LAB

05 – Cloudera Data Warehouse (Part 2)

Data Lifecycle on CDP Public Cloud

Data Warehouse Lab

This is the 2nd part of the 'Data Warehouse Lab' is divided into 2 Parts.

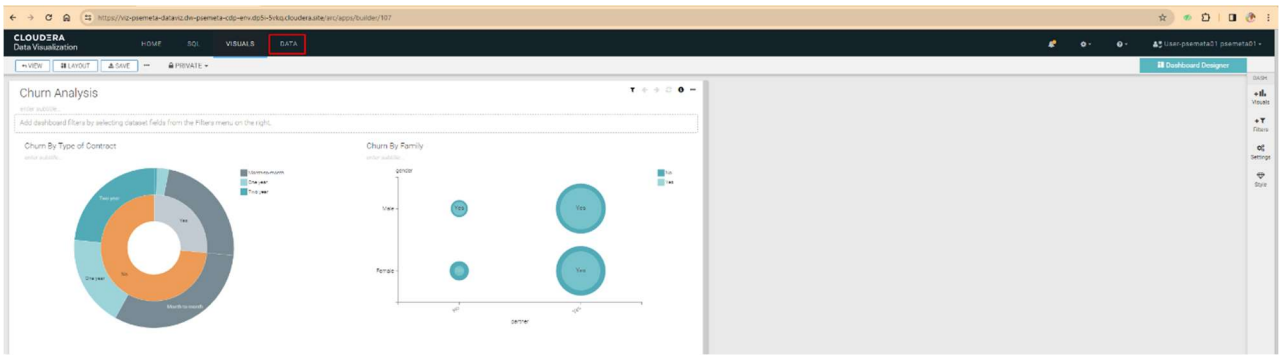
Let's work on the 2nd Part now.

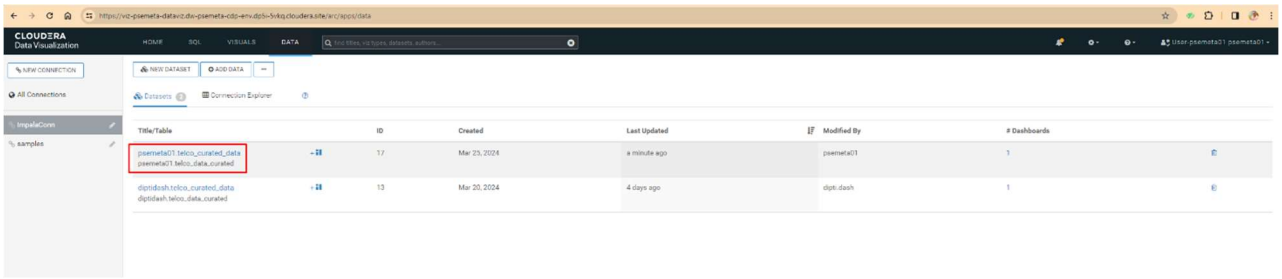
Part 2: Integrating Data Viz & CML

Goals:

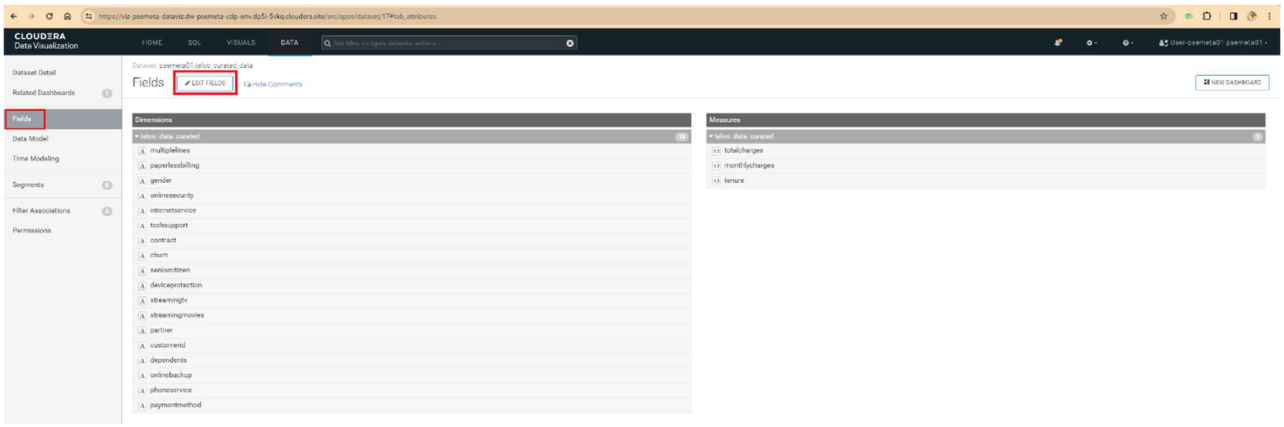
- Add a new field that makes calls to the ML model
- Add the new field to the dashboard

