Insurance Claims Analysis Dashboard

An interactive dashboard analyzing insurance claims data to identify high-risk policyholders, optimize pricing strategies, and improve profitability through data-driven risk segmentation.

Overview

This project implements a comprehensive risk analysis solution for insurance claims data, providing actionable insights into policyholder behavior, claims patterns, and geographic risk concentrations. The dashboard enables data-driven decision making for underwriters, claims managers, and business strategists.

Problem Statement

Insurance companies struggle with:

- Identifying high-risk policyholders before significant losses occur
- Understanding geographic and demographic patterns in claims behavior
- Optimizing premium pricing based on actual risk exposure
- Allocating resources efficiently to manage claims across different segments

Dataset

Synthetic Insurance Claims Data containing:

- 10,000 policy records with 368,000 total claims
- Key Features: Policy type, region, age groups, marital status, claims frequency, claim amounts
- Demographic Data: Age brackets, marital status, geographic location (Urban/Suburban/Rural)

- Claims Data: Claims frequency, severity levels (Low/Medium/High), adjustment amounts

Tools & Technology

- Data Analysis: SQL, DAX, Python (Pandas)
- Visualization: Power BI / Tableau
- Data Processing: Excel, Data Modeling
- Version Control: Git & GitHub

Methods

1. **Data Cleaning & Validation**
- Handled missing values and data consistency checks
- Standardized categorical variables (region, marital status, age groups)
2. **Risk Segmentation**

sql

```
CASE
    WHEN Claims_Frequency > 2 THEN 'High'
    WHEN Claims_Frequency > 1 THEN 'Medium'
    ELSE 'Low'
END as Risk_Segment
```

3. **Geographic Analysis**
- Regional concentration analysis (Urban: 49.83%, Suburban: 28.21%, Rural: 21.96%)
- Claims frequency patterns across different population densities
4. **Demographic Profiling**
- Age band analysis (clear correlation between age and claims frequency)

- Marital status impact on risk behavior
- Lead source effectiveness (Agent, Online, Referral)

Key Insights

Risk Concentration

- Urban areas drive 49.83% of all claims despite having similar policy counts
- Policyholders aged 60+ show significantly higher claims frequency
- High-severity claims (4K records) require focused management despite being fewer in number

Business Impact

- Average claim frequency: 0.50 claims per policyholder
- Average claim amount: $73.97 per claim
- Balanced portfolio distribution across geographic regions

Demographic Patterns

- Married individuals represent the largest policyholder segment
- Clear age-based risk progression from young to senior policyholders
- Urban centers show concentrated high-frequency claims

Dashboard Output

The interactive Power BI dashboard includes:

Key Metrics Display

- Total Policies: 10,000
- Total Claims: 368,000
- Average Claim Frequency: 0.50

- Average Claim Amount: $73.97

Visual Components

- Claims Severity Distribution (Low/Medium/High)
- Geographic Risk Heatmap by region
- Age vs Claims Severity correlation matrix
- Marital Status Analysis by segment
- Lead Source Effectiveness chart

Interactive Features

- Policy type filters
- Demographic segmentation toggles
- Regional focus selectors
- Risk category highlighting

Results & Conclusion

Business Outcomes

1. Risk Identification: Successfully identified high-risk segments (urban seniors, high-frequency claimers)
2. Pricing Optimization: Provided data foundation for risk-based premium adjustments
3. Resource Allocation: Enabled targeted claims management in high-concentration areas
4. Strategic Insights: Revealed demographic patterns for product development

Quantitative Results

- 368,000 claims analyzed across 10,000 policies
- 72K+ low-severity claims identified for automated processing
- 4K high-severity claims flagged for specialized handling

- 49.83% urban concentration highlighting geographic risk focus areas

Conclusion

This dashboard demonstrates how data analytics can transform raw insurance claims data into actionable business intelligence. By identifying clear risk patterns and demographic correlations, insurance companies can make informed decisions about pricing, underwriting, and claims management. The solution provides a scalable framework for ongoing risk monitoring and strategic planning in the insurance industry.

The project successfully addresses the core challenges of risk segmentation and profitability optimization while providing an intuitive interface for business users to explore data and derive insights.