

# PROJECT DEVELOPMENT PHASE

## SPRINT DELIVERY -1

**TEAM ID : PNT2022TMID28169**

**PROJECT : Analytics for Hospitals Health-Care Data**

**DATE :29 OCT 2022**



The background features a light blue gradient with various medical icons scattered throughout. These include several capsules in different colors (brown, pink, teal), a blister pack of white and pink tablets, a pink digital thermometer showing 36.7, a teal stethoscope, and a pink clipboard with a white grid and a teal stethoscope resting on it. Small pink plus signs are also interspersed among the icons.

# **ANALYTICS FOR HOSPITAL'S** **HEALTH - CARE DATA**

PROJECT CONTRIBUTORS

SWARNALATHA S  
SARDAR SOGHRA SAMREEN  
POOJA K  
VINEETA.M.R

TEAM ID-PNT2022TMID01180

# Project Development Phase

<b>Sprint-1:</b> <ul style="list-style-type: none"><li>✓ <b>Data Collection</b></li><li>✓ <b>Loading of data</b></li><li>✓ <b>Data Preparation</b></li></ul>	<b>Sprint-3:</b> <ul style="list-style-type: none"><li>✓ <b>Dashboard Creation</b></li></ul>
<b>Sprint-2:</b> <ul style="list-style-type: none"><li>✓ <b>Data Exploration</b></li><li>✓ <b>Data Visualization</b></li></ul>	<b>Sprint-4:</b> <ul style="list-style-type: none"><li>✓ <b>Report Creation</b></li><li>✓ <b>Story Creation</b></li></ul>

## 1.HOSP TAL DATASET L NK

[https://drive.google.com/drive/folders/1qozDgfVw04hGS8Di\\_SFhUNdUKUdjkhRT?usp=sharing](https://drive.google.com/drive/folders/1qozDgfVw04hGS8Di_SFhUNdUKUdjkhRT?usp=sharing)

## .2.UNDERSTANDING THE DATA

This project is based on understanding the Health Analytics dataset .  
The data is spread across 3 data files (csv) and one data dictionary enclosed.

- 1.Test\_data.csv
- 2.Sample\_sub.csv
- 3.Train\_data\_dictionary.csv
- 4.Train\_data.csv

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

### 3. LOADING OF DATA

- Connect the data to IBM Cognos.
- Cognos supports connecting to a wide variety of data, stored in a variety of places.
- Loading of the data set in IBM Cognos is important for analyzing.

The screenshot displays the IBM Cognos Analytics 'Content' perspective. The browser address bar shows the URL 'us3.ca.analytics.ibm.com/bi/?perspective=content'. The page title is 'Content'. Below the title, there are tabs for 'My content', 'Team content', and 'Samples'. A table lists three CSV files: 'test\_data.csv', 'train\_data\_dictionary.csv', and 'train\_data.csv'. Each file is marked as an 'Uploaded file' with its last accessed date and time.

Name	Type	Last Accessed
<input type="checkbox"/> test_data.csv CSV	Uploaded file	11/1/2022, 10:46 AM
<input type="checkbox"/> train_data_dictionary.csv CSV	Uploaded file	10/17/2022, 10:22 PM
<input type="checkbox"/> train_data.csv CSV	Uploaded file	10/17/2022, 10:23 PM



# DATASET PREPARATION

IBM Cognos Analytics with Watson

\* New data module

304

Search content

?

