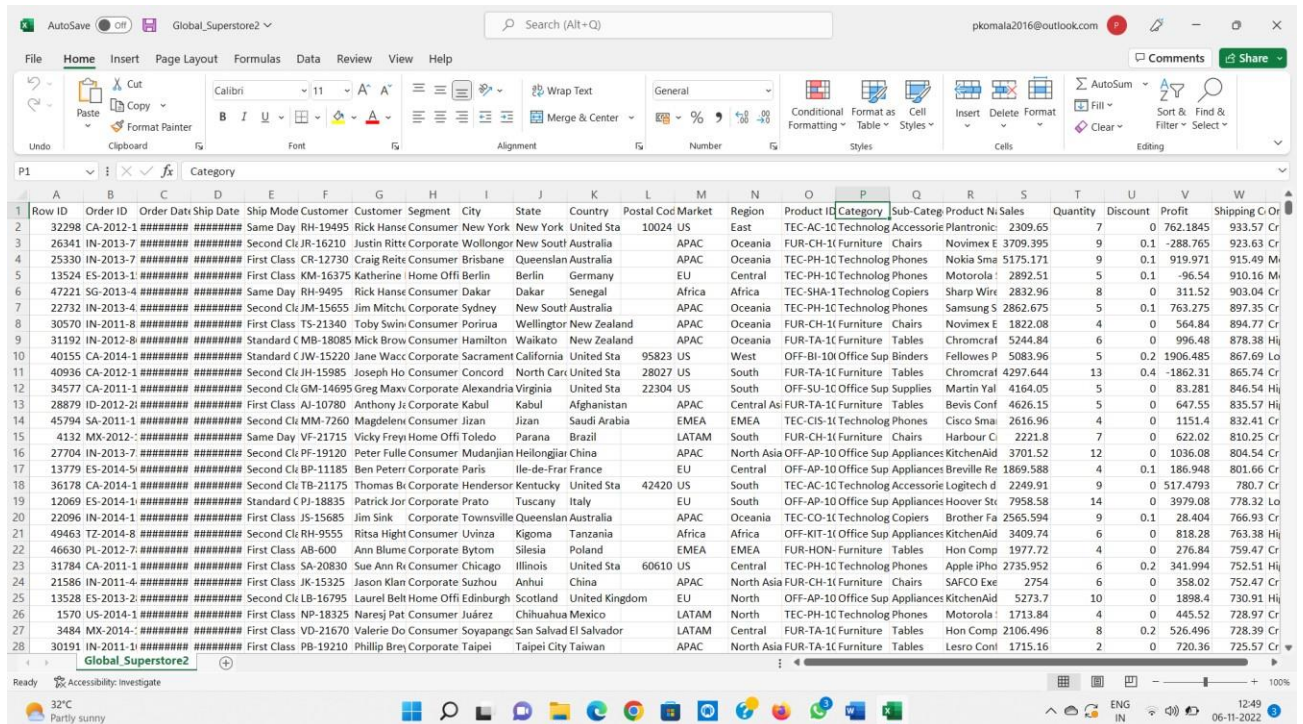


# EXPLORATORY DATA ANALYSIS

Team ID	PNT2022TMID13846
Project Name	Global Sales Data Analytics

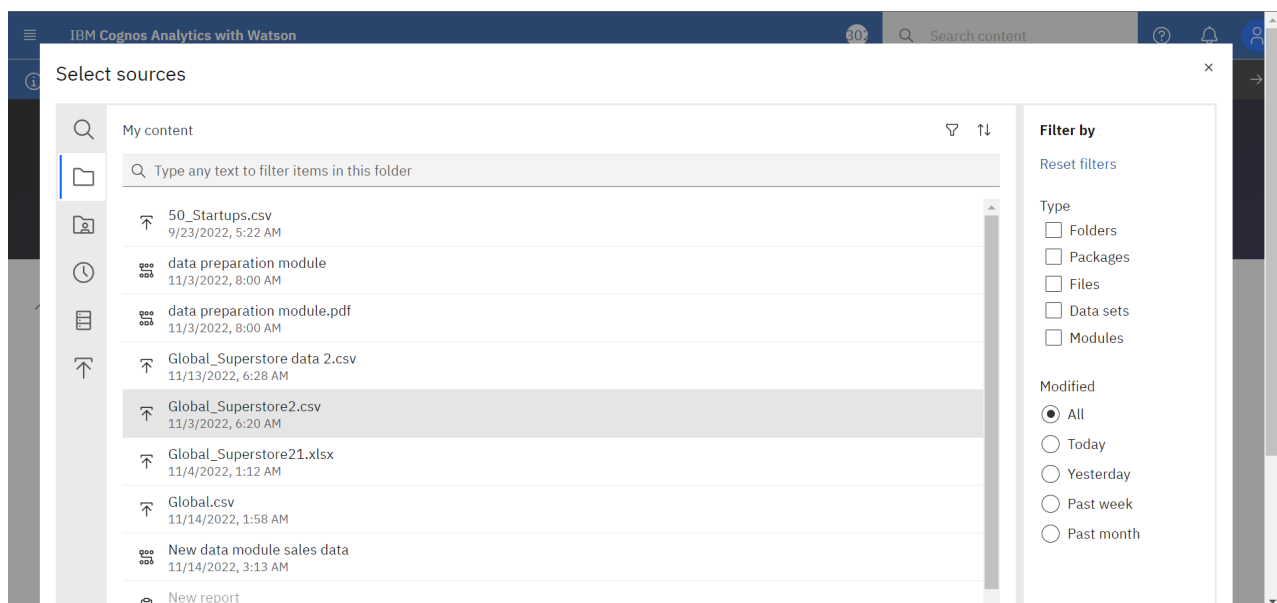
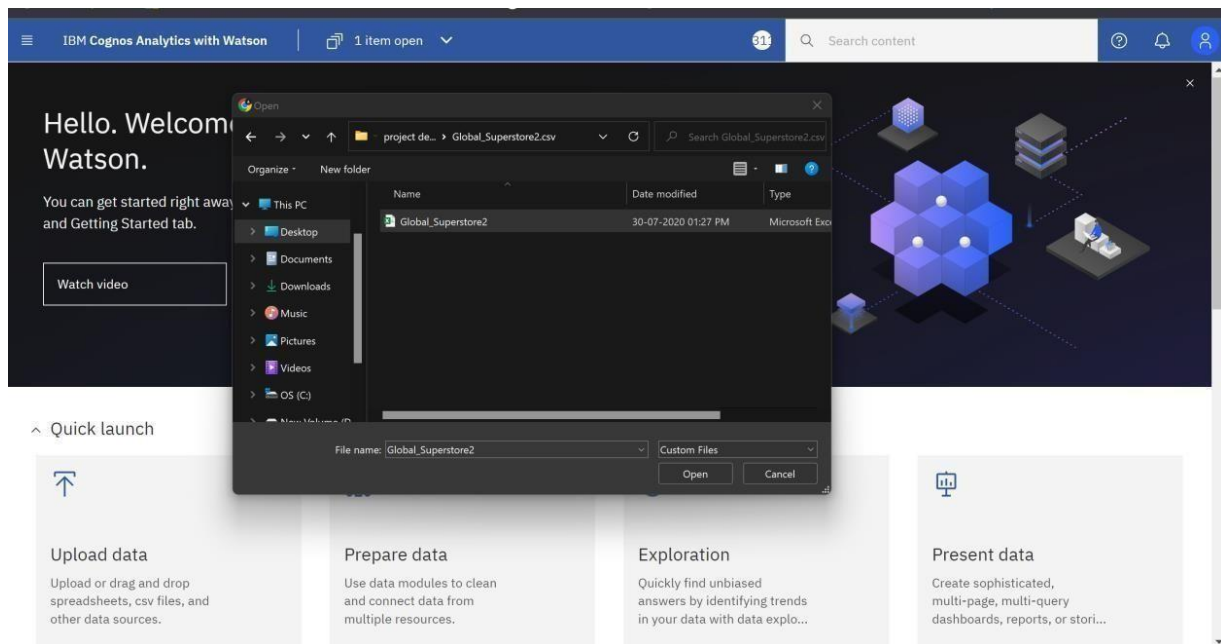
## Dataset:



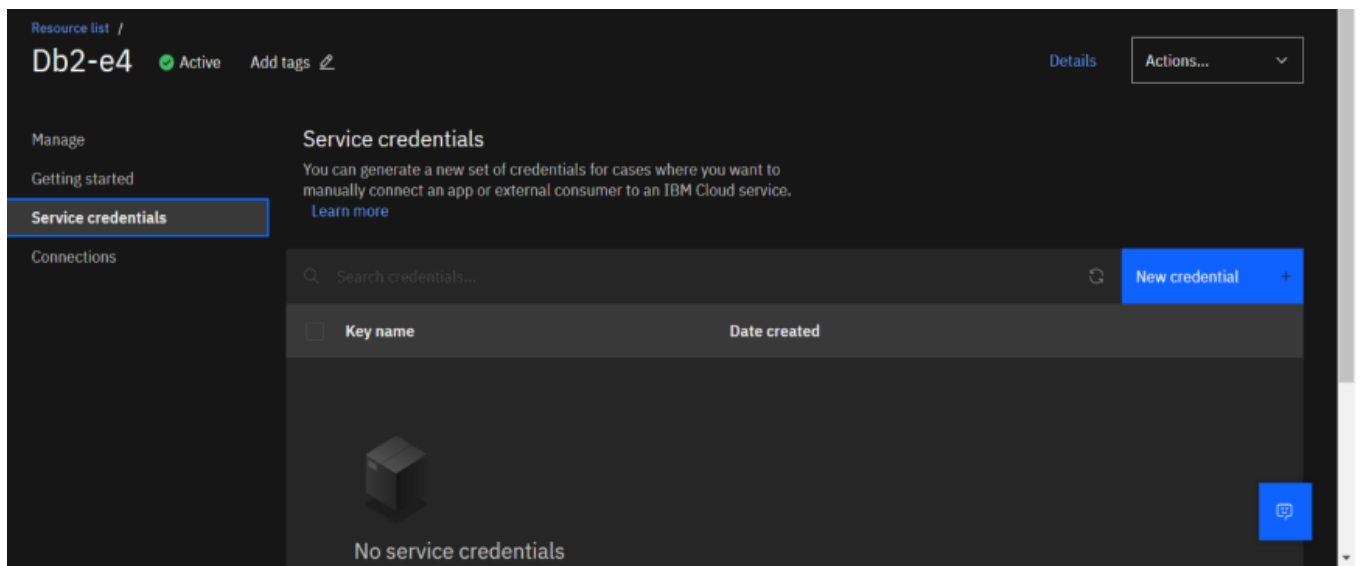
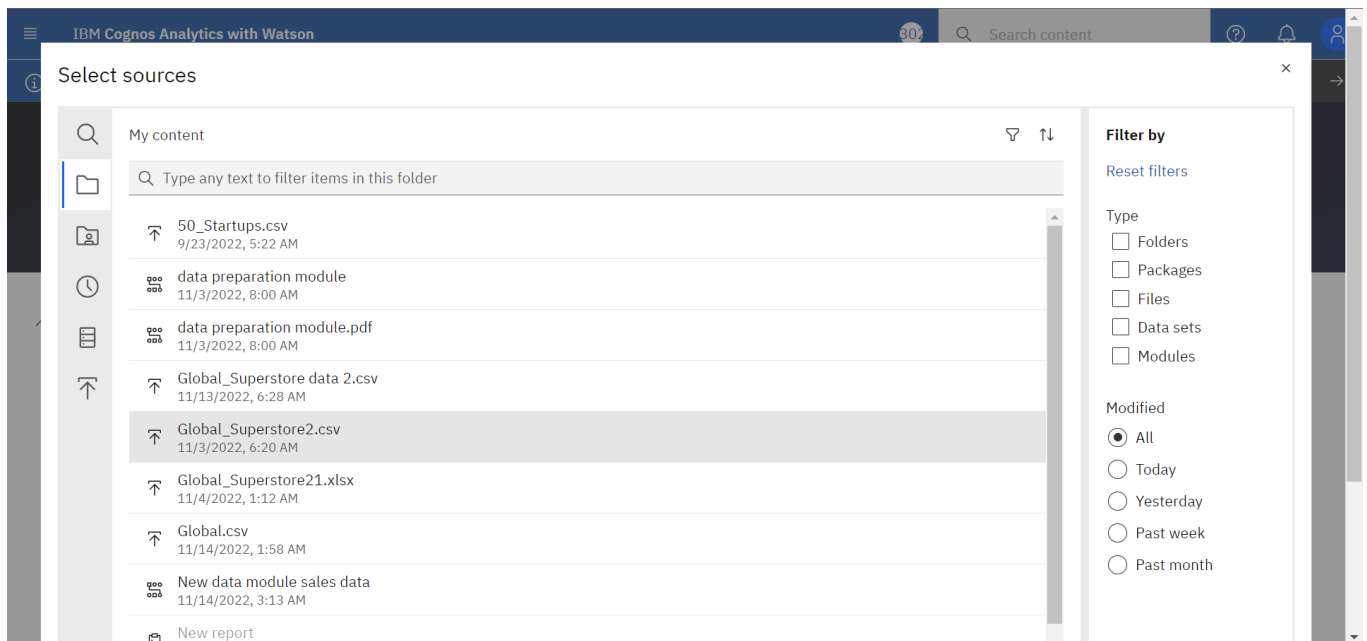
Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer	Customer Segment	City	State	Country	Postal Code	Market	Region	Product ID	Category	Sub-Category	Product Name	Sales	Quantity	Discount	Profit	Shipping Cost
1	32298	CA-2012-1	2012-01-01	Standard	Same Day	RH-19495	Rick Hansc Consumer	New York	New York	United States	10024	US	TEC-AC-1C	Technology	Accessories	Plantronics	2309.65	7	0	762.1845	933.57
2	26341	IN-2013-7	2013-07-01	Second Class	CR-12730	Justin Ritts Corporate	Wollongong	New South Wales	Australia		APAC	Oceania	FUR-CH-1C	Furniture	Chairs	Novimex E	3709.395	9	0.1	-288.765	923.63
3	25330	IN-2013-7	2013-07-01	First Class	KM-16375	Katherine Home Office	Berlin	Berlin	Germany		EU	Central	TEC-PH-1C	Technology	Phones	Motorola	2892.51	5	0.1	-96.54	910.16
4	47221	SG-2013-4	2013-04-01	Same Day	RH-9495	Rick Hansc Consumer	Dakar	Dakar	Senegal		AFRICA	Africa	TEC-SHA-1C	Technology	Copiers	Sharp Wirc	2832.96	8	0	311.52	903.04
5	22732	IN-2013-4	2013-04-01	Second Class	JM-15655	Jim Mitchi Corporate	Sydney	New South Wales	Australia		APAC	Oceania	TEC-PH-1C	Technology	Phones	Samsung S	2862.675	5	0.1	763.275	897.35
6	30570	IN-2011-8	2011-08-01	First Class	TS-21340	Toby Swin Consumer	Porirua	Wellington	New Zealand		APAC	Oceania	FUR-CH-1C	Furniture	Chairs	Novimex E	1822.08	4	0	564.84	894.77
7	31192	IN-2012-8	2012-08-01	Standard	CMB-18085	Mick Brown Consumer	Hamilton	Waikato	New Zealand		APAC	Oceania	FUR-TA-1C	Furniture	Tables	Chromcraft	5244.84	6	0	996.48	878.38
8	40155	CA-2014-1	2014-01-01	Standard	CMB-15220	Jane Wacc Corporate	Sacramento	California	United States	95823	US	West	OFF-BI-10X	Office Supplies	Binders	Fellowes P	5083.96	5	0.2	1906.485	867.69
9	40936	CA-2012-1	2012-01-01	Second Class	JM-15985	Joseph Ho Consumer	Concord	North Carolina	United States	28027	US	South	FUR-TA-1C	Furniture	Tables	Chromcraft	4297.644	13	0.4	-1862.31	865.74
10	34577	CA-2011-1	2011-01-01	Second Class	GM-14695	Greg Maxi Corporate	Alexandria	Virginia	United States	22304	US	South	OFF-SU-1C	Office Supplies	Supplies	Martin Yal	4164.05	5	0	83.281	846.54
11	28879	ID-2012-2	2012-02-01	First Class	AI-10780	Anthony J Corporate	Kabul	Kabul	Afghanistan		APAC	Central Asia	FUR-TA-1C	Furniture	Tables	Bevis Conf	4626.15	5	0	647.55	835.57
12	45794	SA-2011-1	2011-01-01	Second Class	MM-7260	Magdalen Consumer	Jizan	Jizan	Saudi Arabia		EMEA	EMEA	TEC-CIS-1C	Technology	Phones	Cisco Sma	2616.96	4	0	1151.4	832.41
