

Analytical Model to Understand the Performance of Investment Advisory – Bayesian Approach

S. Mythreyi Koppur¹, Dr. B. Senthilkumar²

¹ Research Scholar, PG Research & Department of Statistics, Thanthai Periyar Govt. Arts and Science college (Autonomous), Trichy – 23.

(Affiliated to Bharathidasan University, Trichy – 24).

² Assistant Professor, PG Research & Department of Statistics, Thanthai Periyar Govt. Arts and Science college (Autonomous), Trichy – 23.

(Affiliated to Bharathidasan University, Trichy – 24).

Corresponding Author: mythreyi.phd@gmail.com¹, senthilkumarbalan@gmail.com²

Received: 2022 March 15; **Revised:** 2022 April 20; **Accepted:** 2022 May 10

Abstract

An investment advisory is a service-oriented organisation that buys and sells shares, on behalf of their customers. The goal of this study is to identify a profit/loss pattern for four quarters in call recommendations (Buy/Sell). However, there seems to be a very few or no effort to generate performance indicators for investment advisors based on historical data on their investment advice. The inherent advantage of dealing with Bayesian modelling has been attempted using the necessary models applying adequate underlying parameter transformation. The variation in four quarters could be achieved through the measure of heterogeneity and the odds ratio is used to quantify the measure of associations between the variables of interest. To do Bayesian analysis, the complete exercise was conducted using R and R-Stan.

Keywords: Investment advisory, Bayesian, Call recommendation, Heterogeneity, R stan.

Introduction

People buy stocks to build their wealth over time. While some people assume that stocks are a high-risk investment, several studies have shown that holding the correct shares for a long period of time (five to ten years) may

give inflation-beating returns, making them a better alternative than real estate or gold. When it comes to investing in the stock market, many have short-term strategies. While stocks may be incredibly volatile in short bursts, buying in the correct stocks can help traders earn quickly.

A stockbroker is a financial market agent or middleman who deals in securities. Their major goal is to buy and sell stocks on behalf of their clients. Stock brokers serve as a link or point of contact between stock exchanges and customers. Share brokers might work alone or as part of a brokerage business. The vast majority of stockbrokers work for a brokerage firm, dealing with a wide range of individual and institutional clients. Stockbrokers are typically paid on a commission basis; however, this varies depending on the work. Brokerage firms and broker-dealers are referred to as stockbrokers.

Various research academics conducted several studies on this issue, including behavioural patterns of investors toward investment, capital market, secondary market, equity market, stock broking, and stock brokers. According to Srinivasan and Hanssens (2009), analysts' recommendations have considerable influence on organisations' long-term performance because analysts' recommendations are more compelling than other information channels for public investors due to their specialised reputation. Mohanraj and Kowsalya (2017) conducted research on investor satisfaction with stock broker service quality in the Coimbatore District. Hou, Zhao, and Yang (2018) explore

the link between analyst information connection and earnings forecasting performance and build an information transmission network based on all research reports produced on Shanghai and Shenzhen Stock Exchange equities from 2008 to 2016.

Kalaiselvi and Sangeetha (2018) conducted research on the ratio analysis of chosen stock brokerage organisations. Wang and Jiang (2019) conducted research that suggests financial analysts can act as an information bridge between brand equity and a firm's long-term performance by making expert stock investment recommendations to public investors. Pan and Xu (2020) investigated whether analysts' cash flow estimates boost the profitability of their stock recommendations and if this positive effect changes with business earnings quality.

