

# PREPARING THE DATASETS

Team ID	PNT2022TMID51329
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 action cards: 'Upload data' (with an upload icon), 'Prepare data' (with a puzzle piece icon), 'Exploration' (with a magnifying glass icon), and 'Present data' (with a bar chart icon). Each card has a brief description of its function. Below these cards, there are tabs for 'Get started' and 'Recent'. The 'Recent' tab is active, showing a list of uploaded files. The list has columns for 'Name', 'Location', 'Type', and 'Last Accessed'. Two files are listed: 'Global\_Superstore2.csv' and '50\_Startups.csv', both of type 'Uploaded file' and located in 'My content'. A clipboard notification is visible in the bottom right corner.

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 view of the 'Global\_Superstore2.csv' dataset. The left sidebar shows the 'Data module' and a list of fields including 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 table with 10 rows of data. The columns are Row ID, Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID.

Row ID	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
1	32298	CA-2012-124891	2012-07-11	2012-07-11	Same Day	RM-19495
2	26341	RM-2013-77078	2013-02-05	2013-02-07	Second Class	JR-16210
3	25330	RM-2013-71249	2013-05-17	2013-05-18	First Class	CR-12730
4	13524	IS-2013-1579342	2013-01-28	2013-01-30	First Class	KM-16378
5	47221	IS-2013-43320	2013-11-05	2013-11-06	Same Day	RM-9495
6	22732	RM-2013-42360	2013-06-28	2013-07-01	Second Class	JM-15455
7	30570	RM-2011-81838	2011-11-07	2011-11-09	First Class	TS-21340
8	31192	RM-2012-86369	2012-04-14	2012-04-18	Standard Class	MB-18385
9	40155	CA-2014-10909	2014-05-14	2014-05-21	Standard Class	JW-15220
10	87016	CA-2013-116478	2013-01-18	2013-01-19	Standard Class	ML-16881

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



Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RM-19495
26341	RM-2013-77978	2013-02-05	2013-02-07	Second Class	RM-16210
25330	RM-2013-71249	2013-02-17	2013-02-18	First Class	CA-12730
13124	ES-2013-1579342	2013-01-28	2013-01-30	First Class	RM-16375
47221	SO-2013-4320	2013-11-05	2013-11-06	Same Day	RM-9495
22732	RM-2013-42360	2013-06-18	2013-07-01	Second Class	RM-15655
30570	RM-2011-81826	2011-11-07	2011-11-09	First Class	TS-21340
31192	RM-2012-86369	2012-04-14	2012-04-18	Standard Class	RM-18085
40355	CA-2014-13909	2014-10-14	2014-10-21	Standard Class	RM-15220
40936	CA-2012-116638	2012-01-28	2012-01-31	Second Class	RM-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: Order\_Year

