Volume 7 Issue 2 - July 2023 - pp. 69-80
https://doi.org/10.32734/jocai.v7.i2-11434

DATA SCIENCE Journal of Computing and Applied Informatics

Contents lists available online at TALENTA Publisher

DATA SCIENCE: JOURNAL OF COMPUTING AND APPLIED INFORMATICS (JoCAI)

Journal homepage: https://talenta.usu.ac.id/JoCAI

Price Prediction with Bayesian Inference and Visualization: Empirical Evidence in India Real Estate

Harry Patria

Newcastle University, School of Computing, United Kingdom

harry.patria@sbm-itb.ac.id

ARTICLE INFO

Article history:
Received 11 June 2023
Revised 29 June 2023
Accepted 27 July 2023
Published online 31 July 2023

Keywords:
Bayesian Linear Regression
Bayesian inference
Model
Housing Price

Corresponding Author:
harry.patria@sbm-itb.ac.id

ABSTRACT

Classical regression serves two primary purposes: evaluating the structure and strength of the relationship between variables. However, while classical regression provides only a point estimate and confidence interval, Bayesian regression offers a comprehensive range of inferential solutions. This study demonstrates the suitability of the Bayesian approach for regression tasks and its advantage in incorporating additional a priori information, which can strengthen research. To illustrate, we utilized data from the Indian Housing dataset provided by the Kaggle Repository. We found that prior distributions produce analytical, closed-form conclusions, eliminating the need for numerical techniques like Markov Chain Monte Carlo (MCMC). Furthermore, this study provides software implementations, along with formulas for the posterior outcomes that are explained and presented clearly. In the third step, Bayesian tools were employed to evaluate the assumptions that underlie the proposed approach. Specifically, the essential processes of Bayesian inference - prior elicitation, posterior calculation, and robustness to prior uncertainty and model sufficiency - were assessed.

IEEE style in citing this article:
H. Patria, " Price Prediction with Bayesian Inference and Visualization: Empirical Evidence in India Real Estate," *DATA SCIENCE: JOURNAL OF COMPUTING AND APPLIED INFORMATICS (JoCAI)*, vol. 7, no. 2, pp. 69-80, 2023.

1. **Introduction**

The goal of science is to comprehend phenomena and systems to predict and control their advancement. Models, which are advanced mathematical or algorithmically representations in the language of quantitative sciences, play a vital role in the scientific process of knowledge elaboration found in a wide range of industrial applications such as energy [1,2,3], finance [4,5,6,7,8].

Renting and purchasing houses are becoming more popular as cities become more crowded. As a result, predicting and modeling house prices is a crucial subject in a way for determining a more robust method of calculating house prices [9,10]. Furthermore, this is going to help both sellers and buyers find the best price for their homes [11], strengthening the economy during a recent global recession. This result, which accurately reflects market price become a hot topic in the real estate sector.

This study aims to shed some light on Bayesian linear regression used to model house prices. One advantage of using a Bayesian model for predicting housing prices is that it allows for the incorporation of prior knowledge or belief about model parameters. This can be especially useful in situations where there is limited data available or the data is noisy. Secondly, Bayesian models also allow for the incorporation of regularization techniques, such as shrinkage priors, which can help to reduce overfitting and improve the generalization performance of the model. In addition, Bayesian models can naturally handle missing data, by using appropriate probability models to represent the missing data process.

The advice is organized around the steps of a Bayesian approach. Section 1 outlines the background and purpose. Section 2 describes the literature review while section 3 describes the elicitation of the dataset and method used in this study such as a prior distribution and the estimated posterior distribution. The distributions of the regression parameters, the regression function, and the model, as well as their estimates and uncertainties, are given and explained in Section 4. Lastly, section 5 summarizes the findings and discussion in a way to generalize the findings and application.

2. Literature Review

Real estate market data serves as a treasure trove of invaluable insights for both home buyers and sellers, and contemporary data analytics methodologies, such as machine learning techniques, have emerged as indispensable tools for extracting practical knowledge from this wealth of information [9,10,11,12]. For instance,[9] conducted an insightful analysis of historical Australian real estate transactions, uncovering significant price disparities between homes in Melbourne's most expensive and least expensive suburbs. Their study demonstrated the competitive strength of the Stepwise and Support Vector Machine combo, which was assessed based on mean squared error. Similarly, in China, Yu et al. (2018[10]) harnessed the power of deep learning-based prediction models, attaining enhanced accuracy in predicting current real estate data. This innovative approach enabled them to discern and assess the intricate effects of diverse factors on housing prices. Meanwhile, [11] delved into the subject of house price analysis through an array of machine learning algorithms, encompassing simple linear regression (SLR), multiple linear regression (MLR), and neural networks (NN).

Amidst the array of methodologies available in modern data science, Bayesian regression methods shine as a beacon of intellectual and methodological prowess, proffering a novel perspective on uncertainty quantification [13,14,15]. Unlike traditional point estimators, Bayesian approaches empower analysts by yielding an entire distribution over regression parameters, thereby embracing the full spectrum of model possibilities [16,17,18,19,20,21]. The crux of Bayesian epistemology lies in its capacity to accommodate prior knowledge or beliefs about model parameters, enabling a seamless integration of existing domain expertise into the modeling process. Moreover, the Bayesian paradigm gracefully handles challenges of regularization and missing data, ensuring that the analytical endeavor remains robust and resilient, even in the face of incomplete information [16,17,18,12].

