

NICD Data Scientist - Task A3

Caitlin Halfacre

August 3, 2025

I have chosen to use a Bayesian regression model to fit the clients desire for a model that both describes more of the variance in the data than their existing model and maintains predictive accuracy. While Bayesian modelling is computationally and technically more complex than a standard linear regression model, the approach to prediction and inference is more intuitive, and enables the option to present percentage likelihoods in explanations of valuations. In my analysis of the model I present visualisations and phrasing that could be used to explain valuation decisions to clients, regulators, and internal stakeholders.

1 Model Design

1.1 Priors

Intercept

- normal distribution
- $\mu = 0$
- $\sigma = \text{half of maximum likely greatest value}^1 \text{ (95\% credible interval)} = 2,000,000/2 = 1,000,000$

1.2 Model Selection

highest bayes_R2 - model explains more variance.

highest elpd_diff in your table - best expected predictive accuracy based on your negated version.

2 Model Analysis

x

¹<https://www.chroniclelive.co.uk/news/north-east-news/most-expensive-north-east-houses-30988236>
the most expensive house sold in the North East in 2024