

Preparing data

Team ID	PNT2022TMID13572
Project Name	Global Sales Data Analytics

Data Preparation

(Uploaded DataSet)

The screenshot shows the IBM Cognos Analytics with Watson interface. The top navigation bar includes a menu icon, the text "IBM Cognos Analytics with Watson", a user profile icon, and a search bar. Below the navigation bar, there are four main sections: "Upload data" (with an upload icon and description), "Prepare data" (with a puzzle piece icon and description), "Exploration" (with a magnifying glass icon and description), and "Present data" (with a bar chart icon and description). Below these sections, there are two tabs: "Get started" and "Recent". The "Recent" tab is active, showing a list of uploaded files. The list has columns: Name, Location, Type, and Last Accessed. Two files are listed: "Global_Superstore2.csv" and "50_Startups.csv", both marked as "CSV" files and "Uploaded file". A notification banner at the bottom right says "1 of 24 - Clipboard Item not Collected: Delete items".

Name	Location	Type	Last Accessed
Global_Superstore2.csv	My content	Uploaded file	11/7/2022, 7:41 AM
50_Startups.csv	My content	Uploaded file	9/15/2022, 10:27 AM

DataSet

The screenshot shows the IBM Cognos Analytics interface with a data grid. The left sidebar shows a "Data module" with a search bar and a list of fields: Row ID, Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, City, State, Country, and Postal Code. The main area displays a grid of data with columns: Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. The grid contains 10 rows of data. A "Properties" panel is visible on the right side of the grid.

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
1	CA-2012-124891	2012-07-31	2012-07-31	Same Day	BH-19495
2	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210
3	IN-2013-71249	2013-10-17	2013-10-18	First Class	CR-12730
4	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-16375
5	95-2013-4320	2013-11-05	2013-11-06	Same Day	BH-9495
6	IN-2013-42360	2013-06-28	2013-07-01	Second Class	JM-15655
7	IN-2011-01826	2011-11-07	2011-11-09	First Class	TS-21340
8	IN-2012-86369	2012-04-14	2012-04-18	Standard Class	MB-18085
9	CA-2014-135909	2014-10-24	2014-10-21	Standard Class	JW-15220
10	CA-2015-116618	2015-01-18	2015-01-15	Standard Class	BL-15085

For Prepare the data, we verify whether the null values are present or not.

