PROJECT REPORT

Project Title:

ESTIMATION AND PREDICTION OF
HOSPITALIZATION AND MEDICAL CARE COSTS.

Team ID **: LTVIP2023TMID00120**

Team Size **: 3**

Team Leader **: GANJI DURGA BHAVANI**

Team Member **: DONKADA CHARISMA**

Team Member **: MANDAPAKA SREE DURGA**

Faculty Mentor **: K.NAVEENA**

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 aboutfuture 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 receivenecessary 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 ofthe 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 themedical 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 severalimportant 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, becausevarious 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, ourmethodology provides a predictive method.

2.1 Existing problem:

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

2.2 PROPOSED SOUTION

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.

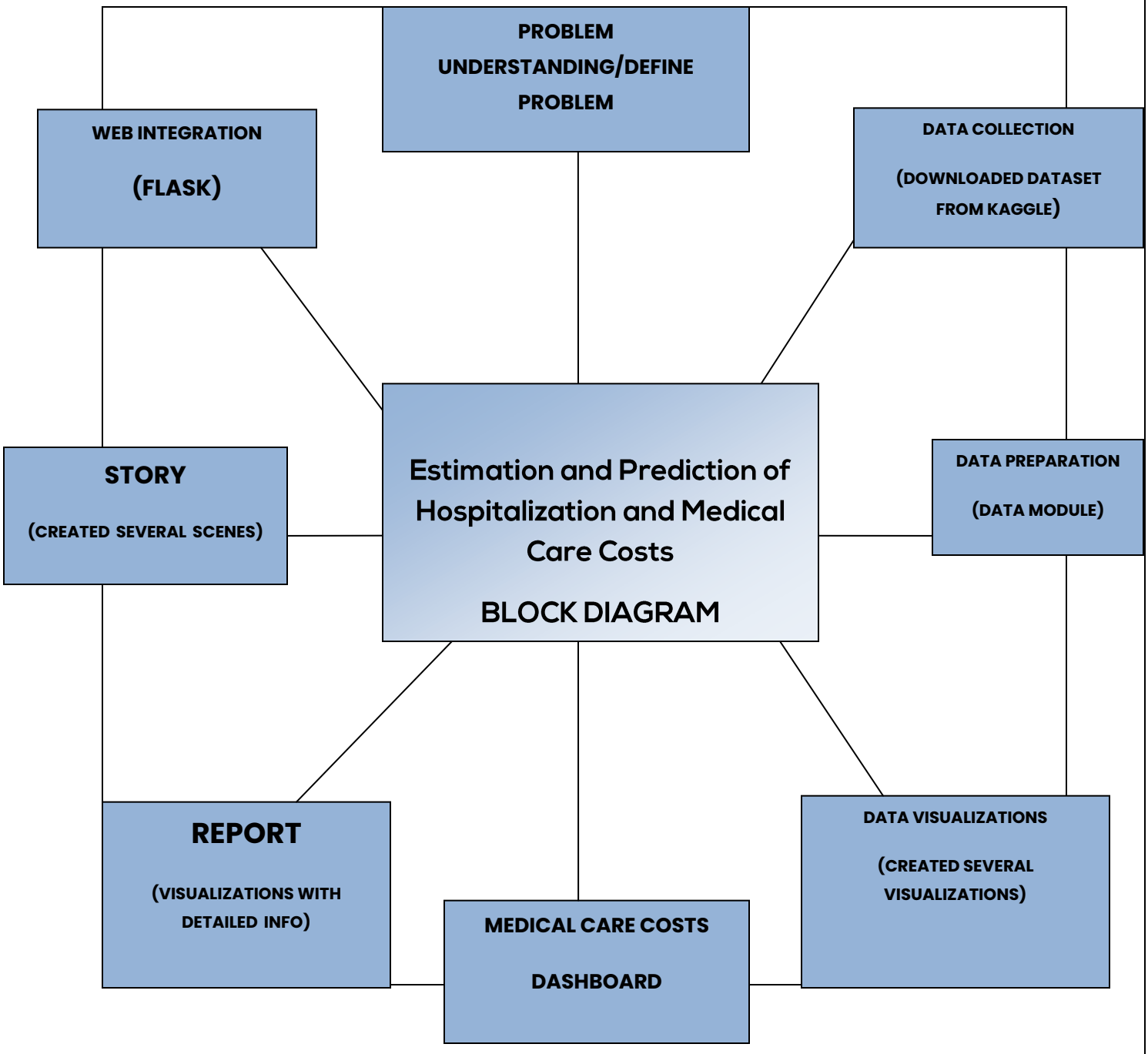
3.THEORITICAL ANALYSIS

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:

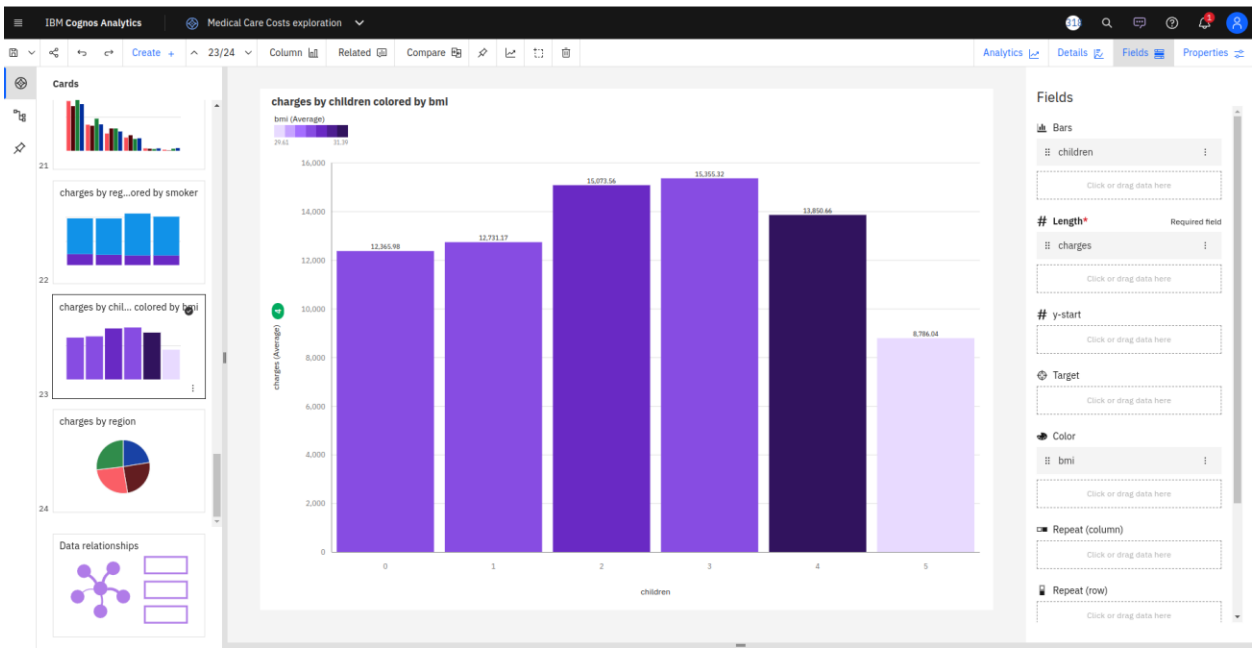
- IBM Cognos Analytics
- Flask
- Integrated Development Environment(IDE)-Spyder