There is a scarcity of study on the performance of stock brokers and their impact on investing decisions. The goal is to have a good knowledge of the factors described in the data and to develop a model to understand the performance in every Quarter, based on the stock broker's recommendations using a Bayesian statistical technique. The odds ratio was used to achieve association measures between the variables of interest, and the measure of heterogeneity was determined as the significant measure of this study. The

posterior distribution is generated using MCMC, which is essentially leveraging Markov chains to address multiple issues using a common operating system. Andrieu et al., (2003) go into further depth on the development and use of various MCMC approaches. Gelman et al (2017) STAN's language stan, a probabilistic programming language, was utilised with R to do Bayesian analysis. All computations were performed in R, a publicly accessible software tool designed specifically for statistics (Chambers, J. (2008)). Section 2 presents a representative data set with a full explanation of the variables; Section 3 includes the methods and models used in the research; Section 4 presents data treatment and interpretation; and Section 5 presents the comments and conclusions drawn from the analysis.

Dataset Description

The data set used in this study was obtained from a stock trading business that was granted permission solely for academic purposes. They are involved in making trading recommendations based on careful study in technical and proven models. They developed

an app, which is a superior equity market analytics and research tool that utilises NSE and BSE data, with analytics and visualisation alerts. It is a piece of the mission of educating ordinary investors on the opportunities to generate wealth in the equity market. This is a typical rectangular data set with 6111 observations and 23 variables from the time period (January 2020 - December 2020).

Whenever there is an opportunity to investigate a data set, the focus is on identifying the variables involved in the process and the relationships between the variables from which the uncertainty may be deduced. On the other hand, there is an interest in forecasting future occurrences related to the subject. Such an endeavour needs a significant amount of prior information or understanding about the process, as well as a science to learn from data using a suitable technology. The aim lies in understanding the Quarterly performance of the investment advisory across call recommendations. The details of the variables which are considered in this study is as follows:

Table 1: Description of the variables involved in this study

S. No	Variable Name	Description	Nature	Details
-------	---------------	-------------	--------	---------

1	Call_Status	Buy / Sell status when the call is initiated	Factor	2 levels
2	Call_Result	Outcome status - P or L	Factor	2 levels
3	Quar	Quarter (Jan - Dec) of the transaction	Factor	4 levels

When a call is launched, the call status is either buy or sell, as recommended by the advice. The accompanying figure 1 shows that there was a total of 3023 buy and 3088 sell calls for the year 2020.

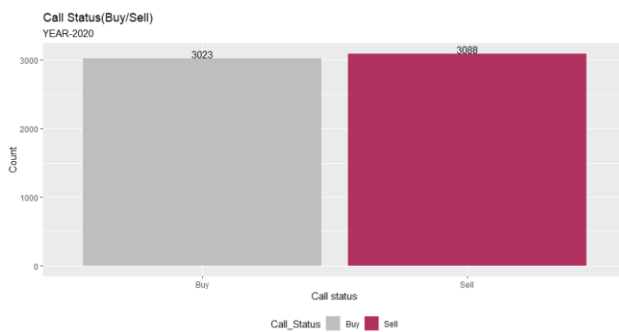


Figure 1: Distribution of Call Status (Buy/Sell) in Recommendations.

The outcome value of a call is the profit or loss determined before deductions. Figure 2 shows that a total of 4180 calls resulted in profit, and around 69 percent, with the remaining calls resulting in loss, being approximately 31 percent.

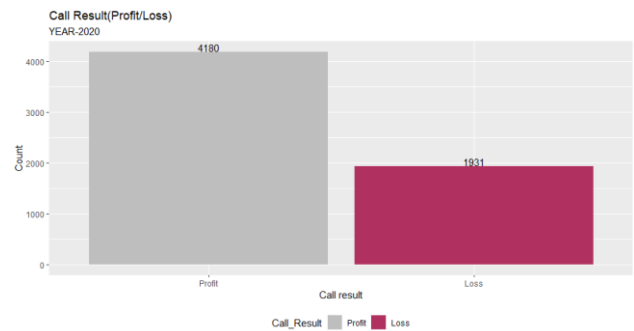


Figure 2: Distribution of Call Result (Profit/Loss).

When looking at the comparison of Call Result and Call Status, it is clear from the Figure 3, that profit calls are quite high in sell call recommendations compared to Buy call recommendations. Considering the loss status, it is quite high in Buy call recommendations compared to sell call recommendations.

