

DATA MODULE

Team ID	PNT2022TMID13846
Project Name	Global Sales Data Analytics

1. DATA DOWNLOAD

```
!pip install kaggle
!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 DATA AND ENCODING TO UTF-8 FORMAT

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
32298	CA-2012-1	2012-01-01	2012-01-01	Standard	Rick Hansi	Consumer	New York	New York	United States	10024	US	East	TEC-AC-1C	Technology	Accessories	Plantronics	2309.65
26341	IN-2013-7	2013-07-01	2013-07-01	Second Class	Justin Ritti	Corporate	Wollongong	New South Wales	Australia		APAC	Oceania	FUR-CH-1C	Furniture	Chairs	Novimex E	3709.395
25330	IN-2013-7	2013-07-01	2013-07-01	First Class	Craig Reitz	Consumer	Brisbane	Queensland	Australia		APAC	Oceania	TEC-PH-1C	Technology	Phones	Nokia Smz	5175.171
13524	ES-2013-1	2013-01-01	2013-01-01	First Class	Katherine Home	Officer	Berlin	Berlin	Germany		EU	Central	TEC-PH-1C	Technology	Phones	Motorola	2892.51
47221	SG-2013-4	2013-04-01	2013-04-01	Same Day	Rick Hansi	Consumer	Dakar	Dakar	Senegal		Africa	Africa	TEC-SHA-1C	Technology	Copiers	Sharp Wir	2832.96
22732	IN-2013-4	2013-04-01	2013-04-01	Second Class	Jim Mitche	Corporate	Sydney	New South Wales	Australia		APAC	Oceania	TEC-PH-1C	Technology	Phones	Samsung t	2862.675
30570	IN-2011-8	2011-08-01	2011-08-01	First Class	Toby Swin	Consumer	Porirua	Wellington	New Zealand		APAC	Oceania	FUR-CH-1C	Furniture	Chairs	Novimex E	1822.08
31192	IN-2012-8	2012-08-01	2012-08-01	Standard	Mick Brox	Consumer	Hamilton	Waikato	New Zealand		APAC	Oceania	FUR-TA-1C	Furniture	Tables	Chromcra	5244.84
40155	CA-2014-1	2014-01-01	2014-01-01	Standard	Jane Wacc	Corporate	Sacramento	California	United States	95823	US	West	OFF-BI-10	Office Supplies	Binders	Fellowes F	5083.96
40936	CA-2012-1	2012-01-01	2012-01-01	Second Class	Joseph Ho	Consumer	Concord	North Carolina	United States	28027	US	South	FUR-TA-1C	Furniture	Tables	Chromcra	4297.644
34577	CA-2011-1	2011-01-01	2011-01-01	Second Class	Greg Max	Corporate	Alexandria	Virginia	United States	22304	US	South	OFF-SU-1C	Office Supplies	Supplies	Martin Yal	4164.05
28879	ID-2012-2	2012-02-01	2012-02-01	First Class	Anthony J	Corporate	Kabul	Kabul	Afghanistan		APAC	Central As	FUR-TA-1C	Furniture	Tables	Bevis Conf	4626.15
45794	SA-2011-1	2011-01-01	2011-01-01	Second Class	Magdelen	Consumer	Jizan	Jizan	Saudi Arabia		EMEA	EMEA	TEC-CIS-1C	Technology	Phones	Cisco Sma	2616.96
4132	MX-2012-2	2012-02-01	2012-02-01	Same Day	Vicky Frey	Home Officer	Toledo	Parana	Brazil		LATAM	South	FUR-CH-1C	Furniture	Chairs	Harbour C	2221.8
27704	IN-2013-7	2013-07-01	2013-07-01	Second Class	Peter Full	Consumer	Mudanjiang	Heilongjiang	China		APAC	North Asia	OFF-AP-1C	Office Supplies	Appliances	KitchenAic	3701.52
13779	IN-2014-5	2014-05-01	2014-05-01	Second Class	Ben Peter	Corporate	Paris	Ile-de-France	France		EU	Central	OFF-AP-1C	Office Supplies	Appliances	Breville Re	1869.588
36178	CA-2014-1	2014-01-01	2014-01-01	Second Class	Thomas B	Corporate	Henderson	Kentucky	United States	42420	US	South	TEC-AC-1C	Technology	Accessories	Logitech d	2249.91
12069	ES-2014-1	2014-01-01	2014-01-01	Standard	Patrick Jor	Corporate	Prato	Tuscany	Italy		EU	South	OFF-AP-1C	Office Supplies	Appliances	Hoover Sti	7958.58
