

PREPARING THE DATASETS

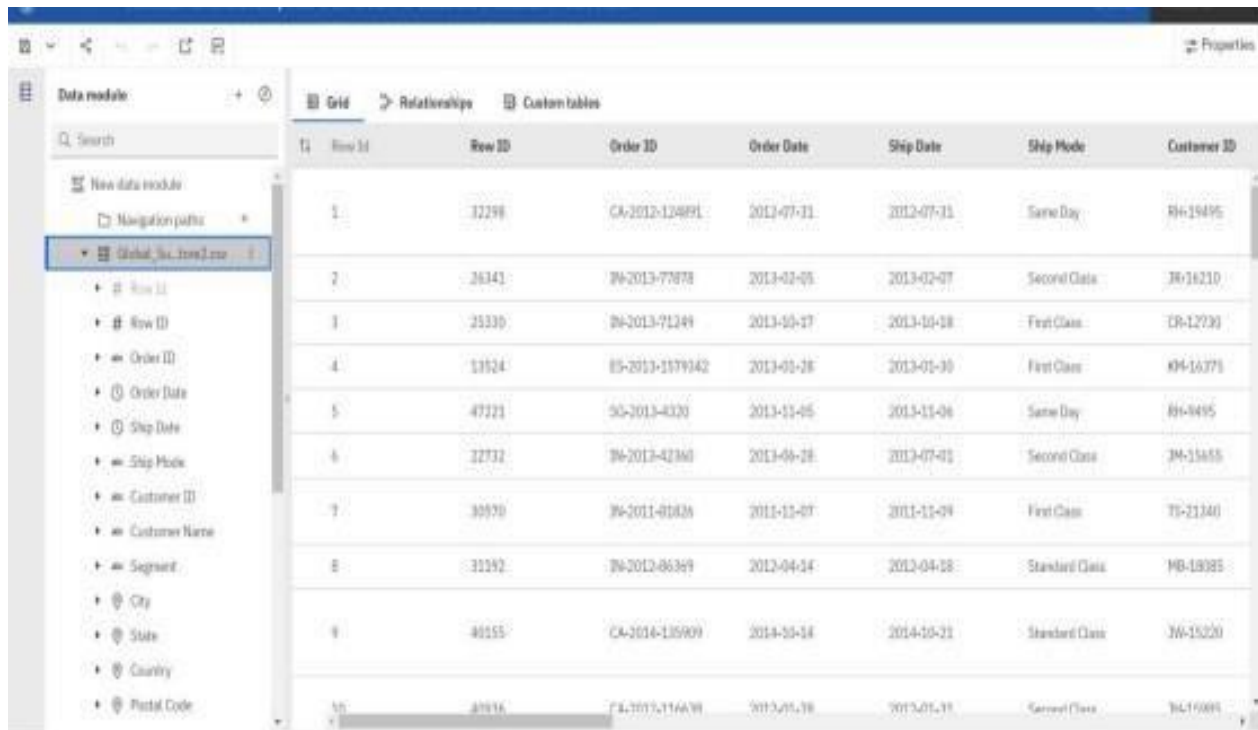
Team ID	PNT2022TMID09265
Project Name	Project – Global Sales Data Analytics

Data Preparation

The screenshot displays the IBM Cognos Analytics with Watson interface. The top navigation bar includes the title 'IBM Cognos Analytics with Watson', a search bar, and user profile icons. Below the navigation bar, there are four main action cards: 'Upload data', 'Prepare data', 'Exploration', and 'Present data'. The 'Recent' tab is selected, showing a list of uploaded files. The list has columns for Name, Location, Type, and Last Accessed. Two files are visible: 'Global_Superstore2.csv' and '50_Startups.csv', both uploaded as CSV files to 'My content'.

