

**Київський національний університет імені Тараса Шевченка**

**Економічний факультет**

**Кафедра економічної кібернетики**

**КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА**

"Моделювання процесу сегментації клієнтів за допомогою RFM-аналізу для підтримки маркетингових рішень"

студентки 4 курсу

спеціальності 051 «Економіка»

ОПП «Економічна кібернетика»

денної форми навчання

Хижняк Анастасії Олександрівни

**Науковий керівник:**

Доктор економічних наук, професор

Чорноус Галина Олександрівна

Засвідчую, що в цій роботі немає запозичень із праць інших авторів без відповідних посилань

Студент \_\_\_\_\_

Роботу допущено до захисту перед ЕК

рішенням кафедри економічної кібернетики

від 12 червня 2025 р., протокол № 15

Завідувач кафедри:

доктор економічних наук, професор

Ляшенко Олена Ігорівна \_\_\_\_\_

КИЇВ – 2025

## РЕФЕРАТ

Кваліфікаційна робота бакалавра містить: 70 ст., 18 рис., 2 табл., 47 джерел, додатки.

Ключові слова: сегментація клієнтів, RFM-аналіз, електронна комерція, кластеризація, маркетингова аналітика, персоналізація, поведінкові змінні, CRM, клієнтська база, цифровий маркетинг.

Об'єкт дослідження: процес сегментації клієнтів в електронній комерції.

Мета дослідження: розробити вдосконалену модель сегментації клієнтів для підтримки маркетингових рішень на основі розширеного RFM-аналізу.

Методи дослідження: описова статистика, кластерний аналіз, нормалізація даних, метод «лікоть», Silhouette Score, машинне навчання (алгоритм K-means), Python (pandas, matplotlib, seaborn, sklearn).

Наукова новизна, теоретична значимість дослідження: полягає у розробці моделі RFMP-DOV+AIC, що поєднує класичні RFM-показники з додатковими поведінковими та демографічними змінними, що дозволяє підвищити точність сегментації та адаптацію маркетингових стратегій.

Практична цінність: запропонована модель може бути використана в CRM-системах підприємств електронної комерції для підвищення ефективності взаємодії з клієнтами, зниження відтоку та оптимізації персоналізованих комунікацій.

## RESUME

Taras Shevchenko National University of Kyiv,  
Faculty of Economics, Department of Economic Cybernetics

Key words: customer segmentation, RFM analysis, e-commerce, clustering, customer analytics, digital marketing, personalization, CRM, user behavior, data analysis.

The graduation research of student Khyzhniak Anastasiia deals with the development and implementation of an extended customer segmentation model based on RFM analysis adapted for the digital environment. The RFMP-DOV+AIC model

integrates behavioral, promotional, and demographic metrics to enhance marketing decisions.

The work is interesting for specialists in marketing analytics, CRM development, and digital strategy planning.

Pages: 70, figures: 18, tables: 2, bibliography: 47 sources, appendices: included.

## ЗМІСТ

ВСТУП .....	5
РОЗДІЛ 1. СЕГМЕНТАЦІЯ КЛІЄНТІВ ЯК СКЛАДОВА МАРКЕТИНГОВОЇ ДІЯЛЬНОСТІ В ЕЛЕКТРОННІЙ КОМЕРЦІЇ.....	8
1.1. Електронна комерція як середовище трансформації маркетингової діяльності .....	8
1.2. Сегментація клієнтів як об'єкт маркетингового аналізу: цілі, сутність та економічна доцільність.....	18
1.3. Переваги та ризики клієнтської сегментації в умовах цифрової економіки	21
ВИСНОВКИ ДО РОЗДІЛУ 1 .....	24
РОЗДІЛ 2. МЕТОДОЛОГІЧНІ ОСНОВИ ПОБУДОВИ МОДЕЛІ СЕГМЕНТАЦІЇ КЛІЄНТІВ НА ОСНОВІ RFM-АНАЛІЗУ.....	25
2.1. Сутність RFM-аналізу та його роль у клієнтській аналітиці.....	25
2.2. Різновиди підходів до реалізації RFM-аналізу: класифікація, особливості застосування. ....	28
2.3. Порівняння модифікацій RFM та розробка комплексної моделі сегментації .....	35
ВИСНОВОК ДО РОЗДІЛУ 2 .....	41
РОЗДІЛ 3. ПРАКТИЧНА РЕАЛІЗАЦІЯ ТА РЕЗУЛЬТАТИ СЕГМЕНТАЦІЇ КЛІЄНТІВ НА ОСНОВІ RFM-АНАЛІЗУ.....	42
3.1. Проведення RFM-сегментації клієнтської бази.....	42
3.2. Аналіз отриманих результатів та управлінська інтерпретація сегментів ...	51
3.3. Оцінка ефективності моделі та можливості її подальшого використання ..	59
ВИСНОВКИ ДО РОЗДІЛУ 3 .....	63
ВИСНОВКИ.....	64
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	65
Додаток А. Перелік змінних для моделі сегментації RFMP-DOV+AIC .....	71
Додаток Б. Описова статистика.....	72
Додаток В. Середні значення показників за кожним кластером моделі RFMP-DOV+AIC .....	73

## ВСТУП

**Актуальність теми дослідження.** У сучасному бізнес-середовищі ефективно залучення клієнтів є ключовою вимогою для досягнення сталого зростання компанії, особливо в умовах цифрової трансформації. Сегментація клієнтів стала одним із ключових інструментів маркетингової аналітики, який допомагає підвищити точність прийнятих рішень, персоналізувати комунікацію та зменшити витрати на залучення клієнтів. Згідно з дослідженнями [11, 13], підвищення рівня персоналізації комунікацій може призвести до зростання продажів до 20-30% при тому ж маркетинговому бюджеті.

Класичний метод сегментації - RFM-аналіз (Recency, Frequency, Monetary) все ще популярний завдяки своїй простоті, але втрачає свою ефективність, коли має справу з багатовимірною поведінкою клієнтів. У відповідь на це дослідники запропонували розширені моделі, що включають демографічні, поведінкові та каналні метрики [33 - 41]. Модель RFMP-DOV+AIC, розроблена в рамках цього дослідження, була створена для задоволення потреби в інтегрованому підході до сегментації, який поєднує сильні сторони класичної аналітики з сучасними методами великих даних.

Тому це дослідження є актуальним як з точки зору теорії маркетингу, так і з точки зору практичного застосування для компаній, які прагнуть використовувати дані про клієнтів для підвищення ефективності своїх маркетингових стратегій.

**Об'єктом дослідження** є процес сегментації клієнтів в електронній комерції.

**Предметом дослідження** є економіко-математичні моделі та інформаційні технології сегментації покупців на основі RFM-аналізу та його розширення з використанням додаткових поведінкових та демографічних змінних.

**Метою дослідження** є розробка та впровадження вдосконаленої моделі сегментації клієнтів для підтримки маркетингових рішень на основі розширеного RFM-аналізу з урахуванням цифрових, соціальних та демографічних факторів.

Для досягнення поставленої мети планується виконати наступні **завдання дослідження**:

- Дослідити вплив електронної комерції на трансформацію маркетингових інструментів і підходів до роботи з клієнтами;
- Проаналізувати природу сегментації клієнтів як елемента маркетингового аналізу;
- Визначити переваги та ризики впровадження сегментації клієнтів з урахуванням цифрових технологій;
- Розкрити теоретичну суть RFM-аналізу як інструменту клієнтської аналітики в маркетингових дослідженнях;
- Систематизувати підходи до реалізації RFM-аналізу та проаналізувати специфіку їх застосування в практиці сегментації;
- Провести порівняльний аналіз методів RFM-аналізу;
- Розробити розширену модель RFMP-DOV+AIC та обґрунтувати вибір змінних;
- Підготувати дані та описовий аналіз вибірки;
- Застосувати кластерний аналіз для формування сегментів;
- Інтерпретувати отримані результати та надати управлінських рекомендацій;
- Оцінити ефективність моделі порівняно з традиційними методами;
- Запропонувати напрямки практичного впровадження моделі в маркетингову діяльність.

У роботі використанні наступні **методи дослідження**: методи статистичного аналізу (описова статистика, нормалізація, кластеризація), алгоритми машинного навчання (K-means кластеризація), інструменти Python (pandas, matplotlib, seaborn, sklearn), методи візуалізації даних, логічні структури та порівняльний аналіз, елементи маркетингового моделювання та управлінської інтерпретації.

**Наукова новизна**: полягає у створенні та тестуванні розширеної моделі сегментації споживачів RFMP-DOV+AIC, яка поєднує класичні RFM-метрики з додатковими метриками (маркетинговий відгук, канал купівлі, тип продукту, вік, дохід, діти). Модель дозволяє точніше визначати цільові групи, адаптувати

маркетингові активності та прогнозувати поведінку клієнтів у цифровому середовищі.

**Практична значимість:** розроблена модель може бути використана для побудови системи персоналізованої комунікації у сфері електронної комерції. Її використання може підвищити ефективність маркетингових кампаній, зменшити відтік клієнтів та оптимізувати рекламні бюджети. Модель може бути інтегрована як аналітичний модуль в CRM-системи або платформи email-маркетингу. Ключові ідеї реалізовано в рамках проектного аналізу із використанням датасету [42] із відкритої платформи Kaggle.

**Інформаційною базою дослідження** стали наукові праці провідних фахівців у сфері маркетингової аналітики, статистичні дані з офіційних джерел (Eurostat тощо), міжнародні звіти з електронної комерції та практичні кейси RFM-аналізу, публікації наукових журналів, датасет “Marketing Campaign Dataset” з платформи Kaggle.

**Структура дослідження.** Кваліфікаційна робота складається зі вступу, трьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг роботи становить 70 сторінок, з них 2 таблиці, 18 рисунків та 3 додатки.

## РОЗДІЛ 1. СЕГМЕНТАЦІЯ КЛІЄНТІВ ЯК СКЛАДОВА МАРКЕТИНГОВОЇ ДІЯЛЬНОСТІ В ЕЛЕКТРОННІЙ КОМЕРЦІЇ

### 1.1. Електронна комерція як середовище трансформації маркетингової діяльності

Відповідно до Закону України «Про електронну комерцію», електронна комерція визначається наступним чином: відносини, що виникають у процесі укладення, зміни або припинення цивільно-правових зобов'язань із використанням засобів електронного зв'язку, мають дистанційний характер та спрямовані на досягнення економічного результату, внаслідок чого між сторонами встановлюються майнові права та обов'язки [1].

В економічному дискурсі електронна комерція (e-commerce) трактується як форма господарської діяльності, що здійснюється з використанням інформаційно-комунікаційних технологій (ІКТ). Вона охоплює широкий спектр економічних операцій, зокрема електронну торгівлю товарами та послугами, укладення угод у цифровому форматі, використання систем електронних платежів, а також комплекс заходів з електронного маркетингу. Електронна комерція, на відміну від традиційного бізнесу, працює у віртуальному середовищі, що допомагає оптимізувати бізнес-процеси, знизити транзакційні витрати та розширити ринкові горизонти.

За визначенням Крауса К. М., електронна комерція - це підприємницька діяльність, яка складається з надання товарів/послуг за допомогою електронних технологій, та включає в себе чотири головні сфери: корпоративні закупівлі, прямі продажі товарів або послуг, безпечне зберігання інформації, банківська справа та фактурування (платіжні системи) [2].

Існує п'ять основних типів електронної комерції, а саме: B2B (бізнес для бізнесу), B2C (бізнес для споживача), C2C (споживач для споживача), B2A (бізнес для адміністрації) та C2A (споживач для адміністрації). Вони охоплюють як комерційні транзакції між компаніями, так і взаємодію бізнесу або громадян з органами влади. Кожен тип має свої особливості, рівень поширення та сферу застосування в сучасному цифровому середовищі [2].

Електронна комерція стала набирати обертів під час COVID-19, коли всі офлайн магазини були зачинені, підприємства переорієнтувалися на онлайн-канали збуту, а споживачі, в свою чергу, відкрили для себе світ онлайн-шопінгу, та й далі продовжує стрімко розвиватись, з кожним роком кількість інтернет-користувачів в Україні постійно зростає.

За даними звіту «E-commerce Europe (2024)» у 2023 році загальний обсяг роздрібної електронної комерції в сегменті B2C європейського ринку демонструє помірне зростання на 3%, з 864 млрд євро у 2022 році до 887 млрд євро у 2023 році. Незважаючи на відносно низькі темпи зростання порівняно з попередніми роками, показник свідчить про стабілізацію ринку та поступове дозрівання електронної комерції в регіоні. Західна Європа традиційно займає найбільшу частку ринку і продовжує залишатися основним драйвером розвитку цифрової комерції на континенті [3]. Примітно, що серед європейських країн Україна має найвищі темпи зростання обсягів торгівлі у сфері електронної комерції у 2023 році - 145%. Таке значне зростання можна пояснити структурною діджиталізацією поведінки споживачів в умовах воєнного стану та посиленням внутрішньої торгівлі через онлайн-канали. Болгарія (+19%) та Польща (+17%) посідають друге та третє місця за темпами зростання відповідно, що також свідчить про поступове розширення бізнесу електронної комерції в країнах Центральної та Східної Європи з раніше низьким рівнем проникнення цифрових каналів продажів [3].

На тлі загальноєвропейських тенденцій статистика Eurostat за 2024 рік свідчить про значне зростання частки населення, яке здійснює покупки онлайн. Так, у 2024 році 79% громадян країн-членів ЄС здійснили щонайменше одну покупку онлайн, що є значним зростанням у порівнянні з 59% у 2014 році. Як результат, за десять років показник зріс на 20 відсоткових пунктів, що відображено на рисунку 1.1. Це підтверджує не лише загальну тенденцію зростання цифрової комерції, а й стрімке зростання довіри споживачів та адаптацію до онлайн форм взаємодії з бізнесом [4].

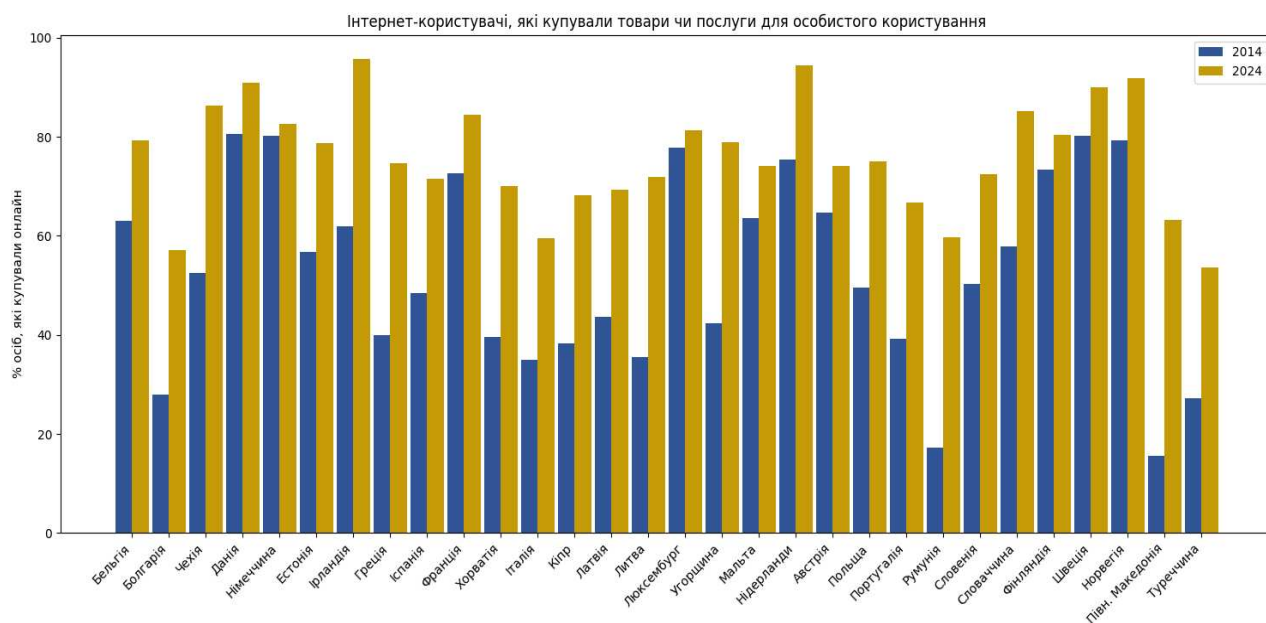


Рис. 1.1. Динаміка електронної комерції в країнах ЄС за 2014 та 2024 рік

Джерело: створено автором в середовищі Python (matplotlib) на основі [4]

Аналіз динаміки окремих країн показує, що лідерами за рівнем проникнення електронної комерції є Ірландія, Нідерланди та Норвегія. Ці країни характеризуються високою цифровою інфраструктурою та стабільною купівельною спроможністю, що сприяє широкому використанню онлайн-каналів у повсякденному житті. Водночас у Румунії спостерігається найбільше відносне зростання частки онлайн-покупців за останнє десятиліття - на 42%, що свідчить про те, що країни з історично низькими стартовими показниками переживають дуже динамічний розвиток цифрового споживчого ринку [4].

На рисунку 1.2 показано розподіл глобальних онлайн-продажів за категоріями товарів у 2024 році. Електронна промисловість, викликана високим попитом на інструменти та інші цифрові пристрої, є лідером ринку, та сягає понад 922 мільярдів доларів, популярність сегменту пов'язана з широким спектром продуктів, від основних пристроїв до спеціалізованого обладнання, такого як розумні системи та дрони [5].

Друге місце займає продукція fashion-індустрії (760 млрд доларів), яка активно просувається через онлайн-канали завдяки розвитку багатоканальних платформ. Їжа займає третє місце (460,1 мільярда доларів), показуючи, що звичка

замовляти їжу онлайн стає все сильніше, а також значні обсяги присутні в інших нішах — косметика, тютюн, розваги та медіа [5].

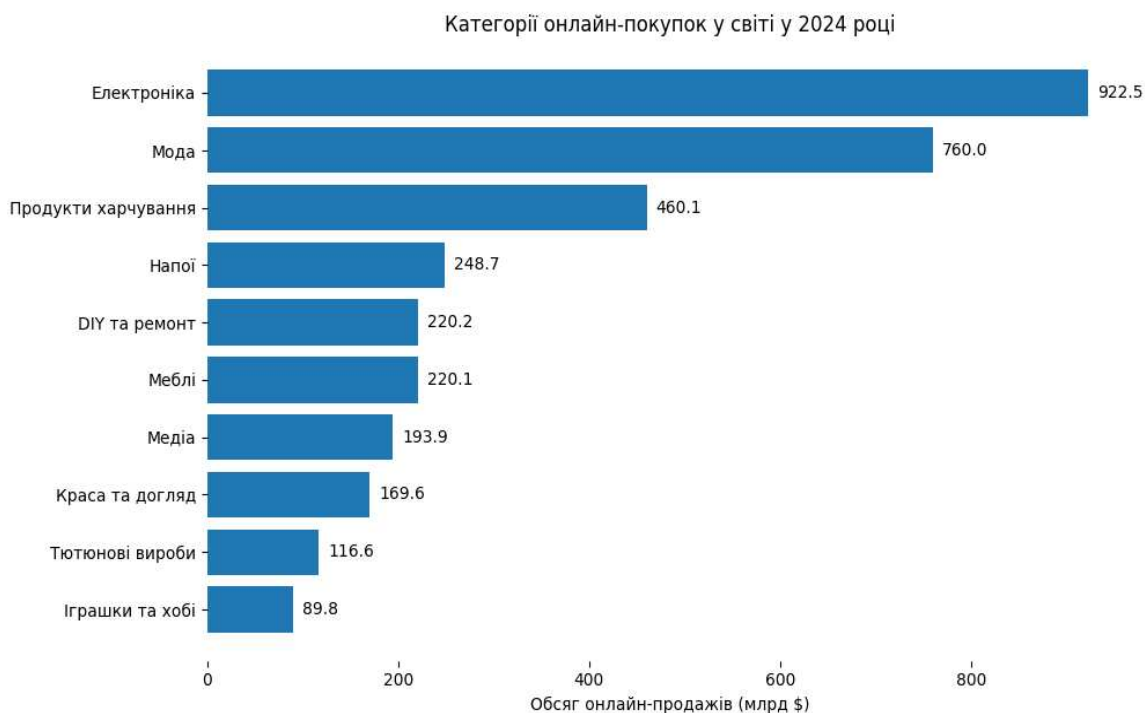


Рис. 1.2. Категорії онлайн-покупок у світі у 2024 році

Джерело: створено автором в середовищі Python (*matplotlib*) на основі [5]

На рисунку 1.3 показано топ-10 онлайн-ритейлерів за ринковою капіталізацією 2024 року. Розглянемо лідерів ринку, Amazon підтримує відкрите проміжне лідерство, зі значенням понад 1,5 трильйона доларів, такі цифри обумовлені не тільки обсягом онлайн-продажів, але і його значною часткою в області хмарних технологій (AWS) і інтенсивним розвитком логістичної інфраструктури і штучного інтелекту для управління ланцюгами поставок [5]. Walmart (428,9 млрд доларів) і Home Depot (337 млрд доларів) зберігають свої високі позиції завдяки активному поєднанню онлайн і офлайн каналів, використанню власних платформ електронної комерції і їх великій присутності в індустрії продуктів швидкого споживання. Costco, Reliance і PDD Holdings досягли стійкого зростання за рахунок ефективної цінової політики, швидкої адаптації до цифрових моделей продажів і розширення бізнесу на регіональних ринках [5].

