AUREUM ACADEMIC JOURNAL
VOL. 2, NO. 3, Sep 2025
https://doi.org/10.48161/aaj.v2n3a91

From Customer Value to Campaign Response: An Explainable Marketing Analytics Framework Using UK Online Retail and Portuguese Bank Marketing Data

Helena Kalp[1*]

[1] Management and Education, Stadio and North West University, Malmesbury, South Africa;

* **Corresponding author:** hkalp@swartlandhs.co.za.

ABSTRACT: This study examines how linking customer value analysis to campaign response prediction can improve marketing decisions. The paper uses two historical public business datasets. The first contains online retail transactions from a UK-based, registered, non-store retailer between 1 December 2010 and 9 December 2011. The retailer sells unique all-occasion gifts, and many customers are wholesalers. The second dataset records direct phone marketing campaigns by an anonymized Portuguese banking institution, where the outcome is customer subscription to a term deposit. The datasets are not merged at the row level because they do not share customer identifiers. Instead, they are linked through one decision logic: identify where customer value is concentrated, then rank campaign contacts by their probability of response. The online retail analysis produces 397,884 clean transactions, 4,338 customers, and total revenue of 8,911,407.90. The top 20% of customers generate 74.6% of revenue. In the bank marketing data, the base response rate is 11.7%. A random forest model without call duration reaches a ROC AUC of 0.801, and the top 20% model-ranked contacts respond at 36.2%, about 3.09 times the base rate. The findings show that marketing return is uneven in both settings. Firms can improve targeting by protecting high-value customers and prioritizing campaign contacts with higher estimated response probability.

Keywords: marketing analytics; RFM segmentation; campaign response; customer value; explainable machine learning; retail analytics.

I. INTRODUCTION

Marketing decisions are often made under budget limits, time pressure, and incomplete customer information. Firms cannot give the same level of attention to every buyer or every campaign contact. Some customers generate a large share of revenue, while many contacts are unlikely to accept an offer. This makes customer prioritization a key part of marketing management.

Marketing analytics has become central to this problem because customer data can help firms move from broad campaign planning to more selective decisions. Earlier studies show that analytics can support segmentation, targeting, customer value measurement, campaign response prediction, and retail decision-making [5], [6], [12], [18]. However, many studies treat customer value analysis and campaign response prediction as separate tasks. One stream focuses on identifying valuable customers, while another predicts whether customers will respond to a campaign.

This paper connects these two tasks within a single marketing decision framework. The aim is not to merge unrelated datasets or claim that the same customers appear in both files. Instead, it shows how two common marketing problems can be linked at the decision level. The first problem is value concentration: which customers carry most of the revenue? The second problem is campaign response: which contacts are more likely to accept an offer?

The study uses two historical public datasets. The online retail dataset is used to analyze customer value, revenue concentration, RFM segments, countries, products, and purchasing patterns. The bank marketing dataset is used to analyze response behavior and build pre-campaign prediction models. The bank model is estimated with and without call duration, but the main operational model excludes duration because this variable is only known after the call starts. This distinction matters because campaign planning should rely only on information available before contact.

The paper contributes in three ways. First, it presents a reproducible workflow that links descriptive customer analytics with predictive campaign targeting. Second, it shows that marketing return is concentrated in both datasets: retail revenue is concentrated among a minority of customers, and bank response is concentrated among higher-ranked contacts. Third, it keeps the analysis managerially realistic by separating pre-campaign variables from post-contact information.

The paper is organized as follows. The literature review discusses customer value, marketing analytics, campaign prediction, retention, retail analytics, and explainable modeling. The hypotheses and research questions then define the empirical focus. The proposed method explains the data sources, preparation steps, statistical tests, and modeling procedure. The results section reports findings from the retail, bank, and cross-dataset analyses. The discussion explains the managerial meaning of the results.

II. LITERATURE REVIEW

Recent marketing studies place these modeling tools inside a broader business context. Customer engagement research in a big data environment argues that data-driven engagement should create value for both customers and firms [19]. Artificial intelligence is expected to change marketing strategy, customer interaction, personalization, and service decisions [27]. Customer analytics capability research in retail shows that firms need organizational capability, not only software, to benefit from data-driven retailing [28]. Digital transformation research also argues that firms need new capabilities, structures, and performance metrics when markets become data-intensive [29]. Strategic AI work links AI to marketing research, segmentation, targeting, positioning, and marketing actions [30], while consumer AI research notes that customers experience AI through data capture, classification, delegation, and social interaction [31].

Explainability matters because the paper uses machine learning to support a managerial decision. LIME was introduced to explain individual predictions from black-box classifiers and help users judge whether to trust a model [15]. SHAP later provided a unified feature-attribution approach for interpreting model predictions [20]. Broader surveys of explainable artificial intelligence show that explanation is not only a technical matter; it affects whether decision makers can understand, check, and use model outputs responsibly [24], [25].

Retention studies also caution that this matters for interpreting predictive models. Research on enhanced retention management stresses that firms must distinguish between at-risk customers and those worth targeting [21]. Another study shows that targeting high-risk customers can be ineffective when the intervention does not change behavior enough to justify the cost [22]. Churn prediction research has also moved toward models that combine predictive performance with interpretability, such as hybrid approaches based on logistic regression and decision trees [23]. These studies support using a pre-campaign model that ranks likely responders while remaining understandable to managers.