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

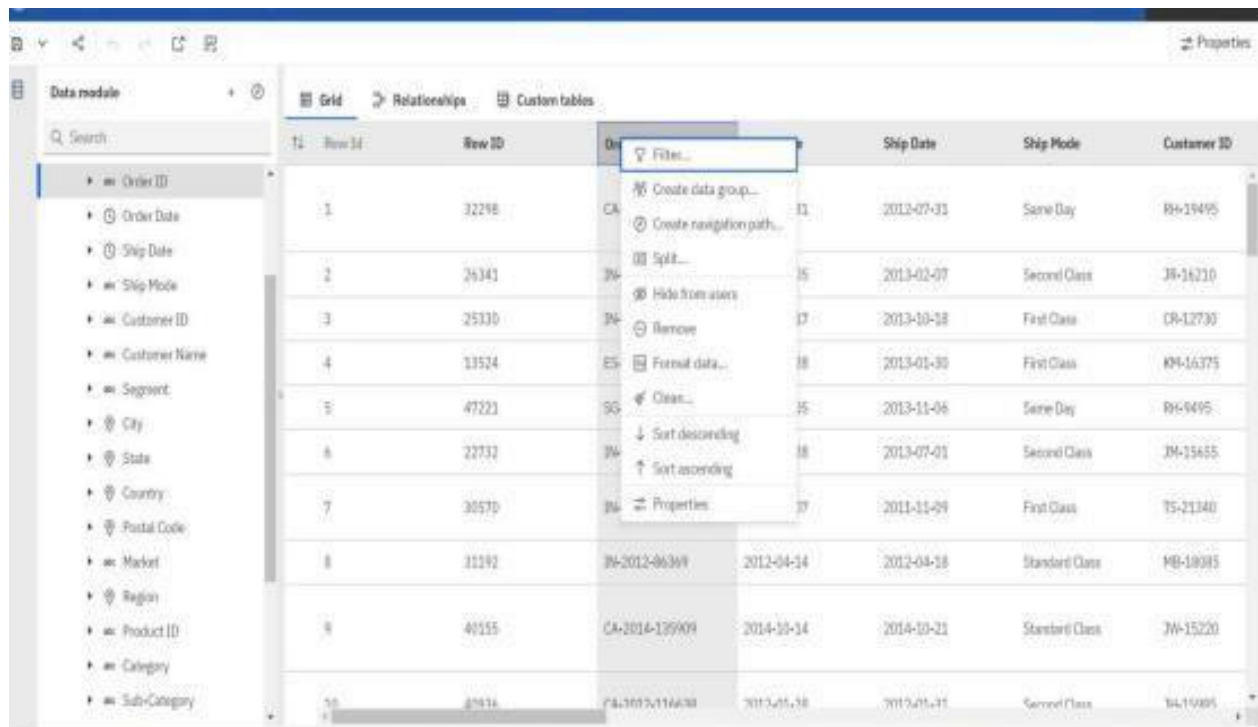
DataSet



The screenshot shows the Power BI Desktop interface with a data table. The left sidebar shows the 'Data module' with a search bar and a list of fields. The main area displays a table with the following columns: Row ID, Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. The table contains 10 rows of data.

