

# Project Development Phase

## Delivery of Sprint-1

Date	03 NOV 2022
Team ID	PNT2022TMID20895
Project Name	Analytics for Hospitals' Health-Care Data

## Project Development Phase:

### Sprint-1:

- Data Collection
- Data Preparation

### Sprint-2:

- Data Exploration

### Sprint-3:

- Dashboard Creation

### Sprint-4:

- Report Creation
- Story Creation

## **Contributors:**

A RAJESH KUMAR

V RAHUL

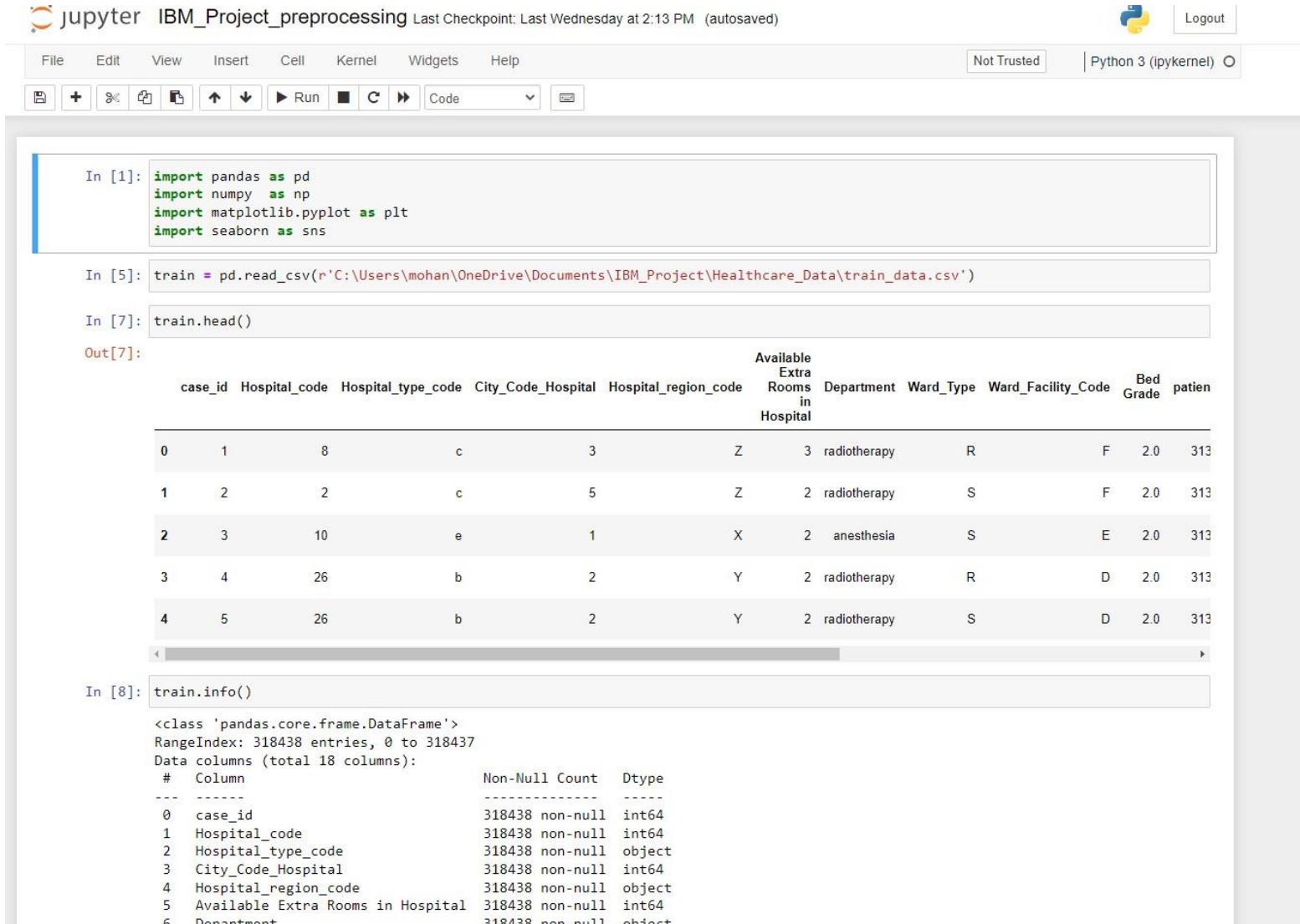
K THILAKRAJ

T AAKASH

# SPRINT-1

## Data Preprocessing:

### Using Jupyter notebook to remove the null values:



The screenshot shows a Jupyter Notebook interface with the following content:

- In [1]:**

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```
- In [5]:**

```
train = pd.read_csv(r'C:\Users\mohan\OneDrive\Documents\IBM_Project\Healthcare_Data\train_data.csv')
```
- In [7]:**

```
train.head()
```
- Out[7]:** A data preview of the 'train' DataFrame. The columns are:

case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available Extra Rooms in Hospital	Department	Ward_Type	Ward_Facility_Code	Bed Grade	patient	
0	1	8	c	3	Z	3	radiotherapy	R	F	2.0	313
1	2	2	c	5	Z	2	radiotherapy	S	F	2.0	313
2	3	10	e	1	X	2	anesthesia	S	E	2.0	313
3	4	26	b	2	Y	2	radiotherapy	R	D	2.0	313
4	5	26	b	2	Y	2	radiotherapy	S	D	2.0	313
- In [8]:**

```
train.info()
```

The output of `train.info()` is as follows:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 318438 entries, 0 to 318437
Data columns (total 18 columns):
 #   Column           Non-Null Count  Dtype  
 --- 
 0   case_id          318438 non-null   int64  
 1   Hospital_code    318438 non-null   int64  
 2   Hospital_type_code 318438 non-null   object 
 3   City_Code_Hospital 318438 non-null   int64  
 4   Hospital_region_code 318438 non-null   object 
 5   Available Extra Rooms in Hospital 318438 non-null   int64  
 6   Department        318438 non-null   object 
```

```
In [17]: train.isnull().sum()
```

```
Out[17]: case_id          0
Hospital_code        0
Hospital_type_code   0
City_Code_Hospital   0
Hospital_region_code 0
Available Extra Rooms in Hospital 0
Department           0
Ward_Type            0
Ward_Facility_Code   0
Bed Grade            0
patientid            0
City_Code_Patient    0
Type of Admission    0
Severity of Illness   0
Visitors with Patient 0
Age                  0
Admission_Deposit    0
Stay                 0
dtype: int64
```

```
In [18]: train = train.to_csv('final_train_data.csv', index=False)
```

```
In [19]: test = pd.read_csv(r'C:\Users\mohan\OneDrive\Documents\IBM_Project\Healthcare_Data\test_data.csv')
```

```
In [20]: test.head()
```

```
Out[20]:
```

	case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available Extra Rooms in Hospital	Department	Ward_Type	Ward_Facility_Code	Bed Grade	patientid	
0	318439	21	c	3	Z	3	gynecology	S		A	2.0	170
1	318440	29	a	4	X	2	gynecology	S		F	2.0	170

```
Stay          0  
dtype: int64
```

```
In [11]: train['Bed Grade'].value_counts()
```

```
Out[11]: 2.0    123671  
3.0    110583  
4.0    57566  
1.0    26505  
Name: Bed Grade, dtype: int64
```

```
In [12]: train['Bed Grade'].unique()
```

```
Out[12]: array([ 2.,  3.,  4.,  1., nan])
```

```
In [13]: train.shape
```

```
Out[13]: (318438, 18)
```

```
In [14]: train.dropna(inplace=True)
```

```
In [15]: train.shape
```

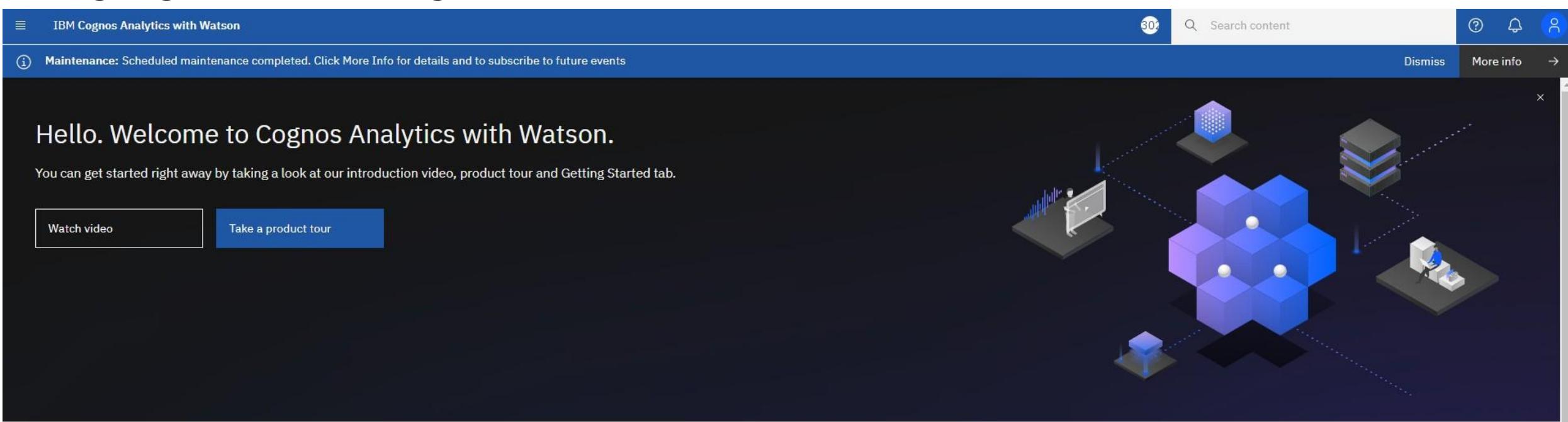
```
Out[15]: (313793, 18)
```

