

# PROJECT REPORT

## Project Title:

ESTIMATION AND PREDICTION OF  
HOSPITALIZATION AND MEDICAL CARE COSTS.

Team ID : **LTVIP2023TMID01163**

Team Size : **5**

Team Leader : **MADDALA SRI KRISHNA  
HARSHA**

Team Member : **ADIREDDY GOPIKRISHNA**

Team Member : **KANIMERAKA HIMABINDU**

Team Member : **ROWTHU LAVANYA**

Team Member : **SHUBHASHREE PUDIPEDDI**

Faculty Mentor : **ROWTHU MURALI KRISHNA**

# 1.INTRODUCTION

## 1.1 OVERVIEW:

Estimation and prediction of hospitalization and medical care costs is a crucial aspect of healthcare planning and management. This process involves analyzing historical data, patient demographics, medical conditions and treatment outcomes to make informed projections about future healthcare expenses. By understanding these costs, healthcare providers, policymakers and insurers can allocate resources effectively and plan for potential financial challenges.

Accurate estimation and prediction of hospitalization and medical care costs are vital for creating sustainable and accessible healthcare systems, ensuring patients receive necessary care while controlling expenses.

### ➤ **Data Collection and Preprocessing:**

A comprehensive dataset was collected from **kaggle**. Which includes age, sex, region, charges, smoker, BMI. The collected data underwent thorough preprocessing to handle missing values, remove inconsistencies, and ensure data quality.

➤ **Exploratory Data Analysis (EDA):**

**EDA** was conducted to gain a deep understanding of the dataset By using **IBM COGNOS ANALYTICS** Tool. Visualizations and summary statistics helped in understanding the characteristics of the data and guided further analysis.

➤ **Creating a Flask web application:**

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

## **1.2 PURPOSE**

The purpose of the Estimation and Prediction of Hospitalization and Medical Care Costs project is to provide valuable insights and support decision-making in the healthcare sector. By developing predictive models to estimate and forecast hospitalization and medical care costs, the project aims to achieve several important goals:

- Cost Optimization
- Improved Financial Planning

- Transparency and Informed Decision-making
- Enhanced Patient Care
- Tailored Insurance Coverage
- Health Care System Performance
- Data-Driven Decision Making

Through accurate estimation and prediction of hospitalization and medical care costs, the project contributes to the broader goal of achieving a more sustainable and efficient healthcare system that benefits both healthcare providers and patients alike.

## **2.LITERATURE SURVEY**

Medical costs are one of the most common recurring expenses in a person's life. Based on different research studies, BMI, ageing, smoking and other factors are all related to greater personal medical care costs. The estimates of the expenditures of health care related to obesity are needed to help create cost-effective obesity prevention strategies. Obesity prevention at a young age is a top concern in global health, clinical practice, and public health. To avoid these restrictions, genetic variants are employed as instrumental variables in this research. Using statistics from public huge datasets, the impact of body mass index (**BMI**) on overall healthcare expenses is predicted.

A multiview learning architecture can be used to leverage BMI information in records, including diagnostic texts, diagnostic IDs and patient traits. A hierarchy perception structure was suggested to choose significant words, health checks, and diagnoses for training phase informative data representations, because various words, diagnoses and previous health care have varying significance for expense calculation. In this system model, linear regression analysis, naive Bayes classifier and random forest algorithms were compared using a business analytic method that applied statistical and machine-learning approaches. According to the results of our forecasting method, linear regression has the maximum accuracy of **97.89** percent in forecasting overall healthcare costs. In terms of financial statistics, our methodology provides a predictive method.

## **2.1 Existing problem:**

Solving the Estimation and Prediction of Hospitalization and Medical Care Costs involves a systematic approach that combines data analysis, model development and evaluation.

## **2.2 PROPOSED SOLUTION**

Proposing a solution for the estimation and prediction of hospitalization and medical care costs involves a combination of data-driven techniques, advanced analytics, and domain expertise. Collect comprehensive and diverse data related to hospitalization and medical care costs from various sources, including electronic health records, insurance claims, and administrative databases.

The success of the proposed solution depends on the availability of quality data, collaboration with healthcare experts, and a commitment to continuous improvement based on real-world feedback. Healthcare cost estimation and prediction are complex tasks, and a multidisciplinary approach is crucial for achieving accurate and reliable results.