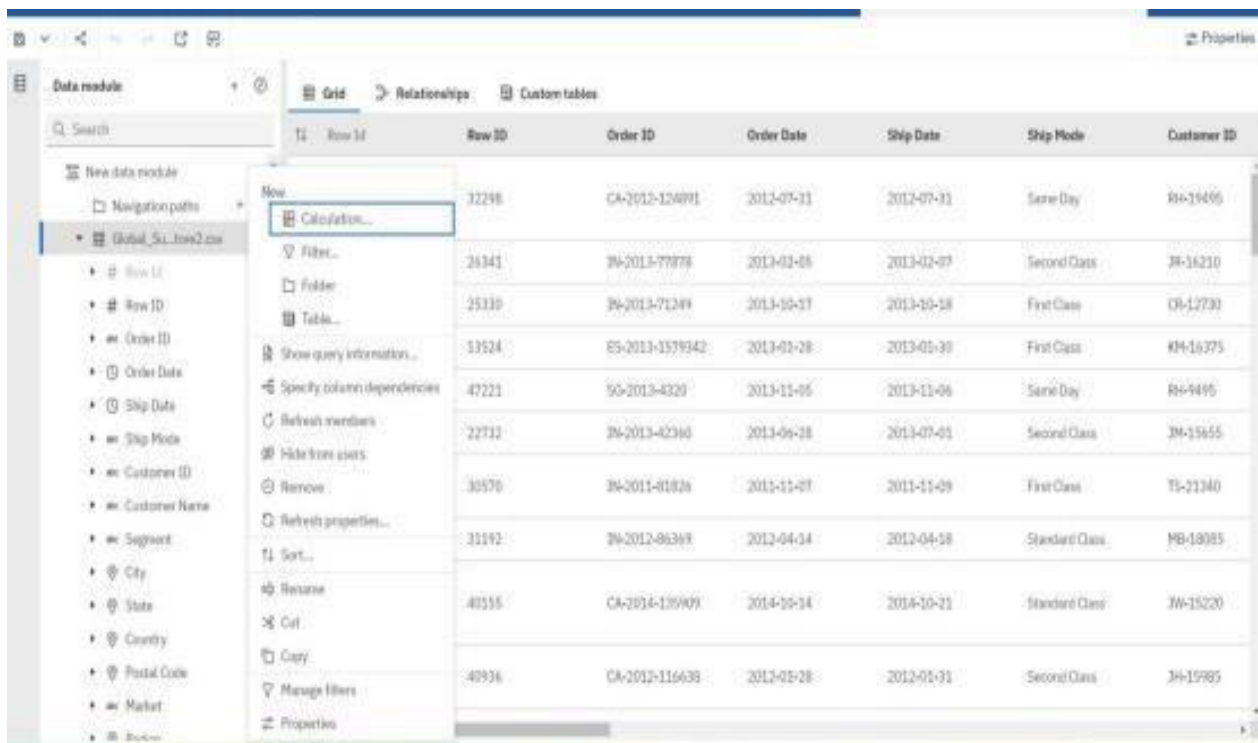
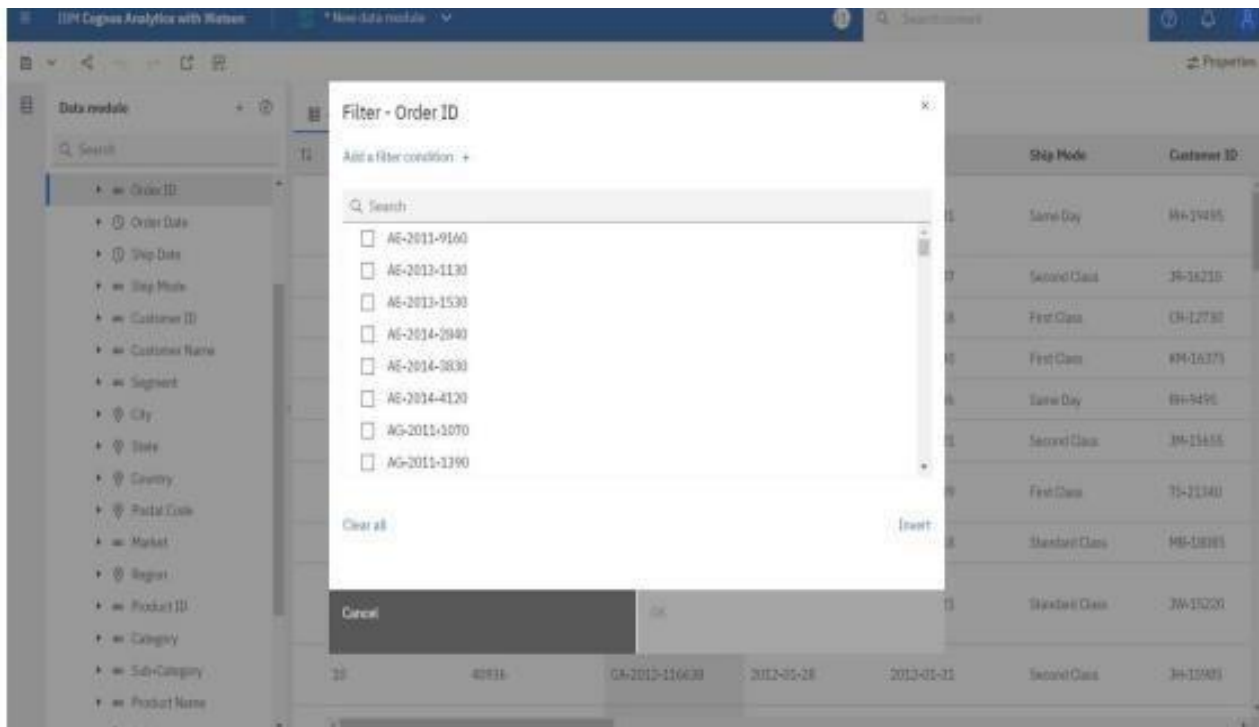
Row ID	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
1	32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495
2	26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210
3	25330	IN-2013-71249	2013-05-27	2013-10-18	First Class	CR-12730
4	13524	ES-2013-1579142	2013-01-28	2013-01-30	First Class	KH-16375
5	47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495
6	22732	IN-2013-42360	2013-06-28	2013-07-01	Second Class	JR-15655
7	30570	IN-2011-01036	2011-11-07	2011-11-09	First Class	TS-21340
8	31192	IN-2012-86369	2012-04-14	2012-04-18	Standard Class	MB-18385
9	40155	CA-2014-135909	2014-10-14	2014-10-21	Standard Class	JR-15220
10	87816	CA-2013-116608	2013-01-18	2013-01-19	Second Class	SL-19895

