

Вінницький національний технічний університет
Факультет інтелектуальних інформаційних технологій та автоматизації
Кафедра автоматизації та інтелектуальних інформаційних технологій

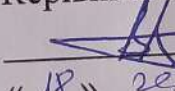
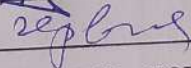
МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему:

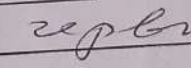
«Система автоматизації прогнозування продажів»

Виконав: студент 2 курсу, групи АКІТ-22мз
спеціальність 151 – «Автоматизація та
комп'ютерно-інтегровані технології»

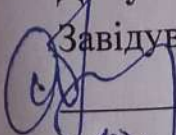
 Марія ВОЙТЮК
Керівник: к.т.н., доц. каф. АІТ

 Владислав КАБАЧІЙ
« 18 »  2024 р.

Опонент: к.т.н., доц. каф. КН

 Володимир ОЗЕРАНСЬКИЙ
« 18 »  2024 р.

Допущено до захисту

 Завідувач кафедри АІТ

д.т.н., проф. Олег БІСІКАЛО
« 19 » 06 2024 р.

Вінниця ВНТУ – 2024 рік

Вінницький національний технічний університет
 Факультет інтелектуальних інформаційних технологій та автоматизації
 Кафедра автоматизації та інтелектуальних інформаційних технологій
 Рівень вищої освіти – другий (магістерський)
 Галузь знань – 15 Автоматизація та приладобудування
 (шифр і назва)
 Спеціальність 151 – «Автоматизація та комп'ютерно-інтегровані технології»
 (шифр і назва спеціальності)
 Освітня програма – Інтелектуальні комп'ютерні системи
 (назва освітньо-професійної програми)

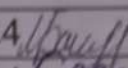
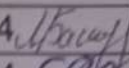
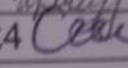
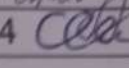
ЗАТВЕРДЖУЮ
 Завідувач кафедри АІТ, проф.
 Олег БІСІКАЛО
 «12» березня 2024 року

ЗАВДАННЯ НА МАГІСТЕРСЬКУ КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ Войтюк Марії Володимирівні

1. Тема роботи: «Система автоматизації прогнозування продажів»
 керівник роботи: Владислав КАБАЧІЙ, к.т.н., доц. каф. АІТ
 затверджені наказом ВНТУ від «11» березня 2024 року №81
2. Термін подання студентом роботи «13» червня 2024 року
3. Вихідні дані до роботи:
 Набір даних про Інтернет-магазин Google за період з 01.01.2016 по 31.01.2019.
4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)
 - 1) Аналіз підходів для автоматизації прогнозування продажів;
 - 2) Дослідження математичних методів для побудови систем автоматизації прогнозування продажів;
 - 3) Розробка системи автоматизації прогнозування продажів;
 - 4) Тестування системи автоматизації прогнозування продажів інтернет-магазину
5. Перелік ілюстративного матеріалу (з точним зазначенням обов'язкових креслень):
 - 1) Таблиця з найбільшими показниками;
 - 2) Кінцевий звіт з деревами декомпозиції, діаграми з областями, тощо;
 - 3) Звіт з діаграмою дерева і результатами розрахунків за допомогою карточки;

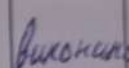
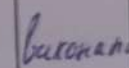
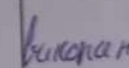
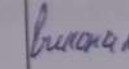
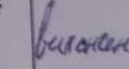
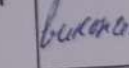
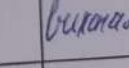
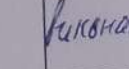
4) Розроблення системи автоматизації прогнозування продажів за допомогою аналітики в Power BI

6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	Завдання прийняв
1-4	Марія БАРАБАН	12.03.2024 	17.05.2024 
5	Володимир КОЗЛОВСЬКИЙ	12.03.2024 	27.05.2024 

7. Дата видачі завдання «12» березня 2024 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів магістерської дипломної роботи	Термін виконання	При мітка
1	Аналіз підходів для автоматизації прогнозування продажів	12.03.2024 – 26.03.2024	
2	Дослідження математичних методів для побудови систем автоматизації прогнозування продажів	27.03.2024 – 15.04.2024	
3	Розробка системи автоматизації прогнозування продажів	16.04.2024 – 02.05.2024	
4	Тестування системи автоматизації прогнозування продажів інтернет-магазину	03.05.2024 – 17.05.2024	
5	Економічний розділ	20.05.2024 – 27.05.2024	
6	Оформлення пояснювальної записки	28.05.2024 – 05.06.2024	
7	Попередній захист роботи	06.06.2024	
8	Остаточний захист роботи	20.06.2024	

Студент

Керівник роботи



Марія ВОЙТЮК
Владислав КАБАЧІЙ

АНОТАЦІЯ

Войтюк М.В. Система автоматизації прогнозування продажів. Магістерська кваліфікаційна робота зі спеціальності 151 – Автоматизація та комп'ютерно-інтегровані технології, освітня програма – Інтелектуальні комп'ютерні системи. Вінниця: ВНТУ, 2024. 85 с.

На укр.мові. Бібліогр.: 30 назва; рис.:33; табл.: 8.

Метою дослідження є покращення процесу прогнозування продажів Інтернет-магазину Google шляхом розробки та впровадження системи автоматизованого прогнозування, що базується на аналізі поведінки клієнтів та їх відповідних дій.

У даній роботі розглянуто ключові аспекти автоматизації прогнозування продажів та досліджено математичні методи, що застосовуються для побудови відповідних систем.

Головною цілю є розробка системи автоматизації за допомогою створення звітів в Microsoft Power BI для кращої візуалізації даних магазинів Google, також досліджено процес вибору оптимальних параметрів для точного автоматизованого прогнозування.

Ключові слова: автоматизація; прогнозування продажів; математичні методи; Microsoft Power BI.

ABSTRACT

Voitiuk M.V. Sales forecasting automation system. Master's thesis on specialty 151 - Automation and computer-integrated technologies, educational program - Intelligent computer systems. Vinnytsia: VNTU, 2024. 85 p.

In the Ukrainian language. Bibliogr.: 30 title; Fig.:33; tab.: 8.

The aim of the thesis is to develop a system for automating sales forecasting for Google's online store by utilizing an automated forecasting system through the analysis of customer behavior in the store and their corresponding actions.

This thesis examines the key aspects of automating sales forecasting and investigates the mathematical methods used to build such systems.

The main goal is to develop an automation system by creating reports in Microsoft Power BI for better visualization of Google Store data, and to study the process of selecting optimal parameters for accurate automated forecasting.

Keywords: automation; sales forecasting; mathematical methods; Microsoft Power BI.

ЗМІСТ

ВСТУП	8
1 АНАЛІЗ ПІДХОДІВ ДЛЯ АВТОМАТИЗАЦІЇ ПРОГНОЗУВАННЯ ПРОДАЖІВ.....	10
1.1 SFA-системи.....	11
1.2 SFMS-системи.....	13
1.3 CBT Demand Forecasting	15
1.4 Streamline	17
1.5 Висновки до розділу	19
2 ДОСЛІДЖЕННЯ МАТЕМАТИЧНИХ МЕТОДІВ ДЛЯ ПОБУДОВИ СИСТЕМ АВТОМАТИЗАЦІЇ ПРОГНОЗУВАННЯ ПРОДАЖІВ.....	20
2.1 Аналіз часових рядів.....	20
2.2 Модель авторегресії	21
2.3 Модель ковзного середнього	23
2.4 Метод декомпозиції	25
2.5 Висновки до розділу	26
3 РОЗРОБКА СИСТЕМИ АВТОМАТИЗАЦІЇ ПРОГНОЗУВАННЯ ПРОДАЖІВ.....	27
3.1 Вибір оптимальних параметрів для прогнозування продажів	27
3.2 Розробка системи автоматизації продажів	29
3.3 Роботи з набором даних.....	35
3.4 Вибір програми для подальшої візуалізації	37
3.5 Висновки до розділу	38
4 ТЕСТУВАННЯ СИСТЕМИ АВТОМАТИЗАЦІЇ ПРОГНОЗУВАННЯ ПРОДАЖІВ ІНТЕРНЕТ-МАГАЗИНУ	39
4.1 Пошук та імпорт даних.....	39
4.2 Побудова дашбордів та використання функцій DAX	41
4.3 Кінцевий дашборд.....	53
4.4 Висновки до розділу	54
5 ЕКОНОМІЧНИЙ РОЗДІЛ.....	56
5.1 Технологічний аудит розробленої системи автоматизації прогнозування продажів	56

5.2 Розрахунок витрат на розроблення системи автоматизації прогнозування обсягів продажів товарів в Інтернет-магазинах	61
5.3 Розрахунок економічного ефекту від можливої комерціалізації нашої розробки	65
ВИСНОВКИ.....	73
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	75
Додаток А (обов'язковий) Технічне завдання	78
Додаток Б (обов'язковий) Ілюстративна частина	81
Додаток В (обов'язковий) Протокол перевірки МКР	85

ВСТУП

Актуальність теми. Актуальність системи автоматизації прогнозування продажів обумовлена кількома факторами. Економічна криза спонукає продавців розвивати онлайн-комерцію. Традиційні безрецептурні продажі коштують дорого через високу орендну плату, витрати на персонал та інші витрати, а також через зміни в поведінці споживачів: все більше українців шукають дешевші товари в Інтернеті перед покупкою. Крім того, спостерігається тенденція згортання або значного скорочення офлайн-майданчиків. Однак більшість ритейлерів все ще намагаються побудувати багатоканальний механізм продажу. Тому розробка системи автоматизації прогнозування продажів є надзвичайно актуальною.

Мета і задачі дослідження. Метою дослідження є покращення процесу прогнозування продажів Інтернет-магазину Google шляхом розробки та впровадження системи автоматизованого прогнозування, що базується на аналізі поведінки клієнтів та їх відповідних дій.

Для досягнення мети необхідно розв'язати такі задачі:

1. проаналізувати підходи для автоматизації прогнозування продажів;
2. дослідити математичні методи для побудови систем автоматизації прогнозування продажів;
3. розробити систему автоматизації прогнозування продажів;
4. протестувати систему автоматизації прогнозування продажів інтернет-магазину.

Об'єктом дослідження є процес автоматизованого прогнозування продажів та оформлення дашбордів за допомогою Microsoft Power BI.

Предметом дослідження є моделі, методи та системи автоматизації прогнозування, на основі яких будуть прийматися рішення стосовно прогнозування продаж в Інтернет-магазині.

Методи дослідження. включають методи статистичного аналізу, а також методи аналізу даних із використанням спеціалізованих програмних засобів, таких як Microsoft Power BI. Статистичний аналіз дозволить провести обґрунтований огляд підходів для автоматизації прогнозування продажів та дослідити математичні методи, що використовуються в цьому процесі. Аналіз даних у програмі Microsoft Power BI допоможе розробити систему автоматизації прогнозування продажів, оцінити її ефективність та впровадити необхідні корективи.

Науково технічний результат роботи полягає у наступному: отримати подальший розвиток методу прогнозування, в основі якого лежить аналіз поведінки клієнтів в магазині та відповідних його дій, що призводить до покращення системи автоматизації прогнозування майбутніх продажів Інтернет-магазину Google.

Практичне значення результатів роботи полягає в розробці системи автоматизації прогнозування продажів Інтернет-магазину Google за рахунок аналізу поведінки клієнтів в магазині та відповідних його дій.

Апробація результатів та публікації. За результати даної роботи опубліковано тези доповіді [1] на конференції «Молодь в науці: дослідження, проблеми, перспективи» факультету інтелектуальних інформаційних технологій та автоматизації Вінницького національного технічного університету (2024).

1 АНАЛІЗ ПІДХОДІВ ДЛЯ АВТОМАТИЗАЦІЇ ПРОГНОЗУВАННЯ ПРОДАЖІВ

В сучасному світі ефективне управління продажами є критично важливим для успіху бізнесу. Забезпечення стабільних і прогнозованих обсягів продажів дозволяє компаніям не лише забезпечити стабільний дохід, а й ефективно планувати виробництво, поставки, а також розробляти маркетингові та стратегічні плани. Однак, навіть при наявності великої кількості інформації про минулі продажі, прогнозування майбутніх обсягів може бути складною задачею через низку факторів, таких як сезонність, випадковість та змінність попиту. У зв'язку з цим, розвиток автоматизованих методів та моделей для прогнозування продажів стає все більш актуальним завданням для бізнесу.

Автоматизація прогнозування продажів може забезпечити значні переваги для компаній, зокрема, зниження людського втручання, підвищення точності прогнозів, зменшення часу, необхідного для аналізу даних, та покращення стратегічного планування. Для цього існують різні підходи та методи, які варто розглянути і порівняти з метою вибору оптимального рішення для конкретного бізнесу. Серед таких підходів можна виділити традиційні статистичні методи, методи машинного навчання, а також гібридні підходи, що комбінують у собі кращі аспекти обох підходів, дозволяючи отримати більш точні та інформативні результати.

В даний розділ буде проведено аналіз різних підходів для автоматизації прогнозування продажів з метою з'ясування їхніх переваг та обмежень, а також вибору найбільш підходящого методу для подальшого дослідження.

1.1 SFA-системи

Для автоматизації рутинних завдань у сфері управління продажами створено спеціальні системи, відомі як SFA (Sales Force Automation) або SFMS (Sales Force Management System). Ці системи автоматизації продажів представляють собою інформаційні платформи, що використовуються для керування відносинами з клієнтами (CRM) та маркетингового менеджменту, а також для автоматизації процесів продажів і управлінських функцій. Основною метою програмного забезпечення SFA є надання потрібних даних потрібним людям у потрібний час і скорочення кількості адміністративних завдань, які продавці та їхні менеджери повинні виконувати вручну.

Системи автоматизації продажів варіюються за своїми можливостями. Їх вибір залежить від потреб організації, розміру компанії, її структури, попиту на нову систему, процесів продажів та кількості користувачів. SFA-системи можуть надавати послуги у двох основних категоріях:

- On-premises – програмне забезпечення, яке встановлюється та запускається на комп'ютерах в офісі користувача або організації, що використовує цю систему, а не на віддалених серверах чи в хмарі.

- On-demand – програмне забезпечення на запит, також відоме як Software as a Service (SaaS), тобто модель ліцензування і доставки програмного забезпечення, де воно ліцензується за підпискою і централізовано розміщується. Робота в системі організована на принципах:

- збереження всієї інформації в центральній базі даних;
- використання КПК мобільними працівниками для повсякденної роботи;
- здійснення обміну даними між КПК і центральною базою даних (БД) в рамках синхронізації через глобальну мережу Інтернет;

– оперативного доступу аналітиків і менеджерів до інформації, що зберігається в центральній базі даних, як на стаціонарних робочих місцях, так і через мобільну версію.

У SFA-системах передбачено два типи робочих місць:

1. Стаціонарне робоче місце – це клієнтське програмне забезпечення встановлюється на робочу станцію користувача і взаємодіє з сервером центральної бази даних безпосередньо в межах локальної мережі або через веб-доступ, при цьому обмін даними здійснюється за протоколом HTTP або HTTPS.

2. Мобільне робоче місце – клієнтське програмне забезпечення встановлюється на КПК співробітника, а обмін даними між КПК та центральною базою даних здійснюється за командою користувача через глобальну мережу Інтернет за протоколом HTTP або HTTPS.

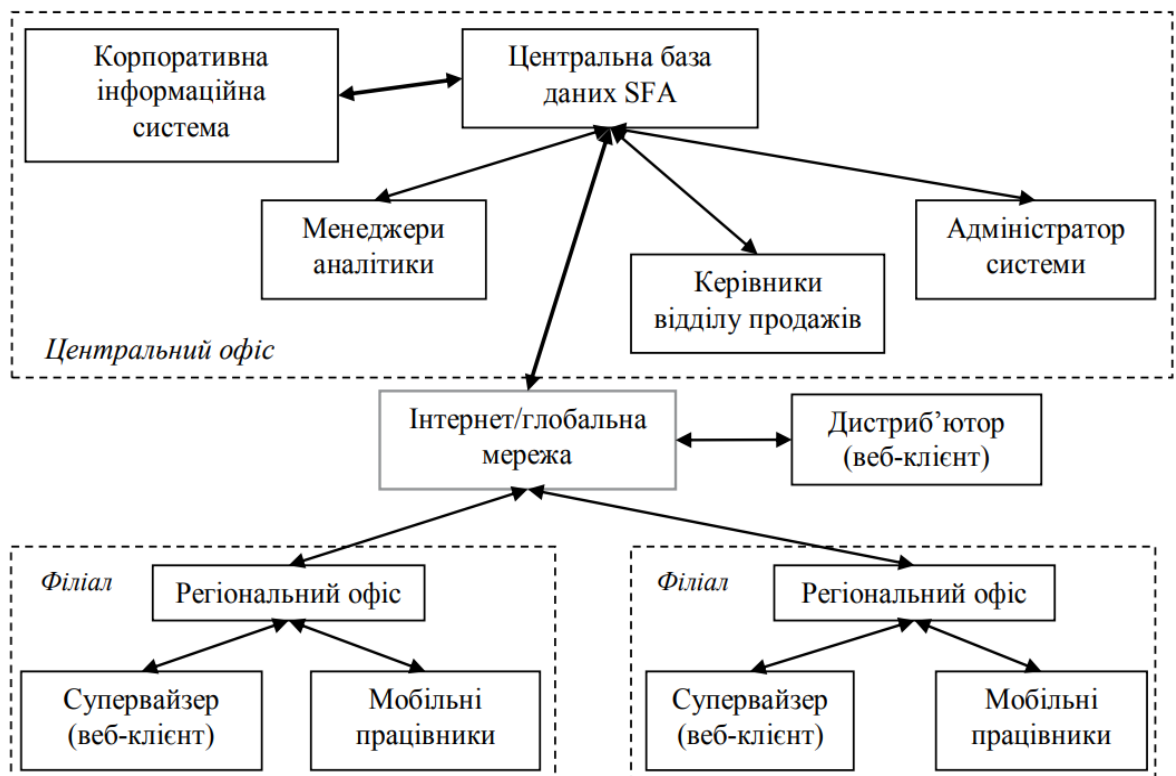


Рисунок 1.1 – Логічна централізована архітектура SFA-системи

На рисунку 1.1 представлена логічна архітектура системи SFA, призначеної для управління збутом компанії регіонального або національного масштабу. Ця архітектура є централізованою, а її основні завдання включають розміщення єдиного модуля SFA в логічному ядрі системи та пряму взаємодію з ним усіх КПК і облікових систем філій.

Дослідження виявило кілька переваг, які отримує підприємство, використовуючи SFA-системи:

- підвищення мобільності продавців;
- покращення взаємодії всередині команд;
- ефективніше використання робочого часу;
- покращення операційної ефективності;
- зміцнення відносин з клієнтами;
- підвищення ефективності та продуктивності працівників відділу продажів;
- можливість оцінювати зовнішнє середовище, займатися прогнозуванням та стратегічним плануванням.

1.2 SFMS-системи

SFMS – це система, призначена для підвищення прибутковості та ефективності процесу продажів. Вона автоматично аналізує дані за допомогою складних статистичних методів і представляє результати в зручному для користувача вигляді. Система генерує таблиці, діаграми та графіки, що дозволяють менеджерам оперативно відстежувати та реагувати на зміни. SFMS сприяє проведенню маркетингових досліджень, допомагаючи визначити лідерів ринку, тренди продажів та поведінку клієнтів. Крім того, система дозволяє координувати діяльність розподілених підрозділів і філій компанії.

SFA (автоматизація продажів) зазвичай є частиною CRM-системи компанії та записує всі етапи продажу, слугуючи першим кроком до впровадження повноцінної CRM-стратегії [2].

SFMS здійснює ряд важливих функцій та має переваги, які визначають її значущість у сфері запобігання та протидії фінансовим злочинам:

- збір і аналіз інформації: SFMS здійснює збір, обробку та аналіз фінансової інформації з метою виявлення потенційних випадків легалізації доходів, фінансування тероризму та інших фінансових злочинів.

- проведення ризик-оцінки: SFMS проводить аналіз національних ризиків, пов'язаних з фінансовими злочинами, для визначення пріоритетів та стратегій запобігання.

- міжнародна співпраця: SFMS співпрацює з міжнародними організаціями та іншими країнами з питань обміну інформацією та координації дій у боротьбі з фінансовими злочинами.

- впровадження технологій: SFMS використовує сучасні технології для виявлення та моніторингу фінансових транзакцій, що дозволяє підвищити ефективність боротьби зі злочинністю.

- співпраця з правоохоронними органами: SFMS співпрацює з правоохоронними та розвідувальними органами для розслідування випадків фінансових злочинів та надання допомоги у виявленні злочинців.

- забезпечення міжнародної представництва: SFMS представляє Україну в міжнародних організаціях, таких як FATF (Financial Action Task Force), з метою спільного розв'язання глобальних проблем фінансової злочинності.

Переваги SFMS полягають у її здатності ефективно виявляти, досліджувати та запобігати фінансовим злочинам, забезпечуючи таким чином безпеку фінансової системи та суспільства.

Служба фінансового моніторингу України (SFMS) відіграє ключову роль у запобіганні та боротьбі з фінансовими злочинами, збираючи, аналізуючи та обмінюючись інформацією про фінансові транзакції. Вона

також сприяє оцінці ризиків і розробці стратегій запобігання фінансовим злочинам. Завдяки співпраці з міжнародними партнерами та використанню сучасних технологій, SFMS ефективно протидіє злочинності та забезпечує безпеку фінансової системи України [3].