By virtue of this remarkable fusion of philosophical depth and analytical rigor, Bayesian linear regression emerges as a potent forecasting tool, meriting consideration over alternative machine learning techniques like simple or multiple linear regression, as well as neural networks, in certain contexts [16,17,18,12]. The ability to quantify uncertainty through Bayesian methods significantly enhances the model's predictive capacity, as it enables the explicit representation of what the model has learned from the available data, providing a comprehensive understanding of the inherent variability in the predictions [19,20,12].

Adopting Bayesian linear regression as a key analytical tool should not be perceived as an antagonistic stance against frequentist approaches; rather, it should be regarded as a harmonious synergy, broadening the data scientist's toolkit and fostering a nuanced appreciation for the context-specific merits of each approach [16,17,18,12]. Embracing both Bayesian and frequentist methodologies equips data scientists with a comprehensive set of analytical instruments, enabling them to discern the most suitable tool for a given task, thereby enhancing analytical productivity and fostering a deeper understanding of the data-driven phenomena they seek to explore. In this spirit of inclusivity, the overarching goal of data science is to embody methodological versatility and adaptability, transcending

 https://doi.org/10.32734/jocai.v7.i2-11434

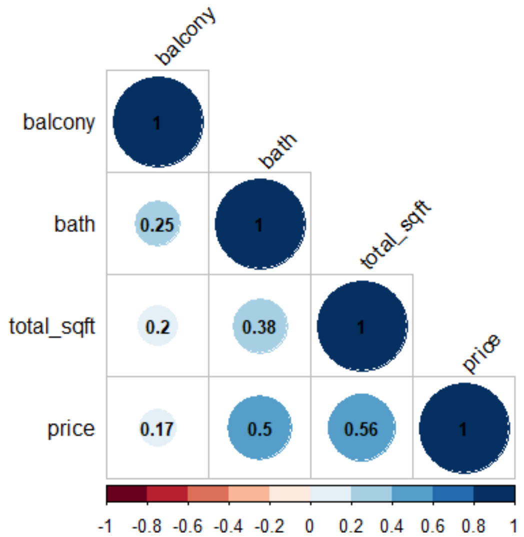
dogmatic adherence to one paradigm and embracing the art of informed selection to empower robust and contextually apt analytical modeling [12].

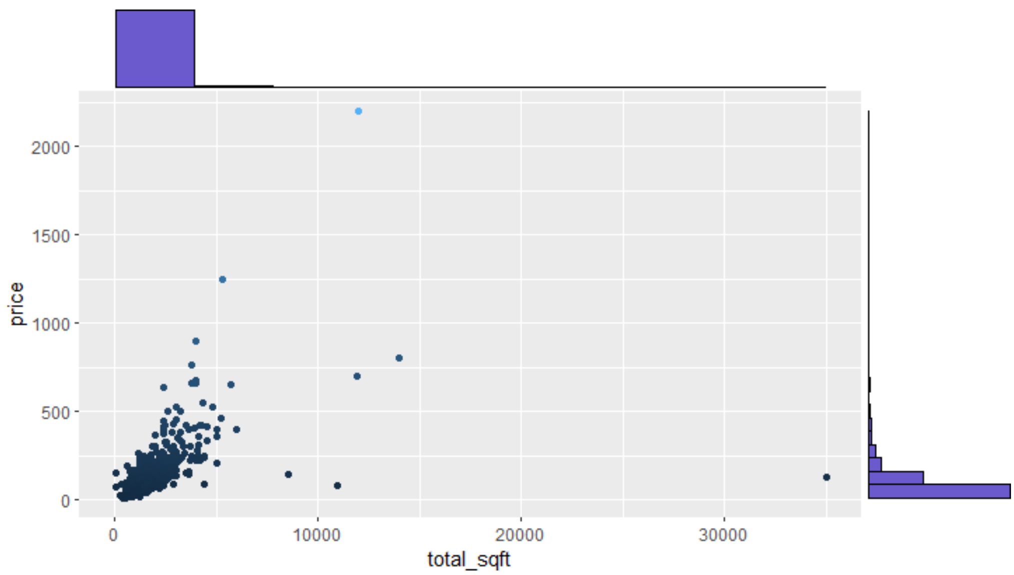
3. Methods

3.1 Data Collection

This study used Indian Housing provided by the Kaggle repository which can be found at https://www.kaggle.com/datasets/aryanfelix/bangalore-housing-prices. Each of the 1,183 entries provides aggregate information about 5 numerical characteristics of residences from different Bangalore areas while the rest are categorical and attributed data, which was uploaded in 2022.