1. Edit the previously created Dataset, in Data -> <user_assigned>.telco_data_curated.

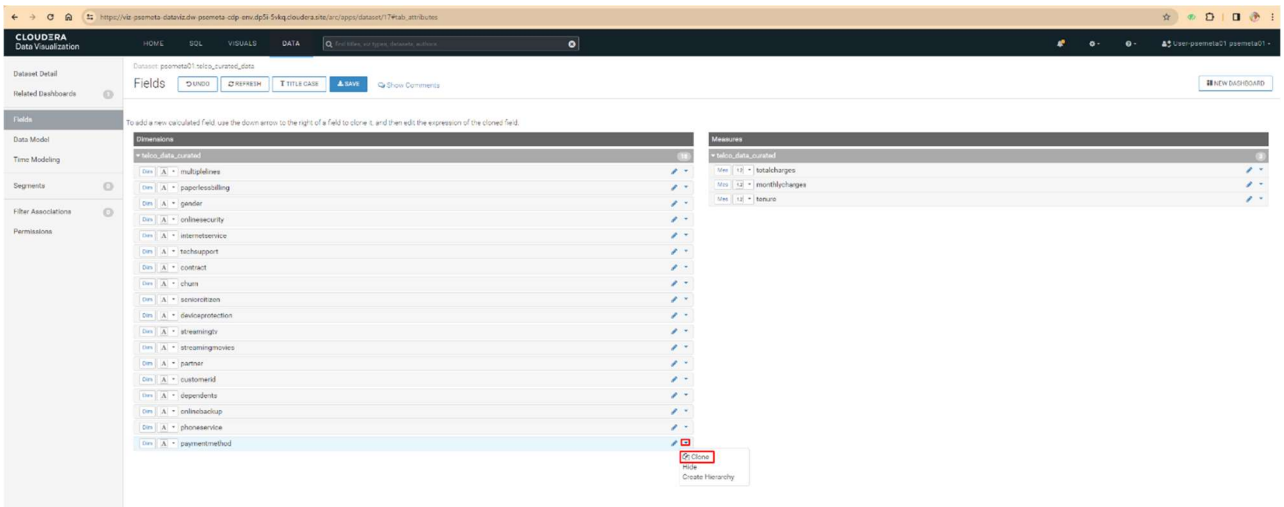




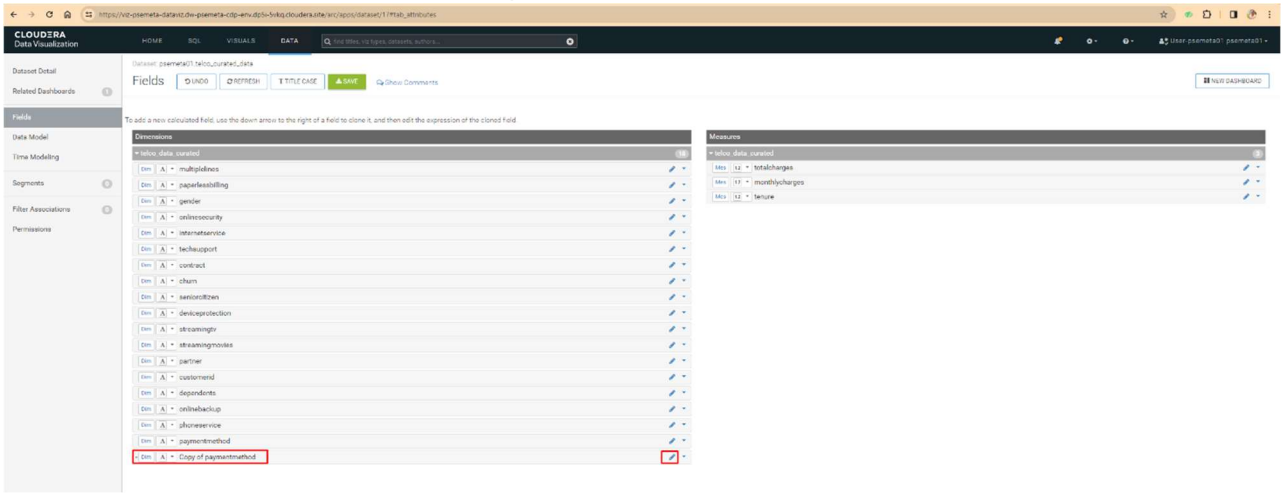
2. Once in the Dataset, go to **Fields** in the left menu and then click on **Edit Field** to edit the fields of your dataset.