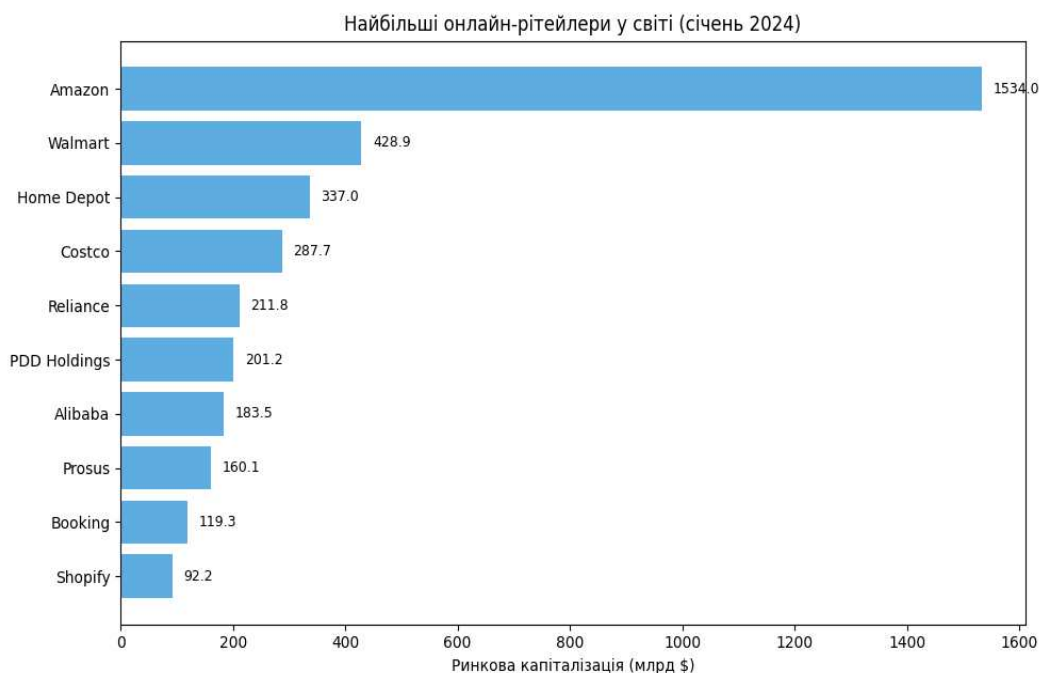


Рис. 1.3. Найбільш онлайн-рітейлери у світі на стан 2024 році

Джерело: створено автором в середовищі Python (*matplotlib*) на основі [5]

Щодо українського ринку електронної комерції, слід відзначити, що на початку повномасштабної війни темпи розвитку ринку уповільнились. У лютому-березні 2022 року спостерігалася низхідна тенденція до зниження виручки на одного покупця, тим не менш, незважаючи на складні обставини, багатьом бізнес-сегментам вдалося досягти рекордних результатів у 2024 році [7]. Ринок електронної комерції в Україні не лише не зупинився, а й продемонстрував стрімке зростання. Ринок електронної комерції збільшився майже втричі порівняно з 2016 року, і аналітики прогнозують, що в найближчі роки він подвоїться. Початок війни став викликом для ринку електронної комерції [7].

У квітні 2025 року щомісячна виручка від онлайн продажів на українському ринку електронної комерції склала 389 млн дол., що більше на 8,8% порівняно з березнем 2025. Протягом останніх 12 місяців найвища частка щомісячних доходів була зафіксована в грудні 2024 року, що становить 10,1% від загальних доходів українського ринку електронної комерції в 2024 році, а найнижча частка була зареєстрована в березні 2025 року, що становить 7% від загальних доходів 2025 року [8].

При аналізі розвитку електронної комерції в Україні, різниця в доходах 2023 та 2024 року позитивно вражає. За даними галузевого звіту [9], загальний тренд електронної комерції в Україні демонструє структурне зростання попиту на онлайн-покупки, що супроводжується розвитком логістики, цифрових платіжних сервісів та маркетингових технологій та показує, що бізнеси поступово адаптуються до умов війни, і динаміка має позитивний характер. Так, у липні 2024 року на рисунку 1.4 доходи електронної комерції зросли на 33,29% порівняно з липнем 2023 року. Це можна пояснити стабілізацією платоспроможного попиту після весняного спаду та початком активного сезону онлайн-продажів [9].

У серпні та вересні темпи зростання сповільнились, знизившись на 10,29% та 7,98% відповідно. Ця тенденція може відображати сезонне зниження споживчої активності, а також вплив інфляційних процесів та скорочення дисконтної активності в окремих секторах, водночас вона також свідчить про відносну стабільність ринку навіть у низький сезон [9].

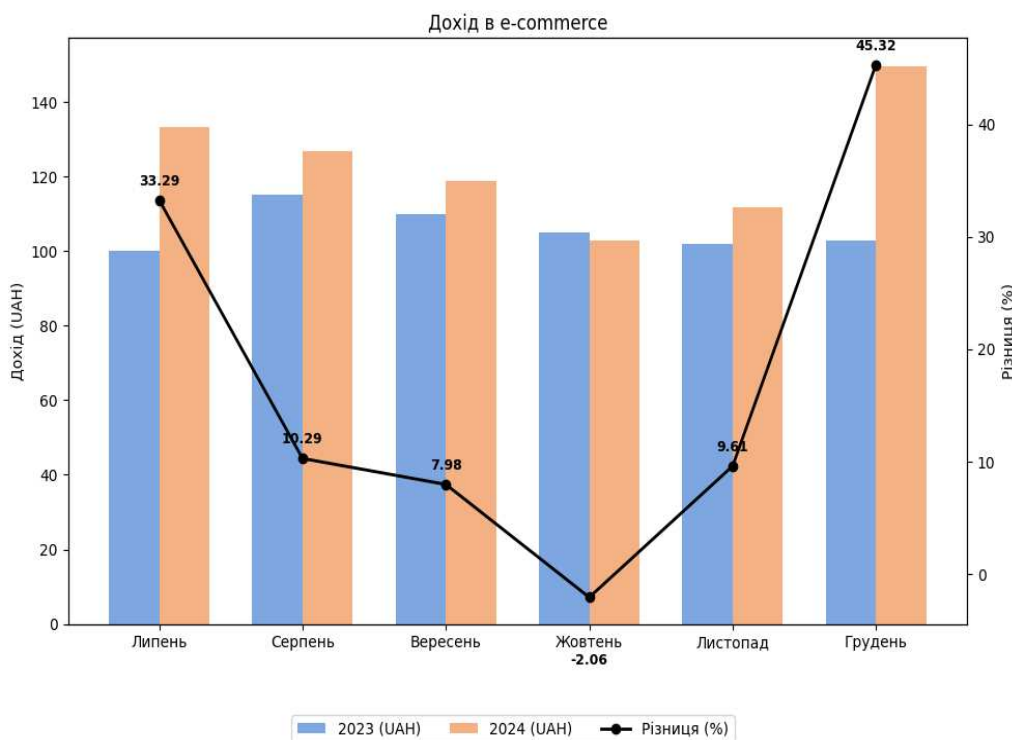


Рис. 1.4. Динаміка доходів у сфері електронної комерції в Україні у 2023-2024 рр  
Джерело: створено автором в середовищі Python (matplotlib) на основі [8]

Винятком стало незначне падіння доходів у жовтні на 2,06%. Можливим поясненням цього є тимчасове зміщення споживчої активності до інших категорій витрат та ефект відкладеного попиту напередодні осінніх розпродажів. Натомість у листопаді доходи знову зросли на 9,61%, що пов'язано з активністю торговців по всій країні під час «чорної п'ятниці» та передсвяткових знижок. Найбільше зростання на 45,32% було зафіксовано в грудні, що відповідає типовому піку покупок напередодні новорічних свят в Україні. Ці значення демонструють високу залежність доходів від сезонних коливань, а також ефективність маркетингових стратегій у пікові періоди [9].

На рисунку 1.5 представлено порівняльний аналіз доходів та інтернет-трафіку провідних ритейлерів у сфері електронної комерції в Україні станом на 2024 рік, де беззаперечним лідером українського ринку стала Rozetka, зайнявши передові позиції як за доходами (45 млрд українських гривень), так і за відвідуваннями (398,5 млн). Друге місце займає Comfy, з 9,7 млрд українських гривень і 80,9 млн трафіку. Компанія відокремлюється від конкурентів своєю високою часткою онлайн-продажів (більше 30%) і активно інвестує в трансформацію логістичних процесів по всій Україні, маючи перевагу, що роблять доставку на наступний день після оплати [10].

Єдиний міжнародний учасник першої п'ятірки - AliExpress, і хоча його доходи коливаються в межах від 7,5 до 11 млрд українських гривень, він як і раніше знаходиться в сильній позиції через високий трафік (96,7 млн). Продажі базуються на низькій вартості товарів, що призводить до високих обсягів транзакцій, але відносно низьких середніх чекових сум. З 2020 року компанія «Епіцентр» активно розвиває свій ринок більш ніж вчетверо збільшивши свій асортимент і оголосивши про плани подвоїти свою частку в українській електронній комерції [10].

У той час як Allo має 5,7 млрд українських гривень доходів і помірні показники трафіку (85,6 млн), компанія активно будує свою дистрибуційну мережу, особливо у співпраці з АЗС, і позиціонує себе як багатofункціональний ринок. У сегменті beauty-індустрії особливо відзначаються невеликі обсяги

переробки MAKEUP і EVA (3,2 млрд українських гривень і 2,4 млрд українських гривень відповідно) з високою лояльністю клієнтів (115,5 млн і 86,9 млн відвідувань відповідно). Це свідчить про те, що вона має сильні позиції у своїй сфері і потребує подальшої оптимізації монетизації трафіку [10].

В результаті структура українського ринку електронної комерції характеризується зростанням декількох багатокатегорних гравців, а також спеціалізованих майданчиків, що займають відкриті лідерські позиції. Основними конкурентними перевагами у воєнний час є адаптивна логістика, комплексна онлайн-присутність, інтеграція технологій ринку та чітка стратегія залучення клієнтів [10].

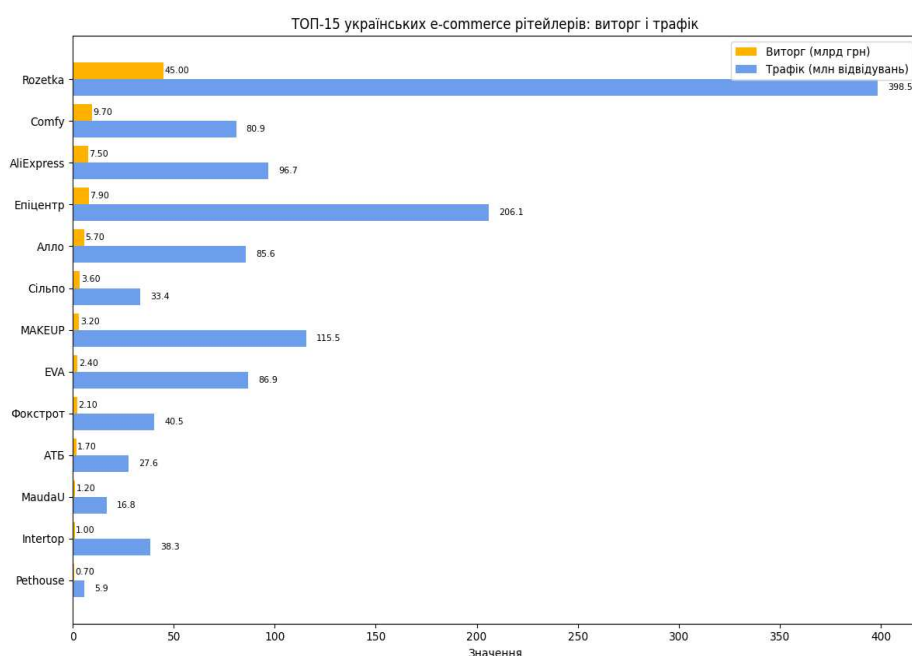


Рис. 1.5. Показники провідних ритейлерів електронної комерції в Україні, 2024 рік  
Джерело: створено автором в середовищі Python (matplotlib) на основі [10]

Проведений аналіз показує, що у другому півріччі 2024 року український ринок електронної комерції демонструватиме переважно позитивну динаміку. Уповільнення в окремому періоді (жовтень) не змінить загального тренду зростання, який забезпечується структурними змінами в поведінці споживачів та розвитком цифрової інфраструктури. Водночас нерівномірність зростання виручки підкреслює необхідність більш точного сегментування клієнтів та

коригування маркетингових стратегій з урахуванням сезонних та поведінкових факторів.

У процесі розвитку електронної комерції маркетингові стратегії зазнали значних змін, зокрема перейшли від універсального підходу до індивідуального взаємодії зі споживачами. Цифрові канали комунікації, такі як таргетована реклама, маркетинг за допомогою електронної пошти та системи рекомендацій, поступово витісняють традиційні форми реклами. Основою сучасного цифрового маркетингу є алгоритми, здатні аналізувати дані про поведінку та транзакції для створення персоналізованих послуг, наприклад, рекомендації продуктів на основі попередніх покупок.

У цьому контексті традиційна модель 4P втрачає свою домінуючу позицію, поступаючись місцем підходам, що базуються на даних і зосереджуються на конкретних умовах окремих клієнтів або сегментів ринку. Еволюція маркетингових інструментів від масових розсилок до персоналізованих стратегій у соціальних мережах зумовлена зростаючою роллю аналізу та штучного інтелекту в управлінні клієнтським досвідом.

Персоналізований маркетинг, заснований на даних про поведінку, уподобання та демографічні характеристики споживачів, допомагає створювати релевантний комунікаційний контент і зміцнювати емоційний зв'язок із брендом. Дослідження McKinsey (2021) показує, що понад 70% споживачів очікують персоналізованого обслуговування, а його відсутність негативно впливає на довіру до компанії. Отже, персоналізація має не тільки комунікативну, а й психологічну функцію і стає ключовим фактором успіху в цифровому середовищі [11].

Окрім покращення клієнтського досвіду, персоналізований маркетинг може принести значні економічні вигоди. За оцінками McKinsey, компанії, які ефективно впроваджують персоналізований підхід, можуть знизити витрати на залучення клієнтів на 10-30% завдяки підвищенню релевантності комунікації та збільшити прибутковість на 5-10% завдяки вищому середньому чеку, частоті покупок і повторних транзакцій. У середовищі електронної комерції, де

конкуренція надзвичайно жорстка, а кількість продуктів надмірна, персоналізація є одним з основних механізмів утримання клієнтів і підвищення ефективності маркетингових кампаній[11].

У сучасних умовах головна мета цифрового маркетингу полягає у залученні цільової аудиторії та забезпеченні стійкої взаємодії споживачів із брендом, що сприяє зростанню прибутку або формуванню постійної групи користувачів. Основні інструменти цифрового маркетингу включають оптимізацію пошукових систем (SEO), контент-стратегію, маркетинг за допомогою електронної пошти, маркетинг у соціальних мережах, комунікацію з впливовими особами та аналіз даних. Комбіноване застосування цих інструментів змінило традиційні методи маркетингу, зробивши підприємства більш гнучкими, мобільними та незалежними від географічних обмежень [12].

Сучасні компанії мають можливість персоналізувати користувацький досвід на всіх етапах взаємодії з клієнтами — від першого контакту через SMS або електронну пошту до поведінки в мобільних додатках, на веб-сайтах або в офлайн-точках продажу. Забезпечення релевантності інформації в кожному каналі може значно підвищити залученість клієнтів та ефективність маркетингових кампаній [12].

Сфера електронної комерції розширяє інструментарій маркетингової діяльності, оскільки є більше можливостей для збору даних та їх обробки. Сучасні інструменти, такі як Google Analytics, SimilarWeb і системи CRM, дозволяють компаніям не тільки збирати дані, але й створювати портрет споживача в реальному часі [13].

Попри очевидні переваги, які дає електронна комерція, присутні недоліки, з якими стикаються маркетологи. З основних можна виділити: інформаційне перенасичення, щодня користувачі стикаються з сотнями рекламних повідомлень, що знижує ефективність навіть ретельно спланованих кампаній та скорочує середню тривалість сесії користувача [13]. В свою чергу маркетологи мають пропрацьовувати надивленість, залучати нові інструменти, наприклад штучний інтелект тощо. По-друге, уніфікація маркетингових стратегій, багато

підприємств роблять всі маркетингові стратегії однорідними та широкомасштабними, що не дозволяє опрацьовувати правильну цільову аудиторію [14]. Відсутність глибокої сегментації, аналітики теж є головними викликами для маркетологів при просуванні товарів на онлайн ринок.

Аналіз дає розуміння, що маркетинг в електронній комерції набуває все більшого розвитку та розширює можливості бізнесів. Компанії, які успішно використовують всі етапи стратегії, сегментацію, та правильну структурування клієнтів, дозволяють не лише утримувати наявних клієнтів, а й оптимізувати витрати на залучення нових.

## 1.2. Сегментація клієнтів як об'єкт маркетингового аналізу: цілі, сутність та економічна доцільність

Сегментація споживачів є важливою складовою аналізу ринку та маркетингового аналізу, оскільки вона дозволяє компаніям краще зрозуміти своїх цільових споживачів. Фірми можуть сегментувати клієнтів на різні групи на основі спільних інтересів і таким чином адаптувати свої маркетингові зусилля для задоволення потреб і бажань кожного сегмента, і робити свої пропозиції більш персоналізовані, щоб клієнти відчували себе унікальними.

Визначати сегментацію ринку також можна за поділом потенційних клієнтів на групи на основі спільних характеристик, таких як демографічні, поведінкові, географічні або психографічні характеристики. Такий підхід дозволяє організаціям розробляти різні маркетингові стратегії (і часто різні продукти) для кожного сегмента, що призводить до підвищення рівня залученості клієнтів і конверсії [15].

Можна виділити декілька основних моделей сегментації клієнтів та підходи до самого процесу сегментування відштовхуючись від моделі. Перші три основні – демографічна, географічна, психографічна сегментація. Загалом ці три типи складають портрет клієнта, і дуже рідко працюють окремо одна від одної. Бізнеси можуть поєднувати дві чи три характеристики, наприклад жінки 30 років, що

проживають у місті Києві, щоб орієнтуватись на нішеву аудиторію в своїй сфері [16].

Технологічна, поведінкова, сегментація за потребами та цінностями хоч і використовуються рідше, проте мають свої прикладні переваги, наприклад, технологічна сегментація процес сегментування споживачів на основі технологій, які вони використовують. Цей метод сегментації може допомогти бізнесу зрозуміти, як різні групи людей використовують такі технології, як мобільні пристрої та операційні системи, і чи використовують вони нові технології, такі як штучний інтелект, що дає розуміння які системи потрібно впроваджувати, робити фокус на мобільну версію сайту чи ні, тощо [16].

Таблиця 1.1.

