

Title: Unveiling Gender Disparities in Medical Insurance Costs: A Comprehensive Analysis

By Gladys Dalla

Executive Summary

Many factors can cause health insurance costs to differ greatly, and it is important for both insurance providers and beneficiaries to be aware of these discrepancies. This analysis explores the complexities of medical insurance costs, providing insight into the various factors that affect these expenses, with a specific focus on gender differences. The information being analyzed offers a thorough understanding of the factors that influence medical costs, enabling a detailed investigation into the disparities in healthcare expenses depending on gender. This study aims to uncover the complex relationship between demographic factors and healthcare costs by asking probing questions about the influence of gender on different BMI categories, the impact of smoking habits on gender-specific medical expenses, and the existence of significant disparities in charges between genders. In addition, the analysis also examines the impact of the number of dependents on medical expenses, and examines potential gender-based differences in healthcare spending. This study delves into the intricacies of healthcare accessibility and cost, providing a more profound comprehension of these issues. Additionally, it lays the groundwork for specific interventions that attempt to address gender-based disparities and promote health equity.

Introduction

Medical expenses are a big part of people's life and have an impact on their general well-being and personal finances. As highlighted by studies such as Taloba et al. (2022) and Jödicke et al. (2019), the continuous rise in healthcare costs presents a significant obstacle for developed countries. Taloba et al. (2022) further emphasizes that medical expenses rank among the most common and ongoing costs people face throughout their lifetimes. However, within these general patterns, gender variations play a significant role. Studies have revealed differences in the experiences of men and women, leading to gender-based discrepancies in medical costs. Given these findings, it is important to understand the variables affecting medical expenses in order to alleviate financial hardships and guarantee fair access to healthcare services irrespective of gender.

Background and Questions

My investigation seeks to answer a number of important questions.

1. Is there any significant difference in medical costs billed between genders?
2. Do beneficiaries with more dependents, categorized by gender, tend to have higher medical costs billed?
3. Is there a significant difference in medical costs billed between male and female smokers compared to male and female non-smokers?
4. Does gender have a significant impact on medical charges among individuals with different BMI categories, particularly focusing on overweight individuals ($25.0 \leq \text{BMI} < 30.0$) and obese individuals ($\text{BMI} \geq 30$)

By addressing these questions, I hope to learn more about the underlying causes of gender-based differences in medical insurance costs and how they affect policy choices that enhance the affordability and equity of healthcare.

Data Selection and structure

URL; [Medical Cost Personal Datasets \(kaggle.com\)](https://www.kaggle.com/datasets/medicaid-medical-costs)

The dataset is composed of 1338 instances (rows) and 7 features (columns). These features include the age of the primary beneficiary, the gender of the insurance contractor (categorized as female or male), the body mass index (BMI), the number of children covered by health insurance or dependents, the smoking status, the residential region in the US (categorized as northeast, southeast, southwest, or northwest), and the individual medical costs billed to health insurance. The dataset offers a thorough overview of numerous health-related and demographic variables that could affect medical costs.

Data Cleaning and preparation

The dataset was put through a number of preprocessing stages that were tailored to the particular variable being studied. To protect the integrity of the data, any null or duplicate values were first

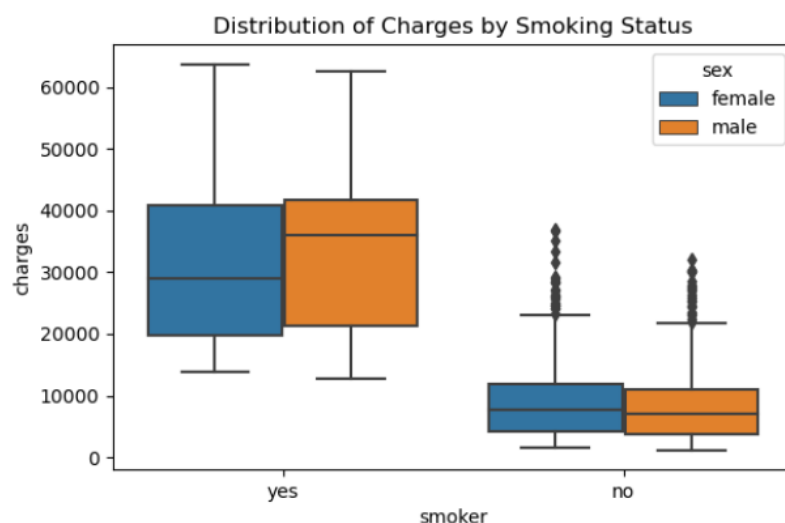
found and then deleted. Following this, preprocessing steps varied according to the type of variable under investigation. Encoding methods like one-hot encoding were used to transform categorical variables into a numerical format that could be analyzed, such as gender and smoking status. Scaling or normalizing procedures were also used for numerical variables in order to normalize their ranges and enable meaningful comparisons, such as medical costs. The goal of these preparation procedures was to guarantee that the dataset was consistent, of high quality, and compatible with the selected analytical methods.