The bank marketing part of this paper builds on studies of campaign response and customer retention. The Portuguese bank telemarketing study showed that data mining can predict the success of term-deposit marketing calls and that campaign, customer, and economic variables can help identify more responsive contacts [8]. Churn and retention research also shows that prediction models should be judged by their business value, not only by statistical accuracy [4]. This is important for the present paper because top-20% lift is a campaign measure that reflects the practical question of whom to contact first.

Retail research also supports the paper's focus on uneven customer value and timing. Customer experience research shows that firms should understand customer behavior across the journey rather than at one isolated transaction point [14]. Digital marketing research highlights that customer touchpoints increasingly shape targeting and response opportunities [16]. Retailing studies further show that technology, customer engagement, big data collection, and analytics have become central parts of retail competition [17],

while predictive analytics in retail depends on customer, product, time, location, and channel information [18]. Digital transformation in retail also reshapes the value chain by shifting the customer interface and creating new value sources [26].

The customer-value literature directly supports this study's retail component. RFM analysis estimates customer lifetime value and groups customers by purchase recency, frequency, and monetary value [1]. A closely related online retail study used RFM-based customer segmentation to help a retailer understand its customer base and support customer-centric marketing [2]. Researchers have also used retail transaction data for lifestyle segmentation, showing that purchase history can reveal meaningful customer groups even when rich demographic data are not available [3]. Later work on RFM-based customer lifetime value also supports the use of behavioral variables to rank customers and guide next-best service or product decisions [11].

A second group of studies explains why this paper treats analytics as a marketing process rather than only a modeling exercise. The marketing mix literature discusses big data analytics and shows how data can support decisions about product, price, promotion, and distribution [10]. A wider review of marketing analytics in data-rich environments shows that customer databases, digital traces, and external data can support segmentation, targeting, personalization, and performance measurement [12]. Consumer analytics research also argues that big data's value comes from extracting hidden insights into customer behavior and turning them into usable marketing action [13].

Studies published between 2011 and 2021 show a clear shift from simple reporting to customer-level data-driven decision support. Business intelligence and analytics research argues that large structured and unstructured datasets can improve managerial decisions when the analysis is connected to a business problem [5]. Marketing analytics has also been linked with stronger firm performance, especially when firms use analytical outputs in routine managerial decisions rather than treating them as isolated technical reports [6]. This is consistent with work on digital marketing challenges, which shows that firms need stronger data skills, better customer insight, and clearer links between analytics and action [7].

Table 1. Comparison Between Studies.

Study group	Related studies	Main focus	Method / approach	Link with this paper
RFM and customer value	[1], [2], [11]	Customer lifetime value, recency, frequency, and monetary value	RFM scoring, segmentation, customer value ranking	Supports the use of RFM to identify valuable online retail customers.
Online retail segmentation	[2], [3]	Understanding retail customers from transaction history	Customer segmentation using purchase behavior	Supports the retail part of this paper, especially customer grouping and value concentration.
Campaign response prediction	[8]	Predicting success of bank telemarketing campaigns	Data mining and classification models	Directly supports the Portuguese bank marketing analysis.
Churn and retention management	[4], [21], [22], [23]	Customer retention, churn risk, and targeting value	Predictive models, retention evaluation, interpretable models	Supports the idea that prediction should be linked with business value, not accuracy only.
Business intelligence and marketing analytics	[5], [6], [7], [10], [12], [13]	Using data to improve marketing decisions	Big data analytics, marketing analytics, decision support	Supports the paper's position that analytics should guide practical marketing action.
Customer engagement and digital marketing	[14], [16], [19]	Customer experience, engagement, and digital touchpoints	Customer journey and digital interaction analysis	Helps explain why customer behavior and

Study group	Related studies	Main focus	Method / approach	Link with this paper
				timing matter in marketing decisions.
Retail analytics and digital retailing	[17], [18], [26], [28]	Retail technology, predictive analytics, and customer analytics capability	Retail analytics, customer/product/time/channel analysis	Supports the use of retail transaction data for managerial decision-making.
Artificial intelligence in marketing	[27], [29], [30], [31]	AI, digital transformation, and marketing strategy	Strategic AI, digital transformation, AI-consumer interaction	Supports the broader framing of the paper as a data-driven marketing framework.
Explainable machine learning	[15], [20], [24], [25]	Explaining black-box prediction models	LIME, SHAP, explainable AI reviews	Supports the use of explainable modeling for campaign targeting decisions.

III. HYPOTHESES / QUESTIONS

The study uses both research questions and testable hypotheses because the analysis combines description, statistical testing, and prediction.

- RQ1: How concentrated is customer value in the online retail transaction data?
- RQ2: Which customer, timing, and campaign variables are associated with bank marketing response?
- RQ3: Can a pre-campaign prediction model rank likely responders better than the base response rate?
- RQ4: What decision link connects retail customer value and bank campaign response?
- H1: Retail revenue is concentrated among a minority of customers.
- H2: Prior behavior and campaign timing are significantly associated with bank campaign response.
- H3: A pre-campaign model that excludes call duration produces a top-contact group with a response rate above the base rate.
- H4: The same managerial pattern appears in both datasets: marketing return is concentrated, not evenly distributed.

