Portfolio Risk Assessment and Predictive Modeling Insights for HELP Insurance

Dream Actuaries

Denis Li Ping Kam, Euichan "Daniel" Kim, Jeremy Tan, Patuan Purba

Table of Contents

1. Executive Summary

This report evaluates HELP's student insurance portfolio through data analysis, predictive modeling, regional segmentation, and scenario testing to identify key risk drivers and strengthen pricing and underwriting decisions.

The portfolio includes 10,000 students and over 40,000 coverage lines, with an overall claim frequency of 4.54%. A logistic regression model (AUC = 0.76) shows that residence-related factors are the strongest predictors of claim activity. Greek housing increases claim odds by 2.26 times, Personal Property coverage significantly increases risk, and each rise in risk tier increases claim odds by 1.15 times. Academic and demographic factors show minimal influence.

Regional analysis indicates that claim frequency and severity rise with distance from campus. While on-campus students make up nearly 90% of the portfolio, those living farther away have higher average and median claim amounts, while on-campus students exhibit heavier-tailed loss risk.

Scenario testing highlights heightened exposure in non-sprinklered residences, where stressed fire events substantially increase claim severity. This identifies a critical concentration of catastrophe risk.

Based on these findings, we recommend that HELP refine its risk tiers at the coverage level, limit insurance approvals for non-sprinklered buildings, expand low-severity coverage offerings, invest in advanced actuarial analytics, and enhance client-level data collection. These steps will support more accurate pricing, reduced exposure, and improved portfolio stability.
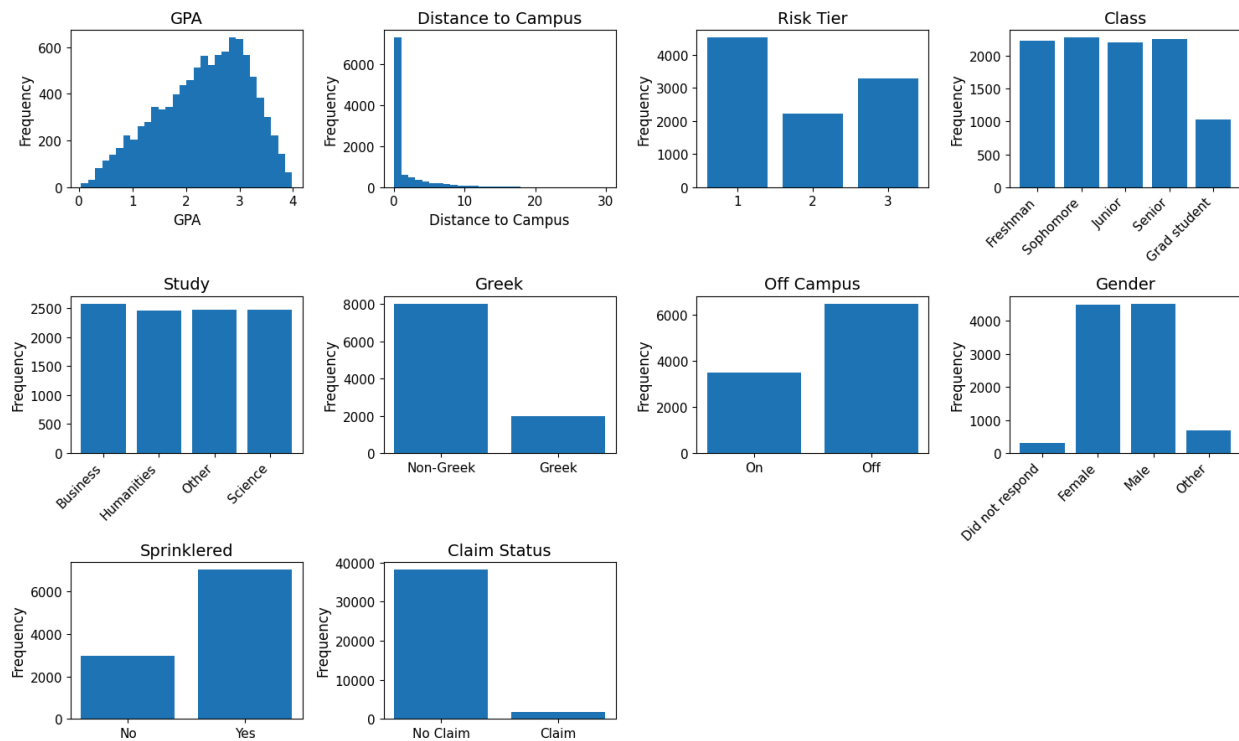
2. Key Findings & Portfolio Insights