Table 1 Descriptive statistics





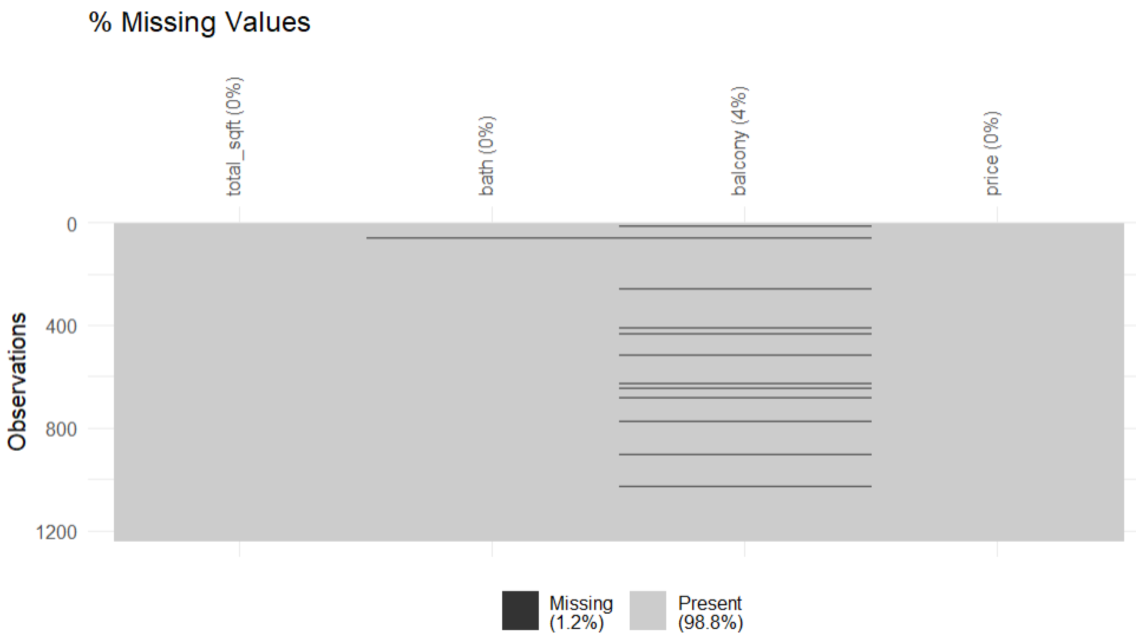


Figure 1. Correlation analysis using R code in the given housing dataset
Source: Author

Table 1 found that the correlation analysis in the property data examined the relationship between the variables balcony, bath, total square feet, and property price. The analysis utilized a half-diagonal correlation heatmap to display the strength of the correlations. Based on the heatmap, it can be observed that the total square feet of property area and its property price have moderate positive correlations with each other, with a correlation coefficient of 0.56. Similarly, bath and price have a moderate positive correlation of 0.5, while the correlation between balcony and property price is slightly weaker at 0.17.

The "total square feet" variable ranges from 11 to 35,000, with a mean value of 1547. The number of bathrooms ranges from 1 to 9, with a mean value of 2.62. The number of balconies ranges from 0 to 3, with a mean value of 1.59. The price of the properties ranges from 11 lakhs to 2200 lakhs, with a mean value of 104 lakhs. The first quartile for total square feet is 1097, meaning 25% of the properties have a total square feet value below this number. The median value is 1278, indicating that 50% of the properties have a total square feet value below this number. The third quartile is 1634, meaning that 75% of the properties have a total square feet value below this number.

In terms of bathrooms, the first quartile is 2, the median is 2, and the third quartile is 3. This indicates that 25% of the properties have two or fewer bathrooms, 50% have two or fewer bathrooms, and 75% have three or fewer bathrooms. The first quartile for balconies is 1, the median is 2, and the third quartile is 2. This means that 25% of the properties have one or fewer balconies, 50% have two or fewer balconies, and 75% have two or fewer balconies. Finally, the first quartile for the price is 49.3 lakhs, the median is 70 lakhs, and the third quartile is 110.5 lakhs. This means that 25% of the properties are priced below 49.3 lakhs, 50% are priced below 70 lakhs, and 75% are priced below 110.5 lakhs.

According to the description and correlation, total square feet will be more important in predicting the price of the property, so this study focuses on the highest connection between house features and their estimated price following the Bayesian approach.

Figure 1 found that the data set contains missing values in approximately 1.2% of the observations. Despite the result found the plausible correlated missing patterns between the balcony and bath, but the proportion of missing data is relatively small, so dropping these observations is a viable option. The MCAR (Missing Completely at Random) test can be used to check if the missingness is random or not. If the test indicates that the missing values are completely random, then dropping them is reasonable. However, if the missing values are not MCAR, then dropping them can introduce bias and affect the validity of the analysis.

In addition to the MCAR test, it's important to consider the reason for the missing data. If the missing data is due to data entry errors or other types of random errors, then the MCAR assumption is likely to hold. However, if the missing data is due to systematic reasons, such as non-response bias, dropping the missing values may introduce bias and affect the validity of the analysis.

Therefore, before deciding whether to drop the missing values, it's important to investigate the reason for the missing data and perform the MCAR test. If the MCAR test supports the assumption of random missingness and the reason for missing data is not systematic, then dropping the missing values can be a reasonable option.

3.2 Bayesian Linear Regression

Linear regression, a venerable statistical approach, assumes a pivotal role in modeling the intricate relationship between a dependent variable and one or more independent variables. At its core lies the elegant utilization of Bayes' theorem, serving as a fundamental principle that guides the updating of our beliefs regarding model parameters in light of the data we observe [22]. Endowed with this Bayesian foundation, linear regression has proliferated across diverse fields, manifesting its prowess in prediction, estimation, and inference tasks.

Notably, the domain of finance has stood witness to the transformative impact of Bayesian linear regression, engendering a multitude of innovative applications. Within this realm, financial data about stock prices, exchange rates, and interest rates have all succumbed to the analytical prowess of Bayesian linear regression, bestowing upon analysts a formidable toolset for exploring and understanding the intricate dynamics of the financial landscape (Chen & Dunson, 2013). Through this approach, financial analysts can obtain not only point estimates but also a full distribution of model parameters, unraveling invaluable insights into the uncertainty that shrouds financial phenomena, thus enhancing the depth of analysis and fostering more informed decision-making.