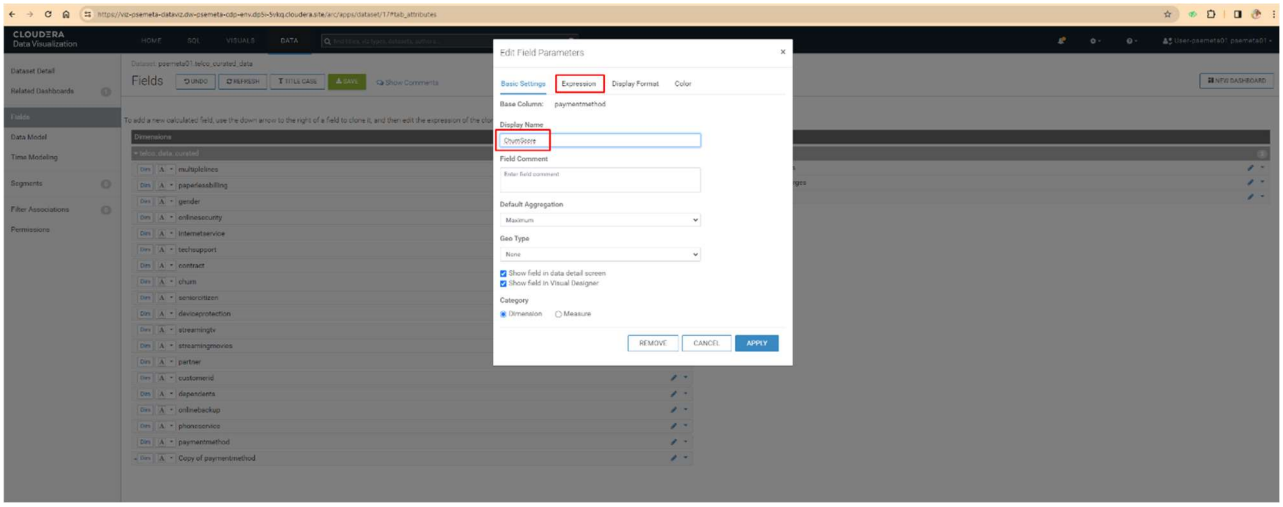
3. In the list of **Dimensions**, click the down arrow of the last field in the list, and select the option **Clone**.



