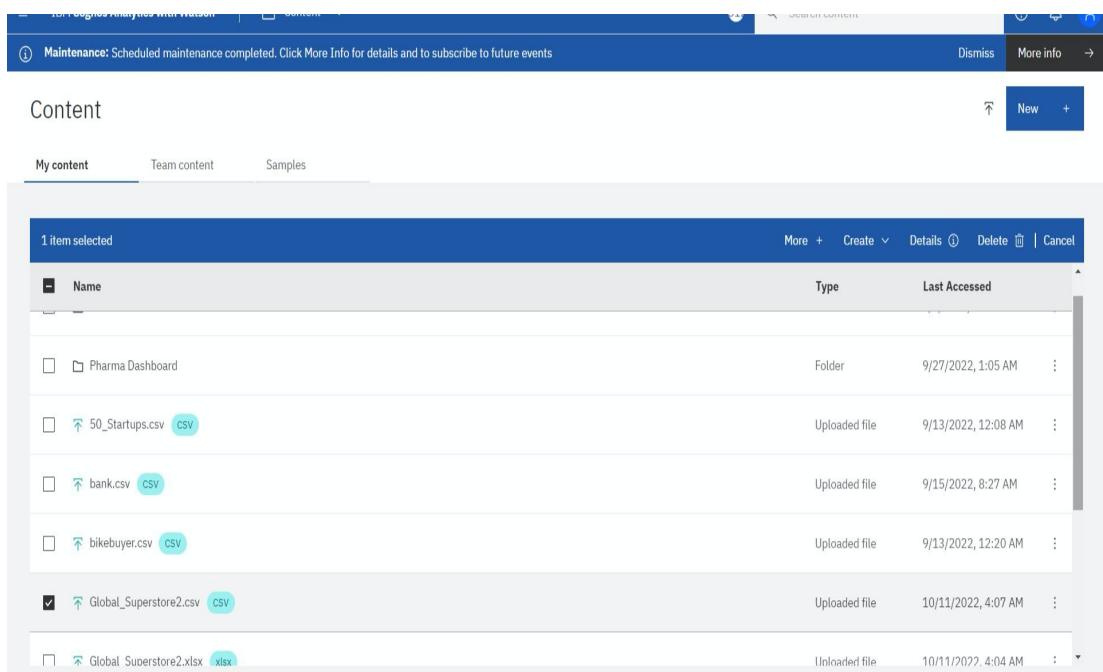


PREPARE THE DATASETS

Team ID	PNT2022TMID13158
Project Name	Global Sales Data Analytics

DATA PREPARATION

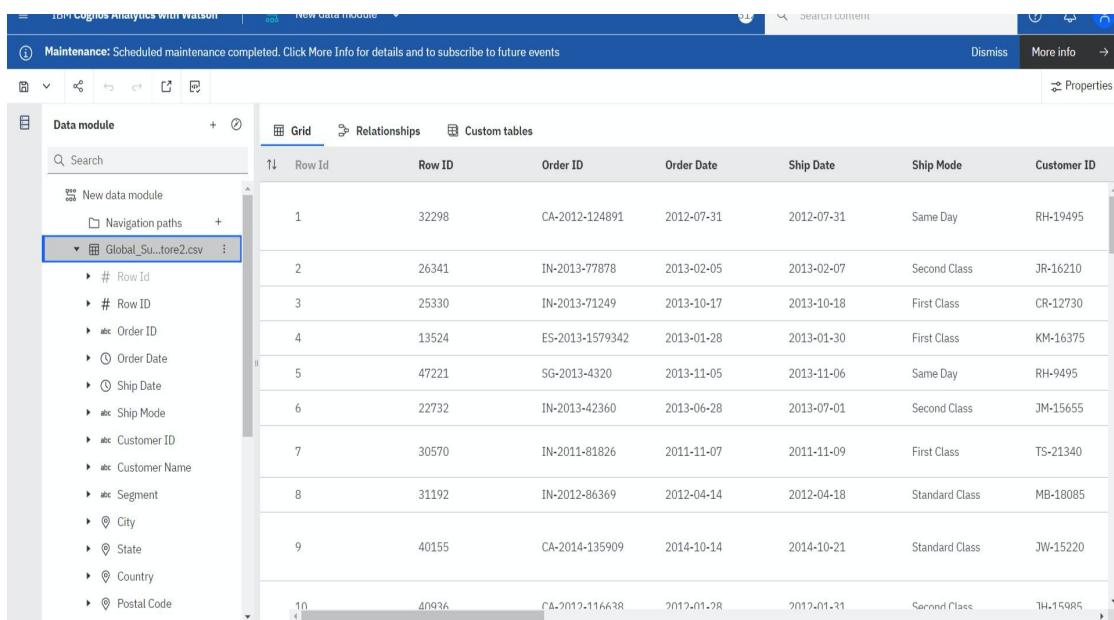
1. First understand and load the data.



