

PROJECT DEVELOPMENT PHASE

SPRINT – 1

Team ID	PNT2022TMID37727
Project Name	Analytics for Hospitals' Health-Care Data
Title	Preparation of data

DATASET LINK :

https://drive.google.com/drive/folders/1qozDgfVw04hGS8Di_SFhUNdUKUdjkhT

DATASET :

The healthcare dataset has three datafiles which was csv files and another one is data dictionary enclosed

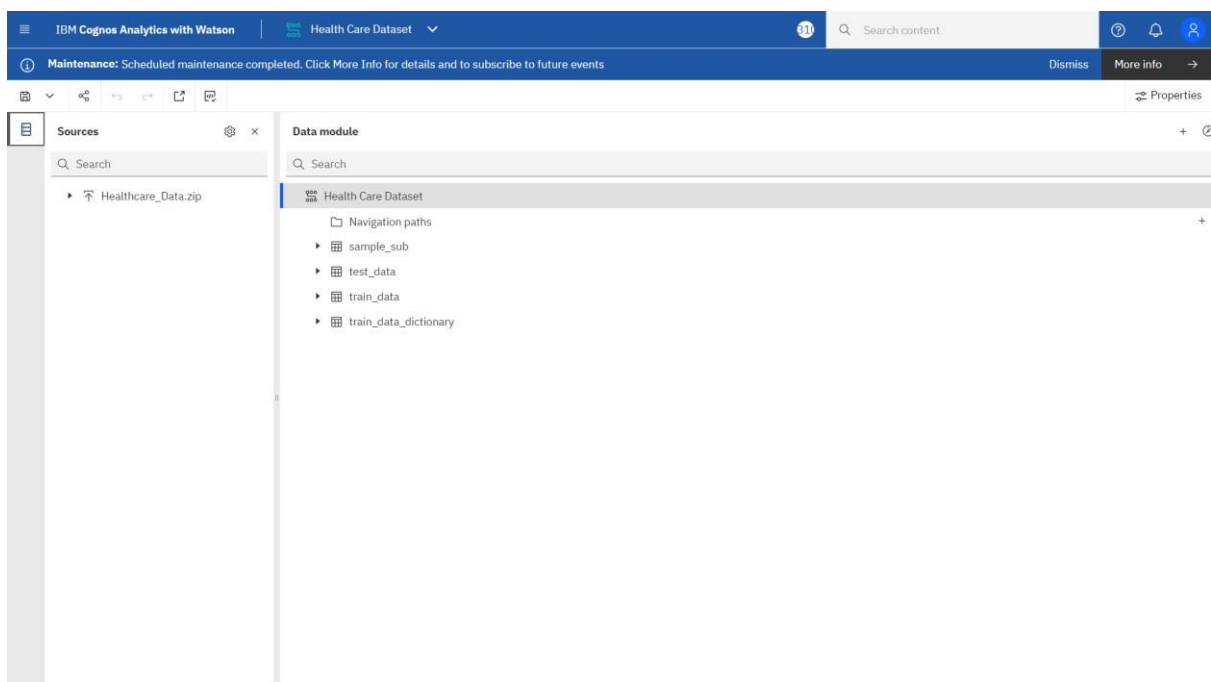
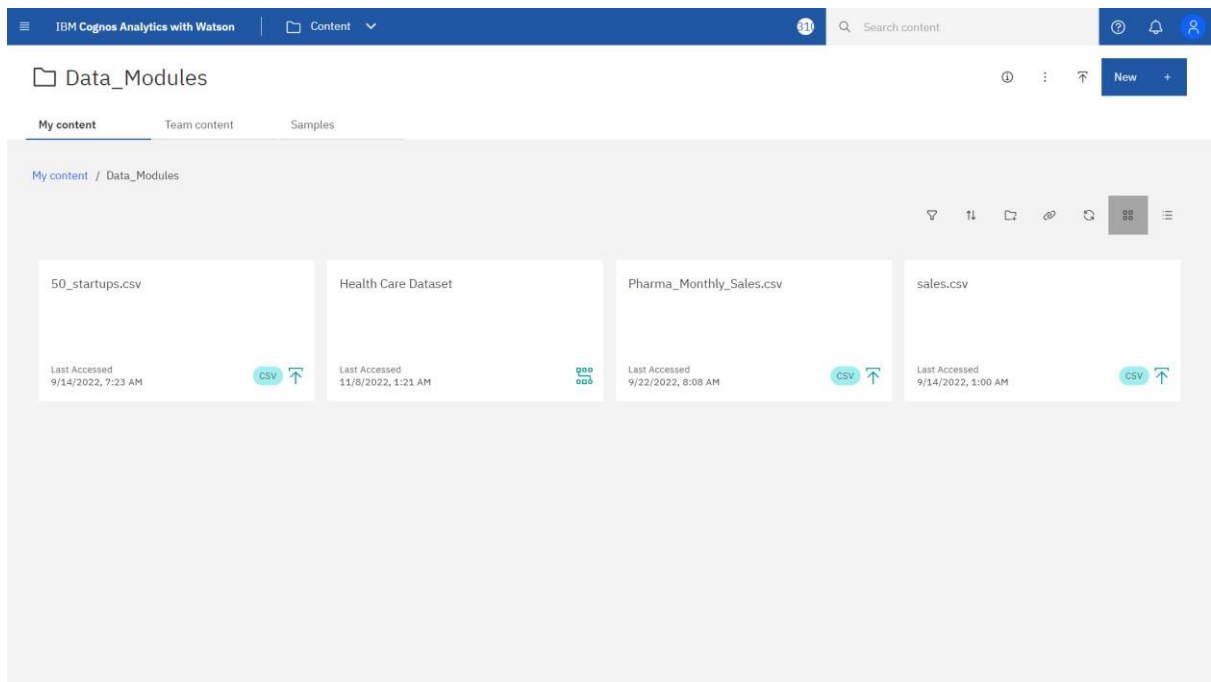
- sample_sub.csv
- test_data.csv
- train_data.csv
- train_data_dictionary.csv

