

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
імені ВАДИМА ГЕТЬМАНА

**Навчально-науковий інститут
«Інститут інформаційних технологій в економіці»**

Кафедра **системного аналізу та кібербезпеки**
(назва кафедри)

ОСВІТНЬО-ПРОФЕСІЙНА ПРОГРАМА

«Системний аналіз»

СПЕЦІАЛЬНІСТЬ

124 «Системний аналіз»

Форма навчання: денна
очна (денна), заочна, дистанційна

КВАЛІФІКАЦІЙНА МАГІСТЕРСЬКА РОБОТА

на тему « Системний аналіз бізнес-процесів у виробничій сфері »
(назва теми)

здобувача Ситника Романа Олександровича
(ПІБ, підпис)

Науковий керівник: д. ф.-м. н., проф. Колечкіна Л. М.
(науковий ступінь, учене звання, ПІБ)

(*nidnuc*)

Робота допущена до захисту перед екзаменаційною комісією з атестації здобувачів вищої освіти (ЕК)

Завідувач кафедри: д.ф.-м.н., проф. Джалладова І.А
(науковий ступінь, учене звання, ПІБ)

(*nidnuc*)

Київ 2024

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
імені ВАДИМА ГЕТЬМАНА**

**Навчально-науковий інститут
«Інститут інформаційних технологій в економіці»**

Кафедра системного аналізу та кібербезпеки
(назва кафедри)

ОСВІТНЬО-ПРОФЕСІЙНА ПРОГРАМА

«Системний аналіз»

СПЕЦІАЛЬНІСТЬ

124 «Системний аналіз»

ПОГОДЖЕНО

Керівник проектної групи (гарант) освітньо-
професійної програми

ЗАТВЕРДЖУЮ

Завідувач кафедри

(підпис) (ініціали, прізвище)
_____ 20__ р.

(підпис) (ініціали, прізвище)
_____ 20__ р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

здобувачу вищої освіти _____ Ситника Романа Олександровича
(прізвище, ім'я, по батькові)

_____ **денної форми навчання**
очної (денної), заочної, дистанційної

на підготовку кваліфікаційної магістерської роботи
на тему «_____ Системний аналіз бізнес-процесів у виробничій сфері _____»
(назва теми)

Тему затверджено наказом ректора Університету від "_____" _____ 20__ р. № _____

Кваліфікаційна магістерська робота виконується на матеріалах

План кваліфікаційної магістерської роботи

Розділ 1	Теоретичні засади системного аналізу та моделювання даних.
Розділ 2	Аналіз практичних засад та інструментів моделювання даних і бізнес-аналітики
Розділ 3	Прикладна обробка даних і створення моделі даних з використання Power BI та Excel

Об'єкт дослідження:	Процес розробки та створення моделі даних з урахуванням засад самого моделювання, а також особливостей датасетів
Предмет дослідження:	Теоретичні та практичні аспекти формування, розробки та створення моделей даних
Мета кваліфікаційної магістерської роботи:	Здійснити покроковий аналіз бізнес-процесу обробки та аналізу даних, розробки та створення моделі даних з використанням технологій програмного забезпечення Microsoft Power BI, а також визначення їх впливу на ефективність та результативність управління підприємством

Конкретні завдання, які здобувач повинен виконати для досягнення поставленої мети:

У розділі 1

Визначити основні поняття даних та системного аналізу

Описати організацію та аналіз життєвого циклу даних у виробничій сфері

Огляд концепцій та методів моделювання даних у виробничій сфері

У розділі 2

Визначити архітектуру моделювання та зберігання даних при роботі з моделями даних

Розглянути та описати Excel як засіб для маніпулювання даними

Розглянути та описати Power BI як інструмент побудови аналітичних моделей даних

У розділі 3

Виконати передобробку та зведення даних за допомогою Excel

Розробити модель даних з використанням засобів Microsoft Power BI

Створити візуальне представлення даних та здійснити аналіз отриманих результатів

**Завдання підготував
науковий керівник**

(підпис)

Колєсчкіна Л. М.

(ініціали, прізвище)

«_____» _____ 2024 р.

**Завдання одержав
здобувач**

(підпис)

Ситник Р. О.

(ініціали, прізвище)

«_____» _____ 2024 р.

Реферат

Кваліфікаційна магістерська робота містить 101 сторінку, 19 рисунків, перелік джерел посилання з 55 найменувань.

«Системний аналіз бізнес-процесів у виробничій сфері»

Об'єктом дослідження є процес розробки та створення моделі даних з урахуванням засад самого моделювання, а також особливостей датасетів.

Предметом дослідження є теоретичні та практичні аспекти формування, розробки та створення моделей даних.

Мета кваліфікаційної магістерської роботи – здійснення покрокового аналізу бізнес-процесу обробки та аналізу даних, розробки та створення моделі даних з використанням технологій програмного забезпечення Microsoft Power BI, а також визначення їх впливу на ефективність та результативність управління підприємством.

Відповідно до поставленої мети були визначені такі *завдання*:

- визначити основні поняття даних та системного аналізу
- описати організацію та аналіз життєвого циклу даних у виробничій сфері
- огляд концепцій та методів моделювання даних у виробничій сфері
- визначити архітектуру моделювання та зберігання даних при роботі з моделями даних
- розглянути та описати Excel як засіб для маніпулювання даними
- розглянути та описати Power BI як інструмент побудови аналітичних моделей даних
- виконати передобробку та зведення даних за допомогою Excel
- розробити модель даних з використанням засобів Microsoft Power BI
- створити візуальне представлення даних та здійснити аналіз отриманих результатів

Практичне значення отриманих результатів. На базі принципів аналізу та моделювання була розроблена функціонуюча модель даних з використанням Power

ВІ, яка містить дані продажів кондитерських виробів на ринку Болгарії як локального виробництва, так і товарів імпортованих закордонними виробниками.

Описана і застосована методика аналізу даних, розробки та створення моделі даних може бути використана на практиці у виробничій сфері, щоб забезпечити підґрунття для формування важливої аналітичної звітності.

Рік виконання кваліфікаційної магістерської роботи 2024.

Рік захисту роботи 2024.

Ключові слова: дані, системний аналіз, модель даних, сховище даних, датасети, обробка датасетів, зведена таблиця.

Відгук
про кваліфікаційну магістерську роботу
здобувача освітньо-професійної програми «Системний аналіз»
навчально-наукового інституту «Інформаційних технологій в економіці»
Ситника Романа Олександровича
на тему «Системний аналіз бізнес-процесів у виробничій сфері»

1. Актуальність теми: дипломна робота та тематика, яка розкривається в ній, є надзвичайно актуальною в умовах, коли виробничі підприємства стикаються з великим обсягом даних, що генеруються щодня. Використання системного аналізу дозволяє ефективно управляти цими даними, оптимізуючи ресурси та підвищуючи продуктивність. У світі, де рішення все більше базуються на аналітичних даних, впровадження моделей даних та системного підходу стає критично важливим для конкурентоспроможності підприємств. Застосування інструментів, таких як Microsoft Power BI, демонструє практичну значущість і необхідність роботи.

2. Позитивні риси кваліфікаційної роботи: у роботі було дано чітке визначення термінології та складових бізнес-процесу аналізу та моделювання даних, розкрито їх методи, інструменти, принципи, недоліки та переваги. Якість проведеного дослідження підтверджується реальними прикладами використання Power BI для моделювання та візуалізації даних. Практична частина роботи є відмінно документованою, що дозволяє легко відтворити результати і використовувати їх у подальших дослідженнях.

3. Наявність самостійних розробок автора: Були проведені самостійні дослідження та експерименти з моделювання даних на базі реальних датасетів, що підтверджує високий рівень його компетентності. Важливою розробкою є створення моделі даних з використанням Microsoft Power BI, що дозволяє ефективно візуалізувати та аналізувати дані виробничих процесів.

4. Цінність теоретичних висновків та практичних рекомендацій: В роботі систематизуються знання про моделювання даних у виробничій сфері та системний аналіз. Практичні рекомендації, надані автором, мають високу прикладну цінність і можуть бути безпосередньо впроваджені на виробничих підприємствах.

5. Наявність недоліків: Зауваження в роботі стосуються лише можливості більш докладного розгляду набору візуалізацій для кращої демонстративності результатів опрацьованих даних.

6. Загальна оцінка кваліфікаційної магістерської роботи та її допущення до захисту перед ЕК: тема роботи розкрита, основні положення та висновки повністю обґрунтовані. Робота Ситника Романа на тему «Системний аналіз бізнес-процесів у виробничій сфері» може бути допущена до захисту перед ЕК і заслуговує на високу оцінку, а її автору може бути присвоєно ступінь магістра зі спеціальності «Системний аналіз».

Науковий керівник: д. ф.-м. н., професор кафедри системного аналізу та кібербезпеки

(підпис)

Колєчкіна Л. М.

«____» _____ 2024р.

Рецензія

на кваліфікаційну магістерську роботу

здобувача вищої освіти

Ситника Романа Олександровича

Тема: «Системний аналіз бізнес-процесів у виробничій сфері»

Актуальність теми кваліфікаційної роботи і доцільність її розроблення. Тема системного аналізу бізнес-процесів у виробничій сфері є надзвичайно актуальною в умовах сучасного економічного розвитку, коли ефективне управління даними стає ключовим фактором успіху. Виробничі підприємства постійно шукають шляхи підвищення продуктивності та оптимізації процесів, і системний аналіз надає інструменти для досягнення цих цілей. Розроблення даної теми дозволяє знайти нові підходи до управління виробництвом, що підвищує конкурентоспроможність підприємств. Це дослідження є своєчасним і важливим для практичного застосування в різних виробничих галузях.

Якість проведеного дослідження підтверджується детальним аналізом та систематизацією теоретичних знань у сфері системного аналізу. Автор грамотно поєднав теоретичні аспекти з практичними прикладами, що демонструє його глибоке розуміння теми. Використання сучасних інструментів, таких як Microsoft Power BI, для моделювання даних свідчить про здатність автора працювати з актуальними технологіями. Результати дослідження є добре структурованими та логічно викладеними, що полегшує їх розуміння та використання.

До позитивних рис роботи можна віднести глибокий теоретичний аналіз та чітку структуру викладення матеріалу. Автор вдало поєднав теоретичні знання з практичними дослідженнями, що робить роботу не тільки інформативною, але й практично значущою. Детальний опис моделювання даних за допомогою Power BI демонструє компетентність автора у використанні таких сучасних технологій. Практична частина роботи є добре документованою і має високу прикладну цінність.

Зауваження до роботи стосуються переважно можливості більш детального розгляду та розширення набору візуалізацій для кращої демонстрації результатів опрацьованих даних.

Практична значимість висновків і рекомендацій. Запропоновані методи і моделі можуть бути безпосередньо використані для формування важливої аналітичної звітності. В роботі продемонстровані інструменти, які можуть бути адаптовані та застосовані в різних виробничих контекстах, що підтверджує її універсальність та практичну значимість.

Рецензент: Начальник відділу інформаційного забезпечення та центральної звітності Аналітичного управління ДП "КК "РОШЕН"

Мельник В. П.
(підпис, ПІБ)

ЗМІСТ

ВСТУП.....	3
РОЗДІЛ 1 ТЕОРЕТИЧНІ ЗАСАДИ СИСТЕМНОГО АНАЛІЗУ ТА МОДЕЛЮВАННЯ ДАНИХ.....	6
1.1 Огляд основних понять даних та системного аналізу	6
1.2 Організація та аналіз життєвого циклу даних у виробничій сфері.....	15
1.3 Огляд концепцій та методів моделювання даних у виробничій сфері.....	22
РОЗДІЛ 2 АНАЛІЗ ПРАКТИЧНИХ ЗАСАД ТА ІНСТРУМЕНТІВ МОДЕЛЮВАННЯ ДАНИХ І БІЗНЕС-АНАЛІТИКИ	32
2.1 Архітектура моделювання та зберігання даних при роботі з моделями даних	32
2.2 Excel як засіб для маніпулювання даними	44
2.3 Power BI як інструмент побудови аналітичних моделей даних	55
РОЗДІЛ 3 ПРИКЛАДНА ОБРОБКА ДАНИХ І СТВОРЕННЯ МОДЕЛІ ДАНИХ З ВИКОРИСТАННЯ POWER BI ТА EXCEL	70
3.1 Передобробка та зведення даних за допомогою Excel.....	71
3.2 Створення моделі даних з використанням засобів Microsoft Power BI.....	74
3.3 Візуальне представлення даних та аналіз отриманих результатів.....	87
ВИСНОВКИ	93
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	95

ВСТУП

У сучасному світі виробничі підприємства стикаються з величезним обсягом даних, що генеруються щодня в ході їхньої діяльності. Ці дані містять безцінну інформацію, яка може бути використана для управління такими бізнес-процесами, як оптимізація ресурсів та підвищення продуктивності. Однак, для того щоб зробити корисні висновки з великих обсягів даних, необхідно мати ефективні інструменти для їхнього аналізу та обробки.

Системний аналіз даних є одним з ключових підходів до вирішення цієї проблеми. Він передбачає комплексне дослідження та аналіз даних з метою виявлення закономірностей, трендів та важливих залежностей між ними. Основною метою системного аналізу даних є отримання цінної інформації, яка може служити основою для прийняття обґрунтованих управлінських рішень.

У даній дипломній роботі досліджуються можливості такого бізнес-процесу як системний аналіз даних у виробничій сфері. Особлива увага буде приділена процесу моделювання даних, який є важливим етапом в аналізі та обробці інформації. Моделювання даних дозволяє створювати абстракції реального світу, що спрощує їхнє подання та аналіз.

Головною частиною дослідження є використання програмного засобу Microsoft Power BI для створення моделей даних та візуалізації результатів аналізу. Power BI надає широкі можливості для обробки та аналізу даних, що робить його ідеальним інструментом для вирішення завдань системного аналізу даних у виробничій сфері.

У розділах даної роботи будуть розглянуті теоретичні засади бізнес-процесів обробки даних та системного аналізу, моделювання даних та конкретні кроки по створенню моделі даних за певними датасетами з використанням Microsoft Power BI.

Ця дипломна робота спрямована на вивчення та практичне застосування сучасних методів аналізу даних у виробничій сфері з метою підвищення ефективності управління та прийняття обґрунтованих управлінських рішень.

Актуальність теми системного аналізу даних та створення моделей даних в наш час важко переоцінити через їх величезний вплив на різні сфери життя. Завдяки аналізу даних, компанії можуть зрозуміти свої операції краще, виявити потенційні проблеми та знайти ефективні способи їх вирішення. Це може включати оптимізацію процесів виробництва, підвищення якості продукції та послуг, а також управління ризиками. Також аналіз даних допомагає у прийнятті обґрунтованих рішень на основі фактичних даних замість власних припущень або інтуїції. Це особливо важливо в умовах непередбачуваних змін та конкурентного середовища.

Аналіз останніх досліджень та публікацій. Останні дослідження та публікації в галузі системного аналізу даних у виробничій сфері підкреслюють важливість застосування аналітичних методів для виявлення паттернів, трендів та найбільш важливих факторів, що впливають на виробничі процеси. Дослідження також акцентують увагу на потребі розробки та застосування моделей даних, які відображають складні взаємозв'язки між різними параметрами та факторами у виробничій сфері.

Метою даної дипломної роботи є детальний покроковий аналіз бізнес-процесу обробки та аналізу даних, розробки та створення моделі даних з використанням технологій програмного забезпечення Microsoft Power BI, а також визначення їх впливу на ефективність та результативність управління підприємством.

Об'єктом дослідження є процес розробки та створення моделі даних з урахуванням засад самого моделювання, а також особливостей датасетів.

Предметом дослідження є теоретичні та практичні аспекти формування, розробки та створення моделей даних.

Методи дослідження включають теоретичний аналіз засад системного аналізу даних, проведення практичних експериментів з моделювання даних за реальними датасетами та аналіз результатів з використанням Microsoft Power BI. Також буде використано методи візуалізації даних для наглядного представлення результатів аналізу.

Теоретична, методична та практична значущість отриманих результатів полягає у розширенні розуміння та поглибленні знань щодо системного аналізу даних у виробничій сфері. Робота висвітлить основні концепції, методи та підходи до аналізу великих обсягів інформації на прикладі даних підприємства. Вона дозволить уточнити та систематизувати знання про моделювання даних, його роль та значення у галузі аналізу даних у виробничій сфері.

Результати роботи мають велике практичне значення для фахівців у галузі аналітики, інформаційного забезпечення та центральної звітності, маркетингової аналітики, аналітики підтримки продажів. Також значення полягає у їхній можливості бути безпосередньо використаними на виробничих підприємствах для покращення ефективності виробничих процесів та прийняття обґрунтованих управлінських рішень. Розроблені методи та моделі даних можуть бути впроваджені для постійного моніторингу та аналізу даних у реальному часі, що дозволить підприємствам оперативно реагувати на зміни у виробничих процесах та ринкових умовах. Таким чином, результати дослідження мають практичне застосування та можуть приносити конкретну користь для підприємств виробничої сфери.

Робота складається з вступу, трьох розділів, висновків, списку використаних джерел. Вступ обґрунтовує актуальність теми, визначає мету, завдання, об'єкт та предмет дослідження, методи дослідження, практичне значення та структуру роботи. Розділ 1 розглядає теоретичні аспекти системного аналізу та моделювання даних. Розділ 2 аналізує практичні засади моделювання даних, а також описує програмне забезпечення Microsoft Power BI. А Розділ 3 детально демонструє розробку та створення моделі даних Power BI та аналіз отриманих результатів.

РОЗДІЛ 1

ТЕОРЕТИЧНІ ЗАСАДИ СИСТЕМНОГО АНАЛІЗУ ТА МОДЕЛЮВАННЯ ДАНИХ

Бізнес-процеси у виробничій сфері охоплюють широкий спектр діяльності, спрямованої на створення продукції, оптимізацію ресурсів та забезпечення ефективного функціонування підприємства. До основних бізнес-процесів належать: планування виробництва, закупівля матеріалів, виробництво, контроль якості, логістика та управління запасами, продажі та маркетинг, управління персоналом, фінансовий менеджмент. І все це надзвичайно залежить від ще одного бізнес-процесу – аналізу даних. До нього входить правильне розуміння даних, системного аналізу, моделювання даних та аналітичних звітностей. Тому щоб повністю розкрити це питання, варто правильно визначити що таке дані та системний аналіз.

1.1 Огляд основних понять даних та системного аналізу

Виробнича сфера, як важлива галузь економіки, залежить від ефективного управління і прийняття рішень на основі різноманітних даних, які вона генерує та обробляє. Розуміння та правильне використання понять "дані" та "системний аналіз" є критичними для досягнення успішності та конкурентоспроможності у цій галузі.

Спочатку варто розглянути те, що означає термін "дані". У виробничій сфері це можуть бути різноманітні інформаційні потоки, що включають в себе дані про виробничі процеси, обсяги виробництва, якість продукції, витрати, ланцюжок постачання, інвентаризацію та інші аспекти. Вони можуть бути представлені у формі чисел, тексту, графіків, таблиць тощо. Дані виробничої сфери можуть бути

як структурованими (наприклад, дані з обліку виробництва), так і неструктурованими (наприклад, коментарі робітників чи результати випробувань продукції) [1].

Виробнича сфера є однією з найбільш наповнених даними галузей, оскільки вона залежить від великого обсягу і різноманітності інформації для оптимізації виробничих процесів та прийняття управлінських рішень. Основні джерела та типи даних, що використовуються в цій галузі, можна класифікувати наступним чином:

Дані з механізмів виробництва, обладнання та систем моніторингу дозволяють виробникам відстежувати й аналізувати ефективність та продуктивність виробничих процесів. Їх аналіз допомагає виявляти та усувати проблеми, що можуть виникати, та вдосконалювати виробничі процеси.

Дані про постачальників, транспортування, запаси сировини та готової продукції тощо. Вони допомагають виробникам ефективно керувати ланцюжком постачання, оптимізувати витрати та постачання продукції, щоб уникати затримок і недостачі матеріалів.

Дані з процесів контролю якості, результатів випробувань та зворотного зв'язку від клієнтів дозволяють виробникам забезпечити високу якість продукції та задоволення потреб споживачів.

Аналіз історичних даних продажів, ринкових тенденцій, та інших факторів допомагає виробникам адаптувати свою продукцію до потреб ринку, розробляти маркетингові стратегії та забезпечувати задоволеність серед клієнтів, прогнозувати попит та ефективно планувати виробництво.

Дані про витрати та фінанси, сировину, працю, енергію та інші ресурси, а також фінансові показники підприємства. Вони дозволяють виробникам ефективно управляти витратами та фінансами, знаходити шляхи для зменшення витрат та підвищення прибутковості, а також шукати можливості для оптимізації процесів на виробництві.

Усі ці аспекти демонструють, що дані у виробничій сфері є ключовим елементом для досягнення успіху та конкурентоспроможності на ринку. Вони дозволяють виробникам приймати обґрунтовані рішення, реагувати на зміни у

середовищі та постійно вдосконалювати свою діяльність. Важливою задачею для підприємства є збір, обробка та аналіз цих даних з метою отримання корисних інсайтів для подальшого розвитку та успіху [2].

Значення системного аналізу для оптимізації виробничих процесів

Системний аналіз в контексті виробничої сфери – це методологія, спрямована на вивчення та оптимізацію виробничих процесів за допомогою аналізу даних. Він базується на принципах системного підходу до управління, які передбачають розгляд системи (у даному випадку – виробничого підприємства) як комплексної єдності її складових елементів, взаємозв'язків та взаємодій [3].

У виробничій сфері, де обсяги даних можуть бути величезними та різноманітними, системний аналіз даних стає незамінним інструментом для вирішення стратегічних та оперативних завдань, таких як покращення продуктивності виробництва та якості товарів, зниження витрат, а також управління ризиками та прийняття рішень. Він дозволяє розуміти динаміку виробничих процесів, виявляти проблемні місця, прогнозувати їх розвиток і розробляти стратегії для їх вирішення [4].

Значення системного аналізу полягає у впровадженні системного підходу до управління виробництвом з метою досягнення оптимальних результатів і покращення продуктивності підприємства.

Призначення системного аналізу включає:

Комплексний огляд виробничих процесів: Системний аналіз дозволяє розглядати виробництво як інтегровану систему зі взаємозв'язаними елементами. Це дозволяє ідентифікувати взаємозв'язки між різними складовими виробничого процесу та визначати ключові фактори, які впливають на ефективність виробництва.

Виявлення проблем та можливостей для оптимізації: Системний аналіз допомагає виявляти проблемні аспекти виробництва, такі як зайві витрати, затримки у виробничому процесі, неефективне використання ресурсів тощо. Він

також дозволяє виявляти можливості для покращення, такі як автоматизація процесів, оптимізація ланцюжка постачання, підвищення якості продукції тощо.

Розробка оптимальних стратегій управління: На основі результатів системного аналізу розробляються стратегії та тактики управління виробництвом, спрямовані на досягнення поставлених цілей підприємства, серед яких покращення технологічної складової, змін в організаційній структурі, оптимізацію робочих процесів тощо.

Моніторинг та контроль за реалізацією стратегій: Системний аналіз допомагає виробникам моніторити виробничі процеси та контролювати виконання стратегій управління. Це дозволяє швидко виділяти певні зміни в середовищі та вчасно коригувати стратегії, якщо це необхідно [5].

Методи та інструменти системного аналізу, які застосовують у виробничій сфері.

У виробничій сфері використовуються різноманітні методи та інструменти системного аналізу для оптимізації виробничих процесів та підвищення продуктивності. Деякі з найпоширеніших методів та інструментів включають:

Системний аналіз процесів: Цей метод передбачає детальний аналіз виробничих процесів з урахуванням їх структури, взаємозв'язків та впливу на різні аспекти діяльності підприємства. Він дозволяє ідентифікувати проблемні зони та визначити шляхи їх покращення.

Моделювання систем: Цей інструмент дозволяє створювати математичні моделі виробничих процесів з метою аналізу їх роботи та ефективності. Моделі дозволяють проводити експерименти та прогнозувати результати впровадження різних стратегій управління [6].

SWOT-аналіз: Цей метод допомагає визначити сильні та слабкі сторони підприємства, а також можливості та загрози, що стоять перед ним на ринку. Він дозволяє виробникам розробляти стратегії на основі їх внутрішніх можливостей та зовнішнього середовища [7].

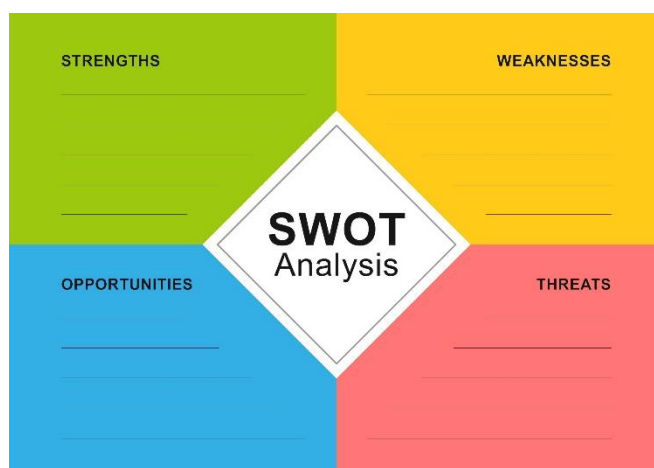


Рисунок 1.1 – Схема для SWOT-аналізу

(Джерело розроблено автором на основі[7])

LEAN-методологія: Цей підхід спрямований на мінімізацію витрат та максимізацію ефективності виробничих процесів шляхом виявлення та усунення зайвих операцій та процедур.

Картографування процесів: Цей інструмент використовується для візуалізації виробничих процесів та їх залежностей. Картографування дозволяє виробникам легше розуміти структуру та потік роботи виробничих процесів та виявляти можливості для їх оптимізації.

Програмне забезпечення для аналізу даних: До цього типу інструментів належать програмні засоби, такі як Microsoft Excel, Power BI, Tableau та інші, які дозволяють виробникам обробляти, аналізувати та візуалізувати великі обсяги даних для прийняття управлінських рішень.

Застосування цих методів та інструментів дозволяє виробникам досягати ефективності, підвищувати якість продукції та зменшувати витрати виробництва. Вони є важливими інструментами для досягнення конкурентної переваги на ринку та забезпечення успішної діяльності підприємства.

Взаємозв'язок між даними та системним аналізом виробничій діяльності є ключовим для ефективного управління та досягнення оптимальних результатів на підприємстві. Дані виступають основою для проведення системного аналізу, в той час як системний аналіз надає інструменти для аналізу та інтерпретації цих даних.

Дані про виробничі процеси, якість продукції, ланцюжок постачання, витрати, фінансові показники, ринкові дослідження та багато іншого можуть бути отримані з різних джерел, таких як датчики на обладнанні, системи моніторингу, бази даних підприємства, зовнішні джерела тощо.

Системний аналіз виробничих процесів базується на використанні цих даних для розуміння та оптимізації роботи підприємства. Цей аналіз допомагає виявити взаємозв'язки та патерни в даних, розглядається їх взаємозв'язок та взаємозалежність. Наприклад, аналіз виробничих даних може показати, які чинники впливають на якість продукції або ефективність виробничих процесів.