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



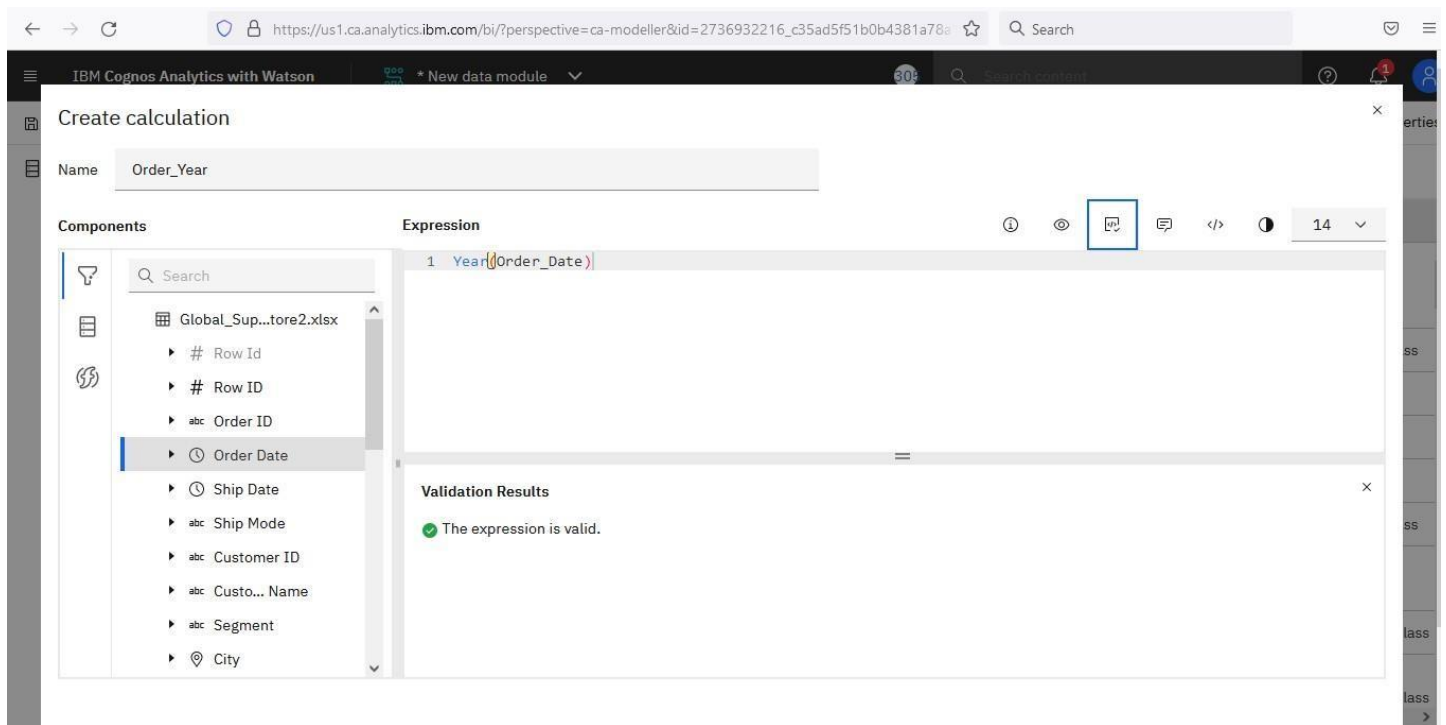
The screenshot shows the Power BI Desktop interface with a data table. The left sidebar shows the 'Data module' with a search bar and a list of fields. The main area displays a table with the following columns: Row ID, Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. A context menu is open over the 'Order ID' column, showing options like 'Filter...', 'Create data group...', 'Create navigation path...', 'Split...', 'Hide from users', 'Remove', 'Format data...', 'Clean...', 'Sort descending', 'Sort ascending', and 'Properties'.

Row ID	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
1	32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495
2	26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210
3	25330	IN-2013-71249	2013-05-27	2013-10-18	First Class	CR-12730
4	13524	ES-2013-1579142	2013-01-28	2013-01-30	First Class	KH-16375
5	47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495
6	22732	IN-2013-42360	2013-06-28	2013-07-01	Second Class	JR-15655
7	30570	IN-2011-01036	2011-11-07	2011-11-09	First Class	TS-21340
8	31192	IN-2012-86369	2012-04-14	2012-04-18	Standard Class	MB-18385
9	40155	CA-2014-135909	2014-10-14	2014-10-21	Standard Class	JR-15220
10	87816	CA-2013-116608	2013-01-18	2013-01-19	Second Class	SL-19895