Grid

Relationships

Custom tables

Search

Order ID

Order Date

Ship Date

Ship Mode

Customer ID

Customer Name

Segment

City

State

Country

Postal Code

Market

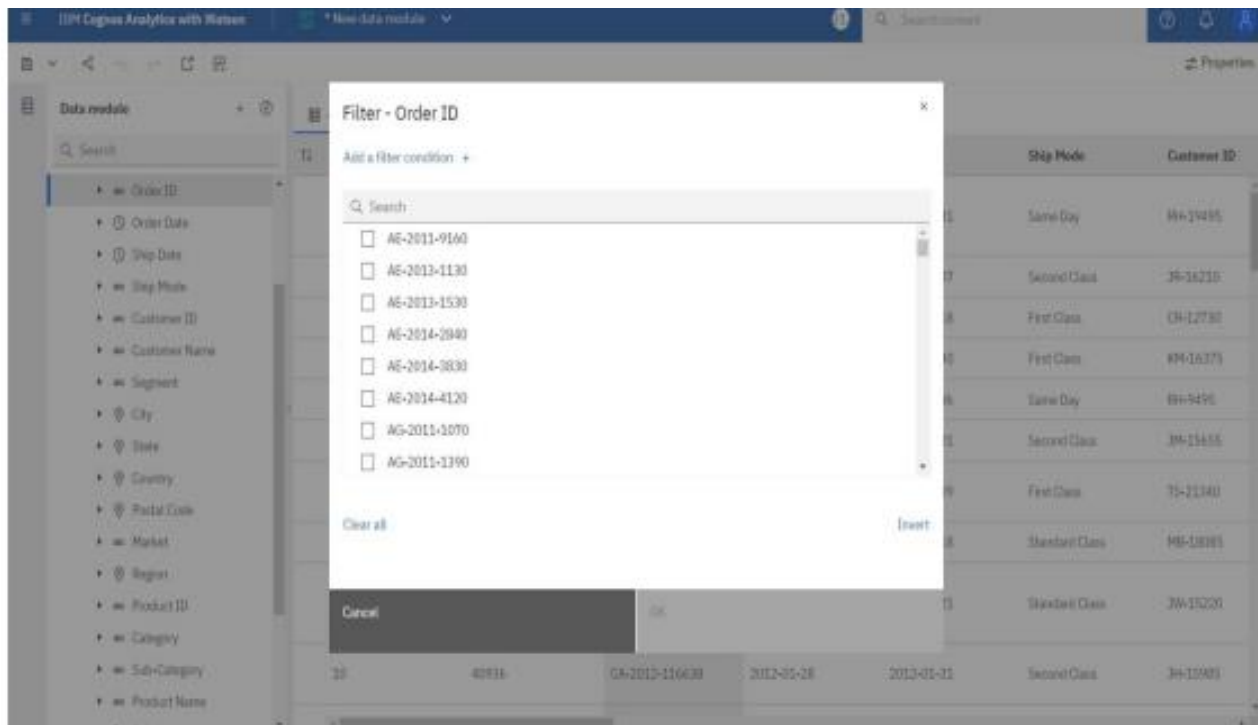
Region

Product ID

Category

Sub-Category

Row ID	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
1	32298	CA	2012-07-31	Same Day	RM-19495	
2	26341	IN	2013-02-07	Second Class	JR-16210	
3	25130	IN	2013-10-18	First Class	CR-12730	
4	13524	ES	2013-01-30	First Class	KM-16375	
5	47221	SG	2013-11-06	Same Day	RM-9495	
6	22732	IN	2013-07-01	Second Class	JM-15655	
7	30570	IN	2011-11-09	First Class	TS-21340	
8	11192	IN-2012-06309	2012-04-14	2012-04-18	Standard Class	MB-18085
9	40155	CA-2014-139009	2014-10-14	2014-10-21	Standard Class	JM-15220
10	40916	CA-2013-116638	2013-05-18	2013-05-21	Second Class	JM-10985



The screenshot shows the 'Filter - Order ID' dialog box in Power BI Desktop. The dialog has a search bar and a list of Order IDs with checkboxes. The 'Add a filter condition' button is at the top. The 'Clear all' button is at the bottom left, and the 'Insert' button is at the bottom right. The background shows the same data table as the previous screenshot.

Order ID
<input type="checkbox"/> AE-2011-9160
<input type="checkbox"/> AE-2013-1130
<input type="checkbox"/> AE-2013-1530
<input type="checkbox"/> AE-2014-2040
<input type="checkbox"/> AE-2014-3830
<input type="checkbox"/> AE-2014-4120
<input type="checkbox"/> AG-2015-1070
<input type="checkbox"/> AG-2015-1190

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
32298	CA-2012-126891	2012-07-31	2012-07-31	Same Day	BA-19406
26341	BA-2013-77878	2013-02-08	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-2013-42360	2013-06-38	2013-07-01	Second Class	BA-15655
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-2014-18909	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:

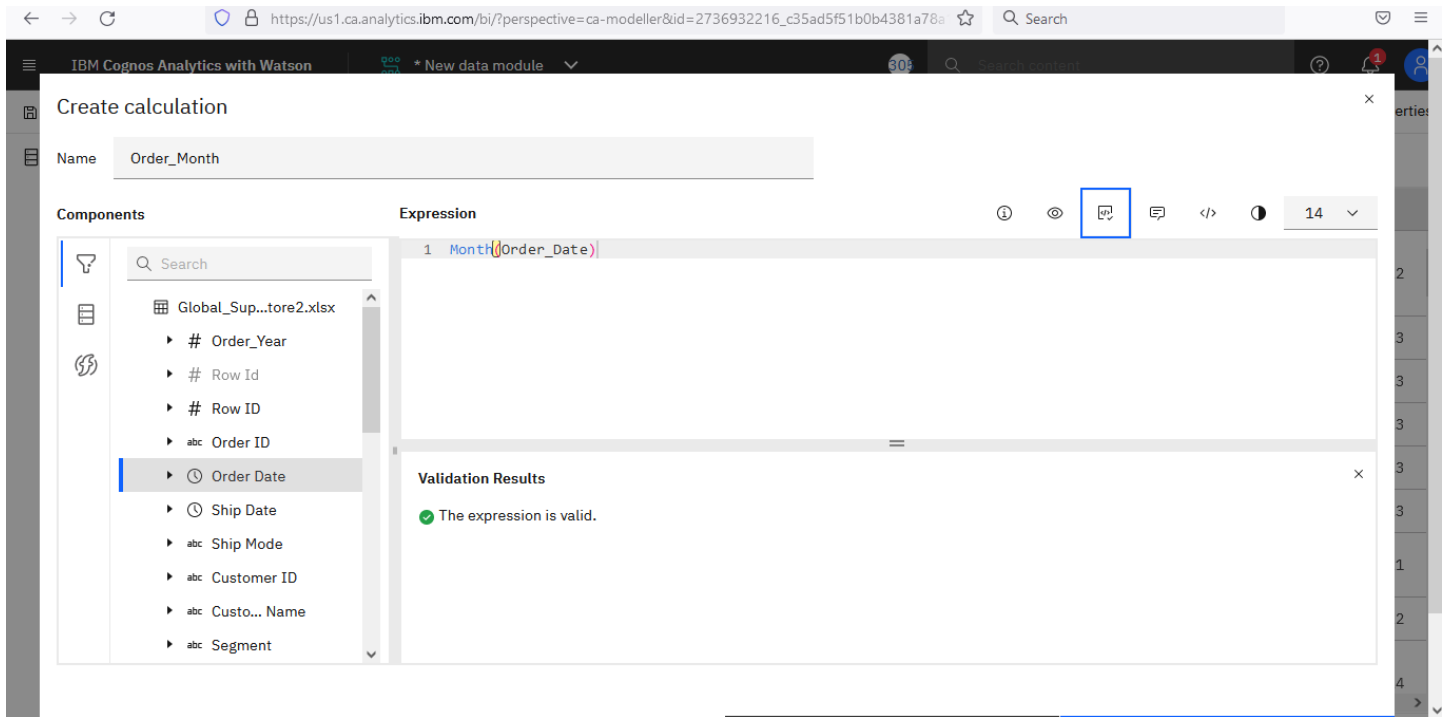
- Global_Sup...tore2.xlsx
 - # Row ID
 - # Row ID
 - abc Order ID
 - 🕒 Order Date
 - 🕒 Ship Date
 - abc Ship Mode
 - abc Customer ID
 - abc Custo... Name
 - abc Segment
 - 📍 City

