

# DATA MODULE

Team ID	PNT2022TMID00819
Project Name	Global Sales Data Analytics

## 1. Downloading the Data Set

```
+ Code + Text
Reconnect Editing
[ ] import os
os.environ["KAGGLE_CONFIG_DIR"] = "/content"

[ ] ! kaggle datasets download -d apoorvaappz/global-super-store-dataset

Warning: Your Kaggle API key is readable by other users on this system! To fix this, you can run 'chmod 600 /content/kaggle.json'
Downloading global-super-store-dataset.zip to /content
45% 5.00M/11.1M [00:00<00:00, 46.1MB/s]
100% 11.1M/11.1M [00:00<00:00, 68.2MB/s]

[ ] ! unzip *.zip && rm *.zip

Archive: global-super-store-dataset.zip
  inflating: Global_Superstore2.csv
  inflating: Global_Superstore2.xlsx
/bin/bash: rm.zip: command not found

[ ] import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
```

## 2. Removal Of Unwanted, Unnecessary Data and Encoding to UTF-8 Format

File	Home	Insert	Draw	Page Layout	Formulas	Data	Review	View	Help	Share								
<div>Undo   Clipboard   Font   Alignment   Number   Styles   Cells   Editing</div>																		
K6   Senegal																		
1	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer	Customer Segment	City	State	Country	Postal Code	Market	Region	Product Category	Sub-Category	Product Name	Sales	Quantity
2	32298	CA-2012-1	#####	#####	Same Day	RH-19495	Rick Hanss Consumer	New York	New York	United States	10024	US	East	TEC-AC-1C	Technology Accessories	Plantronics	2309.65	1
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	3709.395
4	25330	IN-2013-7	#####	#####	First Class	CR-12730	Craig Reitz Consumer	Brisbane	Queensland	Australia		APAC	Oceania	TEC-PH-1C	Technology	Phones	Nokia	5175.171
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
6	47221	SG-2013-4	#####	#####	Same Day	RH-9495	Rick Hanss Consumer	Dakar	Dakar	Senegal		Africa	Africa	TEC-SHA-1C	Technology	Copiers	Sharp	2832.96
7	22732	IN-2013-4	#####	#####	Second Class	JM-15655	Jim Mitchell Corporate	Sydney	New South Wales	Australia		APAC	Oceania	TEC-PH-1C	Technology	Phones	Samsung	2862.675
8	30570	IN-2011-8	#####	#####	First Class	TS-21340	Toby Swin Consumer	Porirua	Wellington	New Zealand		APAC	Oceania	FUR-CH-1C	Furniture	Chairs	Novimex	1822.08
9	31192	IN-2012-8	#####	#####	Standard	MB-18085	Mick Brown Consumer	Hamilton	Waikato	New Zealand		APAC	Oceania	FUR-TA-1C	Furniture	Tables	Chromcraft	5244.84
10	40155	CA-2014-1	#####	#####	Standard	JW-15220	Jane Warr Corporate	Sacramento	California	United States	95823	US	West	OFF-BI-10I	Office Supplies	Binders	Fellowes	5083.96
11	40936	CA-2012-1	#####	#####	Second Class	JH-15985	Joseph Ho Consumer	Concord	North Carolina	United States	28027	US	South	FUR-TA-1C	Furniture	Tables	Chromcraft	4297.644
12	34577	CA-2011-1	#####	#####	Second Class	GM-14695	Greg Maxwell Corporate	Alexandria	Virginia	United States	22304	US	South	OFF-SU-1C	Office Supplies	Supplies	Martin	4164.05
13	28879	ID-2012-2	#####	#####	First Class	AJ-10780	Anthony J Corporate	Kabul	Kabul	Afghanistan		APAC	Central Asia	FUR-TA-1C	Furniture	Tables	Bevis	4626.15
14	45794	SA-2011-1	#####	#####	Second Class	MM-7260	Magdalena Consumer	Jizan	Jizan	Saudi Arabia		EMEA	EMEA	TEC-CIS-1C	Technology	Phones	Cisco	2616.96
15	4132	MX-2012-7	#####	#####	Same Day	VF-21715	Vicky Frey Home Office	Toledo	Parana	Brazil		LATAM	South	FUR-CH-1C	Furniture	Chairs	Harbour	2221.8
16	27704	IN-2013-7	#####	#####	Second Class	PF-19120	Peter Fulle Consumer	Mudanjar	Heilongjiang	China		APAC	North Asia	OFF-AP-1C	Office Supplies	Appliances	KitchenAid	3701.52
17	13779	ES-2014-5	#####	#####	Second Class	BP-11185	Ben Peters Corporate	Paris	Ile-de-France	France		EU	Central	OFF-AP-1C	Office Supplies	Appliances	Breville	1869.588
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
19	12069	ES-2014-1	#####	#####	Standard	PJ-18835	Patrick Jor Corporate	Prato	Tuscany	Italy		EU	South	OFF-AP-1C	Office Supplies	Appliances	Hoover	7958.58
20	22096	IN-2014-1	#####	#####	First Class	JS-15685	Jim Sink Corporate	Townsville	Queensland	Australia		APAC	Oceania	TEC-CO-1C	Technology	Copiers	Brother	2565.594
Global Superstore data																		
Ready   Accessibility: Unavailable																		

