

EXPLORATORY DATA ANALYSIS

Team ID	PNT2022TMID00819
Project Name	Global Sales Data Analytics

Dataset:

File Home Insert Page Layout Formulas Data Review View Help																			Comments		Share		
Undo Paste Cut Copy Format Painter			Calibri 11 A+ A- Wrap Text			General			Conditional Formatting Format as Table Cell Styles			Insert Delete Format			Autosum Fill Sort & Filter Find & Select								
Clipboard			Font			Alignment			Number			Styles			Cells			Editing					
P1	Category																						
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	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	Order
2	32298	CA-2012-1	#####	#####	Same Day	RH-19495	Rick Hans Consumer	New York	New York	United States	10024	US	East	TEC-AC-1C	Technology	Accessories	Plantronics	2309.65	7	0	762.1845	933.57	Cr
3	26341	IN-2013-7	#####	#####	Second Class	JR-16210	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	Cr	
4	25330	IN-2013-7	#####	#####	First Class	CR-12730	Craig Reite Consumer	Brisbane	Queensland	Australia	APAC	Oceania	TEC-PH-1C	Technology	Phones	Nokia Sma	5175.171	9	0.1	919.971	915.49	Ma	
5	13524	ES-2013-1	#####	#####	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	Ma	
6	47221	SG-2013-4	#####	#####	Same Day	RH-9495	Rick Hans Consumer	Dakar	Dakar	Senegal	Africa	Africa	TEC-SHA-1C	Technology	Copiers	Sharp Wirr	2832.96	8	0	311.52	903.04	Cr	
7	22732	IN-2013-4	#####	#####	Second Class	JM-15655	Jim Mitch Corporate	Sydney	New South Wales	Australia	APAC	Oceania	TEC-PH-1C	Technology	Phones	Samsung S	2862.675	5	0.1	763.275	897.35	Cr	
8	30570	IN-2011-8	#####	#####	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	Cr	
9	31192	IN-2012-8	#####	#####	Standard Class	MB-18085	Mick Brown Consumer	Hamilton	Waikato	New Zealand	APAC	Oceania	FUR-TA-1C	Furniture	Tables	Chromcraft	5244.84	6	0	996.48	878.38	Cr	
10	40155	CA-2014-1	#####	#####	Standard Class	JW-15220	Jane Wacc Corporate	Sacramento	California	United States	95823	US	West	OFF-BI-10C	Office Sup	Binders	Fellowes P	5083.96	5	0.2	1906.485	867.69	Lo
11	40936	CA-2012-1	#####	#####	Second Class	JH-15985	Joseph Ho Consumer	Concord	North Car	United States	28027	US	South	FUR-TA-1C	Furniture	Tables	Chromcraft	4297.644	13	0.4	-1862.31	865.74	Cr
12	34577	CA-2011-1	#####	#####	Second Class	GM-14695	Greg Max Corporate	Alexandria	Virginia	United States	22304	US	South	OFF-SU-1C	Office Sup	Supplies	Martin Yal	4164.05	5	0	83.281	846.54	Cr
13	28879	ID-2011-2	#####	#####	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	Cr	
14	4132	MX-2012-1	#####	#####	Second Class	MM-7260	Magdalena Consumer	Jizan	Jizan	Saudi Arabia	EMEA	EMEA	TEC-CIS-1C	Technology	Phones	Cisco Sma	2616.96	4	0	1151.4	832.41	Cr	
15	4132	MX-2012-1	#####	#####	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	Cr	
16	27704	IN-2013-7	#####	#####	Second Class	PF-19120	Peter Fulle Consumer	Mudanjiang	Heilongjiang	China	APAC	North Asia	OFF-AP-10C	Office Sup	Appliances	KitchenAid	3701.52	12	0	1036.08	804.54	Cr	
17	13779	ES-2014-5	#####	#####	Second Class	BP-11185	Ben Peter Corporate	Paris	Ile-de-Franc	France	EU	Central	OFF-AP-10C	Office Sup	Appliances	Breville Re	1869.588	4	0.1	186.948	801.66	Cr	
18	36178	CA-2014-1	#####	#####	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	Cr
19	12069	ES-2014-1	#####	#####	Standard Class	CP-18835	Patrick Jor Corporate	Prato	Tuscany	Italy	EU	South	OFF-AP-10C	Office Sup	Appliances	Hoover Sta	7958.58	14	0	3979.08	778.32	Lo	
20	22096	IN-2014-1	#####	#####	First Class	JS-15685	Jim Sink Corporate	Townsville	Queensland	Australia	APAC	Oceania	FUR-CO-1C	Technology	Copiers	Brother Fa	2565.594	9	0.1	28.404	766.93	Lo	
21	49463	TZ-2014-8	#####	#####	Second Class	RH-9555	Ritsa Hight Consumer	Uvinza	Kigoma	Tanzania	Africa	Africa	OFF-KIT-1C	Office Sup	Appliances	KitchenAid	3409.74	6	0	818.28	763.38	Cr	
22	46630	PL-2012-7	#####	#####	First Class	AB-600	Ann Blume Corporate	Bytom	Silesia	Poland	EMEA	EMEA	FUR-HOM-1C	Furniture	Tables	Hon Comp	1977.72	4	0	276.84	759.47	Cr	
23	31784	CA-2011-1	#####	#####	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	Cr
24	21586	IN-2011-4	#####	#####	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	Cr	
25	13528	ES-2013-2	#####	#####	Second Class	LB-16795	Laurel Belt Home Office	Edinburgh	Scotland	United Kingdom	EU	North	OFF-AP-10C	Office Sup	Appliances	KitchenAid	5273.7	10	0	1898.4	730.91	Cr	
26	1570	US-2014-1	#####	#####	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	Cr	
27	3484	MX-2014-1	#####	#####	First Class	VD-21670	Valerie Do Consumer	Soyapango	San Salvad	El Salvador	LATAM	Central	FUR-TA-1C	Furniture	Tables	Hon Comp	2106.496	8	0.2	526.496	728.39	Cr	
28	30191	IN-2011-1	#####	#####	First Class	PB-19210	Phillip Brey Corporate	Taipei	Taipei City	Taiwan	APAC	North Asia	FUR-TA-1C	Furniture	Tables	Lesro Conl	1715.16	2	0	720.36	725.57	Cr	
Global Superstore2																							
Accessibility: Investigate																							

