

By Miles Data Scientist | Interview Task

Background

In order to optimise our premiums for sales and profit we can use data about other quotes offered in the market. Attached is a (fictional) example of such data, it includes information which is supplied by the customer and an average competitor premium.

Task

Using the attached data, build a model to predict our competitors' premiums. It is worth stressing that rather than judging on a correct answer, we would really like to understand how you *approach* the problem including

1. understanding the data and its quality
2. preparing the data
3. building a model
4. validating the model
5. interpreting the model

Considerations

- This model should be suitable to be implemented in live quoting which means it should be compliant with FCA regulations for insurance pricing. We don't expect you to know or research the full FCA regulation, just think about which data you are using in the model.
- There is no restriction on the type of model that can be implemented in our pricing platform.
- Ideally we would like to see the data preparation and model build written in Python.
- You can present your approach using your preferred format(s), eg. slides, notebook, or anything else that works for you.
- Don't worry about spending time optimising the model performance, instead focus on the process and how you explain your approach to us; we don't expect you to spend more than 2-3 hours on this task.