Figure 1: Portfolio Distribution Overview

Figure 1 provides an overview of HELP's student insurance portfolio. Each record reflects one coverage line for an insured student, and the portfolio includes 10,000 students and more than 40,000 total coverage entries. Coverage offerings include Liability, Personal Property, Additional Living Expense, and Guest Medical.

GPA values are concentrated in the mid-to-high range, with only a small number of students falling into the lowest academic bands. Distance patterns show that most exposures are within the first kilometre of campus, and the number of students declines sharply as distance increases. This indicates that the portfolio is heavily concentrated in near-campus housing.

Risk Tier 1 is the largest underwriting segment with 4,512 students, followed by Tier 3 and Tier 2. Class year is well balanced, with similar representation across Freshman, Sophomore, Junior, and Senior groups and an additional 1,026 graduate students. Academic programs are evenly distributed across Business, Humanities, Science, and Other fields of study.

The Greek residential classification contains 1,965 students, while 8,035 are classified as Non-Greek. A total of 6,505 students live off campus compared to 3,495 on campus, and most insured reside in sprinkler-protected buildings. Gender distribution is balanced across male and female students with smaller groups identifying as other or not responding.

3

Across all coverage lines, there are 1,819 claims and 38,252 non-claim records, resulting in an overall claim frequency of 4.54 percent. Although relatively low, this level of activity provides meaningful data for identifying risk concentrations and supports deeper analysis in later sections of this report.
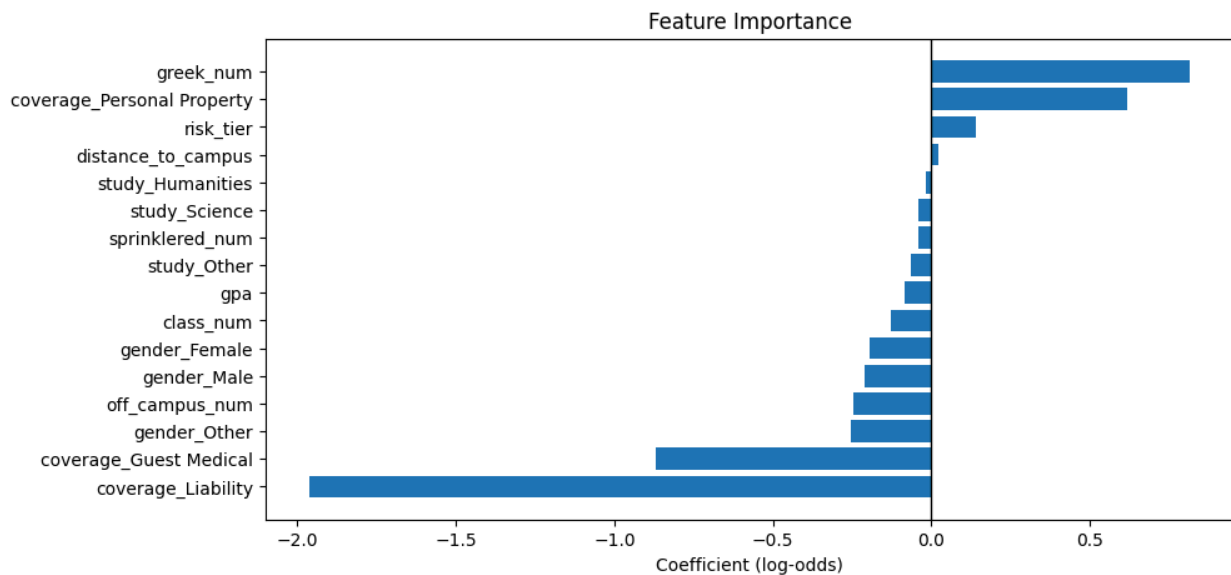
3. Predictive Model Insights

