# PROJECT DEVELOPMENT PHASE SPRINT DELIVERY -1

**TEAM ID** : PNT2022TMID28983

BATCH NO: B5-5M1E

**PROJECT** : Analytics for Hospitals Health-Care Data

**DATE** : 10 NOV 2022

# Project Development Phase

## Sprint-1:

- Data Collection
- Loading of data
- Data Preparation

## **Sprint-2:**

- Data Exploration
- Data Visualization

## **Sprint-3**:

Dashboard Creation

# **Sprint-4:**

- Report Creation
- Story Creation

#### 1. DATA COLLECTION

## **HOSPITAL DATA SET**

https://drive.google.com/drive/folders/1qozDgfVw04hGS8Di\_SFhUNdUKUdjkhrT?usp=sharing

## **UNDERSTANDING THE DATA**

This project is based on understanding the Hospital dataset.

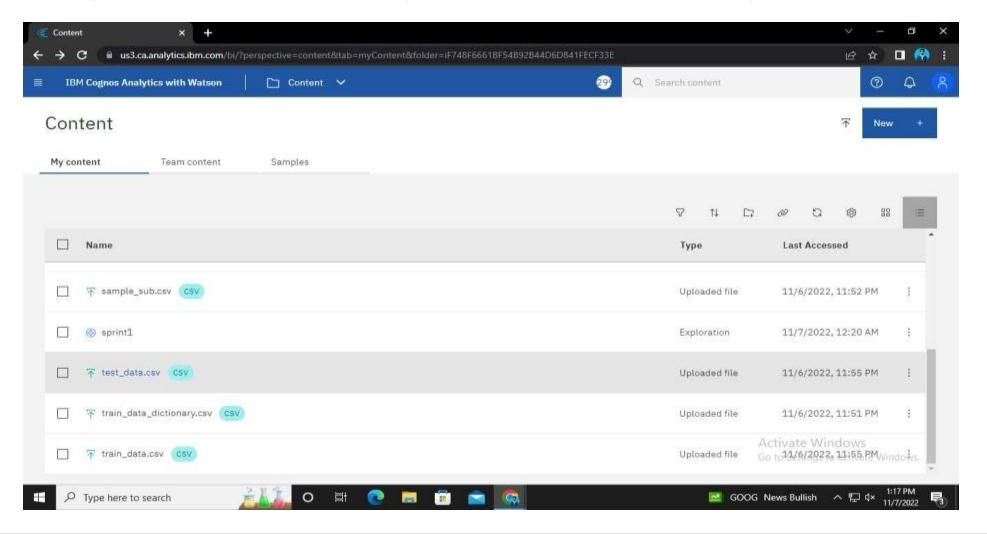
The data is spread across 3 data files (csv) and one data dictionary enclosed.

- 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.

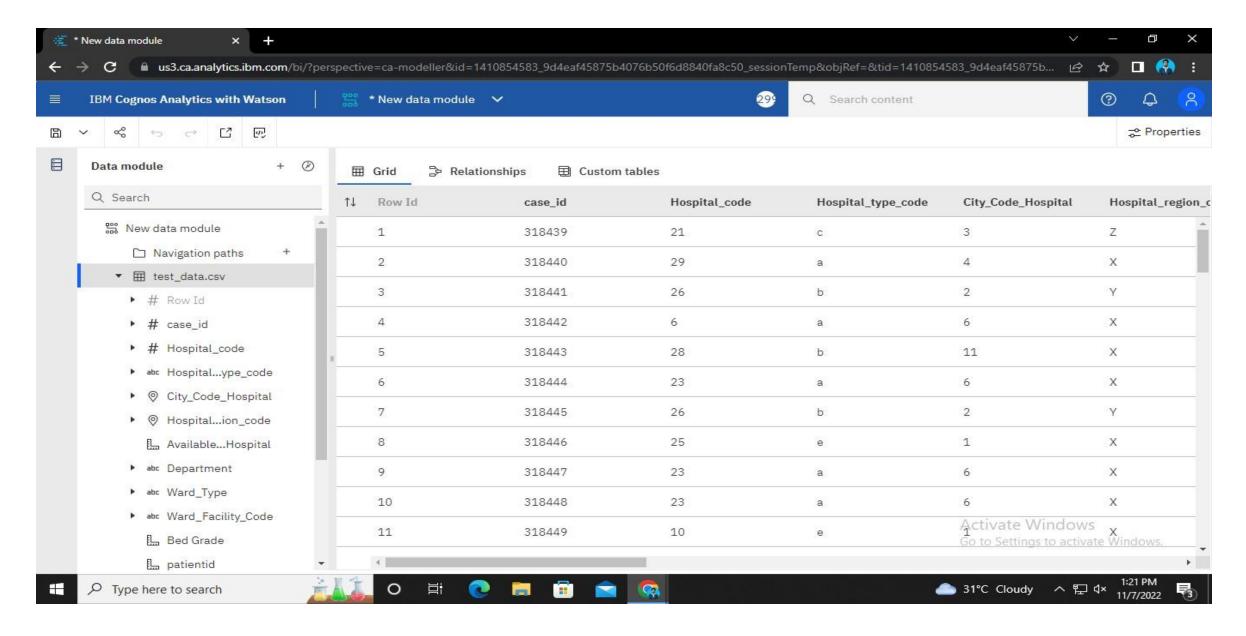
#### 2. LOADING OF DATA

- 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.



## 3.

## **DATA PREPARATION**



# Replacing NULL into values

