

WORKING WITH DATASET

TEAM ID	PNT2022TMID51053
PROJECT NAME	Analytics for Hospitals' Health-Care Data

Understanding the Dataset

The primary data file we use is test_data.csv consist of 17 Columns with 137057 Rows.

Loading the Dataset

The screenshot displays the IBM Cognos Analytics interface. The top navigation bar includes the IBM Cognos Analytics with Watson logo, a 'Content' dropdown menu, a search bar, and user profile icons. The main content area is titled 'Content' and shows a list of files and folders under the 'My content' tab. The files listed are: Reports (Folder), Data Exploration (Folder), train_data.csv (CSV), test_data.csv (CSV), train_data_dictionary.csv (CSV), and sample_sub.csv (CSV). The 'test_data.csv' file is highlighted.

Below the file list, the 'test_data.csv' file is selected, and the 'Data module' view is shown. The 'Data module' view includes a search bar, a list of columns, and a preview of the data. The columns listed are: Row Id, case_id, Hospital_code, Hospital_type_code, City_Code_Hospital, Hospital_region_code, Department, Ward_Type, Ward_Facility_Code, Bed Grade, patientid, and City_Code_Patient. The preview shows the first 12 rows of data.

Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code
1	318439	21	c	3	Z
2	318440	29	a	4	X
3	318441	26	b	2	Y
4	318442	6	a	6	X
5	318443	28	b	11	X
6	318444	23	a	6	X
7	318445	26	b	2	Y
8	318446	25	e	1	X
9	318447	23	a	6	X
10	318448	23	a	6	X
11	318449	10	e	1	X
12	318450	4	a	4	X

Exploration of Data



