PROJECT DEVELOPMENT PHASE
SPRINT DELIVERY -1

TEAM ID : PNT2022TMID32341

BATCH NO : B5-5M1E

PROJECT : Analytics for Hospitals Health-Care Data

DATE : 29 OCT 2022

Project Development Phase

Sprint-1: **✓Data Collection** **✓ Loading of data** **✓ Data Preparation**	**Sprint-3:** **✓ Dashboard Creation**
Sprint-2: ✓ Data Exploration ✓Data Visualization	Sprint-4: ✓ Report Creation ✓ Story Creation

1.HOSP TAL DATASET L NK

https://drive.google.com/drive/folders/1qozDgfVw04hGS8Di_SFhUNdUKUdjkhrT?usp=sharing

2.**UNDERSTANDING THE DATA**

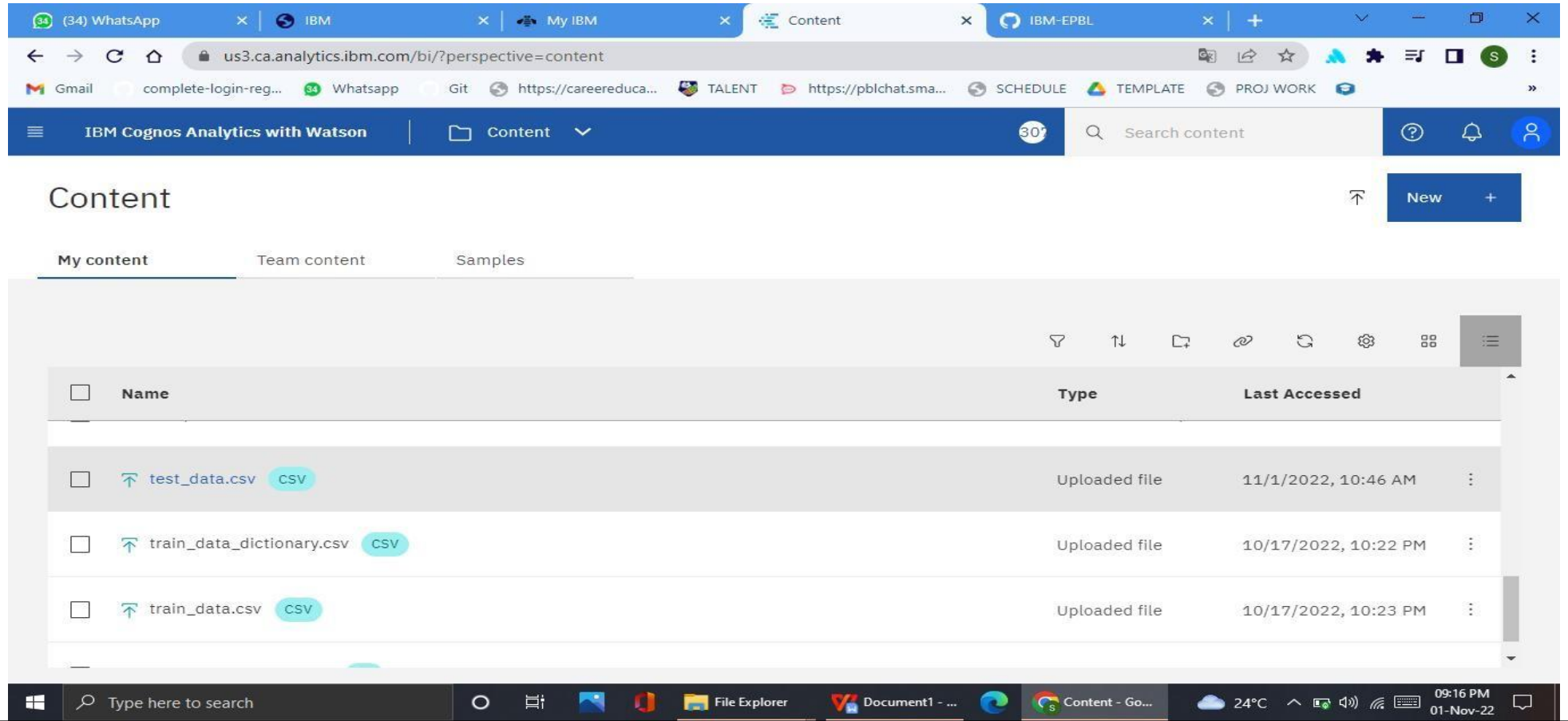
This project is based on understanding the Health Analytics dataset .
The data is spread across 3 data files (csv) and one data dictionary enclosed.