IV. PROPOSED METHOD

1. DATA SOURCES AND STUDIES

The online retail data is dated 2010 to 2011, so no retail observations fall outside the study period. The bank marketing file supplied for this analysis has month and day fields but no year column; because the public source is historical and no later records were introduced, all records were retained.

The source description for the online retail data is important. The data describe transnational transactions between 01/12/2010 and 09/12/2011 for a UK-based, registered non-store online retailer. The company mainly sells unique all-occasion gifts, and many customers are wholesalers. For this reason, this paper defines "customer" to include both retail and wholesale buyers. [32]

The bank marketing dataset represents phone-based direct marketing by an anonymized Portuguese banking institution. The public data does not disclose the bank name. The paper therefore focuses on response behaviour, not on the identity or reputation of a named bank. [33]

Table 2. Dataset screening.

Dataset	Original rows	Rows after 2020 filter	Clean analysis rows	Date handling
Online Retail	541,909	541,909	397,884	InvoiceDate parsed; all rows are before 2020.
Bank Marketing	45,211	45,211	45,211	No year column in file; retained all rows and did not add post-2020 data.

Both datasets fit the study window. The retail file is clearly dated before 2020; the bank file contains no post-2020 records in the supplied data.

2. DATA PREPARATION AND VARIABLES

For the online retail data, we removed canceled invoices, missing customer identifiers, and non-positive quantities and prices. We calculated revenue as quantity multiplied by unit price. We then calculated customer-level recency, frequency, and monetary value.

For the bank marketing data, the target variable is whether the customer subscribed to the term deposit. The analysis uses categorical association tests, numeric group comparisons, and supervised classification. Models are estimated with and without call duration. The no-duration specification is the main pre-campaign model because duration is only known after the call starts.

3. STATISTICAL AND MODELLING PROCEDURE

The statistical work was run in Python. We used pandas and NumPy for data handling, SciPy for statistical testing, and scikit-learn for modeling. We tested categorical relationships with chi-square statistics. Numeric group differences were examined with non-parametric tests because several variables are skewed. We compared logistic regression and random forest models using ROC AUC, precision-recall AUC, F1 score, and top-20% lift.

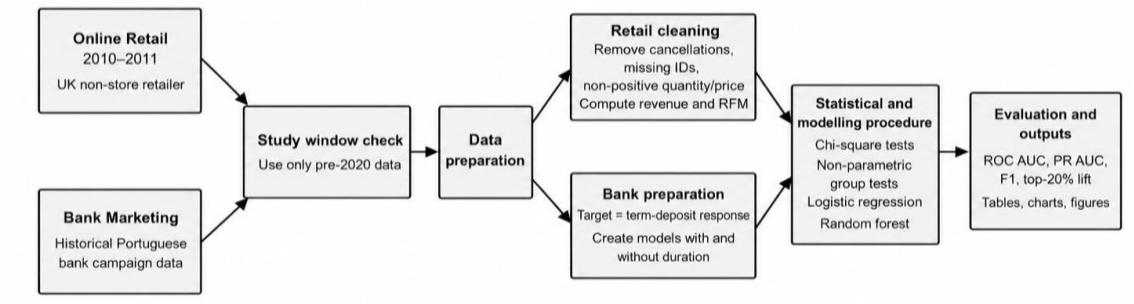


FIGURE 1. Workflow of data sources, preparation, and analysis.

The figure shows how we screened, prepared, analyzed, and evaluated the two historical datasets within a single marketing decision framework.

V. RESULTS

1. ONLINE RETAIL RESULTS

The retail results focus on customer value. The wholesaler note matters because a small number of buyers can naturally account for a large share of revenue in a gift and wholesale-oriented business.

Table 3. Online retail analysis base.

Measure	Value
Clean transactions	397,884
Unique customers	4,338
Unique invoices	18,532
Countries	37
Total revenue	8,911,407.90
Top 20% customer revenue share	74.6%

