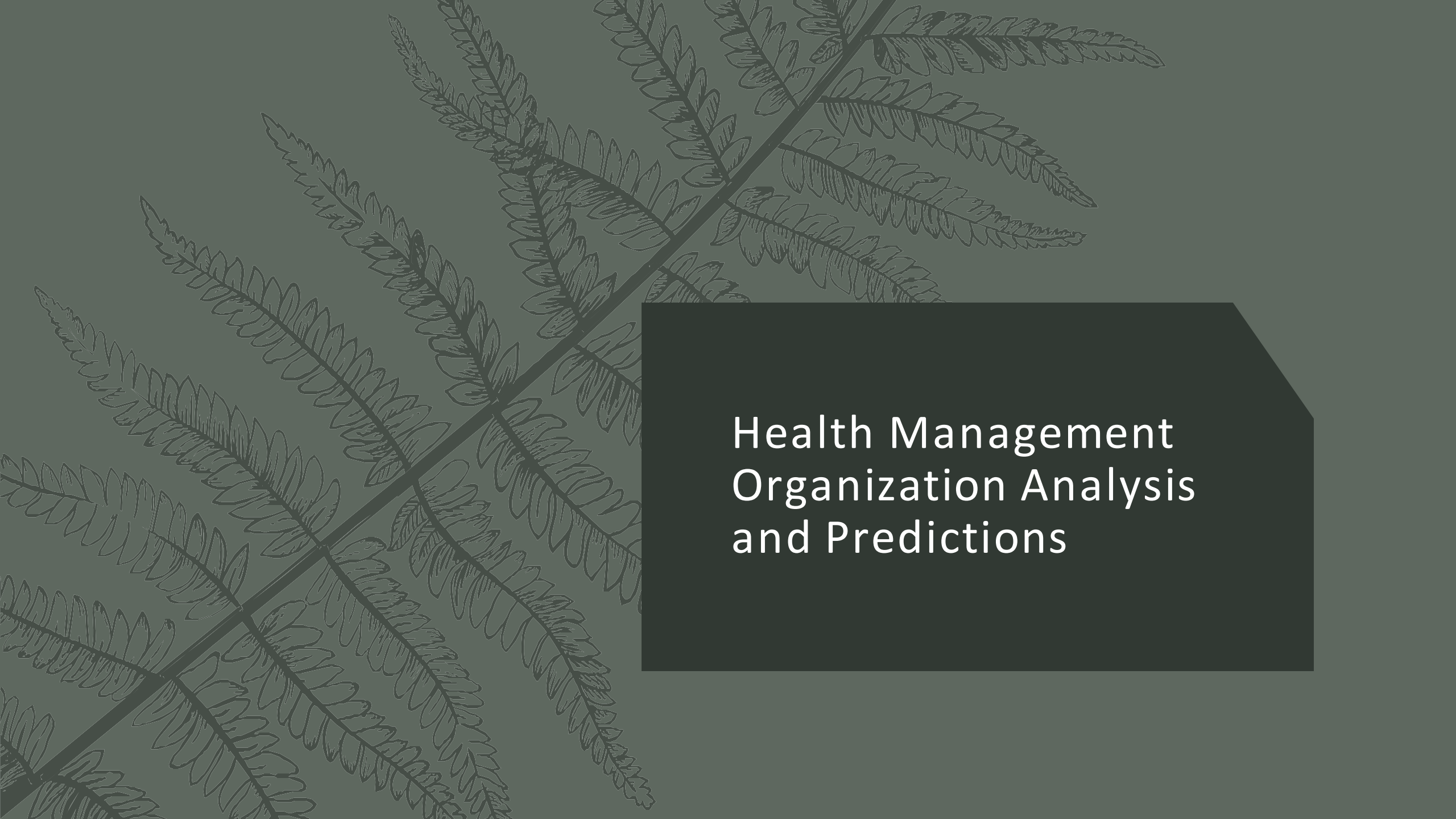


Health Management Organization Analysis and Predictions

Project Approaches & Goals

Exploring the descriptive statistics of the data and visualizing them.

Derive relationship between variables.

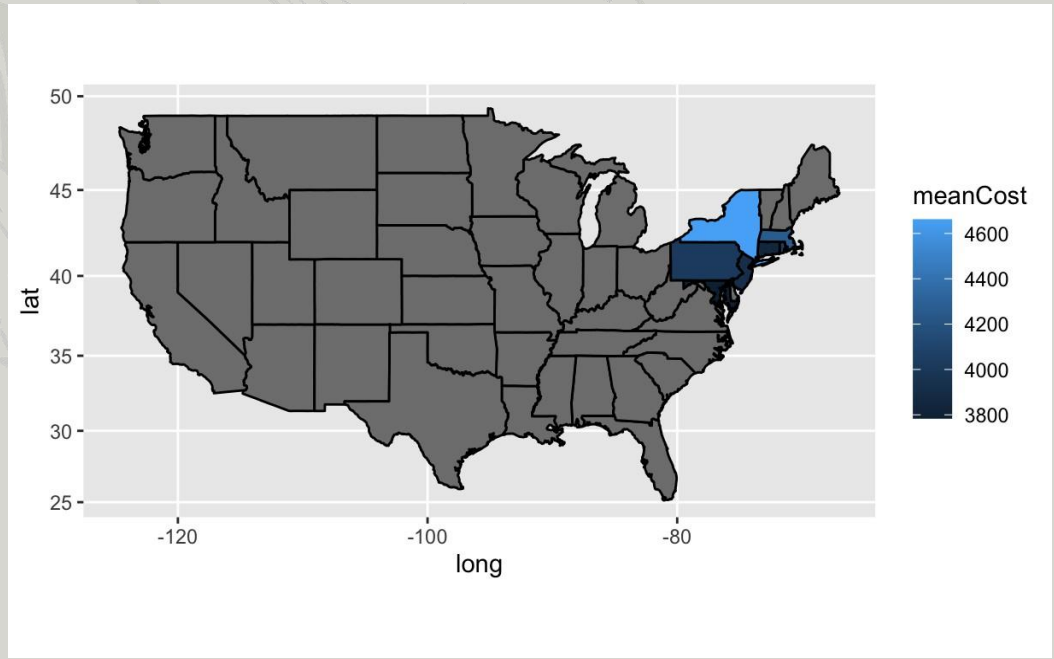
Provide actionable insights and recommendations to the HMO:

- Predict customers that will have high healthcare costs next year;
- How to lower total health care costs.