13	4132	MX-2012-1	2012-01-01	Same Day	VF-21715	Vicky Frey Home Office	Toledo	Parana	Brazil		LATAM	South	FUR-CH-1C	Furniture	Chairs	Harbour C	2221.8	7	0	622.02	810.25
14	27704	IN-2013-7	2013-07-01	Second Class	PF-19120	Peter Fulle Consumer	Mudanjiang	Heilongjiang	China		APAC	North Asia	OFF-AP-10	Office Supplies	Appliances	KitchenAid	3701.52	12	0	1036.08	804.54
15	13779	ES-2014-5	2014-05-01	Second Class	BP-11185	Ben Peterr Corporate	Paris	Ile-de-France	France		EU	Central	OFF-AP-10	Office Supplies	Appliances	Breville Re	1869.588	4	0.1	186.948	801.66
16	36178	CA-2014-1	2014-01-01	Second Class	TB-21175	Thomas B Corporate	Henderson	Kentucky	United States	42420	US	South	TEC-AC-1C	Technology	Accessories	Logitech d	2249.91	9	0	517.4793	780.7
17	12069	ES-2014-1	2014-01-01	Standard	CP-18835	Patrick Jor Corporate	Prato	Tuscany	Italy		EU	South	OFF-AP-10	Office Supplies	Appliances	Hoover St	7958.58	14	0	3979.08	778.32
18	22096	IN-2014-1	2014-01-01	First Class	JS-15685	Jim Sink Corporate	Townsville	Queensland	Australia		APAC	Oceania	TEC-CO-1C	Technology	Copiers	Brother Fa	2565.594	9	0.1	28.404	766.93
19	49463	TZ-2014-8	2014-08-01	Second Class	RH-9555	Ritsa Hight Consumer	Uvlinza	Kigoma	Tanzania		AFRICA	Africa	OFF-KIT-1C	Office Supplies	Appliances	KitchenAid	3409.74	6	0	818.28	763.38
20	46630	PL-2012-7	2012-07-01	First Class	AB-600	Ann Blume Corporate	Bytom	Silesia	Poland		EMEA	EMEA	FUR-HON-1C	Furniture	Tables	Hon Comp	1977.72	4	0	276.84	759.47
21	31784	CA-2011-1	2011-01-01	First Class	SA-20830	Sue Ann R Consumer	Chicago	Illinois	United States	60610	US	Central	TEC-PH-1C	Technology	Phones	Apple iPho	2735.952	6	0.2	341.994	752.51
22	21586	IN-2014-4	2014-04-01	First Class	JK-15325	Jason Klan Corporate	Suzhou	Anhui	China		APAC	North Asia	FUR-CH-1C	Furniture	Chairs	SAFCO Exe	2754	6	0	358.02	752.47
23	13528	ES-2013-2	2013-02-01	Second Class	LB-16795	Laurel Belt Home Office	Edinburgh	Scotland	United Kingdom		EU	North	OFF-AP-10	Office Supplies	Appliances	KitchenAid	5273.7	10	0	1898.4	730.91
24	1570	US-2014-1	2014-01-01	First Class	NP-18325	Naresj Pat Consumer	Juarez	Chihuahua	Mexico		LATAM	North	TEC-PH-1C	Technology	Phones	Motorola	1713.84	4	0	445.52	728.97
25	3484	MX-2014-1	2014-01-01	First Class	VD-21670	Valerie Do Consumer	Soyapang	San Salvador	El Salvador		LATAM	Central	FUR-TA-1C	Furniture	Tables	Hon Comp	2106.496	8	0.2	526.496	728.39
26	30191	IN-2011-1	2011-01-01	First Class	PB-19210	Phillip Brey Corporate	Taipei	Taipei City	Taiwan		APAC	North Asia	FUR-TA-1C	Furniture	Tables	Lesro Cont	1715.16	2	0	720.36	725.57