Figure 2. Logistic Regression Coefficients

To better understand the drivers of claim activity, we developed a logistic regression model using HELP's full coverage-level dataset. This method provides transparent, interpretable results consistent with actuarial modeling practices. The model achieved an AUC of 0.76, indicating reasonable discriminatory power, and an overall accuracy of 0.63, which is appropriate given the imbalance between claim and non-claim records.
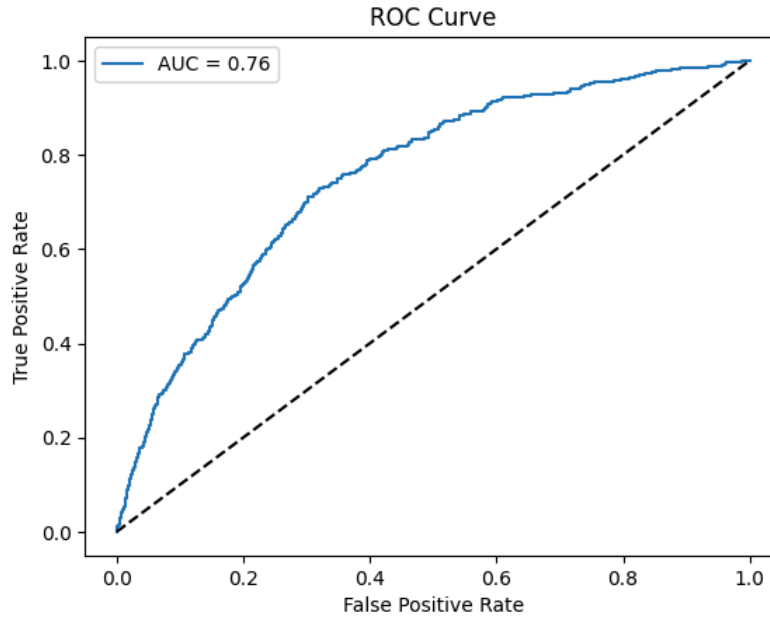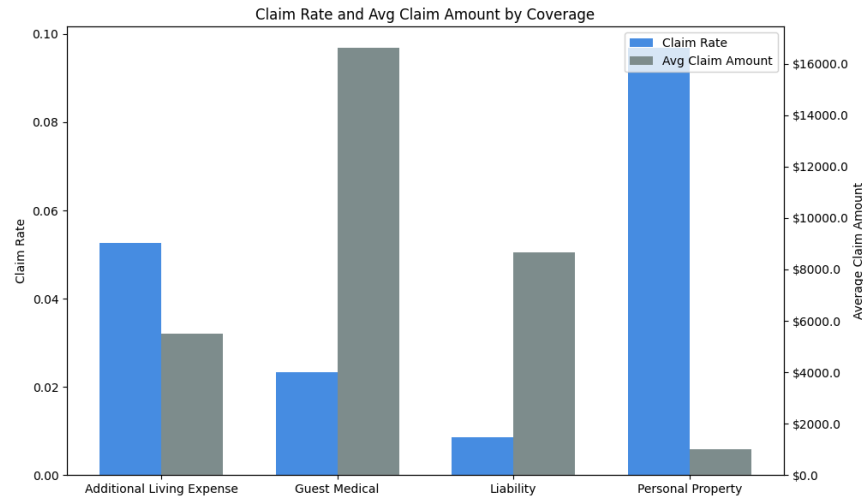
Figure 3. ROC Curve (AUC = 0.76)

Figure 2 presents the feature importance results. The strongest predictor is the Greek residential classification, which increases the odds of a claim by approximately 2.26 times compared to Non-Greek residences. Coverage type is also highly influential. Personal Property coverage raises the odds of a claim by 1.86 times, while Liability and Guest Medical significantly reduce claim likelihood. Risk tier remains a meaningful structural factor: each increase in tier corresponds to a 1.15-fold increase in predicted claim odds.