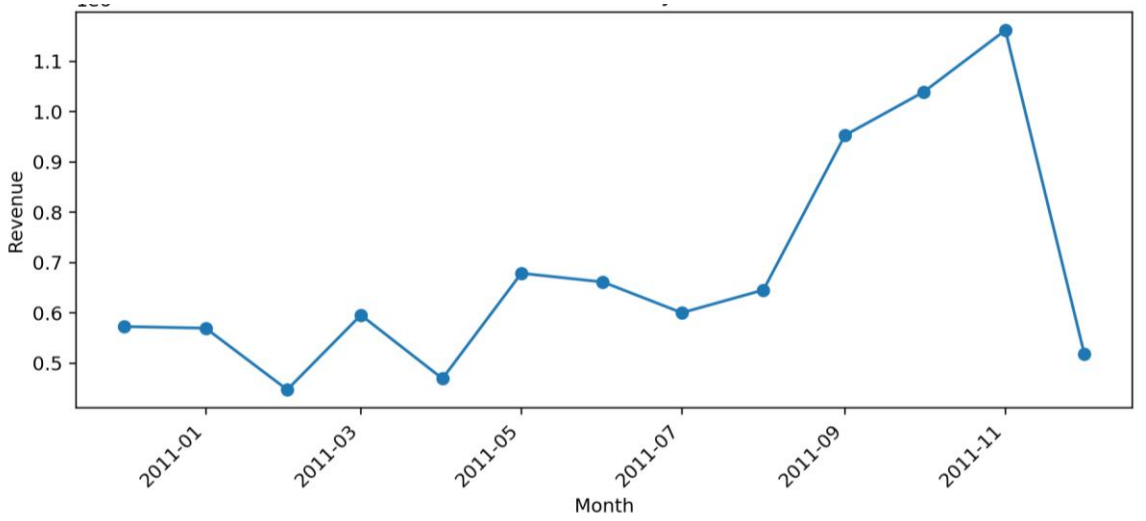


FIGURE 2. Monthly revenue trend.

Revenue rises strongly near the end of 2011. As a gift retailer, consider seasonality when scheduling campaigns.

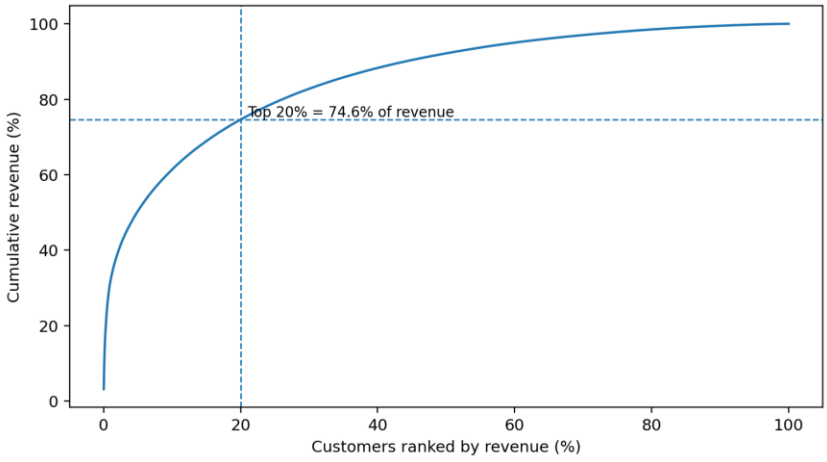


FIGURE 3. Customer revenue concentration.

The top 20% of customers produced 74.6% of revenue. This supports targeted retention rather than equal treatment for all buyers.

https://doi.org/10.48161/aaj.v2n3a91

Table 4. RFM customer segments.

Segment	Customers	Share	Mean recency days	Mean frequency	Mean monetary	Revenue share
Champions	1,314	30.3%	16.97	9.39	4,942.32	72.9%
Recent moderate buyers	655	15.1%	24.86	1.45	757.57	5.6%
Occasional at-risk buyers	922	21.3%	126.71	3.42	1,596.28	16.5%
Dormant low-value	1,447	33.4%	170.02	1.45	310.45	5.0%

The champion segment is the main retention target. It represents less than one-third of customers but accounts for nearly three-quarters of total revenue.

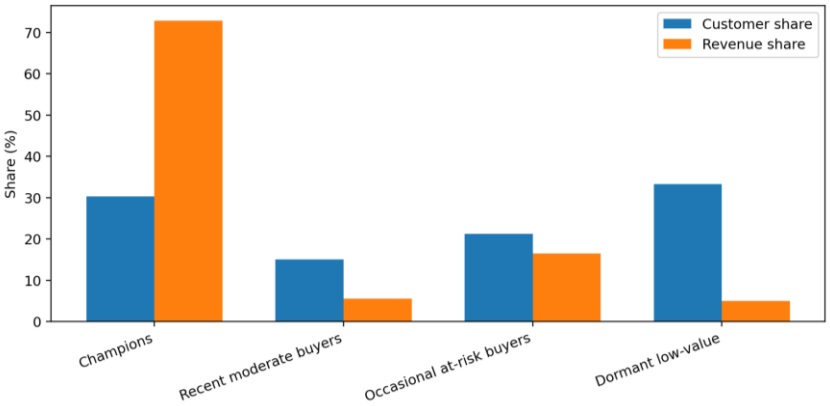


FIGURE 4. Segment size and revenue share.

The chart shows the gap between customer count and revenue. Some segments are large but low-value; the champion segment is smaller but much more profitable.

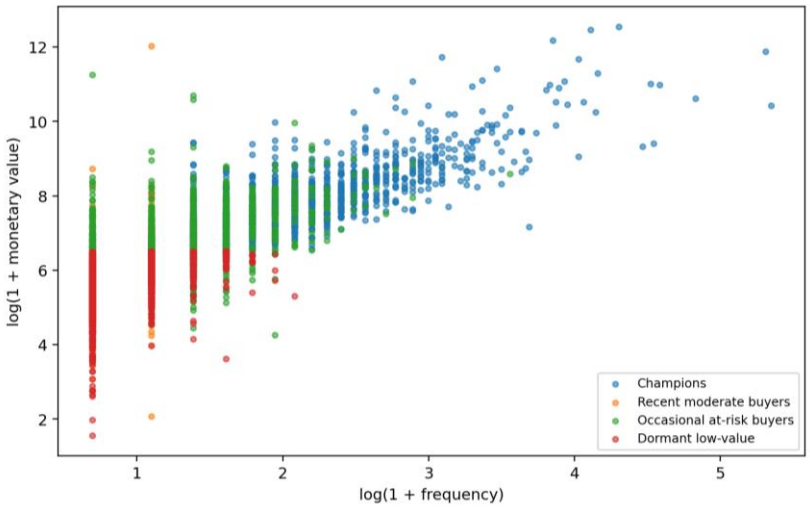


FIGURE 5. RFM cluster map.