Uploading the dataset:

Data uploading can be done via upload option or by logging in to the cloud account and retrieving the DB2 meta data.

The screenshot shows the 'Content' page in IBM Cognos Analytics. The browser address bar displays the URL: `https://us3.ca.analytics.ibm.com/bi/?perspective=content&tab=myContent&folder=IF57C4544AE2C47AE9F28B4A0F5A3AFD8`. The page has a search bar and a 'New' button. Below the navigation tabs, a table lists the content items:

Name	Type	Last Accessed
50_Startups.csv (CSV)	Uploaded file	16/09/2022, 12:15
Assignment-2_Dashboard	Dashboard	24/09/2022, 07:06
Data_Module	Data module	18/11/2022, 08:10
Global Sales Data Analytics Final Dashboard	Dashboard	18/11/2022, 06:20
Global Sales Data Analytics Final Dashboard 2	Dashboard	19/11/2022, 09:58
Global Sales Data Analytics Report	Report	19/11/2022, 08:07
Global_Superstore2_PreparedDataFinal	Data module	18/11/2022, 07:25
Global_Superstore2.csv (CSV)	Uploaded file	15/11/2022, 21:46
Pharma_Monthly_Sales.csv (CSV)	Uploaded file	24/09/2022, 01:43
REPORT	Report	19/11/2022, 08:09
Sprint-4_Story	Story	19/11/2022, 00:09

IBM DB2 can be used as an alternative as well.

The screenshot shows the 'Home' page in IBM Cognos Analytics. The browser address bar displays the URL: `https://us3.ca.analytics.ibm.com/bi/?perspective=home`. The page features a welcome message, a maintenance notice, and a 'Quick launch' section with four options: Upload data, Prepare data, Exploration, and Present data. A user profile dropdown menu is visible on the right side of the page.

Quick launch options:

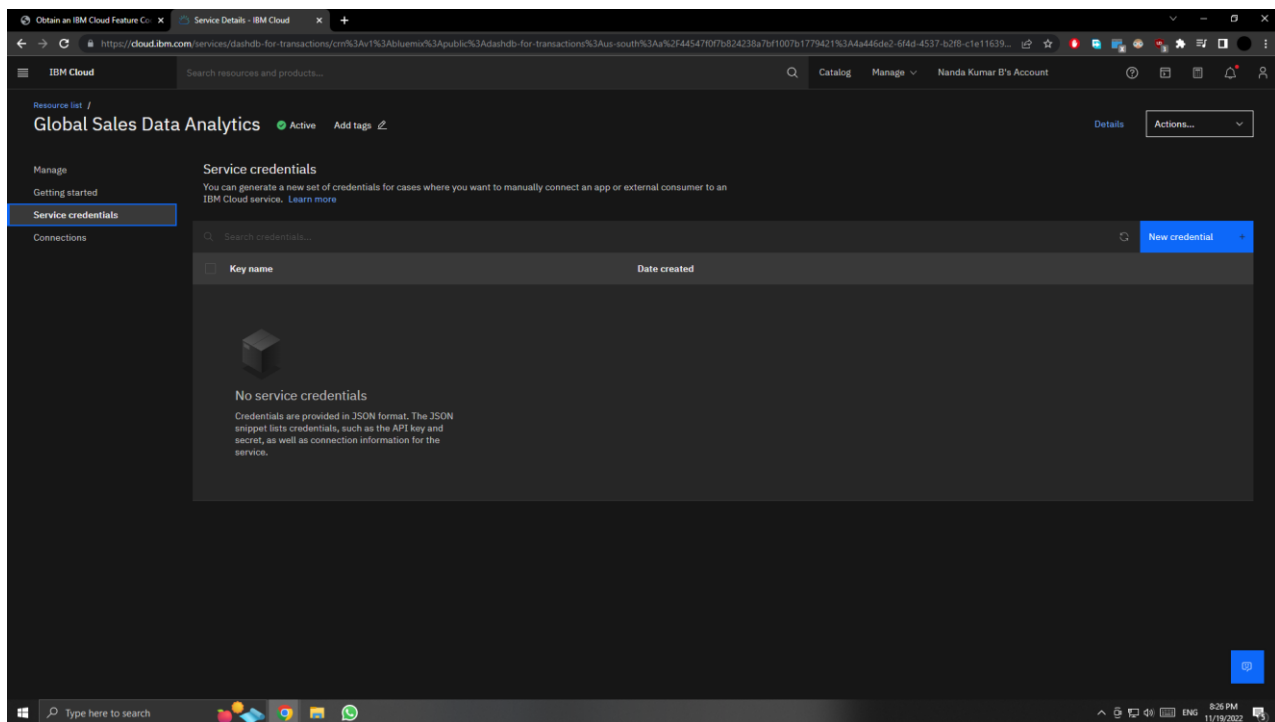
- Upload data:** Upload or drag and drop spreadsheets, csv files, and other data sources.
- Prepare data:** Use data modules to clean and connect data from multiple resources.
- Exploration:** Quickly find unbiased answers by identifying trends in your data with data exploration.
- Present data:** Create sophisticated, multi-page, multi-query dashboards, reports, or stories.

