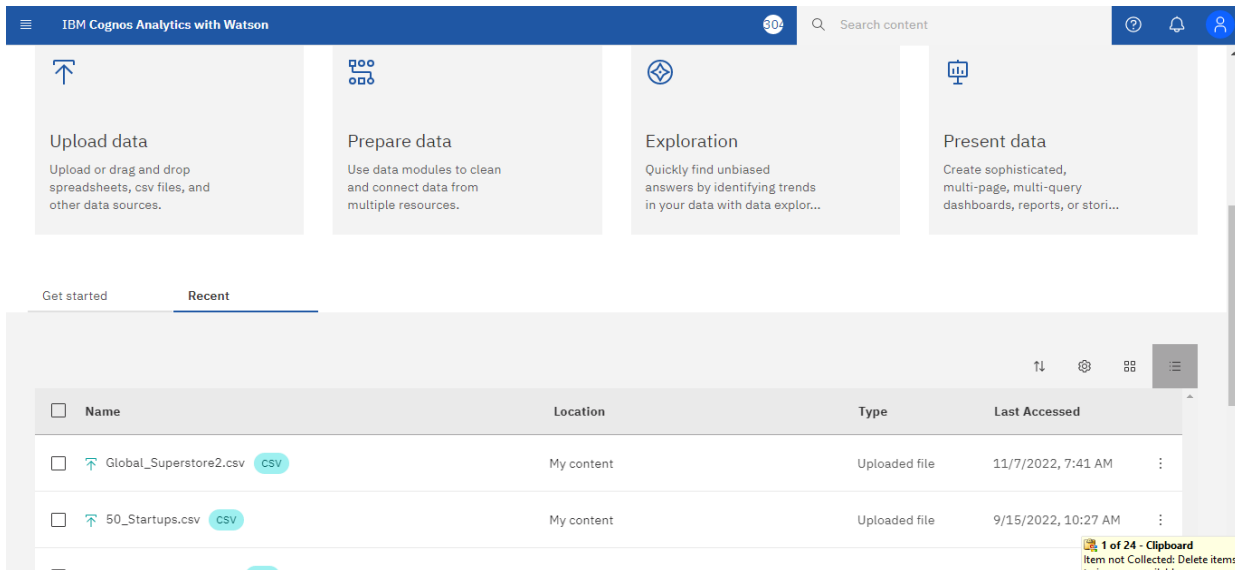


PREPARING THE DATASETS

Team ID	PNT2022TMID31162
Project Name	Global Sales Data Analytics

Data Preparation

(Uploaded DataSet)



DataSet

The screenshot shows the IBM Cognos Analytics Data module interface. The left sidebar contains a 'Data module' section with a search bar and a list of data modules. The 'Global_Superstore2.csv' module is selected. The main area displays a grid view of the dataset. The grid has columns: 'Row ID', 'Row ID', 'Order ID', 'Order Date', 'Ship Date', 'Ship Mode', and 'Customer ID'. The data is organized into rows, with the first row having a 'Row ID' of 1 and an 'Order ID' of CA-2012-124891. The second row has a 'Row ID' of 2 and an 'Order ID' of IN-2013-77878. The third row has a 'Row ID' of 3 and an 'Order ID' of IN-2013-71249. The fourth row has a 'Row ID' of 4 and an 'Order ID' of ES-2013-1579342. The fifth row has a 'Row ID' of 5 and an 'Order ID' of SO-2013-4320. The sixth row has a 'Row ID' of 6 and an 'Order ID' of IN-2013-42360. The seventh row has a 'Row ID' of 7 and an 'Order ID' of IN-2011-81826. The eighth row has a 'Row ID' of 8 and an 'Order ID' of IN-2012-86369. The ninth row has a 'Row ID' of 9 and an 'Order ID' of CA-2014-135909. The tenth row has a 'Row ID' of 10 and an 'Order ID' of IN-2013-116438. The grid also shows 'Order Date', 'Ship Date', 'Ship Mode', and 'Customer ID' for each row.

We check to see whether there are any null values before we prepare the data.

Properties

Data module

Grid Relationships Custom tables

Search

Order ID

Order Date

Ship Date

Ship Mode

Customer ID

Customer Name

Segment

City

State

Country

Postal Code

Market

Region

Product ID

Category

Sub-Category

Row ID

Filter...

