**Insurance Data Analysis Report**

**Executive Summary**

This report presents a comprehensive analysis of simulated insurance data relevant to SIMU Insurance. The data encompasses key variables such as policyholder demographics, policy details, risk factors, claims history, and customer feedback. The analysis aims to extract actionable insights that could guide decision-making processes related to policy offerings, risk management, and customer satisfaction strategies.

**Introduction**

The dataset comprises 1,000 entries, each representing an individual policyholder's information. It includes a diverse range of policy types, coverage amounts, and customer demographics. The integrity of the dataset is robust with no missing values or duplicate entries, providing a solid foundation for analysis.

**Data Summary**

The dataset provided a balanced representation of policyholders across different age groups, genders, and locations. The insurance types included Automotive, Home, and Life, with Life insurance being slightly more prevalent. The risk factor was simulated using a beta distribution, reflecting a variety of risk profiles.

**Exploratory Data Analysis**

1. **Age and Gender Distribution**

The age of policyholders is evenly distributed between 18 and 69 years, with a mean age of approximately 43 years. The gender distribution shows a slight predominance of male policyholders.

1. **Policy Details**

Coverage amounts and premiums are right-skewed, indicating that most policies are concentrated around lower to moderate coverage amounts and premiums, with fewer high-value policies.

1. **Risk Factor**

The distribution of risk factors suggests a skew towards lower risk, which is beneficial from a risk management perspective.

1. **Claims Analysis**

Claims count and total claim cost are positively correlated. While most policyholders have zero or a few claims, a few outliers indicate instances of higher claims frequency and cost.

1. **Customer Feedback**

The average customer feedback score is 5.5 out of 10, suggesting room for improvement in customer satisfaction.

**Correlation Analysis**

A key finding is the positive correlation between coverage amount and premium, which is expected. There is no significant correlation between the risk factor and claims cost, which could be due to the limitations of the simulated risk metric.

**Insights and Recommendations**

* Policy Customization: Tailoring policy offerings based on the demographic data and risk profiles could enhance customer satisfaction and retention.
* Risk Management: The lack of a strong correlation between the risk factor and claims suggests the need for a more robust risk assessment model.
* Customer Service: With the average feedback score at the midpoint, there is potential to improve customer service, which could, in turn, lead to higher customer loyalty and better feedback scores.

**Conclusion**

The analysis of the insurance data has highlighted several areas where strategic interventions could enhance policy offerings and customer service. By focusing on data-driven insights, SIMU Insurance can better serve its constituents and strengthen its market position.

**Appendices**

Additional charts, tables, and detailed statistical analysis as presented in the Exploratory Data Analysis section.

A graph of different sizes and colors

Description automatically generated with medium confidenceA graph of a bar graph

Description automatically generated with medium confidence

Figure - Claims Analysis: A combined histogram and box plot showing the distribution and outliers of claims count and total claim cost.

A screenshot of a graph

Description automatically generated

Figure - Correlation Matrix Heatmap: A heatmap showing the correlation coefficients between different numerical variables.