Other characteristics show much smaller effects. GPA, class year, gender, program of study, and sprinkler protection display minimal influence on claim probability, indicating that student demographics contribute little to claim outcomes. In contrast, variables tied to living environment and coverage choice consistently show the strongest relationships.

Overall, the predictive model reinforces that claim behaviour within HELP's portfolio is shaped primarily by residence type, location characteristics, and coverage selection, rather than by demographic or academic attributes. These insights support more targeted pricing and underwriting strategies explored later in this report.

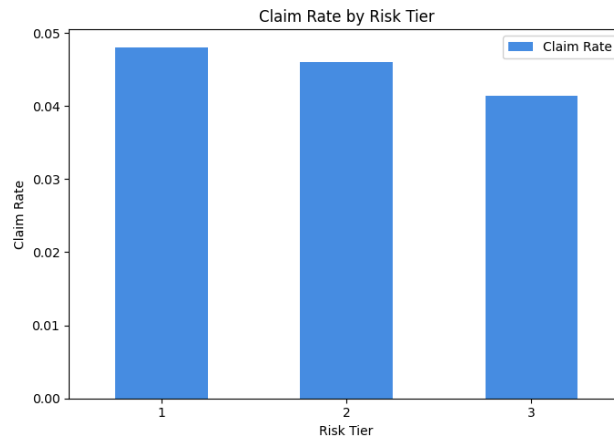# 4. Claim Behaviour and Frequency Patterns

Claim Rate by Coverage

Coverage Type	Historical Claim Rate (Frequency)	Average Claim Amount (Severity)
Liability	0.85%	$8,675.42
Guest Medical	2.32%	$16,611.81
Additional Living Expense	5.27%	$5,513.86
Personal Property	9.68%	$1,023.08
Average Claim Rate	**4.54%**	**$4,681**

A review of claim frequency and severity across HELP's four coverage types highlights clear differences in risk performance. Personal Property shows the highest claim rate at 9.68 percent, which is consistent with the frequency of small but common loss events such as theft or accidental damage. Additional Living Expense follows at 5.27 percent. Guest Medical has a modest claim rate of 2.32 percent, while Liability exhibits the lowest frequency at 0.85 percent.

Although Liability and Guest Medical claims occur less often, their average claim amounts are significantly higher. Guest Medical has the highest severity at approximately 16,600 dollars per claim, reflecting the high cost of medical expenses even when incidents are infrequent. Liability claims average roughly 8,675 dollars, which aligns with the potential for higher-cost third-party exposures. In contrast, Personal Property claims average approximately 1,023 dollars, and Additional Living Expense averages about 5,514 dollars.

On aggregate, the portfolio's overall claim rate across all coverage lines is 4.54 percent, based on 1,819 claims out of 40,071 total coverage entries. The average claim amount across all coverages is approximately 4,681 dollars. These patterns indicate that HELP's portfolio combines a large volume of low-severity property claims with a smaller number of higher-severity liability and medical claims. Understanding this balance between frequency and severity helps clarify where pricing adjustments or product refinements may offer the greatest impact.
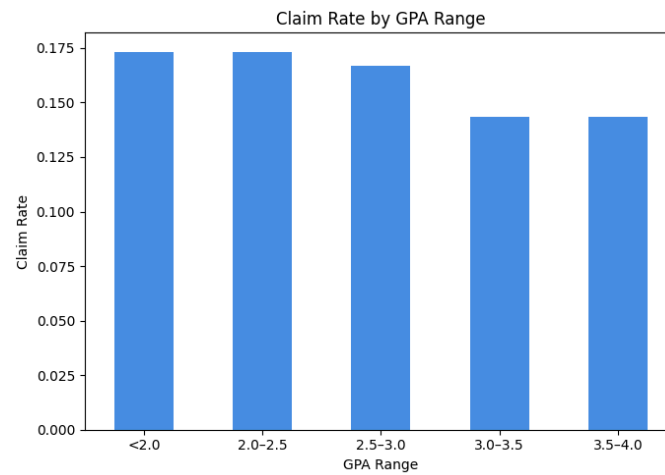
Claim Rate by Risk Tier

The observed claim rates across HELP's underwriting tiers show relatively small differences in overall frequency. Tier 1 exhibits a claim rate of 4.81 percent, followed by Tier 2 at 4.60 percent and Tier 3 at 4.13 percent. Although Tier 1 is slightly higher, the three tiers remain close in magnitude, indicating that claim frequency in this portfolio does not vary sharply by tier.