4. Once the field is cloned, click on the pencil next to the field to edit it.



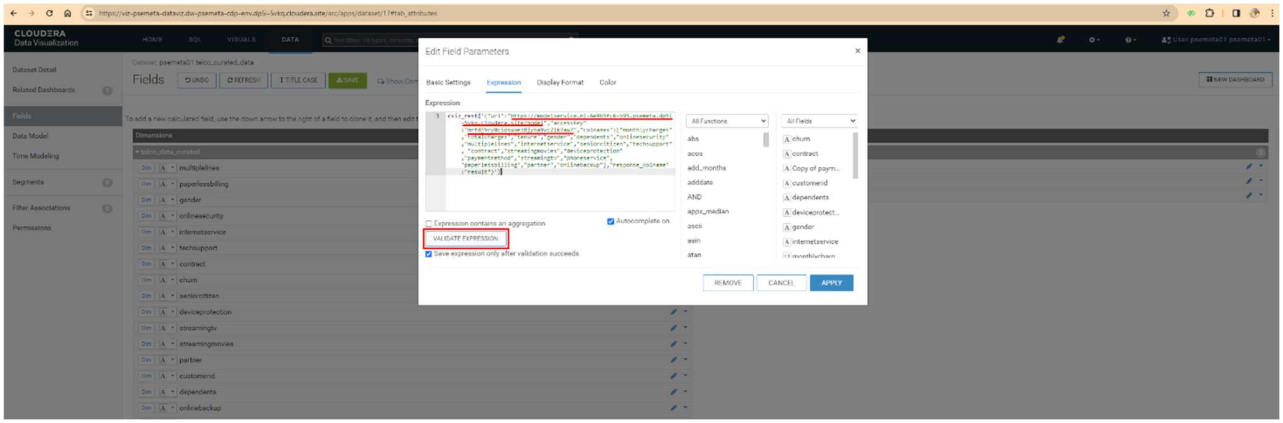
5. In the popup window that appears, enter the name of the new field in **Display Name**. We suggest that you enter *ChurnScore*.



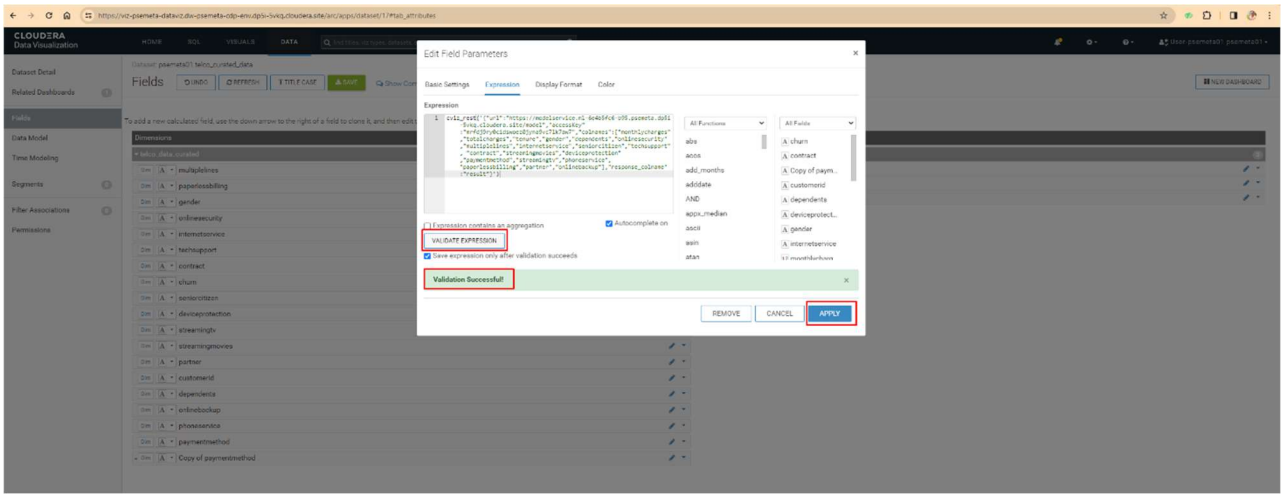
6. Go to the Expressions tab and enter the following value in the Expression field. This will allow you to call the REST API of the Model you have previously deployed.

Note: The url and accessKey is something that you would have noted down as the last step in the CML lab.