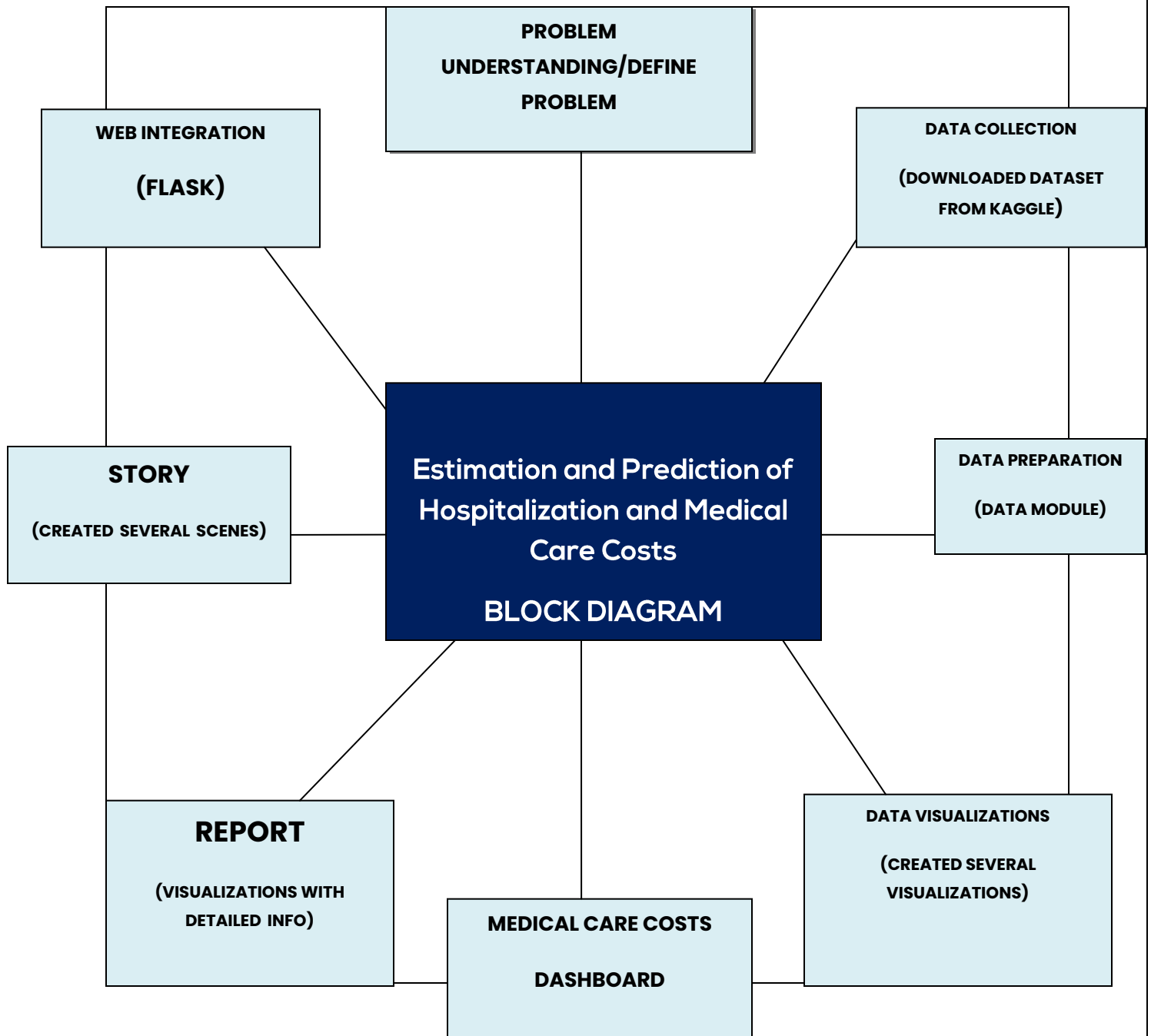
## **3. THEORETICAL ANALYSIS**

### **BLOCK DIAGRAM:**

The block diagram illustrates the end-to-end process of estimating and predicting hospitalization and medical care costs, starting from data collection and preprocessing to deploying the final models for cost estimation and future cost prediction.

# Estimation and Prediction of Hospitalization and Medical Care Costs

## BLOCK DIAGRAM



## 3.2 SOFTWARE OR HARDWARE DESIGNING

### SOFTWARE REQUIREMENTS:

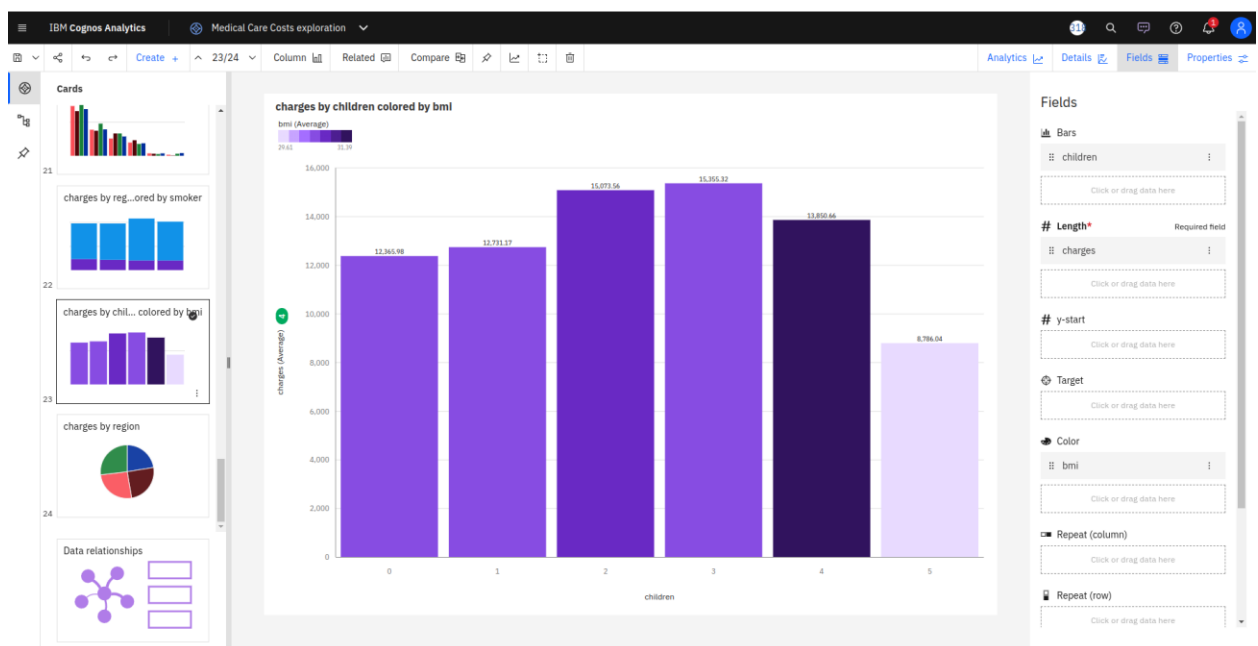
- IBM Cognos Analytics
- Flask
- Integrated Development Environment(IDE)-Spyder

