PROJECT REPORT

- **PROJECT TITLE_:** Estimation and prediction of hospitalisation and medical care costs.
- **PROJECT ID :** LTVIP2023TMID01124
- **TEAM SIZE :** 5
- **TEAM LEADER** : Gudivada likhitha
- **TEAM MEMBER :** Gudivada vamsi krishna
- **TEAM MEMBER :** Guntreddi yaswanth ram
- **TEAM MEMBER :** Katari Deepika
- **TEAM MEMBER :** Manga Jhansi
- **FACULTY MENTOR :** Rowthu murali krishna

1. INTRODUCTION

1.1 OVERVIEW

Estimation and prediction of hospitalization of medical care costs are critical aspects of healthcare planning, resource allocation , and financial management . These processes models, and other techniques to project future expenses related to hospital stays and medical treatments .

Estimation is the process of calculating or approximating the potential costs associated with hospitalization and medical care . This typically involves analyzing historical data on patient admissions, treatment, and expenses to identify patterns and trends . Estimation is useful for budgeting purposes , setting insurance premiums, and determining reimbursement rates for healthcare providers.

Prediction involves forecasting future hospitalization and medical care costs based on existing data and trends . It takes estimation a step further by looking ahead and projecting potential expenses .

Predictive models utilize historical data , patient outcomes etc…other relevant variable to make future predictions .

- **Data collection and preprocessing**

A raw dataset was collected from **Kaggle** which contains the variables of **Sex, Region, Smoker, Charges, Bmi, Children..**

And then we have to convert the raw dataset into a data module by removing the **Inconsistent data, Null values, Alter the variable names etc…**

- **EDA(Exploratory data analysis)**

EDA is the process of transforming the dataset into visualsations which helps to understanding the status **diagrammatically , graphically** and also helps to make future predictions by taking the resultant outcomes.

We are using **IBM COGNOS ANALYTICS** tool to create such types of visualsations.

We can results the outcome in the form of **DASHBOARD , REPORT** and **STORY** .

- **WEB INTEGRATION BY CREATING FLASK**

For Estimating and Prediction of Hospitalization of Medical Care Costs data involves building an interface where we input the relevant data , and the application will use the predictive model to estimate the medical care costs.

1.2 PURPOSE

Estimation is the process of calculating or approximating the potential costs associated with hospitalization and medical care . This typically involves analyzing historical data on patient admissions, treatment, and expenses to identify patterns and trends . Estimation is useful for budgeting purposes , setting insurance premiums, and determining reimbursement rates for healthcare providers.

Prediction involves forecasting future hospitalization and medical care costs based on existing data and trends . It takes estimation a step further by looking ahead and projecting potential expenses .

Predictive models utilize historical data , patient outcomes etc…other relevant variable to make future predictions

KEY OUTCOMES

- **Health care Budgeting and Financial Planning**
- **Resource allocation**
- **Insurance premium setting**
- **Clinical decision making**
- **Enhanced patient care**

2. LITERATURE SURVEY

The prevalence of obesity , which is defined as a body mass index (BMI) greater than 30, has increased dramatically in the **US** since 1990s. So much so that recently obesity has been officially recognized as a disease by the American Medical Association , an action that could put more emphasis on the health condition by doctors and insurance companies to minimize its adverse effects.

Currently, rates of obesity exceeds of 30% in the ost sex and adult groups, whereas its prevalence among children and adolescents, defined as a BMI of more than the 95 th percentile, has reached 17% .

2.1 Existing Problem

Solving the Estimation and Prediction of Hospitalization of Medical care costs involves the systematic approach .

2.2 Proposed Solution

- **Data Collection and Data Preprocessing**
  - Removing inconsistent data
  - Removing null values
  - Removing the duplicates
- **EDA(Exploratory Data Analysis)**
  - DASH BOARD
  - REPORT
  - STORY
- **Web Integration**
  - Using flask

