

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Моделі планування рекламних інвестицій на основі
інформаційних технологій»

на здобуття освітнього ступеня магістра
зі спеціальності 122 Комп'ютерні науки
(код, найменування спеціальності)
освітньо-професійної програми Комп'ютерні науки
(назва)

*Кваліфікаційна робота містить результати власних досліджень.
Використання ідей, результатів і текстів інших авторів мають посилання
на відповідне джерело*

(підпис)

Віктор ШЕРСТЮК
(Ім'я, ПРІЗВИЩЕ здобувача)

Виконав:
здобувач вищої освіти
група КНДМ-61

Віктор ШЕРСТЮК

Керівник:
науковий ступінь,
вчене звання

Андрій ЧИЧУР
доктор філософії

Рецензент:
науковий ступінь,
вчене звання

(Ім'я, ПРІЗВИЩЕ)

Київ 2024

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ**

Навчально-науковий інститут інформаційних технологій

Кафедра Комп'ютерних наук

Ступінь вищої освіти Магістр

Спеціальність 122 Комп'ютерні науки

Освітньо-професійна програма Комп'ютерні науки

ЗАТВЕРДЖУЮ

Завідувач кафедру Комп'ютерних наук

_____ Віктор ВИШНІВСЬКИЙ
«_____» _____ 2024 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

_____ Шерстюк Віктор Вячеславович

(прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи: «Моделі планування рекламних інвестицій на основі інформаційних технологій»

керівник кваліфікаційної роботи Андрій ЧИЧУР доктор філософії,

(Ім'я, ПРІЗВИЩЕ науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологій від «15» жовтня 2024р. №320

2. Строк подання кваліфікаційної роботи «15» грудня 2024р.

3. Вихідні дані до кваліфікаційної роботи: науково-технічна література, методи статистичного аналізу та економетричного моделювання, методи візуалізації даних.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

4.1. Дослідження економетричних моделей для планування рекламних інвестицій

4.2. Побудова та перевірка економетричної моделі

4.3. Аналіз ефективності каналів комунікації.

5. Перелік графічного матеріалу: *презентація*

5.1 Основні методи візуалізації даних за допомогою діаграм та таблиць.

5.2 Приклади програмного коду.

5.3 Результати аналізу економетричної моделі за допомогою Rstudio.

6. Дата видачі завдання «16» вересня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Аналіз наявної науково-технічної літератури	16.09-27.09.24	Виконано
2	Дослідження економетричних моделей та методів статистичного аналізу	30.09-11.10.24	Виконано
3	Постановка задачі та розробка концепції моделі	14.10-25.10.24	Виконано
4	Підготовка вхідних даних для побудови економетричної моделі	28.10-08.11.24	Виконано
5	Розробка економетричної моделі для оптимізації рекламних витрат	11.11-22.12.24	Виконано
6	Проведення аналізу ефективності каналів комунікації	25.11-29.11.24	Виконано
7	Оформлення роботи: вступ, висновки, реферат	02.12-06.12.24	Виконано
8	Розробка демонстраційних матеріалів	09.12-13.12.24	Виконано

Здобувач вищої освіти

(підпис)

Віктор ШЕРСТЮК

(Ім'я, ПРІЗВИЩЕ)

Керівник

кваліфікаційної роботи

(підпис)

Андрій ЧИЧУР

(Ім'я, ПРІЗВИЩЕ)

РЕФЕРАТ

Текстова частина дипломної роботи: 88 с., 3 таб., 33 рис., 1 додаток, 20 джерел.

Наукове завдання – розробка та аналіз моделей планування рекламних інвестицій із використанням інформаційних технологій для підвищення ефективності маркетингових стратегій у сучасному середовищі.

Мета роботи — розробка економетричної моделі, яка дозволяє оптимізувати витрати на рекламу та підвищити ефективність маркетингових кампаній бренду Кардіомагніл.

Об’єкт дослідження — процес планування інвестицій у рекламу фармацевтичного бренду Кардіомагніл.

Предмет дослідження — вплив рекламних інвестицій на обсяги продажів бренду Кардіомагніл у фармацевтичній галузі, враховуючи особливості каналів комунікації та конкурентне середовище.

Короткий зміст роботи: у роботі проведено аналіз ефективності рекламних кампаній у різних каналах комунікації, включаючи телебачення, радіо та цифрові платформи. Побудовано економетричну модель, яка оцінює внесок кожного з цих каналів у загальний приріст продажів. Виявлено, що телевізійна реклама забезпечує найбільший ефект, тоді як радіо і цифрові канали мають менший ефект.

Запропоновано рекомендації щодо розподілу рекламного бюджету з урахуванням показників ROI (рентабельності інвестицій). Модель дозволяє прогнозувати обсяги продажів залежно від рівня інвестицій у різні рекламні активності та впливу конкурентного середовища.

Отримані результати можуть бути використані для вдосконалення стратегій медіапланування, оптимізації рекламних витрат та підвищення ефективності просування фармацевтичних брендів.

Результати роботи мають практичну цінність для маркетингових департаментів фармацевтичних компаній, що працюють у висококонкурентному середовищі, а також для дослідників у галузі маркетингової аналітики.

Ключові слова: реклама, економетрична модель, фармацевтична галузь, інвестиції, канали комунікації, телевізійна реклама, цифровий маркетинг, аналіз даних, регресійні моделі.

ABSTRACT

The text part of the qualification work for obtaining the master's degree: 88 pages, 3 tables, 33 figures, 1 appendix, 20 sources.

The scientific task is to develop and analyse advertising investment planning models using information technologies to enhance the effectiveness of marketing strategies in the modern environment.

The purpose of the work is to develop an econometric model that optimizes advertising expenditures and improves the efficiency of marketing campaigns for the Cardimagnyl brand.

The object of research is the process of planning advertising investments for the CardiMagnyl pharmaceutical brand.

The subject of research is the impact of advertising investments on the sales volumes of the CardiMagnyl brand in the pharmaceutical industry, considering the peculiarities of communication channels and the competitive environment.

Summary of the work: The study analyzes the effectiveness of advertising campaigns across various communication channels, including television, radio, and digital platforms. An econometric model was built to evaluate the contribution of each channel to overall sales growth. It was found that television advertising provides the highest impact, while radio and digital channels exhibit lesser effects.

Recommendations for advertising budget allocation are proposed based on ROI (Return on Investment) indicators. The model allows forecasting sales volumes depending on the level of investment in various advertising activities and the influence of the competitive environment.

The findings can be utilized to enhance media planning strategies, optimize advertising expenditures, and increase the efficiency of promoting pharmaceutical brands.

The results hold practical value for marketing departments of pharmaceutical companies operating in highly competitive environments and for researchers in the field of marketing analytics.

Keywords: advertising, econometric model, pharmaceutical industry, investments, communication channels, television advertising, digital marketing, data analysis, regression models.

ЗМІСТ

ВСТУП.....	11
1 ЗНАЧУЩІСТЬ ЕКОНОМЕТРИЧНОГО МОДЕЛЮВАННЯ У МІДІАПЛАНУАННІ ДЛЯ ФАРМАЦЕВТИЧНИХ БРЕНДІВ.....	12
1.1 Традиційні підходи в медіаплануванні рекламних інвестицій.....	12
1.2 Принципи визначення бюджету рекламних кампаній.....	20
1.3 Основні підходи в економетричному моделюванні: лінійні та нелінійні регресійні моделі.....	24
1.4 Наукове завдання, мета та методологічна основа дослідження	37
1.5 Препарати на основі ацетилсаліцилової кислоти. Значення в сучасній фармакології.....	39
1.6 Використання RStudio для статистичного аналізу та моделювання.....	43
2 РОЗРОБКА МОДЕЛІ РЕКЛАМНИХ ІНВЕСТИЦІЙ БРЕНДУ КАРДІОМАГНІЛ.....	47
2.1 Формулювання гіпотез дослідження.....	47
2.2 Опис даних телевізійної рекламної активності.....	48
2.3 Опис даних діджитал-рекламної активності.....	52
2.4 Опис даних моніторингу радіореклами.....	55
2.5 Опис даних продажів для дослідження.....	55
2.6 Постановка вимог до моделі.....	57
3 РОЗРОБКА ЕКОНОМЕТРИЧНОЇ МОДЕЛІ ДЛЯ БРЕНДУ КАРДІОМАГНІЛ.....	59
3.1 Розробка моделі враховуючи всі наявні незалежні змінні.....	59
3.2 Тестування економетричної моделі на відповідність теоретичним принципам економетрики і подальша оптимізація.....	63
3.3 Результати моделювання рекламних інвестицій бренду Кардіомагніл.....	92
ВИСНОВКИ.....	96
ПЕРЕЛІК ПОСИЛАНЬ.....	99
ДОДАТОК А.....	102

ВСТУП

Сучасна конкуренція в бізнес-середовищі вимагає від компаній стратегічного підходу до розподілу ресурсів, зокрема рекламних інвестицій. Особливу актуальність це питання набуває для фармацевтичних брендів, які працюють у складних ринкових умовах, де ефективність медіапланування є ключовим фактором успіху. Постійне зростання обсягів даних та розвиток інформаційних технологій відкривають нові можливості для оптимізації маркетингових процесів, дозволяючи більш точно прогнозувати результати та досягати поставлених цілей.

Медіапланування визначається як процес вибору найефективніших каналів комунікації та часових інтервалів для розміщення рекламних повідомлень з урахуванням обмежених бюджетів та стратегічних маркетингових завдань. Вибір між телевізійною рекламою, цифровими платформами, друкованими виданнями чи іншими медіа є викликом для маркетологів, які прагнуть досягти балансу між охопленням, частотою та ефективністю витрат. У цьому контексті економетричні моделі надають можливість проводити кількісний аналіз впливу різних факторів на результати рекламних кампаній, що дозволяє приймати обґрунтовані рішення.

Особливо важливою є адаптація методів економетричного моделювання до потреб фармацевтичного ринку, де маркетингова ефективність залежить від багатьох змінних, включаючи демографічні особливості цільової аудиторії, регуляторні вимоги та конкурентне середовище. У роботі також розглядаються підходи до побудови та використання регресійних моделей для прогнозування продажів та визначення оптимального розподілу рекламного бюджету між різними каналами комунікації.

Таким чином, актуальність дослідження полягає у впровадженні сучасних підходів до планування рекламних інвестицій із застосуванням інформаційних технологій та економетричних моделей. Ці підходи дозволяють не лише підвищити ефективність маркетингових кампаній, але й забезпечити стійке зростання брендів у довгостроковій перспективі.

1 ЗНАЧУЩІСТЬ ЕКОНОМЕТРИЧНОГО МОДЕЛЮВАННЯ У МІДІАПЛАНУАННІ ДЛЯ ФАРМАЦЕВТИЧНИХ БРЕНДІВ

1.1 Традиційні підходи в медіаплануванні рекламних інвестицій

Медіапланування — це процес вибору найефективніших медіа каналів і часових проміжків для розміщення реклами з метою досягнення цільової аудиторії з урахуванням заданих бюджетів і маркетингових цілей. Основне завдання медіапланування полягає в тому, щоб визначити, які рекламні канали (телебачення, радіо, друковані видання, цифрові платформи та інші) і які часові проміжки забезпечать максимальну ефективність кампанії. Під час медіапланування фахівці прагнуть збалансувати охоплення, частоту та вартість для оптимізації результатів.

Історія та розвиток медіапланування

Медіапланування як напрям маркетингової діяльності зародилося ще на початку XX століття з появою масових комунікацій. Спочатку основними каналами для просування були газети та журнали, проте з появою радіо та телебачення медіапланування почало активно розвиватися. Телебачення, завдяки широкому охопленню та можливості показу рекламного контенту в прайм-тайм, стало основним рекламним каналом. До 80-х років більшість компаній витратила значні суми на телевізійну рекламу, вважаючи її найефективнішим способом охоплення масової аудиторії.

У 90-х роках, із появою інтернету, ситуація почала змінюватися. З кожним роком цифрові канали комунікації, такі як веб-сайти, пошукові системи, соціальні мережі та мобільні додатки, ставали дедалі важливішими. Це призвело до розвитку нових методів медіапланування, що враховують складну поведінку споживачів в онлайні, можливості таргетингу та персоналізації реклами. Сьогодні цифрові медіа складають значну частку рекламного бюджету більшості компаній, адже дозволяють більш точно досягати обраної аудиторії та вимірювати ефективність у реальному часі.

Ключові елементи медіапланування:

1. Цільова аудиторія. Визначення цільової аудиторії є одним з головних кроків у медіаплануванні. Аудиторію сегментують за віком, статтю, інтересами, географічним положенням та іншими характеристиками, щоб краще розуміти, які канали будуть для неї найбільш привабливими.

2. Охоплення та частота. Охоплення (кількість людей, які побачать рекламу хоча б один раз) і частота (кількість разів, коли цільова аудиторія контактує з рекламою) є основними показниками, що визначають ефективність кампанії. Медіапланувальники прагнуть знайти баланс між охопленням і частотою, щоб досягти максимального впливу.

3. Бюджет. Рекламний бюджет визначає, які медіаканали та формати можна використовувати. Часто бюджет розподіляють пропорційно між різними каналами залежно від їхньої вартості та очікуваної ефективності.

4. Часовий фактор. Вибір правильного часу для запуску рекламної кампанії має значення. Наприклад, для товарів сезонного характеру (новорічні подарунки, туристичні путівки) важливо врахувати сезонність. Для інших продуктів час може залежати від активності аудиторії на певних платформах у різний час доби або тижня.

5. Аналіз та коригування. Одним із ключових аспектів медіапланування є постійний моніторинг і аналіз результатів кампанії. Збір даних про ефективність реклами дозволяє вчасно вносити корективи та оптимізувати розподіл бюджету між каналами.

Основні стратегії медіапланування:

1. Орієнтація на максимальне охоплення. Ця стратегія застосовується для нових брендів або продуктів, які прагнуть швидко здобути впізнаваність. Компанії інвестують у канали, які дозволяють охопити якомога більше людей за короткий час.

2. Збільшення частоти контактів. Для брендів, що прагнуть підвищити лояльність аудиторії, медіапланувальники акцентують увагу на частоті. Це передбачає збільшення кількості контактів з однією і тією ж аудиторією, щоб закріпити повідомлення в свідомості споживачів.

3. Цільовий таргетинг. Цифрові медіа дозволяють використовувати індивідуальні параметри користувачів для точного таргетингу. Це підвищує ефективність реклами, оскільки вона потрапляє саме до тих, хто з найбільшою ймовірністю зацікавиться продуктом.

4. Баланс короткострокових і довгострокових кампаній. Поєднання короткострокових акцій з довгостроковими кампаніями дозволяє досягти як миттєвого зростання продажів, так і тривалого підвищення впізнаваності бренду.

Отже медіапланування постійно змінюється під впливом нових технологій та зміни поведінки споживачів. Завдяки появі цифрових каналів рекламодавці мають більше можливостей для таргетингу та персоналізації реклами. У той час як традиційні медіа, такі як телебачення та радіо, і надалі залишаються важливими інструментами, цифрові платформи дозволяють краще зрозуміти ефективність кожного вкладеного долара і вчасно коригувати кампанії.

Традиційні підходи до медіапланування в контексті визначення бюджету та розподілу між медіаканалами

Історично процес визначення рекламного бюджету та його розподілу серед медіаканалів був менш гнучким і переважно орієнтованим на інтуїтивні методи та правила, що вже показали себе в минулому. За відсутності аналітичних даних та інструментів точного таргетування, які доступні сьогодні, маркетологи використовували загальні моделі, що базувалися на частці ринку, планах продажів та споживчих тенденціях. Нижче наведено огляд найбільш поширених підходів до визначення медіабюджету та його розподілу, які використовувались у традиційних моделях медіапланування.

Відсоток від продажів

Метод «відсотка від продажів» є одним з найбільш поширених і простих у використанні традиційних підходів до визначення бюджету на рекламу. У цій моделі компанія встановлює рекламний бюджет як фіксований відсоток від обсягу продажів попереднього року або прогнозованого обсягу продажів на майбутній рік.

Переваги цього підходу полягають у простоті, оскільки компанії не потрібно здійснювати складний аналіз для визначення розміру рекламного бюджету. Також

цей метод дозволяє забезпечити певну стабільність, оскільки бюджет зростає або зменшується відповідно до обсягів продажів.

Проте недоліки цього підходу очевидні. Метод не враховує ринкових умов, таких як конкуренція або зміни у поведінці споживачів. Оскільки бюджет визначається на основі обсягу продажів, у періоди зниження продажів рекламні витрати також скорочуються, що може створити цикл зниження видимості бренду на ринку і подальшого зменшення продажів. Крім того, цей метод не дозволяє оптимізувати розподіл бюджету серед різних медіаканалів, оскільки він встановлює лише загальну суму витрат, не враховуючи відмінностей в ефективності каналів [20, с. 70].

Конкурентний паритет.

Метод «конкурентного паритету» базується на принципі встановлення рекламного бюджету таким чином, щоб він відповідав витратам конкурентів. Компанії, що використовують цей підхід, аналізують рекламну активність своїх прямих конкурентів та намагаються встановити бюджет, який дорівнює або перевищує витрати на рекламу конкурентів. Ідея полягає в тому, що компанія може утримувати свою частку ринку або навіть збільшити її, якщо виділятиме на рекламу такі ж чи більші суми.

Переваги цього підходу полягають у тому, що він дозволяє компаніям залишатися конкурентоспроможними на ринку. Завдяки використанню бюджету, аналогічного до конкурентів, компанія забезпечує свій «голос» у галузі, що важливо для збереження позицій на ринку.

Основні недоліки конкурентного паритету полягають у відсутності індивідуального підходу. Бюджет визначається на основі дій інших гравців, а не власних потреб чи цілей компанії, що робить цей підхід малоефективним у довгостроковій перспективі. Окрім того, метод ігнорує унікальні особливості бренду, як-от його сильні сторони, потреби аудиторії та особливості каналів, які можуть виявитися більш ефективними для конкретного випадку. В результаті компанія може витрачати бюджет на канали, які є менш ефективними, лише через те, що конкуренти використовують ці канали [20, с. 58].

Метод частки голосу (Share of Voice).

Підхід Share of Voice (частка голосу) став класичним методом у традиційному медіаплануванні. Цей метод передбачає, що компанія повинна встановити свій рекламний бюджет таким чином, щоб його частка у загальних рекламних витратах на ринку відповідала або перевищувала частку ринку бренду. Частка голосу означає частку рекламних повідомлень бренду серед усіх рекламних повідомлень у галузі.

Share of Voice є особливо ефективним для тих компаній, які прагнуть закріпити свою частку ринку або збільшити її, адже існує припущення, що вища частка голосу сприяє вищому рівню впізнаваності та продажів.

Переваги підходу Share of Voice включають його здатність підтримувати або розширювати ринкову частку бренду. Цей підхід дозволяє брендам зберігати високу видимість на ринку, особливо у конкурентних галузях, де важливо, щоб аудиторія постійно сприймала бренд як активного гравця.

Проте недоліками цього підходу є відсутність адаптивності до індивідуальних потреб компанії, а також високі витрати. Підтримка частки голосу на рівні ринкової частки може вимагати значних інвестицій, особливо у галузях з високою рекламною активністю. Крім того, Share of Voice не враховує різницю в ефективності каналів та специфічні цілі бренду. Існує ризик перевитрат на канали, які не приносять бажаних результатів, що може знизити рентабельність інвестицій [11, с. 32].

Метод завдань і цілей.

Метод завдань і цілей (Objectives and Task) є одним із найбільш стратегічно обґрунтованих підходів до планування бюджету. У цьому підході компанія спочатку визначає маркетингові та комунікаційні цілі, які вона прагне досягти, а потім оцінює вартість завдань, необхідних для досягнення цих цілей. Це може включати охоплення певної частки аудиторії, досягнення певної частоти контактів, покращення впізнаваності бренду або стимулювання продажів.

Переваги методу завдань і цілей полягають у його гнучкості та стратегічному підході. Цей метод дозволяє компаніям адаптувати свої рекламні бюджети до

конкретних потреб і завдань, що робить його більш точним і цілеспрямованим. Метод також дозволяє ефективно розподіляти ресурси серед різних каналів залежно від їхньої вартості та очікуваного впливу на досягнення цілей.

Проте недоліком методу завдань і цілей є його складність і вартість, оскільки він вимагає глибокого аналізу та прогнозування. Компанії, які використовують цей метод, повинні мати доступ до значних обсягів даних і аналітичних ресурсів, щоб точно визначити вартість кожного завдання. У минулому цей метод рідко використовувався через обмеження у доступі до необхідних даних і складність реалізації [20, с. 67–70].

Еволюція каналів комунікації: від домінування телебачення до зростання цифрових медіа

Упродовж багатьох десятиліть телебачення залишалося домінуючим каналом для медіапланування та комунікації брендів з аудиторією. Це було зумовлено його можливістю досягати великих мас аудиторії за допомогою візуально насичених, емоційно привабливих і запам'ятовуваних повідомлень. Реклама на телебаченні, особливо у прайм-тайм, забезпечувала не тільки широке охоплення, але й високу частоту контактів зі споживачами, що робило телебачення незамінним інструментом для збільшення обізнаності про бренд і стимулювання попиту. Такі кампанії мали значний вплив на масового споживача, оскільки до початку цифрової ери телебачення фактично було єдиним масовим джерелом відеоконтенту, що об'єднувало різні групи населення [11, с. 23].

Перехід від традиційних медіа до цифрових платформ.

З кінця 1990-х років, із масовим поширенням Інтернету, почали з'являтися нові рекламні канали, що поступово почали змінювати медіапростір. Веб-сайти, онлайн-банери, пошукова реклама та перші соціальні мережі відкрили нові можливості для комунікації, пропонуючи рекламодавцям як альтернативу традиційним каналам, так і можливості взаємодії зі споживачами. Однак на початку Інтернет не міг повноцінно конкурувати з телебаченням за охопленням і впливом, і його розглядали лише як додатковий інструмент для посилення брендової присутності. Телебачення залишалося основою великих рекламних кампаній, тоді

як Інтернет використовувався для підтримки кампанії або залучення більш молодшої аудиторії [6, с. 12].

Проте, із часом Інтернет не тільки розширився, але й змінився, поступово надаючи рекламодавцям більш точні інструменти таргетингу, персоналізації та аналітики. Зі зростанням популярності соціальних мереж, відео на YouTube, пошукової реклами, а також програматик-технологій, які дозволяли автоматизовано купувати рекламу для конкретної аудиторії, цифрові платформи стали повноцінною альтернативою телебаченню.

Цифрові медіа: переваги та зростаюча ефективність.

Цифрові медіа відрізняються від традиційних каналів комунікації низкою унікальних переваг, які сприяли їх швидкому зростанню та популярності серед брендів. Ось кілька ключових переваг цифрових медіа:

1. Точний таргетинг і персоналізація. Однією з головних переваг цифрових платформ є можливість точного таргетингу, що дозволяє звертатися до обраних груп споживачів залежно від демографічних, географічних, психографічних та поведінкових характеристик. Наприклад, соціальні мережі надають можливість рекламодавцям визначати аудиторію з урахуванням її інтересів і попередньої взаємодії з брендом. Це значно підвищує ймовірність того, що реклама буде доречною для споживача, підвищуючи рівень конверсії та знижуючи втрати на нецільові контакти [6, с. 46].

2. Можливість вимірювання результатів у реальному часі. Цифрові платформи надають рекламодавцям доступ до точних метрик ефективності, таких як перегляди, кліки, час перебування на сторінці, конверсії та інші показники. Це дозволяє брендам швидко отримувати дані про ефективність своїх кампаній та вносити корективи в режимі реального часу, що робить рекламу більш адаптивною та економічно вигідною [14, с. 47].

3. Інтерактивність і залученість аудиторії. Цифрові канали надають можливість залучати аудиторію до активної взаємодії з брендом. Наприклад, соціальні мережі дозволяють користувачам лайкати, коментувати, поширювати контент, що підвищує рівень залученості та створює міцніший зв'язок зі

споживачами. Це є значною перевагою над традиційним телебаченням, де комунікація є односторонньою [14, с. 47].

4. Мобільний доступ і глобальне охоплення. Зі зростанням використання мобільних пристроїв цифрові платформи отримали можливість взаємодіяти зі споживачами у будь-який час і будь-де. Це значно розширює охоплення реклами, дозволяючи брендам комунікувати з аудиторією на глобальному рівні. За допомогою інструментів геолокації рекламодавці можуть навіть налаштувати показ оголошень залежно від місцезнаходження користувача, що створює додаткові можливості для локальних акцій та пропозицій [6, с. 46].

Відхід від телебачення: чи завжди цифрові платформи кращі?

Хоча цифрові медіа мають значні переваги, важливо зазначити, що телебачення продовжує залишатися ефективним каналом для брендів, що прагнуть охопити широку аудиторію з мінімальними затратами часу. Телебачення, як і раніше, має здатність створювати високий рівень довіри до бренду, оскільки сприймається як більш надійне джерело інформації, особливо серед старшої аудиторії.

Дослідження показують, що великі бренди, які поєднують телебачення та цифрові платформи, досягають кращих результатів у плані охоплення та впливу, ніж ті, що використовують лише один канал. Поєднання каналів дозволяє брендам максимізувати переваги кожного з них, досягаючи як широкого охоплення (через телебачення), так і точного таргетингу та залучення (через цифрові медіа) [6, с. 46].

Використання мультимедійних стратегій.

Сьогодні багато брендів використовують мультимедійні стратегії, поєднуючи різні канали комунікації для досягнення максимальної ефективності. Наприклад, великі компанії можуть запускати телевізійні рекламні кампанії, щоб досягти широкого охоплення та закріпити ключове повідомлення, тоді як цифрові платформи використовуються для персоналізації та ретаргетингу, залучення молодшої аудиторії та глибшої взаємодії. Такий підхід дозволяє брендам оптимально використовувати рекламний бюджет, досягаючи як короткострокових, так і довгострокових цілей [14, с. 46].

Висновок.

Перехід від домінування телебачення до зростання цифрових медіа відображає зміну в поведінці споживачів та розвиток технологій. Цифрові медіа відкрили нові можливості для точного таргетингу, залучення аудиторії та вимірювання ефективності, що дозволяє брендам досягати кращих результатів з більшою рентабельністю. Проте, телебачення, як і раніше, залишається важливим інструментом у медіаплануванні, особливо для кампаній, які орієнтовані на широку аудиторію та створення високого рівня довіри до бренду. Оптимальним підходом для багатьох брендів є використання мультимедійних стратегій, що поєднують можливості традиційних і цифрових каналів.

1.2 Принципи визначення бюджету рекламних кампаній

Сучасні підходи до визначення рекламного бюджету базуються на принципах, що враховують ефективність різних каналів для досягнення цілей кампанії. Такий підхід дозволяє брендам використовувати бюджет стратегічно, адаптуючи його до ринкових умов і змін у поведінці споживачів. Серед основних підходів, що стали широко застосовуваними в медіаплануванні, вирізняють Excess Share of Voice (ESOV), баланс короткострокових і довгострокових кампаній, а також економетричні моделі, які дозволяють прогнозувати потреби в бюджеті для досягнення конкретних бізнес-цілей.