1.3 CBT Demand Forecasting

CBT Demand Forecasting ChatGPT є інструментом для швидкого узгодження та зручної візуалізації прогнозу попиту на продукцію за допомогою набору математичних моделей та швидкого коригування на основі наявних оперативних даних. Багато компаній оцінюють майбутні продажі, виходячи з минулого показника продажів, наприклад, за минулий місяць чи аналогічний період минулого року. Хоча цей підхід може бути ефективним для невеликих компаній або високорівневого стратегічного планування, проте для розширюючихся компаній з великою кількістю товарних позицій та підрозділів чи при деталізації планів на різних рівнях, необхідні програмні засоби. І справді, якщо у вашій компанії відсутній систематизований процес планування, ви все одно здійснюєте планування.

Точність результатів планування безпосередньо залежить від повноти інформації, факторів, що впливають, прозорості процесу та взаємодії між підрозділами тощо.

У дев'яти з десяти випадком при впровадженні бізнес-систем у клієнтів, ми стикаємося з кейсом запуску новинок. І перше, що ми чуємо від фахівців – це запит, який надходить від їх керівництва, такого змісту: «Як будуть продаватися ці новинки?».

Це буквально запит на графік продажів в майбутньому, в розрізі періодів та регіонів. В майбутньому – це значить планування. При вирішенні даного кейса з одного боку, нам доводиться стикатися з відсутністю історії факту продажу по новинці, і в даному випадком ми можемо підібрати схожі за

характеристиками товари або групу товарів, а якщо таких немає – динаміку продажів можна підглянути у конкурентів. Але потім, це все нам потрібно скомпонувати, розділити по регіонах і за періодами. А з іншого боку нам потрібно розуміти, який результат в зростанні продажів ми отримаємо, якщо проведемо акції на підтримку запуску за п'ять днів до запуску, в день запуску, через п'ять днів після запуску, через п'ятнадцять і т.д.

До цього додається ще кілька важливих аспектів. Потрібно враховувати сезонність, економічні фактори та поведінкові особливості клієнтів у кожному регіоні. Крім того, використання аналітичних інструментів та моделей прогнозування допомагає більш точно передбачити майбутні продажі.

Прогнозування продажів з урахуванням всіх необхідних факторів, що впливають для великого інформаційного потоку неможливо без автоматизованої системи. А якщо немає системи – немає і відповіді на питання: «Як будуть продаватися ці новинки?»

Можливості CBT DF:

1. Аналізувати фактичні продажі;
2. Очищати факт продажу від сплесків промоактивності та інших аномальних значень;
3. Розраховувати тренди за фактом продажу;
4. Розраховувати сезонність за фактом продажу;
5. Шукати залежність продажів до зміни будь-яких факторів, що впливають на ринку: ціни, курс, асортимент, акції, частки в мережах і т.д;
6. Перевіряти автокореляції і тісноту зв'язку між факторами;
7. Розраховувати регресійний аналіз;
8. Моделювати прогноз продажів з урахуванням факторів, що впливають;
9. Порівнювати моделі прогнозу;
10. Надавати рекомендації щодо промо-активностей [4];

1.4 Streamline

Streamline є найбільш швидкозростаючим у світі постачальником програмного забезпечення для прогнозування продажів Unleashed для підприємств середнього та корпоративного бізнесу.

Streamline зі штаб-квартирою в Нью-Йорку має понад 200 партнерів, тисячі середніх і корпоративних клієнтів по всьому світу, які покладаються на її платформу для свого бізнесу. Платформа на базі штучного інтелекту допомагає виробникам, роздрібним торговцям, оптовикам і дистриб'юторам ефективно розвиватися, збільшуючи таким чином їхні прибутки. Вигляд програмного забезпечення Streamline зображено на рисунку 1.2:

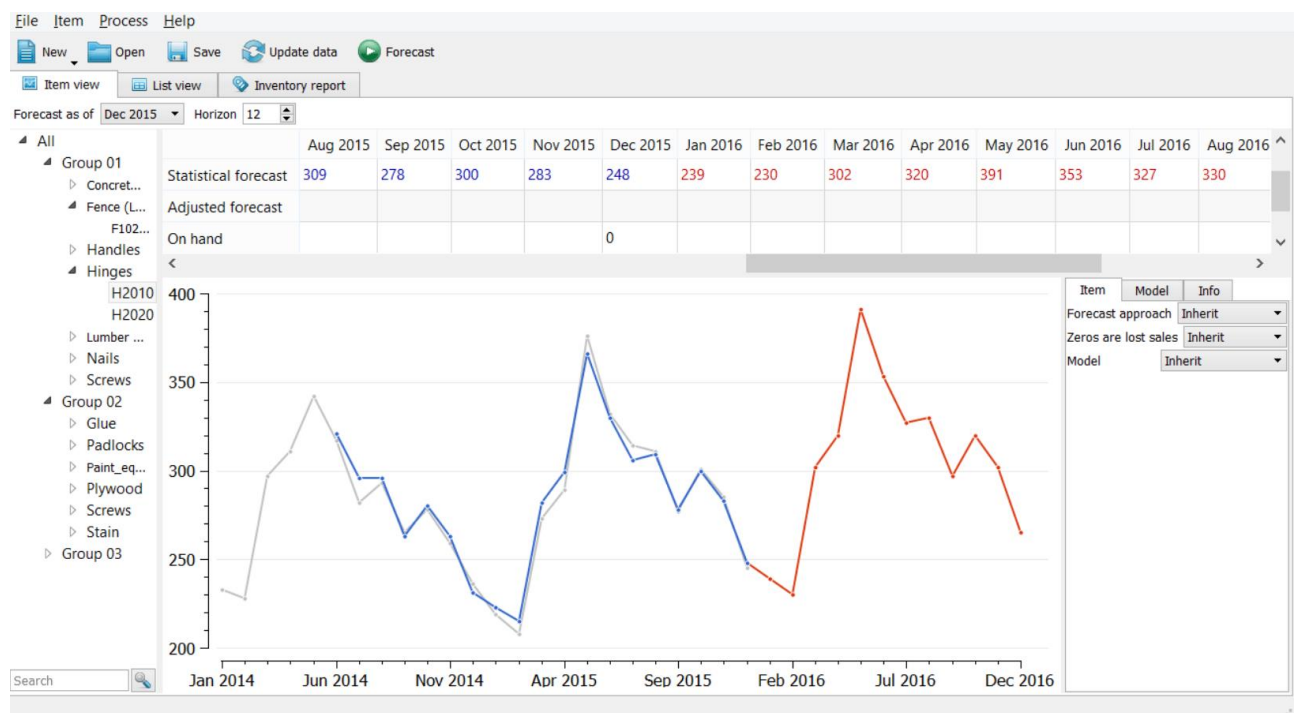


Рисунок 1.2 – Вигляд програмного забезпечення Streamline

Переваги рішення для прогнозування продажів Streamline:

- Швидкий та інтуїтивно зрозумілий користувацький інтерфейс
- Безшовна інтеграція джерел даних компанії
- Плавний та швидкий процес впровадження

- Ідеальне поєднання бізнес -процесу у вашій компанії
- Синхронізація дат замовлення між артикулами
- Заміна формул на моделювання дискретних подій
- Використання AI (штучного інтелекту) для прогнозування попиту
- Груповий EOQ (Економічний розмір замовлення) [5].

Power BI, або Power Business Intelligence, може бути розглянутий як частина стрімлайн-процесу в аналітиці даних та прийнятті рішень у бізнесі. Streamline означає оптимізацію та спрощення процесів з метою підвищення ефективності та зниження витрат часу та ресурсів. Power BI допомагає в цьому, забезпечуючи швидкий доступ до даних, їх візуалізацію та аналіз, що дозволяє бізнес-користувачам приймати обґрунтовані рішення на основі актуальної та достовірної інформації. Таким чином, Power BI допомагає підвищити ефективність та точність процесів прийняття рішень у бізнесі, що відповідає принципам стрімлайнінгу.

Power BI – це потужний інструмент для аналізу та візуалізації даних, розроблений корпорацією Microsoft. Він дозволяє користувачам легко збирати дані з різних джерел, таких як бази даних, файли Excel, хмарні сервіси (наприклад, Azure, Google Analytics) та інші, і створювати з них інтерактивні звіти та панелі керування.

Одним з основних переваг Power BI є його простота використання та можливість швидкої побудови візуалізацій. Користувачі можуть легко створювати різноманітні графіки, діаграми, карту, таблиці та інші типи візуалізацій, які допомагають краще зрозуміти дані та знаходити в них важливі зв'язки та залежності.

Крім того, Power BI пропонує різноманітні можливості для обробки даних, включаючи фільтрацію, сортування, групування, розрахунки та інші операції, що дозволяють отримувати більш детальні та корисні результати.

Ще однією важливою особливістю Power BI є його можливість інтеграції з іншими продуктами та сервісами Microsoft, такими як Office 365, SharePoint,

Teams, що робить його досить універсальним інструментом для аналізу даних в різних сферах бізнесу та галузях [6].

1.5 Висновки до розділу

У даному розділі було проведено аналіз різних підходів до автоматизації прогнозування продажів. Розглянуті такі системи, як SFA і SFMS, їхні переваги та недоліки. Також досліджено CBT Demand Forecasting та Streamline. Результати цього аналізу допоможуть вибрати оптимальний підхід для підвищення ефективності прогнозування продажів.

2 ДОСЛІДЖЕННЯ МАТЕМАТИЧНИХ МЕТОДІВ ДЛЯ ПОБУДОВИ СИСТЕМ АВТОМАТИЗАЦІЇ ПРОГНОЗУВАННЯ ПРОДАЖІВ

2.1 Аналіз часових рядів

Часовий ряд – це послідовність точок даних, впорядкованих у хронологічному порядку. Часто такі ряди складаються з даних, зібраних у рівновіддалені моменти часу. Таким чином, часові ряди є дискретними послідовностями. Прикладами часових рядів можуть бути висоти океанських припливів, кількість сонячних плям або щоденне середньозважене значення індексу ПФТС на момент закриття торгів.

Часові ряди часто зображуються за допомогою лінійних графіків, що дозволяє наочно спостерігати зміни даних у часі. Вони використовуються в багатьох сферах, включаючи статистику, обробку сигналів, розпізнавання образів, економетрику, фінансову математику, прогнозування погоди, розумний транспорт, передбачення траєкторій, землетрусів, електроенцефалографію, автоматичне керування, астрономію, технології зв'язку та інші галузі прикладної науки та інженерії, де здійснюються часові вимірювання.

Аналіз часових рядів включає методи дослідження даних для витягування значимих статистичних характеристик. Прогнозування часових рядів передбачає застосування моделей для передбачення майбутніх значень на основі раніше спостережених даних. Регресійний аналіз часто використовується для перевірки теорій про вплив поточних значень одного чи більше незалежних часових рядів на поточне значення іншого часового ряду. Однак, аналіз часових рядів зосереджується на порівнянні значень одного чи багатьох залежних часових рядів у різні моменти часу [7].

Дані часових рядів мають природну послідовність у часі, що відрізняє їх від поперечних досліджень, де порядок спостережень не має значення

(наприклад, при аналізі заробітної плати людей на основі їхнього рівня освіти, де дані можуть вводитися у будь-якому порядку). Аналіз часових рядів також відрізняється від аналізу просторових даних, де спостереження прив'язані до географічних розташувань (наприклад, аналіз цін на будинки в залежності від їх місця розташування та характеристик). У стохастичних моделях часових рядів зазвичай враховується, що спостереження, близькі в часі, мають сильніший зв'язок, ніж ті, що знаходяться на більшій відстані одне від одного. Також моделі часових рядів зазвичай передбачають односпрямоване впорядкування часу, де значення для даного періоду залежать від попередніх, а не від майбутніх значень.

Метод аналізу часових рядів визначається як цілями аналізу, так і ймовірнісною природою формування його значень. Найпоширеніші методи аналізу часових рядів можна розділити на три основні категорії:

- Аналіз на основі суджень: прогнозування, засноване на суб'єктивних оцінках, інтуїції, глибоких знаннях конкретної області та іншій інформації, що має відношення до прогнозованого процесу, тобто передбачення.

- Методи аналізу однієї змінної: включають авторегресію, авторегресію з ковзним середнім, авторегресію з ковзним середнім з додаванням трендової моделі, спектральний і кореляційний аналіз.

- Методи прогнозування на основі кількох змінних: використання часових рядів, які враховують декілька змінних.

Аналіз часових рядів є одним із найпопулярніших і ефективних підходів до прогнозування. У наступному підрозділі розглянемо кілька варіацій регресійних моделей, які можна використовувати для прогнозування продажів на підприємстві [8].

2.2 Модель авторегресії

Авторегресійна (АР) модель – це модель часових рядів, де кожне значення залежить лінійно від попередніх значень цього ж ряду.

У статистиці, економетриці та обробці сигналів авторегресійна модель представляє тип випадкового процесу, що використовується для опису змінних у часі процесів у природі, економіці та інших сферах. Вона визначає, що вихідна змінна лінійно залежить від своїх попередніх значень і стохастичного члена.

Авторегресійна модель є окремим випадком моделей авторегресії – ковзного середнього (ARMA) та авторегресійної інтегрованої ковзної середньої (ARIMA), що мають складнішу стохастичну структуру. Вона також є окремим випадком векторної авторегресійної моделі (VAR), що складається з системи пов'язаних стохастичних різницевих рівнянь у більш ніж одній змінній.

Авторегресійна модель, на відміну від моделі ковзного середнього, не завжди є стаціонарною, оскільки може містити одиничний корінь. Загальний вигляд авторегресійного процесу порядку p :

$$X_t = c + \sum_{i=1}^p a_i * X_{t-1} + \varepsilon_t,$$

де a_i – коефіцієнти авторегресії;

c – константа;

ε_t – білий шум або похибка моделі.

Модель AR(3) матиме наступний вигляд:

$$y(k) = a_0 + a_1 y(k-1) + a_2 y(k-2) + a_3 y(k-3).$$

Порядок використання авторегресійної складової визначається через ступінь автокореляції змінної. Значення автокореляційної функції $\gamma(s)$ обчислюється за наступною формулою:

$$r_y(i) = r_{y(k)y(k-i)} = \frac{1}{n-1} \frac{\sum_{k=i+1}^n \{[y(k) - \bar{y}][y(k-i) - \bar{y}]\}}{\sigma_y^2}$$

де n – розмір вибірки;

σ_y^2 – дисперсія змінних $y(k)$ на вибірці;

$i = 1, 2, 3, \dots$

Після отримання значень АКФ, вибирається граничне значення функції залежно від вимог до моделі, і визначаються лаги, для яких це значення перевищене. Кожен такий лаг буде відповідати значущому порядку АР складової в моделі.

Точніший порядок АР складової визначається за допомогою часткової автокореляційної функції (ЧАКФ), яка визначає прямий вплив одного значення ряду на інше, виключаючи кореляційний вплив проміжних значень. Подібно до АКФ, розмірність АР складової моделі може визначатися відповідно до ненульового коефіцієнта ЧАКФ з найбільшим лагом [9].

2.3 Модель ковзного середнього

Ковзне середнє (КС) – це набір значень, що відображають середнє значення ряду на певному проміжку часу. Використовується для відображення тенденцій процесу і згладжування його несуттєвих коливань. Основна ідея полягає у тому, щоб розрахувати середнє значення певного періоду, яке «ковзає» по даних у часі. Це дозволяє зменшити випадкові коливання та виділити загальний напрямок часового ряду.

Однією з основних моделей ковзної середньої є проста модель ковзної середньої (SMA), яка полягає у розрахунку середнього значення певного періоду шляхом усереднення значень спостережень за цей період. Вона допомагає згладити випадкові флуктуації та виявити тренди в даних. Це один

із найбільш базових методів аналізу часових рядів, що широко використовується для визначення загального напрямку руху даних.

Зважена модель ковзної середньої (Weighted Moving Average, WMA) враховує різну вагу окремих спостережень при обчисленні середнього значення. Ця модель є корисною, коли певні дані вважаються більш важливими або релевантними для прогнозування. Наприклад, більш нові дані можуть мати більшу вагу у порівнянні зі старішими даними, що дозволяє точніше реагувати на зміни у часових рядах.

Експоненціально згладжена ковзна середня (Exponential Moving Average, ЕМА) приділяє більше уваги останнім спостереженням порівняно зі старішими. Це робиться шляхом застосування експоненційного коефіцієнта, який поступово зменшує вагу старіших даних. Ця модель допомагає швидше реагувати на зміни у даних та виявляти тренди в реальному часі. ЕМА часто використовується в фінансовому аналізі для аналізу цінних паперів та біржових індексів.

Всі ці моделі є важливими інструментами для аналізу часових рядів та прогнозування майбутніх значень на основі минулих даних. Вони можуть бути застосовані в різних галузях, таких як фінанси, економіка, маркетинг, логістика та багато інших сфер діяльності, де важливо виявляти тренди та передбачати майбутні зміни [10].

Загальна формула для ковзного середнього:

$$MA(k) = \frac{\sum_{i=1}^N y(k - i + 1)}{N},$$

де N – розмір вікна КС ;

y – часовий ряд вхідних даних

Власне КС компонента моделі матиме вигляд:

$$y(k) = mv(k) + \sum_{j=1}^n b_j * mv(k - j),$$

де $mv(k)$ – значення ковзного середнього для k -го елементу часового ряду;

b_j – коефіцієнт моделі [11]

2.4 Метод декомпозиції

Метод декомпозиції допомагає аналізувати коливання попиту за певний період і щоквартально прогнозувати продажі на весь рік. Для цього використовуються чотири основні фактори: тенденція, циклічність, сезонність та випадковість, які можуть впливати на зміни попиту. Використання цих факторів дозволяє більш точно прогнозувати продажі та уникнути помилок у плануванні обсягів продажів.

Існує безліч формул для прогнозування обсягу продажів. Вибір формули залежить від різноманітних факторів, таких як тип продукту або послуг, які ви пропонуєте, та інформації, яку враховуєте.

Базова формула, використовуйте наступне:

обсяг продажів = кількість товарів/послуг, проданих за певний період $\times$ ціна кожного товару/послуги.

Для більш точного прогнозування обсягу продажів можна використовувати більш складну модель, яка має наступну формулу:

$$F = T + S \pm E,$$

F – прогнозоване значення;

T – тренд;

S – сезонна складова;

E – помилка моделі.

Основна формула для методу декомпозиції часового ряду може бути виражена наступним чином:

$$Y_t = T_t * S_t * R_t$$

де Y_t – значення часового ряду в момент часу t ;

T_t – трендова складова;

S_t – сезонна складова;

R_t – складова випадкових коливань.

Кожна з цих складових може бути розрахована окремо:

- Трендова складова T_t : тренд може бути оцінений, наприклад, за допомогою методу найменших квадратів або експоненційного згладжування.
- Сезонна складова S_t : сезонність може бути виявлена шляхом виявлення повторюваних циклічних патернів в часовому ряді.
- Складова випадкових коливань R_t : ця складова може бути визначена як відношення фактичного значення до (тренду * сезонності). Вона відображає випадкові відхилення від утвореної моделі [12].

2.5 Висновки до розділу

У другому розділі проведено дослідження математичних методів, які можуть бути використані для побудови систем автоматизації прогнозування продажів. Було проаналізовано такі методи, як аналіз часових рядів, модель авторегресії, модель ковзного середнього та метод декомпозиції. Отримані результати стануть основою для подальших досліджень та розробки ефективної системи прогнозування продажів.

3 РОЗРОБКА СИСТЕМИ АВТОМАТИЗАЦІЇ ПРОГНОЗУВАННЯ ПРОДАЖІВ

3.1 Вибір оптимальних параметрів для прогнозування продажів

Сьогодні Інтернет все більше перетворюється з абстрактної глобальної комп'ютерної мережі в повсякденний інформаційний канал. Багато компаній відкривають офіси на своїх веб-сайтах, а інші активно займаються онлайн-бізнесом. Інтернет продовжує впливати на економіку: люди все частіше обирають покупки в Інтернет-магазинах, спілкування з друзями онлайн, читання онлайн-газет і журналів. Швидко зростаюча Інтернет-аудиторія представляє новий ринок для компаній усіх видів, адже реклама, продаж і залучення нових компаній в Інтернет-бізнес не мають географічних обмежень.