З іншого боку, системний аналіз надає інструменти для обробки та інтерпретації цих даних. Він використовує методи математичного моделювання, статистичного аналізу, оптимізації та інші підходи для вивчення взаємозв'язків між різними факторами та прогнозування їх впливу на виробничі процеси.

Отже, дані і системний аналіз у виробничій діяльності взаємопов'язані та взаємодоповнюються. Дані надають базову інформацію для аналізу, в той час як системний аналіз надає інструменти для розуміння та використання цих даних для прийняття управлінських рішень та оптимізації виробничих процесів [8].

Виробнича аналітика грає важливу роль у контексті системного аналізу даних, надаючи додаткові можливості для розуміння, аналізу та оптимізації виробничих процесів на підприємстві. Її роль та значення можна розглядати з кількох аспектів:

Збір та обробка даних: Виробнича аналітика допомагає виробникам збирати, організовувати та обробляти великі обсяги даних, які надходять з різних джерел на підприємстві. Вона використовується для аналізу даних та виявлення патернів, тенденцій та взаємозв'язків між різними факторами виробництва. Це дозволяє виробникам краще розуміти причинно-наслідкові зв'язки та визначати стратегії для оптимізації виробничих процесів.

Виробнича аналітика також допомагає виробникам прогнозувати попит, визначати оптимальні рівні виробництва та розробляти плани на майбутнє на основі аналізу даних про ринок та внутрішні фактори підприємства. Ще вона надає

виробникам необхідну інформацію та аналітичні засоби для прийняття управлінських рішень. Вона допомагає ідентифікувати пріоритетні напрямки розвитку, виявляти можливості для ефективності та оптимізації виробничих процесів [9].

Загалом, виробнича аналітика є важливим елементом системного аналізу даних, який допомагає виробникам отримувати цінні інсайти з великих обсягів даних та використовувати їх для прийняття обґрунтованих управлінських рішень. Вона допомагає підприємствам стати більш гнучкими, конкурентоспроможними та успішними на ринку.

У системному аналізі виробничих процесів використовуються різні ключові поняття та методи аналізу даних, які сприяють зрозумінню та оптимізації роботи підприємства.

Продуктивність – визначає ефективність використання ресурсів підприємства для виробництва продукції.

Якість – відображає ступінь відповідності продукції встановленим стандартам та вимогам клієнтів.

Ефективність – відображає співвідношення між витратами та результатами у виробничих процесах.

Оптимізація – поліпшення процесів з метою досягнення найкращого можливого результату при мінімальних витратах.

Методи аналізу даних:

Статистичний аналіз: Використовується для вивчення розподілу даних, виявлення тенденцій та залежностей.

Кластерний аналіз: Допомагає групувати дані у класи або кластери на основі їх схожості.

Часовий ряд аналіз: Використовується для аналізу динаміки даних у часі та прогнозування майбутніх значень.

Методи оптимізації: Включають лінійне програмування, динамічне програмування та інші методи для знаходження оптимальних рішень у виробничих процесах.

Під час системного аналізу виробничих процесів, ці поняття та методи допомагають виробникам розуміти та оптимізувати виробничі процеси, забезпечуючи високу продуктивність, якість продукції та ефективність управління ресурсами. Їх використання дозволяє підприємствам досягати конкурентної переваги та стабільного розвитку на ринку [10].

Інтеграція даних та системного аналізу виробничих процесів є критично важливою для прийняття релевантних необхідних та управлінських рішень у цій сфері. Щоб це підтвердити варто виділити деякі аспекти важливості цієї інтеграції.

Збільшення доступності даних: Інтеграція даних у єдину систему дозволяє забезпечити доступність всієї необхідної інформації для управлінського персоналу. Це включає дані про виробничі процеси, фінансові показники, ланцюжок постачання та інші аспекти діяльності підприємства.

Підвищення якості прийнятих рішень: Інтеграція даних дозволяє проводити комплексний аналіз і отримувати повний обсяг інформації перед прийняттям рішення. Це допомагає уникнути затримок, помилок та невдач у управлінні, що можуть виникнути при використанні обмеженої або неповної інформації.

Забезпечення точності та надійності даних: Інтеграція даних дозволяє уникнути дублювання та розбіжностей у даних, оскільки всі дані зберігаються в єдиній системі. Це забезпечує більш високу точність та надійність інформації, на основі якої приймаються управлінські рішення.

Можливість аналізу в реальному часі: Інтеграція даних дозволяє отримувати доступ до інформації в реальному часі, що дозволяє управлінцям швидко реагувати на зміни у виробничих процесах та ринкових умовах.

Отже, інтеграція даних та системного аналізу є необхідною складовою для прийняття ефективних управлінських рішень у виробничій сфері, що надає підґрунтя підприємствам отримати повний обсяг інформації та використовувати її для найкращого управління ресурсами та оптимізації виробничих процесів [11]. Перспективи використання системного аналізу даних для оптимізації виробничих процесів.

У майбутньому системний аналіз даних матиме великий потенціал для оптимізації виробничих процесів. Розвиток технологій, зокрема штучного інтелекту (ШІ), машинного навчання та аналітики великих даних, відкриває нові можливості для вдосконалення методів аналізу та управління виробничими процесами.

З використанням ШІ та автоматизації виробничих процесів системний аналіз може бути вдосконалений для швидкого виявлення аномалій та відповідного реагування на них без необхідності людського втручання.

За допомогою аналізу великих даних та машинного навчання системний аналіз може стати більш точним у прогнозуванні попиту на продукцію та визначенні оптимальних рівнів виробництва.

ШІ та аналітика даних дозволять вдосконалити системний аналіз управління ланцюжком постачання шляхом автоматизації процесів замовлення, прогнозування потреб та виявлення оптимальних постачальників.

За допомогою аналізу даних та машинного навчання можна розробляти прогностичні моделі, які дозволять передбачити виникнення проблем у виробничих процесах та запобігти їм заздалегідь.

Системний аналіз даних також може бути використаний для оптимізації виробничих процесів з урахуванням екологічних аспектів, що дозволить підприємствам стати більш сталими та екологічно чистими [12].

У майбутньому системний аналіз даних й надалі матиме значний вплив на виробничі процеси, допомагаючи підприємствам стати більш ефективними, конкурентоспроможними та сталими. Використання сучасних технологій та розробка нових методів аналізу даних дозволять досягти великих успіхів у цій галузі.

1.2 Організація та аналіз життєвого циклу даних у виробничій сфері

Життєвий цикл даних – це концепція, що описує всі аспекти життя даних від їх створення та збору до обробки, аналізу, використання та видалення. Ця концепція визначає всі етапи, через які дані проходять у своєму шляху від початкового збору до кінцевого використання або видалення.

Важливість життєвого циклу даних полягає в тому, що він дозволяє підприємствам ефективно управляти даними на протязі їх існування. Це означає забезпечення якості, цілісності та доступності даних, а також їх безпеки та конфіденційності. Система життєвого циклу даних допомагає виробникам розуміти, як їхні дані використовуються, як вони можуть бути оптимізовані та як можна використати їх для досягнення стратегічних цілей підприємства.

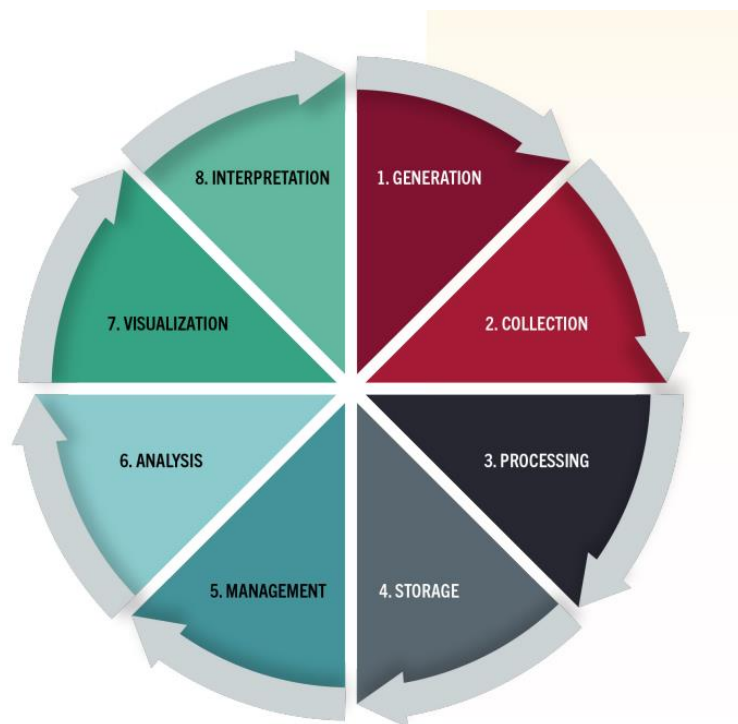


Рисунок 1.2 – Життєвий цикл даних [13, с. 8]

У виробничій сфері життєвий цикл даних має велике значення. Він допомагає виробникам збирати та аналізувати дані про виробничі процеси, ресурси, якість продукції, а також процеси постачання та збуту. Це дозволяє підприємствам покращувати свою роботу, ефективність та якість продукції,

знижувати витрати та ризики, а також реагувати на зміни у внутрішньому та зовнішньому середовищі.

Отже, життєвий цикл даних є ключовим аспектом управління даними у виробничій сфері. Він допомагає підприємствам забезпечувати ефективне використання та управління даними протягом їхнього життя, що є важливим для досягнення стратегічних цілей та успіху бізнесу [13].

1. Генерація даних

Етап "Генерація даних" є першим у такому циклі. На цьому етапі дані створюються або виникають у процесі виробництва. Це може включати в себе автоматичне формування даних за допомогою сенсорів, вимірювального обладнання або систем моніторингу. Також дані можуть створюватися вручну операторами під час ручних операцій або процесів контролю якості.

Генерація даних виробничих процесів може бути постійною або періодичною залежно від характеру виробництва та обсягу даних, які необхідно зібрати. Вона може відбуватися в реальному часі або мати затримку, залежно від технологій та обладнання, що використовуються у виробництві. Забезпечення правильного формування та документування даних на цьому етапі є ключовим для подальшого успішного аналізу та використання виробничих даних [14].

2. Збір даних

На другому етапі життєвого циклу – зборі даних - інформація, яка була згенерована на попередньому етапі, збирається та записується для подальшої обробки та аналізу. Збір даних може відбуватися автоматично за допомогою спеціалізованих датчиків, вимірювального обладнання або програмних засобів, або ж ручним способом, коли оператори заносять дані в базу даних або електронні таблиці.

Цей етап важливий, оскільки від точності та повноти збору даних залежить подальша аналітична обробка та прийняття управлінських рішень. Під час збору даних важливо враховувати не лише кількість інформації, а й її якість та

репрезентативність. Забезпечення однорідності та стандартизації процесу збору даних дозволяє уникнути помилок та забезпечити надійність результатів.

Технічні рішення можуть варіюватися від простих датчиків до складних систем зі збірниками даних та системами зберігання. При цьому важливо враховувати специфіку виробництва, характер даних та обсяг інформації, яка потребується для подальшого аналізу. Успішне виконання цього етапу дозволить забезпечити достовірність та актуальність даних для наступних процесів аналізу та прийняття рішень у виробничому середовищі [15].

3. Обробка даних

На третьому етапі – обробці даних – зібрана інформація піддається різноманітним операціям та трансформаціям з метою очищення, структурування та підготовки для подальшого аналізу та використання. Цей етап включає в себе такі процеси, як фільтрація даних, видалення дублікатів, об'єднання таблиць, створення нових атрибутів та агрегація даних для отримання більш високорівневих показників.

Обробка даних грає важливу роль у підготовці інформації для аналізу та прийняття рішень. Вона дозволяє виявляти та виправляти помилки в даних, забезпечує їхню структурування та нормалізацію для однорідності та однорідності. Крім того, обробка даних може включати в себе застосування алгоритмів машинного навчання для автоматичного виявлення патернів або аномалій у даних.

У виробничій сфері обробка даних може включати в себе також аналіз параметрів виробництва, визначення ефективності процесів та ідентифікацію можливостей для оптимізації. Наприклад, шляхом виявлення та аналізу зайвих витрат або оптимізації часу виробничих циклів.

Успішна обробка даних забезпечує якість та достовірність даних для подальшого використання у виробничих процесах. Вона є ключовим етапом у підготовці даних для аналітичних досліджень та прийняття управлінських рішень у виробничому середовищі [16].

4. Зберігання даних

На четвертому етапі життєвого циклу даних у виробничій сфері – зберіганні даних – інформація, яка пройшла обробку, зберігається для подальшого використання та доступу. Цей етап включає в себе розміщення даних у спеціально призначених сховищах даних, таких як бази даних, дата-сховища або хмарних сервісів.

Зберігання даних є критично важливою складовою управління інформацією, оскільки воно забезпечує доступність, цілісність та безпеку даних протягом тривалого часу. Для забезпечення надійності та відновлюваності даних, їх можна архівувати та резервувати з регулярністю.

У виробничій сфері обсяги даних можуть бути великими, тому для зберігання використовуються спеціалізовані системи з великою місткістю та швидким доступом до інформації. Крім того, важливо враховувати вимоги щодо безпеки даних, зокрема шифрування та контроль доступу до інформації.

Забезпечення ефективного зберігання даних дозволяє підприємствам зберігати великі обсяги інформації, забезпечуючи її доступність для подальшого використання та аналізу. Цей етап є важливим для забезпечення неперервності виробничих процесів та збереження історичних даних для аналізу та прийняття рішень [17].

5. Управління даними

На п'ятому етапі життєвого циклу даних у виробничій сфері – управлінні даними – забезпечується ефективне управління інформацією з метою забезпечення її якості, доступності та використання відповідно до потреб підприємства. Цей етап включає в себе ряд дій та процедур, спрямованих на контроль та координацію всіх аспектів життєвого циклу даних.

Управління даними охоплює такі аспекти, як стандартизація даних, управління версіями, контроль доступу, а також управління метаданими. Воно також включає в себе розробку та впровадження політик та процедур управління

даними, що допомагають забезпечити консистентність та цілісність даних, а також визначають правила їх використання та доступу [18].

Управління даними є ключовим аспектом виробничої діяльності, оскільки забезпечує ефективне використання інформації для прийняття рішень та оптимізації виробничих процесів. Воно дозволяє підприємствам зберігати, організовувати та контролювати дані з метою забезпечення їхньої актуальності та достовірності. Крім того, управління даними сприяє підвищенню продуктивності праці та зменшенню ризиків, пов'язаних з некоректним використанням чи втратою даних.

6. Аналіз даних

На шостому етапі життєвого циклу даних у виробничій сфері – аналізі даних – здійснюється виявлення залежностей, трендів та взаємозв'язків у накопиченій інформації з метою отримання нових знань та інсайтів для підтримки прийняття рішень.

Під час аналізу даних виробничі показники включають виробничу потужність, витрати на матеріали та працю, якість продукції, а також ефективність виробничих процесів. Застосування різноманітних методів статистичного аналізу, машинного навчання та штучного інтелекту дозволяє виявити складні залежності та закономірності в даних, що може сприяти вдосконаленню виробничих процесів та оптимізації бізнесу.

Аналіз даних є ключовим етапом у виробничій сфері, оскільки він дозволяє виявити потенційні можливості, які можуть зацікавити підприємство, як з хорошої точки зору, так і з поганої. Отримані в результаті аналізу інсайти можуть використовуватись для розробки стратегій розвитку підприємства. Таким чином, аналіз даних відіграє важливу роль у забезпеченні конкурентоспроможності та успішності виробничої діяльності [19].

7. Візуалізація даних

На сьомому етапі життєвого циклу даних у виробничій сфері – візуалізації даних – інформація аналізується та представляється у вигляді графіків, діаграм, карт та інших візуальних елементів з метою зрозумілого та ефективного сприйняття результатів аналізу. Візуалізація даних дозволяє відображати складну інформацію у зрозумілій формі, що сприяє швидшому розумінню та прийняттю рішень.

Під час візуалізації даних виробничі параметри, тренди та залежності можуть бути представлені у вигляді різних типів графіків, таких як лінійні графіки, стовпчасті діаграми, кругові діаграми тощо. Крім того, використання географічних карт дозволяє візуалізувати розташування виробничих об'єктів, поширення ринків та інші просторові залежності.

Візуалізація даних уможлиблює виокремлення зв'язків та закономірностей, які може бути важко помітити за допомогою числових значень. Крім того, вона дозволяє виробникам визначати потенційні проблеми або можливості для оптимізації виробничих процесів [20].

Важливість візуалізації даних полягає в тому, що вона допомагає підвищити рівень розуміння та сприйняття важливої інформації. Ефективне її використання дозволяє підприємствам швидше виявляти проблеми та приймати обґрунтовані рішення для подальшого розвитку та оптимізації виробничих процесів.

8. Інтерпретація даних

На восьмому етапі життєвого циклу даних у виробничій сфері – інтерпретації даних – проводиться розбір результатів аналізу та візуалізації з метою зрозуміння їхнього значення та прийняття обґрунтованих рішень. Інтерпретація даних включає в себе оцінку виявлених патернів, трендів та взаємозв'язків із залученням експертного досвіду та знань про виробничі процеси.

Під час інтерпретації даних важливо враховувати контекст виробничої сфери та особливості підприємства. Інтерпретація дозволяє зрозуміти, які фактори

впливають на виробничі показники та які можливості для оптимізації виробничих процесів можуть бути виявлені.

Одним з ключових аспектів інтерпретації даних є виявлення причинно-наслідкових зв'язків між різними факторами та виробничими показниками. Це допомагає зрозуміти, які дії можуть призвести до покращення ситуації або, навпаки, що може призвести до погіршення результатів.

Інтерпретація даних є важливим завершальним етапом аналізу, оскільки вона дозволяє виявити ключові висновки та рекомендації для подальших дій. Вона є основою для прийняття управлінських рішень та розробки стратегій розвитку підприємства виробничої сфери. Інтерпретація даних допомагає підприємствам зробити обґрунтовані вибори та забезпечити успішність їхньої діяльності. [21]

У контексті виробничої сфери життєвий цикл даних виявляється критичним елементом для досягнення успіху та конкурентоспроможності підприємства.

Згідно з розглянутим життєвим циклом даних, виробничі підприємства стикаються з комплексними завданнями щодо збору, обробки, зберігання та управління даними. Кожен етап має своє значення, що впливає на загальний успіх виробничої діяльності.

Генерація даних є вихідною точкою, яка визначає обсяг та якість інформації, що буде доступна для подальшого аналізу. Збір, обробка та зберігання даних створюють фундамент для подальшого управління та аналізу, забезпечуючи доступність та цілісність інформації протягом тривалого часу.

Управління даними допомагає забезпечити ефективне використання інформації та контроль за її якістю, що є ключовим для прийняття обґрунтованих рішень. Аналіз та візуалізація даних надають можливість отримати нові знання та інсайти з накопиченої інформації, що допомагає виявити тенденції та залежності.

Нарешті, інтерпретація даних визначає значення та роль інформації для підприємства, надаючи основу для прийняття стратегічних рішень та розвитку бізнесу.

В цілому, розуміння та ефективне використання життєвого циклу даних у виробничій сфері є ключовими для досягнення успішності та конкурентоспроможності на ринку. Тільки завдяки цілеспрямованій організації та аналізу даних підприємства можуть зробити обґрунтовані рішення та досягти стратегічних цілей.

1.3 Огляд концепцій та методів моделювання даних у виробничій сфері

Моделювання даних — це не створення діаграм заради документації. Йдеться про створення спільного розуміння між бізнесом і командами обробки даних, зміцнення довіри та забезпечення цінності даних. Це також інвестиція. Інвестиції в стабільність, надійність і адаптивність ваших систем даних. Як і всі цінні ініціативи, це потребуватиме додаткових зусиль наперед. Проте в довгостроковій перспективі віддача від інвестицій у зв'язок, ефективність, якість даних і масштабованість є значною [22].

В контексті виробничої сфери, модель є абстрактним представленням реального виробничого процесу або системи за допомогою математичних або комп'ютерних конструкцій. Основними поняттями, пов'язаними з моделями даних у виробничій сфері, є предметна область, змінні, параметри та взаємозв'язки між ними.

Моделі виробничої сфери можуть бути використані для прогнозування виробничих показників, аналізу та оптимізації процесів виробництва, вирішення проблем ефективності та вдосконалення бізнес-стратегій. Вони дозволяють виробникам робити інформовані рішення, аналізуючи вплив різних факторів на виробничі процеси та прогножуючи їхні результати.

Моделі виробничої сфери можуть мати різну структуру та рівень абстракції. Вони можуть бути представлені у вигляді математичних рівнянь, симуляційних

моделей, графіків, діаграм або комп'ютерних програм, що дозволяють відобразити взаємозв'язки між різними аспектами виробничого процесу.

У виробничій сфері моделі даних можуть використовуватися для прогнозування виробничих обсягів, визначення оптимальних стратегій управління запасами, аналізу впливу різних факторів на якість продукції, а також для виявлення та вирішення проблем у виробничому процесі [23].

Отже, моделі даних виробничої сфери є важливим інструментом для аналізу та оптимізації виробничих процесів, що допомагає підприємствам приймати обґрунтовані рішення та досягати стратегічних цілей.

Класифікація моделей даних у виробничій сфері може ґрунтуватися на різних критеріях, включаючи структуру, призначення, рівень абстракції та інші ключові параметри.

По-перше, моделі можуть бути класифіковані за їхньою структурою. Наприклад, моделі можуть бути детермінованими або стохастичними, лінійними або нелінійними, статичними або динамічними. Кожен тип моделі має свої особливості та область застосування у виробничому середовищі.

По-друге, моделі можуть бути класифіковані за їхнім призначенням. Наприклад, існують моделі для прогнозування виробничих показників, оптимізації виробничих процесів, аналізу ризиків та багато інших цілей. Кожна з цих моделей спеціалізується на конкретному завданні та використовує відповідні методи та техніки.

По-третє, моделі можуть бути класифіковані за рівнем абстракції. Наприклад, низькорівневі моделі можуть бути докладними та деталізованими, відображаючи конкретні аспекти виробничого процесу, тоді як високорівневі моделі можуть бути більш узагальненими та загальними.

Крім того, моделі можуть бути класифіковані за їхнім методом побудови, наприклад, аналітичні моделі, що базуються на математичних рівняннях, або симуляційні моделі, що використовують комп'ютерні програми для відтворення виробничих процесів [24].

В результаті класифікації моделей даних за різними критеріями виробничі підприємства можуть вибрати найбільш підходящий тип моделі для вирішення конкретної задачі, що допомагає їм ефективно використовувати дані та приймати обґрунтовані рішення [31-33].

Математичні методи та інструменти для побудови моделей даних у виробничій сфері є важливими інструментами для аналізу, прогнозування та оптимізації виробничих процесів. Основні методи та інструменти включають лінійну регресію, дерева рішень та нейронні мережі.

Лінійна регресія є одним з найпоширеніших методів у статистичному аналізі даних. Цей метод часто використовується для прогнозування виробничих обсягів на основі історичних даних та виявлення впливу різних факторів на виробничий процес [25].

Дерева рішень є графічною моделлю, яка відображає прийняття рішень на основі ряду умовних змінних. Вони можуть бути використані для аналізу та класифікації даних у виробничій сфері, а також для виявлення закономірностей та взаємозв'язків між різними аспектами виробничого процесу.

Нейронні мережі є математичним моделюванням, інспірованим внутрішньою структурою нейронної системи людини. Вони можуть бути використані для аналізу складних виробничих процесів та прогнозування майбутніх тенденцій на основі великих обсягів даних.

Ці методи та інструменти надають виробничим підприємствам можливість ефективно аналізувати та використовувати доступні дані для прийняття обґрунтованих управлінських рішень. Вони допомагають виробникам зрозуміти внутрішню динаміку виробничих процесів, виявити тенденції та прогнозувати майбутні події, що є ключовим для досягнення успіху у сучасній конкурентній виробничій сфері.

Використання методів класифікації та кластерного аналізу є важливим етапом в моделюванні та аналізі даних у виробничому середовищі, оскільки вони

дозволяють розподілити дані на групи або категорії в залежності від їхніх характеристик та властивостей.

Метод класифікації полягає у визначенні відповідності між об'єктами та категоріями на основі певного набору характеристик або ознак. Наприклад, у виробничому середовищі метод класифікації може бути використаний для класифікації виробів за їхніми технічними характеристиками, виробничими параметрами чи споживацькими властивостями. Це дозволяє виробникам краще розуміти свою аудиторію та споживачів, а також управляти асортиментом продукції [26].

Кластерний аналіз, з іншого боку, дозволяє групувати схожі об'єкти разом у кластери на основі їхніх схожих характеристик. У виробничому середовищі це може бути корисно для ідентифікації подібних виробничих процесів, продуктів або клієнтських сегментів. За допомогою кластерного аналізу можна виявити групи, які можуть потребувати різних стратегій управління або маркетингових підходів.

Використання цих методів дозволяє підприємствам виробничої сфери краще розуміти свої дані, здійснювати з ними більш точний та ефективний аналіз, а також приймати обґрунтовані рішення на основі цих даних. Вони допомагають покращити ефективність виробничих процесів, виявити нові можливості для оптимізації та підвищити конкурентоспроможність підприємства.

Використання статистичних методів у моделюванні даних для прогнозування та аналізу виробничих процесів вирішує низку ключових завдань у виробничій сфері. Статистичні методи дозволяють аналізувати структуру даних, виявляти закономірності, розпізнавати та передбачати тенденції, а також робити висновки на основі статистичних зв'язків.

Один з основних аспектів використання статистичних методів – це прогнозування виробничих процесів. Завдяки аналізу історичних даних і використанню різних моделей, таких як часові ряди або регресійні моделі, статистичні методи дозволяють прогнозувати майбутні показники виробництва,

витрат, попиту на продукцію та інші параметри, що допомагає підприємствам у плануванні та управлінні ресурсами.

Додатково, статистичні методи використовуються для аналізу виробничих процесів. Вони дозволяють виявляти взаємозв'язки між різними факторами виробництва, оцінювати їхній вплив на результати та ідентифікувати можливі проблемні аспекти виробництва. Це дозволяє підприємствам вдосконалювати свої процеси, підвищувати ефективність та конкурентоспроможність.

Статистичні методи також використовуються для оцінки ризиків у виробничій сфері. Аналіз варіації, кореляційний аналіз, аналіз часових рядів та інші статистичні методи дозволяють виявити потенційні загрози та ризики, пов'язані з виробничими процесами, що допомагає підприємствам управляти ними та зменшувати ймовірність негативних наслідків.

Усе це свідчить про важливість та широкий спектр застосування статистичних методів у моделюванні даних для прогнозування та аналізу виробничих процесів у сучасній виробничій сфері.

Техніки оптимізації та вирішення задач у модельних системах є важливим інструментом для підвищення ефективності виробничих процесів у виробничій сфері. Ці техніки спрямовані на знаходження оптимальних рішень, що дозволяє підприємствам ефективно використовувати ресурси, мінімізувати витрати та максимізувати прибуток.

Одна з ключових технік – це математичне програмування, яке використовується для розв'язання оптимізаційних задач. Методи лінійного програмування, цільового програмування та інші дозволяють знаходити оптимальні значення змінних, що відображають рішення виробничих завдань, таких як планування виробництва, розподіл ресурсів та інші.

