

Project Development Phase

Sprint -1

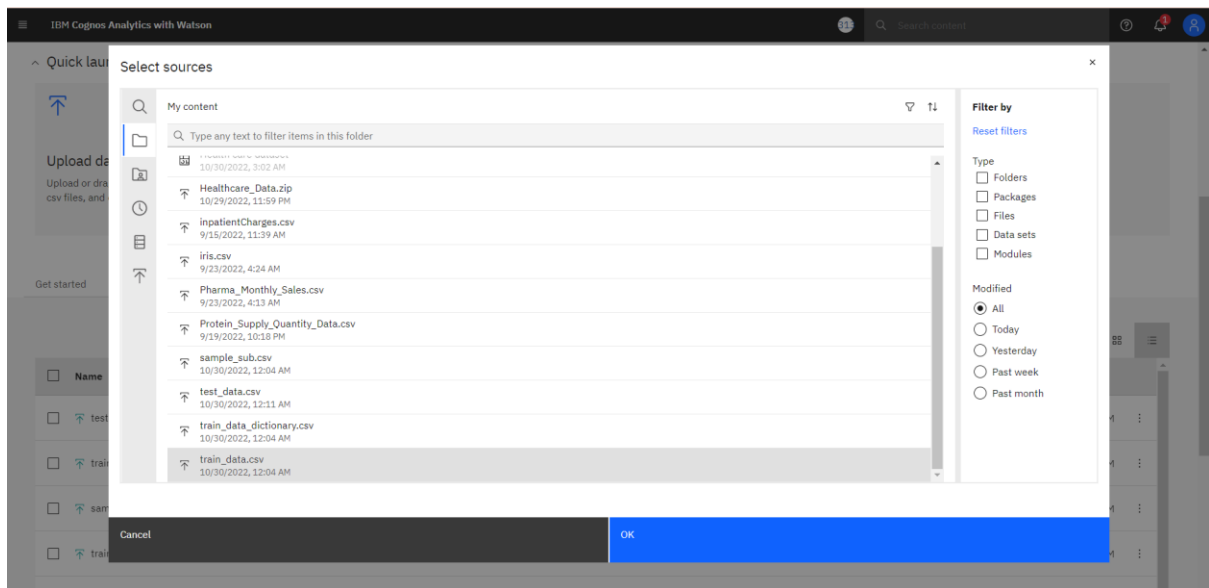
Team ID: PNT2022TMID04371

Step 1:

Collecting the dataset and uploading the dataset into IBM COGNOS ANALYTICS

The screenshot displays the IBM Cognos Analytics with Watson user interface. At the top, there is a navigation bar with the title 'IBM Cognos Analytics with Watson', a search bar labeled 'Search content', and user profile icons. Below the navigation bar, a 'Quick launch' section offers four main actions: 'Upload data' (with an upward arrow icon), 'Prepare data' (with a data flow icon), 'Exploration' (with a magnifying glass icon), and 'Present data' (with a presentation screen icon). Each action includes a brief description of its function. The main content area is divided into two tabs: 'Get started' and 'Recent'. The 'Recent' tab is active, showing a table of recently uploaded files. The table has columns for 'Name', 'Location', 'Type', and 'Last Accessed'. Four CSV files are listed: 'test_data.csv', 'train_data.csv', 'sample_sub.csv', and 'train_data_dictionary.csv', all located in 'My content' and uploaded on 10/30/2022.

Name	Location	Type	Last Accessed
test_data.csv	My content	Uploaded file	10/30/2022, 12:11 AM
train_data.csv	My content	Uploaded file	10/30/2022, 12:04 AM
sample_sub.csv	My content	Uploaded file	10/30/2022, 12:04 AM
train_data_dictionary.csv	My content	Uploaded file	10/30/2022, 12:04 AM



Row Id	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available E...in Hospital	Department
1	1			3	Z	3	radiotherapy
2	2			5	Z	2	radiotherapy
3	3			1	X	2	anesthesia
4	4			2	Y	2	radiotherapy
5	5			2	Y	2	radiotherapy
6	6			6	X	2	anesthesia
7	7			9	Y	1	radiotherapy
8	8			6	X	4	radiotherapy
9	9			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
14	14	9	d	5	Z	3	radiotherapy
15	15	6	a	6	X	4	gynecology
16	16	6	a	6	X	3	gynecology

Step 2:

Data Cleaning and Data Processing using Cognos Analytics:

IBM Cognos Analytics with Watson

New data module

Search content

Properties

Sources

Search

train_data.csv

Data module

Search

Row Id

case_id

Hospital_code

Hospital...pe...s

City_Cod...Hosp...

Hospital...ion...

Available...Hosp...

Department

Ward_Type

Ward_Fa...ty_G...

Bed Grade

patientid

City_Code_Pati...

Type of Admissi...

Severity of Illne...

Visitors ...h Pati...

Age

Admissio...Deposit

Stay

train_data.csv

Clean - Type of Admission

Whitespace

☒ Trim leading and trailing whitespace

Convert case to

☐ UPPERCASE ☐ lowercase ☒ Do not change

Return a substring of characters

Start

Length

Preview

This is a preview

NULL values

☐ Replace this value with NULL

Empty string

☐ Replace NULL values with

Empty string

Cancel

Clean

Age	Admission_Deposit	Stay
71-80	4718	71-80
71-80	3650	0-10
31-40	5038	0-10
31-40	3357	41-50
31-40	4353	51-60
31-40	5152	51-60
31-40	4046	More than 100 Days
51-60	7606	0-10
51-60	6204	51-60
51-60	4638	11-20
51-60	5327	51-60
31-40	3945	11-20
31-40	4253	0-10
31-40	4221	21-30
31-40	4910	21-30
31-40	4628	11-20