Create data group...

Create navigation path...

Split...

Hide from users

Remove

Format data...

Clean...

Sort descending

Sort ascending

Properties

Row ID	Ship Date	Ship Mode	Customer ID
1	2012-07-31	Same Day	BA-19495
2	2013-02-07	Second Class	JB-16210
3	2013-10-18	First Class	CR-12730
4	2013-01-30	First Class	KH-16375
5	2013-11-06	Same Day	BA-9495
6	2013-07-01	Second Class	JM-15655
7	2011-11-29	First Class	TS-21340
8	2012-04-14	Standard Class	MB-18085
9	2014-10-14	Standard Class	JM-15220
10	2013-01-11	Second Class	BA-19495

IBM Cognos Analytics with Watson

New data module

Search

Order ID

Order Date

Ship Date

Ship Mode

Customer ID

Customer Name

Segment

City

State

Country

Postal Code

Market

Region

Product ID

Category

Sub-Category

Product Name

Filter - Order ID

Add a filter condition +

Search

AG-2011-9160

AG-2013-1130

AG-2013-1530

AG-2014-2040

AG-2014-3830

AG-2014-4120

AG-2011-1070

AG-2011-1190

Clear all

Insert

Cancel

OK

Ship Mode	Customer ID
Same Day	BA-19495
Second Class	JB-16210
First Class	CR-12730
First Class	KH-16375
Same Day	BA-9495
Second Class	JM-15655
First Class	TS-21340
Standard Class	MB-18085
Standard Class	JM-15220
Second Class	BA-19495

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	BA-19405
26341	BA-2013-77878	2013-02-05	2013-02-07	Second Class	BA-16210
25330	BA-2013-71249	2013-05-17	2013-05-18	First Class	CA-12730
13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	BA-16375
47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	BA-9495
22732	BA-2013-46290	2013-06-38	2013-07-01	Second Class	BA-15455
30570	BA-2011-81828	2011-11-07	2011-11-09	First Class	TS-21340
31192	BA-2012-86389	2012-04-14	2012-04-18	Standard Class	MB-18085
40155	CA-2014-189009	2014-10-14	2014-10-21	Standard Class	BA-15220
40936	CA-2012-116638	2012-01-28	2012-01-31	Second Class	BA-15985

Calculations of Year, Month, Day fields and also the related Navigation path :
 Right click -> calculation -> type expression -> validate -> Set Name as
 ‘Order_Year’

Create calculation

Name:

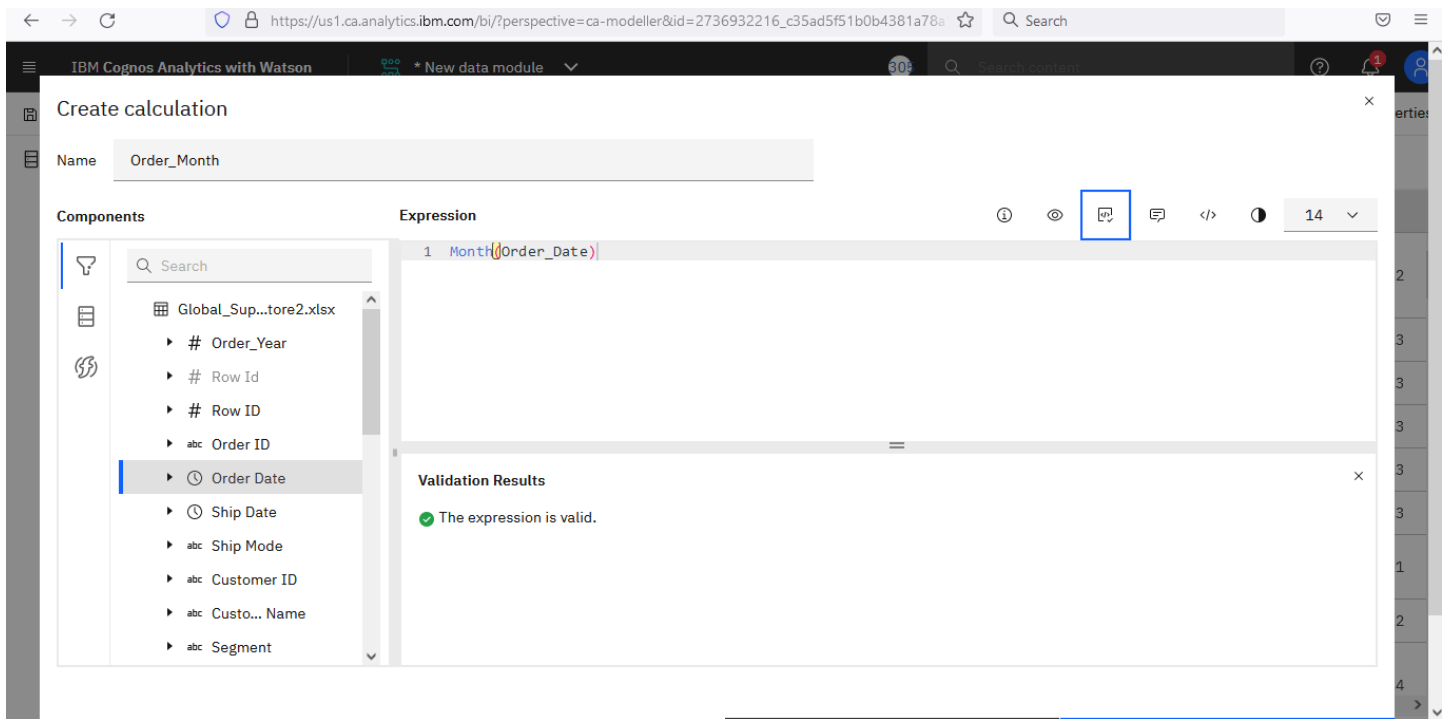
Components:

- Global_Sup...tore2.xlsx
 - # Row ID
 - # Row ID
 - abc Order ID
 - ⌚ Order Date
 - ⌚ Ship Date
 - abc Ship Mode
 - abc Customer ID
 - abc Custo... Name
 - abc Segment
 - 📍 City