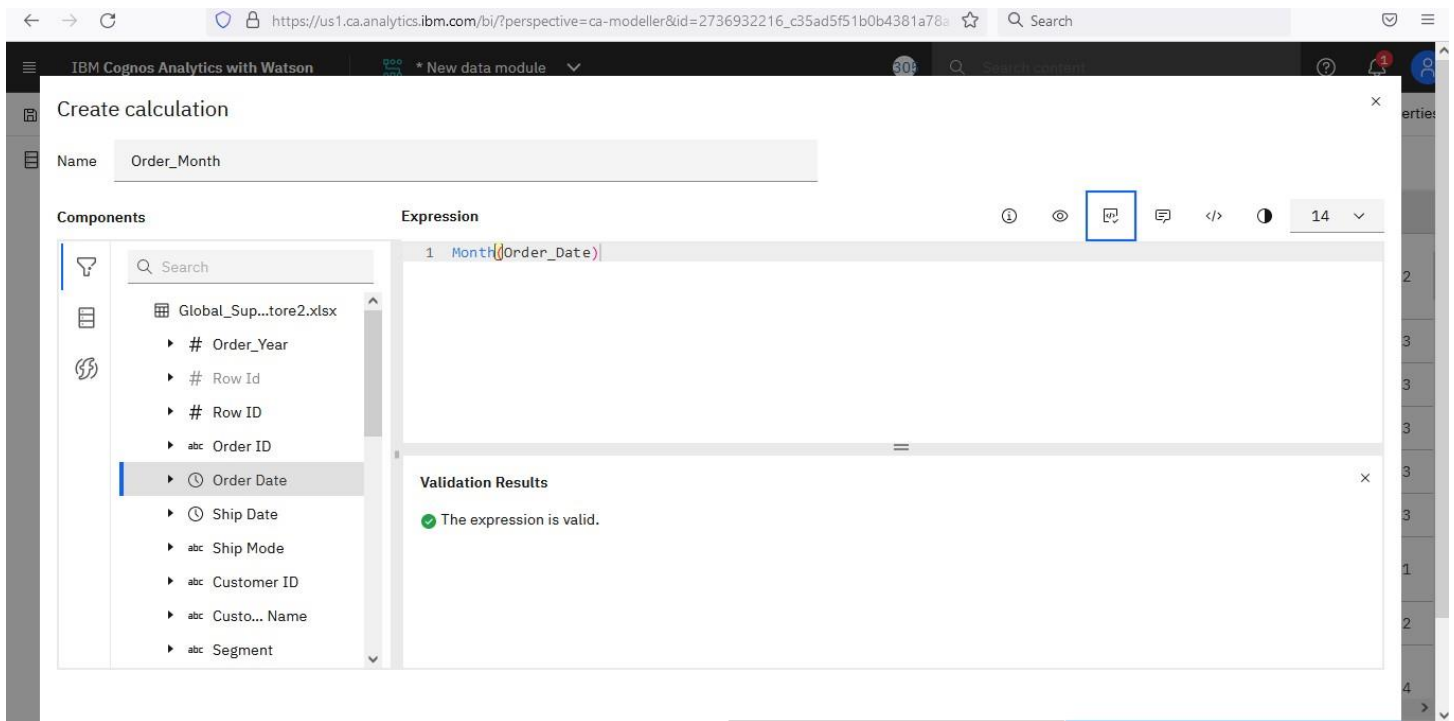
Components:

- 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

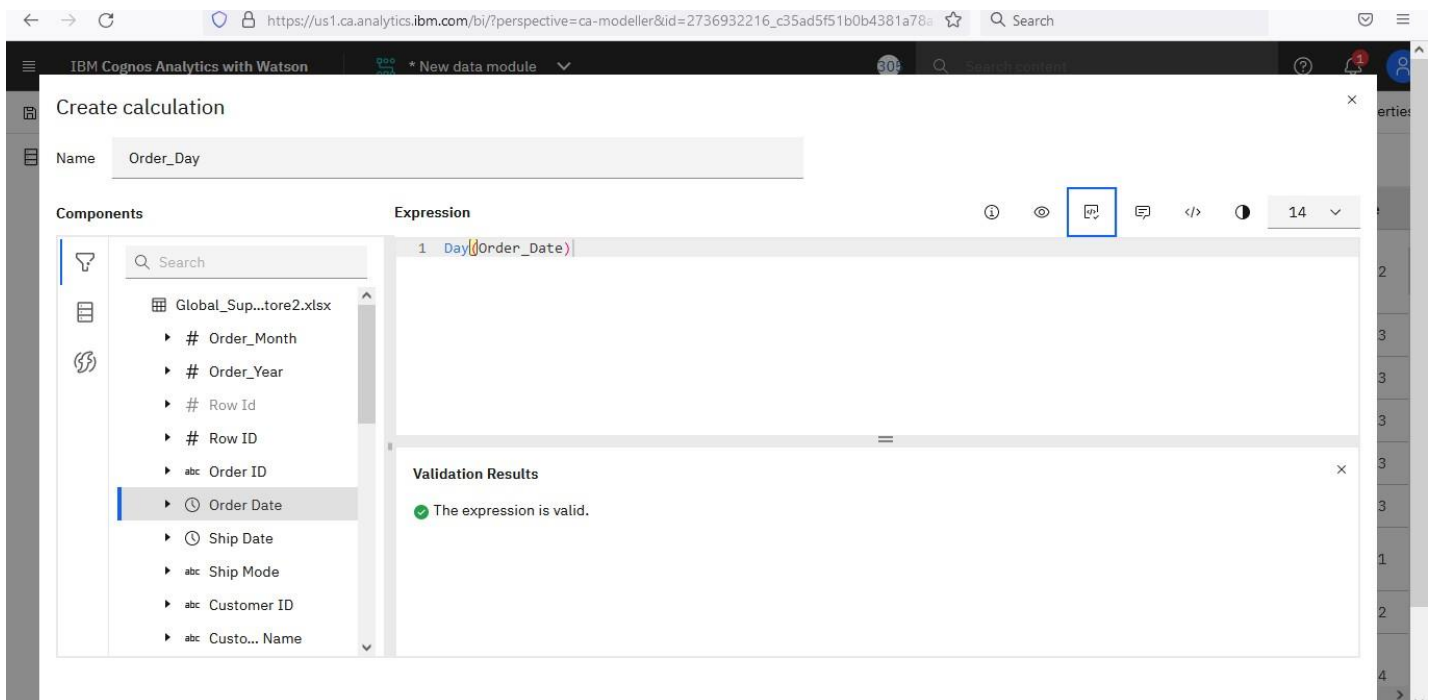
Expression: 1 Year(Order\_Date)

