# Qualifications phase

## Nature of the data compared to normal insurance data

The data provided to you as part of this workshop is simulated data that has been simplified to allow you to be more effective.

The frequency of claims and their severity has also been increased, and multiple policy attributes have been removed, in order to limit the size of the dataset we will be using.

## Context

One of the most important commercial property insurance companies, the Apex Financial Corporation (AFC), recently had to file for bankruptcy. AFC's last year loss history has been made available to the public to incentivise new companies to emerge to fill this new gap in insurance offering.

Your team is seizing the opportunity. You were able to convince multiple investors to gather enough capital to start your own insurance business. The investors are hoping that your company will thrive amongst the numerous other newly formed insurers.

Only property claims are considered for this workshop.

## Files

### AFC_policies.csv

This file contains all the information on the policies insured by AFC during their last year in business. Each policy was effective for a full year.

#### *Columns*

##### BusinessID

This is an identifier that is unique to each specific company that is seeking insurance.

##### BusinessType

AFC was writing 32 different types of business, ranging from accounting firms to ice cream stores. Each business type has its own average limit, revenue and exposition to risk.

Keep in mind that this data is simulated and that our simulated business classes behaviors will differ from the behavior we would expect in the real world.

##### Province

AFC was writing business in two different provinces, Ontario (ON) and Quebec (QC).

### Latitude and longitude

These two (rounded) numbers relate to the geolocalisation of the risks. Geographical correlationwas included in the simulated data, meaning that some regions are riskier than others. You can identify where a risk is located by looking for (Latitude, Longitude) in google maps. For example, try looking for (43.546078, -80.246139)!

### City

This is the city corresponding to the geolocalisation of the risk.

### Amount of insurance

This variable reflects the amount of coverage the policy was insured for. This is usually the sum of the values of the building, equipment, stock, etc. that is insured under the policy.

### Earnings

This is the declared earnings for the policyholder's company, for the year just before the policy year.

## AFC_claims.csv

This file contains information on the claims incurred during the last year.

## *Columns*

### BusinessID

This is an identifier that is unique to each specific company that is seeking insurance. Keep in mind that some companies may have had more than one claim!

### ClaimAmount

This is the incurred amount of the claim, in dollars. You can consider that this amount is fully developed (at ultimate).

## templateForYourPremiums.csv

**!Important : This file is a template that you will have to fill with the proposed premium for each risk. This is the file that will be uploaded to the website at the end of the qualifications phase.**

This file is currently initialized with a premium that is equal to the average loss per policy.

# Task

Your task is to leverage's AFC's loss history data in order to propose a premium (price) to each of their previous insureds that are now looking for a new insurer.

Some insureds are low risk, some insureds are high risk. Some of them won't have any claim, others will have many small claims and some others may have huge claims. Can you predict an accurate expected loss for each of the risk profiles, based on their features?

By charging a price that is too low, you will attract a lot of policyholders, but you will end up not charging enough to pay for all their losses, and your business won't yield any profit. If you charge a premium that is too high, the insureds will prefer doing business with your competitors, and you won't have enough volume to pay for your expenses.

Starting on February 19th, your team will be able to submit, through our website, a two-column CSV file, based on the 'templateForYourPremiums.csv' file.

To help you in your endeavour, please consult the "Basic property insurance ratemaking information" document.

# Scoring

Once the premiums from all the competing teams are received, each insured (BusinessID) will be awarded to one of the teams based on the following decision process:

- Each insured will look at the premium from 5 randomly selected insurers (teams)
- The insured will decide to do business with the team, out of the 5, that is offering him the lowest premium.

Most insureds won't claim in the next year. But some will. By earning an insured's premium, a team will also be committed to paying the insured's claims for the next year. (This information is, of course, unknown to you)

The final scoring will then be done by computingthe total **underwritingprofit** for each team:

UW Profit = Total premium written – Losses – Expenses

**The teams earning the most profit will move to the final round!**

# Useful pricing assumptions

- Allocated Loss Adjustment Expenses : 10% of incurred claim amounts
- Unallocated Loss Adjustment Expenses : 0$
- Expected frequency trend over the next year : 2%
- Expected severity trend over the next year: We expect the trend over the past 12 months on Statistics Canada's building construction prices index to continue for the next year
- Underwriting expenses:
    - Commissions and brokerage: 15% of the final premium
    - Other acquisition: $150,000
    - General expenses: $450,000
    - Taxes, licenses and fees: 12.5% of the finalpremium
- Underwriting profit: The investors would be happy with a 5% underwriting profit, but they are always happy when expectations are exceeded.
- As of today, we estimate that there will be at least 25 new companies competing to write these 250,000 policies
- As specified earlier, you can consider losses to be fully developed.

# Effective dates information

- AFC's data covers policies that were all written on January 1st of 2019. They were all effective for one full year.
- Your pricing structure will be effective starting on January 1st of 2021.
- You can assume that all your policies will be written on January 1st of 2021 and will be effective for one year.

# **Good luck !**