Принцип Excess Share of Voice (ESOV).

Одним із найбільш ефективних підходів до визначення рекламного бюджету є Excess Share of Voice (ESOV). Цей підхід допомагає визначити оптимальний бюджет на основі частки голосу бренду (Share of Voice або SOV) відносно його ринкової частки (Share of Market або SOM). Дослідження показали, що коли частка голосу перевищує частку ринку, тобто $ESOV > SOM$, бренд має більше шансів на зростання ринкової частки в довгостроковій перспективі. Це особливо важливо у висококонкурентних галузях, де ESOV допомагає брендам підтримувати високий рівень видимості та лояльності [11, с. 52].

Баланс між короткостроковими та довгостроковими кампаніями.

Ще одним важливим принципом є баланс між короткостроковими і довгостроковими кампаніями. Короткострокові тактики, як-от розпродажі, спрямовані на негайне зростання продажів, тоді як довгострокові кампанії спрямовані на формування іміджу та довіри до бренду. Як зазначає Sharp (2010), досягнення балансу між цими підходами дозволяє не тільки підвищити ефективність кампаній, а й забезпечити стійке зростання бренду в умовах конкуренції [19, с. 45]. В той самий час дослідження Леса Біне та Пітера Філда підкреслюють, що поєднання цих двох типів кампаній призводить до зростання ринкової частки [14, с. 46].

Економетричний підхід до визначення рекламного бюджету.

Економетричний підхід до визначення рекламного бюджету став новим і надзвичайно ефективним інструментом у медіаплануванні, особливо для великих брендів, що працюють у конкурентних і динамічних ринкових умовах. На відміну від традиційних підходів, економетричні моделі дозволяють розрахувати бюджет, необхідний для досягнення конкретних бізнес-цілей, таких як цільові показники продажів або збільшення частки ринку. Це досягається завдяки застосуванню статистичних методів, які враховують велику кількість факторів, що впливають на ефективність рекламних кампаній.

Ключові компоненти економетричних моделей включають врахування таких змінних, як:

1. Рекламні інвестиції. Модель враховує поточні витрати на рекламу і прогнозує, наскільки ці витрати повинні бути збільшені або зменшені для досягнення цільових показників продажів або ринкової частки.

2. Активність конкурентів. Оскільки дії конкурентів можуть безпосередньо впливати на результати рекламної кампанії, економетрична модель враховує інтенсивність їхньої рекламної активності. Це дозволяє бренду коригувати власний бюджет відповідно до конкурентного середовища, зберігаючи при цьому свою частку ринку.

3. Технічна penetрація. Для того, щоб оцінити, наскільки бренд може досягти певного рівня проникнення на ринок, економетричні моделі враховують дані про технічну доступність продукту на різних ринках, включаючи географічне охоплення та логістичні можливості. Це особливо важливо для глобальних кампаній, де ринки суттєво відрізняються за рівнем розвитку інфраструктури.

4. Ефективність каналів. Кожен медіаканал має різну ефективність залежно від цільової аудиторії та характеристик кампанії. Наприклад, цифрові медіа дозволяють точніше таргетувати рекламу, тоді як телебачення забезпечує високе охоплення. Економетрична модель дозволяє оцінити, який бюджет необхідно виділити для кожного каналу, щоб максимізувати ROI.

5. Зовнішні фактори. Сюди входять макроекономічні показники, сезонність попиту, зміни у поведінці споживачів, зміни регуляторних умов тощо. Включення таких факторів допомагає створити більш реалістичний прогноз і краще зрозуміти, як зовнішні обставини можуть вплинути на результати кампанії.

Переваги економетричних моделей.

Економетричний підхід забезпечує значні переваги порівняно з традиційними моделями бюджетування. Він дозволяє брендам більш точно прогнозувати, скільки коштів необхідно вкласти в рекламу, щоб досягти цілей по продажах або частці ринку. Це досягається завдяки побудові складних статистичних моделей, які враховують не тільки обсяг рекламних інвестицій, але й інші параметри, такі як активність конкурентів, рівень проникнення на ринок та зовнішні економічні фактори. В результаті, бюджет можна адаптувати не лише під поточні потреби, але й під зміну ситуації на ринку.

Оптимізація витрат. Економетричні моделі дозволяють оптимізувати витрати на рекламу шляхом точного визначення того, які канали є найефективнішими для певної аудиторії. Це забезпечує більш раціональний підхід до розподілу бюджету, що допомагає уникнути перевитрат і мінімізувати втрати. Як зазначено Zoltán Lakner та співавторами, використання економетричних моделей у фармацевтичній галузі дозволяє враховувати специфічні характеристики ринку та динаміку попиту.

Це сприяє розробці конкурентних стратегій, які підвищують ефективність маркетингових витрат і забезпечують сталий розвиток бренду [13, с. 45].

Гнучкість і адаптивність. Оскільки моделі враховують численні фактори, вони дозволяють швидко адаптувати бюджет до змін на ринку або в конкурентному середовищі. Наприклад, якщо конкурент запускає масштабну рекламну кампанію, бренд може збільшити інвестиції в рекламу на певний період, щоб зберегти свою частку голосу.

Прогнозування і вимірювання ефективності. Економетричні моделі дозволяють не тільки визначити бюджет, але й прогнозувати очікуваний ефект від рекламних інвестицій. Це допомагає брендам встановлювати реалістичні цілі та слідкувати за їхнім досягненням у процесі реалізації кампанії.

Використання економетричних моделей у практиці.

У сучасній практиці медіапланування економетричні моделі стали основою для визначення бюджету багатьох великих брендів. Наприклад, компанії в галузі FMCG (товарів масового споживання) використовують економетрію для визначення оптимального бюджету, який враховує як рівень конкуренції, так і сезонні коливання попиту. Це дозволяє їм підтримувати стабільний рівень продажів і зберігати лояльність споживачів навіть під час високої активності конкурентів.

Інший приклад — фармацевтичні компанії, які часто використовують економетричні моделі для визначення бюджету на просування продуктів у різних регіонах, беручи до уваги різний рівень доступності медичних послуг, місцеві регуляторні вимоги та активність конкурентів. Завдяки економетричним моделям фармацевтичні бренди можуть точніше визначати, скільки коштів необхідно вкласти в маркетинг у кожному регіоні, щоб забезпечити досягнення цільових показників продажів [20, с. 112].

Отже економетричний підхід до визначення рекламного бюджету дозволяє брендам більш точно адаптувати свої інвестиції до ринкових умов та реальних потреб аудиторії. На відміну від традиційних методів, економетрія забезпечує глибший аналіз і дозволяє врахувати численні фактори, які впливають на

результативність рекламної кампанії. Економетричні моделі дозволяють оптимізувати розподіл бюджету, забезпечуючи підвищення ефективності витрат і досягнення стратегічних цілей бренду. Зокрема, використання таких моделей є особливо важливим для великих компаній, що працюють на динамічних ринках, де конкуренція є високою, а споживчі переваги часто змінюються.

Таким чином, економетричні моделі допомагають брендам не тільки ефективно управляти рекламними витратами, але й створюють основу для довгострокового планування і підтримки конкурентоспроможності. За допомогою регулярного перегляду і корекції моделі можна забезпечити, щоб бюджетування залишалося актуальним і відповідало умовам ринку та цілям компанії.

1.3 Основні підходи в економетричному моделюванні: лінійні та нелінійні регресійні моделі

Економетричне моделювання широко використовується для кількісного аналізу взаємозв'язків між змінними та побудови прогнозів, які враховують динаміку різних факторів. Регресійні моделі є одним із ключових інструментів, що застосовуються в економетрії, для виявлення впливу незалежних змінних на залежні. Лінійні та нелінійні регресії дозволяють по-різному описувати ці залежності залежно від характеру даних, специфіки зв'язків між змінними та мети аналізу. Як зазначено Robin Lock та співавторами у роботі "Statistics: Unlocking the Power of Data", лінійні регресії забезпечують зручність аналізу та дозволяють оцінити припущення моделі, включаючи нормальність розподілу залишків і відсутність гетероскедастичності, що сприяє підвищенню точності прогнозів [7, с. 45]. Регресійні моделі широко застосовуються в дослідженнях ринків, фінансів, маркетингу та інших сфер, де важливо кількісно оцінювати взаємозв'язки та прогнозувати результати.

Лінійні регресійні моделі

Лінійна регресія є одним із базових та широко використовуваних підходів в економетричному моделюванні, оскільки вона надає можливість аналізувати

зв'язок між залежною змінною та однією або кількома незалежними змінними в рамках лінійного зв'язку. Лінійна модель виражається у формі рівняння:

$$Y = \alpha + \beta X + \epsilon, \quad (1.1)$$

де:

- Y — залежна змінна;
- α — константа, яка вказує на початкове значення Y за відсутності змінної X ;
- β — коефіцієнт, що відображає, як зміна X на одиницю впливає на зміну Y ;
- X — незалежна змінна;
- ϵ — залишкова помилка, що враховує інші невраховані фактори.

Лінійна регресія забезпечує простий і зрозумілий спосіб аналізу, і є корисною для побудови базових моделей, які мають лінійні взаємозв'язки між змінними. Такий підхід є зручним і ефективним, коли характер зв'язку простий, і не існує значних відхилень чи інших ускладнень. Наприклад, якщо лінійна регресія використовується для аналізу ринку FMCG, вона може показати, як збільшення витрат на маркетинг впливає на обсяг продажів, за умови, що цей зв'язок є лінійним і не змінюється під впливом інших факторів [12, с. 35; 17, с. 40]. Як зазначає Hansen (2014), лінійна регресія залишається основним інструментом економетричного аналізу завдяки її простоті та можливості розширення для більш складних моделей [18, с. 112].

Лінійна регресія передбачає дотримання ряду умов, таких як лінійність зв'язку, гомоскедастичність (однорідність дисперсії залишків) та незалежність спостережень. Якщо ці умови не дотримуються, це може призвести до спотворення результатів і зниження точності прогнозів. Для перевірки припущень моделі використовують такі діагностичні тести, як тест Дарбіна-Ватсона для автокореляції залишків і графічний аналіз залишків для перевірки на гомоскедастичність [2, с. 52; 15, с. 72; 17, с. 472].

Мультифакторна лінійна регресія.

Мультифакторна регресія є розширенням базової лінійної моделі, яка враховує вплив кількох незалежних змінних на залежну змінну одночасно. У моделі мультифакторної регресії рівняння має вигляд:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon, \quad (1.2)$$

де $X_1, X_2, \dots, X_n$ представляють незалежні змінні, які можуть одночасно впливати на залежну змінну.

Мультифакторна регресія дозволяє забезпечити точніший аналіз та прогнозування у випадках, коли на результат впливає декілька різних факторів. Наприклад, у сфері FMCG мультифакторна регресія може включати такі змінні, як рекламні витрати, сезонність, ціни на продукцію, активність конкурентів тощо. Це підвищує точність моделі та дозволяє зробити більш глибокі висновки щодо системної залежності змінних. Дослідження Naig та співавторів (2022) підкреслюють важливість мультифакторного підходу, особливо для галузей з динамічним середовищем, де взаємозв'язок між кількома факторами може мати значний вплив на кінцевий результат [3, с. 45; 17, с. 123].

Проблеми мультиколінеарності та автокореляції.

Мультиколінеарність виникає, коли незалежні змінні в моделі мають сильний кореляційний зв'язок одна з одною, що ускладнює розрізнення їхнього впливу на залежну змінну. Це може призвести до нестабільності та спотворення оцінок коефіцієнтів, знижуючи точність прогнозів. Наприклад, в аналізі ринку FMCG ціна товару та витрати на просування можуть мати взаємозалежність, що спричиняє мультиколінеарність у моделі. Для її усунення можна застосовувати методи, як-от факторний аналіз або метод головних компонентів, які зменшують кількість змінних і зберігають основну інформацію [17, с. 123].

Автокореляція, або взаємозв'язок залишків, може також впливати на точність моделі, особливо при аналізі часових рядів. Проблему автокореляції можна виявити за допомогою тесту Дарбіна-Ватсона, і для її корекції іноді застосовують

модифікацію лінійної регресії з використанням коригуючих змінних [2, с. 75; 14, с. 82; 17, с. 444-450].

Нелінійні регресійні моделі.

Нелінійні регресійні моделі застосовуються для аналізу складніших залежностей, які не можна описати лінійними співвідношеннями. Нелінійна регресія дозволяє враховувати складні взаємозв'язки, такі як ефекти насичення, циклічність або зменшення ефективності факторів після досягнення певного рівня. Наприклад, у сфері FMCG нелінійна регресія може описати, як збільшення витрат на рекламу спочатку значно підвищує обсяги продажів, однак після досягнення певного рівня приріст продажів стає меншим.

Квадратичні та логарифмічні моделі.

Однією з найпоширеніших форм нелінійних моделей є квадратичні функції, які включають змінні у квадраті для опису криволінійного зв'язку. Квадратичні моделі часто використовуються, коли необхідно відобразити залежність, яка має вигляд параболи, наприклад, коли зростання витрат на маркетинг спочатку збільшує продажі, але після досягнення певного рівня ефективність цих витрат починає знижуватись.

Логарифмічні моделі використовуються для опису ситуацій, коли зміна однієї змінної призводить до зменшення приросту іншої. Наприклад, у дослідженні поведінки споживачів на ринку FMCG логарифмічна модель може описувати ситуацію, коли збільшення доходів споживачів спочатку суттєво підвищує витрати на певні товари, але цей ефект поступово знижується.

Логарифмічні та квадратичні моделі дозволяють краще враховувати специфіку поведінкових залежностей, що є особливо корисним для дослідження ринків із насиченням або зниженням попиту на продукти. Wooldridge (2012) відзначає, що нелінійні моделі значно підвищують точність і надійність економетричного аналізу, особливо коли прості лінійні моделі не здатні повністю відобразити взаємозв'язок між змінними [17, с. 472; 18, с. 74].

Переваги та недоліки нелінійних моделей.

Нелінійні моделі дозволяють точніше описати реальну динаміку залежностей, особливо у випадках, коли ефективність залежить від досягнутого рівня, а подальші інвестиції мають обмежений ефект. Це є важливим для дослідження рекламних кампаній, де ефект від збільшення витрат сповільнюється після певної точки. Як зазначають Bhatt (2020) і Brooks (2008), нелінійні моделі краще описують реалії ринкових процесів, де спостерігається циклічність або сезонність [5, с. 111; 6, с. 46; 12, с. 69].

Проте, нелінійні моделі мають і свої обмеження, зокрема складність у налаштуванні та високі вимоги до якості даних. Вони можуть бути чутливими до вибору вихідних параметрів і вимагати великого обсягу даних для побудови точних прогнозів, що може бути ресурсно затратним. Враховуючи це, нелінійні моделі застосовуються у випадках, де лінійні підходи не забезпечують достатньої точності або не можуть описати складніші взаємозв'язки.

Отже, лінійні та нелінійні регресійні моделі є основою економетричного моделювання, оскільки вони дозволяють аналізувати та прогнозувати вплив різних факторів на змінні. Лінійні моделі є простішими і широко застосовуються для початкового аналізу, тоді як нелінійні забезпечують більшу гнучкість і точність у складніших ситуаціях. Разом ці моделі дозволяють отримати різнопланове уявлення про системи залежностей, що є важливим для аналізу ринку FMCG та багатьох інших сфер.

Моделі часових рядів: Методи ARIMA, SARIMA та їх застосування для сезонних коливань попиту.

Аналіз часових рядів є одним з основних інструментів у економетрії, що допомагає аналізувати й передбачати зміну змінних у часі. Часові ряди знаходять широке застосування у фінансах, економіці, маркетингу та багатьох інших галузях, де існує необхідність прогнозування певних показників на основі їхньої історичної динаміки. Особливу роль часові ряди відіграють у маркетинговому плануванні та в управлінні попитом, де показники, такі як продажі, дохід або витрати, піддаються періодичним сезонним коливанням. Моделі ARIMA (AutoRegressive Integrated Moving Average) та SARIMA (Seasonal ARIMA) є одними з найпоширеніших

методів для аналізу часових рядів, оскільки вони дозволяють моделювати складні залежності та включають як сезонні компоненти, так і довгострокові тренди.

Метод ARIMA

Метод ARIMA є одним із базових інструментів для аналізу часових рядів, який поєднує в собі три складові: авторегресію (AR), інтеграцію (I) та ковзне середнє (MA). Модель ARIMA загального вигляду описується як $ARIMA(p, d, q)$, де:

- p — порядок авторегресії (AR), тобто кількість попередніх значень ряду, які використовуються для передбачення наступного значення;
- d — порядок інтеграції, або кількість диференціювань, необхідних для перетворення ряду в стаціонарний;
- q — порядок ковзного середнього (MA), яке враховує кількість попередніх помилок прогнозу для корекції поточного значення.

Модель ARIMA зручна для застосування у випадках, коли часовий ряд не має яскраво виражених сезонних коливань, але при цьому демонструє тренд та випадкові флуктуації, які можна врахувати за допомогою ковзного середнього. Модель ARIMA має вигляд:

$$Y_t = \alpha + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t, \quad (1.3)$$

де:

- Y_t — значення часового ряду у момент часу t ,
- α — константа,
- β_i — коефіцієнти авторегресійної частини, що визначають вплив попередніх значень на поточне,
- θ_j — коефіцієнти ковзного середнього, які враховують попередні випадкові помилки,
- ϵ_t — випадкова помилка або шум, який не можна пояснити за допомогою інших компонентів моделі.

Процес побудови ARIMA-моделі починається з аналізу часового ряду для визначення необхідності перетворення його у стаціонарний ряд. Стаціонарність означає, що статистичні властивості ряду залишаються незмінними в часі. Якщо часовий ряд є нестаціонарним, його стаціонаризують за допомогою диференціювання, що враховується параметром (d) моделі ARIMA. Така процедура дозволяє зняти тренд і зробити ряд стабільнішим для подальшого моделювання.

Після визначення стаціонарності наступним етапом є вибір параметрів p , d та q . Для цього використовують автокореляційну функцію (ACF) та часткову автокореляційну функцію (PACF), які допомагають визначити оптимальну кількість попередніх значень та рівень впливу попередніх помилок. Після побудови моделі ARIMA результати перевіряють за допомогою діагностичних тестів для оцінки точності прогнозів. ARIMA моделі є дуже корисними для короткострокових прогнозів, де сезонні компоненти відсутні або можуть бути проігноровані [16, с. 104; 17, с. 840–845].

Метод SARIMA.

Метод SARIMA (Seasonal ARIMA) є розширенням базової моделі ARIMA, що включає сезонні компоненти, які враховують коливання, що повторюються через однакові інтервали часу. Це особливо корисно у випадках, коли часовий ряд демонструє стабільні сезонні цикли, наприклад, щомісячні чи щоквартальні коливання попиту на товари. Модель SARIMA позначається як $ARIMA(p, d, q)(P, D, Q)_s$, де:

- p, d, q — параметри базової (несезонної) частини моделі ARIMA,
- P, D, Q — параметри сезонної частини, що відповідають за сезонну авторегресію, сезонне диференціювання та сезонне ковзне середнє,
- s — довжина сезону, яка визначає період повторення коливань (наприклад, для щомісячних даних з річною сезонністю $s = 12$).

Формула SARIMA-моделі виглядає так:

$$Y_t = \alpha + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \sum_{k=1}^P \varphi_k Y_{t-ks} + \sum_{l=1}^Q \psi_l \epsilon_{t-ls} + \epsilon_t \quad (1.4)$$

де додаткові коефіцієнти φ_k і ψ_j враховують сезонні зміни.

Сезонна частина SARIMA дозволяє враховувати періодичність, яка є характерною для багатьох економічних і ринкових показників, наприклад, попиту на споживчі товари (FMCG) під час святкових періодів або знижок. Враховуючи сезонність, SARIMA забезпечує точніший прогноз, що є особливо корисним для середньо- та довгострокового планування в маркетингових та логістичних процесах [16, с. 108; 12, с. 104].

Етапи побудови та використання моделей ARIMA та SARIMA

Процес побудови та використання ARIMA і SARIMA моделей включає кілька основних етапів:

1. Перевірка на стаціонарність і перетворення ряду. Важливо забезпечити стаціонарність ряду, оскільки обидві моделі передбачають стаціонарність. Якщо ряд нестаціонарний, застосовується диференціювання.

2. Аналіз сезонності та вибір моделі. На цьому етапі визначаються сезонні коливання. Якщо присутня чітка сезонність, модель SARIMA є більш доречною.

3. Вибір параметрів. На основі ACF і PACF для ARIMA вибираються оптимальні значення p , d , q , а для SARIMA — значення P , D , Q , що дозволяє моделі врахувати як короткострокові, так і сезонні компоненти.

4. Оцінка точності та діагностика моделі. Після побудови моделі її результати перевіряють на точність за допомогою діагностичних тестів. До основних метрик належать середня абсолютна похибка (MAE), середньоквадратична похибка (MSE) і середня абсолютна відсоткова похибка (MAPE).

5. Прогнозування та інтерпретація. За допомогою обраної моделі здійснюється прогнозування, а результати використовуються для прийняття управлінських рішень.

Застосування моделей ARIMA та SARIMA для прогнозування сезонного попиту.

Моделі ARIMA та SARIMA широко застосовуються для прогнозування попиту на товари, які демонструють сезонні коливання, зокрема на ринку FMCG

(товарів повсякденного попиту). Наприклад, попит на певні види продуктів харчування, напоїв або товарів для дому може зростати під час святкових періодів чи літнього сезону. Використовуючи SARIMA, маркетологи та аналітики можуть точно прогнозувати обсяги продажів і налаштовувати маркетингові активності відповідно до очікуваного попиту.

Для компаній це означає, що вони можуть планувати виробництво та розподіл товарів з урахуванням очікуваних коливань, уникати перевантажень або нестач на складах. У роздрібній торгівлі це дозволяє знижувати ризики втрати прибутку через недостачу товарів або надлишок, що не встигне реалізуватися до кінця сезону. Завдяки точному прогнозуванню на основі SARIMA, компанії можуть гнучко керувати запасами, оптимізувати витрати на рекламу та ефективно розподіляти ресурси.

Крім того, SARIMA ефективно допомагає визначити оптимальні періоди для проведення маркетингових кампаній, таких як знижки чи акції. Враховуючи сезонні піки, можна збільшувати частоту рекламних контактів у періоди підвищеного попиту або знижувати рекламну активність у періоди спадів. Це дозволяє отримати максимальну віддачу від інвестицій у рекламу та підтримувати високий рівень обізнаності про бренд.

Отже, методи ARIMA та SARIMA є потужними інструментами для аналізу часових рядів, які широко використовуються у маркетинговому плануванні, економічному прогнозуванні, фінансах та інших галузях. ARIMA підходить для короткострокових прогнозів без сезонних компонентів, тоді як SARIMA дозволяє точно моделювати ряди з регулярними сезонними коливаннями. Завдяки цим моделям компанії можуть враховувати сезонні тренди, адаптувати свої стратегії під конкретні часові інтервали та оптимізувати витрати на маркетинг і логістику.

Ефективне застосування моделей ARIMA та SARIMA допомагає компаніям отримувати детальні прогнози щодо попиту, коригувати виробничі й маркетингові плани, що сприяє підвищенню ефективності бізнес-процесів і максимізації прибутків. Ці моделі забезпечують більш точне планування, адаптацію до мінливих

ринкових умов та конкурентних змін, що є важливою конкурентною перевагою в сучасному бізнес-середовищі.

Моделі панельних даних: Особливості панельних моделей у прогнозуванні в різних регіонах або сегментах ринку

Моделі панельних даних є важливим інструментом економетричного аналізу, який дозволяє одночасно враховувати часові та просторові аспекти даних. На відміну від звичайних часових або перехресних даних, панельні дані об'єднують обидві складові, що дозволяє аналізувати поведінку різних об'єктів (наприклад, країн, регіонів, підприємств) у часі. Завдяки цьому підходу можна точніше оцінити динаміку змін та виявити специфічні особливості поведінки об'єктів, що є цінним у маркетингових дослідженнях та бізнес-прогнозуванні. Панельні моделі широко застосовуються для аналізу змін у ринкових сегментах, прогнозування продажів у різних регіонах, оцінки ефективності маркетингових кампаній і планування логістичних та виробничих процесів.

Сутність і структура панельних даних

Панельні дані є набором спостережень за кількома об'єктами впродовж певного періоду часу. Типова панельна модель описується як сукупність об'єктів $i=1,2,\dots,N$ і моментів часу $t=1,2,\dots,T$. Кожне спостереження представляє значення змінної для конкретного об'єкта у визначений момент часу. Відповідно, якщо дослідження охоплює 10 регіонів протягом 5 років, то панельні дані матимуть $10 \times 5 = 50$ спостережень. Ця структура дозволяє одночасно враховувати внутрішні характеристики кожного об'єкта та їхню зміну з часом.

Модель панельних даних можна загалом представити так:

$$Y_{it} = \alpha + \beta X_{it} + u_{it} \quad (1.5)$$

де:

- Y_{it} — залежна змінна для об'єкта i у момент часу t ,
- α — константа,
- β — вектор коефіцієнтів,

- X_{it} — набір незалежних змінних для об'єкта i у час t ,
- u_{it} — залишковий компонент, який може бути поділений на індивідуальний ефект для кожного об'єкта i випадкову помилку.

Переваги панельних моделей у прогнозуванні

Панельні моделі мають низку переваг порівняно з моделями, що базуються лише на часових або перехресних даних. Однією з основних переваг є можливість врахування гетерогенності — тобто індивідуальних відмінностей між об'єктами. Це важливо для бізнес-досліджень, де різні регіони або сегменти ринку можуть демонструвати різні закономірності у поведінці споживачів, реакції на маркетингові активності та зміни в попиті. Наприклад, попит на певний товар може коливатися залежно від регіону через відмінності в доходах, культурних особливостях або конкурентному середовищі.

Іншою перевагою панельних моделей є здатність ефективніше працювати з динамічними даними, враховуючи одночасно і часовий аспект, і перехресну складову. Це дозволяє отримати більший обсяг інформації та підвищити точність оцінок. Панельні моделі також допомагають уникати проблеми мультиколінеарності, оскільки окремі індивідуальні характеристики можуть бути враховані через специфічні ефекти.

Панельні моделі забезпечують гнучкість при оцінці короткострокових і довгострокових змін у різних сегментах ринку. У довгостроковій перспективі такі моделі дозволяють відстежувати вплив на залежну змінну, враховуючи структурні зрушення, технологічні інновації, зміни в законодавстві або інші фактори. Наприклад, компанія може відстежувати вплив державних регуляцій на поведінку споживачів у кількох регіонах з часом, використовуючи індивідуальні фіксовані ефекти для врахування регіональних відмінностей. Це допомагає адаптувати стратегії компанії до ринкових умов.

Основні види панельних моделей

У прогнозуванні на основі панельних даних застосовуються два основних види моделей: модель з фіксованими ефектами (FE) та модель з випадковими ефектами (RE).