Data Analysis

Gender plays a significant role in determining medical charges, as evidenced by the dataset where the proportion of males stands at 50.49%, slightly higher than females at 49.51%. Investigating gender-based differences in medical costs is made easier by the underlying knowledge this gender distribution offers. In order to examine how gender affects medical costs, a number of variables are addressed.

1. Smoking habits

Medical expenses are greatly impacted by smoking habits, and the correlation between smoking and healthcare costs is further complicated by the gender difference in smoking popularity. The dataset highlights a significant gender disparity in smoking frequency, with a higher percentage of male smokers (23.52%) than female smokers (17.37%). Given this observed gender gap, the analysis explores the relationship between smoking and health care expenses, taking gender differences into consideration. To address this, an estimation of population mean was conducted, preceded by normality and variance tests. The p-values obtained from the Shapiro-Wilk tests and Levene's tests provide strong evidence against the null hypotheses of normality and equality of variance, respectively.



Mann-Whitney U test, a non-parametric tests was therefore used to compare the medical costs between male and female smokers and non-smokers. The resulting p-values for both male smokers versus male non-smokers and female smokers versus female non-smokers are

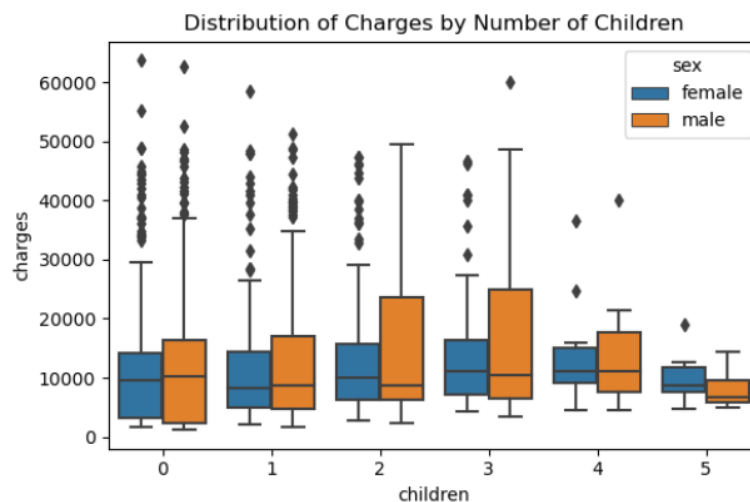
significantly small (far below 0.05), indicating statistically significant differences in medical costs across these groups. Specifically, for both males and females, individuals who smoke tend to have significantly higher medical costs billed compared to those who do not smoke. These findings highlight the financial impact of smoking on medical expenses. It also stress the necessity of focused initiatives to lower smoking rates and enhance public health outcomes.

2. Children

Many people believe that a family's number of children plays a significant role in determining medical costs, with higher healthcare costs generally resulting from having more dependents. The offered code analyzes individuals based on the number of children they have and takes gender differences in medical expenditures into account when examining this relationship.