Expression: `1 Year([Order_Date])`

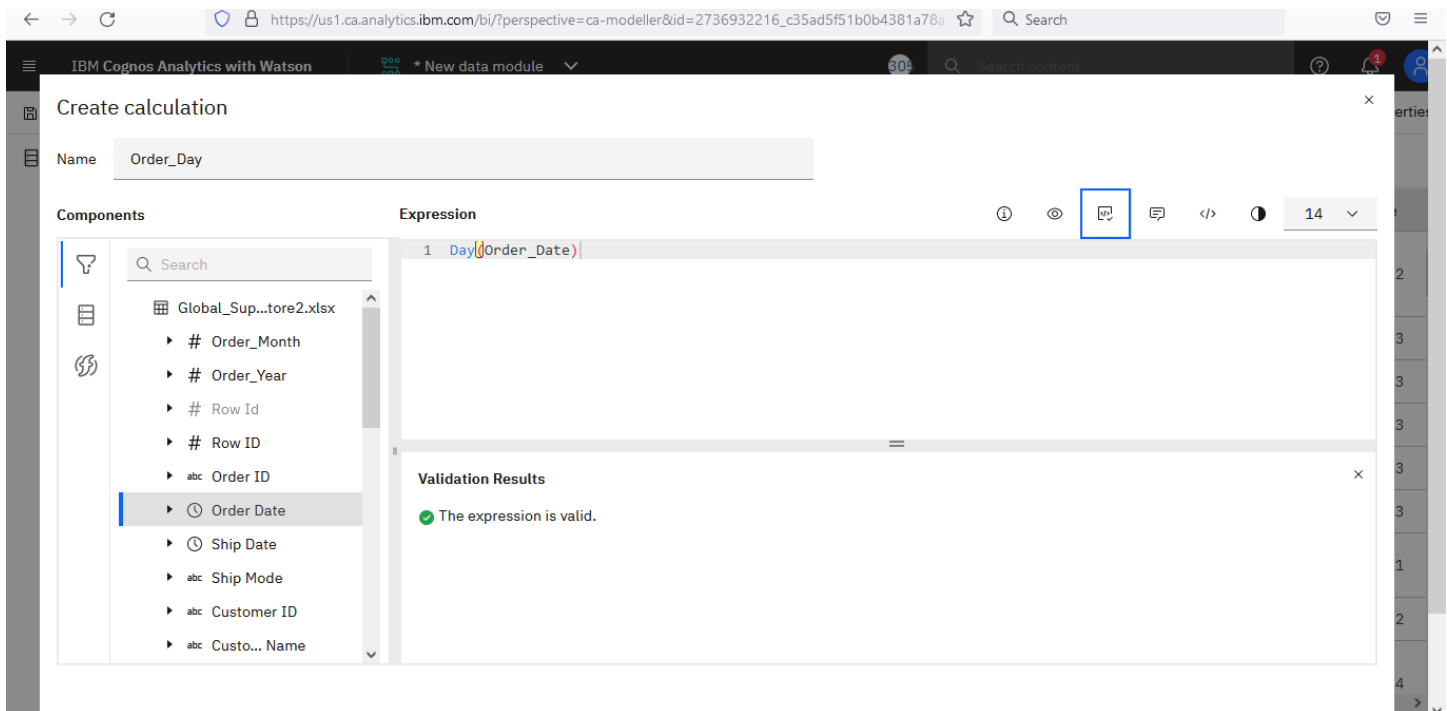
Validation Results

✅ The expression is valid.

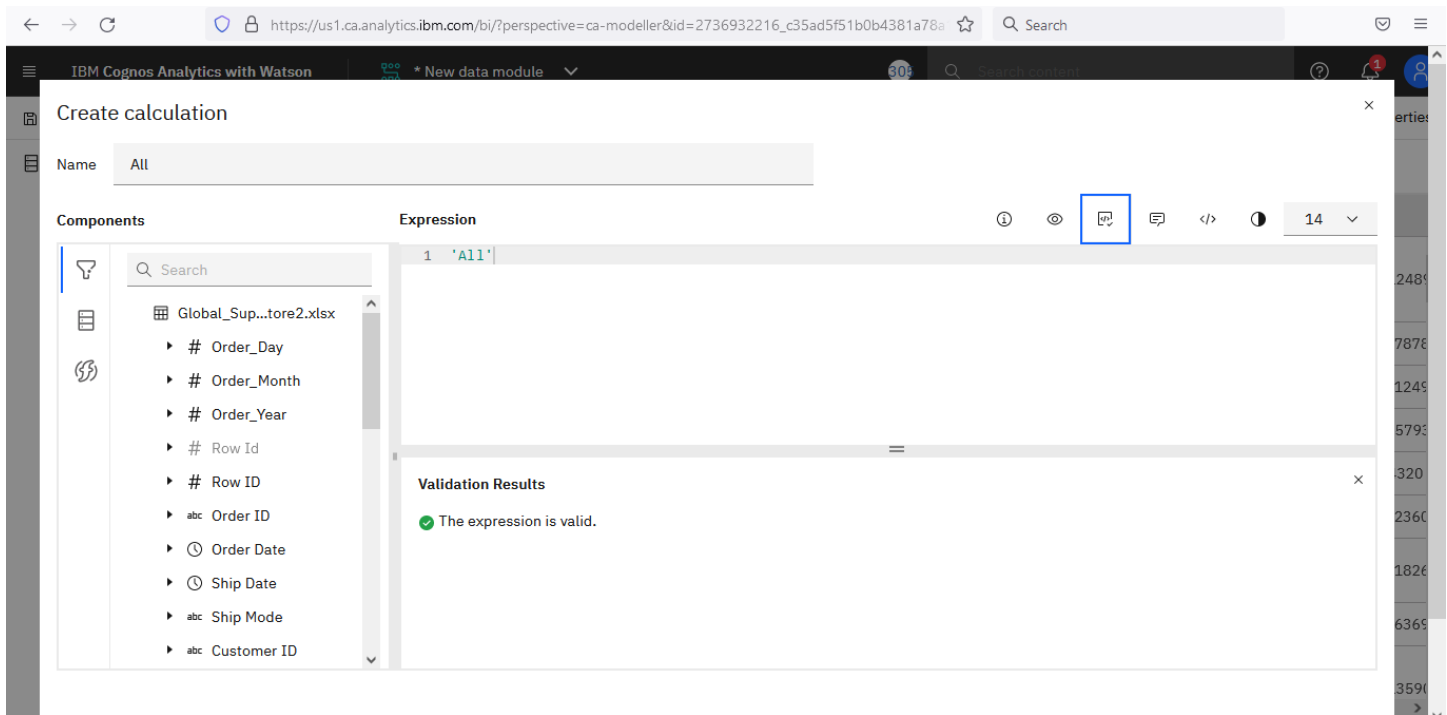
Right click -> calculation -> type expression -> validate -> Set Name as 'Order_Month'