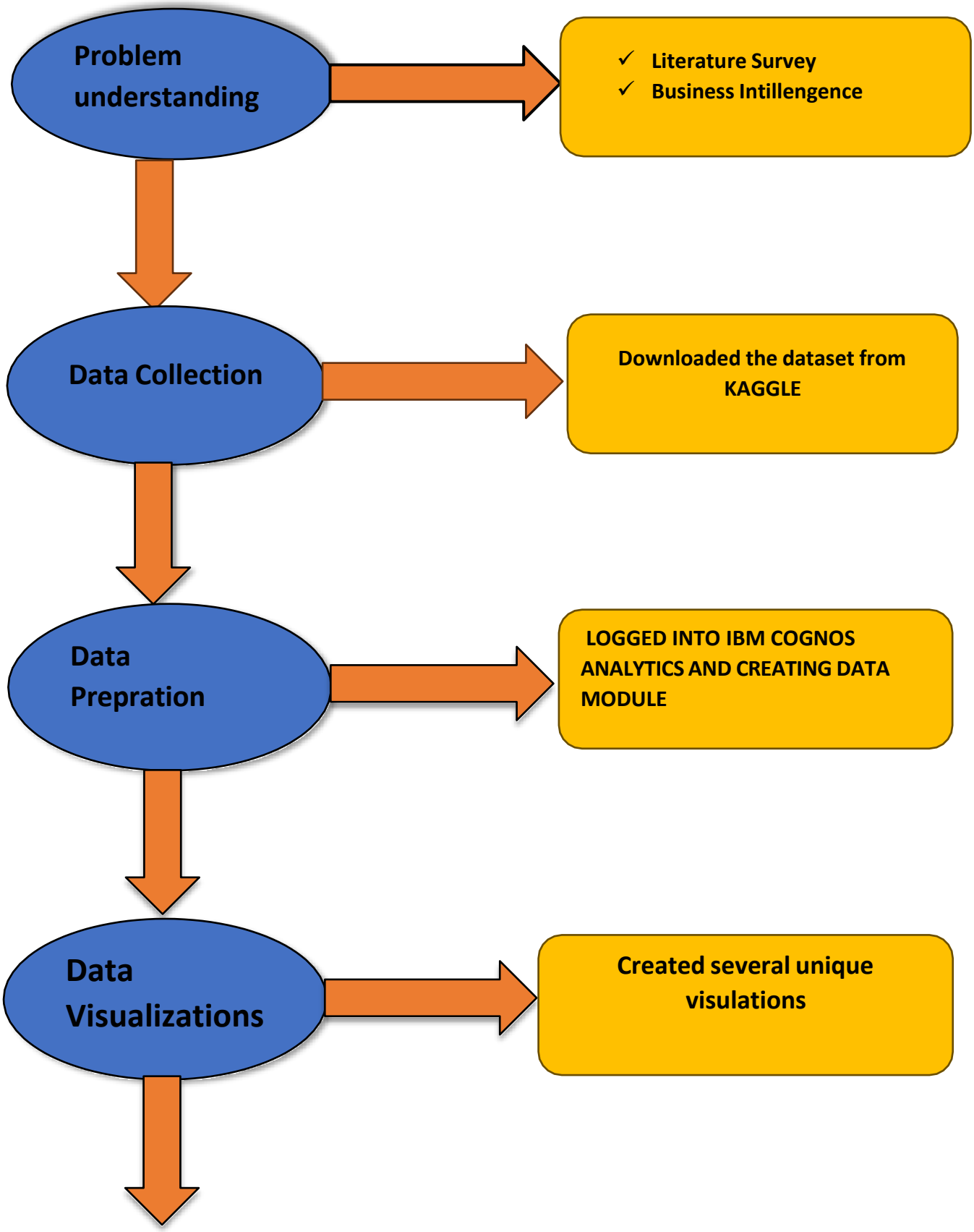
3. THEORTICAL ANALYSIS

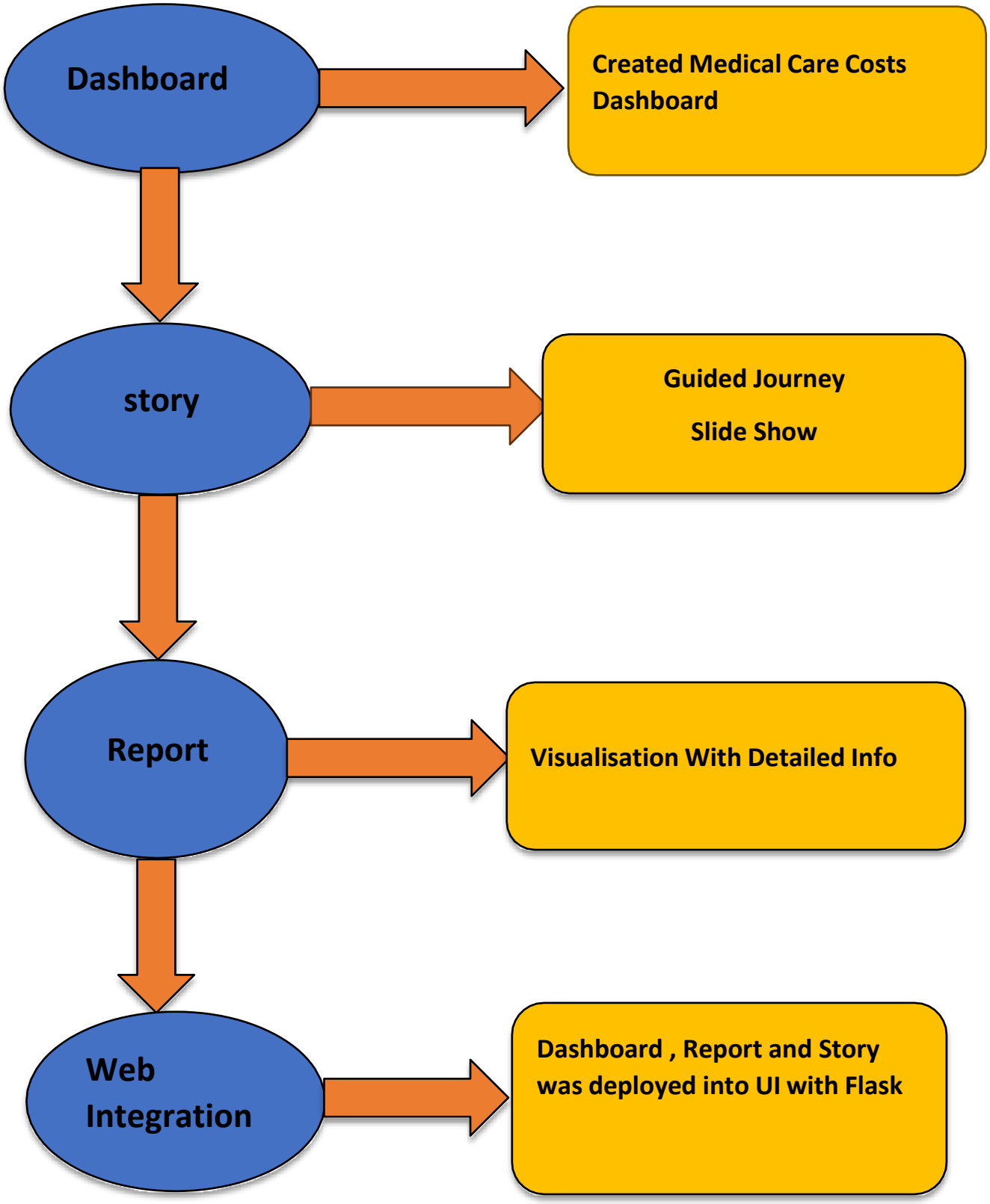
3.1 Block Diagram

Creating a detailed block diagram for estimation and prediction of hospitalization of medical care costs involves breaking down the process into key steps and components. Below there was high-level block diagram outlining the main stages.

The block diagram illustrates the step by step procedure of **ESTIMATION AND PREDICTION OF HOSPTIALIZATION OF MEDICAL CARE COSTS** , starting from data collection and preprocessing to deploying the final models for cost estimation and future prediction.

○ Block Diagram of Estimation and Predictiontion of Hospitalization of Medical Care Costs





3.2 SOFTWARE AND HARDWARE DESIGNING

Software Requirements:

- IBM cognos analytics tool

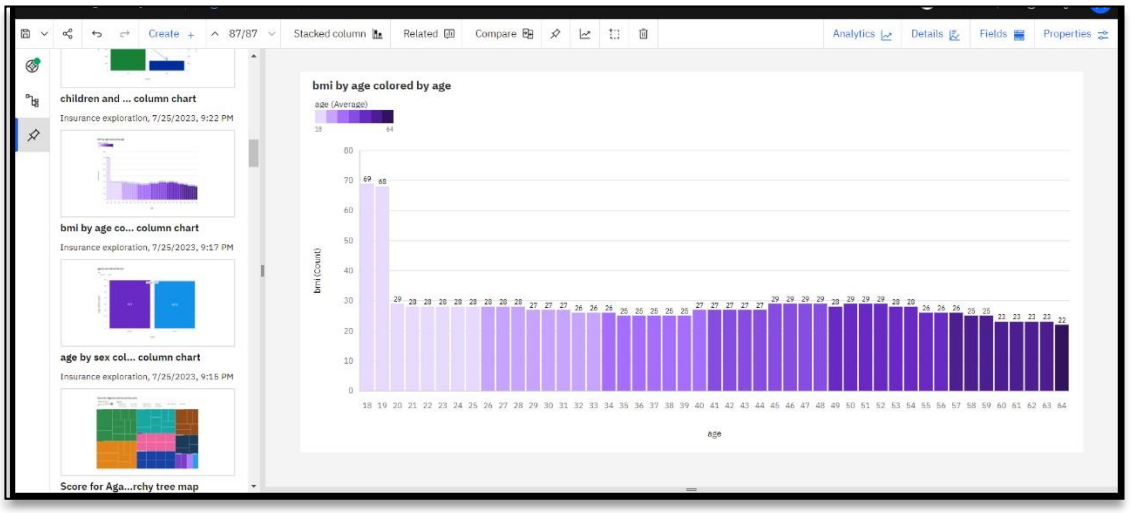
- Flask
- Integrated Development Environment(IDE)-Spyder

