Data Lifecycle en CDP Public Cloud

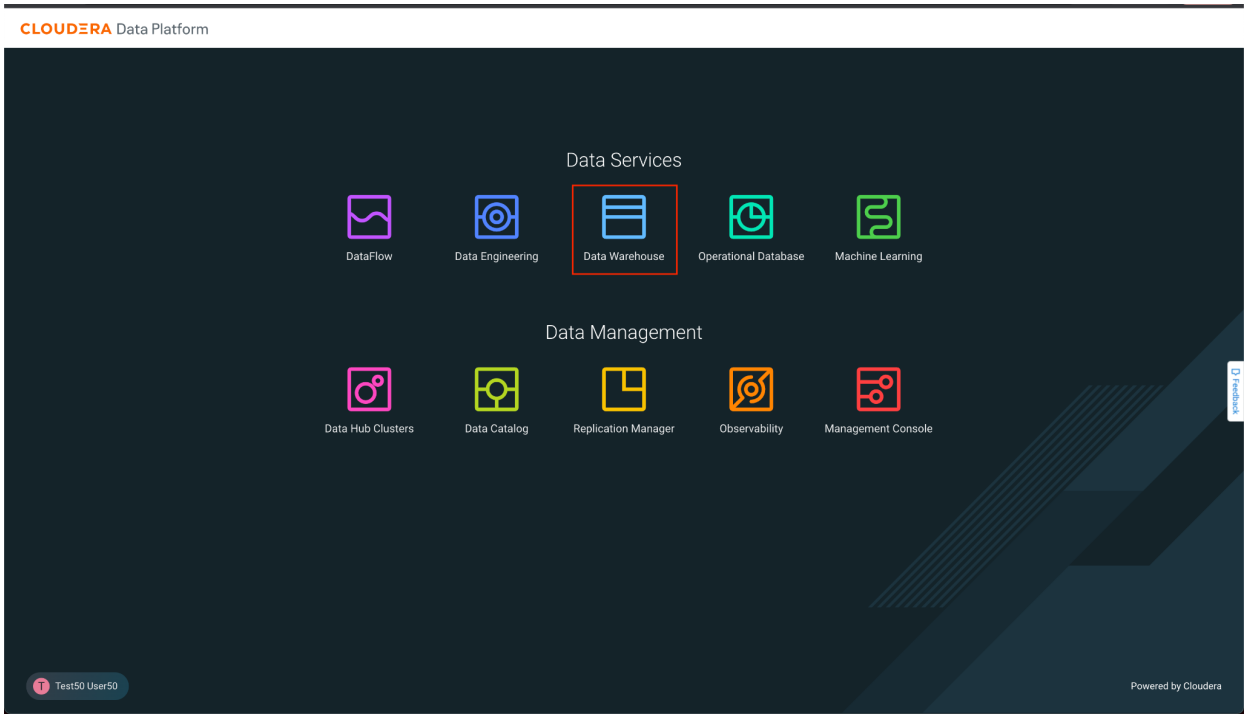
Laboratorio Data Warehouse

Parte 1: Desarrollo dashboard

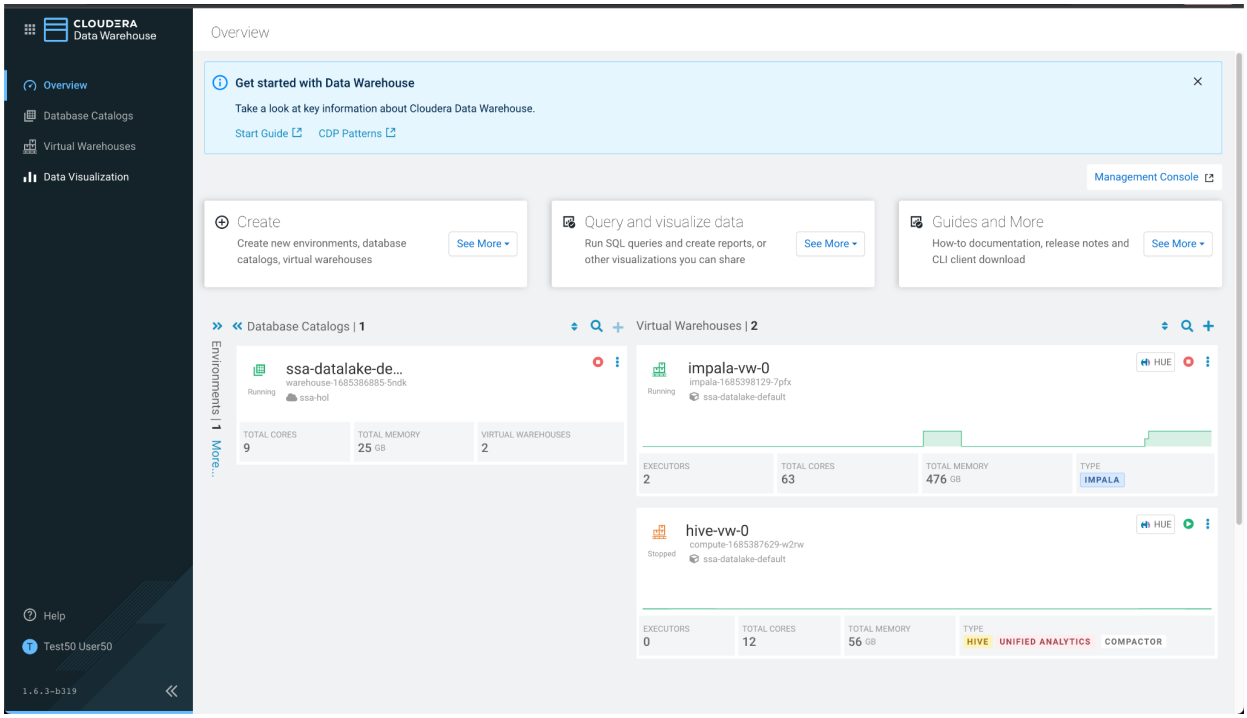
Objetivos:

- Crear un dataset que apunte a la tabla
- Crear un dashboard con métricas y dimensiones

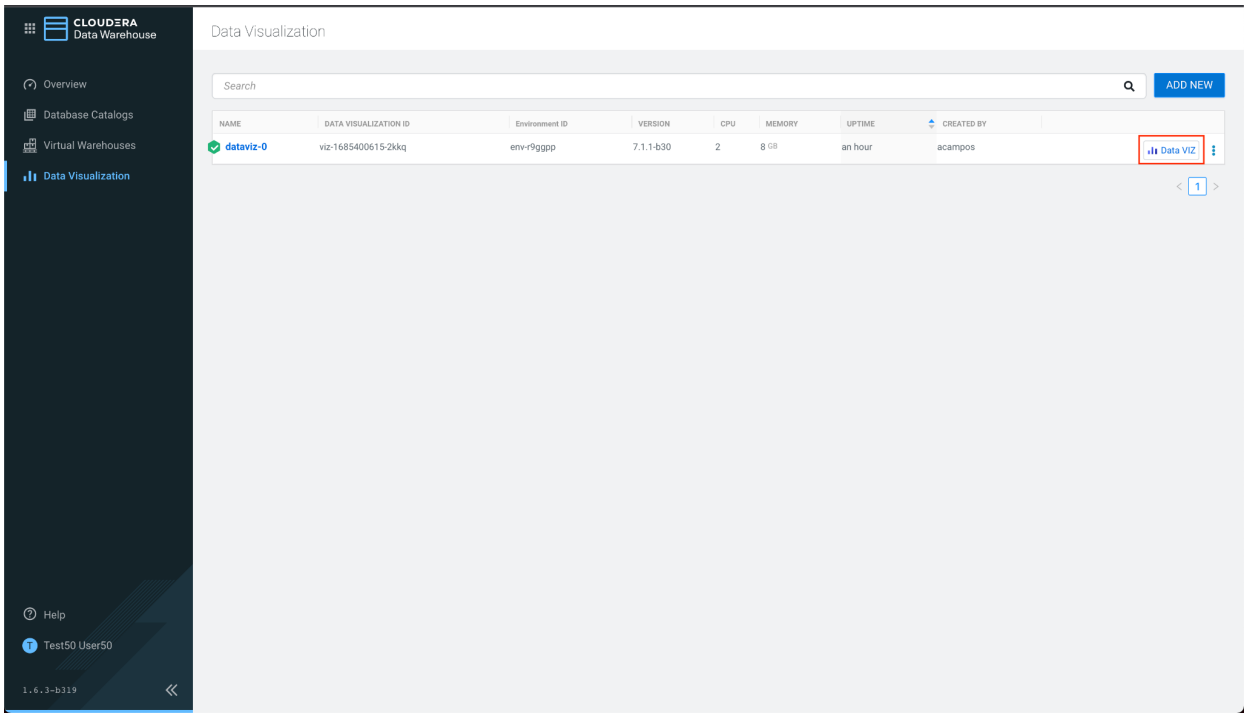
1. Hacer clic en Data Warehouse desde el Home de CDP PC:



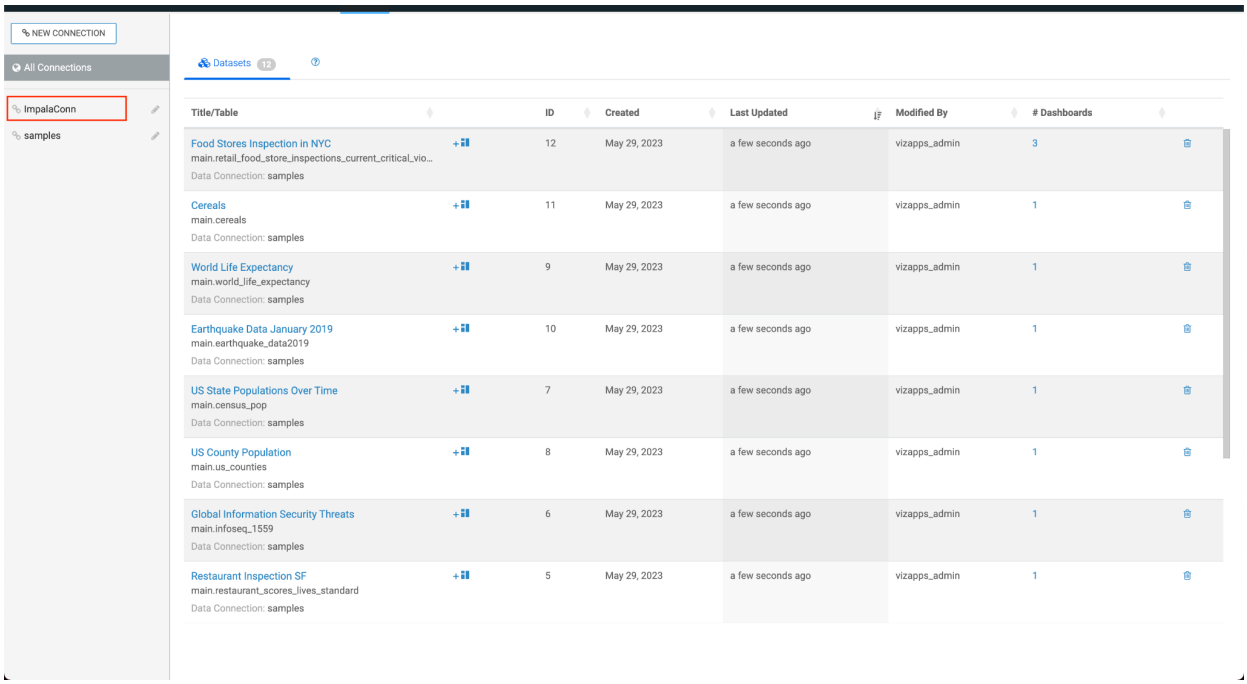
2. Pantalla de bienvenida de Data Warehouse. Hacer clic en Data Visualization en el menú izquierdo.



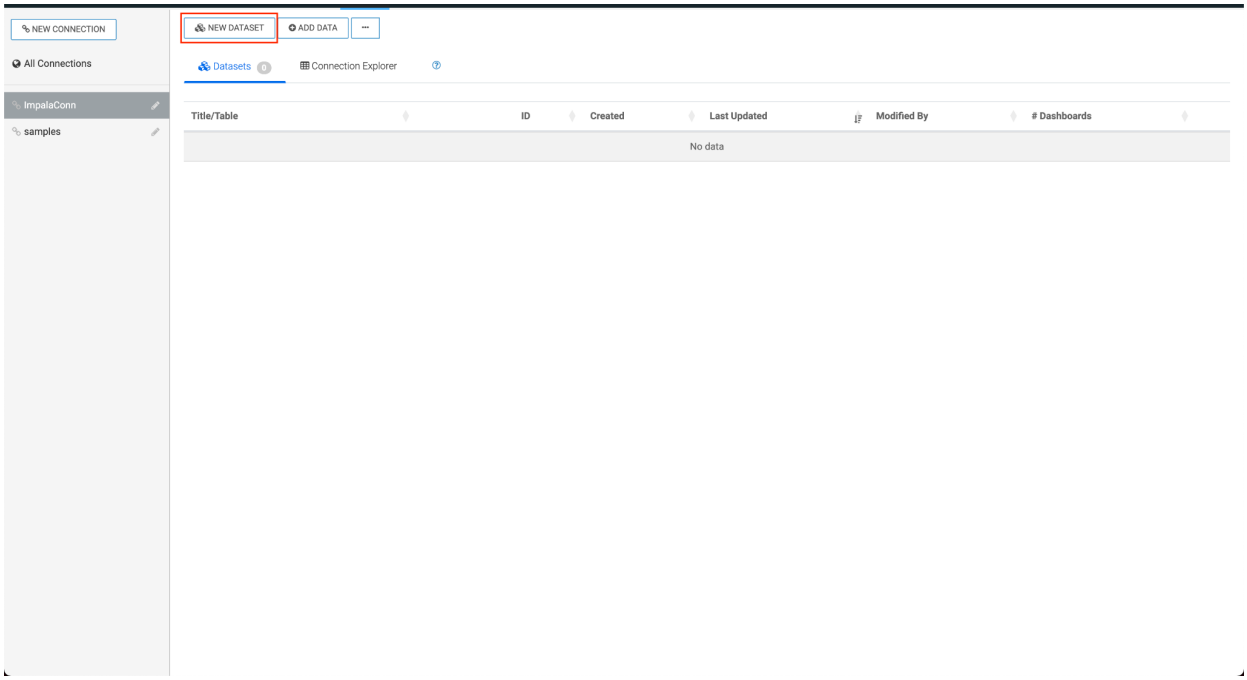
3. En Data Visualization, hacer clic en el botón **Data Viz** del cual les fue asignado.



