

# Swedish Motor Insurance Case Study

---

## Introduction

This case study analyzes regional claim performance using Swedish Motor Insurance data. The goal is to understand claim frequency and severity trends across zones, identify risk segments, and recommend zone-based premium adjustments based on statistical analysis. This approach supports data-driven pricing strategy, minimizes underwriting risk, and ensures profitability across geographic markets.

## Key Business Questions

- What zones have the highest average cost per policy?
- Where are claims most frequent?
- Which zones have high frequency and high severity?
- What premium multipliers are appropriate based on risk segmentation?
- How can we visualize zone-level performance to support pricing decisions?

## Data Overview & Preparation

The dataset includes zone-level claim data, containing fields for exposure (policy-years), total claims, total payments, claim frequency, claim severity, and suggested premium multipliers. All data was pre-aggregated in SAS and exported into Tableau as `zone\_risk\_metrics.csv` for visualization.

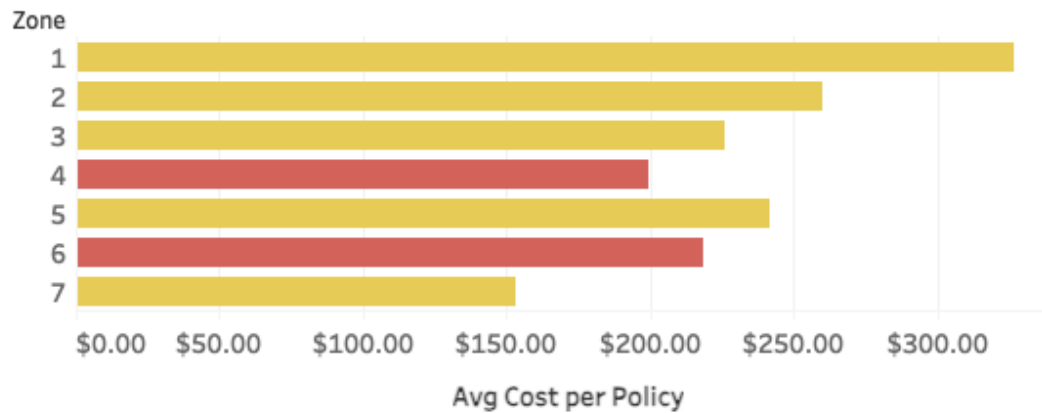
## Analysis and Insights

### 1. Average Cost per Policy by Zone

This metric captures the average financial exposure per policy. Zone 1 has the highest average cost per policy at over \$320, while Zone 7 is the lowest at around \$150. These

figures suggest wide cost variation and pricing imbalance.

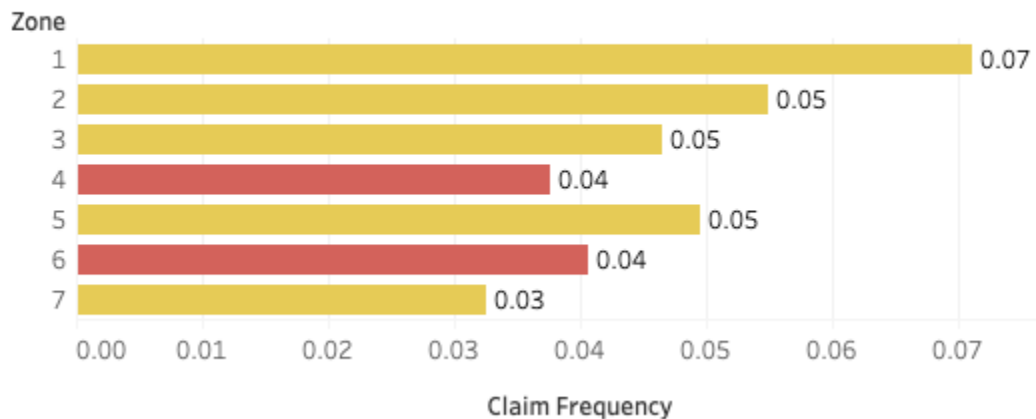
### Avg Cost per Policy by Zone



### 2. Claim Frequency by Zone

Zone 1 also exhibits the highest claim frequency ( $\sim 0.07$ ), indicating higher loss likelihood per policy-year. Zones 4 and 6, while moderate in cost, have relatively elevated claim frequencies, flagging behavioral risk factors.