1. Модель з фіксованими ефектами (Fixed Effects, FE).

Модель з фіксованими ефектами враховує індивідуальні характеристики об'єктів, які не змінюються з часом. Такий підхід дозволяє контролювати внутрішні фактори, специфічні для кожного об'єкта, що можуть впливати на результат, але не залежать від інших змінних у моделі. Модель з фіксованими ефектами виглядає так:

$$Y_{it} = \alpha_i + \beta X_{it} + \epsilon_{it} \quad (1.6)$$

де α_i — індивідуальний ефект для кожного об'єкта i , який є фіксованим, тобто не змінюється в часі.

Моделі з фіксованими ефектами корисні для досліджень, де важливо врахувати унікальні характеристики кожного регіону чи сегмента ринку. Наприклад, якщо аналізується продаж товарів у різних регіонах, то фіксовані ефекти можуть врахувати сталі регіональні фактори, як-от культурні особливості чи рівень доходу, які впливають на попит на певні товари [2, с. 507; 13, с. 72].

Така модель є особливо корисною для аналізу стратегій проникнення на ринок і адаптації продукту до потреб різних ринкових сегментів. Наприклад, фіксовані ефекти дозволяють компаніям тестувати, як зміни у маркетинговій стратегії впливають на продажі у кожному регіоні з урахуванням постійних факторів. Це надає можливість розробляти більш точні та адаптивні стратегії для кожного ринку, орієнтуючись на реальні потреби споживачів та їхню поведінку.

2. Модель з випадковими ефектами (Random Effects, RE).

Модель з випадковими ефектами припускає, що індивідуальні характеристики об'єктів є випадковими та мають нормальний розподіл. Це дозволяє включити ефекти, які змінюються у кожному спостереженні, але при цьому зберігати загальний підхід до аналізу всіх об'єктів. Модель з випадковими ефектами має вигляд:

$$Y_{it} = \alpha + \beta X_{it} + u_i + \epsilon_{it} \quad (1.7)$$

де u_i — індивідуальний ефект, який має випадковий розподіл.

Моделі з випадковими ефектами зручні для прогнозування в умовах, коли індивідуальні характеристики об'єктів мають випадкову природу або не є головною дослідницькою метою. Випадкові ефекти використовують у випадках, коли кількість об'єктів у вибірці є великою і для кожного з них доступний обмежений обсяг даних. Наприклад, така модель підходить для аналізу поведінки клієнтів різних магазинів у межах країни, де немає потреби оцінювати унікальні характеристики кожного магазину, а більше важливо зрозуміти загальні закономірності [3, с. 102; 12, с. 340].

Використання панельних моделей для прогнозування в різних регіонах або сегментах ринку

Панельні моделі є цінними інструментами для прогнозування у різних регіонах або сегментах ринку, оскільки вони дозволяють аналізувати вплив певних змінних у динаміці та порівнювати ефективність маркетингових стратегій або змін попиту в окремих регіонах. Такий підхід є корисним у маркетингових дослідженнях, де компанії прагнуть розробити індивідуалізовані стратегії для різних ринків, беручи до уваги специфіку кожного з них.

Наприклад, у галузі роздрібної торгівлі панельні моделі можуть допомогти спрогнозувати вплив цінових змін або нових рекламних стратегій на попит у різних регіонах. Фіксовані ефекти можуть враховувати сталі регіональні особливості, такі як рівень життя чи демографічні характеристики, тоді як випадкові ефекти дозволяють виявити загальні тенденції, які спостерігаються в усіх регіонах.

У фармацевтичній галузі панельні моделі також використовуються для аналізу та прогнозування попиту на ліки в різних країнах або регіонах, де значний вплив мають такі фактори, як регуляторні норми, рівень доходів та специфіка системи охорони здоров'я. Наприклад, фармацевтична компанія може застосувати модель з фіксованими ефектами, щоб оцінити, як постійні фактори кожного ринку впливають на продажі. Водночас модель з випадковими ефектами може допомогти

врахувати випадкові коливання попиту, які є спільними для всіх ринків і пов'язані з глобальними трендами.

Панельні моделі можуть використовуватися і для аналізу поведінки споживачів на різних сегментах ринку, таких як преміум-сегмент, економ-сегмент або масовий ринок. Наприклад, у сфері FMCG можна використовувати панельні моделі для прогнозування ефективності рекламних кампаній у різних цінових сегментах, що дозволяє адаптувати рекламну стратегію відповідно до купівельної спроможності різних груп споживачів.

Отже, панельні моделі є потужними інструментами, які надають можливість одночасно аналізувати часові та перехресні залежності, що робить їх особливо цінними для досліджень у різних регіонах або сегментах ринку. Моделі з фіксованими та випадковими ефектами забезпечують гнучкість і точність у врахуванні індивідуальних характеристик кожного об'єкта дослідження, а також дозволяють виявляти загальні закономірності, що властиві для різних об'єктів. Це робить панельні моделі надзвичайно корисними у маркетинговому прогнозуванні, роздрібній торгівлі, фармацевтиці, споживчих дослідженнях та інших сферах, де важливе розуміння як специфіки окремих ринків, так і глобальних трендів.

1.4 Наукове завдання, мета та методологічна основа дослідження

Наукове завдання

Наукове завдання цього дослідження полягає у розробці економетричної моделі для підвищення ефективності інвестицій у рекламні бюджети бренду Кардіомагніл. Модель має допомогти оптимально розподіляти рекламні витрати, враховуючи продажі бренду та його конкурентів за попередні періоди, цінову політику, рівень проникнення на ринку, активність конкурентів і витрати на рекламу в різних каналах комунікації. Дослідження передбачає використання багатозмінної лінійної регресії, що забезпечить точність інтерпретації результатів і дозволить оцінити внесок кожного з цих факторів у формування оптимального рекламного бюджету.

Мета дослідження

Метою дослідження є розробка економетричної моделі, яка забезпечить оптимізацію рекламного бюджету бренду Кардіомагніл для досягнення цільових показників. Модель дозволить на основі даних про продажі бренду та його конкурентів, витрат на комунікаційні канали та інших маркетингових показників прогнозувати оптимальні обсяги інвестицій у рекламу. Вона також сприятиме точнішому розподілу бюджету між каналами з орієнтацією на максимальну віддачу від інвестицій (ROI). Це особливо важливо для бренду Кардіомагніл, оскільки така модель дозволить зосередити ресурси на найбільш ефективних напрямках, що сприятиме досягненню цільових показників за обмеженого бюджету.

Методологічна основа

Методологічна основа дослідження включає застосування багатозмінної лінійної регресії для побудови моделі, яка враховує основні фактори, що впливають на рекламний бюджет бренду Кардіомагніл. Вона оцінить вплив витрат на рекламу, цінової політики, рівня проникнення продукту, активності конкурентів та продажів конкурентів.

Багатозмінна лінійна регресія

Багатозмінна лінійна регресія є ключовим методом у побудові моделі, оскільки вона дозволяє оцінити вплив кількох незалежних змінних, зокрема витрат на рекламу, цінової політики, активності конкурентів і продажів конкурентів, на обсяг продажів бренду Кардіомагніл. Цей підхід дозволяє побудувати модель, у якій кожна змінна матиме свій коефіцієнт, що вказує на її вплив на результат. Це сприятиме точнішому розумінню того, які фактори найсильніше впливають на досягнення маркетингових цілей бренду.

Врахування продажів конкурентів

Продажі конкурентів є важливим показником, оскільки їхній обсяг відображає ситуацію на ринку, а також рівень інтенсивності конкурентної боротьби. У моделі вони виступають змінною, що дозволяє прогнозувати необхідний рівень витрат для утримання або зростання частки ринку, враховуючи дії інших брендів. Включення цього показника дає змогу адаптувати стратегії

Кардіомагнілу відповідно до активності конкурентів, що особливо важливо в умовах висококонкурентного ринку.

Кореляційний аналіз

Кореляційний аналіз допоможе виявити взаємозв'язок між витратами на рекламу, ринковими факторами та обсягом продажів Кардіомагнілу. Попереднє дослідження кореляцій між змінними дає змогу виділити основні фактори для моделі, а також підтвердити значущість тих чи інших змінних. Це дозволить налаштувати модель для більш точного прогнозування продажів і розподілу рекламного бюджету.

Врахування додаткових факторів: цінова політика, технічна проникливість та активність конкурентів

Врахування таких факторів, як цінова політика, технічна проникливість і активність конкурентів, є важливою складовою моделі. Цінова політика впливає на сприйняття продукту на ринку, оскільки ціна є важливим фактором у фармацевтичній галузі, де споживачі можуть обирати між оригінальними препаратами та дженериками. Технічна проникливість визначає доступність продукту в аптеках і на онлайн-ресурсах. Включення цих факторів до моделі дозволить бренду Кардіомагніл коригувати витрати на рекламу залежно від наявності товару та його цінової позиції відносно конкурентів.

Загалом, запропонована модель забезпечить інструмент для оптимізації витрат на рекламу, підвищення ефективності маркетингових інвестицій і адаптації стратегії бренду Кардіомагніл до ринкових умов.

1.5 Препарати на основі ацетилсаліцилової кислоти. Значення в сучасній фармакології

Ацетилсаліцилова кислота (ASA), відома в усьому світі під торговою назвою "аспірин", займає важливе місце серед найуживаніших лікарських засобів, відомих людству. Її відкриття в кінці XIX століття стало революційним кроком, який вплинув на розвиток фармакології як науки. Вперше синтезована Феліксом

Хоффманом у 1897 році в лабораторії компанії Bayer, ASA була створена як знеболювальний та жарознижувальний засіб, ефективність якого виявилася надзвичайно високою.

Протягом XX століття ацетилсаліцилова кислота застосовувалася не тільки для зняття болю, а й як протизапальний засіб при ревматоїдному артриті. Пізніше її унікальні антиагрегантні властивості були визнані науковцями, що відкрили новий напрямок у лікуванні серцево-судинних захворювань. У 1960-х роках з'явилися перші дослідження, які підтвердили здатність ASA знижувати ризик тромбозів, і з того часу вона стала важливою частиною терапії пацієнтів із серцево-судинними захворюваннями. На сучасному фармацевтичному ринку існує безліч препаратів, що містять ASA у різних формах і дозуваннях, зокрема Кардіомагніл, що поєднує ASA з гідроксидом магнію для зниження гастроінтестинальних побічних ефектів.

Сьогодні ASA використовується в багатьох країнах для профілактики серцево-судинних ускладнень, лікування болю і зниження температури. Популярність цього препарату зумовлена його доступністю, ефективністю та широким спектром дії, який робить його одним із найуніверсальніших препаратів на ринку.

Фармакологічні властивості ацетилсаліцилової кислоти.

Ацетилсаліцилова кислота відноситься до класу нестероїдних протизапальних препаратів (НПЗП) і діє шляхом пригнічення ферментів циклооксигенази (ЦОГ-1 та ЦОГ-2), що відіграють важливу роль у синтезі простагландинів — молекул, які беруть участь у регуляції запалення, болю та температури. Простагландини активуються у відповідь на ушкодження тканин, сприяючи больовим відчуттям та запальній реакції організму. ASA, пригнічуючи ЦОГ-1, знижує синтез простагландинів, що призводить до зменшення запалення та зниження температури. Саме цей механізм дії забезпечує знеболювальний та жарознижувальний ефект препарату [1, с. 15].

Антиагрегантна дія ASA пов'язана з її здатністю блокувати синтез тромбоксану, який сприяє агрегації тромбоцитів і утворенню тромбів. У низьких

дозах ASA використовується як профілактичний засіб для запобігання інфарктам та інсультам у пацієнтів із підвищеним ризиком серцево-судинних ускладнень. Така дія препарату зумовлена його незворотним впливом на тромбоцити, який триває до 10 днів (середній життєвий цикл тромбоцита). Це робить ASA одним із ключових препаратів у кардіології, особливо в профілактиці тромбозу та атеротромбозу [8, с. 65].

Основні продукти на основі ацетилсаліцилової кислоти.

На сучасному ринку доступний широкий вибір препаратів на основі ASA, які використовуються для різних медичних потреб і відрізняються формами випуску, дозуванням та складом. Серед них можна виділити:

- Кардіомагніл: один із найпоширеніших препаратів, що поєднує ASA з гідроксидом магнію для зниження ризику подразнення слизової оболонки шлунка. Завдяки цій комбінації Кардіомагніл часто призначають для профілактики серцево-судинних ускладнень у пацієнтів, які потребують тривалої терапії ASA.
- Аспірин: класичний препарат, що випускається в дозуваннях 100 мг та 500 мг, використовується для зниження температури та зняття болю. Сучасні форми включають таблетки з кишковорозчинною оболонкою для захисту шлунка.
- Інші продукти: існують також препарати ASA у формі швидко розчинних таблеток, капсул з модифікованим вивільненням та ін'єкційних форм, що розширюють можливості використання препарату в медичній практиці [9, с. 112].

Показання до застосування продуктів на основі ацетилсаліцилової кислоти

ASA має широкий спектр показань і застосовується при лікуванні болю різного генезу, зниженні температури, а також для профілактики тромбозів. Основні показання включають:

- Біль і лихоманка: ASA ефективна при головному болю, м'язових та суглобових болях, а також при лихоманкових станах, викликаних інфекційними захворюваннями.

- Запальні захворювання: ASA застосовується при ревматоїдному артриті, остеоартрозі та інших хронічних захворюваннях, що супроводжуються запаленням.
- Профілактика серцево-судинних захворювань: у низьких дозах ASA призначається для профілактики інфарктів, інсультів та інших тромбозів у пацієнтів з ішемічною хворобою серця або високим ризиком серцево-судинних ускладнень.

Рекомендоване дозування залежить від конкретних показань та стану здоров'я пацієнта. У профілактичній кардіології найчастіше призначають дози 75-100 мг на добу, тоді як для зняття болю дози можуть бути значно вищими [1, с. 25].

Побічні ефекти та протипоказання.

Незважаючи на численні переваги, ASA має ряд побічних ефектів, які обмежують її використання у певних пацієнтів. Основні побічні ефекти пов'язані зі шлунково-кишковим трактом. Препарат може спричинити диспепсію, гастрит і навіть виразки. У деяких випадках можливі шлунково-кишкові кровотечі, особливо у пацієнтів з підвищеною чутливістю до НПЗП або при тривалому застосуванні препарату.

Інші побічні ефекти включають бронхоспазми, особливо у пацієнтів з астмою або алергією на НПЗП. Через ризик кровотеч ASA не рекомендується для пацієнтів з виразковою хворобою шлунка, гіпертензією, а також у випадках, коли передбачається хірургічне втручання [8, с. 74].

Сучасні дослідження та тенденції у розробці препаратів на основі ацетилсаліцилової кислоти

Сьогодні ASA залишається об'єктом численних досліджень, спрямованих на розширення її терапевтичного потенціалу. Одним із таких напрямів є вивчення ASA як засобу профілактики колоректального раку. Деякі дослідження показують, що регулярне застосування ASA може знижувати ризик рецидиву пухлин, що пояснюється її здатністю зменшувати запалення в організмі.

Ще одним важливим напрямом досліджень є поєднання ASA з іншими активними речовинами для посилення терапевтичного ефекту та зниження

побічних дій. Наприклад, поєднання ASA з антиоксидантами або інгібіторами протонної помпи дозволяє значно знизити ризик шлунково-кишкових ускладнень, не втрачаючи ефективності препарату. Також нові препарати з ASA створюються у формах з уповільненим вивільненням, що сприяє більш рівномірному всмоктуванню активної речовини, забезпечуючи стабільний рівень препарату в крові.

Перспективним є дослідження потенційного впливу ASA на нейродегенеративні захворювання, такі як хвороба Альцгеймера та Паркінсона, завдяки її протизапальним властивостям. Дослідження показують, що хронічне запалення може бути однією з причин нейродегенерації, і тому ASA потенційно може сповільнити прогресування цих захворювань.

Висновок.

ASA залишається одним із найпопулярніших та найефективніших препаратів, що мають широкий спектр застосування в медицині. Продукти на її основі продовжують відігравати ключову роль у профілактиці серцево-судинних захворювань, лікуванні болю, запалення та лихоманки. Висока ефективність ASA поєднується з певними ризиками для шлунково-кишкового тракту та інших органів, однак сучасні розробки дозволяють значно знизити ці ризики.

Сьогоднішні інновації та нові напрямки досліджень свідчать про потенціал ASA у профілактиці не тільки серцево-судинних захворювань, а й онкологічних та нейродегенеративних хвороб. Подальші дослідження та розробка нових форм препаратів на основі ASA обіцяють забезпечити безпечніше і ще більш ефективно застосування цього препарату в медичній практиці.

1.6 Використання RStudio для статистичного аналізу та моделювання

RStudio — це інтегроване середовище розробки для роботи з мовою програмування R, яке надає потужний інструментарій для аналізу даних, побудови моделей і візуалізації. Його популярність серед дослідників, студентів і професіоналів пояснюється широким функціоналом, зручністю інтерфейсу та

здатністю автоматизувати складні аналітичні процеси. У цьому розділі ми детально розглянемо переваги RStudio, посилаючись на досвід роботи, описаний у *Statistics: Unlocking the Power of Data – RStudio User's Guide* (Robin Lock та ін.).

Інтуїтивний і зручний інтерфейс

Однією з найбільших переваг RStudio є його інтуїтивний інтерфейс, який забезпечує легкість роботи з мовою R навіть для новачків. Основні функції розділені між чотирма панелями:

- Консоль для виконання команд;
- Редактор коду, який дозволяє створювати, редагувати та запускати скрипти;
- Середовище об'єктів, яке показує поточні змінні та їхні значення;
- Візуалізація, де виводяться графіки та результати.

Ця структура сприяє швидкому доступу до необхідних інструментів і мінімізує кількість помилок, пов'язаних із виконанням команд вручну [4, с. 5]. Взаємодія між компонентами інтерфейсу дозволяє користувачам одночасно писати код, спостерігати за результатами та здійснювати налагодження.

2. Можливості інтерактивної візуалізації

RStudio пропонує унікальні можливості для побудови графіків і візуалізації даних. Завдяки бібліотеці *ggplot2* користувачі можуть створювати складні графіки, які підтримують інтерактивність та багатошаровість. Наприклад, діаграми розсіювання можуть бути доповнені регресійними лініями, а гістограми — маркерами критичних точок [4, с. 10].

Як зазначено у *RStudio User's Guide*, інструменти RStudio дозволяють створювати графіки, які можуть бути легко інтегровані у звіти чи презентації. Це робить його незамінним для аналізу великих наборів даних та комунікації результатів [4, с. 15].

Автоматизація процесів та репродуктивність

Використовуючи скрипти в RStudio, аналітики можуть легко автоматизувати повторювані процеси. Це особливо корисно під час роботи з великими наборами даних, коли необхідно виконати однакові розрахунки для кількох підгруп.

Репродуктивність — ключова перевага для наукових досліджень. Усі етапи аналізу можуть бути збережені у скрипті, що дозволяє іншим дослідникам перевірити результати. Як зазначено у *RStudio User's Guide*, використання Markdown значно полегшує створення інтерактивних звітів, які можуть містити текст, код та результати аналізу в одному документі [4, с. 20].

Розширені можливості для роботи з великими даними

RStudio підтримує роботу з великими наборами даних завдяки використанню бібліотек, таких як `data.table` та `dplyr`. Ці інструменти оптимізовані для швидкої обробки даних і забезпечують високу продуктивність навіть при роботі з мільйонами записів [4, с. 25].

Крім того, середовище RStudio інтегрується з такими системами, як Hadoop і Spark, що дозволяє використовувати його для аналізу великих даних у хмарних середовищах [4, с. 30].

Підтримка машинного навчання

Окрім стандартного статистичного аналізу, RStudio пропонує потужний інструментарій для машинного навчання. Завдяки бібліотекам `caret`, `randomForest` і `xgboost` користувачі можуть реалізовувати алгоритми класифікації, регресії та кластеризації. Ці бібліотеки мають інтуїтивно зрозумілий інтерфейс, що дозволяє дослідникам швидко налаштовувати моделі та оцінювати їхню точність [4, с. 35].

Підтримка спільної роботи

RStudio дозволяє інтегрувати Git та інші системи контролю версій, що робить його ідеальним для спільної роботи над проектами. Як зазначено у *RStudio User's Guide*, можливість відслідковувати зміни у скриптах і документах забезпечує прозорість процесу розробки моделей [4, с. 40].

Висновки

RStudio — це інструмент, який об'єднує зручність, продуктивність та інноваційність. Завдяки його можливостям, дослідники можуть автоматизувати аналіз, створювати якісні візуалізації та забезпечувати прозорість результатів своїх

досліджень. Це робить його невід'ємною частиною сучасного статистичного аналізу.

2 РОЗРОБКА МОДЕЛІ РЕКЛАМНИХ ІНВЕСТИЦІЙ БРЕНДУ КАРДІОМАГНІЛ

2.1 Формулювання гіпотез дослідження

Успішне проведення дослідження передбачає наявність гіпотез, які модель має підтвердити або спростувати. Формулювання гіпотез дозволяє сфокусувати аналіз на ключових аспектах досліджуваного явища та визначити основні взаємозв'язки, які потребують оцінки. У цьому контексті, дослідження передбачає наступні гіпотези щодо впливу рекламної активності, конкурентних факторів та загальної динаміки ринку на продажі бренду Кардіомагніл.

1. Гіпотеза про вплив рекламної активності:

- Основне припущення полягає в тому, що існує позитивний і значущий зв'язок між обсягами рекламної активності на телебаченні, в діджитал-каналах, на радіо і інших медіаканалах і обсягами продажів бренду Кардіомагніл. Відповідно, очікується, що підвищення рівня рекламної активності корелюватиме зі збільшенням продажів.

- Альтернативна гіпотеза стверджує, що рекламна активність не має істотного впливу на продажі бренду, що може пояснюватися особливостями ринку, особливостями поведінки споживачів або високою конкуренцією.

2. Гіпотеза про вплив цінової політики та доступності продукції:

- Гіпотеза передбачає, що нижча цінова позиція бренду Кардіомагніл та широка дистрибуція позитивно впливають на обсяги продажів порівняно з конкурентами. Таким чином, ці фактори повинні стимулювати продажі бренду у порівнянні з аналогічними продуктами на ринку.

- Альтернативне припущення полягає в тому, що ціна і доступність мають обмежений вплив на обсяги продажів у категорії, тоді як більш вагомий вплив справляють рекламна активність і загальні ринкові фактори.

3. Гіпотеза про вплив загальної динаміки ринку:

- Очікується, що продажі бренду Кардіомагніл корелюють із загальними обсягами продажів у категорії. Це означає, що зміни на ринку в цілому (зростання або зниження попиту) мають прямий вплив на динаміку продажів бренду.
- Альтернативна гіпотеза передбачає, що загальна ринкова динаміка не є визначальним фактором для продажів бренду Кардіомагніл, оскільки його обсяги продажів залежать переважно від специфічних факторів, таких як рекламна активність, дистрибуція та цінова стратегія.

Важливість формулювання гіпотез

Формулювання гіпотез дозволяє сфокусувати дослідження на ключових питаннях і деталізувати, які саме фактори повинні бути перевірені за допомогою моделювання. Це, у свою чергу, забезпечує більш структурований підхід до аналізу результатів, дозволяє проводити цілеспрямовану оцінку основних змінних та робить висновки більш обґрунтованими.

2.2 Опис даних телевізійної рекламної активності

У дослідженні аналізується вплив телевізійної рекламної активності на продажі бренду Кардіомагніл. Для цього використовуються дані, зібрані компанією Nielsen Ukraine – одним із провідних постачальників даних про медіа-аудиторію та рекламну активність. Компанія Nielsen Ukraine надає детальні та репрезентативні дані про телевізійні рейтинги, що дозволяє здійснювати комплексний аналіз ефективності рекламних кампаній та їхнього впливу на аудиторію. За інформацією, доступною на офіційному сайті компанії (<https://www.nielsen.com/about-us/locations/ukraine/>), Nielsen Ukraine використовує сучасні методи та технології для збору даних, які охоплюють поведінку глядачів та забезпечують точність і надійність вимірювань.

Основні характеристики даних телевізійної рекламної активності.
Періодичність та актуальність даних

Дані телевізійної рекламної активності, надані Nielsen Ukraine, збираються регулярно з високою періодичністю, що дозволяє аналізувати рекламну активність

як у короткостроковій, так і в довгостроковій перспективі. У межах дослідження використовуються дані за період з 2022 по вересень 2024 року, що дозволяє врахувати різні рекламні стратегії, які змінювались протягом цього часу, та адаптацію бренду до мінливих ринкових умов.

Дані 2022 року зібрані безпосередньо від телевізійних каналів і представлені у вигляді кількості виходів (тобто кількості разів, коли рекламний ролик бренду транслювався на екранах). Оскільки у 2022 році компанія Nielsen Ukraine не проводила вимірювань, які б включали рейтинги телевізійних програм, дані за цей рік мають відмінний формат порівняно з даними 2023 і 2024 років, де використовуються рейтинги.

Дані за 2023 і 2024 роки зібрані за допомогою піплметричних вимірювань, що дозволяють об'єктивно фіксувати поведінку аудиторії на рівні окремих домогосподарств. Ця інформація забезпечує точність у вимірюванні перегляду рекламного контенту, адже рейтинги показують відсоток глядачів, які переглянули конкретний рекламний ролик під час його трансляції. Використання цих рейтингів дозволяє оцінити ефективність рекламних вкладень, оскільки чим вищий рейтинг, тим більше глядачів було охоплено.

Типи показників телевізійної рекламної активності.

Основні показники телевізійної рекламної активності, які використовуються у моделі, включають:

Рейтинги рекламних кампаній – це ключова метрика, що відображає відсоток аудиторії, яка переглянула конкретний рекламний ролик у певний час. Рейтинг є надійним індикатором охоплення рекламною кампанією, оскільки дозволяє оцінити реальну кількість глядачів, що взаємодіяли з контентом. Зокрема, дані рейтингу надаються для кожного рекламного ролика з розбивкою за часовими інтервалами.

Кількість виходів – використовується як показник для 2022 року, коли рейтингові дані були недоступні. Кількість виходів дозволяє оцінити загальну інтенсивність рекламної активності, хоча цей показник не відображає реальної кількості глядачів, які побачили рекламу.

Комбінація цих показників забезпечує комплексний підхід до оцінки телевізійної активності, зважаючи на обмеження даних за різні роки.

Методологія збору даних.

Компанія Nielsen Ukraine використовує технологію піплметрів для вимірювання аудиторії телевізійних каналів. Піплметри є електронними пристроями, які встановлюються в домогосподарствах, що погодилися на участь у дослідженні. Ці пристрої фіксують інформацію про кожне ввімкнення телевізора, включаючи час перегляду та канал. Такий підхід забезпечує високу точність у вимірюванні поведінки глядачів, оскільки дані збираються автоматично без потреби в активній участі користувача.