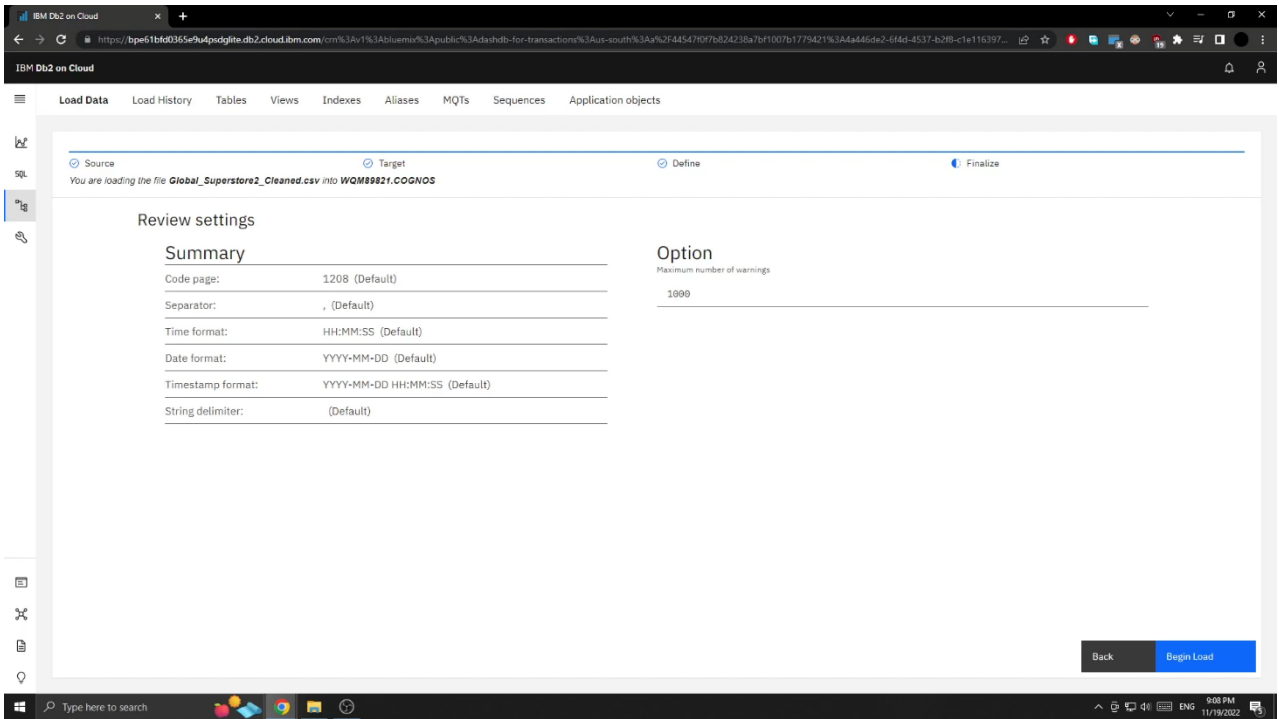
User Profile: Nanda Kumar B (758bb2...), nandakumar778@gmail.com, Profile and settings

Log out options: My schedules and subscriptions, Log my session, About, Manage product subscription, Privacy, Cookie Preferences, IBM Cognos Analytics Mobile, Log out

Recent items table:

Name	Location	Type	Last Accessed
Global Sales Data Analytics Final Dashboard 2	My content	Dashboard	19/11/2022, 08:20





Select a type

IBM DB2 x



IBM Db2



IBM Db2 for i



IBM Db2 Warehouse



IBM_DB2_GSDA



Owner

Nanda ... 58c42b)

Created: 19/11/2022, 09:03

Modified: 19/11/2022, 09:03

Type: Connection

General

Settings

Schemas

Permissions

Connection details

Edit >

Authentication method



Connect anonymously



Prompt for the user ID and password



Use an external namespace



Use the following signon:



Test



Success



Save



Type here to search



IBM Cognos Analytics with Watson

New data module

304

Search content

?