Moreover, the domain of engineering has also embraced the formidable potential of Bayesian linear regression, unraveling novel avenues for addressing complex problems. In the field of engineering, this methodology has found application in diverse realms, ranging from predicting surface roughness with remarkable precision (Chen & Xu, 2012) to forecasting short-term travel speed, a critical aspect of contemporary transportation planning and management.
This section outlines linear regression using probability distributions rather than point estimates in light of a Bayesian perspective. The response, y, is assumed to be drawn from a probability distribution rather than being estimated as a single value. The Bayesian Linear Regression model with a response sampled from a normal distribution is illustrated by Equation 1:

$$y \sim \mathrm{N}(\beta^{T}X, \sigma^{2}I) \quad (1)$$

A notable symbol of y is defined as an output of a normal (Gaussian) distribution with a mean and variance. For linear regression, the mean is calculated by multiplying the weight matrix by the predictor matrix. The variance is equal to the standard deviation squared (multiplied by the Identity matrix because this is a multi-dimensional formulation of the model).

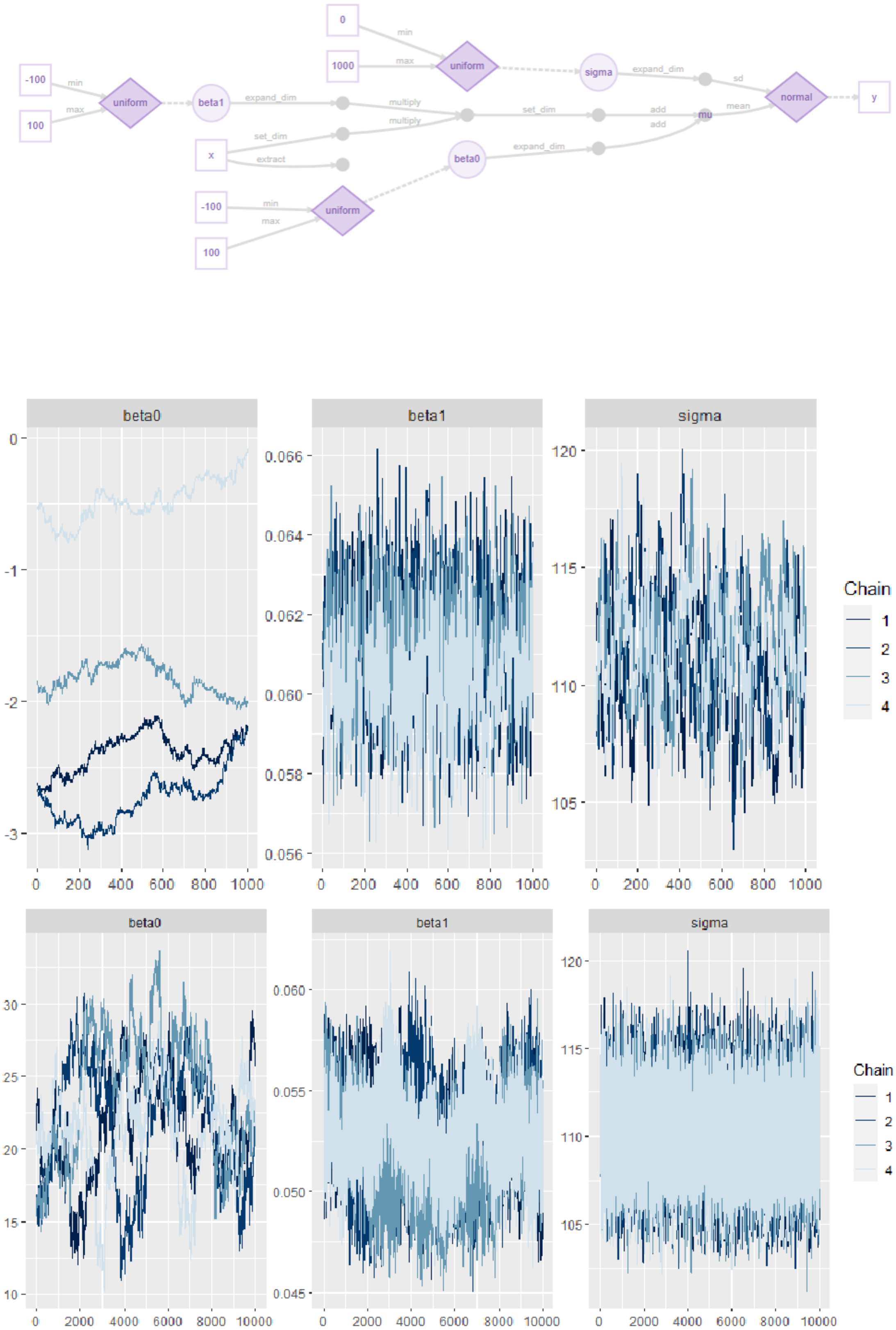


Figure 2. Distribution and Correlation, the prior and posterior probability distribution

Figure 2 described the outcomes of 1000 MCMC steps, which means that 10,000 steps were taken from the posterior distribution. Based on the findings, it appears that a Bayesian regression model was applied to estimate the parameters beta0, beta1, and sigma throughout 4 chains. The findings are useful to check for model convergence. The results of the Bayesian regression model indicate that the parameters beta0, beta1, and sigma were estimated using 10,000 steps from the posterior distribution of 1000 MCMC steps.

Additionally, Figure 2 shows the convergence of the model by displaying the outcomes of the MCMC steps. The convergence of the model is indicated by the stability of the estimated parameters throughout the 4 chains. Furthermore, the Bayesian regression model allows for the incorporation of prior information in the estimation process. This prior information can help improve the accuracy of the parameter estimates and provide a better understanding of the relationship between the variables in the property data.