Validation Results: ✔ The expression is valid.

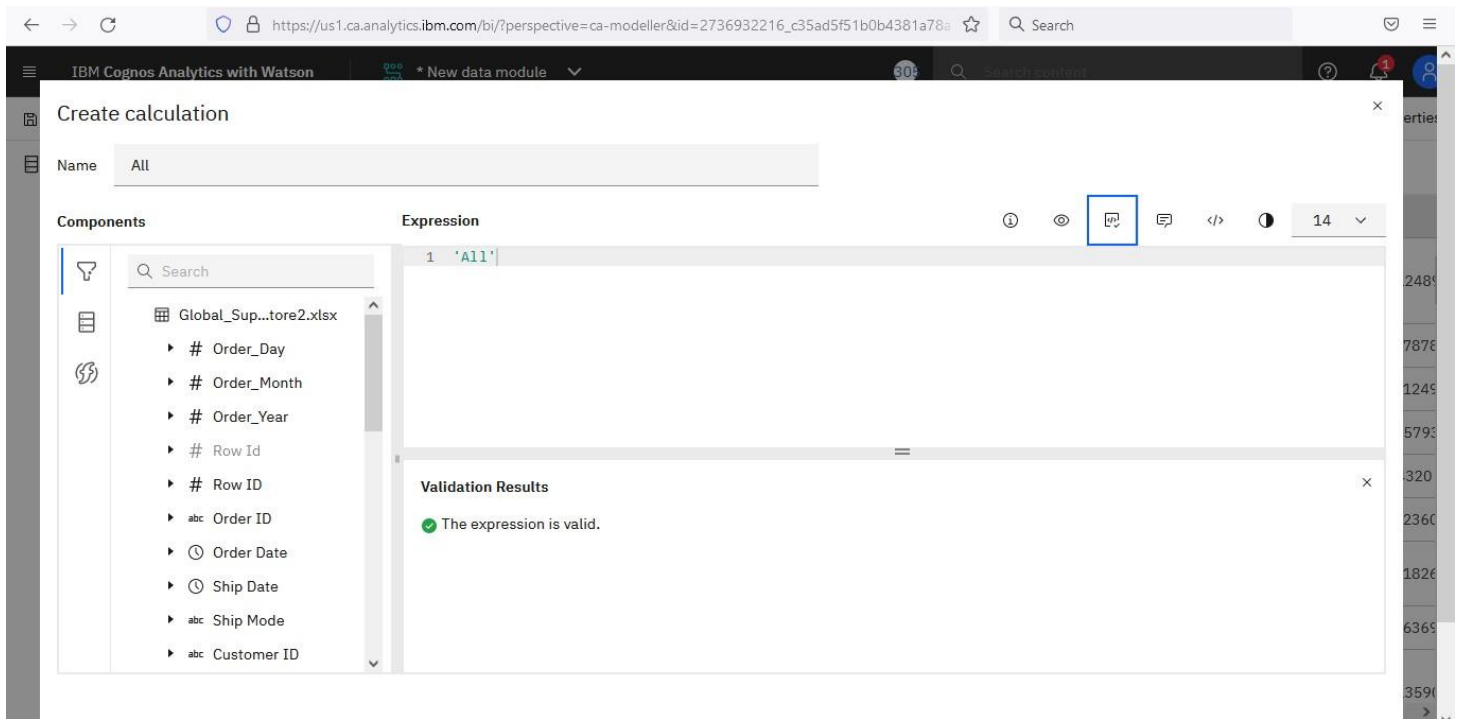
Right click -> calculation -> type expression -> validate -> Set Name as  
'Order\_Month'



