Data Scientist Challenge

Here's a challenge!

We would like to give you the opportunity to show off your skills by completing a short modelling task in either R or Python. We want to see anything you think showcases your skills and abilities in predictive modelling.

For the challenge, AXA is providing some data on claims cost and details from the corresponding motor insurance policy.
Accompanying this document are 2 files:

- *data_train.csv*
- *data_hold_out.csv*

We would like you to build and evaluate a model (or models) to help predict claims cost (the column labelled as "target" in the training data).

Instructions:

1. Use the data in whatever way you see fit to build a predictive model for the cost of claims
2. Evaluate your model as you see fit. **Required Outputs**

Please submit:

1. Your R or Python code (no data or models)
2. Additionally, a csv containing your model's prediction on *data_hold_out.csv* (i.e. "id" column, plus a new "predictions" column) so that we can independently evaluate your model.
3. Be prepared to explain your findings, including limitations of your approach.

Email the above to remi.momo@axa-insurance.co.uk, david.wilson@axa-insurance.co.uk, and james.kelly@axa-insurance.co.uk.

Your code submission should allow us to see how you have built your model and reproduce your model predictions and evaluation metrics.

This Challenge will be used as part of the assessment however it is mainly used for us to understand your skill set and where/if you will need support if successful. Please spend no more than 4 hours on this Challenge.

Good luck!

Know You Can