DATA ANALYSIS CAPSTONE PROJECT

Suhashini D/O K.Surendhran

AGENDA

Business problem

Data source and cleaning

Data analysis - Tableau

Analytic findings and recommendation

BUSINESS PROBLEM

Utilizing existing data, we aim to predict a proper annual premium for customers. While providing affordable insurance, it is important to predict the premiums accurately to avoid company loss.

DATA SOURCE

Data set:
https://www.kaggle.com/datasets/harshsingh2209/medical-insurance-payout

Data Cleaning
Dataset was checked for “NULL” fields and removed

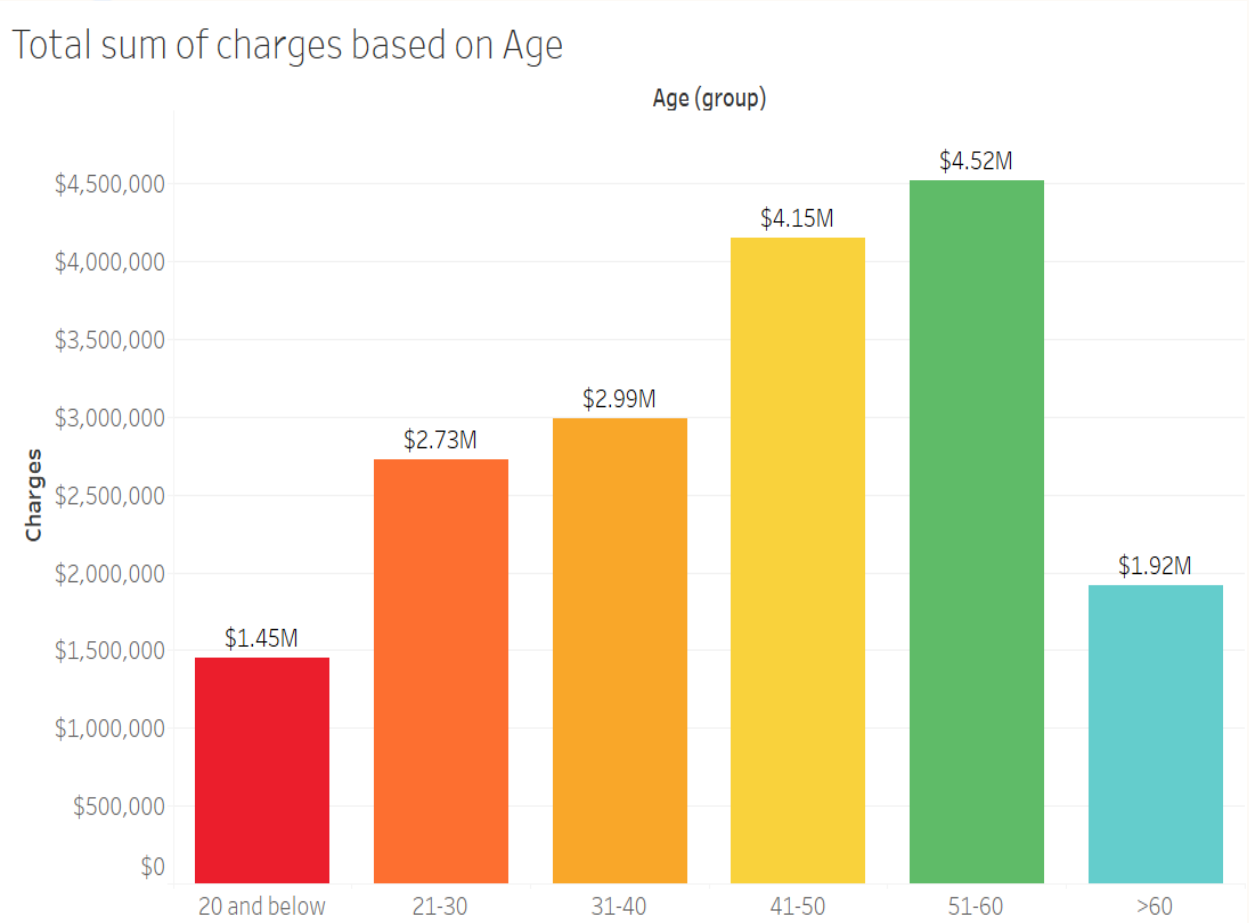
Data Preparation
Grouping were created to enable easier categorizing