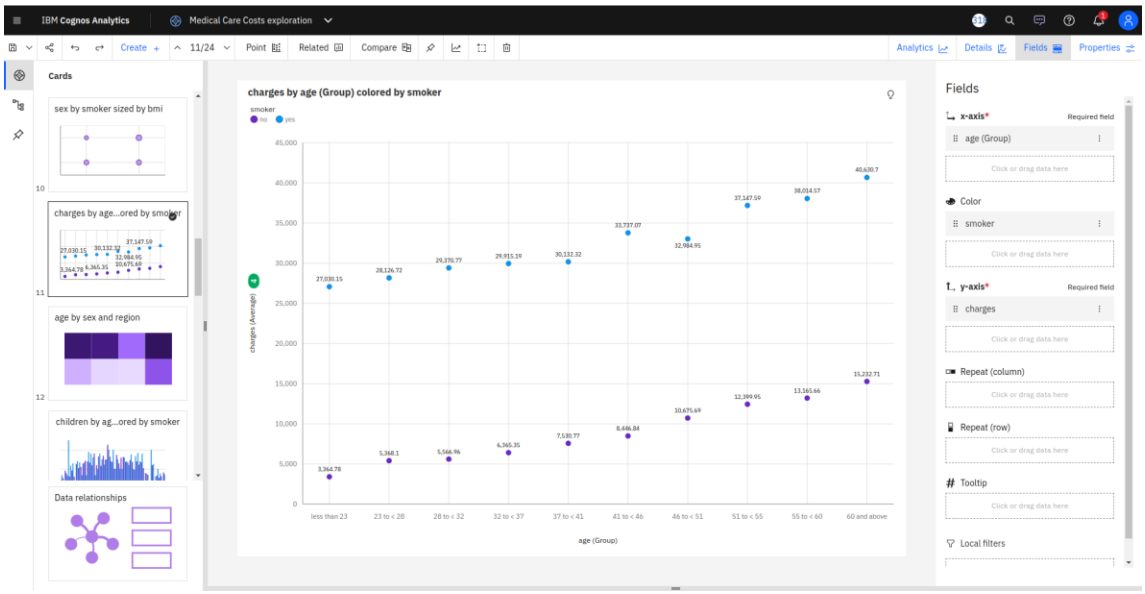
HARDWARE REQUIREMENTS:

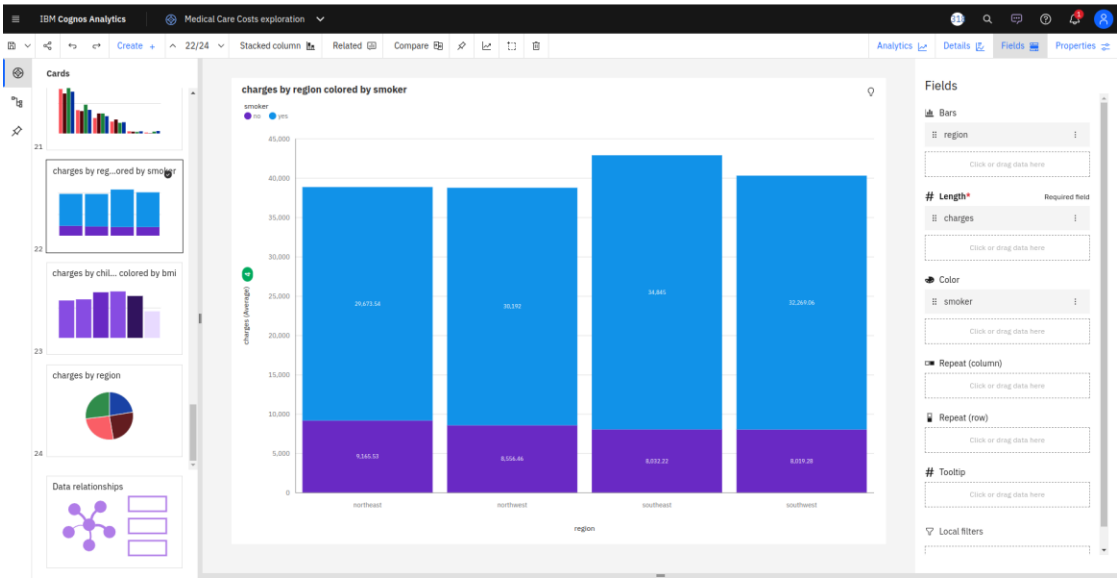
- Minimum System Requirements(RAM-4GB,Quad Core Processor Or Above)

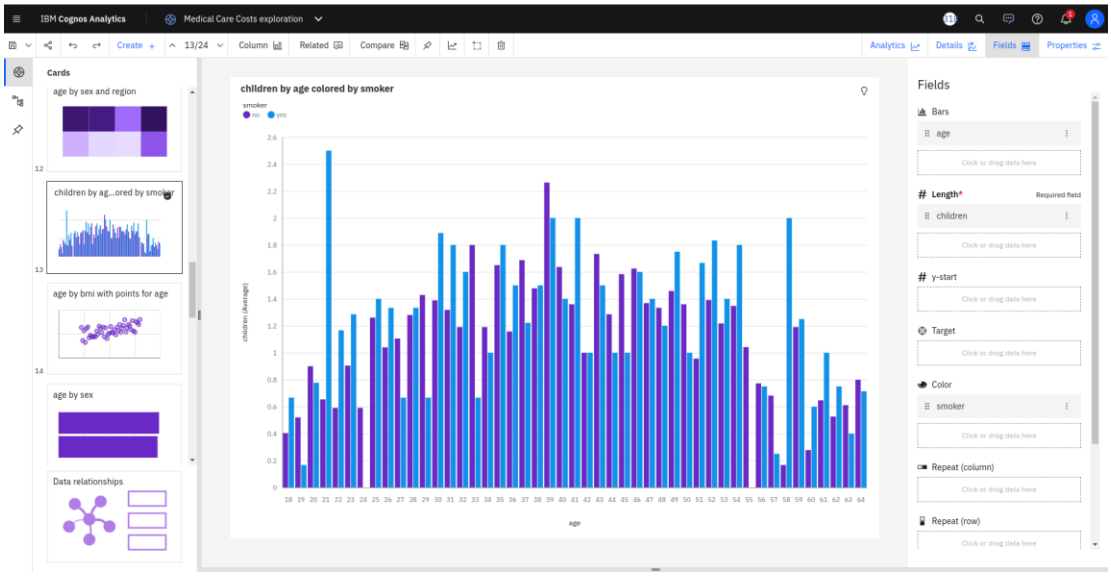
4.RESULT:

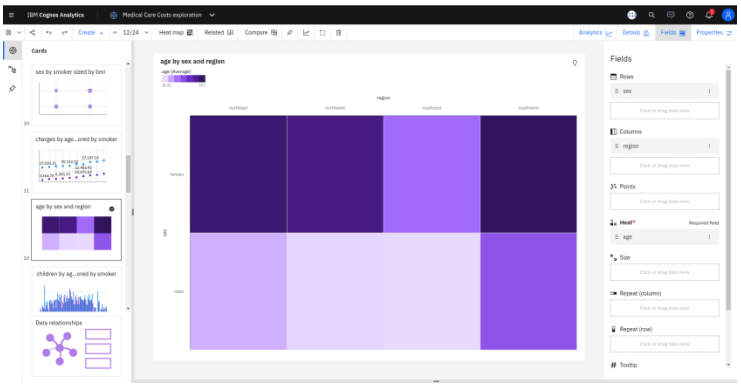
DATA VISUALIZATIONS

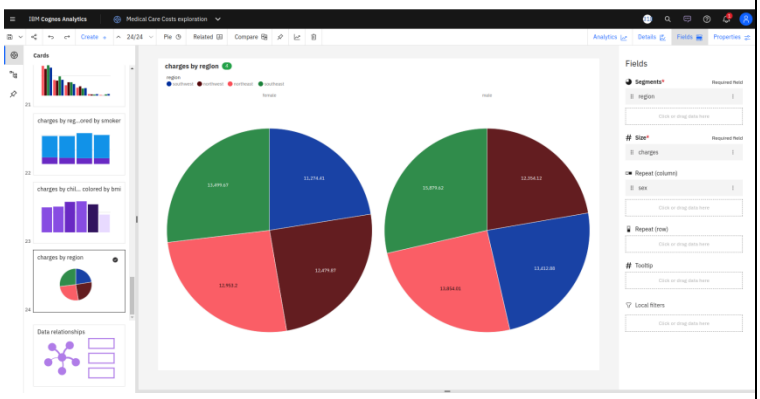


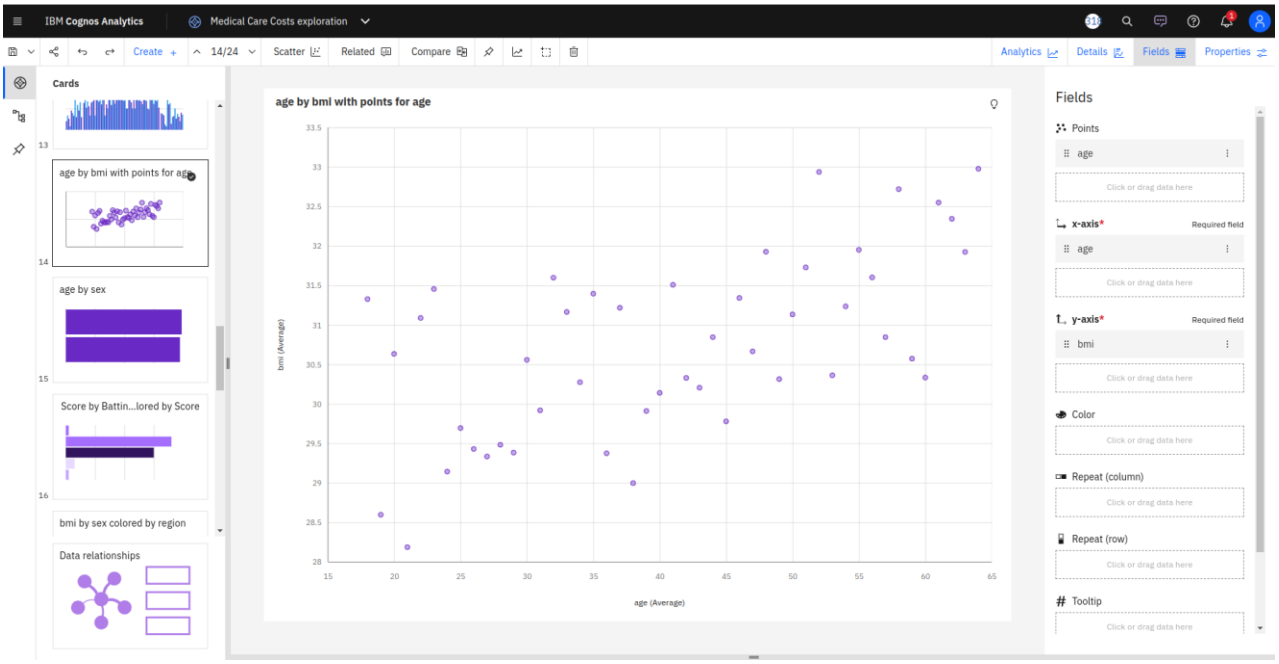

IBM Cognos Analytics
Medical Care Costs exploration
Create
11/24
Point
Related
Compare
Analytics
Details
Fields
Properties
Cards
sex by smoker sized by bmi
charges by age...ored by smoker
age by sex and region
children by ag...ored by smoker
Data relationships
charges by age (Group) colored by smoker
smoker
no
yes
charges (Average)
age (Group)
less than 23
23 to < 28
28 to < 32
32 to < 37
37 to < 41
41 to < 46
46 to < 51
51 to < 55
55 to < 60
60 and above
Fields
x-axis*
Required field
age (Group)
Color
smoker
y-axis*
charges
Repeat (column)
Repeat (row)
Tooltip
Local filters

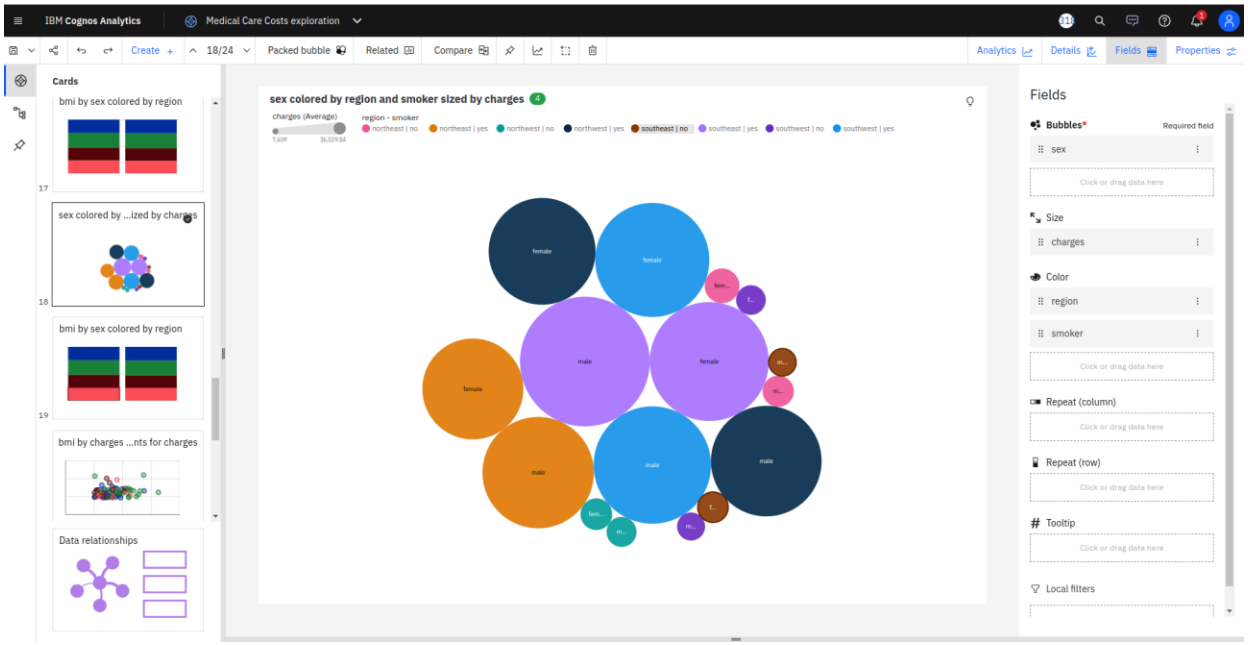

IBM Cognos Analytics
Medical Care Costs exploration
Create
22/24
Stacked column
Related
Compare
Analytics
Details
Fields
Properties
Cards
charges by reg...ored by smoker
charges by chil... colored by bmi
charges by region
Data relationships
charges by region colored by smoker
smoker
charges (Average)
region
northeast
northwest
southeast
southwest
Fields
Bars
region
Length*
Required field
charges
Color
smoker
Repeat (column)
Repeat (row)
Tooltip
Local filters


