

Sprint-I

Date	14 November 2022
Team ID	PNT2022TMID37428
Project Name	Analytics for Hospitals' Healthcare Data

DATA PREPARATION AND CLEANING:

Uploading the dataset

The screenshot shows the IBM Cognos Analytics with Watson interface. At the top, there's a header with the IBM logo and 'IBM Cognos Analytics with Watson'. Below the header, there's a 'Watson.' banner with a 'Reading 2 files...' progress bar and 'Cancel Details' links. The banner also includes a 'Watch video' button and a 'Take a product tour' button. Below the banner, there's a 'Quick launch' section with four tiles: 'Upload data' (with an upload icon), 'Prepare data' (with a data icon), 'Exploration' (with a magnifying glass icon), and 'Present data' (with a presentation icon). Each tile has a brief description of its function.

Select sources

The screenshot shows the 'Select sources' dialog box. It has a search bar at the top with the text 'My content'. Below the search bar, there's a list of items. The first item is '50_Startups.csv' with a timestamp of '9/17/2022, 11:38 AM'. The second item is 'Age wise patient visualisation' with a timestamp of '11/17/2022, 6:10 AM'. The third item is 'bubbles and radial dashboard' with a timestamp of '11/17/2022, 6:19 AM'. The fourth item is 'dashboard list of patients' with a timestamp of '11/17/2022, 6:14 AM'. The fifth item is 'dashboard showing waterfall, pie, bar' with a timestamp of '11/17/2022, 6:51 AM'. The sixth item is 'Data Preparation' with a timestamp of '11/17/2022, 6:51 AM'. The seventh item is 'Number of patients by ward type' with a timestamp of '11/17/2022, 6:03 AM'. The eighth item is 'test_data.csv' with a timestamp of '11/17/2022, 5:48 AM'. The ninth item is 'train_data.csv' with a timestamp of '11/17/2022, 5:52 AM'. On the right side of the dialog, there's a 'Filter by' section with a 'Reset filters' link. Under 'Type', there are checkboxes for 'Folders', 'Packages', 'Files', 'Data sets', and 'Modules'. Under 'Modified', there are radio buttons for 'All', 'Today', 'Yesterday', 'Past week', and 'Past month'. The 'All' radio button is selected.

Cancel

OK

IBM Cognos Analytics with Watson						
New data module						
Search content						
Properties						
Data module						
Grid						
Relationships						
Custom tables						
Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hosp	
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	

Cleaning the empty cells

Create calculation

Name City_Code_Patient

Components

Expression

Search	
train_data.csv	
BedGrade	
City_Cod...Patient	
# Row Id	
# case_id	
# Hospi...code	
abc Hospi...code	
City_C...spital	
1	
2	
3	

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

Validation Results

The expression is valid.

☐ Calculate after aggregation

Cancel

OK

IBM Cognos Analytics with Watson						
New data module						
Search content						
Properties						
Data module						
Grid						
Relationships						
Custom tables						
City_Code_Patient	Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code
7	5	5	26	b	2	Y
7	6	6	23	a	6	X
7	7	7	32	f	9	Y
7	8	8	23	a	6	X
7	9	9	1	d	10	Y
7	10	10	10	e	1	X
7	11	11	22	g	9	Y
7	12	12	26	b	2	Y
7	13	13	16	c	3	Z
7	14	14	9	d	5	Z
8	15	15	6	a	6	X
8	16	16	6	a	6	X
8	17	17	23	a	6	X

Create calculation

Name

BedGrade

Components

Search

train_data.csv

BedGrade

City_Cod...Patient

Row Id

case_id

Hospital_code

Hospital...code

City_C...spital

Hospital...code

Availabl...ospital

Department

Expression

1 IF(BedGrade is missing)

2 THEN (median (BedGrade))

3 ELSE (BedGrade)

Validation Results

The expression is valid.

Calculate after aggregation

Cancel

OK

IBM Cognos Analytics with Watson

New data module

Search content

Properties

Data module

Search

New data module

Navigation paths

train_data.csv

BedGrade

City_Code_Patient

Row Id

case_id

Hospital_code

Hospital...pe_code

City_Cod...Hospital

Hospital...on_code

Availabl...Hospital

Department

Ward_Type

Ward_Fa...ty_Code

Bed Grade

patientid

Grid

Relationships

Custom tables

#	BedGrade	City_Code_Patient	Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital
1	2	7	1	1	8	c	3
2	2	7	2	2	2	c	5
3	2	7	3	3	10	e	1
4	2	7	4	4	26	b	2
5	2	7	5	5	26	b	2
6	2	7	6	6	23	a	6
7	3	7	7	7	32	f	9
8	3	7	8	8	23	a	6
9	4	7	9	9	1	d	10
10	3	7	10	10	10	e	1
11	2	7	11	11	22	g	9
12	1	7	12	12	26	b	2
13	3	7	13	13	16	c	3

IBM Cognos Analytics with Watson

New data module

Search content

Properties

Data module

Search

New data module

Navigation paths

train_data.csv

BedGrade

City_Code_Patient

Row Id

case_id

Hospital_code

Hospital...pe_code

City_Cod...Hospital

Hospital...on_code

Availabl...Hospital

Department

Ward_Type

Ward_Fa...ty_Code

Bed Grade

patientid

Grid

Relationships

Custom tables

#	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available E...in Hospital	Department
1	1	8	c	3	Z	3	radiotherapy
2	2	2	c	5	Z	2	radiotherapy
3	3	10	e	1	X	2	anesthesia
4	4	26	b	2	Y	2	radiotherapy
5	5	26	b	2	Y	2	radiotherapy
6	6	23	a	6	X	2	anesthesia
7	7	32	f	9	Y	1	radiotherapy
8	8	23	a	6	X	4	radiotherapy
9	9	1	d	10	Y	2	gynecology
10	10	10	e	1	X	2	gynecology
11	11	22	g	9	Y	2	radiotherapy
12	12	26	b	2	Y	4	radiotherapy
13	13	16	c	3	Z	2	radiotherapy