Customers are separated mainly by frequency and money spent. This supports segment-based offers instead of one general message for every buyer.

Table 5. Spearman correlations among RFM measures.

Variable	Recency	Frequency	Monetary
Recency	1.00	-0.56	-0.48
Frequency	-0.56	1.00	0.81
Monetary	-0.48	0.81	1.00

Frequency and monetary value move strongly together. Recency moves in the opposite direction, which means recently active customers tend to be more valuable.

Table 6. Top revenue countries.

Country	Revenue	Customers	Invoices
United Kingdom	7,308,391.55	3,920	16,646
Netherlands	285,446.34	9	94
EIRE	265,545.90	3	260
Germany	228,867.14	94	457
France	209,024.05	87	389
Australia	138,521.31	9	57
Spain	61,577.11	30	90
Switzerland	56,443.95	21	51

The United Kingdom dominates the dataset. The findings should therefore be read mainly as evidence from the retailer's main market.

Table 7. Top revenue products.

Stock code	Description	Revenue	Quantity	Invoices
23843	PAPER CRAFT , LITTLE BIRDIE	168,469.60	80,995	1
22423	REGENCY CAKE STAND 3 TIER	142,592.95	12,402	1,703
85123A	WHITE HANGING HEART T-LIGHT HOLDER	100,448.15	36,725	1,971
85099B	JUMBO BAG RED RETROSPOT	85,220.78	46,181	1,600
23166	MEDIUM CERAMIC TOP STORAGE JAR	81,416.73	77,916	195
POST	POSTAGE	77,803.96	3,120	1,099
47566	PARTY BUNTING	68,844.33	15,291	1,379
84879	ASSORTED COLOUR BIRD ORNAMENT	56,580.34	35,362	1,375

A few products carry high revenue. The large single-invoice product values may reflect wholesale ordering, so product totals should be interpreted with care.

2. BANK MARKETING RESULTS

The bank results focus on response to a direct marketing offer. The response rate is low, so the practical task is to rank customers and contact the strongest prospects first.

Table 8. Bank campaign analysis base.

Measure	Value
Campaign contacts	45,211
Positive responses	5,289
Response rate	11.7%
Predictors used in pre-campaign model	16

Measure	Value
Predictor excluded from pre-campaign model	duration

The response rate is 11.7%, so accuracy alone is not enough. Ranking quality and lift are more useful for campaign planning.

Table 9. Response by previous campaign outcome.

Previous outcome	Contacts	Response rate
success	1,511	64.7%
other	1,840	16.7%
failure	4,901	12.6%
unknown	36,959	9.2%

Previous success is the strongest simple signal. Customers with a successful prior outcome respond much more often than customers with unknown history.

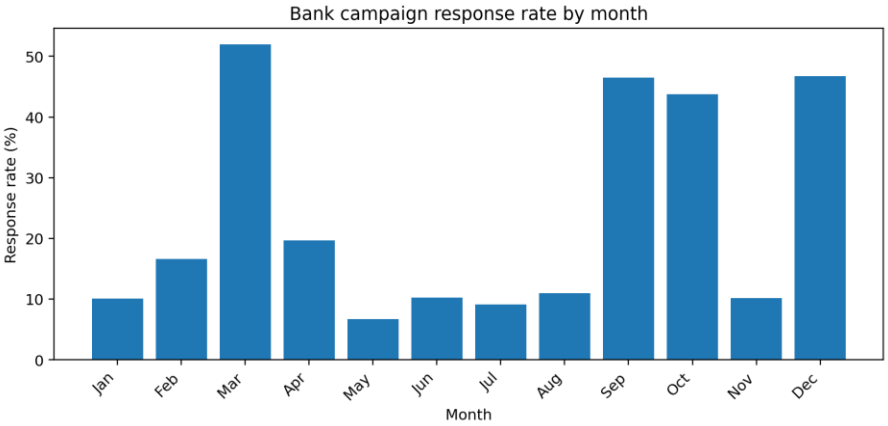


FIGURE 6. Campaign response by month.

The response rate changes by month. Timing is not just a calendar detail; it changes the expected return from a campaign contact.

Table 10. Campaign prediction model performance.

Model	ROC AUC	PR AUC	F1	Top 20% response	Top 20% lift
Logistic regression, no duration	0.773	0.416	0.374	33.1%	2.834
Random forest, no duration	0.801	0.459	0.454	36.2%	3.091
Logistic regression, with duration	0.909	0.537	0.558	45.4%	3.878
Random forest, with duration	0.926	0.609	0.559	47.3%	4.044

The no-duration random forest is the safer pre-campaign model. It reaches an ROC AUC of 0.801 without using information known only after a call starts.