IBM Cognos Analytics
Medical Care Costs exploration
Create
13/24
Column
Related
Compare
Analytics
Details
Fields
Properties
Cards
age by sex and region
children by ag...ored by smoker
age by bmi with points for age
age by sex
Data relationships
children by age colored by smoker
smoker
children (Average)
age
Fields
Bars
age
Length*
Required field
children
y-start
Target
Color
smoker
Repeat (column)
Repeat (row)

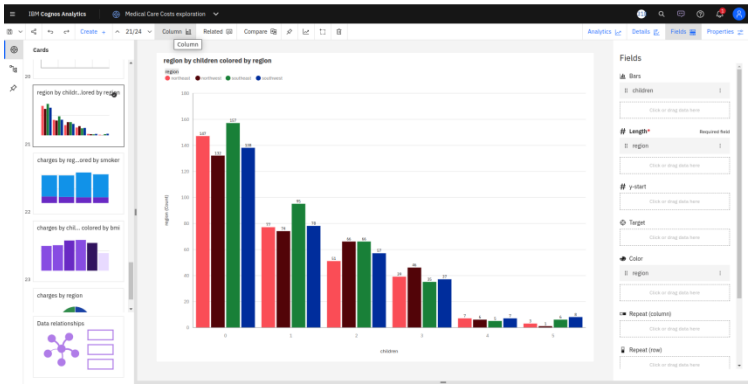


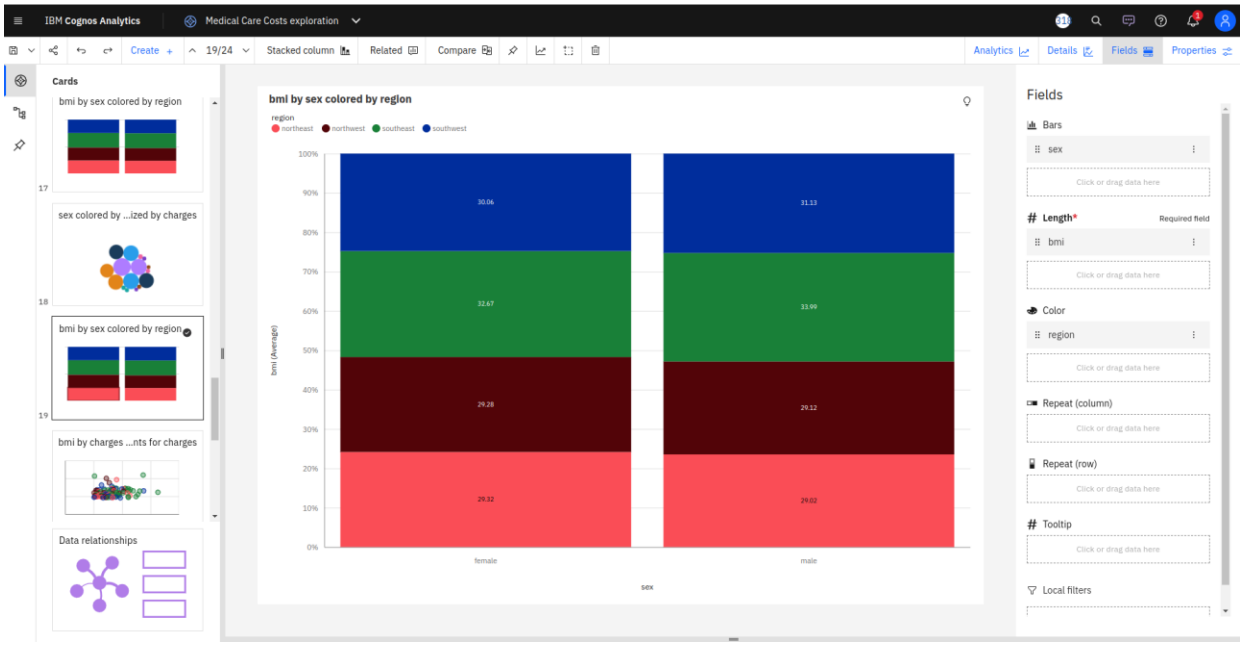


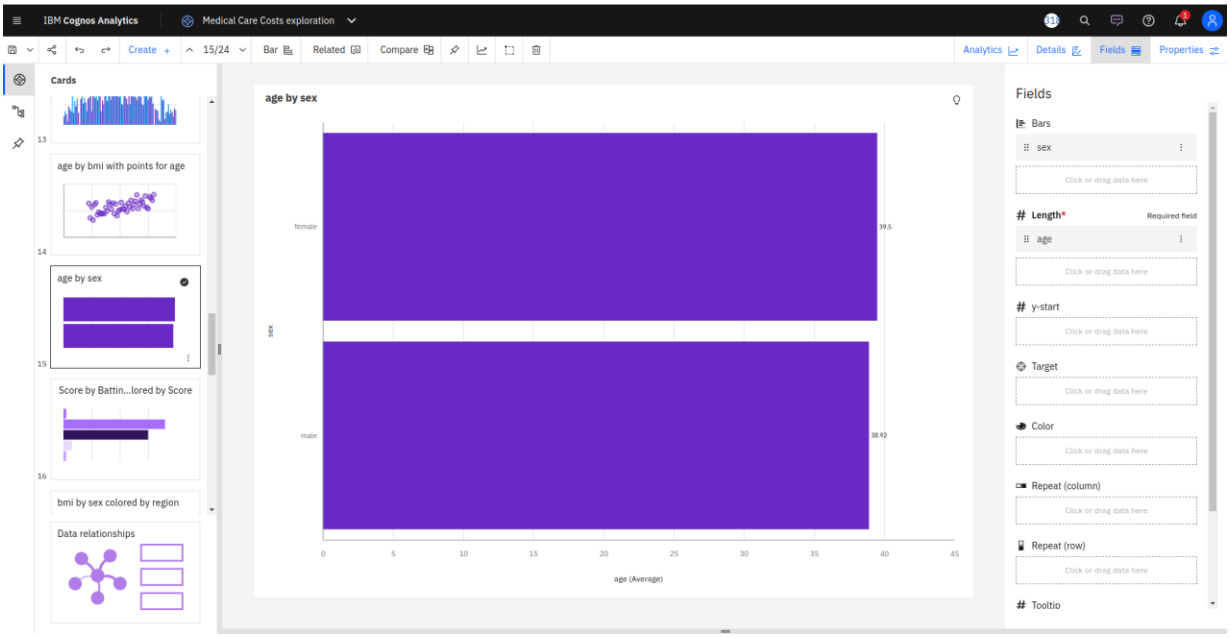

IBM Cognos Analytics
Medical Care Costs exploration
Create
14/24
Scatter
Related
Compare
Analytics
Details
Fields
Properties
Cards
age by bmi with points for age
bmi by sex colored by region
Data relationships
age by bmi with points for age
bmi (Average)
age (Average)
Fields
Points
age
x-axis*
Required field
y-axis*
bmi
Color
Repeat (column)
Repeat (row)
Tooltip
Click or drag data here