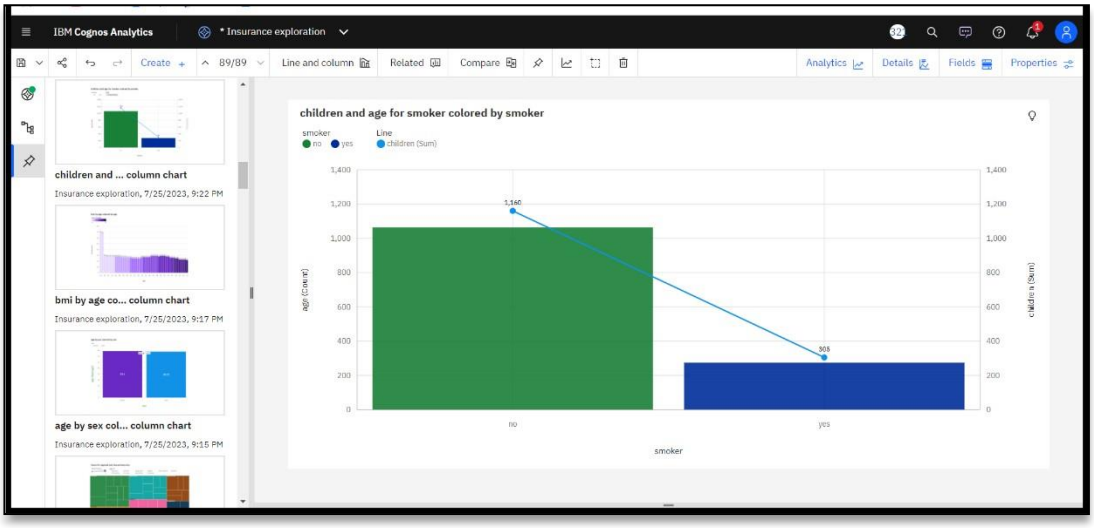
Hardware Requirements:

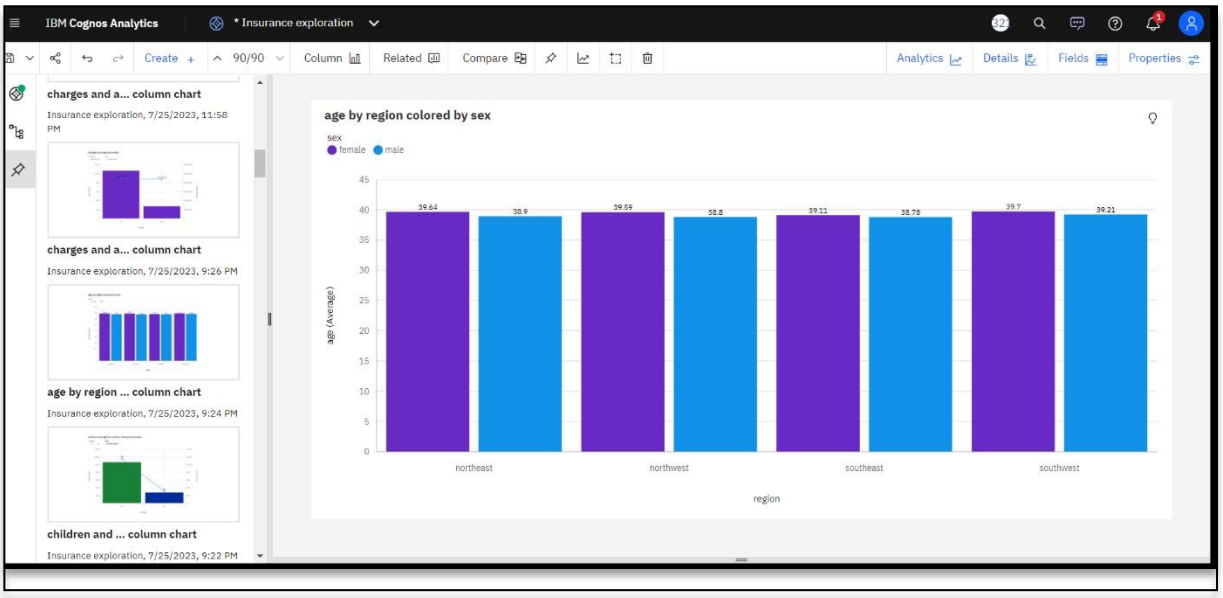
- Minimum System Requirements (RAM-4GB,Quad core processor or Above)

