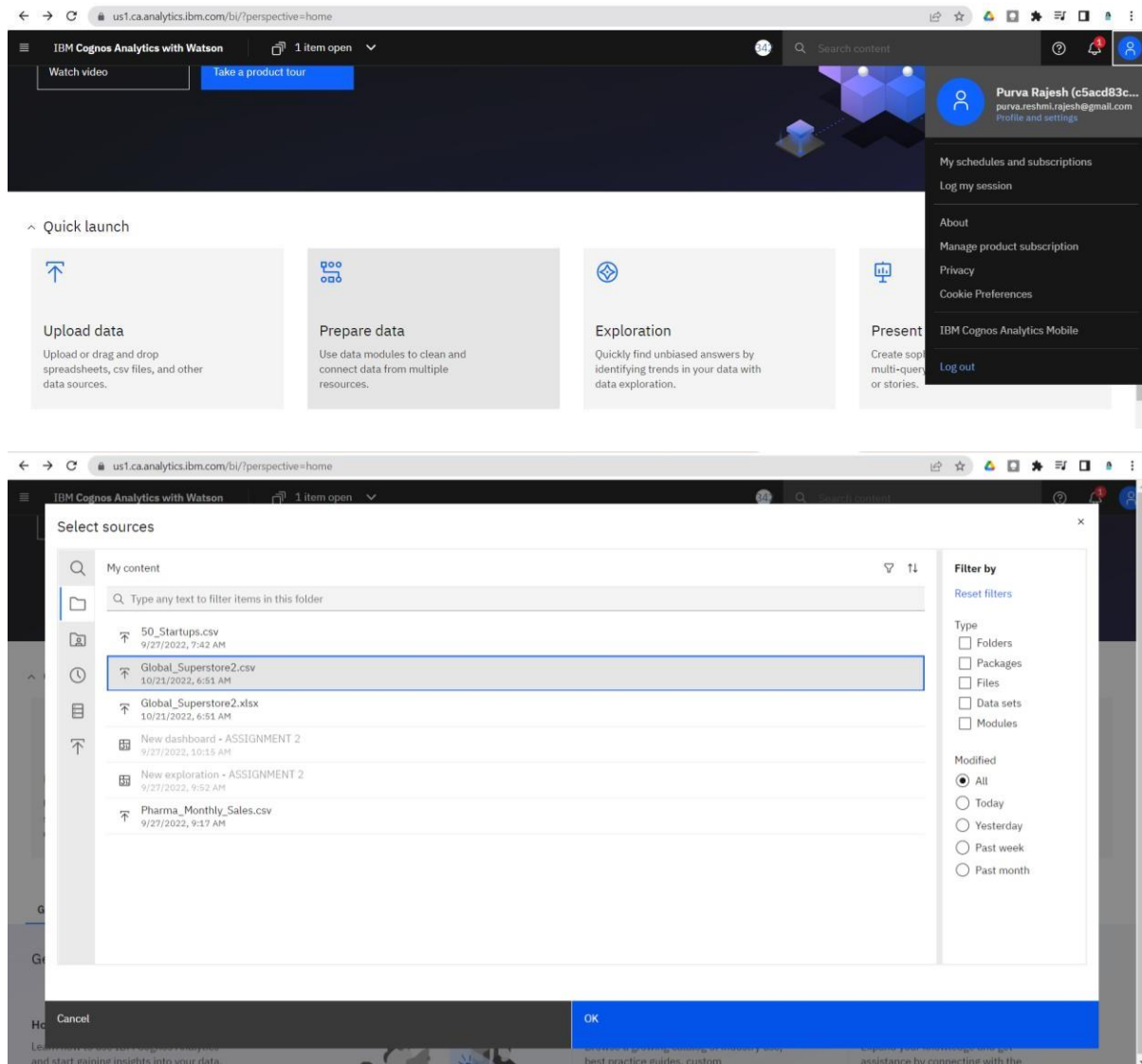


PREPARE THE DATASET

PROJECT: Global Data Sales Analytics

DATA PREPARATION

1. First understand and load the data.



2. Since there is only single file, there need not be relationships or custom tables for the data.

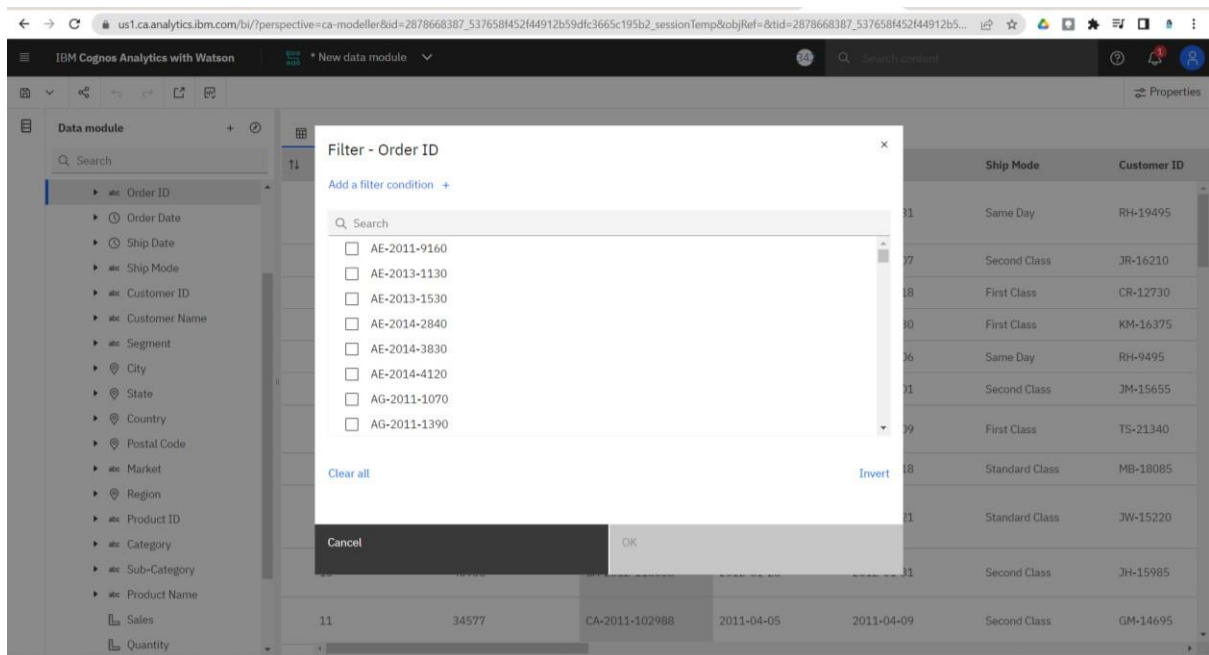
The screenshot shows the IBM Cognos Analytics with Watson interface. The top navigation bar includes the IBM logo, the text "IBM Cognos Analytics with Watson", and a search bar. The main content area is divided into a left sidebar and a central grid. The sidebar contains a "Data module" section with 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, Postal Code, Market, Region, and Product ID. The central grid displays a table with 11 rows and 8 columns: Row Id, Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. The data is as follows:

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	2013-10-17	2013-10-18	First Class	CR-12730
4	13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-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	JM-15655
7	30570	IN-2011-81826	2011-11-07	2011-11-09	First Class	TS-21340
8	31192	IN-2012-86369	2012-04-14	2012-04-18	Standard Class	MB-18085
9	40155	CA-2014-135909	2014-10-14	2014-10-21	Standard Class	JW-15220
10	40936	CA-2012-116638	2012-01-28	2012-01-31	Second Class	JH-15985
11	34577	CA-2011-102988	2011-04-05	2011-04-09	Second Class	GM-14695

3. To Prepare the data, we need to verify it is clean. There must be no null values. If there are null values, there will be empty data.

The screenshot shows the IBM Cognos Analytics with Watson interface, similar to the previous one, but with a context menu open over the "Order ID" column. The menu options are: Filter..., Create data group..., Create navigation path..., Split..., Hide from users, Remove, Format data..., Clean..., Sort descending, Sort ascending, and Properties. The data table is the same as in the previous screenshot.

