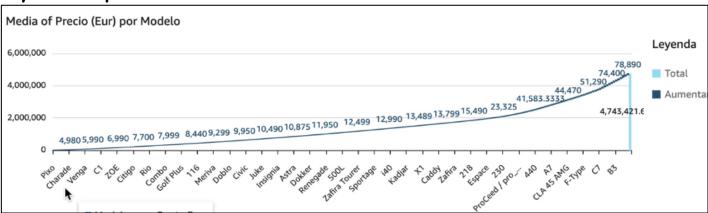


Paso 29. Vamos a hacer un nuevo gráfico, para ver el precio medio por modelo. Añadimos un nuevo gráfico con el botón Añadir/Añadir elemento visual. Y elegimos un gráfico en cascada.

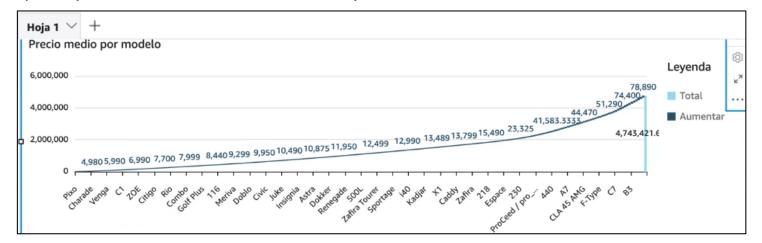
En categoría (eje X), ponemos Modelo y en Valor (eje Y) ponemos Precio. En lugar de sumar, ponemos el valor medio, para saber el precio medio por modelo. También se puede ordenar de mayor a menor o de menor a mayor



Paso 30. Podemos aumentar el tamaño del grafico para que se vea mejor, y nos aparecen los modelos más baratos hasta los mas caros



Podemos cambiar el título y ponerlo en español. Vamos a la pestaña de título y le ponemos Precio medio por modelo, sin subtitulos



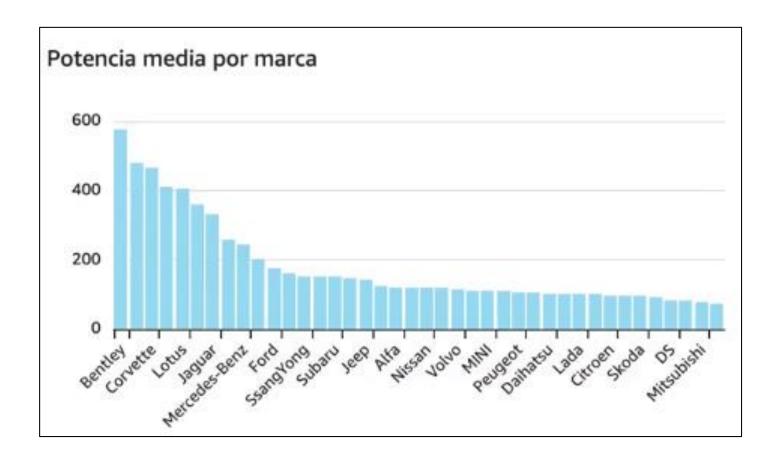
Paso 31. Vamos a utilizar un diagrama de barras verticales, para generar un grafico de la potencia media por marca. Añadimos un nuevo elemento visual. Y elegiremos el gráfico de barras apiladas verticales.



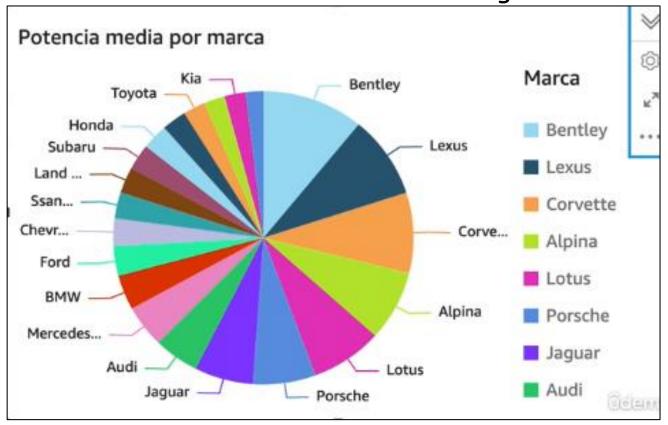
En el eje \times vamos a poner la marca. Y como valor vamos a poner en este caso la potencia. Cambiamos la potencia suma por la potencia media. Lo dejamos ordenado de mayor a menor.



Paso 32. Obtenemos nuestro gráfico de potencia media por marca, donde podemos cambiar el título y eliminar el subtitulo. Los de mayor potencia son Bentley, Corvette, Lotus, etc



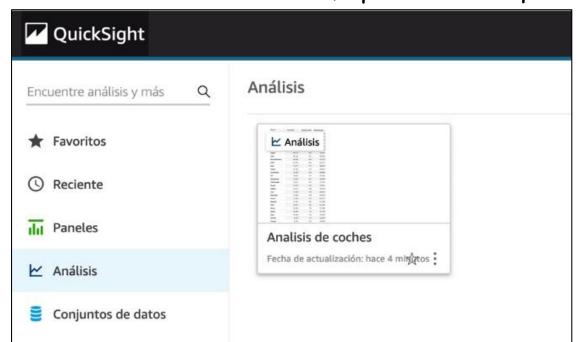
Paso 33. Vamos a generar otro gráfico parecido. Duplicamos este elemento visual y cambiamos el tipo de grafico a un grafico circular. Genera la misma información del diagrama de barras, pero ahora representado de una forma distinta con un diagrama circular.