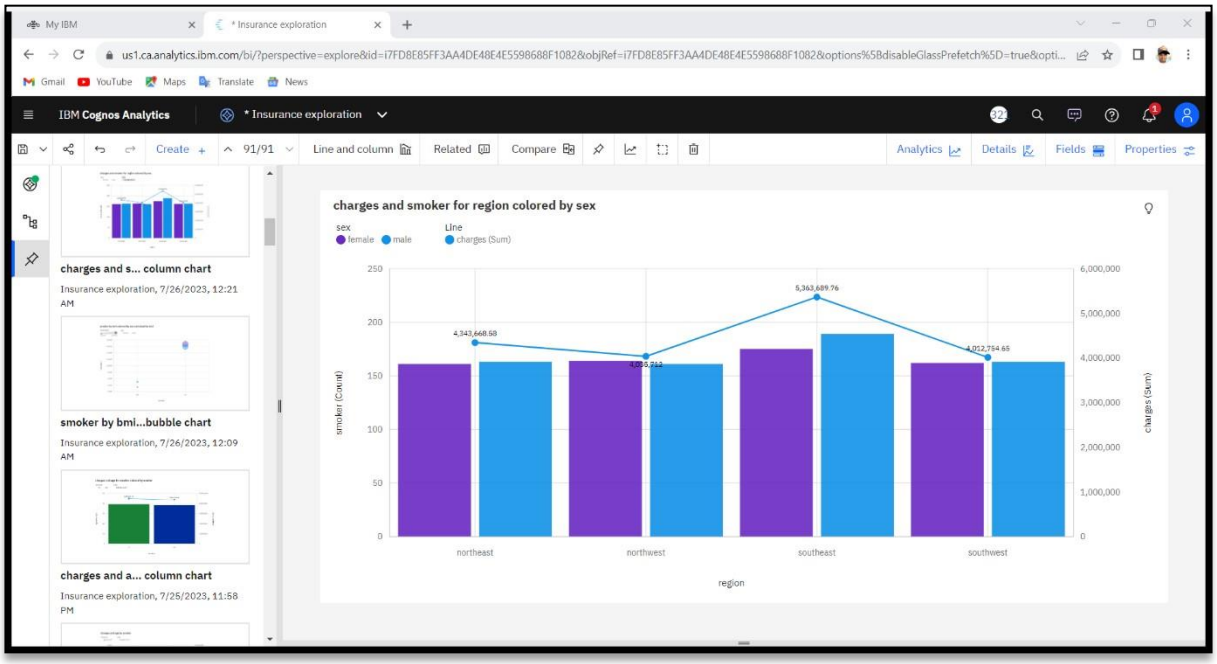
4. RESULT

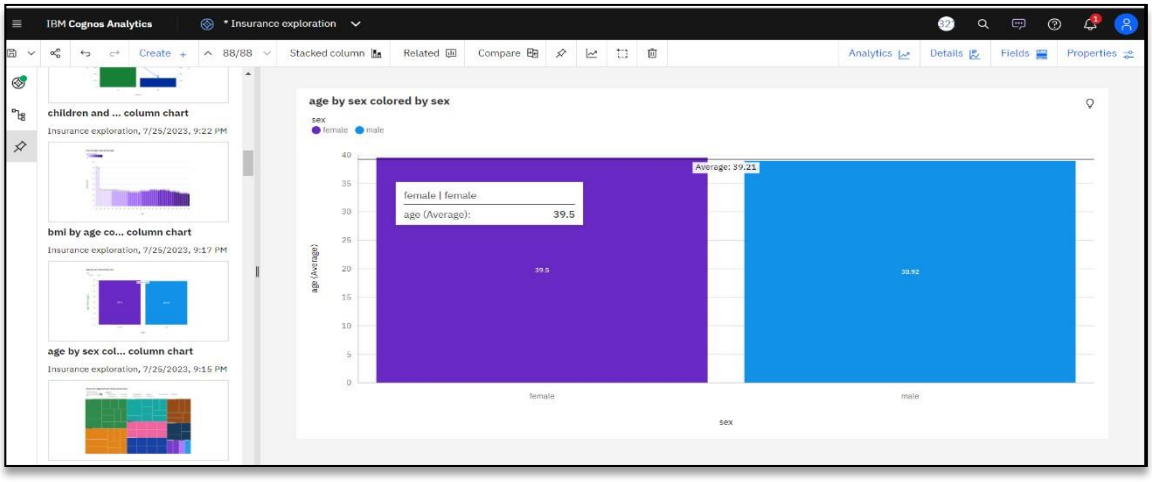
DATA VISUALISATIONS

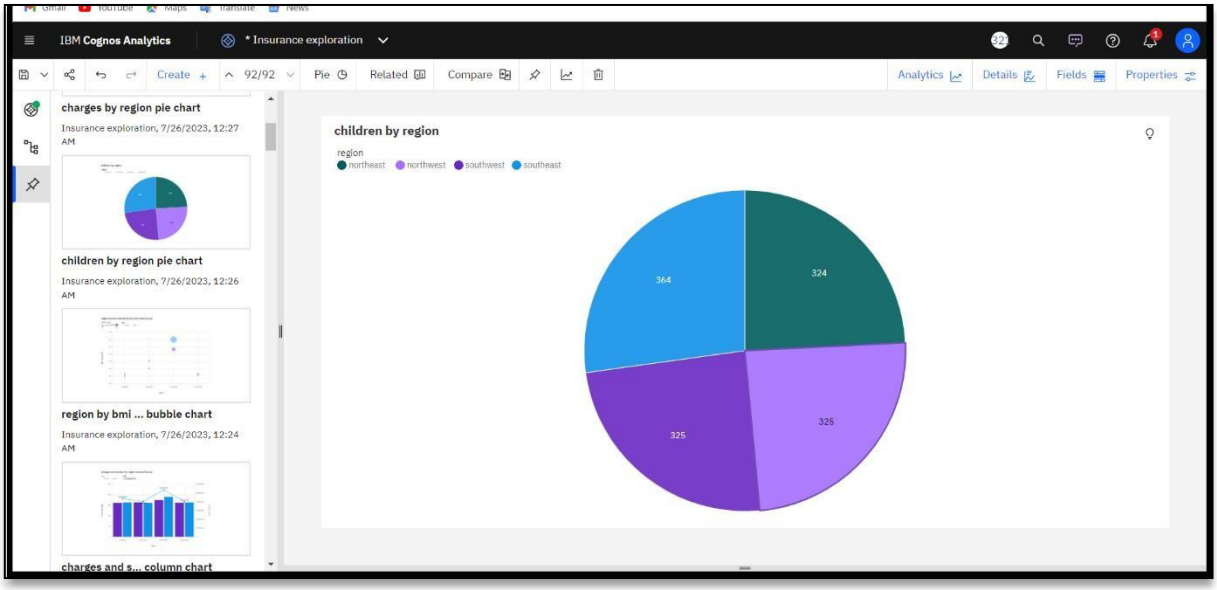


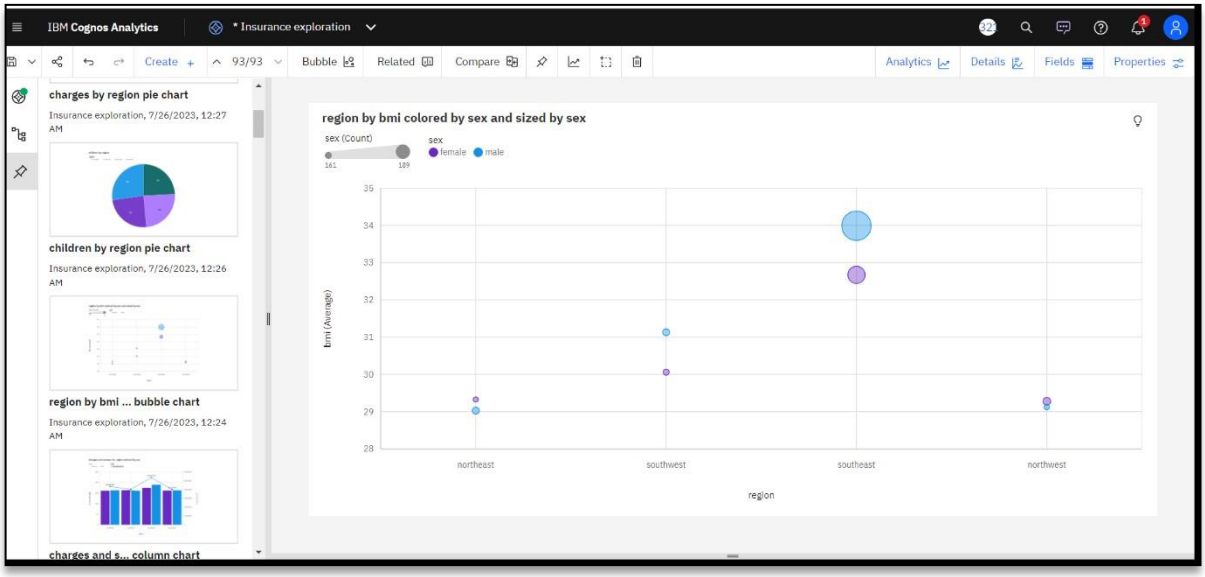


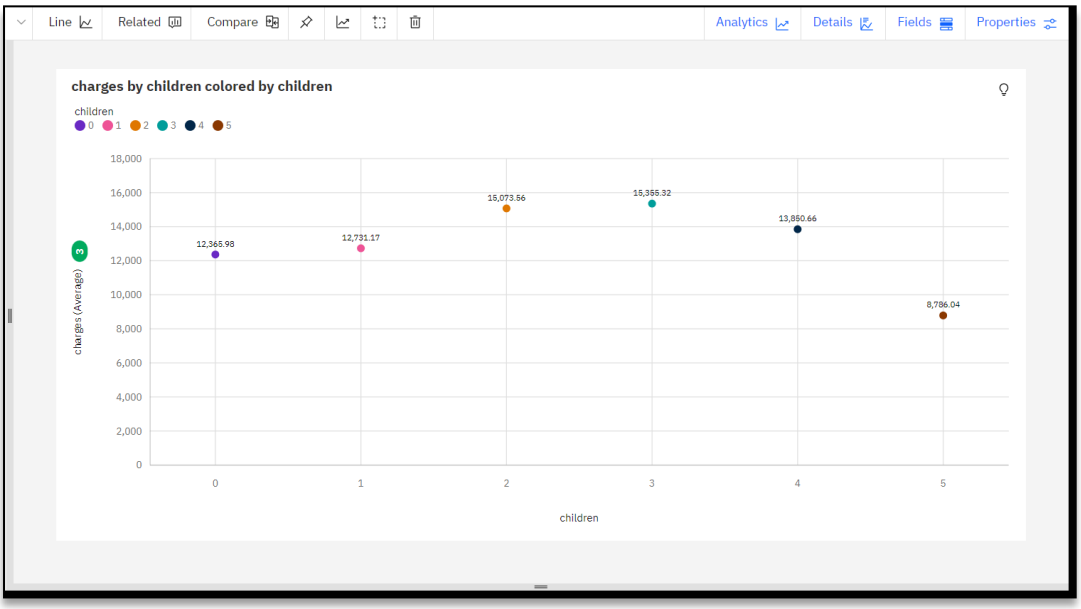

IBM Cognos Analytics
* Insurance exploration
Create
90/90
Column
Related
Compare
Analytics
Details
Fields
Properties
charges and a... column chart
Insurance exploration, 7/25/2023, 11:58 PM
charges and a... column chart
Insurance exploration, 7/25/2023, 9:26 PM
age by region ... column chart
Insurance exploration, 7/25/2023, 9:24 PM
children and ... column chart
Insurance exploration, 7/25/2023, 9:22 PM
age by region colored by sex
sex
female
male
39.64
38.9
39.59
38.8
39.11
38.78
39.7
39.21
age (Average)
northeast
northwest
southeast
southwest
region


My IBM
* Insurance exploration
us1.ca.analytics.ibm.com/bi/?perspective=explore&id=i7FD8E85FF3AA4DE48E4E5598688F1082&objRef=i7FD8E85FF3AA4DE48E4E5598688F1082&options%5BdisableGlassPrefetch%5D=true&opti...
Gmail
YouTube
Maps
Translate
News
IBM Cognos Analytics
* Insurance exploration
Create
91/91
Line and column
Related
Compare
Analytics
Details
Fields
Properties
charges and s... column chart
Insurance exploration, 7/26/2023, 12:21 AM
smoker by bmi...bubble chart