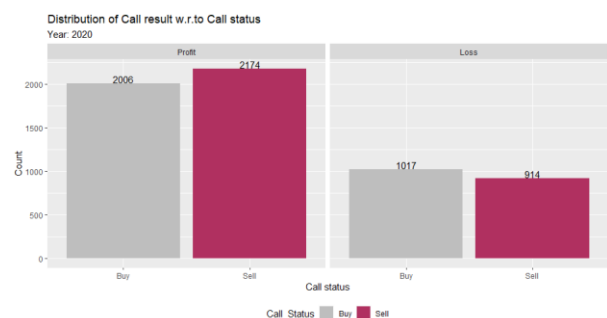


Figure 3: Comparison of Call Status with Call Result.

Quarters signify four quarters in a year, ranging from Q1 to Q4. Considering, the

profit/loss outcomes across four quarters, from Figure 4 it is clear that Profit is quite high in the 1st Quarter and similar way the loss too.

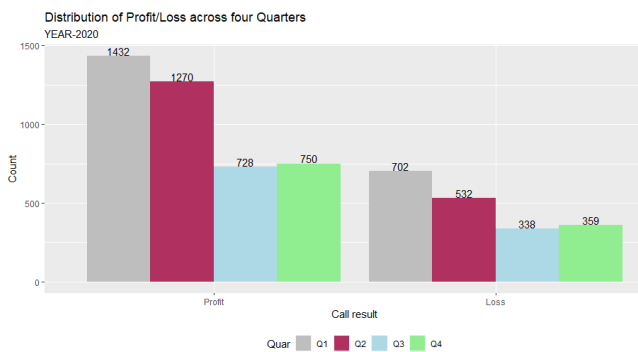


Figure 4: Distribution of Profit/Loss across four quarters

From Figure 5, it is clear that while comparing the Buy/Sell call recommendations across four quarters, calls are quite high in the 1st quarter in both the call recommendations.

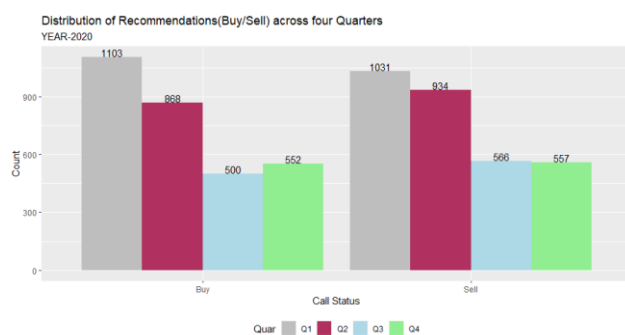


Figure 5: Distribution of Call Recommendations (Buy/Sell) across four Quarters.

In the same way, it is important to consider all the three variables by attempting to connect all the three variables Call status, call result and quarters. By this attempt, it would further give us a better clarity and understand the association between them and the most important notion is “variability” to understand the variation across four quarters.

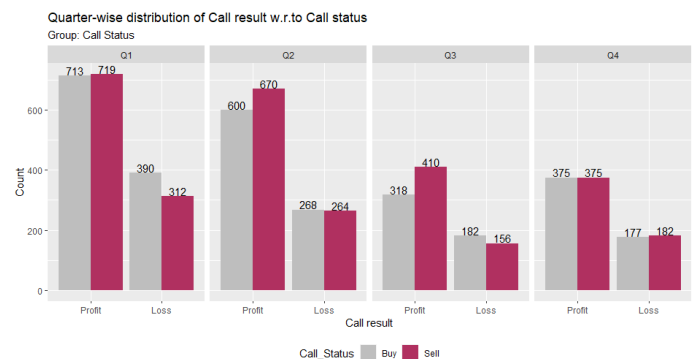


Figure 6: Distribution of Call status and Call Result across four quarters.

