

Hypothesis Testing Report

Insurance Risk Analytics - Task 3

Report Date: December 08, 2025

Executive Summary

This report presents the results of statistical hypothesis testing to validate risk drivers for insurance premium optimization. Four null hypotheses were tested using appropriate statistical methods (Chi-square, Kruskal-Wallis, Mann-Whitney U tests). **3 out of 4 hypotheses were rejected**, indicating significant differences in risk and profitability across geographic and demographic segments. These findings provide statistical validation for risk-based pricing strategies and targeted marketing initiatives.

Hypothesis 1: Risk Differences Across Provinces

Null Hypothesis: There are no risk differences across provinces **Test Results:** • Claim Frequency (Chi-square): $p = 0.000000$ • Claim Severity (Kruskal-Wallis): $p = 0.000000$ • Minimum p-value: 0.000000 **Conclusion:** ✓ REJECT H_0 ($p < 0.05$)

Business Recommendation: We reject the null hypothesis for provinces ($p < 0.01$). Specifically, Gauteng exhibits 0.21 percentage points higher claim frequency than Northern Cape, suggesting a regional risk adjustment to our premiums may be warranted. Premiums should be adjusted upward for high-risk provinces (Gauteng, KwaZulu-Natal, Western Cape) and potentially reduced for low-risk provinces (Northern Cape, Eastern Cape, Limpopo) to attract customers while maintaining profitability.

Hypothesis 2: Risk Differences Between Zip Codes

Null Hypothesis: There are no risk differences between zip codes **Test Results:** • Claim Frequency (Chi-square): $p = 0.000000$ • Claim Severity (Kruskal-Wallis): $p = 0.000004$ • Minimum p-value: 0.000000 **Conclusion:** ✓ REJECT H_0 ($p < 0.05$)

Business Recommendation: We reject the null hypothesis for zip codes ($p < 0.05$). Significant risk differences exist between postal codes, indicating that granular geographic segmentation can improve risk assessment. Consider implementing zip code-level premium adjustments, with higher premiums for high-risk areas and competitive pricing for low-risk zip codes to gain market share.

Hypothesis 3: Margin Differences Between Zip Codes

Null Hypothesis: There is no significant margin (profit) difference between zip codes **Test Results:** • Margin Analysis (Kruskal-Wallis): $p = 0.000000$ **Conclusion:** ✓ REJECT H_0 ($p < 0.05$)

Business Recommendation: We reject the null hypothesis for margin differences between zip codes ($p < 0.001$). Zip code 299 shows ZAR 42.42 higher average margin than zip code 122, indicating significant profitability variation. This suggests that pricing strategies should be optimized at the zip code level, with premium adjustments to improve margins in unprofitable areas while maintaining competitiveness in profitable segments.

Hypothesis 4: Risk Differences Between Women and Men

Null Hypothesis: There is no significant risk difference between Women and Men **Test Results:** • Claim Frequency (Chi-square): $p = 0.935530$ • Claim Severity (Mann-Whitney U): $p = 0.223513$ • Minimum p-value: 0.223513 **Conclusion:** ✗ FAIL TO REJECT H_0 ($p \geq 0.05$)

Business Recommendation: We fail to reject the null hypothesis for gender differences. No statistically significant risk difference was found between men and women, suggesting that gender should not be used as a primary risk factor in premium pricing.

Overall Strategic Recommendations

Based on the hypothesis testing results, the following strategic actions are recommended:

- Implement province-based premium adjustments, with higher premiums for Gauteng, KwaZulu-Natal, and Western Cape
- Develop zip code-level pricing models to capture granular geographic risk differences
- Optimize margins at the zip code level, focusing on improving profitability in unprofitable areas
- Consider gender as a secondary risk factor only if regulatory compliance allows
- Establish continuous monitoring of risk metrics by geographic and demographic segments
- Develop targeted marketing campaigns for low-risk segments (Northern Cape, Eastern Cape, Limpopo)