IBM Cognos Analytics
Medical Care Costs exploration
Create
18/24
Packed bubble
Related
Compare
Analytics
Details
Fields
Properties
Cards
bmi by sex colored by region
sex colored by region and smoker sized by charges
Fields
Bubbles*
Required field
sex
Size
charges
Color
region
smoker
Repeat (column)
Repeat (row)
Tooltip
Local filters
Click or drag data here

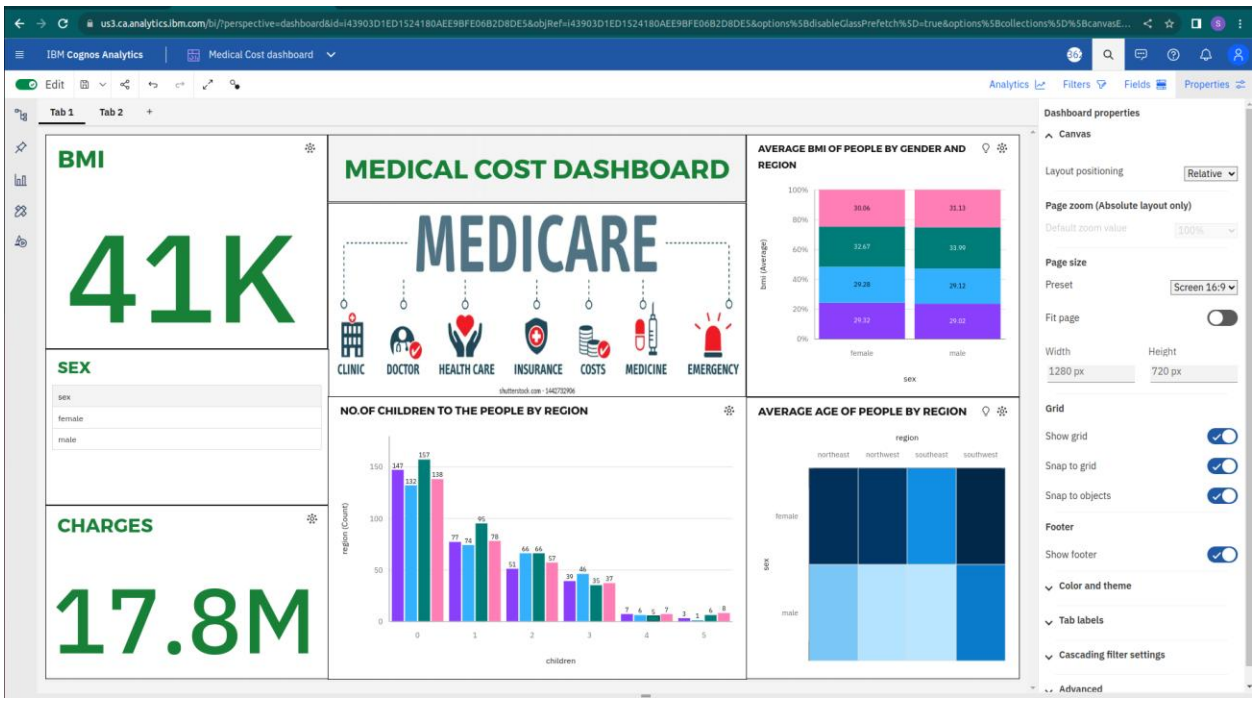
IBM Cognos Analytics
Medical Care Costs exploration
bmi by charges with points for charges
Fields

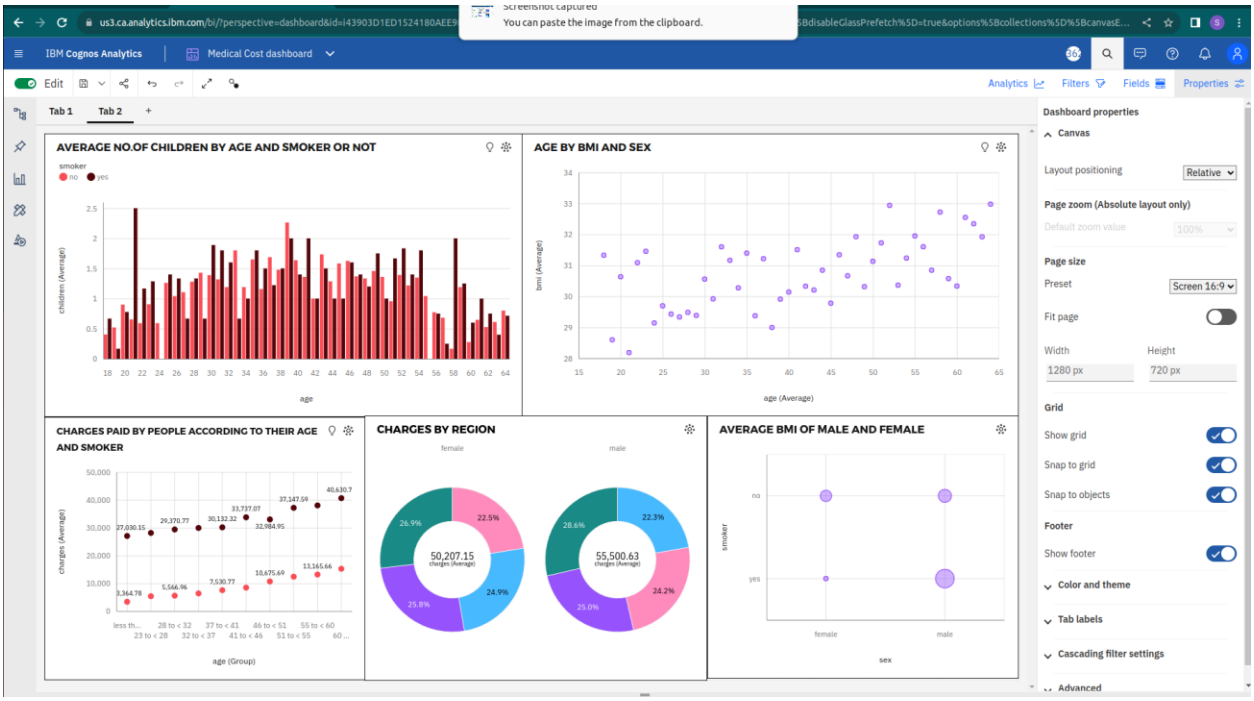

IBM Cognos Analytics
Medical Care Costs exploration
region by children colored by region
Fields


IBM Cognos Analytics
Medical Care Costs exploration
Create
19/24
Stacked column
Related
Compare
Analytics
Details
Fields
Properties
Cards
bmi by sex colored by region
sex colored by ...ized by charges
bmi by sex colored by region
bmi by charges ...nts for charges
Data relationships
bmi by sex colored by region
region
northeast
northwest
southeast
southwest
30.06
31.13
32.67
33.99
29.28
29.12
29.32
29.02
female
male
sex
bmi (Average)
Fields
Bars
sex
Length*
Required field
bmi
Color
region
Repeat (column)
Repeat (row)
Tooltip
Local filters