Exploring the Data

We have 7582 people's data.
People are from the 7 Northeastern States of the US.

All the variables that we have:
Age, BMI, number of children, smoker or non-smoker, state of living, living at country or urban area, education level, go to see doctor yearly or not, exercise or not, married or not, being hypertension or not, gender, and the annual cost on health care.



Examining the missing values:

There are 80 missing values for BMI, 68 missing for hypertension, and 38 for cost.

Exploring the Data

Creating categories for Age & BMI
Easier to conduct and visualize the analysis

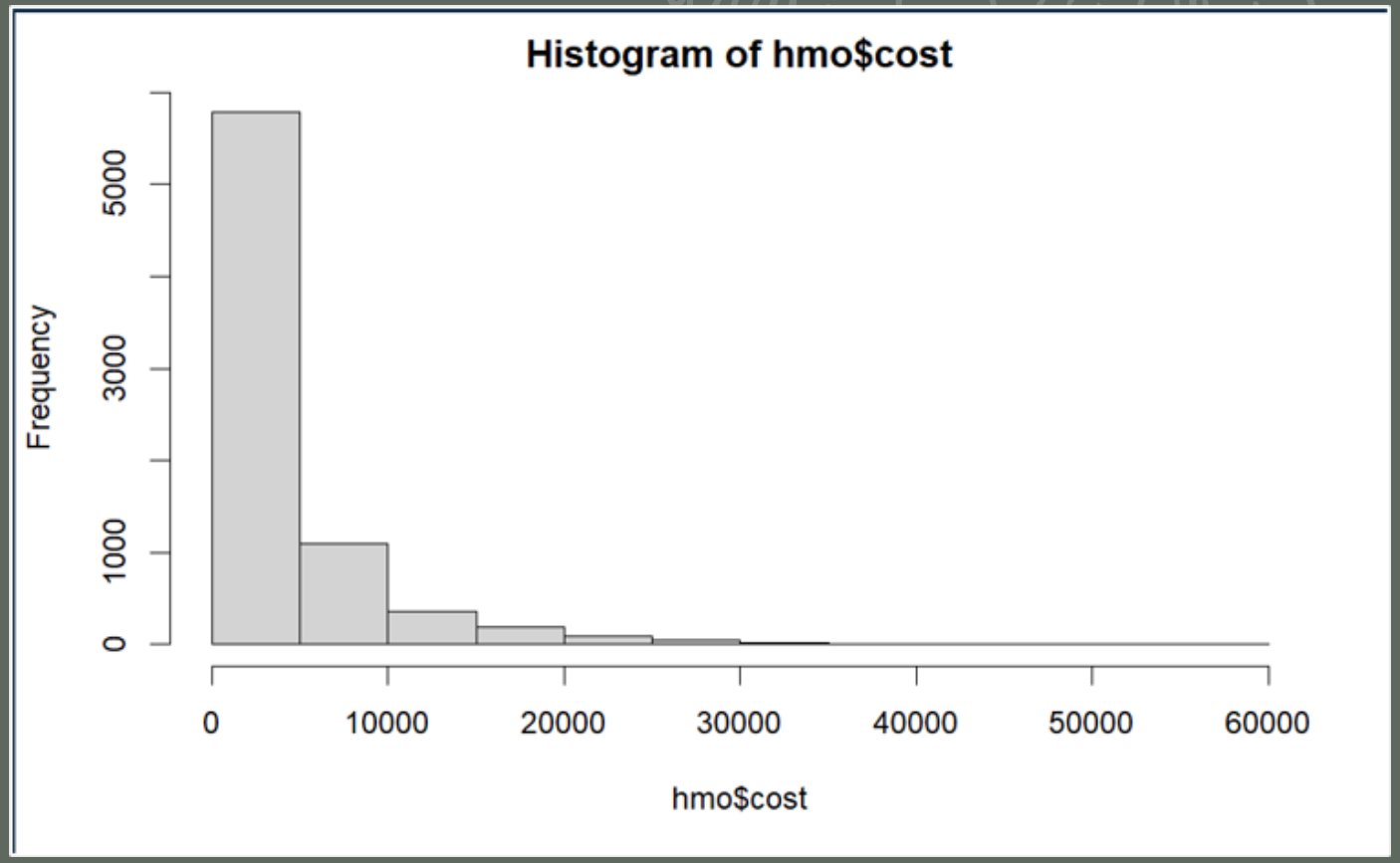
Age Category:

Young Adult	18-34
Middle Aged	35-50
Older Adults	51-66

BMI Category:

Underweight	< 18
Healthy	>=18 & <25
Overweight	>=25 & <30
Obese	>=30 & <40
Extremely Obese	>=40 & <65

Histogram of Cost Variable



- Skewed to the right.
- Most of the data falls between $2 and $5,000, but there are several outliers above $20,000.

- We chose to define expensive as a person that has a healthcare cost of $4,776 or more. The reason we picked this value was because 75% of the people are under $4,775.

