

PROJECT DEVELOPMENT PHASE

SPRINT DELIVERY -1

TEAM ID : PNT2022TMID37134

PROJECT : Analytics for Hospitals Health-Care Data

DATE : 29 OCT 2022



PROJRCT CONTRIBUTORS

MANOJ KUMAR.S

DEEPAN.K

AKASH.M

MARAN.R

GUHAN.S

Project Development Phase

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

1.HOSP TAL DATASET L NK

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 interface. 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 'New' button is visible in the top right corner. The main content area shows a list of files with the following columns: Name, Type, and Last Accessed.

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

The Windows taskbar at the bottom shows the system clock as 09:16 PM on 01-Nov-22, along with various system icons and open applications like File Explorer and Document1.

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...`. The main workspace is divided into three sections:

- Components:** A sidebar on the left containing a search bar and a tree view. The tree view shows a hierarchy starting with "SPRINT 1", followed by "test_dat...Join (1)", "sample_sub.csv", and "test_data.csv". Under "test_data.csv", there are several fields listed: "# Row Id", "# case_id", "# Hosp...code", "abc Hosp...code", "City_...pital", "Hosp...code", and "Availab...ospital".
- Expression:** The central area 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
```
- Validation Results:** A panel at the bottom of the expression editor shows a green checkmark and the text "The expression is valid."

At the bottom of the interface, there is a checkbox labeled "Calculate after aggregation" which is checked. To the right of this checkbox are two buttons: "Cancel" and "OK". The bottom of the screen shows a Windows taskbar with the search bar, several application icons, and the system clock displaying "11:55 AM 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

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

Data module

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

Grid

Relationships

Custom tables

↕	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