The goal of Bayesian Linear Regression is not to find a single "best" value for the model parameters but rather to determine the posterior distribution for the model parameters. Not only is the response generated from a probability distribution, but again the model parameters are also assumed to be generated from a distribution. The posterior probability of the model parameters is completely reliant on the training inputs and outputs:

$$P(\beta|\mathrm{y},\mathrm{X}) = \frac{P(y|\beta,\mathrm{X})\ x\ P(\beta|\mathrm{X})}{P(y|\mathrm{X})} \tag{2}$$

3.3 Posterior Probability Distribution

Given the inputs and outputs, P(∣y, X) represents the posterior probability distribution of the model parameters. This is equal to the prior probability of the parameters multiplied by the likelihood of the data, P(y∣, X), and divided by a normalization constant. This is a straightforward formulation of the Bayes Theorem, which serves as the cornerstone of Bayesian inference. As opposed to OLS, our model's parameters have a posterior distribution that is proportional to the likelihood of the data times the prior probability of the parameters.

Within the purview of priors, Bayesian statistics offers a remarkable avenue to seamlessly incorporate domain knowledge or informed estimates concerning the model parameters, distinguishing itself from the traditional frequentist approach that solely relies on information extracted from the observed data. This pivotal aspect of Bayesian methodology empowers analysts with the flexibility to enrich their models with relevant prior information, thereby ushering in a profound synthesis of expert insights and empirical evidence.

In contrast to the frequentist paradigm, which remains agnostic to external knowledge, Bayesian practitioners can leverage non-informative priors to bolster their analyses in cases where prior estimates are absent. Notably, the normal distribution often assumes the mantle of choice for such scenarios, providing a versatile and widely employed non-informative prior that accords the model with the necessary flexibility to explore a range of potential parameter values.

This capacity to imbue models with informed prior beliefs constitutes a formidable advantage of Bayesian statistics, lending an interpretive depth and contextual richness that transcends the confines of raw data. By judiciously integrating domain knowledge and informed estimates, analysts can bestow upon their models a more nuanced understanding of the underlying phenomena, thus engendering a profound shift in the modeling paradigm, where the data-driven inferences harmoniously intermingle with expert judgment. In this manner, Bayesian statistics amplifies the analytical acumen, liberating analysts from the rigid confines of purely data-centric approaches, and unlocking the true potential of the interplay between human expertise and empirical observation.

4. Results

4.1. Testing and Result Validation

These are the outcomes of 1,000 MCMC steps, which means that 1,000 steps were taken from the posterior distribution by the method. Table 2 and Table 3 show the testing and result validation using 1,000 and 10,000 data points, respectively.

Table 2. Testing and result validation using 1,000 datapoints

	MEAN	SD	NAIVE SE	TIME SER. SE	2.50%	25%	50%	75%	97.50%
$\boldsymbol{\beta_0}$	-1.856	0.8733	0.0138	0.0413	-2.974	-2.562	-2.084	-1.377	-0.283
$\boldsymbol{\beta_1}$	0.061	0.0016	0.00246	0.0437	0.058	0.059	0.061	0.062	0.064
$\boldsymbol{\sigma}$	111.292	2.5076	0.00396	0.174	106.50	109.50	111.26	113.10	116.02

In Table 2, when using 1,000 data points, the mean value of β_0 is -1.856 with a standard deviation of 0.8733. The naive standard error is 0.0138, and the time series standard error is 0.0413. The 2.50% quantile is -2.974, the 25% quantile is -2.562, the 50% quantile is -2.084, the 75% quantile is -1.377, and the 97.50% quantile is -0.283. For β_1, the mean value is 0.061 with a standard deviation of 0.0016. The naive standard error is 0.00246, and the time series standard error is 0.0437. The 2.50% quantile is 0.058, the 25% quantile is 0.059, the 50% quantile is 0.061, the 75% quantile is 0.062, and the 97.50% quantile is 0.064. Finally, the mean value of σ is 111.292 with a standard deviation of 2.5076. The naive standard error is 0.00396, and the time series standard error is 0.174. The 2.50% quantile is 106.50, the 25% quantile is 109.50, the 50% quantile is 111.26, the 75% quantile is 113.10, and the 97.50% quantile is 116.02.

Table 3. Testing and result validation using 10,000 datapoints

	MEAN	SD	NAIVE SE	TIME SER. SE	2.50%	25%	50%	75%	97.50%
$\boldsymbol{\beta_0}$	22.0485	4.0398	0.00202	0.6508	13.94	19.24	22.11	25.09	29.32
$\boldsymbol{\beta_1}$	0.0526	0.0021	0.00105	0.00138	0.049	0.051	0.052	0.054	0.057
$\boldsymbol{\sigma}$	110.0925	2.2725	0.00114	0.01305	105.75	108.54	110.05	111.60	114.65

In Table 3, when using 10,000 data points, the mean value of β_0 is 22.0485 with a standard deviation of 4.0398. The naive standard error is 0.00202, and the time series standard error is 0.6508. The 2.50% quantile is 13.94, the 25% quantile is 19.24, the 50% quantile is 22.11, the 75% quantile is 25.09, and the 97.50% quantile is 29.32. For β_1, the mean value is 0.0526 with a standard deviation of 0.0021. The naive standard error is 0.00105, and the time series standard error is 0.00138. The 2.50% quantile is 0.049, the 25% quantile is 0.051, the 50% quantile is 0.052, the 75% quantile is 0.054, and the 97.50% quantile is 0.057. Finally, the mean value of σ is 110.0925 with a standard deviation of 2.2725. The naive standard error is 0.00114, and the time series standard error is 0.01305. The 2.50% quantile is 105.75, the 25% quantile is 108.54, the 50% quantile is 110.05, the 75% quantile is 111.60, and the 97.50% quantile is 114.65.