This pattern suggests that claim behaviour is distributed fairly evenly across the tier structure. The differences that do exist may reflect variations in coverage mix, residence type, or exposure characteristics within each tier rather than tier alone. Since tier remains a significant predictor in the logistic model, its influence appears to be captured in the interaction of multiple factors rather than frequency alone.
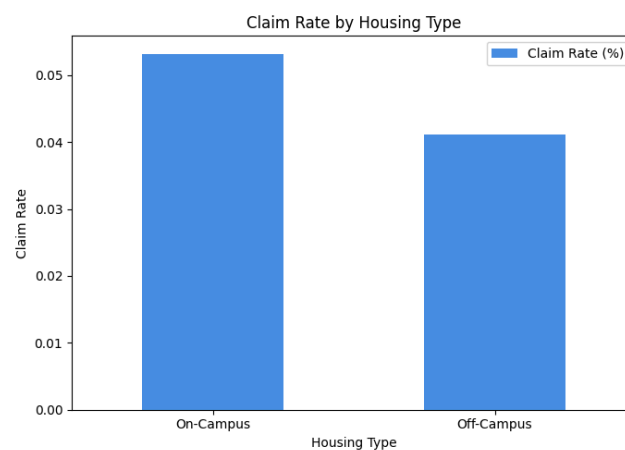
Claim by GPA Range

Claim frequency shows a gradual decline as GPA increases. Students with GPAs below 2.0 have a claim rate of 17.33 percent, which is nearly identical to the 17.30 percent observed in the 2.0 to 2.5 range. Claim rates begin to decrease in the mid-range academic bands. Students with GPAs between 2.5 and 3.0 have a claim rate of 16.70 percent, while those in the 3.0 to 3.5 and 3.5 to 4.0 ranges show lower rates of 14.36 percent and 14.36 percent respectively.

While the differences are not extreme, the pattern indicates a modest downward trend in claim frequency among higher-GPA students. This suggests that academic performance may correlate slightly with lower claim activity, although the effect is limited compared to other structural factors examined in the portfolio.

Claim Rate by Campus Housing

Claim frequency differs modestly between on-campus and off-campus students. On-campus students have a claim rate of 18.66 percent, while off-campus students show a lower rate of 15.27 percent. Although the difference is not large, it suggests that on-campus residences experience slightly higher claim activity relative to off-campus housing.
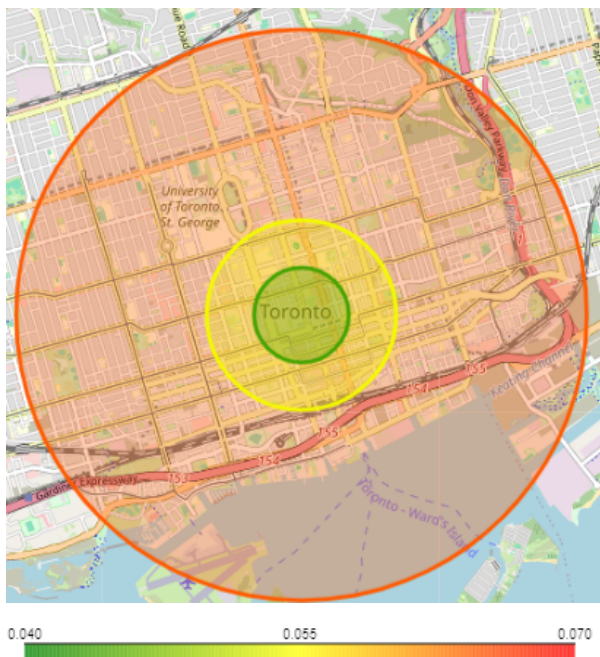
This pattern may reflect differences in living arrangements, property characteristics, or coverage selection across the two groups. However, the effect size remains moderate compared to other structural drivers in the portfolio.