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

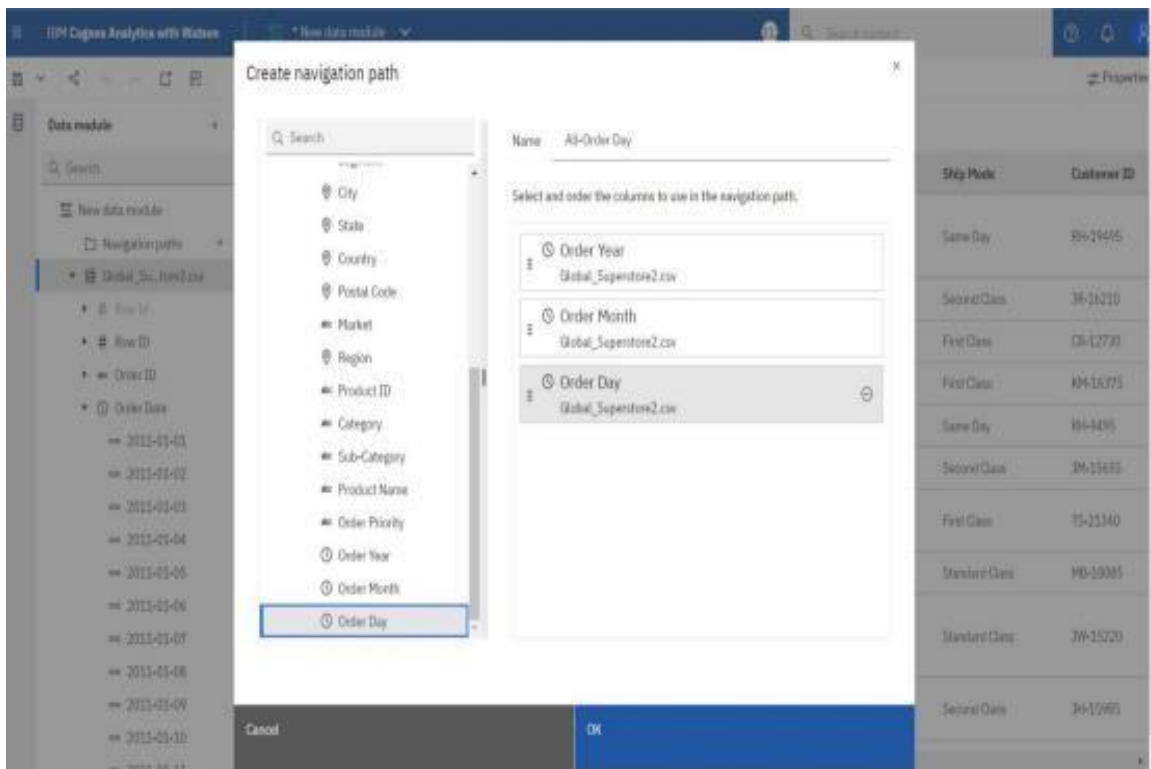
Data Type -> Integer

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

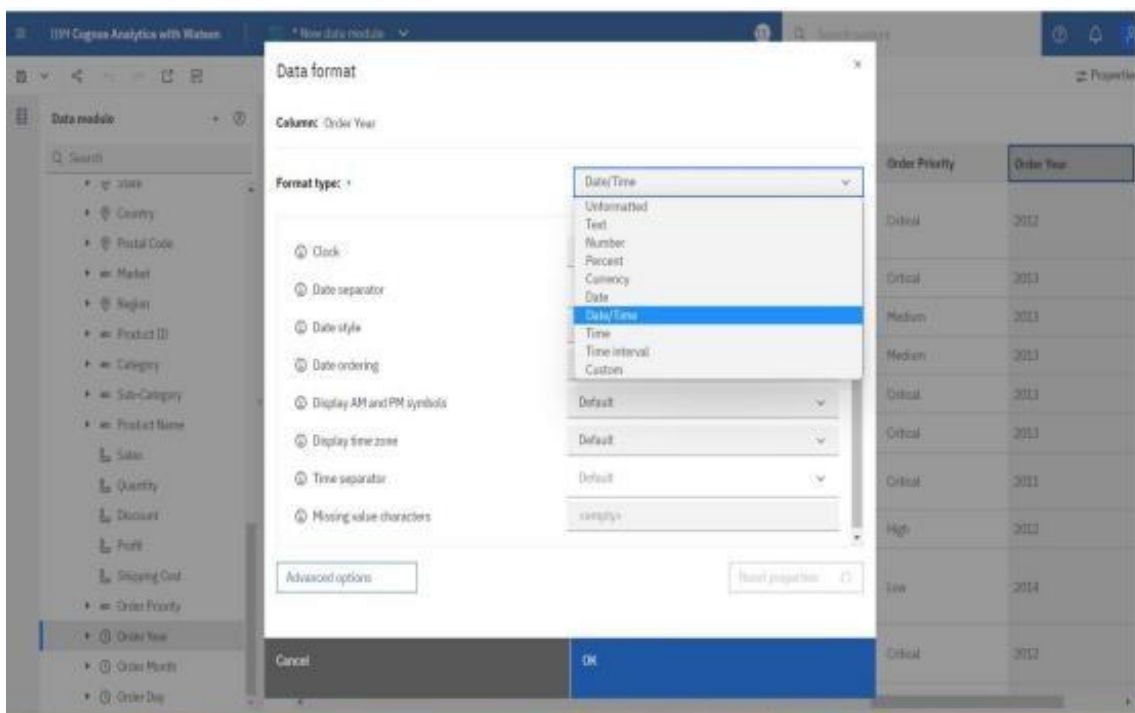
Display Options -> Show Members