Insurance exploration, 7/26/2023, 12:09 AM
charges and a... column chart
Insurance exploration, 7/25/2023, 11:58 PM
charges and smoker for region colored by sex
sex
female
male
Line
charges (Sum)
4,343,668.58
4,009,732
5,363,689.76
4,012,754.65
smoker (Count)
charges (Sum)
northeast
northwest
southeast
southwest
region

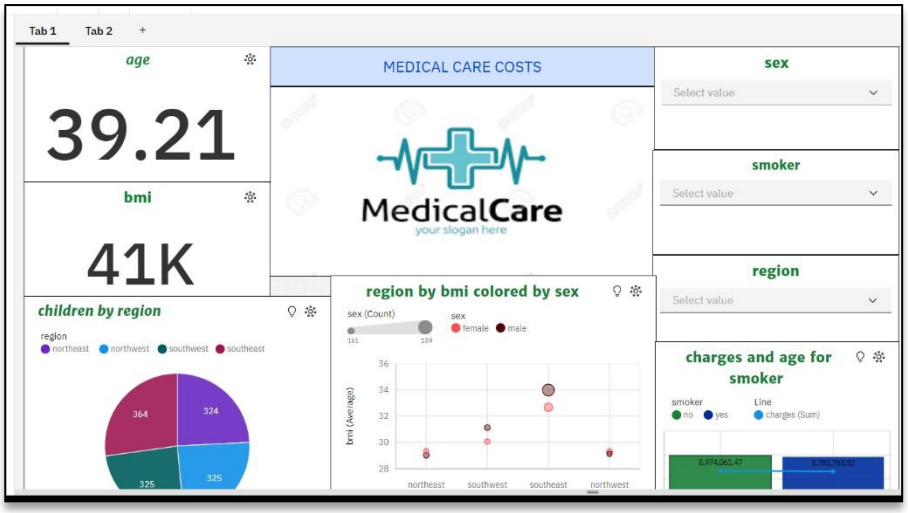

IBM Cognos Analytics
* Insurance exploration
Create
88/88
Stacked column
Related
Compare
Analytics
Details
Fields
Properties
children and ... column chart
Insurance exploration, 7/25/2023, 9:22 PM
bmi by age co... column chart
Insurance exploration, 7/25/2023, 9:17 PM
age by sex col... column chart
Insurance exploration, 7/25/2023, 9:15 PM
age by sex colored by sex
sex
female
male
female | female
age (Average): 39.5
Average: 39.21
39.5
39.92
age (Average)
female
male
sex


IBM Cognos Analytics
* Insurance exploration
Create
92/92
Pie
Related
Compare
Analytics
Details
Fields
Properties
charges by region pie chart
Insurance exploration, 7/26/2023, 12:27 AM
children by region pie chart
Insurance exploration, 7/26/2023, 12:26 AM
region by bmi ... bubble chart
Insurance exploration, 7/26/2023, 12:24 AM
charges and s... column chart
children by region
region
northeast
northwest
southwest
southeast
364
324
325
325