Ще одна важлива техніка – це методи чисельної оптимізації. Ці методи базуються на алгоритмах, які шукають найкраще рішення серед можливих варіантів, використовуючи чисельні методи та обчислювальні алгоритми. Вони застосовуються для вирішення складних оптимізаційних задач, таких як вибір

оптимальних стратегій виробництва або оптимальне розташування виробничих об'єктів.

Додатково, до технік оптимізації також відносяться методи машинного навчання та штучного інтелекту. Ці методи дозволяють автоматизувати процеси прийняття рішень, опрацьовувати масштабної величини данні та виявляти складні взаємозв'язки між різними факторами.

Техніки оптимізації та вирішення задач в модельних системах дозволяють підприємствам ефективно використовувати наявні ресурси, мінімізувати витрати та максимізувати прибуток. Вони є важливим інструментом для досягнення конкурентної переваги та стабільного розвитку в сучасній виробничій сфері.

Роль симуляційних моделей у виробничому аналізі даних та їхнє застосування для прогнозування та вирішення проблем

Симуляційні моделі дозволяють виробникам моделювати та аналізувати різноманітні виробничі сценарії для прогнозування майбутніх тенденцій та вирішення проблем.

Однією з важливих переваг симуляційних моделей є їх здатність відтворювати складні виробничі процеси з точністю та деталізацією. Це дозволяє виробникам аналізувати вплив різних факторів на виробництво та виявляти потенційні ризики та проблемні моменти. Наприклад, симуляційні моделі можуть допомогти виявити проблеми зі збільшенням виробничих обсягів або оптимізувати розклад виробництва для максимізації використання ресурсів.

Крім того, симуляційні моделі можуть бути використані для прогнозування майбутніх тенденцій та розвитку виробничих процесів. Шляхом введення різних сценаріїв та параметрів, виробники можуть оцінити можливі наслідки та прийняти рішення, спрямовані на максимізацію ефективності та прибутковості.

Додатково, симуляційні моделі можуть служити інструментом для вирішення конкретних проблем у виробничому процесі. Наприклад, вони можуть бути використані для виявлення оптимального розташування обладнання,

визначення найбільш ефективних стратегій виробництва або аналізу впливу змін у виробничому процесі на кінцевий результат.

Усі ці фактори підкреслюють важливість та роль симуляційних моделей у виробничому аналізі даних. Вони є потужним інструментом для вирішення різних завдань у виробничій сфері, сприяючи покращенню ефективності, прийняттю обґрунтованих рішень та досягненню конкурентної переваги.

Аналіз часових рядів даних у виробничій сфері та його моделювання для прогнозування та оптимізації виробничих процесів. Цей тип аналізу забезпечує можливість виявлення тенденцій, циклічних змін, сезонності та інших регулярних патернів у даних, що може бути корисним для управління виробничими процесами.

Одним із основних аспектів аналізу часових рядів є моделювання цих даних. Це включає уявлення динаміки процесу за допомогою математичних моделей, які можуть бути використані для прогнозування майбутніх значень. Наприклад, для прогнозування обсягів виробництва, запасів сировини або попиту на продукцію можуть бути використані різні методи, такі як авторегресійні моделі, експоненціальне згладжування або методи скінченних ізольованих автогресійних моделей (SARIMA).

Моделі аналізу часових рядів дозволяють виробникам управляти ризиками та приймати правильні рішення. Наприклад, шляхом прогнозування попиту на продукцію, підприємство може оптимізувати свою виробничу програму та планувати виробництво відповідно до цього попиту, що дозволяє уникнути перевиробництва або нестачі товарів на ринку [27].

Такий аналіз дозволяє підприємствам бути більш гнучкими та реагувати на зміни у виробничих умовах, підвищуючи свою конкурентоспроможність та ефективність виробничих процесів. Він є необхідним етапом у процесі планування та управління виробництвом у сучасній виробничій сфері.

Вплив різних факторів на ефективність та точність моделей даних у виробничій сфері може бути значним і має велике значення для досягнення правильних та достовірних результатів аналізу.

Обсяг даних є одним із ключових факторів, що впливають на ефективність моделей даних. Збільшення обсягу даних може покращити точність моделей шляхом забезпечення більшої репрезентативності та розмаїтості вхідних даних. Однак, при цьому необхідно враховувати можливість перевантаження моделей та виникнення перенавчання, що може призвести до втрати універсальності та загальності результатів.

Якість даних також має значення для точності моделей. Неправильні або неточні дані можуть спричинити викривлення результатів та зробити модель менш надійною. Тому важливо провести перед аналізом дані очистку, перевірку та підготовку для забезпечення їхньої якості.

Вибір методів аналізу та моделювання також може суттєво вплинути на результати. Різні методи можуть мати різний ступінь ефективності та точності в конкретних умовах виробничої сфери. Наприклад, для прогнозування виробничих потоків можуть бути використані різні алгоритми, такі як експоненційне згладжування, авторегресійні моделі або методи машинного навчання.

Інші фактори, такі як складність виробничих процесів, динаміка ринку та технологічні інновації, також можуть впливати на ефективність моделей даних у виробничій сфері. Тому важливо проводити аналіз та оцінку всіх цих факторів при розробці та використанні моделей даних для забезпечення їхньої найвищої точності та релевантності [28].

Переваги та обмеження різних методів моделювання даних у виробничій сфері та їхнє порівняння.

Різні методи моделювання даних у виробничій сфері мають свої переваги та обмеження, і їх вибір повинен базуватись на певних вимогах від підприємства.

Лінійна регресія, наприклад, є одним із найпоширеніших методів, який зазвичай використовується для моделювання залежності між вхідними та

вихідними змінними. Її перевагою є простота та інтерпретованість результатів. Однак цей метод може бути неефективним у випадку складних взаємозв'язків між даними.

Дерева рішень є ще одним популярним методом, який використовується для класифікації та прогнозування. Вони легко інтерпретуються та можуть враховувати нелінійні взаємозв'язки між змінними. Проте дерева рішень схильні до перенавчання, особливо при великій кількості вхідних змінних [29, 30].

Нейронні мережі, зокрема глибокі нейронні мережі, стали популярними у виробничому аналізі даних завдяки їхній здатності розпізнавати складні залежності та патерни у даних. Однак ці методи розраховані на масштабні тривалі масиви даних, а також великого обсягу обчислювальних ресурсів.

Порівняльний аналіз різних методів моделювання даних виробничій сфері дозволяє вибрати найбільш підходящий метод з урахуванням конкретних потреб та обмежень підприємства. Важливо також враховувати фактори, такі як складність реалізації, інтерпретованість результатів, обчислювальні витрати та доступність даних. Обраний метод повинен забезпечити найкращу комбінацію точності та практичності для досягнення максимальної користі для підприємства.

Висновок до першого розділу

Розділ надає глибокий огляд основних аспектів системного аналізу даних у виробничій сфері. Починаючи з першого пункту, зроблено огляд основних понять даних та системного аналізу. Розкривши визначення даних та системного аналізу у виробничій сфері, виявлено їхню ключову роль у сучасній промисловості. Поняття даних, як основний будівельний блок інформаційного аналізу, виявляється незмінно важливим для виробничих підприємств, оскільки вони постійно збирають, обробляють та використовують великі обсяги даних для прийняття управлінських рішень.

У другому пункті розглянута організація та аналіз життєвого циклу даних у виробничій сфері. З'ясовано, що життєвий цикл даних включає важливі етапи, такі

як генерація, збір, обробка, зберігання, управління, аналіз, візуалізація та інтерпретація. Кожен з цих етапів має велике значення для виробничих процесів, а їхнє правильне виконання забезпечує ефективний аналіз та високу якість прийнятих управлінських рішень.

Нарешті, у третьому пункті надано огляд модельних концепцій та методів моделювання даних у виробничій сфері. Розглянуто класифікацію моделей за різними критеріями та проаналізовано їхній вплив на виробничі процеси. Детально проаналізовано математичні методи та інструменти моделювання, їхню застосовність та ефективність у виробничих процесах. Обговорені техніки оптимізації та вирішення завдань у модельних системах, роль симуляційних моделей та використання методів класифікації та кластерного аналізу для розподілу даних у виробничому середовищі.

Завершуючи цей розділ, важливо визначити, що ретельне дослідження основних понять, організація та аналіз життєвого циклу даних, а також огляд модельних концепцій та методів моделювання є ключовими складовими успішного впровадження системного аналізу даних у поєднанні з глибоким розумінням конкретних виробничих процесів допоможе підвищити ефективність та конкурентоспроможність підприємства на ринку.

РОЗДІЛ 2

АНАЛІЗ ПРАКТИЧНИХ ЗАСАД ТА ІНСТРУМЕНТІВ МОДЕЛЮВАННЯ ДАНИХ І БІЗНЕС-АНАЛІТИКИ

2.1 Архітектура моделювання та зберігання даних при роботі з моделями даних

Сховище даних є фундаментальною концепцією в сфері бізнес-аналітики та управління даними. У виробничій практиці найпоширенішою назвою сховищ даних є Data Warehouse, або ж скорочено DWH. Воно представляє собою централізоване сховище, призначене для зберігання великих обсягів даних, що надходять з різних джерел. Основною метою сховищ даних є забезпечення єдиного місця для накопичення, зберігання та аналізу даних, що дозволяє підприємствам приймати обґрунтовані рішення на основі точних та актуальних даних [34,35].

DWH консолідує дані з різних джерел, таких як бази даних, текстові файли, системи CRM, ERP тощо. Це дозволяє уникнути розрізненості даних та забезпечує єдину версію правильності даних, що є критично важливим для точного аналізу.

На відміну від оперативних баз даних, що зосереджені на поточних операціях, сховище даних зберігає історичні дані. Це дозволяє здійснювати аналіз трендів і змін протягом часу, що є важливим для стратегічного планування та прогнозування [36].

Дані у сховищі часто містять часові мітки, що дозволяє виконувати аналіз на основі часу, як-от річні, квартальні або місячні звіти. Це забезпечує контекст для розуміння змін і допомагає виявляти сезонні або інші патерни, що виокремлюються за рахунок часових міток.

Сховища даних забезпечують інтеграцію даних, об'єднуючи інформацію з різних джерел у єдину структуру. Це включає узгодження різних форматів, очищення даних, видалення дублікатів та забезпечення консистентності.

Таблиці фактів (fact tables) – це таблиці, що містять основні бізнес-процеси або транзакції. Таблиці фактів зазвичай містять числові дані, які можна підсумовувати, такі як продажі, доходи або витрати.

Вимірювальні таблиці (dimension tables) – це таблиці, що містять атрибути, які описують факти. Наприклад, вимірювальні таблиці можуть містити інформацію про продукти, клієнтів, місцезнаходження або час. У виробничій практиці більш поширеною назвою вимірювальних таблиці є класифікатори або ж словники.

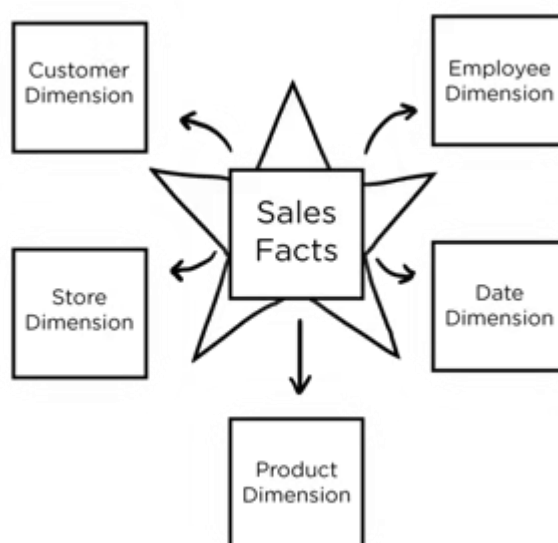


Рисунок 2.1 – Схема моделі даних з вимірювальними таблицями [33]

Сховища даних надають компаніям можливість здійснювати всебічний аналіз даних, що підвищує точність та ефективність прийняття рішень. Консолідація та аналіз великих обсягів даних дозволяє отримувати глибокі інсайти.

Операції, що виконуються на оперативних системах, можуть бути обтяжені аналітичними запитами. Використання сховища даних знижує навантаження на ці системи, що покращує загальну продуктивність.

Сховище даних забезпечує очищення, узгодження та стандартизацію даних, що підвищує їхню достовірність і цілісність. Це критично важливо для точного аналізу та звітності.

Також вони можуть легко адаптуватися до змін у бізнес-вимогах та масштабуватися для обробки зростаючих обсягів даних. Це забезпечує довготривалу ефективність та адаптивність до нових викликів [37].

Сховища даних відіграють ключову роль у сучасній бізнес-аналітиці, забезпечуючи підприємствам можливість здійснювати ефективний та глибокий аналіз даних. Вони об'єднують дані з різних джерел, забезпечують їхню інтеграцію, консистентність та історичність. Це дозволяє компаніям не лише аналізувати минулі події, але й прогнозувати майбутні тенденції, приймати стратегічні рішення та оптимізувати бізнес-процеси. Розуміння основних концепцій і визначень сховищ даних є критично важливим для ефективного використання цієї технології в сучасному бізнес-середовищі [38].

Архітектура сховищ даних визначає, як дані зберігаються, організовуються, інтегруються та доступні для аналітичних цілей. Вона складається з кількох ключових компонентів, кожен з яких виконує специфічні функції для забезпечення ефективної обробки та аналізу великих обсягів даних.

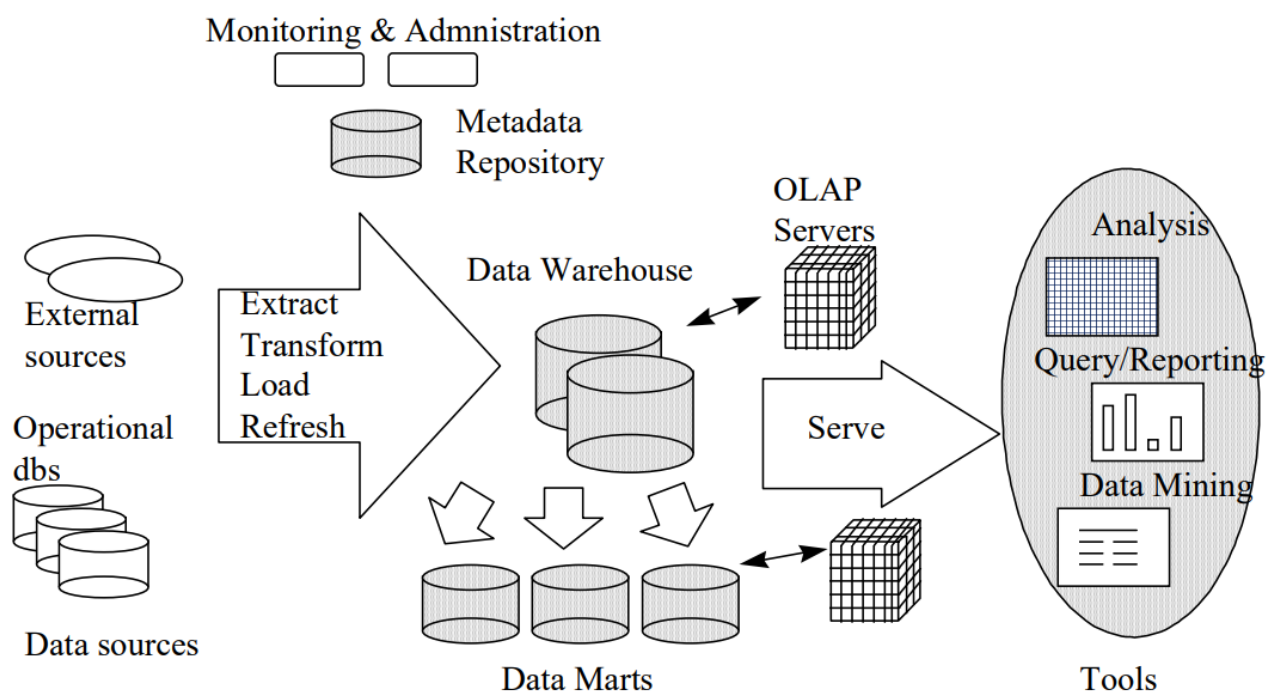


Рисунок 2.2 – Архітектура сховища даних [42]

Джерела даних включають всі системи та платформи, звідки дані збираються для подальшого зберігання у сховищі даних. Це можуть бути

операційні бази даних, ERP-системи, CRM-системи, текстові файли, журнали транзакцій, веб-сервіси та інші зовнішні джерела [39].

Оперативне сховище даних (Operational Data Store, ODS) є проміжним сховищем, яке використовується для короткострокового зберігання та обробки даних перед їх завантаженням у сховище даних. Це дозволяє здійснювати швидкий доступ до оперативних даних для аналізу в режимі реального часу.

Основне або ж централізоване сховище даних — це сховище, де зберігаються всі інтегровані дані в структурованому вигляді. Основне сховище даних підтримує як поточні, так і історичні дані, що дозволяє виконувати аналіз трендів і прогнозування. Дані у сховищі організовані за темами, що відображають ключові бізнес-процеси підприємства.

Март даних (data marts), або ще їх називають вітринами даних, є підмножиною основного сховища даних, оптимізованою для конкретних бізнес-відділів або функцій. Наприклад, можуть існувати окремі марті даних для маркетингу, фінансів, продажів тощо. Це дозволяє спростити доступ до даних та підвищити продуктивність аналітичних запитів.

Метадані (metadata) – це дані про дані. Вони включають інформацію про структуру, формат, джерела, процеси трансформації та інші характеристики даних, що зберігаються у сховищі. Метадані допомагають користувачам розуміти контекст і значення даних, що полегшує їхнє використання та управління.

Інструменти аналітики та звітності надають користувачам можливість виконувати запити, аналізувати дані, створювати звіти та візуалізації. Вони включають OLAP-системи, інструменти бізнес-аналітики (BI) та спеціалізовані програмні продукти для аналізу великих даних [40].

Архітектурні підходи:

У монолітній архітектурі всі компоненти сховища даних зосереджені в єдиній централізованій системі. Цей підхід простий у впровадженні та управлінні, але може бути менш гнучким і масштабованим.

Модульна архітектура як підхід передбачає розподіл компонентів на окремі модулі, які можуть функціонувати незалежно один від одного. Це підвищує

гнучкість і дозволяє легко додавати нові модулі або змінювати існуючі без впливу на всю систему.

Хмарні сховища даних забезпечують масштабованість, гнучкість та зниження витрат на інфраструктуру. Вони дозволяють компаніям швидко адаптуватися до змін у бізнес-вимогах та забезпечують доступ до потужних обчислювальних ресурсів.

Архітектура сховищ даних є складною системою, що включає різні компоненти та процеси для ефективного зберігання, обробки та аналізу даних. Вибір правильної архітектури залежить від специфічних потреб та вимог бізнесу. Ефективна архітектура сховищ даних забезпечує надійність, гнучкість та масштабованість, що дозволяє підприємствам отримувати глибокі інсайти з даних та приймати обґрунтовані рішення.

Процес ETL (Extract, Transform, Load) полягає у зборі, обробці та завантаженні даних з різних джерел у централізоване сховище даних. ETL забезпечує якість даних, їхню інтеграцію та підготовку для аналітичних цілей. Процес складається з трьох основних етапів: екстракції, трансформації та завантаження.

Етап 1: Екстракція (Extract)

Екстракція є першим етапом процесу ETL, під час якого дані вилучаються з різних джерел. Джерела даних можуть бути дуже різноманітними: операційні бази даних, CRM та ERP-системи, файли, веб-сервіси, API, журнали транзакцій та інші. Основна мета екстракції полягає в тому, щоб зібрати необхідні дані для подальшої обробки та завантаження у сховище даних.

Ключові завдання екстракції:

- Ідентифікація джерел даних: Визначення всіх можливих джерел, з яких будуть вилучені дані.
- Вибір підходящої технології: Вибір методів і інструментів для екстракції даних (наприклад, SQL-запити, веб-скрейпінг, API-інтеграції).

- Оптимізація продуктивності: Забезпечення ефективного вилучення великих обсягів даних з мінімальним впливом на продуктивність джерел.

Етап 2: Трансформація (Transform)

Трансформація є другим етапом процесу ETL, під час якого дані обробляються і готуються для завантаження у сховище даних. На цьому етапі дані очищаються, нормалізуються, агрегуються та інтегруються, щоб забезпечити їх консистентність і відповідність бізнес-вимогам.

Ключові завдання трансформації.

- Очищення даних: Виявлення та виправлення помилок, видалення дублікатів, заповнення відсутніх значень.
- Нормалізація: Перетворення даних у стандартизований формат для забезпечення сумісності між різними джерелами.
- Агрегація: Об'єднання даних з різних джерел для створення узагальнених звітів та аналітичних моделей.
- Інтеграція: Об'єднання даних з різних систем для створення єдиного, інтегрованого джерела правди.
- Збагачення даних: Додавання додаткової інформації з зовнішніх джерел для підвищення цінності даних.

Етап 3: Завантаження (Load)

Завантаження є третім і заключним етапом процесу ETL, під час якого трансформовані дані переміщуються у сховище даних або інші системи зберігання. Цей етап забезпечує доступність даних для аналітичних запитів та звітності.

Ключові завдання завантаження:

- Вибір підходящої технології завантаження: Визначення методів завантаження даних (пакетне завантаження, реальне завантаження, мікропакетне завантаження).
- Моніторинг та контроль якості: Перевірка правильності завантажених даних, виявлення та виправлення помилок під час завантаження.
- Оптимізація продуктивності: Забезпечення ефективного завантаження великих обсягів даних без значного впливу на продуктивність системи.

Процес ETL є критично важливим для успішної реалізації сховищ даних та аналітичних систем. Він забезпечує якість, інтеграцію та доступність даних, що є необхідними для прийняття обґрунтованих бізнес-рішень [41]. Ефективна реалізація процесу ETL дозволяє компаніям оптимізувати свої операційні та аналітичні процеси, підвищуючи загальну ефективність і конкурентоспроможність.

OLAP-системи: основи, функціональність та застосування в аналітиці даних.

OLAP (Online Analytical Processing) системи є ключовим елементом сучасної бізнес-аналітики, надаючи засоби для швидкого та інтерактивного аналізу великих обсягів даних. Вони дозволяють користувачам виконувати багатовимірний аналіз даних, що забезпечує гнучкість і швидкість в обробці складних запитів.

Основні концепції OLAP включають:

Багатовимірні куби – структури даних, що представляють інформацію в багатовимірному просторі. Куби складаються з вимірів та фактів. Виміри представляють різні аспекти даних (наприклад, час, географія, продукти), тоді як факти містять числові значення (наприклад, обсяги продажів, доходи).

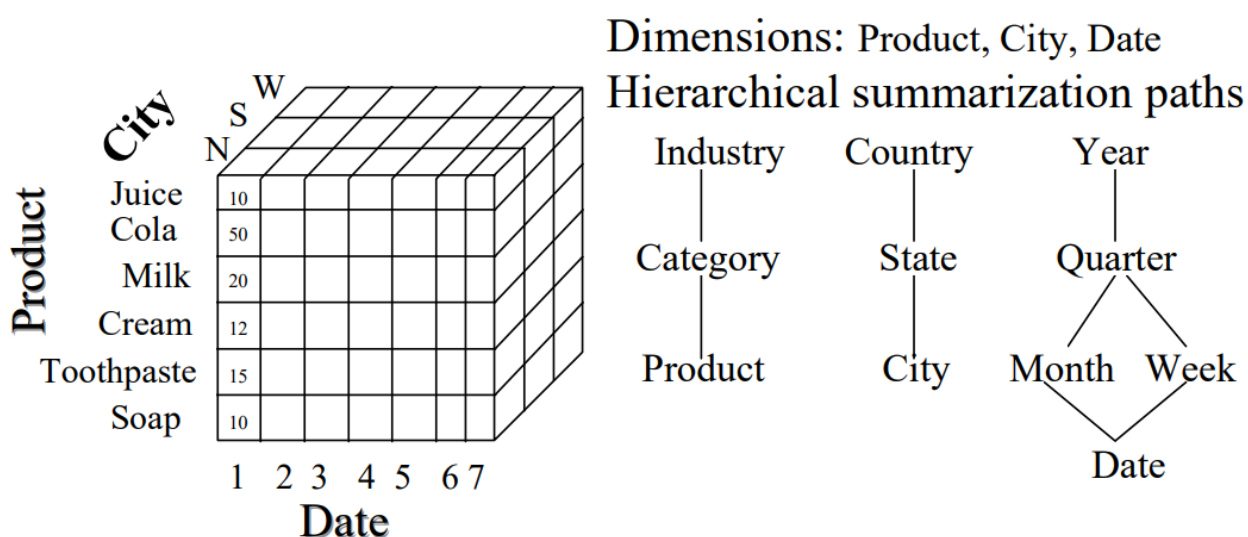


Рисунок 2.3 – Багатовимірні куби даних [42]

Виміри – атрибути, за якими аналізуються дані, такі як час, місце, продукт. Виміри дозволяють користувачам здійснювати різні зрізи даних.

Ієрархії – структури в рамках вимірів, що дозволяють деталізувати дані на різних рівнях (наприклад, рік-місяць-день).

Агрегація: Процес обчислення підсумкових значень за певними вимірами, що дозволяє отримати узагальнену інформацію.

OLAP-системи надають широкий спектр функціональних можливостей, що роблять їх потужним інструментом для аналізу даних.

OLAP забезпечує високопродуктивний доступ до великих обсягів даних завдяки попередній агрегації та оптимізації запитів. За допомогою OLAP також можна виконувати складні аналітичні запити, включаючи розрахунки, порівняння, ієрархічний аналіз та інші багатовимірні операції. Вони підтримують автоматизацію створення та розповсюдження звітів, що дозволяє зменшити час на рутинні завдання.

Користувачі можуть динамічно змінювати запити, додаючи або видаляючи виміри, переглядаючи дані на різних рівнях деталізації.

Drill-Down та Roll-Up: Функції, що дозволяють користувачам переходити від узагальнених даних до детальних (drill-down) та навпаки (roll-up), забезпечуючи глибокий аналіз інформації.

Slicing та Dicing: Методи вибірки даних за певними вимірами (slicing) або зміни перспективи аналізу (dicing), що дозволяють детально досліджувати специфічні аспекти даних.

OLAP-системи широко використовуються у різних сферах для підтримки прийняття рішень та стратегічного планування.

В бізнес-аналітиці вони забезпечує керівництво організацій засобами для аналізу фінансових показників, ефективності продажів, управління запасами та інших ключових аспектів бізнесу. Аналітики використовують OLAP для дослідження ринкових трендів, аналізу поведінки клієнтів, оцінки ефективності маркетингових кампаній.

OLAP-системи дозволяють швидко аналізувати фінансові дані, прогнозувати доходи, витрати та оцінювати фінансові ризики. Також це допомагає

оптимізувати логістичні процеси, знижувати витрати та покращувати управління запасами через детальний аналіз даних постачання та попиту.

OLAP-системи надають численні переваги, такі як швидкий доступ до даних, інтерактивний аналіз, підтримка складних запитів та автоматизація звітності. Водночас, вони мають певні обмеження, включаючи високу вартість впровадження та підтримки, потребу в спеціалізованих навичках для налаштування та використання, а також можливі проблеми з продуктивністю при обробці надзвичайно великих обсягів даних [42].