22096	IN-2014-1	2014-01-01	2014-01-01	First Class	Jim Sink	Corporate	Townsville	Queensland	Australia		APAC	Oceania	TEC-CC-1C	Technology	Copiers	Brother Fe	2565.594

3.LOADING DATA IN DATABASE

The screenshot shows the IBM Db2 on Cloud web interface. The top navigation bar includes links for Load Data, Load History, Tables, Views, Indexes, Aliases, MQTs, Sequences, and Application objects. The main content area displays the 'Load Data' wizard, which is currently in the 'Define' step. The wizard indicates that data is being loaded from the file 'encoded-Global_Superstore data REVISED.csv' into the table 'WTW00999.ENCODDED_DATA_CSV'. The 'Code page' is set to 1208 (UTF-8), and the 'Separator' is a comma. The 'Header in first row' checkbox is checked. The 'Time & date format' is set to 'Detect data types'. The 'Detect data types' checkbox is also checked. Below these settings, a table shows the data being loaded, with columns: ROW_ID, ORDER_ID, ORDER_DATE, SHIP_DATE, SHIP_MODE, CUSTOMER_ID, and CUSTOMER_NAME. The table contains 7 rows of data.

ROW_ID	ORDER_ID	ORDER_DATE	SHIP_DATE	SHIP_MODE	CUSTOMER_ID	CUSTOMER_NAME
1	32298	31-07-2012	31-07-2012	Same Day	RH-19495	Rick H
2	26341	05-02-2013	07-02-2013	Second Class	JR-16210	Justin I
3	25330	17-10-2013	18-10-2013	First Class	CR-12730	Craig R
4	13524	28-01-2013	30-01-2013	First Class	KM-16375	Katheri
5	47221	05-11-2013	06-11-2013	Same Day	RH-9495	Rick H
6	22732	28-06-2013	01-07-2013	Second Class	JM-15655	Jim Mit
7	30570	07-11-2011	09-11-2011	First Class	TS-21340	Tobv S

4.UPLOADNG TO COGNOS ANALYTICS

The screenshot shows the IBM Cognos Analytics web interface. The top navigation bar includes links for Data, Relationships, and Custom tables. The main content area displays the 'Data' module, which is currently in the 'Grid' view. The 'Grid' view shows a table with columns: Row Id, Row ID, Order ID, Order Date, Ship Date, and Ship Mode. The table contains 8 rows of data.

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

The screenshot shows the 'Edit calculation' window in IBM Cognos Analytics. The 'Name' field is set to 'Max sales'. The 'Expression' field contains the formula `Sales*1.20`. The 'Components' panel on the left lists available fields: 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, and abc Ship Mode. A 'Validation Results' panel at the bottom right displays a green checkmark and the message 'The expression is valid.' The window is titled 'us3.ca.analytics.ibm.com/bi/?perspective=dashboard&id=i27591E791B2C4F53A5C77F84CBD4733E&objRef=i27591E791B2C4F53A5C77F84CBD4733...'.

The screenshot shows the 'Edit calculation' window in IBM Cognos Analytics. The 'Name' field is set to 'Middle sales range'. The 'Expression' field contains the formula `Sales*0.90`. The 'Components' panel on the left lists available fields: 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, and abc Ship Mode. An 'Information' panel at the bottom right is visible. The window is titled 'us3.ca.analytics.ibm.com/bi/?perspective=dashboard&id=i27591E791B2C4F53A5C77F84CBD4733E&objRef=i27591E791B2C4F53A5C77F84CBD4733...'.