## Uploading the dataset:

The data set can be uploaded via upload or by connecting with the cloud account and fetch the meta data form DB2



It can be also achieved by the use of the cloud db2



```
IBM Cloud Search resources and products... Catalog Manage Ebin N L's Account

482"
  },
  "hosts": [
    {
      "hostname": "824dfd4d-99de-440d-9991-629c01b3832d.bs2io90108qkb1od8lclg.databases.appdomain.cloud:31",
      "port": 30119
    }
  ],
  "jdbc_url": {
    "jdbc": "db2://824dfd4d-99de-440d-9991-629c01b3832d.bs2io90108qkb1od8lclg.databases.appdomain.cloud:30119/bludb:user=<userid>;password=<your_password>;sslConnection=true;"
  },
  "path": "/bludb",
  "query_options": {
    "authSource": "admin",
    "replicaSet": "replset"
  },
  "replica_set": "replset",
  "scheme": "db2",
  "type": "uri"
},
{
  "instance_administration_api": {
    "deployment_id": "crn:v1:bluemix:public:dashdb-for-transactions:eu-gb:a/0bb0da031a9841ffb510cf39b384a508:cbe0ea64-19ff-4806-86fd-487c96b6e33c::",
    "instance_id": "824dfd4d-99de-440d-9991-629c01b3832d.bs2io90108qkb1od8lclg.databases.appdomain.cloud:30119"
  }
}
```

IBM Cognos Analytics with Watson

sales data analytics

Owner: Ebin N ... (cef8a4) | Created: 11/12/2022, 9:09 AM | Modified: 11/12/2022, 9:09 AM | Type: Data Server

General | **Connections** | Permissions

Name	Modified
sales data analytics	11/12/2022 9:09 AM

sales data analytics

General | Settings | **Schemas** | Permissions

Status	Schema name	Tables loaded
<input type="radio"/>	IBM_RTMON_DATA	
<input type="radio"/>	IBM_RTMON_EVMON	
<input type="radio"/>	IBM_RTMON_METADATA	
<input type="radio"/>	PROCMGMT	
<input type="radio"/>	ST_INFORMTN_SCHEMA	
<input type="radio"/>	WTW00999	14% loaded

Show system schemas

Load metadata  
Load options

IBM Cognos Analytics with Watson

New data module

300

Search content

?