5. Region Analysis

We divide our clients into 3 regions based on the distance from campus:
- "On-Campus": 0 - 5 km
- "Near Campus": 5 - 10 km
- "Far from Campus": 10 - 30 km

Claim Rate

Region	Claim Rate
On-Campus (0-5 km)	4.40%
Near Campus (5-10 km)	5.43%
Far from Campus (10-30 km)	6.47%

The table above presents the claim rate per coverage by region. The claim rate rises as the distance from campus increases. Clients residing on campus have a claim rate of 4.40%, those living near campus exhibit a rate of 5.43%, and those more than 10 kilometers away show the highest claim rate in the portfolio.

Portfolio Composition by Region

0.0   0.5   1.0
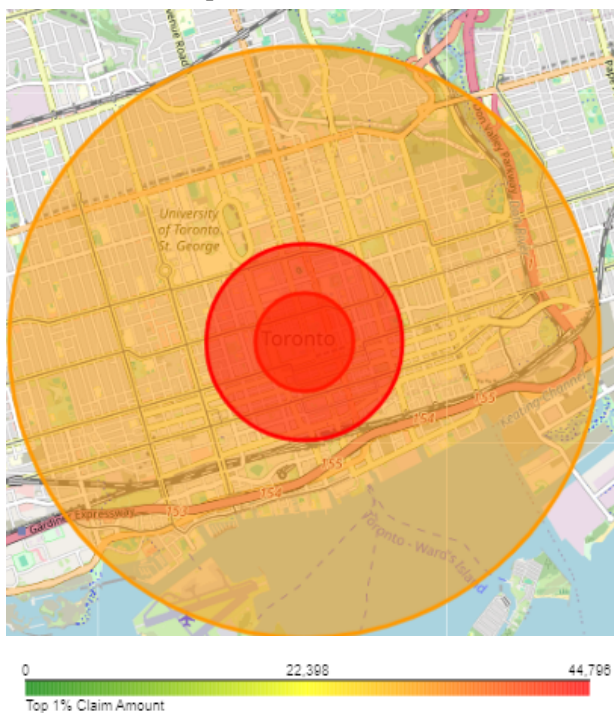Percentage of Insurance

Region	Portfolio composition
On-Campus (0-5 km)	89.76%
Near Campus (5-10 km)	7.30%
Far from Campus (10-30 km)	2.93%

The table above presents the portfolio composition of HELP divided by region. Insured clients residing on campus account for 89.76% of the portfolio. Those living 5 to 10 kilometers from campus represent 7.30%, while clients residing more than 10 kilometers away comprise only 2.93% of the portfolio.

Top 1% Claim Amount

0   22,398   44,796
Top 1% Claim Amount

Region	Top 1% Payout
On-Campus (0-5 km)	$42,551
Near Campus (5-10 km)	$44,796
Far from Campus (10-30 km)	$31,360