#### Моделі та фактори сегментування клієнтів

Модель сегментації	Фактори сегментування клієнтів
Демографічна сегментація	Вік, стать, дохід, освіта, сімейний стан
Географічна сегментація	Країна, область, місто, населений пункт
Психографічна сегментація	Особистість, ставлення, цінності, інтереси
Технологічна сегментація	Використання мобільних пристроїв, комп'ютерів, додатків і програмного забезпечення
Поведінкова сегментація	Схильності та часті дії, використання продуктів чи функцій, звички
Сегментація за потребами	Необхідні продукти або послуги та потреби конкретних груп клієнтів
Сегментація за цінностями	Економічна цінність конкретних клієнтських груп для бізнесу

*Джерело: побудовано автором на основі [16].*

Сегментація ринку є одним з важливих інструментів підвищення ефективності маркетингових заходів та прибутковості підприємства, її використання дозволяє більш точно задовольняти потреби споживачів за допомогою персоналізації продуктів, що сприяє підвищенню рівня задоволеності та лояльності клієнтів. На відміну від універсального підходу, фокус на окремих сегментах ринку забезпечує релевантність комунікації, пропозицій та цінової політики, що підвищує відгук цільової аудиторії та ефективність маркетингових заходів [17].

Додатково сегментація сприяє раціональному розподілу маркетингових ресурсів, дозволяючи зосередити бюджет на найбільш перспективних напрямках, що є важливим в умовах обмеженого фінансування, коли необхідно максимізувати віддачу від кожної інвестиції. Дослідження показують, що орієнтація на цільові групи з високим потенціалом зростання дає перевагу в конкурентному середовищі. Зокрема, в таких конкурентних галузях, як страхування, використання глибокої сегментації дозволяє значно поліпшити фінансові результати в порівнянні з компаніями, що використовують загальні або менш точні методи аналізу [17].

Таким чином, сегментація ринку має не тільки аналітичну, але й стратегічну функцію і є основою для прийняття ефективних рішень у сфері ціноутворення, позиціонування продукту та цільового маркетингу [17].

Сегментація функціонує як інструмент підтримки прийняття рішень в маркетингу та управлінні бізнесом. При прийнятті стратегічних і тактичних рішень сегментація допомагає вибрати цільові ринки, позиціонування продуктів, встановлення цін, проведення рекламних кампаній (компанії, які використовують сегментацію, керуються більш глибоким розумінням своєї клієнтської бази). Це допомагає приймати більш обґрунтовані бізнес-рішення і дає кращі результати. Після того, як вони визначили сегменти, компанія може вибрати, на яких ринках зосередитися (цільові ринки) і розробити для них власні продукти. Цей підхід називається позиціонуванням і тісно пов'язаний з позиціонуванням продуктів і брендів для задоволення потреб конкретних груп. Завдяки сегментації менеджери можуть більш чітко зрозуміти, хто є «ідеальним» клієнтом для кожного продукту або послуги, і налаштувати маркетингові повідомлення з продуктом [18].

У підсумку, ефективна сегментація дозволяє компанії визначити, які групи клієнтів варто обслуговувати і як найкраще позиціонувати свої продукти та послуги для кожної групи. Таким чином, сегментація є важливою частиною розробки маркетингових цілей і стратегій, які зазвичай включають: аналіз того, як продукти повинні продаватися або розроблятися на основі поточних сегментів

клієнтів, або ідентифікацію нових сегментів як цільових для існуючих продуктів або для розробки нових.

### 1.3. Переваги та ризики клієнтської сегментації в умовах цифрової економіки

У цифрову епоху, коли маркетингові стратегії швидко адаптуються до нових реалій та цифровий маркетинг стає популярнішим за традиційний, сегментація клієнтів стала не тільки інструментом розуміння цільової аудиторії, перетворилась на необхідний елемент ефективного функціонування компаній. У той час як використання цифрових технологій, аналітики Big Data, платформ штучного інтелекту (ШІ), CRM тощо, відкриває шлях до нових можливостей, також є недоліки та виклики, які необхідно систематично аналізувати та розуміти перед використанням.

Низка досліджень засвідчила, що інтеграція інструментів data science в діяльність електронної комерції дозволяє суттєво оптимізувати бізнес-процеси, а саме автоматично сегментувати клієнтів, прогнозувати попит, аналізувати вклад кожного маркетингового каналу та впроваджувати системи рекомендацій. Зокрема, в дослідженні [20] показано, що рішення з прогнозування попиту, аналізу ефективності маркетингових каналів, кластеризації клієнтів, рекомендаційні системи й аналіз настроїв користувачів є найефективнішими hard-data інструментами для інтернет-магазинів та було встановлено, що навіть в умовах кризи (COVID-19, війна тощо) ринок електронної комерції адаптується завдяки алгоритмічним рішенням, тобто моделям кластеризації, регресійним оцінкам та нейронним мережам, які дозволяють визначати найбільш впливові драйвери зростання, а системи рекомендацій і логістичні алгоритми забезпечують гнучкість бізнесу [20].

Якщо казати за переваги сегментації клієнтів, то вони охоплюють основні аспекти бізнес-стратегії. Розпочнемо з цільового маркетингу, завдяки сегментації з'являється можливість розробляти маркетингові повідомлення, які чітко відповідають потребам та очікуванням конкретних груп споживачів. Це дозволяє комунікації бути більш зручними, канали продажів змінюватися, а

маркетингові канали від електронної пошти до соціальних медіа ефективно використовуватися. Підвищується задоволеність клієнтів, персоналізований сервіс підвищує задоволення, створюючи відчуття індивідуальності, що є дуже важливим в маркетингу та що розглядалось більш детально раніше, унікальні ціни, бонуси та спеціальні пропозиції підвищують інтерес клієнтів та повторюють рівень покупок [21]. В компанії з'являється можливість економії грошей, оптимізовані рекламні компанії, які мають вузьку цільову аудиторію, контекстні та пошукові оголошення тощо, дозволяють знизити витрати і підвищити прибутковість кожної кампанії. Компанія інвестує свої ресурси в найбільш ефективні напрямки. Та одна з найбільших переваг це утримання клієнтів, адаптивна підтримка, персоналізований зворотний зв'язок, навчання та плани обслуговування сприяють утриманню клієнтів у довгостроковій перспективі, знижуючи витрати на пошук нових клієнтів [21].

За допомогою сегментації компанії можуть створювати відповідні маркетингові повідомлення, використовувати ефективні канали комунікації та адаптувати свої маркетингові канали до характеристик своєї цільової групи. Це дозволяє персоналізувати пропозиції, що і є найголовнішою перевагою сегментації, застосовувати унікальні ціни і бонуси, покращувати користувальницький досвід і створювати нові продукти на основі потреб клієнтів. В результаті компанії досягають підвищення лояльності клієнтів за рахунок підвищення ефективності реклами, зниження витрат на рекламу, а також поліпшення сервісу та адаптивної підтримки [21].

У сучасну цифрову епоху основним джерелом інформації для сегментації клієнтів є Big Data, яка має п'ять основних характеристик: обсяг, різноманітність, цінність, швидкість і достовірність. Ці параметри визначають не тільки кількісні аспекти обробки даних, але і якісні аспекти обробки даних, від джерела і формату до швидкості оновлення і надійності [22].

Отож, оскільки у сучасну цифрову епоху основним джерелом інформації для сегментації клієнтів є великі набори даних, де стандартні інструменти, такі як електронні таблиці, не можуть ефективно опрацьовуватись через обсяг даних,

куди входить статистика з сайтів, історія покупок клієнтів тощо. У цих рамках набуває значення роль науки про дані, яка спеціалізується на виявленні закономірностей і тенденцій у складних послідовностях інформації. Аналітики використовують статистичний аналіз, різні методи та технології для опрацювання величезного обсягу даних і в результаті такого підходу організації можуть розробляти цільові комунікаційні стратегії, які адаптують свій асортимент до потреб окремих сегментів, що призводить до більш високого рівня залучення клієнтів та лояльності [23].

Незважаючи на численні переваги сегментації клієнтів, слід враховувати й деякі недоліки. Одним з перших і найбільш критичних обмежень є відсутність або низька якість інформації про ресурси. У багатьох компаніях дані клієнтів зберігаються фрагментарно: в CRM, частково в системі аналітики, а частково в соціальних мережах або оглядах. Інтеграція цих даних в єдину систему часто вимагає значних ресурсів з точки зору технічних і людських ресурсів. Крім того, відсутність повної картини призводить до спрощених результатів, які впливають на точність створених деталей [24-25].

Наступною помилкою та недоліком є надмірна деталізація, яка призводить до великої кількості мікросегментів, що ускладнює управління кампаніями та збільшує витрати ресурсів. У той час як надто загальні сегменти знижують точність персоналізації [24-25].

Хоча більшість сегментів базуються на кількісному аналізі (кількість продажей, частота покупки, середній час споживання), недооцінка якісних джерел, таких як охоплення, соціальні медіа, поглиблені інтерв'ю або опитування, може позбавити компанії важливих ідей. Ці джерела дозволяють зрозуміти емоційні мотивації, рівні задоволеності або причини втрати клієнтів, які не можна побачити в статистичних діаграмах [24-25].

Клієнти та загалом показники для аналізу не є статичним явищем. Цінності, інтереси та поведінка змінюються залежно від зовнішнього середовища, наприклад економічні кризи, війни, технологічні інновації або зміни способу життя. Якщо компанія регулярно не оновлює свої сегменти, її стратегія може

стати неактуальною, а маркетинг неефективним. Наприклад, сегмент, який побудований в 2021 році, може не відображати поведінку споживачів під час війни або енергетичної кризи в 2025 році [24-25].

Моніторинг ключових показників для визначення ефективності сегментації важливий: коефіцієнт конверсії на сегмент, частота покупки знову, середня LTV (Life Cycle Value), рентабельність кампанії тощо. Якщо компанія не оцінює ці показники, то вона не зможе зрозуміти, ефективна обрана стратегія чи ні, і не зможе оптимізувати стратегію, що призведе до неефективності сегментування. Незважаючи на ці перешкоди, сегментація клієнтів є потужним інструментом для фірм, які прагнуть отримати конкурентну перевагу на ринку [24-25].

## ВИСНОВКИ ДО РОЗДІЛУ 1

У першому розділі досліджено, як електронна комерція змінює способи ведення маркетингової діяльності, зокрема у сфері взаємодії з клієнтами. Встановлено, що з розвитком цифрових технологій компанії все більше відходять від традиційного маркетингу до персоналізованих клієнт орієнтованих стратегій. Це вимагає більш глибокого аналізу споживачів, що посилює актуальність сегментації.

Розглянута природа сегментації клієнтів показала, що вона є не лише технічним інструментом поділу аудиторії, а й важливою основою для прийняття маркетингових рішень. Також було визначено ключові переваги такого підходу - від підвищення ефективності реклами до економії ресурсів і покращення клієнтського досвіду. Водночас аналіз виявив і ризики, зокрема пов'язані з якістю даних, надмірною деталізацією та складністю підтримки актуальності сегментів.

Загалом, сегментація в умовах цифрової економіки стає важливим інструментом для адаптації бізнесу до потреб сучасного споживача, але вимагає грамотного підходу до використання даних і технологій.

## РОЗДІЛ 2. МЕТОДОЛОГІЧНІ ОСНОВИ ПОБУДОВИ МОДЕЛІ СЕГМЕНТАЦІЇ КЛІЄНТІВ НА ОСНОВІ RFM-АНАЛІЗУ

### 2.1. Сутність RFM-аналізу та його роль у клієнтській аналітиці

У дослідженнях ринку ідея сегментації клієнтів на основі їх минулої поведінки сходить до другої половини 20-го століття, підхід RFM (Recency, Frequency, Monetary) був вперше введений в 1990-х роках аналітиками компанії «Hughes Marketing Group». Основа методу заснована на емпіричному спостереженні, що клієнти, які нещодавно зробили покупки, частіше купують і витрачають більше шансів знову торгувати. У своїй класичній презентації модель була вперше систематично представлена Г'юзом з Американського інституту баз даних у 1994 році, в його книзі *Strategic Database Marketing* (1994) [26].

Відповідно до класичної концепції, кожен клієнт отримує оцінку або рейтинг (зазвичай в діапазоні 1-5, тобто існує 5 кластерів, але це йде на вибір бізнесу та чітких правил немає) по наступним трьом критеріям: Recency (остання покупка) - кількість днів з моменту останньої покупки; Frequency (частота покупок) - загальна кількість операцій за даний період часу та Monetary (сума) - загальна або середня вартість витрат клієнта [27].

Ці три показники дозволяють класифікувати клієнтів на основі їх ефективності, лояльності та фінансового значення для компанії. У контексті базової поведінкової економіки така модель дозволяє формалізувати поняття «цінність життєвого циклу клієнта (CLV)» шляхом спрощеної оцінки колективної поведінки.

Аналіз RFM (Recency, Frequency, Amount) відіграє важливу роль у розробці стратегій і дозволяє компаніям приймати обґрунтовані маркетингові рішення на основі реальної поведінки споживачів. Робота з RFM моделями дозволяє не тільки сегментувати свою клієнтську базу за рахунок їх структури і універсальності, але і виявити ризик втрати клієнтів, оптимізувати персоналізовані пропозиції і побудувати довгострокові відносини з цільовими сегментами [28].

Однією з найважливіших переваг аналізу RFM є те, що він може ідентифікувати лояльних і прибуткових клієнтів. Особи, які нещодавно мали високі показники витрат, частоту та грошові витрати, є найбільш лояльними та цінними для бізнесу. Своєчасне їх виявлення може призвести до реалізації особистих стимулів (наприклад, програм лояльності, бонусів і особистих подяк), які підвищують лояльність клієнтів і сприяють подальшій взаємодії [28].

Сегментація клієнтів на основі метрик RFM дозволяє створювати цільові кампанії, які враховують специфічну поведінку кожного сегмента. Клієнтам, які активно взаємодіють з брендом, можуть бути запропоновані спеціальні нові продукти або преміальні пропозиції, в той час як знижки, безкоштовні пробні продукти або неповні покупки рекомендуються для менш активних споживачів. Такий підхід робить рекламну комунікацію більш ефективною через її актуальність [28].

Зниження частоти покупок і збільшення інтервалу між операціями свідчить про зниження лояльності клієнтів. Аналізуючи ці параметри, можна виявити ранні ознаки потенційної втрати клієнта і вжити профілактичних заходів: запропонувати персоналізовані знижки, надати зворотний зв'язок по послугах і поліпшити умови обслуговування. Тому RFM можна використовувати не тільки як інструмент класифікації, але і як інструмент прогнозного аналізу в області клієнтського досвіду [28].

Серед основних обмежень можна виокремити: ігнорування нефінансових факторів, а саме, що RFM аналіз не враховує важливі характеристики клієнтів, такі як соціально-демографічні характеристики (вік, стать, регіон), тип придбаного продукту, відповідь на попередні маркетингові кампанії тощо. В результаті мотивації та переваги двох клієнтів з однаковими метриками RFM можуть значно відрізнятись, що ускладнює глибоку персоналізацію [29].

Наступним недоліком є фокус на застарілих даних, RFM базується на поведінці минулих клієнтів і не враховує поточний контекст або зміни ринку, також немає можливості робити аналіз «в моменті». Відповідно, модель не завжди може реагувати на різкі зміни споживчих переваг або своєчасно

адаптуватися до зовнішніх впливів. Крім того, він не може безпосередньо передбачити поведінку майбутніх клієнтів, а лише описує вже записані дії [29].

Хоча моделі RFM історично були пов'язані з роздрібною торгівлею та електронною комерцією, їх узагальнена логіка дозволяє успішно впроваджувати подібну аналітику в інших областях, де взаємодії користувачів є регулярними та вимірюваними, наприклад: неприбуткові організації, де логіка RFM застосовується до аналізу донорської діяльності. Тут замість покупок враховуються пожертвування, наприклад дата останнього пожертвування, періодичність - кількість пожертвувань в даний період, сума - загальна сума. Таким чином, ви можете виділити активних донорів, спланувати особисті комунікації, щоб знову отримати підтримку, і створити програми подяки для найбільш щедрих донорів [28].

Наступною сферою йдуть медичні послуги, RFM-подібні методи можуть класифікувати пацієнтів за такими критеріями: час з моменту останнього відвідування, кількість відвідувань протягом даного періоду і загальна вартість лікування або консультації. Такі аналізи використовуються для поліпшення планування відвідування, персоналізації медичних нагадувань та оптимізації планів профілактики [28].

Цифрові медіа та онлайн-платформи, для цифрових сервісів (наприклад навчальних порталів) аналіз активності користувачів за допомогою логіки RFM дозволяє оцінити залученість користувачів. Відповідні показники - останній вхід або взаємодія (частота), частота входу або перегляду (частота), загальний час споживання або обсяг контенту (кількість) - дозволяють формувати рекомендаційні системи, оцінювати ризик втрати абонентів, а також підвищити ефективність персоналізованих повідомлень [28].

RFM-аналіз найбільш застосовується в сфері електронної комерції, та є ефективним інструментом для визначення ключових сегментів клієнтів, оптимізації комунікаційних стратегій та підвищення взаємодії з користувачами. Його гнучкість дозволяє адаптувати моделі до особливостей поведінки інтернет-

магазинів і налаштувати стратегії взаємодії під індивідуальні особливості вашої аудиторії.

Як і в загальному ринку, так і в сфері електронної комерції є певні недоліки застосування RFM-аналізу, по-перше, він не підходить для нових підприємств, тобто якщо клієнтська база ще недостатньо велика, результати сегментації можуть бути статистично нестабільними [30].

Наступним пунктом є обмеження на товари тривалого користування, класичне розуміння «повторюваності» не є репрезентативним, коли клієнти купують один і той же товар кожні кілька років (наприклад, побутову техніку). Та відсутність сегментації покупки: стандартні моделі RFM не розрізняють типи продуктів - всі процеси аналізуються однорідно, що може перекреслити реальну картину поведінки клієнтів. У цьому контексті деякі платформи електронної комерції, особливо сучасні CRM або BI системи, дозволяють поєднувати сегментацію RFM з іншими параметрами, такими як категорія продукту, канал взаємодії або джерело трафіку, що значно підвищує точність розробки стратегій персоналізації [30].

Однією з головних цілей кожного бізнесу електронної комерції є збільшення вартості життєвого циклу клієнта (Customer Lifetime Value, CLV). Високий рівень CLV досягається за рахунок збільшення повторних продажів, зниження втрат клієнтів і просування лояльності. Аналіз RFM дозволяє визначити сегменти клієнтів з високим потенціалом CLV і розробити стратегії заохочення регулярних взаємодій для збільшення довгострокової прибутковості [30].

2.2. Різновиди підходів до реалізації RFM-аналізу: класифікація, особливості застосування.

Для вдосконалення маркетингових стратегій та сегментації клієнтів, класичний підхід RFM-аналізу був модифікований, та на його основі створено покращені та ширше версії, які зачіпають не тільки три основні показники (recency, frequency, monetary), а й наприклад тривалість «відносин» з клієнтом, часові інтервали між транзакціями, знижки тощо. Це дозволяє спеціалістам

обирати метод виходячи з мети дослідження, та вхідних даних, допомагаючи зробити сегментацію більш чіткою, залучаючи максимально можливі характеристики в дослідженні.

Класична RFM-модель є гнучким інструментом, який дозволяє компаніям самостійно визначати критерії оцінювання та кількість сегментів. Залежно від обраної кількості рівнів, можливо сформувати різну кількість сегментів, що дає змогу точніше адаптувати маркетингові стратегії. Однак надмірна деталізація при малій клієнтській базі (менше 10 000) може знизити ефективність сегментації. Основна мета - класифікувати клієнтів за показниками R, F та M відповідно до їхньої купівельної поведінки.

Даний аналіз може бути реалізований за допомогою різноманітних інструментів, зокрема Microsoft Excel, у якому первинно розраховуються базові показники клієнтської активності - Recency (R), Frequency (F) та Monetary (M). Для подальшої сегментації клієнтів застосовується функція PERCENTILE, яка дозволяє визначити порогові значення кожного показника на основі їх розподілу.

Формула PERCENTILE, що використовується в даному підході, має вигляд:  
$$P_x = \frac{x(n+1)}{100}$$
,  $P_x$  - порядковий номер у впорядкованій вибірці, який відповідає  $x$ -му перцентилю;  $x$  - заданий рівень перцентилея (наприклад, 25, 50, 75);  $n$  - загальна кількість спостережень [31].