Grid

Relationships

Custom tables

Search

New data module

Navigation paths

Global_Sup...ata 2.csv

Row Id

Row ID

Order ID

Order Date

Ship Date


Ship Mode

Customer ID

Customer Name

Segment

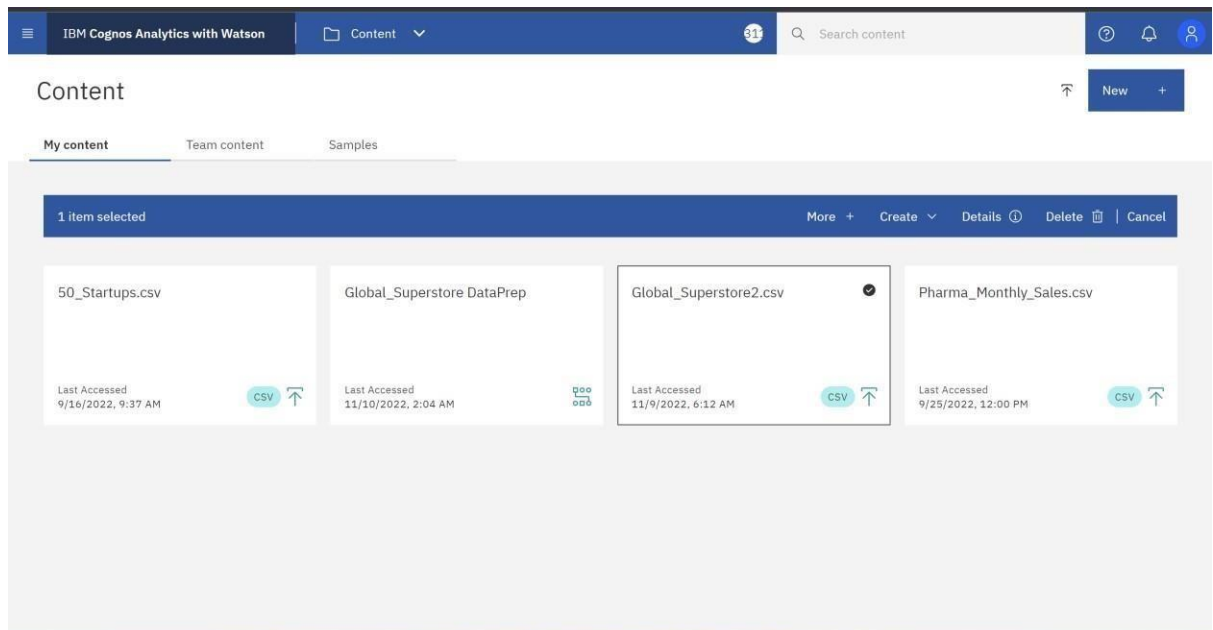
City



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. The top navigation bar includes 'IBM Cognos Analytics with Watson', a 'New data module' dropdown, a search bar, and user icons. A maintenance message is displayed: 'Maintenance: Scheduled maintenance completed. Click More Info for details and to subscribe to future events'. Below the navigation bar, the 'Data module' section has tabs for 'Grid', 'Relationships', and 'Custom tables'. The 'Grid' tab is active, showing a data table with the following columns: Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. The left sidebar shows the 'New data module' section with 'Global_Superstore2.csv' selected under 'Navigation paths'.

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	JH-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, Order Date, 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	Order Date	Ship Date	Ship Mode	Customer ID
1	32298	2012-07-31	2012-07-31	Same Day	RH-19495
2	26341	2013-02-07	2013-02-07	Second Class	JR-16210
3	25330	2013-10-17	2013-10-18	First Class	CR-12730
4	13524	2013-01-28	2013-01-30	First Class	KM-16375
5	47221	2013-11-06	2013-11-06	Same Day	RH-9495
6	22732	2013-07-01	2013-07-01	Second Class	JM-15655
7	30570	2011-11-09	2011-11-09	First Class	TS-21340
8	31192	2012-04-14	2012-04-14	Standard Class	MB-18085
9	40155	2014-10-14	2014-10-14	Standard Class	JW-15220
10	40936	2012-01-28	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.

Order ID
<input type="checkbox"/> AE-2011-9160
<input type="checkbox"/> AE-2013-1130
<input type="checkbox"/> AE-2013-1530
<input type="checkbox"/> AE-2014-2840
<input type="checkbox"/> AE-2014-3830
<input type="checkbox"/> AE-2014-4120
<input type="checkbox"/> AG-2011-1070
<input type="checkbox"/> AG-2011-1390

- Once cleaned save the data.