Дані, зібрані за допомогою піплметрів, відображають реальну поведінку глядачів і дозволяють оцінювати перегляд телевізійних каналів та окремих рекламних роликів. Зібрана інформація є репрезентативною для всього населення, оскільки вибірка складається з домогосподарств, що представляють різні соціально-демографічні групи населення України. Це гарантує, що дані є достатньо точними для аналізу, а також для прийняття рішень щодо оптимізації рекламної активності.

Нормалізація даних для 2022 року

Оскільки у 2022 році дані про телевізійну активність були доступні лише у вигляді кількості виходів, а не рейтингів, було необхідно провести нормалізацію для забезпечення їх порівнянності з даними за 2023 і 2024 роки. Для цього було використано розраховані коефіцієнти на основі даних 2023 року. Ми проаналізували середнє співвідношення між кількістю виходів та рейтингами для кожного рекламного блоку, доступного за 2023 рік. Отримані коефіцієнти дозволили скоригувати дані за 2022 рік і наблизити їх до формату, який більш точно відображає охоплення та інтенсивність рекламних кампаній. Цей метод нормалізації забезпечує узгодженість даних між роками та дозволяє

використовувати показники телевізійної активності 2022 року нарівні з даними наступних періодів у моделі, підвищуючи точність і надійність результатів аналізу.

Охоплення та деталізація даних.

Дані Nielsen Ukraine містять інформацію про демографічні характеристики аудиторії, що дозволяє аналізувати телевізійну активність з погляду охоплення різних цільових груп. Це особливо важливо для брендів, які намагаються залучити певні сегменти ринку, оскільки дозволяє зрозуміти, які саме групи глядачів переглядали рекламу.

Наприклад, дані можуть включати інформацію про глядачів, розподілених за такими категоріями, як:

1. Вікові групи (молодь, дорослі, старше покоління);
2. Стать (чоловіки, жінки);
3. Географічне розташування (регіони України).

Така деталізація дозволяє точно аналізувати ефективність рекламних кампаній та визначати, наскільки успішно бренд досягає своїх цільових аудиторій. Наприклад, можна оцінити, як реклама бренду Кардіомагніл сприймається серед різних вікових груп і чи є потреба в адаптації рекламних матеріалів для охоплення більшої частини аудиторії.

Використання даних телевізійної активності у моделі.

Дані телевізійної рекламної активності є ключовим елементом для моделювання впливу рекламних кампаній на продажі бренду Кардіомагніл. Оскільки рейтинги за 2023 і 2024 роки є точним індикатором переглядів, вони дозволяють моделі визначати, як збільшення охоплення рекламою може впливати на попит і обсяг продажів. Водночас, для 2022 року, де використовуються лише дані про кількість виходів, модель має враховувати обмеження цього показника, оскільки він не дає прямого уявлення про охоплення. Для компенсації цього обмеження застосовано проксі-методи, засновані на середньому співвідношенні

між рейтингами та кількістю виходів, визначеному на основі даних 2023 і 2024 років.

Оцінка переваг використання даних Nielsen Ukraine

Дані Nielsen Ukraine є важливим інструментом для розуміння того, як телевізійна реклама впливає на поведінку глядачів і, як наслідок, на продажі бренду. Доступ до рейтингів та кількості виходів дозволяє не лише визначати рівень рекламної активності, але й проводити порівняння різних рекламних кампаній, оцінюючи їхню ефективність у різних часових періодах. Інтеграція таких показників у модель продажів забезпечує комплексне розуміння впливу телевізійної реклами на ринок та дозволяє робити обґрунтовані висновки щодо оптимізації рекламних вкладень бренду Кардіомагніл.

2.3 Опис даних діджитал-рекламної активності

У дослідженні використовуються дані діджитал-рекламної активності, зібрані компанією Gemius AdReal – провідним постачальником аналітичних даних про онлайн-рекламу, які застосовуються для оцінки ефективності діджитал-кампаній. Компанія Gemius AdReal відома своїм панельним підходом до збору даних та високоточними методами вимірювання рекламної активності в інтернеті. На офіційному сайті компанії (<https://adreal.gemius.com/site/en/>) зазначено, що вона має багаторічний досвід збору та аналізу даних про діджитал-рекламу, зокрема в Європі та країнах Східної Європи, включаючи Україну.

Основні характеристики даних діджитал-рекламної активності:

1. Періодичність та актуальність даних

Дані діджитал-рекламної активності, зібрані Gemius AdReal, характеризуються високою частотою оновлення та актуальністю, що дозволяє аналізувати зміни в рекламній активності на щоденній основі. Така частота збору особливо корисна для аналізу як короткострокових коливань рекламної активності,

так і довгострокових трендів. У рамках даного дослідження використовуються дані за період 2022–2024 років, що дозволяє охопити останні зміни на ринку та їхній вплив на поведінку аудиторії.

2. Основний показник – кількість показів

Для побудови моделі в дослідженні використовується показник кількості показів (impressions) – ключова метрика, що відображає інтенсивність діджитал-рекламної активності бренду. Кількість показів показує, скільки разів рекламний контент був відображений користувачам на різних платформах, таких як новинні портали, відеосервіси та соціальні мережі. Використання показів як основного показника в моделі дозволяє об’єктивно оцінити рекламну активність бренду та її можливий вплив на продажі.

Застосування кількості показів як основного показника має кілька переваг:

- Пряма оцінка інтенсивності реклами: покази точно відображають частоту появи рекламного контенту на екранах користувачів, що важливо для оцінки потенційного впливу реклами на аудиторію.
- Порівнянність з іншими видами реклами: покази є універсальним показником, що дозволяє порівнювати активність в діджитал-каналах із активністю на телебаченні, де також використовується кількість виходів як один із основних показників.

3. Методологія збору даних

Дані про кількість показів реклами збираються Gemius AdReal за допомогою панельного підходу. Компанія формує вибірку користувачів, які добровільно беруть участь у панелі, що дозволяє отримати дані про перегляди реклами на основі анонімного моніторингу їхньої поведінки в інтернеті.

Панельний підхід є одним із найбільш надійних методів для збору інформації про діджитал-рекламу, оскільки він забезпечує безпосередній доступ до даних про перегляди та відображає реальну поведінку користувачів. У випадку Gemius AdReal використовується моніторинг рекламних показів на ПК-трафіку, що дозволяє оцінити рекламну активність у межах використання десктопних пристроїв. Водночас цей підхід має певні обмеження: оскільки дані збираються

лише для ПК, мобільний трафік не враховується, і це може призводити до недооцінки загального обсягу охоплення реклами для аудиторії, що використовує мобільні пристрої.

4. Деталізація даних за платформами та демографічними характеристиками

Дані від Gemius AdReal охоплюють різноманітні платформи, на яких розміщується діджитал-реклама, включаючи новинні портали, відеосервіси, соціальні мережі та інші веб-ресурси. Така деталізація дозволяє дослідити, які канали є найефективнішими для взаємодії з аудиторією, і оцінити вплив кожного каналу на загальну інтенсивність рекламної активності. Зокрема, дані показують, яка частка рекламної активності бренду була спрямована на кожен із платформ, що є важливим для визначення найбільш продуктивних каналів просування.

Окрім цього, дані містять інформацію про демографічні показники аудиторії (вік, стать), яка переглядала рекламу, що дозволяє робити висновки про те, наскільки кампанії досягають цільової аудиторії бренду. Хоча в моделі ці показники безпосередньо не використовуються, наявність такої інформації є корисною для глибшого розуміння загальної ефективності рекламної активності бренду.

Використання показів у моделі

У моделі використовується показник кількості показів (impressions) як основний параметр, що відображає рівень діджитал-рекламної активності. Покази обрані як основна метрика, оскільки вони безпосередньо корелюють з кількістю контактів аудиторії з рекламним контентом, що є важливим для прогнозування впливу реклами на обсяги продажів. У моделюванні кількість показів дає змогу побачити, як змінюється обсяг продажів під впливом інтенсивності діджитал-рекламної активності, дозволяючи виявити взаємозв'язок між вкладеннями у рекламу і кінцевими результатами на ринку.

Оскільки кількість показів є універсальним показником, який легко інтегрується з іншими змінними, такими як продажі та активність конкурентів, вона надає моделі необхідну гнучкість для проведення порівняльного аналізу.

Використання кількості показів без коригування на частоту або охоплення зберігає простоту моделі, одночасно надаючи інформативні результати щодо ефективності рекламних зусиль.

Таким чином, покази є основною змінною в моделі, що дозволяє точно оцінити інтенсивність діджитал-рекламної активності, забезпечуючи адекватне підґрунтя для аналізу впливу діджитал-кампаній на ринок та досягнення цільових показників бренду.

2.4 Опис даних моніторингу радіореклами

Дані моніторингу радіореклами, що використовуються в дослідженні, надані компанією **«Комунікаційний Альянс»** — провідного постачальника послуг з медіамоніторингу та аналізу. Інформацію отримано через платформу MediaMonitor (<https://clip.mediamonitor.com.ua/>), яка забезпечує детальне відстеження рекламних активностей у радіоефірі.

Зібрані дані включають ключові характеристики рекламних кампаній:

1. Часовий слот трансляції — визначає, у який період доби реклама виходила в ефір (ранковий, денний, вечірній прайм-тайм).
2. Частота ротацій — кількість разів, коли рекламне повідомлення транслювалося протягом заданого періоду.
3. Станція трансляції — назва радіостанції, на якій здійснювалася трансляція, з урахуванням її цільової аудиторії.

Завдяки точності та деталізації моніторингу від компанії «Комунікаційний Альянс» дані дозволяють оцінити ефективність радіореклами, зокрема її здатність впливати на обізнаність аудиторії та стимулювати продажі.

2.5 Опис даних продажів для дослідження

У дослідженні використовуються дані про обсяги продажів брендів у категорії, зібрані аналітичною компанією "Моріон". Ця компанія спеціалізується

на дослідженнях та аналітиці ринку фармацевтики в Україні. За інформацією, наведеною на офіційному сайті компанії "Моріон" (<https://www.morion.ua/>), вона надає комплексні дані про продажі фармацевтичних продуктів, що дозволяє отримати точну та актуальну інформацію для аналізу ринкових тенденцій і ефективності брендів.

Структура даних, які надає "Моріон":

1. Періодичність даних:
 - Дані збираються з високою періодичністю, що дозволяє враховувати сезонні та інші коливання ринку. Це надає можливість аналізувати короткострокові й довгострокові тенденції продажів.
2. Охоплення ринку:
 - "Моріон" забезпечує дані про обсяги продажів як бренду Кардіомагніл, так і його основних конкурентів. Це дозволяє оцінити ринкову частку бренду та відслідковувати конкурентні позиції.
3. Методологія збору даних:
 - Дані збираються на основі інформації від ключових аптечних мереж, дистриб'юторів і виробників, що забезпечує точне охоплення фармацевтичного ринку України.
4. Типи показників:
 - До складу даних входять показники обсягів продажів у грошовому та натуральному вираженні, що дозволяє провести глибокий аналіз цінових змін і динаміки попиту. Дані також можуть містити інформацію про дистрибуцію та цінові стратегії брендів.
5. Можливість порівняння з конкурентами:
 - Оскільки "Моріон" надає доступ до інформації про продажі аналогічних препаратів конкурентів, ці дані є цінними для оцінки ринкових часток та розуміння факторів, які впливають на позиціювання Кардіомагніл на ринку.

Використання даних у моделі

Завдяки використанню даних "Моріон" дослідження може забезпечити надійні та обґрунтовані оцінки щодо впливу різних факторів на продажі бренду

Кардіомагніл. Ці дані є основою для оцінки ринкових тенденцій, динаміки попиту та ефективності рекламної активності в порівнянні з конкурентами.

2.6 Постановка вимог до моделі

Модель має відповідати таким ключовим вимогам:

1. Прогнозування продажів на основі рекламної активності

Основна мета моделі – визначення впливу різних видів рекламної активності (телебачення, діджитал, радіо, зовнішня реклама, преса) на продажі бренду Кардіомагніл з урахуванням конкурентного впливу. Це включає:

- Оцінку внеску кожного каналу реклами: модель повинна визначати, який внесок у загальний обсяг продажів забезпечує кожен вид реклами (ТБ, діджитал, радіо, зовнішня реклама, преса).
- Аналіз динаміки рекламної активності протягом досліджуваного періоду з урахуванням коротко- та довгострокових змін у рекламних витратах і стратегіях.

2. Збалансованість історичних даних

Оскільки дані для різних каналів збиралися різними методами, модель повинна враховувати ці відмінності:

- Нормалізація даних для телевізійної реклами: для 2022 року, коли дані доступні лише у вигляді кількості виходів, модель має використати проксі-методи для наближення рейтингів.
- Дані діджитал-реклами лише для ПК: через відсутність інформації про рекламну активність на мобільних пристроях, модель працюватиме з даними ПК без додаткових коригувань.
- Надійність даних радіоактивності до жовтня 2023 року: інформація від більшості радіостанцій дозволяє вважати дані радіореклами достовірними навіть до відновлення моніторингу в жовтні 2023 року, що дозволяє використовувати повні дані радіоактивності без коригувань, оскільки вони відповідають реальній активності на ринку.

3. Врахування конкурентного середовища

Для забезпечення більш точних прогнозів модель повинна враховувати фактори конкурентного середовища:

- Продажі основних конкурентів: включення даних про продажі основних конкурентів дозволить зрозуміти, як змінюється ринкова частка бренду Кардіомагніл під впливом активності інших гравців. Як зазначено в роботі Eduardo Fiuza та Marcelo Lisboa, у фармацевтичній галузі маркетингові інвестиції та здатність брендів створювати довіру є ключовими факторами для збереження конкурентних позицій на насичених ринках [10, с. 203].
- Цінова політика: аналіз середніх цін на продукцію Кардіомагніл і конкурентів дозволить врахувати вплив ціноутворення на продажі.
- Технічне проникнення (дистрибуція): включення показників дистрибуції для бренду Кардіомагніл та конкурентів дасть можливість оцінити вплив охоплення на продажі.

4. Врахування загальних продажів у категорії

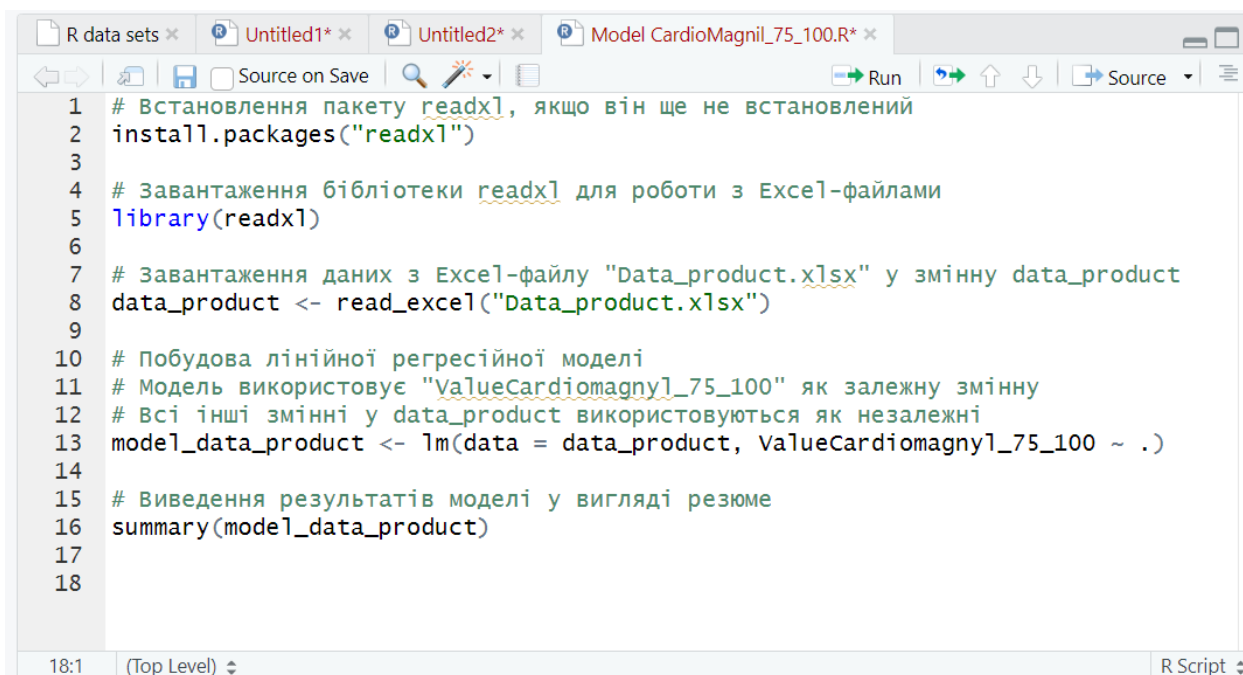
Для розуміння загальних ринкових трендів модель також повинна враховувати:

- Загальні обсяги продажів у категорії: включення цього показника дозволить визначити, як зміни попиту на ринку в цілому впливають на продажі Кардіомагніл. Це забезпечить додатковий контекст, аби відрізнити зміни у продажах, пов'язані зі змінами на ринку, від змін, спричинених специфічною рекламною активністю або іншими факторами.

3 РОЗРОБКА ЕКОНОМЕТРИЧНОЇ МОДЕЛІ ДЛЯ БРЕНДУ КАРДІОМАГНІЛ

3.1 Розробка моделі враховуючи всі наявні незалежні змінні

Для подальшого моделювання ми розробили і застосовуємо наступний код, який дозволяє побудувати модель, в якій залежна змінна ValueCardiomagnil_75_100 відображає продажі продукту від рекламних інвестицій і інших показників категорії продуктів, які містять ацетилсаліцилову кислоту. Опис змінних, які використані в коді, відображено в додатку А. Код зображено на Рисунку 3.1



```

1 # Встановлення пакету readxl, якщо він ще не встановлений
2 install.packages("readxl")
3
4 # Завантаження бібліотеки readxl для роботи з Excel-файлами
5 library(readxl)
6
7 # Завантаження даних з Excel-файлу "Data_product.xlsx" у змінну data_product
8 data_product <- read_excel("Data_product.xlsx")
9
10 # Побудова лінійної регресійної моделі
11 # Модель використовує "ValueCardiomagnyl_75_100" як залежну змінну
12 # Всі інші змінні у data_product використовуються як незалежні
13 model_data_product <- lm(data = data_product, ValueCardiomagnyl_75_100 ~ .)
14
15 # Виведення результатів моделі у вигляді резюме
16 summary(model_data_product)
17
18

```

Рисунок 3.1 – Код розрахунку коефіцієнтів незалежних змін економетричної моделі

В результаті моделювання були отримані коефіцієнти незалежних змінних, які відображені в таблиці 3.1, а на рис 3.2 коефіцієнти якості моделі.

Таблиця 3.1 Коефіцієнти незалежних змінних

Незалежна змінна	Коефіцієнт	Std. Error	t value	Pr(> t)	Pi-нь зна-ті
(Intercept)	2.555e+04	2.760e+04	0.925	0.3599	
ValueCategory	1.000e+00	7.876e-05	12696.919	<2e-16	***
PackCategory	-3.932e-03	7.727e-03	-0.509	0.6135	
ValueMagnicor_75_100	-1.000e+00	1.057e-04	-9462.069	<2e-16	***
ValueMagnicor_75_30	-9.997e-01	2.168e-04	-4610.972	<2e-16	***
ValueCardiomagnyl_Forte_150_100	-1.000e+00	1.709e-04	-5853.153	<2e-16	***
ValueAspirin_100_98	-1.000e+00	1.218e-04	-8206.961	<2e-16	***
ValueCardiomagnyl_75_30	-9.997e-01	2.238e-04	-4466.947	<2e-16	***
ValueAspirin_100_56	-9.999e-01	7.755e-05	-12893.411	<2e-16	***
ValueLowspirin_75_120	-1.000e+00	2.888e-04	-3464.556	<2e-16	***
ValueAspirin_100_28	-9.999e-01	1.680e-04	-5951.725	<2e-16	***
ValueMagnicor_150_100	-1.000e+00	2.999e-04	-3334.389	<2e-16	***
ValueLowspirin_75_30	-1.000e+00	3.722e-04	-2687.730	<2e-16	***
ValueAcard_75_100	-1.001e+00	4.285e-04	-2334.862	<2e-16	***
ValueAsl_75_30	-1.000e+00	4.511e-04	-2216.793	<2e-16	***
ValueAcekor_100_100	-1.000e+00	4.828e-04	-2072.240	<2e-16	***
ValueCardio_Dar_75_100	-9.999e-01	5.068e-04	-1973.076	<2e-16	***
ValueCardiomagnyl_Forte_75_50	-1.000e+00	1.002e-03	-998.453	<2e-16	***
ValueAcard_75_50	-1.000e+00	9.220e-04	-1084.724	<2e-16	***
ValueLowspirin_75_80	-9.993e-01	7.499e-04	-1332.558	<2e-16	***

Продовження таблиці 3.1 Коефіцієнти незалежних змінних

Незалежна змінна	Коефіцієнт	Std. Error	t value	Pr(> t)	Pі-нь зна-ті
ValueAsk_100_30	-9.985e-01	1.144e-03	-872.582	<2e-16	***
ValueAcekor_100_50	-1.002e+00	1.262e-03	-793.947	<2e-16	***
ValueMagnicor_150_30	-9.982e-01	2.134e-03	-467.727	<2e-16	***
ValueCardio_Dar_150_100	-1.000e+00	1.546e-03	-647.048	<2e-16	***
ValueTrombolik_100_20	-9.991e-01	2.644e-03	-377.847	<2e-16	***
ValueCardisave_75_50	-1.002e+00	3.301e-03	-303.581	<2e-16	***
ValueAsaphen_80_90	-9.980e-01	2.944e-03	-339.020	<2e-16	***
ValueCardisave_75_30	-9.992e-01	4.064e-03	-245.850	<2e-16	***
ValueThrombo_100_30	-9.963e-01	3.628e-03	-274.610	<2e-16	***
ValueAspicard_100_20	-1.008e+00	6.677e-03	-150.945	<2e-16	***
ValueAspirin_300_28	-9.876e-01	1.239e-02	-79.701	<2e-16	***
ValueAsaphen_80_30	-9.996e-01	1.458e-02	-68.581	<2e-16	***
CostCardiomagnyl_75_100	-4.388e-01	1.256e+01	-0.035	0.9723	
CostMagnicor_75_100	-6.563e+00	1.961e+01	-0.335	0.7395	
CostMagnicor_75_30	-4.766e+00	5.545e+01	-0.086	0.9319	
CostCardiomagnyl_Forte_150_100	6.164e+00	1.401e+01	0.440	0.6621	
CostAspirin_100_98	2.275e+00	3.591e+00	0.633	0.5298	
CostCardiomagnyl_75_30	-1.139e+01	2.507e+01	-0.454	0.6518	
CostAspirin_100_56	1.080e-01	4.615e-01	0.234	0.8161	

Продовження таблиці 3.1 Коефіцієнти незалежних змінних

Незалежна змінна	Коефіцієнт	Std. Error	t value	Pr(> t)	Рі-нь зна-ті
CostLowspirin_75_120	4.292e+00	1.216e+01	0.353	0.7258	
CostAspirin_100_28	-1.076e+01	8.128e+00	-1.323	0.1927	
CostMagnicor_150_100	-2.247e+01	1.422e+01	-1.580	0.1214	
CostLowspirin_75_30	7.149e+00	3.454e+01	0.207	0.8370	
CostAcard_75_100	1.566e+00	4.138e+00	0.378	0.7069	
CostAsl_75_30	-5.696e+01	3.594e+01	-1.585	0.1203	
CostAcekor_100_100	5.200e-01	4.759e+00	0.109	0.9135	
CostCardio_Dar_75_100	1.434e+01	8.342e+00	1.719	0.0928	
CostCardiomagnyl_Forte_75_50	-1.887e+00	9.345e+00	-0.202	0.8409	
CostAcard_75_50	-2.412e+00	6.856e+00	-0.352	0.7267	
CostLowspirin_75_80	1.252e+00	1.312e+00	0.954	0.3452	
CostAsk_100_30	4.853e-01	1.776e+00	0.273	0.7860	
CostAcekor_100_50	-1.163e+00	3.060e+00	-0.380	0.7059	
CostMagnicor_150_30	2.225e+01	2.360e+01	0.943	0.3510	
CostCardio_Dar_150_100	-4.753e+00	3.095e+00	-1.536	0.1319	
CostTrombolik_100_20	-1.332e+00	4.174e+00	-0.319	0.7511	
CostCardisave_75_50	2.287e+00	6.344e+00	0.360	0.7203	
CostAsaphen_80_90	9.357e-01	2.071e+00	0.452	0.6537	
CostCardisave_75_30	5.256e-01	5.047e+00	0.104	0.9176	

Продовження таблиці 3.1 Коефіцієнти незалежних змінних

Незалежна змінна	Коефіцієнт	Std. Error	t value	Pr(> t)	Рі-нь зна-ті
CostThrombo_100_30	-1.237e+00	1.593e+00	-0.776	0.4419	
CostAspicard_100_20	-1.684e+00	1.706e+01	-0.099	0.9218	
PenCardiomagnyl_75_100	-2.176e+02	3.063e+02	-0.711	0.4811	
PenMagnicor_75_100	-1.952e+01	2.224e+02	-0.088	0.9305	
PenMagnicor_75_30	3.415e+01	1.954e+02	0.175	0.8621	
PenCardiomagnyl_Forte_150_100	-4.035e+01	1.523e+02	-0.265	0.7923	
PenAspirin_100_98	-9.087e+00	1.470e+01	-0.618	0.5397	
PenCardiomagnyl_75_30	-3.332e+01	4.648e+01	-0.717	0.4774	
PenAspirin_100_56	-2.216e+00	8.925e+00	-0.248	0.8051	
PenLowspirin_75_120	-4.896e+01	5.948e+01	-0.823	0.4150	
PenAspirin_100_28	1.758e+01	3.176e+01	0.554	0.5827	
PenMagnicor_150_100	1.871e+01	3.220e+01	0.581	0.5642	
PenLowspirin_75_30	3.885e+01	6.224e+01	0.624	0.5358	
PenAcard_75_100	-7.683e+00	1.798e+01	-0.427	0.6714	
PenAsl_75_30	5.572e+01	6.779e+01	0.822	0.4156	
PenAcekor_100_100	-1.571e+01	3.318e+01	-0.474	0.6382	
PenCardio_Dar_75_100	-1.251e+01	1.451e+01	-0.862	0.3936	
PenCardiomagnyl_Forte_75_50	7.230e-01	1.993e+01	0.036	0.9712	
PenAcard_75_50	1.740e+01	2.353e+01	0.740	0.4636	

Продовження таблиці 3.1 Коефіцієнти незалежних змінних

Незалежна змінна	Коефіцієнт	Std. Error	t value	Pr(> t)	Pі-нь зна-ті
PenLowspirin_75_80	-6.898e+00	1.343e+01	-0.514	0.6101	
PenAsk_100_30	-1.031e+01	2.058e+01	-0.501	0.6189	
PenMagnicor_150_30	1.506e+01	4.378e+01	0.344	0.7325	
PenCardio_Dar_150_100	3.473e+00	1.466e+01	0.237	0.8138	
PenTrombolik_100_20	-5.789e+00	2.719e+01	-0.213	0.8324	
PenCardisave_75_50	3.681e+01	4.286e+01	0.859	0.3952	
PenAsaphen_80_90	-4.170e+01	1.775e+02	-0.235	0.8154	
PenCardisave_75_30	-4.758e+01	5.678e+01	-0.838	0.4066	
PenThrombo_100_30	2.408e+01	7.601e+01	0.317	0.7529	
PenAspicard_100_20	5.482e+01	6.421e+01	0.854	0.3980	
PenAspirin_300_28	-4.072e+01	1.595e+02	-0.255	0.7997	
PenAsaphen_80_30	6.787e+01	9.114e+01	0.745	0.4605	
TVCardiomagnyl	-3.218e-02	3.971e-02	-0.810	0.4222	
TVMagnicor	3.508e-02	5.271e-02	0.666	0.5093	
TVAsk	4.333e-02	8.284e-02	0.523	0.6036	
TVAspirin	1.245e-02	1.752e-01	0.071	0.9437	
TVLowspirin	1.779e-10	7.126e-10	0.250	0.8041	
DigitalMagnicor	-4.554e-04	9.198e-04	-0.495	0.6231	
DigitalCardiomagnyl	-9.703e-05	4.887e-04	-0.199	0.8436	

Продовження таблиці 3.1 Коефіцієнти незалежних змінних

Незалежна змінна	Коефіцієнт	Std. Error	t value	Pr(> t)	Pі-нь зна-ті
DigitalAspirin	2.251e-05	4.594e-05	0.490	0.6266	
DigitalAcekor	-1.142e-04	1.847e-04	-0.618	0.5396	
DigitalAsk0Teva	-1.744e-05	1.556e-05	-1.121	0.2685	
RadioCardiomagnyl	-1.707e-01	4.179e-01	-0.408	0.6850	

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
 Residual standard error: 38 on 43 degrees of freedom
 Multiple R-squared: 1, Adjusted R-squared: 1
 F-statistic: 1.296e+09 on 99 and 43 DF, p-value: < 2.2e-16

Рисунок 3.2 – Коефіцієнти незалежних змінних і ключові показники
економетричної моделі

Коментарі до показників моделі, які зображені на рис 3.2 і в таблиці 3.1

1. Загальна якість моделі:

Модель має дуже високу пояснювальну здатність:

$$R^2 = 1, \text{ Adjusted } R^2 = 1.$$

Ці показники свідчать, що модель практично повністю пояснює варіацію залежної змінної (ValueCardiomagnyl_75_100). Однак, така досконалість може також свідчити про перенавчання моделі на даних, що слід враховувати при оцінці її узагальнювальної здатності.