Також на основі результатів RFM-аналізу можна згрупувати клієнтів за різними алгоритмами кластеризації, найпоширенішим є алгоритм K-means, який групує записи за подібністю властивостей (в даному випадку R, F і M значень) і є обчислювально ефективним навіть у великих наборах даних. Також використовується ієрархічна кластеризація, яка дозволяє створювати дендрограму вкладених кластерів без необхідності попередньо задавати їх число.

В дослідженні [32] при порівнянні алгоритму K-means з ієрархічною кластеризацією пов'язану з сегментами клієнтів електронної комерції на основі поведінкових даних, отримані результати по алгоритму K-means забезпечили трохи кращу відокремленість кластерів, середній коефіцієнт силуету для K-

means склав приблизно 0,29, тоді як для ієрархічного методу 0,25, що свідчить про більш чітко визначені та виразні сегменти.

В дослідженні [33] було порівняно ефективність алгоритмів K-means і K-medoids порівнюється за даними RFM-моделювання для онлайн продажів. Було визначено, що K-means перевершують K-medoids з точки зору ефективності та показників продуктивності, K-means мають значно нижчий середній індекс Девіса-Буллінна (DBI) (0,2962 проти 0,8942 у K-medoids), а час виконання помітно швидшим (0,096 с проти 2,4295 с). Підсумовуючи можна сказати, що кластеризація є більш компактною та дискретною, коли використовується K-means значення, тоді як K-medoids є більш стійкими до викидів, але мають слабшу чіткість. Саме тому вибір алгоритму впливає на якість і інтерпретацію сегментації, K-means добре підходить для швидкої сегментації великої вибірки, тоді як K-medoids може бути розглянутим, коли дані містять викиди, які можуть змістити центроїди кластерів в K-means.

Для підвищення точності сегментації також можуть об'єднувати кілька методів, зокрема в роботі [34] було запропоновано інтегрувати метод головних компонентів (PCA) додатком до K-means та ієрархічної кластеризації. У цьому дослідженні всі змінні були згруповані спочатку для даних транзакцій кредитної компанії, а далі використано PCA для усунення мультиколінеарності, після чого кластеризація була повторена з оновленою кількістю кластерів. Результати показують, що використання PCA, як додатковий етап при дослідженні, покращує якість сегментації та може використовуватися як інструмент перевірки K-means та ієрархічної кластеризації. Цей багатоступеневий метод забезпечує більш стабільні та інтерпретовані розділи, ніж кожен метод окремо.

Після завершення сегментації клієнтів розробляється маркетингова стратегія для кожної групи. Метою компанії є збільшення продажів і середніх перевірок «кращих» клієнтів, просування покупок «середніх» клієнтів і повернення «гірших» клієнтів. На етапі розробки маркетингової стратегії важливо визначитися, які дії робити для кожного сегмента ринку.

Далі першим вдосконаленням класичної моделі було додавання показнику тривалості відносин з клієнтом, L-length, що і стало початком моделі LRFM. Показник L зазвичай визначається як час від початкової покупки клієнта до поточної або остаточної покупки. Насправді L характеризує «досвід» клієнта: як довго він був клієнтом компанії [35].

Також іноді використовують подібну модель, а саме RFM-T, в різних дослідженнях показник тривалості має різну інтерпретацію. Наприклад, існує модель RFM-T (Recency, Frequency, Monetary, Tenure), розраховують як час від початкової покупки до сьогоднішнього дня, також в деяких джерелах цей параметр може називається T (inter-purchase time), що є показує «міжпокупковий час». Додавання L/T дозволяє оцінити лояльність клієнтів та загалом їх поведінку, яка не є чітко видимою лише за критеріями Recency [36].

Іншим важливим напрямком розширення RFM-аналізу є включення метрики, яка характеризує класифікацію покупки або тип клієнта. У класичному RFM, всі кошти (M) об'єднуються в одній сумі, незалежно від того, покупець купує той же продукт або різні продукти. Але що є цінним з точки зору маркетингу, так це те, наскільки різноманітні покупки клієнта, особливо це використовується при таргетованій рекламі, коли кампанія запускається на конкретний продукт, тому було застосовано метрику D (Diversity-різноманітність) для вирішення цього питання [37].

Також показник D дозволяє сегментувати своїх клієнтів на основі широти їх інтересів, клієнти з високим показником D є покупцями з широким спектром покупок різного напрямку, в свою чергу клієнти з низьким значенням D відносяться до спеціалізованих клієнтів, які зацікавлені у вузькому асортименті продукції. Обидва типи можуть мати високі (M), але їх стратегії подолання різні: першому можна запропонувати більш широкий спектр нових каталогів продуктів, тоді як останньому можна запропонувати більш чітку рекламу, фокусуючись на найчастіших категоріях [37].

У роздрібній торгівлі важливо враховувати не тільки суму грошей, які витрачає клієнт, але і кількість продуктів, які він купує. Тобто два клієнти

можуть витратити однакову суму грошей щомісяця, але один клієнт може купити пару дорогих продуктів, а інший може купити багато дешевших продуктів. Ці ситуації відрізняються, і відповідно маркетингова стратегія для обох варіантів має бути, в модель RFM маємо тільки загальну суму покупок, тобто M, тому була запропонована модель LRFMV - Length, Recency, Frequency, Monetary, Volume (Тривалість, Недавність, Частота, Сума, Обсяг). LRFM розширює модель LRFM, додаючи V (обсяг купівель), п'ятий параметр, який безпосередньо відображає кількість придбаних одиниць і дозволяє вивчити взаємозв'язок між обсягом покупки та отриманим прибутком [38].

Тобто його можна інтерпретувати як середню суму продуктів, які купує клієнт. Додавання кількісного параметра дозволяє відслідкувати взаємозв'язок між кількістю купленої продукції і рентабельністю клієнтам та визначити чіткий зв'язок між розміром V і кількістю придбаних продуктів і отриманим прибутком - за раніше невизнаними моделями RFM і навіть LRFM [39].

Наприклад, класифікуючи покупців за LRFMV, магазини можуть встановлювати групу «оптовиків», тобто клієнтів, які купують великі кількості (мають високий V, навіть якщо M такий самий, що і у роздрібних покупців). Це говорить про іншу стратегію: оптовим покупцям можуть бути запропоновані конкретні умови, знижки на великий обсяг та якісне обслуговування для клієнтів з меншим V, але вище M тощо [39].

Класична модель RFM фокусується на кількісних показниках ігноруючи якісні, а саме задоволеність клієнта, для вирішення цієї проблеми була запропонована модель LRFMS, з додатковим показником S (satisfaction – показник задоволеності клієнтів). Даний показник можна отримати через проведення опитувань та отримання відгуків, які клієнти будуть надавати після покупки [40]. Модель LRFMS також передбачає зміну в розрахунку Recency, а саме замість використання показника R, як в класичній моделі, використовується скореговане значення, враховує загальний контекст покупки та зменшує випадковість, бо в класичному RFM-аналізі клієнт, який щойно зробив свою першу покупку, матиме найкращий показник R в моменті, навіть

якщо його цінність ще не доведена, в свою чергу LRFMS-аналіз усуває цю аномалію, беручи до уваги задоволення відносин і довжину [40].

Для збільшення продажів дуже важливою умовою є періодичні знижки, або розпродажі. Вони допомагають залучити нових клієнтів, залучити постійних клієнтів шляхом укріплення взаємовигоди, створення ексклюзивності, покупці відчують, що вони отримали перемогу шляхом купівлі товару дешевше, ніж інші, та допомагає позбавитись від зайвих запасів, що дає можливість оптимізувати запаси. Тому компанії часто використовують знижки та акції для підвищення прибутку. Для цього дуже важливо робити чітку сегментацію, оскільки деякі клієнти будуть купувати тільки тоді, коли є акція, а іншим немає різниці, і вони можуть зробити покупку, як з, так і без знижки. Для того щоб відслідкувати ці тенденції, класична модель була розширена та створена оновлена RFM + DP, де додано метрику DP до стандартних R, F та M, вимірюється у відсотках [41].

Наприклад, якщо клієнт А витрачає 500 000 грн. за певний період часу, а також купує (заощаджує) 25 000 грн. зі знижкою, то  $DP = 5\%$ . Клієнт В в свою чергу може витратити 500 000 грн., зі знижкою - 0 грн., тому  $DP = 0\%$ . В результаті клієнт А буде більш чутливим до знижок, ніж клієнт В та при запуску рекламних компаній треба це враховувати та обирати правильну цільову аудиторію [41].

Окрім модифікованих видів RFM-аналізу за декілька останніх років роль штучного інтелекту сильно зросла в сегментації клієнтів. Сучасні методи штучного інтелекту здатні аналізувати велику кількість різнорідних даних і виявляти глибокі закономірності в поведінці споживачів. В роботі [42] була запропонована інтегрована гібридна модель сегментації, яка поєднує K-means з алгоритмами підкріплення та еволюційної оптимізації. Спочатку здійснюється аналіз кореляції між характеристиками клієнтів та відсіюються надлишки, потім дані очищуються від шуму та скорочується розмірність за допомогою PCA. Наступним кроком модель застосовує диференційний еволюційний алгоритм, керований Q-learning (методом навчання з підкріплення), щоб динамічно

налаштовувати параметр K-means з метою підвищення якості кластеризації. Цей адаптивний підхід значно покращує групування клієнтів, в експериментах над маркетинговим датасетом модель перевершує традиційні підходи, забезпечуючи високий ступінь однорідності в кластерах і чіткі межі між кластерами [42].

Поєднання RFM-аналізу з методами штучного інтелекту (наприклад, навчання з підкріпленням для оптимізації кластеризації) може автоматично знайти найкращі правила сегментації властивостей даних, що неможливо зробити за допомогою статичних алгоритмів з фіксованими налаштуваннями (наприклад, K-означає).

Також штучний інтелект дає змогу оброблювати неструктуровані дані, наприклад відгуки, запити в пошуку, активність в соцмережах тощо, в свою чергу спеціалісти почали це інтегрувати в сегментаційний аналіз, щоб отримати більш повну картину клієнта. В роботі [43] відображено здатність моделі на основі глибокої нейронної мережі та трансформатора (підхід GPT) аналізувати відгуки споживачів. Запропонована мета-інтеграція (трансформатор T5 для вилучення контекстних ознак + капсульна мережа для виявлення ієрархічних настроїв) доповнюється генеративним підходом до збагачення даних для вирівнювання дисбалансу категорій у відгуках. Результатом став рекордний рівень точності емоційної класифікації багатомовних текстових коментарів, близько 97,5%, при цьому F1-міра склав 95,5%.

Даний аналіз показує, що сучасні глибинні моделі здатні враховувати нюанси мови споживчих настроїв і надавати організаціям детальні інсайти про фідбек клієнтів. У контексті RFM-сегментації перспективним напрямком є поєднання таких підходів ШІ з традиційним RFM-аналізом. Наприклад, сегментуючи клієнтів на основі RFM-метрик, компанії можуть додатково застосовувати аналіз настроїв: це дозволить зрозуміти, наскільки задоволений кожен сегмент, і відповідно скоригувати маркетингові повідомлення. Загалом, поєднання ШІ (наприклад, великомасштабних мовних моделей, таких як GPT або глибокі нейронні мережі) з RFM дозволяє враховувати як кількісну цінність

ваших клієнтів, так і якісні виміри поведінки та ставлення до продуктів, тим самим підвищуючи точність та ефективність сегментації.

### 2.3. Порівняння модифікацій RFM та розробка комплексної моделі сегментації

Класична модель RFM з її численними модифікаціями є важливим інструментом у сфері маркетингового аналізу, особливо в задачах, пов'язаних із сегментацією споживачів. Її перевага полягає в тому, що навіть маючи мінімальний набір даних (наприклад частота, сума та давність покупок), можна отримати початкове уявлення про активність споживачів і визначити пріоритетні групи для подальших маркетингових зусиль.

Однак зі зростанням електронної комерції, цифрових каналів комунікації та посиленням конкуренції за увагу споживачів зростає потреба в більш гнучких моделях, які можуть точніше описувати клієнтів. Тому було розроблено низку модифікованих версій класичних моделей RFM, доповнених новими змінними. Ці змінні можуть охоплювати всі аспекти взаємодії з клієнтами: тривалість співпраці (Length), регулярність покупок (Time), широту асортименту (Diversity), обсяг транзакцій (Volume), ступінь задоволеності (Satisfaction), рівень цінової чутливості (Discount Participation), а також соціальні характеристики (Connection).

Порівняльну характеристику моделей представлено в таблиці 2.1. У той же час, моделі, які фокусуються на обсязі або різноманітності (V, D), більше підходять для роздрібної торгівлі та дистрибуції. Варто підкреслити, що ефективність кожної конкретної моделі значною мірою залежить від якості вхідних даних, структури клієнтської бази та поставлених бізнес-цілей. Тому на практиці рекомендується не просто обирати модель за замовчуванням, а комбінувати різні підходи для формування адаптивної системи залежно від конкретної ситуації.

Таблиця 2.1

Порівняльна характеристика різновидів RFM-аналізу та їх практичного застосування

Назва моделі	Додаткові параметри	Основні переваги	Потенційні обмеження	Рекомендовані сфери застосування
RFM (базова)	–	Простота реалізації, достатньо транзакційних даних	Ігнорує стаж, різноманітність, обсяг покупок	Базовий аналіз у більшості секторів
LRFM	L – досвід (Length)	Дає змогу врахувати тривалість взаємодії з клієнтом	Не застосовна при короткостроковій взаємодії	Підписні сервіси, банки, електронна комерція
RFMT	T – інтервал між покупками	Враховує регулярність купівельної активності	Потребує детальної історії транзакцій	Сервіси з регулярними продажами
RFM-D	D – різноманітність	Оцінка широти інтересів покупця	Вимагає класифікації асортименту	Гіпермаркети, ритейл з широким асортиментом
LRFMV	V – обсяг купівлі	Дозволяє виявити оптових клієнтів та порівнювати обсяги	Потребує кількісного виміру товарів у транзакціях	B2B-сегмент, гуртові постачання
LRFMS	S – задоволеність	Враховує рівень сервісу як фактор лояльності	Потрібні дані відгуків або опитувань	Послуги, де важливий клієнтський досвід
RFM+DP	DP – частка знижок	Дозволяє оцінити цінову чутливість клієнта	Потрібні дані про всі типи знижок	Електронна комерція, програми лояльності
RFM-C	C – соціальний вплив	Ідентифікація клієнтів із реферальною активністю	Дані про соціальні зв'язки не завжди наявні	Онлайн-спільноти, брендовий маркетинг

Джерело: розроблено автором

Незважаючи на вищезазначені модифікації, жодна з них не охоплює всі ключові аспекти поведінки покупців, а саме: транзакційні, цифрові, рекламні та демографічні одночасно. Більшість моделей залишаються вузько спеціалізованими, що обмежує їхню адаптивність до сучасних ринкових умов, де покупець взаємодіє з брендами через численні канали, індивідуально реагує на рекламні акції та має унікальні соціально-економічні характеристики.

Виходячи з актуальної ситуації на ринку, було розроблено дану модель, яка є розширенням класичного трикомпонентного RFM-аналізу (давність, частота та сума) шляхом додавання наступних додаткових метрик: PromoScore (P) - показник реакції покупців на маркетингові кампанії, що розраховується як сума позитивних відповідей на п'ять кампаній; NumDealsPurchases (D) - кількість покупок зі знижкою, що дозволяє розпізнати схильність клієнтів реагувати на промоакції; OnlineRatio (O) - частка онлайн-покупок у загальній кількості транзакцій, що вказує на поведінку каналу. Diversity (V) - показник різноманітності, що відображає кількість категорій товарів, які купують покупці; вік, дохід, діти (A, I, C)- демографічні змінні надають додаткову контекстну інформацію для сегментування покупців на основі життєвого циклу та платоспроможності, більш детальний опис та розрахунок даних змін показано в Додаток А.

В результаті розширена модель виглядає RFMP-DOV+AIC, де кожна буква відповідає окремій характеристиці клієнта. Такий підхід дозволяє проводити більш глибоку багатовимірну сегментацію, враховуючи поведінкові, рекламні, цифрові та соціальні компоненти купівельного профілю клієнта та дає можливість створювати чіткі сегменти для рекламних кампаній.

Реалізація моделі передбачає проходження узгодженого аналітичного циклу, що охоплює етапи підготовки, трансформації, кластеризації та інтерпретації даних. Початковий етап передбачає очищення вибірки від технічних змінних, уніфікацію форматів даних та агрегування дотичних показників з метою зниження розмірності та покращення інформативності ознак. Важливу роль відіграє формування нових змінних на основі кількох первинних

атрибутів, що дозволяє спростити структуру та уникнути надлишкової деталізації.

Подальший описовий аналіз дозволяє виявити структурні властивості вибірки, включно з асиметрією, дисперсією, викидами та кореляційними залежностями між змінними. Ці спостереження слугують підґрунтям для прийняття рішень щодо подальших трансформацій ознак, а також дають змогу оцінити потенційну дискримінантну здатність кожного з параметрів моделі.

З огляду на різні шкали вимірювання показників моделі та нерівномірний розподіл, наступним кроком є нормалізація. Застосовується перетворення квантилів до нормального розподілу, що дозволяє збалансувати внесок кожної характеристики в кластеризацію та запобігти спотворенню результатів через домінування змінних з великою дисперсією.

Після формування нормалізованого простору ознак кластеризація проводиться за допомогою методу K-Means. Цей алгоритм реалізується для різних значень параметра  $k$ , щоб знайти оптимальну кількість сегментів. Для оцінки якості кластеризації використовуються чотири показники: SSE, індекс Silhouette, індекс Calinski-Harabasz та індекс Davies-Bouldin. Узгоджені результати цих показників дозволяють обґрунтовано вибрати кількість кластерів, що гарантує внутрішню однорідність та відстань між кластерами.

Після завершення кластеризації кожному клієнту присвоюється мітка кластера, що формує сегментовану структуру вибірки. Останнім етапом є інтерпретація отриманого сегментування, тобто ідентифікація характеристик кожної групи, оцінка їх поведінкових особливостей та потенційної маркетингової цінності. Візуалізація результатів кластеризації за допомогою діаграм розподілу, бокс-діаграм та теплових карт допомагає глибше зрозуміти структуру клієнтської бази.

Модель RFMP-DOV+AIC має високий потенціал для застосування в різних сферах бізнесу, особливо коли взаємодія зі споживачами є багатовимірною і охоплює різні канали комунікації. Її практична цінність зростає в умовах жорсткої конкуренції, необхідності персоналізації маркетингу та широкого

розмаїття товарів і послуг. Модель найкраще підходить для застосування в електронній комерції, роздрібної торгівлі та банківських послуг, а також програм лояльності для великих онлайн-компаній.