- 1895 people's health care cost are expensive

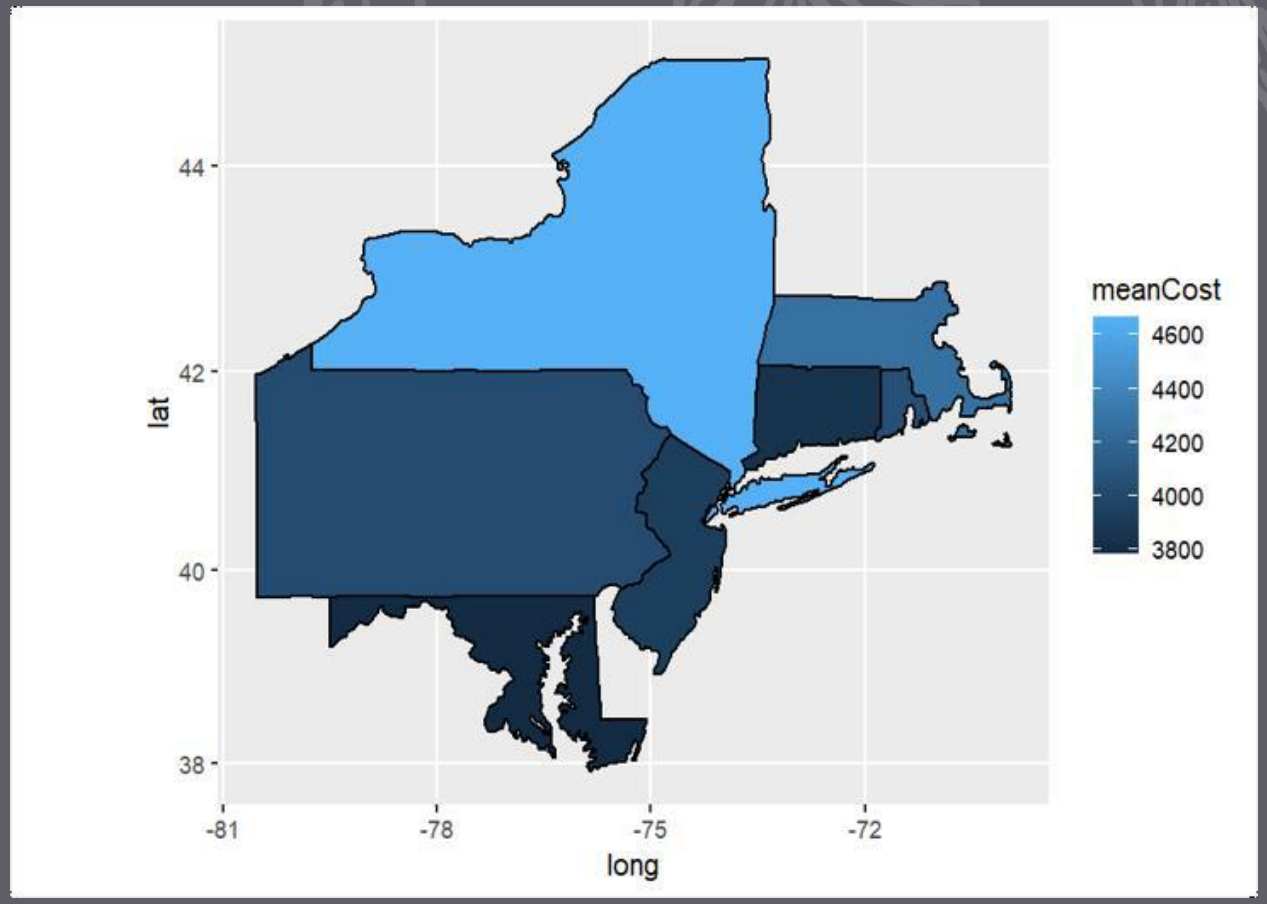
Significant Factors

58.4% of the variation in health cost can be explained by all the 12 factors we have. (Age, BMI, number of children, smoker or non-smoker, state of living, living at country or urban area, education level, go to see doctor yearly or not, exercise or not, married or not, being hypertension or not, gender.)

Significant factors (with more asterisks): smoker(Yes), exercise(Not-Active), Age, BMI, number of children, hypertension

```
.       X
***     age
***     bmi
***     children
***     smokeryes
        locationMARYLAND
        locationMASSACHUSETTS
        locationNEW JERSEY
*       locationNEW YORK
        locationPENNSYLVANIA
        locationRHODE ISLAND
        location_typeUrban
        education_levelMaster
        education_levelNo College Degree
.       education_levelPhD
        yearly_physicalYes
***     exerciseNot-Active
        marriedNot_Married
***     hypertension
        gendermale
        agecategoryolder-adults
***     agecategoryYoung-adults
        bmicategoryHealthy
***     bmicategoryObese
        bmicategoryOverweight
*       bmicategoryUnderweight
```

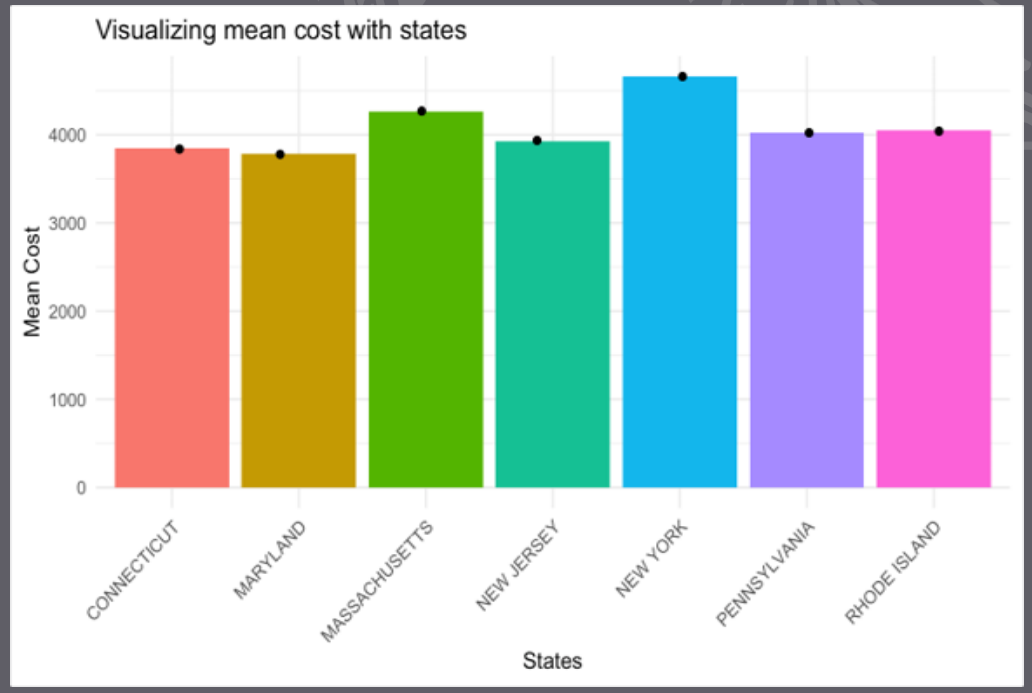
Visualization of mean healthcare costs by states by using a map:



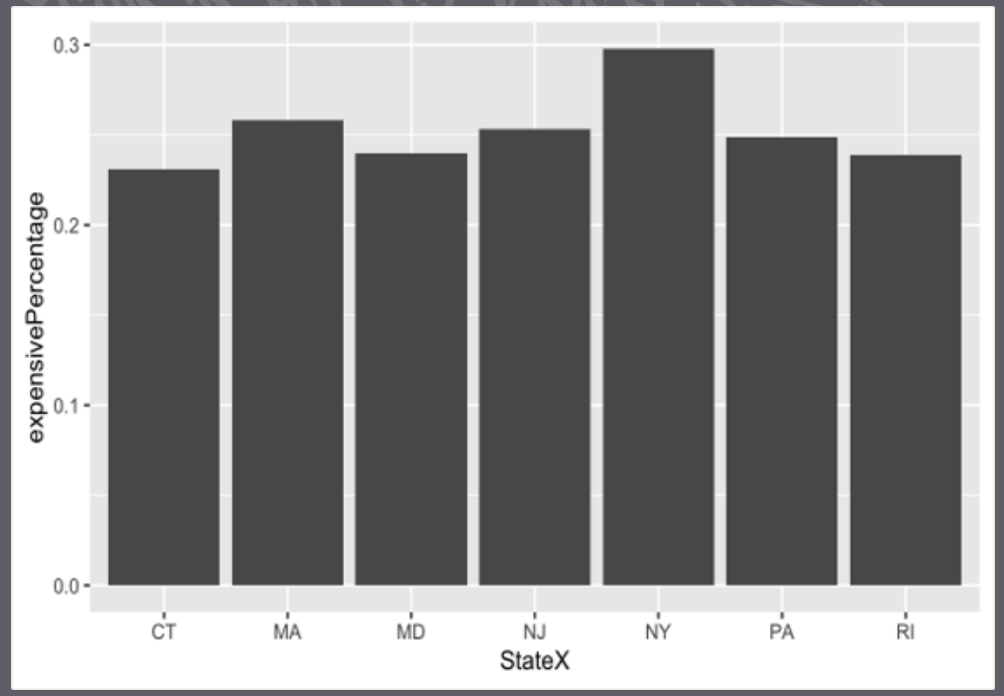
- The lighter the color, the higher the mean cost.
- This shows that people living in New York have the highest healthcare costs, on average. New York is followed by Massachusetts. The remaining states all have negligible differences in the average healthcare costs.

Another 2 visualizations showing the data group by State:

Mean cost of each State:



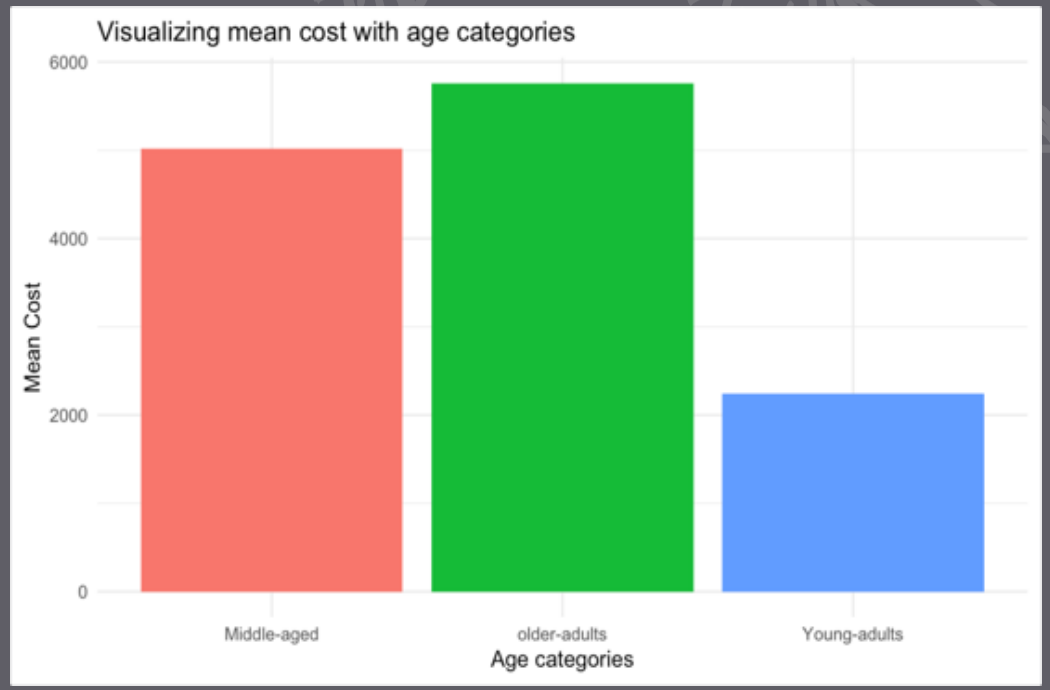
Percentage of being expensive by State:



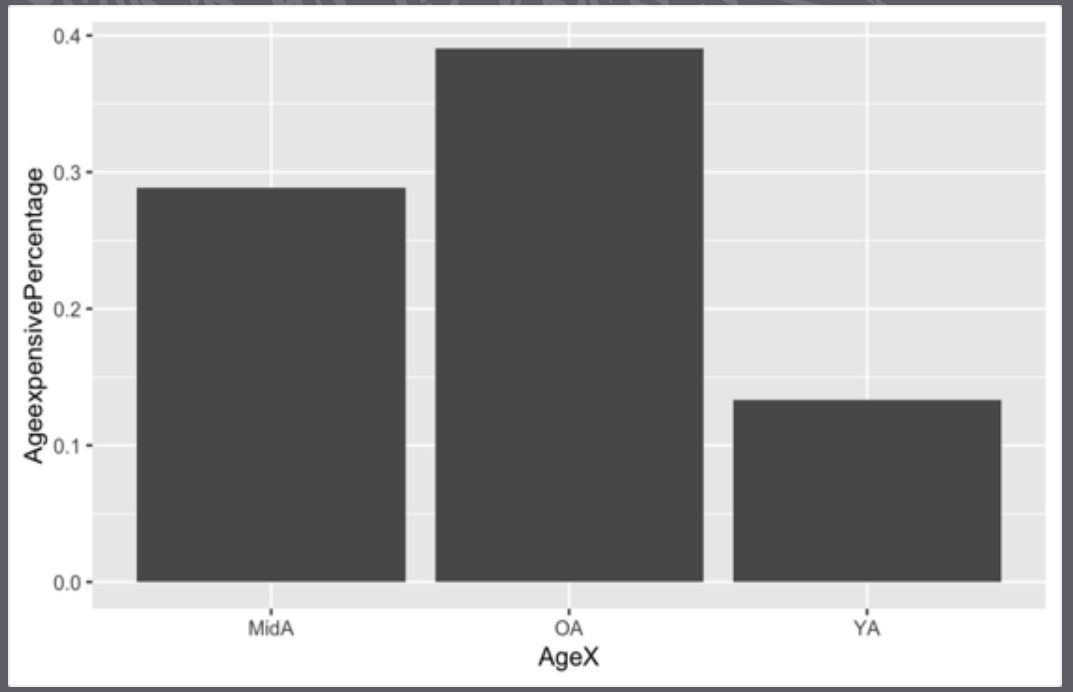
People living in New York have the highest percentage of being expensive on healthcare costs. New York is followed by Massachusetts.