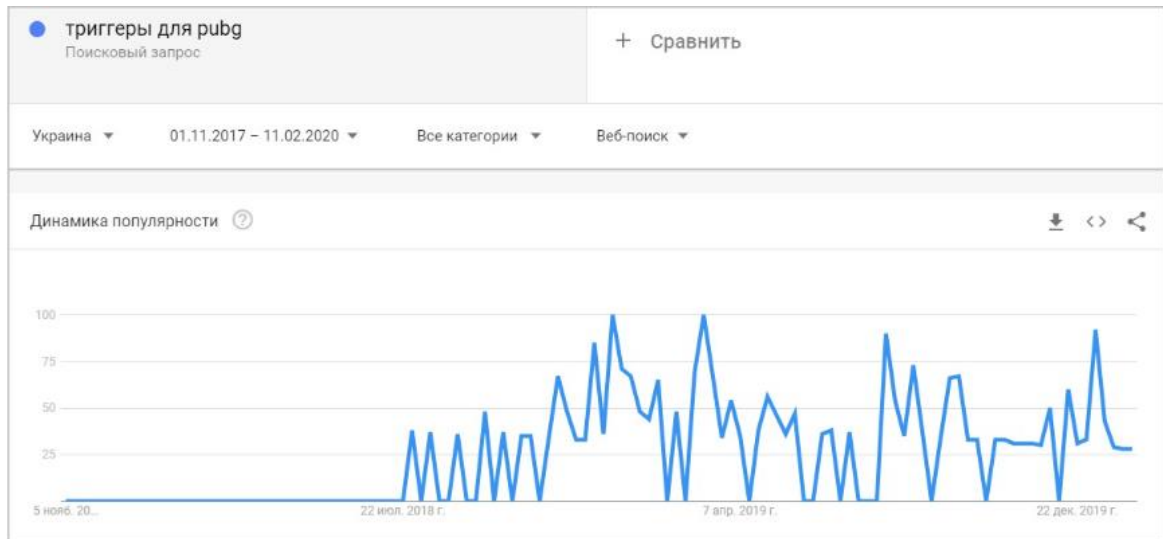
Для більш точного прогнозування продажів необхідно:

1. Вивчити потенційну аудиторію. При виборі ніші профілюйте свою цільову аудиторію. Проводьте опитування, індивідуальні та групові інтерв'ю, збирайте дані із соціальних мереж – це найдоступніші джерела інформації про споживачів. Помилка багатьох стартапів полягає в тому, що вони намагаються знайти універсальний продукт, який сподобається всім. Однак на початкових етапах створення Інтернет-магазинів вигідніше зосередитися на вузькій групі споживачів. Знаючи стать, вік, соціальний статус, регіон проживання та іншу інформацію про потенційних покупців, ви зможете визначити їхні проблеми та запропонувати найкращі рішення.

2. Відстежувати тренди.

Google Trends – це безкоштовний інструмент, який чудово підходить для аналізу сезонності, поточних тенденцій та прогнозування попиту. За допомогою цього сервісу можна відстежувати зростання інтересу до певних товарів і категорій. Наприклад, на графіку Google Trends видно, що в липні 2018 року різко зросли запити на пристрій Trigger Statistics для смартфонів.

Попит на цей товар був високим, досягнувши піку в квітні 2019 року, і досі залишається значним. Статистичні дані Google Trends зображені на рисунку 3.1.



Рисунку 3.1 – Статистика Google Trends

3. Оцінити ризики, пов'язані з товарами.

При виборі ніші враховуйте потенційні загрози, які можуть виникнути з певними продуктами. Переконайтеся, що ви можете забезпечити необхідні умови для зберігання та транспортування товарів. Наприклад, овочі, фрукти, натуральні солодощі та квіти мають обмежений термін зберігання і швидко псуються. Посуд, музичні інструменти, дзеркала та деякі будівельні матеріали – це крихкі предмети, які складно транспортувати без ризику пошкодження. Якщо у вас немає можливості утримувати склад і займатися логістикою, розгляньте варіант дропшипінгу. Це тип партнерства, де ви займаєтеся лише обробкою замовлень і маркетингом, а постачальник бере на себе всі інші процеси. У Khoroshop можна придбати повністю укомплектований Інтернет-магазин з завантаженими товарами.

4. Популярні ніші для Інтернет-магазину.

В нещодавній статті ми поділилися ідеями для бізнесу та назвали перспективні товарні ніші, які частково збіглися з категоріями товарів, які

українці найчастіше купували у 2020 році. Група компаній EVO дослідила ці показники та порівняла їх з даними за 2019 рік:

- Одяг, взуття та аксесуари – 4,6 млн шт., приріст – 13%;
- Побутова техніка та електроніка – 4,4 млн шт., приріст – 28%;
- Товари повсякденного попиту – 4,1 млн шт., приріст – 18%;
- Товари для дому та саду – 4 млн шт., приріст – 43%;
- Косметика і парфумерія – 2,8 млн шт., приріст – 34%.

Отже, лідером онлайн-продажу залишається одяг, найбільше зростання продажу за рік – у товарів для дому і саду.

5. Товари, які продаються найкраще.

За даними сервісу Google Trends, найпопулярнішими товарами 2020 року в Україні:

- iPhone 12;
- маска медична;
- пульсоксиметр;
- респіратор;
- антисептик.

Це статистика за популярними запитами, яка не завжди відображає реальну позицію в онлайн-продажах. Тому насамперед належні гравці, які на великій ринку [13].

3.2 Розробка системи автоматизації продажів

В цьому розділі буде розроблено алгоритм роботи Інтернет-магазину, включаючи діаграму варіантів використання та опис акторів.

IDEF0 – це методологія графічного опису систем і процесів, яка організовує діяльність як сукупність взаємозалежних функцій. Вона дозволяє аналізувати функції організації без прив'язки до об'єктів, що забезпечують їх реалізацію.

У IDEF0 реалізовані три базові принципи моделювання процесів:

- принцип функціональної декомпозиції;
- принцип обмеження складності;
- принцип контексту [14].

Дерево цілей (або ієрархія цілей) — це візуальний інструмент для структуризації та організації цілей, завдань і підзадач у вигляді дерева. Воно допомагає чітко визначити, які кроки потрібно виконати для досягнення головної мети, розподіляючи їх на окремі підцілі та задачі. Кожна підціль підтримує загальну мету, а задачі конкретизують шляхи досягнення цих підцілей. Дерево цілей автоматизованої системи прогнозування продажів зображено на рисунку 3.2.



Рисунку 3.2 – Дерево цілей автоматизованої системи прогнозування продажів

Діаграми варіантів використання – це найпростіші поведінкові діаграми UML, корисні для висвітлення функціональних особливостей програм для користувачів, які не мають глибоких знань в ІТ-галузі, наприклад, для замовників програм. Ці діаграми використовуються для опису цілей, які досягають користувачі ваших програм, будь це люди або інші програми, що

використовують вашу систему. Вони допомагають описати функціональні вимоги до програми або її підсистеми, представляючи загальну картину використання програми.

Основні компоненти діаграми варіантів використання:

- Актори: Зовнішні елементи, які взаємодіють із системою. Це можуть бути люди, інші системи або обладнання.
- Варіанти використання (use cases): Описують специфічні функціональні вимоги системи. Вони представляють завдання або дії, які система повинна виконувати у взаємодії з акторами.
- Система: Відображається як прямокутник, що об'єднує всі варіанти використання. Актори зазвичай розташовуються за межами цього прямокутника.
- Зв'язки: Відображають взаємодію між акторами і варіантами використання. В UML зв'язки позначаються лініями, що з'єднують акторів з відповідними варіантами використання.

Основні типи зв'язків у діаграмі варіантів використання:

- Асоціація: Проста лінія, що з'єднує актора з варіантом використання, показуючи, що актор взаємодіє з цим варіантом.
- Включення (include): Пунктирна лінія зі стрілкою, яка показує, що один варіант використання обов'язково включає функціональність іншого варіанту використання.
- Розширення (extend): Пунктирна лінія зі стрілкою, яка показує, що один варіант використання може бути розширений додатковою поведінкою іншого варіанту використання при певних умовах.
- Загальне використання: Показує загальний сценарій, який використовується кількома акторами або варіантами використання [15].

Для кращої візуалізації створено модель функціонування Інтернет-магазину з використанням структурного підходу на основі стандарту IDEF0, яка зображена на рисунках 3.3-3.4.

На рівні A0 весь процес розглядається як функціональний блок, що включає всі керуючі та робочі об'єкти. Ця діаграма також відображає всю необхідну вхідну інформацію, яка використовується для замовлення товару. Діаграма представлена на рисунку 3.3.

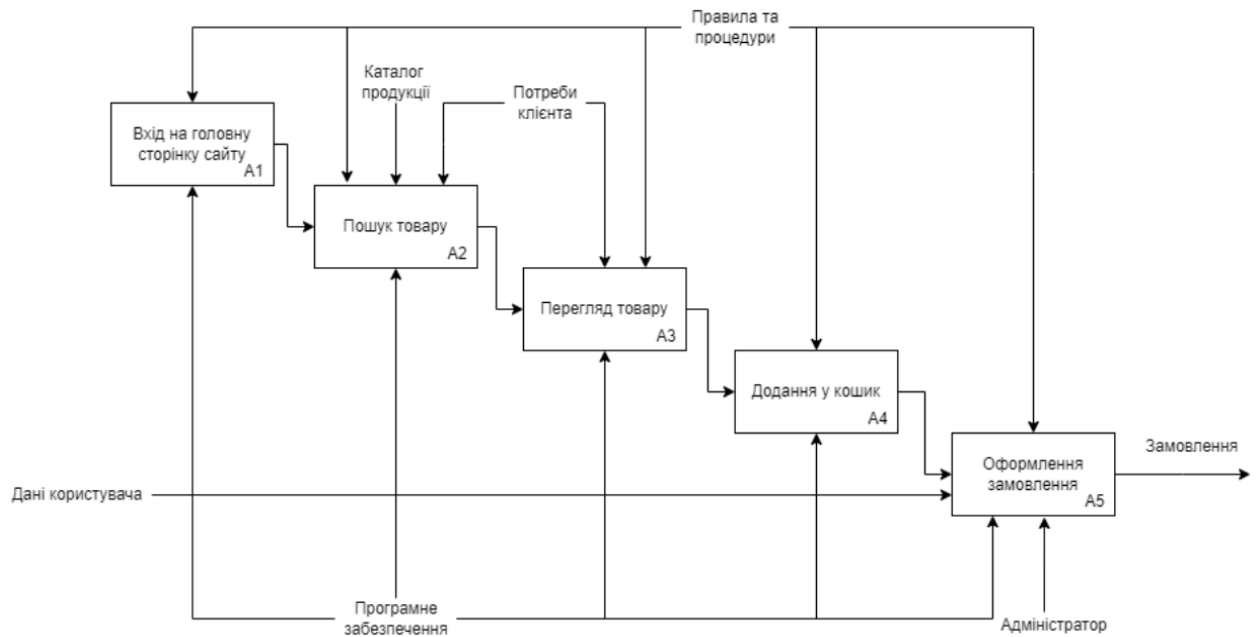


Рисунок 3.3 – Вигляд контекстної моделі

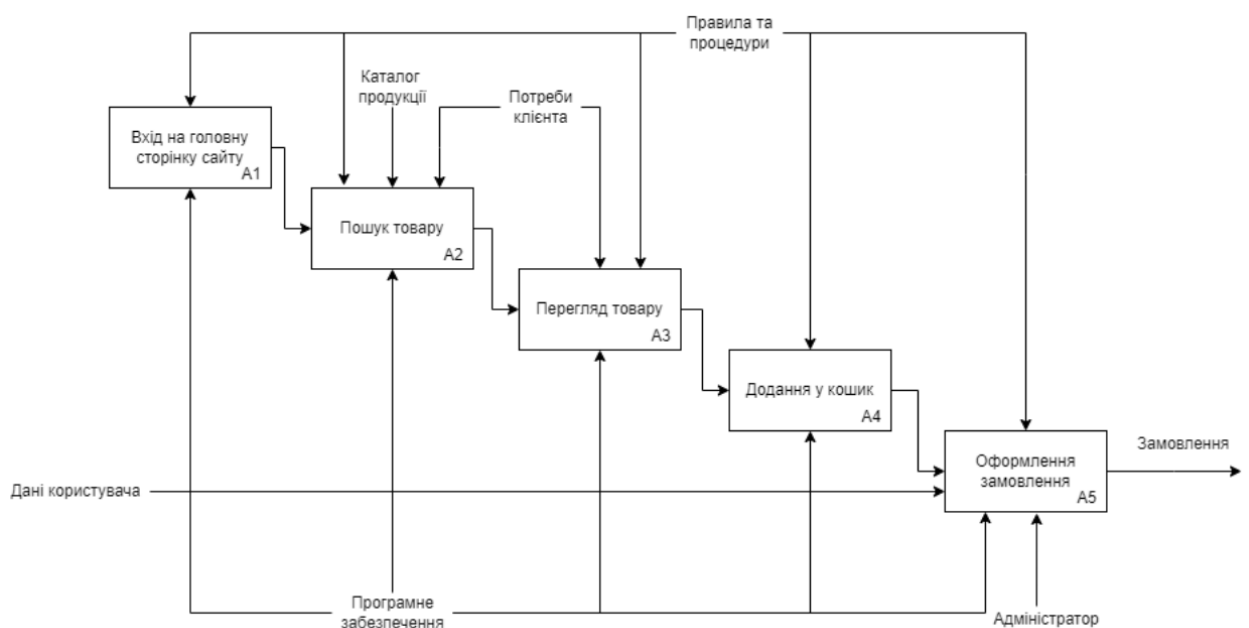


Рисунок 3.4 – Вигляд моделі декомпозиції

Діаграма першого рівня надає опис функції обробки на нульовому рівні та деталізує функціональний блок A0, розкладаючи його на набір взаємопов'язаних підфункцій. Варіант цієї діаграми для даного проекту показаний на рисунку 3.4.

Далі зобразимо приклад діаграми використання для Інтернет-магазину. Ця діаграма демонструє, як різні суб'єкти або сутності взаємодіють із системою через конкретні випадки використання.

Діаграма варіантів використання зображена на рисунку 3.5. Дана діаграма досить добре показує алгоритм роботи магазину і показує самий процес замовлення товару через Інтернет.

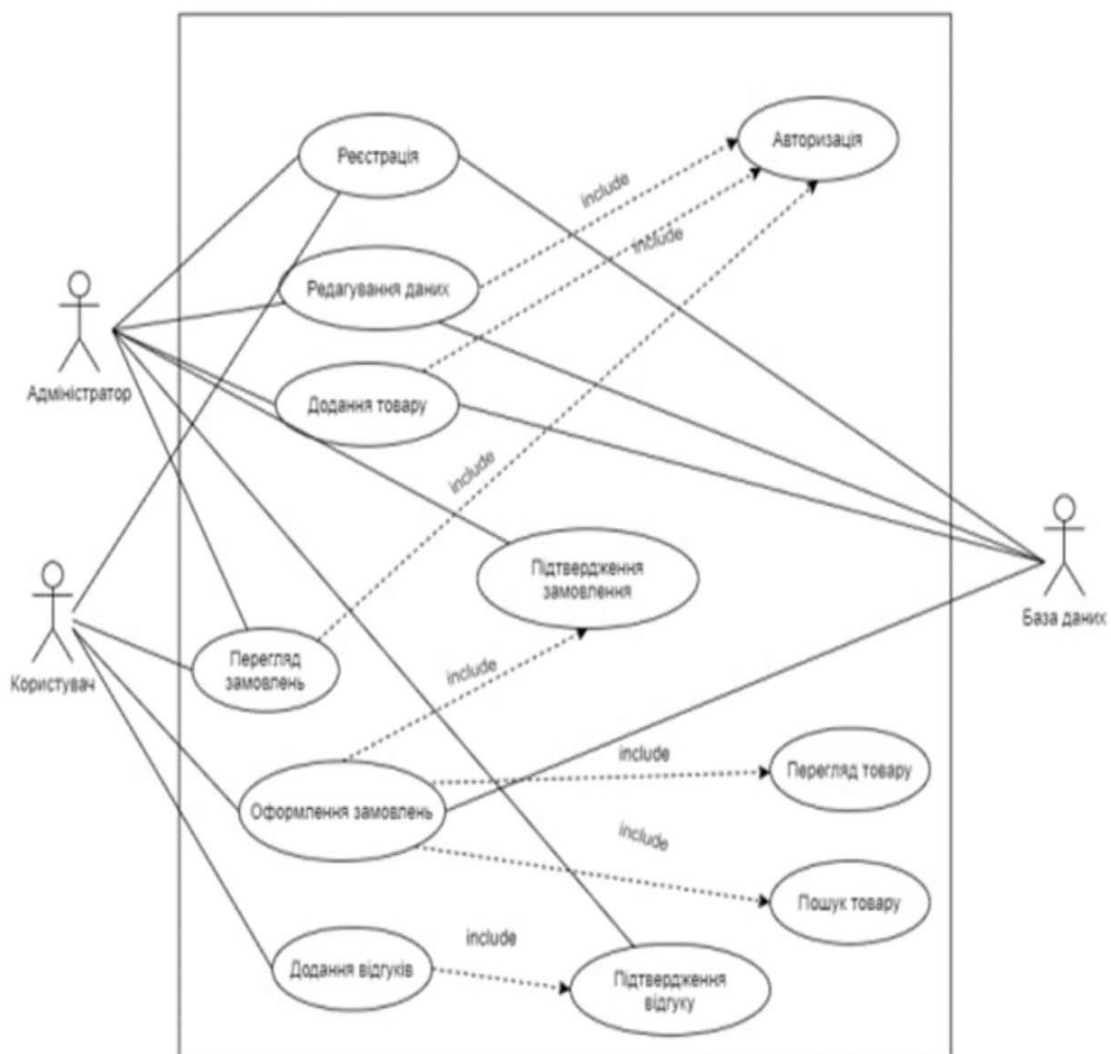


Рисунок 3.5 – Діаграма варіантів використання

Суть діаграми полягає в наступному: розроблена система представлена у вигляді суб'єктів або сутностей, які взаємодіють із системою, використовуючи випадки використання. Учасником є будь-яка організація, яка взаємодіє із системою зовні. Це може бути людина, технічний пристрій, програма або інша система, що впливає на систему моделювання, визначену розробником. Випадки використання описують послуги, які система надає учасникам [16].

Дана діаграма досить добре показує алгоритм роботи магазину і показує самий процес замовлення товару через Інтернет.

Прогнозування продажу – це науково обґрунтований процес передбачення майбутніх обсягів продажу товарів, що має ймовірнісний характер. Він базується на аналізі відповідної інформації за минулі роки з урахуванням змін факторів, що впливають на результати в перспективі. Основна мета прогнозування – визначити такий обсяг продажу на майбутнє, який буде найбільш ефективним для компанії.

Відповідно до цього, система розподіляє свої ресурси для досягнення відповідних показників. Результатом прогнозування є прогноз продажу на майбутній період. Отже, прогнозування – це процес передбачення величини продажу на перспективу, а прогноз – кількісний результат цього процесу. Хоча терміни «прогнозування» і «прогноз» часто використовуються як синоніми, вони мають різний зміст. Глибоке розуміння прогнозування продажів вимагає врахування його основних особливостей:

- Прогнозування базується на реальних умовах діяльності фірми на ринку та чинниках, що впливають на обсяг прогнозу;
- воно містить елементи ймовірності, які стосуються ситуацій у майбутньому, тому результати повинні характеризуватися ступенем ймовірності;

- результати прогнозування слугують орієнтиром при плануванні обсягу продажу;
- при розробці прогнозу не ставляться завдання його деталізації;
- прогноз має багатоваріантний характер.

Розглядаючи сутність прогнозування, важливо розрізняти такі поняття, як прогноз продажу, план продажу, потенційний та прогнозований обсяг продажу. Прогноз продажу – це обсяг, який може бути досягнутий за певних умов або подій, тоді як план продажу – це кількість товарів, яка буде продана за певних дій. [17].

3.3 Роботи з набором даних

Kaggle – платформа для змагань з аналітики та передбачувального моделювання, де статистики та спеціалісти з даних конкурують у створенні найкращих моделей для прогнозування та опису даних, запропонованих компаніями або користувачами. Цей краудсорсинговий підхід ґрунтується на тому, що існує багато стратегій для будь-якого завдання з передбачувального моделювання, і наперед не відомо, яка методика буде найкращою. Крім того, це платформа для численних конкурсів з дослідження даних та машинного навчання.

Існує два основних типи машинного навчання: індуктивне навчання (навчання по прецедентах) і дедуктивне навчання. Індуктивне навчання поділяється на три основні типи:

- контрольоване навчання (supervised learning): використовується, коли є великі обсяги даних з маркерами. Наприклад, алгоритм може навчитися розпізнавати котів і собак на основі тисяч фотографій з мітками, які заздалегідь проставив користувач. У цьому випадку «вчителем» виступає людина, яка визначила маркери;

- неконтрольоване навчання (unsupervised learning): використовується для роботи з даними без міток. Машина шукає закономірності, структурує дані або виконує класифікацію на основі виявлених шаблонів;
- навчання з підкріпленням (reinforcement learning): використовує методи винагороди та покарання для навчання алгоритму на основі досвіду [19].

Класифікація – одна з найбільш популярних задач машинного навчання, аналогічна тому, як дитина вчиться розпізнавати форму і розмір предметів, групуючи їх.

Мова Python є однією з найпростіших у вивченні та використанні серед популярних мов програмування. Програмний код на Python легко читати та писати, і він залишається зрозумілим навіть при високій лаконічності. Це завдяки виразності мови, що дозволяє писати менше рядків коду, ніж у випадку з такими мовами, як C++ або Java.

Python підтримує процедурне, об'єктно-орієнтоване і, меншою мірою, функціональне програмування, хоча в основному це об'єктно-орієнтована мова [20].

Таким чином, було обрано конкурс на Kaggle, а також взято відповідні дані. Вигляд конкурсу зображено на рисунку 3.6 [21].