The data file we use is the test_csv which consist of 17 columns and about 137057 rows

LOADING THE DATA :

Loading the given data in the IBM Cognos platform and we have named as Health Care Dataset and it has the dataset of

- sample_sub
- test_data
- train_data
- train_data_dictionary



PREPARATION OF DATASET :

IBM Cognos Analytics with Watson | Health Care Dataset

Data module

- Health Care Dataset
 - Navigation paths
 - sample_sub
 - test_data

Grid

T1	Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available E...in Hc
1	1	318439	21	c	3	Z	3
2	2	318440	29	a	4	X	2
3	3	318441	26	b	2	Y	3
4	4	318442	6	a	6	X	3
5	5	318443	28	b	11	X	2
6	6	318444	23	a	6	X	3
7	7	318445	26	b	2	Y	2
8	8	318446	25	e	1	X	4
9	9	318447	23	a	6	X	4
10	10	318448	23	a	6	X	3
11	11	318449	10	e	1	X	4
12	12	318450	4	a	4	X	3
13	13	318451	16	c	3	Z	4
14	14	318452	28	b	11	X	3
15	15	318453	19	a	7	Y	1
16	16	318454	26	b	2	Y	5
17	17	318455	10	e	1	X	4

DATA PREPARATION :

IBM Cognos Analytics with Watson | * Health Care Dataset

Edit calculation

Name: City_Code_Patient

Components

- test_data
 - # Row Id
 - # case_id
 - # Hospital_code
 - abc Hospital...code
 - City_Co...spital
 - Hospit...code
 - Availabl...Hospital
 - abc Department
 - abc Ward_Type
 - abc Ward..._Code
 - Bed Grade
 - patientid
 - City_Code_Patient
 - abc Type of Admission

