PROJECT DEVELOPMENT PHASE SPRINT DELIVERY -1

TEAM ID : PNT2022TMID09318

BATCH NO : **B8-2A4E**

PROJECT : Analytics for Hospitals Health-Care Data

DATE : 7 NOV 2022

Sprint-1:

- o Data Collection
- o Loading of data
- o Data Preparation

Sprint-2:

- o Data Exploration
- o Data Visualization

Sprint-3:

o Dashboard Creation

Sprint-4:

- o Report Creation
- o 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.

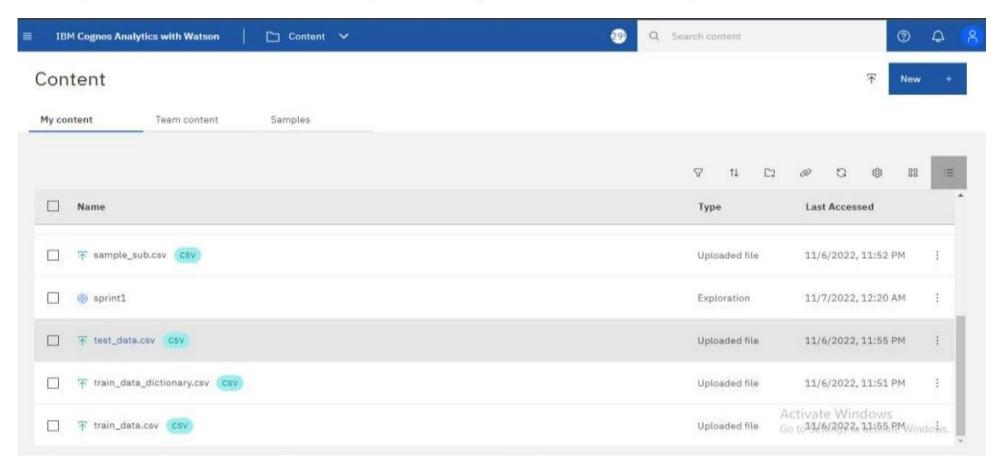
The data is spread across 3 data files (csv) and one data dictionary enclosed.

- 1. Test_data.csv
- 2.Sample_sub.csv
- 3.Train_data_dictionary.csv
- 4.Train_data.csv

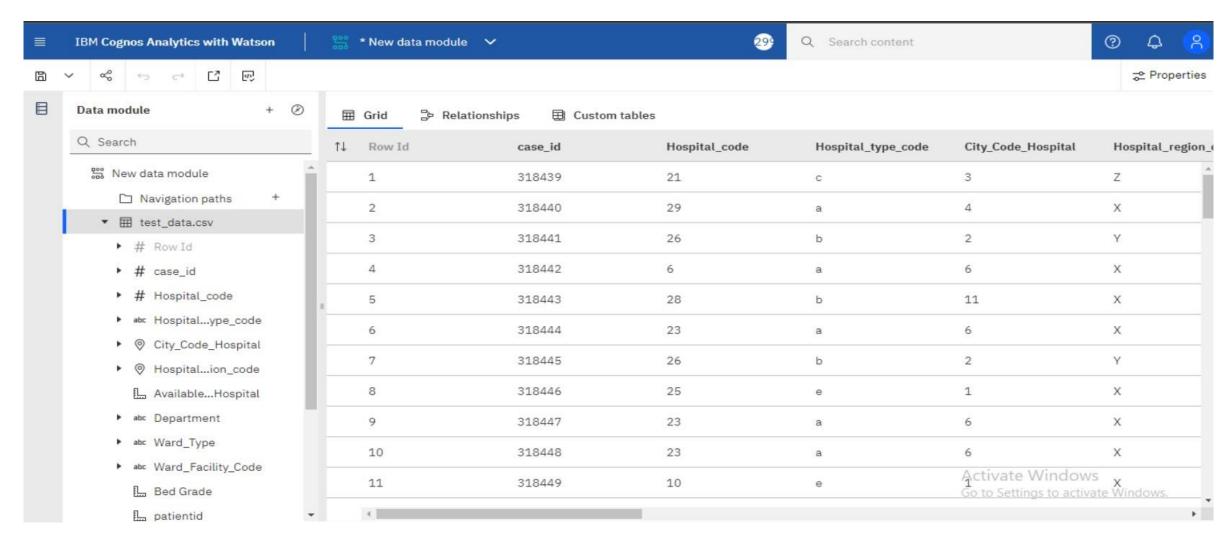
The primary data file we use is **test_data.csv** consist of 17 Columns with 137057 Rows.

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.



DATA PREPARATION



• Replacing NULL into values

