Analytics For Hospitals' Health-Care Data

1.1 ABSTRACT

Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas to focus: HealthcareManagement. While healthcare management has various use cases for using data science, patient length of stay is one critical parameter to observe and predict if one wants to improve the efficiency of the healthcare management in a hospital.This parameter helps hospitals to identify patients of high LOS-risk (patients who will stay longer) at the time of admission. Once identified, patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection. Also, prior knowledge of LOS can aid in logistics such as room and bed allocation planning.Suppose you have been hired as Data Scientist of Health Man – a not for profit organization dedicated to manage the functioning of Hospitals in a professional and optimal manner.

1.2 MOTIVATION & GOAL

The goal is to accurately predict the Length of Stay for each patient on case by case basis so that the Hospitals can use this information for optimal resource allocation and better functioning. The length of stay is divided into 11 different classes ranging from 0-10 days to more than 100 days.

1.3 TECHNICAL ARCHITECTURE

1.4 FINAL DELIVERABLES

1.4.1 SPRINT 1

In First sprint, We took health care dataset from kaggle and preprocessing both in jupiter notebook and IBM cognos.To find out null values present in the dataset and change it into a certain value.

```
In [6]:   d.isnull()
```

Out[6]:

|	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	City_Co
0	False	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...	...
137052	False	False	False	False	False	False	False	False	False	False	False	False
137053	False	False	False	False	False	False	False	False	False	False	False	False
137054	False	False	False	False	False	False	False	False	False	False	False	False
137055	False	False	False	False	False	False	False	False	False	False	False	False
137056	False	False	False	False	False	False	False	False	False	False	False	False

137057 rows × 17 columns

After preprocessing the data the dataset is explored in IBM cognos through various graphs and tabular columns by comparing it with differnet sets of attributes.
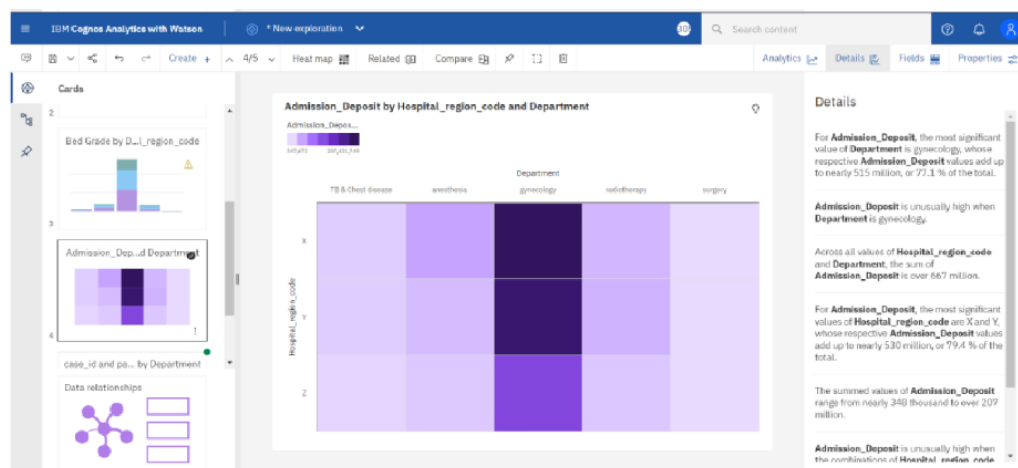
1.4.2 SPRINT 2