4. Una vez en Data Visualization, ir a la opción Data del menú superior, y posteriormente a la Connector **ImpalaConn** del menú izquierdo.



5. Tenemos que crear un nueva fuente de datos, para eso, hacer clic en New Dataset y aparecerá una ventana para ingresar la información de la nueva fuente de datos.

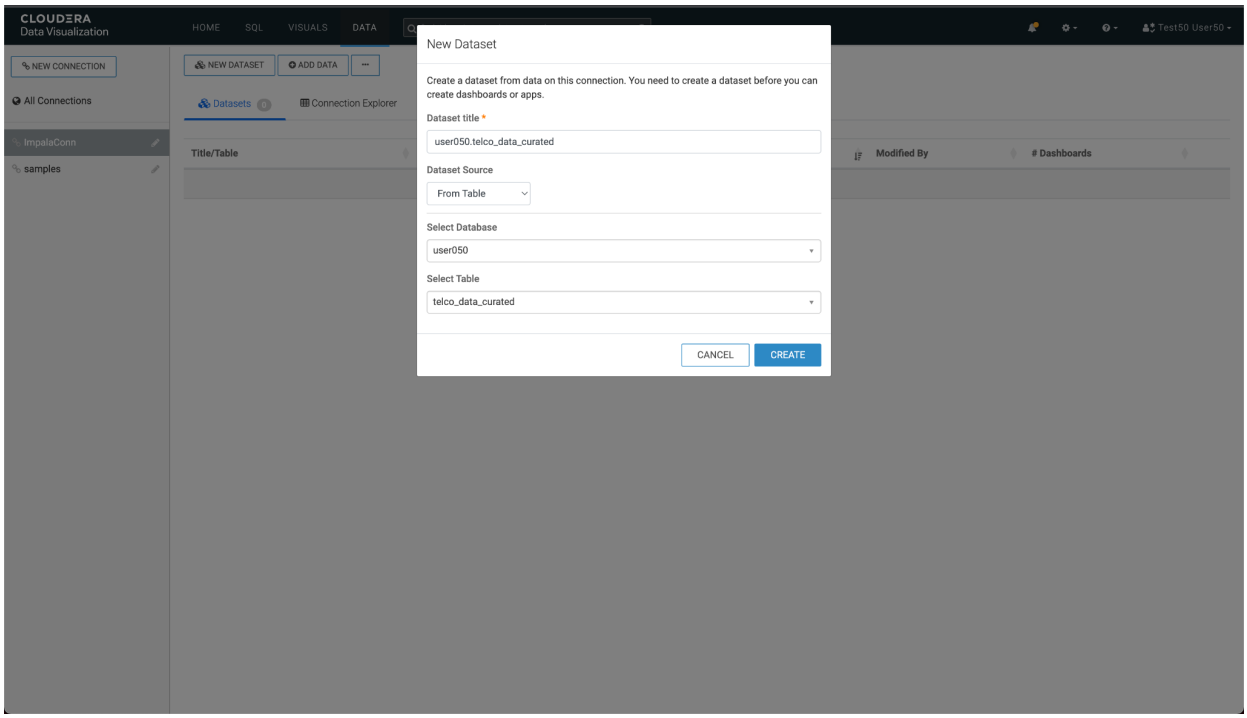


6. Ingresar la información de la nueva fuente de datos:

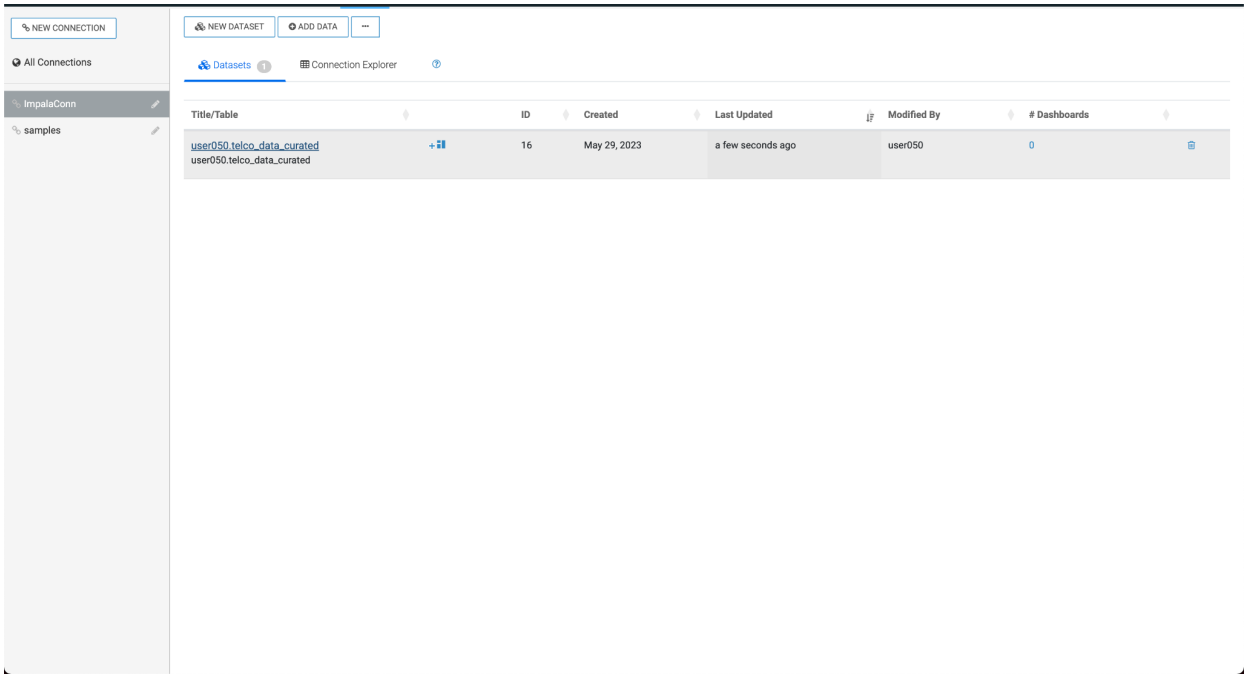
Dataset title: <usuario_asignado>.telco_data_curated
Dataset Source: From table

Select Database: <usuario_asignado>
Select Table: telco_data_curated

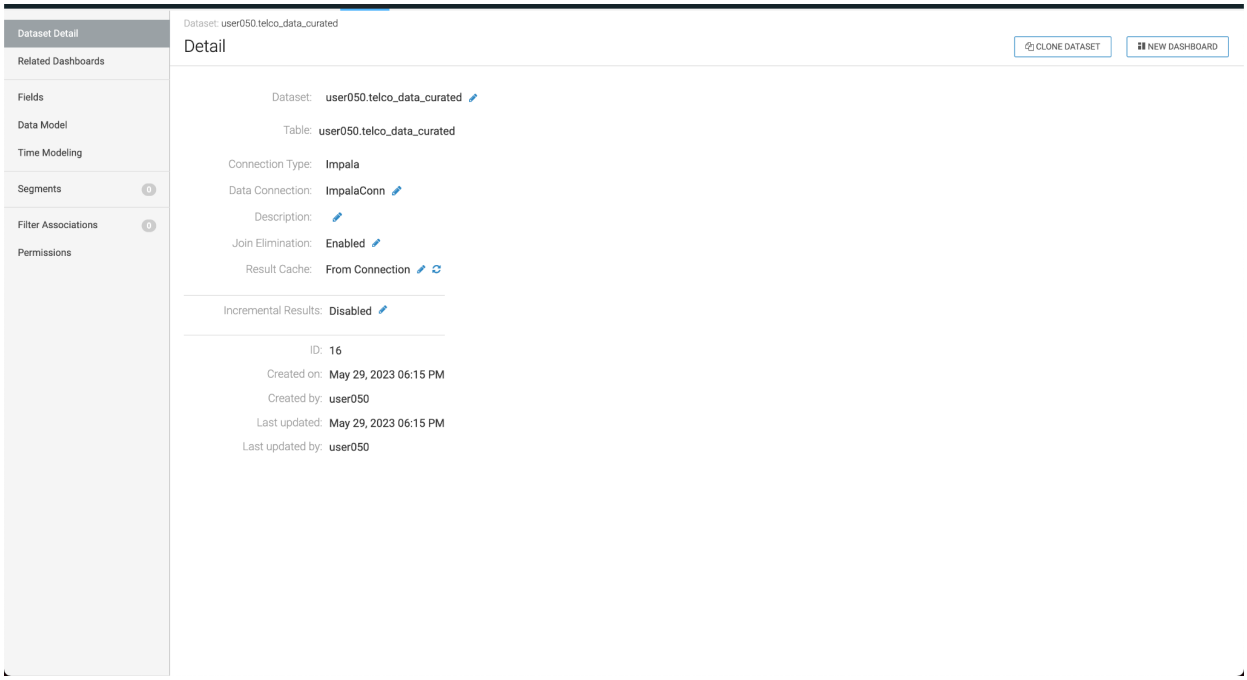
Hacer clic en Create para crear el nuevo Dataset.



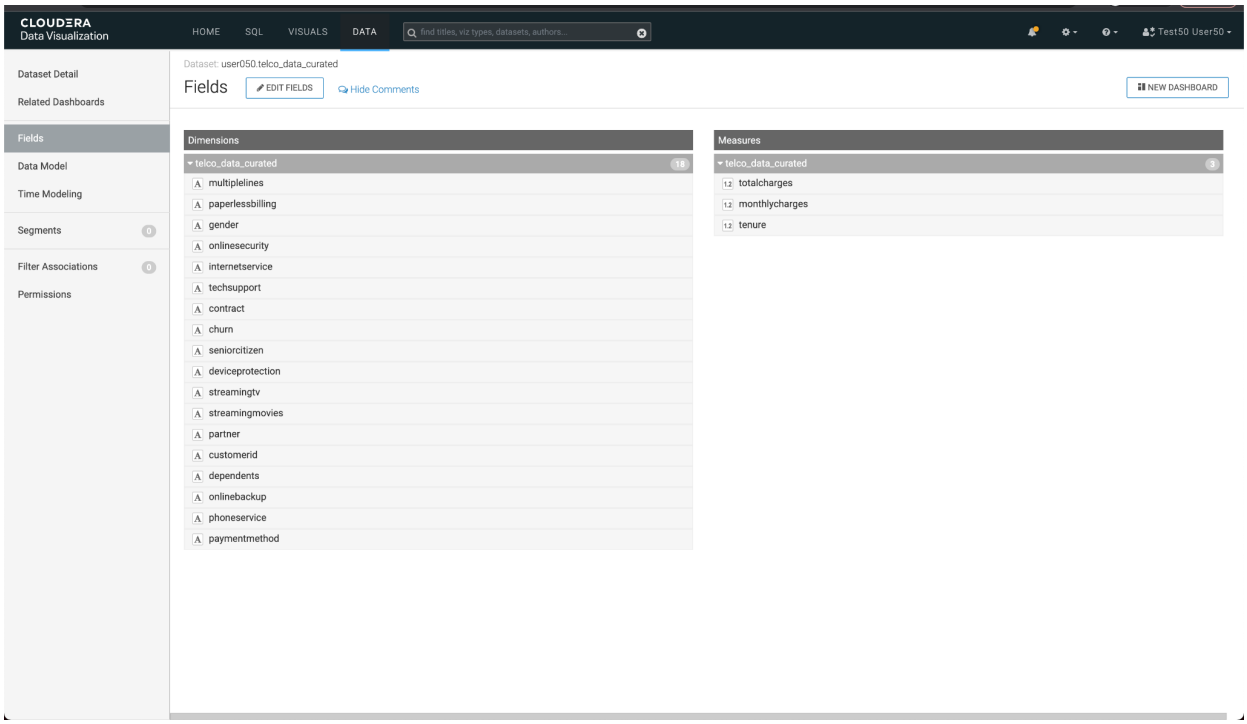
7. El nuevo Dataset deberá aparecer en el listado. Hacer clic sobre el dataset que usted recién creó.