This way it becomes a Numerical data Type

Right click on create navigation path



Data format> Date/Time



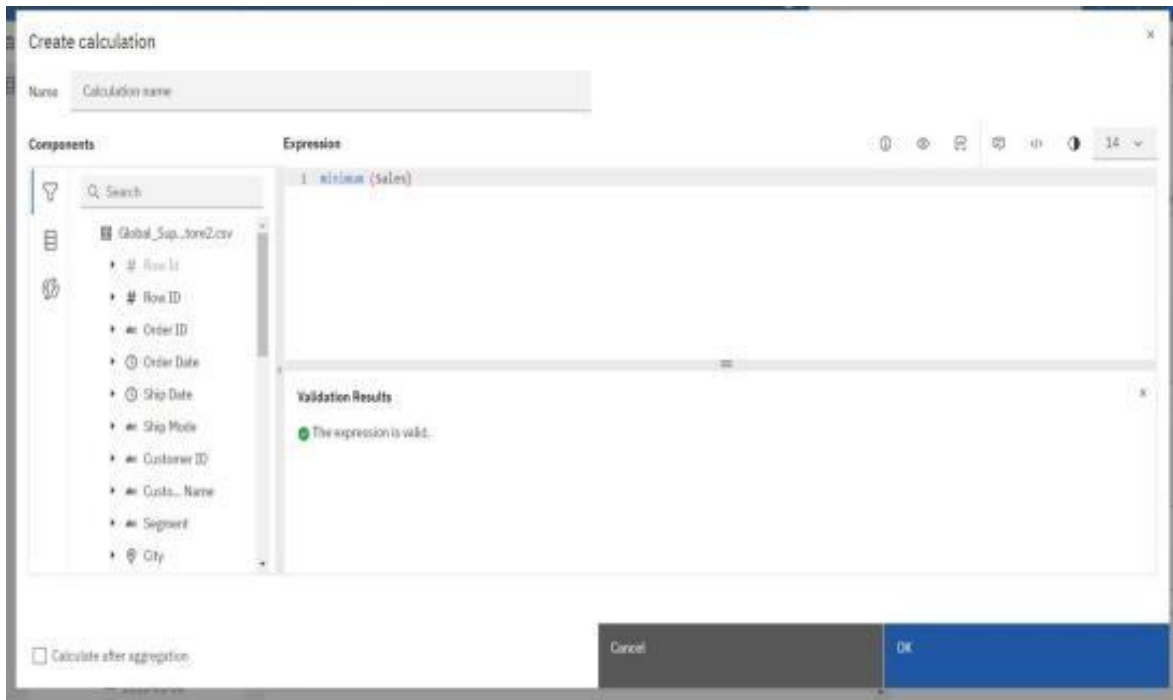
## Calculations for Target sales, Min sales, Max Sales and Middle range sales