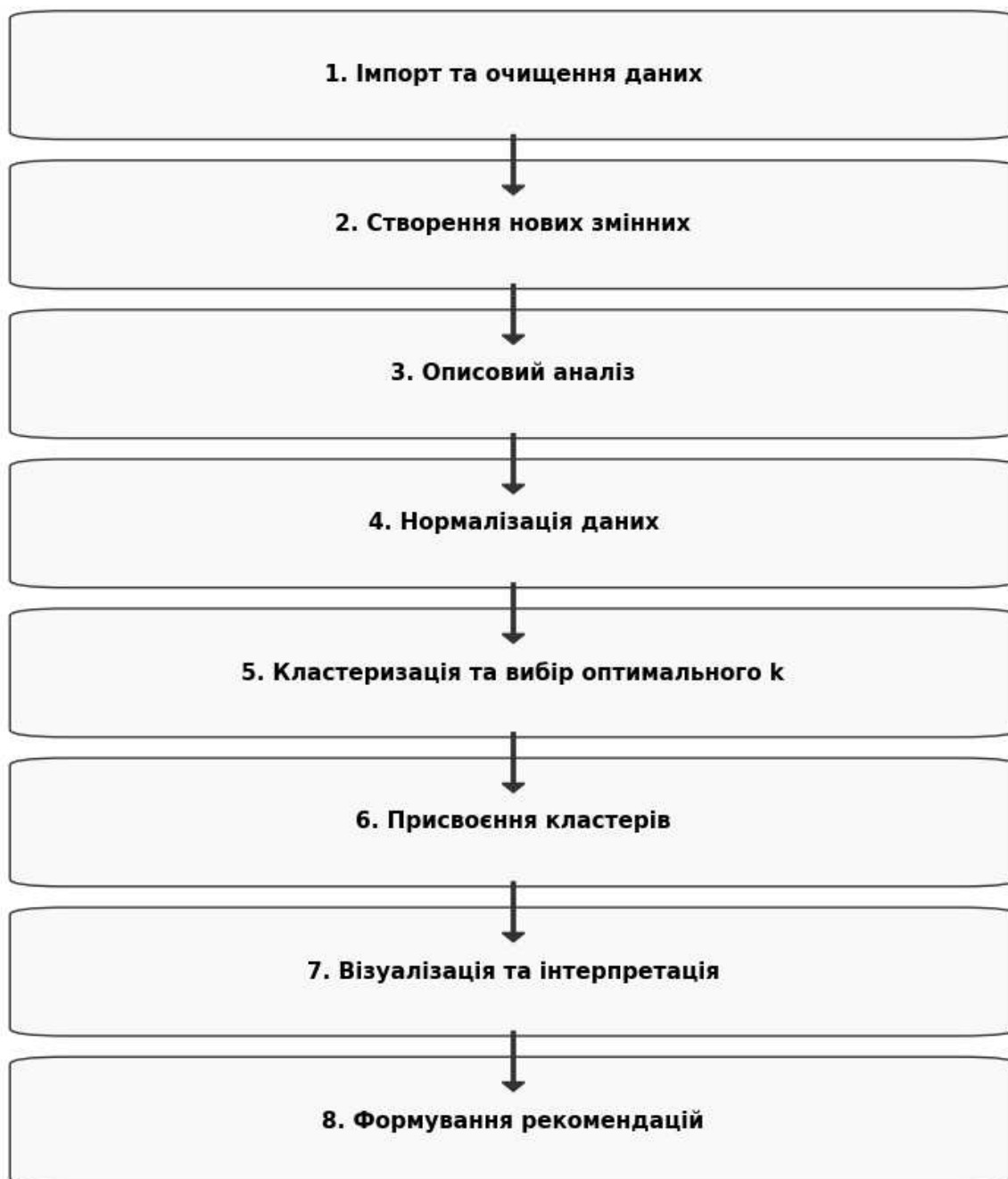


Рис. 2.1. Етапи побудови моделі сегментації клієнтів RFMP-DOV+AIC

*Джерело: розроблено автором*

Найбільш корисною стане у сфері електронної комерції, де широко використовуються різноманітні цифрові канали, часто проводяться акції, а асортимент товарів великий. Що дозволяє компаніям більш точно адаптувати контент платформи під їх цільову аудиторію, яка поділена на більш чіткі та малі

сегменти, автоматизовані електронні листи, алгоритми рекомендацій та програми лояльності.

Також можна взяти до прикладу банківську сферу, в якій ця модель особливо ефективна для роздрібних послуг, оскільки вона поєднує транзакційні та демографічні характеристики, які є ключовими для побудови профілів клієнтів. Завдяки цьому банки можуть ідентифікувати клієнтів, які активно користуються цифровими каналами обслуговування або реагують на різні бонуси або винагороди, і створювати для них відповідні продуктові пропозиції.

У програмах лояльності модель RFMP-DOV+AIC гнучко і багатовимірно групує учасників на основі критеріїв, які виходять за рамки типової купівельної активності. Це дозволяє адаптувати стратегії стимулювання до реальних потреб і мотивацій окремих груп споживачів, підвищуючи ефективність заходів з утримання та заохочуючи до повторних покупок.

Загалом, для ефективного впровадження моделі необхідно забезпечити три важливі умови: наявність інтегрованої CRM-системи, що містить транзакційні, демографічні та маркетингові дані про ваших клієнтів, наявність достатньої кількості спостережень для правильної сегментації, чітке формулювання бізнес-цілей сегментації та визначення фокусу для відбору змінних і параметризації моделі.

Особливу увагу варто приділити питанню адаптивності, залежно від специфіки галузі або доступності окремих змінних, модель можна модифікувати - наприклад, замінити OnlineRatio іншими релевантними метриками (наприклад, канали цифрової взаємодії), або замінити PromoScore індексом загальної залученості клієнтів. Це дозволяє моделі бути гнучкою в застосуванні та піддається поступовому масштабуванню, починаючи з базової версії, розширювати її з часом, у міру зростання аналітичних можливостей та технічної інфраструктури, компанії можуть згодом інтегрувати додаткові змінні.

Якщо казати про впровадження в бізнес дану модель, або загалом подібні підходи, то першочергово сегментацію слід інтерпретувати за категоріями бізнесу, оцінювати за цінністю та ризиком і використовувати як основу для

побудови диференційованої стратегії взаємодії. Моніторинг ефективності таких стратегій у майбутньому дозволить не лише підвищити релевантність комунікації, але й оптимізувати маркетингові витрати.

## ВИСНОВОК ДО РОЗДІЛУ 2

В розділі розглянуто принцип використання RFM-аналізу як інструменту сегментації клієнтів у маркетинговому аналізі. Проаналізовано класичну трикомпонентну модель, засновану на поведінці клієнтів, а також різні модифікації базової моделі, розширені за рахунок додавання інших змінних (таких як досвід клієнтів, частота покупок, класифікація активності, обсяг покупок, чутливість до ціни та інші поведінкові або демографічні характеристики).

Запропонована розширена модель RFMP-DOV+AIC розглядається як логічне продовження концепції RFM. Вона включає параметри, що відображають реакцію на рекламні акції, знижки, цифрову взаємодію, широту асортименту продукції та соціально-демографічні фактори. Така структура дозволяє здійснювати більш точну та інформативну сегментацію, що забезпечує підвищення релевантності маркетингових рішень.

## РОЗДІЛ 3. ПРАКТИЧНА РЕАЛІЗАЦІЯ ТА РЕЗУЛЬТАТИ СЕГМЕНТАЦІЇ КЛІЄНТІВ НА ОСНОВІ RFM-АНАЛІЗУ

### 3.1. Проведення RFM-сегментації клієнтської бази.

У контексті розвитку електронної комерції та персоналізованих маркетингових стратегій існує потреба у створенні більш гнучких та комплексних моделей сегментації, які враховують як класичну купівельну поведінку клієнтів, так і сучасні цифрові характеристики.

Враховуючи обмеження традиційної моделі RFM, для кращої сегментації було розроблено вдосконалену методологію сегментації клієнтів, яка не лише оцінює споживчу цінність, але й аналізує канали взаємодії, реакцію на маркетингові стимули, різноманітність покупок та основні демографічні характеристики та дає можливість зробити сегментацію точнішу, модель RFMP-DOV+AIC.

Для застосування сегментації було обрано набір даних [44], який містить інформацію про клієнтів, їх соціально-демографічні характеристики, фінансовий стан, а також поведінку щодо маркетингових кампаній. Набір даних охоплює період з 30 липня 2012 року по 29 червня 2014 року і включає 2240 запис, кожен з яких представляє унікального клієнта.

Дані включають як звичайних клієнтів (здійснювали покупки офлайн), так і електронних (з високою часткою онлайн-каналів). Набір охоплює лише тих клієнтів, які здійснили щонайменше одну покупку у вказаний період. Жодного випадкового методу вибірки не застосовувалось — аналіз проводився на повній наявній вибірці.

На етапі підготовки даних для побудови моделі, дані були підготовлені та трансформовані для забезпечення коректності подальших розрахунків та підвищення якості моделювання, були вилучені технічні змінні, які не несуть аналітичного навантаження в контексті сегментації, всі дати були приведені до єдиного формату часу для подальших часових розрахунків, проаналізовані та почищені викиди та було створена нова змінна AcceptedCmp, яка підсумовує відповіді споживачів на п'ять маркетингових кампаній (AcceptedCmp1-

AcceptedStr5). Замість того, щоб зберігати п'ять окремих бінарних змінних, формується єдиний категоріальний атрибут, що відображає кількість прийнятих кампаній, або значення нуль, якщо жодна пропозиція не була прийнята. Таке перетворення спростило подальший аналіз і зменшило розмірність простору ознак.

Здійснивши ці кроки, ми отримали вже очищений датасет, який складається з 2209 записів, що на 31 запис менше, та на основі нього можна вже приступити до проведення описового аналізу клієнтської бази, що дозволяє виявити ключові поведінкові патерни та сформувані обґрунтовані припущення щодо очікуваної сегментації. По-перше, змінні, що відображають споживчу поведінку, в тому числі Recency, Frequency та Monetary, демонструють значну варіацію та чітку позитивну асиметрію, що можна відслідкувати в описовій статистиці (див. додаток Б). Наприклад, середнє значення змінної «Monetary» становить 606,66, тоді як її медіана - лише 396,00, а максимальне значення - 2525. це свідчить про те, що існує невелика група клієнтів з високими витратами, які суттєво впливають на середнє значення. Також на рисунку 3.1 можна побачити, що основна маса клієнтів має витрати в діапазоні до 1000, що відповідає міжквартильному розмаху. Водночас наявність візуально окреслених викидів вище 2000 свідчить про існування нечисленної, але стратегічно важливої групи покупців із високим чеком.

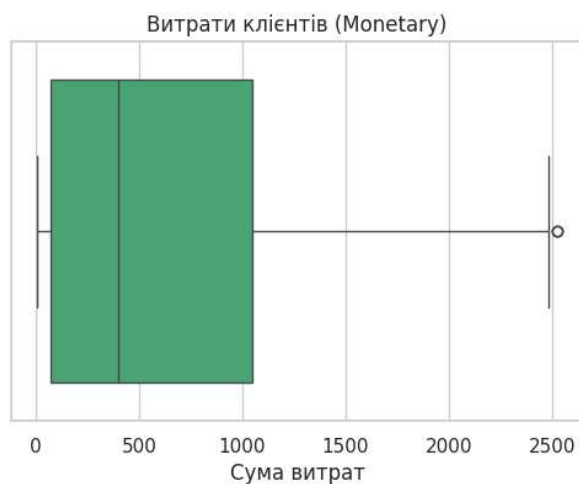


Рис. 3.1. Витрати клієнтів (Monetary)

Джерело: створено автором в середовищі Python (seaborn, matplotlib)

Аналогічна ситуація спостерігається і з Recency (максимум 699, медіана - 356) та Frequency (максимум 32, медіана - 12). Змінні, що характеризують маркетингову реакцію, такі як PromoScore та DiscountPurchases, вказують на нижчу середню активність клієнтів (PromoScore - 0,74, DiscountPurchases - 2,32), але мають вищі максимальні значення (до 6 та 15 відповідно). Це підтверджує існування невеликої, але потенційно цінної групи людей, які позитивно реагують на маркетингові стимули.

Демографічні змінні демонструють стабільну структуру. Зокрема, вік споживачів зосереджений у віковому діапазоні 26-44 роки (25-й та 75-й проценти), що свідчить про переважання дорослої, платоспроможної аудиторії. Рівень доходу має велике стандартне відхилення (приблизно 25 000), що свідчить про нерівномірність купівельної спроможності. Графік показує асиметричний розподіл змінної доходу - переважна більшість клієнтів мають дохід від 20 000 до 80 000, тоді як профіль щільності різко падає після приблизно 100 000. У той же час, існують окремі значення вище 200 000, до 666 666, і їх слід вважати викидами. Такий розподіл підтверджує необхідність попередньої нормалізації даних перед застосуванням методів кластеризації.

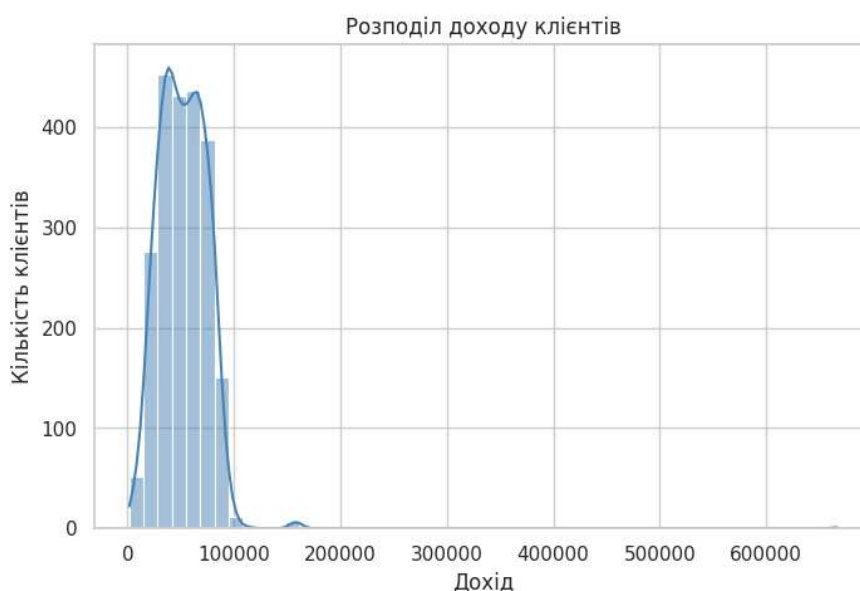


Рис. 3.2. Розподіл доходу клієнта (Income)

Джерело: створено автором в середовищі Python (seaborn, matplotlib)

Змінна «Children» вказує на те, що більшість покупців мають 0 або 1 дитину (медіана - 1,0, максимум - 3,0), що може впливати на типи товарів, які вони споживають.

Загалом, високі значення стандартного відхилення ключових змінних, великий розрив між медіаною та максимальним значенням, а також широкий міжквартильний розмах підтверджують гіпотезу про неоднорідну структуру клієнтської бази. Такий розподіл є методологічно сприятливим для застосування кластерного аналізу, зокрема К-середніх, для визначення сегментів, які є стратегічними для впливу персоналізованого маркетингу.

Для обчислення ступеня лінійного взаємозв'язку між змінними в моделі RFMP-DOV+AIC, було використано коефіцієнт кореляції Пірсона, що визначається за формулою [45]:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}},$$

$r_{xy}$  - коефіцієнт кореляції між змінними  $x$  і  $y$ ;

$x_i, y_i$  - значення спостережень;

$\bar{x}, \bar{y}$  - середні арифметичні значення відповідних змінних;

$n$  - кількість спостережень.

Цей метод дозволяє виявити наявність лінійної залежності між змінними, а також оцінити потенційну загрозу мультиколінеарності, яка може знизити ефективність кластерного моделювання. Висока кореляція між окремими параметрами може спотворити вагу характеристик у процесі сегментації, тому попередній аналіз таких взаємозв'язків є важливим етапом побудови моделі.

Найвища позитивна кореляція спостерігається між частотою змінних і сумою ( $r = 0,82$ ), що є логічним, оскільки частота покупок безпосередньо впливає на загальну суму витрат. Також виявлено значущий взаємозв'язок між доходом і валютою ( $r = 0,67$ ), що свідчить про існування взаємозв'язку між рівнем доходу споживачів і їхньою споживчою активністю. Також було виявлено значний взаємозв'язок між DiscountPurchases і PromoScore ( $r = 0,44$ ), що може свідчити про схильність окремих клієнтів реагувати на маркетингові стимули.

Навпаки, між такими парами змінних, як OnlineRatio і Age, PromoScore і OnlineRatio, було виявлено низькі або близькі до нуля значення коефіцієнтів, що свідчить про їхню незалежність одна від одної. Відсутність значущої кореляції є позитивним сигналом для подальшого кластеризації, оскільки це знижує ризик дублювання інформації в просторі ознак і сприяє стабільності результатів сегментації. Загалом, структура кореляцій підтверджує, що більшість змінних у моделі не є надмірно корельованими, що є бажаним для кластерного аналізу на основі багатовимірних даних.

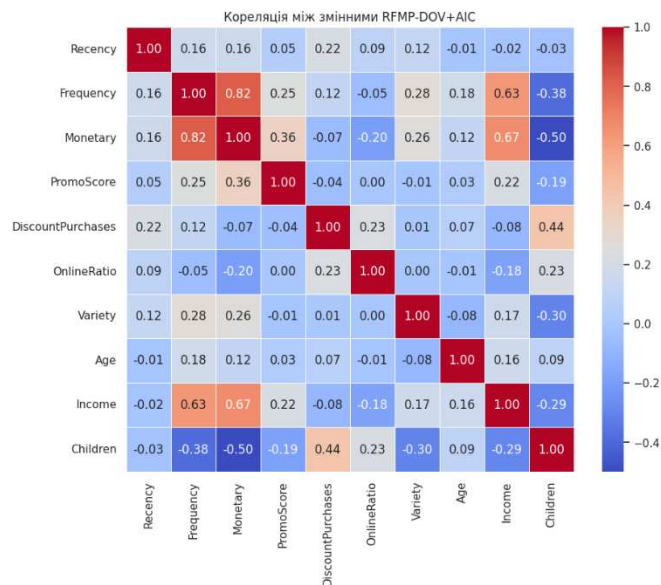


Рис. 3.3. Теплова карта кореляцій між змінними моделі RFMP-DOV+AIC

Джерело: створено автором в середовищі Python (seaborn, matplotlib)

Отже, для побудови моделі сегментації клієнтів RFMP-DOV+AIC було обрано низку поведінкових та демографічних змінних, що відображають основні аспекти взаємодії клієнта з компанією. У таблиці нижче наведено перелік використаних змінних, їх короткі визначення, одиниці виміру, а також метод розрахунку або джерело вхідних даних. Такий підхід дозволив формалізувати методологію побудови моделі та забезпечити її прозорість і відтворюваність.

Для того, щоб забезпечити правильну роботу алгоритмів кластеризації в подальшому, всі змінні, включені в модель RFMP-DOV+AIC, попередньо стандартизуються. Цей крок є необхідним, коли ознаки мають різні масштабні властивості та нерівномірно розподілені. В датасеті присутні деякі змінні

(наприклад, дохід) коливаються від 1 700 до понад 600 000, тоді як OnlineRatio обмежений інтервалом  $[0;1]$ , а Children є дискретною змінною з невеликою кількістю рівнів. Крім того, більшість змінних мають правосторонню асиметрію, що також ускладнює об'єктивну сегментацію.

Для того, щоб вирівняти розподіли та уніфікувати шкали ознак, було використано метод «QuantileTransformer» зі стандартним нормальним розподілом в якості цільового розподілу. Метод трансформує кожен змінну так, щоб її значення були розподілені за законом Гауса. При цьому зберігається ранговий порядок, що є важливою властивістю при роботі з поведінковими показниками.

На рисунку 3.4 можна побачити щільність розподілу змінних до нормалізації. Видно, що переважна більшість ознак мають значну асиметрію, а в деяких випадках спостерігаються викиди, зокрема, Monetary, DiscountPurchases, PromoScore та Income мають значну праву асиметрію, в той час як Children має значну дисперсію.

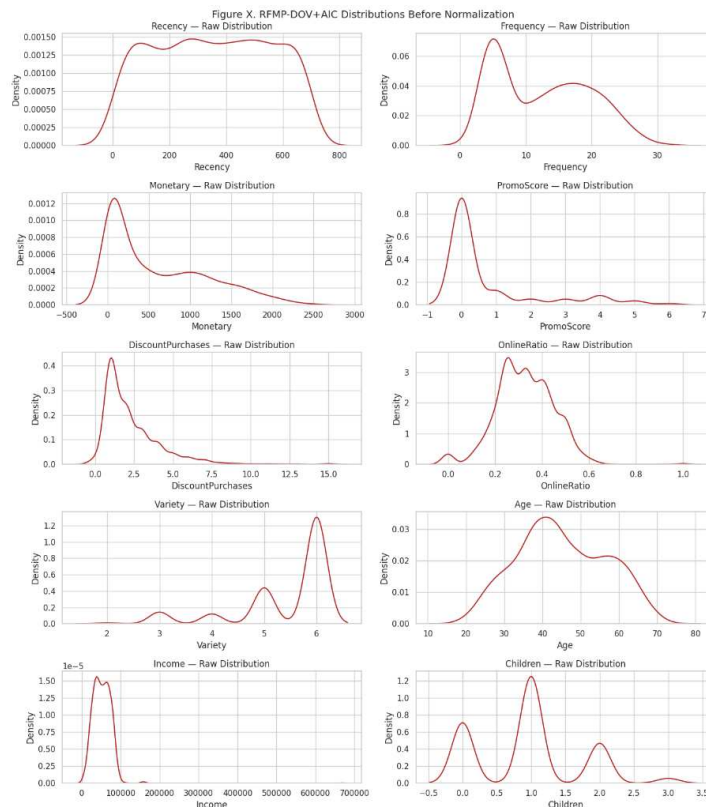


Рис. 3.4. Розподіл змінних до нормалізації

Джерело: створено автором в середовищі Python (seaborn, matplotlib)

Після застосування перетворення, на рисунку 3.5 розподіли набувають симетричного вигляду, здебільшого у вигляді дзвоноподібних кривих. Це дозволяє стверджувати, що статистична стандартизація була досягнута. Результати є очікуваними, оскільки «QuantileTransformer» відображає кумулятивну функцію розподілу кожної змінної на задану функцію розподілу.