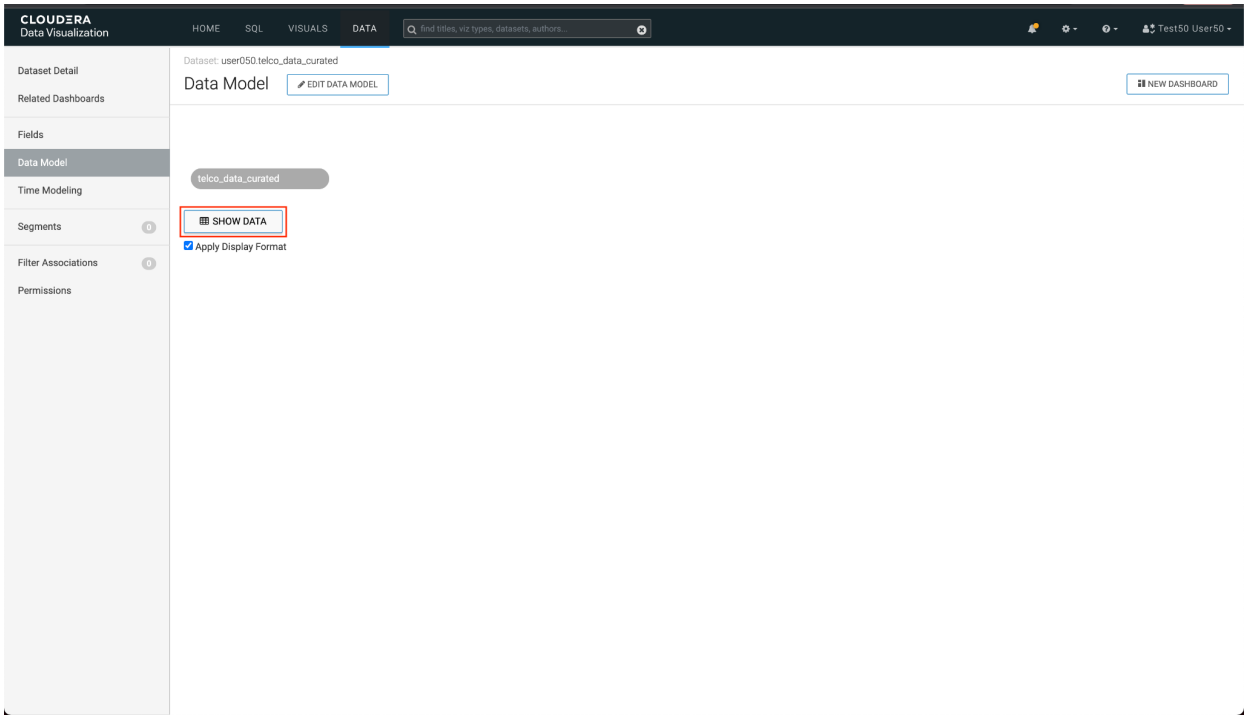
8. Acá verás el detalle del dataset.



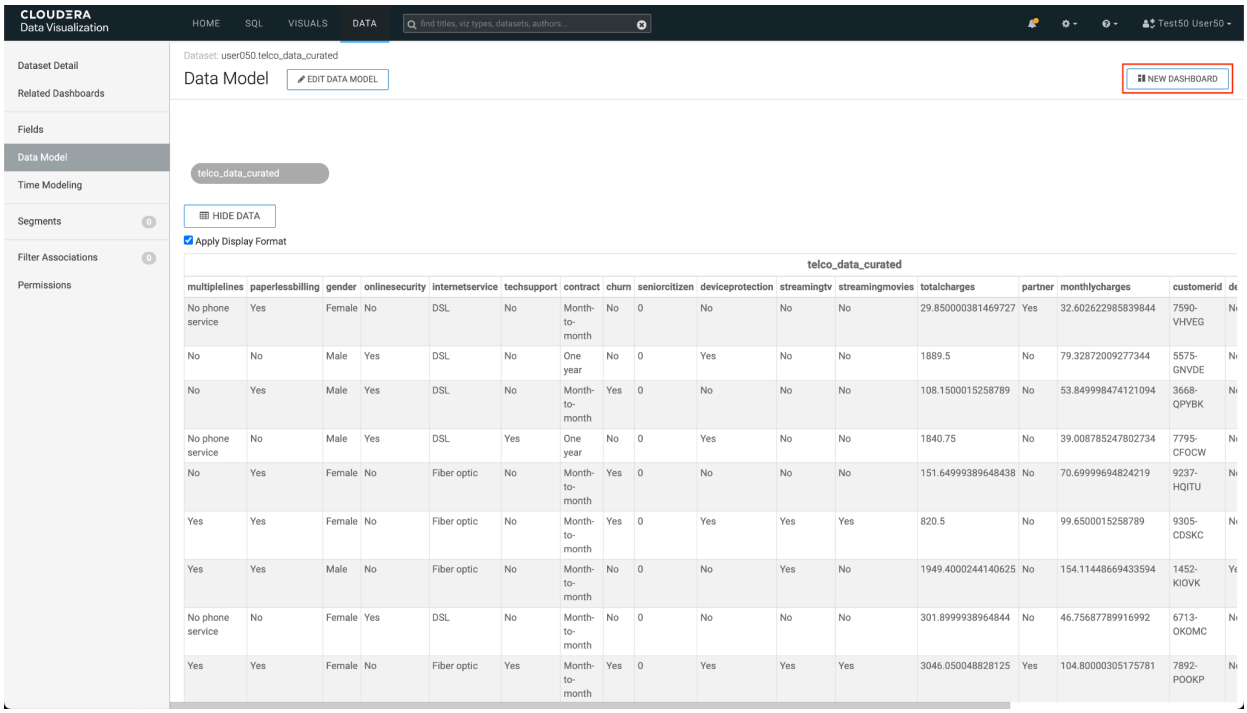
9. Hacer clic en **Fields** (menú izquierdo) para ver los campos automáticamente capturados durante el proceso de creación del dataset.



10. También se puede pre-visualizar los datos desde está pantalla. Hacer clic en **Data Model** (menú izquierdo) y después en el botón **Show Data** que aparece en el centro.



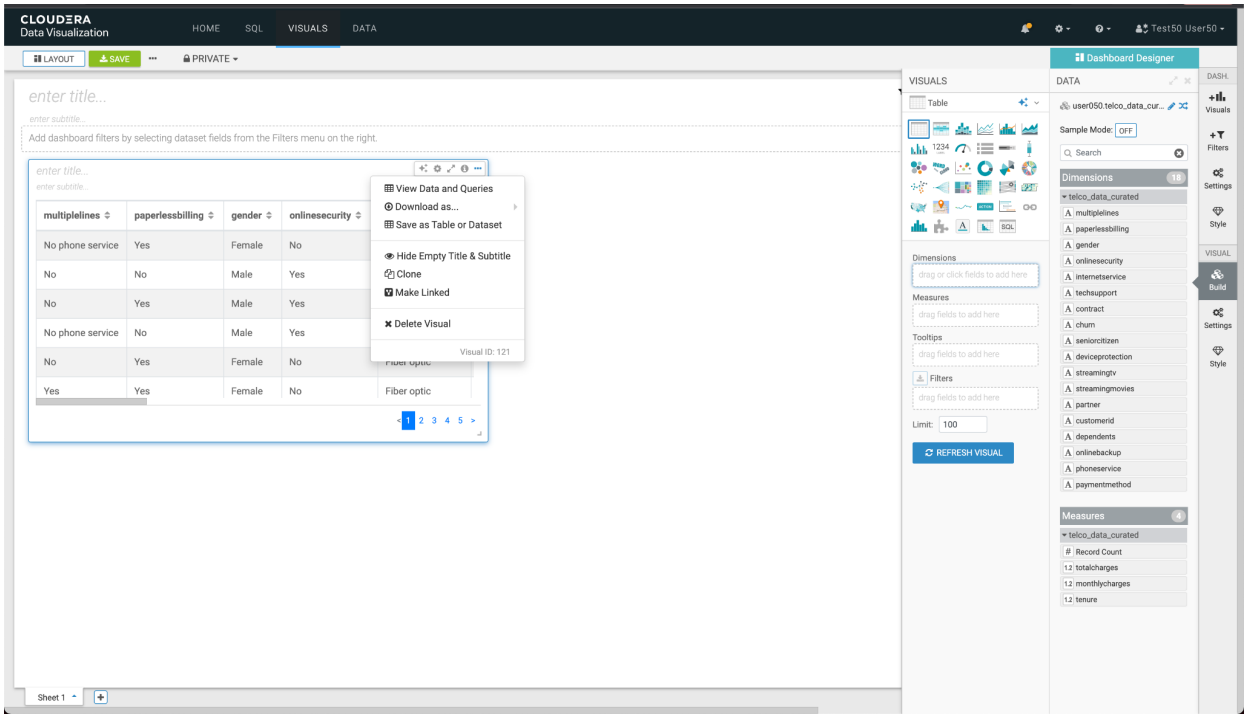
11. En este momento se ejecuta una consulta al Virtual Warehouse para recuperar los datos del data set. Observe las columnas y valores. Hacer clic en New Dashboard para crear un nuevo panel de comando.



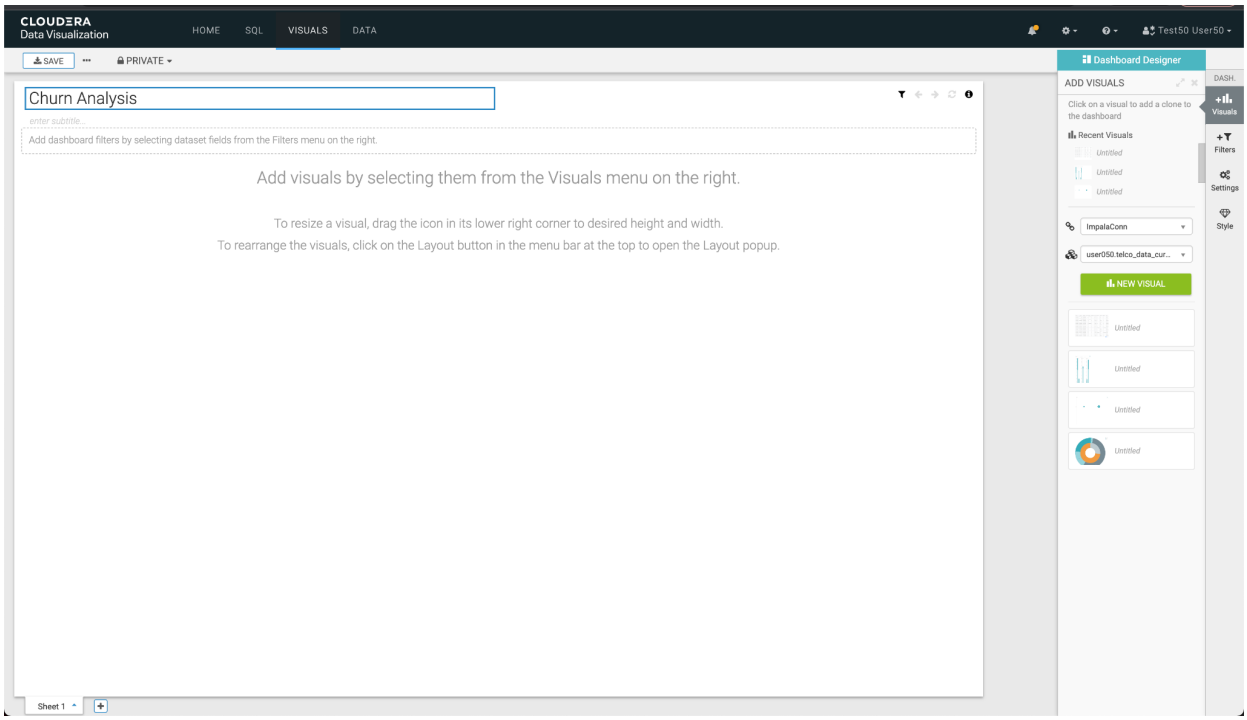
telco_data_curated

multiplelines	paperlessbilling	gender	onlinesecurity	internetservice	techsupport	contract	churn	seniorcitizen	deviceprotection	streamingtv	streamingmovies	totalcharges	partner	monthlycharges	customerid
No phone service	Yes	Female	No	DSL	No	Month-to-month	No	0	No	No	No	29.850000381469727	Yes	32.602622985839844	7590-VHVEG
No	No	Male	Yes	DSL	No	One year	No	0	Yes	No	No	1889.5	No	79.32872009277344	5575-GNVDE
No	Yes	Male	Yes	DSL	No	Month-to-month	Yes	0	No	No	No	108.1500015258789	No	53.849998474121094	3668-QPYBK
No phone service	No	Male	Yes	DSL	Yes	One year	No	0	Yes	No	No	1840.75	No	39.008785247802734	7795-CFOCW
No	Yes	Female	No	Fiber optic	No	Month-to-month	Yes	0	No	No	No	151.64999389648438	No	70.69999694824219	9237-HQITU
Yes	Yes	Female	No	Fiber optic	No	Month-to-month	Yes	0	Yes	Yes	Yes	820.5	No	99.6500015258789	9305-CDSKC
Yes	Yes	Male	No	Fiber optic	No	Month-to-month	No	0	No	Yes	No	1949.4000244140625	No	154.11448669433594	1452-KIOVK
No phone service	No	Female	Yes	DSL	No	Month-to-month	No	0	No	No	No	301.8999938964844	No	46.75687789916992	6713-OKOMC
Yes	Yes	Female	No	Fiber optic	Yes	Month-to-month	Yes	0	Yes	Yes	Yes	3046.050048828125	Yes	104.80000305175781	7892-POOKP