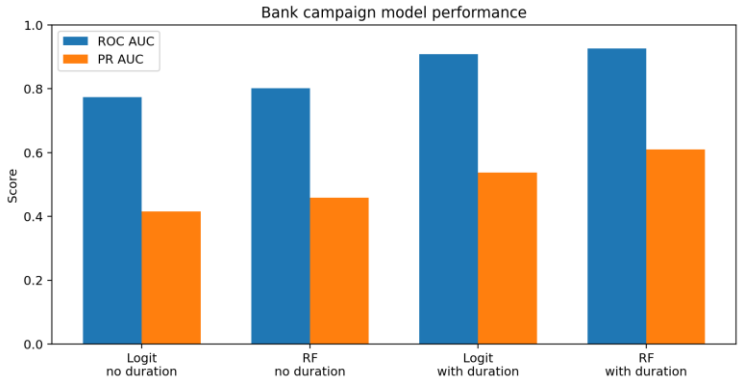


FIGURE 7. Model Performance Comparison.

Adding call duration improves the scores, but that model is better for explaining completed calls than for choosing whom to call in advance.

Table 11. Top Pre-campaign Predictors.

Feature	Coefficient	Odds ratio
cat__poutcome_success	1.67	5.29
cat__month_mar	1.38	3.98
cat__month_nov	-1.05	0.35
cat__month_aug	-0.83	0.43
cat__contact_unknown	-0.73	0.48
cat__month_jan	-0.73	0.48
cat__month_oct	0.73	2.07
cat__poutcome_failure	-0.66	0.52
cat__month_sep	0.64	1.89
cat__month_jul	-0.63	0.53

The most useful signals are available before the call: previous outcome, month, contact type, and basic financial status.

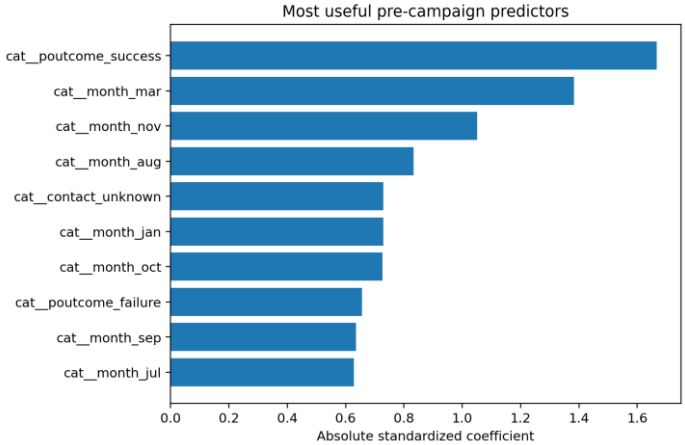


FIGURE 8. Pre-campaign Predictor Importance.

The figure makes the model easier to explain. It shows which pre-campaign signals push the ranking most strongly.

Table 12. Categorical relationships with campaign response.

Variable	Chi-square	p-value	Groups
poutcome	4,391.51	<0.001	4
month	3,061.84	<0.001	12
contact	1,035.71	<0.001	3
housing	874.82	<0.001	2
job	836.11	<0.001	12
education	238.92	<0.001	4
loan	209.62	<0.001	2
marital	196.50	<0.001	3

Campaign response is not randomly spread across customer groups. Prior outcome, month, contact channel, and housing status show strong relationships with response.

Table 13. Numeric relationships with campaign response.

Variable	Mean yes	Mean no	Median yes	Median no	p-value
previous	1.17	0.5	0	0	<0.001
pdays	68.7	36.42	-1	-1	<0.001
balance	1,804.27	1,303.71	733	417	<0.001
campaign	2.14	2.85	2	2	<0.001
day	15.16	15.89	15	16	<0.001
age	41.67	40.84	38	39	0.063

Responders have higher average balances and different contact histories. The size and direction of the differences matter more than p-values alone because the sample is large.

3. CROSS-DATASET INTERPRETATION

The datasets are linked by decision logic rather than by shared records. The retail data show where value is concentrated. The bank data show how to rank responses before a campaign. Together they support a process that starts with customer value and ends with contact prioritization.

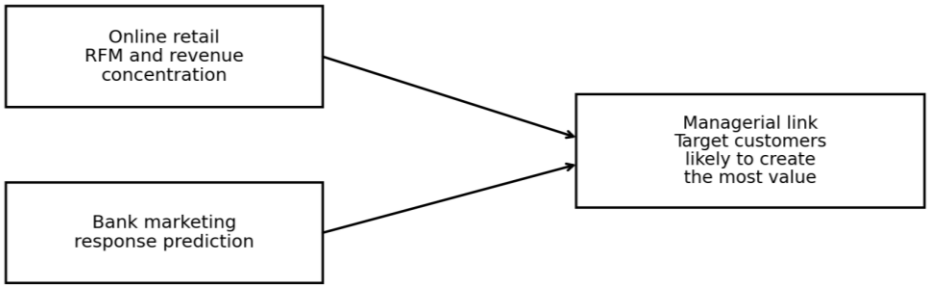


FIGURE 9. Analytical link between the two datasets.