To inspect the effect of the number of data points in the Bayesian model, the author used two attempts using a different number of data points. In the first attempt with 1,000 data points, the resulting fit is shown in Figure 3. The posterior distributions of the model parameters are approximated, and a range of lines representing different estimates of the model parameters are displayed. These lines overlap to a larger extent, indicating higher uncertainty in the model parameters due to the lower number of data points. In the second attempt with 10,000 data points, the resulting fit is shown in Figure 4. The posterior distributions of the model parameters are again approximated, but this time the lines representing different estimates of the model parameters have less overlap. This suggests that with a larger number of data points, the model parameters become less uncertain and the resulting fit becomes more accurate. The effect of the number of data points on the model is evident in the overlapping lines. The use of Bayesian regression modeling is particularly advantageous when dealing with small samples, unbalanced group sizes, and a decreasing number of data points over time.

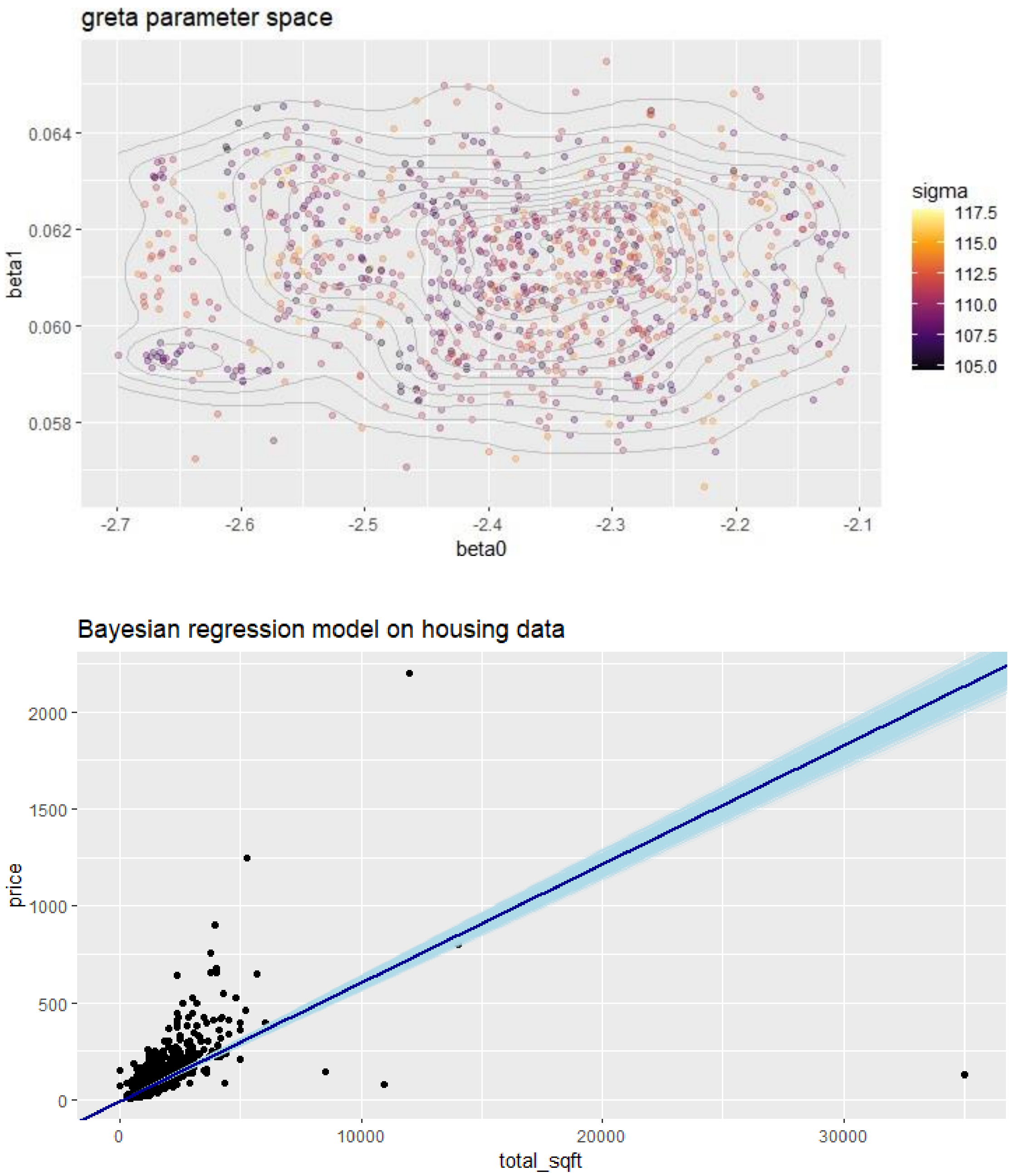
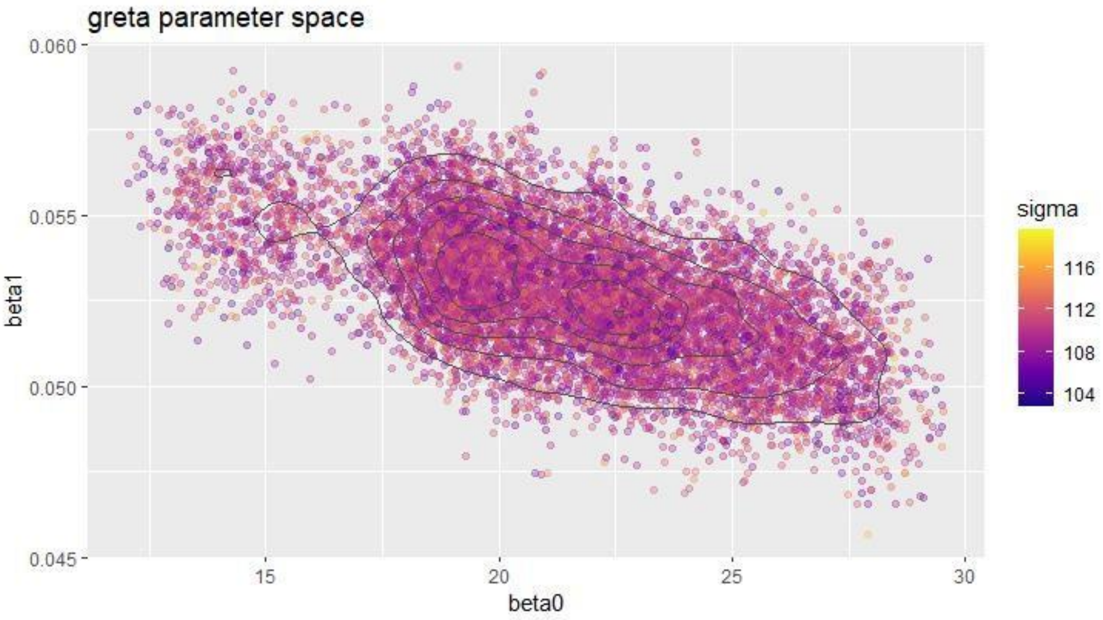


