

Project Development Phase

Delivery of Sprint-1

Date	10 NOV 2022
Team ID	PNT2022TMID04319
Project Name	Analytics for Hospital's Health-care Data

DATASET

Data Collection:

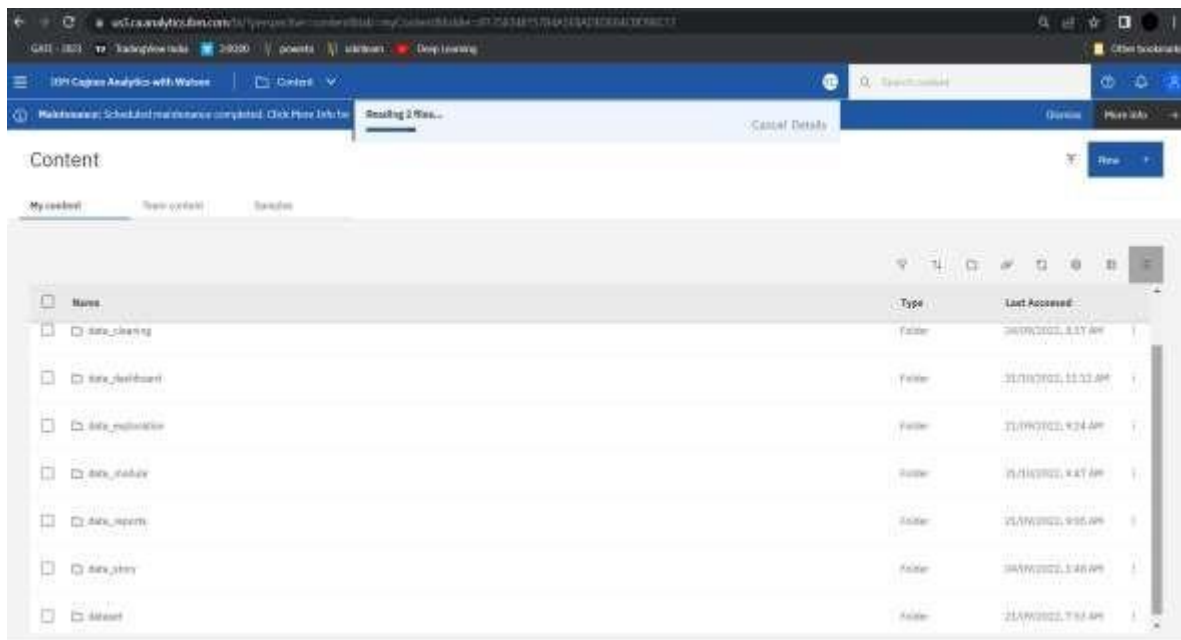
Downloading the Dataset

Dataset :

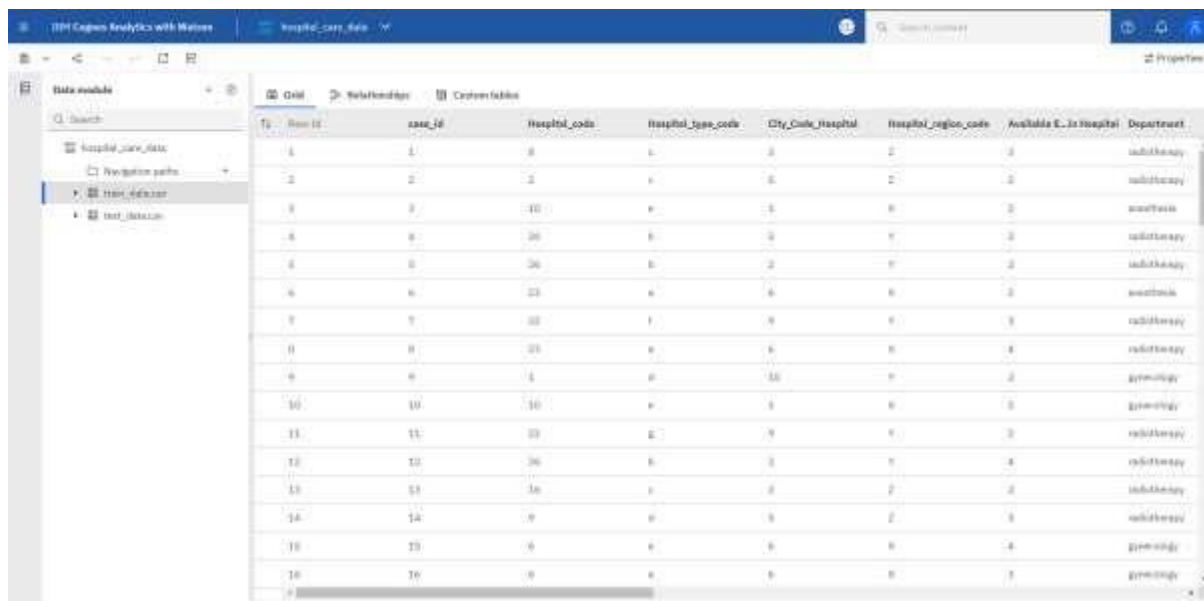
https://drive.google.com/file/d/1AHcg2b_x_pt6LHP2mQVJX-NwxykN4VUx/view?usp=sharing