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

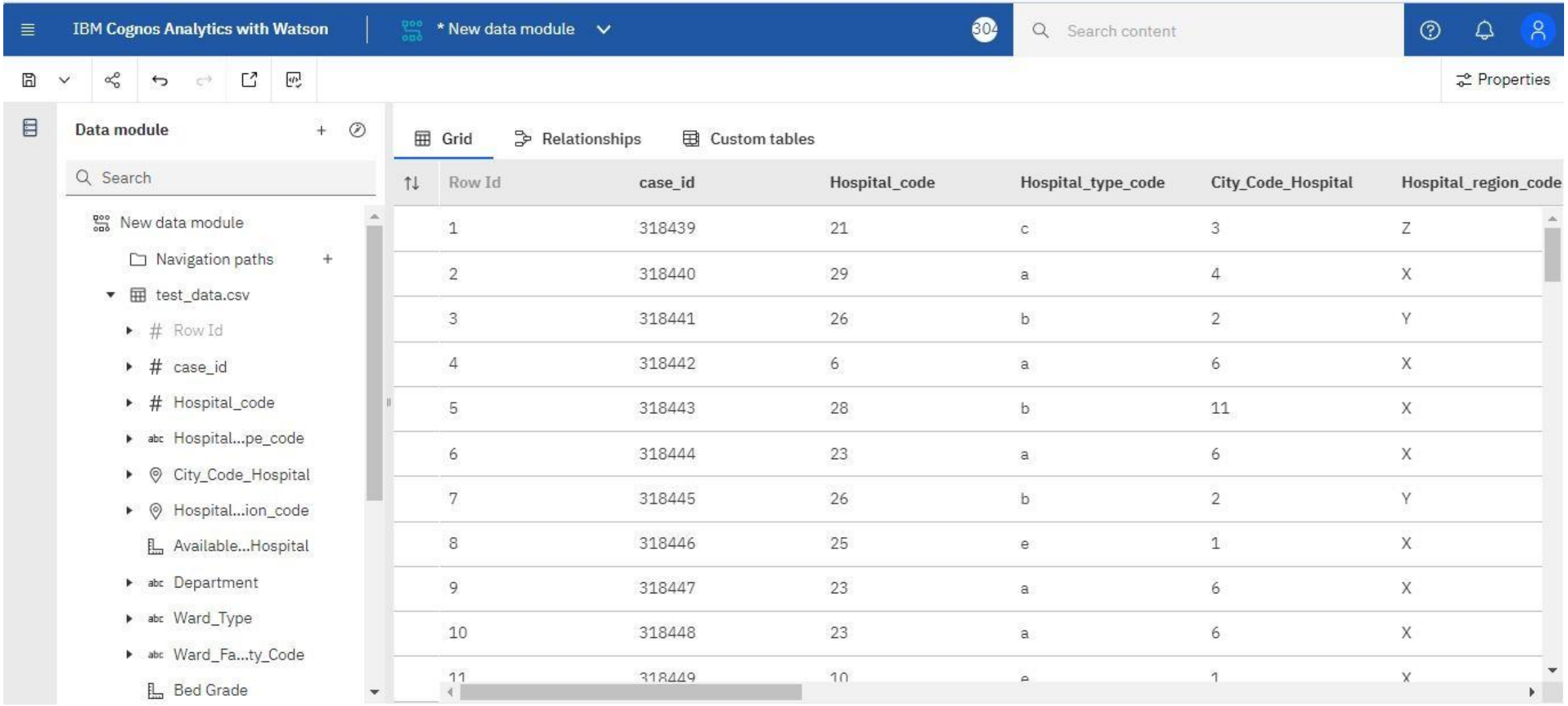
The primary data file we use is test_data.csv consist of 17 Columns with 137057 Rows.