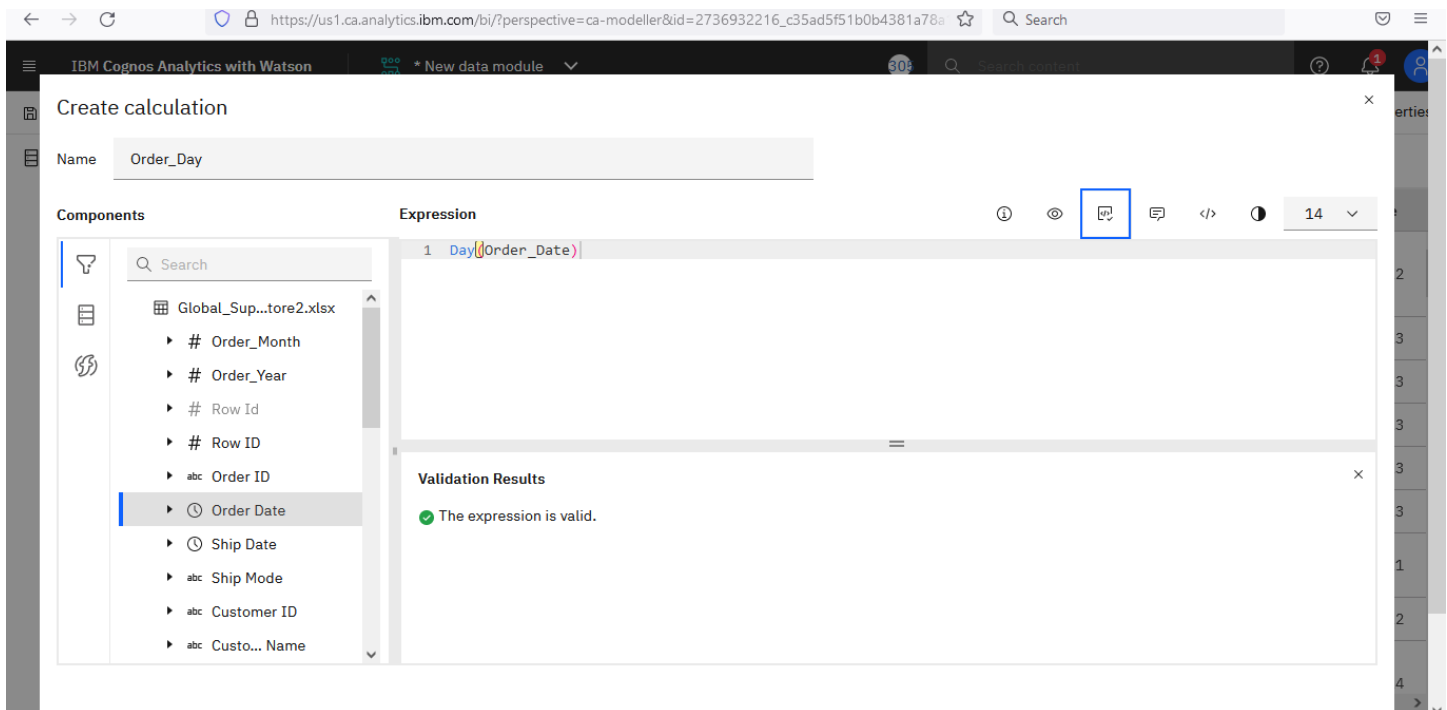
Expression: `Year([Order Date])`

Validation Results: ✔ The expression is valid.

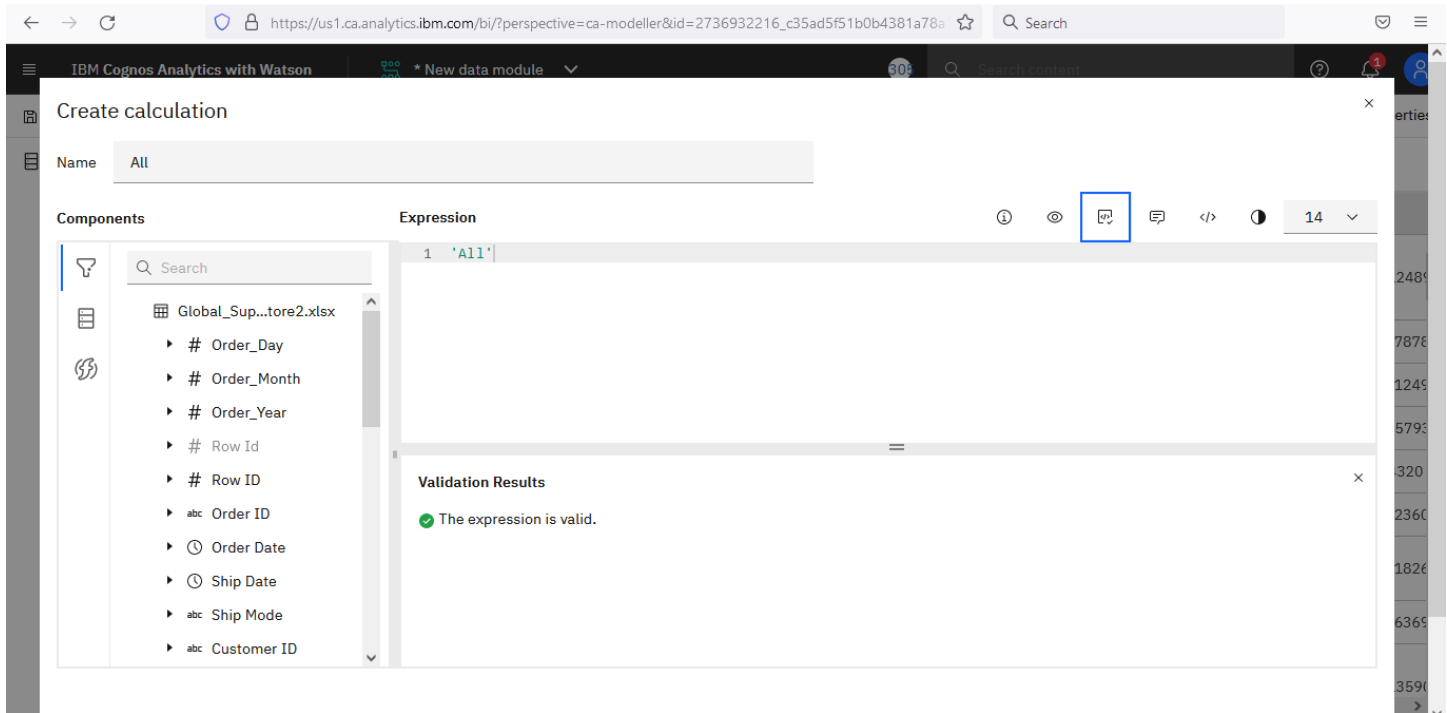
Right click -> calculation -> type expression -> validate -> Set Name as
'Order_Month'