Figure 3. Bayesian Regression of property price and total square feet on the first attempt (1,000 data points)



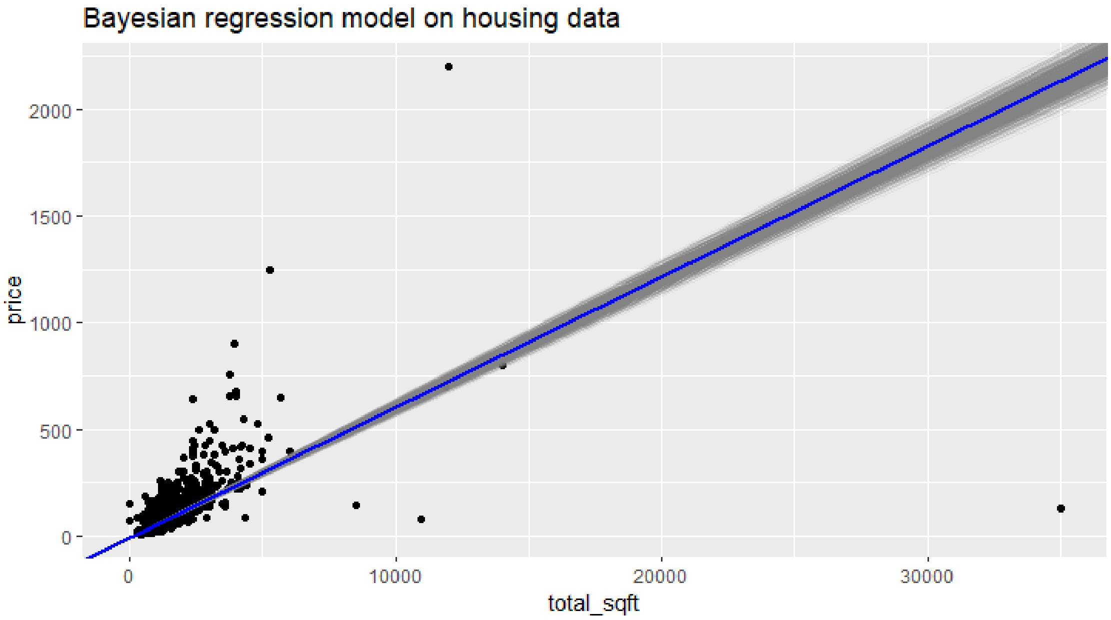


Figure 4. Bayesian Regression of property price and total square feet on the second attempt (10,000 data points)

Based on the findings presented in Figure 3 and Figure 4, it can be inferred that when the number of data points is smaller, it leads to more significant adjustments and greater uncertainty in the model. This implies that the Bayesian Linear Model produces a distribution of potential model parameters for a single data point rather than a single result. The Bayesian Linear Regression approach produces a posterior distribution of possible model parameters based on the data and prior information. With fewer data points, the posterior distribution will be more dispersed, allowing us to quantify our level of uncertainty about the model.

5. Conclusions

It is more beneficial to learn both approaches than to take a stance in the ongoing debate between Bayes and Frequentist, rather than taking sides in the escalating debate between the two. The debate between the frequentist and Bayesian schools of inference can be found in various articles and studies on probability theory, hypothesis testing, and statistical inference. Both paradigms have their strengths and weaknesses, which have led to ongoing discussions and debates within the scientific community. The Bayesian approach is characterized by incorporating prior knowledge and incorporating uncertainty in the model. This allows for a more flexible and nuanced analysis, particularly in situations where limited data or prior information is available.

On the other hand, frequentists argue that inference should be based solely on observed data and aim to make objective conclusions based on the data at hand. They emphasize the importance of avoiding subjectivity and bias in statistical analysis. The frequentist approach is often associated with hypothesis testing and p-values, while the Bayesian approach is known for its use of Bayes' theorem and the incorporation of prior knowledge through prior distributions.

This debate between the two schools of inference has been ongoing for over a century and continues to generate discussions among statisticians and researchers in various scientific fields. One prominent viewpoint is that the Bayesian and frequentist approaches offer complementary views and can be cooperative on a practical level. Rather than taking sides in the ongoing debate between Bayes and Frequentist, it is more beneficial to understand and utilize both approaches in the appropriate context, taking into account their validity, context, and transferability.