Properties

Data module

Search

New data module

Navigation paths

Global\_Sup...ata 2.csv

# Row Id

# Row ID

abc Order ID

Order Date

Ship Date

abc Ship Mode

abc Customer ID

abc Customer Name

abc Segment

City

Grid

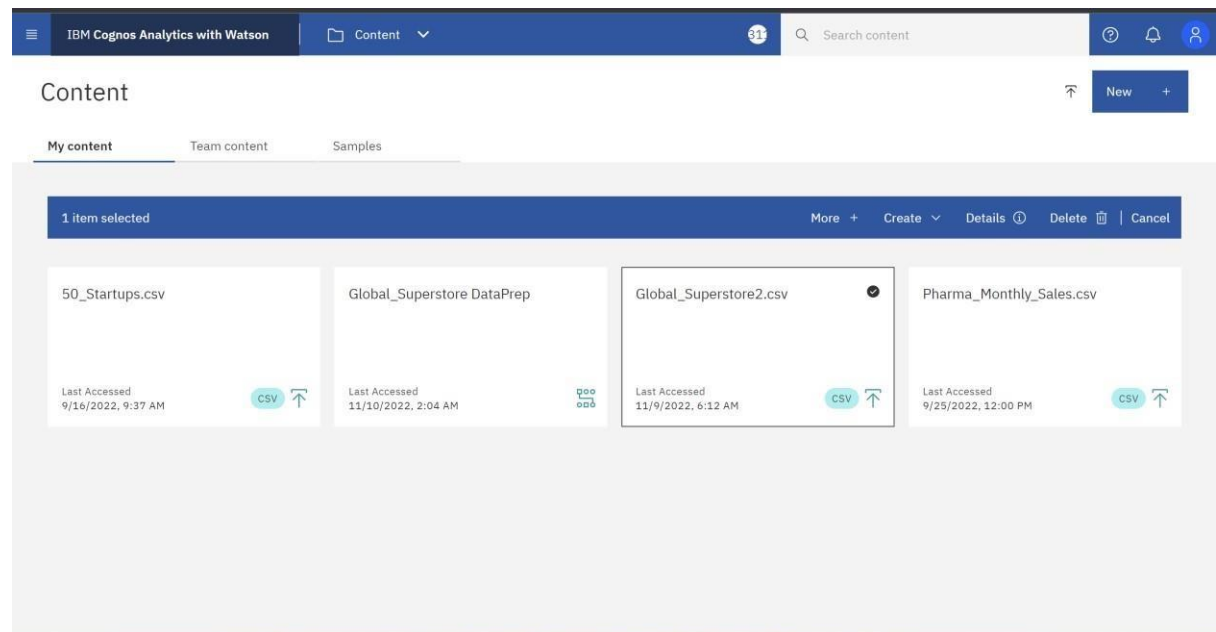
Relationships

Custom tables

**Preview data**  
To preview data, select a table, a column in a table, or a folder that contains columns.

## Data Preparation:

- First understand and load the dataset.



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

The screenshot shows the 'Data module' in IBM Cognos Analytics. A 'Maintenance' banner at the top states 'Scheduled maintenance completed. Click More Info for details and to subscribe to future events'. The 'Data module' section has a search bar and a list of 'New data module' items, including 'Global\_Su...store2.csv'. The 'Grid' view is selected, showing a table with 10 rows and 6 columns: 'Row ID', 'Order ID', 'Order Date', 'Ship Date', 'Ship Mode', and 'Customer ID'. The data is as follows:

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
1	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-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	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495
6	IN-2013-42360	2013-06-28	2013-07-01	Second Class	JM-15655
7	IN-2011-81826	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	2014-10-14	2014-10-21	Standard Class	JW-15220
10	CA-2012-116638	2012-01-28	2012-01-31	Standard Class	TH-15085

- 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 interface. On the left, the 'Data module' pane lists various fields: Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, City, State, Country, Postal Code, Market, Region, Product ID, Category, and Sub-Category. The main area displays a data grid with columns: Row ID, Order ID, Ship Date, Ship Mode, and Customer ID. A context menu is open over the 'Order ID' column, showing options: Filter..., Create data group..., Create navigation path..., Split..., Hide from users, Remove, Format data..., Clean..., Sort descending, Sort ascending, and Properties.

Row ID	Order ID	Ship Date	Ship Mode	Customer ID
1	32298	2012-07-31	Same Day	RH-19495
2	26341	2013-02-07	Second Class	JR-16210
3	25330	2013-10-18	First Class	CR-12730
4	13524	2013-01-30	First Class	KM-16375
5	47221	2013-11-06	Same Day	RH-9495
6	22732	2013-07-01	Second Class	JM-15655
7	30570	2011-11-09	First Class	TS-21340
8	31192	2012-04-14	Standard Class	MB-18085
9	40155	2014-10-14	Standard Class	JW-15220
10	40936	2012-01-28	Second Class	JH-15985

The screenshot shows the 'Filter - Order ID' dialog box. It has a search bar and a list of Order IDs with checkboxes. The 'Clear all' button is at the bottom left, and the 'Invert' button is at the bottom right. The 'Cancel' and 'OK' buttons are at the bottom.

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

Cancel OK

- Once cleaned save the data.

### Data calculations and navigation paths:

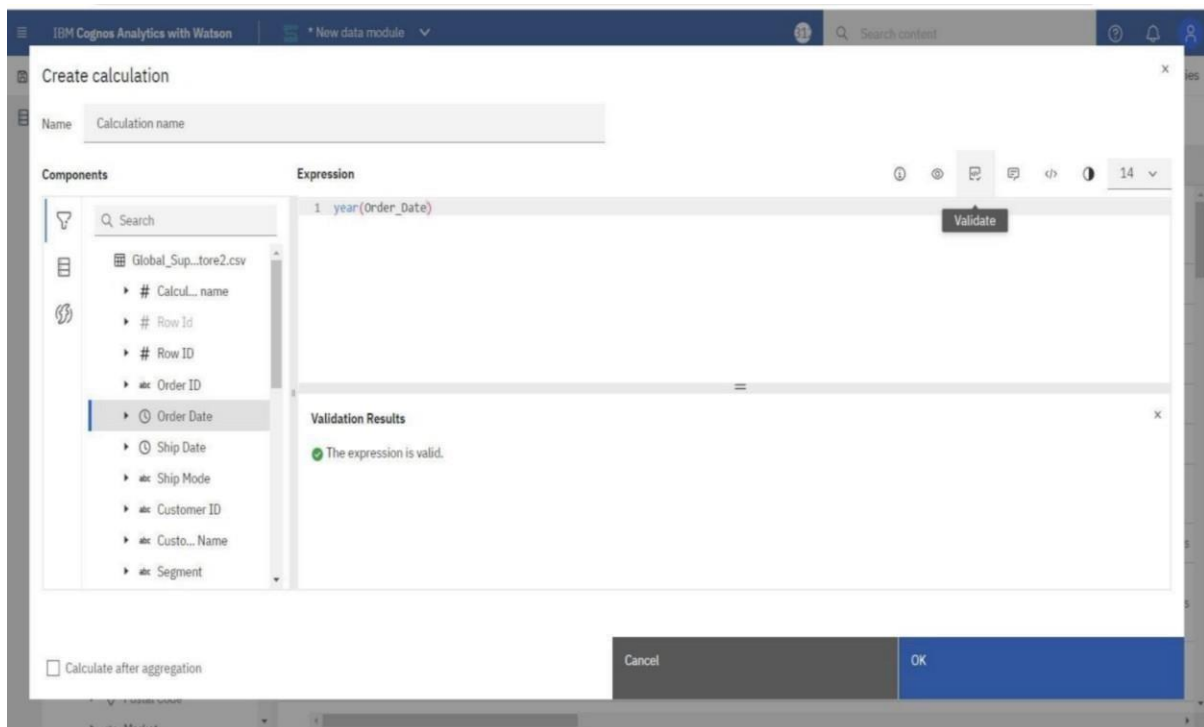
- Once you load the data, we need to prepare the data.
- Prepare calculations of year, month, day fields and also related navigation paths.
- Create a few more calculations – Target sales, Min sales, Max sales, Middle Range sales.