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

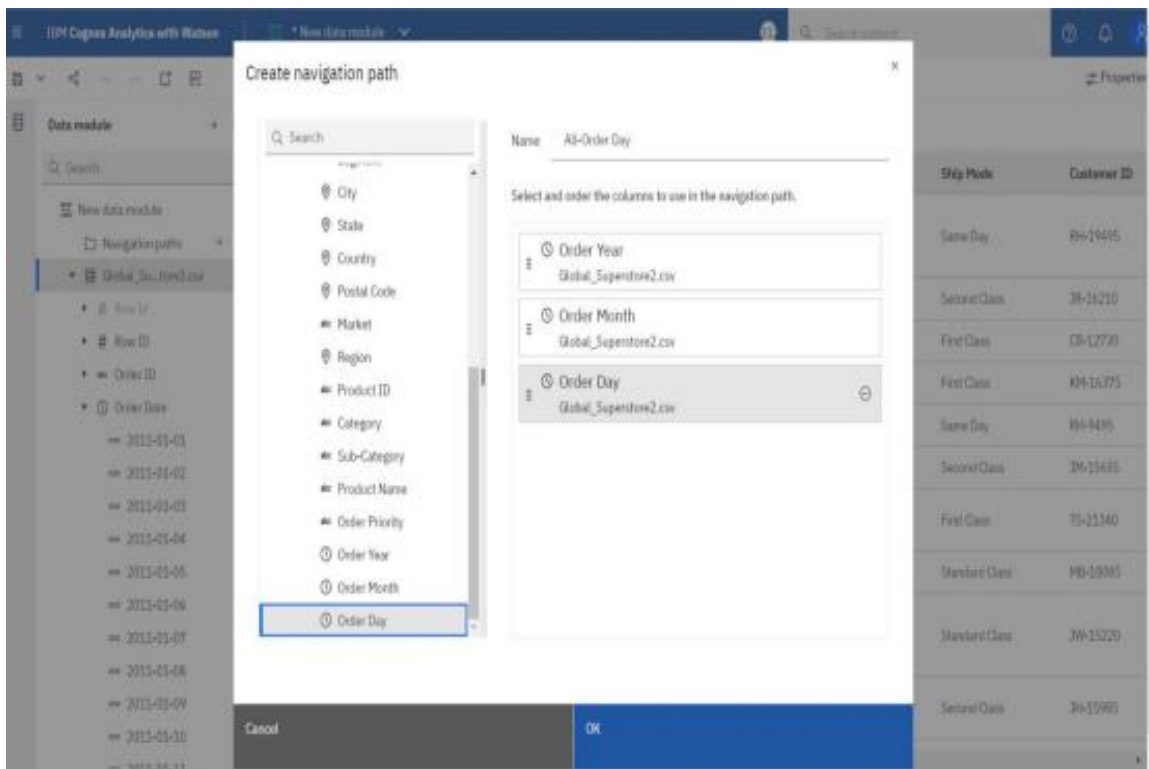
Data Type -> Integer

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

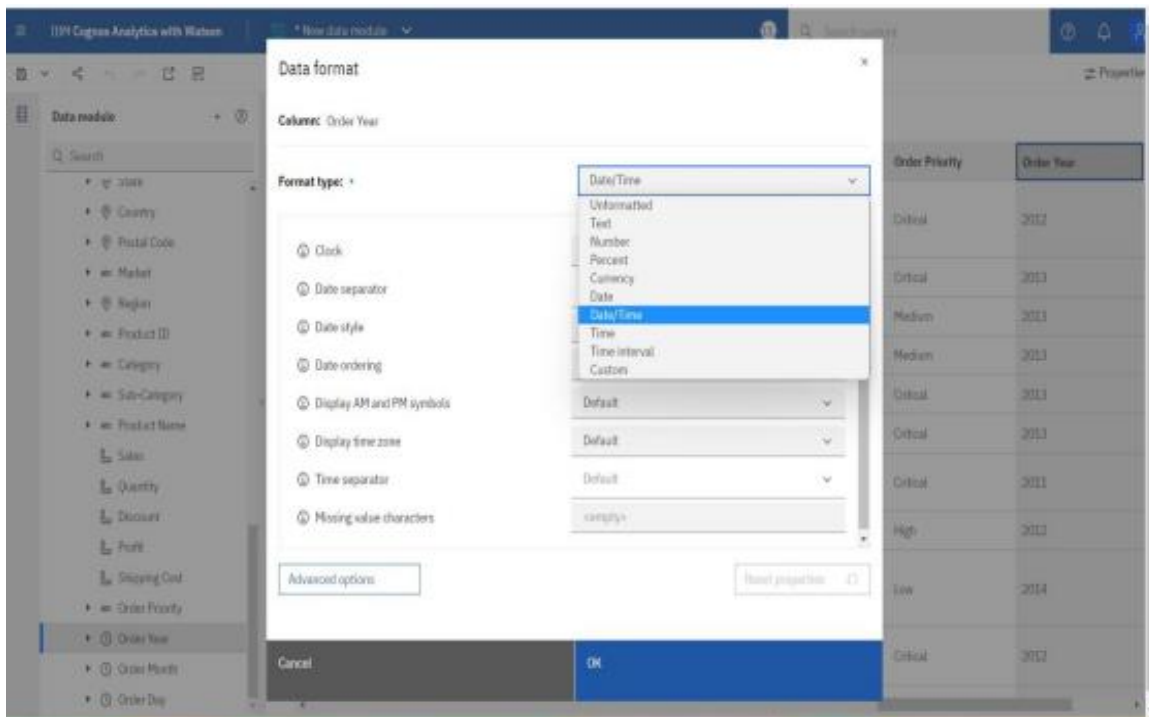
Display Options -> Show Members