### 3. Loading Data in Database

IBM Db2 on Cloud

Load DataLoad HistoryTablesViewsIndexesAliasesMQTsSequencesApplication objects

SourceTargetDefineFinalize

You are loading the file *encoded-Global\_Superstore data REVISED.csv* into *WTW00999.ENCODED\_DATA\_CSV*

Code page (character encoding): 1208 (UTF-8)Separator: ,Header in first row: ☒Time & date format:Detect data types: ☐

ROW_ID	ORDER_ID	ORDER_DATE	SHIP_DATE	SHIP_MODE	CUSTOMER_ID	CUSTOMER_NAME
INTEGER	VARCHAR(15)	DATE	DATE	VARCHAR(14)	VARCHAR(8)	VARCHAR(100)
1	32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day	RH-19495 Rick Hz
2	26341	IN-2013-77878	05-02-2013	07-02-2013	Second Class	JR-16210 Justin I
3	25330	IN-2013-71249	17-10-2013	18-10-2013	First Class	CR-12730 Craig R
4	13524	ES-2013-1579342	28-01-2013	30-01-2013	First Class	KM-16375 Katheri
5	47221	SG-2013-4320	05-11-2013	06-11-2013	Same Day	RH-9495 Rick Hz
6	22732	IN-2013-42360	28-06-2013	01-07-2013	Second Class	JM-15655 Jim Mit
7	30570	IN-2011-81826	07-11-2011	09-11-2011	First Class	TS-21340 Tobv Si