### HARDWARE REQUIREMENTS:

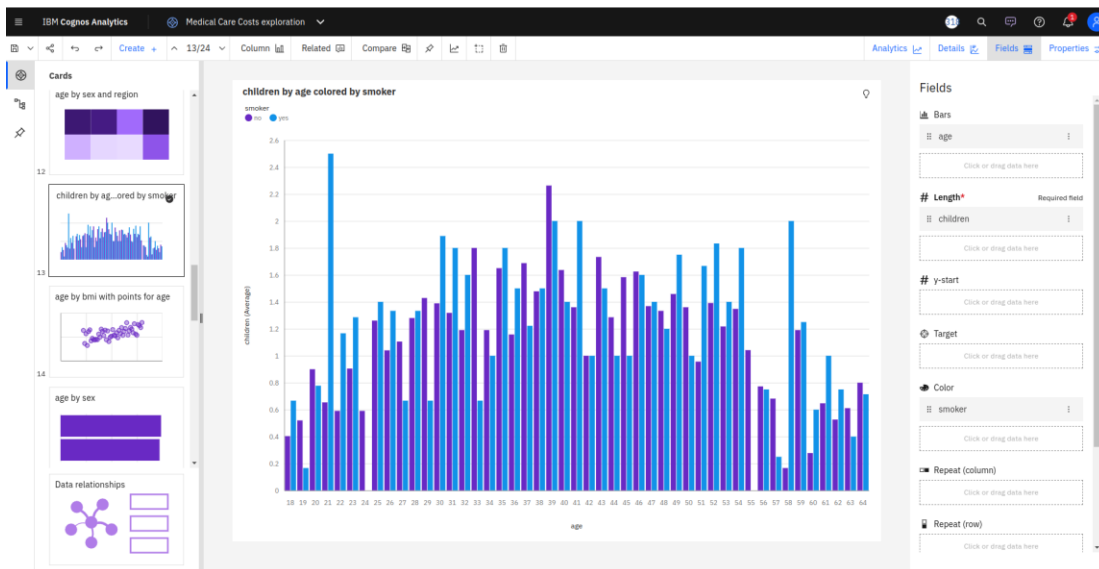
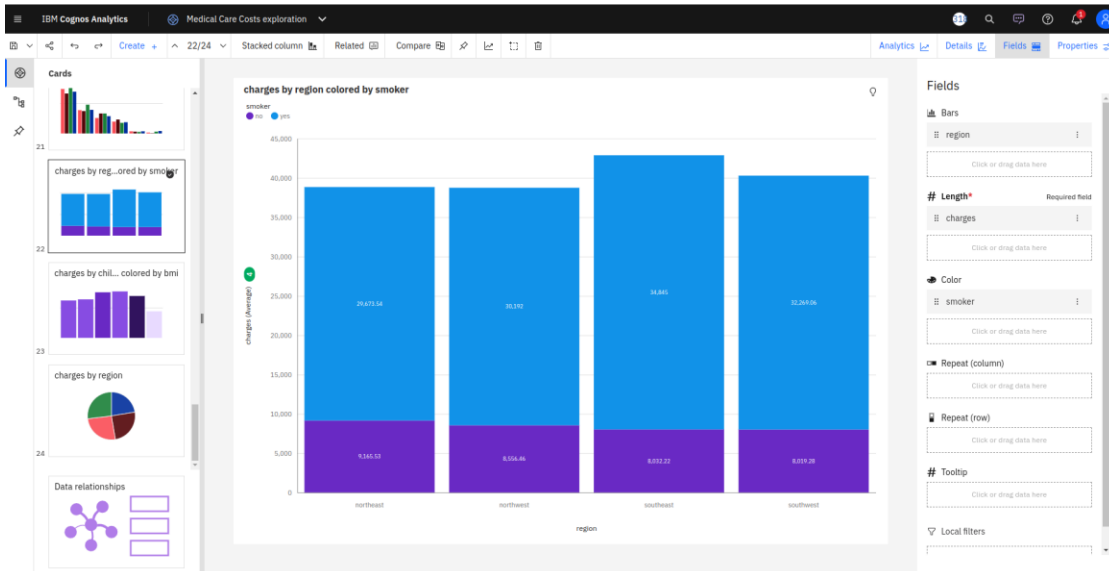
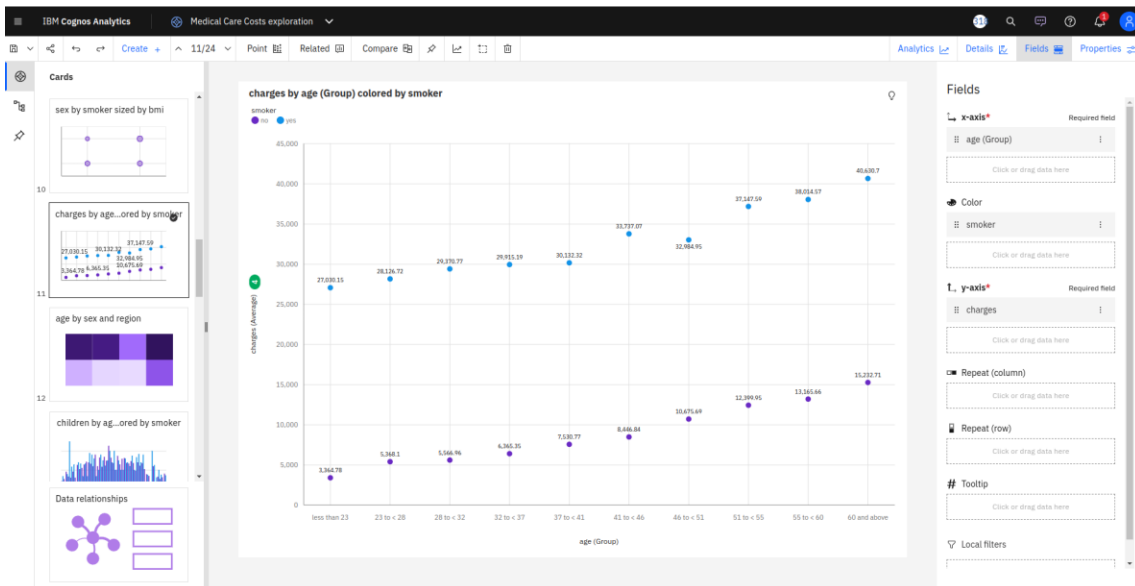
- Minimum System Requirements(RAM-4GB,Quad Core Processor Or Above)