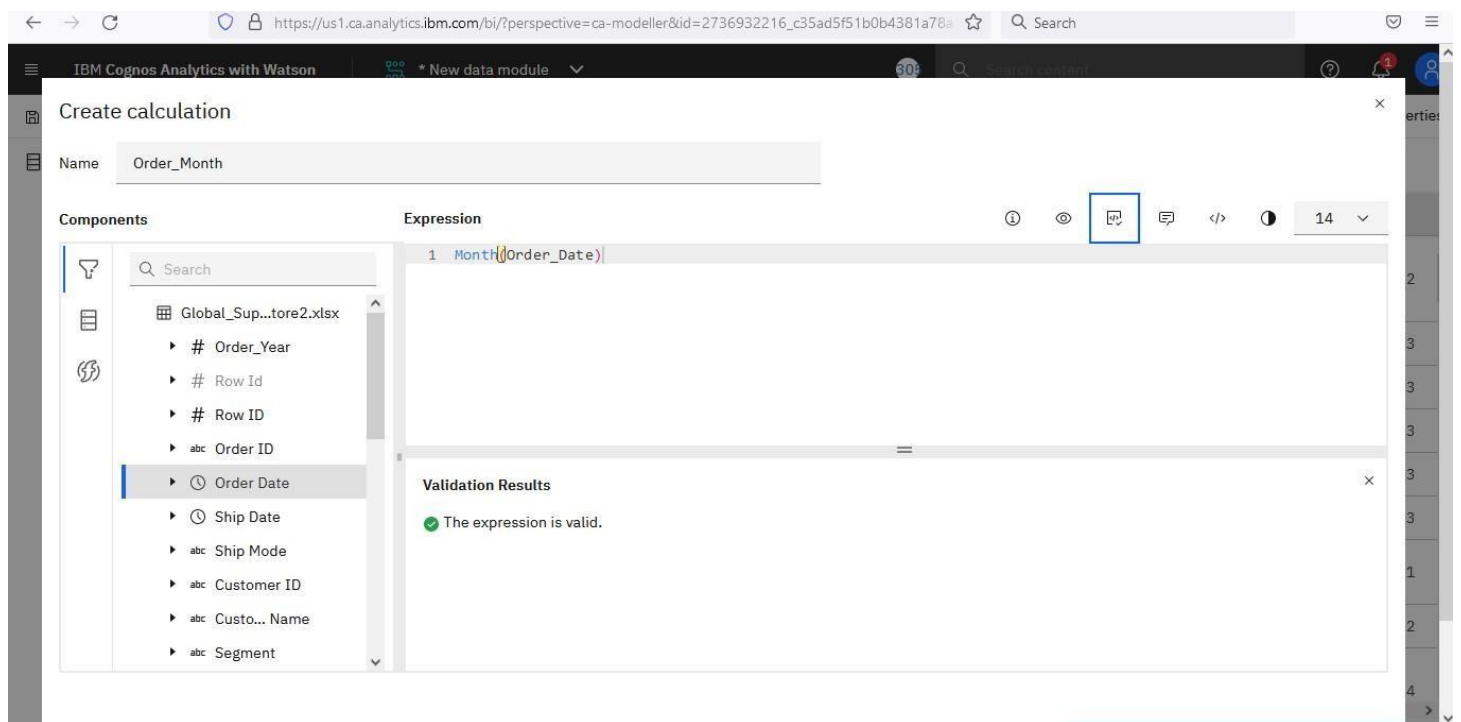


Calculations of Year, Month, Day fields and also the related Navigation path :

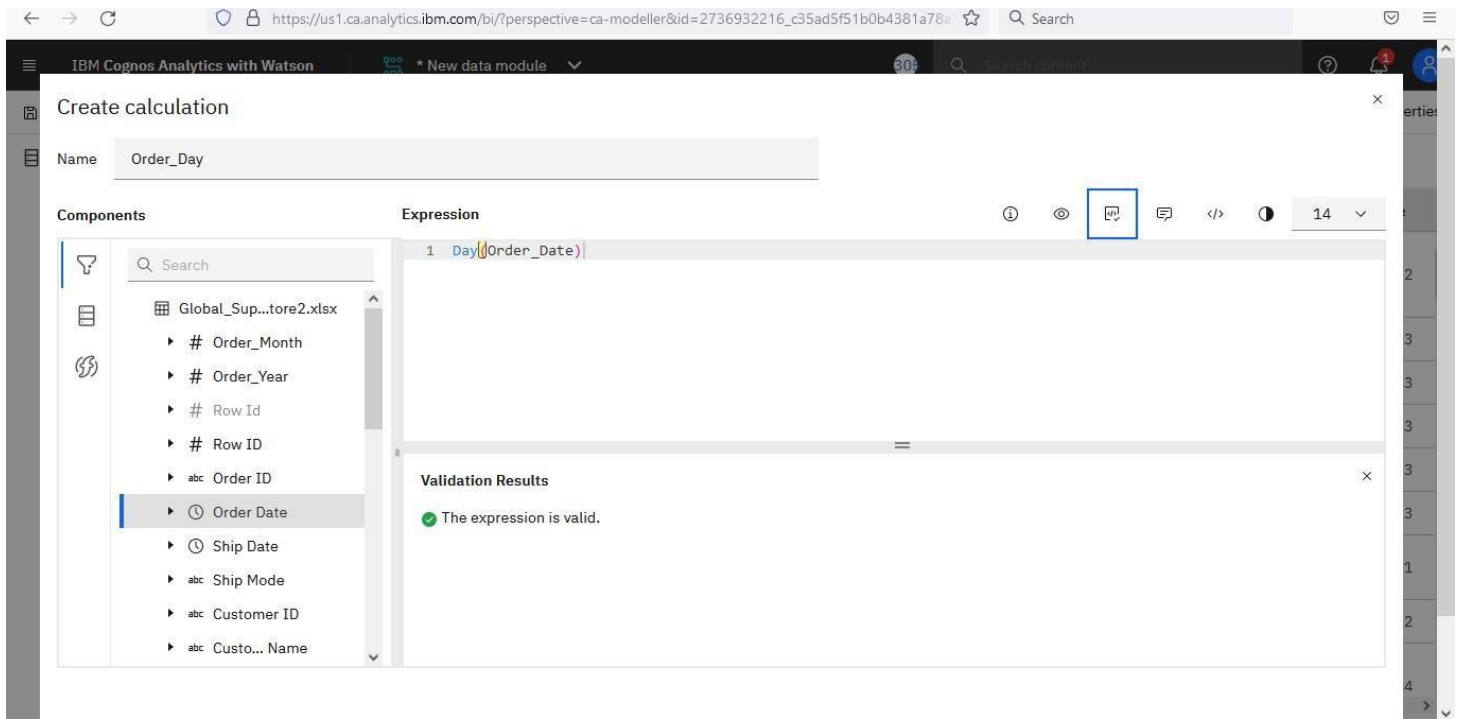
Right click -> calculation -> type expression -> validate -> Set Name as 'Order_Year'