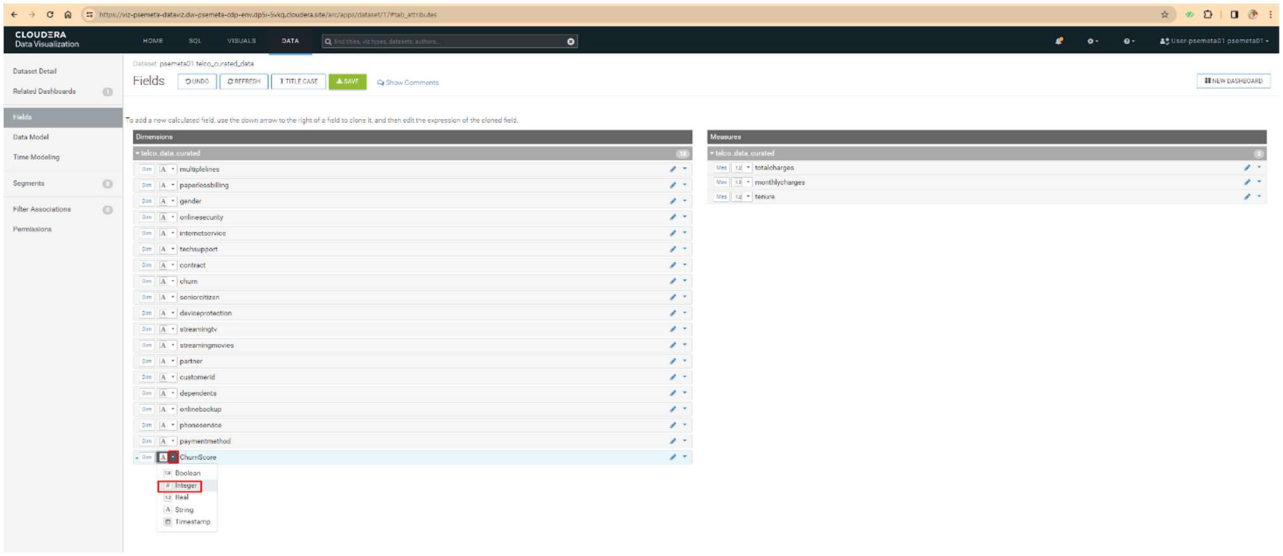
cviz_rest('{"url":"**https://modelservice.ml-6e4b5fc6-b95.psemeta.dp5i-5vkq.cloudera.site/model**","accessKey":"**miqnv6bxiwontz84ozoxfyc5rt6p6grb**","colnames":["monthlycharges","totalcharges","tenure","gender","dependents","onlinesecurity","multiplelines","internetservice","seniorcitizen","techsupport",
"contract","streamingmovies","deviceprotection","paymentmethod","streamingtv","phoneservice",
"paperlessbilling","partner","onlinebackup"],"response_colname":"result"}')



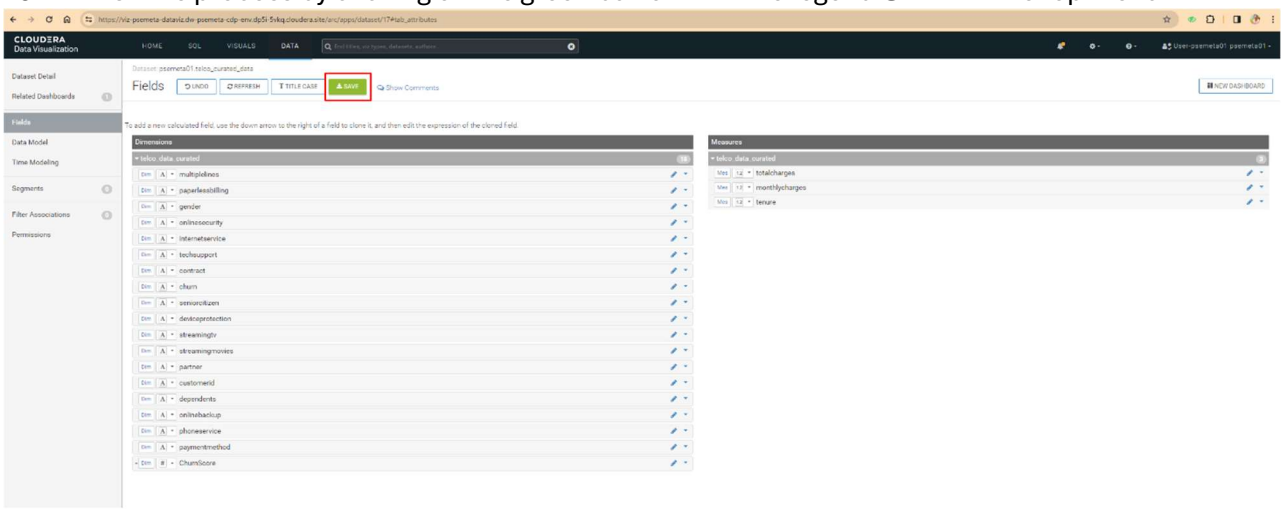
7. Finish the process of copying the ***url*** and the ***accessKey***, click the **Validate Expression** button at the top of the window. If the message appears in green ***Validation Successful***, Click on **Apply** to save the settings made.