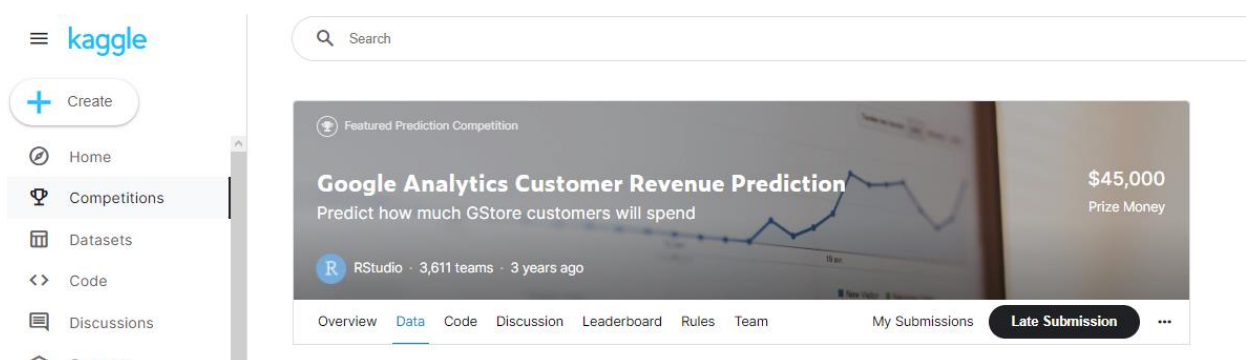


Рисунок 3.6 – Вигляд конкурсу

3.4 Вибір програми для подальшої візуалізації

Для початку потрібно імпортувати дані з MS Excel. Microsoft Excel – це засіб для роботи з електронними таблицями, що надає широкий спектр функцій для обробки даних, аналізу, обробки тексту, створення графіків і роботи з базами даних.

Excel є потужним інструментом для розв'язання різноманітних задач, пов'язаних із великими масивами даних. Його застосовують у різних сферах, від бухгалтерських і складських задач до розрахунків у макроекономіці [22].

Power BI – це набір хмарних служб бізнес-аналітики від Microsoft, який використовується для перетворення необроблених даних в інформативні звіти за допомогою візуальних зображень і таблиць. Приклад дашборду в Microsoft Power BI показаний на рисунку 3.7.

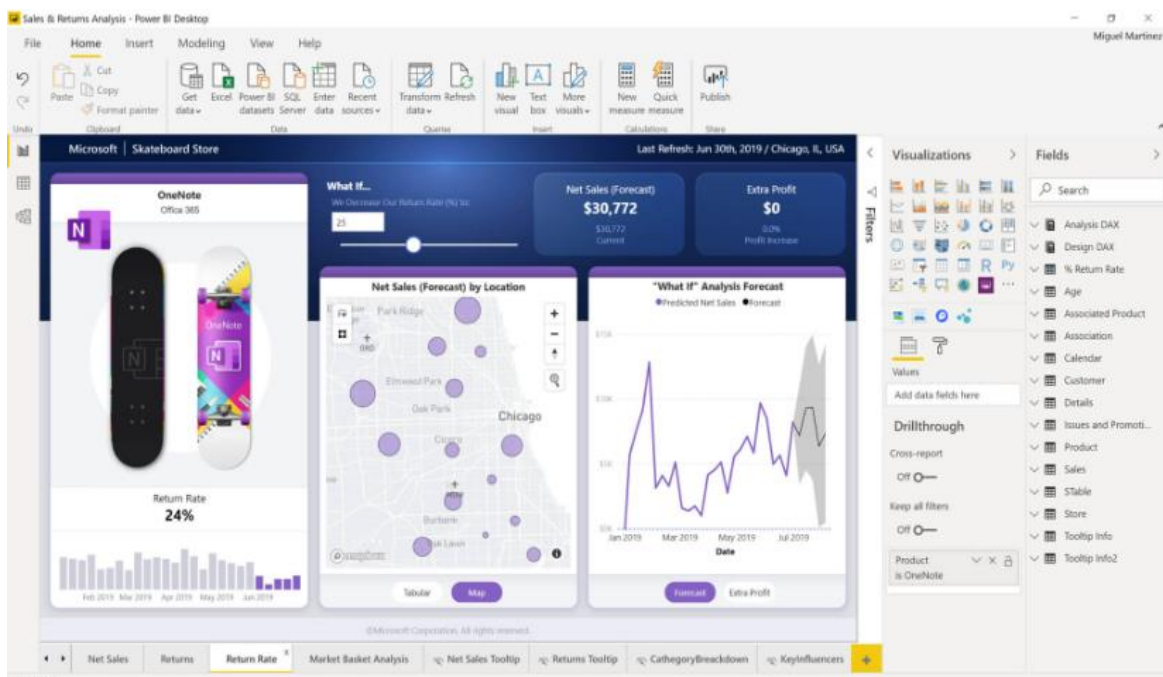


Рисунок. 3.7 – Вигляд прикладу дашборду

Плюси програми:

- Користувальницькі візуалізації;
- Інтеграція з MS Excel;

- Доступність даних;
- Інтерактивні візуалізації.

Мінуси:

- Обробка великих обсягів даних;
- Складно ознаки та освоїти

Після завантаження даних нам будуть доступні вбудовані елементи візуалізації:

- лінійна діаграма з накопиченням;
- гістограма з накопиченням;
- лінійна діаграма з групуванням;
- гістограма з групуванням;
- графік;
- діаграма з областями;
- діаграма з областями з накопиченням;
- лінійна гістограма і гістограма з накопиченням;
- лінійна гістограма і гістограма з групуванням;
- каскадна діаграма;
- точкова діаграма;
- кругова діаграма;
- діаграма дерева;
- таблиця тощо [23].

3.5 Висновки до розділу

В цьому розділі розглянуто оптимальні параметри (показники) для автоматизації прогнозування продажів в Інтернет-магазинах. Далі розроблено алгоритм роботи магазину, пошук та збір даних для прогнозування. Також обрано програму, для подальшої візуалізації даних, ця програма – Power BI.

4 ТЕСТУВАННЯ СИСТЕМИ АВТОМАТИЗАЦІЇ ПРОГНОЗУВАННЯ ПРОДАЖІВ ІНТЕРНЕТ-МАГАЗИНУ

4.1 Пошук та імпорт даних

В даному розділі буде проведено розвідувальний аналіз для автоматизованого прогнозування продаж за допомогою Microsoft Power BI.

Для початку необхідно імпортувати дані з Microsoft Excel, вигляд вікна імпорту зображено на рисунку 4.1.

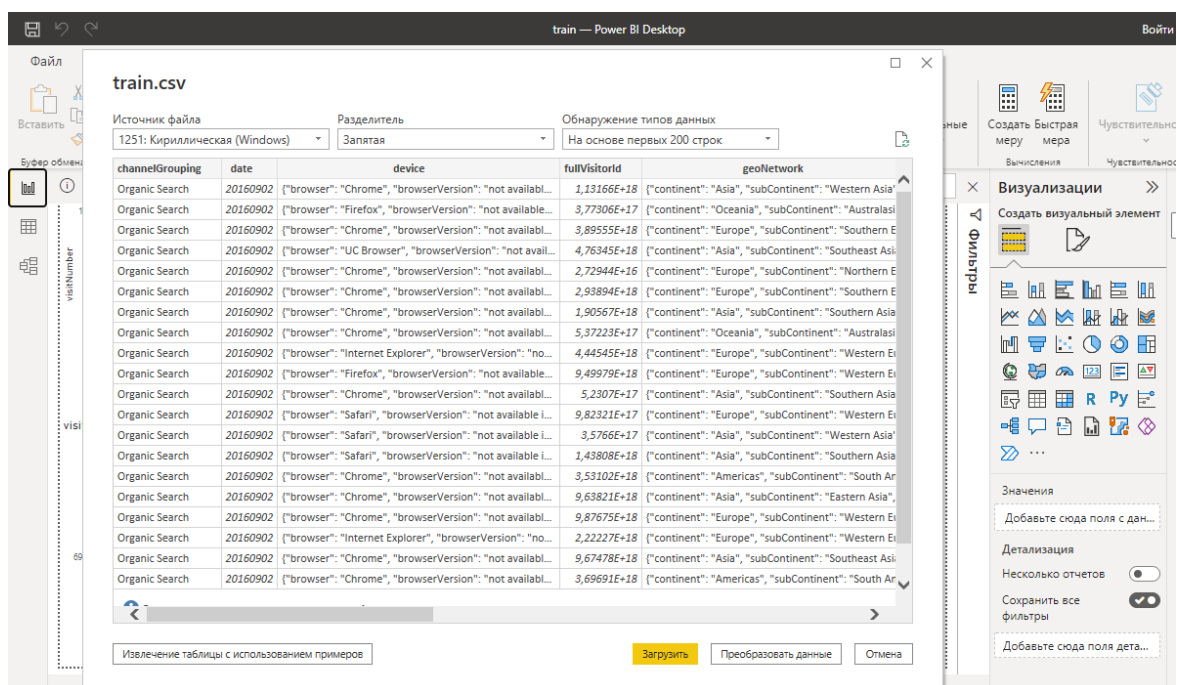


Рисунок 4.1 – Вигляд вікна імпорту

Значення стовбців:

channelGrouping – канал, через який користувач прийшов до Магазину.

fullVisitorId – унікальний ідентифікатор для кожного користувача Google Merchandise Store.

channelGrouping – канал, через який користувач прийшов до Магазину.

date – дата, коли користувач відвідав Магазин.

device – специфікації пристрою, який використовується для доступу до Магазину.

geoNetwork – Цей розділ містить інформацію про географію користувача.

sessionId – унікальний ідентифікатор для цього відвідування магазину.

socialEngagementType – тип залучення, або «Соціально залучений» або «Не соціально залучений».

totals – цей розділ містить зведені значення протягом сеансу.

trafficSource – цей розділ містить інформацію про джерело трафіку, з якого походить сеанс.

visitStartTime – мітка часу (виражається як час POSIX).

visitId – ідентифікатор цього сеансу. Це частина значення, яке зазвичай зберігається як файл cookie_utmb. Щоб отримати повністю унікальний ідентифікатор, ви повинні використовувати комбінацію fullVisitorId і visitId.

visitNumber – номер сеансу для цього користувача. Якщо це перший сеанс, то для нього встановлено значення 1.

Після імпорту даних будуть доступні таблиці, зображені на рисунку 4.2. Ознайомившись з інтерфейсом програми, можна виконувати наступні дії.

channelGrouping	date	device	fullVisitorId	geoNetwork	sessionId
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	3,13752002835417E+18	("continent": "Americas", "subContinent": "Northern Ame	3137520028354173102_1508204646
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	6,77009577251962E+18	("continent": "Americas", "subContinent": "Northern Ame	6770095772519618990_1508201418
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	4,11472577596933E+18	("continent": "Americas", "subContinent": "Northern Ame	4114725775969328107_1508160141
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	3,63340160420526E+18	("continent": "Americas", "subContinent": "Northern Ame	3633401604205263760_1508186392
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	7,29138719185693E+18	("continent": "Americas", "subContinent": "Northern Ame	7291387191856926805_1508143388
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	9,57104598477272E+18	("continent": "Americas", "subContinent": "Northern Ame	9571045984772719982_1508161349
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	558077087682879	("continent": "Americas", "subContinent": "Northern Ame	558077087682879_1508208588
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	3,30945052526305E+18	("continent": "Americas", "subContinent": "Northern Ame	3309450525263048024_1508195520
Organic Search	16 октября 2019 г.	{"browser": "Chrome", "browserVersion": "not available in	6,62854083089109E+18	("continent": "Americas", "subContinent": "Northern Ame	6628540830891089041_1508178277
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	5,82695067475844E+18	("continent": "Americas", "subContinent": "Northern Ame	582695067475844227_1512098799
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	7,9911483740831E+18	("continent": "Americas", "subContinent": "Northern Ame	7991148374083102098_1512079193
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	1,02528016302086E+18	("continent": "Americas", "subContinent": "Northern Ame	1025280163020859873_1512088598
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	3,41504910234899E+18	("continent": "Americas", "subContinent": "Northern Ame	3415049102348987851_1512068458
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	2,30077817192474E+18	("continent": "Americas", "subContinent": "Northern Ame	2300778171924743592_1512079471
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	5,21369860395766E+18	("continent": "Americas", "subContinent": "Northern Ame	5213698603957661056_1512056872
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	3,77005334806856E+18	("continent": "Americas", "subContinent": "Northern Ame	3770053348068561054_1512100608
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	5,18042146424878E+17	("continent": "Americas", "subContinent": "Northern Ame	518042146424878052_1512075096
Organic Search	30 ноября 2017 г.	{"browser": "Chrome", "browserVersion": "not available in	8,70314923846042E+18	("continent": "Americas", "subContinent": "Northern Ame	8703149238460421376_1512104272
Organic Search	15 апреля 2018 г.	{"browser": "Chrome", "browserVersion": "not available in	5,13681918748298E+18	("continent": "Americas", "subContinent": "Northern Ame	5136819187482980969_1523810336
Organic Search	15 апреля 2018 г.	{"browser": "Chrome", "browserVersion": "not available in	1,67447504819235E+18	("continent": "Americas", "subContinent": "Northern Ame	1674475048192349141_1523840386
Organic Search	15 апреля 2018 г.	{"browser": "Chrome", "browserVersion": "not available in	1,93947479171038E+18	("continent": "Americas", "subContinent": "Northern Ame	1939474791710381494_1523807343

Рисунок 4.2 – Вигляд імпортованих даних

Далі показано початкову модель train, зображено на рисунку 4.3.

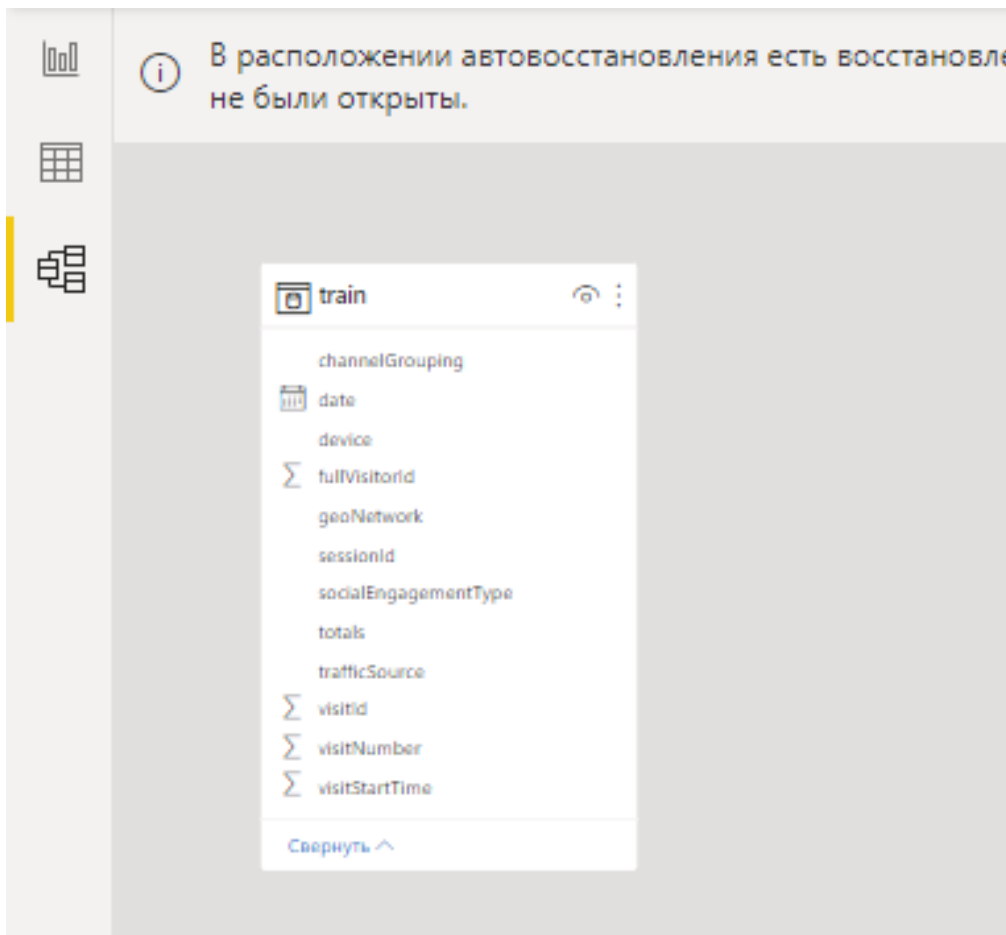


Рисунок 4.3 – Вигляд початкової моделі train

4.2 Побудова дашбордів та використання функцій DAX

Наступним кроком є створення графіків та діаграм у Microsoft Power BI. На рисунку 4.10 представлена кругова діаграма для візуалізації даних про браузери, через які здійснюються покупки. Крім того, ми будемо використовувати набір функцій DAX, який містить функції, оператори та константи для обчислень і створення нових даних на основі існуючих. DAX дозволяє проводити складний аналіз, наприклад, обчислювати річний приріст або аналізувати тенденції ринку.

Володіючи знаннями про DAX можна створювати звіти, які показують цінну інформацію, взагалі не використовуючи формули DAX. Але що, якщо

вам потрібно проаналізувати темпи зростання за категоріями продуктів, також для різних діапазонів дат? Або якщо вам потрібно розрахувати річний приріст у порівнянні з тенденціями ринку? Формули DAX надають цю можливість, як і багато інших важливих функцій. Навчившись створювати ефективні формули DAX, ви зможете максимально використовувати свої дані. Отримавши необхідну інформацію, ви можете почати вирішувати реальні бізнес-проблеми, які впливають на продуктивність. Це перевага Power BI і DAX, щоб допомогти вам отримати його.

В основі своєї, всі функції DAX можна розділити на кілька категорій:

- текстові;
- логічні;
- математичні;
- статистичні (функції агрегування);
- функції дати та часу;
- функції логіки операцій з датою (time intelligence);
- табличні;
- інформаційні;
- та інші, які не підпадають під категорії вище [24].

Використовуючи функцію SUM (SUMX) – це операції додавання, підсумуємо кількість відвідувачів магазину. Результати розрахунків показані на рисунках 4.4 і 4.5.

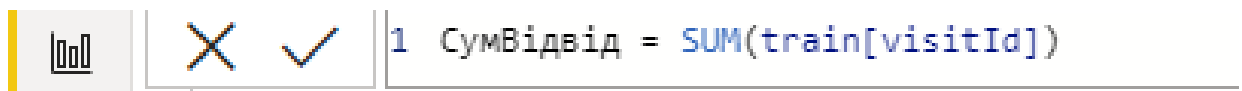


Рисунок 4.4– Вигляд розрахункового поля



Рисунок 4.5 – Вигляд результату розрахунку

Для перевірки на пусті значення використовується функція ISBLANK, яка дозволяє перевірити міру «СумВідвід» на нульові значення. Результати представлені на рисунках 4.6-4.7.



Рисунок 4.6 – Вигляд розрахункового поля



Рисунок 4.7 – Вигляд результату розрахунку

Наступним кроком буде створення міри за допомогою функції FILTER, яка відфільтрує номери сеансів, що перевищують 1. Результати показані на рисунках 4.8-4.9.

1 Таблица = FILTER(train, train[visitNumber] >1)

Рисунок 4.8 – Вигляд розрахункового поля

sessionid	socialEngagementType	totals	trafficSource	visitId	visitNumber
420761633459_1493628221	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1493628221	2
356365071340_1493647469	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1493647469	2
470278689788_1497234304	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1497234304	2
128964214468_1488348062	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1488348062	2
176444721125_1474131213	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1474131213	2
957291559163_1481898109	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1481898109	2
70741221130_1481942606	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1481942606	2
362925336232_1471718237	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1471718237	2
726358186888_1497644659	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1497644659	2
290163080825_1489881192	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1489881192	2
014431607350_1480556036	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1480556036	2
715925604485_1477842273	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1477842273	2
861195708649_1499182983	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1499182983	2
807368840640_1487034726	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1487034726	2
120553935653_1494528470	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1494528470	2
079752867435_1494546477	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1494546477	2
793611346433_1494541648	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1494541648	2
683859032933_1470450666	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1470450666	2
977597589112_1476812206	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1476812206	2
99525270004_1476858677	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1476858677	2
137544679740_1490026702	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1490026702	2
1564259969_1491860995	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1491860995	2
212923249469_1477507034	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "n")	1477507034	2

Рисунок 4.9 – Вигляд результату розрахунку

Секторна діаграма або кругова діаграма із статистичної графіків у вигляді одного кола, який поділений на сегменти, ілюструють чисельні моделі. У секторній діаграмі, довжина кривої кожного сегменту(а отже і його центральний кут та площа), відповідає числу, яке вона зображує. Не зважаючи на те, що англійський варіант (кругла діаграма) отримав свою назву через звичайну схожість з пирогом, який ділять на шматки, зовнішній вигляд секторної діаграми може бути різним. Найпершу секторну діаграму створив Вільям Плейферу .

Секторні діаграми дуже широко застосовуються у бізнесі та засобах масової інформації. Хоча багато хто їх визначає та кількість фахівців, які

рекомендують уникати їх, підсилюючи на дослідження, що показало, як важко зрівняти сегменти секторної діаграми, або зіставляти з різних секторних діаграм. Секторні діаграми в таких випадках можуть бути замінені на інші графіки, наприклад, як стовпчикові діаграми, коробкові графіки та точкові графіки [25].