2. Значущість змінних:

Більшість змінних є статистично значущими (рівень значущості p-value < 0.05). Це свідчить про сильний вплив незалежних змінних на залежну змінну.

Коефіцієнти змінних показують, яким чином кожен фактор впливає на обсяг продажів. Позитивні коефіцієнти вказують на збільшення продажів при зростанні відповідного фактору, а негативні - на зменшення.

3. Інтерпретація результатів моделі:

Категорійні змінні (ValueCategory, PackCategory): Ці змінні є найбільш значущими, що підтверджує, що категорія продукту та пакування відіграють важливу роль у формуванні результатів.

Вартісні змінні (Cost..., Value...): Загальні витрати та продажі в попередні періоди мають вагомий вплив на результативність. Це може свідчити про те, що історичні дані допомагають прогнозувати майбутні результати.

4. Можливі проблеми моделі:

Негативні впливи: Хоча деякі змінні мають негативні коефіцієнти, це може бути викликано:

Непрямими ефектами.

Кореляцією між незалежними змінними (мультиколінеарність).

Проблема перенавчання: Надто високий R^2 може свідчити про те, що модель надмірно адаптована до тренувальних даних і може показувати гірші результати на нових даних.

3.2 Тестування економетричної моделі на відповідність теоретичним принципам економетрики і подальша оптимізація

Для перевірки на мультиколінеарності ми розрахуємо фактор інфляції дисперсії, або VIF (Variance Inflation Factor) і таким чином визначимо змінні, які мають високий рівень кореляції.

Надалі ми усунемо змінні з високим VIF (наприклад, >10), залишивши найбільш значущі.

На рисунку 3.3 відображено код для визначення фактору інфляції дисперсії, а в таблиці 3.2 результат обчислення.

```

R 4.4.1 · ~/Work/Diploma/Data/
> # 1. Перевірка мультиколінеарності
> library(car)
> vif_values <- vif(model_data_product) # Розрахунок VIF для всіх змінних

```

Рисунок 3.3 – обчислення фактору інфляції дисперсії

Таблиця 3.2 Результат обчислення фактору інфляції дисперсії

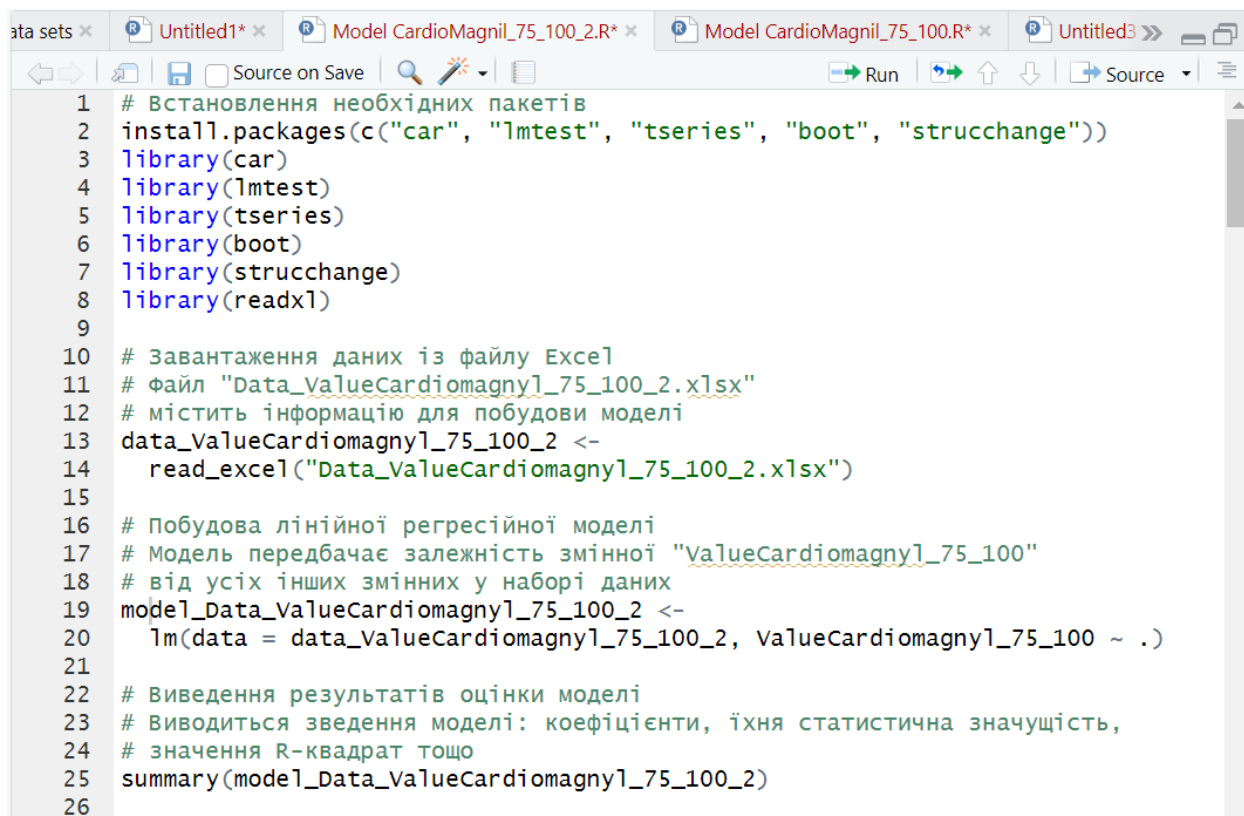
Змінна	Ф-р	Змінна	Ф-р
ValueCategory	7273	CostAsk_100_30	95
PackCategory	10467	CostAcekor_100_50	7
ValueMagnicor_75_100	426	CostMagnicor_150_30	162
ValueMagnicor_75_30	374	CostCardio_Dar_150_100	71
ValueCardiomagnyl_Forte_150_100	215	CostTrombolik_100_20	176
ValueAspirin_100_98	769	CostCardisave_75_50	584
ValueCardiomagnyl_75_30	485	CostAsaphen_80_90	83
ValueAspirin_100_56	802	CostCardisave_75_30	108
ValueLowspirin_75_120	157	CostThrombo_100_30	74
ValueAspirin_100_28	193	CostAspicard_100_20	6
ValueMagnicor_150_100	72	PenCardiomagnyl_75_100	4043
ValueLowspirin_75_30	158	PenMagnicor_75_100	6972
ValueAcard_75_100	88	PenMagnicor_75_30	28848
ValueAsl_75_30	65	PenCardiomagnyl_Forte_150_100	4035
ValueAcekor_100_100	83	PenAspirin_100_98	36118
ValueCardio_Dar_75_100	109	PenCardiomagnyl_75_30	301
ValueCardiomagnyl_Forte_75_50	279	PenAspirin_100_56	15159
ValueAcard_75_50	39	PenLowspirin_75_120	26069
ValueLowspirin_75_80	204	PenAspirin_100_28	9743
ValueAsk_100_30	50	PenMagnicor_150_100	528
ValueAcekor_100_50	614	PenLowspirin_75_30	9175
ValueMagnicor_150_30	63	PenAcard_75_100	935
ValueCardio_Dar_150_100	60	PenAsl_75_30	17419
ValueTrombolik_100_20	52	PenAcekor_100_100	2987
ValueCardisave_75_50	70	PenCardio_Dar_75_100	251
ValueAsaphen_80_90	10	PenCardiomagnyl_Forte_75_50	1359
ValueCardisave_75_30	68	PenAcard_75_50	442
ValueThrombo_100_30	8	PenLowspirin_75_80	10747

Продовження таблиці 3.2 Результат обчислення фактору інфляції дисперсії

Змінна	Ф-р	Змінна	Ф-р
ValueAspicard_100_20	33	PenAsk_100_30	1824
ValueAspirin_300_28	121	PenMagnicor_150_30	1615
ValueAsaphen_80_30	5	PenCardio_Dar_150_100	1436
CostCardiomagnyl_75_100	1799	PenTrombolik_100_20	1844
CostMagnicor_75_100	1333	PenCardisave_75_50	7165
CostMagnicor_75_30	1275	PenAsaphen_80_90	4022
CostCardiomagnyl_Forte_150_100	319	PenCardisave_75_30	37397
CostAspirin_100_98	2360	PenThrombo_100_30	1669
CostCardiomagnyl_75_30	1207	PenAspicard_100_20	7673
CostAspirin_100_56	14	PenAspirin_300_28	814
CostLowspirin_75_120	530	PenAsaphen_80_30	1355
CostAspirin_100_28	815	TVCardiomagnyl	48
CostMagnicor_150_100	343	TVMagnicor	52
CostLowspirin_75_30	1027	TVAsk	136
CostAcard_75_100	492	TVAspirin	103
CostAsl_75_30	678	TVLowspirin	3
CostAcekor_100_100	418	DigitalMagnicor	104
CostCardio_Dar_75_100	330	DigitalCardiomagnyl	19
CostCardiomagnyl_Forte_75_50	97	DigitalAspirin	53
CostAcard_75_50	456	DigitalAcekor	13
CostLowspirin_75_80	38	DigitalAskOTeva	12
		RadioCardiomagnyl	120

В результаті перевірки на мультиколінеарність ми прийшли до висновку, що багато незалежних змінних не проходять тест. Також, попередній аналіз говорить про ризик перенавчання моделі. Тому було прийнято рішення зменшити кількість незалежних змінних, а саме залишити ті змінні, продукти яких складають 67% всіх продажів категорії і провести додаткове економетричне моделювання з подальшим аналізом. Внизу список брендів, показники яких будуть використані для подальшого аналізу: Кардіомагніл 75 мг 100 пігулок, Магнікор 75 мг 100 пігулок, Магнікор 75 мг 30 пігулок, Кардіомагніл Форте 150 мг 100 пігулок, Аспірін-Кардіо 100 мг 98 пігулок. Як і раніше для незалежних змінних використовуються продажі продуктів в грошах, ціна і дистрибуція продуктів, а також рекламна активність.

Завантажуємо данні в RStudio і будуємо модель з допомогою коду, який ми відобразили на рисунку 3.4:



```

1 # Встановлення необхідних пакетів
2 install.packages(c("car", "lmtest", "tseries", "boot", "strucchange"))
3 library(car)
4 library(lmtest)
5 library(tseries)
6 library(boot)
7 library(strucchange)
8 library(readxl)
9
10 # Завантаження даних із файлу Excel
11 # Файл "Data_ValueCardiomagnyl_75_100_2.xlsx"
12 # містить інформацію для побудови моделі
13 data_ValueCardiomagnyl_75_100_2 <-
14   read_excel("Data_ValueCardiomagnyl_75_100_2.xlsx")
15
16 # Побудова лінійної регресійної моделі
17 # Модель передбачає залежність змінної "ValueCardiomagnyl_75_100"
18 # від усіх інших змінних у наборі даних
19 model_Data_ValueCardiomagnyl_75_100_2 <-
20   lm(data = data_ValueCardiomagnyl_75_100_2, ValueCardiomagnyl_75_100 ~ .)
21
22 # Виведення результатів оцінки моделі
23 # Виводиться зведення моделі: коефіцієнти, їхня статистична значущість,
24 # значення R-квадрат тощо
25 summary(model_Data_ValueCardiomagnyl_75_100_2)
26

```

Рисунок 3.4 – код побудови моделі з обмеженою кількістю незалежних змінних

На рисунку 3.5 відображено розраховані коефіцієнти незалежних змінних і ключові показники ефективності економетричної моделі.

```
> summary(model_Data_ValueCardiomagnyl_75_100_2)

Call:
lm(formula = ValueCardiomagnyl_75_100 ~ ., data = data_ValueCardiomagnyl_75_100_2)

Residuals:
    Min       1Q   Median       3Q      Max
-673841  -57495    6179    70390   533007

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.595e+07  6.055e+06   2.633  0.00961 **
ValueCategory  2.118e-01  3.014e-02   7.030  1.53e-10 ***
ValueMagnicor_75_100  5.430e-02  1.163e-01   0.467  0.64140
ValueMagnicor_75_30 -4.043e-01  2.049e-01  -1.973  0.05082 .
ValueCardiomagnyl_Forte_150_100  1.355e+00  2.177e-01   6.227  7.84e-09 ***
ValueAspirin_100_98 -1.241e-01  1.209e-01  -1.027  0.30655
CostCardiomagnyl_75_100 -1.190e+03  1.135e+04  -0.105  0.91667
CostMagnicor_75_100 -2.310e+04  3.325e+04  -0.695  0.48858
CostMagnicor_75_30  4.174e+03  1.037e+05   0.040  0.96797
CostCardiomagnyl_Forte_150_100  5.969e+03  1.406e+04   0.425  0.67189
CostAspirin_100_98 -2.455e+03  1.480e+03  -1.659  0.09989 .
PenCardiomagnyl_75_100 -2.349e+05  1.122e+05  -2.094  0.03841 *
PenMagnicor_75_100  1.093e+05  5.619e+04   1.945  0.05419 .
PenCardiomagnyl_Forte_150_100 -1.853e+04  4.990e+04  -0.371  0.71113
PenCardiomagnyl_75_30 -7.795e+03  3.190e+04  -0.244  0.80738
PenAspirin_100_98  5.527e+03  2.525e+03   2.189  0.03058 *
TVCardiomagnyl  1.794e+02  6.949e+01   2.582  0.01106 *
TVMagnicor  1.271e+02  8.891e+01   1.430  0.15545
TVAsk -1.301e+02  8.651e+01  -1.504  0.13527
TVAspirin  1.197e+02  1.525e+02   0.785  0.43410

TVLowspirin -7.058e+02  2.558e+02  -2.759  0.00674 **
DigitalMagnicor  1.958e-01  1.550e+00   0.126  0.89967
DigitalCardiomagnyl  2.807e-01  6.319e-01   0.444  0.65777
DigitalAspirin  5.952e-02  4.976e-02   1.196  0.23411
DigitalAcekor -2.451e-01  3.928e-01  -0.624  0.53395
DigitalAsk0Teva  3.695e-02  2.829e-02   1.306  0.19407
RadioCardiomagnyl -6.017e+02  6.145e+02  -0.979  0.32955
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 154400 on 116 degrees of freedom
Multiple R-squared:  0.9851,    Adjusted R-squared:  0.9817
F-statistic: 294.7 on 26 and 116 DF, p-value: < 2.2e-16
```

Рисунок 3.5 – коефіцієнти незалежних змін і ключові показники моделі

Детальний аналіз результатів моделювання:

1. Загальна характеристика моделі:

- Кількість змінних у моделі: 26 предикторів.
- Рівень пояснюваної дисперсії:
 - Multiple R^2 : 0.9851. Модель пояснює 98.51% варіації залежної змінної, що є дуже високим показником.
 - Adjusted R^2 : 0.9817. Залишається високим навіть після врахування кількості змінних, що свідчить про адекватність моделі.

- Стандартна похибка залишків: 154400. Це свідчить про певний розкид залишків, який потрібно інтерпретувати з урахуванням масштабу залежної змінної.

2. Оцінка значимості моделі:

- F-статистика: 294.7 з $p\text{-value} < 2.2 \cdot 10^{-16}$, що свідчить про високу статистичну значимість моделі. Імовірність отримання таких результатів випадково є практично нульовою.

3. Інтерпретація змінних:

- Змінні з високою статистичною значимістю ($p < 0.05$):
 - ValueCategory ($p < 2 \cdot 10^{-16}$): Позитивний коефіцієнт (0.2118) вказує на значний вплив загальної категорії товару на обсяг продажів.
 - ValueCardiomagnyl_Forte_150_100 ($p = 7.84 \cdot 10^{-9}$): Значний позитивний ефект на продажі, коефіцієнт 0.1355.
 - PenCardiomagnyl_75_100 ($p = 0.03841$): Негативний вплив ($-2.349 \cdot 10^4$), що потребує подальшого аналізу.
 - PenAspirin_100_98 ($p = 0.03058$): Невеликий позитивний вплив ($5.527 \cdot 10^3$).
 - TVCardiomagnyl ($p = 0.01106$) Позитивний ефект ($1.794 \cdot 10^2$), що підтверджує вплив телевізійної реклами на продажі.
 - TVLowspirin ($p = 0.00674$): Негативний вплив ($-7.058 \cdot 10$).
- Змінні з $p > 0.05$ ($p > 0.05$) (незначущі):
 - Змінні, які відображають витрати на рекламу (Digital, Radio), показали здебільшого несуттєві коефіцієнти, наприклад:
 - DigitalMagnicor ($p = 0.89967$),
 - DigitalAspirin ($p = 0.23411$),
 - RadioCardiomagnyl ($p = 0.32955$).

4. Аналіз рекламних змінних:

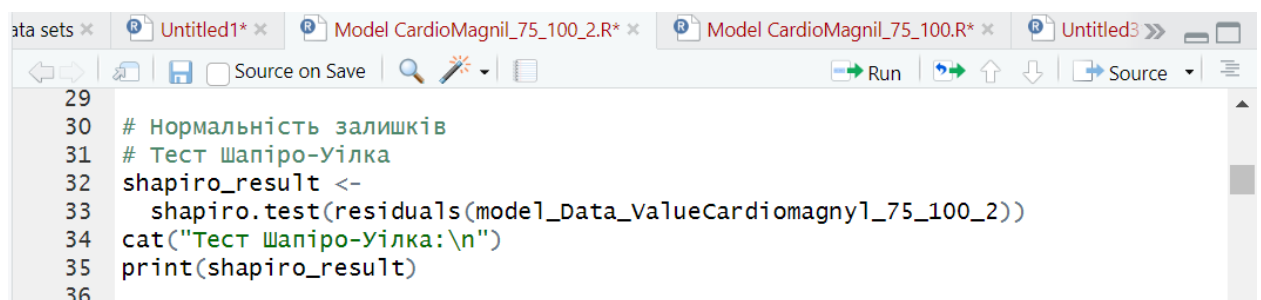
- TV: Телевізійна реклама має змішані результати: деякі змінні є значущими (TVCardiomagnyl, TVLowspirin), що може свідчити про вибірковий вплив на різні продукти.

- Digital: Цифрові рекламні витрати переважно виявилися незначущими, можливо, через слабку асоціацію з продажами або недостатні інвестиції.
- Radio: Радіо має низькі коефіцієнти та статистичну незначущість, що вказує на неефективність цього каналу для розглянутих продуктів.

5. Оцінка залишків:

- Розподіл залишків: Стандартна похибка залишків становить 154400, що є високим значенням. Це може свідчити про наявність інших важливих факторів, які не враховані в моделі.
- Наявність мультиколінеарності: Попередній аналіз VIF показав високі значення для деяких змінних, наприклад:
 - ValueCategory,
 - ValueMagnicor_75_100,
 - ValueCardiomagnyl_75_100. Це може впливати на стабільність оцінок коефіцієнтів.

Надалі ми провели тест Тест Шапіро-Уїлка – рисунок 3.6, який використовується для перевірки нормальності розподілу залишків у регресійній моделі. Це важливо, оскільки нормальність залишків є однією з ключових передумов лінійної регресії. Вона впливає на точність оцінок коефіцієнтів і коректність статистичних тестів, таких як t-тест і F-тест.



```

29
30 # Нормальність залишків
31 # Тест Шапіро-Уїлка
32 shapiro_result <-
33   shapiro.test(residuals(model_Data_ValueCardiomagnyl_75_100_2))
34 cat("Тест Шапіро-Уїлка:\n")
35 print(shapiro_result)
36
  
```

Рисунок 3.6 – код для тесту Шапіро-Уїлка

Ненормальність розподілу залишків може свідчити про систематичні помилки моделі, невраховані фактори або викиди. Це може також вказувати на проблеми гетероскедастичності (змінна дисперсія залишків) або автокореляції (залежність між залишками).

Результати тесту Шапіро-Уїлка зображені на рисунку 3.7 дозволяють діагностувати ці проблеми та обрати необхідні кроки для оптимізації моделі. Це може включати трансформацію змінних, використання альтернативних моделей чи робастних методів оцінки.

```
Тест Шапіро-Уїлка:
> print(shapiro_result)

Shapiro-wilk normality test

data: residuals(model_Data_ValueCardiomagnyl_75_100_2)
W = 0.89356, p-value = 1.083e-08
> |
```

Рисунок 3.7 – результат тесту Шапіро-Уїлка

Результат тесту Шапіро-Уїлка показує, що у представленому результаті значення статистики $W = 0.89356$ та дуже малий рівень значущості ($p\text{-value} = 1.083e-08$) свідчать про те, що нульова гіпотеза про нормальність залишків відхиляється.

Це вказує на те, що розподіл залишків істотно відрізняється від нормального. Такий результат має важливі наслідки для інтерпретації моделі:

1. Проблеми статистичної інтерпретації: Більшість статистичних тестів, що використовуються в лінійній регресії (наприклад, t-тест для коефіцієнтів), припускають нормальність залишків. Порушення цієї передумови може призводити до некоректних висновків.

2. Імовірна присутність систематичних помилок: Ненормальність залишків може бути ознакою того, що модель не враховує деякі важливі фактори або існують нелінійні залежності між змінними.

3. Неоднорідність розсіювання: Можливо, у моделі є гетероскедастичність (змінна дисперсія залишків), що також може впливати на достовірність оцінок.

4. Необхідність трансформації: Можливо, деякі змінні потребують логарифмічної, степеневої чи іншої трансформації для поліпшення відповідності моделі.

Подальші дії - перевірити графіки залишків:

- Побудувати графік залишків (residuals) проти передбачених значень (fitted values).
- Побудувати Q-Q графік для аналізу відхилення від нормального розподілу.

Код для побудови залишків (residuals) проти передбачених значень (fitted values) зображено на рисунку 3.8.

```

37 # Побудова графіка залишків проти передбачених значень
38
39 # Розрахунок залишків і передбачених значень
40 residuals <- residuals(model_Data_ValueCardiomagnyl_75_100_2) # залишки
41 fitted_values <- fitted(model_Data_ValueCardiomagnyl_75_100_2) # передбачені значення
42
43 # Побудова графіка
44 plot(
45   fitted_values, residuals,
46   main = "Графік залишків проти передбачених значень",
47   xlab = "передбачені значення",
48   ylab = "Залишки",
49   pch = 19, col = "blue"
50 )
51 abline(h = 0, col = "red", lwd = 2, lty = 2) # Додавання горизонтальної лінії на рівні 0
52
53 # Додавання додаткових пояснень
54 grid() # Додавання сітки
55

```

Рисунок 3.8 – код для побудови залишків проти передбачених значень

Графік залишків проти передбачених значень, який зображено на рисунку 3.9, показує наступне:

1. Залишки розподілені навколо горизонтальної лінії на рівні 0, що свідчить про відсутність суттєвих систематичних помилок у передбаченнях. Це хороший знак для лінійної моделі.
2. Патерни або систематичні структури в залишках:
 - Якщо точки не демонструють чітких узорів (наприклад, параболічного чи конусоподібного розширення), це означає, що модель добре враховує залежність між змінними.
 - У нашому випадку деякі точки віддалені від основної хмари (ймовірні викиди або екстремальні значення), але загалом залишки виглядають доволі рівномірно розподіленими.
3. Рівномірність розсіювання залишків:

- Рівномірний розподіл залишків свідчить про те, що припущення гомоскедастичності (однорідної варіації залишків) виконується.
- На графіку спостерігається деяке збільшення розсіювання залишків при зростанні передбачених значень.