The screenshot shows a Google Analytics interface with a blue header bar. The bar contains the text "101 Google Analytics with Watson" and "Content". Below the bar, a message says "Maintenance: Scheduled maintenance completed. Click More Info for details and to subscribe to future events". There are "Dismiss", "More info", and a right arrow buttons. The main area is titled "Content" and has tabs for "My content", "Team content", and "Samples". A sub-header "1 item selected" is visible. The list view shows the following items:

Name	Type	Last Accessed
Pharma Dashboard	Folder	9/27/2022, 1:05 AM
50_Startups.csv	Uploaded file	9/13/2022, 12:08 AM
bank.csv	Uploaded file	9/15/2022, 8:27 AM
bikebuyer.csv	Uploaded file	9/13/2022, 12:20 AM
Global_Superstore2.csv	Uploaded file	10/11/2022, 4:07 AM
Global_Superstore2.xlsx	Unloaded file	10/11/2022, 4:04 AM

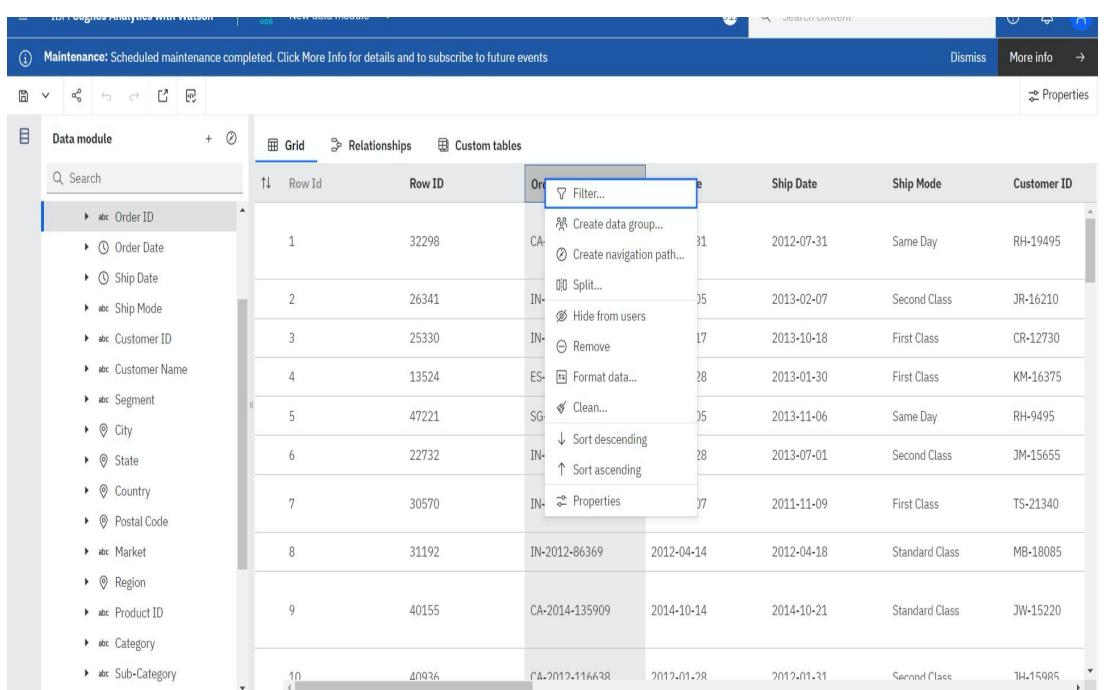
2. Since there is only single file, there need not be relationships or custom tables for the data.



The screenshot shows the IBM Cognos Analytics - Data module interface. On the left, the 'Data module' pane lists a single file named 'Global_Superstore2.csv'. The main area displays a grid of data with columns: Row Id, Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. The data consists of 10 rows of shipping information. A context menu is open over the first row's Order ID column.