Інформація про пристрій, зображена на рисунку 4.10:

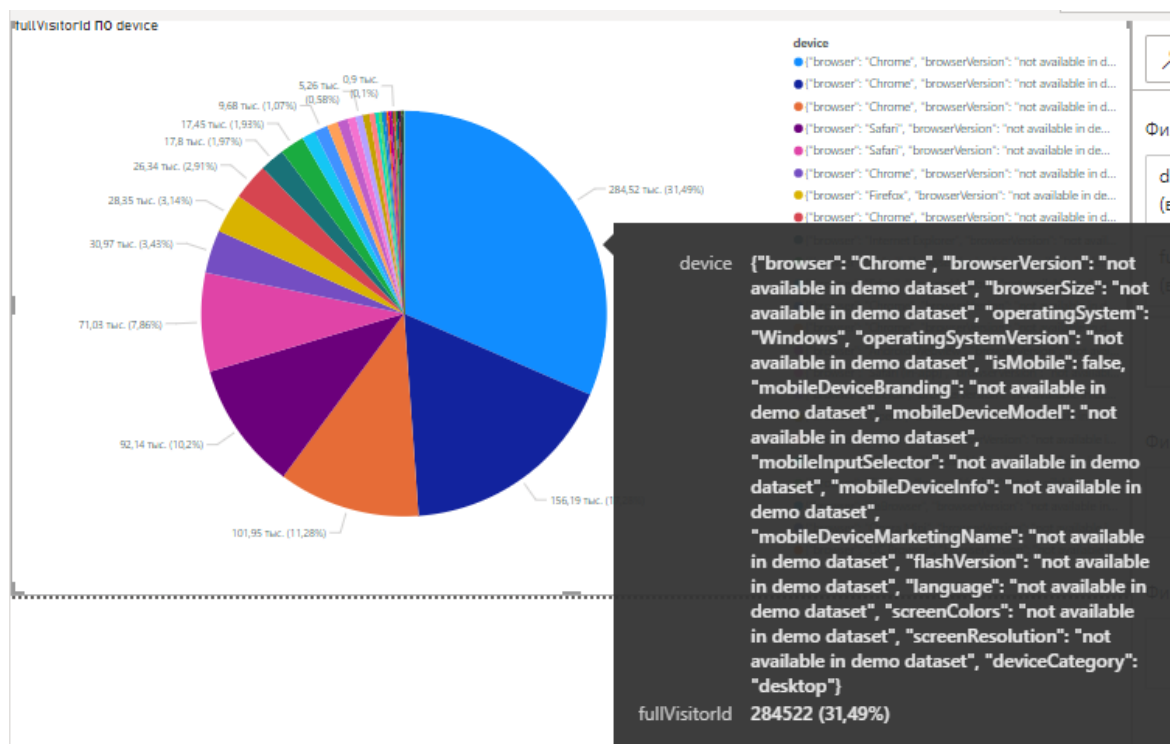


Рисунок 4.10 – Кругова діаграма

Аналізуючи операційні системи пристроїв, можна зробити висновок, що хоча Windows має більше підрахунків, Macintosh має більше підрахунків з ненульовим доходом. Також Chrome OS показує високий відсоток ненульового доходу. У мобільних ОС iOS перевершує Android за цим показником.

Географічні дані візуалізовані за допомогою гістограми на рисунку 4.11, яка демонструє кількість відвідувачів по континентах і відсоток ненульового доходу.

Гістограма – це графічний спосіб представлення табличних даних у вигляді діаграми, яка складається з прямокутників без розривів між ними. Кожен прямокутник пропорційний кількісним значенням певного показника. Для зручності ширина всіх прямокутників зазвичай однакова, а їх висота відповідає значенням відображуваного параметра.

Гістограма будується на основі даних, упорядкованих у рядки або стовпці на аркуші. Вона використовується для ілюстрації порівняння окремих елементів. На гістограмі категорії відкладаються по горизонтальній осі (осі абсцис), а значення – по вертикальній осі, осі ординат [26] .

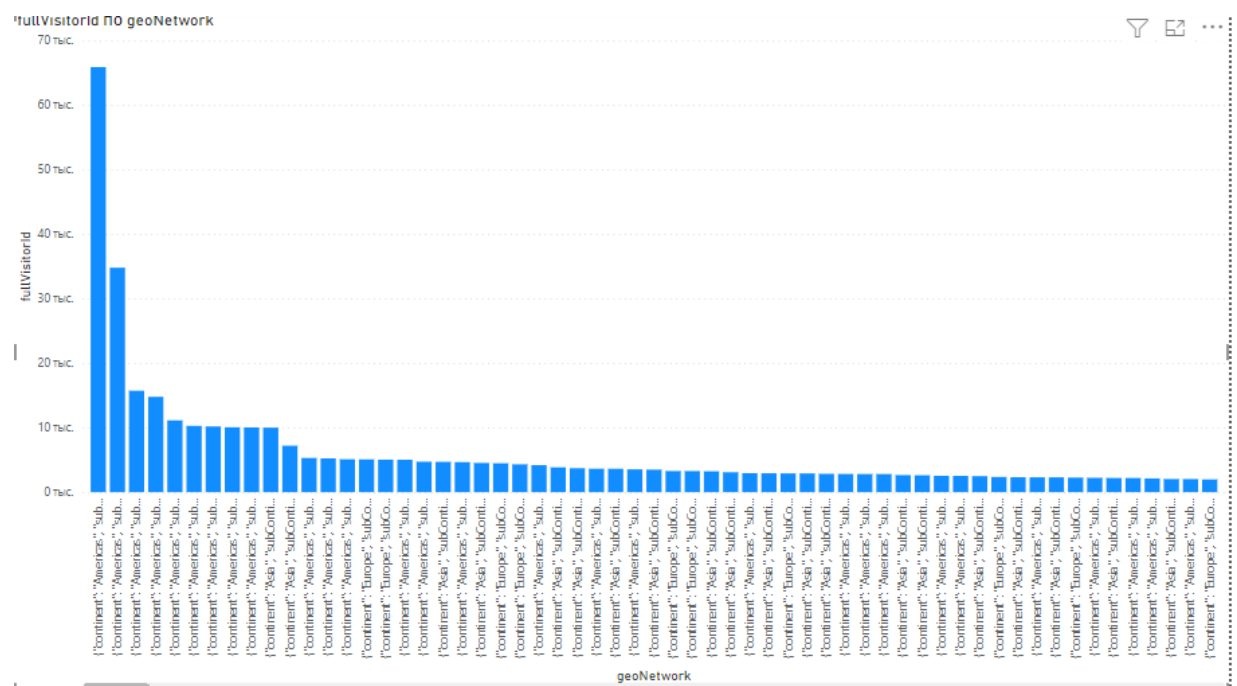


Рисунок 4.11 – Гістограми з групуванням

На цьому рисунку можемо побачити дохід, які приносять різні континенти. Америка має як більшу кількість підрахунків, так і найбільшу кількість підрахунків, де дохід відмінний від нуля. Незважаючи на те, що в Азії та Європі є велика кількість рахунків, кількість ненульових доходів на цих континентах порівняно невелика.

Також можемо зробити висновок перших двох пунктів із графіка субконтинентів. Якщо доменом мережі є "unknown.unknown", а не "(не

виконана за допомогою діаграми з областями і з накопиченням, зображена на рисунку 4.13.

Діаграма з областями з накопиченням – це плоскі діаграми з областями з накопиченням відображають тенденцію зміни відношення шкірного значення до цілого із годиною або залежно від даних інших категорій.

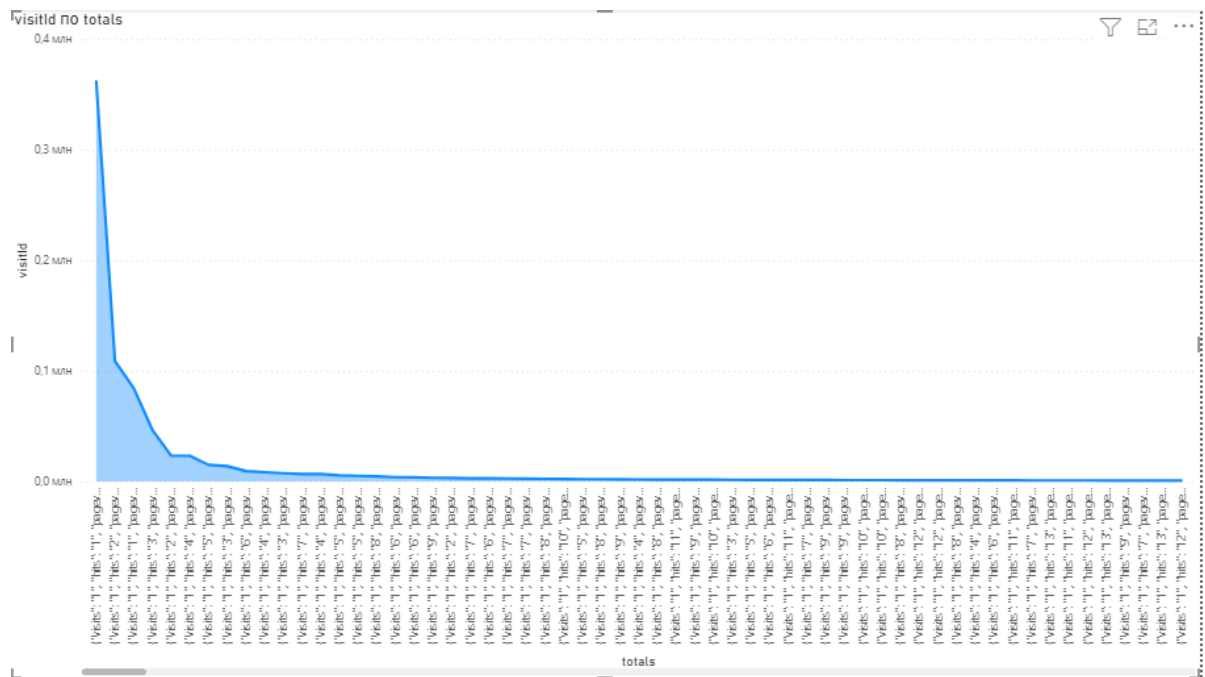


Рисунок 4.13 – Діаграма з областями із накопиченням

Можна зробити висновок, що змінні мають передбачуваний характер. Графік підрахунків показує спадний тренд: висока кількість звернень і переглядів сторінок на одну транзакцію зменшується зі збільшенням кількості звернень.

З іншого боку, ми чітко бачимо, що коли кількість переглядів сторінок на транзакцію відвідувача збільшується, ми бачимо, що існує велика кількість ненульових показників доходу.

Діаграма з областями – це плоскі діаграми з областями показують тенденцію зміни значення із часом або залежно від даних інших категорій. Замість діаграм з областями без накопичення краще використовувати

лінійчатую діаграму, тому що з одного ряду можна перекривати даними з інших рядів [27].

Наступним кроком є візуалізація номера відвідування та дати першого сеансу у вигляді діаграми з областями, що показано на рисунку 4.14.

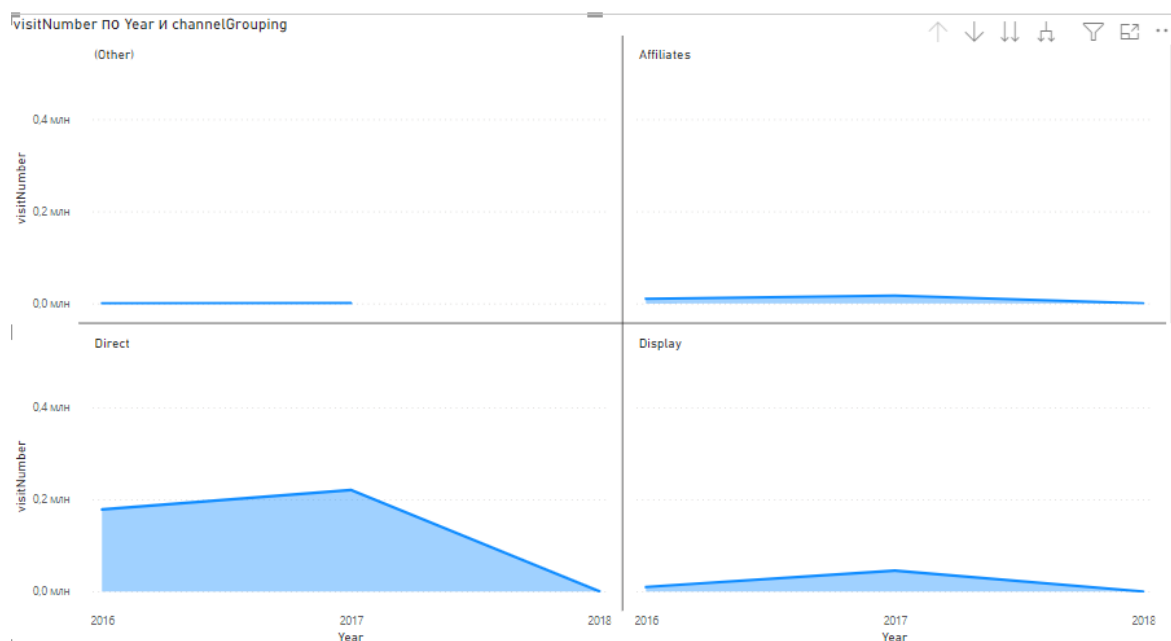


Рисунок 4.14 – Діаграма з областями

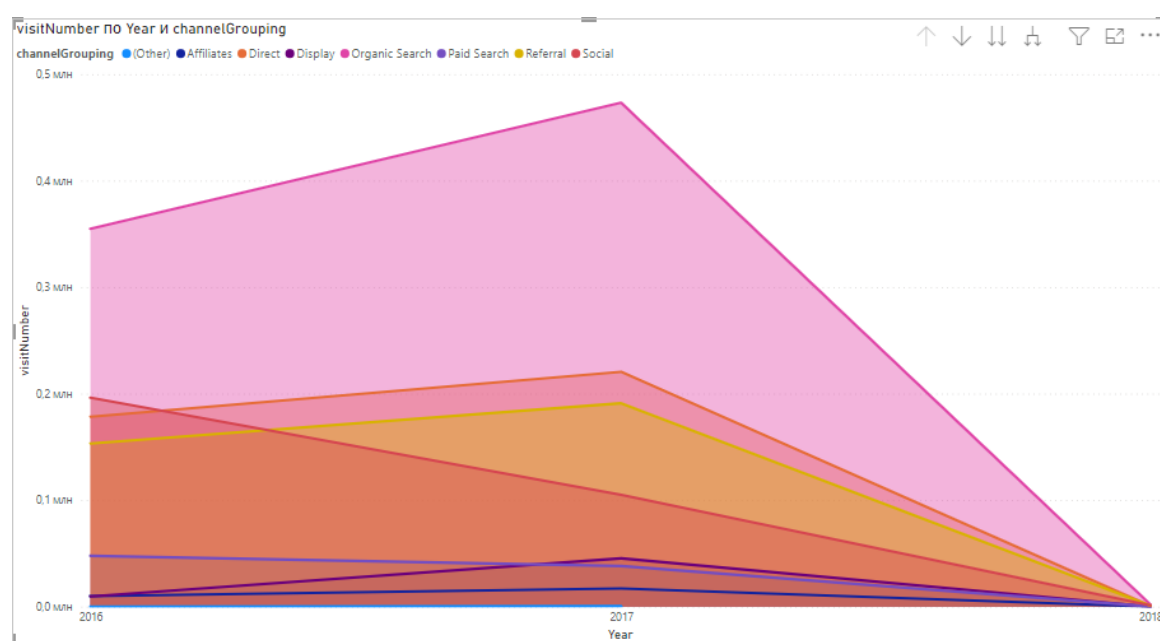


Рисунок 4.15 – Діаграма з областями

Для зручності можна об'єднати всі дані на одній діаграмі, що дозволить побачити, що звичайний пошук є найбільш популярним серед користувачів. Це відображено на рисунку 4.15.

На рисунку 4.16 зображено кількість відвідувань сайтів за континентами, де Північна Америка має найбільшу кількість відвідувань. Візуалізація виконана у вигляді діаграми з областями.

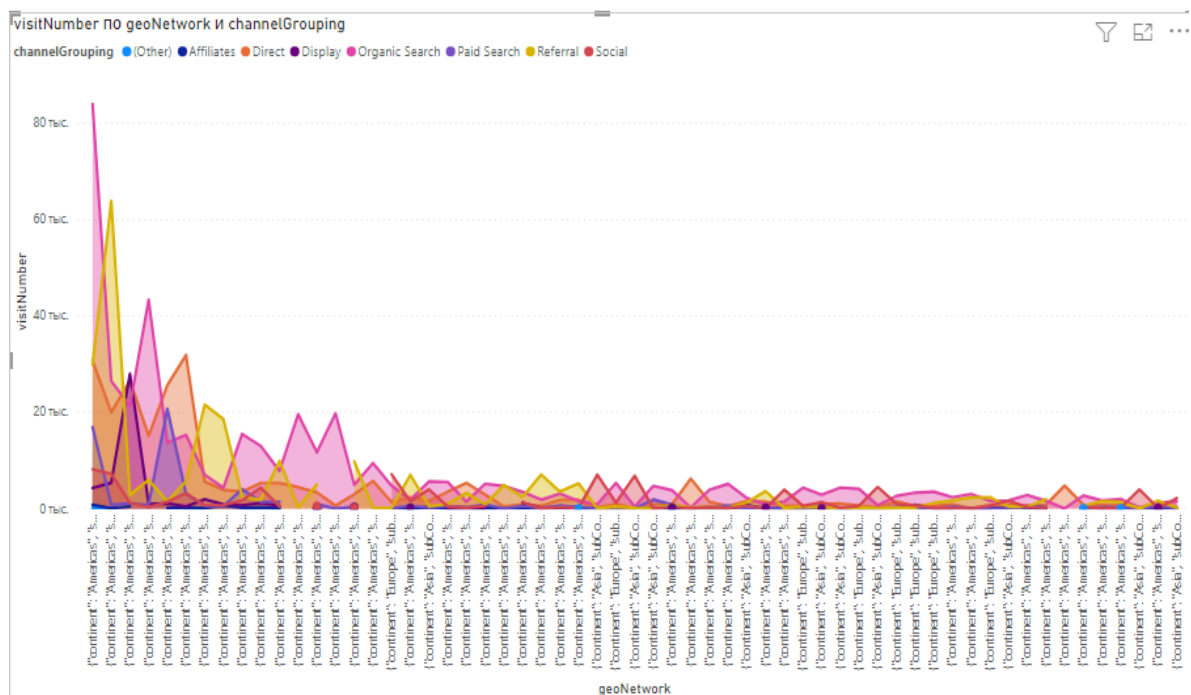


Рисунок 4.16 – Діаграма з областями

Візуалізація у вигляді дерева декомпозиції Power BI дозволяє відображати дані в кількох вимірах. Вона автоматично виконує статистичну обробку даних і дозволяє деталізувати вимірювання в будь-якому порядку. Це інструмент штучного інтелекту, який може знаходити нові виміри для деталізації на основі заданих критеріїв. Така візуалізація є цінним інструментом для глибокого аналізу та виявлення першопричин [28].

Далі розроблено дерево декомпозиції, для початку відобразимо дані, які будуть використані за допомогою таблиці, зображеної на рисунку 4.17.

fullVisitorId	visitId	visitNumber	visitStartTime	channelGrouping	Year	Quarter	Month	Day
270	270	313	404758225337	Social	2017	Qtr 3	July	3
197	197	402	295341236997	Social	2017	Qtr 3	July	4
474	474	520	710660796823	Social	2017	Qtr 3	July	5
366	366	392	548767903594	Social	2017	Qtr 3	July	6
280	280	470	419846249656	Social	2017	Qtr 3	July	7
275	275	299	412373373331	Social	2017	Qtr 3	July	8
231	231	252	346413885935	Social	2017	Qtr 3	July	9
251	251	438	376427786694	Social	2017	Qtr 3	July	10
224	224	250	335955292165	Social	2017	Qtr 3	July	11
267	267	621	400470022144	Social	2017	Qtr 3	July	12
264	264	314	395992070595	Social	2017	Qtr 3	July	13
270	270	319	405015689858	Social	2017	Qtr 3	July	14
235	235	524	352533234540	Social	2017	Qtr 3	July	15
202	202	248	303047050727	Social	2017	Qtr 3	July	16
272	272	428	408086086191	Social	2017	Qtr 3	July	17
243	243	585	364597515472	Social	2017	Qtr 3	July	18
194	194	793	291094586615	Social	2017	Qtr 3	July	19
218	218	499	327125385633	Social	2017	Qtr 3	July	20
177	177	204	265616787057	Social	2017	Qtr 3	July	21
184	184	227	276137853594	Social	2017	Qtr 3	July	22
201	201	492	301667449953	Social	2017	Qtr 3	July	23
195	195	220	292680065391	Social	2017	Qtr 3	July	24
222	222	330	333223274003	Social	2017	Qtr 3	July	25
180	180	756	270196536031	Social	2017	Qtr 3	July	26
213	213	752	319750963943	Social	2017	Qtr 3	July	27
187	187	300	280736582789	Social	2017	Qtr 3	July	28
157	157	486	235712272195	Social	2017	Qtr 3	July	29
188	188	379	282271170932	Social	2017	Qtr 3	July	30
215	215	431	322827296470	Social	2017	Qtr 3	July	31
213	213	397	319844606512	Social	2017	Qtr 3	August	1
951	951	1178	1400669825808	Social	2018	Qtr 3	September	2
903653	903653	2046681	1341931292684163					

Рисунок 4.17 – Вигляд таблиці

На декомпозиційному дереві видно найбільшу кількість відвідувачів за день, що припадає на 24.01.2017. Це зображено на рисунку 4.18.