The table above presents the top 1% of claim amounts by region. The results indicate that the top claim amounts are highest on clients near campus, followed by clients on-campus, and lower on clients far from campus.
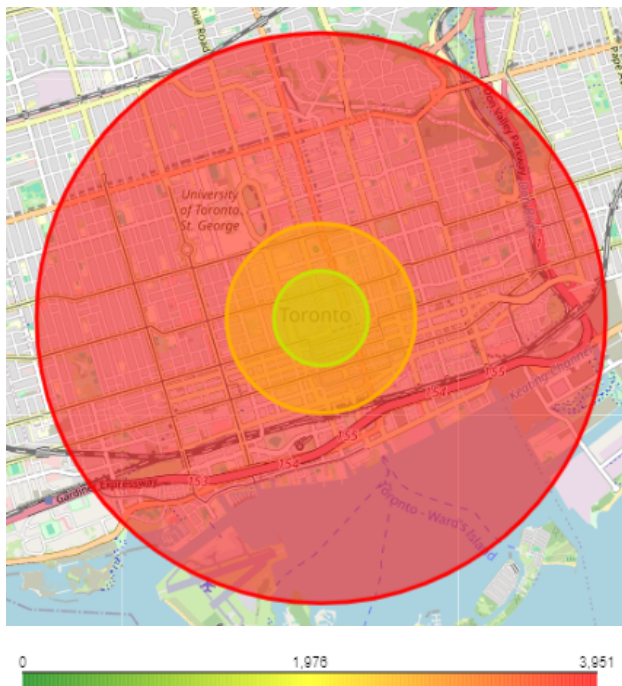
Average Claim Amount

Region	Mean Claim Amount/Coverage
On-Campus (0-5 km)	$192.03
Near Campus (5-10 km)	$370.83
Far from Campus (10-30 km)	$444.56

The table above presents the mean claim amount per coverage. The average claim amount rises as the client region strays further from campus. Clients on-campus have an average claim amount of $192.03/coverage. Clients near campus have an average claim amount of $370.83/coverage. Clients far from campus have an average claim amount of $444.56/coverage.

Median Claim Amount

Region	Median Claim Amount
On-Campus (0-5 km)	$1,535.30
Near Campus (5-10 km)	$2,582.30
Far from Campus (10-30 km)	$3,951.37

Given that a claim has been filed, the table above shows the median claim amounts. The data indicates that median claim values increase as the distance from campus increases.
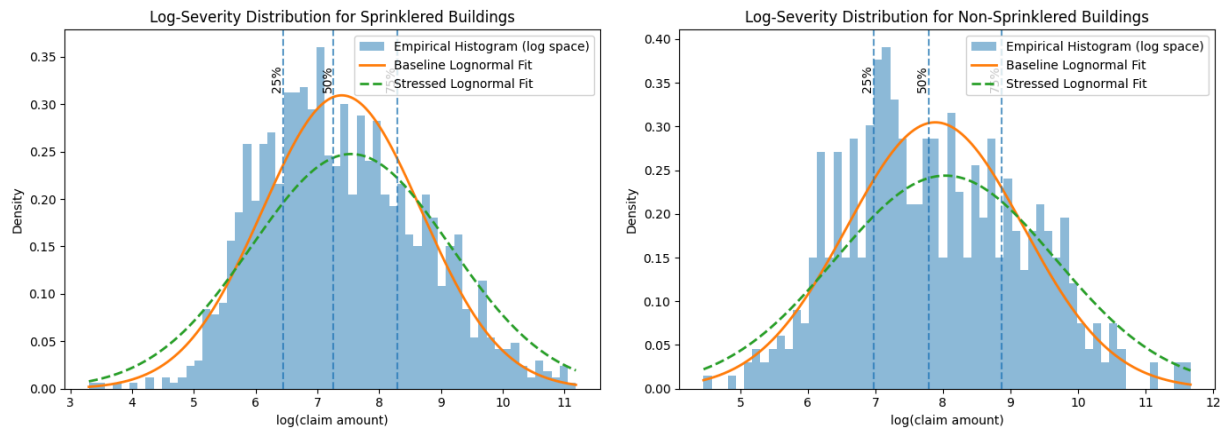
The top 1% of claim amounts is highest for the on-campus group, whereas both the median and average claim amounts are higher for individuals living far from campus. This suggests that the on-campus group exhibits a heavier tail in claim amount distribution, indicating a higher severity risk for

11

this region. Given this pattern, the insurer may need to reevaluate the risk tiers assigned to policyholders residing on campus.

6. Scenario and Stress Testing Considerations

To evaluate the portfolio's vulnerability to low frequency but high severity events, scenario and stress testing of key exposure characteristics present in the portfolio is required. Stress scenarios were constructed around variables in the dataset which are most likely to be catastrophe drivers, which includes Sprinkler Protection.

