Medical Costs Analysis Dashboard

This project showcases an interactive dashboard built in Power BI to analyze medical costs in relation to demographic factors and risk indicators. The goal is to explore patterns in medical charges based on factors such as age, BMI, smoking status, and region. This analysis could support insurance companies in understanding cost drivers among different customer groups.

Table of Contents

Project Overview

This dashboard is designed to provide an in-depth look at factors influencing medical costs. By analyzing different demographic characteristics and health-related factors, the dashboard offers insights into which groups generate higher healthcare costs, including:

- Age groups with the highest average medical costs
- Impact of BMI on charges
- Cost differences between smokers and non-smokers
- Regional variations in medical charges

Dataset

The dataset used in this project is the **Medical Cost Personal Datasets** from Kaggle. It contains information about individual medical charges and various features, including:

- age: The age of the individual
- sex: Gender (male or female)
- bmi: Body mass index, a measure of body fat based on height and weight
- smoker: Smoking status (yes or no)
- region: Region of residence (northeast, northwest, southeast, southwest)
- charges: Medical costs billed to health insurance

This dataset offers a clear and concise view of how personal characteristics can influence medical costs, making it ideal for exploring factors relevant to the health insurance industry.

Data Preparation

Before building the dashboard, the data was processed and cleaned to ensure accurate analysis. Key steps included:

1. Converting data types to appropriate formats (e.g., numeric for charges and bmi, categorical for sex and smoker).
2. Creating age group categories (e.g., 18-25, 26-35, etc.) for easier comparison.
3. Categorizing bmi into ranges: underweight, normal, overweight, and obesity.
4. Assigning states to U.S. regions for mapping purposes.

Dashboard Structure

The dashboard consists of four pages, each focusing on a different aspect of the data:

Age and Gender Analysis

- **Pie Chart**: "Ratio of Sex" – Displays the gender ratio within the dataset.
- **Bar Chart**: "Average Charges for Age Groups" – Compares average medical costs across age groups.
- **Pie Chart**: "Ratio of Age Groups" – Shows the distribution of individuals in each age group.

BMI and Medical Costs

- **Scatter Plot**: "Relationship Between BMI and Medical Costs" – Shows the correlation between BMI and medical costs, with a trend line to highlight the general trend.
- **Bar Chart**: "Average Charges by BMI Category" – Displays the average costs for underweight, normal weight, overweight, and obesity categories.

Smoking Status Impact

- **Pie Chart**: "Ratio of Smokers" – Displays the proportion of smokers and non-smokers.
- **Bar Chart**: "Average Charges for Smokers and Non-Smokers" – Highlights the difference in average costs between smokers and non-smokers.

Regional Analysis

- **Map Chart**: "Average Charges by Region" – Shows the average medical costs across different U.S. regions.

Insights

Key findings from the analysis include:

- **Age Group Costs**: Older age groups tend to have higher average medical charges.
- **BMI Influence**: Higher BMI categories are associated with increased medical costs.
- **Impact of Smoking**: Smokers incur significantly higher medical charges on average compared to non-smokers.
- **Regional Variation**: Medical costs vary by region, which may be due to local healthcare costs or lifestyle differences.

How to Use

1. **Access the Dashboard**:
    - Download the .pbix file and open in Power BI Desktop
2. **Filter Options**:
    - Use the filters for Sex, BMI Category, Region, and Smoker across all pages to explore specific segments of the data.
3. **Interact with Visualizations**:
    - Hover over charts for additional data details, and click on segments to filter other visualizations accordingly.

Future Improvements

To enhance the dashboard, the following improvements could be considered:

- **Additional Data**: Including data on medical conditions, income levels, or healthcare plans for more nuanced analysis.
- **Advanced Calculations**: Adding predictive modeling to forecast costs based on demographic factors.
- **Further Geographic Breakdown**: Using state-level data for more detailed regional insights