

PROJECT DEVELOPMENT PHASE

SPRINT DELIVERY -1

TEAM ID :PNT2022TMID29950

BATCH NO : B10-4A6E

PROJECT : Analytics for Hospitals Health-Care Data

DATE :29 OCT 2022

ANALYTICS FOR HOSPITAL HEALTH-CARE DATA

PROJECT CONTRIBUTORS

CHIDANANDA M
VIGNESH A
DINESH S
SRINATH S

TEAM ID:PNT2022TMID29950

Project Development Phase

Sprint-1:

- ✓ **Data Collection**
- ✓ **Loading of data**
- ✓ **Data Preparation**

Sprint-3:

- ✓ **Dashboard Creation**

Sprint-2:

- ✓ **Data Exploration**
- ✓ **Data Visualization**

Sprint-4:

- ✓ **Report Creation**
- ✓ **Story Creation**

1. HOSPITAL DATASET LINK:

https://drive.google.com/drive/folders/1qozDgfVw04hGS8Di_SFhUNdUKUdjkhRT?usp=sharing

2. UNDERSTANDING THE DATA

This project is based on understanding the Health Analytics dataset .

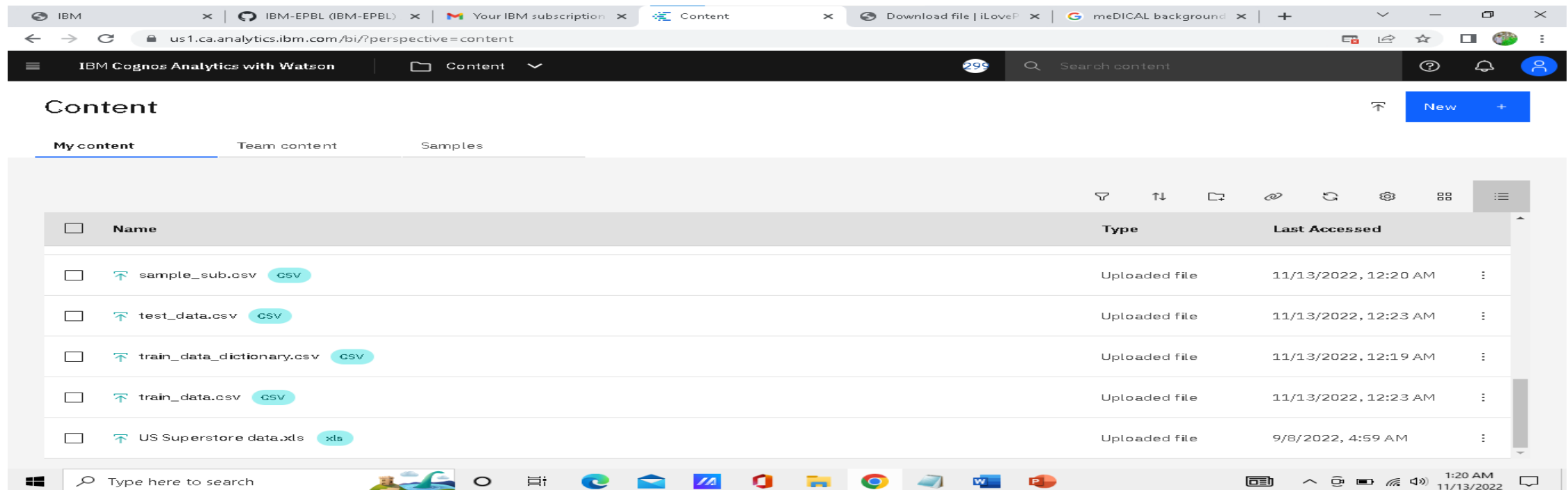
The data is spread across 3 data files (csv) and one data dictionary enclosed.

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' page. The browser address bar shows the URL `us1.ca.analytics.ibm.com/bi/?perspective=content`. The page header includes the 'IBM Cognos Analytics with Watson' logo, a 'Content' dropdown menu, a search bar, and a 'New' button. Below the header, there are tabs for 'My content', 'Team content', and 'Samples'. The main area shows a list of files with columns for 'Name', 'Type', and 'Last Accessed'. The files listed are:

Name	Type	Last Accessed
sample_sub.csv	Uploaded file	11/13/2022, 12:20 AM
test_data.csv	Uploaded file	11/13/2022, 12:23 AM
train_data_dictionary.csv	Uploaded file	11/13/2022, 12:19 AM
train_data.csv	Uploaded file	11/13/2022, 12:23 AM
US Superstore data.xls	Uploaded file	9/8/2022, 4:59 AM

The taskbar at the bottom shows the Windows Start button, a search bar, and several application icons including Edge, File Explorer, and Microsoft Office applications. The system clock indicates 1:20 AM on 11/13/2022.