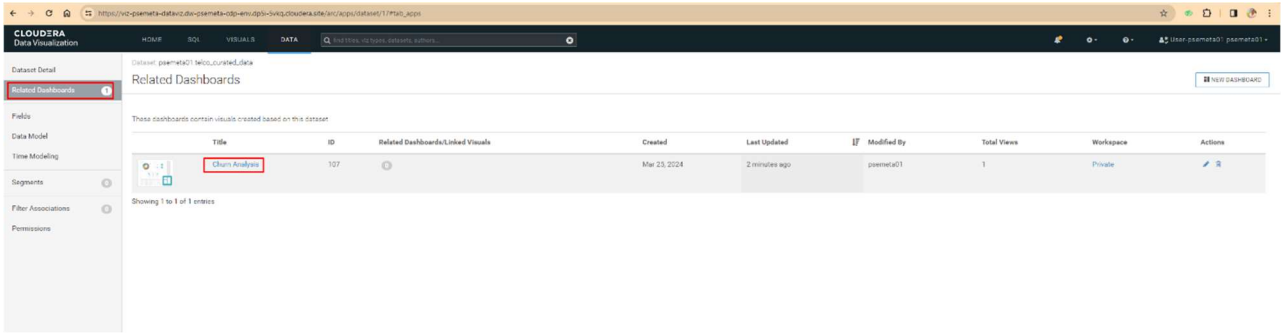
8. The new field should appear in the list of fields. Change the data type, selecting the type ***Integer***, which is represented by the symbol **#**



9. Finish the process by clicking on the green button with the legend **SAVE** in the top menu.

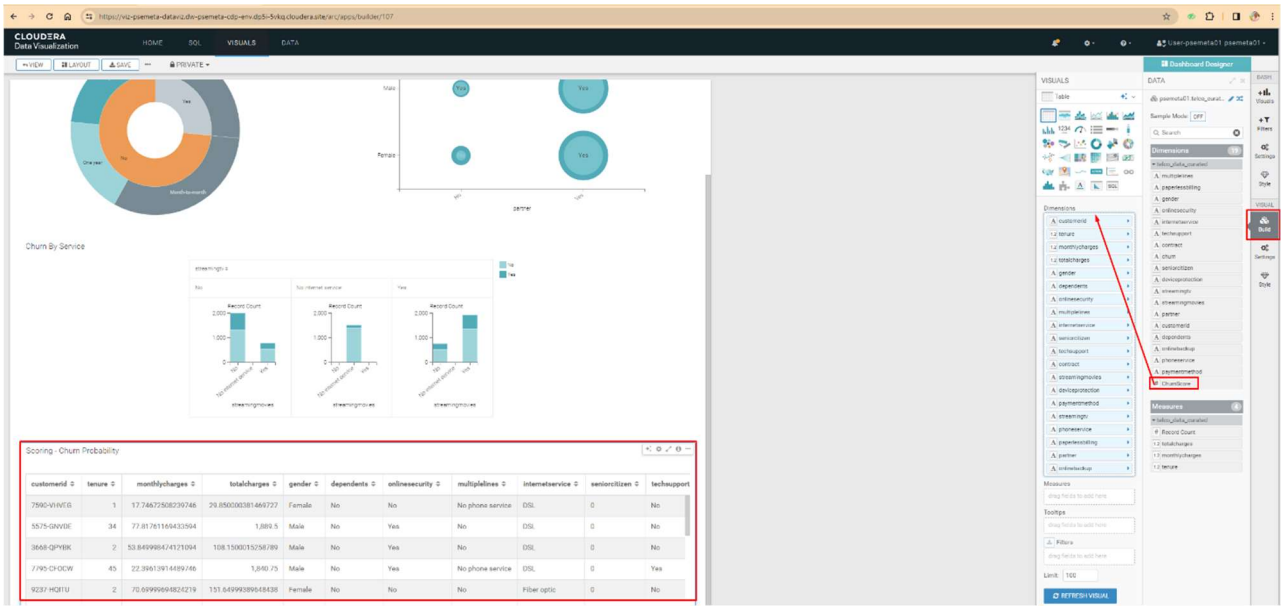


10. Return to the dashboard, select the option **Related Dashboards** from the top left menu, and clicking on the name of the dashboard (**Churn Analysis**) that was previously created.