This way it becomes a Numerical data Type

Right click on create navigation path



Data format> Date/Time



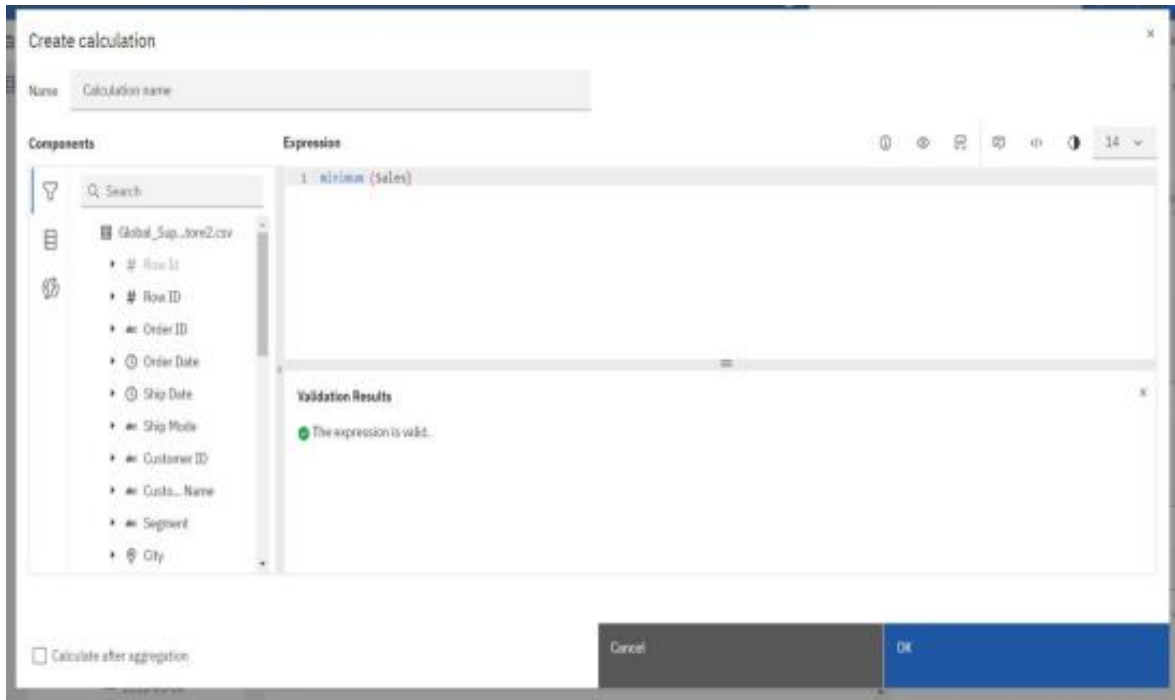
Calculations for Target sales, Min sales, Max Sales and Middle range sales