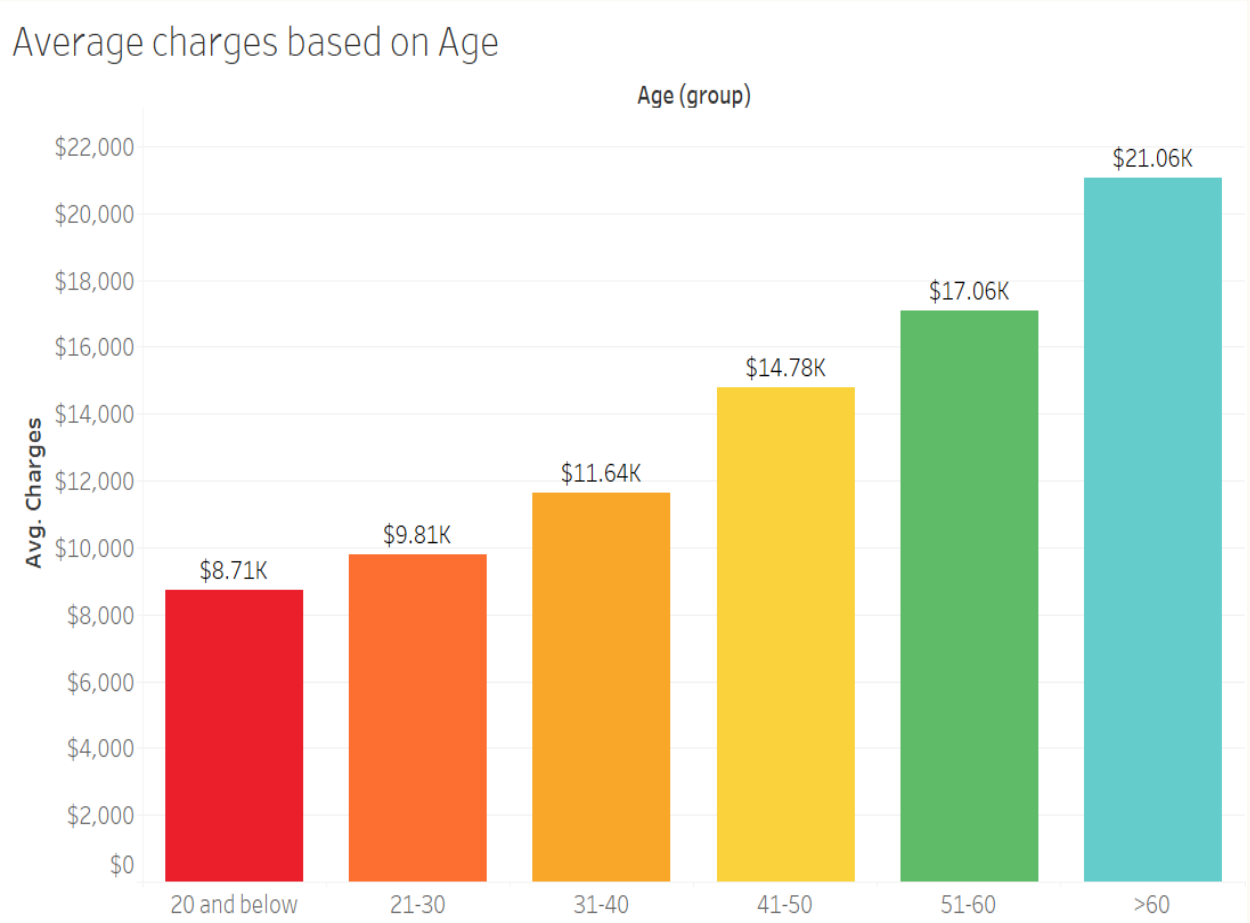
Tableau Link
https://public.tableau.com/app/profile/suhashini1994/viz/Capstoneproject_16622976659920/AVG

COLUMNS

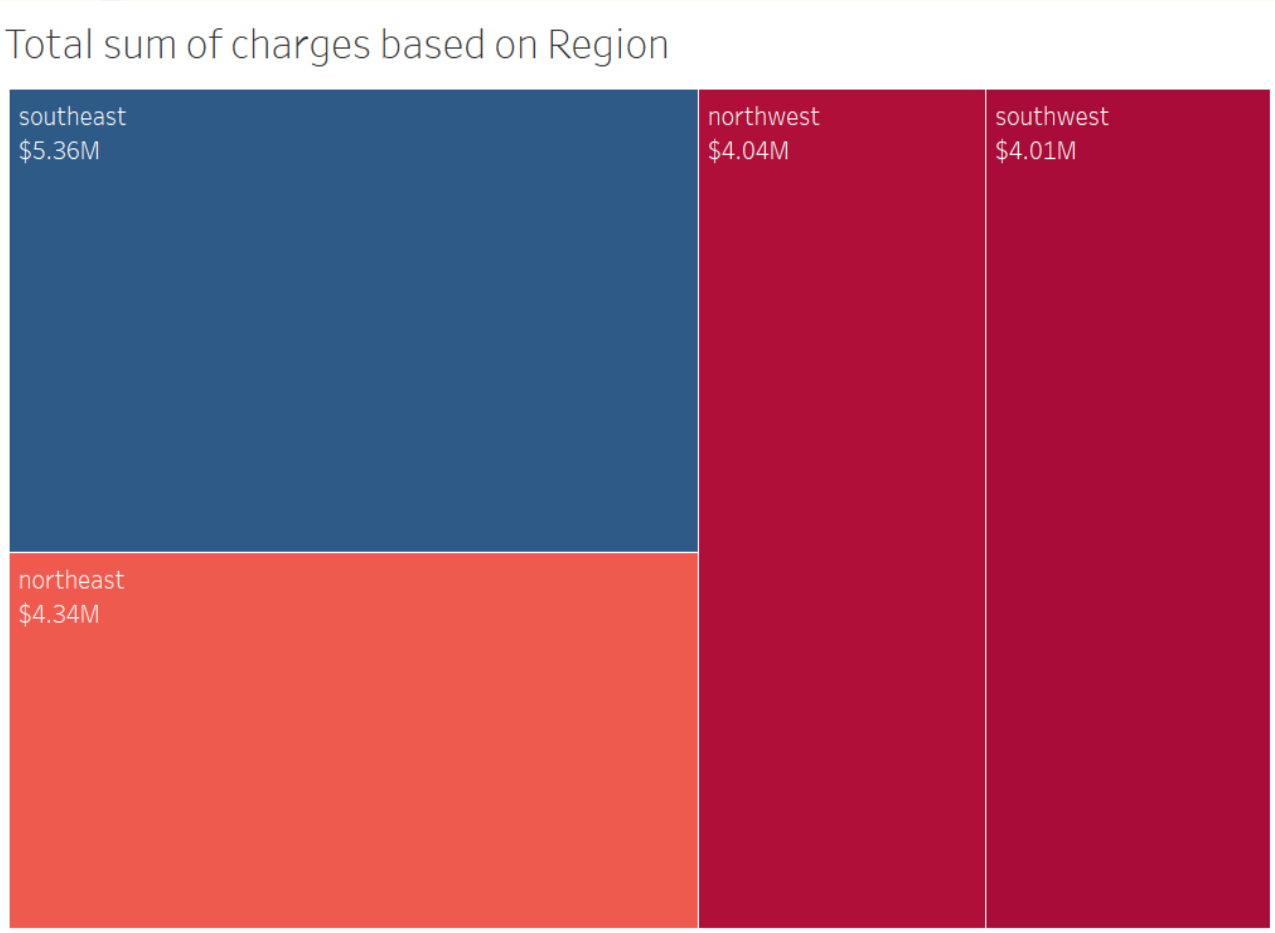
- Age: 20 and below, 21-30, 31-40, 41-50, 50-51, >60
- Sex: female, male
- Body mass index (BMI): Underweight, Normal, Overweight, Obese and Extremely Obese
- Children: Number of children covered by health insurance / Number of dependents
- Smoker: Smoking or Non-Smoking
- Region: Customer's residential area in the US; northeast, southeast, southwest, northwest.
- Charges: Medical costs billed by insurance