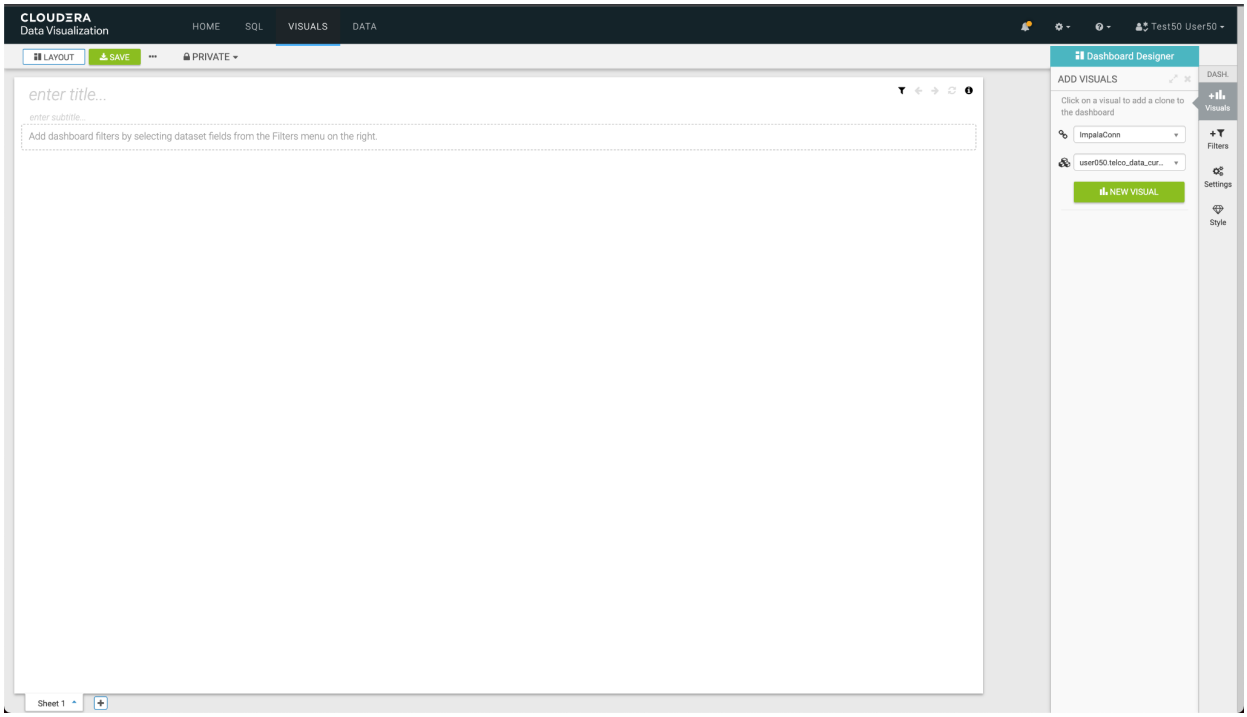
12. Al abrir el canvas de diseño de un nuevo panel, eliminar el elemento que se agrega por defefcto, haciendo clic en en el botón de tres puntos (...) que hay en la parte superior derecha del elemento, y después clic en la opción **Delete Visual**



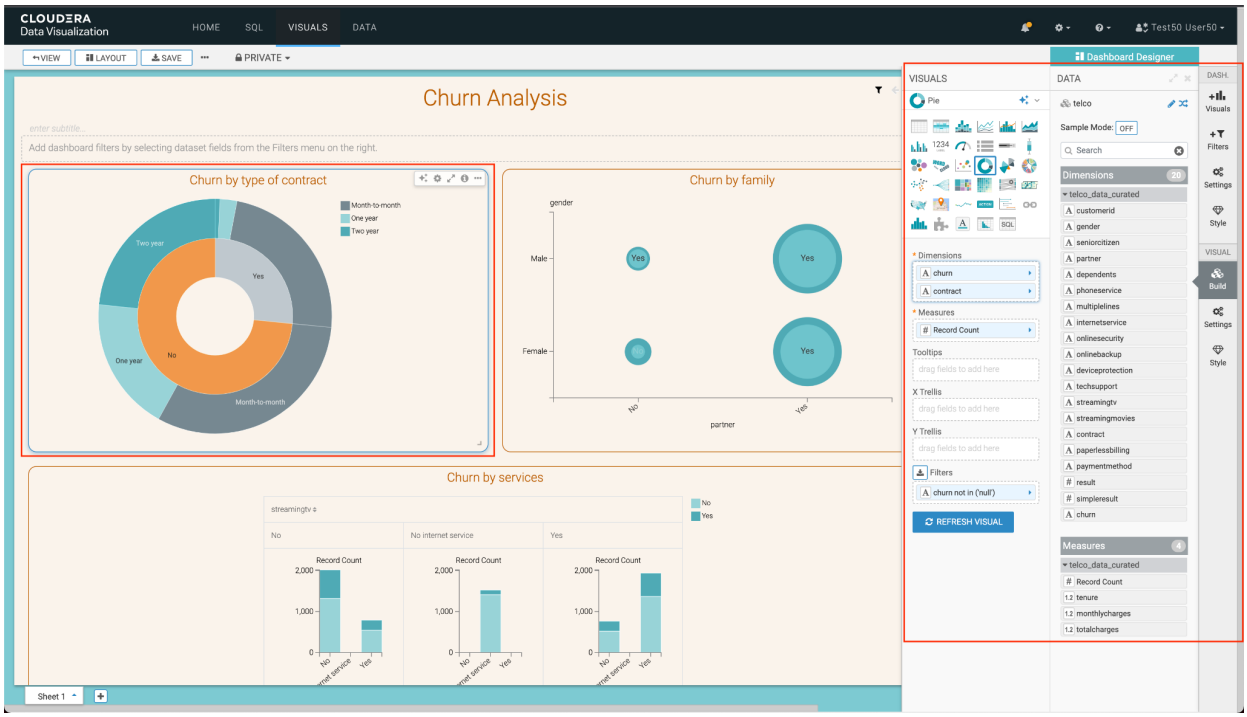
En la parte superior del canvas, en el campo enter title, ingresar el nombre *Churn Analysis* para el identificar el dashboard.



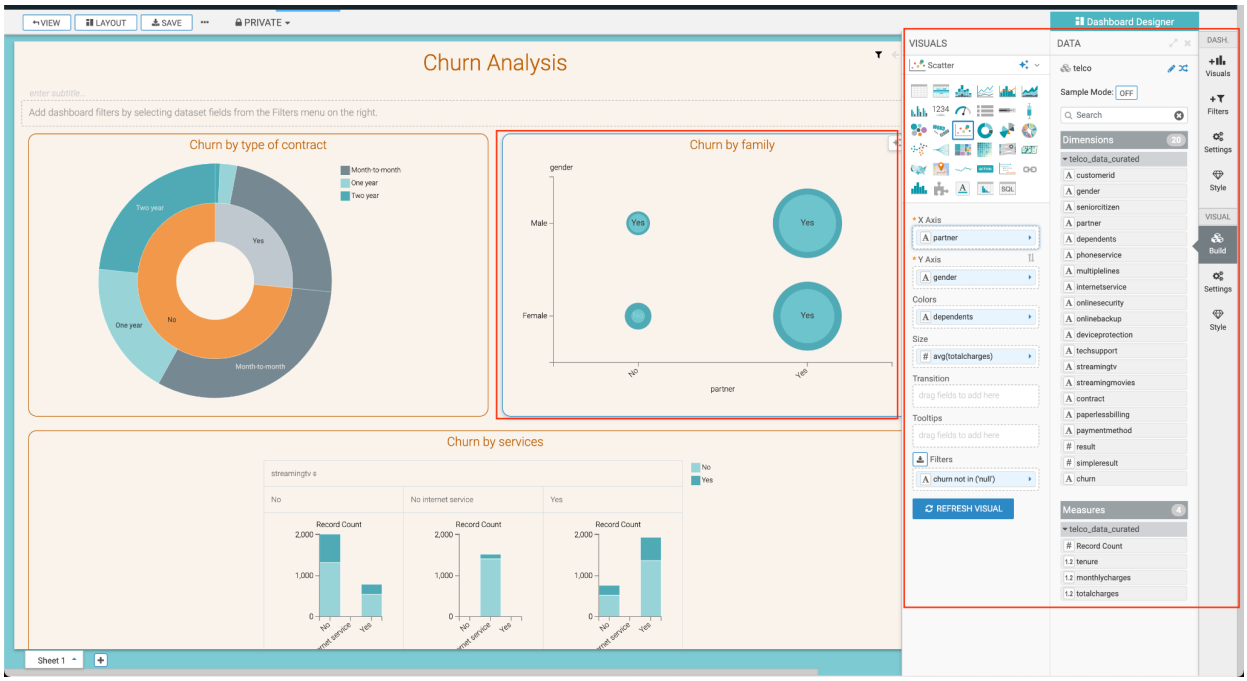
13. Para agregar un nuevo elemento visual, hacer clic en el botón **Visuals** del menú derecho, seleccionar el dataset que les corresponda, y hacer clic en el botón **New Visual.**



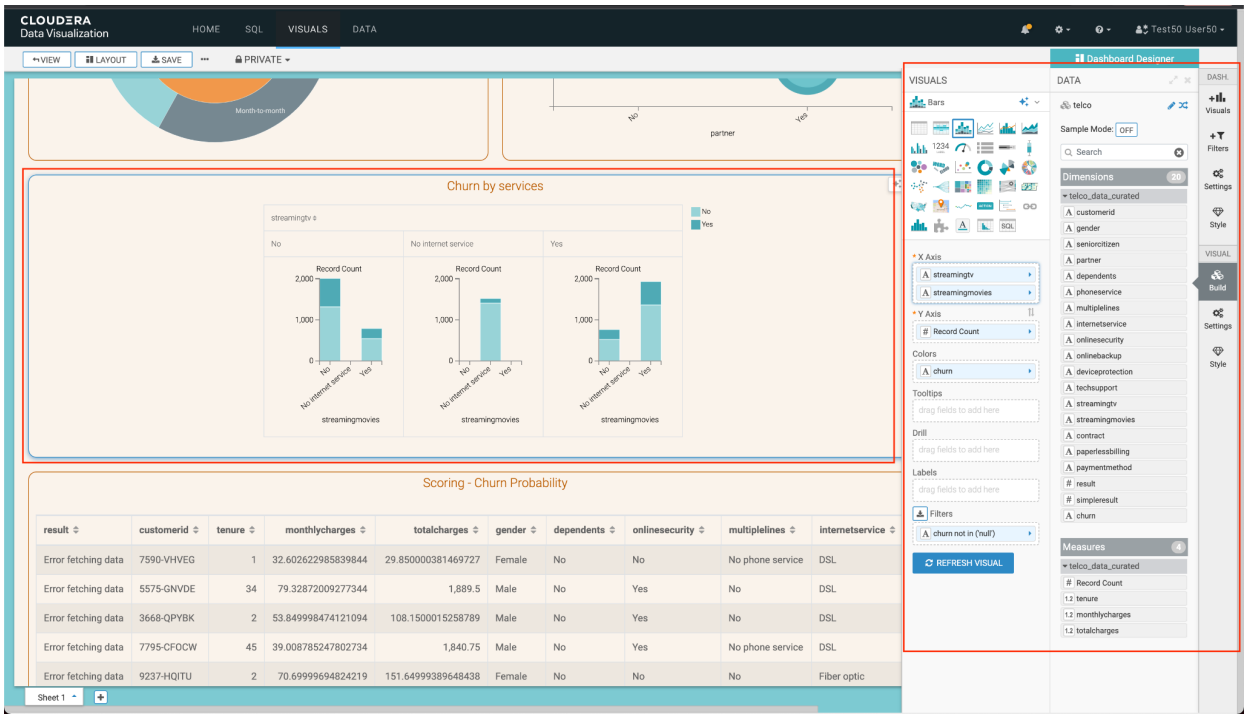
14. Agregar el primer elemento visual, que es un gráfico de torta con las dimensiones **churn** y **contract**, con la métrica de **Record count**. Una vez terminado, hacer clic en el botón **Refresh Visual**.



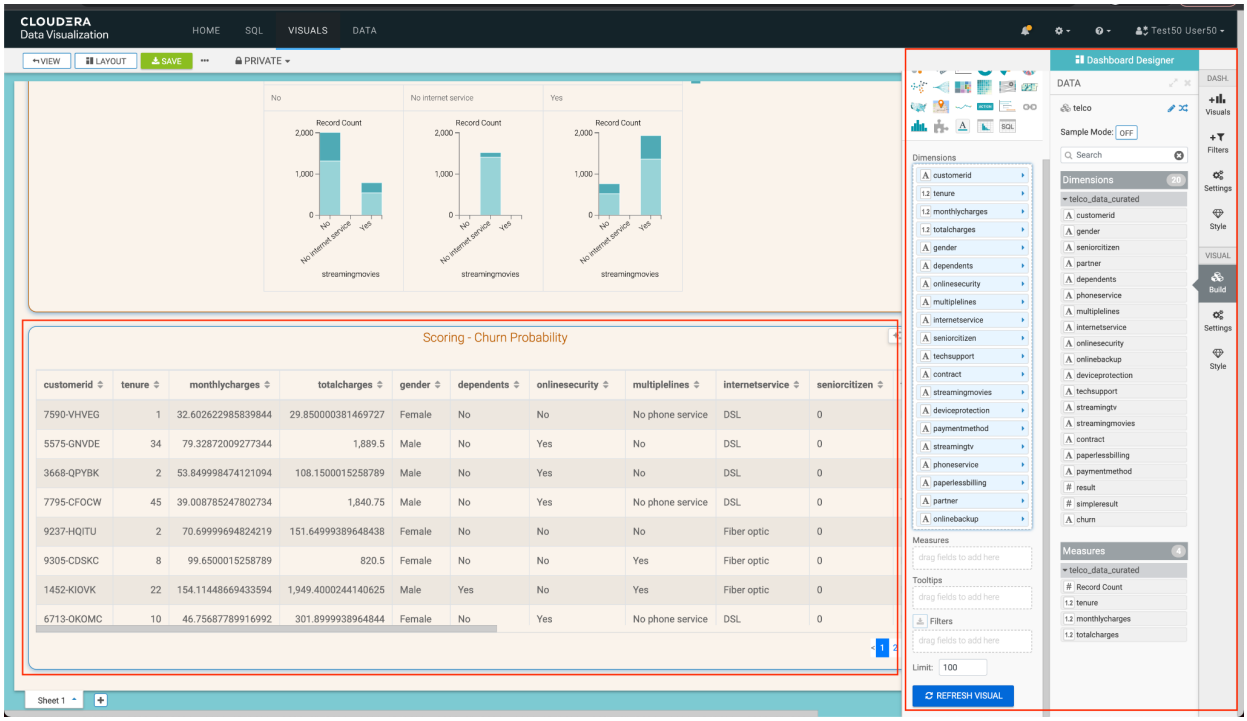
15. Agregar el segundo elemento visual, que es un gráfico tipo Scatter con la dimensión **partner** como X Axis, **gender** cómo Y Axis, **dependents** como Colors y **avg(totalcharges)** como Size. Una vez terminado, hacer clic en el botón **Refresh Visual**.