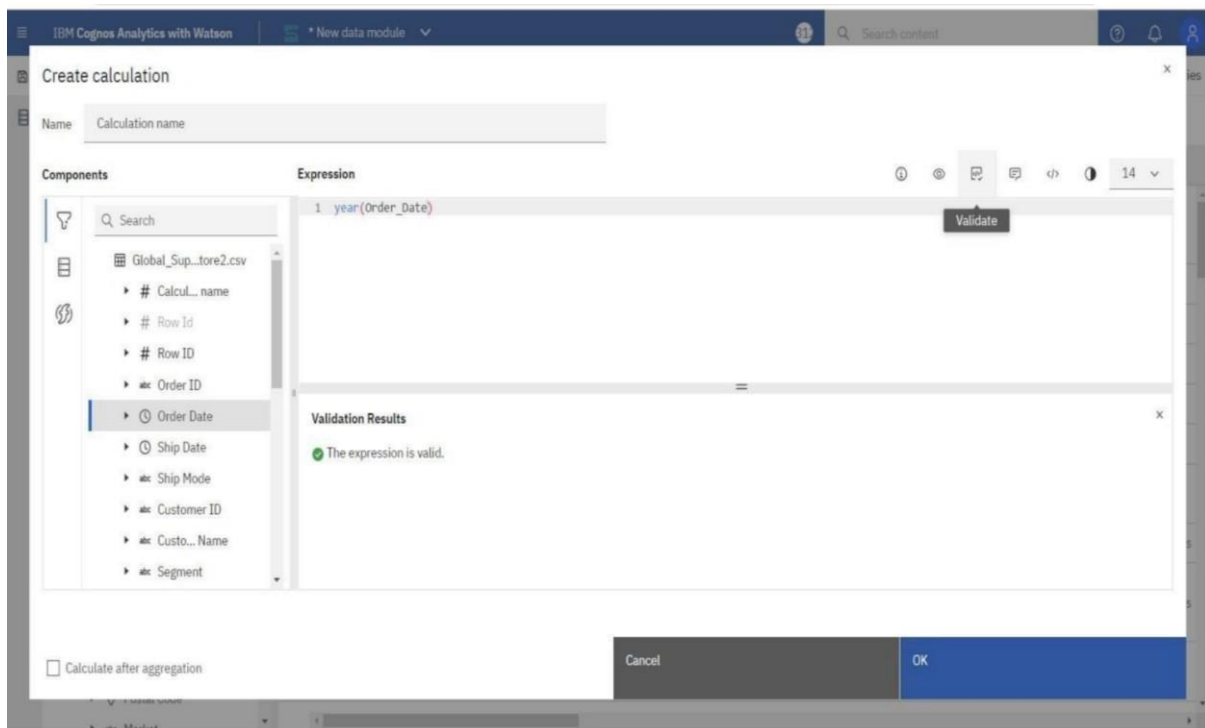
Data calculations and navigation paths:

- 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.

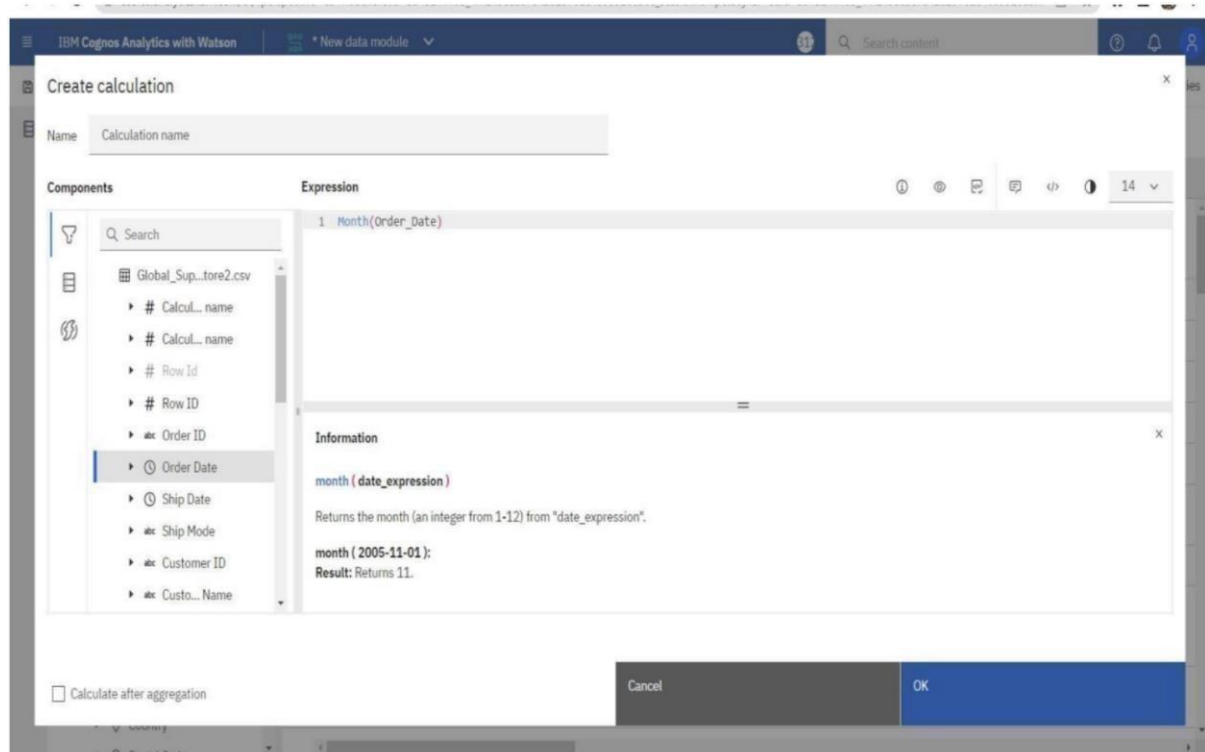
The screenshot shows the IBM Cognos Analytics interface. The top bar indicates 'IBM Cognos Analytics with Watson' and 'New data module'. The left sidebar shows a 'Data module' with a search bar and a list of fields: Row ID, Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, City, State, Country, Postal Code, Market, and Location. The main area displays a grid of data with columns: Row ID, Order ID, Order Date, Ship Date, Ship Mode, and Customer ID. A context menu is open over the grid, showing options like 'New', '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