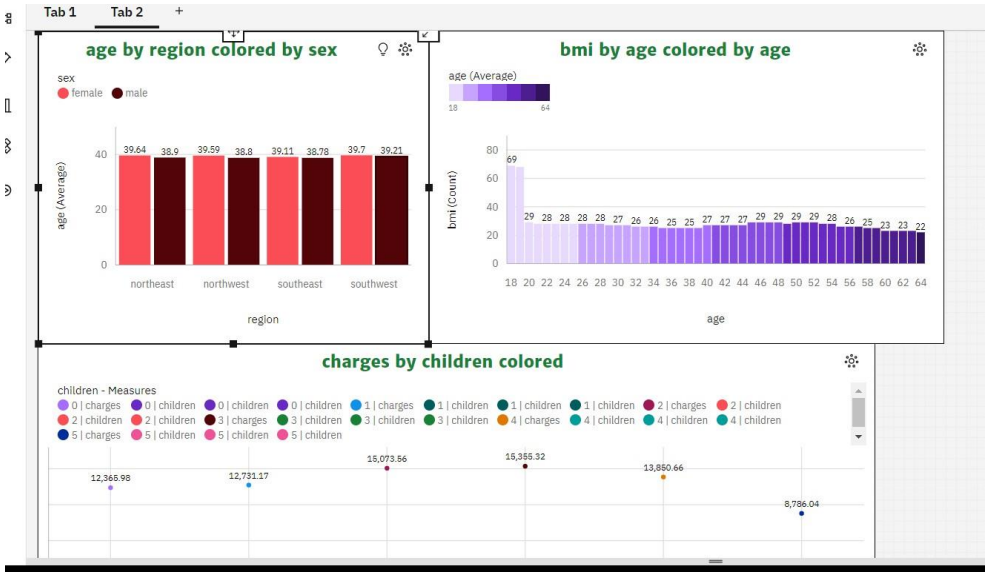

IBM Cognos Analytics
* Insurance exploration
Create
93/93
Bubble
Related
Compare
Analytics
Details
Fields
Properties
charges by region pie chart
Insurance exploration, 7/26/2023, 12:27 AM
children by region pie chart
Insurance exploration, 7/26/2023, 12:26 AM
region by bmi ... bubble chart
Insurance exploration, 7/26/2023, 12:24 AM
charges and s... column chart
region by bmi colored by sex and sized by sex
sex (Count)
161
189
sex
female
male
bmi (Average)
35
34
33
32
31
30
29
28
northeast
southwest
southeast
northwest
region


Line
Related
Compare
Analytics
Details
Fields
Properties
charges by children colored by children
children
0
1
2
3
4
5
18,000
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
charges (Average)
12,365.98
12,731.17
15,073.56
15,355.32
13,850.66
8,786.04
0
1
2
3
4
5
children

MEDICAL CARE COSTS DASHBOARD



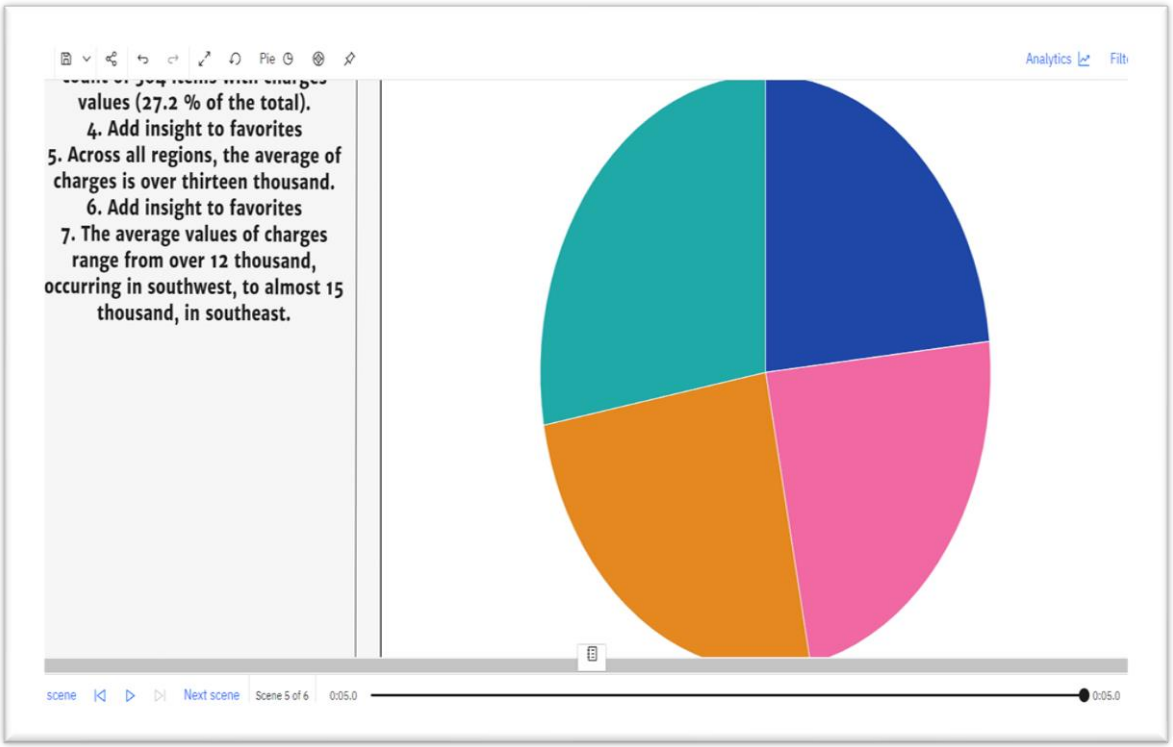
Dashboard(Tab-1)