References

[1] H. Patria and V. Adrison, "Oil Exploration Economics: Empirical Evidence from Indonesian Geological Basins," *Economics and Finance in Indonesia*, vol. 61, no. 3, p. 196, Dec. 2015, doi: 10.7454/efi.v61i3.514.

[2] H. Patria, "The Role of Success Rate, Discovery, Appraisal Spending, and Transitioning Region on Exploration Drilling of Oil and Gas in Indonesia in 2004–2015," *Economics and Finance in Indonesia*, vol. 67, no. 2, p. 183, Dec. 2021, doi: 10.47291/efi.v67i2.952.

[3] Harry Patria, "Predicting the Oil Investment Decision through Data Mining," *Data Science: Journal of Computing and Applied Informatics*, vol. 6, no. 1, pp. 1–11, Jan. 2022, doi: 10.32734/jocai.v6.i1-7539.

[4] C. D. Mariana and H. Patria, "Are Electric Vehicle Stocks in ASEAN Countries Investible during the Covid-19 Pandemic?," *IOP Conf Ser Earth Environ Sci*, vol. 997, no. 1, p. 012002, Feb. 2022, doi: 10.1088/1755-1315/997/1/012002.

[5] D. Anggraeni, K. Sugiyanto, M. Irwan, Z. Zam, and H. Patria, "STOCK PRICE MOVEMENT PREDICTION USING SUPERVISED MACHINE LEARNING ALGORITHM: The Konstanz Information Miner (KNIME)".

[6] F. Zulfikri, D. Tryanda, A. Syarif, and H. Patria, "Predicting Peer to Peer Lending Loan Risk Using Classification Approach," *International Journal of Advanced Science Computing and Engineering*, vol. 3, no. 2, pp. 94–100, Oct. 2021, doi: 10.30630/ijasce.3.2.57.

[7] H. Patria, "Predicting Fraudulence Transaction under Data Imbalance using Neural Network (Deep Learning)," *Data Science: Journal of Computing and Applied Informatics*, vol. 6, no. 2, pp. 67–80, Jul. 2022, doi: 10.32734/jocai.v6.i2-8309.

[8] A. Kurniawan, A. Rifa'i, M. A. Nafis, N. S. Andriaswuri, H. Patria, and D. Purwitasari, "Feature Selection and Sensitivity Analysis of Oversampling in Big and Highly Imbalanced Bank's Credit Data," in *2022 10th International Conference on Information and Communication Technology (ICoICT)*, IEEE, Aug. 2022, pp. 35–40. doi: 10.1109/ICoICT55009.2022.9914889.

[9] T. D. Phan, "Housing Price Prediction Using Machine Learning Algorithms: The Case of Melbourne City, Australia," in *2018 International Conference on Machine Learning and Data Engineering (iCMLDE)*, IEEE, Dec. 2018, pp. 35–42. doi: 10.1109/iCMLDE.2018.00017.

[10] L. Yu, C. Jiao, H. Xin, Y. Wang, and K. Wang, "Prediction on Housing Price Based on Deep Learning," *World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering*, vol. 12, no. 2, 2018.

[11] N. Vineeth, M. Ayyappa, and B. Bharathi, "House Price Prediction Using Machine Learning Algorithms," 2018, pp. 425–433. doi: 10.1007/978-981-13-1936-5_45.

[12] H. Patria, "Bayesian Regression for Predicting Price Empirical Evidence in American Real Estate," *Data Science: Journal of Computing and Applied Informatics*, vol. 7, no. 1, pp. 15–23, Jan. 2023, doi: 10.32734/jocai.v7.i1-10082.

[13] M. West, "Outlier models and prior distributions in Bayesian linear regression. ," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 46, no. 3, pp. 431–439, 1984.

[14] T. J. Mitchell and J. J. Beauchamp, "Bayesian Variable Selection in Linear Regression," *J Am Stat Assoc*, vol. 83, no. 404, p. 1023, Dec. 1988, doi: 10.2307/2290129.

[15] D. J. C. Mackay, "BAYESIAN NON-LINEAR MODELING FOR THE PREDICTION COMPETITION," 1994.

[16] A. E. Raftery, D. Madigan, and J. A. Hoeting, "Bayesian Model Averaging for Linear Regression Models," *J Am Stat Assoc*, vol. 92, no. 437, p. 179, Mar. 1997, doi: 10.2307/2291462.

[17] T. P. Minka, "Bayesian linear regression Thomas P. Minka," 1998.

[18] S. M. Lynch, *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York, NY: Springer New York, 2007. doi: 10.1007/978-0-387-71265-9.

[19] P. Pérez, G. Campos, J. Crossa, and D. Gianola, "Genomic-Enabled Prediction Based on Molecular Markers and Pedigree Using the Bayesian Linear Regression Package in R," *Plant Genome*, vol. 3, no. 2, p. plantgenome2010.04.0005, Sep. 2010, doi: 10.3835/plantgenome2010.04.0005.

[20] I. Castillo, J. Schmidt-Hieber, and A. van der Vaart, "Bayesian linear regression with sparse priors," *The Annals of Statistics*, vol. 43, no. 5, Oct. 2015, doi: 10.1214/15-AOS1334.

[21] A. Gelman, B. Goodrich, J. Gabry, and A. Vehtari, "R-squared for Bayesian Regression Models," *Am Stat*, vol. 73, no. 3, pp. 307–309, Jul. 2019, doi: 10.1080/00031305.2018.1549100.