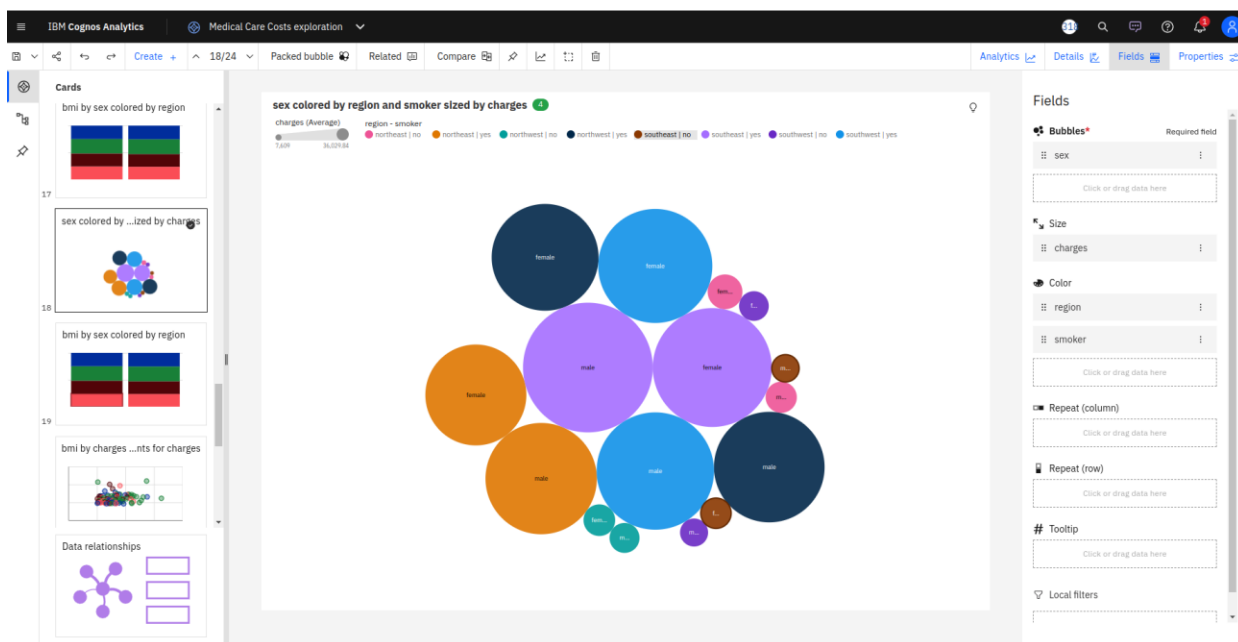
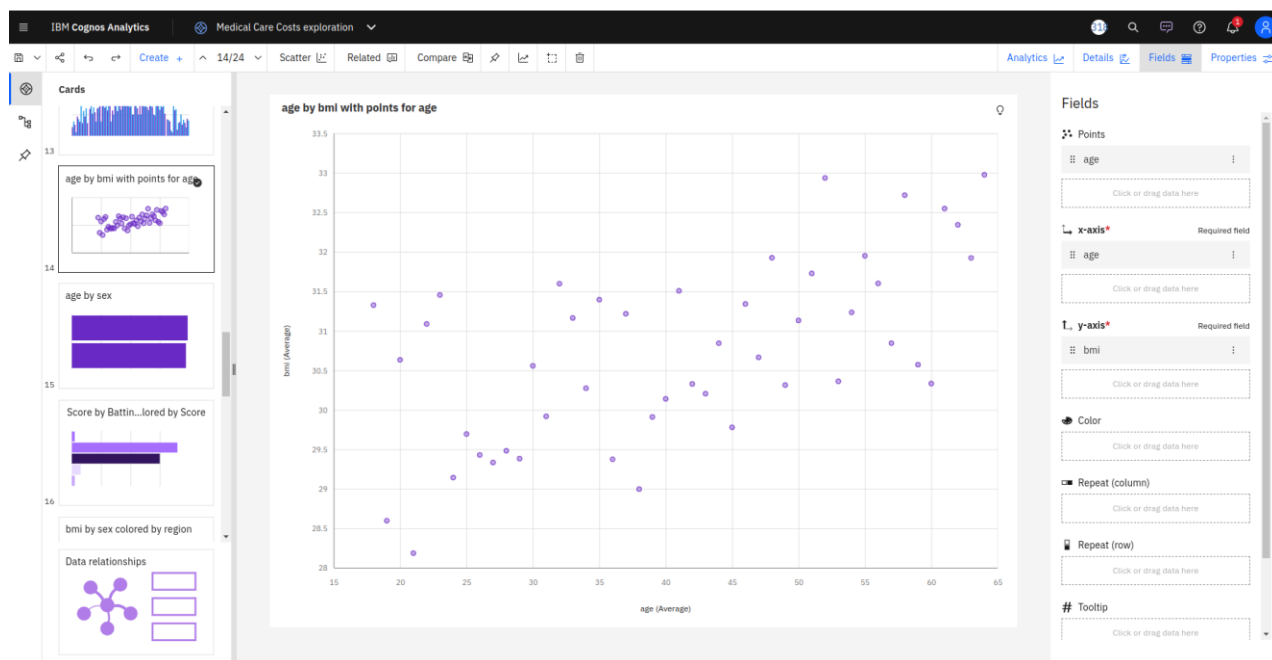
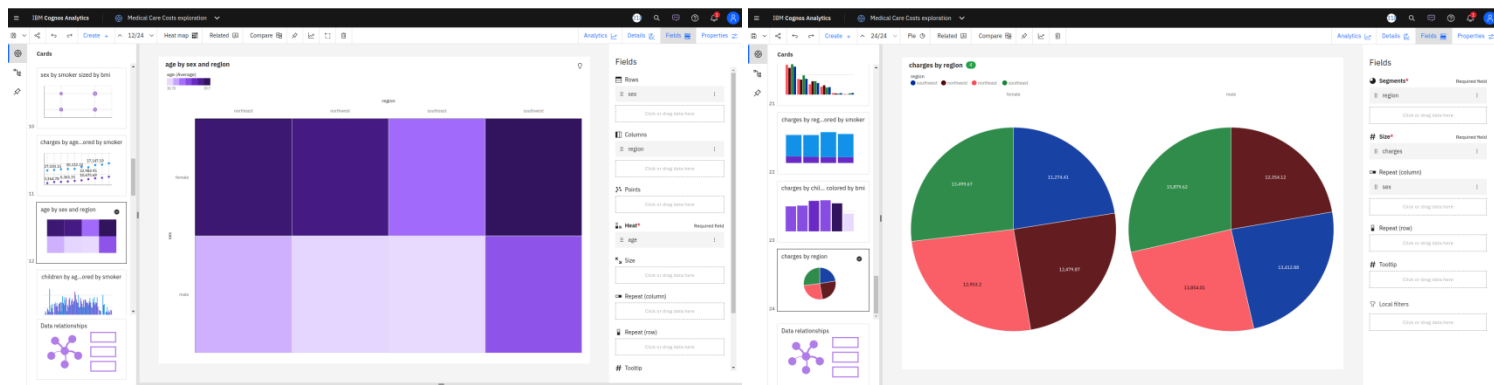
## 4.RESULT:

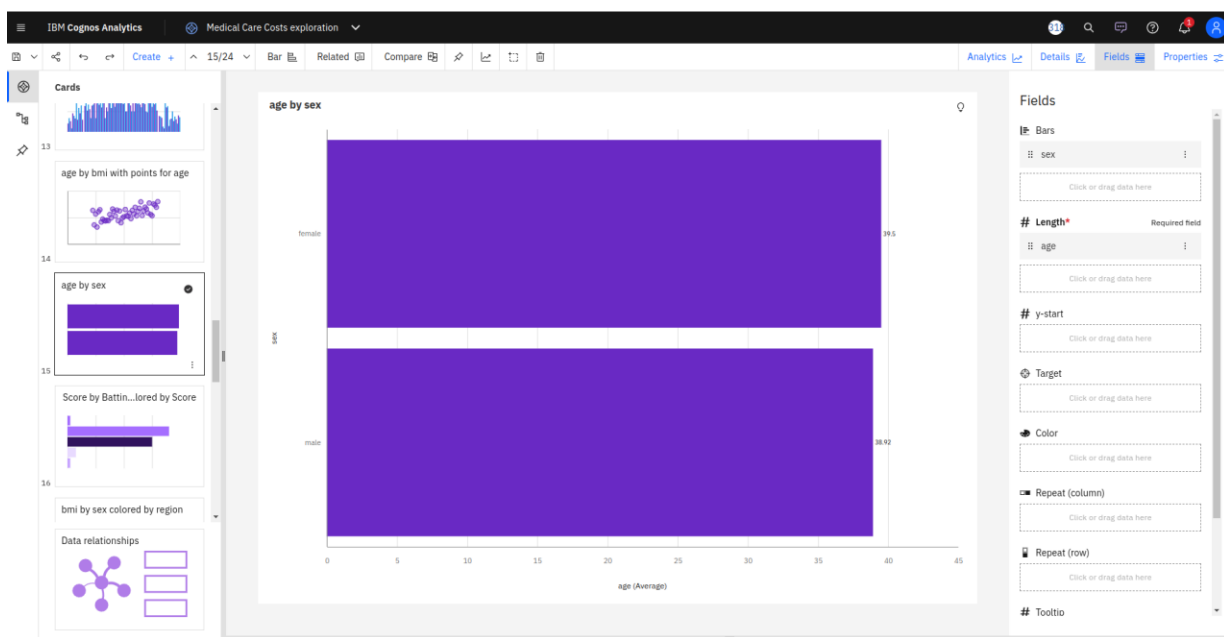
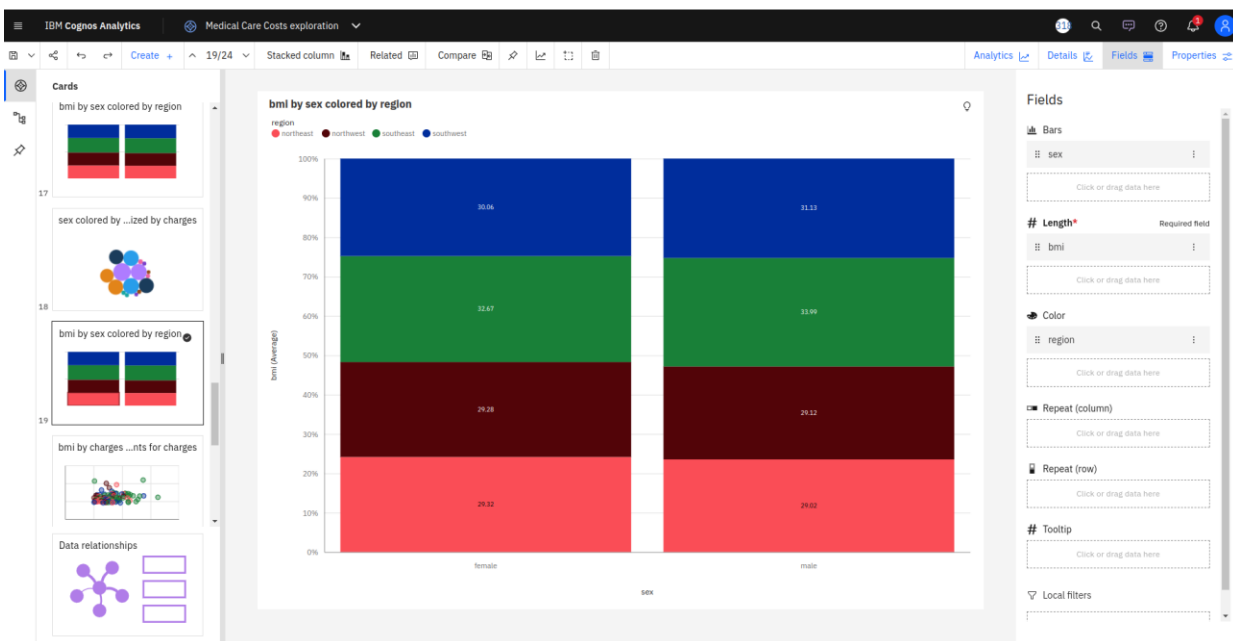
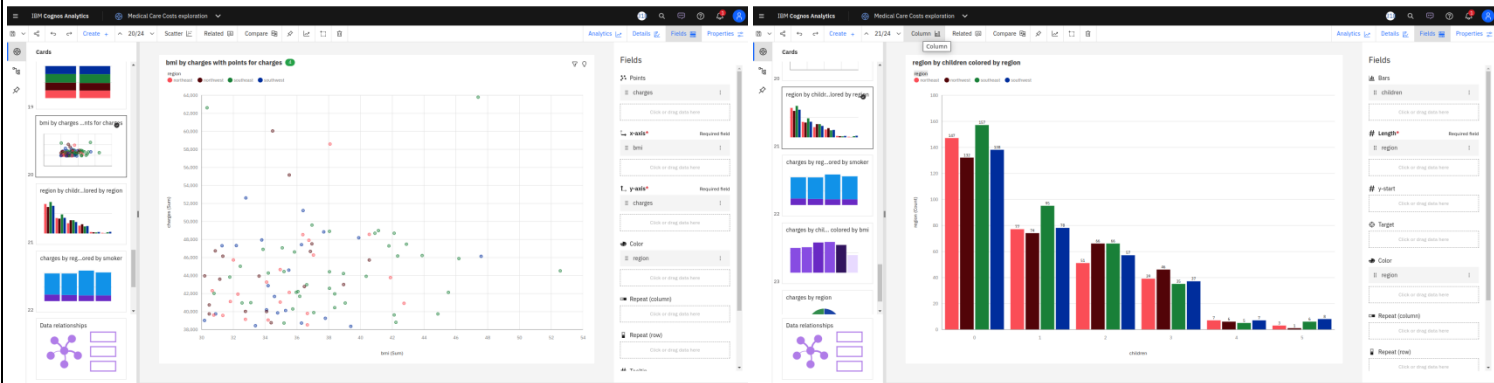
## DATA VISUALIZATIONS



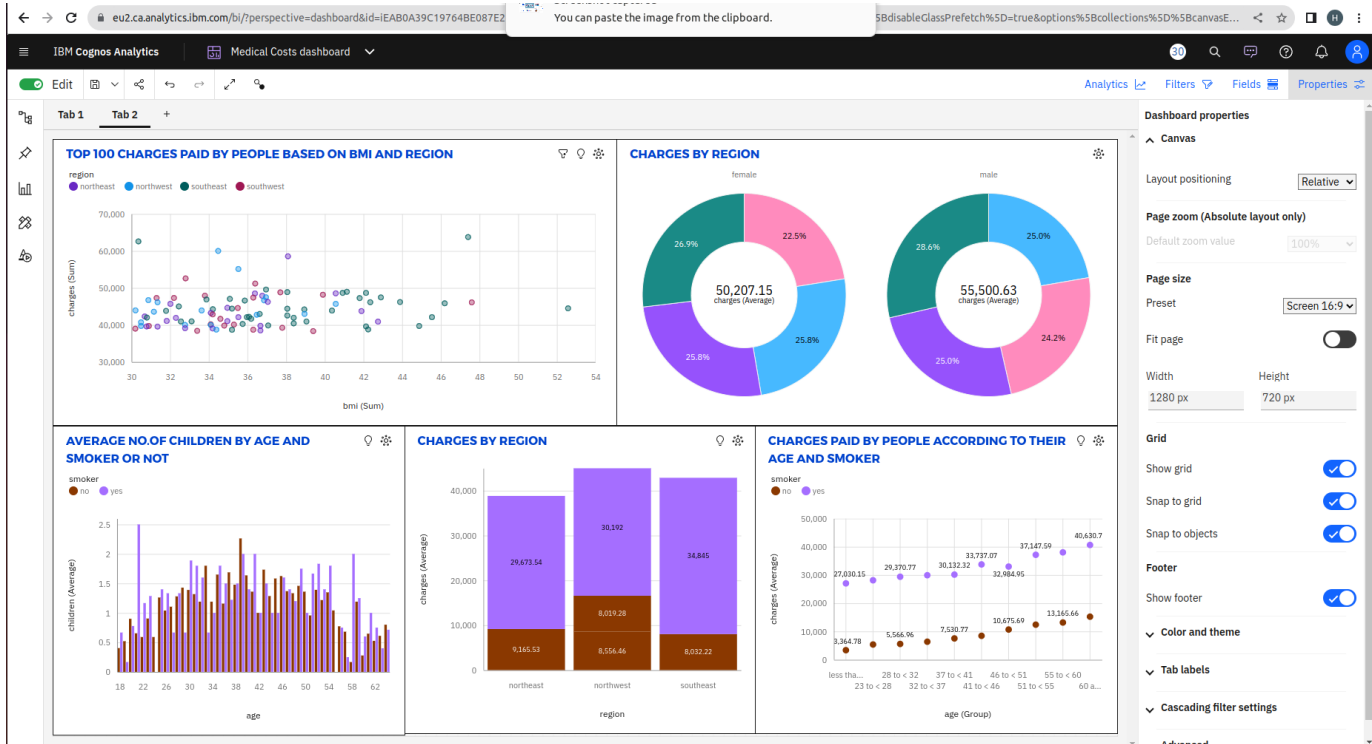
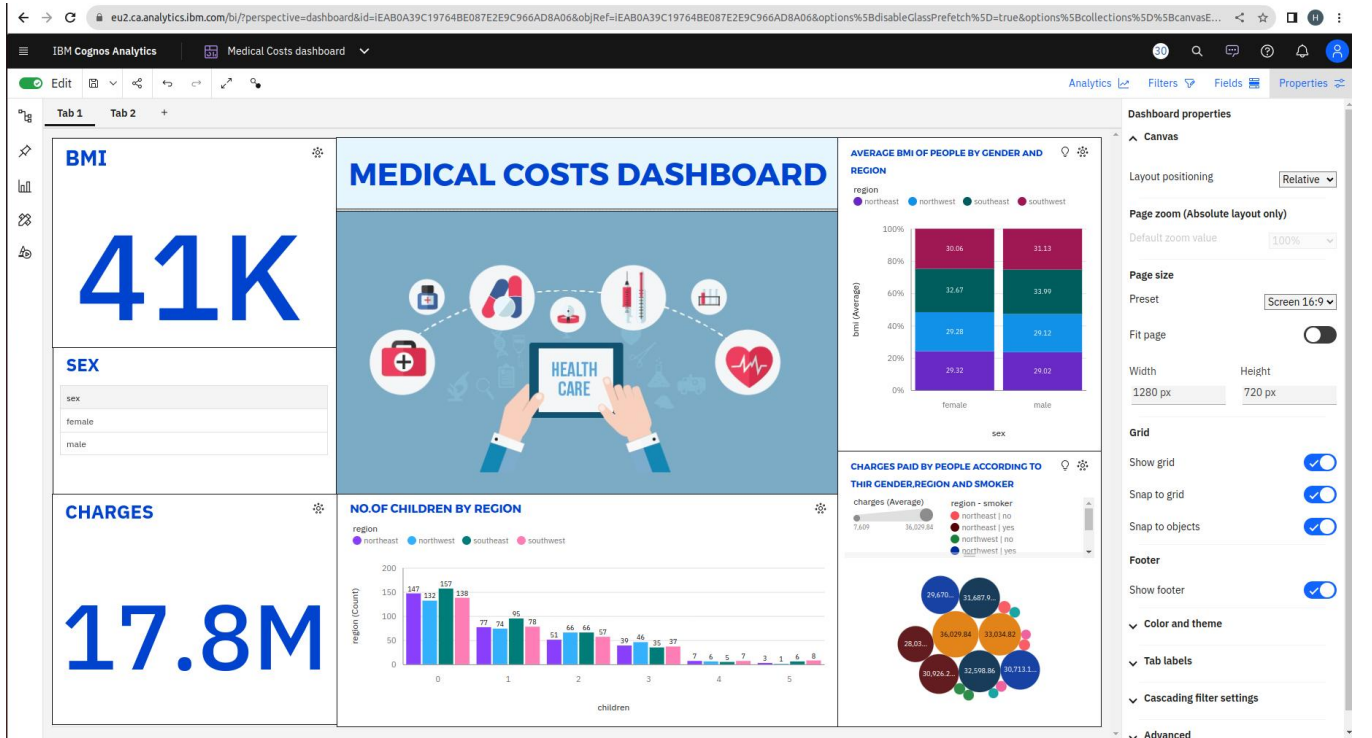