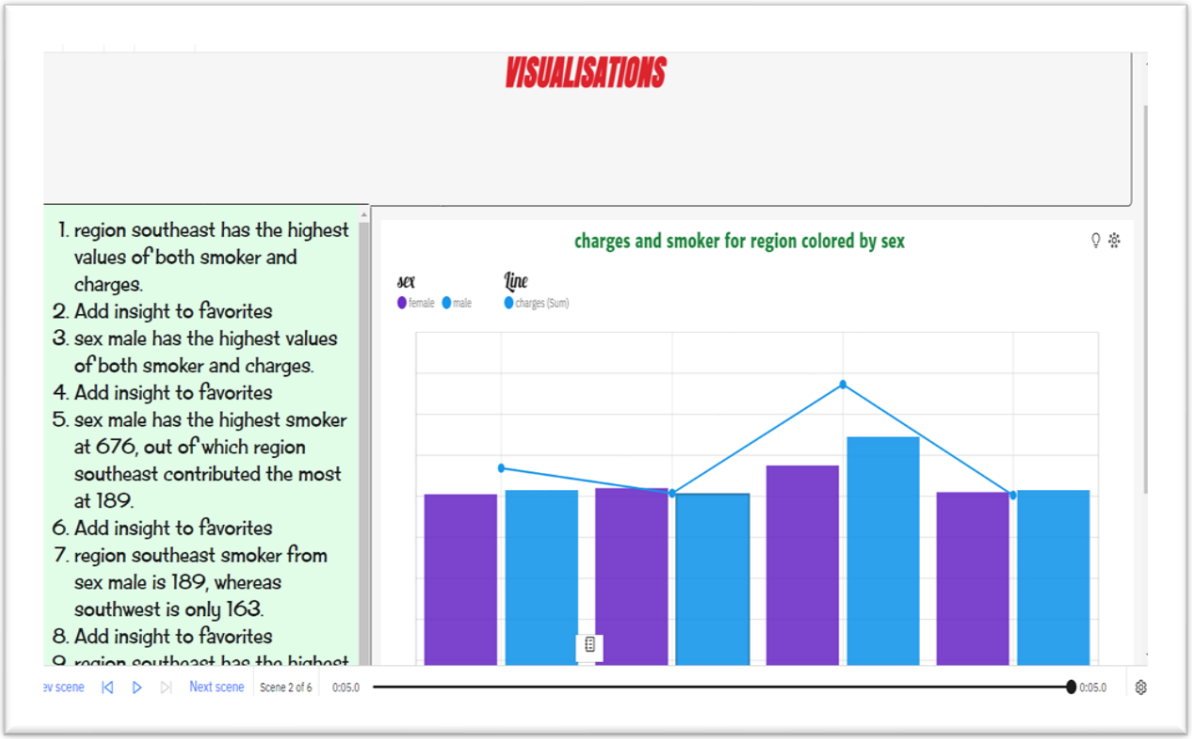


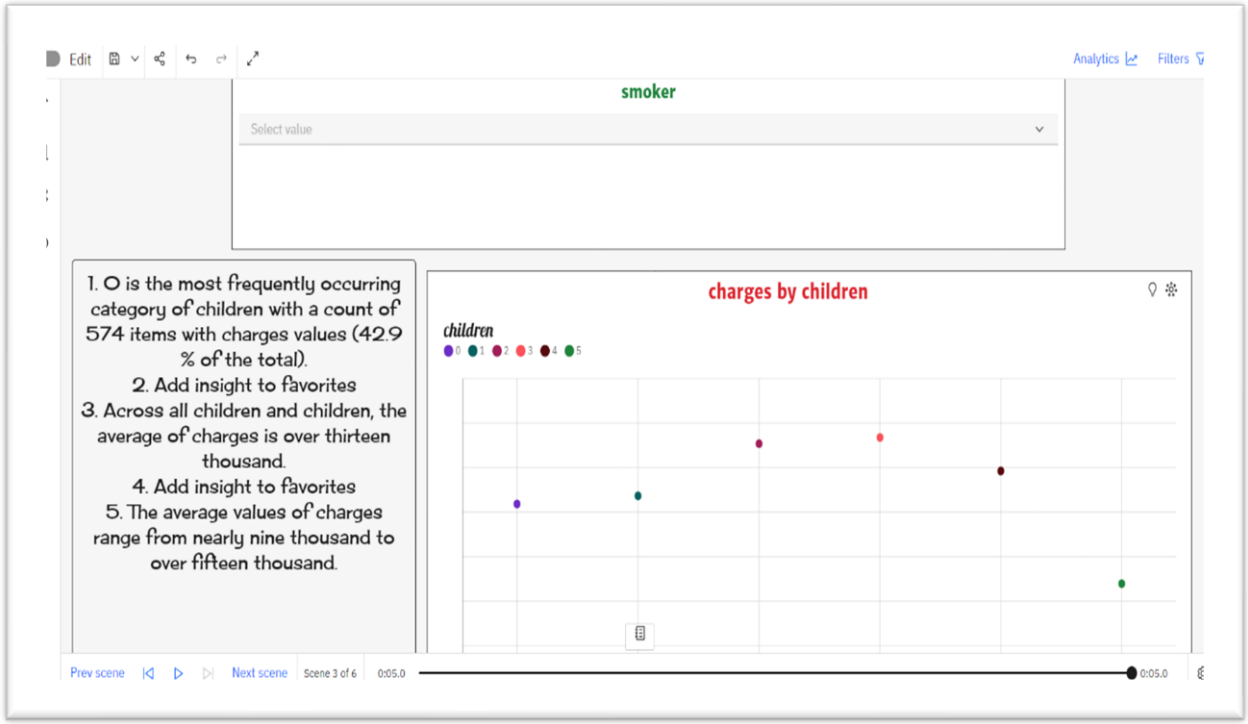
Dasboard(Tab-2)

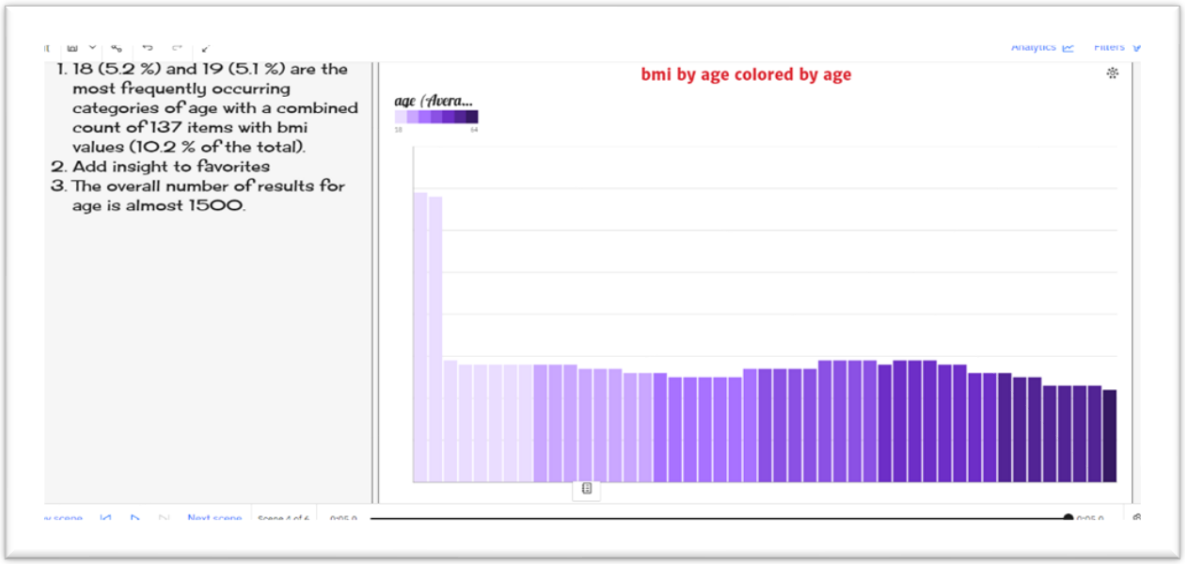
STORY(Slide Show)

STORY(Guided Journey)