The screenshot shows the IBM Cognos Analytics interface. The top bar indicates 'IBM Cognos Analytics with Watson' and 'New data module'. 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, Postal Code, Market, and Location. The main area displays a table with columns: Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. A context menu is open over the 'Order Date' column, showing options: 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.

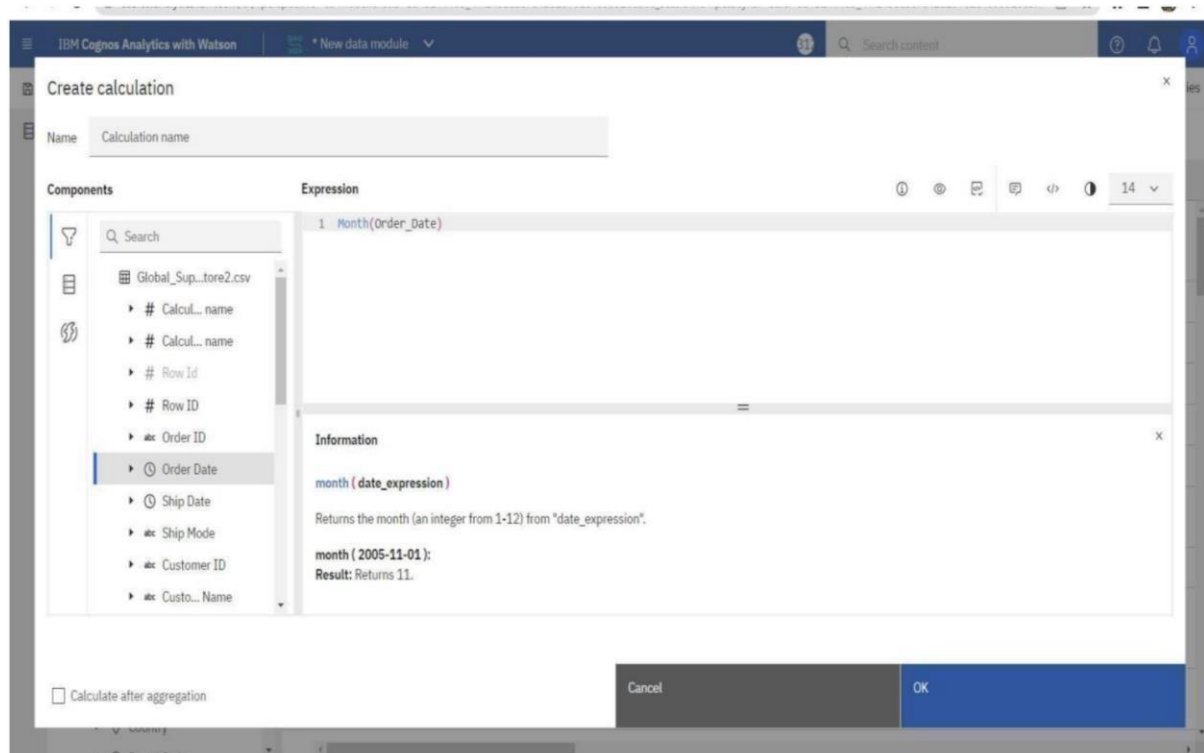
Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495
26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210
25330	IN-2013-71249	2013-10-17	2013-10-18	First Class	CR-12730
13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	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

- Right click -> calculation -> type expression -> validate -> set new name as Order Year.