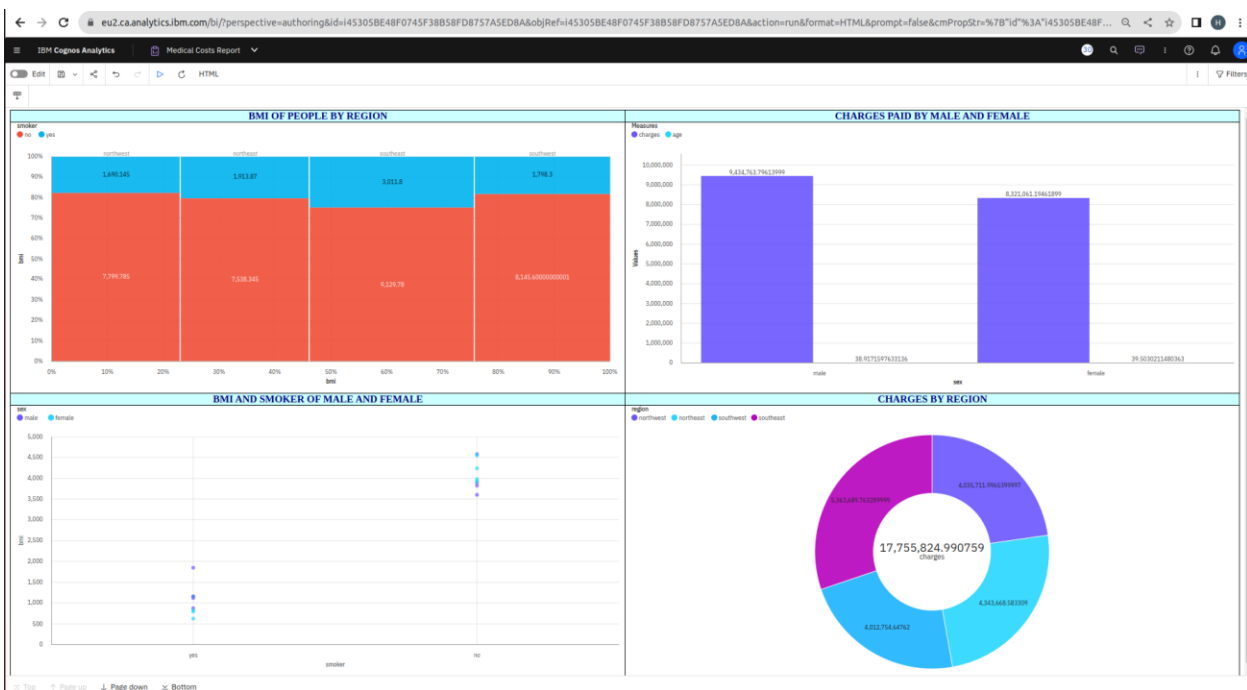
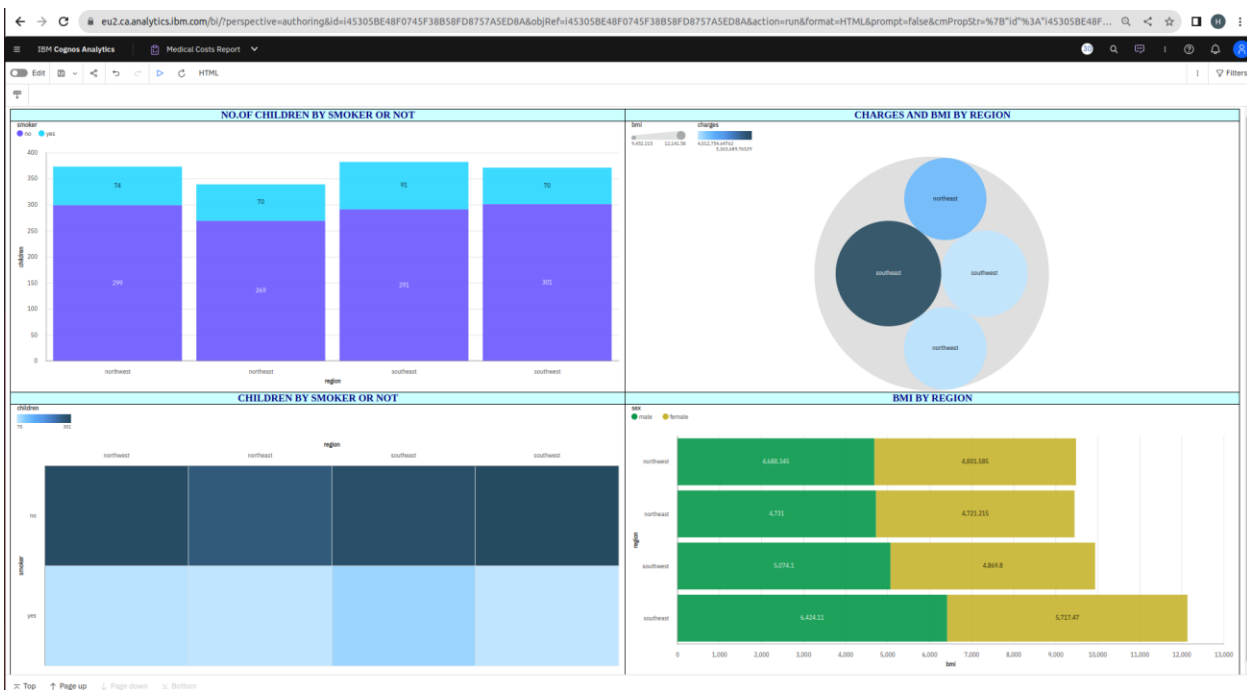