Min sales -> minimum (Sales)

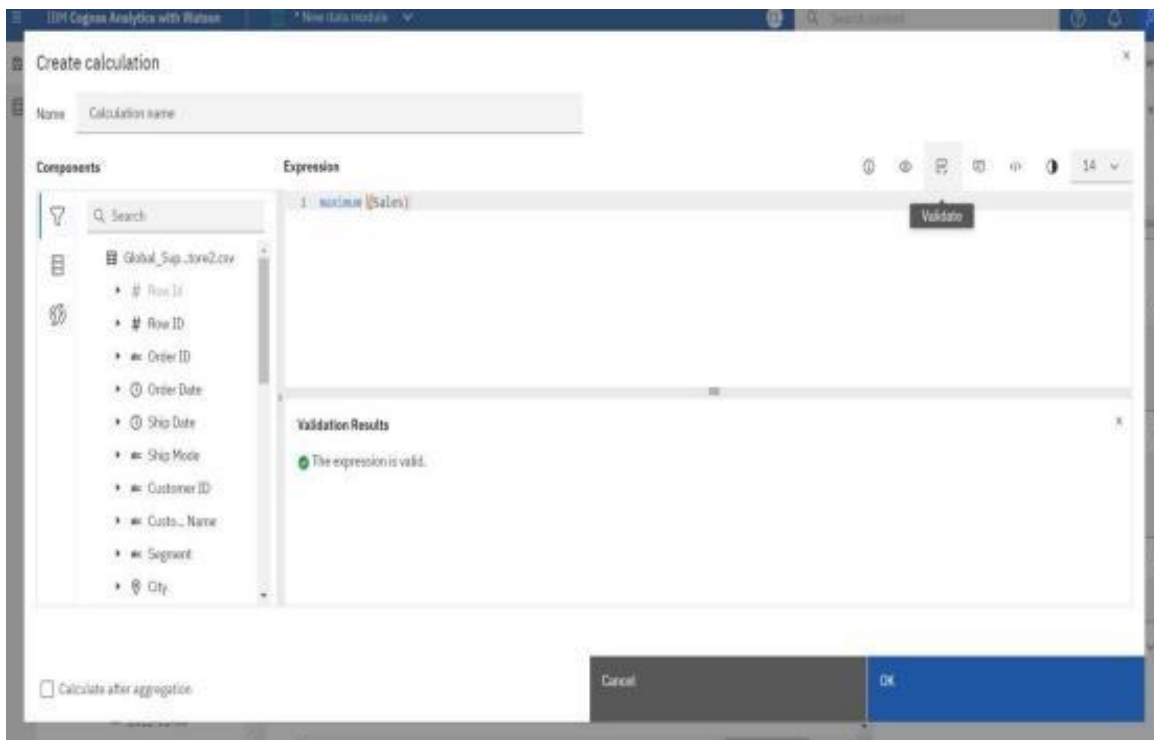
Max sales, Target sales -> maximum (Sales)

Middle range sales -> average (Sales)

Minimum sales :



Maximum Sales :



## Middle Range sales :