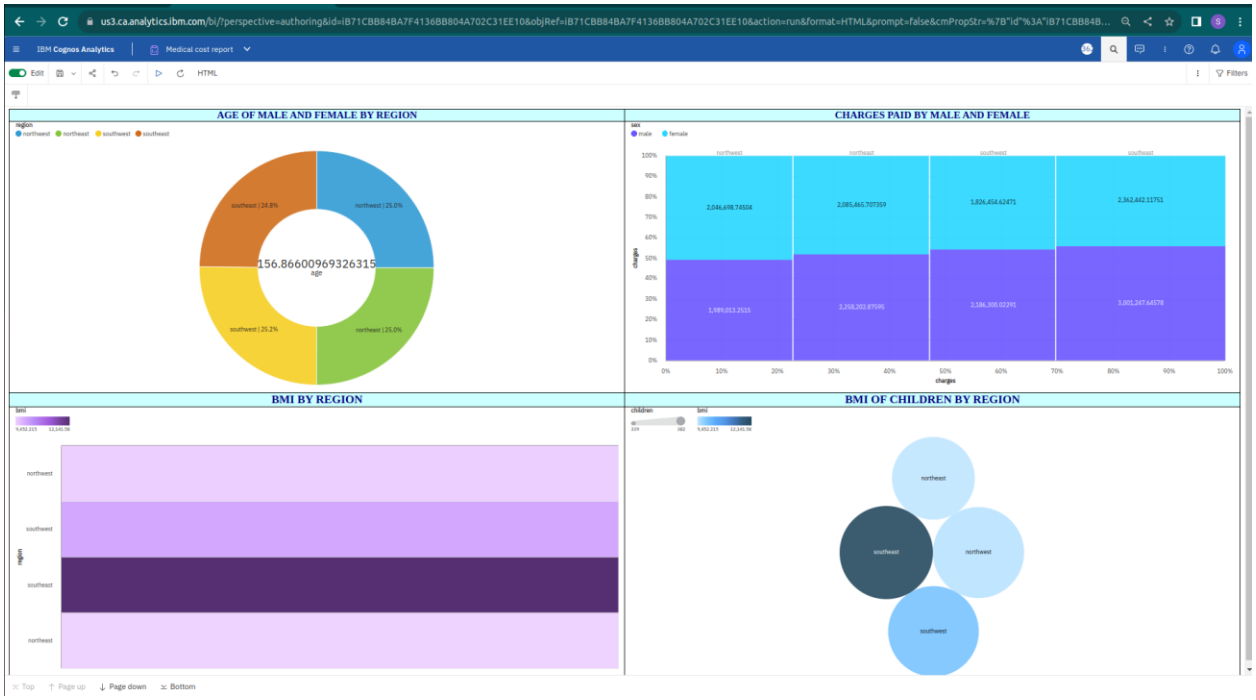

IBM Cognos Analytics
Medical Care Costs exploration
Create
15/24
Bar
Related
Compare
Analytics
Details
Fields
Properties
Cards
age by bmi with points for age
age by sex
Score by Battin...lored by Score
bmi by sex colored by region
Data relationships
age by sex
female
male
39.5
38.92
sex
age (Average)
Fields
Bars
sex
Length*
Required field
age
y-start
Target
Color
Repeat (column)
Repeat (row)
Tooltip

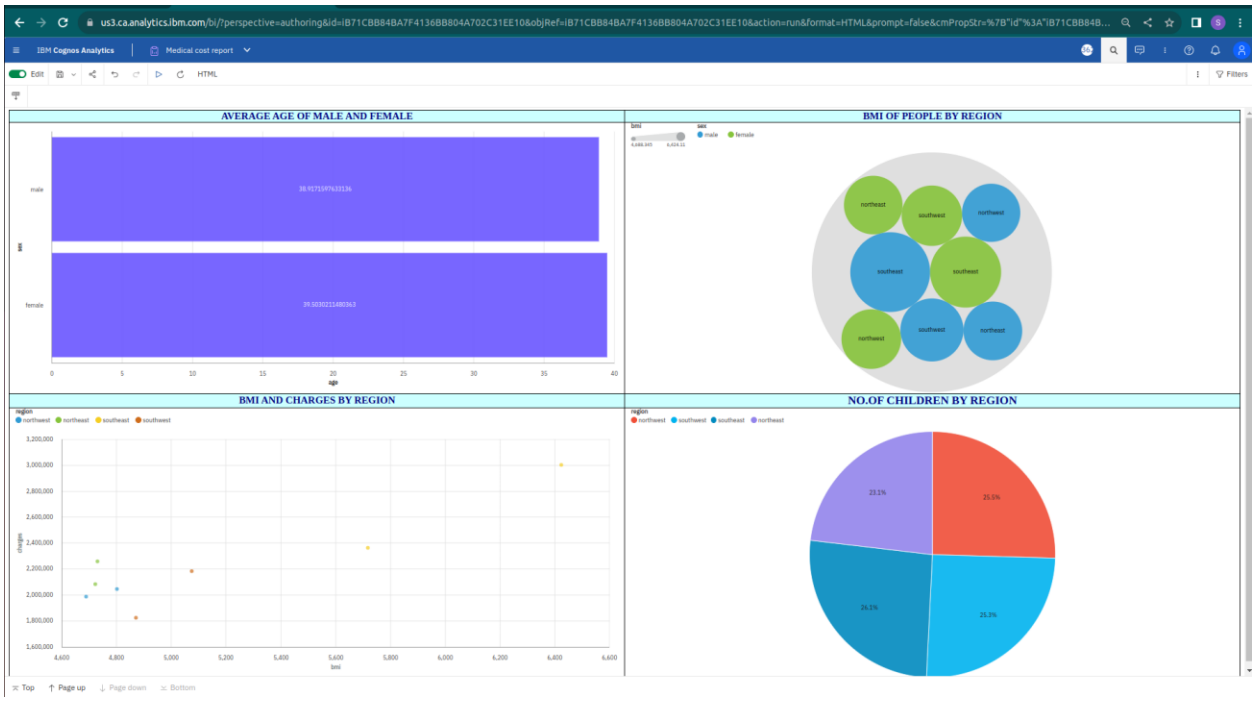
DASHBOARD


IBM Cognos Analytics
Medical Cost dashboard
Edit
Analytics
Filters
Fields
Properties
Tab 1
Tab 2
BMI
41K
MEDICAL COST DASHBOARD
MEDICARE
CLINIC
DOCTOR
HEALTH CARE
INSURANCE
COSTS
MEDICINE
EMERGENCY
AVERAGE BMI OF PEOPLE BY GENDER AND REGION
SEX
sex
female
male
CHARGES
17.8M
NO.OF CHILDREN TO THE PEOPLE BY REGION
region (Count)
children
AVERAGE AGE OF PEOPLE BY REGION
region
northeast
northwest
southeast
southwest
Dashboard properties
Canvas
Layout positioning
Relative
Page zoom (Absolute layout only)
Page size
Preset
Screen 16:9
Fit page
Width
Height
1280 px
720 px
Grid
Show grid
Snap to grid
Snap to objects
Footer
Show footer
Color and theme
Tab labels
Cascading filter settings
Advanced

