

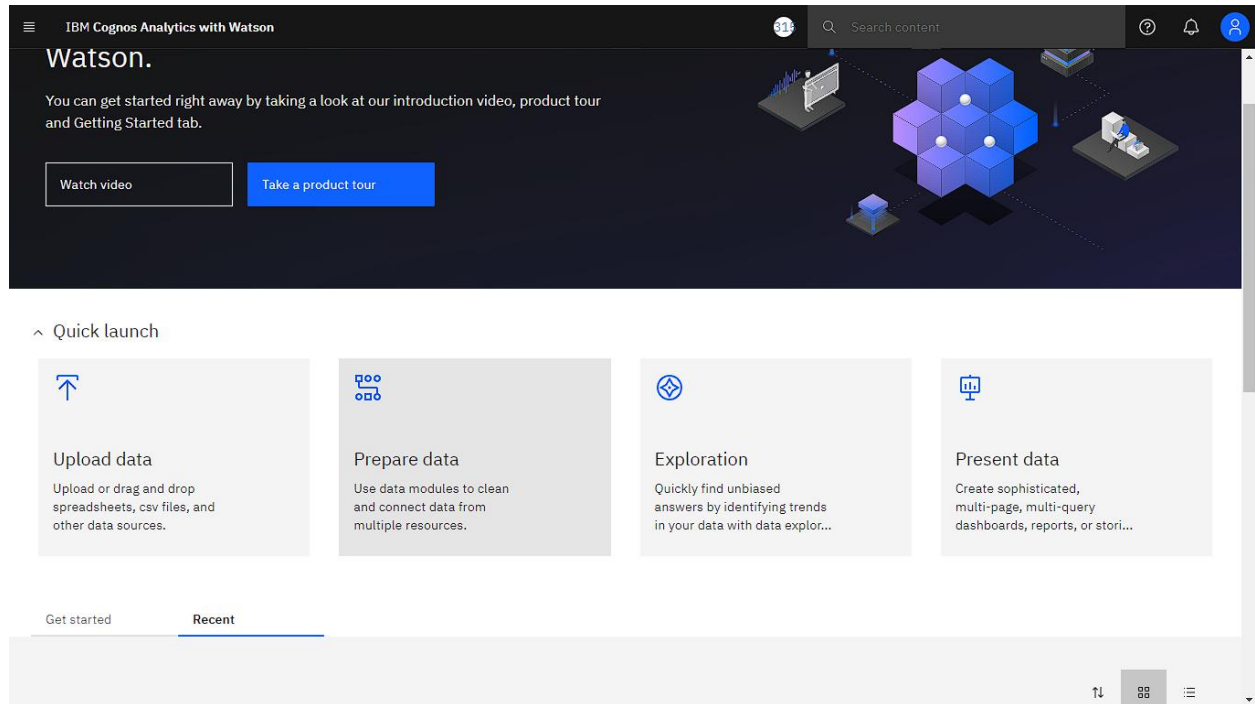
WORKING WITH DATASET

PREPARING THE DATASET

Date	29 OCTOBER 2022
TEAM ID	PNT2022TMID00847
PROJECT NAME	GLOBAL SALES DATA ANALYTICS

PREPARE DATA:

STEP 1: Click on the “Prepare Data” option in the quick launch menu.



STEP 2: Select the .csv files which we want to prepare.

The screenshot shows the IBM Cognos Analytics with Watson interface. The top navigation bar includes the IBM logo, the text 'IBM Cognos Analytics with Watson', a dropdown menu for 'New data module', a search bar, and user profile icons. The main workspace is divided into a left sidebar and a central grid view. The sidebar, titled 'Data module', contains a search bar and a list of data modules: 'New data module', 'Navigation paths', 'Pharma_Mon...Sales.csv' (highlighted with a blue border), 'manager_sur...y_data.csv', and 'general_data.csv'. The central grid view displays a table with 14 rows and 7 columns: 'Row Id', 'Month', 'Customer', 'Period', 'Product', and 'Location'. The data is organized by month (March and January) and customer (Mobil). The table is scrollable, with a scrollbar visible on the right side.

Row Id	Month	Customer	Period	Product	Location
1	March	Mobil	Q1	Afinitor Tab	Port Harcourt
2	March	Mobil	Q1	Brinerdine SCT	Port Harcourt
3	March	Mobil	Q1	Coartem Tab	Port Harcourt
4	March	Mobil	Q1	Codiovan FCT	Port Harcourt
5	March	Mobil	Q1	Femara FCT	Port Harcourt
6	March	Mobil	Q1	Galvusmet FCT	Port Harcourt
7	March	Mobil	Q1	Lescol HGC	Port Harcourt
8	March	Mobil	Q1	Ritalin Tab	Port Harcourt
9	March	Mobil	Q1	Tegretol Tab	Port Harcourt
10	March	Mobil	Q1	Zaditen SRO FCT	Port Harcourt
11	January	Mobil	Q1	Afinitor Tab	Abeokuta
12	January	Mobil	Q1	Brinerdine SCT	Abeokuta
13	January	Mobil	Q1	Coartem Tab	Abeokuta
14	January	Mobil	Q1	Codiovan FCT	Abeokuta

STEP 3: We can see the data tables, here we need to clean the data set using clean option.