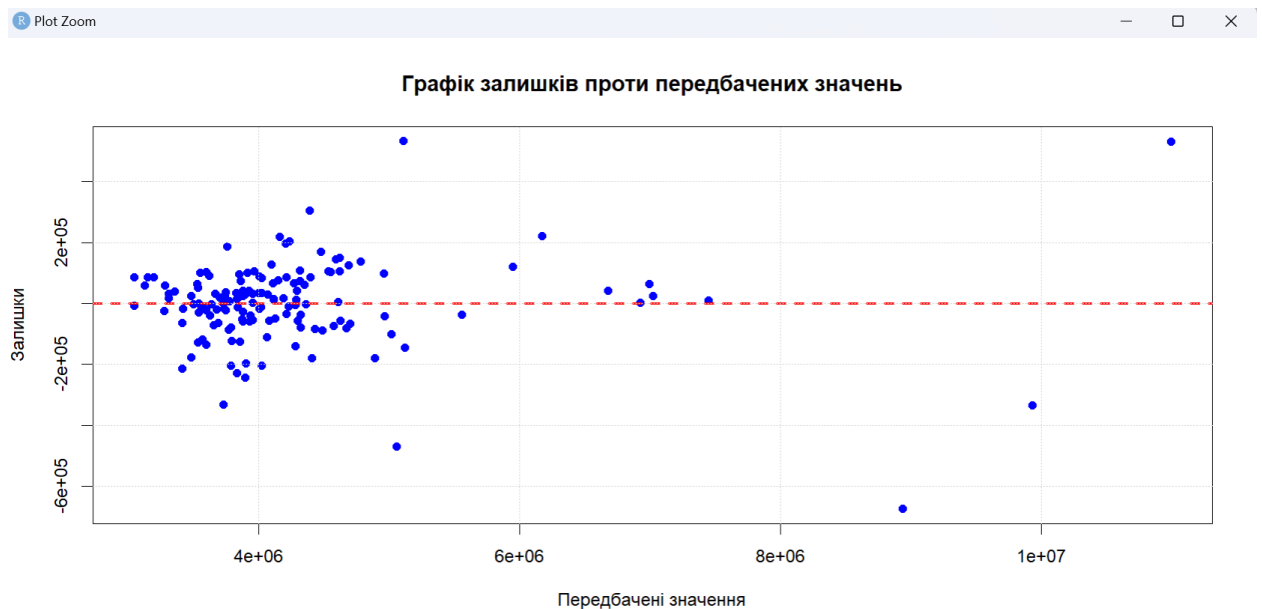


Рисунок 3.9 – Графік залишків проти передбачених значень

Код для Q-Q графіку для аналізу відхилення від нормального розподілу зображено на рисунку 3.10.

```

55
56
57 # Q-Q графік
58 qqnorm(residuals(model_Data_ValueCardiomagnyl_75_100_2))
59 qqline(residuals(model_Data_ValueCardiomagnyl_75_100_2), col = "red", lwd = 2)
60

```

Рисунок 3.10 - Код для Q-Q графіку для аналізу відхилення від нормального розподілу зображено

На Q-Q графіку, що зображено на рисунку 3.11, видно, що значна частина точок лежить на червоній прямій, що свідчить про те, що залишки здебільшого відповідають нормальному розподілу. Проте:

1. Відхилення на кінцях графіка:

- На краях графіка (квантили в діапазоні менше -2 та більше 2) точки значно відхиляються від червоної лінії.

- Це свідчить про аномалії в хвостах розподілу залишків, що можуть вказувати на наявність впливових точок або асиметрію в даних.

2. Висновок щодо нормальності:

- У центрі графіка залишки близькі до нормального розподілу.
- Проте відхилення на краях вказують на те, що розподіл залишків не є повністю нормальним.

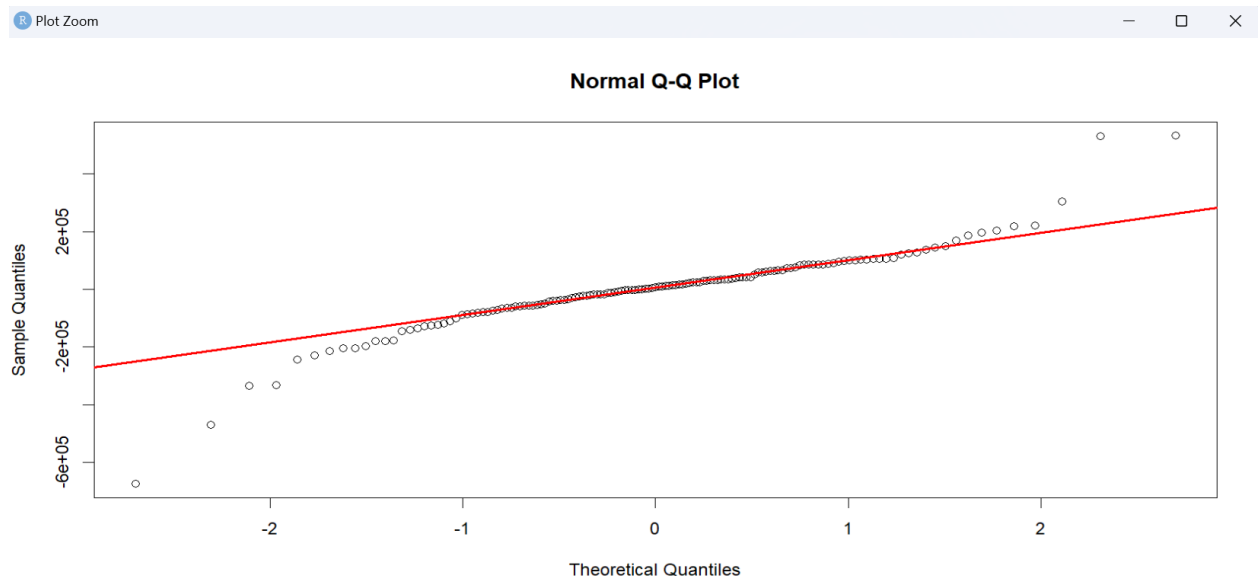


Рисунок 3.11 – Q-Q графік для аналізу відхилення від нормального розподілу

Отже, ми будемо ідентифікувати впливові точки, які можуть значно впливати на результати, використовуючи куханівські відстані для їх виявлення і видалимо їх. На рисунку 3.12 зображено код для виявлення впливових точок і їх видалення:

```

136 cooks_d <- cooks.distance(model_Data_ValueCardiomagnyl_75_100_2)
137 plot(cooks_d, type = "h", main = "Куханівські відстані", ylab = "Cook's Distance")
138 abline(h = 4 / nrow(model_Data_ValueCardiomagnyl_75_100_2$model), col = "red", lwd = 2)
139 influential_points <- which(cooks_d > (4 / nrow(model_Data_ValueCardiomagnyl_75_100_2$model)))
140 cat("Впливові точки:\n")
141 print(influential_points)
142
143
144 data_cleaned <- model_Data_ValueCardiomagnyl_75_100_2$model[-influential_points, ]
145 model_cleaned <- lm(formula(model_Data_ValueCardiomagnyl_75_100_2), data = data_cleaned)
146 qqnorm(residuals(model_cleaned), main = "Q-Q графік (без впливових точок)")
147 qqline(residuals(model_cleaned), col = "red", lwd = 2)
148 shapiro.test(residuals(model_cleaned))
149

```

Рисунок 3.12 – код для виявлення впливових точок

В результаті ідентифікування визначили впливові точки для видалення, які зображені на рисунку 3.13.

```
> cat("Впливові точки:\n")
Впливові точки:
> print(influential_points)
  1  8  9 12 13 17 18 45 53 65 67 103 111
  1  8  9 12 13 17 18 45 53 65 67 103 111
> data_cleaned <- model_Data_ValueCardiomagnyl_75_100_2$model[-influential_points, ]
> model_cleaned <- lm(formula(model_Data_ValueCardiomagnyl_75_100_2), data = data_cleaned)
> qqnorm(residuals(model_cleaned), main = "Q-Q графік (без впливових точок)")
> qqline(residuals(model_cleaned), col = "red", lwd = 2)
> shapiro.test(residuals(model_cleaned))

      Shapiro-Wilk normality test

data:  residuals(model_cleaned)
W = 0.99299, p-value = 0.7679
```

Рисунок 3.13 – Визначення впливових точок і результати тесту Шапіро-Уїлка після їх видалення

Результати тесту Шапіро-Уїлка після очищення моделі (видалення впливових точок), який зображено на рисунку 3.13, показують такі значення:

- $W = 0.99299$: Значення близьке до 1, що свідчить про те, що розподіл залишків є майже нормальним.
- $p\text{-value} = 0.7679$: Значення значно перевищує 0.05, що означає, що немає підстав відхиляти нульову гіпотезу про нормальність розподілу залишків.

Отже, можемо зробити наступні висновки:

Нормальність залишків підтверджена:

- Після очищення моделі залишки розподілені нормально. Це підтверджується високим $p\text{-value}$ та W , близьким до 1.

Модель відповідає припущенню нормальності залишків:

- Це покращує надійність статистичних тестів (значущості коефіцієнтів, довірчих інтервалів) і підтверджує коректність моделі.

Далі ми проведемо тест Бреуша-Пагана на гомоскедастичність, для визначення чи виконується умова рівномірності дисперсії залишків. Код для проведення тесту Бреуша-Пагана зображено на рисунку 3.14.

```

150 # Гомоскедастичність (тест Бреуша-Пагана)
151 bptest_result <- bptest(model_cleaned)
152 cat("Тест Бреуша-Пагана:\n")
153 print(bptest_result)
154

```

Рисунок 3.14 – Код для тесту Бреуша-Пагана на гомоскедастичність

Отже в результаті проведеного тесту Бреуша-Пагана, який зображено на рисунку 3.15, виявлена гетероскедастичність:

- Нульова гіпотеза (H_0): "Дисперсія залишків є постійною" відхиляється, оскільки $p\text{-value} < 0.05$.
- Це означає, що дисперсія залишків змінюється залежно від значень незалежних змінних, і модель страждає від гетероскедастичності.

```

> # Гомоскедастичність (тест Бреуша-Пагана)
> bptest_result <- bptest(model_cleaned)
> cat("Тест Бреуша-Пагана:\n")
Тест Бреуша-Пагана:
> print(bptest_result)

studentized Breusch-Pagan test

data: model_cleaned
BP = 52.201, df = 26, p-value = 0.001701

```

Рисунок 3.15 – Результат тесту Бреуша-Пагана

Отже це значить, що порушення припущення гомоскедастичності може вплинути на коректність стандартних помилок коефіцієнтів, що призводить до ненадійних $p\text{-value}$ і довірчих інтервалів.

Далі ми використаємо робастні стандартні помилки, що дозволять коригувати стандартні помилки для врахування гетероскедастичності. На рисунку 3.16 зображено код

```

159
160 # Робастні стандартні помилки (HC3)
161 robust_se <- coeftest(model_cleaned, vcov = vcovHC(model_cleaned, type = "HC3"))
162 cat("Робастні стандартні помилки:\n")
163 print(robust_se)
164

```

Рисунок 3.16 – Код для розрахунку робастних стандартних помилок

Результат використання робастних стандартних помилок зображено на рисунку 3.17.



t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.2844e+06	4.0818e+06	0.5597	0.5769301
ValueCategory	3.3493e-01	3.0971e-02	10.8141	< 2.2e-16 ***
ValueMagnicor_75_100	-9.7159e-02	1.2465e-01	-0.7795	0.4374807
ValueMagnicor_75_30	-5.5119e-01	1.8876e-01	-2.9200	0.0042999 **
ValueCardiomagnyl_Forte_150_100	4.1305e-01	3.0403e-01	1.3586	0.1772515
ValueAspirin_100_98	-5.7183e-01	1.4527e-01	-3.9364	0.0001505 ***
CostCardiomagnyl_75_100	9.7315e+03	9.5970e+03	1.0140	0.3129501
CostMagnicor_75_100	-1.1353e+04	2.2600e+04	-0.5024	0.6164797
CostMagnicor_75_30	7.8515e+03	6.2067e+04	0.1265	0.8995819
CostCardiomagnyl_Forte_150_100	1.7288e+04	9.8308e+03	1.7586	0.0816161 .
CostAspirin_100_98	-1.7618e+02	2.3180e+03	-0.0760	0.9395614
PenCardiomagnyl_75_100	-1.2152e+05	5.8873e+04	-2.0642	0.0415162 *
PenMagnicor_75_100	6.8858e+04	3.1773e+04	2.1672	0.0325266 *
PenCardiomagnyl_Forte_150_100	3.6753e+03	2.8625e+04	0.1284	0.8980855
PenCardiomagnyl_75_30	-1.1047e+04	1.6204e+04	-0.6817	0.4969297
PenAspirin_100_98	7.9217e+03	2.6309e+03	3.0110	0.0032754 **
TVCardiomagnyl	1.6561e+02	3.3262e+01	4.9789	2.578e-06 ***
TMagnicor	-3.3142e+01	4.3098e+01	-0.7690	0.4436571
TVAAsk	-1.1415e+02	5.1393e+01	-2.2210	0.0285379 *
TVAspirin	2.7914e+02	9.9378e+01	2.8089	0.0059490 **
TVLowspirin	-3.4417e+02	1.3023e+02	-2.6429	0.0095034 **
DigitalMagnicor	6.9701e-01	1.1054e+00	0.6305	0.5297339
DigitalCardiomagnyl	3.6014e-01	2.5169e-01	1.4309	0.1554958
DigitalAspirin	4.1543e-02	3.6757e-02	1.1302	0.2610139
DigitalAcekor	-1.8696e-01	2.5679e-01	-0.7281	0.4682280
DigitalAskOTeva	3.1160e-02	1.6120e-02	1.9330	0.0559837 .
RadioCardiomagnyl	-4.5690e+02	3.9082e+02	-1.1691	0.2450732

Рисунок 3.17 – Розрахунок робастних стандартних помилок

Інтерпретуємо ключові значення:

Можемо зробити висновок, що результати повторюють стандартну регресію, але стандартні помилки скориговані для врахування можливих проблем, таких як:

- Гетероскедастичність: Коли дисперсія залишків залежить від значень предикторів.
- Автокореляція: Коли залишки пов'язані між собою.

Таким чином, реклама на TV (TVCardiomagnyl, TVAspirin) суттєво впливає на залежну змінну. Також, бачимо, що реклама конкурентів, а саме Ask Teva і Lowspirin, має негативний вплив на залежну змінну, в той самий час реклама Aspirin, має позитивний вплив, що імовірно залежить від цінових характеристики брендів конкурентів. Реклама більш дешевших препаратів призводить до падіння продажів бренду Кардіомагніл, тоді як реклама більш дорогих препаратів (Aspirin) призводить до зростання. Вплив вартості (Cost) є здебільшого незначущим ($p > 0.05$). ValueCategory — одна з найсильніших змінних. Змінні, пов'язані з Digital, мають менший вплив порівняно з TV.

Для подальшого усунення гетероскедастичності ми вирішили логарифмувати залежну і незалежні змінні і таким чином трансформувати модель. Для визначення, які незалежні змінні трансформувати ми використовували наступні критерії: широкий діапазон значень, асиметричний розподіл і наявність нульових значень. На основі цих критерії ми вибрали наступні незалежні змінні для логарифмування: ValueCategory, ValueMagnicor_75_100, TVCardiomagnyl, TVMagnicor, RadioCardiomagnyl, DigitalMagnicor", DigitalCardiomagnyl, DigitalAspirin. Також ми логарифмували залежну змінну. На рисунку 3.18 зображено код для логарифмування.

```

58
59 # Логарифмування змінних
60 data_cleaned$log_ValueCategory <- log(data_cleaned$ValueCategory)
61 data_cleaned$log_ValueMagnicor_75_100 <- log(data_cleaned$ValueMagnicor_75_100)
62 data_cleaned$log_ValueCardiomagnyl_75_100 <- log(data_cleaned$ValueCardiomagnyl_75_100)
63 data_cleaned$log_TVCardiomagnyl <- log(data_cleaned$TVCardiomagnyl + 1) # +1 для уникнення log(0)
64 data_cleaned$log_TVMagnicor <- log(data_cleaned$TVMagnicor + 1)
65 data_cleaned$log_RadioCardiomagnyl <- log(data_cleaned$RadioCardiomagnyl + 1)
66 data_cleaned$log_DigitalMagnicor <- log(data_cleaned$DigitalMagnicor + 1)
67 data_cleaned$log_DigitalCardiomagnyl <- log(data_cleaned$DigitalCardiomagnyl + 1)
68 data_cleaned$log_DigitalAspirin <- log(data_cleaned$DigitalAspirin + 1)
69

```

Рисунок 3.18 – код для логарифмування визначених незалежних змінних

Надалі ми побудували нову трансформовану модель з логарифмованими змінними. На рисунку 3.19 зображено код для побудови моделі і демонстрації основних показників.

```

106
107 # Формування формули моделі з урахуванням логарифмованих та оригінальних змінних
108 model_formula <- as.formula(paste(
109   "log_ValueCardiomagnyl_75_100 ~",
110   paste(c(
111     # Логарифмовані змінні
112     "log_ValueCategory", "log_ValueMagnicor_75_100", "log_TVCardiomagnyl",
113     "log_TVMagnicor", "log_RadioCardiomagnyl", "log_DigitalMagnicor",
114     "log_DigitalCardiomagnyl", "log_DigitalAspirin",
115     # Оригінальні змінні
116     "ValueMagnicor_75_30", "ValueCardiomagnyl_Forte_150_100", |
117     "ValueAspirin_100_98", "CostCardiomagnyl_75_100",
118     "CostMagnicor_75_30", "CostCardiomagnyl_Forte_150_100",
119     "CostAspirin_100_98", "PenMagnicor_75_100", "PenCardiomagnyl_75_100",
120     "PenCardiomagnyl_Forte_150_100", "PenCardiomagnyl_75_30",
121     "PenAspirin_100_98", "TVAsk", "TVAspirin", "TVLowspirin",
122     "DigitalAsk0Teva", "RadioCardiomagnyl"
123   ), collapse = " + ")
124 ))
125
126 # Побудова моделі
127 model_log_combined <- lm(model_formula, data = data_cleaned)
128
129 # Результати моделі
130 summary(model_log_combined)

```

Рисунок 3.19 – Код для побудови моделі і демонстрації основних показників

В результаті ми отримали показники, які зображені на рисунку 3.20.

```

R 4.4.1 · ~/Work/Diploma/Data/
> summary(model_log_combined)

Call:
lm(formula = model_formula, data = data_cleaned)

Residuals:
    Min       1Q   Median       3Q      Max
-0.040308 -0.010602  0.001435  0.009021  0.032013

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -4.657e+00  1.294e+00  -3.599  0.000492 ***
log_ValueCategory  1.260e+00  8.757e-02  14.392  < 2e-16 ***
log_ValueMagnicor_75_100 -1.164e-01  5.055e-02  -2.302  0.023311 *
log_TVCardiomagnyl  1.210e-02  2.850e-03   4.244  4.79e-05 ***
log_TVMagnicor    4.838e-03  2.628e-03   1.841  0.068462 .
log_RadioCardiomagnyl  5.542e-03  2.356e-03   2.352  0.020562 *
log_DigitalMagnicor -3.695e-03  1.200e-03  -3.078  0.002664 **
log_DigitalCardiomagnyl  2.206e-03  8.176e-04   2.698  0.008152 **
log_DigitalAspirin -1.663e-03  7.802e-04  -2.132  0.035394 *
ValueMagnicor_75_30 -1.065e-07  2.922e-08  -3.644  0.000421 ***
ValueCardiomagnyl_Forte_150_100 4.929e-08  4.334e-08   1.137  0.258022
ValueAspirin_100_98 -7.698e-08  2.240e-08  -3.437  0.000846 ***
CostCardiomagnyl_75_100 -9.791e-04  1.556e-03  -0.629  0.530667
CostMagnicor_75_30 -5.394e-03  5.475e-03  -0.985  0.326816
CostCardiomagnyl_Forte_150_100 6.376e-03  1.807e-03   3.529  0.000623 ***
CostAspirin_100_98 -1.923e-04  2.486e-04  -0.773  0.441003
PenMagnicor_75_100  1.393e-02  6.525e-03   2.135  0.035092 *
PenCardiomagnyl_75_100 -7.246e-03  1.167e-02  -0.621  0.535856
PenCardiomagnyl_Forte_150_100 4.163e-03  5.580e-03   0.746  0.457334
PenCardiomagnyl_75_30 -9.788e-03  3.655e-03  -2.678  0.008608 **
PenAspirin 100 98  2.028e-03  4.536e-04   4.471  1.99e-05 ***

TVAsk          -1.472e-05  9.707e-06  -1.516  0.132551
TVAspirin       7.342e-06  1.592e-05   0.461  0.645610
TVLowspirin    -8.213e-05  3.833e-05  -2.143  0.034440 *
DigitalAsk0Teva  8.501e-09  3.069e-09   2.770  0.006639 **
RadioCardiomagnyl -9.638e-05  5.231e-05  -1.843  0.068228 .

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01652 on 104 degrees of freedom
Multiple R-squared:  0.9922,    Adjusted R-squared:  0.9903
F-statistic: 526.2 on 25 and 104 DF,  p-value: < 2.2e-16
> |

```

Рисунок 3.20 – Коефіцієнти незалежних змін після логарифмування

Отже, ми можемо зробити наступні висновки:

Якість моделі:

- Значення коефіцієнта детермінації ($R^2 = 0.992$) показує, що 99.2% варіації залежної змінної пояснюється незалежними змінними моделі.
- Коригований коефіцієнт детермінації ($\text{Adjusted } R^2 = 0.9903$) підтверджує, що включені змінні суттєво покращують модель, навіть після врахування кількості предикторів.

Рівень помилки:

- Стандартна похибка залишків (Residual standard error = 0.01652) вказує на те, що розбіжності між фактичними і передбаченими значеннями є мінімальними.

Залишки:

- Мінімальні та максимальні залишки моделі (від -0.040308 до 0.032013) свідчать про те, що немає аномальних або великих похибок у прогнозуванні.
- Симетричний розподіл залишків свідчить про відсутність серйозних порушень передумов лінійної регресії.

Загальні висновки:

- Модель є високоякісною з точки зору пояснення варіації залежної змінної.
- Деякі змінні мають низьку статистичну значущість і, можливо, їх можна видалити для спрощення моделі.
- Висока точність моделі дозволяє вважати її придатною для подальшого використання, однак слід перевірити її на гетероскедастичність.

Для перевірки на гетероскедастичність ми проведемо тест Бреуша-Пагана, результати якого зображені на рисунку 3.21.

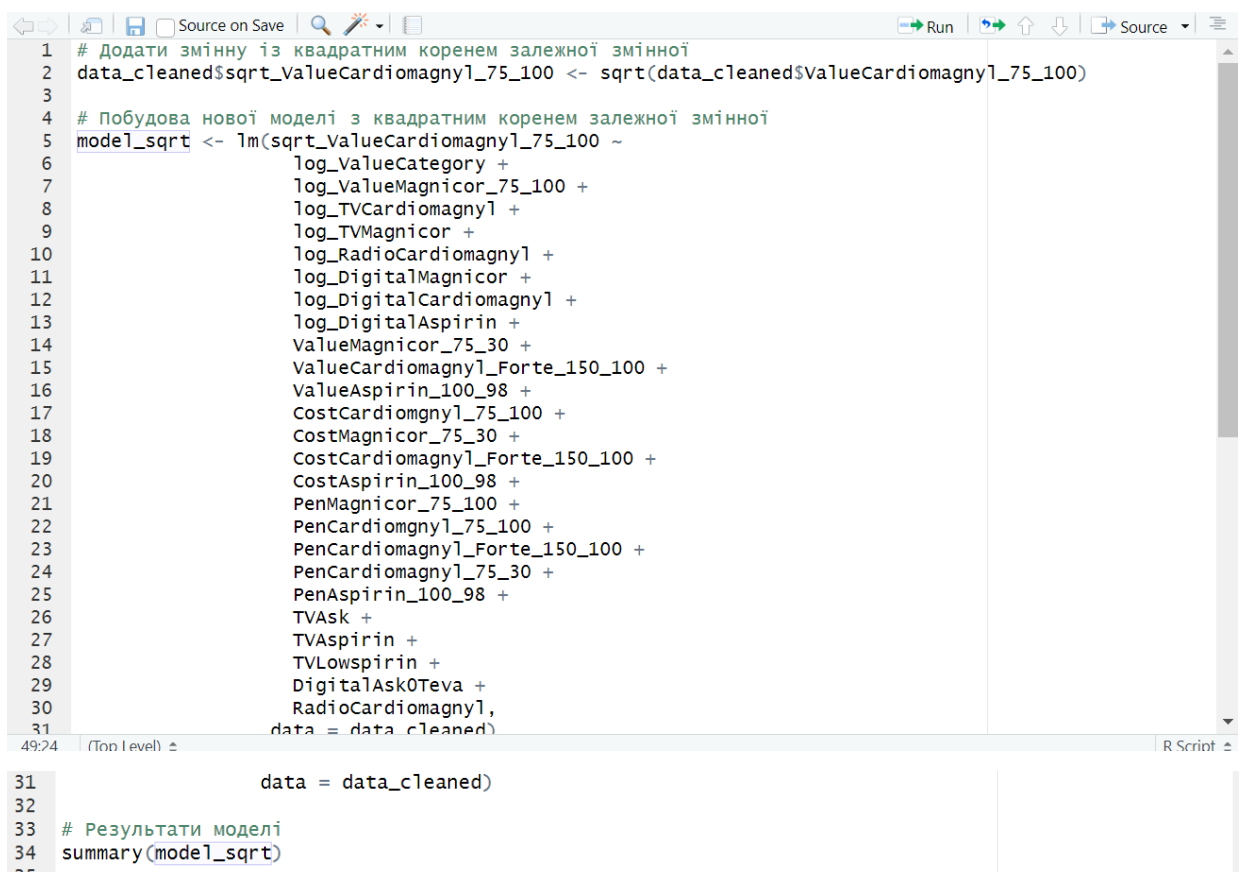
```
> print(bp_test_log_combined)
      studentized Breusch-Pagan test
data:  model_log_combined
BP = 43.16, df = 25, p-value = 0.01344
> |
```

Рисунок 3.21 – Тест Бреуша-Пагана після логарифмування вибраних змінних

На основі результатів тесту Бреуша-Пагана для моделі `model_log_combined`, отримано:

- Значення статистики тесту (BP) = 43.16
- Ступені свободи (df) = 25
- Рівень значущості (p-value) = 0.01344

Оскільки p-value (0.01344) менше стандартного рівня значущості (0.05), ми відхиляємо нульову гіпотезу про гомоскедастичність. Це свідчить про те, що в моделі присутня гетероскедастичність. Модель потребує подальшого доопрацювання. Для цього ми спробуємо трансформувати модель додаючи змінну з квадратним коренем залежної змінної. На рисунку 3.22 зображено код для розрахунку трансформованої моделі.



```

1 # Додати змінну із квадратним коренем залежної змінної
2 data_cleaned$sqrt_valueCardiomagnyl_75_100 <- sqrt(data_cleaned$valueCardiomagnyl_75_100)
3
4 # Побудова нової моделі з квадратним коренем залежної змінної
5 model_sqrt <- lm(sqrt_valueCardiomagnyl_75_100 ~
6     log_valueCategory +
7     log_valueMagnicor_75_100 +
8     log_TVCardiomagnyl +
9     log_TVMagnicor +
10    log_RadioCardiomagnyl +
11    log_DigitalMagnicor +
12    log_DigitalCardiomagnyl +
13    log_DigitalAspirin +
14    ValueMagnicor_75_30 +
15    ValueCardiomagnyl_Forte_150_100 +
16    ValueAspirin_100_98 +
17    CostCardiomagnyl_75_100 +
18    CostMagnicor_75_30 +
19    CostCardiomagnyl_Forte_150_100 +
20    CostAspirin_100_98 +
21    PenMagnicor_75_100 +
22    PenCardiomagnyl_75_100 +
23    PenCardiomagnyl_Forte_150_100 +
24    PenCardiomagnyl_75_30 +
25    PenAspirin_100_98 +
26    TVAsk +
27    TVAspirin +
28    TVLowspirin +
29    DigitalAskOTeva +
30    RadioCardiomagnyl,
31    data = data_cleaned)
32
33 # Результати моделі
34 summary(model_sqrt)
35