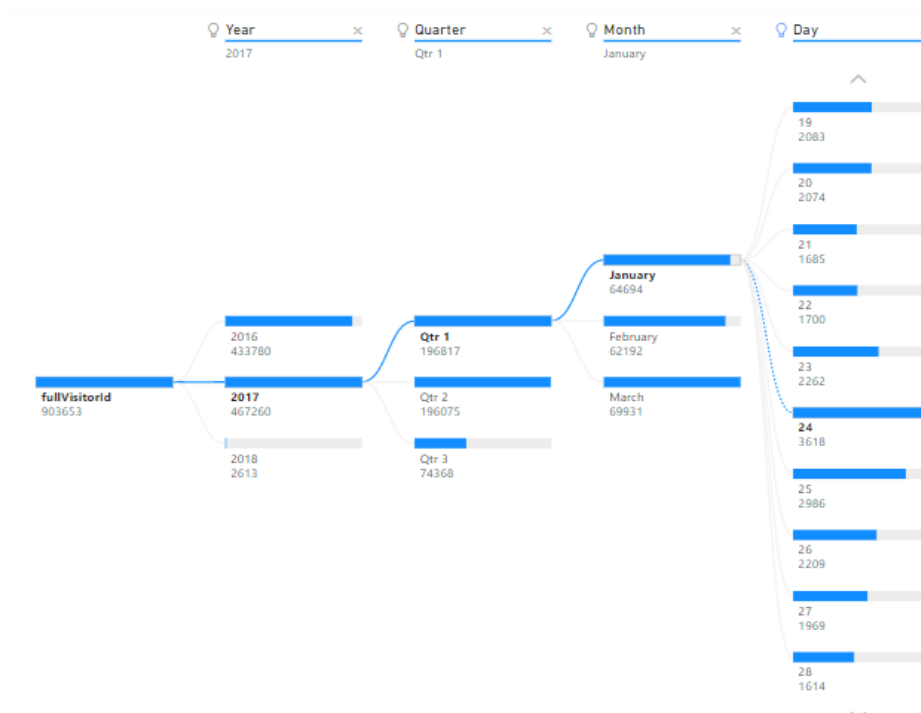


Рисунок 4.18 – Вигляд дерева декомпозиції з найбільшим значенням

Діаграма – графічне представлення числових даних. Діаграми воронки часто використовуються для відображення даних у потоках і процесах. З точки зору відображення, він схожий на кругову діаграму. Діаграма може відображати висоту/ширину сегмента або площу сегмента, пропорційну даним. Діаграми також можуть містити сегменти однакової висоти/ширини без посилання на початок координат. На рисунку 4.19 зображена статистика відвідувань за 2016-2018 роки.

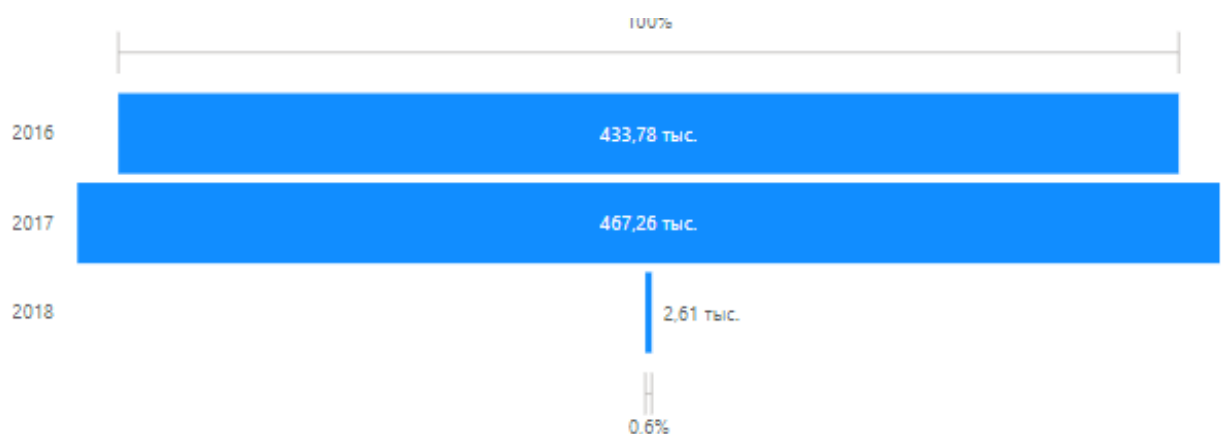


Рисунок 4.19 – Вигляд діаграми воронки

Деревоподібна діаграма відображає ієрархічні дані у вигляді вкладених прямокутників. Кожен рівень ієрархії представлений кольоровим прямокутником (гілкою), що містить менші прямокутники (листя). Power BI розподіляє простір у кожному прямокутнику на основі вимірювань, сортуючи їх від найбільшого у верхньому лівому куті до найменшого у нижньому правому.

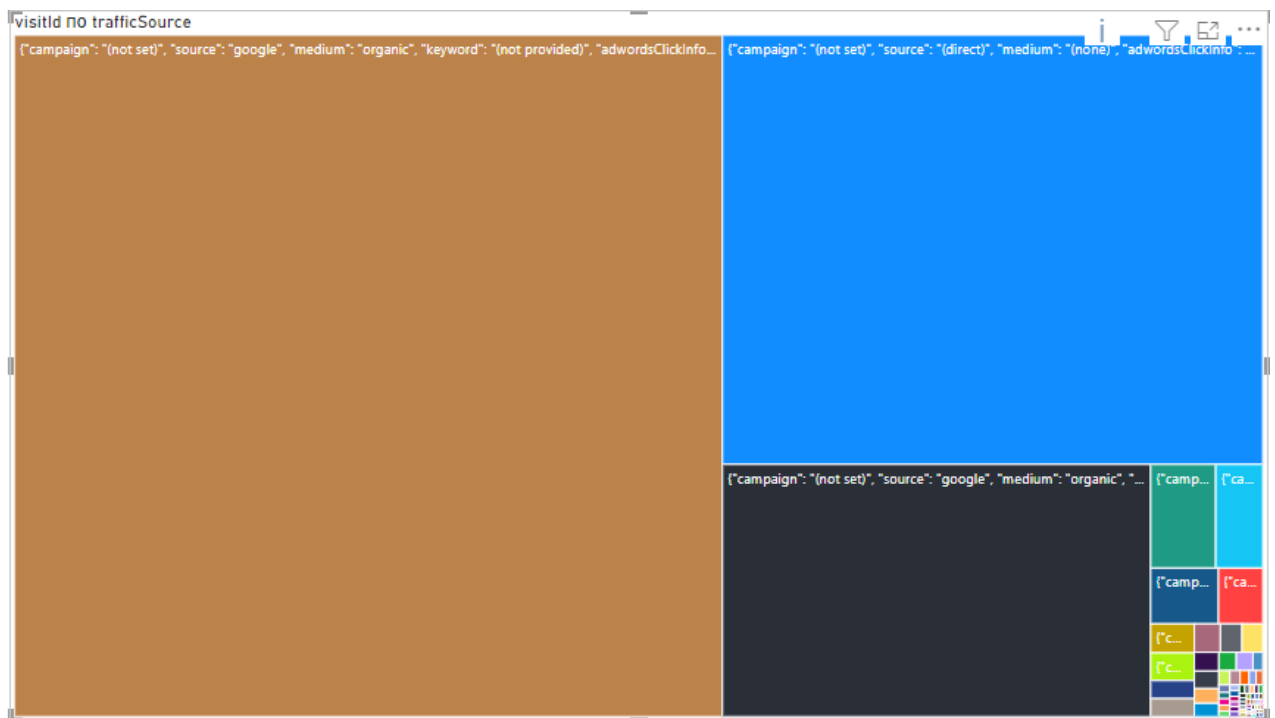


Рисунок 4.20 – Вигляд діаграми дерева

Деревоподібна діаграма показує відношення джерела трафіку до ідентифікатора користувача. Найбільше джерело трафіку – Google. Це візуалізовано на рисунку 4.20.

4.3 Кінцевий дашборд

У результаті створено кінцевий дашборд, який добре візуалізує дані, як показано на рисунку 4.21. Цей звіт містить діаграму з областями із

накопиченням, каскадну діаграму, гістограму з групуванням та кругову діаграму. Другий звіт, зображений у додатку В, містить додаткові візуалізації.

Дашборд – це програмне рішення, яке збирає, аналізує та візуалізує оброблені дані. На практиці це виглядає як панель з простим інтерфейсом, на якій у вигляді діаграм відображаються показники, що змінюються в реальному часі. Програма порівнює одержувані цифри і дає користувачеві «живі» графіки. Для зручності використання і візуалізації можна побачити в дашборді з урахуванням індивідуальних переваг [29].

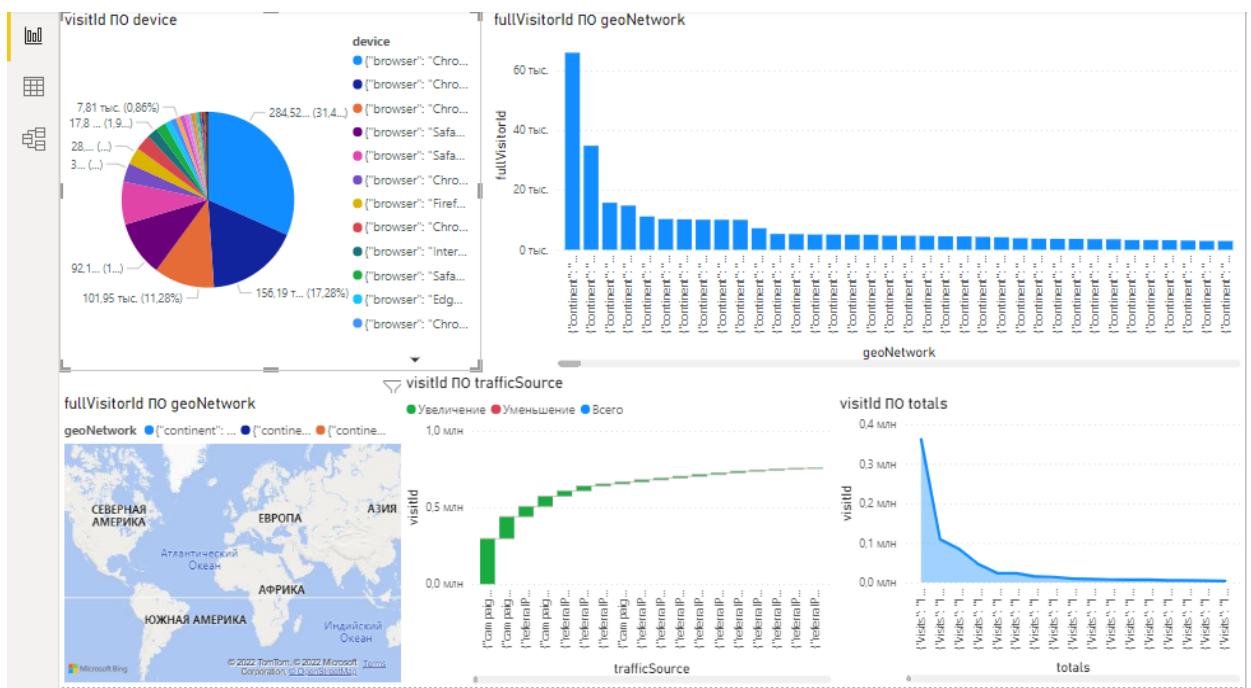


Рисунок 4.21 – Вигляд кінцевого дашборду

Цей звіт містить діаграму з областями із накопиченням, каскадну діаграму, гістограму з групуванням та кругову діаграму. Другий звіт, зображений у додатку В, містить додаткові візуалізації.

4.4 Висновки до розділу

У цьому розділі створено два дашборди для кращого прогнозування даних. Перший дашборд містить діаграму з областями із накопиченням, каскадну діаграму, гістограму з групуванням та кругову діаграму. Другий звіт візуалізує: дерево декомпозиції з найменшим та найбільшим значенням, таблицю та діаграму з областями. Дані прогнозовано за допомогою аналітики Power BI, яка візуалізує ключові характеристики, такі як totals.pageviews та totals.hits. Виявлено, що браузер Chrome є найпопулярнішим для здійснення покупок, на другому місці – Firefox. Візуалізації також показують, що кількість відвідувань збільшується з кожним роком, що чітко видно на порівнянні даних за 2016-2017 роки.

5 ЕКОНОМІЧНИЙ РОЗДІЛ

5.1 Технологічний аудит розробленої системи автоматизації прогнозування продажів

Як було зазначено раніше, сучасний стан економіки багатьох країн світу спонукає продавців товарів вкладати кошти в розвиток так званої онлайн-комерції, оскільки традиційні методи і канали продажу товарів вимагають постійно зростаючих витрат на оренду приміщень, оплату праці торговельного персоналу та посередників тощо. Сучасні фахівці з маркетингу відзначають значні зміни в поведінці споживачів: все більше українців спочатку шукають більш дешеві товари в Інтернеті, а тільки потім приймають рішення про їх покупку. Крім того, спостерігається тенденція згортання або значного скорочення кількості так званих офлайн-майданчиків, де здійснюється реалізація товарів. Все це вимагає докорінних змін в організації продажів товарів різної номенклатури і асортименту.

Тому метою виконаної нами магістерської кваліфікаційної роботи було розроблення системи автоматизації прогнозування продажів Інтернет-магазину шляхом використання сучасних систем автоматизованого прогнозування, які дозволяють аналізувати поведінку клієнтів (покупців) в магазині та їх конкретні дії.

Для досягнення поставленої мети були розв'язані такі задачі: проаналізовано підходи для автоматизації прогнозування продажів товарів; досліджено математичні методи для побудови систем автоматизації прогнозування продажів; розроблено систему автоматизації прогнозування продажів; протестовано систему автоматизації прогнозування продажів інтернет-магазину.

В результаті було розроблено моделі, методи і системи автоматизації прогнозування (далі – програмне забезпечення), на основі яких можна

ухвалювати рішення щодо прогнозування обсягів продажів товарів в Інтернет-магазині.

Для встановлення комерційного потенціалу розробленої нами системи автоматизації продажів було запрошено 3-х відомих експертів – кандидата технічних наук, професора Папінова В.М., кандидата технічних наук, доцента Кабачого В.В. та кандидатку технічних наук, доценткиню Барабан М.В.

Встановлення комерційного потенціалу розробленої системи автоматизації прогнозування продажів було здійснено за критеріями, наведеними в таблиці 5.1,

Таблиця 5.1 – Рекомендовані критерії оцінювання комерційного потенціалу будь-якої розробки і їх бальна оцінка

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
	0	1	2	3	4
Технічна здійсненність концепції:					
1	Достовірність концепції не підтверджується	Концепція підтверджена експертними висновками	Концепція підтверджена розрахунками	Концепція перевірена на практиці	Перевірено роботоздатність продукту в реальних умовах
Ринкові переваги (недоліки):					
2	Багато аналогів на малому ринку	Мало аналогів на малому ринку	Кілька аналогів на великому ринку	Один аналог на великому ринку	Продукт не має аналогів на великому ринку
3	Ціна продукту значно вища за ціни аналогів	Ціна продукту дещо вища за ціни аналогів	Ціна продукту приблизно дорівнює цінам аналогів	Ціна продукту дещо нижче за ціни аналогів	Ціна продукту значно нижче за ціни аналогів
4	Технічні та споживчі властивості продукту	Технічні та споживчі властивості продукту	Технічні та споживчі властивості продукту на	Технічні та споживчі властивості продукту	Технічні та споживчі властивості продукту

	значно гірші, ніж в аналогів	трохи гірші, ніж в аналогів	рівні аналогів	трохи кращі, ніж в аналогів	значно кращі, ніж в аналогів
Ринкові перспективи					
5	Експлуатаційні витрати значно вищі, ніж в аналогів	Експлуатаційні витрати дещо вищі, ніж в аналогів	Експлуатаційні витрати на рівні експлуатаційних витрат аналогів	Експлуатаційні витрати трохи нижчі, ніж в аналогів	Експлуатаційні витрати значно нижчі, ніж в аналогів
6	Ринок малий і не має позитивної динаміки	Ринок малий, але має позитивну динаміку	Середній ринок з позитивною динамікою	Великий стабільний ринок	Великий ринок з позитивною динамікою
7	Активна конкуренція великих компаній на ринку	Активна конкуренція	Помірна конкуренція	Незначна конкуренція	Конкуренція немає
Практична здійсненність					
8	Відсутні фахівці як з технічної, так і з комерційної реалізації ідеї	Необхідно наймати фахівців або витратити значні кошти та час на навчання наявних фахівців	Необхідне незначне навчання фахівців та збільшення їх штату	Необхідне незначне навчання фахівців	Є фахівці з питань як з технічної, так і з комерційної реалізації ідеї

Продовження таблиці 5.1

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
	0	1	2	3	4
9	Потрібні значні фінансові ресурси, які відсутні. Джерела фінансування ідеї відсутні	Потрібні незначні фінансові ресурси. Джерела фінансування відсутні	Потрібні значні фінансові ресурси. Джерела фінансування є	Потрібні незначні фінансові ресурси. Джерела фінансування є	Не потребує додаткового фінансування
10	Необхідна розробка нових матеріалів	Потрібні матеріали, що використовуються у військово-промислому комплексі	Потрібні дорогі матеріали	Потрібні досяжні та дешеві матеріали	Всі матеріали для реалізації ідеї відомі та давно використовуються у виробництві
11	Термін реалізації ідеї більший за 10 років	Термін реалізації ідеї більший за 5 років. Термін окупності інвестицій більше 10-ти років	Термін реалізації ідеї від 3-х до 5-ти років. Термін окупності інвестицій більше 5-ти років	Термін реалізації ідеї менше 3-х років. Термін окупності інвестицій від 3-х до 5-ти років	Термін реалізації ідеї менше 3-х років. Термін окупності інвестицій менше 3-х років
12	Необхідна розробка регламентних документів та отримання великої	Необхідно отримання великої кількості дозвільних документів на виробництво	Процедура отримання дозвільних документів для виробництва та реалізації	Необхідно тільки повідомлення відповідним органам про виробництво та реалізацію продукту	Відсутні будь-які регламентні обмеження на виробництво та

	кількості дозвільних документів на виробництв о та реалізацію продукту	о та реалізацію продукту, що вимагає значних коштів та часу	продукту вимагає незначних коштів та часу		реалізацію продукту
--	---	---	---	--	------------------------

Запрошені експерти оцінили розроблену нами систему автоматизації прогнозування продажів таким чином (див. таблицю 5.2):

Таблиця 5.2 – Результати технологічного аудиту розробленої системи автоматизації прогнозування продажів (за шкалою оцінювання 0-1-2-3-4)

Критерії	Прізвище, ініціали експертів		
	Кабачій В.В.	Папінов В.М.	Барабан М.В.
	Бали, що їх виставили експерти:		
1	4	3	3
2	3	3	4
3	3	3	3
4	3	4	3
5	3	3	3
6	4	4	4
7	3	3	3
8	4	3	3
9	3	4	3
10	3	3	3
11	4	3	3
12	3	3	3
Сума балів	СБ ₁ = 40	СБ ₂ = 39	СБ ₃ = 38
Середньоарифметична сума балів $\overline{СБ}$	$\overline{СБ} = \frac{\sum_{i=1}^3 СБ_i}{3} = \frac{40 + 39 + 38}{3} = \frac{117}{3} = 39,00$		

Встановлення комерційного потенціалу розробленої системи автоматизації прогнозування продажів будемо здійснювати на основі рекомендацій, наведених в таблиці 5.3 [30].

Таблиця 5.3 – Рівні комерційного потенціалу будь-якої наукової розробки

Середньоарифметична сума балів $\overline{СБ}$, розрахована на основі висновків експертів	Рівень комерційного потенціалу розробки
0 – 10	Низький
11 – 20	Нижче середнього
21 – 30	Середній
31 – 40	Вище середнього
41 – 48	Високий

Оскільки середньоарифметична сума балів, що їх виставили експерти, складає 39 балів, то це свідчить, що розроблена система автоматизації прогнозування продажів має рівень комерційного потенціалу, який вважається «вище середнього».

Це пояснюється тим, що розроблена нами система автоматизації прогнозування продажів базується на найсучасніших методах прогнозування, що дозволяє більш ефективно організовувати процес та контролювати результати прогнозування обсягів продажів товарів в Інтернет-магазинах.

5.2 Розрахунок витрат на розроблення системи автоматизації прогнозування обсягів продажів товарів в Інтернет-магазинах

При розробленні системи автоматизації прогнозування обсягів продажів в Інтернет-магазинах були зроблені певні витрати. Зокрема:

А). Основна заробітна плата Z_o розробників, яка визначається за формулою:

$$Z_o = \frac{M}{T_p} \cdot t \text{ грн}, \quad (5.1)$$

де M – місячний посадовий оклад розробника, грн; приймемо, що

$M = (8000 \dots 30000) \text{ грн/місяць};$

T_p – число робочих днів в місяці; прийmemo $T_p = 21$ день;

t – число днів роботи розробників.