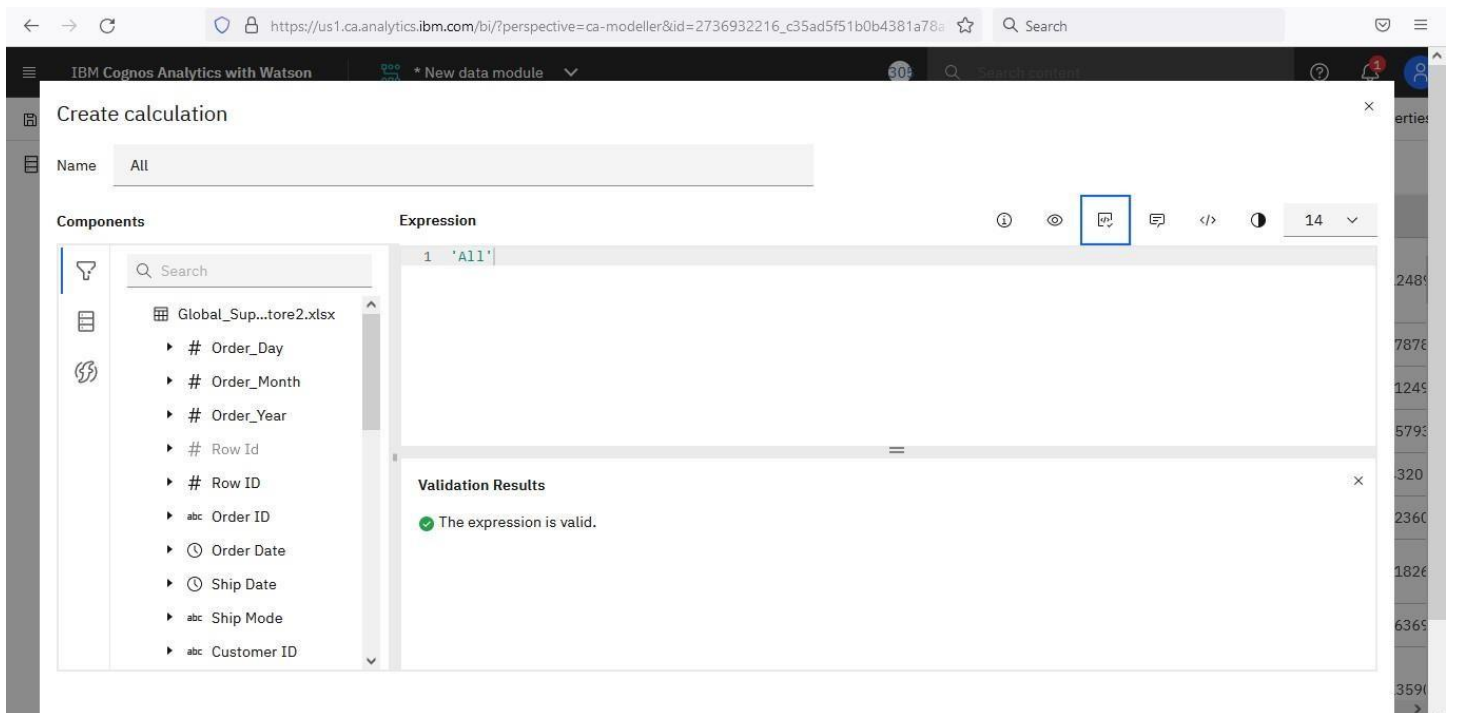
Right click -> calculation -> type expression -> validate -> Set Name as
'Order_Month'