У підсумку, OLAP-системи є потужним інструментом для бізнес-аналітики, забезпечуючи ефективний аналіз та обробку великих обсягів даних. Вони сприяють покращенню прийняття рішень та оптимізації бізнес-процесів, надаючи керівникам та аналітикам гнучкі та ефективні засоби для роботи з даними.

Моделювання даних у сховищах даних є критично важливим етапом у створенні ефективних і надійних систем зберігання та аналізу даних. Цей процес включає проектування структури сховища даних, яка дозволяє оптимально організувати та керувати великими обсягами інформації, що надходять з різних джерел. Основна мета моделювання даних – забезпечити логічну і фізичну організацію даних таким чином, щоб вони були легко доступними для аналізу і звітності.

Перший крок у моделюванні даних полягає у визначенні вимог до даних. Це включає ідентифікацію джерел даних, розуміння бізнес-процесів, які будуть підтримуватися сховищем даних, а також визначення ключових показників ефективності (KPI), які будуть використовуватися для аналізу.

На етапі концептуального моделювання створюється концептуальна модель даних, яка представляє основні об'єкти та їх взаємозв'язки на високому рівні. Ця модель є абстрактним представленням структури даних і слугує основою для подальшого детального моделювання. Популярними інструментами для концептуального моделювання є діаграми сутність-зв'язок (ERD).

Логічна моделювання деталізує концептуальну модель і визначає структуру даних більш детально. Вона включає атрибути сутностей, типи даних, ключові обмеження та взаємозв'язки між таблицями. Логічна модель є незалежною від конкретної системи управління базами даних (СУБД) і служить для розуміння бізнес-вимог на більш детальному рівні.

На етапі фізичного моделювання логічна модель перетворюється на фізичну, яка враховує специфіку вибраної СУБД. Фізичне моделювання включає створення таблиць, визначення індексів, ключів, переглядів та інших елементів бази даних. Метою фізичного моделювання є реалізація всієї моделі та оптимізація продуктивності та забезпечення ефективного зберігання даних.

Основні принципи моделювання даних:

Цілісність даних – моделі даних повинні забезпечувати цілісність даних, включаючи цілісність сутностей та цілісність посилань. Це гарантує, що всі дані у сховищі є правильними і взаємопов'язаними.

Продуктивність – моделі даних повинні бути оптимізовані для швидкого доступу і обробки великих обсягів даних. Це включає вибір правильних індексів, нормалізацію або денормалізацію таблиць, та використання ефективних запитів.

Масштабованість – сховище даних повинно бути здатне обробляти збільшення обсягів даних і навантаження без значного зниження продуктивності.

Простота та зрозумілість – модель даних повинна бути простою для розуміння і використання, що дозволяє аналітикам і користувачам легко орієнтуватися в структурі даних.

Моделювання даних у сховищах даних є фундаментальним процесом, що визначає ефективність системи зберігання та аналізу даних. Правильне проектування концептуальної, логічної та фізичної моделей забезпечує високу якість, продуктивність та надійність даних, що в свою чергу підтримує прийняття обґрунтованих бізнес-рішень. Моделі даних мають бути гнучкими, масштабованими та легко доступними для забезпечення безперервної підтримки бізнес-потреб.

Інтеграція сховищ даних з аналітичними інструментами дозволяє організаціям отримувати глибокі інсайти з даних, оптимізувати бізнес-процеси та приймати обґрунтовані рішення. Розглянемо детально основні аспекти інтеграції сховищ даних з аналітичними інструментами.

Інтеграція сховищ даних з аналітичними інструментами дозволяє забезпечити доступ до консолідованих даних. Сховища даних об'єднують інформацію з різних джерел, що полегшує доступ до повного набору даних для аналізу.

Також це дозволяє підвищити точність аналізу. Центральне зберігання даних забезпечує їхню узгодженість, що важливо для точного аналізу та побудови надійних аналітичних моделей.

Аналітичні інструменти надають керівникам можливість швидко отримувати необхідну інформацію для прийняття рішень на основі даних.

Серед популярних аналітичних інструментів, які часто інтегруються зі сховищами даних, можна виділити:

Power BI: Інструмент від Microsoft, що дозволяє створювати інтерактивні звіти та панелі, легко інтегрується з SQL Server, Azure, Excel та іншими джерелами.

Tableau: Потужний інструмент для візуалізації даних, який підтримує інтеграцію з різними базами даних, включаючи MySQL, Oracle, PostgreSQL та інші.

Qlik Sense: Інструмент для самостійного створення візуалізацій та аналізу даних, який підтримує роботу з різними типами сховищ даних.

SAS: Програмний комплекс для статистичного аналізу, який широко використовується у великих організаціях для глибокого аналізу даних та моделювання.

Інтеграція сховищ даних з аналітичними інструментами є важливим кроком для ефективного управління та аналізу даних. Вона забезпечує доступ до консолідованих та узгоджених даних, підвищує точність аналізу та оптимізує процеси прийняття рішень. Використання сучасних аналітичних інструментів та методів інтеграції дозволяє організаціям ефективно використовувати свої дані для досягнення стратегічних цілей.

Розвиток технологій, зміна вимог бізнесу та зростання обсягів даних впливають на еволюцію сховищ даних. Розглянемо основні тренди та перспективи розвитку в цій галузі.

Хмарні сховища даних. Одним із головних трендів є перехід до хмарних сховищ даних. Хмарні платформи, такі як Amazon Redshift, Google BigQuery та Microsoft Azure SQL Data Warehouse, надають значні переваги, включаючи масштабованість, гнучкість та економію витрат. Хмарні сховища дозволяють компаніям швидко масштабувати свої ресурси в залежності від потреб, що особливо важливо для бізнесів зі змінним навантаженням.

Розподілені сховища даних, такі як Hadoop та Apache Spark, також стають все більш популярними. Вони забезпечують обробку великих обсягів даних через розподіл завдань на безліч серверів. Це дозволяє ефективно обробляти великі масиви даних у реальному часі та зменшувати час на виконання запитів.

Інтеграція сховищ даних зі штучним інтелектом та машинним навчанням відкриває нові можливості для аналізу та прогнозування. ШІ та МН можуть використовувати дані зі сховищ для тренування моделей, виявлення закономірностей та створення прогнозів. Це дозволяє бізнесам отримувати глибші інсайти та приймати більш обґрунтовані рішення.

Перехід до гібридних архітектур. Вони поєднують локальні та хмарні сховища даних, стають популярними серед організацій, які хочуть зберегти контроль над критично важливими даними, одночасно використовуючи переваги хмарних технологій. Такий підхід забезпечує гнучкість та дозволяє оптимізувати витрати, зберігаючи високий рівень безпеки та доступності даних.

Вдосконалення безпеки даних, що залишається дуже важливим аспектом для будь-якої організації. Розвиток технологій захисту даних, таких як шифрування, багатофакторна автентифікація та моніторинг аномалій, сприяє підвищенню рівня безпеки сховищ даних. Це важливо в умовах зростання кількості кібератак та посилення регуляторних вимог.

Автоматизація процесів управління даними, включаючи ETL та адміністрування баз даних, дозволяє знизити витрати на підтримку сховищ даних та підвищити їхню продуктивність. Інструменти автоматизації допомагають швидко виявляти та виправляти помилки, оптимізувати використання ресурсів та забезпечувати безперервність бізнес-процесів.

Створення інтегрованих екосистем для управління даними, що включають інструменти для зберігання, обробки, аналізу та візуалізації, стає все більш актуальним. Платформи, що забезпечують комплексний підхід до роботи з даними, дозволяють організаціям ефективніше управляти своїми інформаційними ресурсами та отримувати максимальну користь від даних.

У підсумку, розвиток сховищ даних спрямований на підвищення продуктивності, гнучкості та безпеки, що дозволяє організаціям ефективніше використовувати свої дані для прийняття стратегічних рішень. Нові технології та підходи, такі як хмарні рішення, інтеграція з ІІ та МН, реальний час обробки даних та автоматизація, сприяють створенню більш потужних та ефективних систем управління даними.

2.2 Excel як засіб для маніпулювання даними

Microsoft Excel є одним із найпоширеніших та найпотужніших інструментів для роботи з даними. Виробнича сфера, як і багато інших галузей, активно використовує Excel для різноманітних завдань, включаючи аналіз даних, планування, створення звітів та прийняття рішень.

Основна функціональність Excel полягає в його здатності створювати та редагувати електронні таблиці. Робоча книга Excel складається з одного або декількох аркушів, кожен з яких містить комірки, організовані у вигляді рядків і стовпців. Це дозволяє користувачам легко організовувати, зберігати та маніпулювати даними [43].

Excel надає широкий спектр вбудованих функцій для виконання математичних, статистичних, логічних та інших видів обчислень. Користувачі можуть створювати власні формули для автоматизації обчислень, що значно спрощує аналіз даних. Наприклад, функції SUM, AVERAGE, COUNT, IF, VLOOKUP та багато інших дозволяють швидко отримувати необхідні результати.

Щоб зробити таблиці більш зрозумілими та привабливішими, Excel пропонує широкий спектр інструментів для форматування даних. Користувачі можуть змінювати шрифти, кольори, об'єднувати комірки, використовувати умовне форматування для виділення певних значень, а також створювати стилі таблиць для покращення зовнішнього вигляду та читабельності даних.

Функції валідації даних дозволяють контролювати тип та формат даних, що вводяться в комірки. Це допомагає забезпечити точність та консистентність даних, що є критично важливим у виробничій сфері, де помилки можуть призвести до значних витрат. Валідація даних може включати обмеження на введення чисел, дат, тексту певного формату та багато іншого.

Excel забезпечує можливість захисту даних від несанкціонованого доступу та випадкових змін. Користувачі можуть встановлювати паролі на робочі книги та аркуші, обмежувати доступ до редагування окремих комірок або діапазонів, а також використовувати функцію відстеження змін для контролю за модифікаціями.

Можливості для роботи з даними/

Excel надає різноманітні інструменти для візуалізації даних, включаючи діаграми та графіки. Це дозволяє легко представляти дані у візуально привабливій формі, що сприяє кращому розумінню інформації. Користувачі можуть створювати стовпчасті, лінійні, кругові, гістограми та інші типи діаграм, налаштовувати їх вигляд та додавати анотації для покращення інтерпретації даних [44].

Excel дозволяє об'єднувати дані з різних джерел у єдину таблицю для зручного аналізу. Користувачі можуть імпортувати дані з текстових файлів, баз даних, веб-сторінок та інших джерел, а також використовувати функції об'єднання та консолідації для створення єдиної звітності.

Зведені таблиці (PivotTables) є одним з найпотужніших інструментів Excel для аналізу даних. Вони дозволяють швидко агрегувати великі обсяги даних, виконувати багаторівневий аналіз, створювати детальні звіти та виявляти приховані закономірності. Зведені таблиці можуть бути легко налаштовані та адаптовані під потреби користувача, що робить їх незамінним інструментом для аналізу великих обсягів даних.

Додатково Excel підтримує можливість спільної роботи над документами, що дозволяє кільком користувачам одночасно редагувати одну робочу книгу. Це особливо корисно у великих компаніях та виробничих підприємствах, де часто необхідно співпрацювати над великими проектами. Користувачі можуть залишати коментарі, рецензувати зміни та використовувати функції спільного доступу для ефективної командної роботи.

Підсумовуючи, Excel є надзвичайно потужним та гнучким інструментом для роботи з даними, що пропонує широкий спектр функцій для аналізу, візуалізації, консолідації та спільної роботи. Його можливості дозволяють ефективно організовувати та аналізувати дані, приймати обґрунтовані рішення та підвищувати продуктивність у виробничій сфері.

Microsoft Excel надає користувачам можливість імпортувати дані з різних джерел та експортувати їх для обробки, аналізу та подальшого використання. Цей процес імпорту та експорту даних може бути виконаний у різних форматах файлів, відкриваючи широкі можливості для обміну інформацією між Excel та іншими програмами.

Формати файлів для імпорту та експорту даних

1. Excel (XLSX, XLS). Формати файлів Excel, такі як XLSX та XLS, є найпоширенішими для обміну даними між Excel та іншими програмами. Вони підтримують всі типи даних, форматування, діаграми та формули, що дозволяє зберігати дані в їхньому оригінальному вигляді.

2. CSV (Comma-Separated Values). Формат CSV є простим та універсальним способом для зберігання табличних даних у текстовому форматі, де значення

розділені комою. CSV-файли можуть бути легко створені та читабельні для багатьох програм, що робить їх ідеальними для імпорту даних у Excel.

3. TXT (Plain Text). Формат TXT є ще одним простим форматом для зберігання текстових даних у текстовому форматі без форматування. Вони можуть бути імпортовані у Excel для подальшого оброблення, хоча цей формат не підтримує складні типи даних.

4. XML (eXtensible Markup Language). XML є мовою розмітки, яка дозволяє структурувати дані за допомогою тегів. Використання XML для імпорту та експорту даних у Excel дозволяє зберігати структуру та відносини між різними елементами даних.

5. JSON (JavaScript Object Notation). JSON є легким форматом обміну даними, який використовується для зберігання та обміну даними у вигляді пар "ключ-значення". Використання JSON для імпорту та експорту даних у Excel дозволяє зберігати структуровані дані та обробляти їх за допомогою спеціальних засобів.

Методи інтеграції даних з Excel:

1. Імпорт даних через вбудовані функції. Excel надає вбудовані функції для імпорту даних з різних джерел, таких як бази даних, текстові файли, веб-сторінки та інші. Користувачі можуть скористатися функціями, такими як "Дані зовнішнього джерела", "Дані з текстового файлу" або "Дані з веб-запиту", щоб легко імпортувати дані безпосередньо у свою робочу книгу Excel.

2. Експорт даних через збереження у файл. Користувачі можуть експортувати дані з Excel, зберігаючи їх у різних форматах файлів, таких як XLSX, CSV або TXT. Це дозволяє легко обмінюватися даними з іншими програмами або імпортувати їх для подальшого використання.

3. Використання зовнішніх інструментів. Деякі зовнішні програми та сервіси надають інтеграцію з Excel через API або інші інтерфейси програмування. Це дозволяє автоматизувати процеси імпорту та експорту даних, використовуючи зовнішні інструменти або послуги.

4. Використання додатків та додаткових засобів. Для полегшення імпорту та експорту даних у Excel можуть використовуватися різноманітні додатки та додаткові засоби, що розширюють функціональність програми. Наприклад, деякі додатки дозволяють автоматично оновлювати дані з зовнішніх джерел або виконувати складні операції імпорту та експорту.

5. Ручне копіювання та вставлення. Незважаючи на наявність різних інструментів для імпорту та експорту даних, іноді найпростіший спосіб - це просто скопіювати дані з одного джерела та вставити їх у Excel, або навпаки. Цей метод може бути швидким та ефективним для невеликих обсягів даних.

Загалом, Excel надає користувачам широкі можливості для імпорту та експорту даних, що дозволяє легко обмінюватися інформацією з іншими програмами та джерелами даних. Різноманітність форматів файлів та методів інтеграції робить Excel потужним інструментом для роботи з даними у виробничій сфері.

Робота з великими обсягами даних у Excel може бути великим викликом через обмеження програми на обробку великої кількості інформації. Проте, існують різні поради та стратегії, які допоможуть зробити цей процес більш ефективним та продуктивним.

Створення таблиць з правильною структурою є ключовим аспектом роботи з великими обсягами даних. Потрібно уникати зайвих комірок та стовпців, а також використовувати функції Excel для автоматизації обробки даних.

Excel надає можливості фільтрації та сортування даних, що дозволяє швидко знаходити потрібну інформацію. Важливо вміти користуватися цими функціями для швидкого доступу до потрібних даних.

Зведені таблиці дозволяють швидко агрегувати та аналізувати великі обсяги даних. Функції агрегування, такі як Сума, Середнє, Мінімум та Максимум, можуть бути використані для швидкого розрахунку підсумків.

Надання імені діапазонам даних дозволяє швидко та легко використовувати їх у формулах та функціях Excel. Це полегшує читабельність та редагування формул, особливо у великих робочих книгах з багатьма діапазонами.

Макроси та мова програмування VBA можуть бути використані для автоматизації рутинних завдань та оптимізації продуктивності. Вони дозволяють створювати власні функції та процедури для роботи з даними у Excel.

Для підвищення продуктивності можна розподілити дані на різні аркуші за темами чи категоріями. Це дозволить зменшити обсяг даних на кожному аркуші та полегшить навігацію між різними частинами робочої книги.

Для роботи з великими обсягами даних можна використовувати зовнішні з'єднання з базами даних або іншими джерелами даних, а також розширені додатки та плагіни, які доповнюють функціональність Excel.

Робота з великими обсягами даних у Excel вимагає використання правильних стратегій для забезпечення ефективності та продуктивності. Правильна організація даних, використання вбудованих функцій та фільтрів, а також вміння користуватися зовнішніми джерелами та макросами допоможуть оптимізувати процес роботи з даними у Excel [45].

Щоб зробити інформацію більш зрозумілою та доступною для аналізу, потрібно використовувати візуалізації. Для цього можна використовувати інструменти форматування та організації даних, які надає Excel. Важливо забезпечити чітке розташування даних у вигляді рядків і стовпців, а також використовувати заголовки для кожного стовпця, що дозволить краще розуміти зміст таблиці.

Excel пропонує різноманітні типи діаграм, такі як стовпчасті, кругові, лінійні тощо. Вибір типу діаграми залежить від природи даних і мети візуалізації. Для створення діаграми необхідно виділити дані у таблиці, а потім вибрати відповідний тип діаграми з панелі інструментів Excel. Після цього Excel автоматично побудує діаграму на основі вибраних даних.

Після створення діаграми можна налаштувати її вигляд та стиль. В Excel доступні різноманітні опції форматування, такі як зміна кольорів, шрифтів, розмірів тощо. Також можна змінювати типи ліній, додавати легенду, підписи до осей та інші деталі.

Для поліпшення розуміння даних можна додавати додаткові елементи до діаграми, такі як анотації, текстові поля, стрілки тощо. Це допомагає виділити ключову інформацію та зробити діаграму більш зрозумілою для аудиторії.

Необхідно надавати кожній діаграмі чіткі заголовки та підписи. Це допомагає зрозуміти, що саме вона відображає, а також полегшує сприйняття інформації.

Після створення та налаштування діаграми важливо зберегти її в потрібному форматі. Excel дозволяє зберігати діаграми у форматах зображень або вбудовувати їх у інші документи.

Зведені таблиці є потужним інструментом у Excel для швидкого та ефективного аналізу великих обсягів даних та створення звітів на їх основі. Для їх створення у Excel використовується функція "Зведена таблиця". Ця функція дозволяє об'єднати дані з різних джерел та виконувати різноманітні операції з ними, такі як сумування, підрахунок кількості, обчислення середнього значення тощо.

Перш ніж створити зведену таблицю, необхідно вибрати дані, з якими буде проводитися аналіз. Дані можуть бути вже введені в таблицю Excel або імпортовані з інших джерел. Після вибору даних потрібно визначити, які поля будуть використовуватися для аналізу. В зведеній таблиці можна використовувати різні поля для групування, фільтрації та обчислень.

Однією з ключових можливостей зведених таблиць є можливість виконувати різні операції з даними, такі як сумування, підрахунок кількості, обчислення середнього значення тощо. Ці операції дозволяють отримати необхідну статистичну інформацію про дані. Також в зведених таблицях можна виконувати

фільтрацію та сортування даних за різними критеріями. Це дозволяє здійснювати більш детальний аналіз та отримувати звіти за певними категоріями.

Після створення зведеної таблиці та виконання необхідних операцій з даними можна створювати звіти на її основі. Excel надає можливості форматування та оформлення звітів, а також додавання графіків та діаграм для кращого візуального представлення результатів аналізу.

Після генерації звітів важливо перевірити їх на достовірність та коректність. При необхідності можна вносити коригування у зведені таблиці або у вихідні дані для отримання більш точних результатів.

Створення та використання зведених таблиць у Excel є важливим етапом у аналізі даних та генерації звітів. Цей інструмент дозволяє швидко та ефективно обробляти великі обсяги даних, виконувати різноманітні операції з ними та створювати звіти, які допомагають у прийнятті управлінських рішень.

Excel надає широкі можливості для фільтрації та сортування даних, що робить процес аналізу більш ефективним та продуктивним.

Фільтрація даних у Excel дозволяє відображати лише ті рядки даних, які задовольняють певні критерії. Для цього можна використовувати автоматичні фільтри, які вбудовані у Excel, або вручну налаштовувати фільтри за різними параметрами, такими як значення, текст, дата тощо. У Excel доступні розширені можливості фільтрації, такі як фільтрація за кольором, додаткові умови фільтрації та складні комбіновані фільтри. Це дозволяє точніше налаштовувати фільтри для відбору необхідних даних.

Сортування даних дозволяє впорядковувати рядки даних за певними критеріями, такими як значення у стовпцях, алфавітний порядок або дата. Це дозволяє зручно організувати дані та знайти необхідну інформацію. Excel підтримує і мультирівневе сортування, що дозволяє впорядковувати дані за кількома критеріями одночасно. Це дозволяє проводити більш детальний аналіз даних та знаходити необхідну інформацію за різними параметрами.

Excel надає можливість фільтрувати дані за значеннями у певних стовпцях. Це дозволяє відфільтрувати дані, які відповідають певним критеріям, і виконувати аналіз лише на підмножині даних. У Excel доступні динамічні фільтри, які автоматично оновлюються при зміні даних. Це дозволяє зберігати актуальність результатів аналізу та спрощує процес роботи з даними [46].

Інструменти фільтрації та сортування даних у Excel надають широкі можливості для проведення ефективного аналізу даних та отримання потрібної інформації. Вони дозволяють швидко та зручно відбирати, організовувати та аналізувати великі обсяги даних, що робить Excel потужним інструментом для роботи з даними в різних галузях та сферах діяльності.

Для збільшення ефективності та автоматизації рутинних завдань можна використовувати макроси та мову програмування VBA (Visual Basic for Applications).

Макроси — це набір інструкцій, які записуються для автоматичного виконання певних дій у Excel. Вони можуть бути записані користувачем або згенеровані автоматично під час записування дій користувача. VBA — це мова програмування, яка дозволяє створювати складніші макроси з використанням умов, циклів та інших конструкцій.

У Excel можна записувати макроси за допомогою вбудованого засобу "Макроси". Після запису Excel зберігає його у вигляді VBA-коду. Записаний макрос можна редагувати та вдосконалювати вручну за допомогою VBA-редактора. У цьому редакторі можна створювати нові макроси, редагувати існуючі, додавати коментарі, встановлювати змінні та виконувати різні дії з кодом.

За допомогою VBA можна створювати складні макроси, які виконують різноманітні завдання, такі як автоматизація введення даних, генерація звітів, оптимізація аналітичних процесів тощо. Використання VBA дозволяє значно зменшити час, необхідний для виконання цих завдань, та знизити ймовірність помилок. Макроси можуть бути використані для автоматизації рутинних операцій, таких як копіювання, вставка даних, форматування таблиць, обробка даних тощо.

Це дозволяє зосередитися на більш складних завданнях та підвищує продуктивність роботи.

За допомогою макросів та VBA можна створювати інтерактивні додатки у Excel, які дозволяють користувачам взаємодіяти з даними та виконувати різноманітні завдання, такі як фільтрація, сортування, аналіз тощо, використовуючи вбудовані елементи програмування інтерфейсу аркуша (кнопки. Скролбари тощо).

Використання макросів та VBA для автоматизації завдань у Excel є потужним інструментом для підвищення продуктивності та ефективності роботи з даними. Вони дозволяють автоматизувати рутинні операції, виконувати складні аналітичні завдання та створювати інтерактивні додатки, що значно спрощує та прискорює роботу з Excel.

Excel може бути використаний як частина більшої екосистеми програм та баз даних для розширеного аналізу даних. Інтеграція Excel з іншими програмами та базами даних надає можливість працювати з різноманітними джерелами даних та ефективно аналізувати їх.

Excel має вбудовані засоби імпорту даних з різних джерел, таких як бази даних, текстові файли, CSV файли, веб-сторінки та інші. Це дозволяє легко отримувати дані з різних джерел для подальшого аналізу. Можна також використовувати зовнішні з'єднання для отримання даних з інших програм, таких як бази даних SQL, SAP, CRM-системи та інші. Це дозволяє підключатися до великих обсягів даних та виконувати складні аналітичні завдання безпосередньо в Excel.

Excel може бути інтегрований з іншими програмами за допомогою макросів та VBA. Це дозволяє автоматизувати обмін даними між Excel та іншими програмами, виконувати складні обчислення та обробку даних, а також створювати інтерактивні додатки для спільної роботи з даними.

У Excel є можливість встановлювати різні додатки та плагіни, які розширюють його функціональність. Наприклад, додатки для інтеграції з

соціальними медіа, сервісами хмарного сховища, онлайн-банкінгом та іншими сервісами. Це дозволяє розширити можливості Excel та зручно працювати з різними джерелами даних.

Excel може бути інтегрований з хмарними сервісами, такими як Google Sheets, Microsoft 365, Dropbox та інші. Це дозволяє працювати з даними у реальному часі та спільно редагувати їх з іншими користувачами.

Інтеграція Excel з іншими програмами та базами даних є важливим аспектом роботи з даними. Це дозволяє отримувати доступ до різноманітних джерел даних, розширює функціональність Excel та спрощує аналіз та обробку даних. Використання цих можливостей дозволяє підвищити ефективність та продуктивність роботи з даними у Excel.

Щоб підсумувати все описане вище про Microsoft Excel, можна провести SWOT-аналіз цього програмного засобу, продемонструвавши його сильні та слабкі сторони, можливості та загрози при його використанні при аналізі даних, використовуючи відповідну діаграму із цими показниками (рис. 2.4).



Рисунок 2.4 – SWOT-аналіз програмного засобу Excel

(Джерело розроблене автором)

2.3 Power BI як інструмент побудови аналітичних моделей даних

Power BI є сучасним інструментом від Microsoft, що забезпечує широкі можливості для збору, обробки, аналізу та візуалізації даних. Його головна мета — допомогти організаціям приймати обґрунтовані рішення на основі даних, що генеруються в процесі їхньої діяльності. Power BI дозволяє зібрати дані з різних джерел, об'єднати їх в єдину модель і створити інформативні звіти та панелі, які можна легко зрозуміти і використовувати для подальших дій.

Основні структурні елементи Power BI:

Datasets (набори даних): Набори даних є основою для створення звітів та панелей. Вони включають дані з різних джерел, які були оброблені та підготовлені для аналізу. Power BI дозволяє створювати набори даних з використанням різноманітних джерел, таких як SQL Server, Excel, текстові файли, веб-сервіси та інші.

Tiles (плитки): Плитки представляють окремі візуалізації або значення, які додаються на панелі для представлення ключових показників. Користувачі можуть налаштовувати плитки відповідно до своїх потреб, щоб зосередитися на найбільш важливих аспектах бізнесу.

