6th Edition of the Annual McMaster/Co-operators Workshop – 2025

Qualification Round Problem

**Scope:**

The newly appointed Chief Actuary of Blue Skies Insurance Co. is reviewing the company’s auto book in the Greater Toronto Area and has grown concerned about several emerging trends in the last two years.

Over that two year period, a few things have occurred:

1. Retention and closing rates have worsened;
2. New competitors in the market have grown their business faster than expected;
3. Competitors have adopted greater use of AI and machine learning in their pricing;
4. Blue Skies average premium is increasing faster than the market; and
5. Blue Skies has only performed indications up to now and their segmentation model is based on two simple variables, namely *Age of Insured* and *Credit Score*.

The Chief Actuary believes that the company is growing its share of unprofitable risks, a phenomenon called adverse selection.

To fix the issue, the Chief Actuary has tasked the pricing department (you) to improve the current segmentation model.

**Your task:**

1. Use **modelling\_data.xlsx** to build a regression claim count model**1** using exposure as an offset
2. Then use your model to predict on **evaluation\_data.xlsx**
3. Submit your predictions by January 31st to **mcmaster\_workshop@cooperators.ca2**
   1. Actual values withheld and will be used to rank competitors. Use this format:

|  |  |  |
| --- | --- | --- |
| ROW\_ID | EXPOSURE | PREDICTION |
| *(test set row identifier)* |  | *(Model\_Prediction)* |

\*All other columns in test set are to be dropped.

Using this provided data the Chief Actuary will evaluate your predictions using Root Mean Squared Error (RMSE).

Good luck!

1. *Note on claim count model: Including exposure as an offset is exactly equivalent to a frequency model that includes exposure as a weight (but not as an offset)—that is, they will yield the same predictions, relativity factors and standard errors*
2. *Multiple submissions can be made until January 31st, but only the last one will be used for scoring*