Min sales -> minimum (Sales)

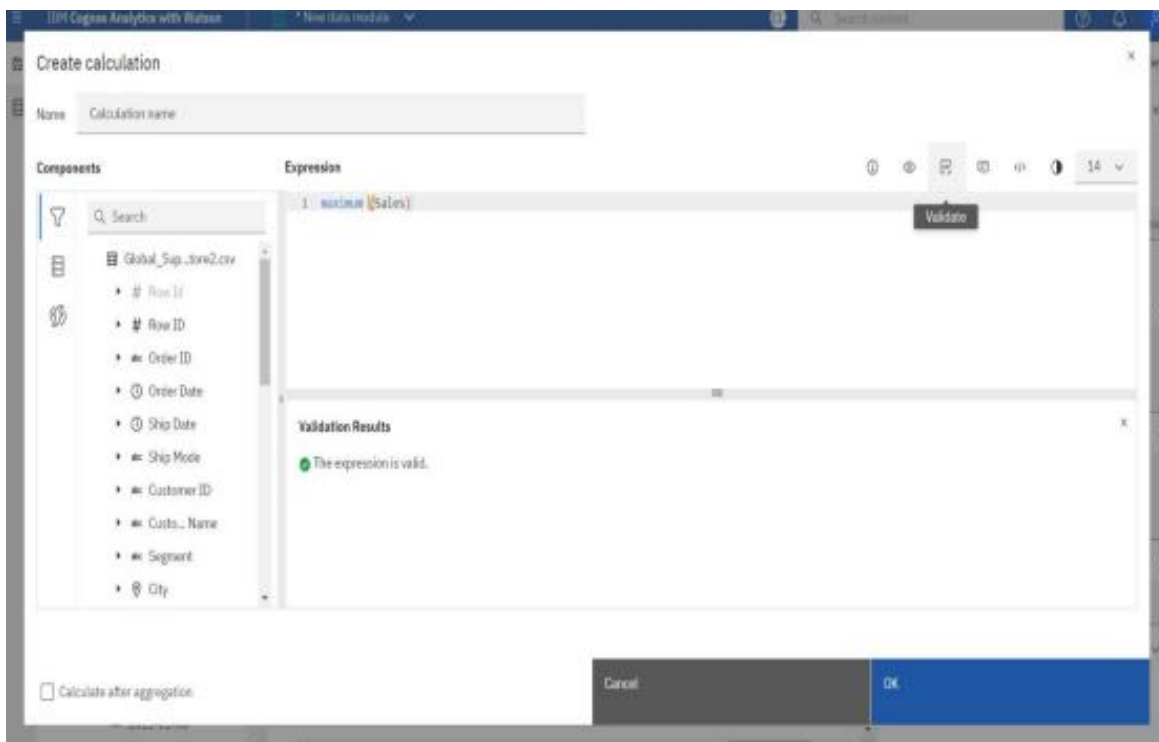
Max sales, Target sales -> maximum (Sales)

Middle range sales -> average (Sales)

Minimum sales :



Maximum Sales :



Middle Range sales :