Dashboards (панелі): Панелі є основними інструментами для візуалізації даних в Power BI. Вони представляють собою набір візуальних компонентів, таких як графіки, діаграми, таблиці та карти, які забезпечують огляд ключових показників та трендів. Користувачі можуть налаштовувати панелі відповідно до своїх потреб, що дозволяє швидко отримувати потрібну інформацію.

Reports (звіти): Звіти в Power BI складаються з декількох сторінок, кожна з яких містить різні візуалізації та дані. Звіти дозволяють глибше аналізувати дані та виявляти деталі, які можуть бути приховані на панелях. Вони забезпечують можливість інтерактивного аналізу, де користувачі можуть фільтрувати, сортувати та досліджувати дані за різними параметрами.

Можливості Power BI:

Інтеграція з різними джерелами даних: Power BI підтримує підключення до широкого спектру джерел даних, що дозволяє збирати та об'єднувати інформацію з різних систем. Це включає бази даних, хмарні сервіси, файли та інші джерела. Завдяки цьому користувачі можуть створювати комплексні моделі даних, які відображають всі аспекти їхнього бізнесу.

Очищення та трансформація даних: З допомогою Power Query, вбудованого інструменту для обробки даних, користувачі можуть очищати та трансформувати дані перед аналізом. Це включає видалення дублікатів, заповнення пропусків, об'єднання таблиць та багато іншого.

Візуалізація даних: Power BI надає широкий вибір візуалізацій, включаючи стовпчасті, лінійні, кругові діаграми, теплові карти, географічні карти та інші. Це дозволяє користувачам вибирати найбільш підходящий спосіб представлення даних залежно від їхніх аналітичних потреб.

Можливості для розширеного аналізу: За допомогою DAX (Data Analysis Expressions) користувачі можуть створювати складні обчислення та аналізувати дані на більш глибокому рівні. DAX дозволяє створювати розрахункові стовпці, міри та інші елементи, що допомагають виявити взаємозв'язки та тренди в даних.

Інтерактивність та взаємодія: Звіти та панелі в Power BI є інтерактивними, що дозволяє користувачам взаємодіяти з даними в реальному часі. Вони можуть фільтрувати, сортувати та досліджувати дані, щоб знайти відповіді на свої питання. Це підвищує ефективність аналізу та дозволяє приймати більш обґрунтовані рішення.

Публікація та спільний доступ: Power BI дозволяє легко публікувати створені звіти та панелі, а також надавати доступ до них іншим користувачам. Це сприяє колективній роботі та забезпечує прозорість та доступність аналітичної інформації для всіх зацікавлених сторін.

Мобільний доступ: Power BI має мобільні додатки для різних платформ, що дозволяє користувачам отримувати доступ до своїх даних та звітів з будь-якого місця. Це забезпечує гнучкість та мобільність у роботі з даними.

Power BI є потужним інструментом для аналізу та візуалізації даних, який надає користувачам широкі можливості для збору, обробки та аналізу інформації. Завдяки інтеграції з різними джерелами даних, можливостям очищення та трансформації даних, багатому вибору візуалізацій та розширеним аналітичним можливостям, Power BI допомагає організаціям приймати обґрунтовані рішення на основі даних [47-50].

Однією з ключових переваг Power BI є його здатність інтегруватися з різноманітними джерелами даних, що дозволяє користувачам збирати, об'єднувати та аналізувати інформацію з різних систем і форматів. Нижче наведено огляд основних джерел даних, з якими можна працювати в Power BI.

Excel є одним з найпоширеніших інструментів для роботи з даними, і Power BI має чудову інтеграцію з цим форматом. Користувачі можуть імпортувати дані з Excel-таблиць, а також оновлювати ці дані в режимі реального часу. Це забезпечує можливість легко інтегрувати існуючі дані з Excel у більш складні аналітичні моделі Power BI [51].

Імпорт даних з CSV-файлів також є поширеним форматом для зберігання та обміну даними. Користувачі можуть налаштовувати імпорт, щоб враховувати різні формати та роздільники.

Текстові файли з фіксованою або змінною шириною полів також підтримуються Power BI. Це дозволяє імпортувати дані з різних систем, які використовують текстові формати для експорту.

Інтеграція з SQL Server дозволяє підключатися до реляційних баз даних, виконувати запити та імпортувати дані. Це забезпечує можливість роботи з великими обсягами даних і створення комплексних аналітичних моделей.

Підключення до Oracle Database дозволяє користувачам отримувати доступ до даних з корпоративних баз даних та використовувати їх для аналізу.

Такі популярні системи управління базами даних як MySQL та PostgreSQL також підтримуються в Power BI, що дозволяє підключатися до них, виконувати запити та імпортувати дані для подальшого аналізу.

Power BI інтегрується з різними хмарними сервісами та онлайн-платформами Azure, такими як Azure SQL Database, Azure Blob Storage та Azure Data Lake. Це дозволяє користувачам легко підключатися до даних, збережених у хмарі, і використовувати їх для аналізу.

Підтримується підключення до Google Analytics, що дозволяє отримувати дані про веб-трафік, поведінку користувачів та інші важливі метрики для аналізу.

Інтеграція з Salesforce дозволяє користувачам Power BI підключатися до даних про продажі, клієнтів та інші комерційні метрики, що зберігаються в цій платформі.

Power BI дозволяє підключатися до даних, збережених у SharePoint, що забезпечує доступ до корпоративної інформації та документів для аналізу.

Інструмент може імпортувати дані з веб-сторінок, що дозволяє використовувати дані з інтернет-ресурсів для створення аналітичних моделей. Це може включати таблиці та іншу структуровану інформацію, доступну на веб-сторінках.

Power BI підтримує протокол OData, що дозволяє підключатися до різних веб-служб і витягувати дані для аналізу. Це забезпечує можливість інтеграції з багатьма сучасними веб-додатками та сервісами.

Використання REST API дозволяє Power BI підключатися до сторонніх додатків і сервісів, що надають дані через API. Це відкриває широкі можливості для інтеграції та використання даних з різних джерел.

Power BI підтримує розширення через додаткові конектори, що дозволяє підключатися до спеціалізованих або рідкісних джерел даних. Це забезпечує гнучкість та розширює можливості інтеграції.

Як ми вже переконались, Power BI надає користувачам широкий вибір інструментів та можливостей для підключення до різних джерел даних. Це забезпечує гнучкість та масштабованість аналізу, дозволяючи об'єднувати дані з різних систем і створювати комплексні аналітичні моделі. Завдяки інтеграції з різноманітними джерелами даних, Power BI допомагає організаціям отримувати

цілісне уявлення про їхню діяльність та приймати обґрунтовані рішення на основі актуальної інформації [52].

Power Query — це потужний інструмент, інтегрований у Power BI, який дозволяє користувачам завантажувати, очищати та трансформувати дані з різних джерел перед їх аналізом. Цей інструмент відіграє ключову роль у процесі підготовки даних, забезпечуючи їхню якість та узгодженість, що є основою для подальшого аналізу.

Очищення даних є критичним етапом у процесі підготовки даних, що включає видалення або виправлення неточностей та інших аномалій. Power Query пропонує широкий набір функцій для цього завдання:

Power Query дозволяє легко знаходити та видаляти дублікати записів, що забезпечує унікальність даних.

Інструмент надає можливість знаходити та заповнювати пропуски у даних. Це може включати заповнення порожніх полів певними значеннями або видалення рядків з пропусками.

Часто дані з різних джерел мають різні типи, які необхідно узгодити. Power Query дозволяє перетворювати типи даних, наприклад, з тексту на число або дату.

Інструмент надає можливість очищати текстові дані, видаляючи зайві пробіли, змінюючи регістр тексту або виконуючи пошук і заміну.

Трансформація даних передбачає зміну їх структури або формату, щоб підготувати їх для аналізу. Power Query пропонує широкий спектр функцій для трансформації даних.

Об'єднання даних з різних джерел – Power Query дозволяє об'єднувати дані з різних джерел, таких як таблиці Excel, бази даних або веб-сервіси, використовуючи операції злиття та приєднання.

Розділення та злиття стовпців – інструмент дозволяє розділяти стовпці на декілька частин або зливати кілька стовпців в один. Це може бути корисно для створення нових змінних або приведення даних до бажаного формату.

Фільтрація даних – Power Query дозволяє фільтрувати дані на основі певних умов, таких як значення у стовпці, діапазони дат або логічні вирази. Це допомагає зосередитися на релевантних даних для подальшого аналізу.

Групування даних – функція групування дозволяє агрегувати дані, обчислюючи суму, середнє, мінімальне або максимальне значення для певних груп. Це корисно для підготовки зведених даних або розрахунку метрик.

Створення обчислювальних стовпців – Power Query дозволяє створювати нові стовпці на основі обчислень, виконаних над існуючими даними. Це може включати арифметичні операції, логічні вирази або використання функцій.

Power Query підтримує автоматизацію процесів очищення та трансформації даних за допомогою мовного редактора M. Це дозволяє записувати та зберігати послідовність трансформацій як скрипт, який можна повторно використовувати та змінювати. Автоматизація допомагає зменшити ручну працю та забезпечує узгодженість результатів.

Також Power Query надає користувачам інтерактивний інтерфейс, що дозволяє в режимі реального часу бачити результати трансформацій. Це забезпечує гнучкість та дозволяє експериментувати з різними підходами до очищення та трансформації даних, швидко оцінюючи їх вплив на кінцевий результат.

Після завершення процесу очищення та трансформації даних, їх можна безпосередньо використовувати для створення візуалізацій, звітів та панелей управління в Power BI. Це забезпечує безперебійний процес від збору даних до їх аналізу та візуалізації.

Power Query є незамінним інструментом для підготовки даних у Power BI. Його потужні можливості з очищення та трансформації даних дозволяють забезпечити якість та узгодженість даних, що є основою для точного та ефективного аналізу. Завдяки інтерактивному інтерфейсу та можливостям автоматизації, Power Query значно спрощує процес підготовки даних, роблячи його більш ефективним та зручним для користувачів [53].

DAX (Data Analysis Expressions) — це формульна мова, використовувана в Power BI, Power Pivot, та Analysis Services, що дозволяє здійснювати гнучкі та потужні обчислення над даними. DAX містить функції, оператори та константи, які можна використовувати в формулах або виразах для обчислення та маніпулювання даними. Цей інструмент є невід'ємною частиною аналітичної роботи в Power BI, особливо коли мова йде про створення розрахункових стовпців та мір. Загалом розроблений для роботи з даними в таблицях моделі даних.

DAX включає понад 200 функцій, які поділяються на різні категорії, такі як функції часу, логічні функції, інформаційні функції, математичні функції тощо.

Ця формульна мова підтримує стандартні арифметичні оператори (додавання, віднімання, множення, ділення), логічні оператори (AND, OR, NOT) та оператори порівняння. DAX дозволяє використовувати константи, такі як числа, текстові значення, дати тощо, у формулах.

Розрахункові стовпці створюються на основі існуючих даних у таблицях і використовуються для додавання нових стовпців із обчисленими значеннями. Вони визначаються за допомогою DAX-формул і зберігаються в таблицях моделі даних. Розрахункові стовпці обчислюються для кожного рядка таблиці окремо.

Приклад: Припустимо, є таблиця "Продажі" з колонками з числовими даними "Кількість" та "Ціна за одиницю". Для обчислення загальної вартості кожного продажу можна створити розрахунковий стовпчик з наступною формулою мовою DAX:

$$\text{ЗагальнаВартість} = [\text{Кількість}] * [\text{Ціна за одиницю}]$$

Цей стовпчик буде автоматично обчислений для кожного рядка таблиці "Продажі".

Міри (Measures) використовуються для обчислень, які виконуються на рівні агрегатів. Вони не зберігаються в таблицях як стовпці, а обчислюються динамічно під час аналізу даних, наприклад, у зведених таблицях або звітах. Міри визначаються DAX-формулами, які зазвичай включають функції агрегування, такі як SUM, AVERAGE, MIN, MAX тощо.

Приклад: Для обчислення загальної суми продажів можна створити міру з наступною формулою:

ЗагальнаСумаПродажів = SUM('Продажі'[ЗагальнаВартість])

Ця міра буде обчислювати загальну вартість продажів для кожного елемента, в контексті якого вона використовується (наприклад, для кожного місяця або продукту).

Однією з найпотужніших категорій функцій DAX є функції часу, які дозволяють виконувати обчислення на основі дат. Вони включають функції для обчислення поточного місяця, кварталу, року, створення рухомих середніх періодів, порівняння періодів тощо.

Приклад: Для обчислення накопичувальної суми продажів з початку року можна використати наступну формулу:

НакопичувальнаСума = CALCULATE(SUM('Продажі'[ЗагальнаВартість]),
DATESYTD('Календар'[Дата]))

Ця формула використовує функцію CALCULATE для зміни контексту обчислення та функцію DATESYTD для вибору дат з початку поточного року.

Контекст обчислень є важливою концепцією в DAX. Існують два типи контексту: рядковий та фільтровий. Рядковий контекст відноситься до обчислень, які виконуються для кожного окремого рядка, тоді як фільтровий контекст включає фільтри, застосовані до даних, наприклад, у зведених таблицях або звітах. Функції, такі як CALCULATE, можуть змінювати фільтровий контекст для виконання обчислень у заданих умовах.

DAX — це потужний інструмент для створення розрахункових стовпців та міри у Power BI, що дозволяє виконувати складні обчислення та аналіз даних. Розуміння та вміння використовувати DAX відкриває безмежні можливості для глибокого аналізу даних, побудови аналітичних моделей та генерації інформативних звітів. Використання DAX забезпечує точність, гнучкість та ефективність аналізу, що є критично важливим у сучасному бізнесі.

Power BI дозволяє об'єднувати дані з різних джерел, створюючи комплексні моделі для глибокого аналітичного опрацювання. Однією з ключових можливостей Power BI є моделювання зв'язків між даними. Це дозволяє користувачам створювати взаємозв'язки між таблицями даних, що забезпечує ефективний аналіз та інтеграцію даних з різних джерел.

Зв'язки між таблицями у Power BI встановлюються на основі спільних полів, таких як ідентифікатори або ключі. Наприклад, таблиця продажів може мати поле "ProductID", яке зв'язується з полем "ProductID" у таблиці продуктів. Ці зв'язки можуть бути різних типів:

- **один-до-одного (One-to-One):** Кожен рядок однієї таблиці відповідає одному рядку іншої таблиці. Цей тип зв'язку використовується рідко, але може бути корисним для об'єднання даних з двох таблиць, що мають однакову структуру.
- **один-до-багатьох (One-to-Many):** Один рядок однієї таблиці може відповідати багатьом рядкам іншої таблиці. Це найпоширеніший тип зв'язку, наприклад, одна категорія продуктів може мати багато продуктів.
- **багато-до-багатьох (Many-to-Many):** Рядки з однієї таблиці можуть відповідати багатьом рядкам іншої таблиці і навпаки. Цей тип зв'язку використовується рідше, але підтримується Power BI для складних сценаріїв.

Щоб створити зв'язки між таблицями в Power BI, необхідно перейти до режиму моделювання, та вказівником миші перетягнути поле з однієї таблиці до відповідного поля іншої таблиці.

Power BI автоматично визначає тип зв'язку та встановлює його. Користувачі можуть змінювати властивості зв'язку, такі як напрямок фільтрації та активність зв'язку.

Напрямок фільтрації визначає, як фільтри застосовуються до зв'язаних таблиць. У Power BI існує два типи напрямків фільтрації:

одностороння фільтрація (Single-direction) – фільтри застосовуються в одному напрямку, наприклад, від таблиці продуктів до таблиці продажів.

двостороння фільтрація (Both-direction) – фільтри застосовуються в обох напрямках, що дозволяє фільтрувати дані в обох таблицях. Це корисно для складних моделей даних, але може впливати на продуктивність.

Активний зв'язок використовується за замовчуванням для обчислень і візуалізацій. Неактивні зв'язки можуть використовуватися у формульних виразах за допомогою функції USERRELATIONSHIP для вибіркового обчислення.

Зв'язки між таблицями дозволяють створювати комплексні звіти та аналізувати дані з їх урахуванням. Наприклад, зв'язок між таблицею продажів та таблицею продуктів дозволяє аналізувати продажі за категоріями продуктів, брендами, або іншими атрибутами. Це дозволяє отримати більш глибоке розуміння даних та виявити приховані взаємозв'язки.

Моделювання зв'язків між даними у Power BI є ключовою складовою для створення комплексних та інформативних аналітичних звітів. Правильне налаштування зв'язків дозволяє здійснювати глибокий аналіз, виявляти тренди та кореляції, а також оптимізувати процес прийняття рішень. Використання зв'язків значно розширює можливості аналізу даних та робить Power BI потужним інструментом для бізнес-аналітики.

Power BI — це потужна платформа для бізнес-аналітики, яка надає широкий спектр інструментів для візуалізації даних. Завдяки різноманітності типів візуалізацій користувачі можуть ефективно представляти складну інформацію у зрозумілій і наочній формі. Розглянемо детальніше основні види візуалізацій, доступних у Power BI, та можливості налаштування цих візуалізацій.

Стовпчикові діаграми – використовуються для порівняння величин між різними категоріями. Доступні варіанти: прості стовпчикові діаграми, групові стовпчикові діаграми, складені стовпчикові діаграми та 100% складені стовпчикові діаграми.

Лінійні діаграми – ідеально підходять для відображення змін даних протягом певного періоду. Дозволяють виявляти тренди та сезонні коливання.

Кругові діаграми – використовуються для відображення частки кожної категорії в загальному обсязі. Можуть бути сегментованими для кращого розуміння структури даних.

Гістограми – допомагають у розподілі частот для числових даних. Відмінно підходять для аналізу розподілу значень та виявлення аномалій.

Географічні діаграми – використовуються для відображення даних, прив'язаних до географічних координат. Доступні різні типи: карти з бульбашками, карти з тепловими точками, прості карти.

Таблиці та матриці – допомагають представити детальну інформацію в табличному форматі. Таблиці відображають дані у вигляді рядків і стовпців, а матриці дозволяють групувати дані за різними критеріями.

Секторні діаграми – подібні до кругових діаграм, але з додатковим центральним простором. Дозволяють відображати кілька рівнів інформації одночасно.

Воронкоподібні діаграми – використовуються для відображення процесів або шляхів, де кількість даних поступово зменшується на кожному етапі. Застосовуються для аналізу воронок продажів, процесів обробки та інших подібних процесів.

Графіки розсіювання – використовуються для відображення зв'язків між двома або більше числовими змінними. Дозволяють виявляти кореляції та взаємозв'язки між даними.

Комбіновані діаграми – поєднують різні типи діаграм, такі як стовпчикові та лінійні діаграми, для відображення комплексних даних.

Power BI надає різні аспекти для налаштування візуалізацій.

- налаштування зовнішнього вигляду – користувачі можуть налаштовувати кольори, шрифти, стилі та інші візуальні елементи діаграм. Доступні різні теми та шаблони для швидкого застосування стилів.

- форматування даних – можливість налаштування форматів чисел, дат, валют та інших типів даних. Підтримка умовного форматування для виділення важливих даних.

- інтерактивність – візуалізації у Power BI інтерактивні, що дозволяє користувачам фільтрувати дані та деталізувати інформацію на основі їхніх запитів. Інтерактивні елементи, такі як слайдери, кнопки та фільтри, забезпечують гнучкість аналізу.

- анімація – деякі візуалізації підтримують анімацію, що дозволяє відслідковувати зміни даних у динаміці. Це особливо корисно для презентацій та звітів, де важливо підкреслити зміни у часі.

- користувацькі візуалізації – Power BI підтримує додавання користувацьких візуалізацій з Microsoft AppSource або створення власних. Це дозволяє розширити можливості стандартних візуалізацій та адаптувати їх під специфічні потреби.

- інтеграція з іншими інструментами – візуалізації можуть бути інтегровані з іншими інструментами та платформами, такими як Excel, SharePoint та Teams. Це забезпечує більш широкі можливості для співпраці та обміну даними.

Різноманітність візуалізацій у Power BI надає користувачам потужні інструменти для аналізу та представлення даних. Завдяки гнучким можливостям налаштування, інтерактивності та підтримці користувацьких візуалізацій, Power BI дозволяє ефективно працювати з даними та отримувати цінні інсайти для прийняття обґрунтованих рішень. Візуалізації є ключовим елементом у процесі аналітики, допомагаючи перетворити сирі дані на зрозумілі та наочні графічні представлення [54].

Power BI пропонує широкий спектр можливостей для публікації та розповсюдження аналітичних звітів і панелей, забезпечуючи ефективне співробітництво та обмін даними між користувачами. Розглянемо детально основні аспекти цього процесу.

Сама ж публікація звітів є одним із останніх етапів роботи в Power BI, який дозволяє зробити аналітичні дані доступними для ширшої аудиторії. Після створення звіту в Power BI Desktop, його можна опублікувати в Power BI Service — хмарній платформі Microsoft для зберігання та спільного використання звітів.

Для публікації звіту потрібно натиснути кнопку «Опублікувати» в Power BI Desktop. Звіт завантажується на Power BI Service, де його можна переглядати та аналізувати в режимі онлайн. Після публікації звіт стає доступною через веб-інтерфейс або мобільні додатки Power BI.

Можна налаштувати рівні доступу для різних користувачів, визначаючи, хто може переглядати або редагувати звіти. Використання груп безпеки та ролей допомагає забезпечити конфіденційність даних та контроль за їхнім використанням.

Однією з ключових переваг Power BI є можливість спільного доступу до звітів між користувачами, що дозволяє команді ефективно співпрацювати та обговорювати аналітичні дані. Звітами можна ділитися з окремими користувачами або групами всередині організації.

Також Power BI дозволяє вбудовувати звіти у внутрішні портали або інші веб-додатки, використовуючи API. Це забезпечує інтеграцію аналітичних даних у різні системи компанії, роблячи їх більш доступними для користувачів.

Power BI підтримує функції автоматичного оновлення даних та сповіщень, що дозволяє забезпечити актуальність аналітичних звітів і панелей. Можна налаштувати регулярні оновлення даних, що дозволяє автоматично завантажувати нові дані в звіти та панелі. Це особливо важливо для динамічних бізнес-середовищ, де дані постійно змінюються. Також можна налаштовувати сповіщення для ключових показників, щоб користувачі отримували повідомлення про зміни даних. Це допомагає оперативно реагувати на важливі події та зміни у показниках.

Power BI інтегрується з іншими сервісами Microsoft, такими як Teams та SharePoint, що забезпечує ще більшу гнучкість у спільній роботі та обміні даними. Можна ділитися звітами та панелями безпосередньо в Teams, що дозволяє обговорювати аналітичні дані у робочих групах. Це сприяє ефективному співробітництву та прийняттю колективних рішень на основі даних.

Публікація та спільний доступ до звітів і панелей у Power BI є важливими аспектами аналітичного процесу, які забезпечують ефективне співробітництво та обмін даними між користувачами. Завдяки можливостям налаштування рівнів

доступу, автоматичного оновлення даних, сповіщень та інтеграції з іншими сервісами Microsoft, Power BI дозволяє організаціям ефективно використовувати аналітичні дані для прийняття обґрунтованих рішень.

Щоб підсумувати все описане вище про Microsoft Excel, можна провести SWOT-аналіз цього програмного засобу, продемонструвавши його сильні та слабкі сторони, можливості та загрози при його використанні при аналізі даних, використовуючи відповідну діаграму із цими показниками (рис. 2.5).

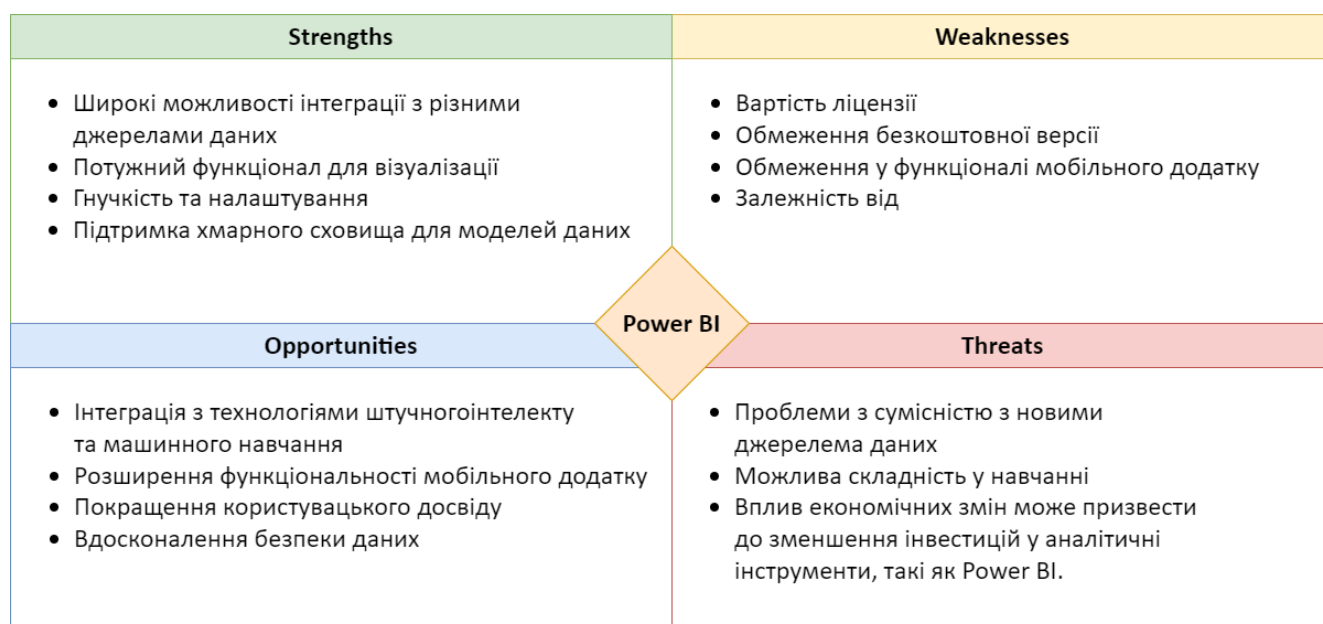


Рисунок 2.5 – SWOT-аналіз програмного засобу Excel

(Джерело розроблене автором)

Висновок до другого розділу

Розділ присвячений теоретичним та практичним аспектам моделювання даних та бізнес-аналітики з використанням сучасних інструментів, таких як Excel, Power BI та сховища даних (Data Warehouse). Також розглядаються можливості кожного інструментарію, їх застосування та важливість у процесах аналізу даних.

Сховища даних відіграють важливу роль у зберіганні та управлінні великими обсягами даних. Вступ до сховищ даних охоплює основні поняття та визначення, підкреслюючи їх важливість для бізнес-аналітики. Архітектура

сховищ даних включає різні компоненти та рівні, що забезпечують ефективне зберігання та обробку даних.

Моделювання у сховищах даних включає створення логічних і фізичних моделей, що дозволяють ефективно організовувати та аналізувати дані. Використання сучасних технологій та інструментів для створення сховищ даних забезпечує їх надійність та продуктивність.

Excel є одним із найбільш поширених та універсальних інструментів для роботи з даними. Вступ до Excel охоплює основні функції та можливості програми, включаючи базові операції з даними, створення таблиць та діаграм, а також функції обчислення. Одним із ключових аспектів є можливість імпорту та експорту даних у різних форматах, що забезпечує гнучкість при роботі з різноманітними джерелами інформації.