```

Рисунок 3.22 – Код для розрахунку трансформованої моделі (залежну змінну перераховано в змінну з квадратним коренем)

В результаті подальшої трансформації моделі отримали показники моделі, що зображені на рисунку 3.23.

```

Console Terminal Background Jobs
R 4.4.1 - ~/Work/Diploma/Data/
> # Результати моделі
> summary(model_sqrt)

Call:
lm(formula = sqrt_ValueCardiomagnyl_75_100 ~ log_ValueCategory +
    log_ValueMagnicor_75_100 + log_TVCardiomagnyl + log_TVMagnicor +
    log_RadioCardiomagnyl + log_DigitalMagnicor + log_DigitalCardiomagnyl +
    log_DigitalAspirin + ValueMagnicor_75_30 + ValueCardiomagnyl_Forte_150_100 +
    ValueAspirin_100_98 + CostCardiomagnyl_75_100 + CostMagnicor_75_30 +
    CostCardiomagnyl_Forte_150_100 + CostAspirin_100_98 + PenMagnicor_75_100 +
    PenCardiomagnyl_75_100 + PenCardiomagnyl_Forte_150_100 + PenCardiomagnyl_75_30 +
    PenAspirin_100_98 + TVAsk + TVAspirin + TVLowspirin + DigitalAskOTeva +
    RadioCardiomagnyl, data = data_cleaned)

Residuals:
    Min       1Q   Median       3Q      Max
-41.566 -10.145   1.134   10.147   30.265

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.982e+04  1.276e+03 -15.536 < 2e-16 ***
log_ValueCategory  1.378e+03  8.633e+01  15.957 < 2e-16 ***
log_ValueMagnicor_75_100 -1.294e+02  4.983e+01 -2.597 0.010751 *
log_TVCardiomagnyl  8.597e+00  2.810e+00  3.060 0.002819 **
log_TVMagnicor  6.610e+00  2.590e+00  2.552 0.012171 *
log_RadioCardiomagnyl  7.408e+00  2.323e+00  3.189 0.001887 **
log_DigitalMagnicor -3.907e+00  1.183e+00 -3.302 0.001316 **
log_DigitalCardiomagnyl  2.543e+00  8.060e-01  3.154 0.002103 **
log_DigitalAspirin -8.077e-01  7.691e-01 -1.050 0.296088
ValueMagnicor_75_30 -1.867e-04  2.881e-05 -6.481 3.10e-09 ***
ValueCardiomagnyl_Forte_150_100 9.978e-05  4.272e-05  2.336 0.021422 *

ValueAspirin_100_98 -8.037e-05  2.208e-05 -3.640 0.000426 ***
CostCardiomagnyl_75_100 -1.127e+00  1.534e+00 -0.734 0.464432
CostMagnicor_75_30  3.766e+00  5.398e+00  0.698 0.486859
CostCardiomagnyl_Forte_150_100 4.742e+00  1.781e+00  2.663 0.008986 **
CostAspirin_100_98 -2.328e-01  2.451e-01 -0.950 0.344405
PenMagnicor_75_100  1.254e+01  6.432e+00  1.950 0.053902 .
PenCardiomagnyl_75_100 -5.931e-01  1.150e+01 -0.052 0.958971
PenCardiomagnyl_Forte_150_100 2.302e+00  5.501e+00  0.418 0.676480
PenCardiomagnyl_75_30 -1.066e+01  3.603e+00 -2.957 0.003839 **
PenAspirin_100_98  2.099e+00  4.471e-01  4.695 8.17e-06 ***
TVAsk -7.029e-03  9.569e-03 -0.735 0.464247
TVAspirin -3.428e-03  1.569e-02 -0.218 0.827517
TVLowspirin -3.738e-02  3.778e-02 -0.989 0.324799
DigitalAskOTeva  7.142e-06  3.025e-06  2.361 0.020105 *
RadioCardiomagnyl -1.064e-01  5.156e-02 -2.063 0.041647 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16.29 on 104 degrees of freedom
Multiple R-squared:  0.9935,    Adjusted R-squared:  0.9919
F-statistic: 632.4 on 25 and 104 DF,  p-value: < 2.2e-16

```

Рисунок 3.23 – Розраховані коефіцієнти для незалежних змінних і основні показники моделі після її трансформації (використання квадратного кореню для залежної змінної)

Значення коефіцієнта детермінації $R^2 = 0.9935$ свідчить про те, що модель пояснює 99.35% варіації залежної змінної. Виправлений коефіцієнт детермінації $R^2_{adj} = 0.9919$ враховує кількість змінних і підтверджує високу якість моделі. F-статистика (632.4, $p < 2.2e-16$) вказує на значущість моделі загалом. Стандартна помилка залишків становить 16.29, що свідчить про низьке відхилення фактичних

значень залежної змінної від передбачених моделлю. Залишки варіюються від -41.566 до 30.265 із середнім значенням, близьким до нуля, що свідчить про адекватність моделі. Надалі проведемо тест Бреуша-Пагана аби перевірити модель на наявність гетероскедастичності, код якого зображено на рисунку 3.24.

```
> bp_test_sqrt <- bptest(model_sqrt)
> print(bp_test_sqrt)

studentized Breusch-Pagan test

data:  model_sqrt
BP = 33.041, df = 25, p-value = 0.13
```

Рисунок 3.24 – Код Бреуша-Пагана

На основі результатів тесту Бреуша-Пагана (Breusch-Pagan test), модель `model_sqrt` не має проблем із гетероскедастичністю.

- Значення $p\text{-value} = 0.13 > 0.05$ вказує на те, що нульова гіпотеза про відсутність гетероскедастичності не може бути відхилена на рівні значущості 5%.
- Отже, варіація залишків моделі є стабільною, і немає систематичних змін дисперсії залишків, залежно від значень незалежних змінних.

Це підтверджує коректність припущення гомоскедастичності моделі.

Далі ми перевіримо розподіл залишків на предмет нормального розподілу і відсутності трендів або структури.

На основі графіків залишків, які зображені на рисунку 3.25, можна зробити наступні висновки:

1. Гістограма залишків:
 - Розподіл залишків виглядає симетричним і наближеним до нормального. Це підтверджує припущення нормальності залишків, необхідне для застосування методу найменших квадратів.
2. Залишки vs прогнозовані значення:
 - Залишки розподілені випадково навколо осі 0, що є хорошою ознакою. Відсутність явної структури або трендів у цьому графіку свідчить про коректність моделі.

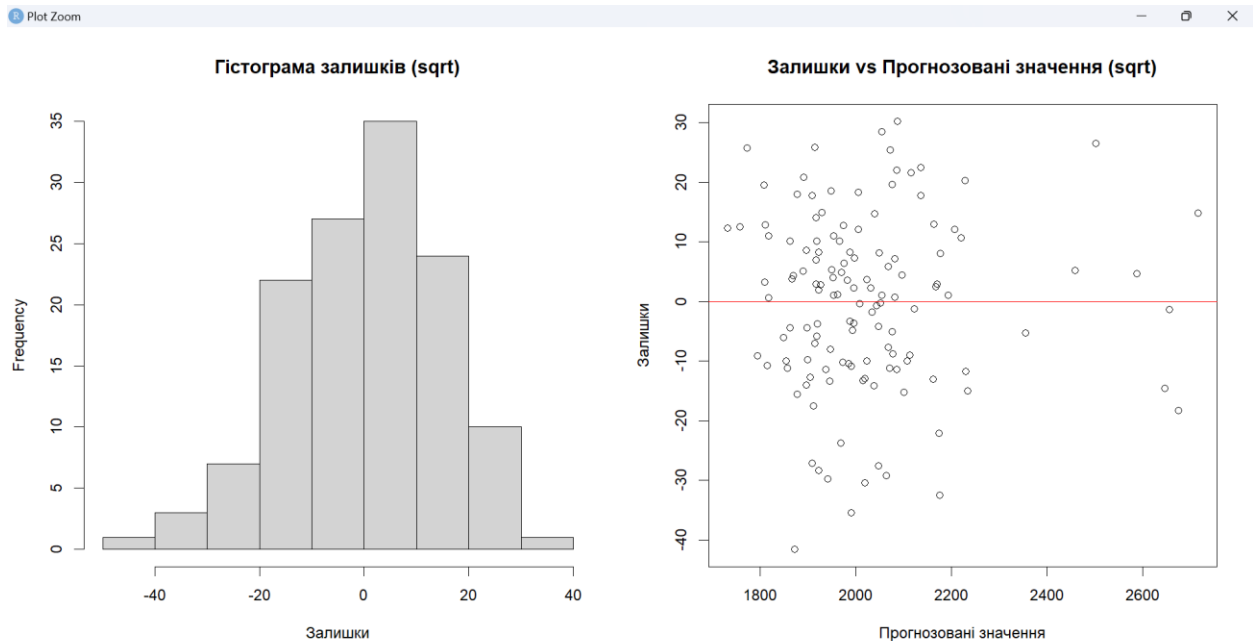


Рисунок 3.25 – Графіки залишків

Таким чином ми приходимо до наступних висновків:

- Результати вказують на те, що модель адекватно пояснює варіацію даних і задовольняє основні припущення щодо залишків (нормальність та відсутність автокореляції).

Надалі ми провели тест Шапіро-Вілка, результати якого зображені на рисунку 3.26, для перевірки гіпотези, що залишки мають нормальний розподіл.

```
Shapiro-wilk normality test
data: residuals(model_sqrt)
W = 0.98602, p-value = 0.2062
```

Рисунок 3.26 – тест Шапіро-Вілка

На основі результатів тесту Шапіро-Вілка:

- $W = 0.98602$, $p\text{-value} = 0.2062$:
 - Нульова гіпотеза тесту Шапіро-Вілка полягає в тому, що залишки мають нормальний розподіл.
 - Оскільки значення $p\text{-value} > 0.05$, ми не можемо відхилити нульову гіпотезу. Це означає, що залишки моделі можна вважати нормально розподіленими.

Отже, можемо зробити висновок, що результати тесту підтверджують, що залишки моделі відповідають припущенню нормальності, що є важливою умовою для інтерпретації моделі і проведення подальших статистичних перевірок.

Надалі ми проведемо VIF тест на мультиколінеарність, результат якого зображено на рисунку 3.27.

```
> vif(model_sqrt)
log_ValueCategory          log_ValueMagnicor_75_100
      71.017192                23.147092
log_TVCardiomagnyl         log_TVMagnicor
      13.137585                11.079567
log_RadioCardiomagnyl      log_DigitalMagnicor
      9.660985                12.302764
log_DigitalCardiomagnyl    log_DigitalAspirin
      3.656737                7.974255
ValueMagnicor_75_30 ValueCardiomagnyl_Forte_150_100
      29.439665                21.235402
ValueAspirin_100_98      CostCardiomagnyl_75_100
      121.587585            123.523654
CostMagnicor_75_30      CostCardiomagnyl_Forte_150_100
      58.603110            22.122856
CostAspirin_100_98      PenMagnicor_75_100
      53.846491            22.674849
PenCardiomagnyl_75_100  PenCardiomagnyl_Forte_150_100
      22.481181            22.548066
PenCardiomagnyl_75_30  PenAspirin_100_98
      8.926191            159.463773
TVAsk                    TVAspirin
      8.851969            3.979711
TVLowspirin              DigitalAskOTeva
      6.275799            2.258045
RadioCardiomagnyl
      9.064824
```

Рисунок 3.27 – тест на мультиколінеарність

На основі результатів тесту VIF (Variance Inflation Factor):

- Значення $VIF > 10$ свідчать про потенційну мультиколінеарність серед змінних.

- У моделі є змінні з високими значеннями VIF:

- log_ValueCategory: 71.017192
- ValueAspirin_100_98: 121.587585
- CostCardiomagnyl_75_100: 123.523654
- PenAspirin_100_98: 159.463773
- CostMagnicor_75_30: 58.603110

Ці змінні мають найбільшу мультиколінеарність і можуть впливати на стабільність оцінки моделі.

Для уникнення мультиколінеарності ми застосуємо методи регуляризації (Lasso, Ridge, Elastic регресії) і виберемо найкращий із них. На рисунку 3.28 зображено код для застосування регуляризаційних методів.

```

4
5 # Підготовка даних для регуляризації
6 X <- model.matrix(sqrt_ValueCardiomagnyl_75_100 ~ ., data = data_sqrt)[, -1] # Ма
7 y <- data_sqrt$sqrt_ValueCardiomagnyl_75_100 # Залежна змінна
8
9 # Регуляризація Lasso (alpha = 1)
10 lasso_model <- cv.glmnet(X, y, alpha = 1)
11 lasso_min_lambda <- lasso_model$lambda.min
12 lasso_mse <- min(lasso_model$cvm)
13
14 # Регуляризація Ridge (alpha = 0)
15 ridge_model <- cv.glmnet(X, y, alpha = 0)
16 ridge_min_lambda <- ridge_model$lambda.min
17 ridge_mse <- min(ridge_model$cv)
18
19 # Регуляризація Elastic Net (комбінація alpha = 0.5)
20 elastic_model <- cv.glmnet(X, y, alpha = 0.5)
21 elastic_min_lambda <- elastic_model$lambda.min
22 elastic_mse <- min(elastic_model$cv)
23
24 # Порівняння моделей
25 comparison <- data.frame(
26   Model = c("Lasso", "Ridge", "Elastic Net"),
27   Min_Lambda = c(lasso_min_lambda, ridge_min_lambda, elastic_min_lambda),
28   MSE = c(lasso_mse, ridge_mse, elastic_mse)
29 )
30
31 print(comparison)
32
33 # Найкраща модель
34 best_model <- comparison[which.min(comparison$MSE), ]
35 print(paste("Найкраща модель:", best_model$Model, "з MSE =", best_model$MSE))
36

```

Рисунок 3.28 – Код для застосування регуляризаційних методів

Для визначення найкращої моделі ми використали показник середньоквадратичної похибки (MSE), яка зображена на рисунку 3.29.

```

> print(comparison)
  Model Min_Lambda      MSE
1  Lasso  0.09030199 358.5993
2  Ridge 16.92094411 792.1214
3 Elastic Net 0.13662029 358.1340
> # Найкраща модель
> best_model <- comparison[which.min(comparison$MSE), ]
> print(paste("Найкраща модель:", best_model$Model, "з MSE =", best_model$MSE))
[1] "Найкраща модель: Elastic Net з MSE = 358.134020039131"
>

```

Рисунок 3.29 – Показник середньоквадратичної похибки

На рисунку 3.29 представлений порівняльний аналіз трьох моделей регресії: Lasso, Ridge, і Elastic Net, із зазначенням мінімального значення параметра регуляризації (λ) та середньоквадратичної похибки (MSE).

Результати:

1. Lasso:

Мінімальне значення λ : 0.09030199.

MSE: 358.5993.

2. Ridge:

Мінімальне значення λ : 16.92094411.

MSE: 792.1214.

3. Elastic Net:

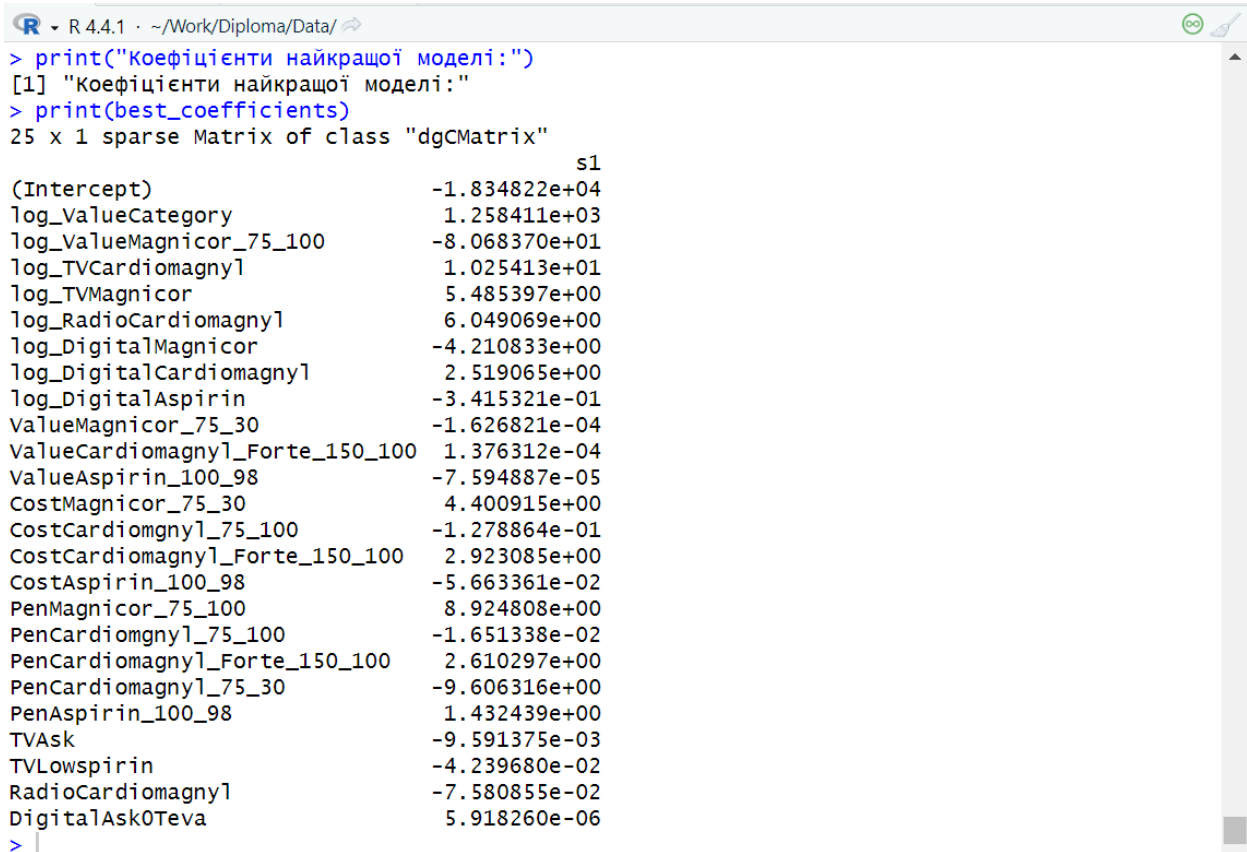
Мінімальне значення λ : 0.13662029.

MSE: 358.1340.

Найкраща модель:

- Elastic Net обрано як найкращу модель, оскільки вона має найнижче значення MSE серед усіх трьох моделей (MSE = 358.1340). Це свідчить, що Elastic Net забезпечує найкращий баланс між помилкою та регуляризацією для цього набору даних.

На рисунку 3.30 представлені коефіцієнти незалежних змінних для Elastic Net моделі.



```

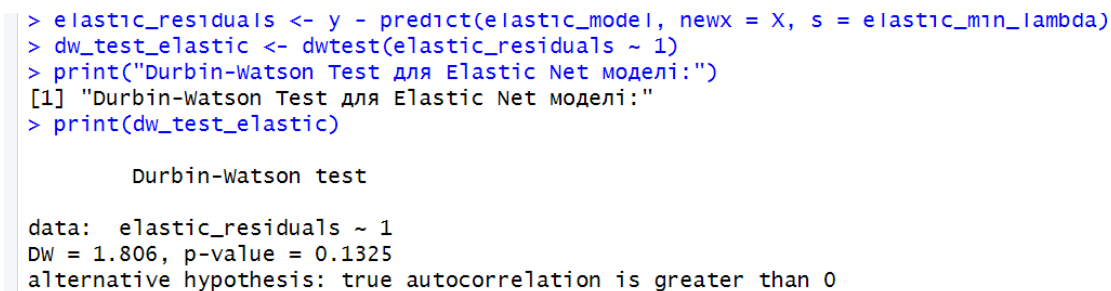
R 4.4.1 · ~/Work/Diploma/Data/
> print("Коефіцієнти найкращої моделі:")
[1] "Коефіцієнти найкращої моделі:"
> print(best_coefficients)
25 x 1 sparse Matrix of class "dgCMatrix"

              s1
(Intercept)    -1.834822e+04
log_ValueCategory  1.258411e+03
log_ValueMagnicor_75_100 -8.068370e+01
log_TVCardiomagnyl  1.025413e+01
log_TVMagnicor    5.485397e+00
log_RadioCardiomagnyl  6.049069e+00
log_DigitalMagnicor -4.210833e+00
log_DigitalCardiomagnyl  2.519065e+00
log_DigitalAspirin -3.415321e-01
ValueMagnicor_75_30 -1.626821e-04
ValueCardiomagnyl_Forte_150_100 1.376312e-04
ValueAspirin_100_98 -7.594887e-05
CostMagnicor_75_30  4.400915e+00
CostCardiomagnyl_75_100 -1.278864e-01
CostCardiomagnyl_Forte_150_100 2.923085e+00
CostAspirin_100_98 -5.663361e-02
PenMagnicor_75_100  8.924808e+00
PenCardiomagnyl_75_100 -1.651338e-02
PenCardiomagnyl_Forte_150_100 2.610297e+00
PenCardiomagnyl_75_30 -9.606316e+00
PenAspirin_100_98  1.432439e+00
TVAsk            -9.591375e-03
TVLowspirin      -4.239680e-02
RadioCardiomagnyl -7.580855e-02
DigitalAskOTeva   5.918260e-06
>

```

Рисунок 3.30 - коефіцієнти незалежних змінних для Elastic Net моделі

Надалі ми проведемо тест Дарбіна-Уотсона, який зображено на рисунку 3.31 і перевіримо модель на автокореляцію.



```

> elastic_residuals <- y - predict(elastic_model, newx = X, s = elastic_min_lambda)
> dw_test_elastic <- dwtest(elastic_residuals ~ 1)
> print("Durbin-Watson Test для Elastic Net моделі:")
[1] "Durbin-Watson Test для Elastic Net моделі:"
> print(dw_test_elastic)

Durbin-Watson test

data:  elastic_residuals ~ 1
DW = 1.806, p-value = 0.1325
alternative hypothesis: true autocorrelation is greater than 0

```

Рисунок 3.31 – Тест Дарбіна-Уотсона

Тест показав, що

- Durbin-Watson статистика (DW): 1.806

Значення статистики DW близьке до 2, що свідчить про відсутність автокореляції залишків (що є бажаним результатом).

- p-value: 0.1325

Значення p-value більше 0.05. Це означає, що немає статистично значущих доказів для відхилення нульової гіпотези (нульова гіпотеза передбачає відсутність автокореляції).

Отже, тест Дарбіна-Уотсона не виявив проблеми автокореляції у залишках. Модель Elastic Net добре підходить за цим критерієм.

Надалі проведемо тест Бреуша-Пагана для перевірки моделі на гетероскедастичність, що зображено на рисунку 3.32:

```
> # Тест Breusch-Pagan для Elastic Net моделі
> bp_test_elastic <- bptest(y ~ predict(elastic_model, X), data = as.data.frame(X))
> print("Breusch-Pagan Test для Elastic Net моделі:")
[1] "Breusch-Pagan Test для Elastic Net моделі:"
> print(bp_test_elastic)

studentized Breusch-Pagan test

data: y ~ predict(elastic_model, X)
BP = 0.44054, df = 1, p-value = 0.5069
```

Рисунок 3.32 – Тест Бреуша-Пагана для моделі Elastic Net

Результати тесту Бреуша-Пагана для моделі Elastic Net:

Інтерпретація:

BP (Breusch-Pagan статистика): 0.44054

Це значення тестової статистики, яке використовується для перевірки гетероскедастичності.

p-value: 0.5069

Значення p-value більше 0.05 свідчить про відсутність статистично значущої гетероскедастичності у залишках моделі.

Отже, можна зробити висновок, що:

Тест Бреуша-Пагана не виявив гетероскедастичності (p-value > 0.05). Це означає, що дисперсія залишків є стабільною, і модель Elastic Net задовольняє припущення гомоскедастичності.

Модель є надійною з точки зору рівномірного розподілу помилок.

Надалі ми провели тест Шапіро-Уїлка, що зображено на рисунку 3.33 для перевірки отриманих відхилень залишків на нормальність.

```
> shapiro_test <- shapiro.test(elastic_residuals)
> print(shapiro_test)
```

Shapiro-Wilk normality test

data: elastic_residuals
W = 0.98874, p-value = 0.3699

Рисунок 3.33 – Тест Шапіро-Уїлка

Результат тесту Шапіро-Уїлка:

Статистика тесту $W=0.98874$: Значення, близьке до 1, означає високу схожість розподілу залишків до нормального.

$p\text{-value}=0.3699$. Це значення вказує на ймовірність, що отримані відхилення від нормального розподілу є випадковими. $p>0.05$ (в даному випадку 0.3699) дає підставу не відхиляти нульову гіпотезу про нормальний розподіл.

Отже, можемо зробити висновок, що розподіл залишків моделі не має суттєвих відхилень від нормального. Це підтверджує, що модель відповідає передумовам нормальності залишків, що є важливою характеристикою для адекватності регресійної моделі.

3.3 Результати моделювання рекламних інвестицій бренду Кардіоманіл

Розробка економетричної моделі для оцінки інвестицій у рекламні активності є важливим кроком для покращення маркетингової стратегії бренду Кардіоманіл. Враховуючи високий рівень конкуренції у фармацевтичній галузі та специфіку ринку лікарських засобів, точний аналіз впливу витрат на різні рекламні канали може суттєво підвищити ефективність маркетингових кампаній і збільшити ринкову частку бренду. Ця модель враховує вплив телевізійної, радіо- та цифрової реклами, що дозволяє сформувати комплексний підхід до управління маркетинговим бюджетом.

Основні результати моделювання

1. Телевізійна реклама

Телевізійна реклама виявилася найбільш ефективним інструментом для стимулювання продажів. З аналізу коефіцієнтів регресії випливає, що збільшення

витрат на телевізійні кампанії забезпечує значний приріст продажів у коротко- і середньостроковій перспективі. Цей канал залишається потужним інструментом для побудови бренду, зокрема через широке охоплення та можливість привернути увагу цільової аудиторії до ключових повідомлень бренду.

2. Радіо-реклама

Радіо також демонструє позитивний вплив на продажі, хоча його ефект менш виражений, ніж у телевізійній рекламі. Цей канал ефективний для охоплення аудиторії, яка регулярно користується автомобілем або проводить час поза домом, де радіо є основним джерелом інформації. Таким чином, радіо-кампанії доповнюють телевізійні, створюючи мультимедіальний ефект.

3. Цифрова реклама

Результати для цифрової реклами свідчать про її зростаючу роль. Хоча ефективність цифрових каналів, як-от соціальні мережі, контекстна реклама та таргетовані email-розсилки, наразі поступається традиційним медіа, її потенціал суттєво збільшується завдяки інтерактивності та можливостям точного таргетингу. Варто зазначити, що цифрова реклама також сприяє формуванню лояльності клієнтів завдяки персоналізованому підходу.