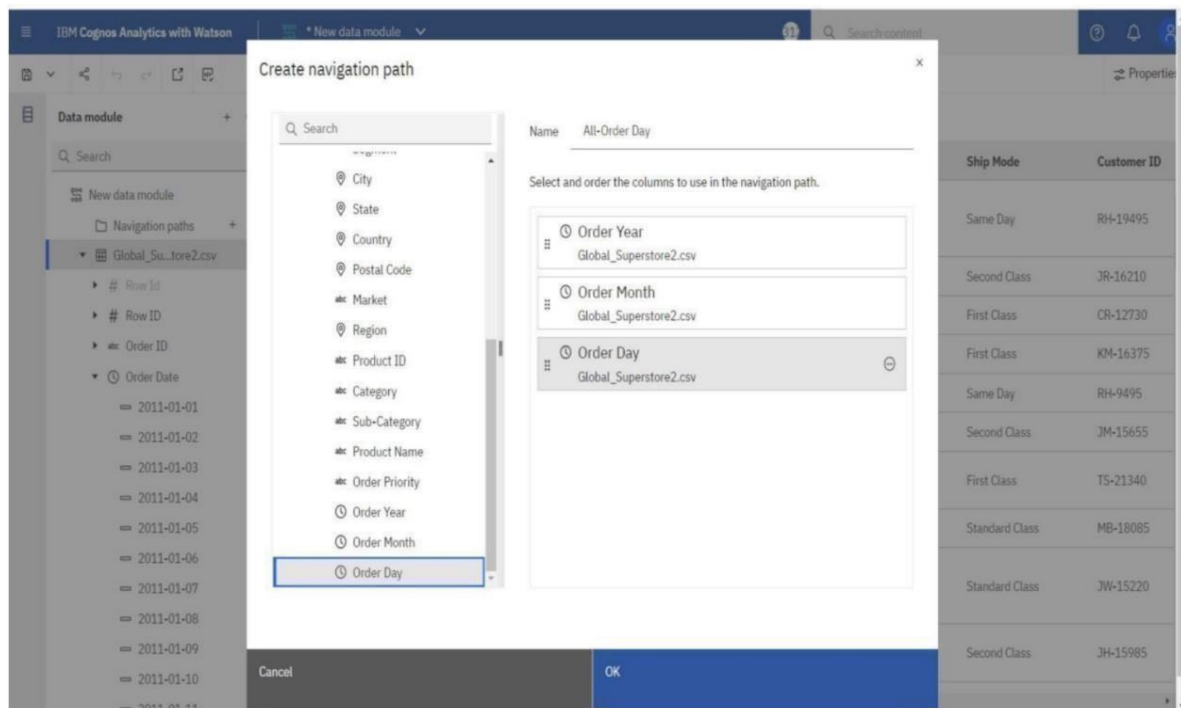
- Right click -> calculation -> type expression -> validate -> set new name as Order Year.