DATASET PREPRATION

IBM EPBL (IBM-EPBL) | Your IBM subscription | * SPRINT 1 | Download file | iLoveP | meDICAL background

us1.ca.analytics.ibm.com/bj/?perspective=ca-modeller&id=iCA7BEE54A52A4979AAACE8050F0EDF30F&objRef=iCA7BEE54A52A4979AAACE8050F0EDF30F&tid=9572083...

IBM Cognos Analytics with Watson | * SPRINT 1 | Search content

Properties

Data module

Search

SPRINT 1

- Navigation paths
- Calculation name
- Calculation name
- train_data.csv
- sample_sub.csv
- test_data.csv
- Calculation name
- Calculation name
- Calculation name
- # Row Id
- # case_id
- # Hospital_code
- abc Hospital...ype_code
- City_Code_Hospital

Grid Relationships Custom tables

n name	Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Ho
	1	318439	21	c	3	
	2	318440	29	a	4	
	3	318441	26	b	2	
	4	318442	6	a	6	
	5	318443	28	b	11	
	6	318444	23	a	6	
	7	318445	26	b	2	
	8	318446	25	e	1	
	9	318447	23	a	6	
	10	318448	23	a	6	
	11	318449	10	e	1	

Type here to search

1:46 AM 11/13/2022

DATASET PREPRATION

IBM EPBL (IBM-EPBL) Your IBM subscription SPRINT1 Download file | iLoveP meDICAL background

us1.ca.analytics.ibm.com/bi/?perspective=ca-modeller&id=iCA7BEE54A52A4979AAACE8050F0EDF30F&objRef=iCA7BEE54A52A4979AAACE8050F0EDF30F&tid=9572083...

Name Calculation name

Components

test_data.csv

- Calculation name
- Calculation name
- # Row Id
- # case_id
- # Hospital_code
- abc Hospit..._code
- City_Co...spital
- Hospit..._code
- Availabl...Hospital
- abc Department

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
```

Validation Results

✓ The expression is valid.

☒ Calculate after aggregation

Cancel OK

Type here to search

1:36 AM 11/13/2022

Create calculation

Name Calculation name

Components

Expression

14

Search

SPRINT 1

Calculation name

train_data.csv

sample_sub.csv

test_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)
```

Validation Results

✓ The expression is valid.

☒ Calculate after aggregation

Cancel

OK

Replacing NULL into values

IBM EPBL (IBM-EPBL) Your IBM subscription SPRINT 1 Download file | iLoveP meDICAL background

us1.ca.analytics.ibm.com/bi/?perspective=ca-modeller&id=iCA7BEE54A52A4979AAACE8050F0EDF30F&objRef=iCA7BEE54A52A4979AAACE8050F0EDF30F&tid=9572083...

IBM Cognos Analytics with Watson SPRINT 1 299 Search content

Data module

- Available... Hospital
 - abc Department
 - abc Ward_Type
 - abc Ward_Facility_Code
- Bed Grade
- patientid
- City_Code_Patient**
- abc Type of Admission
- abc Severity of Illness
- Visitors ...h Patient
- abc Age
- Admission_Deposit
- abc Stay
- sample_sub.csv
- test_data.csv

Grid Relationships Custom tables

int	Ward_Type	Ward_Facility_Code	Bed Grade	patientid	City_Code_Patient	Typ
py	R	F	2	31397	7	
py	S	F	2	31397	7	
i	S	E	2	31397	7	
py	R	D	2	31397	7	
py	S	D	2	31397	7	
i	S	F	2	31397	7	
py	S	B	3	31397	7	
py	Q	F	3	31397	7	
y	R	B	4	31397	7	
y	S	E	3	31397	7	
py	S	B	2	31397	7	

Type here to search 1:43 AM 11/13/2022

Data module

Search

- Available... Hospital
- abc Department
- abc Ward_Type
- abc Ward_Facility_Code
- Bed Grade**
- patientid
- City_Code_Patient
- abc Type of Admission
- abc Severity of Illness
- Visitors ...h Patient
- abc Age
- Admission_Deposit
- abc Stay
- sample_sub.csv
- test_data.csv

Grid Relationships Custom tables

int	Ward_Type	Ward_Facility_Code	Bed Grade	patientid	City_Code_Patient	Type
py	R	F	2	31397	7	
py	S	F	2	31397	7	
i	S	E	2	31397	7	
py	R	D	2	31397	7	
py	S	D	2	31397	7	
i	S	F	2	31397	7	
py	S	B	3	31397	7	
py	Q	F	3	31397	7	
y	R	B	4	31397	7	
y	S	E	3	31397	7	
py	S	B	2	31397	7	

PREPARED DATA LINK

https://us1.ca.analytics.ibm.com/bi/?perspective=ca-modeller&pathRef=.public_folders%2FPROJECT%2BDEVELOPMENT%2FSPRINT%2B1