2 visualizations showing the data group by Age Category:

Mean cost of each Age Category:

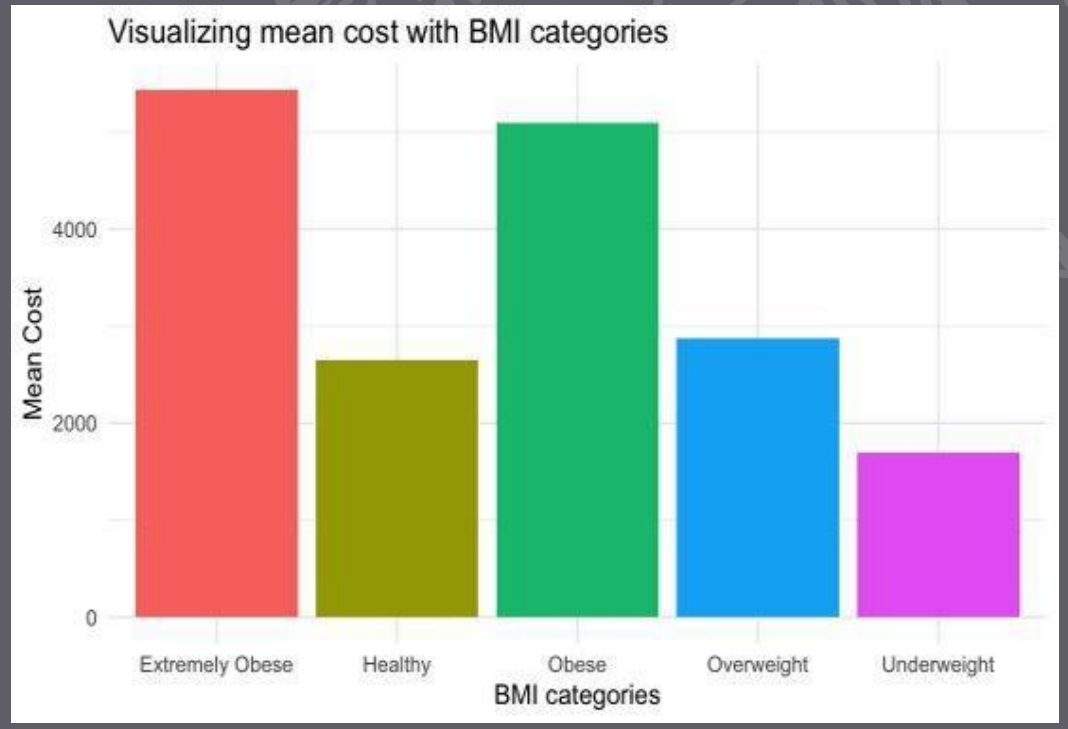


Expensive Percentage by Age Category:

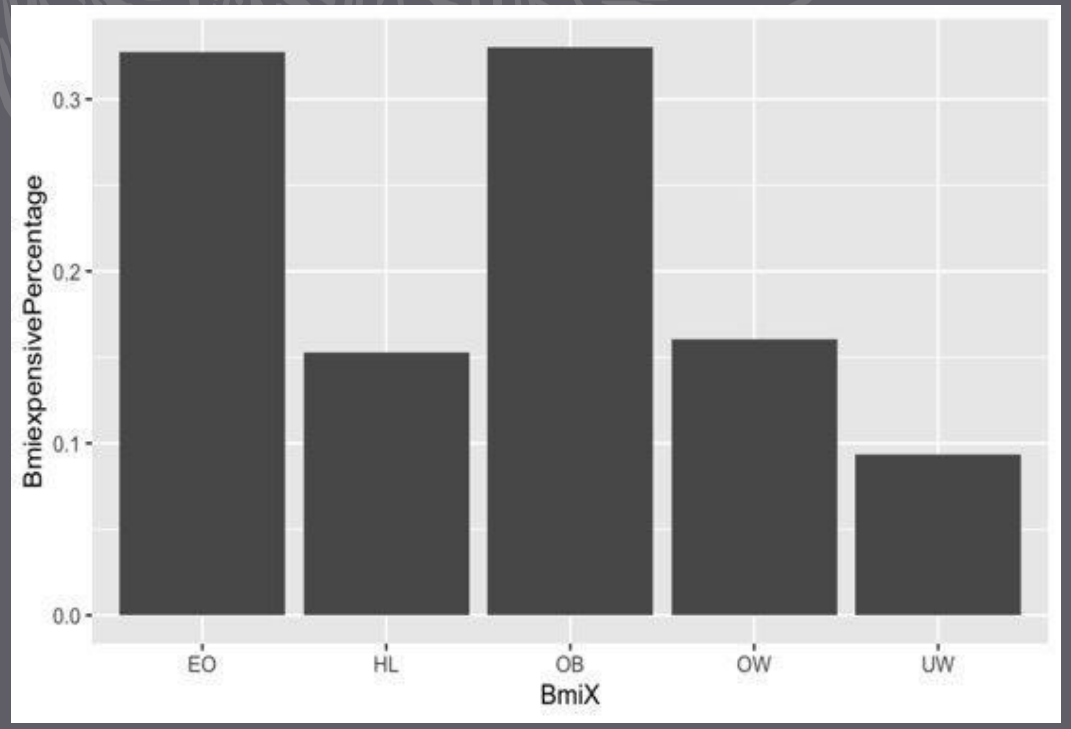


The mean cost and the percentage of being expensive are higher for older adults, and are significantly lower for young adults.