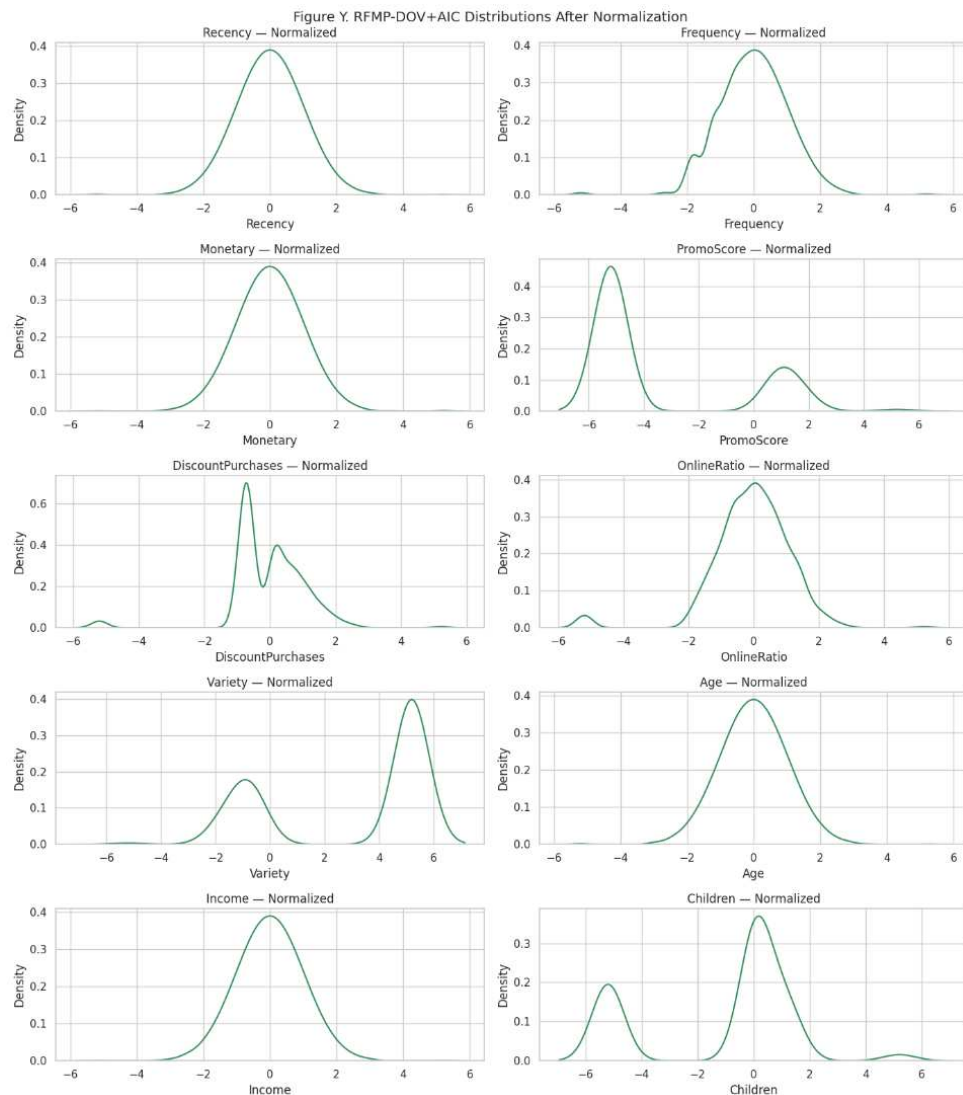


Рис. 3.5. Розподіл змінних після нормалізації

*Джерело: створено автором в середовищі Python (seaborn, matplotlib)*

Наступним етапом є визначення оптимальної кількості кластерів  $k$ , яка дозволяє отримати максимальну інформацію про сегментацію. Було застосовано декілька метрик, які дозволяють максимально оцінити підхід з різних сторін.

Першою метрикою виступає SSE (Sum of Squared Errors), що вимірює загальну внутрішньокластерну дисперсію - суму квадратів відстаней кожної

точки до центроїда свого кластера та визначається за формулою [46]:  $SSE = \sum_{i=1}^n (x_i - \bar{x})^2$ ,  $x_i$  - значення  $i$ -ої точки;  $\bar{x}$  - центроїд (середнє значення) кластера;  $n$  - кількість спостережень у кластері. Зменшення цього показника зі зростанням кількості кластерів є очікуваним, однак важливо виявити момент, коли подальше зростання  $k$  не призводить до суттєвого покращення. На рисунку 3.6 синя крива показує зміну суми квадратів відстаней (SSE) зі збільшенням кількості кластерів, як видно з рисунка, крива вигинається при  $k = 7$  тобто швидкість зменшення помилки значно сповільнюється. Це свідчить про те, що досягнуто балансу між кількістю кластерів та якістю сегментації.

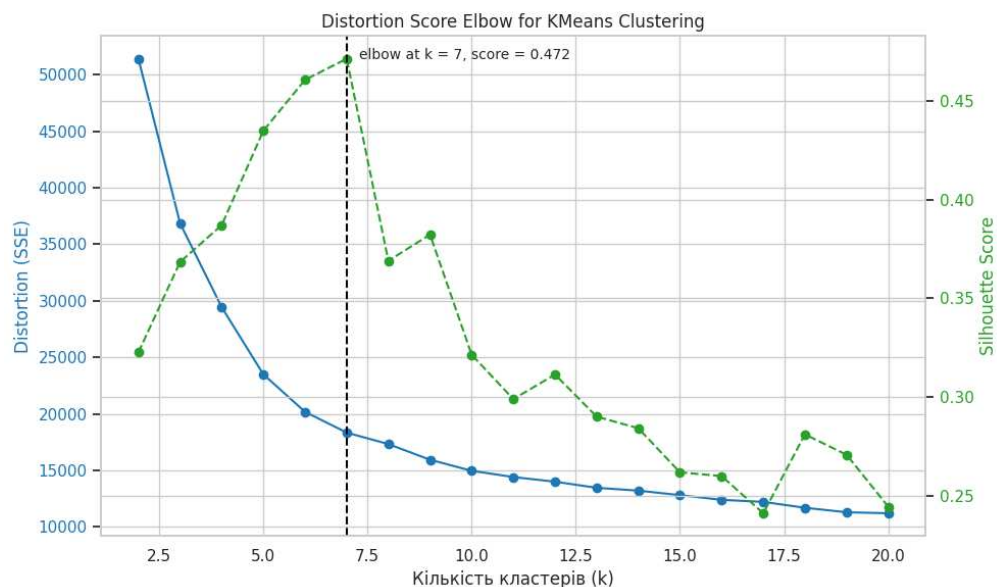


Рис. 3.6. Метод «лікоть» (Elbow Method) та метод силуетного коефіцієнта (Silhouette Score)

*Джерело: створено автором в середовищі Python (seaborn, matplotlib)*

Друга метрика на рисунку 3.6 Silhouette Score дозволяє оцінити ступінь подібності об'єктів до власного кластера в порівнянні з іншими. У проведеному аналізі Silhouette Score досягає максимуму при  $k = 7$ , що свідчить про найкращу якість кластерного поділу саме на цьому рівні. Подальше збільшення кількості кластерів призводить до зниження оцінки, тобто зменшення якості сегментації.

Третім критерієм виступає індекс Калінського-Харабаша (Calinski-Harabasz Index), який обчислюється як відношення дисперсії між кластерами до дисперсії всередині кластерів, за формулою [47]:  $CH = \frac{Tr(B_k)}{Tr(W_k)}$ ,  $Tr(B_k)$  - міжкластерна

дисперсія;  $Tr(W_k)$  - внутрішньокластерна дисперсія;  $n$  - загальна кількість спостережень;  $k$  - кількість кластерів.. Чим вище значення індексу, тим краще визначені кластери. На рисунку 3.7 можна побачити, що найвище значення СН-індексу спостерігається при  $k = 5$ , однак для  $k = 6$  та  $k = 7$  воно залишається на високому рівні. Це дозволяє вважати ці значення конкурентними, особливо в контексті результатів інших метрик.

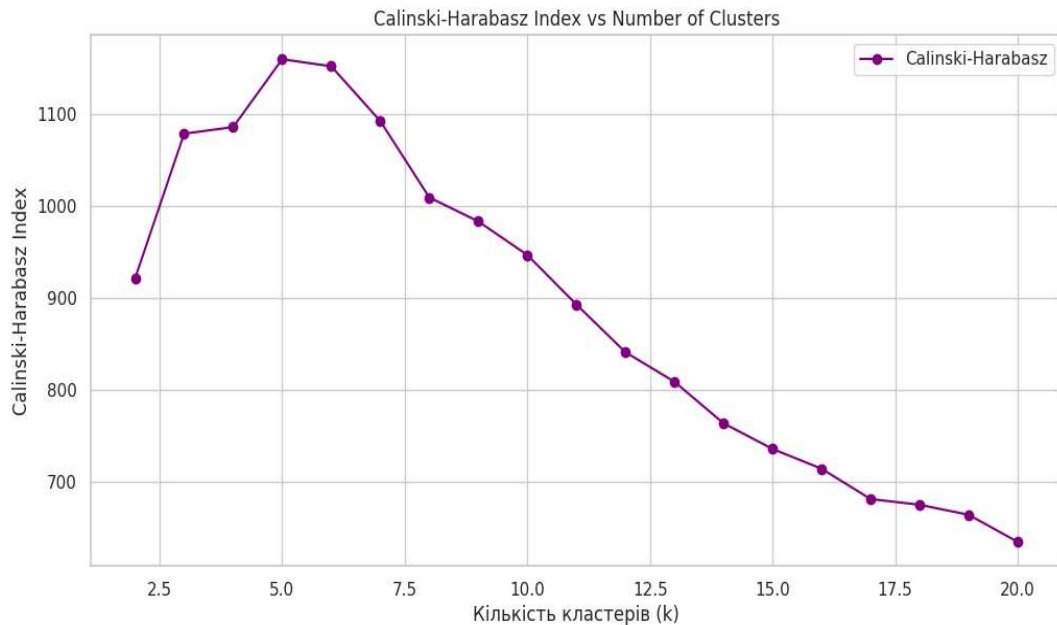


Рис. 3.7. Індекс Калінського-Харабаша

Джерело: створено автором в середовищі Python (*seaborn, matplotlib*)

Четверта метрика - індекс Девіса-Болдіна (Davies-Bouldin Index) враховує як розсіювання точок у межах кластера, так і відстані між центроїдами різних кластерів за формулою:

$$DB = \frac{1}{K} \sum_{i=1}^k \max_{j \neq i} \left( \frac{\Delta(x_i) + \Delta(x_j)}{\partial(x_i, x_j)} \right),$$

де  $\Delta(x_i)$  - середня відстань точок у кластері  $i$  до його центроїда,

$\partial(x_i, x_j)$  - міжкластерна відстань між центрами кластерів  $i$  та  $j$ .

Цей показник необхідно мінімізувати: чим нижчим є його значення, тим краща сегментація. На рисунку 3.8 значення DB-індексу також зафіксовано при  $k = 7$ , що додатково підтверджує доцільність вибору саме цього значення.

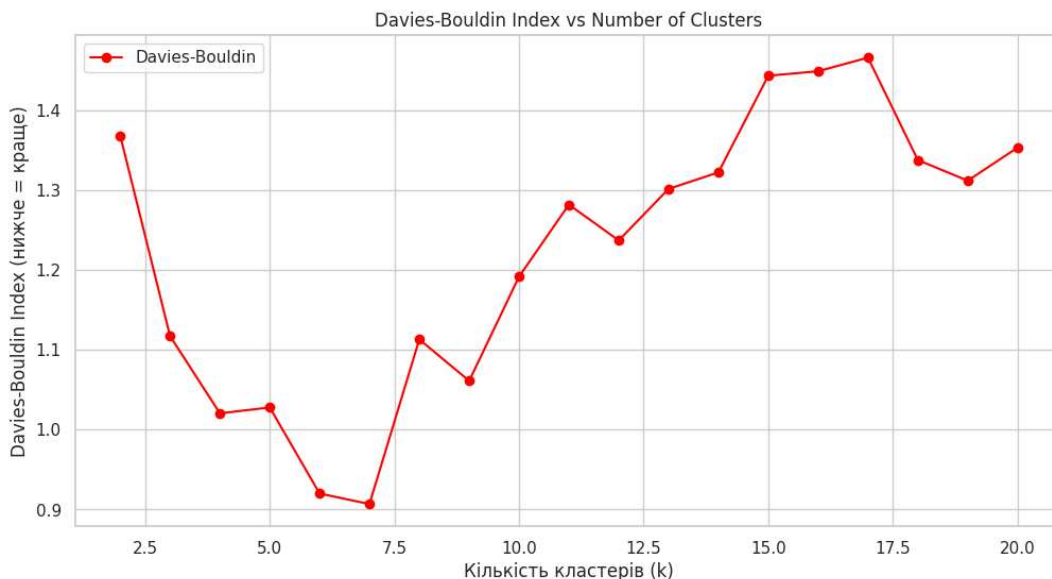


Рис. 3.8. Індекс Девіса-Болдіна

*Джерело: створено автором в середовищі Python (seaborn, matplotlib)*

Проаналізувавши всі метрики, можна прийти до висновку, що оптимальне значення кластерів  $k=7$ , для остаточної кластеризації було використано алгоритм К-середніх. Кожному споживачеві у вибірці було присвоєно кластерну мітку на основі найближчого центру маси, розрахованого за стандартизованими змінними на основі моделі RFMP-DOV+AIC.

### 3.2. Аналіз отриманих результатів та управлінська інтерпретація сегментів

За результатами кластеризації клієнтів за допомогою моделі RFMP-DOV+AIC було отримано сім сегментів на основі метрик Recency (як давно була остання покупка), Frequency (як часто здійснювалася покупка), Monetary (скільки було витрачено) та інших метрик PromoScore, OnlineRatio, DOV та AIC. Кластеризація моделі дозволяє виявити однорідні групи клієнтів, що допомагає адаптувати маркетингові стратегії до специфіки кожного сегмента.

Для наочного уявлення про розподіл клієнтів за кластерами, ми побудували кругову діаграму, яка показує частку кожної групи в загальній структурі. Як видно з діаграми, найбільша частка припадає на кластер 1 (33,3%), що свідчить про його домінування серед споживачів. Кластер 0 охоплює 23,6% споживачів, тоді як інші групи представлені менше: зокрема, на кластери 3 та 6 припадає

лише 2,7% та 7,8% відповідно. Такий дисбаланс дозволяє зробити висновки про неоднорідність сегментів споживачів і дає можливість сфокусувати маркетингову стратегію на найбільш перспективних групах.

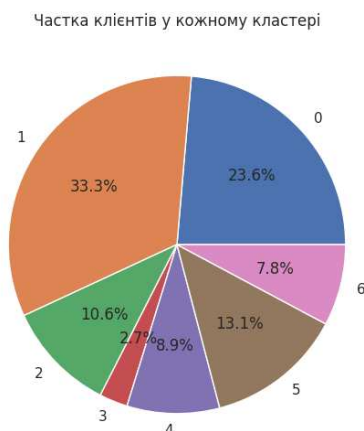


Рис. 3.9. Частка клієнтів у кожному кластері

*Джерело: створено автором в середовищі Python (seaborn, matplotlib)*

Для детального аналізу поведінкових характеристик клієнтів ми побудували діаграму розмаху (boxplot), що відображає розподіл витрат клієнтів (грошові показники) у межах семи кластерів).

На рисунку видно, що кластери демонструють значні відмінності в рівні споживчих витрат: наприклад кластер 2 характеризується найвищим середнім значенням витрат серед усіх груп. Цей кластер також має широкий міжквартильний розмах, що свідчить про значні відмінності у витратах всередині сегмента. Це дає підстави віднести цю групу до «преміум» клієнтів, які роблять великі покупки.

Кластери 0 та 1 мають найнижчі медіанні значення витрат, у той же час, у кластері 1 є ряд винятків - окремі покупці, чиї витрати значно перевищують середні витрати у своєму кластері, це може свідчити про наявність підсегментів з високим потенціалом, які можливо в наступному аналізі будуть відноситись до іншого кластеру.

Кластери 4, 5 та 6 продемонстрували збалансовані показники з медіанними значеннями в межах середнього рівня витрат. Це дає підстави віднести їх до стабільних груп, на які можна орієнтуватися через стратегії утримання клієнтів.

Кластер 3 також має досить широкий розподіл витрат, але медіанний рівень є нижчим за середнє значення вибірки. Це може свідчити про неоднорідність клієнтської бази в цьому кластері.

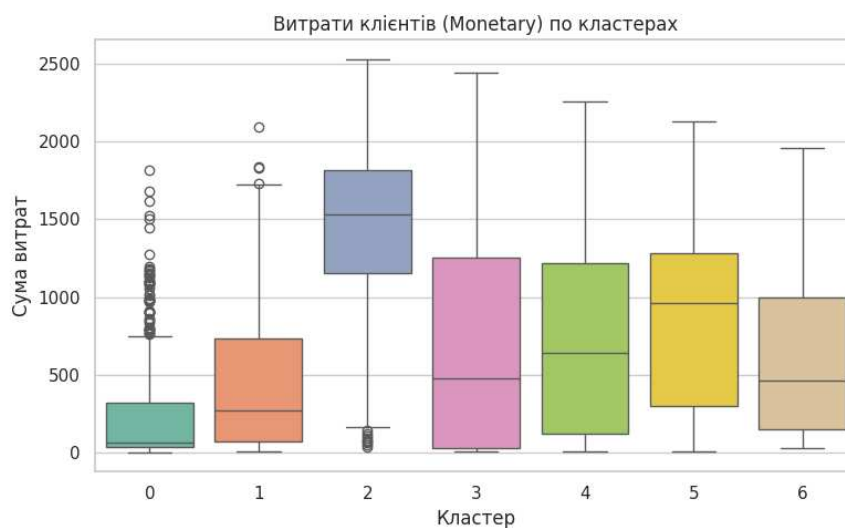


Рис. 3.10. Витрати клієнтів (Monetary) за кластерами

*Джерело: створено автором в середовищі Python (seaborn, matplotlib)*

Щоб оцінити ефективність маркетингових комунікацій, було проаналізовано середнє значення PromoScore у кожному кластері (див. додаток В). Цей показник відображає реакцію споживачів на рекламні акції та загальну маркетингову чутливість. Можна сказати, що кластери 2, 4 та 6 мають найвищі середні значення (понад 2,5), що свідчить про високу готовність взаємодіяти з акціями та спеціальними пропозиціями. Для цих кластерів рекомендується використовувати персоналізовані акції, програми лояльності або індивідуальні знижки.

Інші кластери, навпаки, мають нульовий або дуже низький показник PromoScore, що свідчить про обмежену маркетингову реакцію. У цьому випадку рекомендується використовувати інші стратегії залучення - наприклад, покращення сервісу, брендування або пряму взаємодію з клієнтами через інші канали.