Power BI є потужним інструментом для створення аналітичних моделей та інтерактивних звітів. Вступ до Power BI охоплює основні концепції та можливості платформи, підкреслюючи її значення для сучасної бізнес-аналітики. Однією з головних переваг Power BI є можливість підключення до різних джерел даних, включаючи бази даних, онлайн-сервіси, Excel та текстові файли.

РОЗДІЛ 3

ПРИКЛАДНА ОБРОБКА ДАНИХ І СТВОРЕННЯ МОДЕЛІ ДАНИХ З ВИКОРИСТАННЯ POWER BI ТА EXCEL

Життєвий цикл даних, як було представлено в Розділі 1, починається зі збору даних. Але дуже часто у виробничій сфері дані збираються не самим підприємством, а зовнішньою компанією, яка спеціалізується на маркетингових дослідженнях, аналітичних звітах та консалтингових послугах. Тому етап збору в нашому випадку полягає в купівлі таких даних підприємством. Серед особливостей, які можуть надавати такі компанії, є дані про продажі товарів певних категорій від конкурентів, а також на закордонних ринках.

Проходячи переддипломну практику на Дочірньому підприємстві «Кондитерській корпорації «Рошен», та маючи при цьому певний досвід роботи Молодшим фахівцем з аналізу даних, можу з упевненістю сказати, що такий варіант несе за собою дуже багато нюансів при обробці даних та створенні моделей даних для подальшого аналізу результатів таких звітів. З кожним наступним етапом розробки моделі даних виникають проблемні питання, які можуть сповільнювати та гальмувати процес аналітики загалом. Тому щоб цього уникнути варто розглянути на конкретному прикладі нюанси аналізу та моделювання даних. Звичайно всі дані, які я наводжу, є випадковими та не відповідають дійсності, а слугують лише для демонстрації процесу обробки та моделювання. Реальні датасети та інформація, що міститься в них є абсолютно конфіденційними.

На системній основі нам надходять дані про продажі кондитерських виробів у Болгарії. Дані зібрані та підготовлені аналітичною компанією Нільсен. Потрібно обробити та проаналізувати дані з усіх датасетів аналітичних звітів, а також додати дані про власні продажі товарів Рошен на ринках Болгарії. Для цього потрібно побудувати модель, яка буде містити усі фактичні дані, та надасть можливість гнучкої класифікації товарів для зручного формування аналітичної звітності продажів, та додати функціонал фільтрування даних за часовою ознакою.

3.1 Передобробка та зведення даних за допомогою Excel

Перевірка датасетів на спорідненість

Перш ніж розпочати підготовку даних, важливо чітко зрозуміти мету проекту та завдання, які необхідно вирішити. Це допомагає визначити, які саме дані нам потрібні та як їх обробити. Самі ж дані можуть бути отримані з різних джерел, але в нашому випадку це будуть дані листів у Microsoft Excel. На цьому етапі важливо відслідковувати джерело кожного запису та збирати додаткову інформацію про контекст і якість даних. Перше на що варто звернути увагу це є спорідненість джерел даних, себто, якщо ми використовуємо датасети із файловим розширення Excel, то всі наступні датасети (таблиці фактів) повинні бути такого ж типу.

Окрім цього критично важливим аспектом є спільна структура датасетів. Під час порівняння шапок (рядків заголовків стовпців) датасетів важливо звернути увагу на такі аспекти:

- структура даних – перевіряємо, чи всі необхідні поля присутні у кожному датасеті. Чи мають вони однаковий формат та тип даних?
- якість даних – оцінюємо якість даних у кожному датасеті. Чи є пропущені значення, дублікати, аномальні або неочищені дані?
- кількість записів – переконайтеся, що кожен датасет містить достатню кількість записів для адекватного аналізу та моделювання.
- репрезентативність вибірки – перевірте, наскільки представницьким є кожен датасет відносно загальної популяції даних. Чи відображає він різноманітність даних?

Порівнюємо ці аспекти у кожному датасеті, а також звертаємо увагу на будь-які інші відмінності, які можуть вплинути на їхню придатність для конкретного проекту чи дослідження.

Для виконання всього цього зазвичай створюють перевірковий файл шапок датасетів (рис. 3.1). Формується список усіх потрібних файлів, копіюються всі

шапки із усіх датасетів, а потім виділяємо всі поля, які є важливими для нас. Для покращення читабельності такого набору рекомендую кожне спільне поле для певних датасетів позначати своїм кольором. Наприклад поле MARKET присутнє в кожному датасеті, позначаємо його блідо-зеленим кольором, та переходимо до наступного.

Коли всі поля позначені, можна легко виокремити необхідне. Враховуючи те, що нам потрібно зберегти якомога більше даних, то й відбирати поля будемо за логічним принципом диз'юнкції, або як його ще називають операція «або» – себто при порівнянні структури двох шапок беремо всі поля з першої та всіх з другої. Іноді поля мають іншу назву, проте містять дані потрібної нам структури, тому можемо вважати їх однорідними. Наприклад, як можемо спостерігати на рис. 3.1, колонка TYPE OF CANDIES у файлі Hard Candies Bulgaria_Dec'19 містить дані поля PRODUCT TYPE як у інших таблицях, саме тому ці поля позначені одним кольором, та вважаються однорідними.

А	В	С	Д	Е	F	G	H	I	J	K	L	M
1	КОМПРОМІС	MARKET	PERIOD	ITEM	LDESC	CATEGORY	MANUFACTURER	BRAND	SUB-BRAND	PRODUCT TYPE	PRODUCT CLASS	CHOCOLATE TYPE
2	ПОЛЕ											
3												
4	СТАРИ ШАПКИ	ПОЛЕ 1	ПОЛЕ 2	ПОЛЕ 3	ПОЛЕ 4	ПОЛЕ 5	ПОЛЕ 6	ПОЛЕ 7	ПОЛЕ 8	ПОЛЕ 9	ПОЛЕ 10	ПОЛЕ 11
5	Chocolate Assortments Bulgaria_Jan'22	MARKET	PERIOD	year	month	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PACK TYPE	PACKAGE SIZE
6	Chocolate Countlines Bulgaria_Dec'22	SDESC	SDESC	SDESC	LDESC	_CATEGORY VIEW	PRIVATE LABEL	MANUFACTURER	BRAND	MULTIPACK	PACK SIZE	COUNTRY OF ORIG
7	Chocolate Countlines Bulgaria_Jan'23	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	MULTIPACK	PACK SIZE	COUNTRY OF ORIGIN	IF LIGHT/DIET PRO
8	Chocolate Countlines Bulgaria_Jul'24	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	MULTIPACK	PACK SIZE	COUNTRY OF ORIGIN	IF LIGHT/DIET PRO
9	Chocolate Novelties Bulgaria 2022	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PRODUCT CLASS	MULTIPACK	PACK SIZE	COUNTRY OF ORIG
10	Chocolate Tablets Bulgaria_Dec'22	SDESC	SDESC	SDESC	_CATEGORY VIEW	MANUFACTURER	BRAND	SUB-BRAND	PRODUCT TYPE	VARIETIES	ADDED INGR/FILL/TOPPIN-CONFECT	PACKAGE SIZE
11	Chocolate Tablets Bulgaria_Jan'22	MARKET	MECRIL	ГОД	PERIOD	ITEM	CATEGORY	MANUFACTURER	BRAND	SUB-BRAND	PRODUCT TYPE	VARIETIES
12	Chocolate Tablets Bulgaria_Jan'24	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	SUB-BRAND	PRODUCT TYPE	VARIETIES	ADDED INGR/FILL/TOPPIN-CONFECT
13	Chocolate Tablets Bulgaria_Jul'23	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	SUB-BRAND	PRODUCT TYPE	VARIETIES	ADDED INGR/FILL/TOPPIN-CONFECT
14	Chocolate Wafers Bulgaria Jan'24	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	SUB-BRAND	PRODUCT TYPE	VARIETIES	ADDED INGR/FILL/TOPPIN-CONFECT
15	Chocolate Wafers Bulgaria Jul'23	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	SUB-BRAND	PRODUCT TYPE	VARIETIES	ADDED INGR/FILL/TOPPIN-CONFECT
16	Copy of BG_Total BG with Discounters extrac	MARKET	PRALINES	LEVEL	MANUFACTURER	BRAND	PACK TYPE	PACKAGE SIZE	CHOCOLATE TYPE	ADDED INGREDIENT	PERIODS	Value (in 1 000 BG
17	Copy of BG_Total BG with Discounters extrac	MARKET	TABLETS	LEVEL	MANUFACTURER	BRAND	SUBBRAND	PACKAGE TYPE	PACKAGE SIZE	CHOCOLATE TYPE	ADDED INGR/FILL/TOPPIN-CONFECT	IF SEASONAL
18	Hard Candies Bulgaria_Dec'19	MARKET	PERIOD	SEQ	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	TYPE OF CANDIES	PACKAGE TYPE	FLAVOUR
19	Hard Candies Bulgaria_Feb'24	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PRODUCT CLASS	LEVEL OF SUGAR	PACKAGE TYPE	MULTIPACK
20	Non-Coated Wafers Bulgaria Jan'24	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PACKAGE TYPE	PACKAGE SIZE	MULTIPACK	FLAVOUR
21	Non-Coated Wafers Bulgaria Jul'23	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PACKAGE TYPE	PACKAGE SIZE	MULTIPACK	FLAVOUR
22	Soft Candies Bulgaria_Dec'19	MARKET	PERIOD	SEQ	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	TYPE OF DESERT	FLAVOUR	PACKAGE TYPE
23	Soft Candies Bulgaria_Feb'24	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	TYPE OF DESERT	FLAVOUR	PACKAGE TYPE	DESCRPTIVE SIZE
24	Sweet Biscuits Bulgaria_Dec'22	SDESC	SDESC	SDESC	_CATEGORY VIEW	PRIVATE LABEL	MANUFACTURER	BRAND	IF COATED	TYPE OF DESERT	PACK TYPE	PACKAGE SIZE
25	Sweet Biscuits Bulgaria_Jan'23	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	IF COATED	TYPE OF DESERT	PACK TYPE	PACKAGE SIZE
26	Sweet Biscuits Bulgaria_Jul'23	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	IF COATED	TYPE OF DESERT	PACK TYPE	PACKAGE SIZE
27	Choco Wafers Bulgaria_Dec'20	MARKET	PERIOD	Year	Month	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	LEVEL OF SUGAR	MULTIPACK
28	Chocolate Countlines Bulgaria_Sep'22	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	MULTIPACK	PACK SIZE	COUNTRY OF ORIGIN	IF LIGHT/DIET PRO
29	Non-Coated Wafers Bulgaria Jan'21	MARKET	PERIOD	Year	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PACKAGE TYPE	PACKAGE SIZE	MULTIPACK
30	Sweet Biscuits Bulgaria_Jan'21	MARKET	PERIOD	Year	ITEM	Топ виробник	Вид печива	Тип печива	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND
31	Sweet Biscuits Bulgaria_Sep'22	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	IF COATED	TYPE OF DESERT	PACK TYPE	PACKAGE SIZE
32	Sweet Biscuits Bulgaria_Sep'22	MARKET	PERIOD	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	MULTIPACK	PACKAGE TYPE	SIZE - ACTUAL	IF ORGANIC

Рисунок 3.1 – Файл із шапками усіх датасетів

(Джерело розроблене автором)

Список усіх спільних полів (компроміс для зведення датасетів):

- MARKET
 - PERIOD
 - ITEM
 - LDESC
 - CATEGORY
 - MANUFACTURER
 - BRAND
 - SUB-BRAND
 - PRODUCT TYPE
 - PRODUCT CLASS
 - CHOCOLATE TYPE
 - TYPE OF DESERT
 - PACKAGE TYPE
 - PACKAGE SIZE
- MULTIPACK
 - PRIVATE LABEL
 - VARIETIES
 - COUNTRY OF ORIGIN
 - IF LIGHT/DIET PRODUCT
 - IF ORGANIC
 - IF SEASONAL
 - ADDED INGREDIENT
 - TYPE OF PROMOTION
 - FLAVOUR
 - LEVEL OF SUGAR
 - PACKAGING MATERIAL
 - IF COATED
 - TYPE OF COATING
- SHAPE
 - DESCRIPTIVE SIZE
 - FLAVOUR-CONFECTIONARY/DESRTS
 - TYPE - CUT OF PRODUCT
 - WEIGHT RANGE
 - VOLUME SALES
 - VALUE SALES
 - ITEMS SALES
 - NUMERIC SELLING DISTRIBUTIUTION (MAX)
 - WEIGHTED SELLING DISTRIBUTIUTION (MAX)

Зведення датасетів

Після того, як ми знайшли так званий спільний знаменник полів, можемо вважати їх компромісом для зведення даних. Тепер потрібно кожен файл датасету розширити та підігнати під структуру цього компромісного рішення. Для цього потрібно кожен датасет редагувати шляхом переміщення колонок з усім їх вмістом під їх відповідники зі спільної шапки (рис. 3.2).

AC82	FILE	ОСНОВЕ	ВСТАВЛЕННЯ	МАЛЮВАННЯ	МАКЕТ СТОРІНКИ	ФОРМУЛИ	ДАНИ	РЕЦЕНЗУВАННЯ	ПОДАННЯ	АВТОМАТИЗАЦІЯ	РОЗРОБНИК	ДОВІДКА	POWER PIVOT	СІЛКІНГ РОЗМІР	ПРИМІТКИ	СІЛКІНГ ДОСТУП
1	Спільна (компромісна) шапка	PERIOD	ITEM	LDESC	CATEGORY	MANUFACTURER	BRAND	SUB-BRAND	PRODUCT TYPE	PRODUCT CLASS	CHOCOLATE TYPE	TYPE OF DESERT	PACKAGE TYPE	PACKAGE SIZE	MULTI	
2	Шапка датасету (розфарбована)	year	month	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PACK TYPE	PACKAGE SIZE	LEVEL OF SUGAR	COUNTRY OF ORIGIN	CHOCOLATE TYPE	ADDEC		
3	MARKET	year	month	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PACK TYPE	PACKAGE SIZE	LEVEL OF SUGAR	COUNTRY OF ORIGIN	CHOCOLATE TYPE	ADDEC		
4	MARKET	year	month	ITEM	CATEGORY	PRIVATE LABEL	MANUFACTURER	BRAND	PACK TYPE	PACKAGE SIZE	LEVEL OF SUGAR	COUNTRY OF ORIGIN	CHOCOLATE TYPE	ADDEC		
5	TOTAL BULGARIA	2019	1	MASHA CHOCOLATE F 200G BOX BOOK	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	5STAR GASTRO	MASHA (5STAR GASTRO)	BOX	200g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	WITH (		
6	TOTAL BULGARIA	2019	1	MASHA CHOCOLATE F 85G BOX BOOK	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	5STAR GASTRO	MASHA (5STAR GASTRO)	BOX	Up to 100 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	WITH (		
7	TOTAL BULGARIA	2019	1	PERGALE COFFEE M 225G BAG	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BAGS		201 g - 249 g	WITH SUGAR / REGULAR	IMPORTED BRAND	DARK CHOCOLATE	MONO		
8	TOTAL BULGARIA	2019	1	PERGALE SAUDONOV AGDONA M 225G BAG	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BAGS		201 g - 249 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	MONO		
9	TOTAL BULGARIA	2019	1	PERGALE CHERRYBERRYCOLLECTION A 130G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
10	TOTAL BULGARIA	2019	1	PERGALE DARK CLASSIC A 114G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	DARK CHOCOLATE	ASSOR		
11	TOTAL BULGARIA	2019	1	PERGALE DARKCHOCOLATECOLLECTION A 125GBX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
12	TOTAL BULGARIA	2019	1	PERGALE DESERTCOLLECTION MILK A 125G BX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
13	TOTAL BULGARIA	2019	1	PERGALE GOURMETDESSERTCOLLECTION A 125G	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
14	TOTAL BULGARIA	2019	1	PERGALE MILK BERRIES A 117G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
15	TOTAL BULGARIA	2019	1	PERGALE MILK CHOCCOLLECTION A 125G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
16	TOTAL BULGARIA	2019	1	PERGALE MILK CLASSIC A 114G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
17	TOTAL BULGARIA	2019	1	PERGALE MILK DESSERTS A 113G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
18	TOTAL BULGARIA	2019	1	PERGALE MILK HAZELNUT F 119G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	WITH (		
19	TOTAL BULGARIA	2019	1	PERGALE MILK MANGO A 114G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
20	TOTAL BULGARIA	2019	1	PERGALE SPRING MILK COLLECT A 125G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
21	TOTAL BULGARIA	2019	1	PERGALE WINTER MASQUERADE DARK A 125G BX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	DARK CHOCOLATE	ASSOR		
22	TOTAL BULGARIA	2019	1	PERGALE WINTER MASQUERADE MILK A 125G BX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
23	TOTAL BULGARIA	2019	1	PERGALE WINTER MILKCOLLECTION A 125G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
24	TOTAL BULGARIA	2019	1	PERGALE WINTERCOLLECTION A MILK 125G BOX	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
25	TOTAL BULGARIA	2019	1	PERGALE WINTERISCOMINGTOWTOWN DARK A 125G	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	DARK CHOCOLATE	ASSOR		
26	TOTAL BULGARIA	2019	1	PERGALE WINTERISCOMINGTOWTOWN MILK A 125G	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
27	TOTAL BULGARIA	2019	1	PERGALE WINTERISCOMINGTOWTOWN MILK A 125G	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
28	TOTAL BULGARIA	2019	1	PERGALE WINTERISCOMINGTOWTOWN MILK A 125G	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		
29	TOTAL BULGARIA	2019	1	PERGALE WINTERISCOMINGTOWTOWN MILK A 125G	CHOCOLATE ASSORTMENTS	NO (NOT PRIVATE LABEL)	AB VINIAUS PERGALE	PERGALE (AB VINIAUS PERGALE BOX		101 g - 149 g	WITH SUGAR / REGULAR	IMPORTED BRAND	MILK CHOCOLATE	ASSOR		

Рисунок 3.2 – Процес зведення датасету до спільної шапки
(Джерело розроблене автором)

На цьому етапі важливо бути максимальною зосередженим при перенесеннях колонок у файлі, оскільки випадково дані можуть накладись одні на одні, що дуже

сильно пошкодить наш датасет. Ще одним важливим аспектом є те, що назви аркушів у Excel, а вони повинні бути однаковими, також є критично важливими у портуванні даних до моделі, оскільки вони є шляхом для пошуку та завантаження датасетів.

Провівши ці маніпуляції з колонками у кожному файлі можемо сказати, що зведення виконане, а наші датасети мають однакову структуру. Хоч деякі нові поля порожні за рахунок їх розширення, але вони не завжди повинні бути усі заповнені, оскільки датасети містять дані про продажі кондитерських товарів різних категорій, і не кожен товар має спільні типи характеристик з іншими.

3.2 Створення моделі даних з використанням засобів Microsoft Power BI

Завантаження даних до моделі Power BI

Це перший етап створення моделі даних, який полягає в завантаженні інформації з різних джерел до Power BI для подальшого аналізу та використання. Платформа підтримує різноманітні джерела даних, включаючи бази даних, файли Excel, хмарні сервіси, веб-ресурси та інші.

При початковому імпорті даних користувач має можливість вибрати джерело даних та визначити параметри підключення. Power BI надає інтуїтивний інтерфейс для вибору джерел та налаштування параметрів підключення, що робить процес імпорту даних доступним навіть для неспеціалістів.

Хоч наші датасети є книгами Excel, та вони мають однакову структуру. Враховуючи це, доцільніше буде як джерело вибрати Папку (рис. 3.3). До того ж в майбутньому будуть надходити нові дані за наступні періоди на постійній основі. Такий підхід дозволить кардинально спростити подальшу обробку всіх файлів.

Після вибору джерела даних та введення необхідних параметрів, Power BI автоматично виконує процес імпорту, завантажуючи дані в середовище обробки.

При цьому кожне джерело даних є окремим запитом. Якщо опція імпорту це папка, то її файли об'єднуються в одне джерело і стають одним запитом.

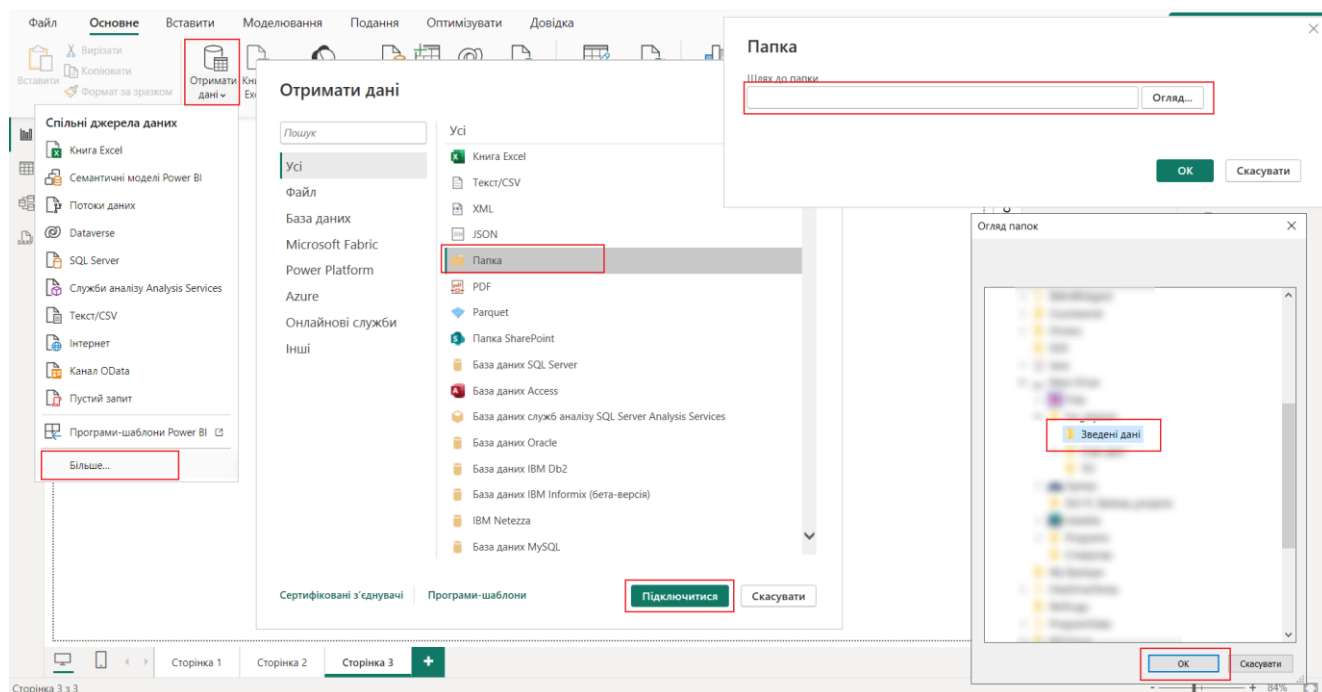


Рисунок 3.3 – Сценарій імпорту датасетів із підключенням до папки

(Джерело розроблено автором)

Перетворення даних в Power Query

Однією з ключових можливостей Power BI є Power Query Editor, який дозволяє користувачам виконувати різноманітні операції з даними шляхом створення кроків трансформації. Цей інструмент надає зручний та інтуїтивно зрозумілий інтерфейс для редагування та очищення даних.

Під час перетворення даних користувачі можуть використовувати ряд вбудованих функцій, а також писати власні скрипти мовою M для більш специфічних операцій з обробки інформації.

Прямо під час перетворень користувач має можливість переглянути та перевірити зміни, внесені до даних. Більше того, кожна операція перетворення дорівнює одному кроку в загальному наборі процесів обробки запиту, що дозволяє бачити як кожна операція поступово впливає на дані із завантажених датасетів. Провівши аналогії з програмуванням, можна це назвати процесом відлагодження або ж дебагу (Debug). Цей етап дозволяє забезпечити якість даних та підготувати їх для подальшого аналізу та використання у моделі даних Power BI.

Також цікавим та корисним функціоналом Power Query є керування всіма кроками трансформації одночасно за допомогою Розширеного редактора, який надає можливість бачити усі кроки у вигляді коду мовою M, а також редагувати їх. На своєму досвіді я переконався у зручності цього засобу, коли потрібно було скопіювати всі кроки трансформації із одного запиту до іншого. Ще таким чином можна власноруч робити резервні копії запитів, замінивши зламаний код бекапом.

Для правильної обробки датасетів, нам потрібно виконати кілька кроків:

- перейменування стовпців
- видалення зайвих стовпців (за потреби)
- відфільтрування рядків
- змінення типу даних в колонках
- інші важливі технічні операції

В Розширеному редакторі ці перетворення виконуються таким кодом:

```
let
    Джерело      =      Folder.Files("C:\Users\roman.sytnyk\Downloads\Інше\For_diplom\Зведені
дані"),
    #"Викликати спеціальну функцію1" = Table.AddColumn(Джерело, "Перетворити файл (3)", each
#"Перетворити файл (3)"([Content])),
    #"Перейменовані стовпці1" = Table.RenameColumns(#"Викликати спеціальну функцію1",
{"Name", "Джерело.Ім'я"}),
    #"Видалені інші стовпці1" = Table.SelectColumns(#"Перейменовані стовпці1",
{"Джерело.Ім'я", "Перетворити файл (3)"}),
    #"Розгорнутий стовпець таблиці1" = Table.ExpandTableColumn(#"Видалені інші стовпці1",
"Перетворити файл (3)", Table.ColumnNames(#"Перетворити файл (3)"("#Файл-зразок (3)")),
    #"Відфільтровані рядки" = Table.SelectRows(#"Розгорнутий стовпець таблиці1", each
[PERIOD] <> "MAT"),
    #"Змінений тип" = Table.TransformColumnTypes(#"Відфільтровані рядки",{{"Джерело.Ім'я",
type text}, {"MARKET", type text}, {"PERIOD", type date}, {"ITEM", type text}, {"LDESC",
type text}, {"CATEGORY", type text}, {"MANUFACTURER", type text}, {"BRAND", type text},
{"SUB-BRAND", type text}, {"PRODUCT TYPE", type text}, {"PRODUCT CLASS", type text},
{"CHOCOLATE TYPE", type text}, {"TYPE OF DESERT", type text}, {"PACKAGE TYPE", type text},
{"PACKAGE SIZE", type text}, {"MULTIPACK", type text}, {"PRIVATE LABEL", type text},
{"VARIETIES", type text}, {"COUNTRY OF ORIGIN", type text}, {"IF LIGHT/DIET PRODUCT", type
text}, {"IF ORGANIC", type text}, {"IF SEASONAL", type text}, {"ADDED INGREDIENT", type
text}, {"TYPE OF PROMOTION", type text}, {"FLAVOUR", type text}, {"LEVEL OF SUGAR", type
```

```

text}, {"PACKAGING MATERIAL", type text}, {"IF COATED", type text}, {"TYPE OF COATING",
type text}, {"SHAPE", type text}, {"DESCRIPTIVE SIZE", type text}, {"FLAVOUR-
CONFECTIONARY/DESRTS", type text}, {"TYPE - CUT OF PRODUCT", type text}, {"VOLUME SALES",
type number}, {"VALUE SALES", type number}, {"ITEMS SALES", type number}, {"NUMERIC SELLING
DISTRIBIUTION (MAX)", type number}, {"WEIGHTED SELLING DISTRIBIUTION (MAX)", type number},
{"WEIGHT RANGE", type text}}),
  #"Перейменовані стовпці" = Table.RenameColumns("#Змінений тип",{{"Джерело.Ім'я",
"ДЖЕРЕЛО.ФАЙЛ"}}),
  #"Додане настроюване значення" = Table.AddColumn("#Перейменовані стовпці", "Mark", each
"Nielsen"),
  #"Перевпорядковані стовпці" = Table.ReorderColumns("#Додане настроюване
значення",{ "ДЖЕРЕЛО.ФАЙЛ", "Mark", "MARKET", "PERIOD", "ITEM", "LDESC", "CATEGORY",
"MANUFACTURER", "BRAND", "SUB-BRAND", "PRODUCT TYPE", "PRODUCT CLASS", "CHOCOLATE TYPE",
"TYPE OF DESERT", "PACKAGE TYPE", "PACKAGE SIZE", "MULTIPACK", "PRIVATE LABEL", "VARIETIES",
"COUNTRY OF ORIGIN", "IF LIGHT/DIET PRODUCT", "IF ORGANIC", "IF SEASONAL", "ADDED
INGREDIENT", "TYPE OF PROMOTION", "FLAVOUR", "LEVEL OF SUGAR", "PACKAGING MATERIAL", "IF
COATED", "TYPE OF COATING", "SHAPE", "DESCRIPTIVE SIZE", "FLAVOUR-CONFECTIONARY/DESRTS",
"TYPE - CUT OF PRODUCT", "VOLUME SALES", "VALUE SALES", "ITEMS SALES", "NUMERIC SELLING
DISTRIBIUTION (MAX)", "WEIGHTED SELLING DISTRIBIUTION (MAX)"}),
  #"Відфільтровані рядки1" = Table.SelectRows("#Перевпорядковані стовпці", each
([ДЖЕРЕЛО.ФАЙЛ] <> "!_shapka.xlsx"))
in
  #"Відфільтровані рядки1"

```

Після виконання всіх перетворень із датасетами Нільсена, отримали запит RAW Nielsen Datasets.

