Insurance Loss Analysis: Key Vehicle Risk Factors

This analysis identifies vehicle attributes that are most associated with higher normalized insurance losses. Important factors include engine size, horsepower, curb weight, and vehicle price, as these features are closely related to vehicle performance and the cost of repairs. Cars from manufacturers such as Porsche, BMW, and Peugeot show higher insurance losses on average, whereas brands like Honda and Toyota typically show lower losses.

Body style also influences losses. Convertibles and hatchbacks tend to have higher losses, which may be due to their appeal to younger or more performance-driven drivers. Drive wheel configuration affects losses as well. Rear-wheel and four-wheel drive vehicles generally experience higher losses than front-wheel drive models. Vehicles with turbocharged or high-displacement engines have increased losses, mainly due to performance use and higher repair complexity.

Fuel type and the number of doors have limited impact on insurance losses. However, fuel efficiency does show a modest inverse relationship with losses, suggesting that more efficient, economy-class cars are generally involved in fewer or less costly claims. Vehicle dimensions, including width and weight, also show a moderate correlation with losses. Engine aspiration and drive system types offer additional predictive value.

Several nonlinear patterns were identified. For example, horsepower correlates with insurance losses up to a point, after which the relationship levels off. The distribution of losses is skewed, with most vehicles falling in the low to moderate range but a small number of high-loss vehicles significantly affecting the overall average. These outliers are often luxury or performance cars with extreme values in price or engine power.

Cluster analysis helped identify groups of vehicles with similar risk profiles. For instance, one group includes affordable, fuel-efficient cars with low losses, while another includes high-end vehicles with higher risk. Insurers can use these groupings to better tailor premiums and coverage options. It is also critical to manage missing data and outliers during modeling, as these can distort results.

In conclusion, insurers should consider a range of vehicle features when setting premiums. These include performance-related specifications, brand, body type, aspiration method, and drive system. Together, these factors offer a strong basis for developing accurate and fair insurance pricing models.

RECOMMENDATIONS

The strongest predictors of high insurance losses are performance-related features such as engine size, horsepower, and car price. Instead of using a flat-rate premium model, insurers should implement a tiered premium system that adjusts based on these key vehicle characteristics. This means higher premiums for high-powered, luxury, and performance vehicles (e.g., BMW, Porsche) and lower premiums for economy-class models (e.g., Honda, Toyota). Such a model will more accurately reflect the true risk, reduce underwriting losses, and encourage safer vehicle choices among customers.

Several body styles (convertibles, hatchbacks), aspiration types (turbocharged), and drive configurations (rear-wheel drive) are associated with more aggressive driving and higher loss outcomes. These vehicle types often correlate with younger, riskier drivers or high-speed usage. The company should consider incorporating vehicle usage patterns—derived from make, model, and build—into a behavioral risk index. Combining this with telematics or customer driving history, if available, can offer a more dynamic and personalized premium structure, enhancing fairness and profitability.

Car make and model are strong indicators of average loss levels. Developing a dynamic, annually updated risk scoring system for each brand and model would allow insurers to better manage pricing strategies. For example, brands with a history of high claims (e.g., Porsche, Peugeot) should trigger increased scrutiny and higher premiums, while safer brands may be rewarded with lower rates. This strategy will align premiums with actual risk and prevent underpricing of high-risk vehicles.

While vehicle attributes are essential, the absence of driver-related variables—such as age, location, and usage purpose—limits the depth of risk assessment. High insurance losses may not stem solely from the car's characteristics but from who drives the vehicle and how it's used. For example, a sports car driven by an experienced, risk-averse driver may generate fewer claims than a modest vehicle driven aggressively by an inexperienced driver. Insurers should integrate demographic and contextual data into their risk models, potentially through external datasets, surveys, or telematics. This holistic approach will provide a more accurate risk picture and lead to better-aligned premiums and policy conditions.

Clusters of vehicles with high insurance losses were identified using combinations of features like high horsepower, turbo aspiration, luxury brand, and specific body styles. The company should actively segment these high-risk clusters and track them separately. Special monitoring can include more stringent underwriting checks, higher deductible requirements, or even customized safety incentive programs. This targeted approach balances risk control with customer service, ensuring that the highest-risk profiles are managed proactively without disrupting the broader client base.