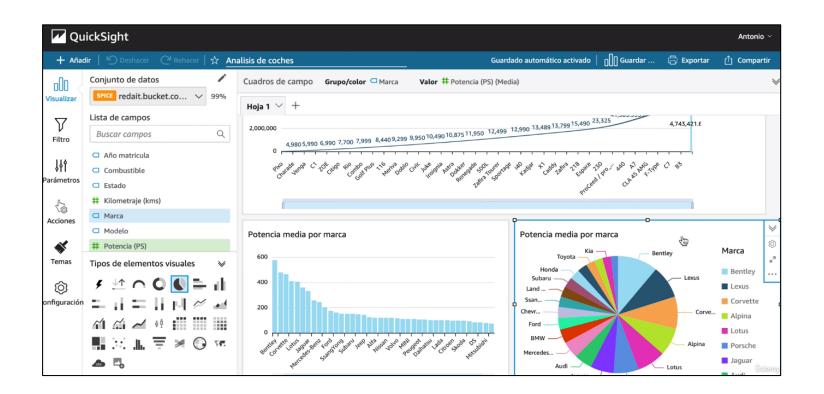
Paso 34. Una vez que tenemos todos nuestros gráficos creados, podemos guardar los análisis realizados en el botón Guardar, con el nombre Análisis de coches QuickSight



Si vamos a la pagina principal de QuickSight, haciendo click encima del icono de QuickSight, aparece nuestro análisis de coches. Si tuviéramos más análisis realizados, aparecerían aquí



Paso 35. Si pulsamos en nuestro análisis, volvemos a cargar la información del bucket y los resultados de las gráficas que habíamos construido. De manera sencilla siempre tenemos disponible aquí toda la información de estos gráficos.



Paso 36. Si quisiéramos exportar esta información en un PDF, vamos a Exportar/Descargar como PDF. Después de 1 o 2 min, aparece un enlace para descargar el archivo. Genera un PDF con la información de los gráficos creados,

Descargar como PDF

[Î] Compart

Exportar

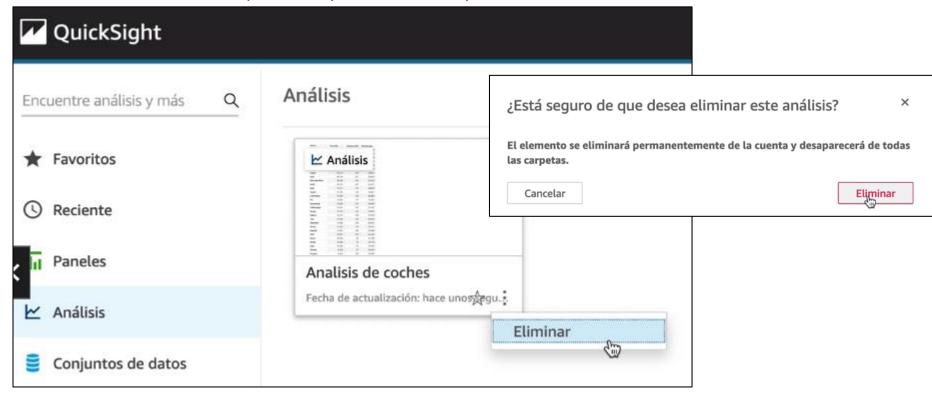
Imprimir



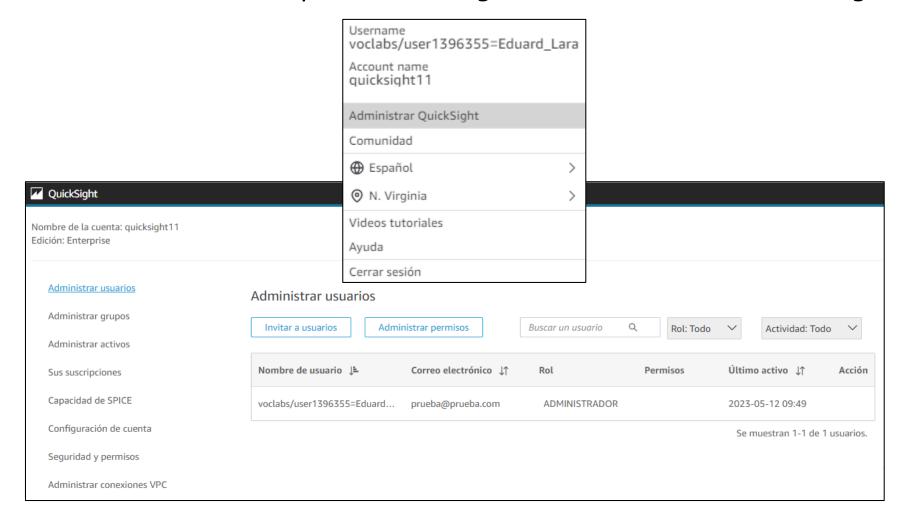
Paso 37. Vamos a ver cómo podemos ir eliminando gráficos, análisis, incluso darnos de baja de QuickSight para que no nos cobren nada. Si queremos eliminar un gráfico, hacemos click sobre los tres puntos, y seleccionamos. Asi podemos borrar un gráfico de manera muy sencilla.