3. LOADING OF DATA

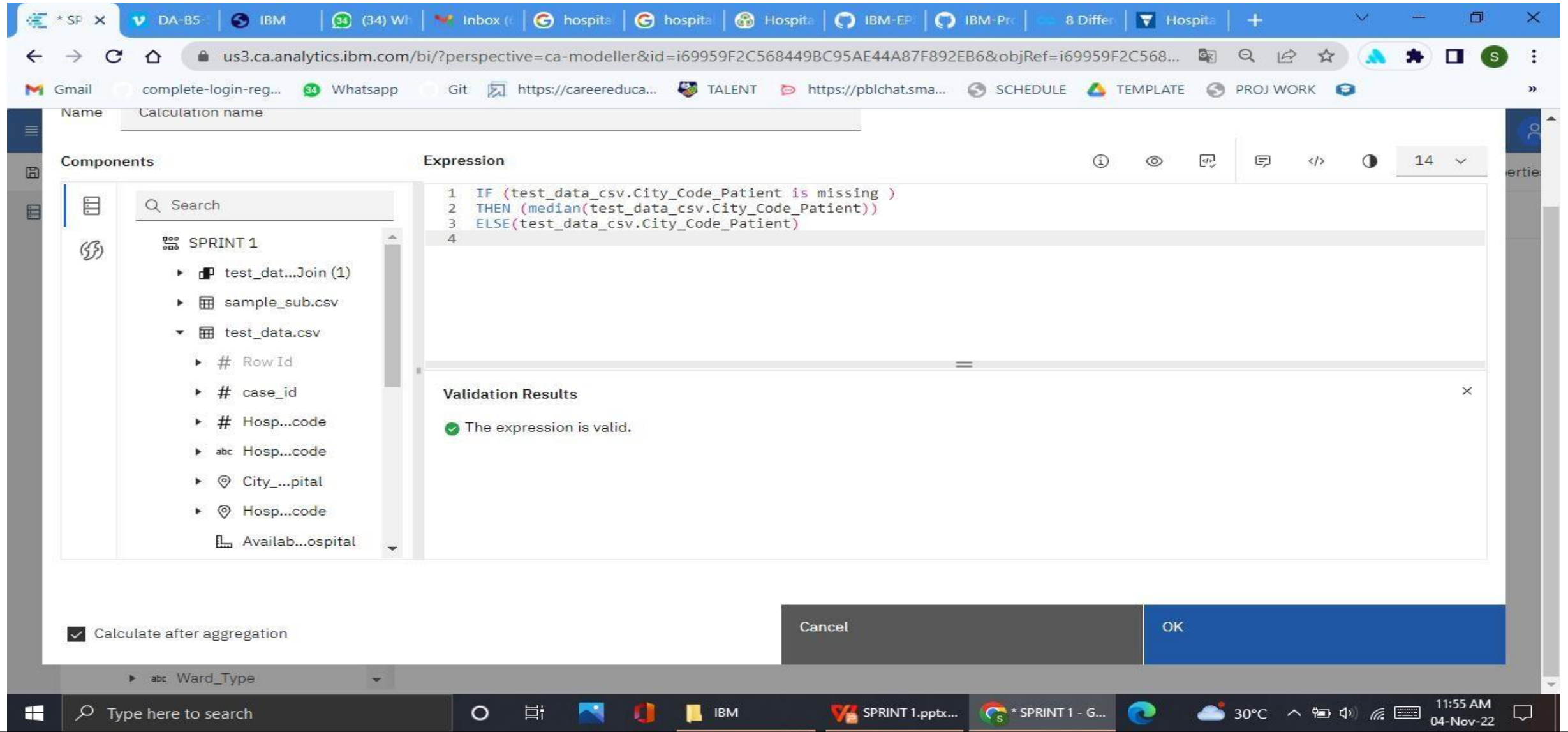
- Connect the data to IBM Cognos.
- Cognos supports connecting to a wide variety of data, stored in a variety of places.
- Loading of the data set in IBM Cognos is important for analyzing.