From the Figure 6, it is clear evident to understand the association between all the three variables. Calls are quite high in 1st and 2nd quarter. Almost, equal number of buy and sell call recommendations in profit in the 4th quarter. A scientific approach exists for finishing and/or assisting the investigation process. This might aid in measuring variability and understanding the evolving nature of a significant variable.

Methods and Materials

The odds ratio can be used to evaluate the strength of the link between the variables of

interest. It has evolved into one of the most important tools for comparing two proportions. It's a straightforward 2×2 contingency table analysis measure. Agresti(2012) may be referenced for a better understanding of the odds ratio and its applications, as well as acceptance of it as a superior summary measure in the random effect model when comparing binomial proportions. The odds ratio, which is commonly employed in medical research, can be an effective tool for detecting if categorical forecasts and observations are connected. Furthermore, as mentioned in Stephenson, D. B., significance tests on the logarithm of the odds ratio can be used to assess if the talent is purely due to random sampling (2000). Finley's standard set of tornado forecast is used in this study to demonstrate the principles. For a probability η of a success, the odds are defined to be $\tau = \frac{\eta}{(1-\eta)}$.

The ratio of the odds η_1 and η_2 in the two rows,

$\omega = \frac{\tau_1}{\tau_2} = \frac{\eta_1/(1-\eta_1)}{\eta_2/(1-\eta_2)}$ is called the odds ratio.

When the odds ratio is greater than one, the subject in the first row is more likely to have the first response than the subject in the second row. If the odds ratio is between zero and one, the subject in the first row is less

likely to have the first response than the subject in the second row.

Bayesian approaches have been used to solve issues in a variety of domains, including archaeology, as cited in Buck, C. E., et al (1996). Dorfman, J. H. proposed a Bayesian method to econometrics (1997). Pollard, W. E.(1986) used Bayesian statistics in his assessment studies. The goal is to demonstrate the essential concept of the Bayesian method, which begins by assigning priors to all unknown parameters, then defines the probability of the data given the parameters, and lastly determines the posterior distribution of the parameters given the data. Gelman, A., et al., provide a detailed explanation of the practical approach to data analysis and issue resolution using Bayesian approaches (2013). Gelman, A., and C. R. Shalizi (2013) provide an essential but not exhaustive list of materials discussing the theory and practise of Bayesian statistics. Another component of Bayesian method is the need for calculations, which has witnessed a significant increase in algorithms such as Markov Chain Monte Carlo (MCMC) Smith, B. J. (2007) and Denwood, M. J. (2016) might be considered for R implementation. Stan supports comprehensive Bayesian inference for continuous-variable models using Markov chain Monte Carlo techniques, as mentioned

in Carpenter, B., et al., (2017) and Gelman, A., et al., (2015).

Though the intent and purpose of various strategies varies, the underlying statistical model is the random effects model. The random effect model is essentially a two-stage hierarchical structure based on a normal distribution. The majority of research, including the most recent (Langan et al., 2017), use normally distributed models with proper adjustment of underlying parameters.

In Random Effect Model, if λ_i is an effect size estimate of a corresponding true effect size θ_i with the within-study variance σ_1^2 , then we could **estimate** θ_i , $i=1,2, \dots, k$ from the sample data; let us denote these estimated values as λ_i , $i=1,2, \dots, k$; that is $\hat{\theta}_i = \lambda_i$, $i=1,2, \dots, k$. Here we estimate Odds Ratio based on the independent binomial distribution for the two rows.