2 visualizations showing the data group by BMI Category:

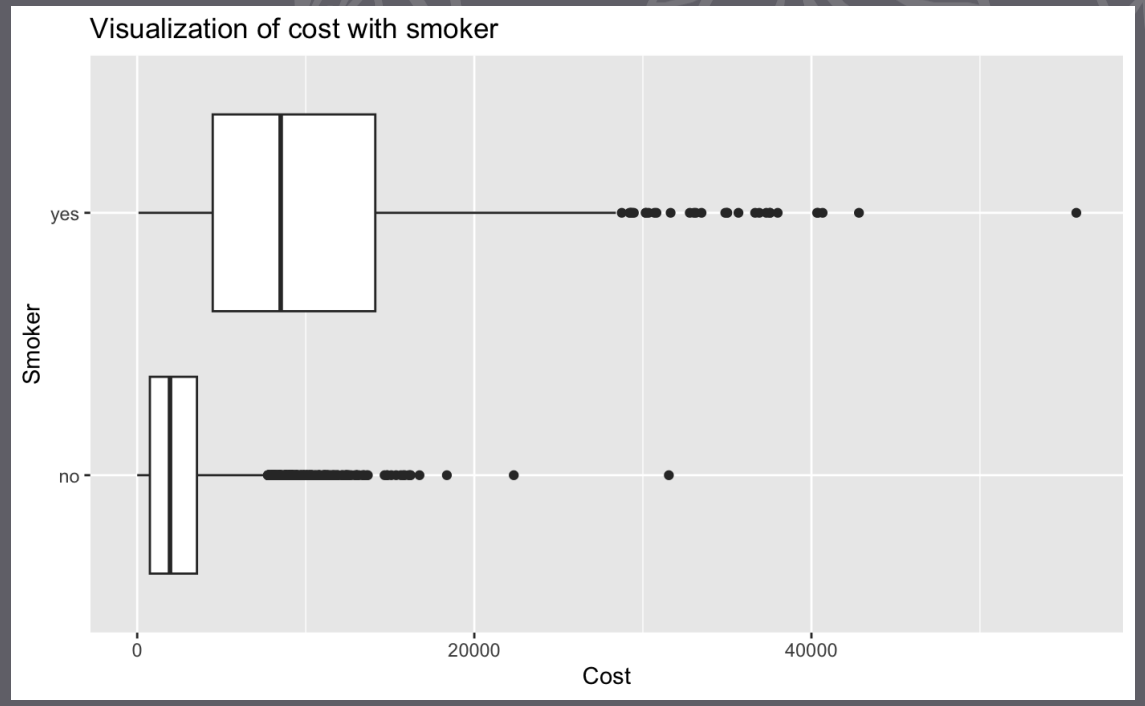


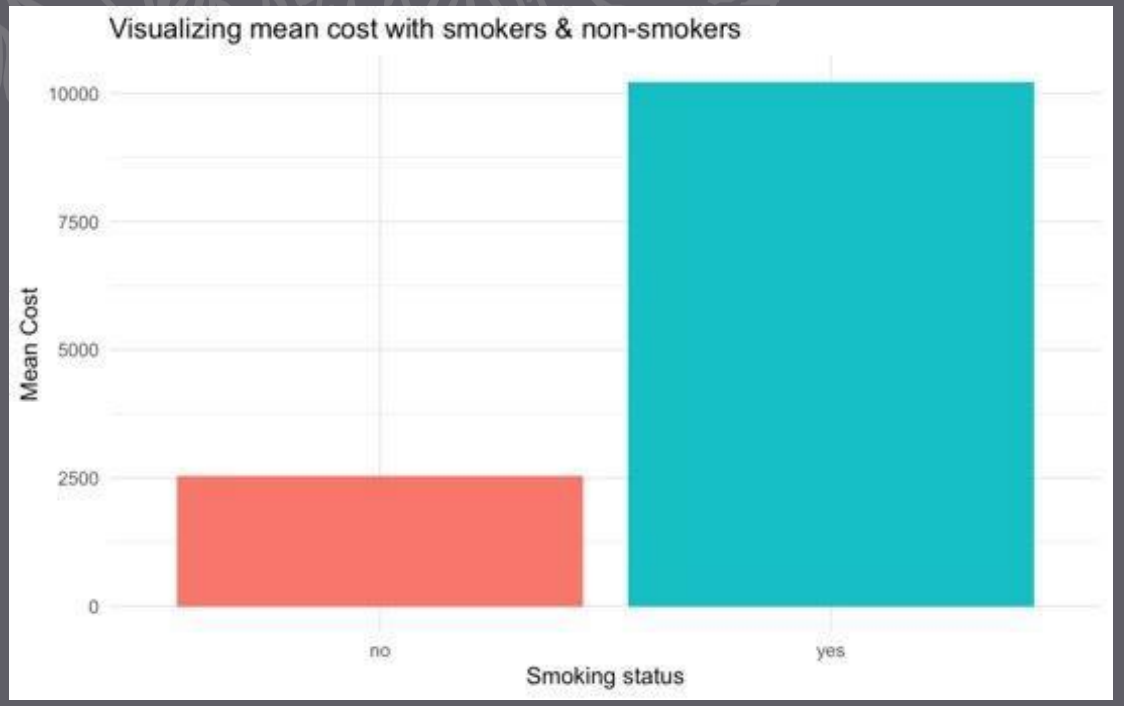
Average cost for extremely obese and obese people is well over $4,000.
The cost for overweight and healthy people is similar while the cost for underweight is the lowest.



The percentage of being expensive grouped by the BMI category, the cost for extremely obese and obese people is more likely to cost more on health care.