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	40926	CA-2012-116638	2012-01-28	2012-01-31	Second Class	JH-15085

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 same IBM Cognos Analytics - Data module interface. The context menu is now open over the 'Order ID' column header of the grid. The menu options include: Filter..., Create data group..., Create navigation path..., Split..., Hide from users, Remove, Format data..., Clean..., Sort descending, Sort ascending, Properties, and Properties.

Filter - Order ID

Add a filter condition +

Search

- AE-2011-9160
- AE-2013-1130
- AE-2013-1530
- AE-2014-2840
- AE-2014-3830
- AE-2014-4120
- AG-2011-1070
- AG-2011-1390

Clear all Invert

3. Once cleaned, the data is saved.

DATE CALCULATIONS AND NAVIGATION PATHS

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

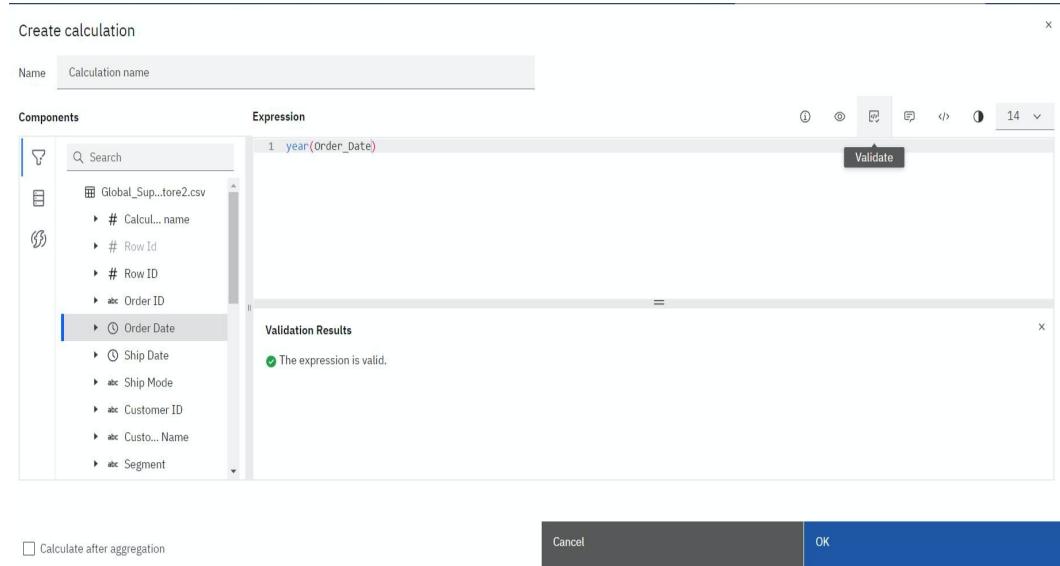
a. Prepare Calculations of Year, Month, Day fields and also the related Navigation path

b. Create a Few more Calculations – Target Sales, Min Sales, Max Sales, Middle Range Sales.

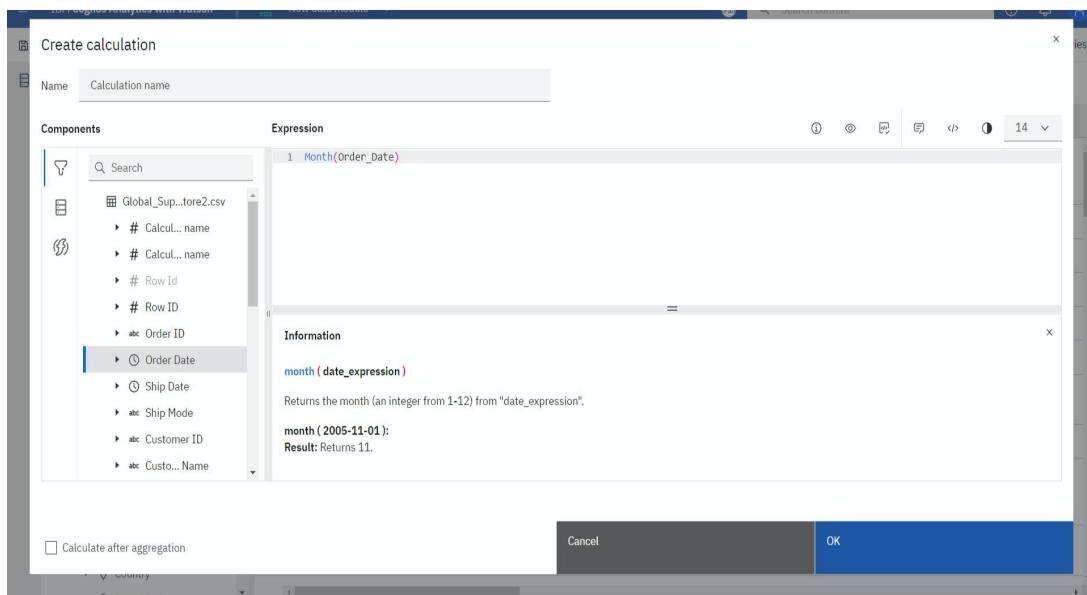
The screenshot shows the Power BI Data module interface. On the left, there's a sidebar with a tree view of data modules, navigation paths, and other items. A context menu is open over a 'Global_Su...tore2.csv' navigation path item. The menu options include 'New', 'Calculation...', 'Filter...', 'Folder', 'Table...', 'Show query information...', 'Specify column dependencies', 'Refresh members', 'Hide from users', 'Remove', 'Refresh properties...', 'Sort...', 'Rename', 'Cut', 'Copy', 'Manage filters', and 'Properties'. The 'Calculation...' option is highlighted with a blue border. The main grid area displays a table with columns: Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. The data consists of several rows with various values.