A primary scenario considered was the event of a severe fire happening in student accomodations. In this case, sprinkler protection could provide a better overview of the exposure and loss considerations in this type of event.

Sprinkler Protection	Expected Claim Amount	Scenario Claim Amount
Sprinklered Buildings	$3,740	$6,938
Non-Sprinklered Buildings	$6,272	$11,798

From the above, the mean claim amount for students living in sprinkler-protected residences is $3,740, while for non sprinkler-protected residences is $6,272. These numbers are expected because sprinklers provide much better protection in case of fires, limiting property losses and any medical coverage claims.

Suppose a hypothetical fire event increases tail risk and mean claim amounts (severity) by 25% and 10% respectively. From the above stressed distribution, the mean claim amount for sprinkler-protected residences increases to $6,938, while for non sprinkler-protected residences claim amounts increase to $11,798. This represents an increase of around 85% and 90% respectively, with

non-sprinklered residences more affected because of the increased tail-risk present due to higher possible property losses. Hence, this hypothetical scenario provides insights into where the business is most exposed in case of a catastrophe happening. A review of exposure concentration for the non-sprinklered group is recommended in order to more carefully allocate insurance coverages and minimize exposure.

7. Product or Coverage Opportunity Evaluation

From the above analysis of the portfolio, there are several product or coverage opportunities HELP could implement that would increase premiums while also minimizing loss exposure. Currently, liability coverage and guest medical coverage have the lowest frequency and severity of the other types of coverage. HELP could consider increasing insurance coverage allocation to these types of coverage or develop new product types that include medical and liability coverage. This would increase premiums collected while keeping exposure low in the portfolio.

Another coverage opportunity that HELP could implement would be reevaluating coverage claims contracts by increasing legal spending and improving underwriting standards. These measures would help decrease claim amounts paid out, thus decreasing severity in the portfolio, particularly in personal property claims, since they have the highest frequency and severity out of the types of coverages.

8. Recommendations: What HELP should do next

Based on our model's predictions and industry research, we recommend assigning tailored risk ratings at the individual coverage level rather than applying a uniform rating across each client[1]. This approach enables HELP to price policies and evaluate insurance risk more dynamically and accurately, reflecting both coverage-specific and client-specific factors.

Based on our scenario testing, we recommend that HELP rebalance its portfolio by limiting insurance approvals for individuals residing in non-sprinklered buildings. This adjustment would help reduce overall claim risk and improve portfolio stability against a fire.

We also propose integrating advanced actuarial data analytics into HELP's operations. Strengthening analytics capabilities can support data-driven decision-making that follow actuarial standards, improve operational efficiency, and enhance pricing and revenue optimization while minimizing loss exposure. Implementing this vision will require investment in modern data tools including cloud-based storage platforms (e.g AWS, Google Cloud, Azure), data visualization solutions (e.g Tableau, Power BI), and actuarial modelling tools (e.g AXIS). Equally important is cultivating a data-driven culture within HELP's internal teams to ensure these tools are used effectively[2].

Finally, we recommend expanding and improving data collection efforts, particularly around client-level information, to further increase the accuracy of pricing models, strengthen risk assessments, and deepen portfolio insights. This can be achieved by leveraging industry data platforms and third-party data providers to enrich HELP's internal datasets[3].

9. References

1. Predictive modeling for insurance pricing: A comprehensive guide using R | by Temiloluwa Emmanuel Olabiyi | Medium. Accessed December 1, 2025. https://medium.com/@temiloluwa.e.olabiyi/predictive-modeling-for-insurance-pricing-a-comprehensive-guide-using-r-3a9c328c1b03.

2. Nikhil. "Data Analytics in Insurance Industry 2025 Guide - CARMATEC." Carmatec Inc - Mobile App Development Company, April 1, 2025. https://www.carmatec.com/blog/data-analytics-in-insurance-industry-guide/.

3. "Insurance Data Tools." Insurance Bureau of Canada - Association Representing Canada's private home, car & business insurers. Accessed November 30, 2025. https://www.ibc.ca/industry-resources/insurance-data-tools.