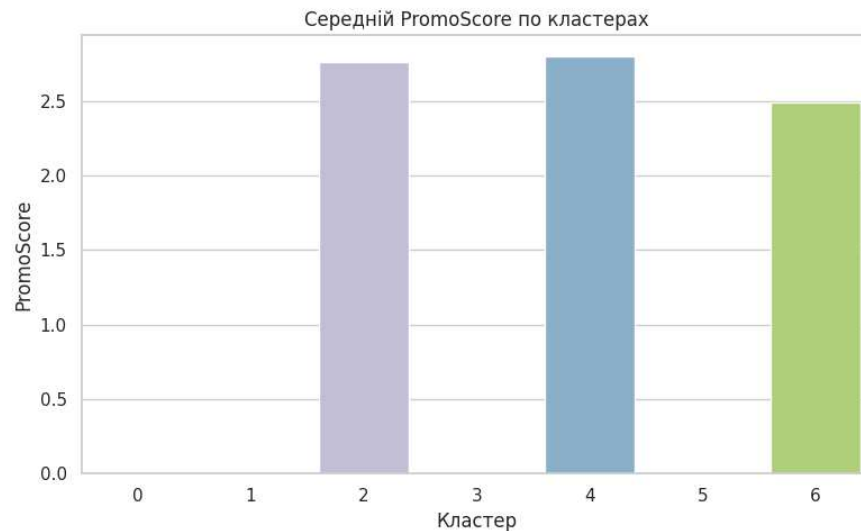


Рис. 3.11. Середній показник *PromoScore* для кожного сегмента

*Джерело: створено автором в середовищі Python (seaborn, matplotlib)*

Для того, щоб краще зрозуміти поведінкові характеристики клієнтів було проаналізовано середні значення двох ключових показників - Frequency (частота покупок) та OnlineRatio (відсоток онлайн-покупок). Можна побачити, що найвища частота покупок спостерігається в кластері 2 із середнім значенням Frequency 18,97, що свідчить про те, що клієнти лояльні до компанії і часто здійснюють транзакції. Такі клієнти є стратегічно важливими для підтримки стабільного доходу і допомагають у впровадженні програм лояльності.

Водночас кластер 6 має найвищий середній рівень OnlineRatio (0,39), що свідчить про те, що в сегменті переважають онлайн-канали продажів. Це відкриває потенціал для стратегій цифрового маркетингу, таких як ремаркетинг, персоналізовані електронні розсилки або таргетована реклама в соціальних мережах. Хоча кластери 3 і 5 мають помірну частоту покупок (близько 12-15), їхній OnlineRatio все ще нижчий за 0,30, що може свідчити про більшу залежність від традиційних каналів продажів (наприклад, традиційні магазини або каталоги).

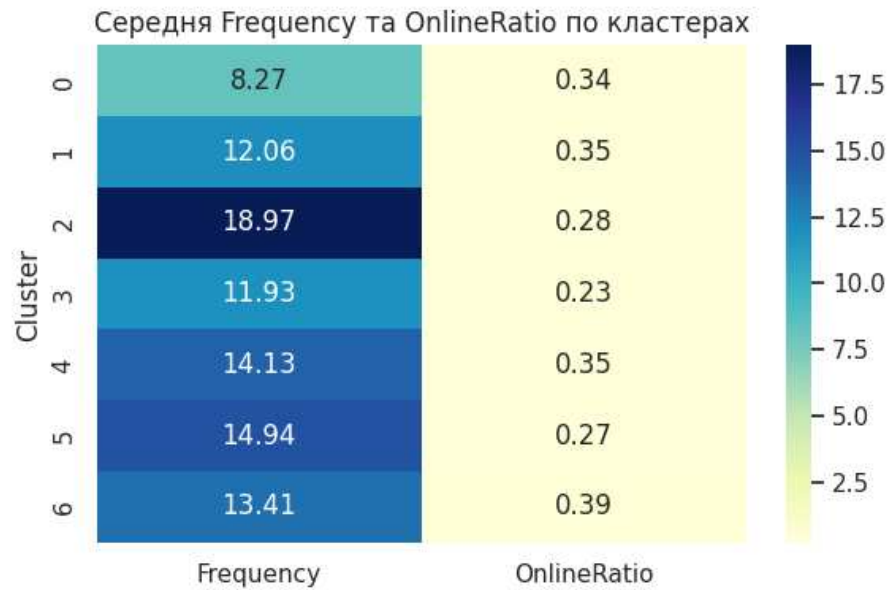


Рис. 3.12. Теплова карта середніх значень Frequency та OnlineRatio за кластерами  
Джерело: створено автором в середовищі Python (seaborn, matplotlib)

Таким чином, поєднання показників частоти та типу каналу дозволяє чітко розмежувати сегменти, орієнтовані на цифрові технології, та сегменти, орієнтовані на офлайн, що дає змогу адаптувати комунікаційні стратегії до відповідних уподобань кожної з груп.

Для наступного аналізу отриманих груп було розраховано середні значення всіх 11 ключових змінних моделі в кожному кластері, а також визначено кількість споживачів у кожному кластері та їх відсотковий розподіл у загальній структурі вибірки.

Результати показують, що існують значні відмінності між кластерами з точки зору поведінкових, рекламних та демографічних характеристик. Наприклад, кластер 2 демонструє найвищі показники частоти покупок, витрат та рекламних акцій, що свідчить про прибуткову та зацікавлену групу споживачів. Водночас кластер 3 має найнижчу частку покупців (2,72%) і нижчі значення показників PromoScore та «Діти», що може свідчити про менш зацікавлену аудиторію. Подальший аналіз кластерів дозволив нам розробити таргетовану маркетингову стратегію з урахуванням профілю кожної групи

Сім кластерів були якісно описані на основі візуалізації та середніх показників моделі та пророблено ключові атрибути та пояснення маркетингових

кампаній та менеджменту для кожного сегмента: сегмент 1 (високоцінні лояльні клієнти), мають найвищі грошові та частотні показники, а також короткий період часу з моменту останньої покупки. Це дуже цінні постійні клієнти, які регулярно роблять великі замовлення. Їхній показник PromoScore середній або високий, що означає, що вони активно беруть участь у промоакціях. Впровадження програм лояльності, персоналізованих знижок і преміум-сервісів добре працює для них, підвищуючи їхню задоволеність і подовжуючи життєвий цикл клієнта.

Сегмент 2 (лояльні клієнти з середнім рівнем доходу), мають високу частоту та середню грошову вартість, де покупки відбуваються досить часто, але на помірну суму за одне замовлення. Ці клієнти постійно повертаються, але витрачають менше, ніж сегмент 1. PromoScore показує, що вони зацікавлені в акціях, потрібно заохочувати їх збільшувати кількість покупок (перехресні продажі та продажі), наприклад, пропонуючи додаткові товари у вигляді пакетів або персоналізованих пропозицій. Важливо підтримувати їхню лояльність і поступово підвищувати середню заробітну плату.

Сегмент 3 (нові активні клієнти), показник нещодавніх покупок дуже низький, частота покупок відносно низька (зазвичай це перша або друга покупка), а сума витрачених коштів невелика. OnlineRatio і PromoScore можуть сильно відрізнятись від клієнта до клієнта. Це, швидше за все, нові або недавні покупці, які ще не виробили постійної поведінки. Основне завдання - перетворити їх на постійних клієнтів, для цього слід використовувати привітальні пропозиції, бонуси або знижки на наступну покупку, щоб стимулювати подальшу активність.

Сегмент 4 (часті ошадливі покупці), роблять покупки часто, але витрачають помірну або низьку суму грошей. Вартість однієї покупки невелика, але інтенсивність покупок висока, наприклад, вони часто купують дешеві товари або користуються підписками. Клієнти цієї категорії зазвичай мають високий показник PromoScore, оскільки орієнтуються на знижки. Необхідно спробувати збільшити їхній середній чек: впроваджувати перехресні продажі, «комбіновані»

пропозиції, нараховувати бонуси. Тепер, коли вони залучені, важливо не втратити їхній інтерес і надавати часті оновлення.

Сегмент 5 (активні онлайн-клієнти), клієнти з високою частотою та високою присутністю в мережі. Вони регулярно здійснюють покупки, переважно через онлайн-канали (сайт, мобільні додатки). Зазвичай вони мають середній або високий рівень валюти і низький рівень новизни. Це молоді або технологічно просунуті покупці, для них добре працюють діджитал-стратегії: push-повідомлення, email-маркетинг, програми винагород у мобільних додатках. Щоб утримати цей сегмент і підвищити їхню лояльність, можна створити спеціальні промо-програми, присвячені онлайн-каналам (наприклад, знижки за завантаження додатків).

Сегмент 6 (пасивні «сплячі» клієнти), клієнти, які є дуже недавніми (не здійснювали покупки протягом тривалого часу), нечастими і мають низьку цінність. Вони ледь активні, робили покупки в минулому, але не повертаються. Зазвичай вони мають низький показник PromoScore і погано реагують на традиційні стимули. Це типовий сегмент ризику: існує потреба в кампаніях з реактивації, таких як персоналізовані пропозиції «Ми сумуємо за тобою», нагадування або спеціальні одноразові знижки. Навіть невеликі стимули можуть повернути цих клієнтів до каналу купівлі. Також важливо виявити причини відтоку (опитування, фокус-групи) та усунути бар'єри (якість обслуговування, вартість доставки тощо).

Сегмент 7 (малоцінні, нерегулярні клієнти), низька частота, низька вартість і часто високий рівень новизни. Ці клієнти мають дуже малий досвід покупок - одноразові або випадкові покупки і витрачають дуже мало. Зазвичай вони мають найнижчий PromoScore і змінний OnlineRatio. Вони не є пріоритетним джерелом доходу з точки зору маркетингу, але їх можна залучити за допомогою масових стимулів - загальних знижок, рекламних електронних листів, взаємодії через соціальні мережі. Якщо сегмент великий, можна додати йому цінності, включивши в програму лояльності або реферальну програму.

Кожен сегмент вимагає окремого підходу. Наприклад, для найбільш прибуткових клієнтів (сегмент 1) рекомендується зосередитися на програмах лояльності та персоналізації, а для пасивних клієнтів (сегмент 6) - реактивувати за допомогою спеціальних пропозицій. Аналізуючи сегменти таким чином, ми можемо чітко бачити, які сегменти клієнтів потребують збільшення витрат і лояльності, а які шукають певних стимулів або взагалі перебувають у «групі ризику».

На основі характеристик кожного сегмента ми розробили управлінські рекомендації та відповідні маркетингові заходи, наприклад підтримка лояльності, для сегментів з високими частотними та грошовими показниками (особливо для сегментів 1 та 2) необхідно розробити персоналізовані програми лояльності, VIP-сервіси та ексклюзивні знижки. Це допоможе утримати їх і заохотити до повторних покупок. Нараховування бонусів, міні-подарунки з кожною покупкою, повернення «сплячих» клієнтів, кампанії з повторного залучення використовуються для сегментів з низькою активністю (особливо для сегмента 6). Це може бути персоналізовані email-розсилки зі спеціальними знижками, подарунками або безкоштовною доставкою, щоб змусити клієнтів повертатися. Цей метод допомагає «повернути» втрачених глядачів і запобігає повному зникненню клієнтської бази. Збільшення середнього чеку, для постійних клієнтів з невеликими чеками (сегмент 4) можуть бути ефективними стратегії допродажу та перехресних продажів. Наприклад, пропонувати комплекти товарів, додавати супутні продукти або заохочувати купувати дорожчі варіанти - це допоможе підвищити Monetary без зниження Frequency.

Онлайн-канали та діджитал-промо, покупці сегменту 5 часто здійснюють покупки онлайн, і їх слід залучати через цифрові канали - налаштування цільових кампаній в інтернеті, надсилання сповіщень у додатках та промо-акції в соціальних мережах. Також рекомендується надавати спеціальні промокоди для онлайн-замовлень, щоб зберегти їхню лояльність до онлайн-платформи.

Надання цінності «новим» клієнтам, сегмент 3 вимагає заходів, спрямованих на формування звички до бренду. Це можуть бути привітальні

бонуси для перших покупців, інструкції про те, як користуватися продуктом, або реферальна програма, щоб вони приводили своїх друзів і поверталися самі.

Акційні пропозиції для любителів промоакцій, кластери з високим показником PromoScore (наприклад, сегмент 4) реагують на знижки. Для них ви можете проводити регулярні акції «купи два, отримай один безкоштовно» або пропонувати накопичувальні знижки залежно від кількості покупок. Використовуючи цей інструмент, важливо не знижувати маржу занадто сильно, оскільки таких клієнтів важко утримати без додаткових стимулів.

Детальний кластерний аналіз і стратегії, що базуються на ньому, мають велику практичну цінність. Перш за все, вони дозволяють ефективніше розподіляти маркетингові ресурси: фокусуватися на найцінніших сегментах (максимізуючи ROI) та індивідуально працювати з «ризиковими» клієнтами, щоб повернути їх. Також сегментація допомагає персоналізувати комунікацію: кожен клієнт отримує пропозиції, які відповідають його поведінці. Це підвищує ефективність кампанії та задоволеність клієнтів. Відділи продажів та маркетингу отримують чіткі інсайти, а саме хто чітко реагує на знижки, хто любить швидку доставку, хто є фанатом новинок тощо. Ці дані допомагають розробляти довгострокову стратегію роботи з клієнтами та приймати рішення на основі аналізу їхньої поведінки.

Таким чином, визначення семи сегментів за допомогою методології RFMP-DOV+AIC надає компаніям інструменти для точного налаштування своїх маркетингових рішень. Поєднання кількісних показників (RFM) з поведінковими (PromoScore, OnlineRatio) дає комплексне уявлення про клієнтську базу. Результати аналізу візуалізуються у вигляді графіків та діаграм, які забезпечують прозору інтерпретацію, менеджери бачать чіткі відмінності між групами і можуть швидко скоригувати свою стратегію для кожної групи.

### 3.3. Оцінка ефективності моделі та можливості її подальшого використання

Після отримання результатів кластеризації на основі моделі RFMP-DOV+AIC та розробки маркетингових рішень, ми оцінили її ефективність,

порівняємо з іншими методами сегментації та визначимо перспективи практичного застосування в бізнесі. Такий аналіз дозволяє виявити сильні та слабкі сторони запропонованої моделі, підтвердити її конкурентні переваги над класичною моделлю RFM та її модифікаціями, а також сформулювати рекомендації щодо впровадження моделі в реальний маркетинговий процес.

Основною перевагою моделі RFMP-DOV+AIC є її здатність комплексно охоплювати транзакційні, поведінкові та демографічні аспекти активності клієнтів. На відміну від класичної RFM-моделі, яка враховує лише тривалість, частоту та суму покупок, розширена модель інтегрує інші змінні, такі як цифрове залучення (OnlineRatio), реакція на маркетингові кампанії (PromoScore), чутливість до знижок (DiscountPurchases), різноманітність покупок (Variety), а також вік, дохід та наявність дітей.

Це дає змогу глибше проаналізувати поведінку споживачів. Зокрема, сегментація каналів дозволяє розрізняти між собою офлайн- та онлайн-покупців, а показники активності зі знижками допомагають відокремити «мисливців за знижками» від елітних покупців. Додавання демографічного контексту дозволяє точніше адаптувати маркетингову стратегію до життєвого циклу клієнта та його купівельної спроможності. Крім того, така багатовимірність підвищує точність кластеризації і дозволяє виявити приховані закономірності, які традиційні методи не можуть зафіксувати. Це допомагає створювати релевантні сегменти та підвищувати ефективність персоналізованих комунікацій.

Незважаючи на очевидні переваги, модель RFMP-DOV+AIC має певні обмеження, які слід враховувати при її застосуванні. Одним з основних недоліків є те, що через розширення моделі шляхом додавання змінних підвищує вимоги до якості та повноти даних. Для розрахунку таких показників, як PromoScore, DiscountPurchases, OnlineRatio, Variety та демографічних параметрів, необхідно мати детальну інформацію про кожного покупця. У багатьох випадках ці дані частково відсутні або збираються нерегулярно, що ускладнює точне моделювання сегментації. За відсутності повного набору даних зростає ризик

викривлення результатів або необхідність зменшення кількості змінних, що, в свою чергу, знижує аналітичну цінність моделі.

Крім того, збільшення кількості змінних значно ускладнює процес інтерпретації результатів. Якщо класичні RFM-моделі можна інтерпретувати в термінах кількох простих категорій, то сегментація, побудована на основі десяти параметрів, вимагає більш глибокого аналізу та візуалізації. Велика кількість ознак збільшує когнітивне навантаження на користувача, особливо коли модель використовується в крос-функціональному середовищі, де результати мають бути зрозумілими для стейкхолдерів, які не є експертами. Використання багатовимірної кластеризації також супроводжується певною методологічною складністю: зростає ймовірність неочевидної сегментації бізнес-контенту, що потребуватиме додаткової інтерпретаційної роботи.

Ще одна складність - технічна підготовка даних. Через неоднорідність змінних (як за масштабом, так і за типом вимірювання), необхідним кроком перед аналізом є їхня нормалізація. Без належного масштабування певні змінні можуть мати непропорційний вплив на результати. Крім того, виникає питання вагових коефіцієнтів: оскільки не всі фактори мають однакову управлінську цінність, компаніям може знадобитися диференціювати їхній вплив, що вимагає експертних методів або статистичного зважування. Без цього результати сегментації можуть бути упередженими або непридатними для практичного використання.

Варто також враховувати ризик надмірної деталізації. Чим більше вимірів містить модель, тим більша ймовірність створення занадто дрібних або фрагментованих сегментів, що ускладнює їх подальше застосування в маркетингових кампаніях. Для компаній, які прагнуть вести бізнес з обмеженою кількістю чітко визначених цільових груп, надмірна деталізація може бути скоріше перешкодою, ніж перевагою. У цьому випадку необхідна подальша агрегація або оптимізація кількості сегментів.

Хоча модель RFMP-DOV+AIC є набагато ширшою, ніж класичний підхід RFM, вона все ще не охоплює всі поведінкові змінні. Зокрема, модель не включає

метрики задоволеності, лояльності або тривалості взаємодії з брендом, які можуть бути критично важливими в деяких сферах (наприклад, послуги або підписка на послуги). За потреби модель можна було б доповнити цими параметрами, але це, в свою чергу, призвело б до ускладнення структури та збільшення попиту на доступну інформацію. Таким чином, незважаючи на свою глибину та аналітичну цінність, модель RFMP-DOV+AIC є складним інструментом, який вимагає ретельної підготовки даних, методологічної обґрунтованості та технічних можливостей компанії для ефективного впровадження.

Порівняно з вже існуючими модифікаціями моделі RFM, які були розглянуті вище, підхід RFMP-DOV+AIC вирізняється своєю комплексністю та універсальністю. На відміну від точкових розширень, де до класичного RFM додається один параметр (наприклад, довжина, тип або відсоток знижки), запропонована модель поєднує в собі одразу кілька додаткових вимірів - поведінковий, каналний та демографічний, що дозволяє не тільки зафіксувати цінність клієнта для компанії, як у класичному RFM, але й краще зрозуміти причини цієї цінності, зокрема: через канал взаємодії, реакцію на промоакції, структуру покупок та соціально-економічні характеристики клієнта.

Розглянувши конкретно на прикладах, то можна помітити наступну тенденцію: на відміну від моделі LRFM, які враховують лише тривалість взаємодії, RFMP-DOV+AIC частково охоплює цей аспект, поєднуючи Recency з демографічними змінними (вік, сімейний стан). Так само, на відміну від RFMT, яка фокусується на часових інтервалах, ця модель фокусується на поточній поведінці, що є більш цінним у динамічному середовищі.

У той час як модель RFM-D (або RFM-V) враховує різноманітність товарних категорій, але не охоплює канали та промоакції, RFMP-DOV+AIC дозволяє сегментувати клієнтів, які мають навіть однаковий рівень різноманітності покупок, але різну чутливість до промоакцій або цифрових каналів. Аналогічно, модель RFM+DP фокусується на ціновій чутливості, тоді як RFMP-DOV+AIC не

лише розрізняє «мисливців за знижками», але й дозволяє охарактеризувати їх: наприклад, це молоді онлайн-покупці з широким розмаїттям замовлень.

На відміну від LRFMS, яка підвищує задоволеність, запропонована модель менше покладається на суб'єктивні опитування, а більше - на об'єктивні дані про поведінку покупців. Таким чином, RFMP-DOV+AIC можна розглядати як інтегровану модель, яка поєднує сильні сторони кожного з альтернативних підходів, мінімізуючи їхні обмеження. Головною перевагою є більш глибоке та реалістичне уявлення про клієнта, а головним викликом - зростаючий попит на аналітичні дані та ресурси. Якщо компанії матимуть доступ до якісної інформації про клієнтів, ця модель надасть можливості для більш точної, персоналізованої та ефективної маркетингової сегментації.

### ВИСНОВКИ ДО РОЗДІЛУ 3

У цьому розділі було реалізовано практичну частину дослідження, в межах якої побудовано розширену модель RFMP-DOV+AIC, що враховує як класичні параметри RFM-аналізу, так і додаткові демографічні та поведінкові характеристики клієнтів. Проведено підготовку та очищення даних, виконано описовий аналіз вибірки, а також здійснено кластеризацію з подальшим визначенням оптимальної кількості сегментів. Отримані результати дозволили виокремити цільові групи споживачів і надати управлінські рекомендації щодо підвищення ефективності маркетингових заходів. Порівняння з традиційними підходами підтвердило доцільність використання вдосконаленої моделі у практиці електронної комерції.