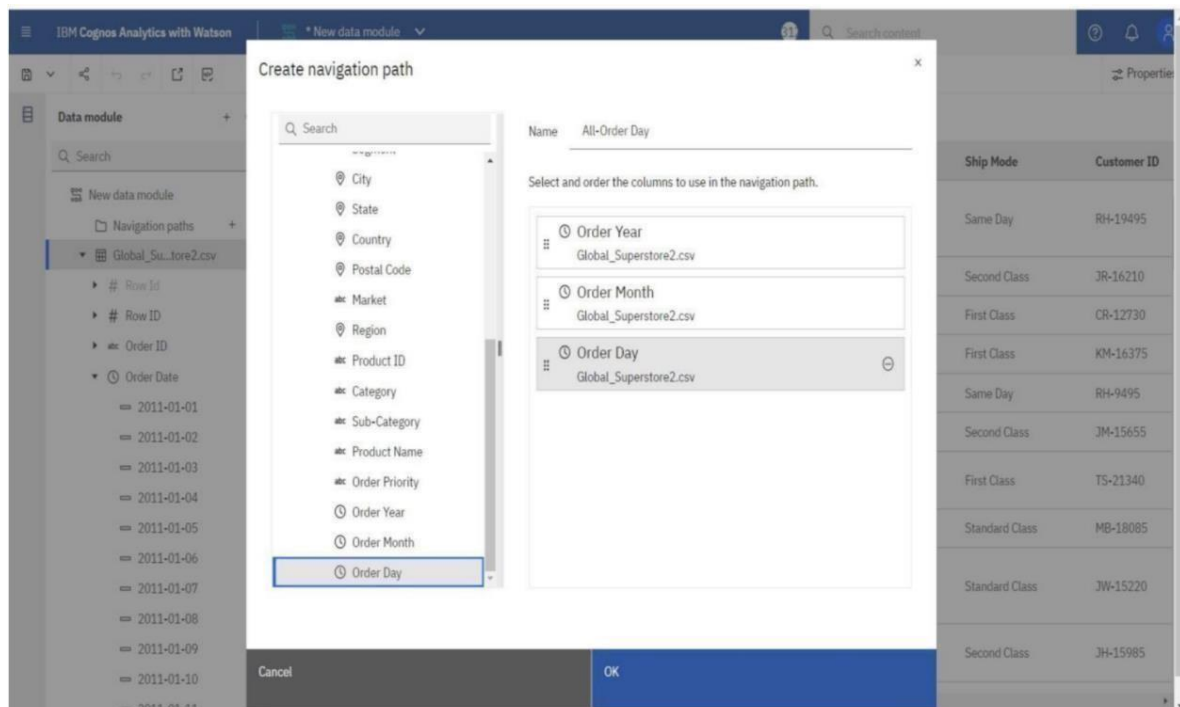




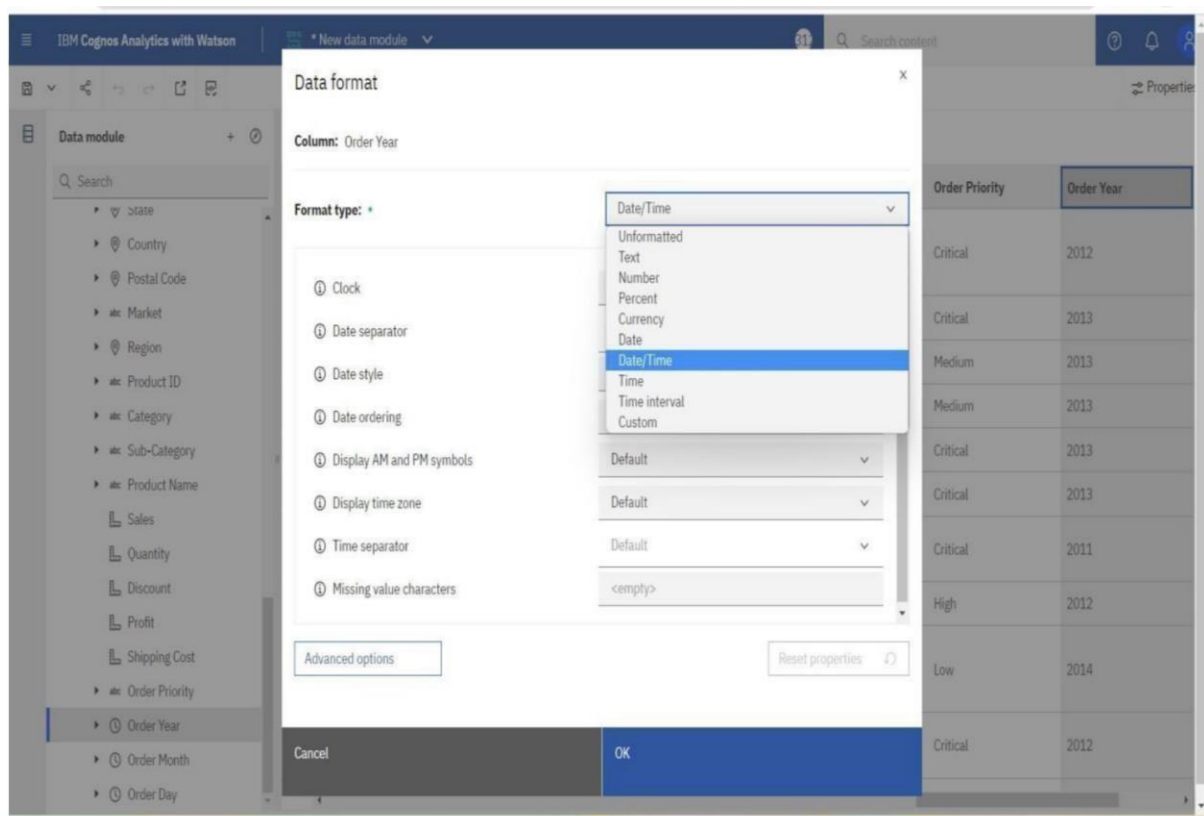
- Similarly, 'Order Month', 'Order Day' and 'All'.



- Right click on create navigation path.



- Data format -> date/Time



- Properties -> usage -> Attribute  
Aggregate -> Count Distinct

Data Type -> Integer

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

Display options -> show members

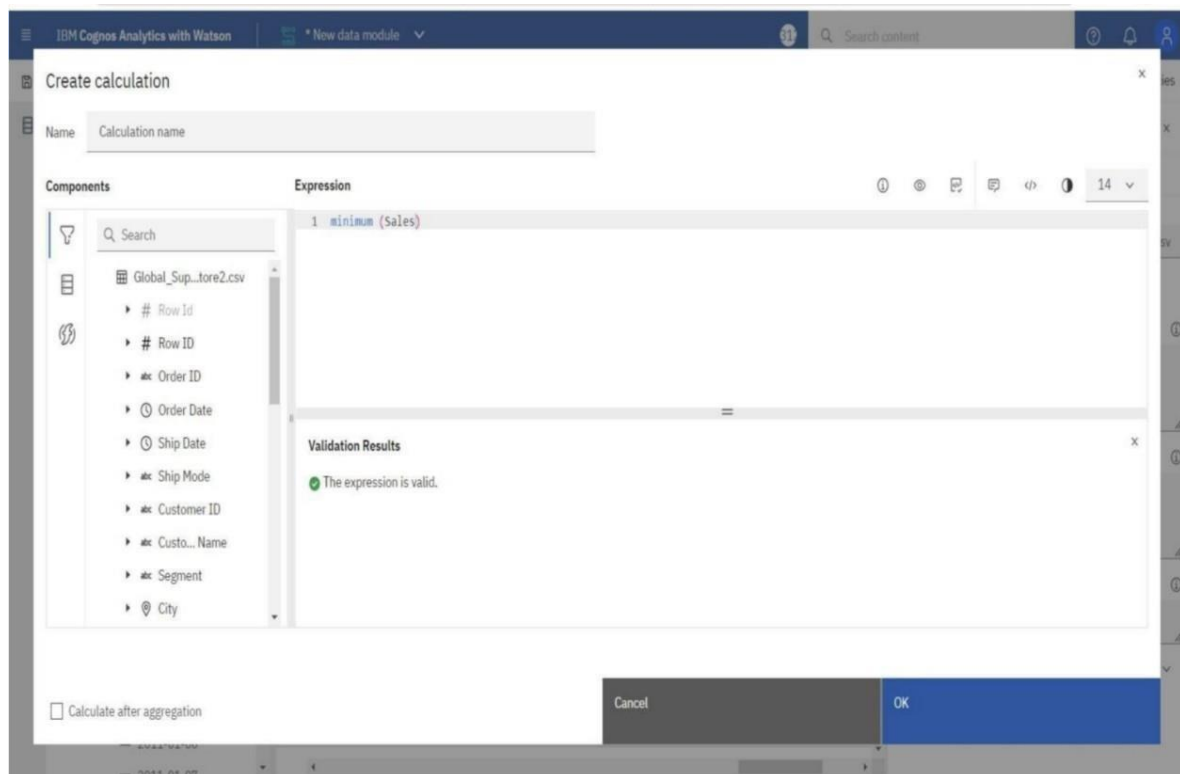
This way it becomes a Numerical Data Type