This study includes a normal model for a log transformed Binomial parameter. The models are presented in a schematic way as given below:

Scheme I:

$$X_1 \sim \text{Bin}(n_1, \theta_1)$$

$$X_2 \sim \text{Bin}(n_2, \theta_2)$$

Then defining, $\Omega = \text{logit}(\theta_1) - \text{logit}(\theta_2)$ which is log-odds ratio. For k tables, we

$$\text{define } \mu = \frac{\text{logit}(\theta_1) + \text{logit}(\theta_2)}{2}$$

$$\theta_1 = \text{logit}^{-1}(\mu + \Omega/2)$$

$$\theta_2 = \text{logit}^{-1}(\mu - \Omega/2)$$

Scheme II:

$$\mu \sim N(\mu_0, \sigma_0^2)$$

$$\Omega \sim N(d, \tau^2)$$

Scheme III:

$$d \sim N(m_d, s_d)$$

$$\tau^2 \sim \text{inv-gamma}(\tau_1, \tau_2)$$

All hyper parameters $\mu_0, \sigma_0^2, \tau_1, \tau_2, m_d, s_d$ are appropriately chosen and the choices are presented as for Scheme II, $N(0, 10^4)$ are for mean parameter and in Scheme III, the hyper priors in inverse gamma distribution are retained with (3,1).

The entire exercise has been carried out in R and R-Stan has been used to do Bayesian analysis.

Data treatment and interpretation

The primary goal is to provide a model or a quality indicator to stock brokers based on their call recommendations. The stock trading business has provided a data set for one full year (2020). The following are the many stages of data treatment:

This study is presented in the form of a two-fold contingency table that categorises two dichotomous variables. In this scenario, the two levels of X_1 are Buy and Sell in terms of call status, and X_2 are Profit and Loss in terms

of call result. Each cell count is the number those accounts for the combination of X_1 and X_2 . Quarters are the stratifying factors. Changing a rectangle dataset obtained from the brokerage into a 2 x 2 data set with k levels (Quarters) which is clearly seen from the table below.

	X_1/X_2	Profit	Loss
Q1	Buy	713	390
	Sell	719	312
Q2	Buy	600	268
	Sell	670	264
Q3	Buy	318	182
	Sell	410	156
Q4	Buy	375	177
	Sell	375	182

After the data has been analysed and the proper linked variables identified, the final result will be a $k \times 2 \times 2$ data set that will be retrieved for the model. The purpose of this effort was to harness the benefits of Bayesian modelling in dealing with such models. The basic purpose is to build posterior distributions using the best approach available. Bayesian analysis for complicated models may be made relatively simple by using MCMC to simulate posterior distributions. A Markov chain is founded on the fundamental idea that the future is exclusively determined by the present and not

by the past. So, the most important aspects of MCMC are understanding where to begin, where to stop, and how to allow them to mix. The most important component is selecting independent samples.

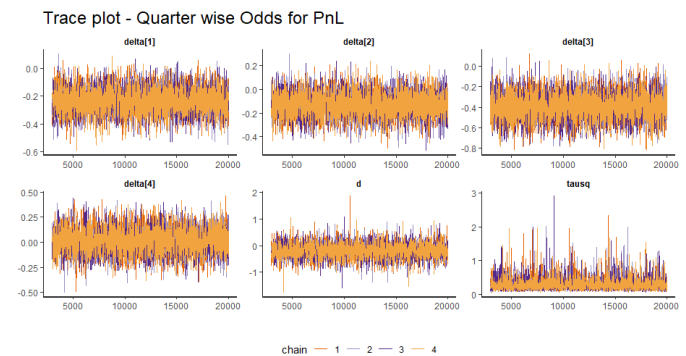


Figure 7: Trace plot for the Quarter wise analysis

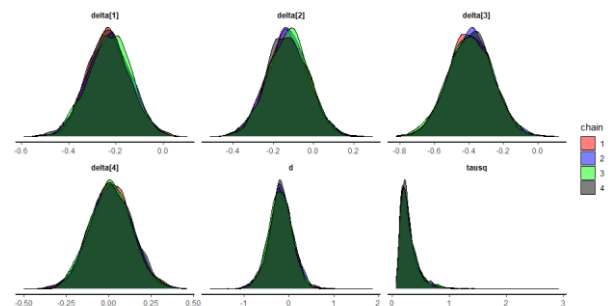


Figure 8: Density plots for Quarter wise analysis

A wide range of MCMC convergence measurement, posterior summarization, and data visualisation applications are accessible when R is used as the MCMC home environment. Before analysing the model's outputs, many diagnostics must be checked to determine whether the Markov chain has