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	2013-10-17	2013-10-18	First Class	CR-12730
4	13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-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	JM-15655
7	30570	IN-2011-81826	2011-11-07	2011-11-09	First Class	TS-21340
8	31192	IN-2012-86369	2012-04-14	2012-04-18	Standard Class	MB-18085
9	40155	CA-2014-135909	2014-10-14	2014-10-21	Standard Class	JW-15220
10	40936	CA-2012-116638	2012-01-28	2012-01-31	Second Class	JH-15985
11	34577	CA-2011-102988	2011-04-05	2011-04-09	Second Class	GM-14695

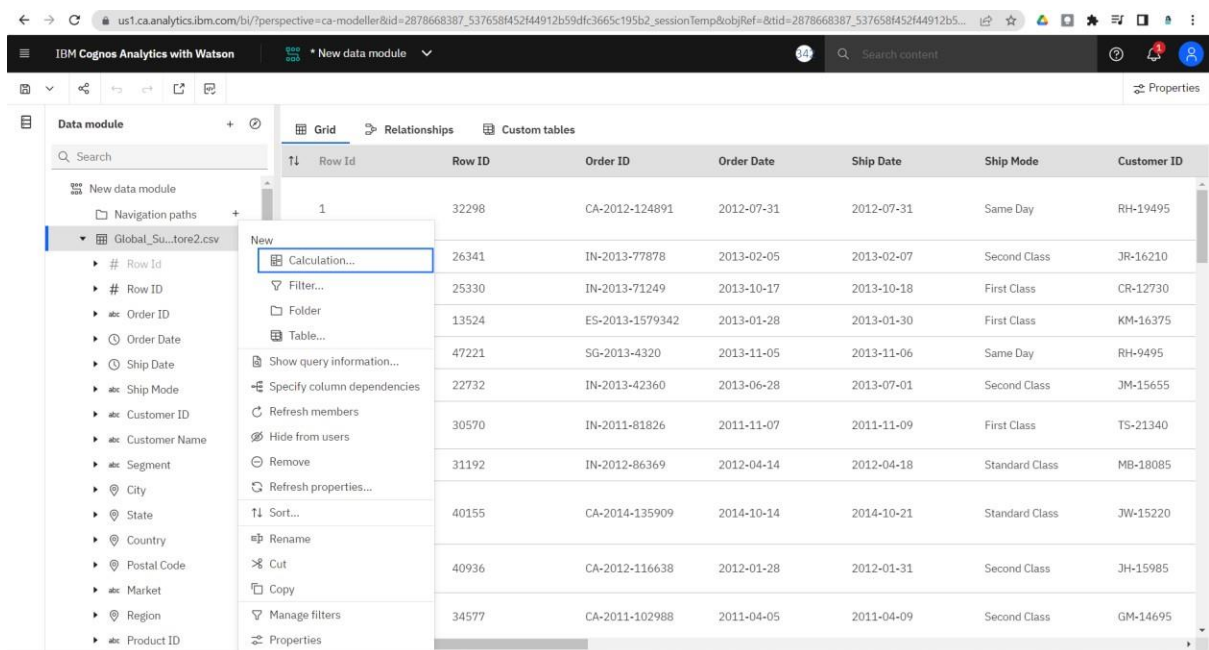


4. Once cleaned, the data is saved.

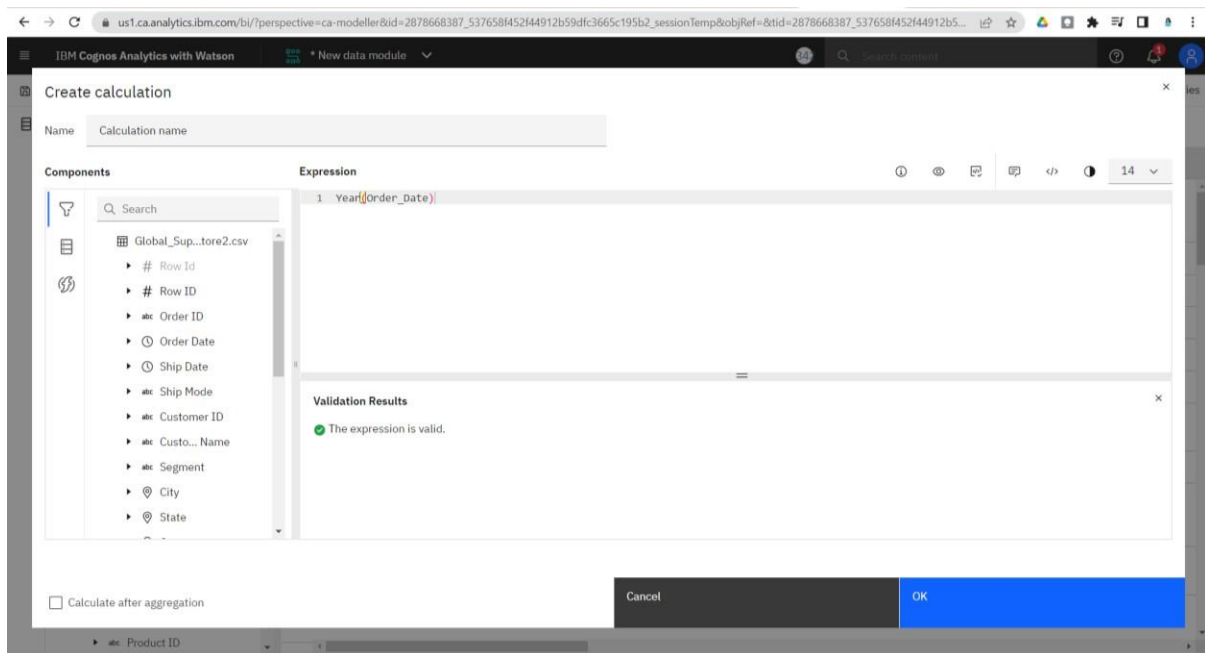
DATE CALCULATIONS AND NAVIGATION PATHS

Once you load the data, we need to Prepare the data.

- Prepare Calculations of Year, Month, Day fields and also the related Navigation path**
- Create a Few more Calculations – Target Sales, Min Sales, Max Sales, Middle Range Sales.**



1. Right click> calculation> type expression> validate> Set Name as 'Order Year'



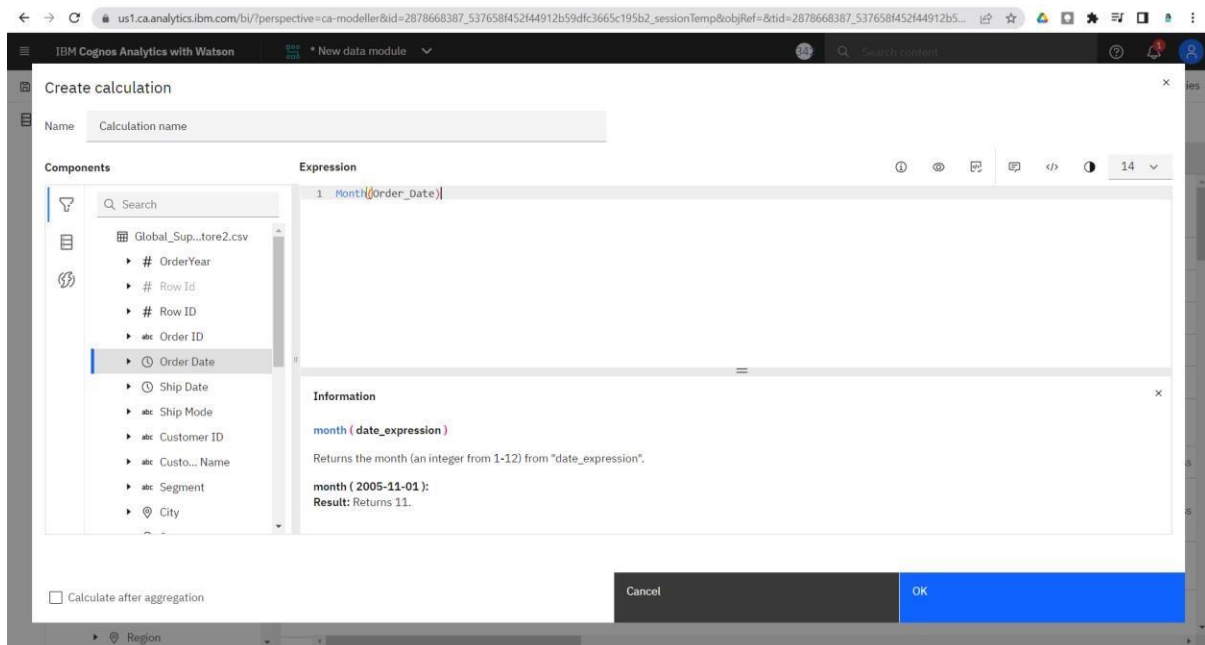
2. Similarly, 'OrderMonth', 'OrderDay' and 'All'