Load the Dataset:

Tool used: IBM Cognos Analytics

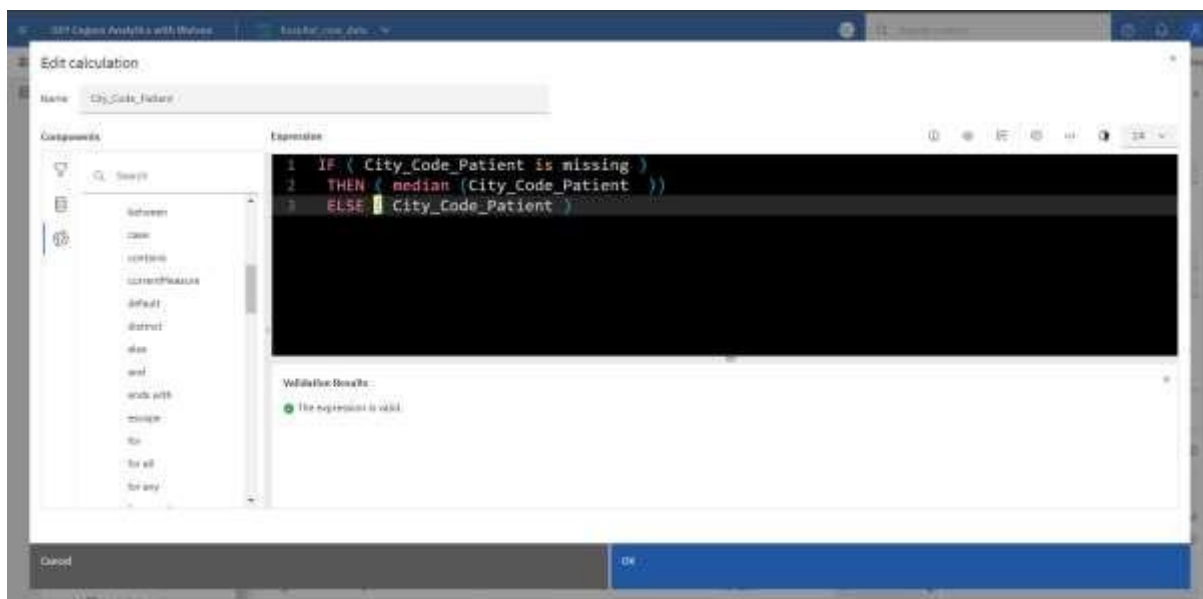


Data Preparation :



Ts	Row ID	case_id	Hospital_code	Hospital_type_code	City_Code_hospital	Hospital_region_code	Available E..to Hospital	Department
1	1	1	8	1	3	2	3	radiology
2	2	2	2	1	5	2	2	radiology
3	3	3	10	1	4	3	2	radiology
4	4	4	26	1	3	1	2	radiology
5	5	5	26	1	2	1	2	radiology
6	6	6	23	1	4	3	2	radiology
7	7	7	22	1	3	1	2	radiology
8	8	8	23	1	5	3	2	radiology
9	9	9	4	1	10	1	2	radiology
10	10	10	10	1	3	3	2	radiology
11	11	11	23	1	3	1	2	radiology
12	12	12	26	1	2	1	2	radiology
13	13	13	26	1	2	2	2	radiology
14	14	14	9	1	3	2	1	radiology
15	15	15	9	1	3	3	2	radiology
16	16	16	9	1	3	3	2	radiology

Null Values cleaning Process :



Name: City_Code_Patient

Expression:

```
1 IF ( City_Code_Patient is missing )  
2 THEN ( median (City_Code_Patient ) )  
3 ELSE City_Code_Patient )
```

Validation Results:

The expression is valid.