converged to its stationary distribution. The above trace plot for the model with several predictors evaluated in this investigation does not appear to be divergent. None of these graphs reveal any concerning departures from normality as a gauge of convergence. Convergence, on the other hand, does not show any evidence of considerable divergence. The posterior sample can be

displayed using a trace plot, which shows samples over iterations. A trace graphic depicts the sampled data per chain and node during iterations. It allows us to visually assess the convergence and mixing of the chains. Convergence, on the other hand, reveals no significant divergence for any of the models.

Table 2: Combined Odds Ratio for month wise analysis; LL and UL are lower and upper limits of 97.5% confidence interval.

Quarter	Estimate	LL	UL
Q1	0.794106	0.664526	0.950154
Q2	0.878969	0.717971	1.078474
Q3	0.676632	0.523507	0.870569
Q4	1.014262	0.793479	1.304908
Overall	0.83	0.469904	1.437622
Heterogeneity	0.3	0.111371	0.796744

From Table 4, it is clear evident that all the estimates are less than 1, except for the 4th Quarter. Overall estimate is less than 1, indicating that the Sell call recommendations has higher odds for profit than buy call recommendations in all the four quarters. Especially, to note that the individual OR value is greater than 1 in the 4th Quarter indicating that the Buy call recommendations has higher odds for profit than sell call recommendations. It is important to note that there is a positive measure of heterogeneity of

about 30% across four quarter of the year 2020.

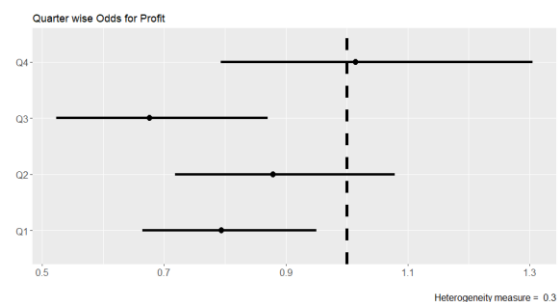


Figure 9: Forest plot of the point and interval estimates of individual odds ratio corresponding to various Quarters.

From the forest plot it is clear than odds ratio estimate is greater than one only in the 4th Quarter, showing higher odds for profit in Buy call recommendations. There seems to be a positive measure of heterogeneity of about 30% across four quarters.

Conclusion

Investment advice is an essential component of individual and/or corporate financial planning. Bodies involved in such consultation use a variety of strategies for a wide range of financial items. This is accomplished in the setting of a very volatile and uncertain global financial market and environment. Many organisations, including regulatory bodies, use very complicated analytical methodologies to report on business or sector performance.

An investment consultant advises clients on which equities to add to their financial portfolios and looks for profitable businesses. Accounting, economic, and technical study on various organisations and stocks is performed by competent stock brokers. They understand the client's requirements and financial goals and, as a consequence, advise them on the best investment possibilities.

In this study, the advisory's suggestions and the results based on those recommendations are examined throughout four quarters. This

model shown that the adviser generates a higher profit on sell call recommendations than on buy call recommendations. The R stan functions were used to swiftly fit Bayesian models.

Working in this statistical framework requires comprehension of a new set of basic ideas relevant to model creation, estimation, and inference. It is advantageous for a number of reasons, including the inclusion of prior distributions, which allows useful structure to be incorporated at each level of a model, and regularisation, which benefits in model maintenance. The posterior predictive distribution is easy to use and effectively propagates parameter uncertainty into predictions; and the language we may use to discuss our findings is straightforward to comprehend and allows us to make probability assumptions regarding parameters of interest.

- It is obvious from the preceding model that the advisory generates a substantial profit on sell call recommendations.
- With the exception of the fourth quarter, the profit in sell call suggestion is higher.
- The buy call recommendation has a better chance of profiting solely in the fourth quarter.