Наступним кроком потрібно буде опрацювати дані про власні продажі товарів Рошен на ринках Болгарії. Отримуємо потрібні файли (ТО_Болгарія і подібні) із корпоративних сховищ даних. Як можемо бачити на рис. 3.4, їх структура відрізняється від файлів Нільсену. Тому виконаємо зведення подібним чином, але вже починаємо обробляти їх відразу в Power Query. Імпорт даних виконуємо також з підключенням до папки із файлами, як на рис. 3.3.

	1	2	3	4	5	6	7	8	9	10
	Ритейлер	Початок місяця	Ном позиція			Вес	Сумма документа с НДС, нац	Кво		
2	300	01.11.2023	Assortment classic			21,437	458,89	139,200		
3	300	01.11.2023	Assortment compliment			19,256	434,36	132,800		
4	300	01.11.2023	Assortment elegant			27,376	613,58	188,800		
5	300	01.11.2023	Bonny-Fruit berry mix			1,600	12,13	8,000		
6	300	01.11.2023	Bonny-Fruit summer mix			0,320	2,43	1,600		
7	300	01.11.2023	Chocolateria 194+			14,589	504,07	75,200		
8	300	01.11.2023	Chocolateria 256+			18,022	614,62	70,400		
9	300	01.11.2023	Coffeelike			18,400	122,36	18,400		
10	300	01.11.2023	Ko-ko choco white			1,600	21,21	1,600		
11	300	01.11.2023	Konafetto с молочной начинкой			43,344	514,94	309,600		
12	300	01.11.2023	Konafetto с начинкой крем-какао			52,416	627,58	374,400		
13	300	01.11.2023	Krock			3,200	30,37	3,200		
14	300	01.11.2023	Lacmi crisper пористый с молочной начинкой с крипи			11,424	245,08	134,400		
15	300	01.11.2023	Lacmi nutter-butter пористый с шоколадно-ореховой начинкой			-1,292	-21,35	-15,200		
16	300	01.11.2023	Lacmi молочный			11,520	233,51	128,000		
17	300	01.11.2023	Lacmi с арахисом и карамельно-арахисовой начинкой			0	0	0		
18	300	01.11.2023	Lacmi с молочной и шоколадной начинкой, вафлей с какао 300			11,448	207,84	43,200		
19	300	01.11.2023	Lacmi с молочной начинкой и вафлей			7,992	161,92	88,800		
20	300	01.11.2023	Lacmi с молочной начинкой и печеньем			17,280	315,27	172,800		
21	300	01.11.2023	Lacmi с молочной начинкой, карамелью и печеньем с какао			4,800	87,54	48,000		
22	300	01.11.2023	Lacmi с шоколадно-ореховой начинкой и печеньем 300			6,032	100,06	20,800		

Рисунок 3.4 – Приклад структури датову ТО_Болгарія

(Джерело розроблено автором)

Зведення датасетів робимо шляхом перейменування колонок, та додавання нових відповідно до компромісного рішення, яке використовували для зведення попередніх датасетів у пункті 3.1. Проаналізувавши їх, бачимо відповідність деяким полям датасетів Нільсен:

Ритейлер = MARKET

Вес = VOLUME SALES

Початок місяця = PERIOD

Сумма документа с НДС, нац = VALUE SALES

Ном позиція = ITEM

Кво = ITEMS SALES

В результаті в Power Query отримуємо запит ТО_Болгарія.

Далі необхідно об'єднати два запити із різними датасетами в один. Для цього ми якраз виконали зведення всіх датасетів Нільсен та ТО_Болгарія, виділивши спільну шапку таблиць, та виконавши всі перетворення в Power Query. В кожному запиті додатково створюємо новий стовпець поле-маркер для визначення джерел звіту. Для даних Нільсен він буде містити текст "Nielsen", а для ТО_Болгарія – текст "ТО".

Об'єднувати запити будемо шляхом їх додавання (а якщо конкретніше, то доповнення) в новий. На рис. 3.5 продемонстровано цей процес.

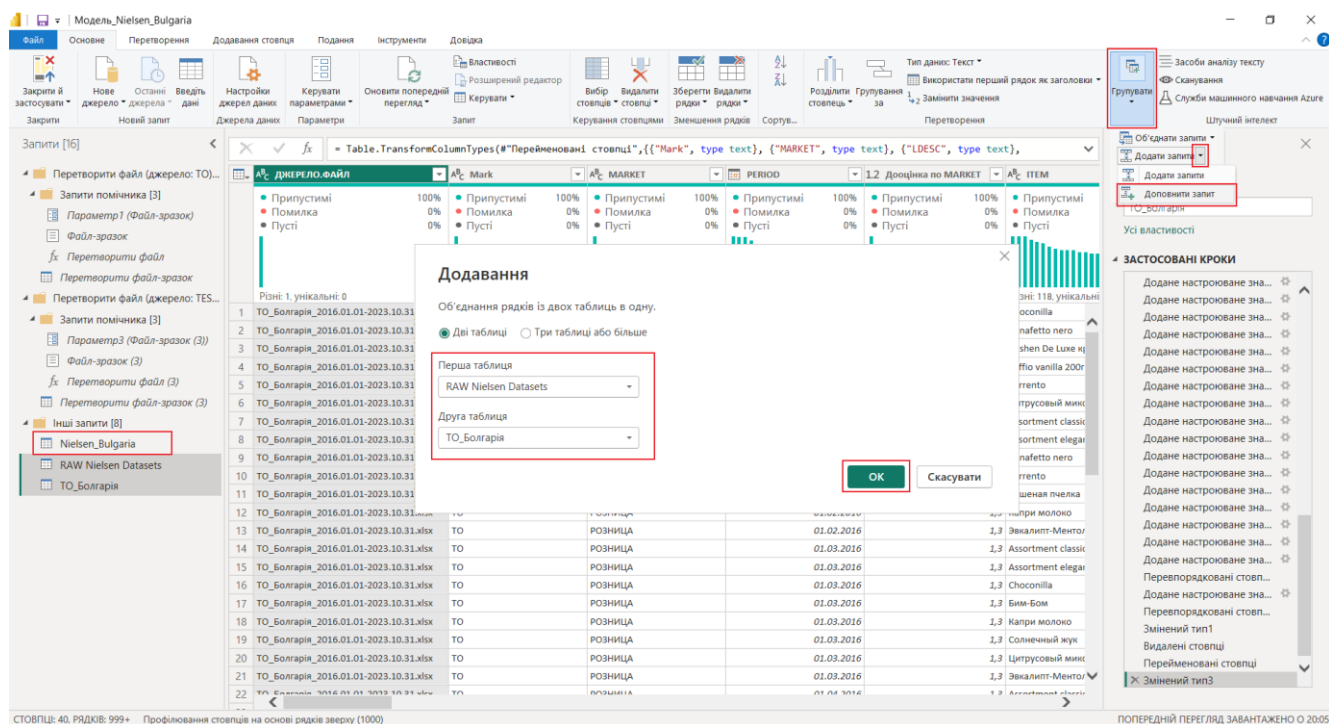


Рисунок 3.5 – Процес додавання запитів у новий
(Джерело розроблено автором)

Після об'єднання, отримали єдиний запит Nielsen_Bulgaria, який містить дані про продажі з обох джерел.

Коли всі необхідні операції з перетворенням даних виконані, то можемо переходити до налаштувань моделі даних. Щоб це зробити та зберегти всі дії, у вікні Power Query потрібно натиснути кнопку «Закрити та застосувати». Після цього оброблені дані з датасетів у форматі таблиць завантажуються у саму модель даних Power BI, та можемо продовжувати працювати з ними.

Проаналізувавши всі поля датасетів, бачимо кілька структур даних, а саме: класифікаційні поля із характеристиками товарів, точки реалізації продажів – ритейлери (поле MARKET), поле з датами PERIOD та числові показники продажів.

На своєму робочому досвіді вже переконався у зручності та геніальності мови DAX. Вона дозволяє Power BI не лише завантажувати готові дані із файлів, але й також створювати та генерувати свої власні таблицьки, використовуючи для цього формульний інструментарій.

Однією з таких згенерованих таблиць є таблиця дат. Вона відіграє роль класифікатора, та містить усі дати із точністю до дня (хоча можна зробити навіть із

ще більшою деталізацією), перераховуючи кожну в певних календарних проміжках, які ми можемо задати.

Використання таблиці дат забезпечує узгодженість у всіх часових обчисленнях і фільтраціях. Наявність заздалегідь визначених стовпців (наприклад, рік, квартал, місяць) полегшує створення виразів DAX для часових обчислень. Таблиця дат дозволяє легко фільтрувати дані за різними часовими періодами.

У нашій задачі потрібна таблиця дат за період від 01.01.2016 року по 31.12.2025 рік. Тобто ця табличка буде містити кожен день із цього 10-річного відтинку (3 653 дні). Ось код мовою DAX для її генерування:

```
Date = ADDCOLUMNS (
    CALENDAR (DATE(2016,1,1), DATE(2025,12,31)),
    "DateAsInteger", FORMAT ( [Date], "YYYYMMDD" ),
    "Year", YEAR ( [Date] ),
    "Monthnumber", FORMAT ( [Date], "MM" ),
    "YearMonthnumber", FORMAT ( [Date], "MM/YYYY" ),
    "YearMonthShort", FORMAT ( [Date], "YYYY/mmm" ),
    "MonthNameShort", FORMAT ( [Date], "mmm" ),
    "MonthNameLong", FORMAT ( [Date], "mmmm" ),
    "Quarter", "Q" & FORMAT ( [Date], "Q" ), "Quarter Year", "Q" & FORMAT ( [Date], "Q.YYYY" ),
    "YearQuarter", FORMAT ( [Date], "YYYY" ) & "/" & FORMAT ( [Date], "Q" ) )
```

Для коректного та зручного формування аналітичних звітів, потрібно керуватись принципом класифікації та розподілення даних на певні групи. Виконується це шляхом створення класифікаційних (вимірювальних) таблиць, або ж словників, та додаванням їх до моделі даних (рис. 2.1). Щоб уможливити їх використання, потрібно мати якийсь унікальний ідентифікатор для кожного товару. Зазвичай кожен з них має якийсь код (штрих-код, EAN тощо). Але в наших датасетах такого немає, а використання назви товару як ідентифікатора, є недоречним, оскільки тоді вся класифікація буде некоректною та не можна буде відслідкувати продажі за різними вимірами.

Серед усіх товарів у таблицях можна помітити назву PRIVATE LABEL. Якщо коротко, то це продукти, які виробляються однією компанією, але продаються під брендом іншої компанії. Зазвичай це великі роздрібні торговці, які

продають такі продукти під власною маркою. Наші ж датасети містять їх для того, щоб приховувати назви товарів, і показувати продажі лише за класифікаційними ознаками, що трішки ускладнює процес аналізу даних. Саме тому нам потрібно зібрати лише ті класифікаційні поля, які міститимуть точні ознаки товарів. А саме: ITEM, MANUFACTURER, BRAND, CATEGORY, PRIVATE LABEL, PRODUCT TYPE, PRODUCT CLASS, TYPE OF DESERT, PACKAGE TYPE, IF SEASONAL, FLAVOUR, IF COATED, TYPE OF COATING, SHAPE.

Враховуючи всі ці особливості, потрібно правильно ідентифікувати кожен товар. Щоб зробити це до таблиці Nielsen_Bulgaria додаємо поле ALL_ITEM_ID, яке буде містити ідентифікатор, що виконує конкатенацію класифікаційних полів.

Код мовою DAX буде наступним:

```
ALL_ITEM_ID = LOWER(IF([MANUFACTURER]="PRIVATE LABEL", [ITEM] &"_"& [MANUFACTURER] &"_"& [BRAND] &"_"& [CATEGORY] &"_"& [PRIVATE LABEL] &"_"& [PRODUCT TYPE] &"_"& [PRODUCT CLASS] &"_"& [TYPE OF DESERT] &"_"& [PACKAGE TYPE] &"_"& [IF SEASONAL] &"_"& [FLAVOUR] &"_"& [IF COATED] &"_"& [TYPE OF COATING] &"_"& [SHAPE], [ITEM]))
```

Далі створюємо таблицю, яка буде містити урізану версію всіх наших датасетів, а саме лише ці класифікаційні поля та створений нами ідентифікатор товару. Зробити це можна за допомогою мови DAX, використовуючи команду SUMMARIZECOLUMNS. Виконуючи її, Power BI згенерує таблицю із цими стовпцями. Повний код команди буде наступним:

```
Новинки СКЮ = SUMMARIZECOLUMNS(
    'Nielsen_Bulgaria'[ITEM],
    'Nielsen_Bulgaria'[MANUFACTURER],
    'Nielsen_Bulgaria'[BRAND],
    'Nielsen_Bulgaria'[CATEGORY],
    'Nielsen_Bulgaria'[PRIVATE LABEL],
    'Nielsen_Bulgaria'[PRODUCT TYPE],
    'Nielsen_Bulgaria'[PRODUCT CLASS],
    'Nielsen_Bulgaria'[TYPE OF DESERT],
    'Nielsen_Bulgaria'[PACKAGE TYPE],
    'Nielsen_Bulgaria'[IF SEASONAL],
    'Nielsen_Bulgaria'[FLAVOUR],
    'Nielsen_Bulgaria'[IF COATED],
    'Nielsen_Bulgaria'[TYPE OF COATING],
```

```
'Nielsen_Bulgaria'[SHAPE],
'Nielsen_Bulgaria'[ALL_ITEM_ID]
)
```

Далі копіюємо цю таблицю в новий файл Excel. Називаємо його Класифікатор_Болгарія, а самому аркушу Excel даємо назву Dictionary. На рис. 3.6 можемо бачити його в Excel. Перенесені нами поля на рисунку зображені синім та червоним, в той час як зеленим – поля додані вручну, що відповідають принципам класифікації товарів у самій компанії. Такий підхід дозволяє власноруч розподіляти товари на власні категорії, яких не було в початкових даних.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Окрім поля MARKET таблиця також містить поле Channel, яке показує тип точок реалізації продажів, та поле Дооцінка полиці, що слугує коефіцієнтом для зрівняння числових показників, нівелюючи похибку закладену у звітах Нільсен.

Також виділяємо таблицю із категоріями товарів, та в які періоди вони зустрічались. Додаємо її на новий аркуш Coef (рис. 3.8) у файлі Класифікатор_Болгарія. Цей класифікатор має таке ж призначення, як і попередній, тому не варто зупинятись на ньому детальніше.

	A	B	C	D	E	F
1	CATEGORY	GROUP	Year	Koef	Exchange rate	
2	CHOCOLATE TABLETS	Tablets	2016	1,96	1,9559	
3	CHOCOLATE TABLETS	Tablets	2017	1,96	1,9559	
4	CHOCOLATE TABLETS	Tablets	2018	1,9	1,9559	
5	CHOCOLATE TABLETS	Tablets	2019	1,9	1,9559	
6	CHOCOLATE TABLETS	Tablets	2020	1,9	1,9559	
7	CHOCOLATE TABLETS	Tablets	2021	1,9	1,9559	
8	CHOCOLATE TABLETS	Tablets	2022	1,9	1,9559	
9	CHOCOLATE TABLETS	Tablets	2023	1,9	1,9559	
10	COUNTLINES	Countlines	2021	1,90	1,9559	
11	COUNTLINES	Countlines	2022	1,90	1,9559	
12	COUNTLINES	Countlines	2023	1,90	1,9559	
13	SWEET BISCUITS	Sweet Biscuits	2016	1,3	1,9559	
14	SWEET BISCUITS	Sweet Biscuits	2017	1,3	1,9559	
15	SWEET BISCUITS	Sweet Biscuits	2018	1,3	1,9559	
16	SWEET BISCUITS	Sweet Biscuits	2019	1,3	1,9559	
17	SWEET BISCUITS	Sweet Biscuits	2020	1,3	1,9559	
18	SWEET BISCUITS	Sweet Biscuits	2021	1,3	1,9559	

Рисунок 3.8 – Вигляд класифікатора категорій для дооцінки

(Джерело розроблено автором)

Оскільки дані про продажів представлені у національній валюті – болгарський лев (BGN), а нам потрібно аналізувати їх у євро, то наступним кроком буде додавання вимірювальної таблиці для коректної конвертації курсів валют. Як можна бачити на рисунку 3.9, вона містить дані про дату курсу, та його коефіцієнт конвертації.

	A	C	D	E	F	G
	Валюта	Дата курса.Початок місяця	Основная валюта	Курс		
6						
7	EUR	01.01.2016	BGN	1,95583000		
8	EUR	01.02.2016	BGN	1,95583000		
9	EUR	01.03.2016	BGN	1,95583000		
10	EUR	01.04.2016	BGN	1,95583000		
11	EUR	01.05.2016	BGN	1,95583000		
12	EUR	01.06.2016	BGN	1,95583000		
13	EUR	01.07.2016	BGN	1,95583000		
14	EUR	01.08.2016	BGN	1,95583000		
15	EUR	01.09.2016	BGN	1,95583000		
16	EUR	01.10.2016	BGN	1,95583000		
17	EUR	01.11.2016	BGN	1,95583000		
18	EUR	01.12.2016	BGN	1,95583000		

Рисунок 3.9 – Вигляд вимірювальної таблиці курсів валют

(Джерело розроблено автором)

Тепер можемо завантажити всі нові класифікатори до моделі даних. Виконуємо портування таблиць із підключенням до файлу Excel Класифікатор

Після налаштування усіх потрібних класифікаторів та зв'язків між таблицями, потрібно правильно обробити числові показники. Серед них є наступні:

- VOLUME SALES — продажі у тонах
- VALUE SALES — продажі у національній валюті (Болгарський лев, BGN)
- ITEMS SALES — продажі у штуках.
- NUMERIC SELLING DISTRIBUTION (MAX) — це показник, який вимірює відсоток торгових точок, що продають певний продукт від загальної кількості можливих торгових точок на ринку. Дає уявлення про те, наскільки широко представлений продукт у мережі роздрібних торговців.
- WEIGHTED SELLING DISTRIBUTION (MAX) — це показник, який враховує не лише кількість торгових точок, де продається продукт, а й їхню відносну важливість або обсяги продажів. Відображає частку загального обсягу продажів усіх товарів, що припадає на торгові точки, де присутній аналізований продукт.

Враховуючи свій досвід роботи та настанови від керівників, виділив для себе наступне правило: Використання всіх числових полів краще перевести на міри. Такий підхід дозволяє зручно та ефективно керувати важливими показниками, а також позбавляє від можливих проблем із невідповідністю типів серед них. Для зручної організації керуванням мір рекомендую створювати спеціальні таблиці, де зберігаються лише вони. Всі міри написані мовою DAX наведено нижче.

Міра для розрахунку продажів у тисячах одиниць:

```
ITEMS SALES (in 1000) = SUMX('Nielsen_Bulgaria', [ITEMS SALES])
```

Міра для розрахунку продажів у тисячах одиниць з урахуванням коефіцієнту дооцінки:

```
ITEMS SALES (in 1000) + coef = SUMX('Nielsen_Bulgaria', [ITEMS SALES]*[Дооцінка Nielsen])
```

Міра для розрахунку об'ємів продажів у тонах:

```
VOLUME SALES (in 1000) = SUMX(Nielsen_Bulgaria, [VOLUME SALES])
```

Міра для розрахунку об'ємів продажів у тонах з урахуванням коефіцієнту дооцінки:

VOLUME SALES (in 1000) + coef = SUMX('Nielsen_Bulgaria',[VOLUME SALES]*[Дооцінка Nielsen])

Міра для грошового розрахунку продажів у BGN:

VALUE SALES (in 1000) = SUMX(Nielsen_Bulgaria, [VALUE SALES])

Міра для грошового розрахунку продажів у BGN з урахуванням коефіцієнту дооцінки:

VALUE SALES (in 1000) + coef = SUMX('Nielsen_Bulgaria',[VALUE SALES]*[Дооцінка Nielsen])

Міра для грошового розрахунку продажів у EURO:

VALUE SALES (in 1000) in EURO = SUMX(Nielsen_Bulgaria,[VALUE SALES])/[Середній_Курс]

Міра для грошового розрахунку продажів у EURO з урахуванням коефіцієнту дооцінки:

VALUE SALES (in 1000) in EURO + coef = SUMX('Nielsen_Bulgaria',[VALUE SALES]*[Дооцінка Nielsen])/[Середній_Курс]

Міра для підрахунку унікальних товарів:

Кі-сть СКЮ = DISTINCTCOUNT('Nielsen_Bulgaria'[ALL_ITEM_ID])

Міра для розрахунку середнього курсу BGN до EURO:

Середній_Курс = AVERAGEX('Курси валют',[Курс])

Міра для розрахунку NUMERIC SELLING DISTRIBUTION (MAX)

NUMERIC SELLING DISTRIBUTION (Max) = MAX('Nielsen_Bulgaria'[NUMERIC SELLING DISTRIBUTION (MAX)])

Міра для розрахунку WEIGHTED SELLING DISTRIBUTION (MAX)

WEIGHTED SELLING DISTRIBUTION (Max) = MAX('Nielsen_Bulgaria'[WEIGHTED SELLING DISTRIBUTION (MAX)])

Серед усіх налаштувань моделі також існує функціонал приховування полів таблиць для забезпечення зручності користування та підвищення ефективності аналізу. Приховуючи непотрібні поля, можна зробити модель більш зрозумілою та зручною для кінцевих користувачів, а деякі поля можуть містити чутливі дані, які не повинні бути доступними для всіх. Так загалом менша кількість полів знижує ймовірність випадкового використання неправильних у звітах і візуалізаціях.

Для нас важливо приховати усі числові поля, адже ми замінюємо їх мірами, також всі ідентифікаційні та класифікаційні поля, і загалом технічні поля, які не несуть користі для аналізу даних та кінцевих користувачів моделі. Приховати поля можна у Табличному поданні або Поданні моделі.

Після всіх налаштувань моделі потрібно опублікувати її у хмарному просторі, щоб мати змогу аналізувати дані не лише в Power BI, а й у сторонніх програмах. Зробити це можна натиснувши кнопку "Опублікувати" на вкладці Основне. Далі модель стане доступно в хмарному сховищі.

3.3 Візуальне представлення даних та аналіз отриманих результатів

Наступним етапом у всьому процесі обробки даних є їх аналіз. Виконувати який, ми будемо за допомогою Excel та наявного в ньому функціоналі.

Зведені таблиці в Excel є потужним інструментом для аналізу та узагальнення великих обсягів даних. Вони дозволяють швидко обробляти, фільтрувати та впорядковувати дані з однієї або декількох таблиць, створюючи зручні для аналізу звіти. Основна функція зведених таблиць полягає в тому, щоб надавати користувачам можливість групувати дані за категоріями, виконувати обчислення, такі як суми, середні значення, кількість записів, мінімальні та максимальні значення, а також відображати результати у вигляді зрозумілих звітів.

Користувач може легко налаштовувати зведену таблицю, перетягуючи поля до різних областей, таких як рядки, стовпці, значення та фільтри. Це дозволяє швидко змінювати спосіб відображення даних без необхідності редагування вихідних таблиць. Зведені таблиці підтримують можливість деталізації даних, що дозволяє користувачам глибше аналізувати окремі елементи. Крім того, вони можуть автоматично оновлюватися при зміні вихідних даних, що робить їх дуже корисними для роботи з динамічними даними.

Однією з ключових переваг зведених таблиць є їхня здатність працювати з великими наборами даних і перетворювати їх у зручний для аналізу формат за кілька кліків. Це робить їх незамінним інструментом для бізнес-аналітики, фінансових звітів, аналізу продажів та інших видів діяльності, де потрібне швидке

та ефективно узагальнення інформації. Зведені таблиці в Excel надають користувачам гнучкість і зручність, дозволяючи швидко отримувати корисні інсайти з великих обсягів даних.

Оскільки наші дані є дуже розрізнені та представлені у кількох вимірах одночасно, то доцільно буде використовувати саме функціонал зведених таблиць для аналізу та візуалізації.

Щоб створити зведені таблиці, потрібно сформулювати підключення до створеної нами моделі, яка знаходиться на хмарному просторі Power BI. Зробити це можна в Excel, відкривши меню зведеної таблиці на вкладці «Вставлення» та вибрати потрібну зі списку доступних моделей Power BI. Або ж відкрити саме хмарне сховище у браузері [55], перейти на потрібну модель та натиснути кнопку «Аналізувати в Excel», тоді створюється необхідний нам файл із підключення.

Створимо просту зведену таблицю, яка покаже продажі у грошовому вираженні в Євро за всі наявні роки із деталізацією до країни походження товару. В рядки додаємо CATEGORY та COUNTRY OF ORIGIN із таблиці Nielsen_Bulgaria, в стовпці Year із таблиці Date, в значення VALUE SALES (in 1000) in EURO, а у фільтри маркер джерела даних Mark, щоб залишити лише дані Нільсену. Зведена таблиця продемонстрована на рисунку 3.11 тепер містить дані про продажі в Євро в тисячах в розрізі років та походження. Також всі категорії відсортовані за спаданням від найприбутковішої до найменш прибуткової.