The screenshot shows the Databricks workspace interface. The top navigation bar includes the Databricks logo, the workspace name 'Databricks Analytics with Watson', the current schema 'hospital_data', and a search bar. The left sidebar shows the 'Data Explorer' with a tree view of the 'hospital_data' schema. The main area displays the 'hospital_data' table in a table view. The table has 10 columns: ID, Name, Age, Sex, City, Hospital, Available, Department, and Stay. The data is displayed in a table with 10 rows.

ID	Name	Age	Sex	City	Hospital	Available	Department	Stay
1	John Doe	35	M	New York	St. Mary's	Yes	Cardiology	5
2	Jane Smith	42	F	Los Angeles	St. Mary's	Yes	Cardiology	3
3	Michael Johnson	28	M	Chicago	St. Mary's	Yes	Cardiology	7
4	Emily Davis	31	F	San Francisco	St. Mary's	Yes	Cardiology	4
5	Robert Brown	45	M	Seattle	St. Mary's	Yes	Cardiology	6
6	Sarah White	38	F	Portland	St. Mary's	Yes	Cardiology	2
7	David Miller	50	M	Denver	St. Mary's	Yes	Cardiology	8
8	Lisa Wilson	25	F	Phoenix	St. Mary's	Yes	Cardiology	1
9	James Taylor	33	M	San Diego	St. Mary's	Yes	Cardiology	9
10	Amanda Lee	40	F	San Jose	St. Mary's	Yes	Cardiology	6