Зроблені розрахунки зведемо до таблиці 5.4:

Таблиця 5.4 – Основна заробітна плата розробників

Найменування посади виконавця	Місячний посадовий оклад, грн	Оплата за робочий день, грн	Число днів роботи	Витрати на оплату праці, грн
1. Науковий керівник магістерської роботи	26000	1238,09	20 годин	≈ 4127
2. Магістрант-студент-виконавець	Беремо мінімальну оплату 8000 грн	380,95	85	≈ 32381
3. Консультант з економічної частини	20750	988,09	1,5 години	≈ 247 (при 6-годинному робочому дні)
Загалом				$3_o = 36755$ грн

Б). Додаткова заробітна плата $З_d$ розробників розраховується як (10...12)% від величини їх основної заробітної плати, тобто:

$$З_d = \alpha \cdot З_o = (0,1...0,12) \cdot З_o. \quad (5.2)$$

Прийmemo, що $\alpha = 0,1$. Тоді для нашого випадку отримаємо:

$$З_d = 0,1 \times 36755 = 3675,5 \approx 3676 \text{ грн.}$$

В). Нарахування на заробітну плату $НЗП_{зп}$ розробників (дослідників) розраховуються за формулою:

$$НЗП_{зп} = (З_o + З_d) \cdot \frac{\beta}{100}, \quad (5.3)$$

де β – ставка обов'язкового єдиного внеску на державне соціальне страхування, %. $\beta = 22\%$. Тоді:

$$НЗН_{зп} = (36755 + 3676) \times 0,22 = 8894,82 \approx 8895 \text{ грн.}$$

Г). Амортизація основних засобів А, які використовувались під час виконання цієї роботи:

$$A = \frac{Ц \cdot H_a}{100} \cdot \frac{T}{12} \text{ грн}, \quad (5.4)$$

де Ц – загальна балансова вартість основних засобів, грн;

H_a – річна норма амортизаційних відрахувань. Для нашого випадку можна прийняти, що $H_a = (2,5...25)\%$;

T – термін використання основних засобів, місяці.

Зроблені розрахунки зведено в таблицю 5.5.

Таблиця 5.5 – Розрахунок амортизаційних відрахувань

Найменування обладнання, приміщень тощо	Балансова вартість, грн.	Норма амортизації, %	Термін використання, міс.	Величина амортизаційних відрахувань, грн
1. Комп'ютерна техніка, обладнання, принтери тощо	55000	25	3,2 (при 80% використанні)	2933,33 $\approx$ 2934
2. Приміщення університету, кафедри	22000	3,0	3,2 при 25% використанні	44
Всього				A = 2978 грн

Д). Витрати на матеріали М розраховуються за формулою:

$$M = \sum_1^n H_i \cdot Ц_i \cdot K_i - \sum_1^n B_i \cdot Ц_b \text{ грн.}, \quad (5.5)$$

де H_i – витрати матеріалу i -го найменування, кг; $Ц_i$ – вартість матеріалу i -го найменування; K_i – коефіцієнт транспортних витрат, $K_i = (1,1...1,15)$; B_i – маса відходів матеріалу i -го найменування; $Ц_b$ – ціна відходів матеріалу i -го найменування; n – кількість видів матеріалів.

Е). Витрати на комплектуючі К розраховуються за формулою:

$$K = \sum_{i=1}^n H_i \cdot C_i \cdot K_i \text{ грн}, \quad (5.6)$$

де H_i – кількість комплектуючих i -го виду, шт.; C_i – ціна комплектуючих i -го виду; K_i – коефіцієнт транспортних витрат, $K_i = (1,1 \dots 1,15)$; n – кількість видів комплектуючих.

Під час виконання магістерської кваліфікаційної роботи загальні витрати на матеріали та комплектуючі склали приблизно 1500 грн.

Ж). Витрати на силову електроенергію V_e розраховуються за формулою:

$$V_e = \frac{B \cdot P \cdot \Phi \cdot K_{\pi}}{K_d}, \quad (5.7)$$

де B – вартість 1 кВт-год. електроенергії, в 2024 р. $B \approx 4,5$ грн/кВт;

P – установлена потужність обладнання, кВт; $P = 1,1$ кВт;

Φ – фактична кількість годин роботи обладнання, годин.

Прийmemo, що $\Phi = 315$ годин;

K_{π} – коефіцієнт використання потужності; $K_{\pi} < 1 = 0,82$.

K_d – коефіцієнт корисної дії, $K_d = 0,77$.

Тоді витрати на силову електроенергію будуть дорівнювати:

$$V_e = \frac{B \cdot P \cdot \Phi \cdot K_{\pi}}{K_d} = \frac{4,5 \cdot 1,1 \cdot 315 \cdot 0,82}{0,77} = 1660,5 \approx 1661 \text{ грн.}$$

И). Інші витрати $V_{\text{інш}}$ можна прийняти як $(50 \dots 300)\%$ від основної заробітної плати розробників, тобто:

$$V_{\text{інш}} = (0,5 \dots 3) \times Z_o. \quad (5.8)$$

Для нашого випадку отримаємо:

$$V_{\text{інш}} = 1,50 \times 36755 = 55132,5 \approx 55133 \text{ грн.}$$

К). Сума всіх попередніх статей витрат складає витрати на виконання роботи безпосередньо розробником-магістрантом – B .

$$B = 36755 + 3676 + 8895 + 2978 + 1500 + 1661 + 55133 = 110598 \text{ грн.}$$

Л). Загальні витрати на розробку та можливе впровадження розробленої нами системи автоматизації прогнозування обсягів продажів $V_{\text{заг}}$ розраховуються за формулою:

$$V_{\text{заг}} = \frac{V}{\beta}, \quad (5.9)$$

де β – коефіцієнт, який характеризує етап (стадію) виконання цієї роботи. Можна прийняти, що, $\beta \approx 0,7$ [30].

$$\text{Тоді: } V_{\text{заг}} = \frac{110598}{0,7} = 157997,14 \text{ грн або приблизно 158 тисяч грн.}$$

Тобто прогнозовані загальні витрати на розробку нашої системи автоматизації прогнозування обсягу продажів товарів в Інтернет-магазинах становлять приблизно 158 тисяч грн.

5.3 Розрахунок економічного ефекту від можливої комерціалізації нашої розробки

Економічний ефект від впровадження та можливої комерціалізації розробленої системи автоматизації прогнозування обсягу продажів товарів пояснюється її значно кращими функціональними можливостями. Тому нашу розробку можна реалізовувати на ринку дещо дорожче, ніж аналогічні за функціями розробки. Так, якщо подібна за функціями автоматизована система прогнозування коштує на ринку приблизно 500 тисяч грн, то нашу розробку можна буде реалізовувати на ринку приблизно за 530 тисяч грн або на 30 тисяч грн дорожче.

Аналіз місткості ринку показує, що на сьогодні в Україні кількість потенційних споживачів розробленої нами системи автоматизації прогнозування обсягів продажу може становити приблизно 50 осіб. Це, насамперед, великі Інтернет-магазини (підприємства, що мають значні обсяги продажів та велику кількість товарів); ритейлери (магазини, що прагнуть

оптимізувати свої продажі та зменшити витрати); бізнес-аналітики та маркетингологи (фахівці, що займаються аналізом продажів та поведінки споживачів) та інші потенційні споживачі. Тому можна очікувати зростання попиту на нашу розробку принаймні протягом 3-х років після її впровадження.

Тобто, якщо наша розробка буде впроваджена з 1 січня 2025 року, то її результати будуть виявлятися протягом 2025-го, 2026-го та 2027-го років.

Прогноз зростання попиту на нашу розробку складає по роках:

- а) 2025 р. – приблизно +10 шт. до базового року;
- б) 2026 р. – +15 шт. до базового року;
- в) 2027 р. – +20 шт. до базового року.

Можливе збільшення чистого прибутку $\Delta\Pi_i$, що його може отримати потенційний інвестор від комерціалізації нашої розробки, тобто виведення нашої розробки на ринок, становитиме:

$$\Delta\Pi_i = \sum_1^n (\Delta\Pi_o \cdot N + \Pi_o \cdot \Delta N)_i \cdot \lambda \cdot \rho \cdot \left(1 - \frac{\nu}{100}\right), \quad (5.10)$$

де $\Delta\Pi_o$ – покращення основного якісного показника від впровадження результатів нашої розробки у цьому році. Для нашого випадку це є збільшення ціни реалізації нашої розробки $\Delta\Pi_o = (530 - 500) = + 30$ тисяч грн;

N – основний кількісний показник, який визначає обсяг діяльності у році до впровадження результатів розробки; $N = 50$ шт.;

ΔN – покращення основного кількісного показника від впровадження результатів розробки. Таке покращення становитиме по роках, відповідно: у 2025 році – + 10 шт., у 2026 році +15 шт, та у 2027 році + 20 шт. (відносно базового 2024 року);

Π_o – основний якісний показник (тобто ціна), який визначає обсяг діяльності у році після впровадження результатів розробки, грн; $\Pi_o = 530$ тисяч грн;

n – кількість років, протягом яких очікується отримання позитивних результатів від впровадження розробки; для нашого випадку $n = 3$;

λ – коефіцієнт, який враховує сплату податку на додану вартість;
 $\lambda = 0,8333$;

ρ – коефіцієнт, який враховує рентабельність продукту. Рекомендується приймати $\rho = (0,2 \dots 0,5)$; візьмемо $\rho = 0,4$;

υ – ставка податку на прибуток. У 2024 $\upsilon = 18\%$. У 2025 році і в наступних роках також очікуємо ставку $\upsilon = 18\%$.

Тоді можливе зростання чистого прибутку $\Delta\Pi_1$ для потенційного інвестора протягом першого року від можливого впровадження (комерціалізації) нашої розробки (2025 р.) складе:

$$\Delta\Pi_1 = [30 \cdot 50 + 530 \cdot 10] \cdot 0,8333 \cdot 0,4 \cdot \left(1 - \frac{18}{100}\right) \approx 1859 \text{ тис. грн.}$$

Можливе зростання чистого прибутку $\Delta\Pi_2$ для потенційного інвестора від можливого впровадження (комерціалізації) нашої розробки протягом другого (2026 р.) року складе:

$$\Delta\Pi_2 = [30 \cdot 50 + 530 \cdot 15] \cdot 0,8333 \cdot 0,4 \cdot \left(1 - \frac{18}{100}\right) \approx 2583 \text{ тис. грн.}$$

Можливе зростання чистого прибутку $\Delta\Pi_3$ для потенційного інвестора від можливого впровадження (комерціалізації) нашої розробки протягом третього (2027 р.) року становитиме:

$$\Delta\Pi_3 = [30 \cdot 50 + 530 \cdot 20] \cdot 0,8333 \cdot 0,4 \cdot \left(1 - \frac{18}{100}\right) \approx 3307 \text{ тис. грн.}$$

Приведена вартість зростання всіх чистих прибутків потенційного інвестора від можливого впровадження нашої розробки становитиме:

$$\Pi\Pi = \sum_{i=1}^t \frac{\Delta\Pi_i}{(1 + \tau)^i}, \quad (5.11)$$

де $\Delta\Pi_i$ – збільшення чистого прибутку у кожному із років, протягом яких виявляються результати виконаної та впровадженої роботи, грн;

t – період часу, протягом якого виявляються результати впровадженої роботи, роки. Для нашого випадку $t = 3$ роки;

τ – ставка дисконтування. Прийmemo $\tau = 0,10$ (10%);

t – період часу від моменту початку розроблення системи автоматизації прогнозування продажів до моменту отримання можливих чистих прибутків потенційним інвестором.

Тоді приведена вартість зростання всіх можливих чистих прибутків ПП, що їх може отримати потенційний інвестор від комерціалізації нашої розробки, складе:

$$\text{ПП}^{10\%} = \frac{1859}{(1+0,1)^2} + \frac{2583}{(1+0,1)^3} + \frac{3307}{(1+0,1)^4} \approx 1537 + 1941 + 2259 = 5737 \text{ тис. грн.}$$

Теперішня вартість інвестицій PV , що можуть бути вкладені для реалізації нашої розробки: $PV = (1,0 \dots 5) \times B_{\text{заг.}}$

Для нашого випадку прийmemo:

$$PV = (1,0 \dots 5) \times 158 = 5 \times 158 = 790 \text{ тис. грн.}$$

Розраховуємо абсолютний ефект від можливих вкладених інвестицій $E_{\text{абс.}}$

$$E_{\text{абс}} = \text{ПП} - PV, \quad (5.12)$$

де ПП – приведена вартість збільшення всіх чистих прибутків для інвестора від можливого впровадження нашої розробки, грн;

PV – теперішня вартість інвестицій $PV = 790$ тисяч грн.

Абсолютний ефект від можливого впровадження нашої розробки (при прогнозованому ринку збуту) за три роки складе:

$$E_{\text{абс}} = 5737 - 790 = 4947 \text{ тисяч грн.}$$

Оскільки $E_{\text{абс}} > 0$, то комерціалізація нашої розробки може бути доцільною.

Далі розрахуємо внутрішню дохідність E_v вкладених інвестицій:

$$E_v = \sqrt[T_{\text{ж}}]{1 + \frac{E_{\text{абс}}}{PV}} - 1, \quad (5.13)$$

де $E_{\text{абс}}$ – абсолютний ефект вкладених інвестицій; $E_{\text{абс}} = 4947$ тис. грн;

PV – теперішня вартість початкових інвестицій $PV = 790$ тис. грн;

$T_{\text{ж}}$ – життєвий цикл розробки, роки.

$$T_{\text{ж}} = 4 \text{ років (2024-й, 2025-й, 2026-й, 2027-й роки)}$$

Для нашого випадку отримаємо:

$$E_{\text{в}} = \sqrt[4]{1 + \frac{4947}{790}} - 1 = \sqrt[4]{1 + 6,2620} - 1 = \sqrt[4]{7,2620} - 1 = 1,642 - 1 = 0,642 = 64,2\%.$$

Далі визначимо ту мінімальну дохідність, нижче за яку потенційному інвестору не вигідно буде займатися комерціалізацією нашої розробки.

Мінімальна дохідність або мінімальна (бар'єрна) ставка дисконтування $\tau_{\text{мін}}$ визначається за формулою:

$$\tau_{\text{мін}} = d + f, \quad (5.14)$$

де d – середньозважена ставка за депозитними операціями в комерційних банках; в 2024 році в Україні $d = (0,14...0,18)$;

f – показник, що характеризує ризикованість вкладень; $f = (0,05...0,40)$.

Для нашого випадку отримаємо:

$$\tau_{\text{мін}} = 0,15 + 0,35 = 0,50 \text{ або } \tau_{\text{мін}} = 50\%.$$

Оскільки величина $E_{\text{в}} = 64,2\% > \tau_{\text{мін}} = 50\%$, то потенційний інвестор у принципі може бути зацікавлений у фінансуванні та комерціалізації розробленої нами системи автоматизації прогнозування продажів.

Далі розраховуємо термін окупності коштів, вкладених у можливу комерціалізацію розробленої системи автоматизації прогнозування продажів.

Термін окупності $T_{\text{ок}}$ розраховується за формулою:

$$T_{\text{ок}} = \frac{1}{E_{\text{в}}}. \quad (5.15)$$

Для нашого випадку термін окупності $T_{\text{ок}}$ коштів становитиме:

$$T_{\text{ок}} = \frac{1}{0,642} = 1,56 \text{ років} < 3 \text{ років},$$

що свідчить про потенційну доцільність комерціалізації розробленої нами системи автоматизації прогнозування продажів в Інтернет-магазинах.

Далі проведено моделювання залежності величини внутрішньої дохідності вкладених потенційних інвестицій від рівня інфляції в країні (який в умовах війни може зростати).

Прийнявши рівень інфляції у 20%, отримаємо:

$$\text{ПП}^{20\%} = \frac{1859}{(1+0,2)^2} + \frac{2583}{(1+0,2)^3} + \frac{3307}{(1+0,2)^4} \approx 1291 + 1495 + 1595 = 4381 \text{ тис. грн.}$$

Тоді абсолютний ефект від можливого впровадження нашої розробки за три роки складе:

$$E_{\text{абс}} = 4381 - 790 = 3591 \text{ тисяч грн.}$$

Внутрішня дохідність E_v вкладених інвестицій становитиме:

$$E_v = T_{\text{ж}} \sqrt[4]{1 + \frac{E_{\text{абс}}}{PV}} - 1,$$

де $E_{\text{абс}}$ – абсолютний ефект вкладених інвестицій; $E_{\text{абс}} = 3591$ тисяч грн;

PV –теперішня вартість початкових інвестицій $PV = 790$ тисяч грн.

Для нашого випадку отримаємо:

$$E_v = \sqrt[4]{1 + \frac{3591}{790}} - 1 = \sqrt[4]{1 + 4,5456} - 1 = \sqrt[4]{5,5456} - 1 = 1,534 - 1 = 0,534 = 53,4\%.$$

Прийнявши рівень інфляції у 30%, отримаємо:

$$\text{ПП}^{30\%} = \frac{1859}{(1+0,3)^2} + \frac{2583}{(1+0,3)^3} + \frac{3307}{(1+0,3)^4} \approx 1100 + 1176 + 1158 = 3434 \text{ тис. грн.}$$

Тоді абсолютний ефект від можливого впровадження нашої розробки за три роки складе:

$$E_{\text{абс}} = 3434 - 790 = 2644 \text{ тисячі грн.}$$

Внутрішня дохідність E_v вкладених інвестицій становитиме:

$$E_v = T_{\text{ж}} \sqrt[4]{1 + \frac{E_{\text{абс}}}{PV}} - 1,$$

де $E_{\text{абс}}$ – абсолютний ефект вкладених інвестицій; $E_{\text{абс}} = 2644$ тисяч грн;

PV –теперішня вартість початкових інвестицій $PV = 790$ тисяч грн.

Для нашого випадку отримаємо:

$$E_B = \sqrt[4]{1 + \frac{2644}{790}} - 1 = \sqrt[4]{1 + 3,3468} - 1 = \sqrt[4]{4,3468} - 1 = 1,444 - 1 = 0,444 = 44,4\%.$$

Зроблені розрахунки у вигляді графіків наведено на рис. 5.1.

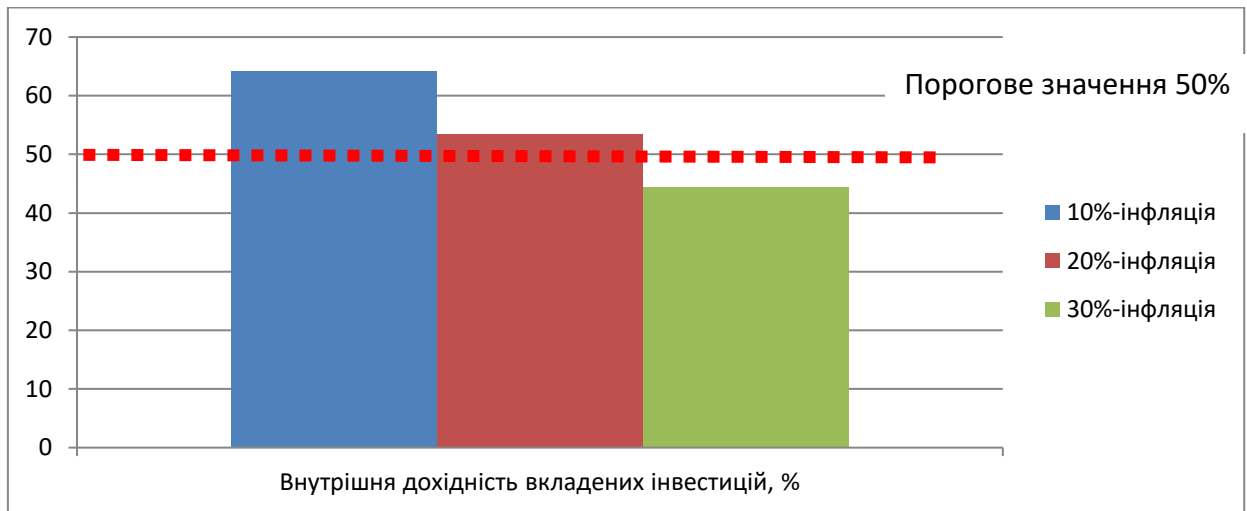


Рисунок 5.1 – Моделювання залежності величини внутрішньої дохідності потенційних інвестицій від рівня інфляції в країні

Аналіз графіка на рис 5.1 показує, що при рівні інфляції в 10% величина внутрішньої дохідності інвестицій становить $E_B = 64,2\%$, що більше порогового значення $\tau_{\min} = 50\%$ і тому комерціалізація нашої розробки може бути доцільною. При рівні інфляції в 20% величина внутрішньої дохідності інвестицій, вкладених в комерціалізацію нашої розробки, становитиме $E_B = 53,4\%$, що також більше порогового значення $\tau_{\min} = 50\%$, і тому комерціалізація нашої розробки (при нинішніх умовах) може бути для інвестора доцільною. При рівні інфляції у 30% величина внутрішньої дохідності інвестицій, вкладених в можливу комерціалізацію нашої розробки, може становити $E_B = 44,4\%$, що дещо менше порогового значення $\tau_{\min} = 50\%$. Для прийняття остаточного рішення потрібно провести додаткові обґрунтування і розрахунки (наприклад, знизити рівень прийнятого ризику, підняти ціну реалізації нашої розробки тощо).