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

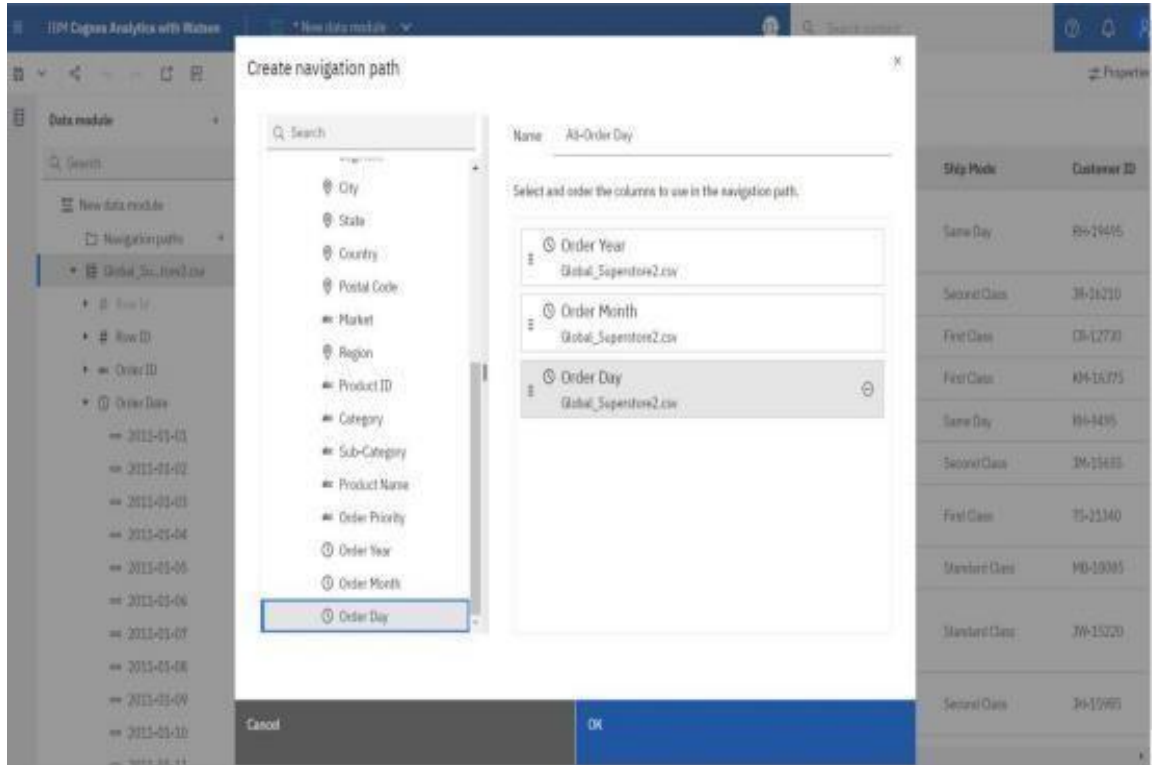
Data Type -> Integer

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

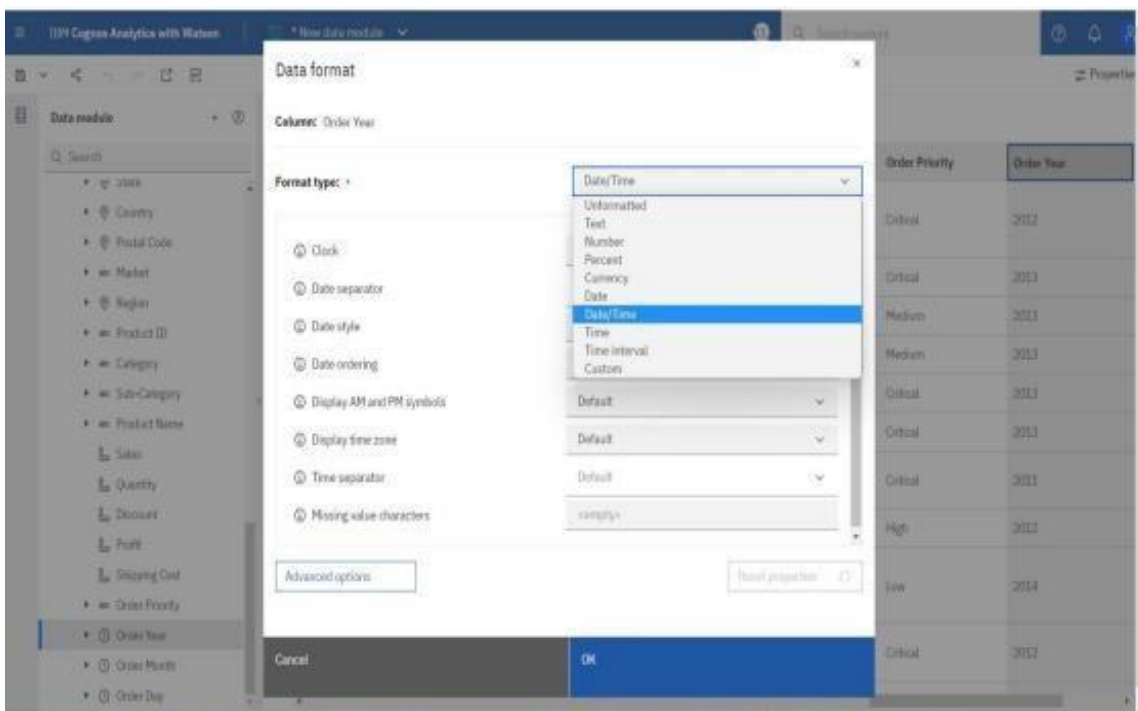
Display Options -> Show Members

This way it becomes a Numerical data Type

Right click on create navigation path