The link is managerial, not technical. One dataset identifies value, and the other identifies the probability of response.

Table 14. Relation between retail value and campaign response.

Marketing relation	Online retail evidence	Bank campaign evidence	Decision use
Value concentration	Top 20% of customers generated 74.6% of revenue.	Top 20% model-ranked contacts had 36.2% response, about 3.09x the base rate.	Spend more effort where the return is concentrated.
Loyalty and response	Champions were 30.3% of customers and produced 72.9% of revenue.	Previous campaign outcome was the strongest categorical relation with response.	Use prior behavior to prioritize retention and next-best offers.
Timing matters	Monthly revenue peaked at 2011-11.	Highest monthly response rate was Mar at 52.0%.	Campaign timing should be planned, not treated as a fixed calendar routine.

VI. DISCUSSION

The findings give a practical message. A firm should not begin with the campaign channel only. It should first identify which customers matter most, then decide which are most likely to respond. In the retail data, value is concentrated among repeat and high-spending customers. In the bank data, response is concentrated among customers with favorable prior signals.

The online retail source note changes how we interpret the retail results. Because many buyers are wholesalers, revenue concentration should not be described as ordinary household shopping behavior. It is better to describe the evidence as retail and wholesale customer behavior inside a UK non-store retail business.

For the bank setting, the safest operational model is the model without call duration. A model with duration helps explain completed calls, but it is not suitable for choosing who to call before the campaign starts. This is why we treat the no-duration random forest as the main targeting model.

The broader implication is that marketing return is uneven. The retail data show uneven value. The bank data show uneven response. This pattern supports a managerial process based on ranking, prioritisation, and explainable decision rules.

VII. CONCLUSION

The findings show that marketing return is not evenly distributed. In the online retail data, a small share of customers creates most of the revenue. In the bank marketing data, a small group of model-ranked contacts produces a much higher response rate than the average campaign contact. This pattern supports a targeted marketing approach rather than treating all customers equally.

The retail findings align with RFM analysis and customer value segmentation. Customers who buy more often and spend more money form the strongest commercial group. In this study, the champion segment represents 30.3% of customers but produces 72.9% of revenue. This result gives managers a clear retention priority: protect the customers who already carry most of the business value.

The online retail context also matters. The retailer sells all-occasion gifts, and many customers are wholesalers. For that reason, the revenue concentration should not be interpreted as ordinary household shopping behavior only. Large orders and high product totals may reflect wholesale purchasing. This makes the results more useful for retail and wholesale customer management, stock planning, and targeted retention.

The bank marketing findings show that response prediction is useful as a ranking tool. The base response rate is 11.7%, but the top 20% of contacts ranked by the no-duration random forest model respond at 36.2%. This means the model not only improves statistical performance; it also helps answer a practical campaign question: whom should the firm contact first?

The difference between models with and without call duration is also important. Models that include duration perform better, but they use information that is only available after the call begins. For campaign planning, this is not suitable. The no-duration model is weaker in pure predictive terms, but it is more realistic for pre-campaign targeting. This makes it the safer model for managerial use.

The cross-dataset result is the main contribution. The datasets share no customers, but they show the same managerial pattern. In retail, value is concentrated. In banking, response is concentrated. Together, they support a marketing process based on ranking, prioritization, and explainable decision rules. Managers should first identify where value lies, then allocate campaign effort to customers or contacts with higher expected returns.

The findings also support the need for explainable modeling. A campaign model is more useful when managers can understand why some contacts are ranked higher than others. Previous outcomes, month, contact type, and financial indicators provide interpretable signals before a campaign starts. This reduces the risk of treating the model as a black box and makes the results easier to use in planning.

Overall, the study shows that customer analytics should not be treated as only a technical exercise. Its value appears when statistical results are connected to a real marketing decision. In this paper, that decision is clear: protect high-value customers and focus campaign effort on contacts with higher estimated response probability.

Author Contributions

The author conducted the conceptualization, methodology, data analysis, investigation, writing, review, editing, and final approval of the manuscript.

Funding

This research received no external funding.

Data Availability

The dataset will be available from the author upon reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

REFERENCES

[1]. Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. Procedia Computer Science, 3, 57-63. https://doi.org/10.1016/j.procs.2010.12.011

[2]. Chen, D., Sain, S. L., & Guo, K. (2012). Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. Journal of Database Marketing & Customer Strategy Management, 19(3), 197-208. https://doi.org/10.1057/dbm.2012.17

[3]. Migueis, V. L., Camanho, A. S., & e Cunha, J. F. (2012). Customer data mining for lifestyle segmentation. Expert Systems with Applications, 39(10), 9359-9366. https://doi.org/10.1016/j.eswa.2012.02.133

[4]. Verbeke, W., Dejaeger, K., Martens, D., Hur, J., & Baesens, B. (2012). New insights into churn prediction in the telecommunication sector: A profit-driven data mining approach. European Journal of Operational Research, 218(1), 211-229. https://doi.org/10.1016/j.ejor.2011.09.031

[5]. Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS Quarterly, 36(4), 1165-1188. https://doi.org/10.2307/41703503