After preprocessing and exploring the health care data.Now upload the preprocessed data to dashboard.In here you can able to create dashboard's for each and single graph representation.This will help to make sure to find the Average availability of beds in the hospitals.
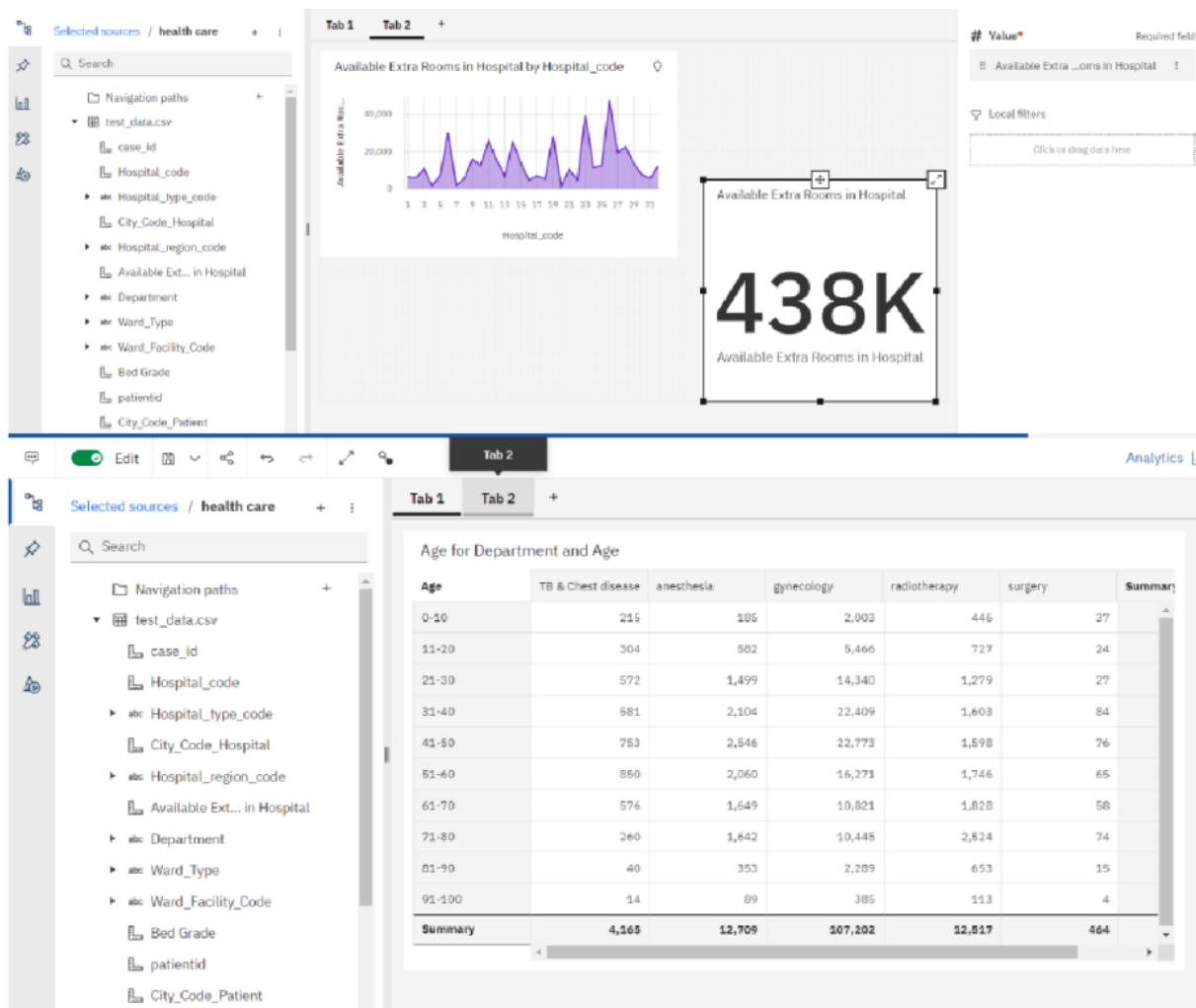
1.4.3 SPRINT 3