15. Agregar el tercer elemento visual, que es un gráfico de barras con la dimensiones **streamingtv** y **streamingmovies** como X Axis, **Record Count** cómo Y Axis y **churn** cómo Colors. Una vez terminado, hacer clic en el botón **Refresh Visual**.



16. Agregar el cuatro y último elemento visual, que es una table con las dimensiones y métricas del dataset. Asegurarse de agregar las 17 dimensiones y 3 métricas a la tabla. Una vez terminado, hacer clic en el botón **Refresh Visual**.



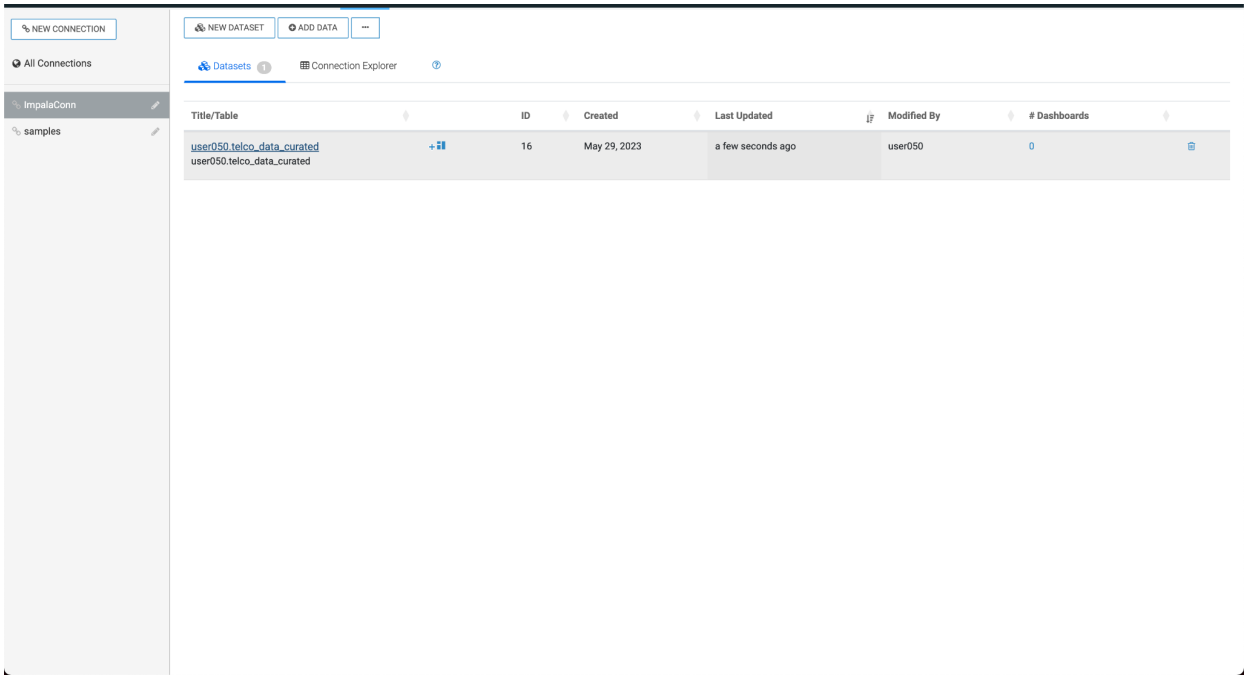
Guarda el dashboard haciendo clic en el botón **Save** del menú superior.

Parte 2: Agregar nuevo campo

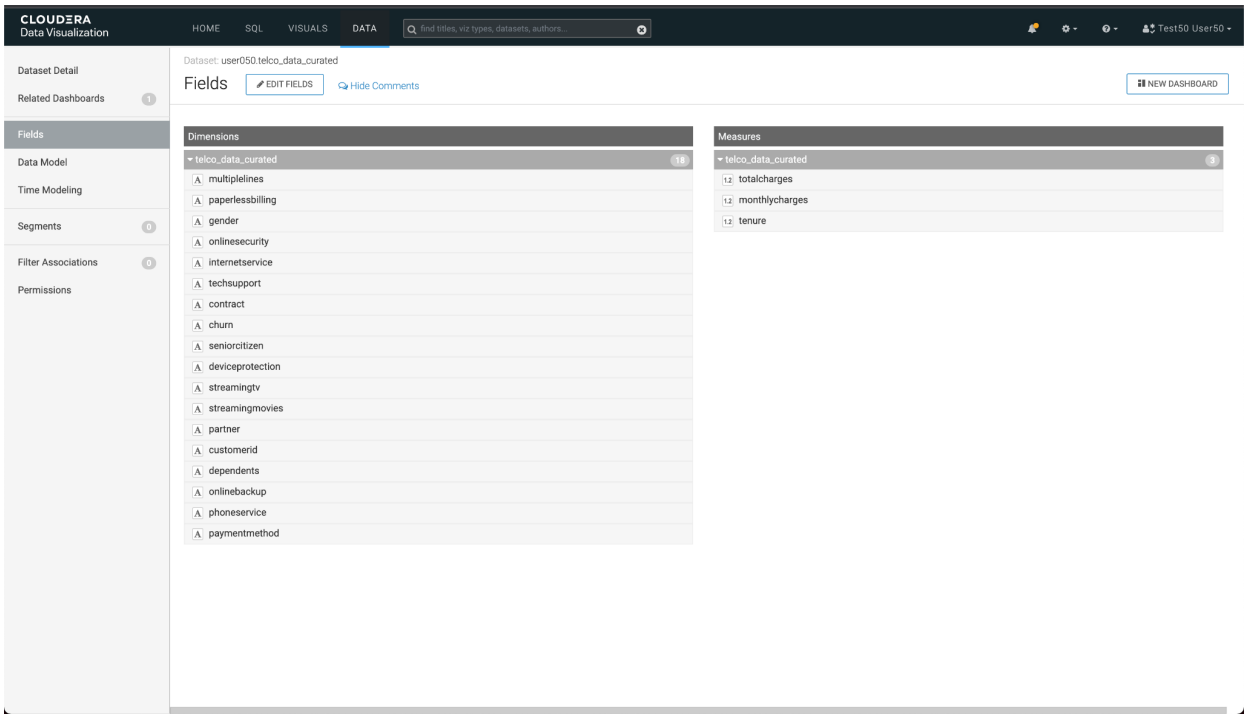
Objetivos:

- Agregar un nuevo campo que hace llamadas al modelo de ML
- Agregar el nuevo campo al dashboard

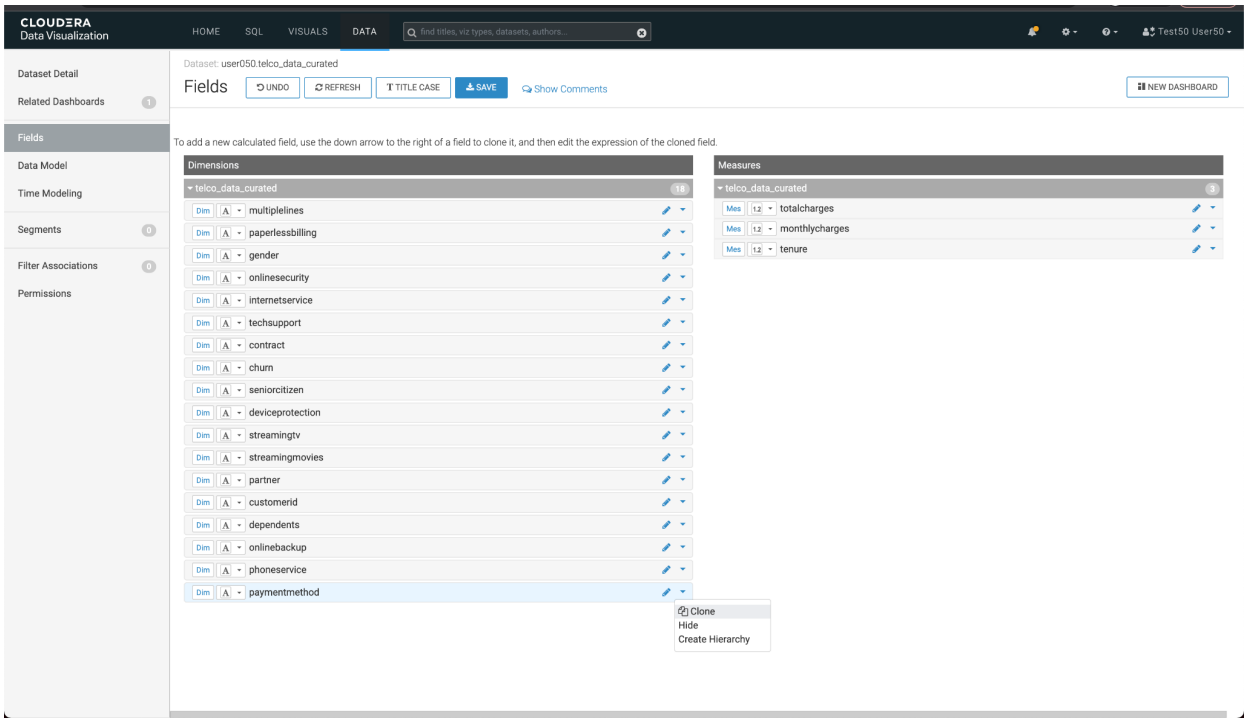
1. Editar el Dataset creado anteriormente, en Data -> <usuario_asignado>.telco_data_curated.



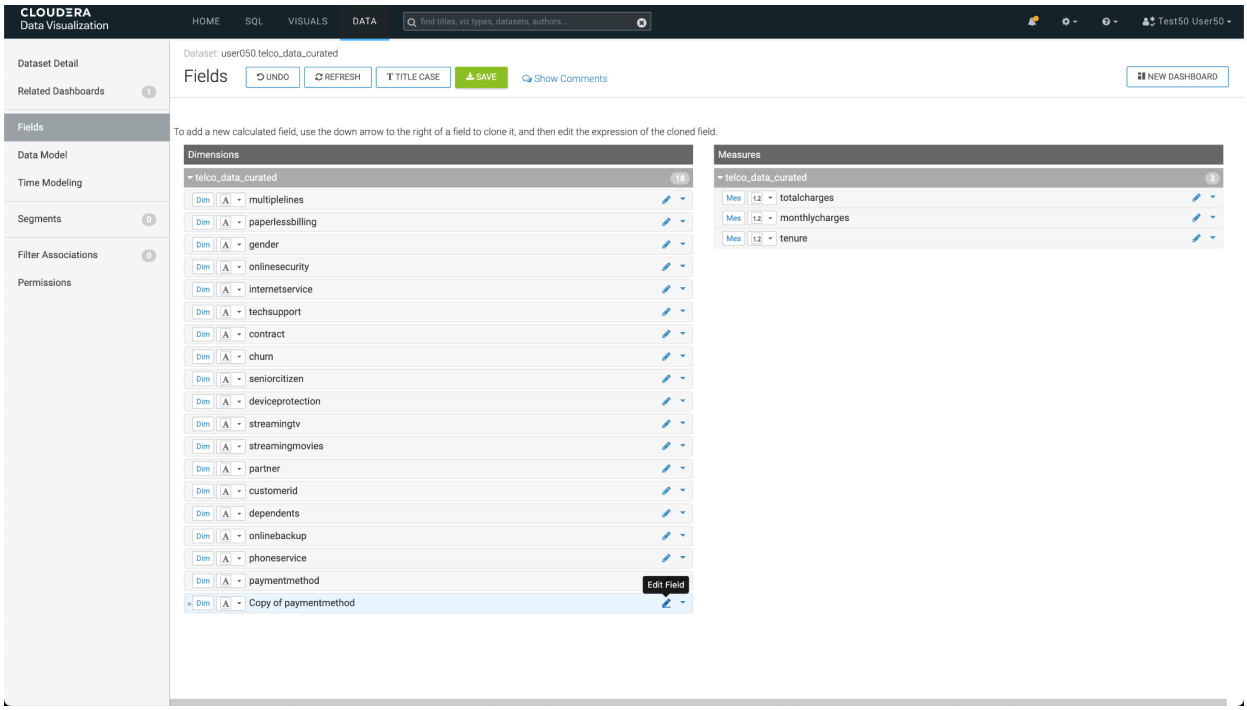
2. Una vez en el Dataset, ir a **Fields** en el menú izquierdo y posteriormente hacer clic en **Edit Field** para editar los campos de su dataset.