## ВИСНОВКИ

У межах даної кваліфікаційної роботи було всебічно досліджено проблему сегментації клієнтів в електронній комерції з урахуванням викликів цифрової трансформації. Актуальність теми зумовлена зростаючою потребою компаній в ефективному управлінні клієнтською базою, персоналізації маркетингових активностей та підвищенні якості комунікації із споживачем.

На основі проведеного теоретичного аналізу було виявлено, що традиційні методи сегментації, зокрема класичний RFM-аналіз, мають обмеження в умовах багатоканальності та мінливої поведінки споживачів. Було доведено доцільність розширення цих моделей за рахунок додаткових змінних, зокрема демографічних, поведінкових та цифрових показників. У рамках другого розділу було досліджено існуючі методики RFM-аналізу, систематизовано їх варіанти реалізації та надано порівняльну характеристику.

Практична частина роботи присвячена розробці та впровадженню вдосконаленої моделі RFMP-DOV+AIC. Було здійснено комплексну підготовку даних, побудову нових змінних, описовий аналіз, стандартизацію та кластеризацію вибірки на основі алгоритму K-means. Обґрунтовано вибір змінних, включених до моделі, що дозволило врахувати не лише історію покупок, а й реакцію клієнтів на промоакції, канали взаємодії, вік, дохід, наявність дітей та інші релевантні параметри.

Отримані результати підтвердили ефективність моделі: сегменти чітко розділені за поведінковими характеристиками, що дозволяє адаптувати маркетингові стратегії під кожну групу. Порівняно з класичним RFM, запропонований підхід виявився точнішим і практично релевантнішим.

Модель може бути інтегрована в CRM-системи компаній для підвищення CLV та зниження відтоку клієнтів, а також адаптована до інших сфер із регулярною взаємодією зі споживачами. Таким чином, поставлену мету досягнуто, завдання виконано, а результати дослідження мають як теоретичне, так і прикладне значення.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Про основні засади забезпечення кібербезпеки України: Закон України від 05.10.2017 р. № 675-VIII. URL:<https://zakon.rada.gov.ua/laws/show/675-19>
2. Краус К.М., Краус Н.М., Манжура О.В. Електронна комерція та Інтернет-торгівля: навчально-методичний посібник. Київ: Аграр Медіа Груп, 2021. 454 с. URL: <https://elibrary.kubg.edu.ua/id/eprint/37044/>
3. Ecommerce Europe, EuroCommerce. European E-commerce Report 2024. Brussels, 2024. 124p. URL:[https://ecommerce-europe.eu/wpcontent/uploads/2024/10/CMI2024\\_Complete\\_light\\_v1.pdf](https://ecommerce-europe.eu/wpcontent/uploads/2024/10/CMI2024_Complete_light_v1.pdf)
4. E-commerce statistics for individuals. Eurostat – Statistics Explained. 2025. URL:[https://ec.europa.eu/eurostat/statistics-explained/index.php?title=E-commerce\\_statistics\\_for\\_individuals](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=E-commerce_statistics_for_individuals)
5. Top Online Shopping Categories (2025). Shopify. 2025. URL: <https://www.shopify.com/blog/top-online-shopping-categories>
6. Top Ecommerce Companies in 2024. Oberlo. 2024. URL: <https://www.oberlo.com/statistics/top-ecommerce-companies>
7. Maksymenko I., Akimov A., Bikulov D. Development of the e-commerce market in Ukraine in the context of modern challenges. Baltic Journal of Economic Studies. 2024. 10(2). P. 177–186. doi: <https://doi.org/10.30525/2256-0742/2024-10-2-177-186>
8. eCommerce Industry in Ukraine 2017–2029. ECDB. 2025. URL: <https://ecdb.com/resources/sample-data/market/ua/all>
9. Дослідження українського ринку eCommerce: друге півріччя 2024. Promodo. 2024. URL: <https://www.promodo.ua/ukrayinskiy-ecommerce-2024-2#zahalnyi-trafik-ukrainskoho-onlain-prostoru>
10. Рейтинг найбільших ритейлерів в українському e-commerce – 15 кращих. UA-Retail. 2024. URL: <https://ua-retail.com/2024/03/rejting-najbilshix-ritejleriv-v-ukrainskomu-e-commerce-15-krashhix/>
11. Arora N., Ensslen D., Fiedler L., Liu W.W., Robinson K., Stein E., Schüler G. The value of getting personalization right—or wrong—is multiplying.

- McKinsey & Company. 2021. URL: <https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or-wrong-is-multiplying>
12. Big Data Analytics for Customer Segmentation Success. Maestra.io. 2023. URL: <https://maestra.io/blog/marketing/marketing-data-science-segmentation-of-customers/#WhatisBigDataandHowDoesDataScienceEnhanceCustomerSegmentation>
13. Stein E., Robinson K., Wolfer A., Almeida G., Huang W. Unlocking the next frontier of personalized marketing. McKinsey & Company. 2025. URL: <https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/unlocking-the-next-frontier-of-personalized-marketing>
14. Ma X., Gu X. Big data analytics for customer segmentation success. Heliyon. 2024. 10(5): e29038. doi: <https://doi.org/10.1016/j.heliyon.2024.e29038>
15. Hoenig H. Understanding Market Segmentation: A Comprehensive Guide. Investopedia. 2025. URL: <https://www.investopedia.com/terms/m/marketsegmentation.asp>
16. Customer Segmentation: How to Segment Users & Clients Effectively. HubSpot Blog. 2020. URL: <https://blog.hubspot.com/service/customer-segmentation>
17. Big Data in insurance: a game-changer for the future growth. N-iX. 2023. URL: <https://www.n-ix.com/big-data-in-insurance/>
18. Carpenter A. Market segmentation: definition, types, benefits, & best practices. Qualtrics. 2020. URL: <https://www.qualtrics.com/experience-management/brand/what-is-market-segmentation/>
19. Сегментація ринку – поняття, ознаки, методи і етапи сегментування, приклади. Моя освіта. 2023. URL: <https://moyaosvita.com.ua/ekonomika/segmentaciya-rinku-ponyattya-oznaki-metodi-i-etapi-segmentuvannya-prikladi/>

20. Fedirko O., Zatonatska T., Wołowiec T., Skowron S. Data Science and Marketing in E-Commerce Amid COVID-19 Pandemic. *European Research Studies Journal*. 2021. XXIV(Special Issue 2): 3–16. doi: <https://doi.org/10.35808/ersj/2187>
21. Customer Segmentation in Salesforce: Options and Benefits. Twistellar. 2024. URL: <https://twistellar.medium.com/customer-segmentation-in-salesforce-options-and-benefits-049decc7a33a>
22. Varma M.A., Akhil A.S., Anoushka K.S., Maddukuri P.S. Use of Big Data in the Process of Customer Segmentation in the Retail Sector. *Technoarete Transactions on Advances in Data Science and Analytics*. 2022. 1(2): 8–13. doi: <https://doi.org/10.36647/TTADSA/01.02.A002>
23. Big Data Analytics for Customer Segmentation Success. Maestra.io. 2023. URL: <https://maestra.io/blog/marketing/marketing-data-science-segmentation-of-customers/>
24. Customer Segmentation: How to Segment Users & Clients Effectively. HubSpot Blog. 2025. URL: <https://blog.hubspot.com/service/customer-segmentation>
25. Mehta J. Segmenting customers by behavior: pros and cons. Abmatic AI. 2023. URL: <https://abmatic.ai/blog/segmenting-customers-by-behavior-pros-and-cons>
26. Hughes A.M. *Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program*. 4th ed. New York: McGraw-Hill, 2012. 608 p. ISBN 978-0-07-177348-5
27. A Comprehensive Guide to the RFM Model: Recency, Frequency & Monetary Value. Omniconvert. 2025. URL: <https://www.omniconvert.com/blog/rfm-model/>
28. Lee, Ernesto. RFM Analysis and Clustering for Customer Segmentation. DrLee.io. 2025. URL: <https://drlee.io/rfm-analysis-and-clustering-for-customer-segmentation-1087fb2a4eb2>

29. What Is RFM? Definition, Purpose, and Examples. Mailchimp. 2025. URL: <https://mailchimp.com/resources/rfm-analysis/>
30. Sobhi N. What Is RFM Customer Segmentation in E-commerce? Gameball. 2024. URL: <https://www.gameball.co/blog/rfm-customer-segmentation-in-ecommerce>
31. Biswal A. What Is a Percentile in Statistics and How to Calculate It? Simplilearn. 2024. URL: <https://www.simplilearn.com/tutorials/data-analytics-tutorial/percentile-in-statistics>
32. Kumar S., Rani R., Pippal S.K., Agrawal R. Customer segmentation in e-commerce: K-means vs hierarchical clustering. TELKOMNIKA Telecommunication Computing Electronics and Control. 2025. Vol. 23, No. 1. P. 119–128. doi: <https://doi.org/10.12928/telkomnika.v23i1.26384>
33. Fahrudin N.F., Rindiyan R. Comparison of K-Medoids and K-Means Algorithms in Segmenting Customers Based on RFM Criteria. E3S Web of Conferences. 2024. Vol. 484. Article No. 02008. P. 1–17. doi: <https://doi.org/10.1051/e3sconf/202448402008>
34. Abdulhafedh A. Incorporating K-means, Hierarchical Clustering and PCA in Customer Segmentation. Journal of City and Development. 2021. 3(1): 12–30. doi: <https://doi.org/10.12691/jcd-3-1-3>
35. Alizadeh Zoeram A., Karimi Mazidi A.R. A New Approach for Customer Clustering by Integrating the LRFM Model and Fuzzy Inference System. Iranian Journal of Management Studies. 2018. Vol. 11, No. 2. P. 351–378. doi: <https://doi.org/10.22059/ijms.2018.242528.672839>
36. Seyhan M. KMeans Clustering on RFM-T Segmentation with Python for Online Retail Data. Analytics Vidhya. 2021. URL: <https://medium.com/analytics-vidhya/kmeans-clustering-on-rfm-t-segmentation-with-python-for-online-retail-data-ad9a2a7ab59>
37. Rana A.D., Hadisantoso Q.C.M., Girsang A.S. RFM-T Model Clustering Analysis in Improving Customer Segmentation. International Journal of

- Computing and Digital Systems. 2025. 17(1). P. 1–11. doi: <https://doi.org/10.12785/ijcds/1571024871>
38. Smaili M.Y., Hachimi H. New RFM-D classification model for improving customer analysis and response prediction. *Ain Shams Engineering Journal*. 2023. 14(6): 102254. doi: <https://doi.org/10.1016/j.asej.2023.102254>
39. Mahfuza R., Islam N., Toyeb M., Emon M.A.F., Chowdhury S.A., Alam M.G.R. LRFMV: An efficient customer segmentation model for superstores. *PLoS ONE*. 2022. 17(12): e0279262. doi: <https://doi.org/10.1371/journal.pone.0279262>
40. Wang Sh., Sun L., Yu Y. A dynamic customer segmentation approach by combining LRFMS and multivariate time series clustering. *Scientific Reports*. 2024. 14(1): 17491. doi: <https://doi.org/10.1038/s41598-024-68621-2>
41. Antonius V.H., Fitriyah D. Enhancing Customer Segmentation Insights by using RFM + Discount Proportion Model with Clustering Algorithms. *International Journal of Advanced Computer Science and Applications*. 2024. 15(3): Article 90. doi: <https://doi.org/10.14569/IJACSA.2024.0150390>
42. Wang G. Customer segmentation in the digital marketing using a Q-learning based differential evolution algorithm integrated with K-means clustering. *PLoS ONE*. 2025; 20(2): e0318519. doi: <https://doi.org/10.1371/journal.pone.0318519>
43. Kim K., Jo M., Ra I., Park S. RFMVDA: An Enhanced Deep Learning Approach for Customer Behavior Classification in E-Commerce Environments. *IEEE Access*. 2025. (Published January 2025) Article ID 10.1109/ACCESS.2025.3529023. doi: <https://doi.org/10.1109/ACCESS.2025.3529023>
44. Saldanha R. Marketing Campaign Dataset. Kaggle (електронний файл: CSV). 2018. 2 файли, загальний розмір приблизно 360 КБ. URL: <https://www.kaggle.com/datasets/rodsaldanha/arketing-campaign/data>

45. Коефіцієнт кореляції Пірсона: правила обчислення та інтерпретації. Studfile.net (освітній портал). 2018. URL: <https://studfile.net/preview/9290218/page:7/>
46. Error Sum of Squares (SSE): базова характеристика та приклади використання в кластерному аналізі. H-Lab, Stanford University. URL: [https://hlab.stanford.edu/brian/error\\_sum\\_of\\_squares.html](https://hlab.stanford.edu/brian/error_sum_of_squares.html)
47. Lee S. Expert Comparison: 6 Clustering Metrics Featuring Calinski-Harabasz Index. Number Analytics. 2025. URL: <https://www.numberanalytics.com/blog/expert-comparison-6-clustering-metrics-calinski-harabasz-index>

## Додаток А. Перелік змінних для моделі сегментації RFMP-DOV+AIC

Назва змінної	Опис	Одиниця виміру	Джерело в даних / Формула
Recency (R)	Кількість днів від останньої покупки	Дні	Recency
Frequency (F)	Загальна кількість покупок	Кількість	NumWebPurchases + NumCatalogPurchases + NumStorePurchases
Monetary (M)	Загальна сума витрат на всі категорії товарів	Грошова одиниця	Сума: MntWines, MntFruits, MntMeatProducts, MntFishProducts, MntSweetProducts, MntGoldProds
PromoScore (P)	Кількість прийнятих маркетингових кампаній	Кількість	AcceptedCmp1 + ... + AcceptedCmp5
DiscountPurchases (D)	Кількість покупок зі знижками	Кількість	NumDealsPurchases
OnlineRatio (O)	Частка онлайн-покупок	Частка (0–1)	NumWebPurchases / Frequency
Variety (V)	Кількість категорій, у яких були покупки	Кількість	Кількість полів Mnt... > 0
Age (A)	Вік клієнта	Роки	2014 - Year_Birth
Income (I)	Річний дохід	Грошова одиниця	Income
Children (C)	Кількість дітей у домогосподарстві	Кількість	Kidhome + Teenhome

*Джерело: розроблено автором*

### Додаток Б. Описова статистика

Variable	Count	Mean	Std. Dev	Min	25%	Median	75%	Max
Recency	2209.00	353.48	202.39	0.00	180.00	356.00	529.00	699.0
Frequency	2209.00	12.56	7.21	0.00	6.00	12.00	18.00	32.00
Monetary	2209.00	606.66	602.75	5.00	69.00	396.00	1047.00	2525.0 0
PromoScore	2209.00	0.74	1.45	0.00	0.00	0.00	1.00	6.00
DiscountPurchases	2209.00	2.32	1.92	0.00	1.00	2.00	3.00	15.00
OnlineRatio	2209.00	0.33	0.12	0.00	0.25	0.33	0.40	1.00
Variety	2209.00	5.43	0.92	2.00	5.00	6.00	6.00	6.00
Age	2209.00	34.09	11.70	7.00	26.00	33.00	44.00	63.00
Income	2209.00	52221.80	25193.00	1730. 00	35196. 00	51373.00	68487.00	66666 6.00
Children	2209.00	0.95	0.75	0.0	0.00	1.00	1.00	3.00

*Джерело: побудовано автором.*

**Додаток В. Середні значення показників за кожним кластером моделі RFMP-DOV+AIC**

<b>Показник</b>	<b>Cluster 0</b>	<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>	<b>Cluster 4</b>	<b>Cluster 5</b>	<b>Cluster 6</b>
<b>Recency</b>	319.31	356.92	363.36	324.98	353.83	355.08	435.92
<b>Frequency</b>	8.27	12.6	18.97	11.93	14.13	14.94	13.41
<b>Monetary</b>	236.46	443.85	1464.39	680.35	764.34	859.34	627.53
<b>PromoScore</b>	0.00	0.00	2.76	0.00	2.80	0.00	2.49
<b>DiscountPurchases</b>	2.39	2.88	1.00	0.17	2.62	1.27	3.42
<b>OnlineRatio</b>	0.34	0.35	0.28	0.24	0.35	0.27	0.39
<b>Variety</b>	4.36	6.00	5.99	4.78	4.32	6.00	6.00
<b>Age</b>	45.43	44.31	45.58	46.37	47.01	44.83	44.56
<b>Income</b>	42143.04	48840.31	76075.06	49119.02	55970.99	59876.61	48711.78
<b>Children</b>	1.43	1.27	0.00	0.60	0.13	0.00	1.24
<b>Count (осіб)</b>	522.00	735.00	234.00	60.00	196.00	290.00	172.00
<b>Percentage (%)</b>	23.63	33.27	10.59	2.72	8.87	13.13	7.79

Джерело: побудовано автором.

## ЗАВДАННЯ

### на кваліфікаційну роботу бакалавра

студентки 4 курсу спеціальності 051 «Економіка», ОПП «Економічна кібернетика»

Хижняк Анастасії Олександрівни

1. Тема роботи: Моделювання процесу сегментації клієнтів за допомогою RFM-аналізу для підтримки маркетингових рішень.
2. Термін завершення роботи: 23.05.2025.
3. Попередній захист роботи: 12.06.2025.
4. Об'єкт дослідження: процес сегментації клієнтів в електронній комерції.
5. Предмет дослідження: економіко-математичні моделі та інформаційні технології сегментації покупців на основі RFM-аналізу та його розширення з використанням додаткових поведінкових та демографічних змінних.
6. Мета дослідження: розробити вдосконалену модель сегментації клієнтів для підтримки маркетингових рішень на основі розширеного RFM-аналізу.
7. Завдання дослідження:
  - 7.1. Дослідити вплив електронної комерції на трансформацію маркетингових інструментів і підходів до роботи з клієнтами;
  - 7.2. Проаналізувати природу сегментації клієнтів як елемента маркетингового аналізу;
  - 7.3. Визначити переваги та ризики впровадження сегментації клієнтів з урахуванням цифрових технологій;
  - 7.4. Розкрити теоретичну суть RFM-аналізу як інструменту клієнтської аналітики в маркетингових дослідженнях;
  - 7.5. Систематизувати підходи до реалізації RFM-аналізу та проаналізувати специфіку їх застосування в практиці сегментації;
  - 7.6. Провести порівняльний аналіз методів RFM-аналізу;
  - 7.7. Розробити розширену модель RFMP-DOV+AIC та обґрунтувати вибір змінних;
  - 7.8. Підготувати дані та описовий аналіз вибірки;

- 7.9. Застосувати кластерний аналіз для формування сегментів;
- 7.10. Інтерпретувати отримані результати та надати управлінських рекомендацій;
- 7.11. Оцінити ефективність моделі порівняно з традиційними методами;
- 7.12. Запропонувати напрямки практичного впровадження моделі в маркетингову діяльність.

Науковий керівник: д.е.н, професор, Черноус Галина Олександрівна

Студент: Хижняк Анастасія Олександрівна

Затверджено на засіданні кафедри економічної кібернетики  
протокол № 6 від 27 листопада 2025 р.

### Календарний план виконання кваліфікаційної роботи бакалавра

№	Етапи роботи	Терміни виконання	Відмітка керівника про виконання
1	Вибір теми кваліфікаційної роботи бакалавра	До 1 листопада 2024 року	
2	Розробка та затвердження завдання кваліфікаційної роботи бакалавра	До 1 грудня 2024 року	
3	Збір інформації, її аналіз, обробка, консультації з науковим керівником	До 10 квітня 2025 року	
4	Підготовка роботи відповідно до вимог оформлення	До 20 квітня 2025 року	
5	Подача роботи для попереднього захисту	До 23 травня 2025 року	
6	Отримання відгуку наукового керівника	До 10 червня 2025 року	
7	Отримання рецензії на кваліфікаційну роботу бакалавра	До 10 червня 2025 року	
8	Перевірка роботи на плагіат	До 18 червня 2025 року	

**Науковий керівник:** д.е.н, професор

Чорноус Галина Олександрівна

**Студент:** Хижняк Анастасія Олександрівна