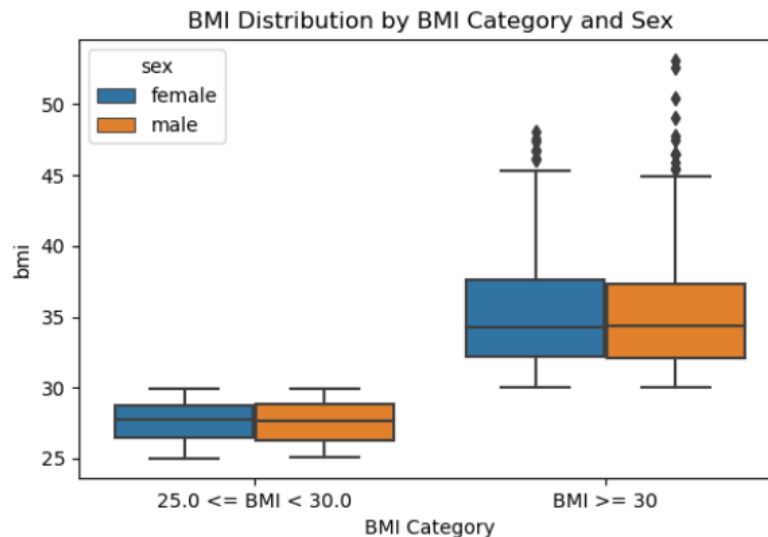
Number of Children	Discrepancy
0	-966.395388
1	-1112.162033
2	-2245.777893
3	-2923.562690
4	155.387208
5	1922.347250

The findings above show a substantial variation in medical expenditures between genders for varying child counts. For example, the negative disparity of -966.395388 indicates that, for childless families, females typically have higher medical costs than males. This discrepancy gets more noticeable as the number of children rises, with women continuously paying more for medical care than men do in all categories. On the other hand, in households with four children, men show a little greater medical expense than women, as shown by a positive discrepancy of 155.387208. These results highlight the complex relationship between the number of children and medical expenses, highlighting the difference in insurance premiums between males and females can be influenced by the number of children. This can also be seen from the boxplots below



3. BMI

A common area of study is the association between Body Mass Index (BMI) and medical expenses because of the possible influence of weight on healthcare costs. Higher BMIs are often associated with an increased chance of developing certain illnesses, which can eventually result in higher medical costs. Through statistical research, this relationship between BMI and medical costs is investigated in further detail. The dataset is filtered to concentrate on people whose BMI falls into two categories: obese ($\text{BMI} \geq 30.0$) and overweight ($25.0 \leq \text{BMI} < 30.0$). The medical costs for these two groups are then compared using a Mann-Whitney U test. With a Mann-Whitney U test statistic of 121626.0 and a p-value of 0.002958017122337365, the test findings show a significant difference in charges between overweight and obese persons. This implies that those who are considered obese typically have substantially higher medical expenses than people who are considered overweight. Gender disparity in this case is not statistically significant in both overweight and obese as seen below.



Discussion and Conclusions

The dataset analysis in this study shed light on the variables affecting medical costs and discrepancies between demographic groups. There were notable gender-based disparities even with a fairly balanced gender distribution, especially when it came to the prevalence of smoking. There was a significant gender difference in smoking habit, with men having a higher percentage of smokers than women. Additionally, it was discovered that smoking status had a major impact on medical costs, with smokers, both male and female, paying more than non-smokers. Furthermore, there was a strong correlation found between the number of dependents and medical expenses, with families with three or more children paying more. The results highlight the significance of taking demographic factors into account in healthcare policy and interventions to promote equitable access to healthcare services and address healthcare cost burdens for vulnerable populations, even though regional disparities in medical charges were not thoroughly investigated.

References

- Choi, Miri. (2018). Medical Insurance Cost. Kaggle. Retrieved from;
<https://www.kaggle.com/datasets/mirichoi0218/insurance>
- Jödicke, A.M., Zellweger, U., Tomka, I.T. et al. Prediction of health care expenditure increase: how does pharmacotherapy contribute?. BMC Health Serv Res 19, 953 (2019).
<https://doi.org/10.1186/s12913-019-4616-x>
- Taloba, A. I., Abd El-Aziz, R. M., Alshanbari, H. M., & El-Bagoury, A. A. H. (2022). Estimation and prediction of hospitalization and medical care costs using regression in machine learning. Journal of Healthcare Engineering, 2022.