Properties

Data module

+

Search

New data module

Navigation paths

test\_data.csv

# Row Id

# case\_id

# Hospital\_code

abc Hospital...pe\_code

City\_Code\_Hospital

Hospital...ion\_code

Available...Hospital

abc Department

abc Ward\_Type

abc Ward\_Fa...ty\_Code

Bed Grade

Grid

Relationships

Custom tables

	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	

# DATA PREPARATION

The screenshot displays the IBM Analytics web interface. At the top, a browser window shows the URL `us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&id=i69959F2C568449BC95AE44A87F892EB6&objRef=i69959F2C568...`. Below the browser, the application window has a header with a search bar and a list of components. The main area is divided into two panes: 'Components' on the left and 'Expression' on the right. The 'Components' pane shows a tree structure under 'SPRINT 1' with items like 'test\_dat...Join (1)', 'sample\_sub.csv', and 'test\_data.csv'. The 'Expression' pane contains a code editor with the following SQL-like expression:

```
1 IF (test_data_csv.City_Code_Patient is missing )
2 THEN (median(test_data_csv.City_Code_Patient))
3 ELSE(test_data_csv.City_Code_Patient)
4
```

Below the code editor, a 'Validation Results' section shows a green checkmark and the text 'The expression is valid.' At the bottom of the window, there is a checkbox labeled 'Calculate after aggregation' which is checked, and two buttons: 'Cancel' and 'OK'. The Windows taskbar at the very bottom shows the time as 11:55 AM on 04-Nov-22.

IBM Cognos Analytics with Watson

SPRINT 1

304

Search content

?

🔔

👤

Create calculation

Name

Calculation name

Components

Expression

🔍 Search

SPRINT 1

- ▶ 📄 test\_data....- Join (1)
- ▶ 📄 sample\_sub.csv
- ▶ 📄 test\_data.csv
- ▶ 📄 train\_data.csv

1 IF (test\_data\_csv.Bed\_Grade is missing )

2 THEN (median(test\_data\_csv.Bed\_Grade))

3 ELSE(test\_data\_csv.Bed\_Grade)

4

Validate

Validation Results

✔ The expression is valid.



# Replacing NULL into values

IBM Cognos Analytics with Watson | \* SPRINT 1 | 304 | Search content | ? | Bell | User

Grid Relationships Custom tables

artment	Ward_Type	Ward_Facility_Code	Bed Grade	patientid	City_Code_Patient
ecology	R	F	2	17006	2
ecology	Q	F	2	17006	2
ecology	Q	D	2	17006	2
ecology	S	E	3	95946	7
ecology	Q	F	3	95946	7
ecology	Q	F	4	95946	7
ecology	Q	E	2	95946	7
ecology	R	F	3	95946	7
ecology	R	A	3	95946	7
ecology	R	F	4	95946	7
ecology	S	C	2	10728	8

Data module

- Hospital...ion\_code
- Available...Hospital
- abc Department
- abc Ward\_Type
- abc Ward\_Fa...ty\_Code
- Bed Grade
- patientid
- City\_Code\_Patient
- abc Type of Admission
- abc Severity of Illness
- Visitors ...h Patient
- abc Age
- Admission\_Deposit
- train\_data.csv

Data module

+

?

Grid

Relationships

Custom tables

Search

Hospital...ion\_code

Available...Hospital

abc Department

abc Ward\_Type

abc Ward\_Fa...ty\_Code

Bed Grade

patientid

City\_Code\_Patient

abc Type of Admission

abc Severity of Illness

Visitors ...h Patient

abc Age

Admission\_Deposit

train\_data.csv

↕	Ward_Type	Ward_Facility_Code	Bed Grade	patientid	City_Code_Patient	Type of
	S	A	2	17006	2	Emerge
	S	F	2	17006	2	Trauma
	Q	D	4	17006	2	Emerge
	Q	F	2	17006	2	Trauma
	R	F	2	17006	2	Trauma
	Q	F	2	17006	2	Trauma
	Q	D	2	17006	2	Trauma
	S	E	3	95946	7	Emerge
	Q	F	3	95946	7	Trauma
	Q	F	4	95946	7	Urgent
	Q	F	2	95946	7	Trauma

## PREPARED DATA LINK

[https://us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&pathRef=.my\\_folders%2FSPRINT%2B1](https://us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&pathRef=.my_folders%2FSPRINT%2B1)