Expression

```

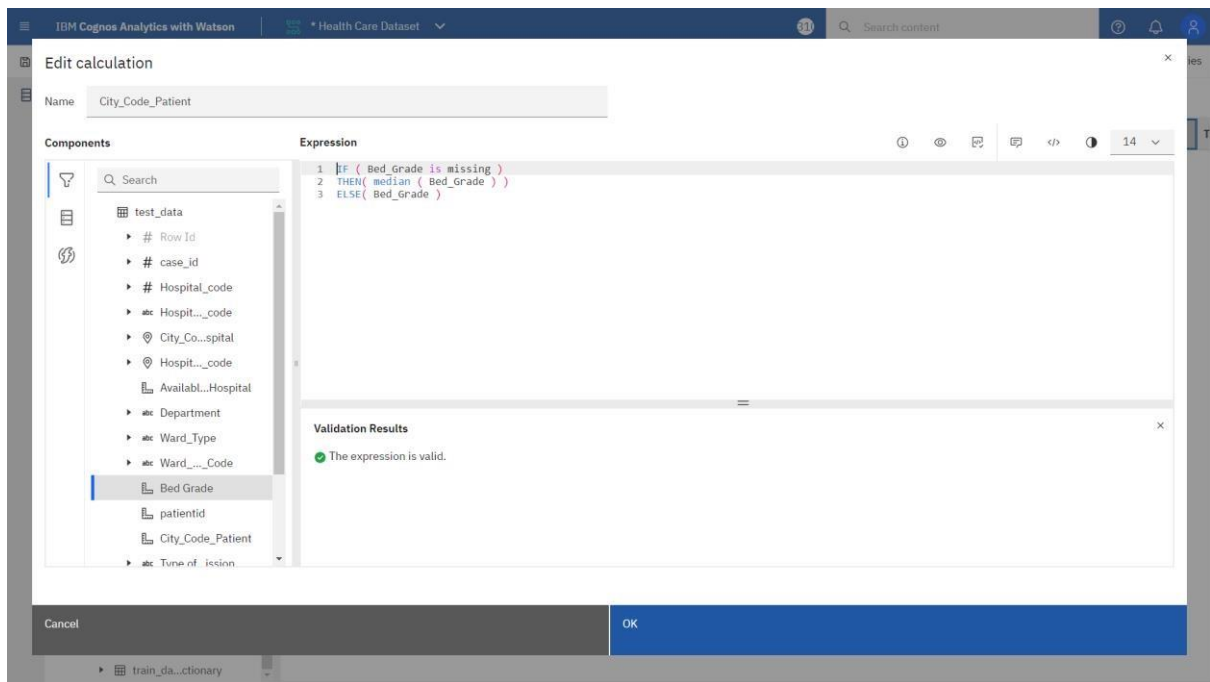
1 IF ( City_code_Patient is missing )
2 THEN( median ( City_code_Patient ) )
3 ELSE( City_Code_Patient )

```

Validation Results

✓ The expression is valid.

Cancel OK



REPLACING THE NULL VALUE :

Data module

Grid

Ward_Facility_Code	Bed Grade	patientid	City_Code_Patient	Type of Admission	Severity of Illness	Visitors with Patient
A	2	17006	2	Emergency	Moderate	2
F	2	17006	2	Trauma	Moderate	4
D	4	17006	2	Emergency	Moderate	3
F	2	17006	2	Trauma	Moderate	3
F	2	17006	2	Trauma	Moderate	4
F	2	17006	2	Trauma	Moderate	2
D	2	17006	2	Trauma	Moderate	2
E	3	95946	8	Emergency	Moderate	2
F	3	95946	8	Trauma	Moderate	2
F	4	95946	8	Urgent	Moderate	2
E	2	95946	8	Trauma	Moderate	2
F	3	95946	8	Emergency	Moderate	6
A	3	95946	8	Trauma	Moderate	3
F	4	95946	8	Urgent	Moderate	2
C	2	40728	8	Emergency	Moderate	4
D	4	40728	8	Emergency	Moderate	4
C	4	40728	8	Emergency	Moderate	2

IBM Cognos Analytics with Watson

Health Care Dataset

11

Search content

Properties

Data module

Search

tra...ta

R...

C...d

H...

abc H...

Cl...l

H...

A...l

abc D...t

abc W...

abc W...

B...

p...d

C...t

abc T...n

abc S...s

V...t

abc Age

A...t

abc Stay

tra...ty

Grid

Relationships

Custom tables

	City_Code_Hospital	Hospital_region_code	Available E...in Hospital	Department	Ward_Type	Ward_Facility_Code	Bed Grade	p
3		Z	3	radiotherapy	R	F	2	
5		Z	2	radiotherapy	S	F	2	
1		X	2	anesthesia	S	E	2	
2		Y	2	radiotherapy	R	D	2	
2		Y	2	radiotherapy	S	D	2	
6		X	2	anesthesia	S	F	2	
9		Y	1	radiotherapy	S	B	3	
6		X	4	radiotherapy	Q	F	3	
10		Y	2	gynecology	R	B	4	
1		X	2	gynecology	S	E	3	
9		Y	2	radiotherapy	S	B	2	
2		Y	4	radiotherapy	R	D	1	
3		Z	2	radiotherapy	R	A	3	
5		Z	3	radiotherapy	S	F	3	
6		X	4	gynecology	Q	F	3	
6		X	3	gynecology	Q	F	3	