OrderYear -> Year (Order_Date)

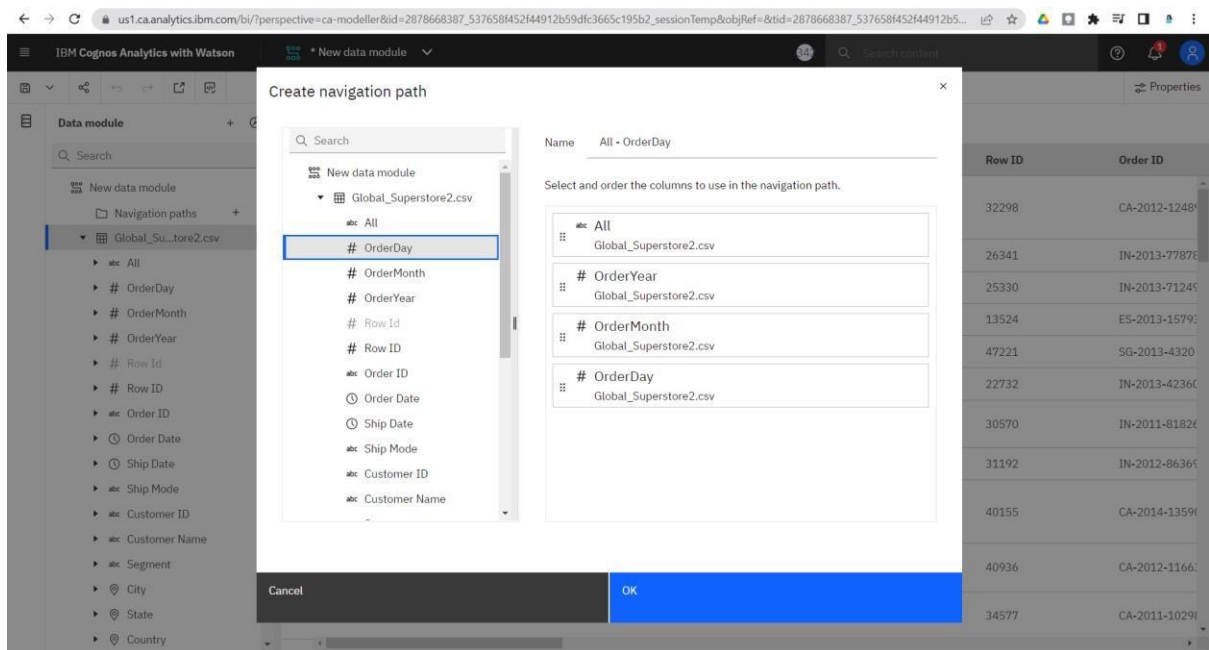
OrderMonth -> Month (Order_Date)

OrderDay -> Day (Order_Date)

