**Splendor Analytics Insurance Data Analysis Project**

**Introduction to Insurance and Project Overview**

**What is Insurance?**

Insurance is a financial product provided by insurance companies to safeguard individuals and entities against potential financial losses. These losses can arise from various unforeseen events such as accidents, natural disasters, theft, or health issues. In exchange for a premium paid by the policyholder, the insurance company commits to covering specific types of risks, thereby providing financial security and peace of mind.

**Project Overview**

Splendor Analytics Insurance is a company specializing in offering a variety of insurance products to its customers. The company's portfolio includes auto insurance, which covers losses or damages related to vehicles. The provided dataset contains anonymized information about the policyholders, their vehicles, and their claim histories. The goal of this project is to analyze this dataset to uncover insights that can help Splendor Analytics Insurance improve its services, optimize premiums, and reduce claim costs.

Splendor Analytics Insurance aims to enhance its auto insurance division by leveraging data analysis to:

1. Identify patterns in claim frequencies and amounts.
2. Understand the demographics and characteristics of high-risk and low-risk policyholders.
3. Optimize premium pricing models based on risk factors.
4. Develop targeted marketing strategies to attract and retain customers.

**Project Brief**

You are tasked with analyzing the provided dataset to help Splendor Analytics Insurance achieve the above objectives. Your analysis should be thorough and provide actionable insights. Use any data analysis tool of your choice (e.g., Python, R, Excel, Tableau, SQL, Power BI etc) to conduct this analysis. The project should be completed within one week.

**Key Questions to Answer**

1. **Claim Frequency and Amount Analysis**
   * What are the average claim frequencies and amounts for different demographic groups (e.g., age, gender, marital status)?
   * Are there any specific vehicle characteristics (e.g., make, model, year) that correlate with higher claim frequencies or amounts?
2. **Risk Assessment**
   * Which factors (e.g., household income, education level, coverage zone) are most indicative of high-risk policyholders?
   * Can we identify any common characteristics among policyholders who make frequent claims?
3. **Premium Optimization**
   * How do current premium amounts relate to the risk profiles of policyholders?
   * Can we recommend adjustments to the premium pricing model to better reflect risk levels and improve profitability?
4. **Customer Segmentation and Marketing**
   * What are the key characteristics of policyholders with low claim frequencies and high household incomes?
   * How can we segment the customer base to identify high-value customers for targeted marketing campaigns?
5. **Demographic Analysis**
   * How does the distribution of policyholders vary across different demographic factors (age, gender, marital status)?
   * Are there any noticeable trends in car usage and ownership among different demographic groups?
6. **Geographical Analysis**
   * How do claim frequencies and amounts vary across different coverage zones?
   * Are there any regional trends or patterns that should be taken into consideration for marketing or risk assessment?
7. **Customer Behavior Insights**
   * Are there any trends or patterns in the behavior of policyholders who have children driving?
   * How does the presence of children driving affect the frequency and number of claims?

**Data Dictionary**

| **Column Name** | **Description** |
| --- | --- |
| ID | Unique identifier for the policyholder |
| Birthdate | Birthdate of the policyholder |
| marital\_status | Marital status of the policyholder (Single, Married, etc.) |
| car\_use | Purpose of car use (Private, Commercial) |
| Gender | Gender of the policyholder (Male, Female) |
| kids\_driving | Number of kids of the policyholder who are driving |
| Parent | Whether the policyholder is a parent (Yes, No) |
| Education | Education level of the policyholder (High School, Bachelors, etc.) |
| car\_make | Make of the policyholder's car |
| car\_model | Model of the policyholder's car |
| car\_color | Color of the policyholder's car |
| car\_year | Year of the policyholder's car |
| claim\_freq | Frequency of claims made by the policyholder |
| coverage\_zone | Area of coverage (Highly Urban, Urban, Suburban, Rural, Highly Rural) |
| claim\_amt | Amount claimed by the policyholder |
| household\_income | Household income of the policyholder |

**Deliverables**

1. Cleaned dataset.
2. Dashboard/Report
3. Detailed report answering the key questions.
4. Presentation slides summarizing key insights and recommendations.

**Submission Guidelines:**

* **Submit on Twitter or LinkedIn by tagging @iam\_uchenna or @splendanalytics**

This project will provide valuable insights to Splendor Analytics Insurance, helping them to optimize their auto insurance operations and enhance customer satisfaction. Good luck!

**Identifying High-Value Segments**

Once you've segmented your customer base, analyze each segment to identify those with the highest potential value. Look for segments with:

1. High average policy premium
2. Low claim frequency
3. High policy count
4. High renewal rates
5. High CLTV

**Targeted Marketing Campaigns**

Tailor marketing campaigns to each high-value segment:

* **Premium Customers:** Offer exclusive benefits, loyalty programs, or personalized services.
* **Low Claim Customers:** Promote preventive measures, discounts, or bundled products.
* **Multi-Policy Holders:** Cross-sell additional products or services.
* **High LTV Customers:** Focus on retention, upselling, and customer satisfaction.