Результати виконаної економічної частини магістерської кваліфікаційної роботи зведено у таблицю:

Показники	Задані у ТЗ	Досягнуті у магістерській кваліфікаційній роботі	Висновок
1. Витрати на розробку	Приблизно 150 тис. грн	158 тис. грн.	Практично досягнуто
2. Абсолютний ефект від впровадження розробки, тисяч грн	В межах 5000 тисяч грн (за три роки)	4947 тисяч грн	Практично виконано
3. Внутрішня дохідність інвестицій, %	не менше 50%	64,2%	Досягнуто
4. Термін окупності інвестицій, роки	до 3-ти років	1,56 років	Виконано

Таким чином, основні техніко-економічні показники розробленої нами системи автоматизації прогнозування продажів в Інтернет-магазинах, визначені у технічному завданні, виконані.

ВИСНОВКИ

В результаті виконання дипломної роботи було доведено актуальність теми системи автоматизації прогнозування продажів.

У першому розділі було проведено аналіз різних підходів до автоматизації прогнозування продажів. Розглянуті такі системи, як SFA і SFMS, їхні переваги та недоліки. Також досліджено CBT Demand Forecasting та Streamline. Результати цього аналізу допоможуть вибрати оптимальний підхід для підвищення ефективності прогнозування продажів.

У другому розділі проведено дослідження математичних методів, які можуть бути використані для побудови систем автоматизації прогнозування продажів. Було проаналізовано такі методи, як аналіз часових рядів, модель авторегресії, модель ковзного середнього та метод декомпозиції. Отримані результати стануть основою для подальших досліджень та розробки ефективної системи прогнозування продажів.

В третьому розділі розглянуто оптимальні параметри (показники) для автоматизованог прогнозування продажів в Інтернет-магазинах. Далі розроблено алгоритм роботи магазину, пошук та збір даних для прогнозування. Також обрано програму, в якій візуалізовано дані, ця програма – Power BI.

Четвертий розділ присвячено конструюванню дашбордів за допомогою функцій DAX, що допомогло прогнозувати поведінку покупців Інтернет-магазину Google. Для кращого прогнозування даних також здійснено аналіз продаж Інтернет-магазину Google на світовому ринку.

В результаті прогнозовано дані за допомогою аналітики в Power BI, на якому візуалізовано важливі характеристики, на першому та другому місці : totals.pageviews та totals hits. Завдяки цьому визначено найпопулярніші серед користувачів браузер через який здійснюється найбільше покупок – це Chrome, а на другому місці Firefox. На візуалізаціях добре видно, що кількість

відвідувань з кожним роком збільшується, це можна спостерігати порівнюючи 2016-2017 роки, на яких це добре показано.

За результатами даної роботи була зроблена доповідь на тему «Система автоматизації прогнозування продажів» на конференції «Молодь в науці: дослідження, проблеми, перспективи» факультету інтелектуальних інформаційних технологій та автоматизації Вінницького національного технічного університету (2024) з публікацією тез.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Кабачій В.В., Войтюк М.В Система автоматизації прогнозування продажів. «Молодь в науці: дослідження, проблеми, перспективи» конференції факультету інтелектуальних інформаційних технологій та автоматизації Вінницького національного технічного університету [Електронний ресурс]. URL: <https://conferences.vntu.edu.ua/index.php/mn/mn2024/paper/view/21117>
2. Крупчатнікова Т.С., Буга Н.Ю. SFA-системи як інструмент підвищення ефективності продаж.
3. About the SFMS [Електронний ресурс]. Режим доступу: <https://fiu.gov.ua/en/pages/pro-sluzhbu/zagalni-vidomosti/pro-derzhfinmonitoring.html>
4. CBT SCM Demand Forecasting [Електронний ресурс]. Режим доступу: <https://cbt.ua/product/cbt-demand-forecasting/>
5. 10 найкращих програм для прогнозування продажів у 2024 році [Електронний ресурс]. Режим доступу: <https://gmdhsoftware.com/ua/sales-forecasting-software/>
6. Power BI [Електронний ресурс]. Режим доступу: https://uk.wikipedia.org/wiki/Power_BI
7. Часові ряди та що з ними робити [Електронний ресурс]. Режим доступу: <https://dou.ua/forums/topic/40751/>
8. Про схвалення Концепції створення системи рейтингового оцінювання регіонів, галузей національної економіки, суб'єктів господарювання: Розпорядження Кабінету Міністрів України від 01.04.2004 №208-р // zakon1.rada.gov.ua
9. Моделі і методи авторегресії [Електронний ресурс]. Режим доступу: https://stud.com.ua/52035/ekonomika/modeli_metodi_avtoregresiyi

10. Моделі ковшної середньої [Електронний ресурс]. Режим доступу: https://stud.com.ua/72679/ekonomika/modeli_kovznoyi_serednoyi
11. Прогнозування продажів: значення та методи [Електронний ресурс]. Режим доступу: <https://www.businessmanagementideas.com/sales/forecastingsales/sales-forecasting-meaning-importance-and-methods/7122>
12. Методики та інструменти для прогнозування продажів: класика та сучасні підходи [Електронний ресурс]. Режим доступу: <https://it-rating.ua/metodiki-ta-instrumenti-dlya-prognozuvannya-prodajiv-klasika-ta-suchasni-pidhodi>
13. Що найкраще продавати в Інтернеті у 2022 році [Електронний ресурс]. URL: <https://horoshop.ua/ua/blog/chto-prodavati-v-internete/>
14. Крижановський Є. М. Яцолт А. Р. Жуков С. О. Козачко О.М. Вінниця Моделювання бізнес-процесів та управління ІТ-проектами: ВНТУ, Вінниця, 2018. 91 с.
15. Застосування UML [Електронний ресурс]. URL: <https://dut.edu.ua/ua/news-1-626-7758-zastosuvannya-uml-v-diplomnih-robotah-kafedra-kompyuternih-nauk-ta-informaciyних-tehnologiy>
16. Теорія та практика UML [Електронний ресурс]. URL: http://www.it-gost.ru/articles/view_articles/94.
17. Дячун О.Д. Прогнозування продажу та його методи в системі управління підприємством: Київ, 2018, 129 с.
18. Kaggle – Вікіпедія [Електронний ресурс]. URL: <https://uk.wikipedia.org/wiki/Kaggle>
19. Machine Learning – Машинне навчання [Електронний ресурс]. URL: <https://www.it.ua/knowledge-base/technology-innovation/machine-learning>
20. Програмування числових методів мовою Python : навч. посіб. за ред. : А. Ю. Дорошенко, С. Д. Погорілий, Я. Ю. Дорогий, Є. В. Глушко, А. В. Анісімова. Київ, 2013. – 463 с.

21. Google Analytics Client Revenue Forecast [Електронний ресурс].
URL: <https://www.kaggle.com/competitions/ga-customer-revenue-prediction/data>

22. Microsoft Excel [Електронний ресурс]. URL:
<https://support.microsoft.com/uk-ua/office/загальні-відомості-про-excel-starter-601794a9-b73d-4d04-b2d4-eed4c40f98be>

23. Alberto Ferrari and Marco Russo Data Analysis with Microsoft Power BI and Power Pivot for Excel. from eng. A. Yu. Ginko. М.: DMK Press, 2020. 288 р.: ill.

24. Кругова діаграма [Електронний ресурс]. URL:
https://uk.wikipedia.org/wiki/Секторна_діаграма

25. Гістограма [Електронний ресурс]. URL:
<https://sites.google.com/site/galereadiagram1/gistogrami>

26. Каскадна діаграма [Електронний ресурс]. URL:
<https://programka.com.ua/instrukcija/excel/kak-postroit-vodopad-v-excel>

27. Діаграма з областями з накопиченням [Електронний ресурс].
URL: <https://sites.google.com/site/galereadiagram1/diagrami-z-oblastami>

28. Дерево декомпозиції в Power BI [Електронний ресурс]. URL:
<https://docs.microsoft.com/power-bi/visuals/power-bi-visualization-decomposition-tree>

29. Дашборд - що це таке і для чого потрібний [Електронний ресурс].
URL: <https://waytobi.com/ua/blog/kpi-dashboards.html>

30. Методичні вказівки до виконання економічної частини магістерських кваліфікаційних робіт. / Укладачі В.О. Козловський, О.Й. Лесько, В.В.Кавецький. – Вінниця : ВНТУ, 2021. – 42 с.

Додаток А
(обов'язковий)

Міністерство освіти і науки України
Вінницький національний технічний університет
Факультет інтелектуальних інформаційних технологій та автоматизації

ЗАТВЕРДЖУЮ

Завідувач кафедри АІТ

_____ проф. Олег БІСІКАЛО

« ____ » _____ 2024 р.

ТЕХНІЧНЕ ЗАВДАННЯ

на виконання магістерської кваліфікаційної дипломної роботи

Система автоматизації прогнозування продажів

08-31.МКР.001.00.000 ТЗ

Керівник: к.т.н., доц.

_____ Владислав КАБАЧІЙ

« __ » _____ 2024 р.

Розробив студент гр. АКІТ-22мз

_____ Марія ВОЙТЮК

« __ » _____ 2024 р.

Вінниця 2024

1. Підстава для проведення робіт

Підставою для виконання роботи є наказ №81 по ВНТУ «11» березня 2024 року, та індивідуальне завдання на МКР, затверджене протоколом №__ засідання кафедри АІТ від «__» _____ 2024р.

2. Джерела розробки

1) Alberto Ferrari and Marco Russo Data Analysis with Microsoft Power BI and Power Pivot for Excel. from eng. A. Yu. Ginko. М.: DMK Press, 2020. 288 p.

2) Google Analytics Client Revenue Forecast [Електронний ресурс]. URL: <https://docs.google.com/document/d/1Gp4qTRymLYteHT4DJd9r5gHfW7BsXWtcMJznEZFXgzA/edit>

3. Мета і призначення роботи.

Метою роботи є розроблення системи автоматизації прогнозування продаж Інтернет-магазину Google за рахунок використання системи автоматизованого прогнозування.

Вихідні дані для проведення робіт:

Набір даних зібраний з Kaggle, який містить інформацію про продажі Google Shopping.

4. Методи дослідження:

В даній роботі використовується конструювання дашбордів для розроблення системи автоматизації прогнозування продажів Інтернет-магазину.

5. Етапи роботи і терміни їх виконання

а) Аналіз підходів для автоматизації прогнозування продажів.....12.03.2024 – 26.03.2024

б) Дослідження математичних методів для побудови систем автоматизації прогнозування продажів.....27.03.2024 – 15.04.2024

с) Розробка системи автоматизації прогнозування продажів.....16.04.2024 – 02.05.2024

- d) Тестування системи автоматизації прогнозування продажів
Інтернет-магазину.....03.05.2024 – 17.05.2024
- e) Економічний розділ..... 20.05.2024 – 27.05.2024
- f) Оформлення пояснювальної записки.....28.05.2024 – 05.06.2024
- g) Попередній захист роботи.....06.06.2024
- h) Остаточний захист роботи.....20.06.2024

7. Очікувані результати та порядок реалізації

Побудувати дашборди за допомогою функцій DAX, використовуючи дані про Інтернет-магазин Google за період з 01.01.2016 по 31.01.2019.

8. Вимоги до розробленої документації

Пояснювальна записка оформлена у відповідності до вимог «Методичних вказівок до виконання та оформлення магістерських кваліфікаційних дипломних робіт для студентів спеціальності 151 – «Автоматизація та комп'ютерно-інтегровані технології».

9. Порядок приймання роботи

Публічний захист	« <u>20</u> » <u>червня</u> 2024 р.
Початок розробки	« <u>12</u> » <u>березень</u> 2024 р.
Граничні терміни виконання МКР	« <u>13</u> » <u>червня</u> 2024 р.
Розробив студент групи АКІТ-22мз _____	Марія ВОЙТЮК

Додаток Б
(обов'язковий)

ІЛЮСТРАТИВНА ЧАСТИНА

Система автоматизації прогнозування продажів

Виконав: студент гр. АКІТ-22мз

_____ Марія ВОЙТЮК

«__» _____ 2024 р.

Керівник: к.т.н., доцент

_____ Владислав КАБАЧІЙ

«__» _____ 2024 р.

sessionId	socialEngagementType	totals	trafficSource	visitId	visitNumber	visitStartTime
028354173102_1508204646	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508204646	1	1508204646
772519618990_1508201418	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508201418	1	1508201418
775969328107_1508160141	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508160141	1	1508160141
604205263760_1508186392	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508186392	1	1508186392
191856926805_1508143388	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508143388	1	1508143388
984772719982_1508161349	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508161349	1	1508161349
87682879_1508208588	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508208588	1	1508208588
525263048024_1508195520	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508195520	1	1508195520
830891089041_1508178277	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1508178277	1	1508178277
674758444227_1512098799	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512098799	1	1512098799
374083102098_1512079193	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512079193	1	1512079193
163020859873_1512088598	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512088598	1	1512088598
102348987851_1512068458	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512068458	1	1512068458
171924743592_1512079471	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512079471	1	1512079471
603957661056_1512056872	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512056872	1	1512056872
348068561054_1512100608	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512100608	1	1512100608
46424878052_1512075096	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512075096	1	1512075096
238460421376_1512104272	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1512104272	1	1512104272
187482980969_1523810336	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1523810336	1	1523810336
048192349141_1523840386	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1523840386	1	1523840386
791710381494_1523807343	Not Socially Engaged	("visits": "1", "hits": "1", "pageviews": "1", "bounces": "1")	("campaign": "(not set)", "source": "google", "medium": "organic")	1523807343	1	1523807343

Рисунок Б.1 – Вигляд імпортованих даних (2)

Name	Name_1	dtypes	Missing	FutImportance	First Value	Second Value	Third Value
0	channelGrouping	object	0	137	Organic Search	Organic Search	Organic Search
1	date	int64	0	0	20160902	20160902	20160902
2	fullVisitorId	object	0	0	1,13166E+18	3,77306E+17	3,89555E+18
3	sessionId	object	0	2	1131660440785968503_1472830385	377306020877927890_1472880147	3895546263509774583_1472865386
4	socialEngagementType	object	0	0	Not Socially Engaged	Not Socially Engaged	Not Socially Engaged
5	visitId	int64	0	0	1472830385	1472880147	1472865386
6	visitNumber	int64	0	585	1	1	1
7	visitStartTime	int64	0	1058	1472830385	1472880147	1472865386
8	device.browser	object	0	71	Chrome	Firefox	Chrome
9	device.browserVersion	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
10	device.browserSize	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
11	device.operatingSystem	object	0	288	Windows	Macintosh	Windows
12	device.operatingSystemVersion	object	0	0	not available in demo dataset	not available in demo dataset	not available in demo dataset
13	device.isMobile	bool	0	2	False	False	False
14	device.mobileDeviceBranding	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
15	device.mobileDeviceModel	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
16	device.mobileInputSelector	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
17	device.mobileDeviceInfo	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
18	device.mobileDeviceMarketingName	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
19	device.flashVersion	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
20	device.language	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
21	device.screenColors	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset
22	device.screenResolution	object	0	1	not available in demo dataset	not available in demo dataset	not available in demo dataset

Рисунок Б.2 – Таблиця з найбільшими показниками

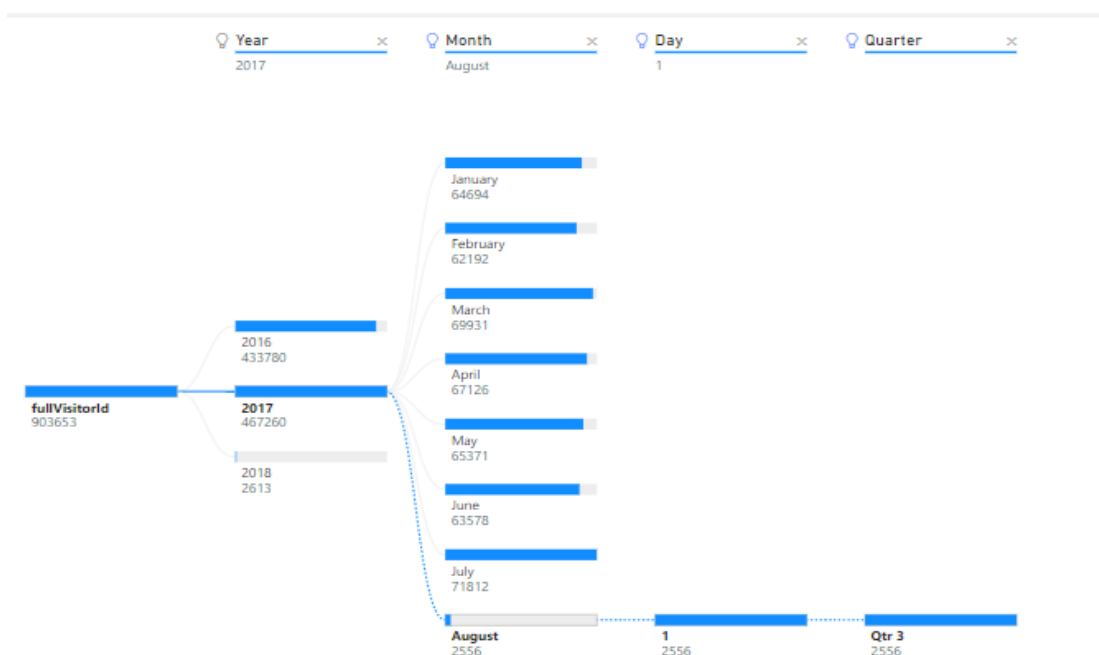


Рисунок Б.5 – Вигляд дерева декомпозиції з найменшим значенням

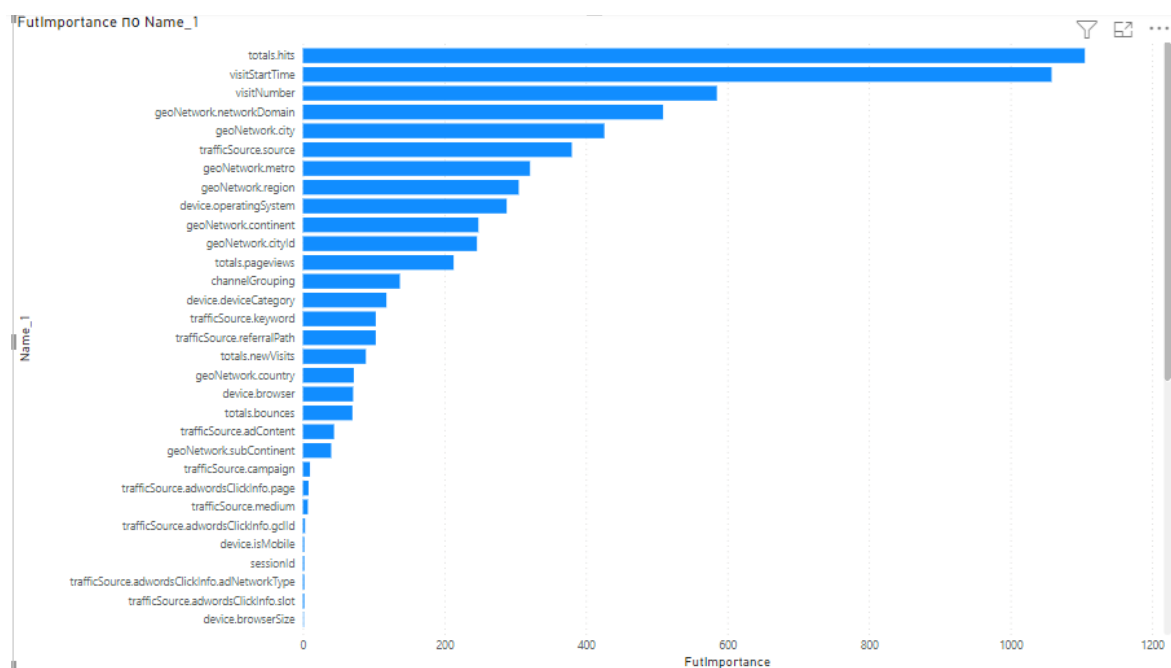


Рисунок Б.6 – Відображення даних за допомогою аналітики в Power BI

Додаток В
(обов'язковий)

Протокол перевірки магістерської кваліфікаційної дипломної роботи

Назва роботи: «Система автоматизації прогнозування продажів»

Тип роботи: магістерська кваліфікаційна дипломна робота

Підрозділ: кафедра АІТ

Науковий керівник: Владислав КАБАЧІЙ, к.т.н., доцент

Показники звіту подібності

Unichек	
Оригінальність	90,1%
Схожість	9,9%

Аналіз звіту подібності (відмітити потрібне)

- ☐ Запозичення, виявлені у роботі, оформлені коректно і не містять ознак плагіату.
- ☐ Виявлені у роботі запозичення не мають ознак плагіату, але їх надмірна кількість викликає сумніви щодо цінності роботи і самостійності її автора. Роботу направити на доопрацювання.
- ☐ Виявлені у роботі запозичення є недобросовісними і мають ознаки плагіату та/або в ній містяться навмисні спотворення тексту, що вказують на спроби приховування недобросовісних запозичень.

Заявляю, що ознайомлений з повним звітом подібності, який був згенерований системою щодо роботи

Автор _____ Марія ВОЙТЮК
(підпис)

Опис прийнятого рішення

Робота допускається до захисту

Особа, відповідальна за перевірку _____ Роман МАСЛІЙ
(підпис)