Таблиця коефіцієнтів моделі

Для кращого розуміння впливу незалежних змінних на результативний показник представлено таблицю 3.3 з коефіцієнтами регресійної моделі:

Таблиця 3.3 Коефіцієнти регресійної моделі

Змінна	Коефіцієнт
(Intercept)	-18348.22
log_ValueCategory	1258.41
log_ValueMagnicor_75_100	-8068.37
log_TVCardiomagnyl	1.03E+01
log_TVMagnicor	5.85E+00
log_RadioCardiomagnyl	5.86E-01
log_DigitalMagnicor	-4.21E+00
log_DigitalCardiomagnyl	-2.52E+00
log_DigitalAspirin	-3.43E-01

Продовження таблиці 3.3 Коефіцієнти регресійної моделі

Змінна	Коефіцієнт
ValueMagnicor_75_30	-1.62E-04
ValueCardiomagnyl_Forte_150_100	-7.60E-04
ValueAspirin_100_98	-5.75E-04
CostMagnicor_75_30	4.40E+00
CostCardiomagnyl_75_100	-2.38E+00
CostCardiomagnyl_Forte_150_100	-2.25E+00
CostAspirin_100_98	-5.07E+00
PenMagnicor_75_100	-6.97E-01
PenCardiomagnyl_75_100	-2.67E-01
PenCardiomagnyl_Forte_150_100	-1.90E-01
PenAspirin_100_98	-5.38E-02
TVAsk	-9.59E-03
TVLowspirin	-4.23E-03
RadioCardiomagnyl	-7.58E-02
DigitalAskOTeva	5.92E-06

Розроблена економетрична модель показала свою надійність завдяки високим показникам R^2 та проходженню діагностичних тестів. Наприклад, тест Шапіро-Уїлка підтвердив нормальний розподіл залишків, а тест Бреуша-Пагана свідчить про відсутність гетероскедастичності. Ці результати забезпечують високу точність прогнозів та надійність інтерпретації результатів.

Отримані дані свідчать про те, що споживачі фармацевтичних продуктів часто піддаються впливу масових медіа, таких як телебачення та радіо. Це обумовлено довірою до інформації, представленої у цих каналах, а також їх здатністю охоплювати широку аудиторію. Водночас цифрова реклама є важливим інструментом для залучення молодшої аудиторії, яка активніше використовує інтернет для пошуку інформації про ліки.

Посилення телевізійної реклами. Враховуючи високий коефіцієнт ефективності цього каналу, варто збільшити інвестиції у телевізійні кампанії, зокрема під час пікових сезонів.

Розширення цифрових активностей. Необхідно переглянути стратегію використання цифрової реклами, зокрема впровадження програматик-кампаній та інтерактивного контенту.

Посилення мультиканального підходу. Комбінування кількох каналів комунікації дозволить максимально ефективно використовувати маркетинговий бюджет і збільшити охоплення.

Довгострокові перспективи

Оцінка довгострокового впливу інвестицій є важливим завданням. Вивчення тенденцій продажів протягом кількох років дозволить оцінити, які канали забезпечують сталий ефект, а які підходять для швидкого досягнення результатів.

Виклики та подальші дослідження

Оцінка конкурентного середовища

Важливо враховувати дії конкурентів, які також активно використовують різні канали реклами. Доцільно відслідковувати рекламні стратегії конкурентів і враховувати їх в подальшому моделюванні рекламних інвестицій.

Адаптація до нових технологій

Зі зростанням ролі штучного інтелекту в маркетингу виникають нові можливості для аналізу даних та автоматизації кампаній. Використання інструментів AI для аналізу поведінки споживачів може суттєво підвищити ефективність моделі.

Вплив макроекономічних факторів

Фактори, такі як зміна купівельної спроможності, економічні кризи або пандемії, можуть суттєво впливати на поведінку споживачів. Включення таких змінних у модель дозволить покращити її прогностичну силу.

ВИСНОВКИ

У процесі виконання магістерської роботи було досліджено теоретичні та практичні аспекти планування рекламних інвестицій у фармацевтичній галузі з використанням економетричних методів. Отримані результати дозволяють сформулювати такі висновки та рекомендації:

1. Економетричні моделі як інструмент оптимізації рекламних інвестицій.

Розроблена модель довела свою ефективність у вирішенні завдань, пов'язаних із плануванням рекламних кампаній. Використання таких методів дає змогу кількісно оцінити вплив різних рекламних каналів на ключові маркетингові показники, такі як зростання продажів та залучення нових клієнтів. На підставі проведеного аналізу встановлено, що економетричні моделі можуть застосовуватись для визначення оптимального розподілу бюджету між різними медіаканалами.

2. Аналіз ефективності телевізійної реклами.

Телевізійна реклама залишається найпотужнішим каналом комунікації, особливо для фармацевтичних брендів, які орієнтовані на масовий ринок. На основі отриманих даних визначено, що телевізійна реклама забезпечує найбільше охоплення аудиторії та найвищу короткострокову та довгострокову віддачу. Вона є ефективною для швидкого донесення інформації до широких верств населення, що особливо важливо під час запуску нових продуктів або акційних кампаній.

3. Роль радіореклами у комунікаційній стратегії.

Радіореклама довела свою значущість як інструмент підтримки регулярного контакту з аудиторією. Вона є особливо ефективною для аудиторії, яка щодня слухає радіо під час роботи чи пересування. Встановлено, що радіореклама забезпечує сталість контактів із цільовою аудиторією, підтримує рівень впізнаваності бренду та підсилює ефективність телевізійних кампаній.

4. Потенціал цифрових медіа у фармацевтичному маркетингу.

Цифрові канали комунікації є перспективними для довгострокових маркетингових цілей завдяки своїй таргетованості, можливості аналізу поведінки користувачів та персоналізації контенту.

5. Практичні рекомендації щодо оптимізації рекламних інвестицій:

- **Для телевізійної реклами:** зосередити основні ресурси на її використанні для досягнення короткострокових та довгострокових маркетингових цілей. Найбільший ефект спостерігається при запуску національних кампаній, спрямованих на широке охоплення аудиторії.
- **Для радіореклами:** використовувати її для підтримання частоти контактів із цільовою аудиторією та доповнення телевізійних кампаній. Рекомендується активніше залучати радіорекламу у регіонах, де її слухання є стабільним.
- **Для цифрової реклами:** інтегрувати її у загальну комунікаційну стратегію як засіб для досягнення довгострокових результатів. Використання цифрових платформ доцільне для персоналізованої взаємодії з цільовою аудиторією та створення контенту, який враховує потреби окремих сегментів ринку.

6. Наукова новизна.

У роботі запропоновано економетричну модель для оцінки ефективності рекламних кампаній у фармацевтичному секторі. Модель дозволяє проводити аналіз ефективності кожного медіаканалу окремо, а також інтегрувати їх у єдину систему для забезпечення раціонального розподілу рекламного бюджету.

7. Практичне значення роботи.

Результати дослідження можуть бути використані маркетинговими відділами фармацевтичних компаній для вдосконалення медіапланування. Запропоновані рекомендації дозволяють адаптувати маркетингові стратегії до змін ринку, підвищити рентабельність рекламних інвестицій та забезпечити довгострокову конкурентоспроможність бренду.

8. Подальші перспективи дослідження.

На підставі отриманих результатів рекомендується зосередити майбутні дослідження на інтеграції додаткових змінних, таких як макроекономічні фактори, демографічні особливості аудиторії та поведінкові моделі споживачів, що дозволить підвищити точність прогнозів та розширити сферу застосування економетричних моделей у маркетингу.

ПЕРЕЛІК ПОСИЛАНЬ

1. Karsten Schrör, Acetylsalicylic Acid
2. Chris Brooks, Introductory Econometrics for Finance
3. Joseph F. Hair Jr., G. Tomas M. Hult, Christian M. Ringle, Marko Sarstedt, Nicholas P. Danks, Soumya Ray, Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R
4. Robin Lock, Patti Frazer Lock, Kari Lock Morgan, Eric Lock, Dennis F. Lock, Statistics: Unlocking the Power of Data – RStudio User's Guide
5. Manoj Bhatt, Econometrics 513 Lecture Notes
6. Les Binet, Peter Field, Marketing Effectiveness in the Digital Era
7. Robin Lock, Patti Frazer Lock, Kari Lock Morgan, Eric Lock, Dennis F. Lock, Statistics: Unlocking the Power of Data
8. Thijs J. Rinsema, One Hundred Years of Aspirin
9. Łukasz Fijałkowski, Magdalena Skubiszewska, Grzegorz Grześk, Frankline Kiptoo Koech, Alicja Nowaczyk, Acetylsalicylic Acid–Primus Inter Pares in Pharmacology
10. Eduardo Fiuza, Marcelo Lisboa, Credence Goods and Market Power: An Econometric Study on the Brazilian Pharmaceutical Industry
11. Les Binet, Peter Field, The Long and the Short of It: Balancing Short and Long-Term Marketing Strategies
12. Jeffrey M. Wooldridge, Introductory Econometrics: A Modern Approach
13. Zoltán Lakner, Anna Kiss, József Popp, Zoltán Zéman, Domicián Máté, Judit Oláh, From Basic Research to Competitiveness: An Econometric Analysis of the Global Pharmaceutical Sector
14. Les Binet, Peter Field, Effectiveness in Context: A Manual for Brand Building
15. Олександра Волошин, Наталія Галайко, Економетрія (Частина 1)
16. W. N. Venables, D. M. Smith, R Core Team, An Introduction to R
17. Damodar N. Gujarati, Basic Econometrics

18. Bruce E. Hansen, Econometrics
19. Byron Sharp, How Brands Grow: What Marketers Don't Know
20. Кочеткова А. В. Медиапланирование Первоисточник: М.: РИП-холдинг

ДОДАТОК А

Таблиця А.1 – опис змінних, які використовуються в коді програмування

Змінні	Опис	Категорія / Бренд / Продукт
ValueCategory	Об'єм продажів в гривнях	Категорія
PackCategory	Кількість упаковок	Категорія
ValueCardiomagnil_75_100	Об'єм продажів в гривнях	Кардіомагніл 75 мг, 100 таб
ValueMagnicor_75_100	Об'єм продажів в гривнях	Магнікор 75 мг, 100 таб
ValueMagnicor_75_30	Об'єм продажів в гривнях	Магнікор 75 мг, 30 таб
ValueCardiomagnyl_Forte_150_100	Об'єм продажів в гривнях	Кардіомагніл Форте 150 мг, 100 таб
ValueAspirin_100_98	Об'єм продажів в гривнях	Аспірин 100 мг, 98 таб
ValueCardiomagnyl_75_30	Об'єм продажів в гривнях	Кардіомагніл_75_30
ValueAspirin_100_56	Об'єм продажів в гривнях	Аспірин 100 мг, 56 таб
ValueLowspirin_75_120	Об'єм продажів в гривнях	Ловспірин 75 мг, 120 таб
ValueAspirin_100_28	Об'єм продажів в гривнях	Аспірин 100 мг, 28 таб
ValueMagnicor_150_100	Об'єм продажів в гривнях	Магнікор 150 мг, 100 таб
ValueLowspirin_75_30	Об'єм продажів в гривнях	Ловспірин 75 мг, 30 таб
ValueAcard_75_100	Об'єм продажів в гривнях	Акар 75 мг, 100 таб
ValueAsl_75_30	Об'єм продажів в гривнях	АСЛ 75 мг, 30 таб
ValueAcekor_100_100	Об'єм продажів в гривнях	Ацекор 100 мг, 100 таб
ValueCardio_Dar_75_100	Об'єм продажів в гривнях	Кардіо Дар 75 мг, 100 таб

Продовження таблиці А.1 – опис змінних, які використовуються в коді програмування

Змінні	Опис	Категорія / Бренд / Продукт
ValueCardiomagnyl_Forte_75_50	Об'єм продажів в гривнях	Кардіомагніл Форте 75 мг, 50 таб
ValueAcard_75_50	Об'єм продажів в гривнях	Акар 75 мг, 50 таб
ValueLowspirin_75_80	Об'єм продажів в гривнях	Ловспірин 75 мг, 80 таб
ValueAsk_100_30	Об'єм продажів в гривнях	АСК 100 мг, 30 таб
ValueAcekor_100_50	Об'єм продажів в гривнях	Ацекор 100 мг, 50 таб
ValueMagnicor_150_30	Об'єм продажів в гривнях	Магнікор 150 мг, 30 таб
ValueCardio_Dar_150_100	Об'єм продажів в гривнях	Кардіо Дар 150 мг, 100 таб
ValueTrombolik_100_20	Об'єм продажів в гривнях	Тромболік 100 мг, 20 таб
ValueCardisave_75_50	Об'єм продажів в гривнях	Кардісейв 75 мг, 50 таб
ValueAsaphen_80_90	Об'єм продажів в гривнях	Асафен 80 мг, 90 таб
ValueCardisave_75_30	Об'єм продажів в гривнях	Кардісейв 75 мг, 30 таб
ValueThrombo_100_30	Об'єм продажів в гривнях	Тромбо 100 мг, 30 таб
ValueAspicard_100_20	Об'єм продажів в гривнях	Аспікард 100 мг, 20 таб
ValueAspirin_300_28	Об'єм продажів в гривнях	Аспірин 300 мг, 28 таб
ValueAsaphen_80_30	Об'єм продажів в гривнях	Асафен 80 мг, 30 таб
CostCardiomgnyl_75_100	Середня ціна бренду	Кардіомагніл 75 мг, 100 таб
CostMagnicor_75_100	Середня ціна бренду	Магнікор 75 мг, 100 таб
CostMagnicor_75_30	Середня ціна бренду	Магнікор 75 мг, 30 таб

Продовження таблиці А.1 – опис змінних, які використовуються в коді програмування

Змінні	Опис	Категорія / Бренд / Продукт
CostCardiomagnyl_Forte_150_100	Середня ціна бренду	Кардіомагніл Форте 150 мг, 100 таб
CostAspirin_100_98	Середня ціна бренду	Аспірин 100 мг, 98 таб
CostCardiomagnyl_75_30	Середня ціна бренду	Кардіомагніл_75_30
CostAspirin_100_56	Середня ціна бренду	Аспірин 100 мг, 56 таб
CostLowspirin_75_120	Середня ціна бренду	Ловспірин 75 мг, 120 таб
CostAspirin_100_28	Середня ціна бренду	Аспірин 100 мг, 28 таб
CostMagnicor_150_100	Середня ціна бренду	Магнікор 150 мг, 100 таб
CostLowspirin_75_30	Середня ціна бренду	Ловспірин 75 мг, 30 таб
CostAcard_75_100	Середня ціна бренду	Акар 75 мг, 100 таб
CostAsl_75_30	Середня ціна бренду	АСЛ 75 мг, 30 таб
CostAcekor_100_100	Середня ціна бренду	Ацекор 100 мг, 100 таб
CostCardio_Dar_75_100	Середня ціна бренду	Кардіо Дар 75 мг, 100 таб
CostCardiomagnyl_Forte_75_50	Середня ціна бренду	Кардіомагніл Форте 75 мг, 50 таб
CostAcard_75_50	Середня ціна бренду	Акар 75 мг, 50 таб
CostLowspirin_75_80	Середня ціна бренду	Ловспірин 75 мг, 80 таб
CostAsk_100_30	Середня ціна бренду	АСК 100 мг, 30 таб
CostAcekor_100_50	Середня ціна бренду	Ацекор 100 мг, 50 таб
CostMagnicor_150_30	Середня ціна бренду	Магнікор 150 мг, 30 таб

Продовження таблиці А.1 – опис змінних, які використовуються в коді програмування

Змінні	Опис	Категорія / Бренд / Продукт
CostCardio_Dar_150_100	Середня ціна бренду	Кардіо Дар 150 мг, 100 таб
CostTrombolik_100_20	Середня ціна бренду	Тромболік 100 мг, 20 таб
CostCardisave_75_50	Середня ціна бренду	Кардісейв 75 мг, 50 таб
CostAsaphen_80_90	Середня ціна бренду	Асафен 80 мг, 90 таб
CostCardisave_75_30	Середня ціна бренду	Кардісейв 75 мг, 30 таб
CostThrombo_100_30	Середня ціна бренду	Тромбо 100 мг, 30 таб
CostAspicard_100_20	Середня ціна бренду	Аспікард 100 мг, 20 таб
PenCardiomgnyl_75_100	Пенетрація бренду	Кардіомагніл 75 мг, 100 таб
PenMagnicor_75_100	Пенетрація бренду	Магнікор 75 мг, 100 таб
PenMagnicor_75_30	Пенетрація бренду	Магнікор 75 мг, 30 таб
PenCardiomagnyl_Forte_150_100	Пенетрація бренду	Кардіомагніл Форте 150 мг, 100 таб
PenAspirin_100_98	Пенетрація бренду	Аспірин 100 мг, 98 таб
PenCardiomagnyl_75_30	Пенетрація бренду	Кардіомагніл_75_30
PenAspirin_100_56	Пенетрація бренду	Аспірин 100 мг, 56 таб
PenLowspirin_75_120	Пенетрація бренду	Ловспірин 75 мг, 120 таб
PenAspirin_100_28	Пенетрація бренду	Аспірин 100 мг, 28 таб
PenMagnicor_150_100	Пенетрація бренду	Магнікор 150 мг, 100 таб
PenLowspirin_75_30	Пенетрація бренду	Ловспірин 75 мг, 30 таб
PenAcard_75_100	Пенетрація бренду	Акар 75 мг, 100 таб
PenAsl_75_30	Пенетрація бренду	АСЛ 75 мг, 30 таб

Продовження таблиці А.1 – опис змінних, які використовуються в коді програмування

Змінні	Опис	Категорія / Бренд / Продукт
PenAcekor_100_100	Пенетрація бренду	Ацекор 100 мг, 100 таб
PenCardio_Dar_75_100	Пенетрація бренду	Кардіо Дар 75 мг, 100 таб
PenCardiomagnyl_Forte_75_50	Пенетрація бренду	Кардіомагніл Форте 75 мг, 50 таб
PenAcard_75_50	Пенетрація бренду	Акар 75 мг, 50 таб
PenLowspirin_75_80	Пенетрація бренду	Ловспірин 75 мг, 80 таб
PenAsk_100_30	Пенетрація бренду	АСК 100 мг, 30 таб
PenMagnicor_150_30	Пенетрація бренду	Магнікор 150 мг, 30 таб
PenCardio_Dar_150_100	Пенетрація бренду	Кардіо Дар 150 мг, 100 таб
PenTrombolik_100_20	Пенетрація бренду	Тромболік 100 мг, 20 таб
PenCardisave_75_50	Пенетрація бренду	Кардісейв 75 мг, 50 таб
PenAsaphen_80_90	Пенетрація бренду	Асафен 80 мг, 90 таб
PenCardisave_75_30	Пенетрація бренду	Кардісейв 75 мг, 30 таб
PenThrombo_100_30	Пенетрація бренду	Тромбо 100 мг, 30 таб
PenAspicard_100_20	Пенетрація бренду	Аспікард 100 мг, 20 таб
PenAspirin_300_28	Пенетрація бренду	Аспірин 300 мг, 28 таб
PenAsaphen_80_30	Пенетрація бренду	Асафен 80 мг, 30 таб
TVCardiomagnyl	Витрати на ТВ рекламу	Кардіомагніл
TVMagnicor	Витрати на ТВ рекламу	Магнікор
TVAsk	Витрати на ТВ рекламу	АСК

Продовження таблиці А.1 – опис змінних, які використовуються в коді програмування

Змінні	Опис	Категорія / Бренд / Продукт
TVAspirin	Витрати на ТВ рекламу	Аспірін
TVLowspirin	Витрати на ТВ рекламу	Лоуспірін
DigitalMagnicor	Витрати на діджитал рекламу	Магнікор
DigitalCardiomagnyl	Витрати на діджитал рекламу	Кардіомагніл
DigitalAspirin	Витрати на діджитал рекламу	Аспірін
DigitalAcekor	Витрати на діджитал рекламу	Асекор
DigitalAsk0Teva	Витрати на діджитал рекламу	Аск
RadioCardiomagnyl	Витрати на радіо рекламу	Кардіомагніл
Log_ValueCategory	Логарифм об'єму продажів в гривнях	Категорія
Log_ValueMagnicor_75_100	Логарифм об'єму продажів в гривнях	Магнікор 75 мг, 100 таб
Log_TVCardiomagnyl	Логарифм телевізійної реклами	Кардіомагніл
Log_TVMagnicor	Логарифм телевізійної реклами	Магнікор
Log_RadioCardiomagnyl	Логарифм телевізійної реклами	Кардіомагніл
Log_DigitalMagnicor	Логарифм телевізійної реклами	Магнікор
Log_DigitalCardiomagnyl	Логарифм телевізійної реклами	Кардіомагніл

Продовження таблиці А.1 – опис змінних, які використовуються в коді програмування

Змінні	Опис	Категорія / Бренд / Продукт
Log_DigitalAspirin	Логарифм телевізійної реклами	Аспірін
sqrt_ValueCardiomagnil_75_100	Корінь квадратний	Кардіомагніл 75 мг, 100 таб



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

ДИПЛОМНА РОБОТА
на ступінь вищої освіти магістр
із спеціальності 122 Комп'ютерні науки

Моделі планування рекламних інвестицій на основі
інформаційних технологій

Виконав: студент 6 курсу, групи КНДМ-61

Шерстюк Віктор Вячеславович

Керівник: Чичур Андрій, доктор
філософії.

Київ - 2024

Економетричне моделювання рекламних інвестицій фармацевтичного бренду

Дослідження присвячене розробці економетричної моделі для оптимізації рекламних інвестицій бренду Кардіомагніл. Модель враховує вплив телевізійної, радіо та цифрової реклами на продажі, що дозволяє сформувати комплексний підхід до управління маркетинговим бюджетом.





Актуальність дослідження

- 1** Висока конкуренція
Фармацевтичний ринок характеризується високим рівнем конкуренції, що вимагає ефективного розподілу рекламних бюджетів
- 2** Зростання обсягів даних
Постійне збільшення доступних даних відкриває нові можливості для оптимізації маркетингових процесів
- 3** Розвиток технологій
Нові інформаційні технології дозволяють точніше прогнозувати результати рекламних кампаній

Мета дослідження

Основна мета

Розробка економетричної моделі для оптимізації рекламного бюджету бренду Кардіомагніл з метою досягнення цільових показників продажів

Завдання

- Аналіз впливу різних рекламних каналів на продажі
- Оцінка ефективності інвестицій у рекламу
- Розробка рекомендацій щодо оптимального розподілу бюджету

Ключові фактори впливу на продажі



Рекомендації щодо оптимізації рекламних інвестицій



Телебачення

Збільшити інвестиції у телевізійні кампанії, особливо під час пікових сезонів



Радіо

Використовувати для підтримки частоти контактів та доповнення телевізійних кампаній



Цифрові медіа

Розширити цифрові активності, впровадити програматик-кампанії та інтерактивний контент

Довгострокові перспективи

Оцінка довгострокового впливу

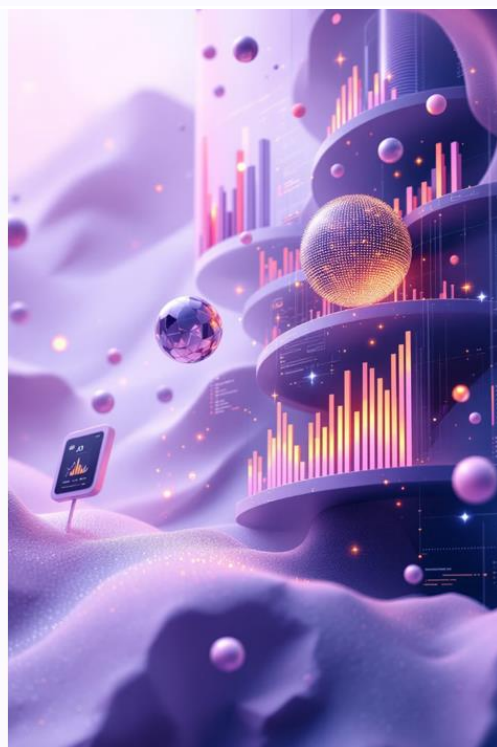
Вивчення тенденцій продажів протягом кількох років дозволить оцінити, які канали забезпечують сталий ефект, а які підходять для швидкого досягнення результатів

Адаптація до змін ринку

Регулярний перегляд та оновлення моделі дозволить враховувати зміни у поведінці споживачів та нові тенденції у рекламі

Виклики та подальші дослідження

- | | |
|--|--|
| <p>1 Оцінка конкурентного середовища</p> <p>Врахування дій конкурентів та їх вплив на ефективність рекламних кампаній</p> | <p>2 Адаптація до нових технологій</p> <p>Інтеграція інструментів штучного інтелекту для аналізу поведінки споживачів</p> |
| <p>3 Вплив макроекономічних факторів</p> <p>Включення змінних, що відображають економічні кризи, пандемії тощо</p> | |





Практичне значення дослідження

Оптимізація бюджету

Модель дозволяє ефективно розподіляти рекламний бюджет між різними каналами

Підвищення ROI

Точніше прогнозування ефективності рекламних кампаній сприяє підвищенню рентабельності інвестицій

Стратегічне планування

Результати дослідження можуть бути використані для довгострокового планування маркетингових стратегій

Висновки

1 Ефективність моделі

Розроблена економетрична модель довела свою ефективність у плануванні рекламних кампаній для фармацевтичного бренду

2 Оптимізація каналів

Телевізійна реклама залишається найпотужнішим каналом, але цифрові медіа мають значний потенціал росту

3 Майбутні перспективи

Подальше вдосконалення моделі дозволить підвищити точність прогнозів та адаптуватися до змін ринку



Дякуємо за увагу!

Це дослідження демонструє потенціал економетричного моделювання для оптимізації рекламних інвестицій у фармацевтичній галузі. Застосування таких моделей дозволяє підвищити ефективність маркетингових кампаній та забезпечити стійке зростання брендів на конкурентному ринку.



Протокол аналізу звіту подібності науковим керівником

Заявляю, що я ознайомився (-лась) з Повним звітом подібності, який був згенерований Системою виявлення і запобігання плагіату щодо роботи:

Автор: Віктор ШЕРСТЮК

Назва: Моделі планування рекламних інвестицій на основі інформаційних технологій

Координатор: Андрій ЧИЧУР

Підрозділ: ННІІТ

Коефіцієнт подібності 1:0.7

Коефіцієнт подібності 2:0

Тривога: 25

Після аналізу Звіту подібності констатую наступне:

☐ виявлені в роботі запозичення є сумлінними і не мають ознак плагіату. Тому робота визнається самостійною і допускається до захисту;

☐ виявлені в роботі запозичення не мають ознак плагіату, але їх надмірна кількість викликає сумніви щодо цінності роботи і самостійності її автора. Роботу направити на доопрацювання;

☐ виявлені в роботі запозичення є недобросовісними і мають ознаки плагіату або в ній містяться навмисні спотворення тексту, що вказують на спроби приховування недобросовісних запозичень. У зв'язку з чим, робота не допускається до захисту.

.....

.....

Дата

Підпис Наукового керівн