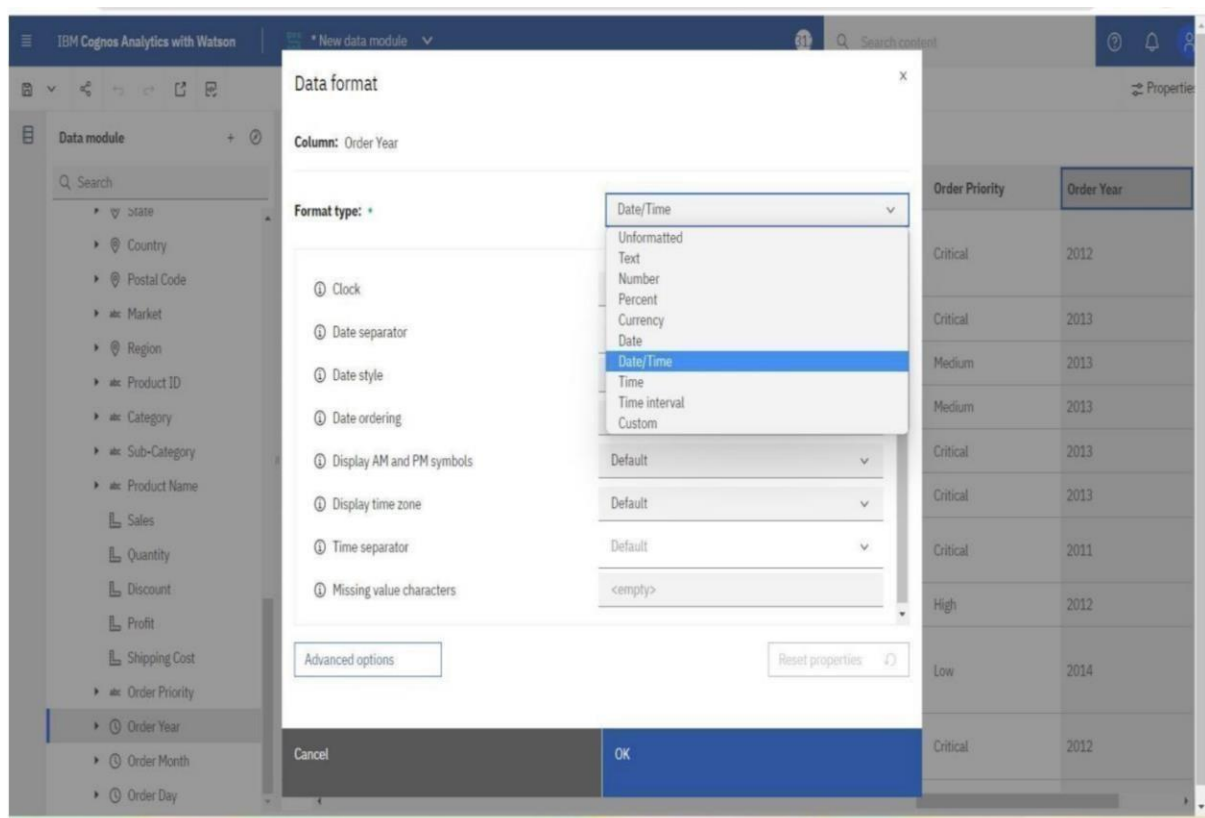
- Similarly, 'Order Month', 'Order Day' and 'All'.



- 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 fields: 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' field is selected. The main area displays a 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 'Expression' is 'View or edit'. The 'Usage' dropdown is set to 'Attribute', and the 'Date' dropdown is set to 'Year'. The 'Description' field is empty.

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

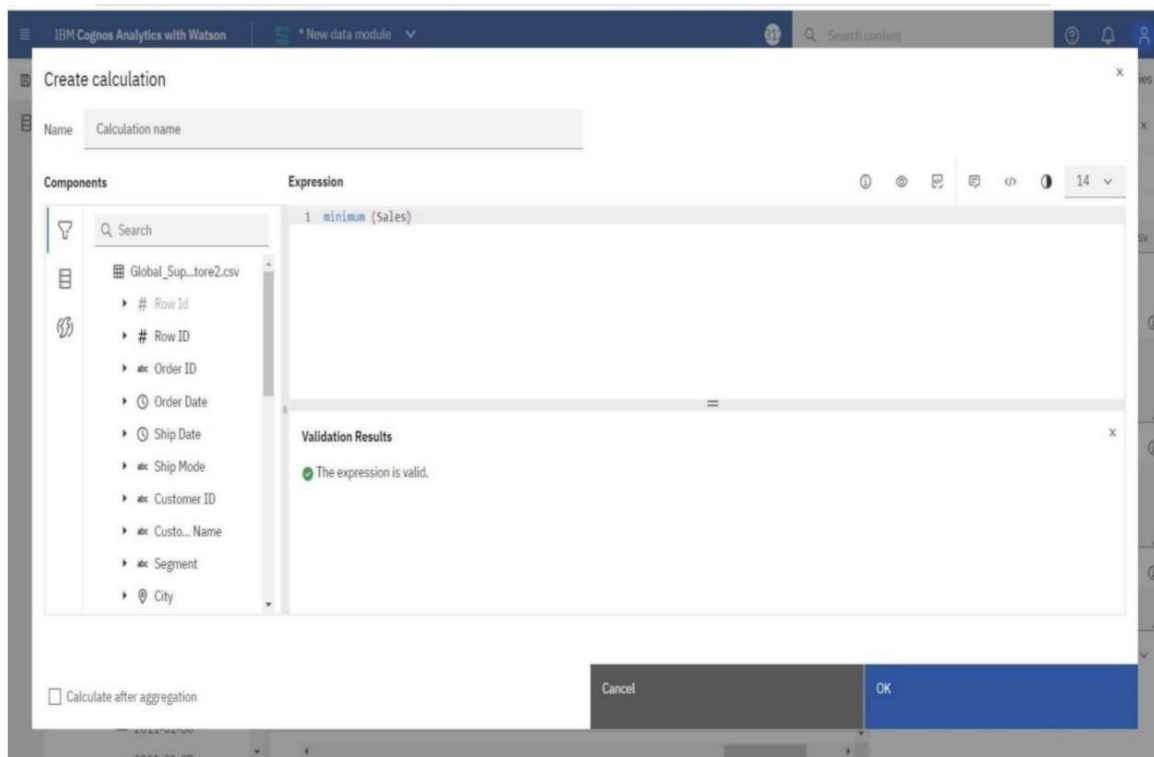
- Hence calculations for day, month and year are completed.
- Similarly calculations for Target sales, Min sales, Max sales And Middle range sales are done.

