Task 3 Report: A/B Hypothesis Testing on Insurance Risk Drivers

AlphaCare Insurance Solutions | June 2025
Uncovering Hidden Risk Patterns: Evidence-Driven Segmentation for Smarter Premium Strategies

---

Introduction

Understanding what drives insurance risk is vital for creating fair pricing models and risk-based customer segmentation. In this task, we performed a statistical hypothesis testing analysis to validate whether **gender**, **geographic region**, or **zip code** significantly affect risk or profitability.

Key KPIs:

- **Claim Frequency**: Proportion of policies with at least one claim.
- **Claim Severity**: Average claim amount, conditional on a claim.
- **Margin**: Difference between Total Premium and Total Claim Amount.

I applied **t-tests** for numeric comparisons and **chi-squared tests** for categorical risk patterns.

🎯 Hypotheses Tested

Hypothesis	Metric	Method	Result	Significant?
H_0: No risk difference across provinces	Claim Frequency	T-Test	*p = NaN*	No
H_0: No risk difference between Postal Codes	Claim Frequency	T-Test	*p = NaN*	No
H_0: No profit difference between Postal Codes	Margin	T-Test	*p = 0.244*	No
H_0: No risk difference between Men and Women	Claim Frequency	T-Test	*p = NaN*	No

Hypothesis	Metric	Method	Result	Significant?
H_0: No risk difference between Men and Women	Claim Occurrence	Chi-Squared	**p = 0.00024**	Yes

---

Geographic Insights

- **Province (e.g., Gauteng vs KwaZulu-Natal)**:
  A t-test on claim frequency revealed **no statistically significant difference** (p-value = NaN). This suggests that province-level geography alone is **not a strong predictor** of risk in our current dataset.
- **Zip Codes (e.g., 2000 vs 122)**:
  Both risk and margin comparisons by postal code also failed to show significance (p-values = NaN and 0.244). Although visual differences may exist, they are **not statistically significant** under our tests.

Business Implication:
We do **not** currently need zip-code- or province-based pricing adjustments. Future segmentation should focus elsewhere unless further local patterns emerge with a larger sample.

Gender-Based Risk Differences

Significant: Gender vs Claim Occurrence

- A **chi-squared test** between gender and whether a claim occurred yielded a **p-value of 0.00024**, which is **highly significant**.
- This suggests that **men and women exhibit statistically different claim behaviors**, even though average claim frequencies were not significantly different via t-test.

Interpreting the Contradiction:

- While the **mean Claim Frequency** is similar across genders, the **distribution of claim occurrence** is **not uniform**, potentially indicating different behavioral profiles or claim thresholds.

Business Implication:
Gender should be **included as a segmentation feature**. For example, underwriting or behavioral models can explore why women appear to have different claim incidence patterns.

Recommendations

1. **Incorporate Gender into Segmentation**:
   Use gender as a **feature in predictive models** or offer differentiated customer engagement to reflect underlying claim behavior.
2. **Hold Off on Geographic-Based Adjustments**:
   Current data **does not justify premium adjustments** based on province or zip code. Future models should still log these features, but more data is needed.
3. **Refine Postal Code Analysis**:
   If possible, **bin postal codes** into socioeconomic clusters or urban vs rural classes, this may yield more meaningful segmentation than raw zip.

Caveats

- **Missing P-values (NaN)** in t-tests likely reflect small or unequal group sizes. Consider bootstrapping or nonparametric methods in future iterations.
- **Confounding variables** were not controlled (e.g., policy type, vehicle age). These should be included in downstream modeling for robustness.

Conclusion

This hypothesis testing step helped us validate assumptions and refine our risk segmentation strategy. The **significant relationship between gender and claim occurrence** is a clear actionable insight, while **location-based features** currently show limited statistical value. These findings lay the foundation for **more targeted risk models** in Task 4 and future pricing experiments.