

PROJECT DEVELOPMENT PHASE

DELIVERY OF SPRINT 1

TEAM ID	PNT2022TMID13523
DATE	29/10/2022
PROJECT NAME	Analytics for Hospital's Health Care Data

Project Development Phase :

Sprint 1 :

- Data Collection
- Data Preparation

Sprint 2 :

- Data Exploration

Sprint 3 :

- DashBoard Creation

Sprint 4:

- Report creation
 - Story creation
-

SPRINT 1:

Data Collection : The data set used is shared in this drive link

https://drive.google.com/file/d/1slC0MhsJHeuODVkhIXrdNOX_aBDySSB h/view

1. Load the dataset : Using IBM cognos analytics

us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&iid=3019265856_7786e234de564f078541990a392c4722_sessionTemp&objRef=&tid=3019265856_778...

Gmail YouTube Maps Microsoft Teams Artificial Intelligenc... Content Other bookmarks

IBM Cognos Analytics with Watson New data module Search content Properties

Data module +

Search

New data module

Navigation paths

train_data.csv

test_data.csv

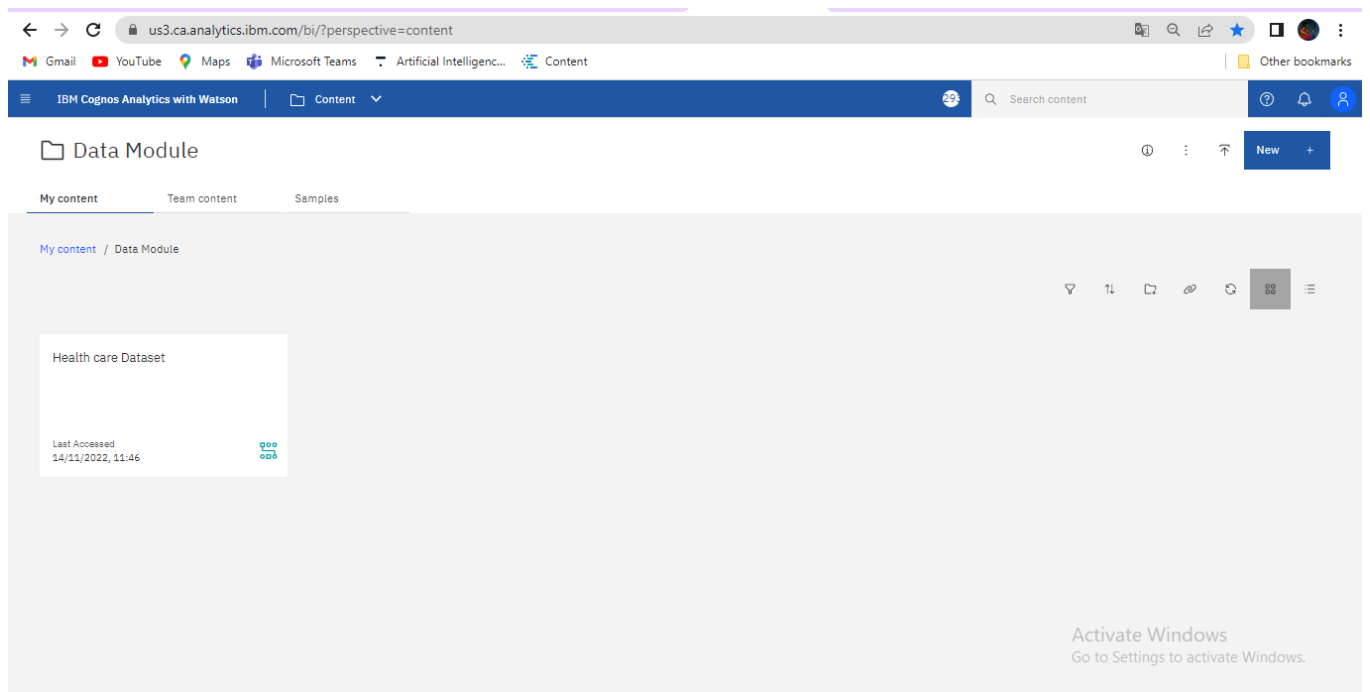
Grid Relationships Custom tables

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

Activate Windows
Go to Settings to activate Windows.

1. Save the datasets inside a **Data Module**:

<https://us3.ca.analytics.ibm.com/bi/?perspective=content>



3. Data Preparation (includes data cleaning) :

Prepare the dataset in the aim to find right insights

us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&id=id4D8239682CB4F4DA139074AFF61A2D9&objRef=id4D8239682CB4F4DA139074AFF61A2D9&tid...