Right click -> calculation -> type expression -> validate -> Set Name as
'Order_Day'



Right click -> calculation -> type expression -> validate -> Set Name as 'All'



Properties -> Usage-> Attribute

Aggregate -> Count Distinct

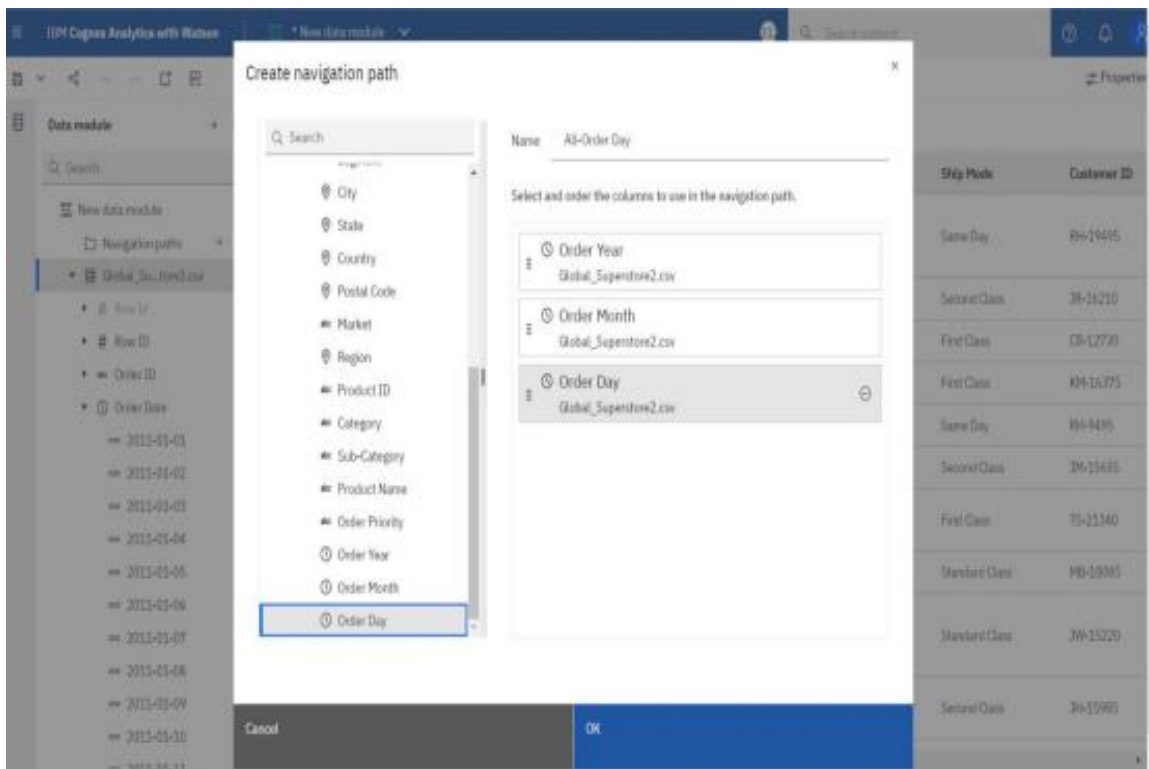
Data Type -> Integer

Represents -> Time, (Year, Month, Day Respectively)