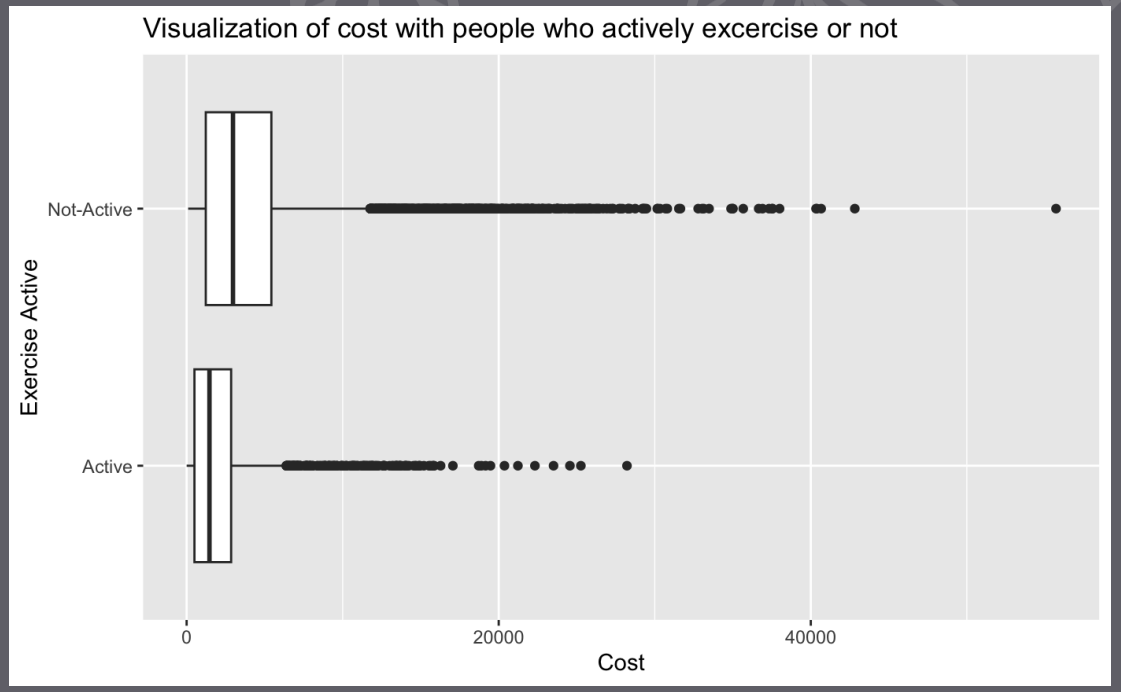
Visualizations group by Smoking status

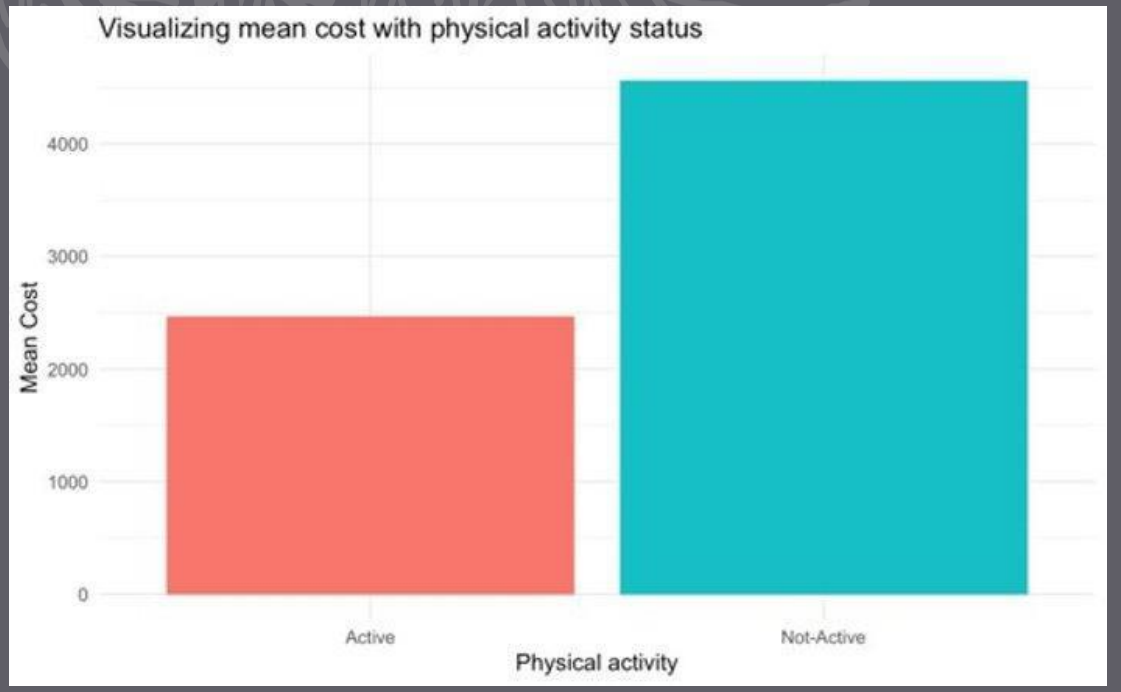




Smokers have significantly higher mean cost than non-smokers.
A smoker has a 73% probability of having an expensive cost on health care

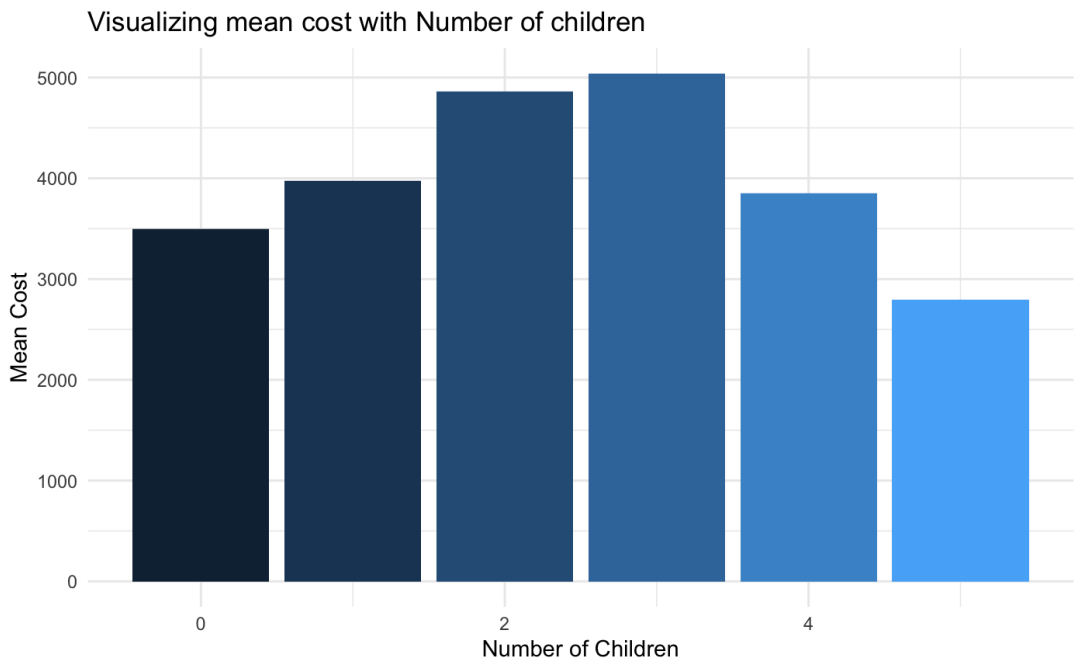
Visualizations group by Physical Exercise status