Paso 38. Si queremos eliminar de QuickSight todo el análisis completo, vamos a la gestión de análisis de QuickSight (haciendo click encima del logo de QuickSight), pulsamos en los tres puntos de nuestro análisis y nos aparece la opción Eliminar



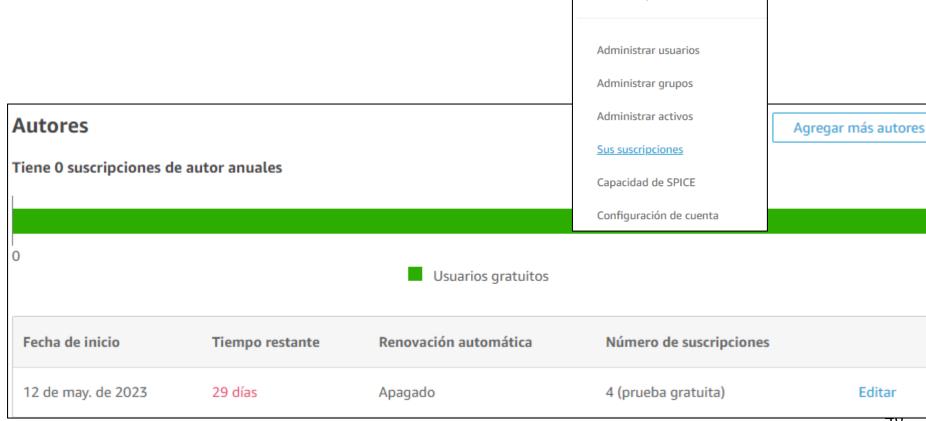
Paso 39. Si queremos darnos de baja de QuickSight para que no nos cobre más, vamos a la parte de configuración, Administrar QuickSight



Paso 40. Si vamos a Sus suscripciones, podemos ver la suscripción que tenemos, que quedan 29 días de gratuit quicksight

Nombre de la cuenta: quicksight11

Edición: Enterprise



Paso 41. Si queremos darnos de baja de forma temporal, vamos a Configuración de Cuenta y pulsamos el botón Administrar que lo que hace es iniciar el proceso de borrado de esta cuenta. Ya no nos van a cobrar por el uso que hagamos de QuickSight. Si el dia de mañana queremos volver a utilizarlo, podemos darnos de alta nuevamente en QuickSight y volver a empezar de cero, cargar los datos, hacer los gráficos, etc



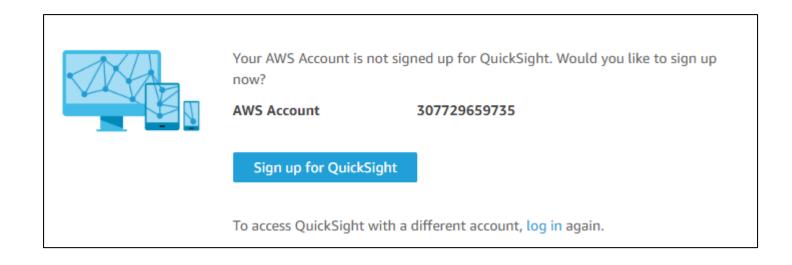
Paso 42. Nos pide que confirmemos la terminación de la cuenta de QuickSight, escribimos confirmar y ya podemos hacer click en el botón Eliminar la cuenta. Nos indica que la terminación ha tenido éxito

Terminación de la cuenta		
Nombre de la cuenta de QuickSight	quicksight11	
Protección contra la terminación de cuentas ① La protección contra la terminación de cuentas es una medida de seguridad adicional para ayudar a evitar la eliminación accidental de las cuentas.		
Estado de la protección contra la terminación de cuentas: off.	Unsubscribe suc	
Eliminar la cuenta La eliminación de esta cuenta es una acción que no se puede desh permanentemente todos los usuarios, paneles, análisis, junto con ot relacionados.	pudimos eliminar. Exar	ntenido se ha eliminado, pero hay algunos elementos que no mine los elementos que se indican a continuación para asegurarse QuickSight no genere costos adicionales. Go to AWS console
Escriba "confirmar" para eliminar esta cuenta confirmar		50

Eliminar la cuent

Cancelar

Paso 43. Podemos ir algo a la consola de AWS y al cabo de un rato ya no podemos entrar en QuickSight y tenemos que volver a configurarlo de nuevo.



51