# DASHBOARD



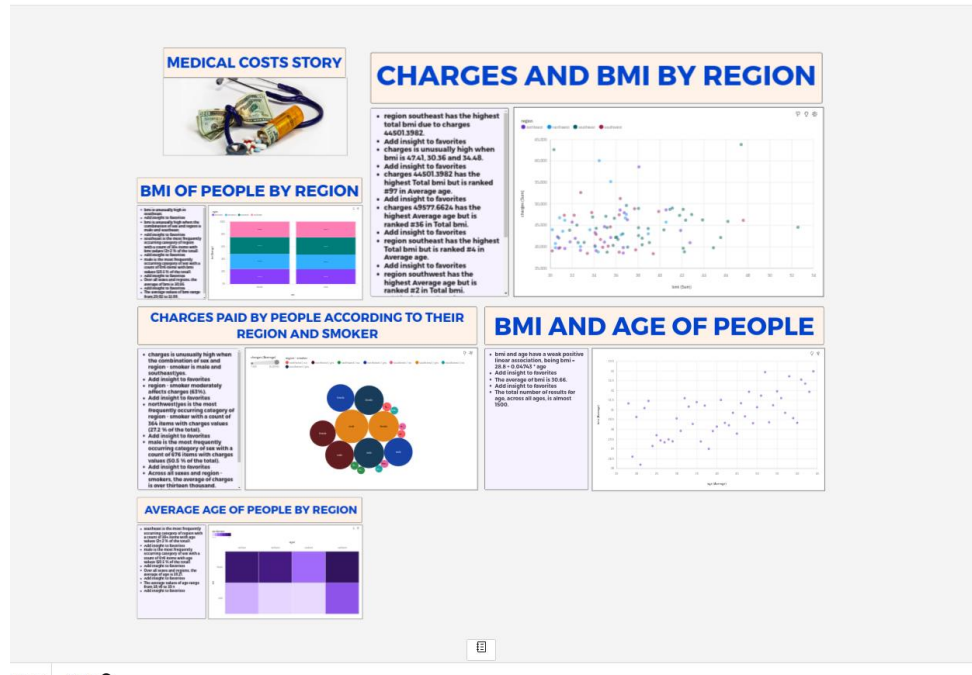
# REPORT



# STORY(GUIDED JOURNEY)

/perspective=story&id=1654CF37672E9457B943980E4833490AA&objRef=1654CF37672E9457B943980E4833490AA&options%5BdisableGlassPrefetch%5D=true&options%5Bcollections%5D%5Bca

Medical Costs Guided Journey

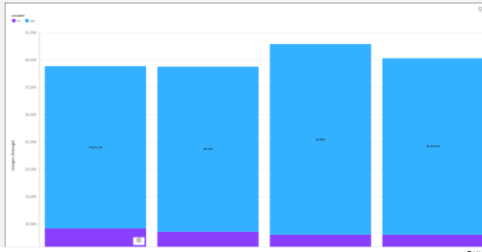


## SLIDESHOW STORY



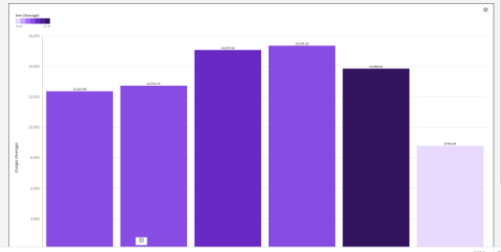
## CHARGES BY REGION AND SMOKER

- charges is unusually high when the combinations of region and smoker are southeast and yes, southwest and yes and northeast and yes.
- Add insight to favorites
- smoker moderately affects charges (62%).
- Add insight to favorites
- southeast is the most frequently occurring category of region with a count of 384 items with charges values (27.2 % of the total).
- Add insight to favorites
- no is the most frequently occurring category of smoker with a count of 1064 items with charges values (79.5 % of the total).
- Add insight to favorites
- Across all regions and smokers, the average of charges is over thirteen thousand.
- Add insight to favorites
- The average values of charges range from over eight thousand to almost 15 thousand.
- Add insight to favorites



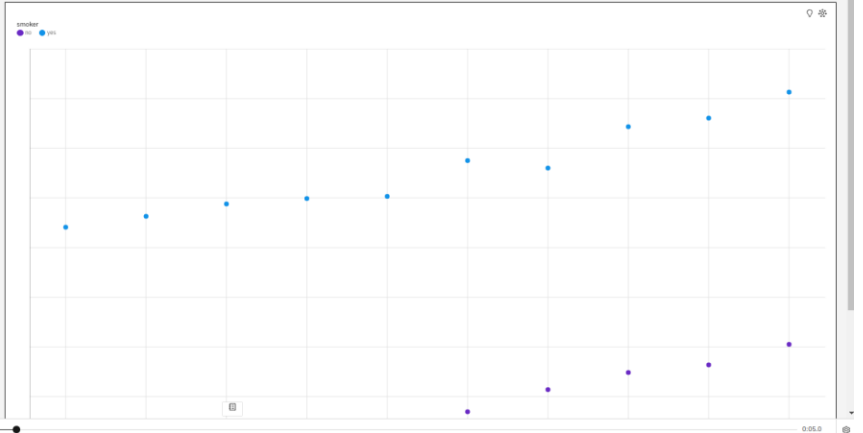
## NO.OF CHILDREN BY BMI

- 0 is the most frequently occurring category of children with a count of 874 items with charges values (42.9 % of the total).
- Add insight to favorites
- Across all children, the average of charges is over thirteen thousand.
- Add insight to favorites
- The overall number of results for bmi is almost 1500.
- Add insight to favorites
- The average values of charges range from nearly nine thousand, occurring when children is 5, to over fifteen thousand, when children is 3.