```
In [16]: train.head()
```

```
Out[16]:
```

case_id	Hospital_code	Hospital_type_code	City_Code_Hospital	Hospital_region_code	Available Extra Rooms in Hospital	Department	Ward_Type	Ward_Facility_Code	Bed Grade	patient	
0	1	8	c	3	Z	3	radiotherapy	R	F	2.0	313
1	2	2	c	5	Z	2	radiotherapy	S	F	2.0	313
2	3	10	e	1	X	2	anesthesia	S	E	2.0	313
3	4	26	b	2	Y	2	radiotherapy	R	D	2.0	313

# Signing in to the IBM Cognos tool:



The screenshot shows the IBM Cognos Analytics with Watson homepage. At the top, there's a blue header bar with the title "IBM Cognos Analytics with Watson". Below it, a banner displays a message about maintenance completed. The main content area features a large, dark background image of a 3D cube structure with a person working at a desk, surrounded by data visualization icons like a video player and a stack of servers. On the left, there's a "Hello. Welcome to Cognos Analytics with Watson." message and two buttons: "Watch video" (white background) and "Take a product tour" (blue background). The bottom section contains four cards with icons and descriptions: "Upload data" (up arrow icon), "Prepare data" (data modules icon), "Exploration" (cube icon), and "Present data" (chart icon).

^ Quick launch



## Upload data

Upload or drag and drop spreadsheets, csv files, and other data sources.



## Prepare data

Use data modules to clean and connect data from multiple resources.



## Exploration

Quickly find unbiased answers by identifying trends in your data with data exploration.



## Present data

Create sophisticated, multi-page, multi-query dashboards, reports, or stories.

Get started

Recent

## Selecting the dataset to upload into the Cognos:

The screenshot shows a 'Select sources' dialog box from a software application. On the left, there's a sidebar with icons for search, folder, history, and other navigation. The main area displays a list of items under 'My content'. A search bar at the top allows filtering by text. The list includes:

- Assignment-1\_Exploration (9/15/2022, 3:58 PM)
- Data\_module1 (9/19/2022, 8:51 AM)
- IBM\_Finalproject (11/3/2022, 10:25 PM)
- 50\_Startups.csv (9/9/2022, 12:13 PM)
- Assignment2\_Dashboard (9/24/2022, 2:14 PM)
- final\_train\_data.csv (11/2/2022, 4:46 AM) - This item is highlighted with a blue selection border.
- IPL Ball-by-Ball 2008-2020.csv (9/5/2022, 9:58 AM)
- IPL data module (9/5/2022, 10:05 AM)
- IPL Matches 2008-2020.csv (9/5/2022, 9:58 AM)
- IPL\_story

On the right side, there are 'Filter by' sections for Type (Folders, Packages, Files, Data sets, Modules) and Modified (All, Today, Yesterday, Past week, Past month). At the bottom, there are 'Cancel' and 'OK' buttons.

# Uploading the dataset to perform the preparation and visualization:

IBM Cognos Analytics with Watson

302 Search content

Maintenance: Scheduled maintenance completed. Click More Info for details and to sub...

Reading 2 files...

Cancel Details

Dismiss More info

Hello. Welcome to Cognos Analytics with Watson.

You can get started right away by taking a look at our introduction video, product tour and Getting Started tab.

Watch video Take a product tour



x

Quick launch



Upload data

Upload or drag and drop spreadsheets, csv files, and other data sources.



Prepare data

Use data modules to clean and connect data from multiple resources.



Exploration

Quickly find unbiased answers by identifying trends in your data with data exploration.



Present data

Create sophisticated, multi-page, multi-query dashboards, reports, or stories.

Get started

Recent