This screenshot is identical to the one above, showing the IBM Cognos Analytics with Watson interface. The 'Data module' panel on the left has 'Pharma_Mon...Sales.csv' selected. The central grid view displays a table with 14 rows and 7 columns: 'Row Id', 'Month', 'Customer', 'Period', 'Product', and 'Location'. The data is organized by month (March and January) and customer (Mobil). The table is scrollable, with a scrollbar visible on the right side.

Row Id	Month	Customer	Period	Product	Location
1	March	Mobil	Q1	Afinitor Tab	Port Harcourt
2	March	Mobil	Q1	Brinerdine SCT	Port Harcourt
3	March	Mobil	Q1	Coartem Tab	Port Harcourt
4	March	Mobil	Q1	Codiovan FCT	Port Harcourt
5	March	Mobil	Q1	Femara FCT	Port Harcourt
6	March	Mobil	Q1	Galvusmet FCT	Port Harcourt
7	March	Mobil	Q1	Lescol HGC	Port Harcourt
8	March	Mobil	Q1	Ritalin Tab	Port Harcourt
9	March	Mobil	Q1	Tegretol Tab	Port Harcourt
10	March	Mobil	Q1	Zaditen SRO FCT	Port Harcourt
11	January	Mobil	Q1	Afinitor Tab	Abeokuta
12	January	Mobil	Q1	Brinerdine SCT	Abeokuta
13	January	Mobil	Q1	Coartem Tab	Abeokuta
14	January	Mobil	Q1	Codiovan FCT	Abeokuta

STEP 4: Right click on the column name and select “Clean” option.

The screenshot shows the IBM Cognos Analytics interface. On the left, the 'Data module' pane lists various data sources, including 'manager_s..._data.csv'. The main area displays a data grid with columns: Row Id, Month, Period, Product, and Location. A right-click context menu is open over the 'Month' column, showing options like 'Filter...', 'Create data group...', 'Create navigation path...', 'Split...', 'Hide from users', 'Remove', 'Format data...', 'Clean...', 'Sort descending', 'Sort ascending', and 'Properties'. The 'Clean...' option is highlighted.

Row Id	Month	Period	Product	Location
1	March	1	Afinitor Tab	Port Harcourt
2	March	1	Brinerdine SCT	Port Harcourt
3	March	1	Coartem Tab	Port Harcourt
4	March	1	Codiovan FCT	Port Harcourt
5	March	1	Femara FCT	Port Harcourt
6	March	1	Galvusmet FCT	Port Harcourt
7	March	1	Lescol HGC	Port Harcourt
8	March	1	Ritalin Tab	Port Harcourt
9	March	Q1	Tegretol Tab	Port Harcourt
10	March	Q1	Zaditen SRO FCT	Port Harcourt
11	January	Q1	Afinitor Tab	Abeokuta
12	January	Q1	Brinerdine SCT	Abeokuta
13	January	Q1	Coartem Tab	Abeokuta
14	January	Q1	Codiovan FCT	Abeokuta

STEP 5: Enter the conditions to clean the data for eg:replace the null values with 0.

The screenshot shows the IBM Cognos Analytics interface with a 'Clean - PerformanceRating' dialog box open. The dialog box has a title bar with a close button. It contains a section for 'NULL values' with two options: 'Replace this value with NULL' (unchecked) and 'Replace NULL values with' (checked). The 'Replace NULL values with' option has a text input field containing '0'. At the bottom of the dialog are 'Cancel' and 'Clean' buttons. The background shows a data grid with columns: Row Id, EmployeeID, JobInvolvement, and PerformanceRating.

Row Id	EmployeeID	JobInvolvement	PerformanceRating
1	1	3	3
2			
3			
4			
5			
6			
7			
8			
9			
10	10	3	3
11	11	2	3
12	12	3	3
13	13	3	3
14	14	2	3
15	15	2	3

STEP 6: Repeat this step for all Data set files.

The screenshot shows the IBM Cognos Analytics interface. On the left, the 'Data module' pane lists several data sources, including 'PerformanceRating'. The main area displays a data grid with columns: Row Id, EmployeeID, JobInvolvement, and PerformanceRating. A dialog box titled 'Clean - PerformanceRating' is open, showing options for handling NULL values. The 'Replace NULL values with' checkbox is checked, and the value '0' is entered in the adjacent field. The 'Clean' button is highlighted in blue.

Row Id	EmployeeID	JobInvolvement	PerformanceRating
1	1	3	3
2			
3			
4			
5			
6			
7			
8			
9			
10	10	3	3
11	11	2	3
12	12	3	3
13	13	3	3
14	14	2	3
15	15	2	3

STEP 7: Draw relationships between tables to create connection from relationships tab.

The screenshot shows the IBM Cognos Analytics interface in the 'Relationships' tab. The 'Data module' pane on the left lists several data sources, including 'employee_survey_data.csv', 'general_data.csv', and 'manager_survey_data.csv'. The main area displays a diagram view with three data tables: 'general_data.csv', 'employee_survey_data.csv', and 'manager_survey_data.csv'. A 'Diagram settings' panel is visible on the right, showing options for 'Cardinality' and 'Focus mode', and a 'Degrees of separation' slider set to 1.

STEP 8: Draw connection as per need.