Finally after creating a dashboard for the health-care dataset.Now build a report and story using same method. Report can be build by importing the data exploration templates or templates which was available in the dashboard.This will make sure to create a organized way of report and story of the given dataset even layman can understand.
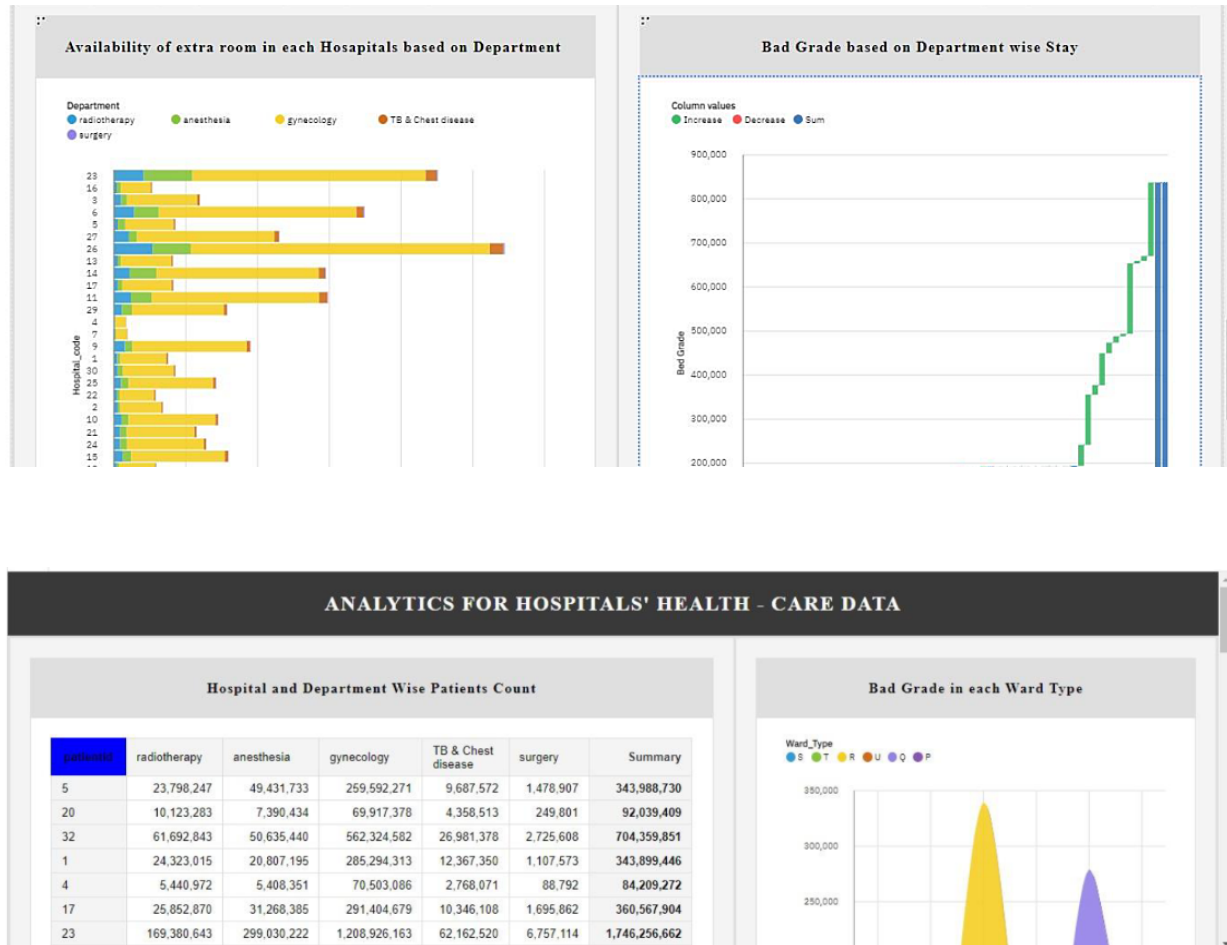
To create a story,which is quite similar to slideshare.We need to build the templates and place it in form of powerpoint and allott them with given time for each slide.So that we can able to create a video representation of the received graph representation

1.4.4 Sprint 4

In final sprint we will document the process we followed thorughout the entire project to give desired results and submit it.
Finally,we came to :
->Know fundamental concepts and can work on IBM Cognos Analytics
->Gain a broad understanding of plotting different visualizations to provide the suitable solution.
->Able to create meaningful Visualizations and the Dashboard(s).