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495
26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210
25330	IN-2013-71249	2013-10-17	2013-10-18	First Class	CR-12730
13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-16375
47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495
22732	IN-2013-42360	2013-06-28	2013-07-01	Second Class	JM-15655
30570	IN-2011-81826	2011-11-07	2011-11-09	First Class	TS-21340
31192	IN-2012-86369	2012-04-14	2012-04-18	Standard Class	MB-18085
40155	CA-2014-135909	2014-10-14	2014-10-21	Standard Class	JW-15220
40936	CA-2012-116638	2012-01-28	2012-01-31	Second Class	JH-15985

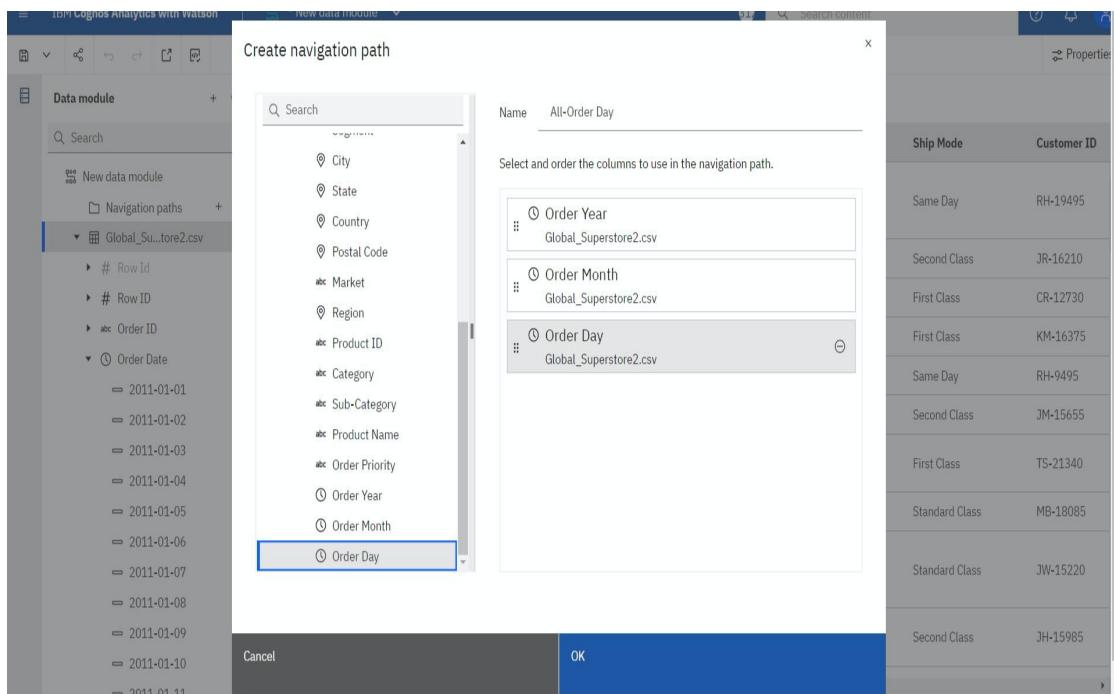
1. Right click > calculation > type expression > validate > Set Name as 'Order Year'



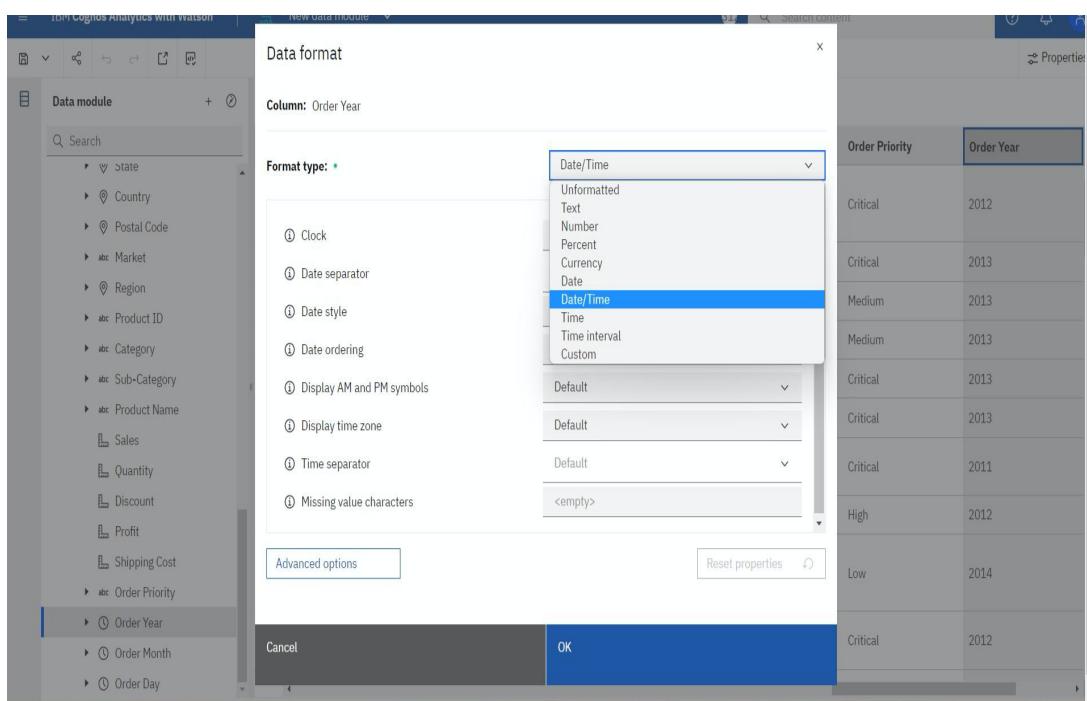
2. Similarly, 'OrderMonth', 'OrderDay' and 'All'
 OrderYear -> Year (Order_Date)
 OrderMonth -> Month (Order_Date)
 OrderDay -> Day (Order_Date)
 All -> All



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

Order	Label	Hide from users
7	Order Year	<input checked="" type="checkbox"/>
2		<input type="checkbox"/>
1		<input type="checkbox"/>
1		<input type="checkbox"/>
1		<input type="checkbox"/>
6		<input type="checkbox"/>
1		<input type="checkbox"/>
4		<input type="checkbox"/>
1		<input type="checkbox"/>
1		<input type="checkbox"/>
1		<input type="checkbox"/>

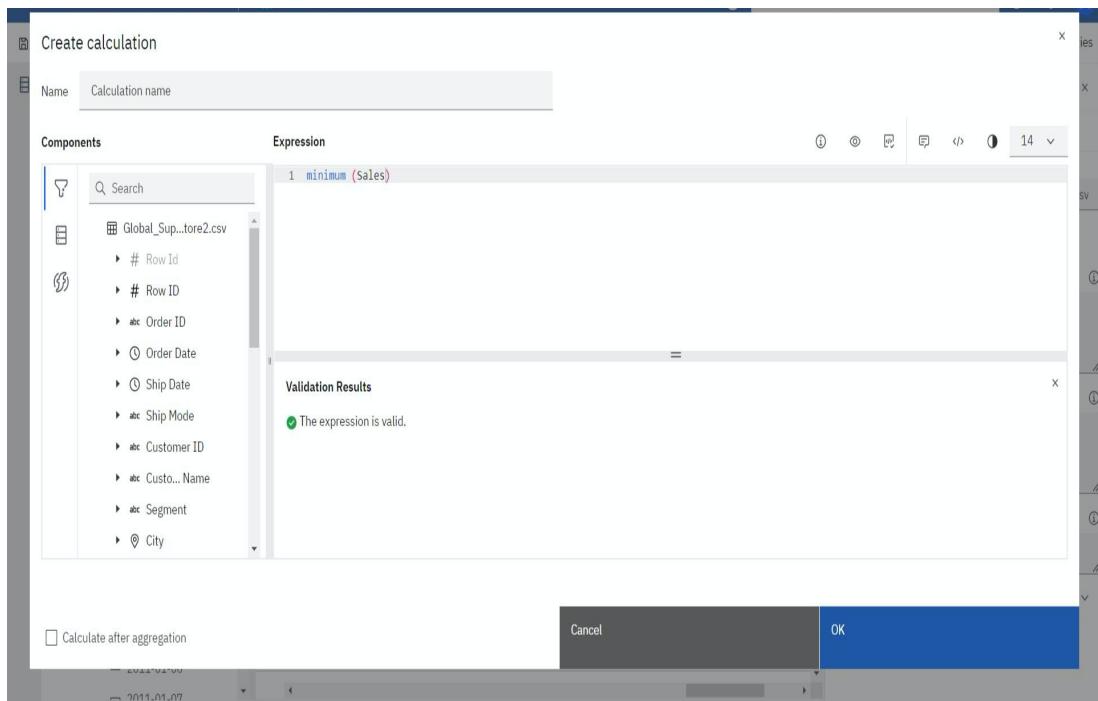
6. Hence calculations for day, month and year are completed.

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

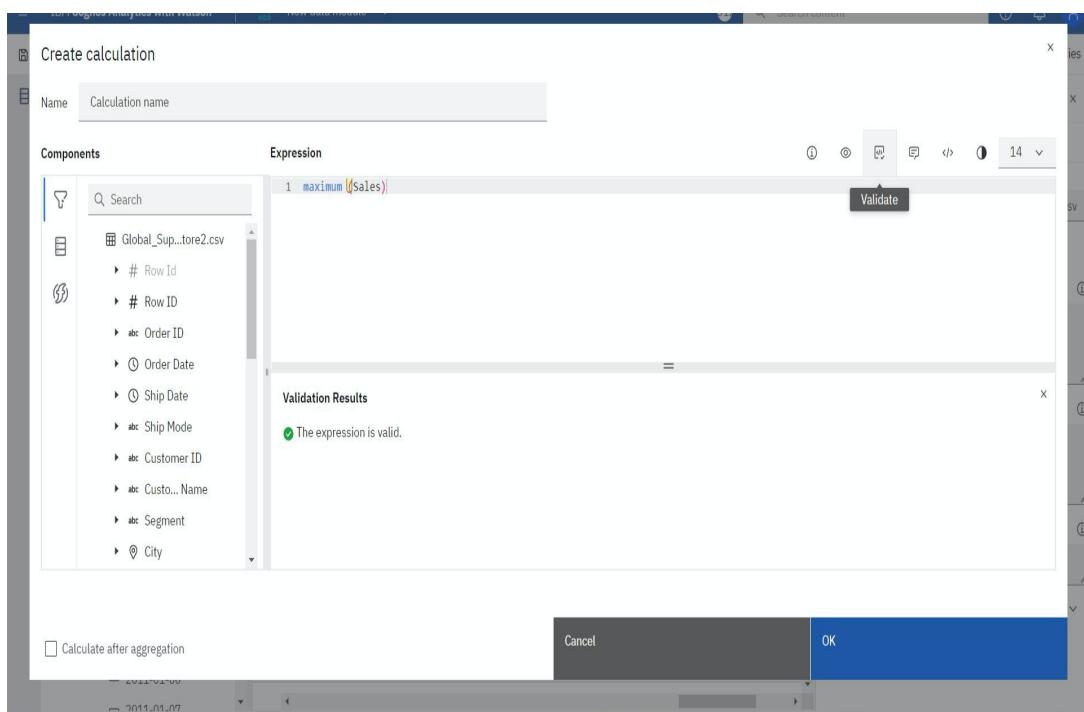
- Prepare Calculations of Year, Month, Day fields and also the related Navigation path
- Create a Few more Calculations – Target Sales, Min Sales, Max Sales, Middle Range Sales.**

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)

Minimum(Sales)



Maximum(Sales)



Middle Range (Sales)