DBT Capgem Analytics with Watson

hospital.com_data

Search hospital

Properties

Data table

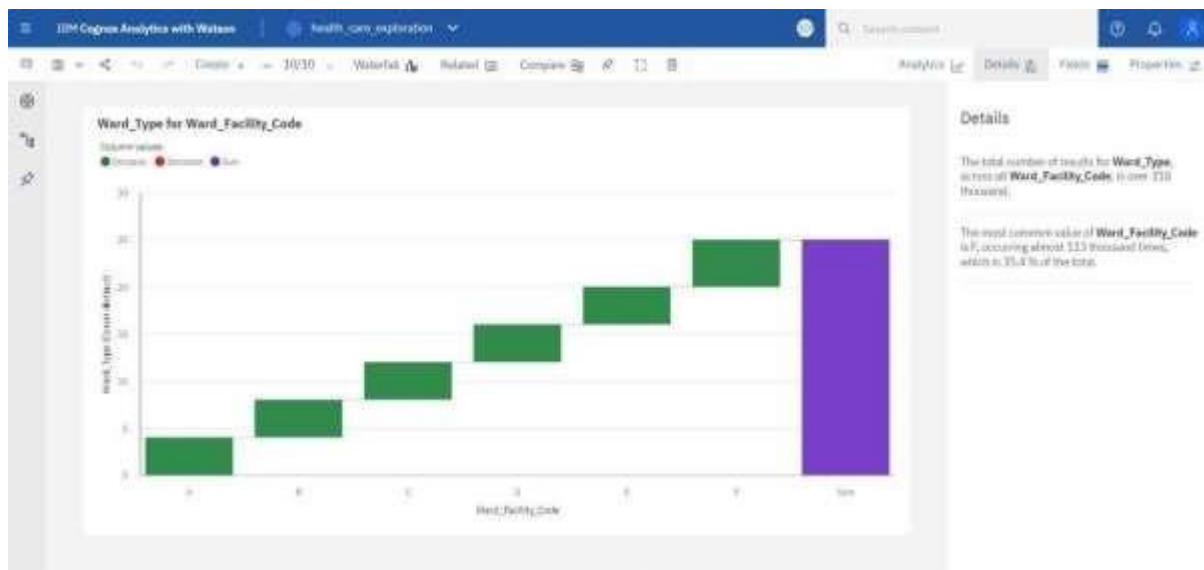
Field Relationships Custom Tables

Search

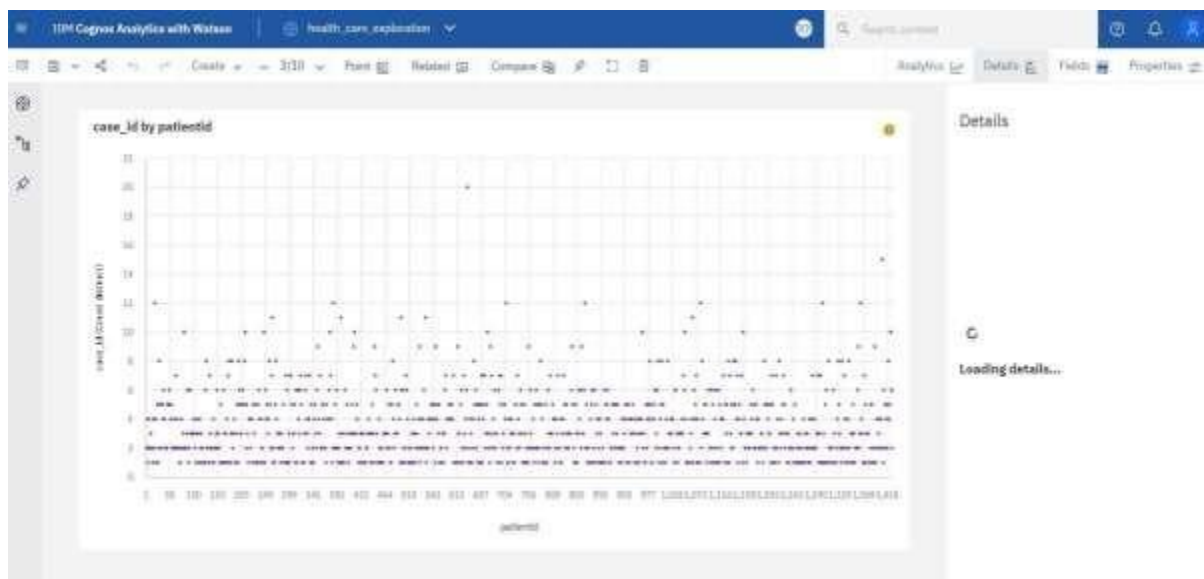
Grade	patientID	City_Code_Patient	Type of Admission	Severity of Illness	Visitors with Patient	Age	Admission_Deposit	Stay
	31281	1	Emergency	Extreme	0	31-40	2921	0-10
	31287	1	Trauma	Extreme	0	31-40	3954	41-50
	31297	1	Trauma	Extreme	0	31-40	4740	31-40
	31397	1	Trauma	Extreme	0	31-40	1273	41-50
	31597	1	Trauma	Extreme	0	31-40	2558	41-50
	31287	1	Trauma	Extreme	0	31-60	4449	31-20
	31287	1	Emergency	Extreme	0	31-60	6107	0-10
	31287	1	Trauma	Extreme	0	31-60	3875	41-50
	31287	1	Trauma	Extreme	0	31-40	7023	31-40
	31287	1	Trauma	Extreme	0	31-40	6056	31-40
	31297	1	Urgent	Extreme	0	31-40	3797	21-30
	31281	1	Urgent	Extreme	0	31-40	2891	31-20
	31297	1	Emergency	Extreme	0	31-40	5141	0-10
	31297	1	Urgent	Extreme	0	31-60	3477	21-30
	41213	0	Emergency	Extreme	0	71-80	2685	0-10
	41213	0	Emergency	Extreme	0	71-80	3098	0-10

DATA EXPLORATION :