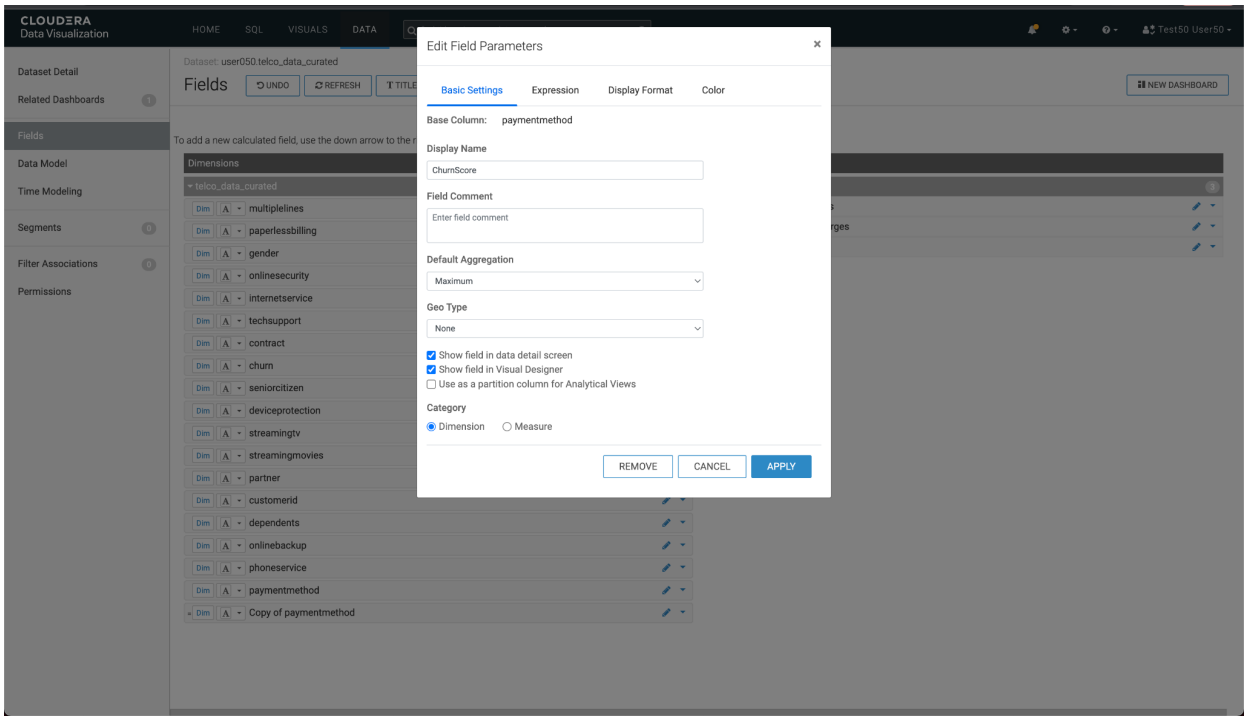
3. En la lista de **Dimensions**, hacer clic en la flecha hacia abajo del último campo de la lista, y seleccionar la opción **Clone**.



4. Una vez clonado el campo, hacer clic en el lápiz al lado del campo para editarlo.

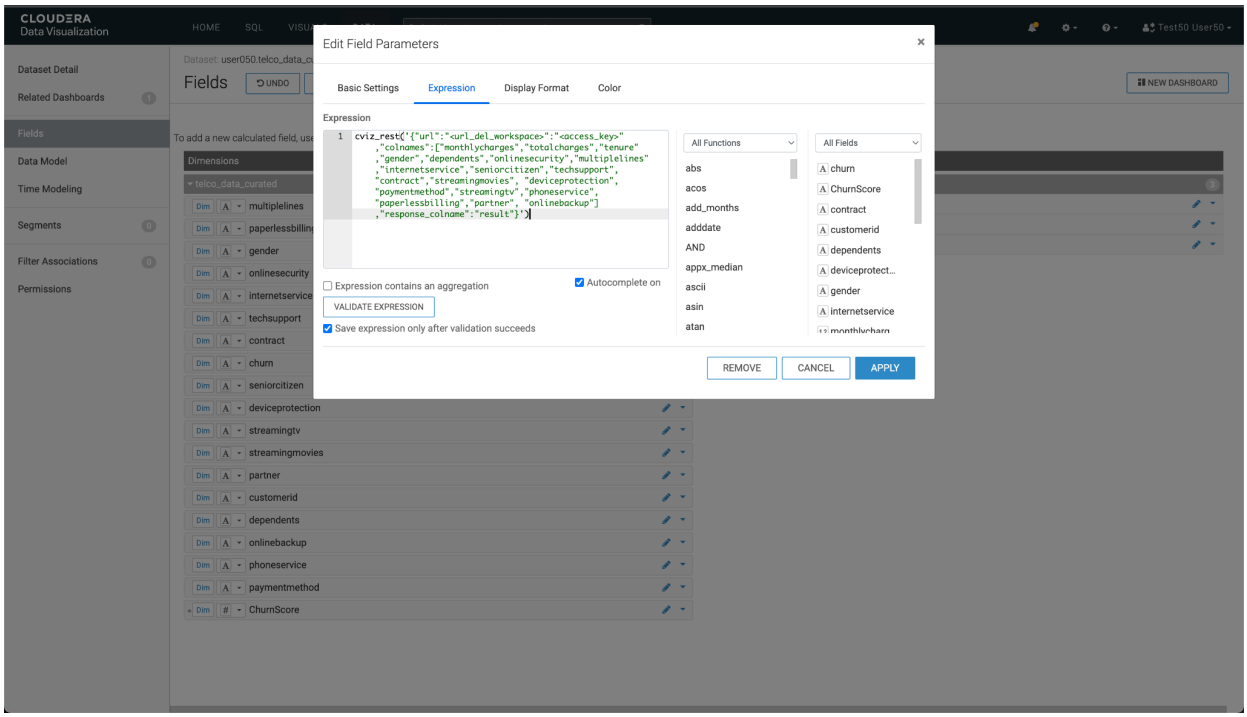


5. En la ventana emergente que aparece, introducir el nombre del nuevo campo en **Display Name**. Sugerimos que introduzcas *ChurnScore.*

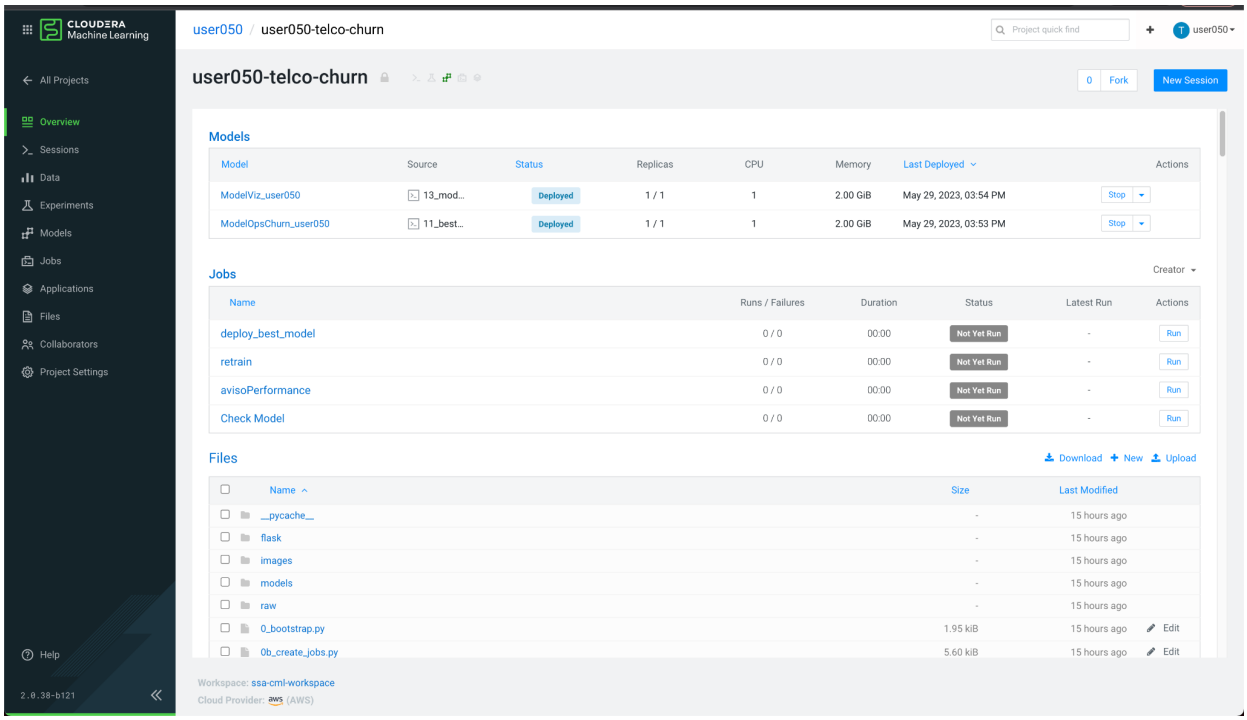


6. Ir a la pestaña Expressions e introducir el siguiente valor en el campo Expression. Esto permitirá hacer llamada a la API REST del Modelo que anteriormente ha desplegado.

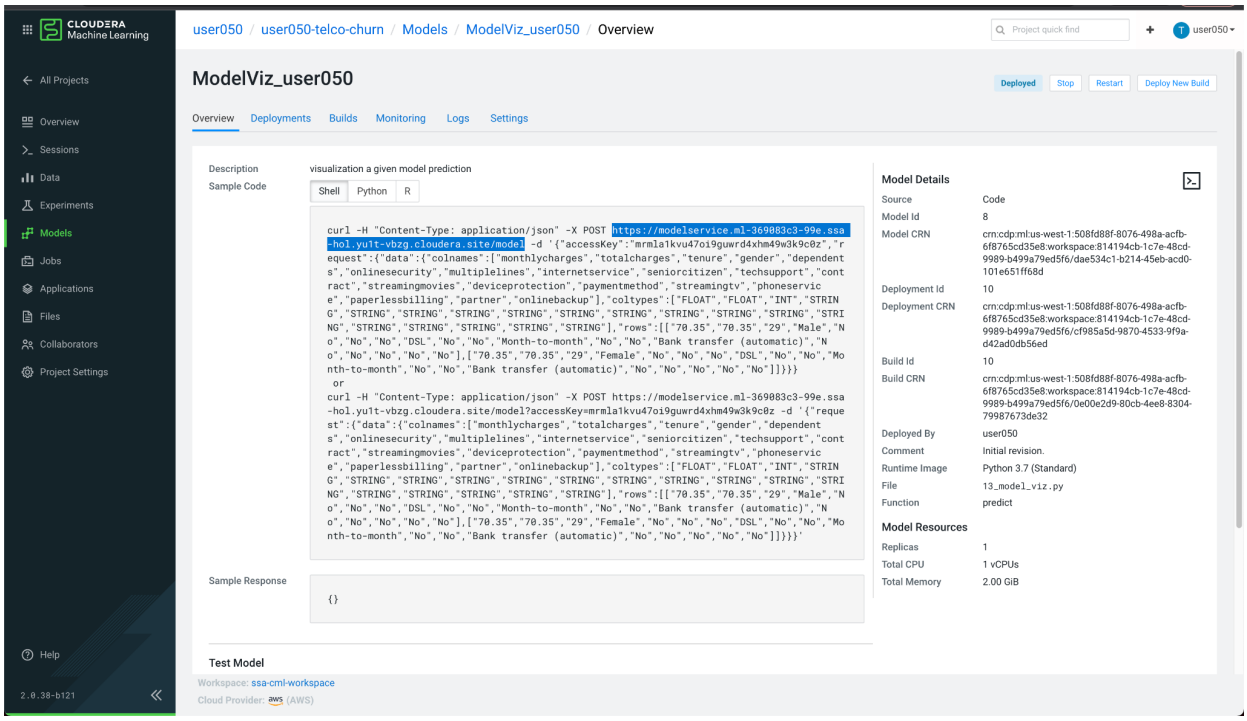
cviz_rest('{"url":"<url_del_workspace>":"<access_key>","colnames":["monthlycharges","totalcharges","tenure","gender","dependents","onlinesecurity","multiplelines","internetservice","seniorcitizen","techsupport", "contract","streamingmovies", "deviceprotection", "paymentmethod","streamingtv","phoneservice", "paperlessbilling","partner", "onlinebackup"],"response_colname":"result"}')



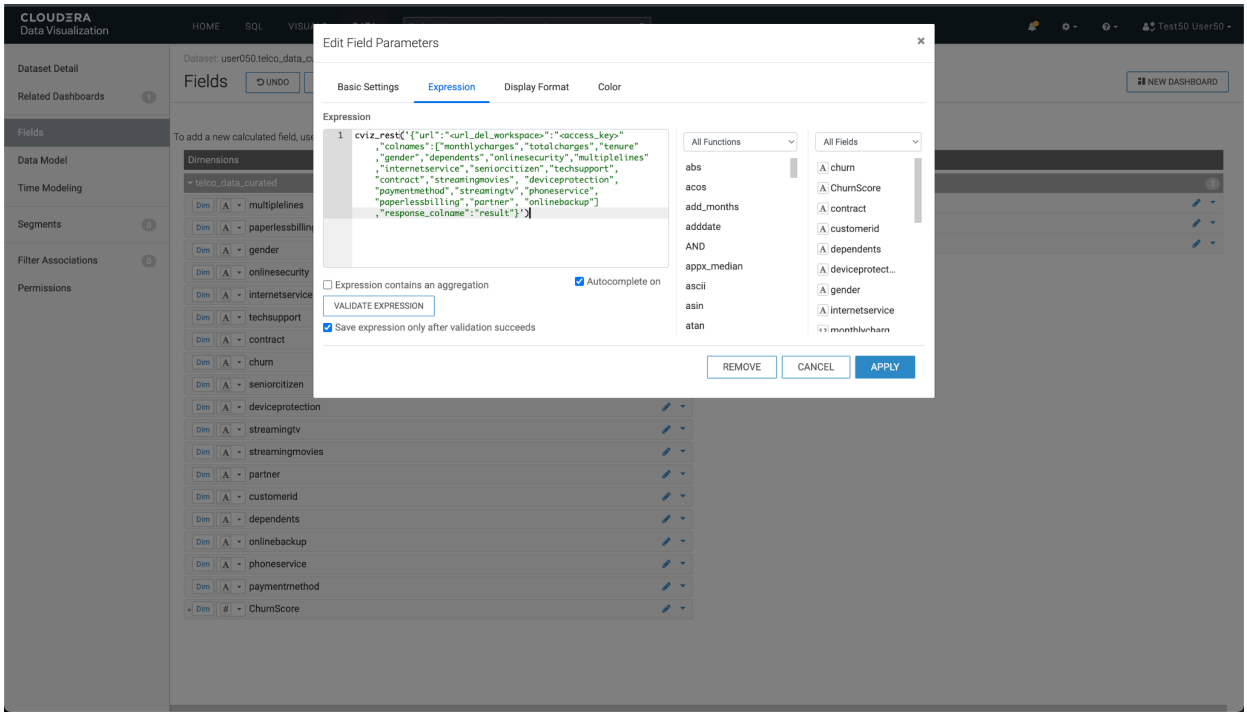
7. Estando en CML en otra pestaña del navegador web, ir a la sección de **Models** de su proyecto, y hacer clic sobre el Model que empieza con el nombre *ModelViz*, seguido de su nombre de usuario asignado.



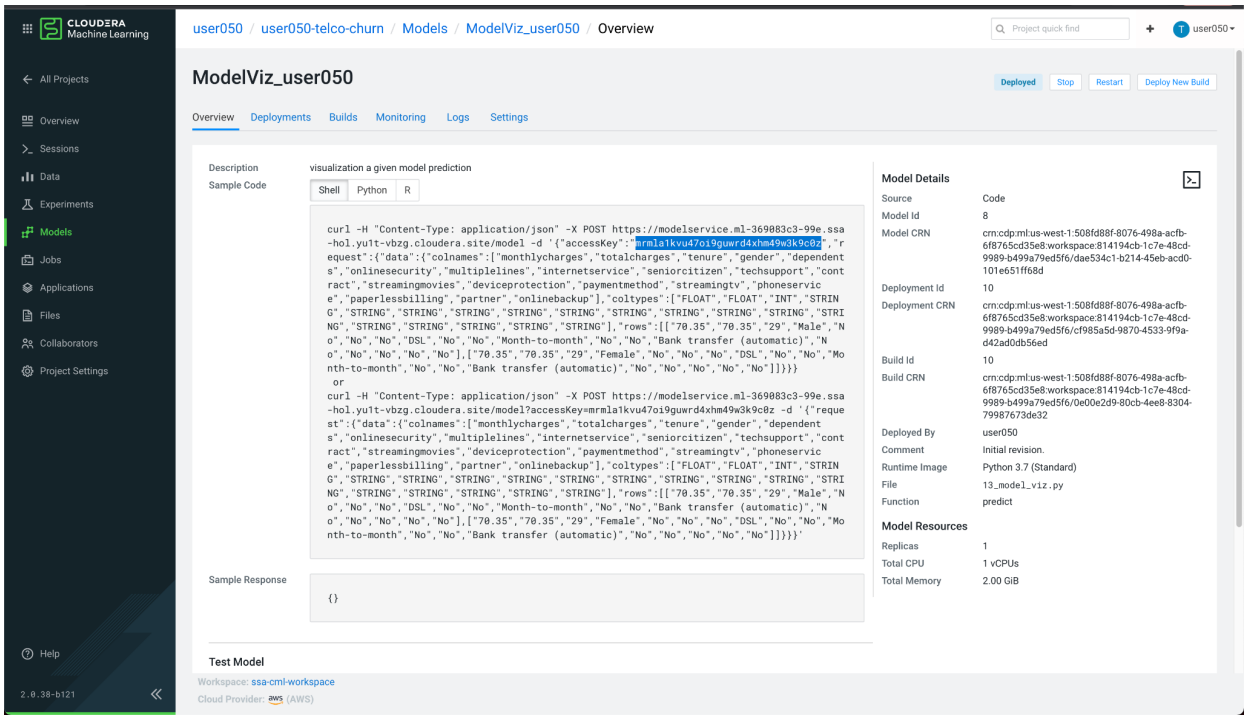
8. En la pestaña de Overview, copiar la URL que permite interactuar y hacer llamada a la API del workspace.



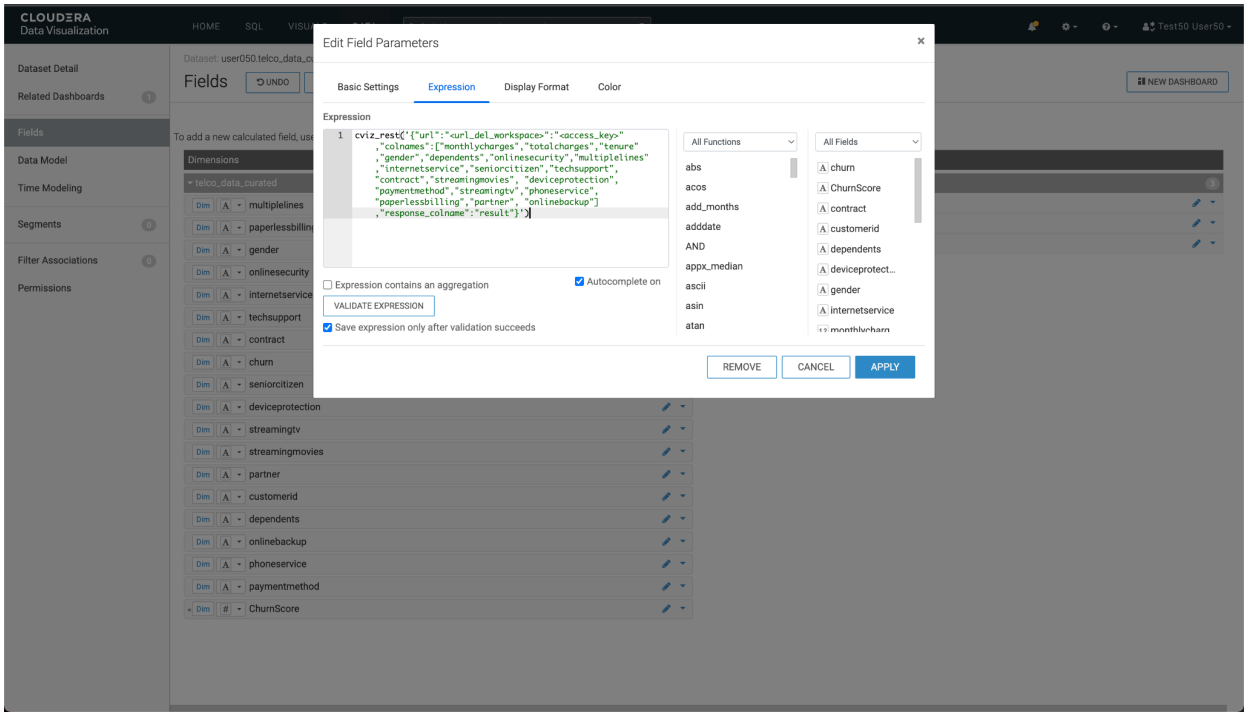
Reemplazar el valor copiado en el atributo **<url_del_workspace>** del campo Expression.



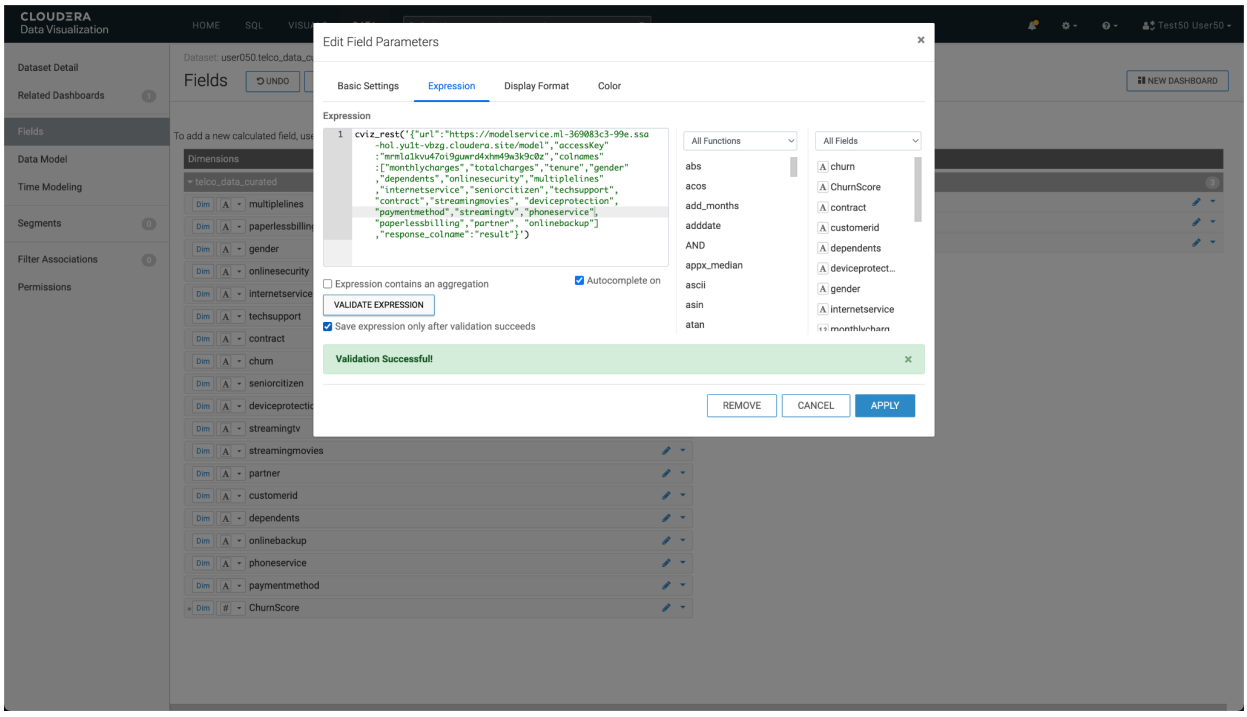
9. Regresando a CML, copiar la accessKey del modelo.



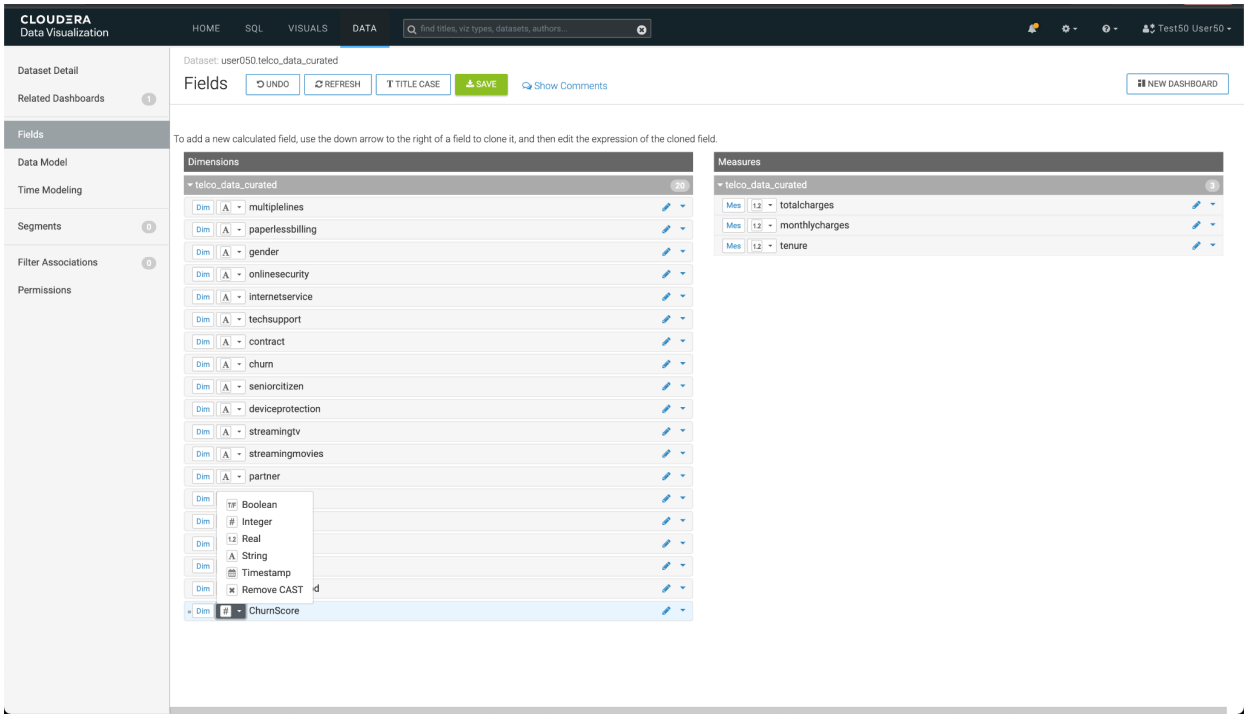
Reemplazar el valor copiado en el atributo **<access_key>** del campo Expression.



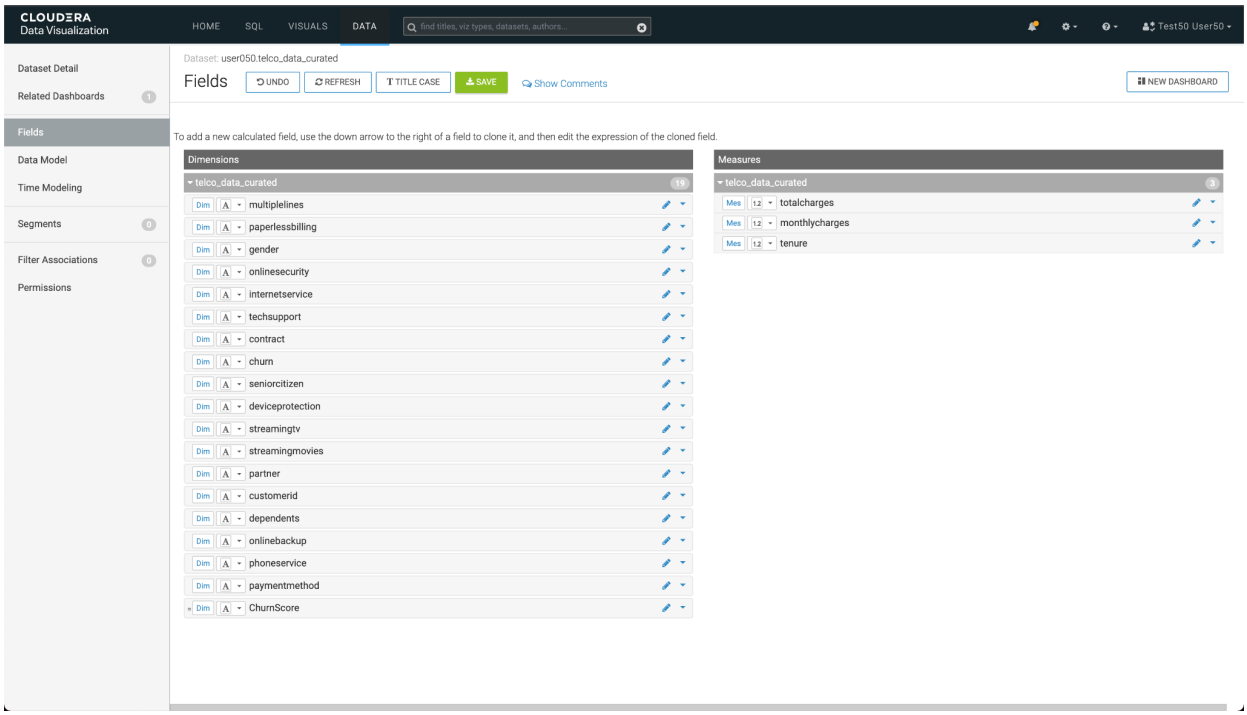
10. Finalizado el proceso de copiar la *url del workspace* y la *accessKey*, hacer clic en el botón Validate Expression, que se encuentra en la parte superior de la ventana. Si aparece el mensaje en verde *Validation Successful*, hacer clic en **Apply** para guardar la configuración realizada.



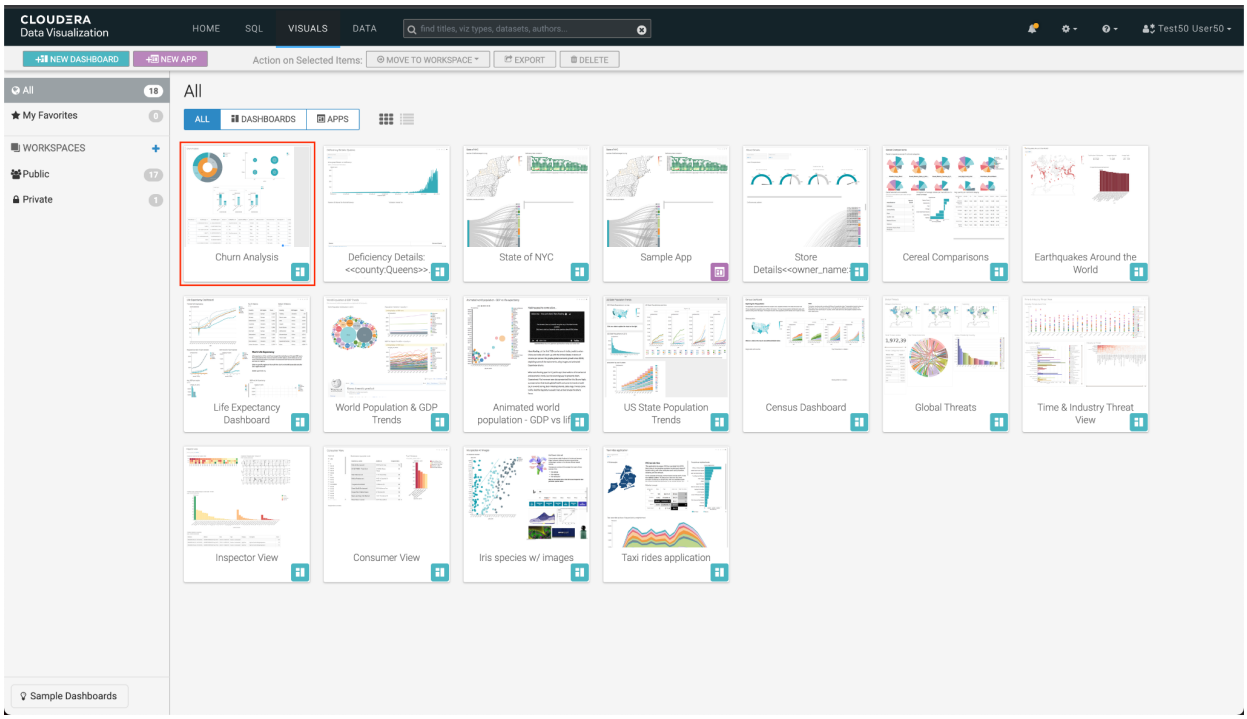
11. El nuevo campo deberá aparecer en el listado de campos. Cambiar el tipo de dato, seleccionando el tipo *Integer*, que se representa por el símbolo **#**



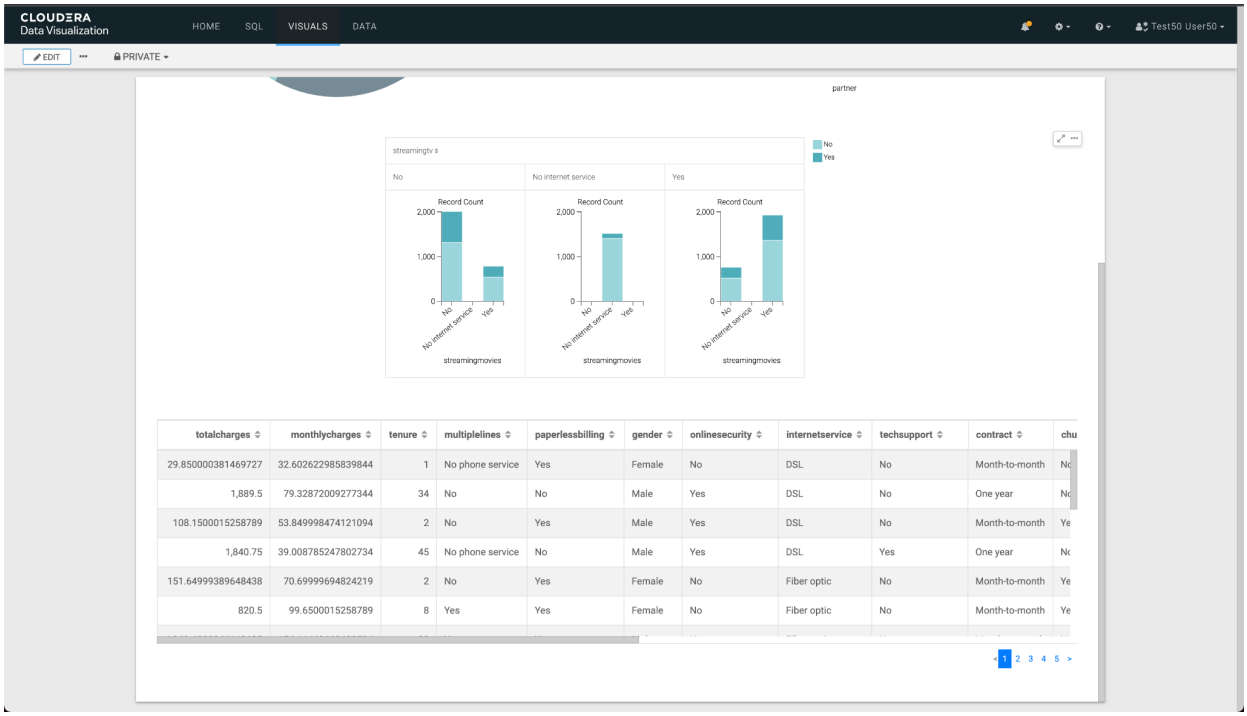
12. Finalizar el proceso haciendo clic en el botón verde con leyenda **SAVE** en el menú superior.



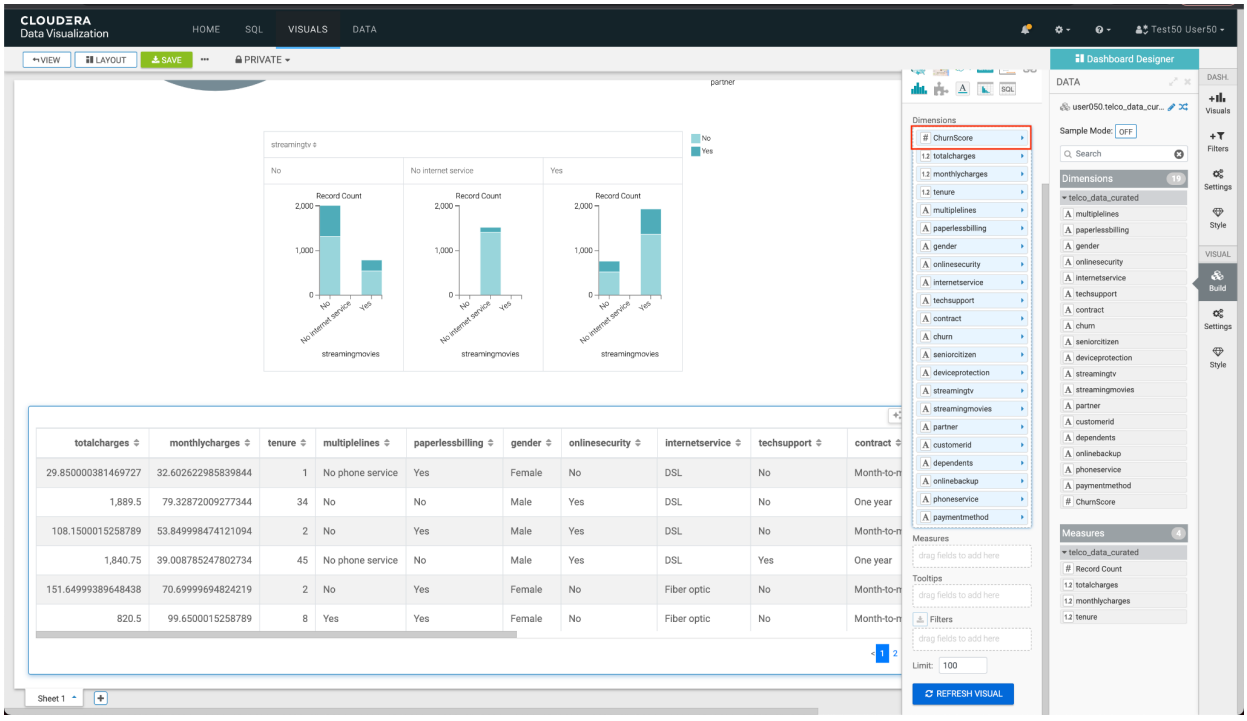
13. Regresar al dashboard, seleccionando la opción **VISUALS** del menú superior, y haciendo clic sobre el nombre del dashboard que se creó anteriormente.



14. Una vez en el dashboard, hacer clic en el botón **Edit** que se encuentra en la parte superior izquierda.



15. Editar la tabla inferior, haciendo clic sobre la misma y posteriormente sobre la opción **Build** del menú vertical derecho. Agregar el nuevo campo, **ChurnScore**, al inicio de la tabla, haciendo clic y arrastrando desde la opción de **Dimensions** disponible.



16. Hacer clic en el botón Refresh Visual para actualizar los datos. Deberá aparecer la nueva columna *ChurnScore* luego al inicio de la tabla, con un valor de tipo numérico. Finalizar el proceso haciendo clic en botón **SAVE** del menú superior izquierdo.