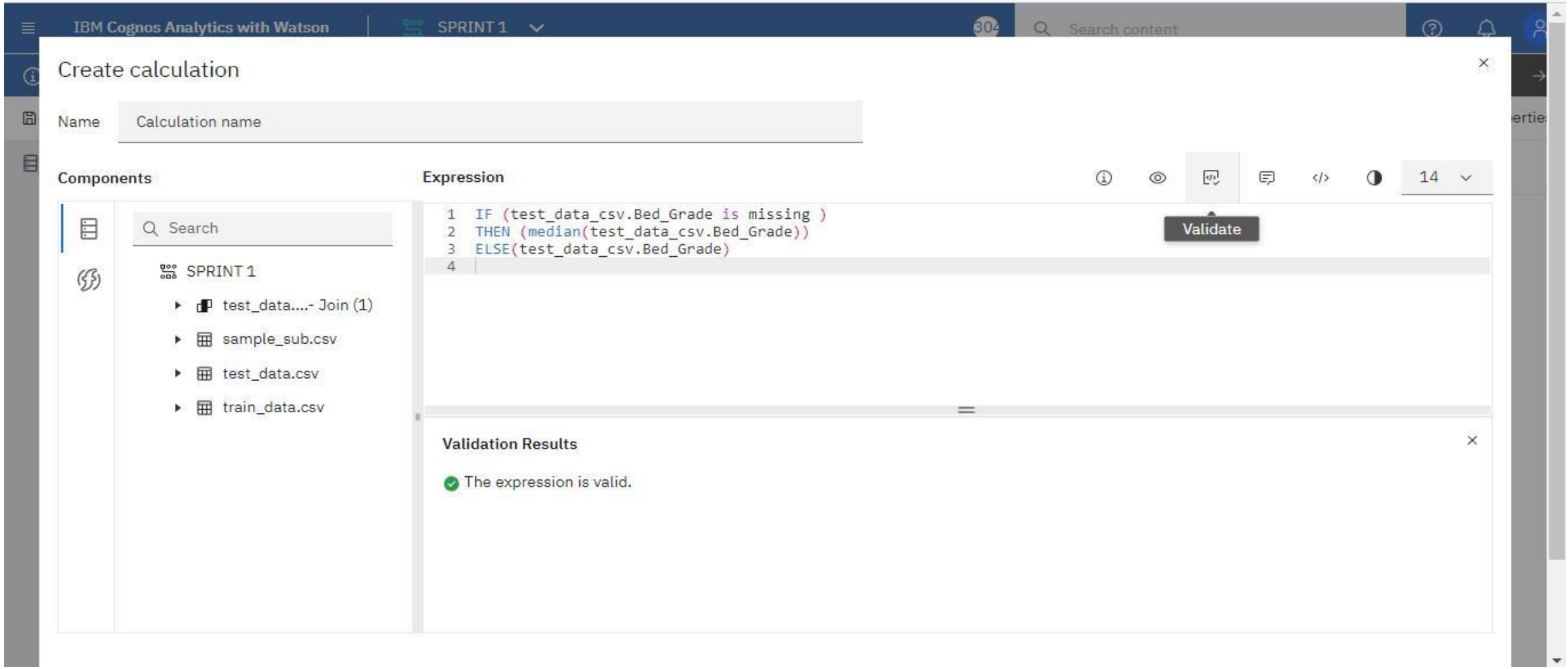


DATASET PREPARATION

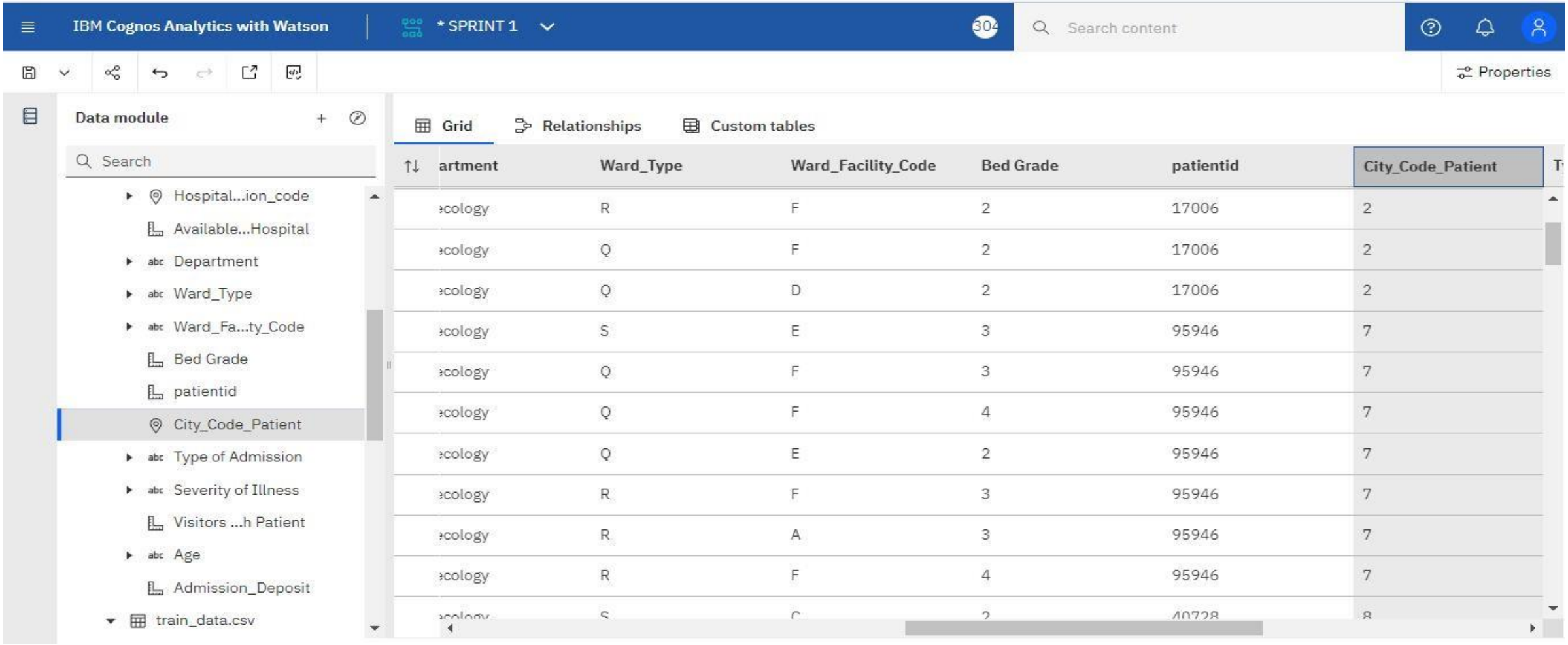


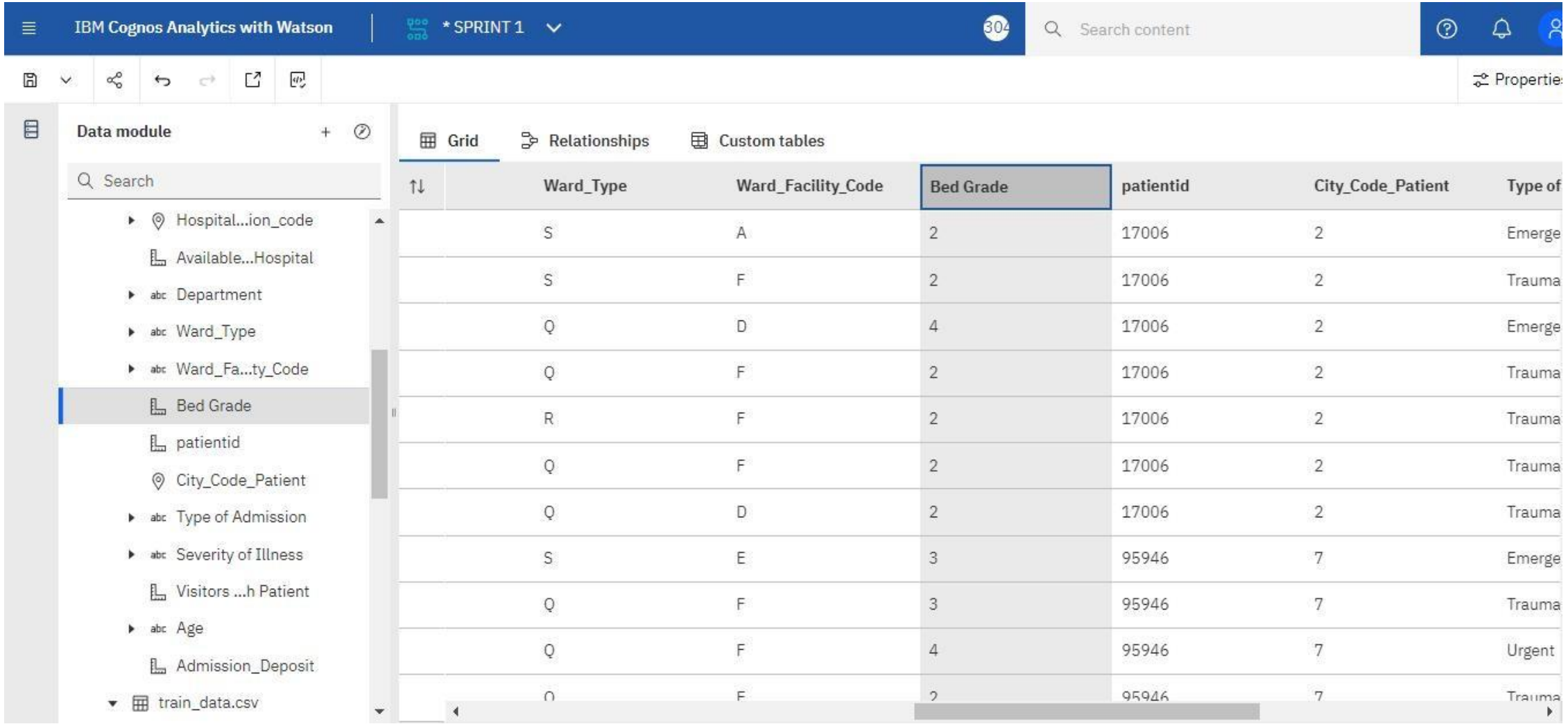
DATA PREPARATION




IBM Cognos Analytics with Watson
SPRINT 1
304
Search content
Create calculation
Name
Calculation name
Components
Expression
14
Search
SPRINT 1
test_data....- Join (1)
sample_sub.csv
test_data.csv
train_data.csv
1 IF (test_data_csv.Bed_Grade is missing)
2 THEN (median(test_data_csv.Bed_Grade))
3 ELSE(test_data_csv.Bed_Grade)
4
Validate
Validation Results
The expression is valid.

Replacing NULL into values





	Ward_Type	Ward_Facility_Code	Bed Grade	patientid	City_Code_Patient	Type of
	S	A	2	17006	2	Emerge
	S	F	2	17006	2	Trauma
	Q	D	4	17006	2	Emerge
	Q	F	2	17006	2	Trauma
	R	F	2	17006	2	Trauma
	Q	F	2	17006	2	Trauma
	Q	D	2	17006	2	Trauma
	S	E	3	95946	7	Emerge
	Q	F	3	95946	7	Trauma
	Q	F	4	95946	7	Urgent
	Q	F	2	95946	7	Trauma

PREPARED DATA LINK

https://us3.ca.analytics.ibm.com/bi/?perspective=ca-modeller&pathRef=.my_folders%2FSPRINT%2B1

.