INSHUR – Pricing Data Scientist Task

**Context:** INSHUR is aiming to grow its UK rideshare book, with the immediate focus being to attract more new business. They have provided you with 3 months of quote data and want you to create a predictive model that will help them target areas of new business where conversions can be improved.

The business understands that the data needs some attention before use, so would like to see any steps you have taken to clean it, and any assumptions you may have made along the way. Additionally, they are happy for you to choose any modelling technique that suits the problem at hand, but would like to be talked through your reasons for your choice.

Your model and findings will be presented to a small cross-functional team via a 10-15 minute presentation. It is worth noting that some of the team you will present to will not have a technical background, so please tailor your presentation to suit the audience.

-----

Please create and deliver a 10-15 minute presentation (in any format) to the hiring team keeping the following points for consideration;

* What steps have you taken to clean and manipulate the data
* What modelling technique you have chosen and why
* What steps you have taken in building your model
* Statistical results and insights from your model
* Any business recommendations you have from working with the data and creating the model

-----

To help you tackle the problem, INSHUR have also provided the following data dictionary;

|  |  |  |
| --- | --- | --- |
| **Column** | **Type** | **Description** |
| INCEPTIONDATE | Date | Date the policy starts |
| QUOTEDATE | Date | Date the quote was made for the policy |
| SALE | Binary | Flag to say if a policy was purchased (1) or not (0) |
| PREMIUM | Numeric | The cost of the policy |
| DRIVERAGE | Numeric | The age of the policy holder |
| LICENCEYEARSHELD | Numeric | The number of years the policy holder has held a UK driving licence |
| NOCLAIMSDISCOUNT | Numeric | The number of years no claims bonus the policy holder has |
| PREVIOUSCLAIMS | Numeric | The number of previous claims the policy holder has |
| REGION | String | The region of the UK the policy holder lives |
| VEHICLEAGE | Numeric | The age of the vehicle being insured |
| VEHICLEVALUE | Numeric | The value of the vehicle being insured |
| VEHICLETYPE | String | The type of vehicle being insured |