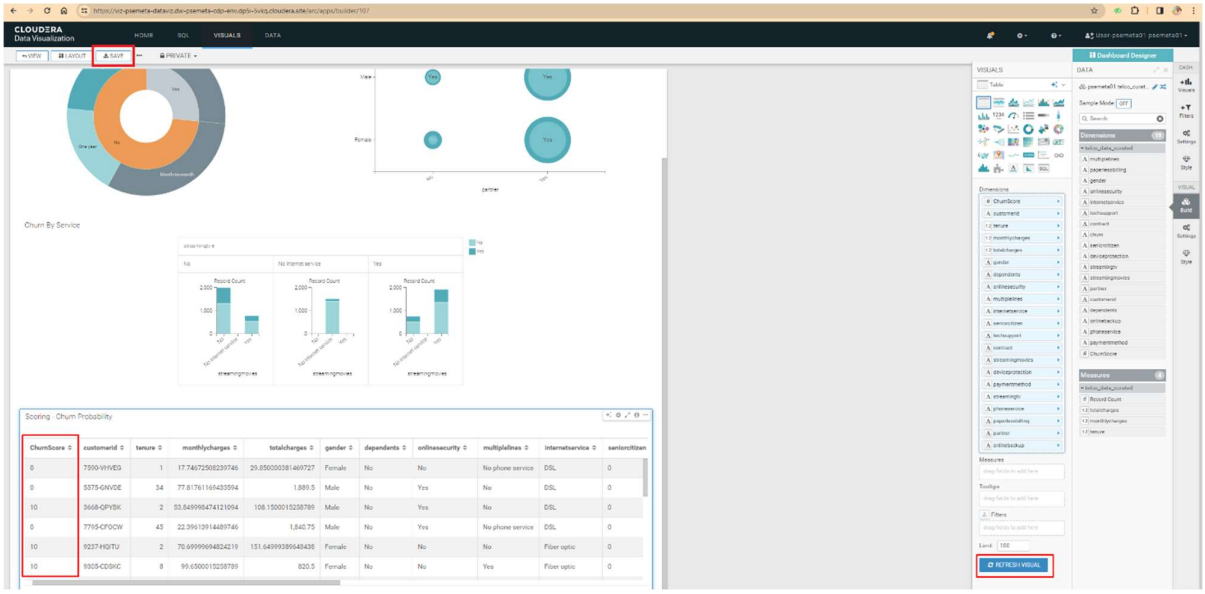


11. Once in the dashboard, click on the button **Edit** which is in the upper left.

12. Edit the lower table by clicking on it and then on the option **Build** from the right vertical menu. Add the new field, **ChurnScore**, at the beginning of the table, by clicking and dragging from the option **Dimensions** available.



13. Click on the Refresh Visual button to update the data. The new column should appear *ChurnScore* then at the beginning of the table, with a value of numeric type. Finish the process by clicking the button **SAVE** from the top left menu.


CLOUDERA
Data Visualization
HOME
SQL
VISUALS
DATA
SAVE
PRIVATE
Churn By Service
Scoring - Churn Probability
ChurnScore
customerid
tenure
monthlycharges
totalcharges
gender
dependents
onlinesecurity
multiplelines
internetservice
seniorcitizen
7590-VHVEG 1 17.746725808239746 29.850000381469727 Female No No No phone service DSL 0
5575-GNVDE 34 77.81761169433594 1,889.5 Male No Yes No DSL 0
3668-QPYBK 2 53.849998474121094 108.15000152587891 Male No Yes No DSL 0
7795-CFOCW 45 22.396139144897461 1,840.75 Male No Yes No phone service DSL 0
9237-HQITU 2 70.699996948242219 151.64999389648438 Female No No No Fiber optic 0
9305-CDSKC 8 99.650001525878906 820.5 Female No No Yes Fiber optic 0
VISUALS
DATA
Dashboard Designer
REFRESH VISUAL