Gmail YouTube Maps Microsoft Teams Artificial Intelligenc... Content

IBM Cognos Analytics with Watson Health care Dataset

Search content

Properties

Data module

Health care Dataset

Navigation paths

train_data.csv

test_data.csv

Row Id

case_id

318439

318440

318441

318442

318443

318444

318445

318446

318447

318448

318449

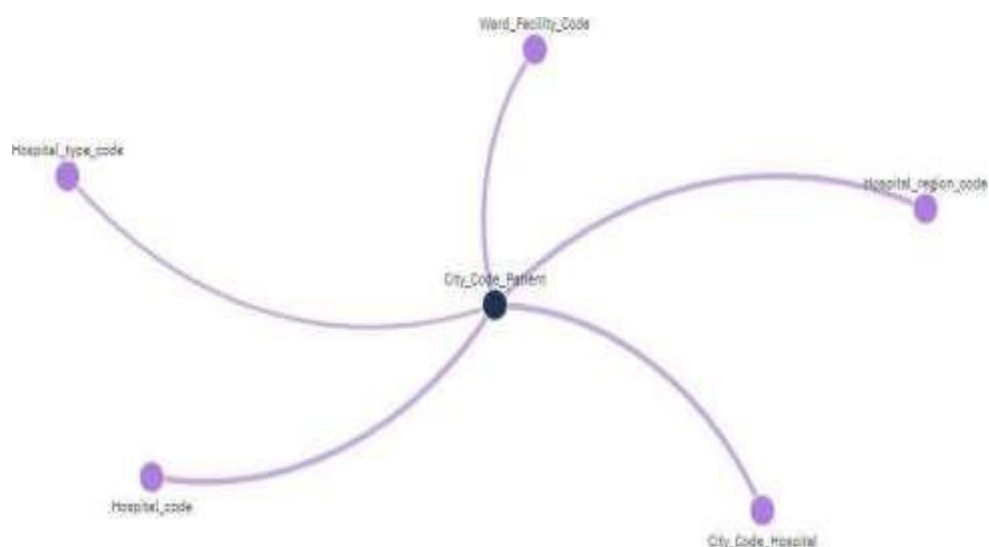
318450

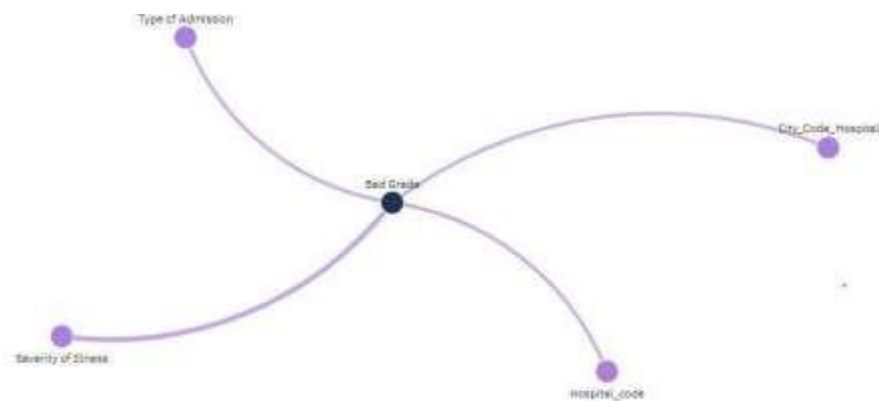
318451

Grid Relationships Custom tables

Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available E...in Hospital	Department
5	318443	28	b	11	X	2	gynecology
6	318444	23	a	6	X	3	gynecology
7	318445	26	b	2	Y	2	gynecology
8	318446	25	e	1	X	4	gynecology
9	318447	23	a	6	X	4	gynecology
10	318448	23	a	6	X	3	gynecology
11	318449	10	e	1	X	4	gynecology
12	318450	4	a	4	X	3	gynecology
13	318451	16	c	3	Z	4	gynecology
14	318452	28	b	11	X	3	gynecology
15	318453	19	a	7	Y	1	gynecology
16	318454	26	b	2	Y	5	gynecology
17	318455	19	a	7	Y	2	gynecology
18	318456	26	b	2	Y	3	gynecology

Activate Windows
Go to Settings to activate Windows.





Relationship diagram ①

1.Null values cleaning process :

Edit calculation

Name: City_Code_Patient

Components

- train_data.csv
 - # Row Id
 - # case_id
 - # Hospital_code
 - # Hospit..._code
 - City_Co...spital
 - Hospit..._code
 - Availabl...Hospital
 - Department
 - Ward_Type
 - Ward..._Code

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.

Prepared Data Set Link : https://us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&id=iD4D82396B2CB4F4DA139074AFF61A2D9&objRef=iD4D82396B2CB4F4DA139074AFF61A2D9&tid=3019265856_7786e234de564f078541990a392c4722_sessionTemp