The screenshot shows the IBM Cognos Analytics interface. On the left, the 'Data module' pane lists various attributes: State, Country, Postal Code, Market, Region, Product ID, Category, Sub-Category, Product Name, Sales, Quantity, Discount, Profit, Shipping Cost, Order Priority, Order Year, Order Month, and Order Day. The 'Order Year' attribute is selected. The main area displays a data grid with columns: Profit, Shipping Cost, Order Priority, Order Year, and Ord. The grid contains several rows of data. On the right, the 'Properties' panel is open, showing the 'General' tab. The 'Label' is 'Order Year', 'Hide from users' is checked, and 'Usage' is set to 'Attribute'. The 'Date' dropdown menu is open, showing options: Year, Quarter, Season, Month, Week, Day, Hour, Minute, and Second. The 'Year' option is selected.

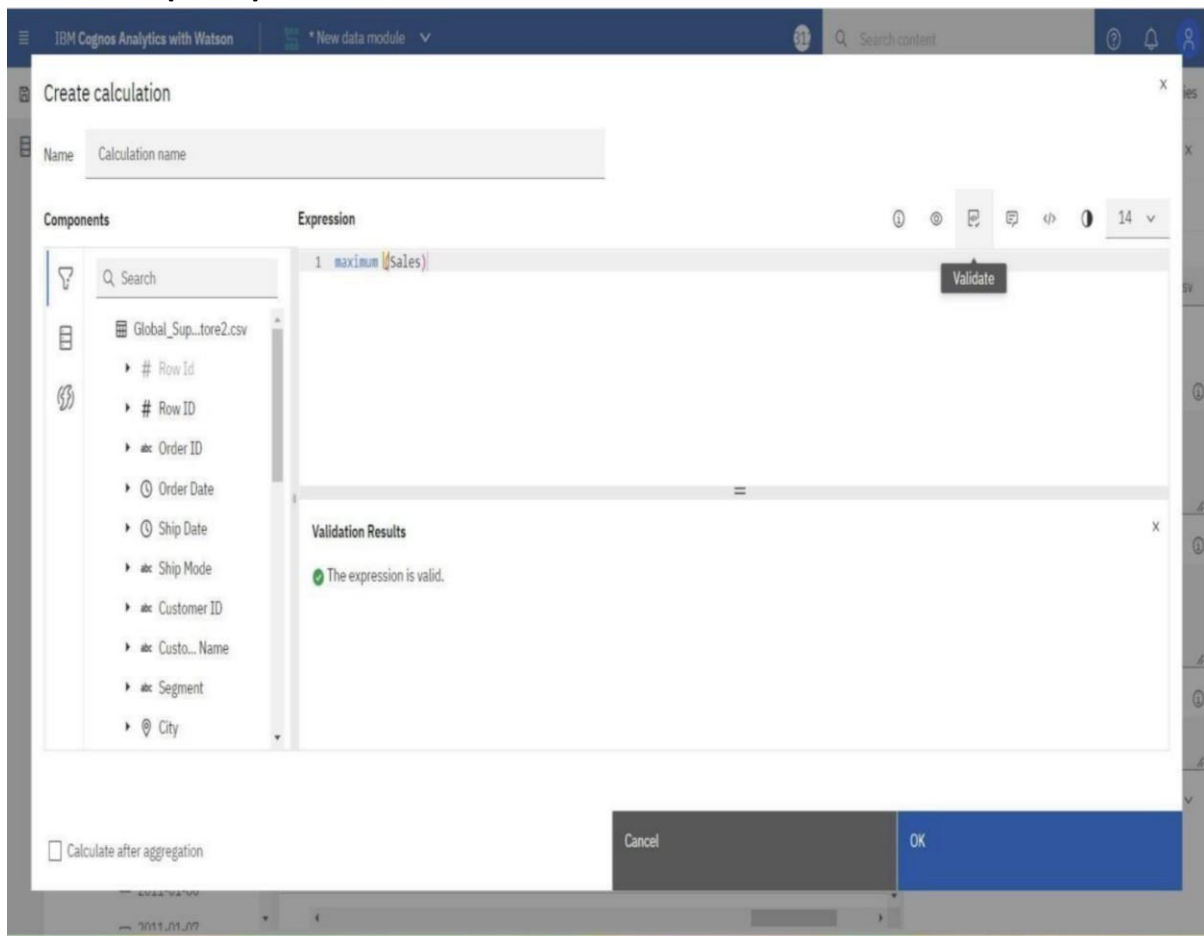
Profit	Shipping Cost	Order Priority	Order Year	Ord
762.1845	933.57	Critical	2012	7
-288.765	923.63	Critical	2013	2
919.971	915.49	Medium	2013	1
-96.54	910.16	Medium	2013	1
311.52	903.04	Critical	2013	1
763.275	897.35	Critical	2013	6
564.84	894.77	Critical	2011	1
996.48	878.38	High	2012	4
1906.485	867.69	Low	2014	1
-1862.3124	865.74	Critical	2012	1

- Hence calculations for day, month and year are completed.
- 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)