Min sales -> Minimum(sales)

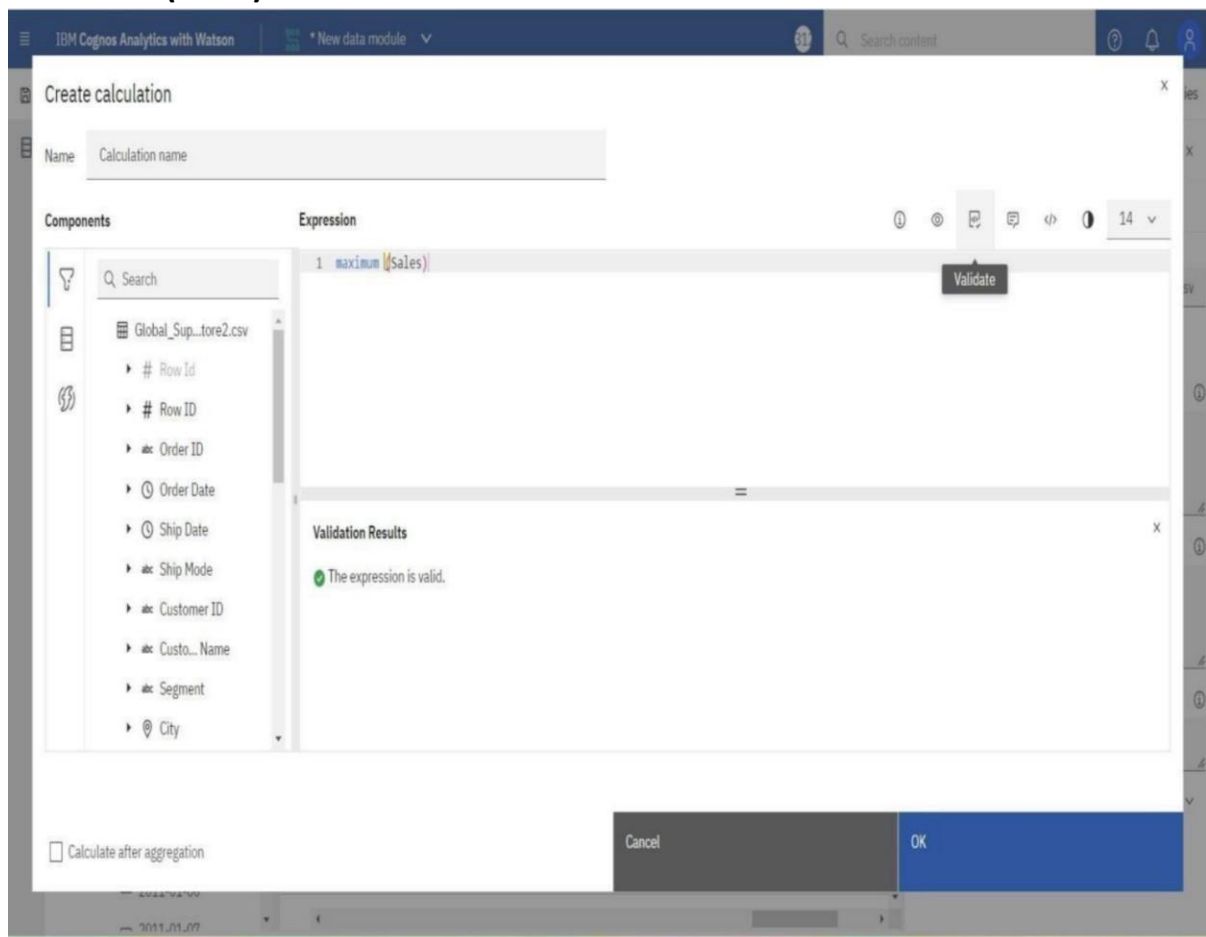
Max sales, Target sales -> Maximum(sales)

Middle range sales -> Average (sales)

Minimum(sales)



Maximum(sales)



Middle Range(sales)