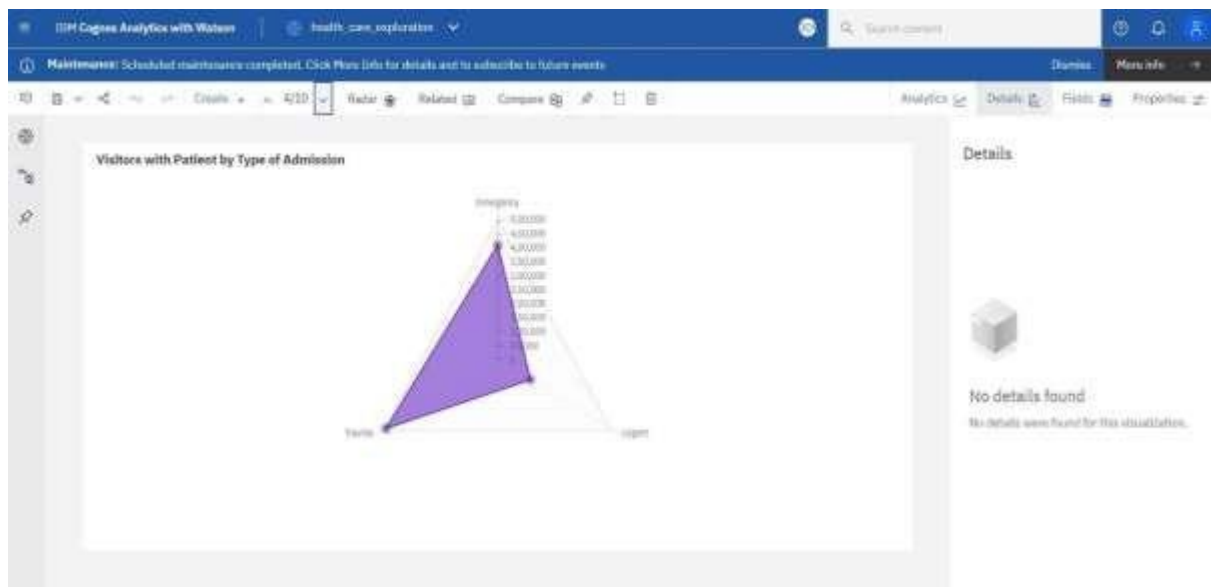
1. Ward_Type for Ward_Facility_Code:



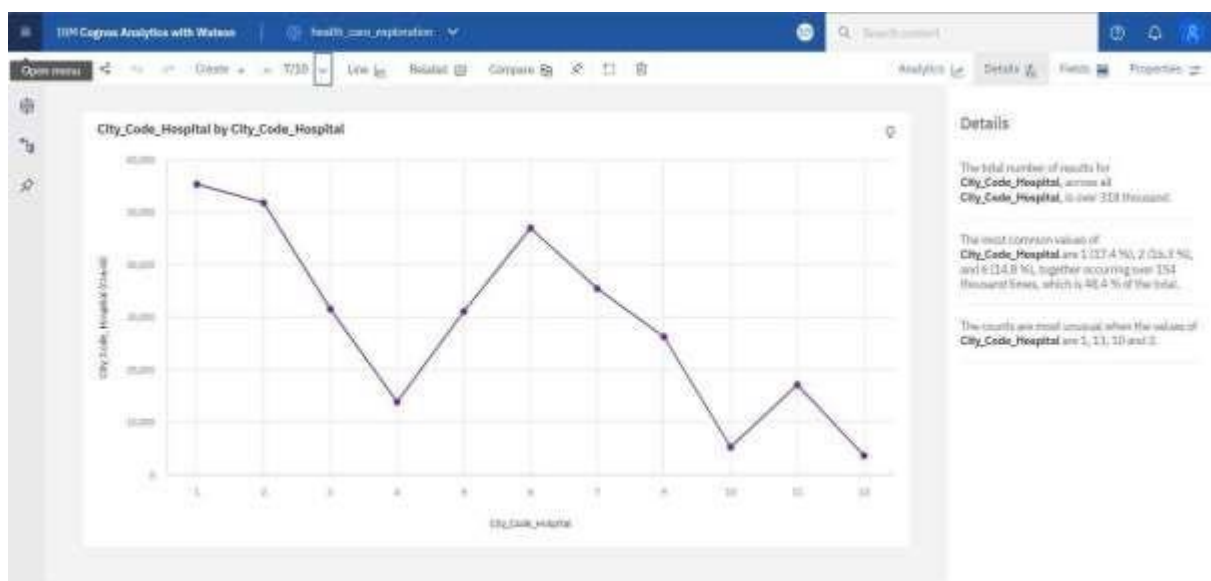
2. Case_id by Patient_id



3. Visitors with Patient by Type of Admission:



4. City_Code_Hospital by City_Code_Hospital:



5.City_Code_Hospital by Department:

