

І. В. Мірошніченко,
к. е. н., доцент кафедри математичного моделювання та статистики,
Київський національний економічний університет імені Вадима Гетьмана
ORCID ID: 0000-0002-1307-7889

В. Д. Алексєєва,
магістрант спеціальності "Економічна кібернетика та Дата Сайанс",
Київський національний економічний університет імені Вадима Гетьмана
ORCID ID: 0000-0002-1001-6933

DOI: 10.32702/2306-6806.2022.1.114

ВИКОРИСТАННЯ RFM-АНАЛІЗУ В СЕГМЕНТАЦІЇ КЛІЄНТІВ

I. Miroshnychenko,
PhD in Economics, Associate Professor of the Department of Mathematical Modeling
and Statistics, Kyiv National Economic University named after Vadym Hetman

V. Aleksieieva,
Master's student of "Economic cybernetics and Data science",
Kyiv National Economic University named after Vadym Hetman

USE OF RFM ANALYSIS IN CUSTOMER SEGMENTATION

Статтю присвячено застосуванню RFM-сегментації та алгоритму Аpriori Аpriori для виявлення сегментів клієнтів та асоціативних правил у споживчих кошиках. У результаті дослідження було вивчено поняття RFM-аналізу, визначено його переваги, етапи побудови моделі, способи застосування на практиці. Надано рекомендації щодо підготовки даних для виконання аналізу, визначено обов'язкові дані. Також ознайомлено з алгоритмом Аpriori для виявлення асоціативних правил у споживчих кошиках. Було визначено сфери застосування, його сильні та слабкі сторони, запропоновано загальний алгоритм його побудови та вивчено, як серед множини результатів знайти ключові правила. Після теоретичного ознайомлення було побудовано RFM-модель та використано алгоритм Аpriori. Отримані результати проаналізовано, запропоновано подальші кроки дій для компаній, а саме: розробка комунікаційної стратегії для збільшення обсягів продажів завдяки email-ретаргетингу та таргетингу в соціальних мережах. Запропоновано, як можна проваджувати отримані результати для створення спеціальних пропозицій та рекламних оголошень.

Digital technologies cease to seem complex and immensity and begin to necessitate their use in business. This is not about IT startups, but about the digitalization of internal business processes. One of the tools of a successful company is a well-developed marketing company that meets modern market conditions. Moreover, during quarantine restrictions, more and more companies are forced to go online and face the need to improve or only implement digital marketing technologies.

The article is devoted to the application of RFM-segmentation and the Apriori Apriori algorithm to identify customer segments and associative rules in consumer baskets. As a result of research the concept of RFM-analysis was studied, its advantages, stages of construction of model, ways of application in practice were defined. Recommendations for preparing data for analysis are provided, mandatory data are defined. Also acquainted with the Apriori algorithm for identifying associative rules in consumer baskets. The areas of application, its strengths and weaknesses were identified, a general algorithm for its construction was proposed, and how to find key rules among the many results was studied. After theoretical acquaintance, the RFM model was built and the Apriori algorithm was used. The obtained results are analyzed, further steps for companies are proposed, namely the development of a communication strategy to increase sales through email retargeting and targeting on social networks. Suggest how you can implement the results to create special offers and advertisements.

Search and development of effective algorithms and their implementation in customer relationship management systems allows to optimize work with customers in all directions and, above all, to conduct sound marketing activities designed, according to the identified benefits, directly for target segments of the customer base, ie to send advertising information is not chaotic, but purposeful for customer segments that are interested in this information and will most likely have a response.

The work is of practical importance for all companies that have a database of transactions (sales) with the specified customer ID, because the paper presents a step-by-step algorithm for data analysis and provides recommendations for further work with the obtained segments. The study also has universal theoretical and methodological significance: the method of combining RFM-analysis and analysis of consumer baskets, providing recommendations for the use of email-retargeting and targeted advertising on social networks can contribute to further in-depth study and improvement of algorithms, not only in marketing but, web-programming, medicine.

Ключові слова: сегментація клієнтів, комунікаційна стратегія, RFM-аналіз, алгоритм Аpriori, споживчий кошик, асоціативні правила, таргетована реклама.

Key words: customer segmentation, communication strategy, RFM-analysis, A priori algorithm, consumer basket, associative rules, targeted advertising.

ПОСТАНОВКА ПРОБЛЕМИ

Цифрові технології перестають здаватися складними і неосяжними та починають викликати необхідність їх використання в бізнесі. Мова йде не про IT-стартапи, а про цифровізацію внутрішніх процесів бізнесу. Одним з інструментів успішної компанії є розвинені маркетингові компанії, які відповідають сучасним умовам ринку. Тим більше, під час карантинних обмежень, все більше підприємств вимушені переходити в онлайн-режим і стикаються з необхідністю вдосконалювати або тільки впроваджувати технології цифрового маркетингу. В світі сучасного маркетингу актуальним є інструмент таргетованої реклами, який має значні переваги, одні з них — зниження вартості рекламних кампаній та підвищення ефективності. Але для налаштування потрібно мати чітке представлення стану своєї аудиторії та їх потреб. Використані в роботі технології допоможуть будь-якій компанії, яка веде базу даних транзакцій, отримати чітке уявлення клієнтської бази та проаналізувати споживчі кошики. Цих даних буде достатньо для складання рекламних кампаній і відстежування їх ефективності.

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Було ознайомлено з іноземними роботами на тему сегментації клієнтської бази та RFM-аналізу, значущим дослідником можна вважати Понлача Ройлєртяна з Вищої школи Бангкокського університету, який в 2019 році випустив незалежний дослідницький рукопис, де розгорнуто дослідив процес сегментації клієнтів. Також важливо зазначити сучасні роботи Мохамеда Абдул Кадіра та Адріана Ачяра, які на практиці продемонстрували повний аналіз великих даних та сегментування; і Рой Волен, який надав ґрунтовну відповідь, чому RFM-аналіз — це зручно і ефективно. Перелічені дослідники внесли вагомий внесок у дослідження, але роботи були або суто теоретичні, або тільки практичні. Не вистачало цілісного аналізу даних з конкретними рекомендаціями подальших дій, також не запропоновано паралельні дослідження, які підсилять користь RFM-аналізу.

МЕТА СТАТТІ

Метою дослідження є отримання сегментації клієнтів, виявлення закономірностей в споживчих кошиках та надання рекомендації щодо подальшої побудови комунікаційної стратегії.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ ДОСЛІДЖЕННЯ

RFM-аналіз — аналіз клієнтської бази на основі історії покупок. Сама назва складається з абrevіатури його складових: Resency, Frequency, Monetary.

1. Resency (укр. новизна, недавність) показує, скільки часу пройшло з останньої покупки або іншої активності клієнта.

2. Frequency (укр. частота) показує кількість покупок клієнта за певний період. Якщо споживач часто

здійснює покупки, його легше стимулювати купувати більше на більшу суму.

3. Monetary (укр. гроші, фінанси) є грошовим показником суми покупок клієнтом за певний проміжок часу. Це важлива частка клієнтів, які можуть купувати дорогі товари або просто у великій кількості [30].

Модель RFM дозволяє підприємствам отримувати ключові уявлення про клієнтів за допомогою зручного збору даних та формувати бізнес-стратегію. Модель дозволяє зрозуміти, що означає бренд компанії для існуючих клієнтів, допомагає компаніям керувати сприйняттям клієнтів [28].

Нецільові маркетингові рекламні кампанії можуть бути дорогими. Зосередження уваги на меншому сегменті клієнтів значно зменшить витрати, дозволить більше експериментувати та приймати рішення на основі даних, а не інтуїції. Завдяки RFM-аналізу можна розподілити споживачів на сегменти і будувати маркетингову політику для кожної окремо [22]. Перевагою RFM-аналізу можна назвати його застосовність для будь-якої сфери і масштабу бізнесу. Процес проведення аналізу буде завжди один, відрізнятимуться масштаби бази клієнтів і маркетингова стратегія. [3].

Основи побудови RFM-моделі. RFM-модель є індивідуальною. Тобто кожна компанія обирає самостійно для себе, за якими конкретно критеріями розглядати елементи та яку градацію використовувати [19; 45].

У RFM-моделі можуть бути різні кількість кластерів. Можна зробити 3 основні групи по трьох елементах RFM, можна 4 та 5. В кінці організація отримує 27, 64, 125 сегментів відповідно. Чим більше сегментів, тим більше аналітичної інформації компанія може краще пристосовуватися до кожної групи, складаючи для них різні стратегічні плани. Але це залежить від обсягів клієнтської бази, якщо вона не є великою (до 10 000 транзакцій), то дрібна сегментація навпаки стане нерезультативною [3; 7].

Основним моментом при сегментуванні є оцінка адекватності розподілу балів R, F та M для кожного клієнта та присвоєння йому сегменту RFM відповідно до поведінки покупки [29].

Після отриманої сегментації клієнтів розробляється маркетингова стратегія для кожної групи. Метою компанії буде збільшення продажів і середнього чеку для "кращих" клієнтів, стимулювання до покупок "середніх" клієнтів, та повернення "гірших". Як етап маркетингової стратегії, потрібно вирішити, які дії будуть направлені на кожний сегмент [28]. Щоб рекомендації були корисними для всіх RFM-аналізів, ми будемо використовувати такі позначення сегментів, як у таблиці 1.

До першої групи можна віднести перші шість сегментів з таблиці 1. Головна мета роботи з такими клієнтами — збільшити обсяги покупок, їх лояльність до компанії. Наприклад, можна запускати e-mail розсилки, щоб нагадувати користувачам оплачувати щомісячні підписки, можна інформувати про новинки товарів, спеціальні пропозиції [15]. Але слід зауважити,

Таблиця 1. Сегментація клієнтської бази

| Сегмент | Пояснення |
|------------------------------|--|
| Чемпіони | Купували недавно, купують часто і витрачають найбільше |
| Лояльні клієнти | Купують регулярно, особливо впливає акційна компанія |
| Потенційні лояльні клієнти | Нещодавні клієнти із середньою частотою |
| Нещодавні клієнти | Купували зовсім недавно, але не часто |
| Перспективні | Недавні покупці, що не витратили багато |
| Клієнти, що потребують уваги | Вище середнього рівня відвідуваності до певного моменту, частоти та грошових значень |
| Сплячі | Середня частота покупок і середній показник останнього відвідування. Якщо їх не буде активовано, їх можна втратити |
| Під загрозою відтоку | Купували часто, але втратили інтерес |
| Важливо не втратити | Були постійними клієнтами, але вже давно нічого не купують. Потрібно повернути їх |
| Втрачені | Остання покупка була тривалою і мала кількість замовлень в минулому |

Джерело: [27].

що покупки акційних товарів, особливо, коли клієнт обмежений в часі, відносяться до імпульсивних, спонтанних, а пошту клієнт може не відкривати певний час [34]. Тому для акційних рекламних кампаній ефективніше використовувати інструмент таргетованої реклами на залучену аудиторію в соціальних мережах [44]. Компанія може зв'язатися з "Чемпіонами" та "Лояльними клієнтами" ще до створення продукту. Вони можуть надати якісні рекомендації, на які характеристики робити акцент, що покращити в старих версіях [2].

До другої групи ми можемо віднести останні чотири сегменти. В цьому випадку головною метою буде нагадати про минулі покупки і змотивувати їх почати купувати (на відміну від першої групи, нам не важливий розмір витраченої суми). В email-розсилках вже неактуально нагадувати про забуті кошики, якщо клієнт останній раз заходив на сайт рік тому. Тому замість цього можна надіслати оголошення про оновлення товару та лист про повторне залучення [41, 44].

Компаніям важливо раціонально оцінити бюджет на повернення таких клієнтів. Споживачі, що "У зоні ризику" та "Сплячі" — єдині два перспективні сегменти, на які потрібно звернути особливу увагу. Якщо ж компанія має достатній бюджет, то компанії ретаргетингу в соціальних мережах та електронній пошті буде достатньо [28; 39].

Поведінка споживача націлена на прийняття рішення щодо придбання і споживання певних товарів. Кожен клієнт має різні потреби, смаки, бюджет, які можуть впливати на його рішення. Однак при цьому поведінка споживачів має подібність в тому, що кожен бажає максимізувати користь від споживання товару X [10].

Роблячи покупки в супермаркеті, споживачі рідко купують один товар, вони набагато частіше купують кошик товарів, переважно з різних категорій [1]. Метою роботи є відокремлення продуктів/груп товарів, які купуються разом у формі правил асоціації [10]. Транзакції визначаються наборами товарів, куплених клієнтом під час одного відвідування магазину [42].

Асоціативні правила використовуються не для прогнозування, а для невідконтрольного пошуку знань у великих базах даних, тобто немає необхідності в навчанні алгоритму, цим вони відрізняються від сегментації та числового прогнозування [24]. Недоліком цього процесу є те, що не існує простого способу об'єктивно виміряти ефективність алгоритму навчання асоціативних правил, крім якісної оцінки користьності [10].

Аналіз споживчого кошика в основному використовується у трьох сферах [24]. Перша — це створення персоналізованих рекомендацій [38; 43]. Друга — це просторовий розподіл в мережах магазинів (мерчандайзинг) [18; 38]. Третя — створення маркетингових стратегій, орієнтованих на знижки та акції, що базується виключ-

но на поведінці клієнта [38]. На основі отриманих правил можна: змінити макет вебсайту та дизайн каталогу, менеджери по продажам можуть використовувати дані для складання скрипта продажу індивідуально під клієнта [43].

Якщо обчислювати правила асоціації за щомісячними даними про продажі, ми можемо спостерігати сезонні зміни, коли певні правила відповідають дійсності приблизно в один і той же місяць кожного року [21].

Розглянемо алгоритм Apriori для пошуку асоціативних правил. Найбільш популярний підхід для ефективного пошуку у великих базах даних правил — це Apriori, який був створений ще в 1994 році [10; 38].

Чи вважати асоціативне правило цікавим, алгоритм визначає за двома статистичними показниками: підтримці і довірі [42].

Підтримка (support) правила — це частота появи цього правила серед даних [43]. Визначається за такою формулою:

$$support(X) = \frac{\{X \in N\}}{count(X)} \quad (1),$$

де N — кількість транзакцій в базі даних; $count(X)$ — кількість транзакцій, які містять безліч елементів X [42].

Довіра (confidence) правила — це міра здатності прогнозування:

$$confidence(X \rightarrow Y) = \frac{support(X, Y)}{support(X)} \quad (2),$$

де $support(X, Y)$ — підтримка безлічі елементів, що містять елементи X і Y ;

$support(X)$ — підтримка безлічі елементів, що містить тільки X [42].

Третій показник ліфт (англ. Lift) показує, наскільки ймовірність придбання одного або декількох товарів вище, якщо відомо, що при цьому був придбаний інший товар або кілька товарів:

$$lift(X \rightarrow Y) = \frac{confidence(X \rightarrow Y)}{support(Y)} \quad (3),$$

де $confidence(X \rightarrow Y)$ — показник довіри безлічі елементів, що містять елементи X і Y ;

$support(Y)$ — підтримка безлічі елементів, що містить тільки X [42].

Щоб провести RFM-аналіз і використати алгоритм Apriori, знадобляться дані всіх транзакцій компанії. Їх можна отримати завдяки CRM-системі, картки лояльності, інформаційної системи вебсайту. Знадобиться унікальний ідентифікатор клієнта (ID, email, ПІБ, номер телефону), унікальний ідентифікатор транзакції, дата транзакції, унікальний ідентифікатор придбаного товару, сума покупки [42]. Першим рекомендуємо проводи-

Таблиця 2. Розподіл RFM-показників за сегментами

| Сегмент | Показник R | Показник F | Показник M |
|------------------------------|------------|------------|------------|
| Чемпіони | 4–5 | 4–5 | 4–5 |
| Лояльні клієнти | 2–4 | 3–4 | 4–5 |
| Потенційні лояльні клієнти | 3–5 | 1–3 | 1–3 |
| Нещодавні клієнти | 4–5 | 1–2 | 1–2 |
| Перспективні | 3–4 | 1–2 | 1–2 |
| Клієнти, що потребують уваги | 3–4 | 3–4 | 3–4 |
| Сплячі | 2–3 | 1–3 | 1–3 |
| Під загрозою відтоку | 1–3 | 2–5 | 2–5 |
| Важливо не втратити | 1–2 | 4–5 | 4–5 |
| Втрачені | 1–2 | 1–2 | 1–2 |

Джерело: [29].

ти RFM-аналіз, бо він дає базову інформацію про стан бази клієнтів, дозволяє побудувати маркетингову таргетовану стратегію по сегментах. А аналіз споживчих кошків доповнює інформацію і конкретизує дії компанії з кожним клієнтом окремо.

Основні кроки побудови RFM-моделі для сегментації:

Крок 1. Збираємо інформацію за показниками, вказаними вище.

Можливо, отриману базу даних треба буде привести в необхідний для нас вигляд. Потрібно самостійно розрахувати суму чеку, якщо буде інформація ціни товару та купленої кількості; можливо, строки будуть відображати не одну транзакцію, а проданий один товар, тому аналітик має згрупувати дані за ID транзакціями [25; 28].

Крок 2. Для кожного клієнта визначаємо суму всіх чеків, кількість виконаних транзакцій та кількість днів з моменту останньої покупки.

Для розрахунку загальної суми треба згрупувати базу даних за ID клієнта і просумувати дані суми чеків. Аналогічно, щоб дізнатись кількість виконаних транзакцій для кожного клієнта, потрібно згрупувати базу даних за клієнтами і розрахувати кількість таких строк [35]. Щоб отримати, скільки днів тому була виконана остання транзакція, для початку в програмному коді потрібно визначити точку відліку. А потім в окремий стовпець розрахувати, яка різниця між максимальною датою транзакцій клієнта та встановленою датою [16; 23].

Крок 3. Кожному клієнтові присвоюємо бали показників: R, F, M.

На цьому етапі ми маємо всі необхідні дані по клієнтах, а також вирішені діапазони для R, F, M. Тому ми можемо отримати три стовпця, де зазначені бали трьох показників. Надалі треба об'єднати три стовпця в один, щоб отримати кінцевий сегмент клієнта, він буде виглядати як трьохзначний код сегмента [2].

Крок 4. Аналіз отриманих сегментів.

Якщо компанія відокремила 3 групи в показниках, то в результаті отримала 27 сегментів. Будувати 27 комунікаційних стратегій буде не раціонально: грані між сусідніми сегментами не такі значні, та кількість клієнтів у деяких сегментах може бути незначною, щоб витратити час на окрему комунікацію з ними. Тому кожна компанія для себе виділяє кінцеві сегменти [33]. В таблиці 2 ми пропонуємо орієнтовний розподіл RFM-показників за сегментами (враховуючи, що оцінка "5" — найкраща, "1" — найгірша).

Крок 5. Побудова комунікаційної стратегії.

На останньому кроці маркетингологи розробляють комунікацію для кожного сегмента.

Розглянемо процес формування асоціативних правил за допомогою алгоритму Apriori. Процес формування правил відбувається в два етапи.

1. Визначення всіх множин товарів, які відповідають мінімальному порогу підтримки, що зазначається вручну спеціалістом [42].

Цей етап виконується за декілька ітерацій. Кожна наступна ітерація включає в себе оцінку підтримки всіх крупніших множин елементів. Результатом кожної ітерації і є множина, що складається з всіх і-елементних множин, що відповідають мінімальному порогу підтримки. Так, якщо $\{A\}$, $\{B\}$ і $\{C\}$ зустрічаються на першій ітерації часто, а $\{D\}$ — нечасто, то на другій ітерації будуть розглядатися тільки $\{A, B\}$, $\{A, C\}$ і $\{B, C\}$. Таким чином, алгоритму залишається оцінити лише три множини елементів, а не шість, які були б сформовані з елементом D. Для наочності формування правил у таблиці 3 представлено ітерації для визначення всіх множин товарів, які відповідають мінімальному порогу підтримки [38; 42].

2. Генерація асоціативних правил для всіх можливих підмножин, отриманих в на першому етапі, які відповідають мінімальному рівню довіри, що визначається людиною вручну [42]. Наприклад, $\{A, B\}$ призведе до створення правил $\{A\} \rightarrow \{B\}$ і $\{B\} \rightarrow \{A\}$. Ці правила оцінюються за мінімальним порогом довіри. Ті, що не відповідають вимогам — виключаються [1].

Кроки отримання набору правил за принципом Apriori:

Крок 1. Збір даних.

Для алгоритму Apriori потрібно записувати ID товарів по мірі їх додавання в кошик. Дослідникам потрібно вирішити, на якому рівні будуть розглядатися товари [17]. Чи достатньо їм отримати правило за групами товарів, наприклад, чи вони хочуть досліджувати окремі одиниці товару по типу.

Крок 2. Підготовка даних.

Для алгоритму дані треба завантажувати у вигляді розрідженій матриці, де кожному товару відповідає окремий стовпець (ознака). Побудувати матрицю допоможе спеціальний пакет "arules" у середовищі R [42].

Крок 3. Навчання моделі на даних.

Крок містить підбір показників довіри та підтримки. Технічно це виконати просто за допомогою однієї строки коду. А щодо визначення показників, то тільки методом проб і помилок можна знайти значення підтримки і довіри, які дозволять отримати розумну кількість асоціативних правил. Якщо встановити ці рівні занадто високими, то правила виявляться занадто елементарними, щоб бути корисними. Надто низький поріг може привести до громіздкому набору правил [42].

Крок 4. Оцінка ефективності моделі.

На даному кроці основною роботою буде поділ асоціативних правил на наступні категорії: корисні, тривіальні, незрозумілі [42]. Корисні правила дають чітку інформацію, що можна використати на практиці. Тривіальні правила — це будь-які правила, які настільки очевидні, що про них не варто згадувати: вони зрозумілі, але не приносять користі [18]. Правила є незрозумілими, якщо взаємозв'язок між елементами неясний, тому зрозуміти, як використовувати цю інформацію, неможливо [42]. При наявності достатньої кількості часу можна аналізувати кожне правило, щоб знайти цінні

Таблиця 3. Ітерації відбору множин алгоритму Apriori

| Номер ітерації | Правила, які алгоритм повинен оцінити | Правила, які алгоритм в результаті оцінює | Множини елементів, які задовольняють мінімальну підтримку | Множини елементів, які не задовольняють мінімальну підтримку |
|----------------|---|---|---|--|
| 1 | {A}, {B}, {C}, {D} | {A}, {B}, {C}, {D} | {A}, {B}, {C} | {D} |
| 2 | {A,B}, {A,C}, {A,D}, {B,C}, {B,D}, {C,D} | {A,B}, {A,C}, {B,C} | {A,B}, {B,C} | {A,C} |
| 3 | {A,B,C} | Алгоритм зупиняється | | |

Джерело: [42].

рідкісні правила. Але якісно аналізувати результати можуть лише експерти в цій галузі [38].

Крок 5. Підвищення ефективності моделі.

На цьому етапі спеціаліст проаналізував отримані правила і відокремив пару груп цікавих. Але потрібно розуміти, на скільки воно важливе для даних товарів. Наприклад, для транзакції {збиті вершки} → {заморожена вишня} дивимось всі правила де в лівій частині вишня. Знаходимо правило {заморожена вишня} → {збиті вершки} з достатнім рівнем ліфта. Таким чином ми упевнюємося, що ця пара працює ефективно в обидві сторони, тому є сенс впроваджувати правило на практиці [42].

У роботі я буду використовувати набір даних Online Retail з Інтернет-джерела Kaggle [40]. Датасет описує клієнтську базу мережевого онлайн-магазину товарів для дому з роздрібних та оптових продажів. Рядки позначають проданий товар за 1 транзакцією певним клієнтом, стовпці в базі даних наступні: ID транзакції, ID товару, назва товару, кількість купленого товару, дата і час покупки, ціна товару, ID клієнта, країна, де відбулась покупка.

Для проведення попереднього аналізу і подальшої роботи за допомогою мови програмування R. Бібліотеки: readxl [32]; dplyr [9]; doby [8]; ggplot2 [6, 12, 13, 20, 37]; forcats [11]; scales [36]; formattable [26]; rcolorbrewer [31]; ggpubr [14, 27].

В базі даних 541909 рядків, що описують 25900 транзакцій з 38 країн; всього 4070 видів товарів в асортименті; 4373 зареєстрованих клієнтів. Транзакції описані за період 01.12.2010 8:26:00 — 09.12.2012 12:50:00. Провівши попередній аналіз за країнами, явним лідером за всіма показниками є Великобританія. За доходами вона має 6767873,394 доларів і посягає перше місце.

Перейдемо до побудови RFM-моделі для сегментації клієнтів.

Для даного пункту ми обрали мову програмування Python, тому що вона дозволяє побудувати наглядні зрозумілі графіки для RFM-сегментації.

База даних онлайн-магазину велика, тому будемо розподіляти показники R, F, M за 5-бальною шкалою (5 балів — найкращий результат).

Етапи побудови RFM-моделі:

1. На початку завантажуюмо бібліотеки і завантажуюмо датафрейм.

2. Видаляємо рядки, де CustomerID не заповнений. У результаті ми видалили 135 080 строк, бо вони не допоможуть у нашому аналізі клієнтів.

3. Визначаємо, за який проміжок часу ми маємо дані. За термін дослідження беремо рік: 365 днів.

4. Обчислюємо загальні витрачені кошти кожного рядка (тобто купленого товару). Для цього помножимо кількість товару на ціну і додаємо окремий стовпець з результатами.

5. В наших даних строки позначають проданий товар з асортименту, а нам потрібно мати інформацію за транзакцію. У випадку RFM-моделі ми шукаємо скільки разів купував клієнт, а не скільки предметів [33]. Тому ми створюємо новий датафрейм замовлень, який згрупований за транзакціями і має суму чеку куплених товарів.

6. Встановимо дату "NOW", як сутки після останньої покупки. Ця дата буде використана як орієнтир для обчислення оцінки R, як точка відліку.

7. Нехай маркетологи онлайн-магазину вважають, що на повернення клієнтів, які купували більше року, тому вони не хочуть витрачати рекламний бюджет, тому ми встановлюємо період = 365 днів.

8. Починаємо розраховувати R і F кожного клієнта. Для зручності розрахунку R ми додаємо ще один стовпець, де буде відображена кількість днів між певною покупкою та точкою відліку, яку встановили раніше. В подальшому потрібно згрупувати дані за клієнтами, знайти мінімум нової колонки та одразу підрахувати, скільки строк згрупувалось під ID клієнта.

Оцінки підраховуються для кожного замовника та зберігатимуться у стовпцях.

Отримали дані про R, F, тепер потрібно знайти M. Для цього ми беремо рядки транзакції згідно з обраним періодом, групуємо дані за клієнтами і сумуємо стовпець коштів.

9. Розподіляємо результати R, F і M за балами від 1 до 5.

Оскільки ручне налаштування границь оцінок може потребувати декілька альтернатив і проб, щоб знайти оптимальне значення, ми будемо використовувати відсоток. Такий спосіб є більш гнучким, оскільки діапазони будуть адаптуватися до даних і працюватимуть у різних галузях промисловості, або якщо є якась зміна очікуваної поведінки клієнтів.

Тому спочатку ми розподіляємо на рівні 5 частин отримані показники R, F, M. А потім присвоюємо кожній отриманій групі бал. Для всіх трьох показників алгоритм аналогічний. У результаті ми отримуємо додаткові три стовпці.

10. Поєднуємо оцінки R, F і M у фінальний бал RFM. Тепер ми отримали повну таблицю з необхідними даними.

11. Ми будемо використовувати назви для сегментації клієнтської бази, які визначили (див. табл. 1).

12. Кінцевим результатом ми отримуємо загальний розподіл клієнтів за визначеними нами сегментами (див. рис. 1).

Спеціально виділили сегменти "чемпіони" та "лояльні клієнти" іншим кольором, тому що вони є найперспективнішими. Отже, 32% бази відносяться до клієнтів, які дають нам найбільший прибуток за останній час. Тому метою комунікаційної стратегії на ці сегменти буде такою: збільшити обсяги покупок та їх лояльність до компанії. В нашому випадку це може бути інформування на пошті про новинки товарів, спеціальні пропозиції, знижки, обмеження в кількості акційного товару або в часі діючої пропозиції. Також можна застосувати картки лояльності (навіть враховуючи, що наш магазин в онлайн), це буде мотивувати клієнтів брати більше товарів, щоб накопити відсоток знижки. Щодо таргетованої реклами в соціальних мережах, то в цьому випадку ми маємо залишатись у центрі уваги даних користувачів. Якщо клієнт купує часто певну продукцію, то алгоритм Facebook за довгий час "спіймав" це уподобання і тому нашим клієнтам може показуватись реклама конку-



Рис. 1. Розподіл клієнтської бази на основні сегменти

Джерело: розроблено авторами.

регентів. Реклама в Facebook показується за логікою аукціону, тобто чим більше рекламодавець платить, тим більше користувачів побачать рекламу. Але в нашому випадку нам не потрібно витратити великі кошти. Ми налаштуємо рекламу лише на цей сегмент, відповідно аудиторія зовсім незначна, тому дуже з малим бюджет для реклами в Facebook (до \$ 3 дол. у день) ми будемо "перемагати" конкурентів, які показують рекламу на широку аудиторію.

Новачків та тих, яких важливо не втратити — 63 та 64 клієнта відповідно. Відносно це дуже малий сегмент, але на них варто звертати увагу. Якщо це інтернет-магазин, то він має call-менеджерів. У такому випадку можна обдзвонити особисто всіх, кого важливо не втратити, отримати зворотній зв'язок від товару, компанії, запропонувати їм особливу пропозицію, яка змотивує їх возобновити покупки.

Якщо рекламний бюджет компанії середній, то над сегментами "сплячі", "під загрозою витоку" та "потребують уваги" можна працювати разом. Їм можна надіслати лист про повторне залучення, зробити його візуально яскравим, а заголовки листа — захоплюючими. В самому листі прописати основні цікаві пропозиції, можна нагадати сильні сторони компанії.

Так само потенційно лояльних та перспективних клієнтів можна згрупувати в один сегмент, бо мета комунікації одна — змусити витратити більше. Якщо це дозволяє політика магазину, можна підготувати такі пропозиції: "купи два товари і отримуй третій у подарунок" або "купуй на суму N і отримай подарунок". Такі оголошення вигідніше робити в Facebook та Instagram, тому що відвідування пошти нижче, ніж соціальних мереж, і якісний креатив оголошення зацікавить більше, ніж текст листа.

Втрачених клієнтів отримано 29% (1 271 людина). Це нормальне значення, бо в кожному бізнесі є користувачі, які купували товар і залишались незадоволеними якістю. З цим сегментом можна нічого не робити, бо є інші сегменти, на яких витратити бюджет є раціональніше.

Виявлення закономірностей (аналіз споживчого кошика за допомогою асоціативних правил алгоритму Apriori):

Цей етап роботи буде виконаний в R з використанням пакетів:

- arules (алгоритм Apriori) [4];
- arulesViz (візуалізація результатів) [5, 6, 12, 37].

Етапи отримання асоціативних правил:

1. Спочатку розглянемо історію покупок за весь період по всіх країнах. Важливо, що ми будемо працювати лише з кодами товарів.

Маємо 24982 транзакції (вчинені покупки), 4054 видів товарів. Значення щільності 0.005236173 ($0,5236\%$) означає частку ненульових осередків матриці. Середня транзакція містила $530304/24982 = 21,23$ товарів.

На наступному кроці ми бачимо куплені товари. Наприклад, товар з кодом 85123A білий підвісний свічник у формі серця. Він зустрічається в 2246 транзакціях, тобто $2246/24982 * 100\% = 8,99\%$

Розглянемо перші три товари і їх частоту в транзакціях. Отримали, що товар 10002-політичний глобус зустрічається в $0,2922\%$ транзакціях, тобто в 73 чеках.

2. Завершивши попередній аналіз, можемо почати шукати залежності між групами товарів. Для цього ми будемо користуватися алгоритмом Apriori.

Support (мінімальна частота покупки набору товару) = 0,02.

Confidence (мінімальна ймовірність такого набору) = 0,2 (20%).

minlen=2 (розглядаємо ті покупки, де більше двох товарів у чеку).

3. Проаналізуємо наше правило:

82 правила має 2 товари, і всього 3 правила — 3.

4. Тепер наше завдання відібрати цікаві правила серед тривіальних (елементарних). Прикладом такого елементарного правила буде представлено в таблиця 4.

Таке правило зустрічалося в 433 транзакціях. Якщо клієнт візьме перший товар, то в $49,48\%$ ймовірністю візьме другий набір. Якщо мова йде про офлайн-магазині, то такі товари вже стоять поруч на полицях, бо схожі, але в нас інтернет-магазин, тому навіть елементарне правило можна використати на сайті: впровадити технологію рекомендації другого товару, якщо в кошик додано перший.

Серед множини правил ми відокремили два, які представлено в таблиці 5.

Тобто перше правило з 3 товарів зустрічалося 412 разів і має ймовірність $76,72\%$. Два товари з другого правила купують в 54% випадках разом. Давайте подивимось, які групи товарів створює товар "Дерев'яна рамка біла, стиль античний". Отримали, що в 77% випадків при покупці білої тумбочки, відвідувачі беруть білу рамку. Тому є сенс розташувати ці товари (біла тумбочка, шафа, біла антична рамка, біла класична рамка) поруч, якщо це офлайн-магазин, або рекомендувати ці покупки на сайті. Для рекламної кампанії є можливість створити акцію "при покупці X тумби безкоштовний набір рамок для фото".

Отже, в результаті виконаного дослідження, ми повністю дослідили базу даних з продажів онлайн-магазину. Провели попередній аналіз, щоб оцінити масштаб даних, мінімальні, максимальні та середні значення показників (наприклад, прибуток від клієнтів). Підготували дані до подальшого проведення RFM-аналізу в сегментації клієнтів за допомогою мови програмування Python. RFM-аналіз дав можливість нам побачити розподіл клієнтів по основним сегментам, які ми самостійно визначили. Ми можемо оцінити ситуацію щодо продажів і почати планувати комунікаційну стратегію для збільшення клієнтської бази та прибутку. Додатково

було вирішено, що можна додатково провести аналіз споживчих кошиків, який можна застосувати для: покращення індивідуального звернення до клієнта, для побудови таргетованої пропозиції клієнтам (на основі їх особистої історії покупок та всіх клієнтів). Рекомендації в основному стосувались інструментів цифрового маркетингу — e-mail-ремаркетингу та таргетованої реклами в соціальних мережах. Після аналізу споживчих кошиків ми знаємо, які групи товарів беруть найчастіше на рівні магазину та клієнтів, також ми можемо знаходити сезонні тренди. Загальні рекомендації по застосуванню результатів прописані, вони будуть корисними для онлайн- та оффлайн-магазинів.

У перспективі дослідження можна робити рекомендації більш конкретними: які спеціальні пропозиції запускати, на яку аудиторію, на який бюджет; прогнозувати дохід магазинів; прогнозувати поведінку клієнтів, не враховуючи впливові фактори у вигляді нашої маркетингової кампанії. Ефективно проводити А/В-тестування запусканих оголошень в Інтернеті.

ВИСНОВКИ

Пошук і розробка ефективних алгоритмів та їх реалізація в системах управління взаємовідносинами з клієнтами дозволяє оптимізувати роботу з клієнтами в усіх напрямках і, перш за все, проводити обґрунтовані маркетингові заходи, призначені, відповідно до виявлених перевагам, безпосередньо для цільових сегментів клієнтської бази, тобто виконувати розсилку рекламної інформації не хаотично, а цілеспрямовано для сегментів клієнтів, яких зацікавить саме ця інформація і буде з великою ймовірністю мати відгук.

Результатом дослідження є розробка стратегії для отриманих сегментів клієнтів. Як ми казали, це дозволить значно економити рекламний бюджет і при цьому підвищувати її результативність, тому що пропозиції будуть цільові для окремих аудиторій. Також це дає можливість робити дуже вигідні спеціальні пропозиції певним сегментам, які компанія не може пропонувати всім клієнтам. І взагалі комунікаційна стратегія є гнучкою та може пристосовуватися до будь-якої сфери бізнесу, його бюджету та цілей. Але важливо зауважити, що розробка стратегії вимагає спеціаліста, який має досвід і повне розуміння ситуації компанії. Багато етапів аналізу повністю залежить від ОНР: кількість груп показників R, F, M, правила їх розбиття, кількість кінцевих сегментів та правила привласнення їх до клієнтів. Процес побудови комунікаційної стратегії є довгим, тому він не відносить до інструментів швидкого збільшення продажів.

Завдяки поясненому алгоритму, дослідження може бути цікавим будь-яким онлайн— оффлайн-магазинам, які хочуть покращити або тільки провадити інструменти цифрового маркетингу, тому що практична частина виконана на датасеті онлайн-магазину з роздрібними та оптовими покупцями. Але робота може бути корисною будь-яким підприємствам/компаніям, які просувають свій товар чи послугу. Більше того, описаний алгоритм Apriori може застосовуватись в інших сферах, наприклад, медицині та страхування.

Література:

1. Ahmad Heru Mujianto, Chamdan Mashuri, Anita Andriani. Consumer customs analysis using the association rule and apriori algorithm for determining sales strategies in retail central. E3S Web of Conferences — ICENIS. 2019. No. 125.
2. Andrew Aziz. Customer segmentation based on behavioural data in e-marketplace. 2017. 42 p. URL: <https://uu.diva-portal.org/smash/get/diva2:1145508/FULLTEXT01.pdf> (дата звернення: 14.01.2022).

Таблиця 4. Характеристики першого правила, відібраного нами

| | |
|------------------------|--|
| Товари | {Набір форм для кексів рожевого кольору. 60 шт.} => {Набір форм для кексів 72 шт.} |
| Підтримка | 0.01733248 |
| Довіра | 0.4948571 |
| Охоплення | 0.03502522 |
| Ліфт | 9.274209 |
| Кількість таких правил | 433 |

Джерело: розроблено авторами.

Таблиця 5. Характеристики двох правил, відібраних нами з множини

| Характеристика | Перше правило | Друге правило |
|------------------------|---|---|
| Товари | {3-рівнева підставка для торта, чайник і блюдец з принтом роз} => {зелений чайник і блюдец} | {Дерев'яна рамка біла, стиль античний} => {Дерев'яна рамка для фото, класична біла} |
| Підтримка | 0.01649187 | 0.02045473 |
| Довіра | 0.7672253 | 0.5488722 |
| Охоплення | 0.02149548 | 0.03726683 |
| Ліфт | 18.15040 | 12.275671 |
| Кількість таких правил | 412 | 511 |

Джерело: розроблено авторами.

3. Anish Nair. RFM analysis for successful customer segmentation. URL: https://www.putler.com/rfm-analysis/#What_is_Recency_Frequency_and_Monetary_analysis (дата звернення: 14.01.2022).
4. Arules: Mining Association Rules and Frequent Itemsets. URL: <https://cran.r-project.org/web/packages/arules/index.html> (дата звернення: 14.01.2022).
5. ArulesViz: Visualizing Association Rules and Frequent Itemsets. URL: <https://cran.r-project.org/web/packages/arulesViz/index.html> (дата звернення: 14.01.2022).
6. Data Visualization in R with ggplot2. URL: <https://ssc.wisc.edu/sscc/pubs/dvr/index.html> (дата звернення: 14.01.2022).
7. Derya Birant. Data mining using RFM analysis. URL: https://cdn.intechopen.com/pdfs/13162/InTech-Data_mining_using_rfm_analysis.pdf (дата звернення: 14.01.2022).
8. DoBy: Groupwise Statistics, LSmeans, Linear Contrasts, Utilities. URL: <https://cran.r-project.org/web/packages/doBy/index.html> (дата звернення: 14.01.2022).
9. Dplyr: A Grammar of Data Manipulation. URL: <https://cran.r-project.org/web/packages/dplyr/index.html> (дата звернення: 14.01.2022).
10. Fachrul Kurniawan, Binti Umayah, Jihad Hammad. Market basket analysis to identify customer behaviors by way of transaction data. Knowledge engineering and data science. 2018. Т. 1, № 1. С. 20—25.
11. Forcats: Tools for Working with Categorical Variables (Factors). URL: <https://cran.r-project.org/web/packages/forcats/index.html> (дата звернення: 14.01.2022).
12. Gert Stulp. R visualization workshop. URL: <https://stulp.gmw.rug.nl/ggplotworkshop/index.html> (дата звернення: 14.01.2022).
13. Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. URL: <https://cran.r-project.org/web/packages/ggplot2/index.html> (дата звернення: 14.01.2022).
14. Ggpubr: 'ggplot2' Based Publication Ready Plots. URL: <https://cran.r-project.org/web/packages/ggpubr/index.html> (дата звернення: 14.01.2022).
15. Habil Gadirli. Segmenting Bank Customers via RFM Model and Unsupervised Machine Learning. URL: <https://>

arxiv.org/ftp/arxiv/papers/2008/2008.08662.pdf (дата звернення: 14.01.2022).

16. Hadi Roshan, Masoumeh Afsharinezhad. The new approach in market segmentation by using RFM model. *Journal of applied research on industrial engineering*. 2017. Т. 4 (17). С. 259—267.

17. Hafsa Jabeen. Market basket analysis using R. URL: <https://www.datacamp.com/community/tutorials/market-basket-analysis-r> (дата звернення: 14.01.2022).

18. Holly Bastow-Shoop, Dale Zetocha, Gregory Passewitz. *Visual merchandising. A guide for small retailers*. Iowa State University, 1991. 68 p.

19. Jo-Ting Wei, Shih-Yen Lin, Hsin-Hung Wu. A review of the application of RFM model. *African journal of business management*. 2010. December Special Review, no. 4 (19). Pp. 4199—4206.

20. Michael Hahsler. Visualizing association rules: introduction to the r-extension package arulesviz. 2015. 24 p.

21. Mining frequent items bought together using Apriori Algorithm. URL: <https://www.analyticsvidhya.com/blog/2017/08/mining-frequent-items-using-apriori-algorithm.com> (дата звернення: 14.01.2022).

22. Mohamad Abdul Kadir, Adrian Achyar. Customer segmentation on online retail using RFM analysis: big data case of bukku.id. *ICEASD*. 2019. Vol. 4, no. 1. Pp. 1—11.

23. Mohammadreza Tavakoli, Mohammadreza Molavi, Vahid Masoumi. Customer segmentation and strategy development based on user behavior analysis, RFM model and data mining techniques: a case study. URL: https://www.researchgate.net/publication/330027350_Customer_Segmentation_and_Strategy_Development_Based_on_User_Behavior_Analysis_RF_M_Model_and_Data_Mining_Techniques_A_Case_Study (дата звернення: 14.01.2022).

24. Nagesh Singh Chauhan. Association Rule Mining. URL: <https://www.kdnuggets.com/2019/12/market-basket-analysis.html> (дата звернення: 14.01.2022).

25. Onur Dogan, Ejder Aycion, Zeki Atil BULUT. Customer segmentation by using RFM model and clustering methods: a case study in retail industry. *International journal of contemporary economics and administrative sciences*. 2018. Vol. 8, no. 1. Pp. 1—19.

26. Package 'formattable'. URL: <https://cran.r-project.org/web/packages/formattable/formattable.pdf> (дата звернення: 14.01.2022).

27. Plot method to visualize association rules and itemsets. URL: <https://rdrr.io/rforge/arulesViz/man/plot.html> (дата звернення: 14.01.2022).

28. Ponlacha Rojlerjtanya. Customer segmentation based on the RFM analysis model using k-means clustering technique: a case of it solution and service provider in thailand. *Bangkok University*, 2019. 103 p. URL: http://dspace.bu.ac.th/bitstream/123456789/4003/1/ponlacha_rojl.pdf (дата звернення: 14.01.2022).

29. Predictive segments using RFM analysis: an in-depth guide. URL: <https://www.moengage.com/blog/rfm-analysis-using-predictive-segments/> (дата звернення: 14.01.2022).

30. Pushpa Makhija. RFM analysis for customer segmentation. URL: <https://clevertap.com/blog/rfm-analysis/> (дата звернення: 14.01.2022).

31. RColorBrewer: ColorBrewer Palettes. URL: <https://cran.r-project.org/web/packages/RColorBrewer/index.html> (дата звернення: 14.01.2022).

32. Readxl: Read Excel Files URL: <https://cran.r-project.org/web/packages/readxl/index.html> (дата звернення: 14.01.2022).

33. Recency, Frequency, Monetary Model with Python — and how Sephora uses it to optimize their Google and Facebook Ads. URL: <https://towardsdatascience.com> (дата звернення: 14.01.2022).

34. Robert C. Blattberg, Byung-Do Kim, Scott A. Neslin. RFM analysis. URL: <https://link.springer.com/chapter/>

10.1007/978-0-387-72579-6_12 (дата звернення: 14.01.2022).

35. Roy Wollen. A modern approach to RFM segmentation. URL: <https://cdn2.hubspot.net/hub/184373/file-41856256-pdf/docs/modern-approach-to-rfm-segmentation-ebook.pdf> (дата звернення: 14.01.2022).

36. Scales: Scale Functions for Visualization URL: <https://cran.r-project.org/web/packages/scales/index.html> (дата звернення: 14.01.2022).

37. Selva Prabhakaran. The complete ggplot2 tutorial — part 2 | how to customize ggplot2. URL: <http://r-statistics.co/Complete-Ggplot2-Tutorial-Part2-Customizing-Theme-With-R-Code.html> (дата звернення: 14.01.2022).

38. Sohaib Zafar Ansari. Market basket analysis: Dissertation. 2019. 56 p. URL: <https://run.unl.pt/bitstream/10362/80955/1/TEG10458.pdf> (дата звернення: 14.01.2022).

39. Vasilis Aggelis, Dimitris Christodoulakis. Customer Clustering using RFM analysis. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.128.7091&rep=rep1&type=pdf> (дата звернення: 14.01.2022).

40. Vijaykumar Ummadisetty. Online retail data set. URL: <https://www.kaggle.com/vijayuv/onlineretail.com> (дата звернення: 14.01.2022).

41. Прокофьева Д. Как улучшить рассылку: советы от маркетинговых платформ и экспертов. URL: <https://www.cossa.ru/trends/276162> (дата звернення: 14.01.2022).

42. Ланц Бретт. Машинное обучение на R: экспертные техники для прогностического анализа. Питер, 2020. 463 с.

43. Разделяем клиентов по лояльности с помощью RFM-анализа. URL: <https://emailsoldiers.ru/blog/rfm-analysis> (дата звернення: 14.01.2022).

44. Таргетована реклама: що це, для чого потрібна, як ефективно використовувати. URL: <https://bestmarketing.com.ua/ua/targetovana-reklama-shcho-tse-i-dlya-choho-vona-potribna> (дата звернення: 14.01.2022).

45. Что такое RFM-анализ и как его применить в рассылке интернет-магазина. *Dacademy*. URL: <https://digital-academy.ru> (дата звернення: 14.01.2022).

References:

1. Mujiyanto, A.H. Chamdan, M.A. and Andriani, A.C. (2019), "Consumer customs analysis using the association rule and apriori algorithm for determining sales strategies in retail central", *E3S Web of Conferences*, vol. 125.

2. Aziz, A. (2017), "Customer segmentation based on behavioural data in e-marketplace", [Online], available at: <https://uu.diva-portal.org/smash/get/diva2:1145508/FULLTEXT01.pdf> (Accessed 14 January 2022).

3. Nair, M. (2015), "RFM analysis for successful customer segmentation", [Online], available at: <https://www.putler.com/rfm-analysis> (Accessed 14 January 2022).

4. R documentantion (2021), "Arules: Mining Association Rules and Frequent Itemsets", [Online], available at: <https://cran.r-project.org/web/packages/arules/index.html> (Accessed 14 January 2022).

5. R documentantion (2021), "ArulesViz: Visualizing Association Rules and Frequent Itemsets", [Online], available at: <https://cran.r-project.org/web/packages/arulesViz/index.html> (Accessed 14 January 2022).

6. R documentantion (2021), "Data Visualization in R with ggplot2", [Online], available at: <https://ssc.wisc.edu/sscc/pubs/dvr/index.html> (Accessed 14 January 2022).

7. Birant, D. (2021), "Data mining using RFM analysis", [Online], available at: https://cdn.intechopen.com/pdfs/13162/InTech-Data_mining_using_rfm_analysis.pdf (Accessed 14 January 2022).

8. R documentantion (2021), "DoBy: Groupwise Statistics, LSmeans, Linear Contrasts, Utilities", [Online], available at: URL: <https://cran.r-project.org/web/packages/doBy/index.html> (Accessed 14 January 2022).

9. R documentantation (2021), "Dplyr: A Grammar of Data Manipulation", [Online], available at: <https://cran.r-project.org/web/packages/dplyr/index.html> (Accessed 14 January 2022).
10. Kurniawan, B.U. and Hammad, J. (2018), "Market basket analysis to identify customer behaviors by way of transaction data", Knowledge engineering and data science, vol. 1.
11. R documentantation (2021), "Forcats: Tools for Working with Categorical Variables (Factors)", [Online], available at: <https://cran.r-project.org/web/packages/forcats/index.html> (Accessed 14 January 2022).
12. R documentantation (2021), "R visualization workshop", [Online], available at: <https://stulp.gmw.rug.nl/ggplotworkshop/index.html> (Accessed 14 January 2022).
13. R documentantation (2021), "Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics", [Online], available at: <https://cran.r-project.org/web/packages/ggplot2/index.html> (Accessed 14 January 2022).
14. R documentantation (2021), "Ggpubr: 'ggplot2' Based Publication Ready Plots", [Online], available at: <https://cran.r-project.org/web/packages/ggpubr/index.html> (Accessed 14 January 2022).
15. Gadirli, H. (2020), "Segmenting Bank Customers via RFM Model and Unsupervised Machine Learning", [Online], available at: <https://arxiv.org/ftp/arxiv/papers/2008/2008.08662.pdf> (Accessed 14 January 2022).
16. Roshan, H. and Afsharinezhad, M. (2017), "The new approach in market segmentation by using RFM model", Journal of applied research on industrial engineering, vol. 17.
17. Jabeen, H. (2021), "Market basket analysis using R", [Online], available at: <https://www.datacamp.com/community/tutorials/market-basket-analysis-r> (дата звернення: Accessed 14 January 2022).
18. Holly, B.S. Zetocha, D.A. and Passewitz, G.C. (1991), "Visual merchandising. A guide for small retailers", pp. 67—70.
19. Wei, J.T. Lin, S.Y. and Wu, H.H. (2010), "A review of the application of RFM model", African journal of business management, vol. 19.
20. Hahsler, M. (2015), "Visualizing association rules: introduction to the r-extension package arulesviz", pp. 24—26.
21. R documentantation (2021), "Mining frequent items bought together using Apriori Algorithm", [Online], available at: <https://www.analyticsvidhya.com/blog/2017/08/mining-frequent-items-using-apriori-algorithm.com> (Accessed 14 January 2022).
22. Kadir, M. A. and Achyar, A. (2019), "Customer segmentation on online retail using RFM analysis", vol. 4, no. 1, pp. 1—11.
23. Tavakoli, M. Molavi, M. and Masoumi, V. (2019), "Customer segmentation and strategy development based on user behavior analysis, RFM model and data mining techniques: a case study", [Online], available at: https://www.researchgate.net/publication/330027350_Customer_Segmentation_and_Strategy_Development_Based_on_User_Behavior_Analysis_RF_Model_and_Data_Mining_Techniques_A_Case_Study (Accessed 14 January 2022).
24. Chauhan, N.S. (2019), "Association Rule Mining", [Online], available at: <https://www.kdnuggets.com/2019/12/market-basket-analysis.html> (Accessed 14 January 2022).
25. Dogan, D. Aycion, A. and Atil, Z. (2018), "Customer segmentation by using RFM model and clustering methods: a case study in retail industry", International journal of contemporary economics and administrative sciences, vol. 8, no. 1, pp. 1—19.
26. R documentantation (2021), "Package 'formattable'", [Online], available at: <https://cran.r-project.org/web/packages/formattable/formattable.pdf> (Accessed 14 January 2022).
27. R documentantation (2021), "Plot method to visualize association rules and itemsets", [Online], available at: <https://rdr.io/rforge/arulesViz/man/plot.html> (Accessed 14 January 2022).
28. Rojlerjanya, P. (2019), "Customer segmentation based on the RFM analysis model using k-means clustering technique: a case of it solution and service provider in thailand", Bangkok University, pp. 103—105.
29. Moengage blog (2021), "Predictive segments using RFM analysis: an in-depth guide", [Online], available at: <https://www.moengage.com/blog/rfm-analysis-using-predictive-segments/> (Accessed 14 January 2022).
30. Makhija, P. (2020), "RFM analysis for customer segmentation", [Online], available at: <https://clever-tap.com/blog/rfm-analysis/> (Accessed 14 January 2022).
31. R documentantation (2021), "RColorBrewer: ColorBrewer Palettes", [Online], available at: <https://cran.r-project.org/web/packages/RColorBrewer/index.html> (Accessed 14 January 2022).
32. R documentantation (2021), "Readxl: Read Excel Files", [Online], available at: <https://cran.r-project.org/web/packages/readxl/index.html> (Accessed 14 January 2022).
33. Data science articles (2019), "Recency, Frequency, Monetary Model with Python — and how Sephora uses it to optimize their Google and Facebook Ads", [Online], available at: <https://towardsdatascience.com> (Accessed 14 January 2022).
34. Blattberg, B. Kim, D. and Neslin, M. (2019), "RFM analysis", [Online], available at: https://link.springer.com/chapter/10.1007/978-0-387-72579-6_12 (Accessed 14 January 2022).
35. Wollen, R. (2019), "A modern approach to RFM segmentation", [Online], available at: <https://cdn2.hubspot.net/hub/184373/file-41856256-pdf/docs/modern-approach-to-rfm-segmentation-ebook.pdf> (Accessed 14 January 2022).
36. R documentantation (2021), "Scales: Scale Functions for Visualization", [Online], available at: <https://cran.r-project.org/web/packages/scales/index.html> (Accessed 14 January 2022).
37. Prabhakaran, S. (2019), "The complete ggplot2 tutorial, how to customize ggplot2", [Online], available at: <http://r-statistics.co/Complete-Ggplot2-Tutorial-Part2-Customizing-Theme-With-R-Code.html> (Accessed 14 January 2022).
38. Ansari, S.Z. (2019), "Market basket analysis", Dissertation, pp. 55—57, [Online], available at: <https://run.unl.pt/bitstream/10362/80955/1/TEG10458.pdf> (дата звернення: Accessed 14 January 2022).
39. Aggelis, V. and Christodoulakis, D. (2019), "Customer Clustering using RFM analysis", [Online], available at: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.128.7091&rep=rep1&type=pdf> (Accessed 14 January 2022).
40. Ummadisetty, V. (2020), "Online retail data set", [Online], available at: <https://www.kaggle.com/vijayuv/onlineretail.com> (Accessed 14 January 2022).
41. Prokofeva, D. (2019), "How to Improve Your Newsletter: Tips from Marketing Platforms and Experts", [Online], available at: <https://www.cossa.ru/trends/276162> (Accessed 14 January 2022).
42. Brett, L. (2020), "Machine Learning in R: Expert Techniques for Predictive Analysis", Piter, pp. 460—465.
43. Email soldiers blog (2019), "Separate customers by loyalty using RFM analysis", [Online], available at: <https://emailsoldiers.ru/blog/rfm-analysis> (Accessed 14 January 2022).
44. Best Marketing Blog (2020), "Targeted advertising: what it is, what it is for, how to use it effectively", [Online], available at: <https://bestmarketing.com.ua/ua/tarheto-vana-reklama-shcho-tse-i-dlya-choho-vona-potribna> (Accessed 14 January 2022).
45. Digital Academy (2021), "What is RFM analysis and how to apply it in the mailing list of an online store", [Online], available at: <https://digital-academy.ru> (Accessed 14 January 2022).

Стаття надійшла до редакції 14.01.2022 р.