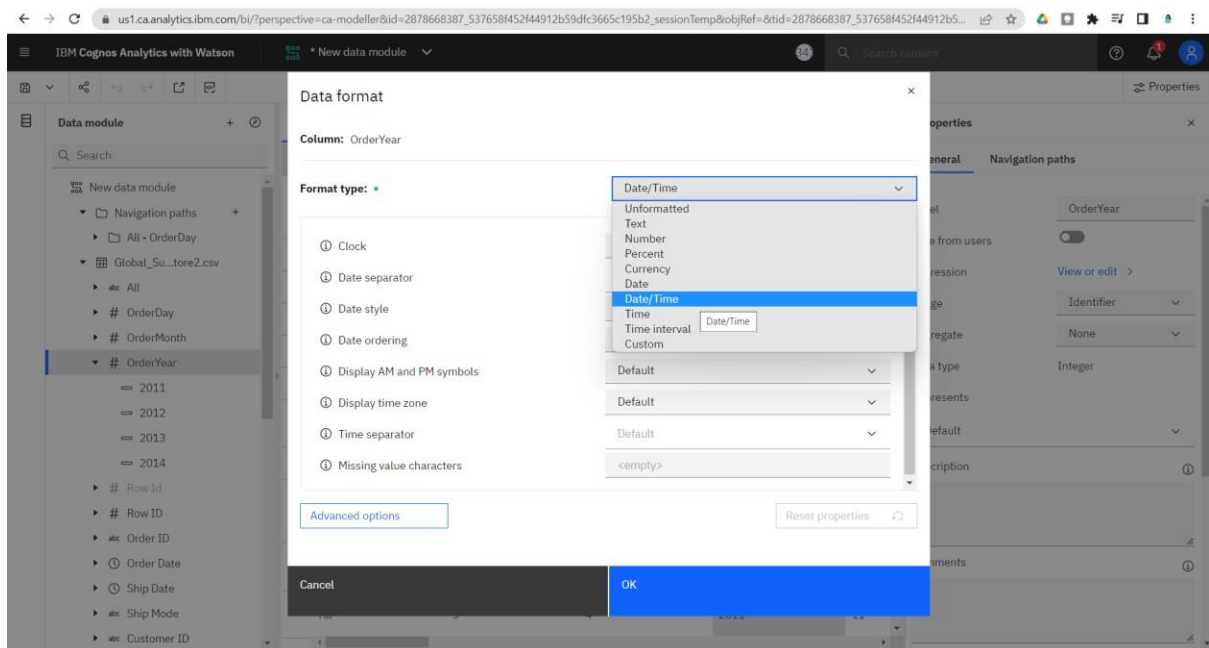
All -> All



3. Right click on create navigation path



4. Data format> Date/Time



5. Properties> Usage-> Attribute

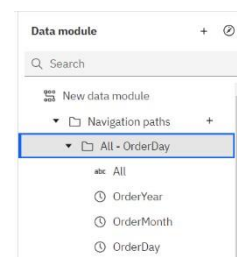
Aggregate-> Count Distinct

Data Type-> Integer

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

Display Options-> Show Members

This way it becomes a Numerical data Type



	OrderDay	OrderMonth	OrderYear	Row Id
All	14	10	2014	9
All	28	1	2012	10
All	5	4	2011	11
All	19	4	2012	12
All	27	12	2011	13
All	13	11	2012	14
All	6	6	2013	15
All	31	7	2014	16
All	3	11	2014	17
All	8	9	2014	18
All	31	1	2014	19
All	5	12	2014	20

6. Hence calculations for day, month and year are completed.

Once you load the data, we need to Prepare the data.

- Prepare Calculations of Year, Month, Day fields and also the related Navigation path
- Create a Few more Calculations – Target Sales, Min Sales, Max Sales, Middle Range Sales.

7. Similarly, calculations for Target sales, Min sales, Max Sales and Middle range sales are done
 Min sales-> minimum (Sales)
 Max sales, Target sales-> maximum (Sales)
 Middle range sales-> average (Sales)

Create calculation

Name: MinSales

Expression: 1 minimum (Sales)

Components:

- Global_Sup...tore2.csv
 - All
 - OrderDay
 - OrderMonth
 - OrderYear
 - Row Id
 - Row ID
 - Order ID
 - Order Date
 - Ship Date
 - Ship Mode
 - Customer ID

Validation Results: The expression is valid.

☐ Calculate after aggregation

Cancel OK

8. Hence the calculations for Target sales, Min sales, Max Sales and Middle range sales are completed.