[6]. Germann, F., Lilien, G. L., & Rangaswamy, A. (2013). Performance implications of deploying marketing analytics. International Journal of Research in Marketing, 30(2), 114-128. https://doi.org/10.1016/j.ijresmar.2012.10.001

[7]. Leeflang, P. S. H., Verhoef, P. C., Dahlstrom, P., & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. European Management Journal, 32(1), 1-12. https://doi.org/10.1016/j.emj.2013.12.001

[8]. Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. Decision Support Systems, 62, 22-31. https://doi.org/10.1016/j.dss.2014.03.001

[9]. Germann, F., Lilien, G. L., Fiedler, L., & Kraus, M. (2014). Do retailers benefit from deploying customer analytics? Journal of Retailing, 90(4), 587-593. https://doi.org/10.1016/j.jretai.2014.08.002

[10]. Fan, S., Lau, R. Y. K., & Zhao, J. L. (2015). Demystifying big data analytics for business intelligence through the lens of marketing mix. Big Data Research, 2(1), 28-32. https://doi.org/10.1016/j.bdr.2015.02.006

[11]. Safari, F., Safari, N., & Montazer, G. A. (2016). Customer lifetime value determination based on RFM model. Marketing Intelligence & Planning, 34(4), 446-461. https://doi.org/10.1108/MIP-03-2015-0060

[12]. Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. Journal of Marketing, 80(6), 97-121. https://doi.org/10.1509/jm.15.0413

[13]. Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. Journal of Business Research, 69(2), 897-904. https://doi.org/10.1016/j.jbusres.2015.07.001
[14]. Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. Journal of Marketing, 80(6), 69-96. https://doi.org/10.1509/jm.15.0420
[15]. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135-1144). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939778
[16]. Kannan, P. K., & Li, H. A. (2017). Digital marketing: A framework, review and research agenda. International Journal of Research in Marketing, 34(1), 22-45. https://doi.org/10.1016/j.ijresmar.2016.11.006
[17]. Grewal, D., Roggeveen, A. L., & Nordfalt, J. (2017). The future of retailing. Journal of Retailing, 93(1), 1-6. https://doi.org/10.1016/j.jretai.2016.12.008
[18]. Bradlow, E. T., Gangwar, M., Kopalle, P., & Voleti, S. (2017). The role of big data and predictive analytics in retailing. Journal of Retailing, 93(1), 79-95. https://doi.org/10.1016/j.jretai.2016.12.004
[19]. Kunz, W. H., Aksoy, L., Bart, Y., Heinonen, K., Kabadayi, S., Ordenes, F. V., Sigala, M., Diaz, D., & Theodoulidis, B. (2017). Customer engagement in a Big Data world. Journal of Services Marketing, 31(2), 161-171. https://doi.org/10.1108/JSM-10-2016-0352
[20]. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In I. Guyon et al. (Eds.), Advances in Neural Information Processing Systems 30 (pp. 4765-4774). Curran Associates.
[21]. Ascarza, E., Neslin, S., Netzer, O., Anderson, Z., Fader, P. S., Gupta, S., Hardie, B. G. S., Lemmens, A., Libai, B., Neal, D., Provost, F., & Schrift, R. (2018). In pursuit of enhanced customer retention management: Review, key issues, and future directions. Customer Needs and Solutions, 5(1-2), 65-81. https://doi.org/10.1007/s40547-017-0080-0
[22]. Ascarza, E. (2018). Retention futility: Targeting high-risk customers might be ineffective. Journal of Marketing Research, 55(1), 80-98. https://doi.org/10.1509/jmr.16.0163
[23]. De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. European Journal of Operational Research, 269(2), 760-772. https://doi.org/10.1016/j.ejor.2018.02.009
[24]. Adadi, A., & Berrada, M. (2018). Peeking inside the black box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138-52160. https://doi.org/10.1109/ACCESS.2018.2870052
[25]. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. ACM Computing Surveys, 51(5), Article 93. https://doi.org/10.1145/3236009
[26]. Reinartz, W., Wiegand, N., & Imschloss, M. (2019). The impact of digital transformation on the retailing value chain. International Journal of Research in Marketing, 36(3), 350-366. https://doi.org/10.1016/j.ijresmar.2018.12.002
[27]. Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. Journal of the Academy of Marketing Science, 48, 24-42. https://doi.org/10.1007/s11747-019-00696-0
[28]. Hossain, M. A., Akter, S., & Yanamandram, V. (2020). Revisiting customer analytics capability for data-driven retailing. Journal of Retailing and Consumer Services, 56, 102187. https://doi.org/10.1016/j.jretconser.2020.102187
[29]. Verhoef, P. C., Broekhuizen, T. L. J., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. Journal of Business Research, 122, 889-901. https://doi.org/10.1016/j.jbusres.2019.09.022
[30]. Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. Journal of the Academy of Marketing Science, 49, 30-50. https://doi.org/10.1007/s11747-020-00749-9
[31]. Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. Journal of Marketing, 85(1), 131-151. https://doi.org/10.1177/0022242920953847
[32]. UCI Machine Learning Repository. (2015). Online Retail dataset. https://archive.ics.uci.edu/dataset/352/online+retail
[33]. UCI Machine Learning Repository. (2012). Bank Marketing dataset. https://archive.ics.uci.edu/dataset/222/bank+marketing