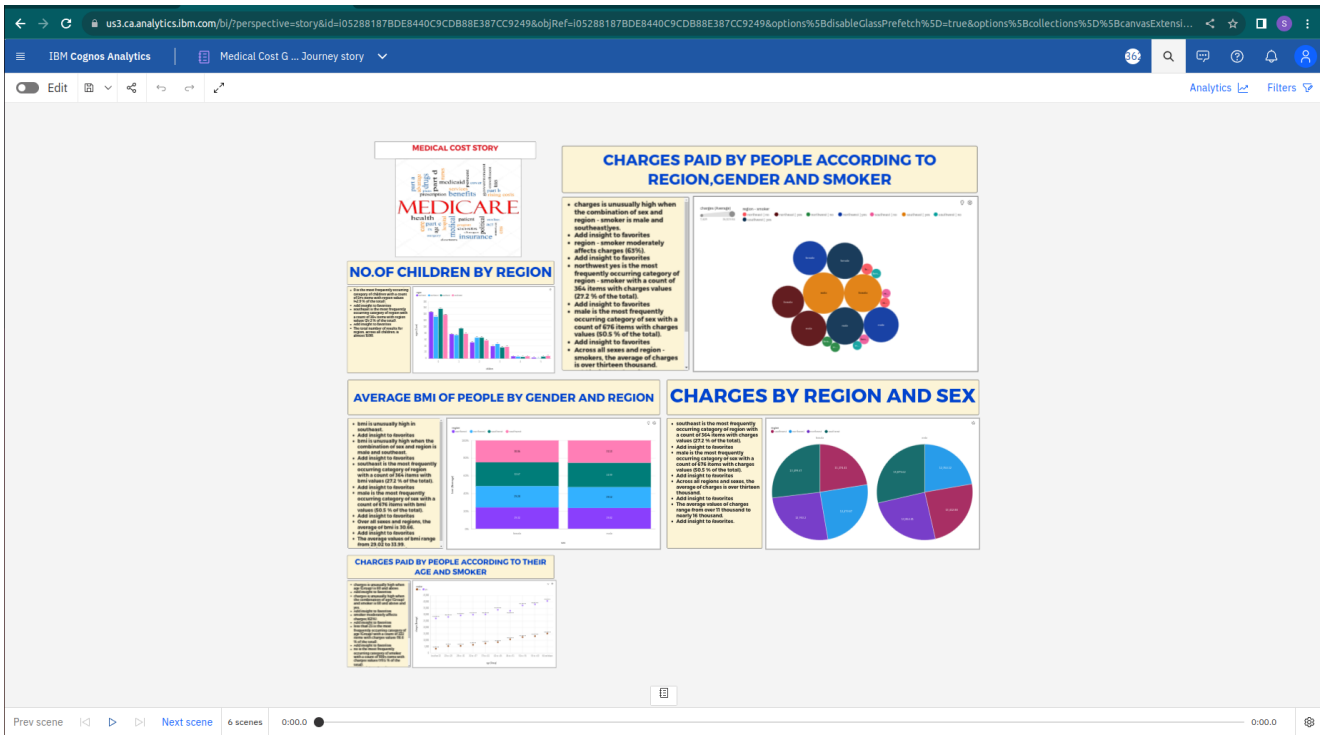

You can paste the image from the clipboard.
IBM Cognos Analytics
Medical Cost dashboard
Edit
Analytics
Filters
Fields
Properties
Tab 1
Tab 2
AVERAGE NO.OF CHILDREN BY AGE AND SMOKER OR NOT
smoker
no
yes
children (Average)
age
AGE BY BMI AND SEX
bmi (Average)
age (Average)
CHARGES PAID BY PEOPLE ACCORDING TO THEIR AGE AND SMOKER
charges (Average)
age (Group)
CHARGES BY REGION
female
male
50,207.15
charges (Average)
55,500.63
charges (Average)
AVERAGE BMI OF MALE AND FEMALE
smoker
sex
Dashboard properties
Canvas
Layout positioning
Relative
Page zoom (Absolute layout only)
Page size
Preset
Screen 16:9
Fit page
Width
Height
1280 px
720 px
Grid
Show grid
Snap to grid
Snap to objects
Footer
Show footer
Color and theme
Tab labels
Cascading filter settings
Advanced

REPORT


AGE OF MALE AND FEMALE BY REGION
CHARGES PAID BY MALE AND FEMALE
BMI BY REGION
BMI OF CHILDREN BY REGION

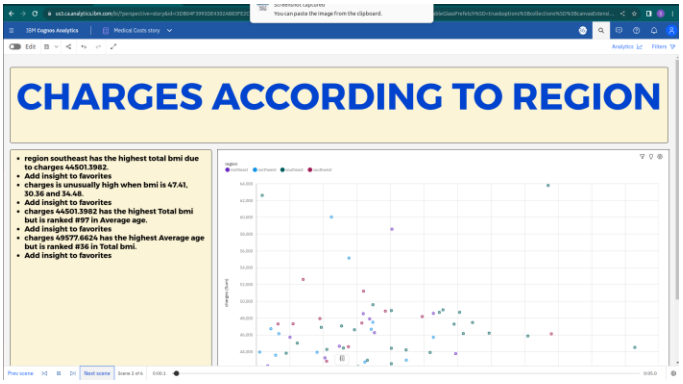

AVERAGE AGE OF MALE AND FEMALE
BMI OF PEOPLE BY REGION
BMI AND CHARGES BY REGION
NO.OF CHILDREN BY REGION

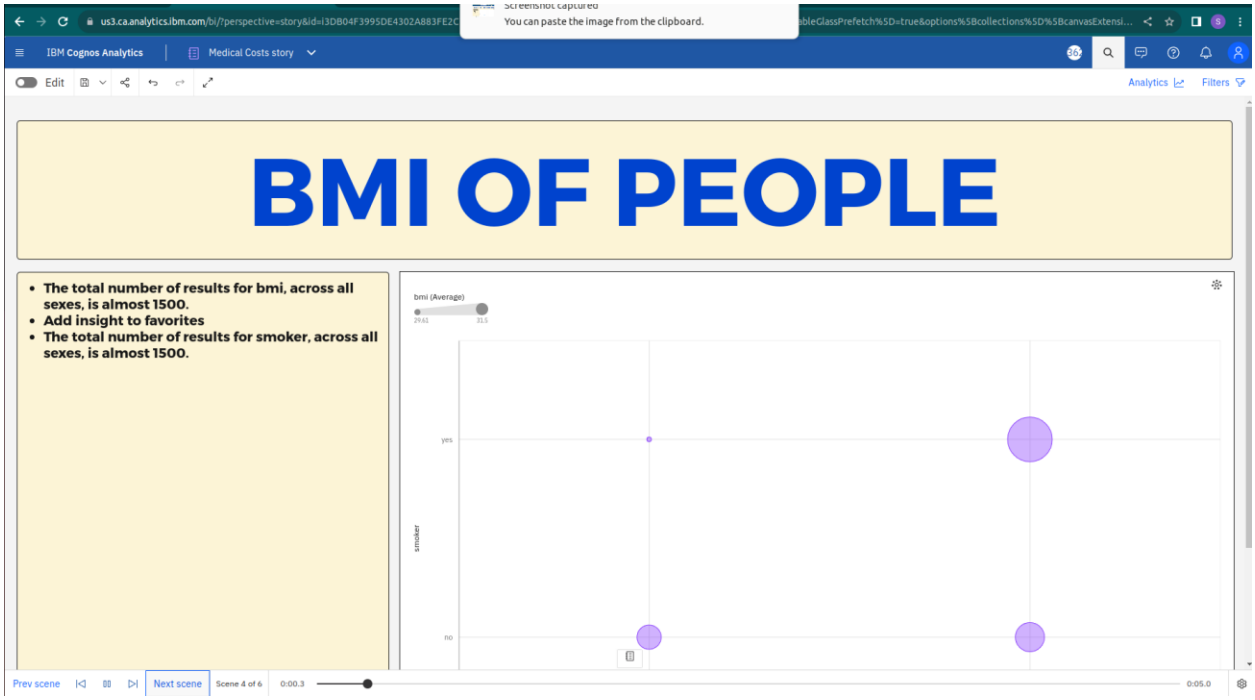
STORY(GUIDED JOURNEY)