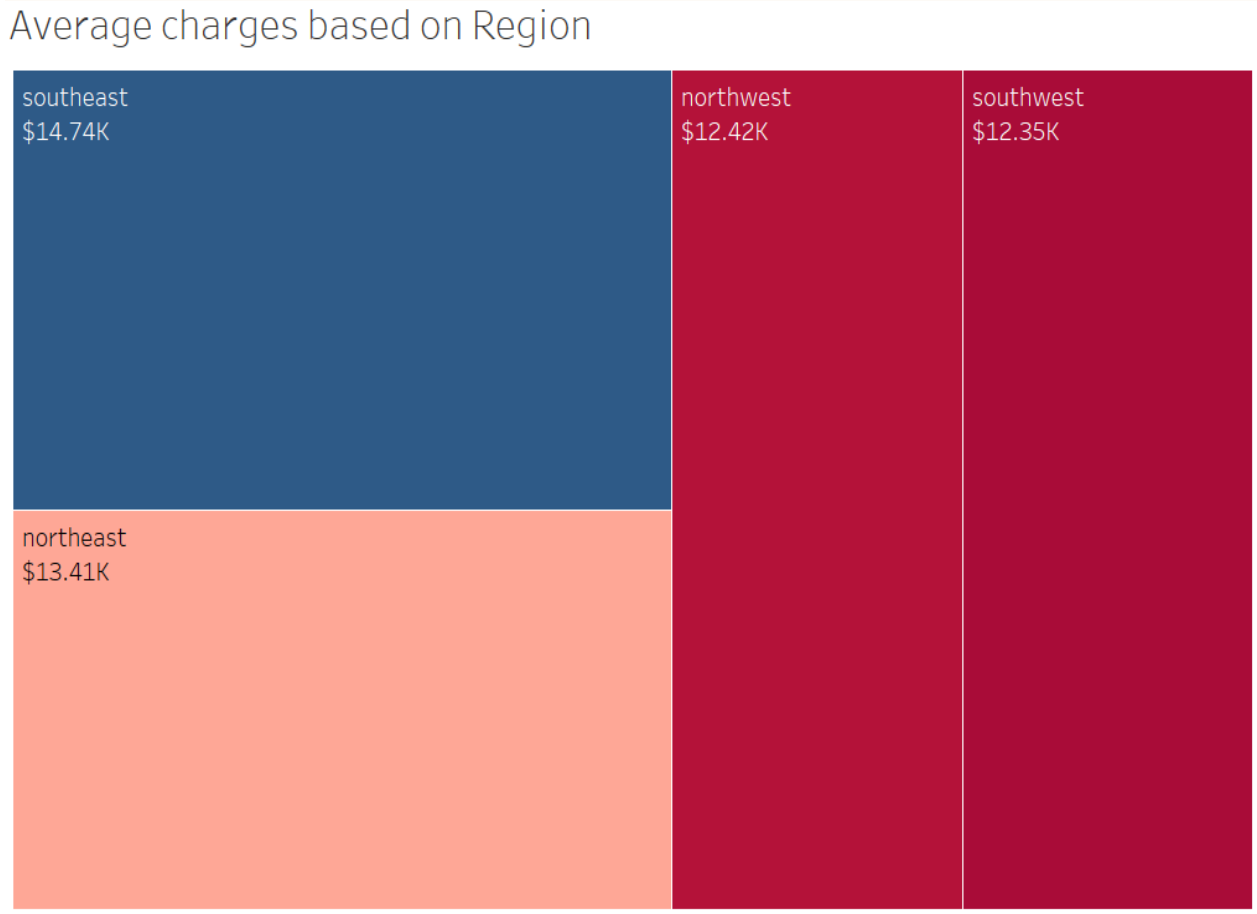
AGE VS CHARGES



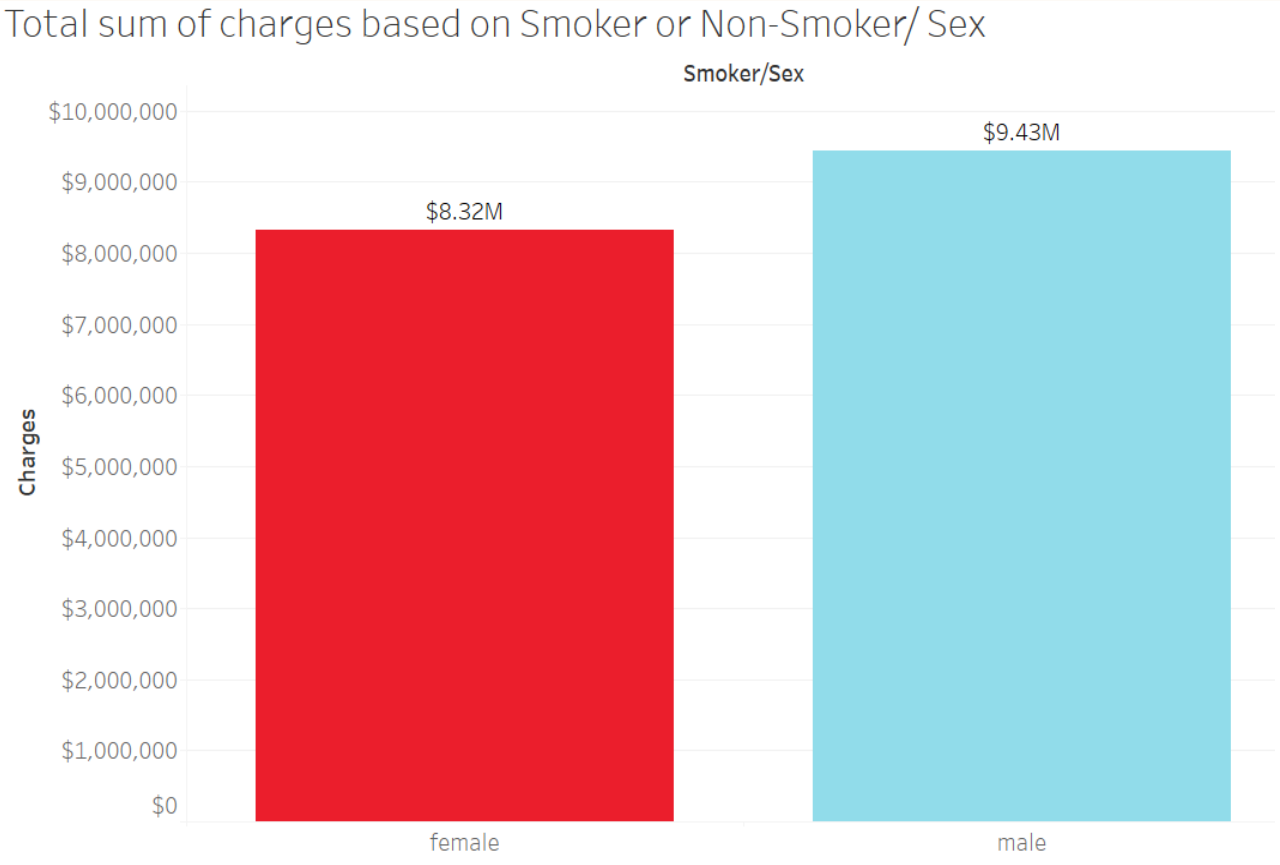


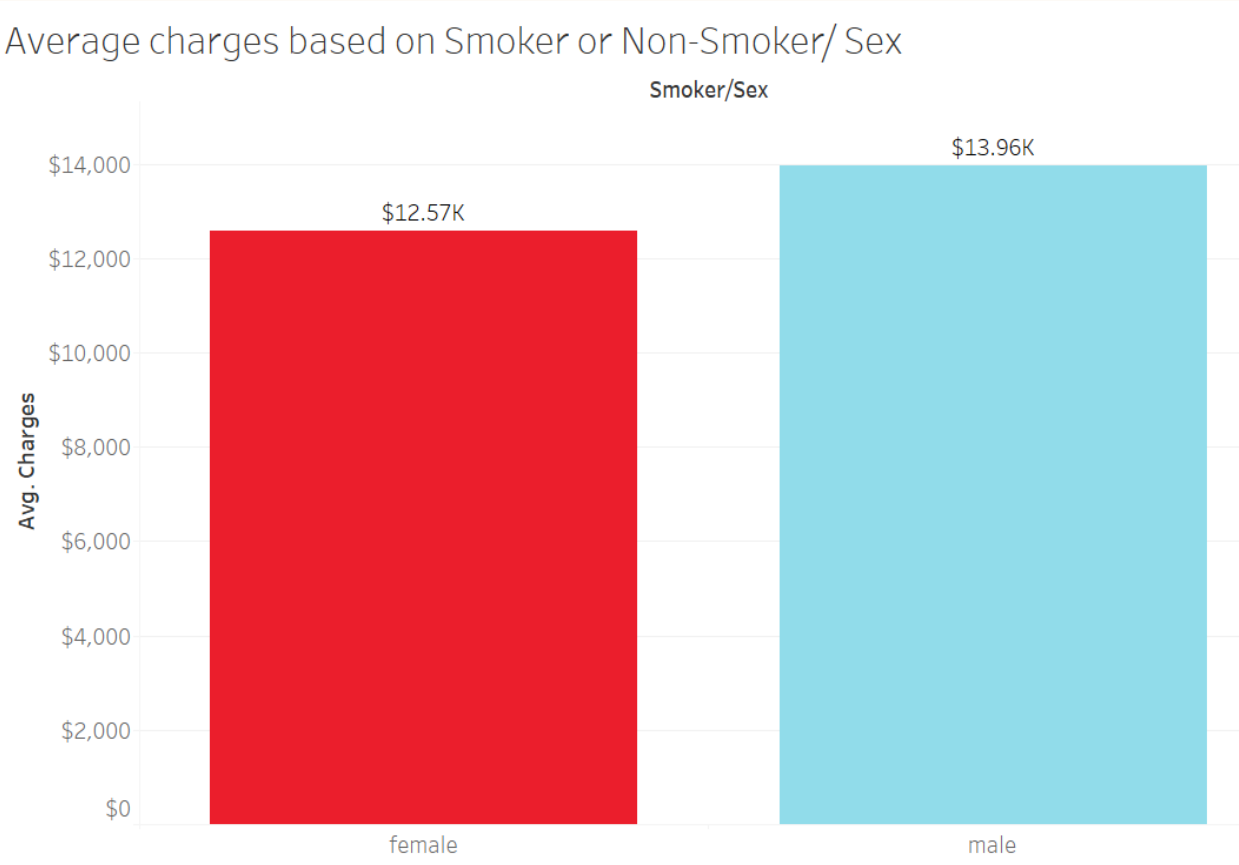
REGION VS CHARGES



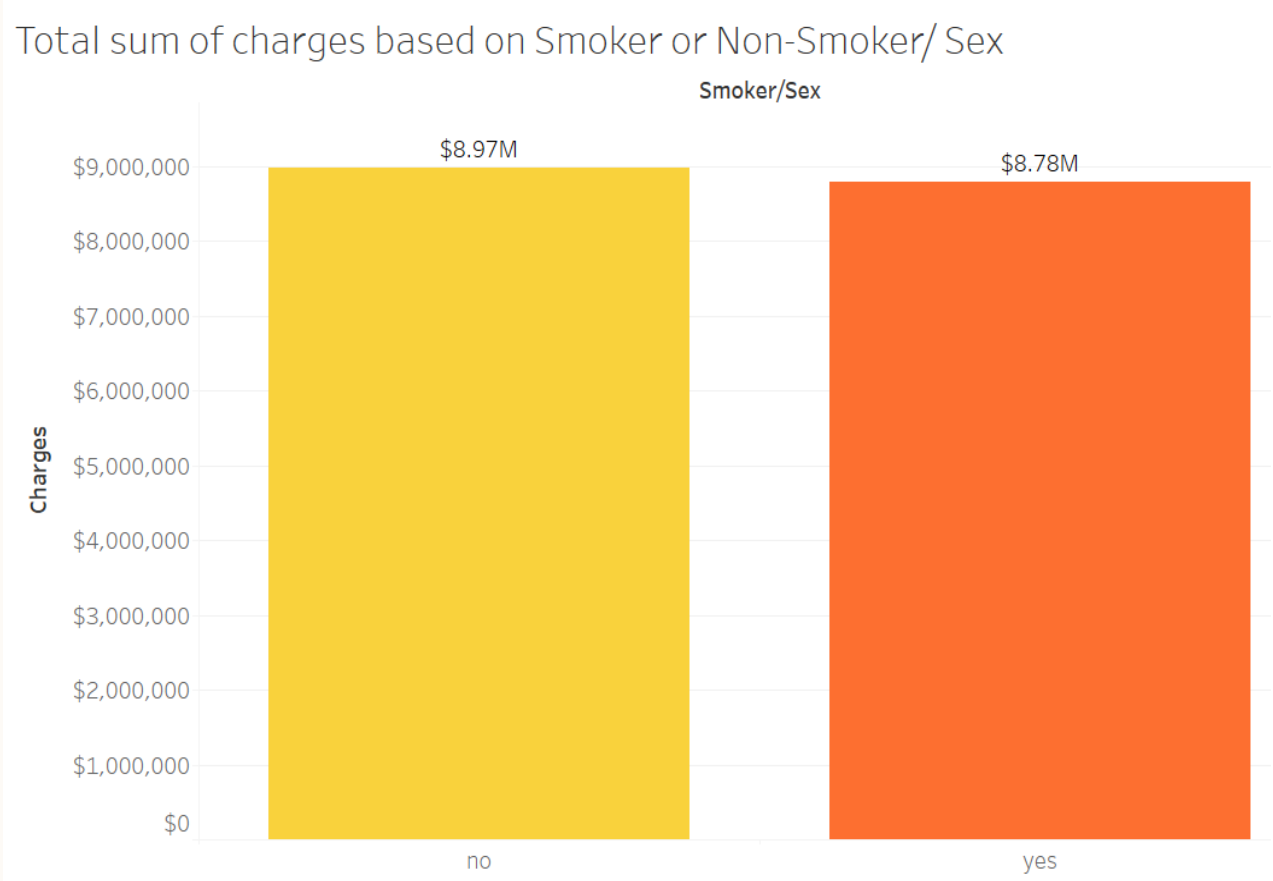


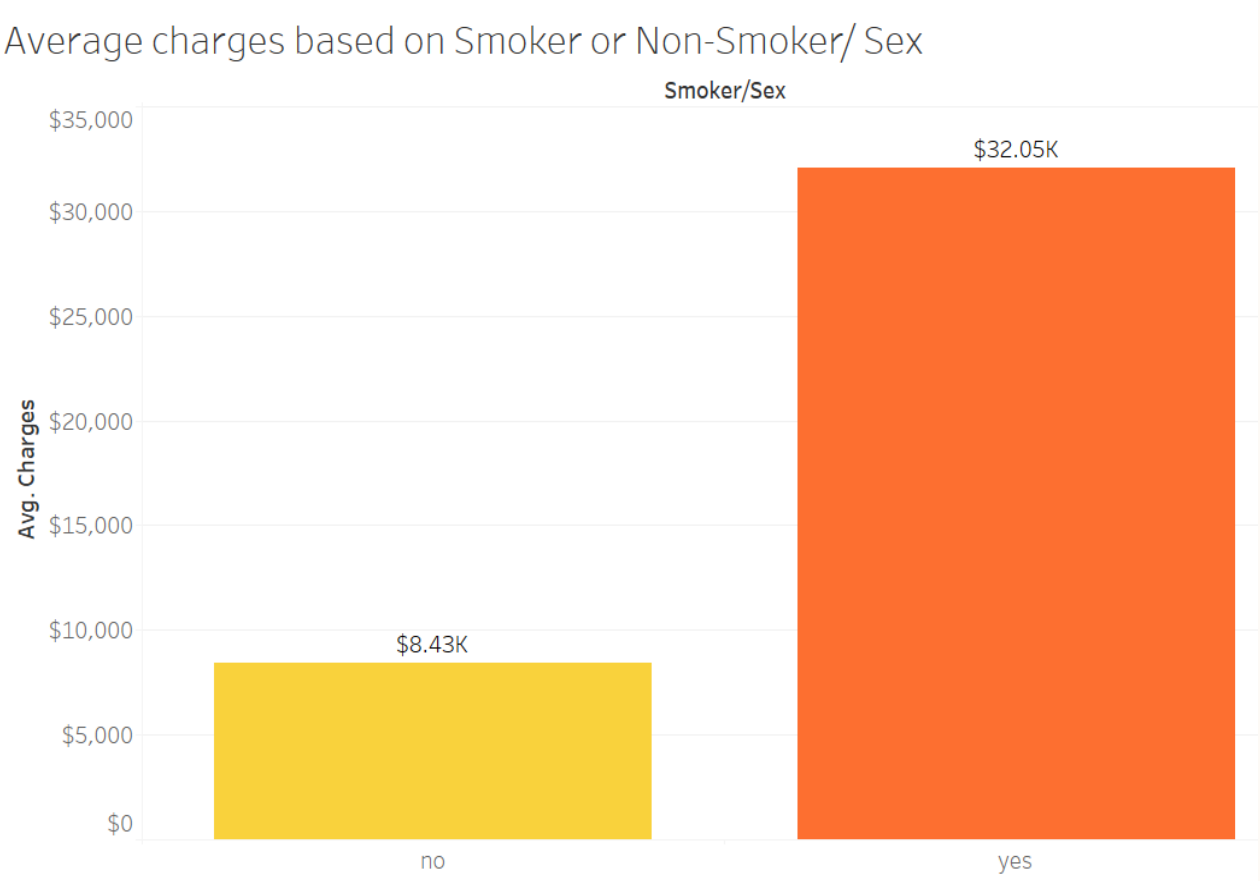
SEX VS CHARGES