## CHARGES PAID BY PEOPLE ACCORDING TO AGE AND SMOKER

- charges is unusually high when age (Group) is 60 and above.
- Add insight to favorites
- charges is unusually high when the combination of age (Group) and smoker is 60 and above and yes.
- Add insight to favorites
- smoker moderately affects charges (62%).
- Add insight to favorites
- less than 23 is the most frequently occurring category of age (Group) with a count of 222 items with charges values (16.6 % of the total).
- Add insight to favorites
- no is the most frequently occurring category of smoker with a count of 1064 items with charges values (79.5 % of the total).
- Add insight to favorites
- Across all values of age (Group) and smoker, the average of charges is over thirteen thousand.
- Add insight to favorites
- The average values of charges range from almost 3500 to nearly 41 thousand.
- Add insight to favorites



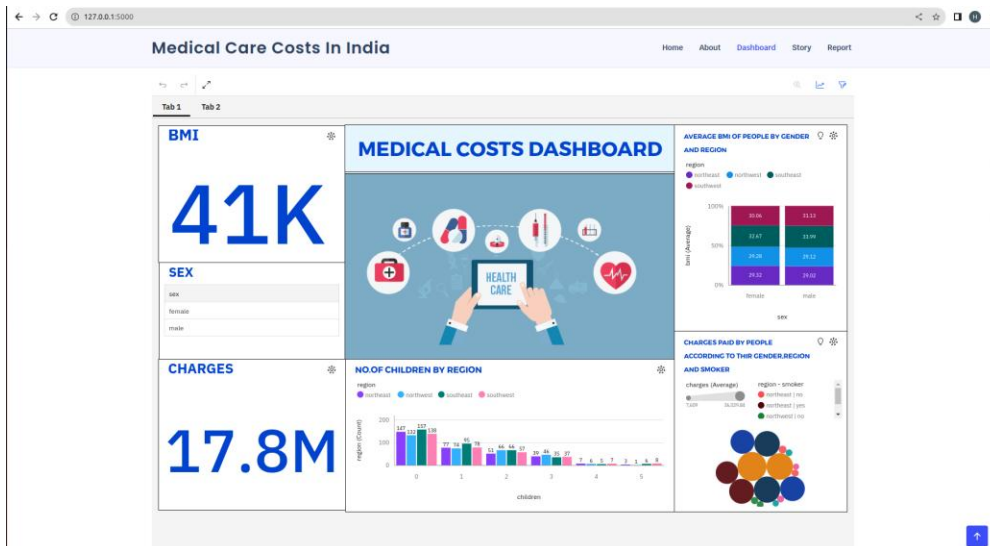
## AVERAGE AGE OF MALE AND FEMALE

- female exceeds male in age by 0.5859.
- Add insight to favorites
- sex male has the highest Total charges but is ranked #2 in Average age.
- Add insight to favorites
- sex female has the highest Average age but is ranked #2 in Total charges.
- Add insight to favorites
- male is the most frequently occurring category of sex with a count of 676 items with age values (50.5 % of the total).

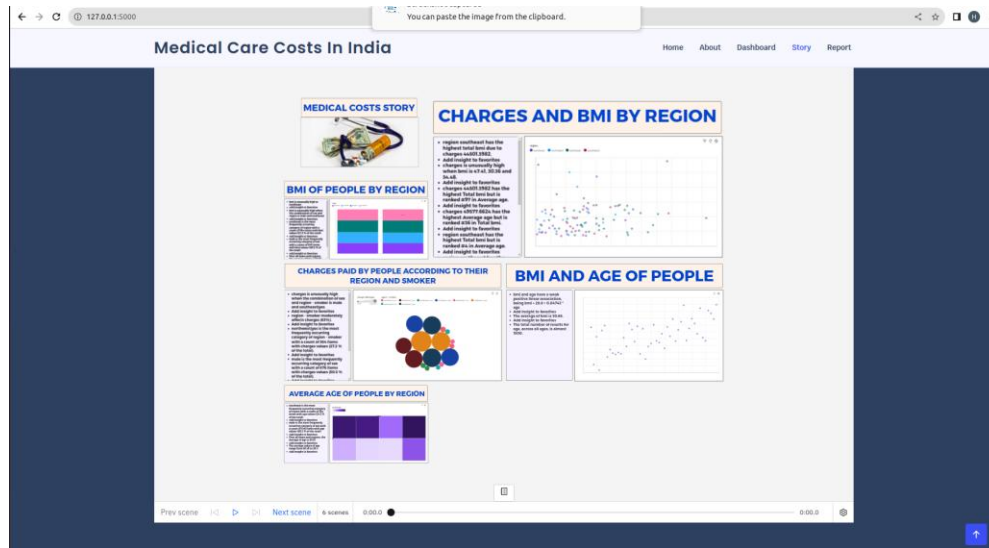
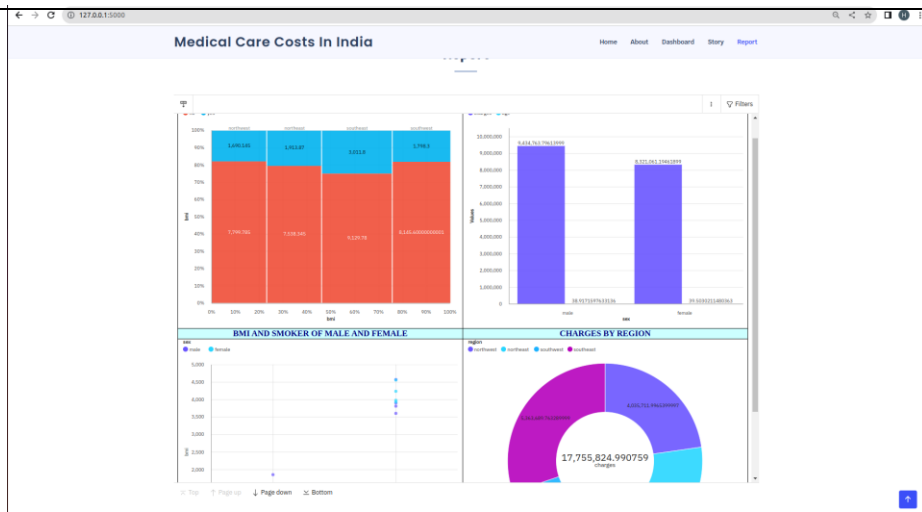




# WEB INTEGRATION







## 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.APPLICATIONS:**

The Estimation and Prediction of Hospitalization and Medical Care Costs project has several valuable applications in the healthcare industry and beyond.

- Healthcare Cost Management
- Financial Planning
- Insurance Pricing and Coverage
- Resource Allocation
- Treatment Decision Support
- Patient Cost Transparency
- Policy Development
- Fraud Detection
- Benchmarking and Performance
- Research and Public Health

- Cost-Effective Healthcare Programs
- 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 analytics, 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 Estimation and Prediction of Hospitalization and Medical Care Costs project is dynamic and transformative. As technology continues to evolve and data-driven decision-making becomes increasingly prevalent, the project's applications have the potential to revolutionize healthcare cost management, resource allocation, and patient care on a global scale.