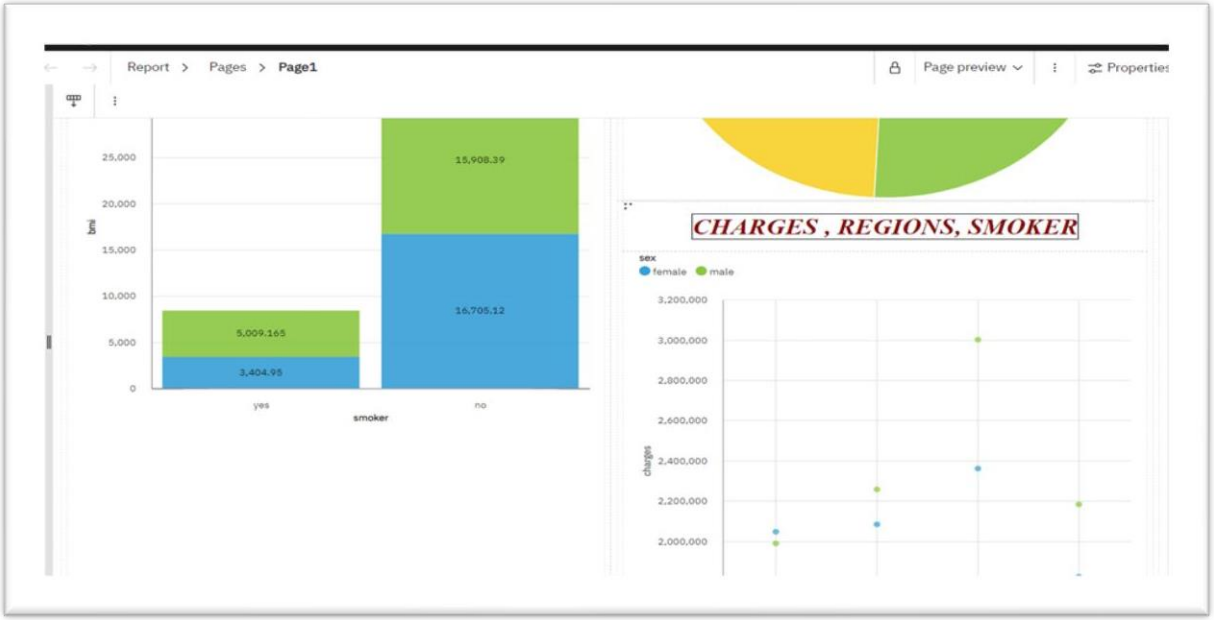

Pie
Analytics
values (27.2 % of the total).
4. Add insight to favorites
5. Across all regions, the average of charges is over thirteen thousand.
6. Add insight to favorites
7. The average values of charges range from over 12 thousand, occurring in southwest, to almost 15 thousand, in southeast.
scene
Next scene
Scene 5 of 6
0:05.0
0:05.0


VISUALISATIONS
1. region southeast has the highest values of both smoker and charges.
2. Add insight to favorites
3. sex male has the highest values of both smoker and charges.
4. Add insight to favorites
5. sex male has the highest smoker at 676, out of which region southeast contributed the most at 189.
6. Add insight to favorites
7. region southeast smoker from sex male is 189, whereas southwest is only 163.
8. Add insight to favorites
charges and smoker for region colored by sex
sex
female
male
Line
charges (Sum)
Next scene
Scene 2 of 6
0:05.0
0:05.0


Edit
Analytics
Filters
smoker
Select value
1. 0 is the most frequently occurring category of children with a count of 574 items with charges values (42.9 % of the total).
2. Add insight to favorites
3. Across all children and children, the average of charges is over thirteen thousand.
4. Add insight to favorites
5. The average values of charges range from nearly nine thousand to over fifteen thousand.
charges by children
children
0
1
2
3
4
5
Prev scene
Next scene
Scene 3 of 6
0:05.0
0:05.0

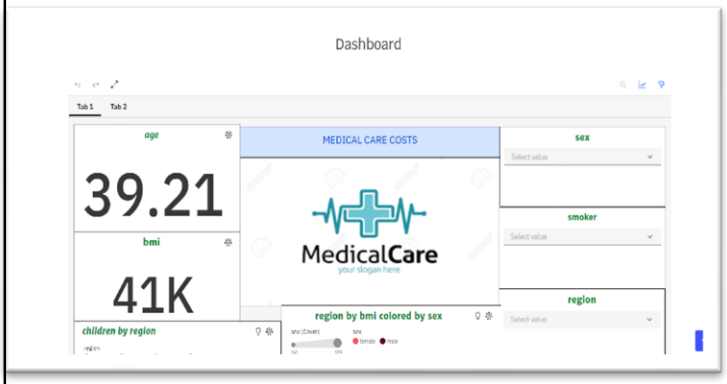

1. 18 (5.2 %) and 19 (5.1 %) are the most frequently occurring categories of age with a combined count of 137 items with bmi values (10.2 % of the total).
2. Add insight to favorites
3. The overall number of results for age is almost 1500.
bmi by age colored by age
age (Avera...

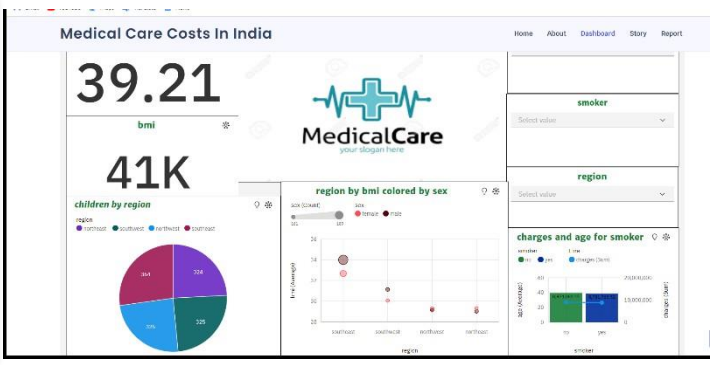
REPORT



WEB INTEGRATION

DASHBOARD





STORY

REPORT

5. ADVANTAGES

- Cost Optimization
- Informed Decision-Making
- Improved Patient Care
- Tailored insurance Coverage
- Fraud Detection
- Research and Policy development

DISADVANTAGES

- Data privacy concerns
- Data quality
- Model complexity
- Limited predictability
- Ethical Considerations
- Overemphasis on costs

6 APPLICATION:

- Healthcare cost management
- Financial Planning
- Insurance pricing and coverage
- Resource Allocation
- Treatment decision support
- Patient cost transparency
- Fraud detection
- Benchmarking and performance
- Research and public health
- Long- term cost control

7 CONCLUSION:

In conclusion, the Estimation And Prediction Of Hospitalization And Medical Care Costs project holds significant value and potential for the healthcare industry. By leveraging data analysis, exploratory data analysis, the project aims to achieve several important outcomes.

8 FUTURE SCOPE:

The future scope of the Estimation and Prediction of Hospitalization and Medical Care Costs project is vast and holds great potential in transforming the healthcare industry.

Overall, the future scope of the project is dynamic and transformative. As technology continues to evolve and data-driven decision- making becomes increasingly prevalent.