SLIDESHOW STORY

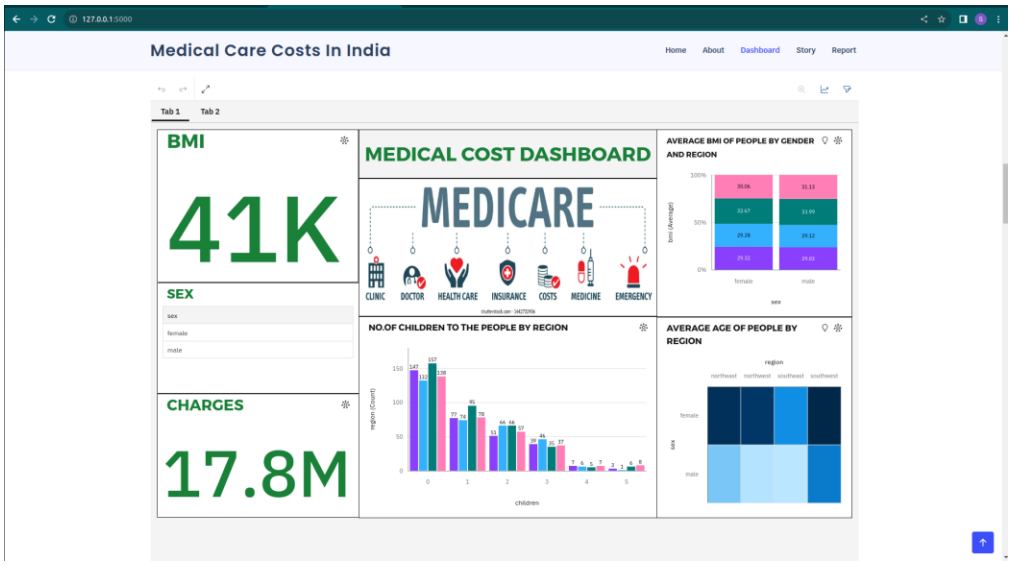
CHARGES PAID BY CHILDREN
• 0 is the most frequently occurring category of children with a count of 574 items with charges values (42.9 % of the total).
• Add insight to favorites
• Across all children, the average of charges is over thirteen thousand.
• Add insight to favorites
• The average values of charges range from nearly nine thousand, occurring when children is 5, to over fifteen thousand, when children is 3.

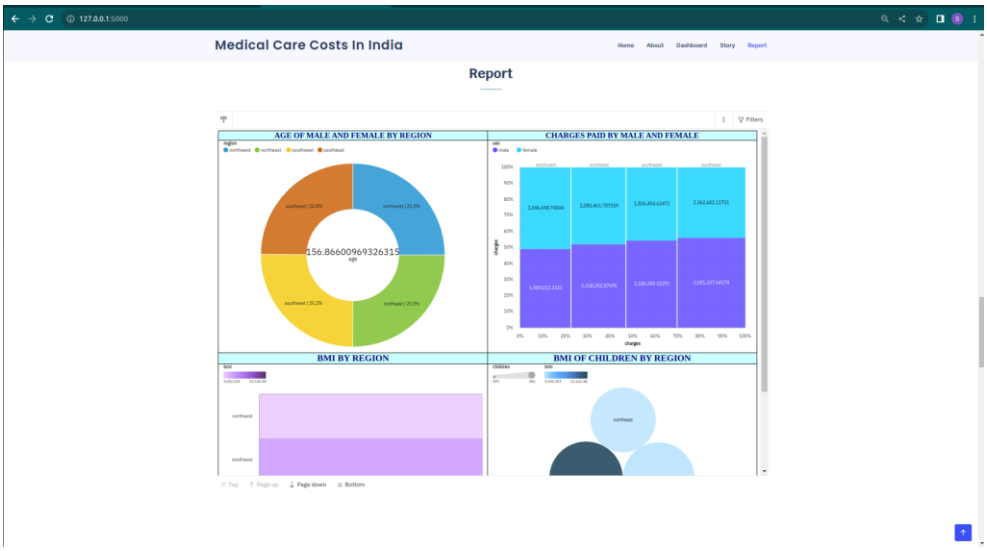

CHARGES ACCORDING TO REGION
• region southeast has the highest total bmi due to charges 44501.3982.
• Add insight to favorites
• charges is unusually high when bmi is 47.41, 30.36 and 34.48.
• Add insight to favorites
• charges 44501.3982 has the highest Total bmi but is ranked #97 in Average age.
• Add insight to favorites
• charges 49577.6624 has the highest Average age but is ranked #36 in Total bmi.
• Add insight to favorites

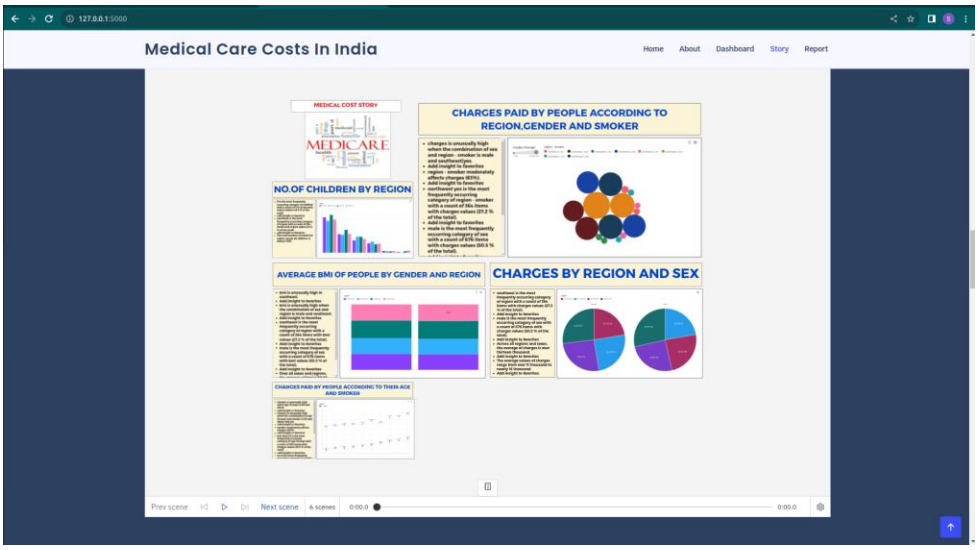

You can paste the image from the clipboard.
IBM Cognos Analytics
Medical Costs story
Edit
Analytics
Filters
BMI OF PEOPLE
• The total number of results for bmi, across all sexes, is almost 1500.
• Add insight to favorites
• The total number of results for smoker, across all sexes, is almost 1500.
bmi (Average)
29.61
31.5
yes
no
smoker
Prev scene
Next scene
Scene 4 of 6
0:00.3
0:05.0

You can paste the image from the clipboard.
IBM Cognos Analytics
Medical Costs story
Edit
Analytics
Filters
CHARGES PAID BY PEOPLE ACCORDING TO REGION AND SMOKER
• charges is unusually high when the combinations of region and smoker are southeast and yes, southwest and yes and northwest 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 364 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 35 thousand.
• Add insight to favorites
smoker
no
yes
45,000
40,000
35,000
30,000
25,000
20,000
15,000
10,000
charges (Average)
29,673.54
30,192
34,845
32,269.06
Prev scene
Next scene
Scene 3 of 6
0:00.4
0:05.0

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.