Data format> Date/Time



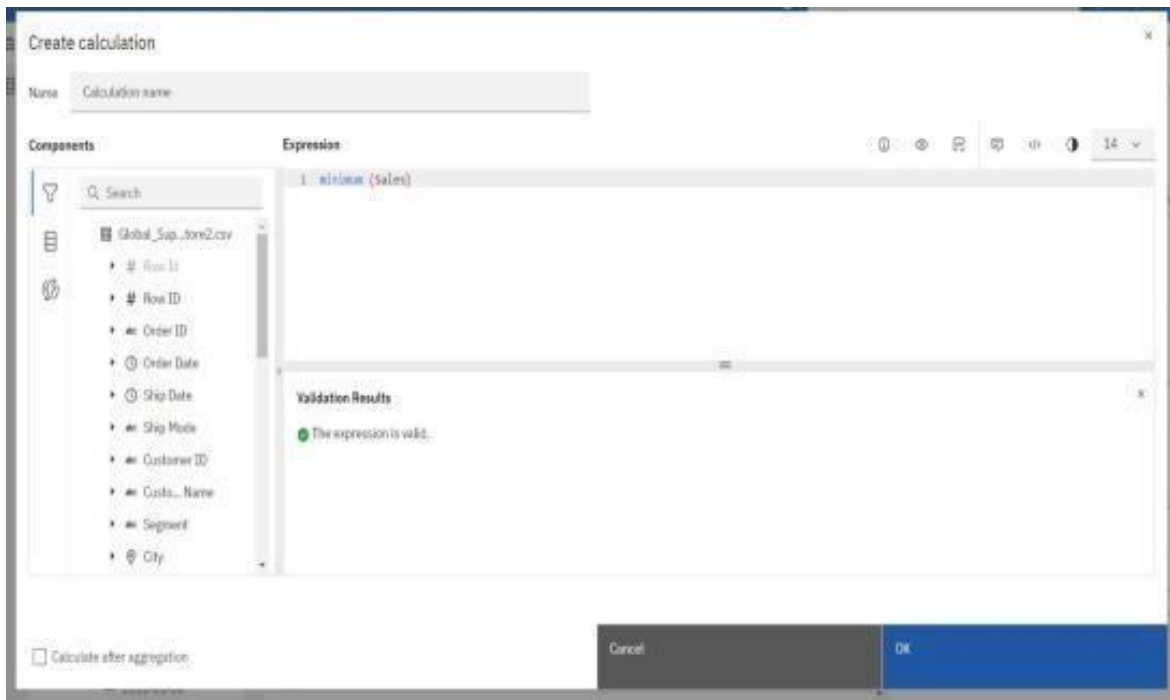
Calculations for Target sales, Min sales, Max Sales and Middle range sales

Min sales -> minimum (Sales)

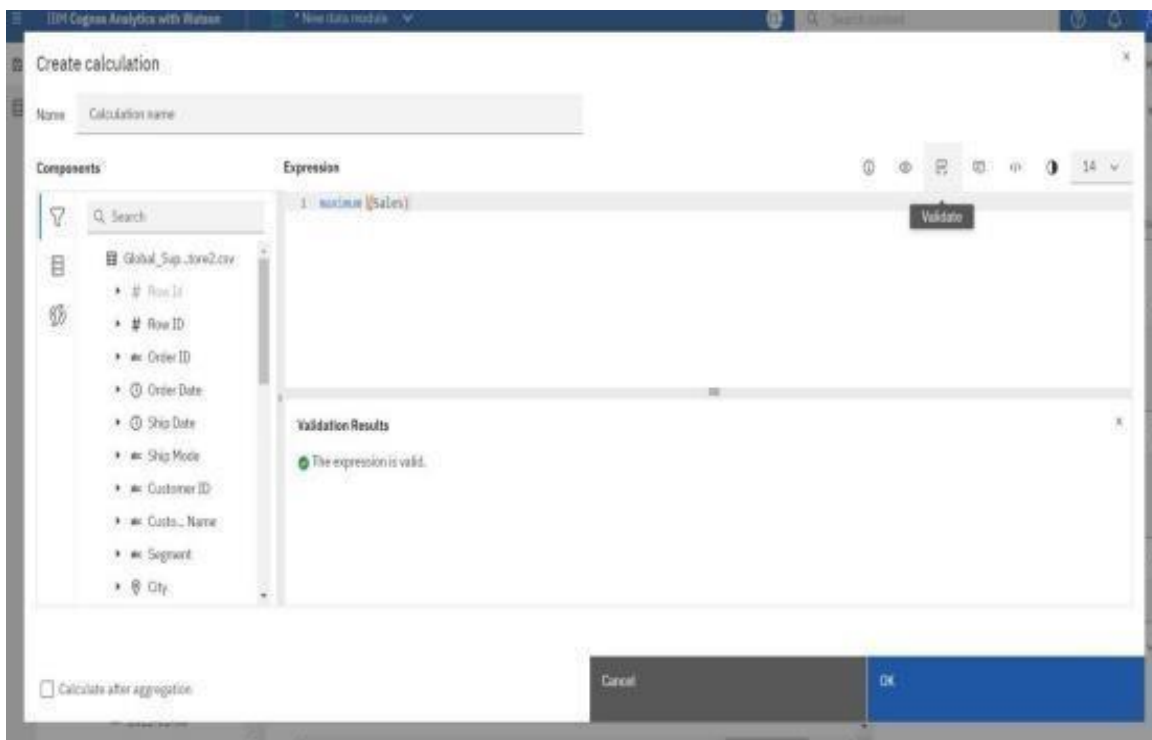
Max sales, Target sales -> maximum (Sales)

Middle range sales -> average (Sales)

Minimum sales :



Maximum Sales :



Middle Range sales :

