Analysis

Caroline Foshee, Sivangi Raychoudhury, and Christian Zaner

We had three main questions for this project: Which independent variables are correlated with each other? Which factors are most closely associated with insurance costs? Why are certain factors most closely associated with insurance costs? We found that the most significant factors that affect charges are smoking and BMI, and the factors with less significance are age, gender, region of the United States, and number of children.

Smoking is the strongest predictor of health insurance costs, because it has been linked to major health issues such as various cancers, respiratory problems, heart disease, lung disease, increased risk of infection, and weakening of the immune system. Because of all these reasons, smoking will always raise the cost of health insurance regardless of other factors, such as age or gender, as demonstrated in slides 12, 13, 14, and 16. This is most pronounced on slide 14, where the mean cost of the smokers is 32050.23 and the non-smoker mean cost is 8,440. That is around a 3.8 times increase.

BMI also significantly raises health insurance costs, because a high BMI can cause type-2 diabetes, heart disease, stroke, high blood pressure, sleep apnea, joint problems, and mental health issues. We see the significance of BMI on slides 13 and 16. On slide 13, regardless of the other categories such as smoking, age, and family size, a BMI of say 25, could be as low as a couple thousands while the upper end could cost over 10,000. With the correlation line of .2, there is a low to moderate positive correlation between BMI and insurance cost.

Other factors such as age do play a role in insurance costs, as seen on slide 16. Young non-smokers with a low BMI tend to pay less than $5,000, while older people with similar BMIs and smoking habits pay slightly above $10,000. However, smoking increases insurance costs across all ages. Young and old smokers that have lower BMIs tend to pay around $15,000 and $25,000, respectively. Age is correlated with insurance costs, because older people tend to have declining health due to weak immune systems and chronic conditions building up over time.

Gender also has a slight effect on insurance costs. Males tend to have higher insurance costs than females (as seen on slide 14). This is because men smoke more frequently and have higher BMIs than women, as seen on slides 11 and 12. Similarly, the southeast region of the US has the highest insurance costs, because they have higher average BMIs and higher smoking rates.

The last factor that could play a role is family size. Non-smoking families pay less than families that do, though non-smoking families' costs increase slightly as the number of children increases. For smoking families, their insurance costs decrease slightly as the number of children increases, see slide 17. This could be because more kids cost more money to raise, so there's less money to spend on cigarettes, and so they don't have as severe health consequences to smoking.