

DATA MODULE

Team ID	PNT2022TMID17402
Project Name	Global Sales Data Analytics

1. Downloading the Data Set

```
+ Code + Text
[ ] ! 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

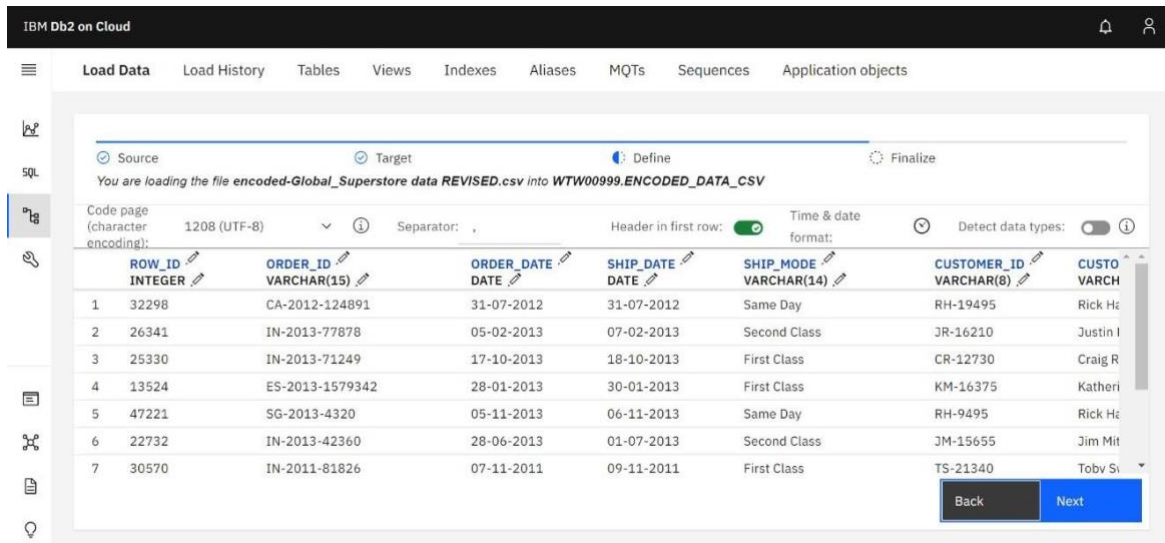
Archive: global-super-store-dataset.zip
  inflating: Global_Superstore2.csv
  inflating: Global_Superstore2.xlsx
/bin/bash: rm.zip: command not found

[ ] 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

FileHomeInsertDrawPage LayoutFormulasDataReviewViewHelp

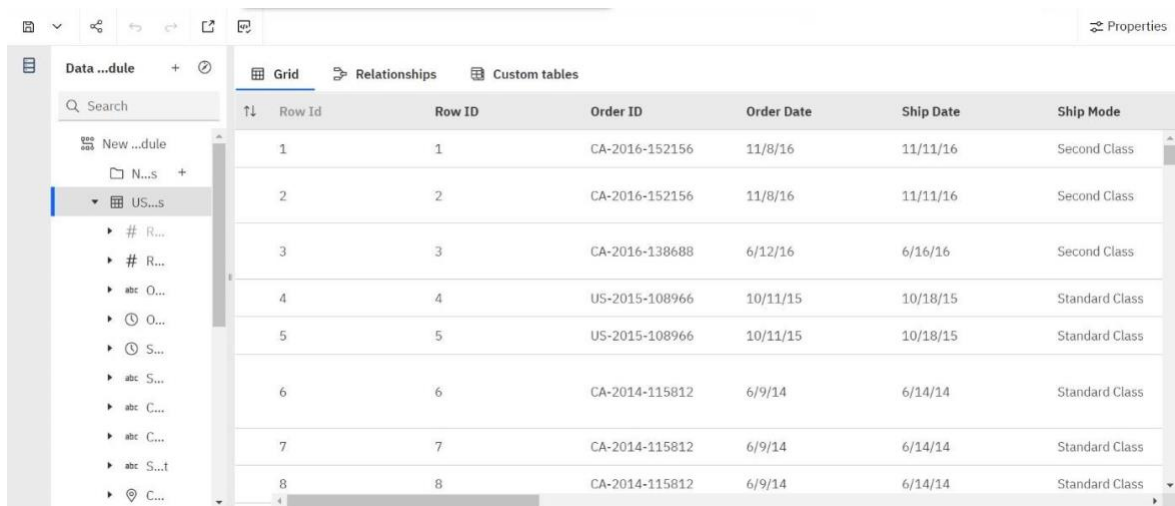
3. Loading Data in Database



The screenshot shows the 'Load Data' interface in IBM Db2 on Cloud. The 'Define' step is active, showing the source file 'encoded-Global_Superstore data REVISED.csv' being loaded into the target table 'WTW00999.ENCODED_DATA_CSV'. The interface displays a preview of the data with columns: ROW_ID, ORDER_ID, ORDER_DATE, SHIP_DATE, SHIP_MODE, CUSTOMER_ID, and CUSTO VARCH. The data is presented in a table with 7 rows and 7 columns. The 'Back' and 'Next' buttons are visible at the bottom right.

ROW_ID	ORDER_ID	ORDER_DATE	SHIP_DATE	SHIP_MODE	CUSTOMER_ID	CUSTO VARCH
1	32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day	RH-19495 Rick Hs
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 Hs
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

4. Uploading to Cognos Analytics



The screenshot shows the 'Data Grid' view in Cognos Analytics. The grid displays data with columns: Row Id, Row ID, Order ID, Order Date, Ship Date, and Ship Mode. The data is presented in a table with 8 rows and 6 columns. The 'Grid' tab is selected, and the 'Properties' panel is visible on the right.

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