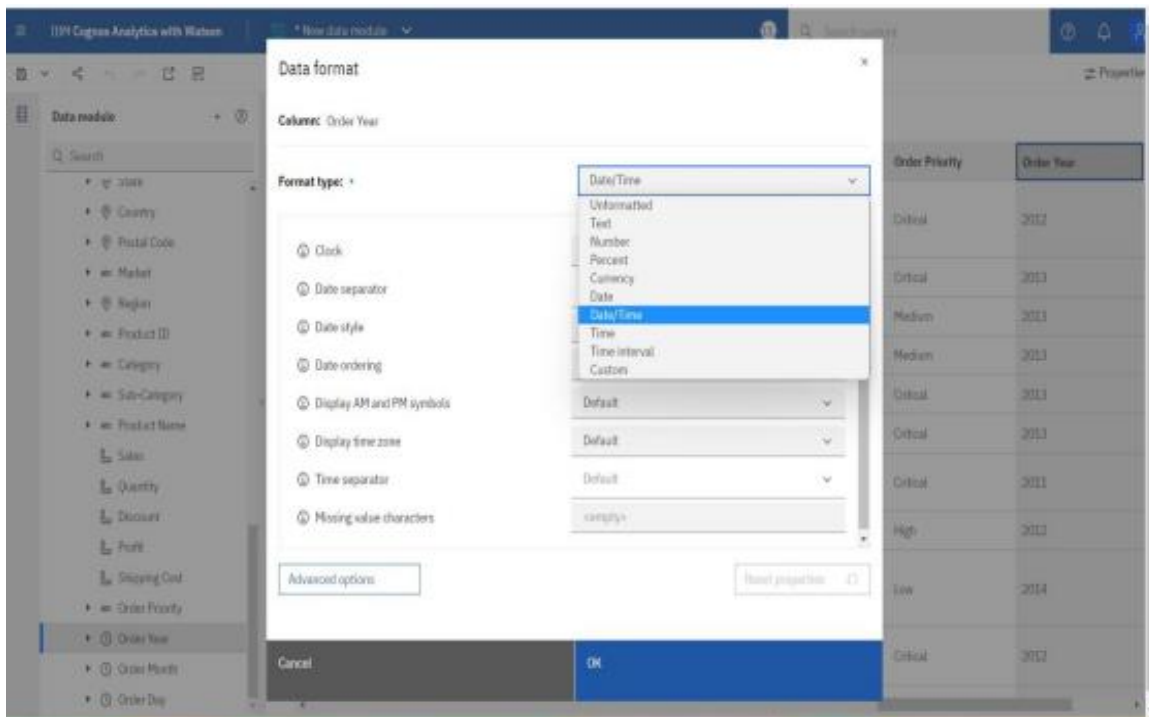
Display Options -> Show Members

This way it becomes a Numerical data Type

Right click on create navigation path



Data format> Date/Time



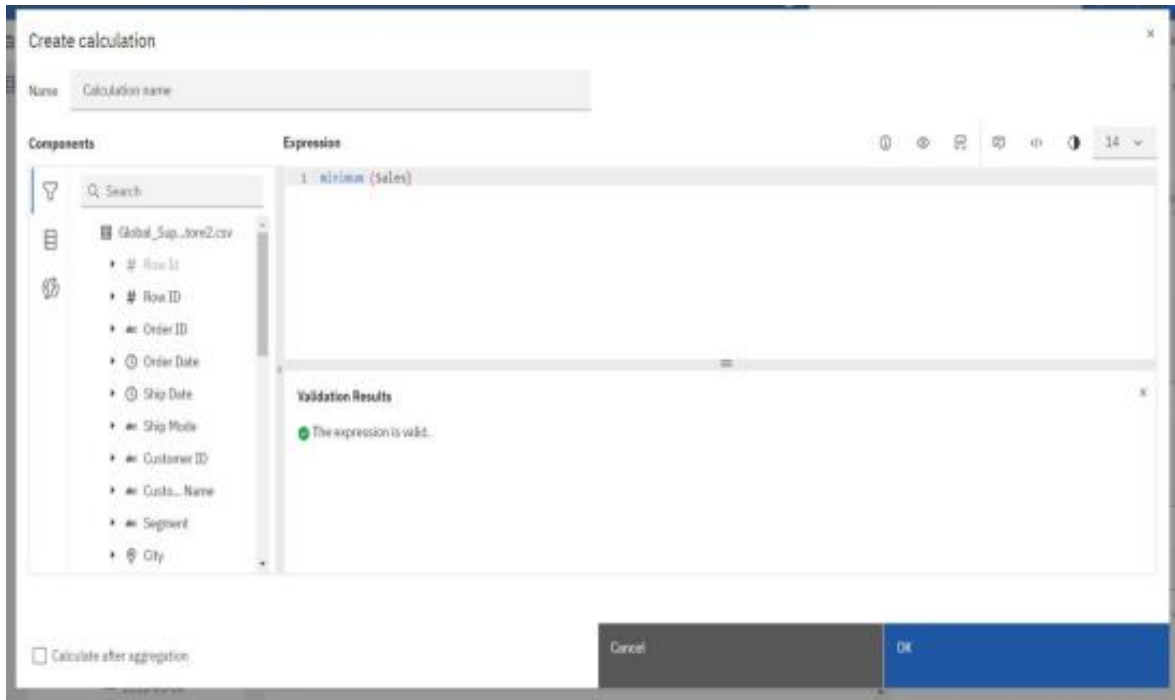
Calculations for Target sales, Min sales, Max Sales and Middle range sales