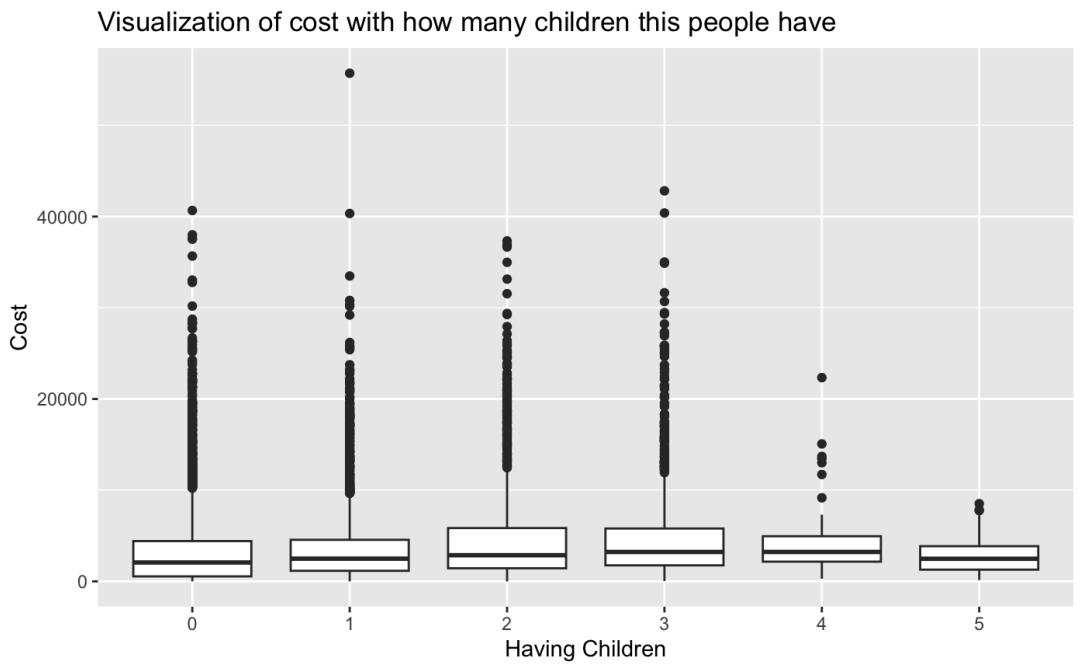




People with Non-Active physical exercise have higher mean cost.

Visualizations group by Number of Children

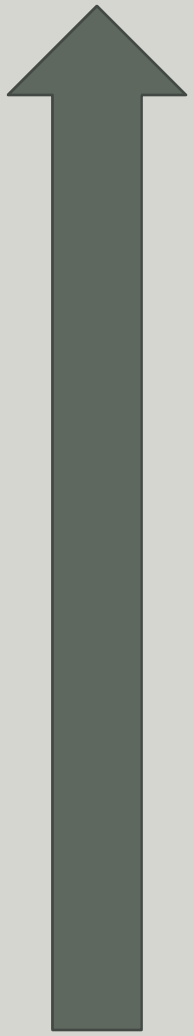




As the number of children increases, the cost of health care increases,
but starts from there are four, the cost of healthcare decreases,
and when there are five children, the cost of healthcare is even lower than if there were no children.

Within the 1895 people whose health care cost are expensive:

Let's pay more attention to these combinations



Factors	Count
{bmicategory=Obese,smoker=yes,yearly_physical=No,exercise=Not-Active}	383
{bmicategory=Obese,smoker=yes,exercise=Not-Active}	500
{bmicategory=Obese,smoker=yes,exercise=Not-Active,hypertension=yes}	394
{bmicategory=Obese,smoker=yes,married=Married}	434
{bmicategory=Obese,smoker=yes}	639
{bmicategory=Obese,smoker=yes,education_level=Bachelor}	400
{bmicategory=Obese,smoker=yes,gender=male}	414
{bmicategory=Obese,smoker=yes,yearly_physical=No}	485
{bmicategory=Obese,smoker=yes,hypertension=yes}	494
{bmicategory=Obese,smoker=yes,location_type=Urban}	476
{smoker=yes,exercise=Not-Active,married=Married,gender=male}	383
{smoker=yes,location_type=Urban,exercise=Not-Active,gender=male}	424
{agecategory=Middle-aged,smoker=yes}	449
{smoker=yes,exercise=Not-Active,gender=male}	557
{smoker=yes,exercise=Not-Active,hypertension=0,gender=male}	437
{smoker=yes,yearly_physical=No,exercise=Not-Active,gender=male}	419

Prediction & Insights

- people that are obese, extremely obese, not active exercise, a smoker, and not going to a yearly physical are most likely to be expensive.
- In addition to this, living in New York increases the chances that a person is expensive compared to people living in the surrounding states in the northeast.
- Additionally, as people get older, they are more likely to be expensive; and an increase in children (but less than 5) leads to an increase in price generally.

Recommendations

how to lower health care costs? - since we cannot control our aging
Plans & programs

- Reduce smoking
- Keep BMI less than 30
- Be active at exercise
- Go to yearly physical-health checkup plans

Thank you

Any Questions