BackNext

### 4. Uploading to Cognos Analytics

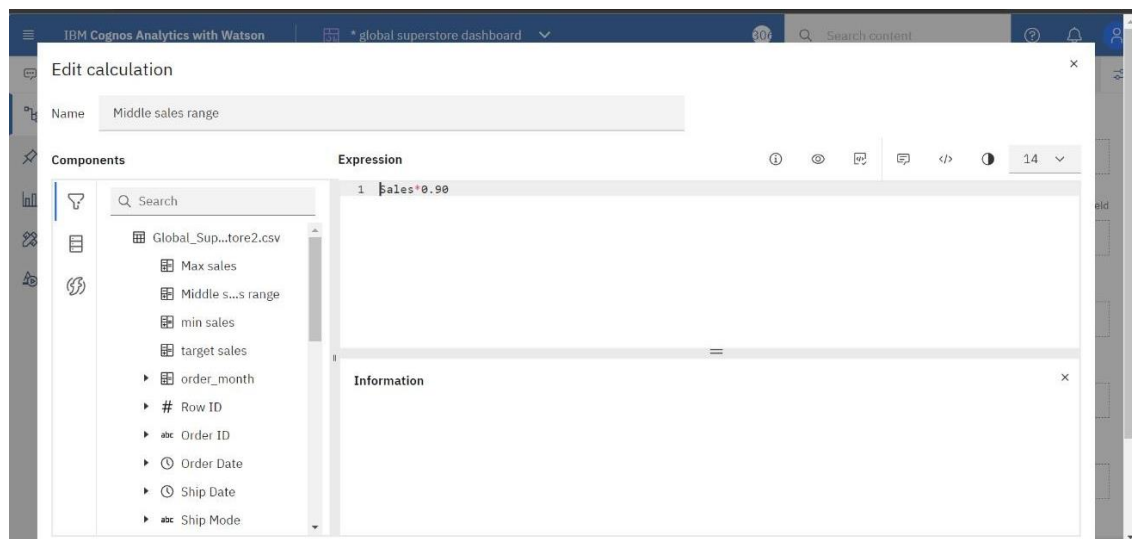
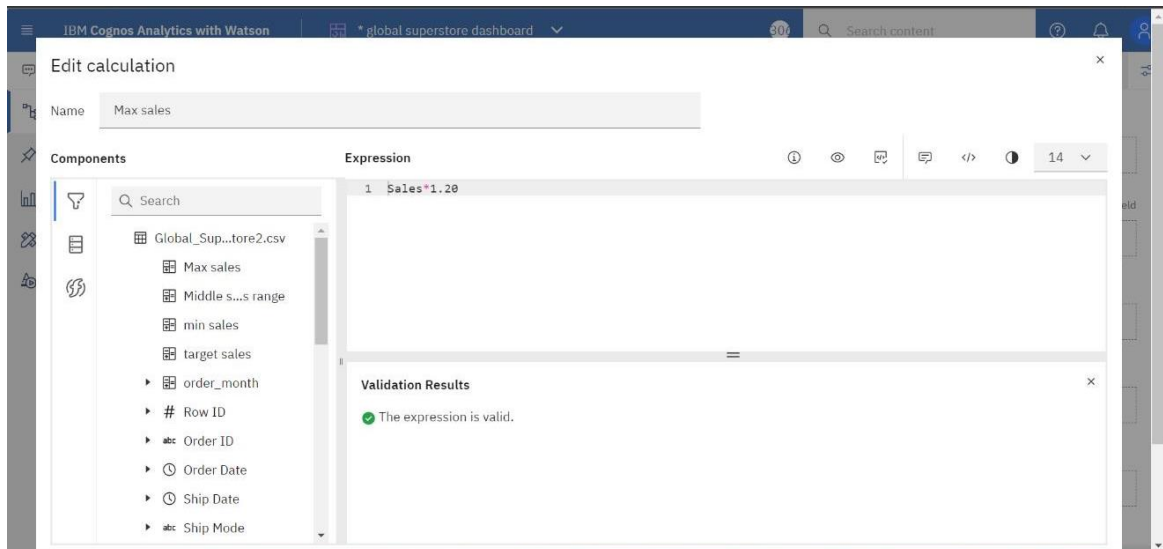
Properties

Data ...dule

GridRelationshipsCustom tables

Row Id	Row ID	Order ID	Order Date	Ship Date	Ship Mode
1	1	CA-2016-152156	11/8/16	11/11/16	Second Class
2	2	CA-2016-152156	11/8/16	11/11/16	Second Class
3	3	CA-2016-138688	6/12/16	6/16/16	Second Class
4	4	US-2015-108966	10/11/15	10/18/15	Standard Class
5	5	US-2015-108966	10/11/15	10/18/15	Standard Class
6	6	CA-2014-115812	6/9/14	6/14/14	Standard Class
7	7	CA-2014-115812	6/9/14	6/14/14	Standard Class
8	8	CA-2014-115812	6/9/14	6/14/14	Standard Class

## 5. Data Wrangling



IBM Cognos Analytics with Watson | \* global superstore dashboard | Search content

### Edit calculation

Name: min sales

Components

- Global\_Sup...tore2.csv
  - Max sales
  - Middle s...s range
  - min sales
  - target sales
  - order\_month
  - # Row ID
  - abc Order ID
  - Order Date
  - Ship Date
  - abc Ship Mode

Expression

1 Sales\*0.80

Information

IBM Cognos Analytics with Watson | \* global superstore dashboard | Search content

### Edit calculation

Name: target sales

Components

- Global\_Sup...tore2.csv
  - Max sales
  - Middle s...s range
  - min sales
  - target sales
  - order\_month
  - # Row ID
  - abc Order ID
  - Order Date
  - Ship Date
  - abc Ship Mode

Expression

1 Sales\*1.10

Information