Min sales -> minimum (Sales)

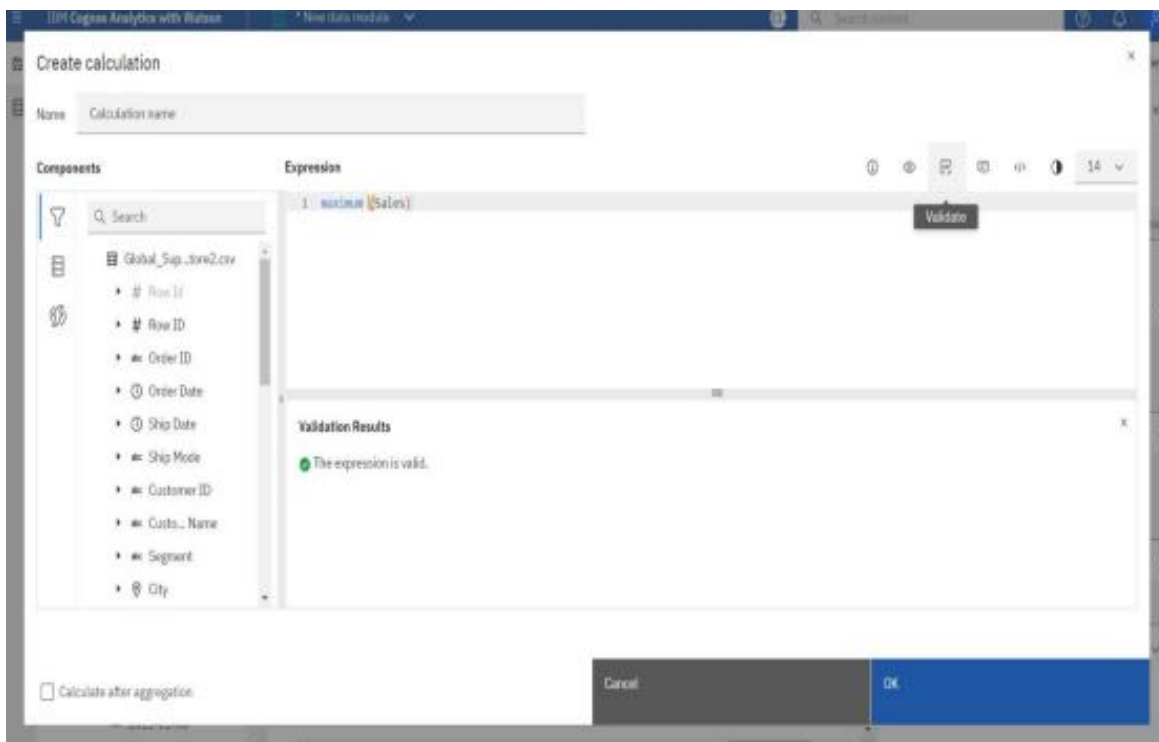
Max sales, Target sales -> maximum (Sales)

Middle range sales -> average (Sales)

Minimum sales :



Maximum Sales :



Middle Range sales :