Як можемо спостерігати найкраще продалась категорія Sweet Biscuits H1 із загальними продажами у розмірі 652 681 тис. євро, приблизно 57% з яких займала продукція внутрішнього виробництва, та 43% імпортованих товарів. Також можемо спостерігати, що в більшості років ця категорія лідирувала за цим показником, поступившись лише раз в 2021 році категорії Chocolate tablets.

Рисунок 3.11 – Зведена таблиця із даними продажів категорій в розрізі років та країн походжень

(Джерело розроблено автором)

Рисунок 3.11 – Зведена таблиця із даними продажів категорій в розрізі років та країн походжень

(Джерело розроблено автором)

Припустимо, що нам потрібно проаналізувати продажі за виробниками та їх брендами по певній категорії (скажімо Countlines). При цьому необхідно виокремити дані за 2022 та 2023 роки, порівнявши їх за показниками об'ємів продажів та грошовими прибутками. Щоб зробити відповідну зведену таблицю вставляємо її (на новий аркуш) з підключенням до тієї ж моделі даних.

Для пришвидшення опрацювання даних моделлю, можна першим кроком відразу встановити фільтри, якими вона повинна керуватись обраховуючи всі дані. Тож до Фільтрів додаємо поле CAT із таблиці Класифікатор та поле Channel із таблиці Канали. Для першого обираємо потрібну категорію – Countlines, а для другого канал Total (прибирає розбиття на точки реалізації, залишаючи лише сумарні показники) для даних про продажі ритейлера.

В Рядки додаємо поля TOP-CAT, TOP-MANUFACTURER та TOP-BRAND із таблиці Класифікатор. Ці поля слугують власним внутрішнім класифікаційним принципом підприємства, та дозволяють правильно розподіляти дані не

побоюючись похибок від даних Нільсен. В стовпці додаємо лише поле Year із таблиці Date.

У значення додаємо два основні показники: VOLUME SALES (in 1000) + coef та VALUE SALES (in 1000) in EURO + coef. Також додаємо ці показники ще раз, але вже змінюючи їх параметри обчислення на Відображення значення як Зведена відмінність зважаючи на поле Year елемента 2022. Таким чином ми отримаємо наступні показники:

- об'ємів продажів у тонах,
- динаміку об'ємів продажів 2023 року у відсотках до 2022,
- продажів у грошовому вираженні в Євро із дооцінкою,
- динаміку продажів у грошовому вираженні із дооцінкою 2023 року у відсотках до 2022.

Додатково можна створити паралельну таблицю обрахунку відсотковості цих продажів від загального ринку. Створені зведені таблиці можна побачити на рисунку 3.12.

BULGARIA. ESTIMATED. COUNTLINES (MANUFACTURER/BRAND). 2022-2023							MARKET SHARE, %						
BULGARIA. COUNTLINES	TOTAL MARKET, VOLUME (tons)			TOTAL MARKET, VALUE (kEUR)			BULGARIA. COUNTLINES	TOTAL MARKET, VOLUME (tons)			TOTAL MARKET, VALUE (kEUR)		
	2022	2023	2023/2022,%	2022	2023	2023/2022, %		2022	2023	2023/2022,%	2022	2023	2023/2022,%
TOP-MANUFACTURERS							TOP-MANUFACTURERS						
MARS	1 663	1 514	-9,01%	17 579	20 308	15,53%	MARS	60%	54%	-5,9%	63%	61%	-2,8%
BOUNTY	342	280	-17,95%	3 979	4 170	4,80%	BOUNTY	12%	10%	-2,3%	14%	12%	-1,9%
MARS	204	198	-3,04%	2 183	2 574	17,91%	MARS	7%	7%	-0,3%	8%	8%	-0,2%
MILKY WAY	22	23	6,97%	336	480	43,00%	MILKY WAY	1%	1%	0,0%	1%	1%	0,2%
SNICKERS	786	717	-8,71%	7 775	9 172	17,97%	SNICKERS	28%	26%	-2,7%	28%	27%	-0,7%
TWIX	310	294	-4,97%	3 307	3 913	18,33%	TWIX	11%	11%	-0,6%	12%	12%	-0,2%
MONDELEZ	285	360	26,32%	3 197	4 272	33,61%	MONDELEZ	10%	13%	2,6%	12%	13%	1,2%
OTHER	1	6	367,07%	14	79	460,45%	OTHER	0%	0%	0,2%	0%	0%	0,2%
REPUBLIKA	284	354	24,71%	3 183	4 192	31,71%	REPUBLIKA	10%	13%	2,4%	11%	12%	1,0%
ROSHEN SELL-OUT	241	331	37,53%	1 820	2 785	53,05%	ROSHEN SELL-OUT	9%	12%	3,2%	7%	8%	1,7%
TIDBIT	16	11	-29,92%	114	80	-29,99%	TIDBIT	1%	0%	-0,2%	0%	0%	-0,2%
Батон ROSHEN	225	320	42,25%	1 706	2 705	58,61%	Батон ROSHEN	8%	11%	3,3%	6%	8%	1,9%
FOBOS	153	151	-1,70%	1 147	1 326	15,58%	FOBOS	6%	5%	-0,1%	4%	4%	-0,2%
OTHER	4	4	3,35%	25	32	29,32%	OTHER	0%	0%	0,0%	0%	0%	0,0%
RODOPA	149	147	-1,83%	1 123	1 294	15,28%	RODOPA	5%	5%	-0,1%	4%	4%	-0,2%
ULKER	94	75	-19,84%	967	951	-1,65%	ULKER	3%	3%	-0,7%	3%	3%	-0,6%
OTHER	3	2	-31,29%	32	20	-38,14%	OTHER	0%	0%	0,0%	0%	0%	-0,1%
ALBENI	77	63	-18,17%	770	787	2,28%	ALBENI	3%	2%	-0,5%	3%	2%	-0,4%
LAVIVA	13	10	-26,97%	165	144	-12,92%	LAVIVA	0%	0%	-0,1%	1%	0%	-0,2%
ELVAN	41	51	25,35%	216	308	42,33%	ELVAN	1%	2%	0,4%	1%	1%	0,1%
COCOS	35	47	33,12%	187	286	53,26%	COCOS	1%	2%	0,4%	1%	1%	0,2%
OTHER	5	4	-28,15%	30	22	-26,70%	OTHER	0%	0%	-0,1%	0%	0%	0,0%
MILMEKS	44	43	-2,31%	582	631	8,40%	MILMEKS	2%	2%	0,0%	2%	2%	-0,2%
OTHER	2	2	-5,71%	26	26	-0,08%	OTHER	0%	0%	0,0%	0%	0%	0,0%
SINEMORETZ	42	41	-2,14%	556	605	8,80%	SINEMORETZ	2%	1%	0,0%	2%	2%	-0,2%
SLADAK	23	35	55,42%	103	147	43,13%	SLADAK	1%	1%	0,4%	0%	0%	0,1%
KREMALINO	23	35	55,42%	103	147	43,13%	KREMALINO	1%	1%	0,4%	0%	0%	0,1%
PRIVATE LABEL	16	39	140,38%	95	268	181,87%	PRIVATE LABEL	1%	1%	0,8%	0%	1%	0,5%
OTHER	210	195	-7,14%	2 059	2 553	23,99%	OTHER	8%	7%	-0,6%	7%	8%	0,2%
Загальний підсумок	2 769	2 793	0,86%	27 765	33 548	20,83%	Загальний підсумок	100%	100%		100%	100%	

Рисунок 3.12 – Зведена таблиця із даними продажів у серед виробників та їх брендів у розрізі 2022 та 2023 років із динамікою (зведеною відмінністю)

(Джерело розроблено автором)

Проаналізувавши дані таблиць можемо зробити висновок, що в категорії Countlines за 2022 та 2023 роки в Болгарії найбільшими продажами (як об'ємами, так і в грошовому вираженні) вирізнявся виробник Mars зі своїми про п'ятьма найпродаванішими брендами такими як Bounty, Mars, Milky Way, Snickers та Twix. Сумарно всі вони продались в розмірі 1 663 тони в 2022 році та 1 514 тонами у 2023-му, що характеризує спад 9,01% за рік.

Проте об'єми продажів не завадили отримати прибуток на 15,53% більше в 2023 році в порівнянні з 2022-им, а саме 20 308 тис. євро до 17 579 тис. євро відповідно.

Аналізуючи дані продажів продуктів компанії Рошен, бачимо стрімке зростання об'ємів із 241 в 2022 році до 331 тон в 2023-му, що характеризує збільшення їх на 37,53% за рік. Кажучи про прибутки, зростання було ще помітнішим, а саме на 53,05%, за рахунок 2 785 тис. євро в 2023 році, до 1 820 тис. євро в 2022-му.

Лева частка у розмірі 60% ринку серед об'ємів продажів у 2022 році припадала на компанію Mars, та в 2023 вже становила 54%. Також частка прибутків зменшилась на 2,8% від загалу (63% до 61%).

Також помітне зростання частки об'ємів продажів компанії Рошен на ринку Болгарії з 9% в 2022-му до 12% в 2023 році, та зростання частки прибутків із 7% до 8%

Висновок до третього розділу

У даному розділі було наведено наочний приклад обробки даних, розробки моделі даних для створення важливої аналітичної звітності.

Було продемонстровано повний процес передобробки датасетів в Excel, шляхом виокремленням спільних полів та їх розширенням до єдиної спільної структури. Також наведено приклад портування та перетворення даних в редакторі Power Query з використанням мови M та користувацького інтерфейсу вікна Перетворення. Наступним етапом було досліджено процес налаштування моделі використовуючи методи класифікації даних, уніфікації та ідентифікації записів баз

даних про товари, їх категорій, виробників, точок реалізації продажів, каналів збуту. Також продемонстровано створення та використання таблиці дат для аналізу даних за часовою ознакою. Було розглянуто створення зв'язків між таблицями, мір та публікація моделі до хмарного сховища Power BI.

Заключною частиною розділу стали формування аналітичної звітності в Excel з використанням функціоналу зведених таблиць з під'єднанням до моделі даних, для якісного багатовимірного опрацювання даних.

Аналіз цієї інформації демонструє стан ринку кондитерських виробів в Болгарії, спираючись на дані аналітичних звітів, та допомагає помітити зростання та падіння певних суб'єктів цього ринку за різними економічними показниками.

ВИСНОВКИ

У першому розділі було детально визначено та розглянуто основні поняття даних та системного аналізу, виявлено їхню ключову роль у сучасному інформаційному забезпеченні. Було продемонстровано організацію та аналіз життєвого циклу даних, що є невід’ємною частиною всіх процесів у виробничій сфері. Було розглянуто теоретичні засади, концепції та методи моделювання даних, які сприяють правильному керуванню ними на підприємстві. Таким чином, даний розділ формує теоретичний фундамент для подальшого дослідження процесу аналізу та моделювання даних.

У другому розділі було розглянуто практичні аспекти моделювання даних та бізнес-аналітики з використанням такого поняття як сховища даних. Було проведено детальний огляд сучасних аналітичних інструментів, таких як Excel та Power BI, розглянуті їх основні характеристики, структура, мета використання, сильні та слабкі сторони, можливості а також загрози їх застосування у виробничій сфері.

У третьому розділі на конкретному прикладі було детально продемонстровано та досліджено процес попереднього аналізу та обробки даних, зведення датасетів у Excel та перетворення даних у Power Query, з метою правильного їх використання у наступних етапах системного аналізу даних. Також було розглянуто процес налаштування моделі даних Power BI, а саме створення та використання класифікаційних таблиць, зв’язків між таблицями та опрацьовано створення мір, публікації моделі до хмарного сховища.

Заключною частиною роботи стало візуалізація даних та формування зведених таблиць для виділення якісних аналітичних висновків, спираючись на економічні показники суб’єктів ринку кондитерських виробів в Болгарії.

Отже, проведене дослідження показало, що все перераховане вище є ключовими складовими успішного впровадження системного аналізу даних у поєднанні з глибоким розумінням конкретних бізнес-процесів, що допоможе

підвищити ефективність та конкурентоспроможність підприємства на ринку. Це лише підтверджує доцільність дослідження методів обробки та моделювання для виконання аналізу даних, а також доцільність розвитку різноманітних програмних засобів, які уможливають та підтримують виконання цих процесів у виробничій сфері.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Шепиленко, В. Теоретичні основи організації бізнес-процесів промислового підприємства. Bulletin of Sumy National Agrarian University, 2022. (1 (91), 39-46). URL: <https://doi.org/10.32782/bsnau.2022.1.6> (дата звернення: 20.04.2024)
2. The Importance of System Analysis in Software Development : веб-сайт. URL: <https://www.linkedin.com/pulse/importance-system-analysis-software-development-tradinos> (дата звернення: 27.04.2024)
3. Прокопенко Т. О. Теорія систем і системний аналіз : навч. посіб. – Черкаси : ЧДТУ, 2019 – 139 с. URL: <https://er.chdtu.edu.ua/handle/ChSTU/986> (дата звернення: 20.04.2024)
4. Grorapid Labs, What is system analysis?- Steps, importance & implementation : веб-сайт. URL: <https://www.grorapidlabs.com/blog/what-is-system-analysis-steps-importance-implementation> (дата звернення: 27.04.2024)
5. System Analysis | System Design : веб-сайт. URL: <https://www.geeksforgeeks.org/system-analysis-system-design/> (дата звернення: 27.04.2024)
6. Gerrit Muller, System Modeling and Analysis: a Practical Approach : веб-сайт. URL: <https://www.gaudisite.nl/SystemModelingAndAnalysisBook.pdf> (дата звернення: 27.04.2024)
7. Christine Namugenyia, Shastri L Nimmagaddab, Torsten Reiners. Design of a SWOT Analysis Model and its Evaluation in Diverse Digital Business Ecosystem Contexts. : веб-сайт URL: <https://demo.dspacedirect.org/items/a94ef210-25e1-4399-b93e-54930f2ba37f> (дата звернення: 20.04.2024)
8. Martin E. Modell, Data Analysis, Data Modeling and Classification : веб-сайт URL: <http://www.martymodell.com/dadmc/dadmc03.html> (дата звернення: 27.04.2024)

9. Jianqing Fan, Fang Han, Han Liu, Challenges of Big Data analysis, National Science Review, Volume 1, Issue 2, June 2014, Pages 293–314,
10. Міца О.В., Лавер В.О. Системний аналіз : навч.-метод. посіб. / О.В. Міца, В.О. Лавер. – Ужгород : вид-во ПП «АУТДОР - ШАРК», 2021. – 63 с.
11. І. Ю. Черепанська, А. Ю. Сазонов. Сучасні інформаційні технології та системний аналіз у наукових дослідженнях. : навч. посіб. – Київ : КПП ім. Ігоря Сікорського, 2021. – 270 с. URL: <https://ela.kpi.ua/server/api/core/bitstreams/ff473831-3602-4088-8f53-4df924733bc5/content> (дата звернення: 20.04.2024)
12. Томчук О.Ф., Кафлевська С.Г. Системний аналіз як метод ухвалення і обґрунтування рішень : Збірник наукових праць ВНАУ. 2012. №3 (69) Том 2. URL: <http://socrates.vsau.org/repository/getfile.php/6276.pdf> (дата звернення: 27.04.2024)
13. A Beginner's Guide to Data & Analytics : веб-сайт. URL: https://online.hbs.edu/Documents/a-beginners-guide-to-data-and-analytics.pdf?_gl=1*108cr72*_gcl_au*MTkwNjUwNzExNS4xNzE1NjIzOTAx (дата звернення: 27.04.2024)
14. Catherine Cote, 7 data collection methods in business analytics : веб-сайт. URL: <https://online.hbs.edu/blog/post/data-collection-methods> (дата звернення: 27.04.2024)
15. Xu, K., Li, Y., Liu, C. et al. Advanced Data Collection and Analysis in Data-Driven Manufacturing Process. Chin. J. Mech. Eng. 33, 43 (2020). URL: <https://doi.org/10.1186/s10033-020-00459-x>
16. Michael G. Rossmann, Cornelis G. van Beek. Data processing. Part of CCP4 Study Weekend 1999. URL: <https://doi.org/10.1107/S09074444999008379> (дата звернення: 20.04.2024)
17. Stravoravdis, S, and Marsh, A. A proposed method for generating, storing and managing large amounts of modelling data using scripts and on-line databases. Canada: N. p., 2005. URL: <https://www.osti.gov/etdeweb/biblio/20685539> (дата звернення: 20.04.2024)

18. David Loshin. Master Data Management. Burlington: книга, 2009. URL: <https://books.google.com.ua/books?id=dJtmVz3nrOUC&lpg=PP1&ots=9IqLVWkW5T&dq=data%20management&lr&hl=uk&pg=PR6#v=onepage&q=data%20management&f=false> (дата звернення: 20.04.2024)
19. Michael R. Berthold, David J Handю Intelligent Data Analysis: An Introduction. New York. : книга, 2007. URL: <https://books.google.com.ua/books?id=den1CAAAQBAJ&lpg=PA1&ots=p93DrVBWtR&dq=data%20analysis&lr&hl=uk&pg=PP4#v=onepage&q=data%20analysis&f=false> (дата звернення: 20.04.2024)
20. Kieran Healy. Data Visualization: A Practical Introduction. : книга, 2019. URL: https://books.google.com.ua/books?id=3XOYDwAAQBAJ&lpg=PR11&ots=I_a26LfjQ0&dq=data%20visualization&lr&hl=uk&pg=PR11#v=onepage&q=data%20visualization&f=false (дата звернення: 20.04.2024)
21. Susan Spiggle, Analysis and Interpretation of Qualitative Data in Consumer Research, Journal of Consumer Research, Volume 21, Issue 3, December 1994, Pages 491–503, <https://doi.org/10.1086/209413> (дата звернення: 20.04.2024)
22. О.Н.Романюк, Т.О.Савчук Організація баз даних і знань : навч. посіб. [Електронний ресурс] Вінниця: УНІВЕРСИУМ-Вінниця, 2003 — 217 с. URL: <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/19527/4.17%20Організація%20баз%20даних%20і%20знань.output.pdf>
23. Gil R. Mosard, Modeling in systems analysis : веб-сайт. URL: [https://doi.org/10.1016/0167-5419\(85\)90017-1](https://doi.org/10.1016/0167-5419(85)90017-1) (дата звернення: 27.04.2024)
24. Varuni Mallikaarachchi, Information systems analysis - IS 6840: Data Modeling for System Analysis : веб-сайт. URL: <https://www.umsl.edu/~sauterv/analysis/Fall2010Papers/varuni/> (дата звернення: 27.04.2024)
25. Zhao L, Chen Y, Schaffner DW.2001.Comparison of Logistic Regression and Linear Regression in Modeling Percentage Data. Appl Environ Microbiol67:. URL: <https://doi.org/10.1128/AEM.67.5.2129-2135.2001>

26. D. R. Dolk. Model management and structured modeling: the role of an information resource dictionary system. : веб-сайт. URL: <https://dl.acm.org/doi/abs/10.1145/62959.62967> (дата звернення: 27.04.2024)
27. Васильєва, Л. В., Кльованик, О. А. Регресійні моделі та аналіз часових рядів. : навч. посіб. – Краматорськ : ДДМА, 2010. – 176 с. URL: <http://dspace.dgma.donetsk.ua:8080/handle/DSEA/252> (дата звернення: 04.05.2024)
28. Lawrence M. Seiford, Modeling undesirable factors in efficiency evaluation. European Journal of Operational Research. Volume 142, Issue 1, 1 October 2002, Pages 16-20 URL: [https://doi.org/10.1016/S0377-2217\(01\)00293-4](https://doi.org/10.1016/S0377-2217(01)00293-4) (дата звернення: 27.04.2024)
29. Черчелюк С.М. Метод «дерево рішень». : зб. доп. 89-ї наукової студентської конференції, 20 травня — 4 червня 2022 р. [Електронний ресурс]. Київ, КНЕУ, 2022. С. 366–368.
30. Yan-yan Song, Ying Lu. Decision tree methods: applications for classification and prediction. Shanghai Archives of Psychiatry, 2015, Vol. 27, No. 2 URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4466856/pdf/sap-27-02-130.pdf> (дата звернення: 04.05.2024)
31. Madison Schott, What is data modeling? Your go-to guide : веб-сайт. URL: <https://www.thoughtspot.com/data-trends/data-modeling> (дата звернення: 04.05.2024)
32. Sonny Rivera, 3 types of data models and when to use them : веб-сайт. URL: <https://www.thoughtspot.com/data-trends/data-modeling/types-of-data-models> (дата звернення: 04.05.2024)
33. Joe Abbott, A complete overview of dimensional data modeling : веб-сайт. URL: <https://www.thoughtspot.com/data-trends/data-modeling/dimensional-data-modeling> (дата звернення: 04.05.2024)
34. Craig S. Mullins, Definition Database management system (DBMS) : веб-сайт. URL: <https://www.techtarget.com/searchdatamanagement/definition/database-management-system> (дата звернення: 04.05.2024)

35. PictureMichael Hammer, PictureDennis Mc Leod, Database description with SDM: a semantic database model : веб-сайт. URL: <https://dl.acm.org/doi/abs/10.1145/319587.319588> (дата звернення: 04.05.2024)
36. Robert J. Muller, Database Design for Smarties: Using UML for Data Modeling : веб-сайт. URL: https://books.google.com.ua/books?id=saBE4jdxw_wC&lpg=PR15&dq=data%20modeling&lr&hl=uk&pg=PR4#v=onepage&q=data%20modeling&f=false (дата звернення: 04.05.2024)
37. Chuck Ballard, Dirk Herreman, Don Schau, Rhonda Bell, Eunsang Kim, Ann Valencic, Data Modeling Techniques for Data Warehousing : веб-сайт. URL: <https://eddyswork.synthasite.com/resources/Data%20Modeling%20Tech%20For%20Data%20Warehouseing.pdf> (дата звернення: 04.05.2024)
38. Shamkant B. Navathem, Evolution of data modeling for databases : веб-сайт. URL: <https://dl.acm.org/doi/abs/10.1145/130994.131001> (дата звернення: 04.05.2024)
39. What Is a Data Warehouse? : веб-сайт. URL: <https://www.oracle.com/database/what-is-a-data-warehouse/> (дата звернення: 04.05.2024)
40. Enterprise Data Warehouse (EDW): Solutions & Benefits : веб-сайт. URL: <https://www.tridant.com/enterprise-data-warehouse/> (дата звернення: 04.05.2024)
41. S. Sharma and R. Jain, "Modeling ETL Process for Data Warehouse: An Exploratory Study," 2014 Fourth International Conference on Advanced Computing & Communication Technologies, Rohtak, India, 2014, с. 271-276, URL: <https://ieeexplore.ieee.org/abstract/document/6783463> (дата звернення: 04.05.2024)
42. Surajit Chaudhuri and Umeshwar Dayal. An overview of data warehousing and OLAP technology. SIGMOD Rec. 26, 1. March 1997, С. 65–74. URL: <https://doi.org/10.1145/248603.248616> (дата звернення: 04.05.2024)

43. Пономаренко І. В., Телеус А. В. Бізнес-аналітика як ефективний інструмент обробки даних. Проблеми інноваційно-інвестиційного розвитку. 2020. № 23. С. 64-70.
44. Jonathan Paul , 4 Best Practices For Creating a Variance Analysis Report in Excel : веб-сайт. URL: <https://www.venasolutions.com/blog/4-best-practices-creating-variance-analysis-report-excel> (дата звернення: 04.05.2024)
45. Joseph Schmuller, Statistical Analysis with Excel For Dummies : веб-сайт. URL: <https://books.google.com.ua/books?id=E5nIk3NK6SoC&lpg=PA4&ots=Vn1Hj1Uo75&dq=Everything%20You%20Need%20To%20Know%20About%20excel&lr&hl=uk&pg=PA4#v=onepage&q=Everything%20You%20Need%20To%20Know%20About%20excel&f=false> (дата звернення: 04.05.2024)
46. Stephen L. Nelson, E. C. Nelson, Excel Data Analysis For Dummies : веб-сайт. URL: <https://books.google.com.ua/books?id=5oxgAwAAQBAJ&lpg=PA141&ots=jnQtSWjtsS&dq=Everything%20You%20Need%20To%20Know%20About%20excel&lr&hl=uk&pg=PR9#v=onepage&q=Everything%20You%20Need%20To%20Know%20About%20excel&f=false> (дата звернення: 04.05.2024)
47. Сирий В.М, Моделювання даних у MS Power BI Desktop : Методичний коментар URL: https://repo.btu.kharkov.ua/bitstream/123456789/44464/1/TDMNPK_Finansova%20arkhitektonika%20ta%20stsenarii%20konkurentnykh%20modelei%20rozvytku_2023-253-255.pdf (дата звернення: 04.05.2024)
48. What is Power BI? : веб-сайт. URL: <https://learn.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview> (дата звернення: 04.05.2024)
49. Avijeet Biswal, What is Power BI?: Architecture, and Features Explained : веб-сайт. URL: <https://www.simplilearn.com/tutorials/power-bi-tutorial/what-is-power-bi> (дата звернення: 04.05.2024)
50. Nicole Diceman, What Is Power BI? - Everything You Need To Know : веб-сайт. URL: <https://www.venasolutions.com/blog/what-is-power-bi> (дата звернення: 04.05.2024)

51. Ліщинська Л. Б. Основні підходи до моделювання даних у Microsoft Power BI. : збірник матеріалів Міжнародної науково-практичної Інтернет конференції 20-21 листопада 2023 р. – Суми/Вінниця: НІКО/ КЗВО «Вінницька академія безперервної освіти», 2023. – С. 144–147
52. Everything You Need To Know About Power BI : веб-сайт. URL: <https://www.tridant.com/everything-need-know-power-bi/> (дата звернення: 04.05.2024)
53. Vena Solutions, How To Improve Your Financial Reporting With Power BI : веб-сайт. URL: <https://www.venasolutions.com/blog/financial-reporting-in-power-bi> (дата звернення: 04.05.2024)
54. Amruta Deshpande, How to create effective public-facing reports in Microsoft Power BI : веб-сайт. URL: <https://www.tridant.com/create-effective-public-facing-reports-power-bi/> (дата звернення: 04.05.2024)
55. Power BI : веб-сайт URL: <https://app.powerbi.com/> (дата звернення: 04.05.2024)

Ім'я користувача: Комп'ютерної математики та інформаційної безпеки...	ID перевірки: 1016289373
Дата перевірки: 28.05.2024 09:21:13 EEST	Тип перевірки: Doc vs Internet + Library
Дата звіту: 28.05.2024 09:22:01 EEST	ID користувача: 100005746

Назва документа: Диплом_Ситник

Кількість сторінок: 94 Кількість слів: 20262 Кількість символів: 156619 Розмір файлу: 2.66 MB ID файлу: 1016082580

2.38% Схожість

Найбільша схожість: 0.19% з Інтернет-джерелом (<https://community.powerbi.com/t5/Desktop/Calculating-for-Quarter/...>)

1.26% Джерела з Інтернету	69	Сторінка 96
---------------------------	----	-------------

1.97% Джерела з Бібліотеки	321	Сторінка 97
----------------------------	-----	-------------

0% Цитат

Вилучення цитат вимкнене

Вилучення списку бібліографічних посилань вимкнене

0% Вилучень

Немає вилучених джерел