- Across four quarters, there appears to be a positive measure of heterogeneity of roughly 30%.

According to the findings of this study, Bayesian modelling offers a lot of potential in terms of offering analytical tools that may integrate historical / domain information via suitable prior considerations. When assessing the relationship of acceptable factors, this research provides several chances for better comprehending the data.

Overall, it is obvious from the results of this study that the advisory produces a solid profit in all of the quarters in sell call recommendation, which is a significant characteristic.

References

- [1] Lucas Jr, H. C., & Spitler, V. K. (1999). Technology use and performance: A field study of broker workstations. *Decision sciences*, 30(2), 291-311.
- [2] Firth, M. (1997). An analysis of the stock market performance of new issues in New Zealand. *Pacific-Basin Finance Journal*, 5(1), 63-85.
- [3] Madden, J. F. (2012). Performance-support bias and the gender pay gap among stockbrokers. *Gender & Society*, 26(3), 488-518.
- [4] Domowitz, I., & Yegerman, H. (2005). Measuring and interpreting the performance of broker algorithms. *Algorithmic Trading: a Buyside handbook*.
- [5] McHugh, M. L. (2009). The odds ratio: calculation, usage, and interpretation. *Biochemia Medica*, 19(2), 120-126.
- [6] Pathak, M., Dwivedi, S. N., Deo, S. V. S., Sreenivas, V., & Thakur, B. (2017). Which is the preferred measure of heterogeneity in meta-analysis and why? A revisit. *Biostat Biometrics Open Acc*, 1, 1-7.
- [7] Song, F., Sheldon, T. A., Sutton, A. J., Abrams, K. R., & Jones, D. R. (2001). Methods for exploring heterogeneity in meta-analysis. *Evaluation & the health professions*, 24(2), 126-151.
- [8] Sutton, A. J., & Higgins, J. P. (2008). Recent developments in meta-analysis. *Statistics in medicine*, 27(5), 625-650.
- [9] Hardy, R. J., & Thompson, S. G. (1998). Detecting and describing heterogeneity in meta-analysis. *Statistics in medicine*, 17(8), 841-856.
- [10] Costantini, D., & Møller, A. P. (2013). A meta-analysis of the effects of geolocator application on birds. *Current Zoology*, 59(6), 697-706.
- [11] Baujat, B., Mahé, C., Pignon, J. P., & Hill, C. (2002). A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. *Statistics in medicine*, 21(18), 2641-2652.
- [12] S. Mythreyi Koppur and Dr. B. Senthilkumar, Random effect model for the causes of recommendation - A

- Bayesian approach, *Advances and Applications in Statistics* 76 (2022), 53-74.
- [13] Froelicher, V. F., Perdue, S., Pewen, W., & Risch, M. (1987). Application of meta-analysis using an electronic spread sheet to exercise testing in patients after myocardial infarction. *The American journal of medicine*, 83(6), 1045-1054.
- [14] L'ABBÉ, K. A., Detsky, A. S., & O'ROURKE, K. E. I. T. H. (1987). Meta-analysis in clinical research. *Annals of internal medicine*, 107(2), 224-233.
- [15] Patil, S. R., Morales, R., Cates, S., Anderson, D., & Kendall, D. (2004). An application of meta-analysis in food safety consumer research to evaluate consumer behaviors and practices. *Journal of food protection*, 67(11), 2587-2595.
- [16] Lavielle, M., & Lebarbier, E. (2001). An application of MCMC methods for the multiple change-points problem. *Signal processing*, 81(1), 39-53.
- [17] Warnes, G. R., Bolker, B., Gorjanc, G., Grothendieck, G., Korosec, A., Lumley, T., ... & Rogers, J. (2014). gdata: Various R programming tools for data manipulation. R package version, 2(3), 35.
- [18] Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: a probabilistic programming language. *Grantee Submission*, 76(1), 1-32.