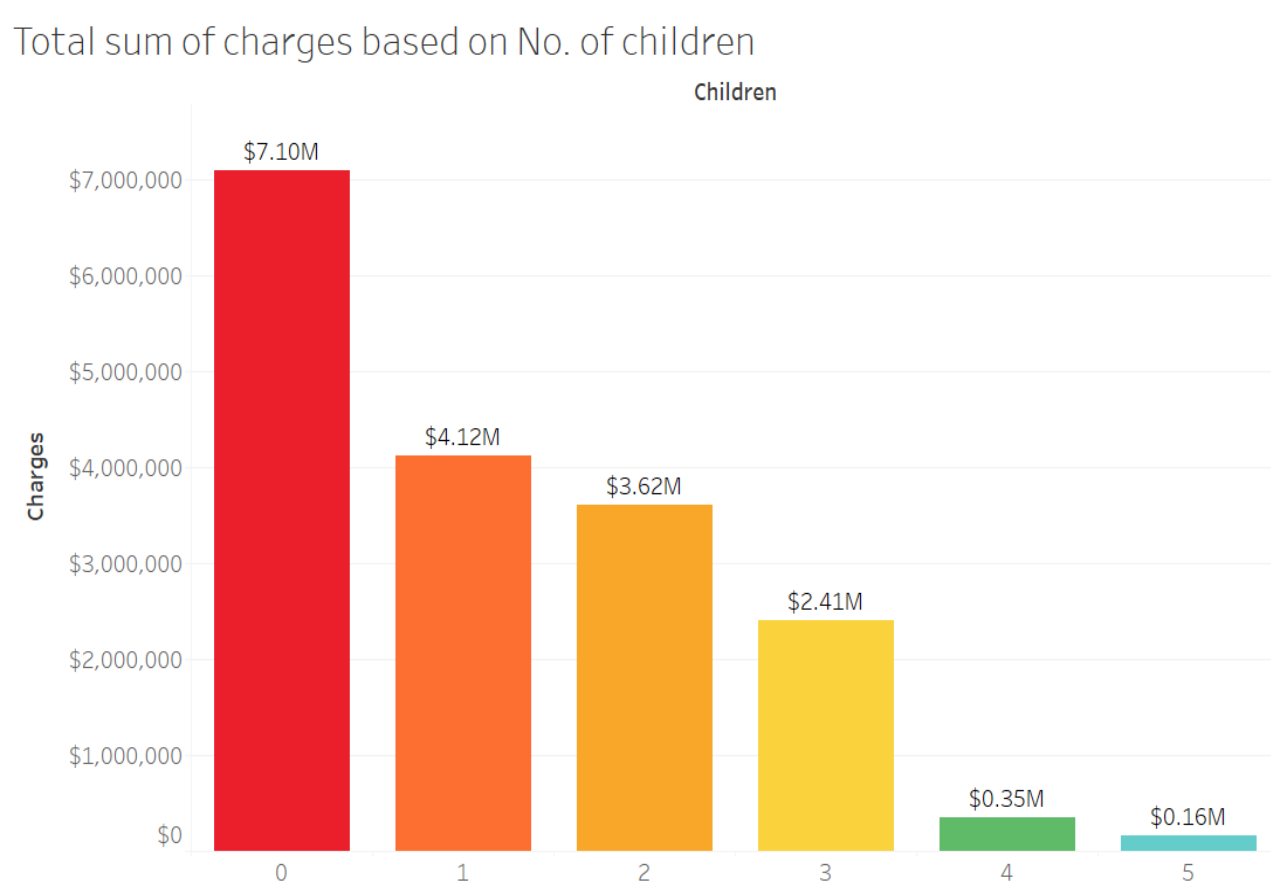


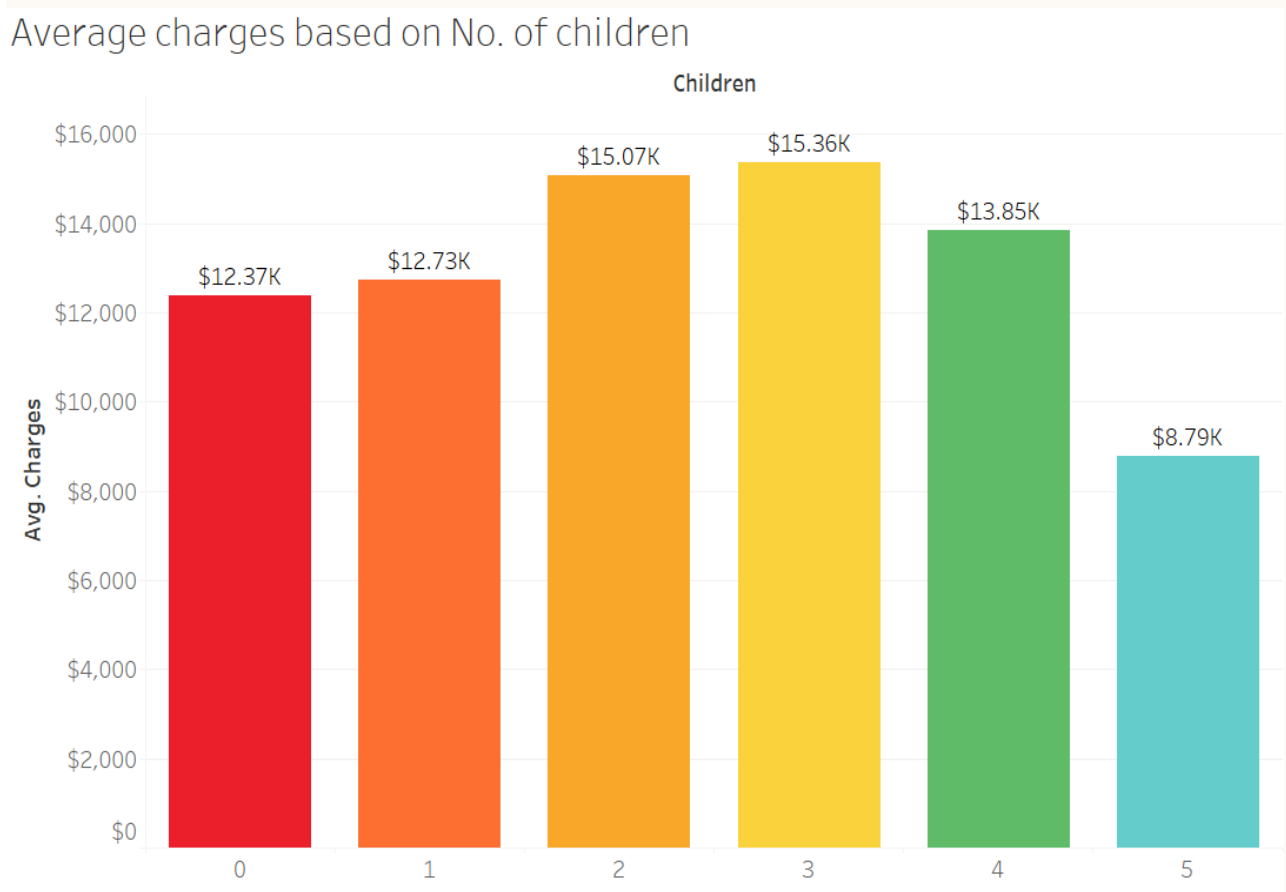
SMOKER/NON-SMOKER VS CHARGES



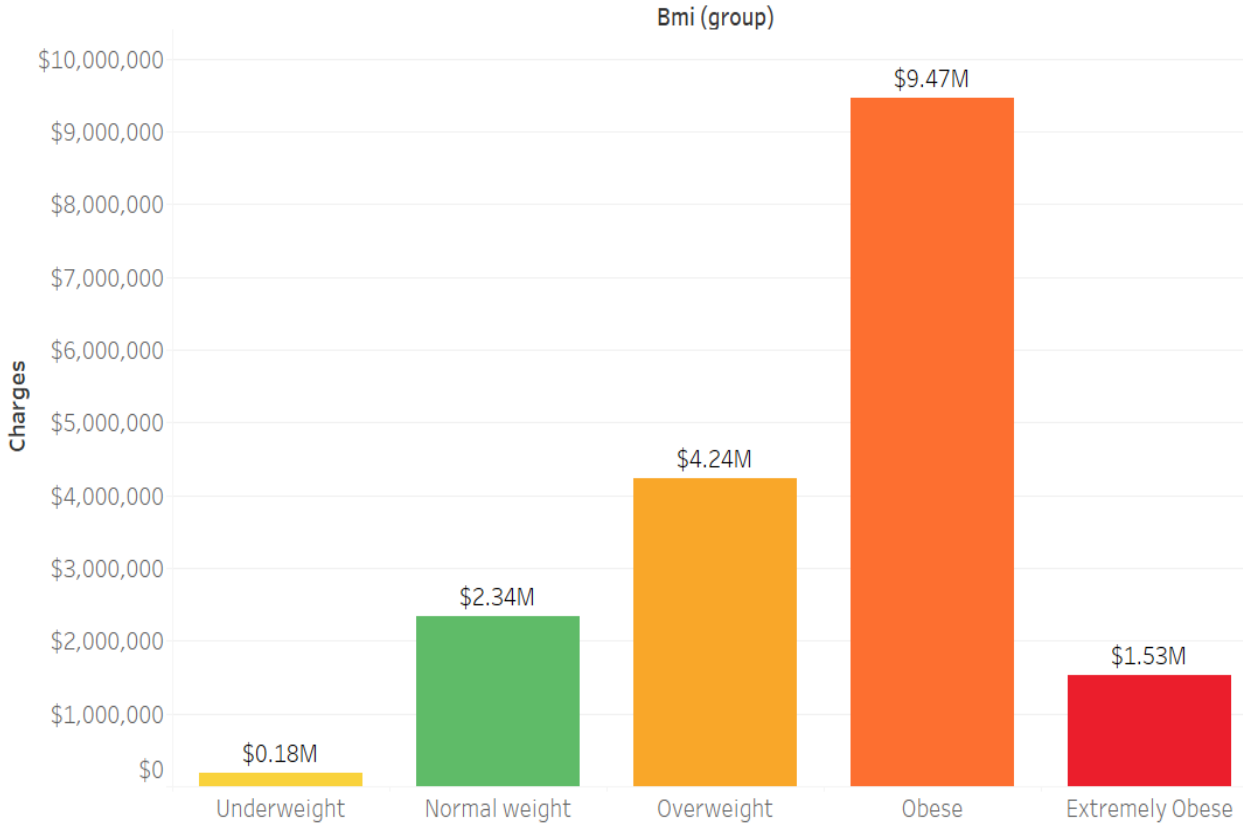


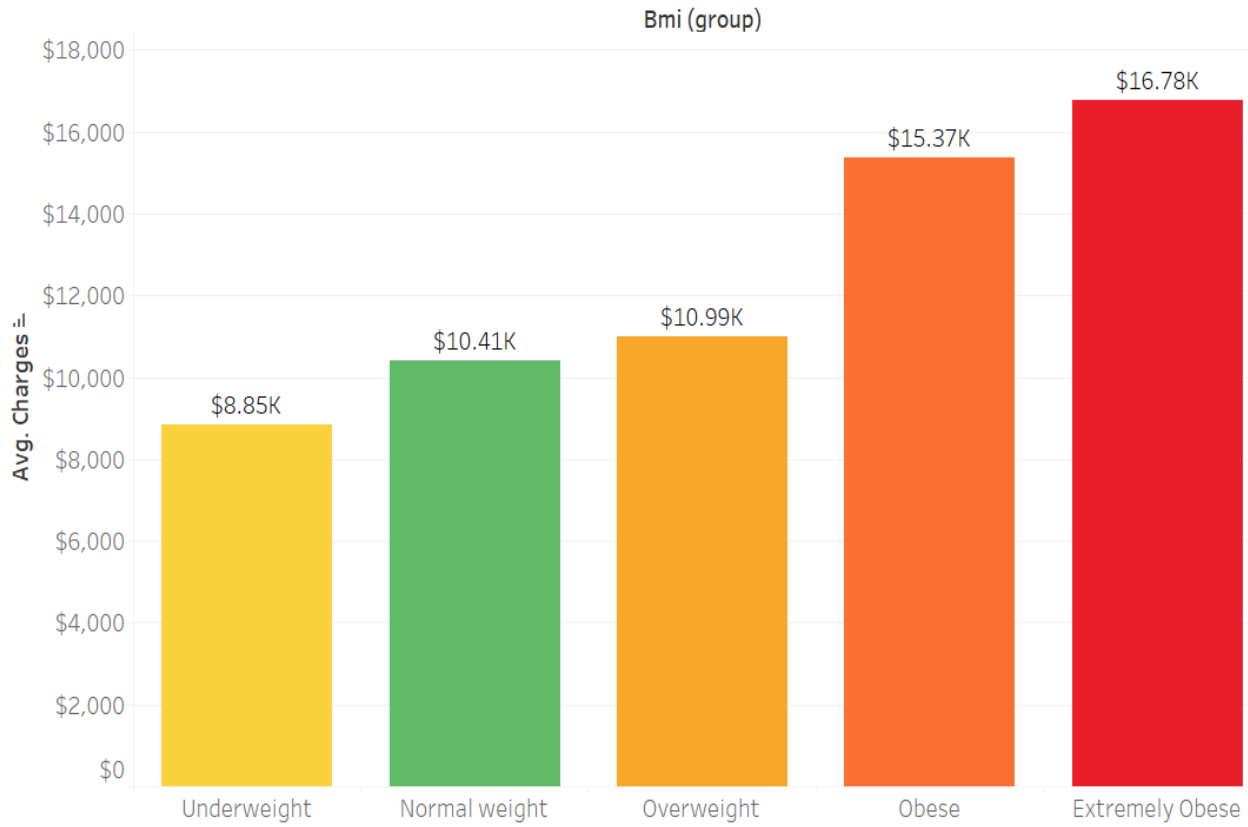
NO. OF CHILDREN VS CHARGES





BMI VS CHARGES





ANALYTIC FINDINGS

Looking at both the total sum and the average charges of medical charges for each customer demographics (Age, Region, Sex, Smoker/Non-Smoker, No. of children and BMI)

Findings:
Age: Charges increases as the age of the customer increases
Regions: Highest charges incurred are by those who reside in the east regions.
Sex: Female customers incurred more charges compared to Male customers.
No. of children: The average charges were incurred the most by parents with 3 and 2 children.
BMI: Highest sum of charges were incurred by those with BMI= Obese. While the highest average charge was incurred by BMI= Extremely obese.

RECOMMENDATIONS

From the analysis (refer to dashboard SUM and AVG) it can be seen that the SEX of the customer plays a significant role in charges incurred. Therefore this can be kept in mind when determining future premium rates.

The company can consider increasing future premiums for customers who are aged >40, male, smokers, customers who reside in the east of US and those who have a BMI of Overweight and more.

For those with children, further investigations would be required to determine premium rates.

THANK YOU