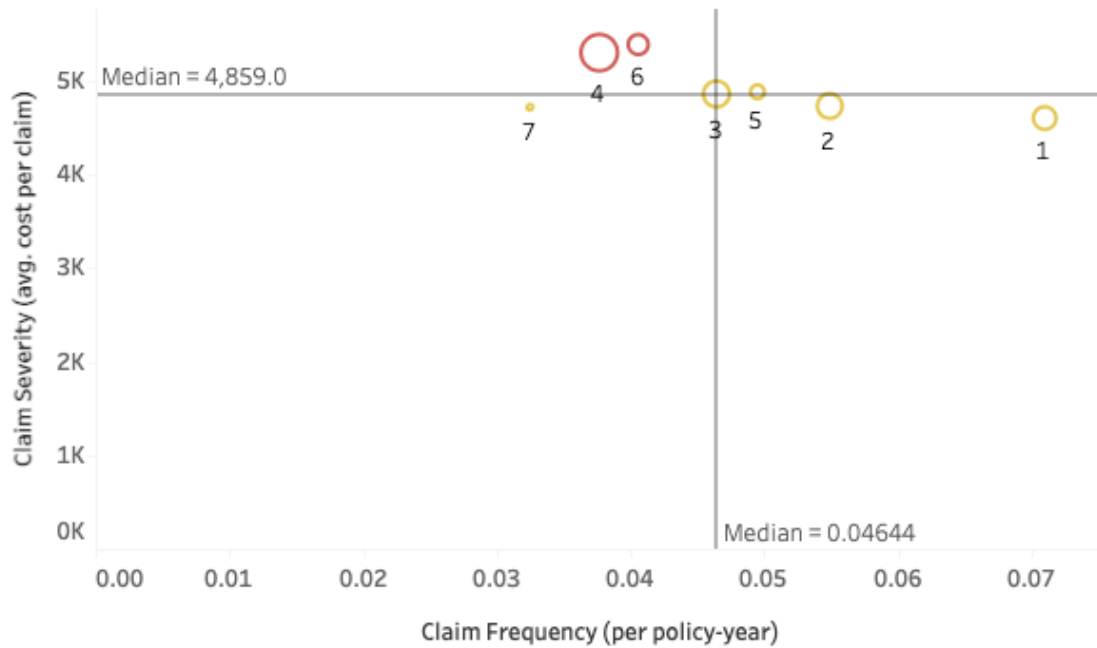
### Claim Frequency by Zone



### 3. Frequency vs. Severity Risk Map

Zones are plotted in quadrants to reflect whether they have high/low claim frequency and severity. Zones 4 and 6 fall into the high-frequency and high-severity quadrant, indicating urgent pricing or underwriting action.

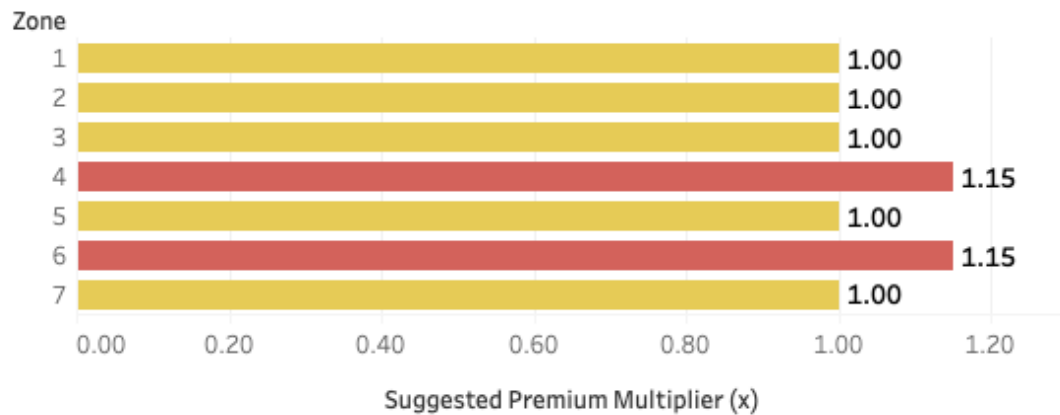
## Frequency vs Severity (Risk)



### 4. Suggested Premium Multipliers

Zones 4 and 6 are assigned a premium multiplier of 1.15 based on their above-average frequency and severity. Other zones remain at the baseline 1.00. These adjustments provide fair, risk-aligned pricing.

### Suggested Premium Multiplier



## Recommendations

Based on the insights from the data, we recommend the following actions:

1. Increase premiums in Zones 4 and 6 by 15% to align with their elevated claim cost and risk.
2. Monitor claim activity in Zone 1, which shows high average cost and frequency.
3. Investigate Zones 5 and 3 for high claim frequency despite moderate severity — this may reflect environmental or fraud exposure.
4. Continue using this data-driven model to reassess pricing on a quarterly basis as more claim data becomes available.

## Limitations & Opportunities for Further Analysis

Although this analysis provides actionable insights based on the available zone-level claim data, several key features were not present in the dataset. Incorporating additional variables could significantly enhance model precision, segmentation quality, and pricing strategy:

- **Driver Demographics:**  
Age, gender, driving history, and tenure could help isolate behavioral risk factors within each zone.
- **Vehicle Data:**  
Vehicle type, usage (personal vs. commercial), and value could refine both frequency and severity risk assessments.
- **Geographic Specificity:**  
Zones are aggregated regions. More granular data (postal code or municipality) would allow for better micro-segmentation.
- **Temporal Data:**  
Including month or season would reveal seasonal trends in claim frequency or severity (e.g., winter spikes).
- **External Risk Factors:**  
Traffic density, crime rates, and weather conditions could explain variations in loss performance that aren't captured by the current variables.

## Conclusion

This analysis successfully segments zone-level insurance risk using claim frequency, severity, and cost data. By quantifying key performance metrics and translating them into actionable premium guidance, we've laid the foundation for a fair and financially sound pricing strategy. Zones 4 and 6 clearly emerged as high-risk areas warranting pricing adjustments, while other zones showed moderate to low risk profiles.

However, the analysis is limited by the granularity and scope of the available dataset. The absence of driver-level, vehicle-level, and time-based features means we are only able to generalize risk at the regional level. While this approach is valid for high-level pricing decisions, integrating richer contextual variables, such as vehicle type, driving behavior, or seasonal trends, would enable a more precise and individualized pricing framework.

Despite these constraints, the current model provides valuable strategic direction and establishes a repeatable methodology for future portfolio reviews. As more detailed data becomes available, this framework can be expanded to drive deeper insights and enhance the insurer's competitive position through smarter, risk-aligned pricing.