

Міністерство освіти і науки України
Національний університет «Львівська політехніка»
Інститут прикладної математики та фундаментальних наук
Кафедра прикладної математики

Звіт
про виконання курсової роботи
з курсу «Надвеликі бази даних»

Варіант 8

Виконав :
студент групи ПМ-42
Пелещишин Володимир Андрійович
Перевірив:
Любінський Богдан Богданович

(дата)

(підпис викладача)

Анотація

У даній курсовій роботі виконано розробку інформаційної системи "Облік товарів", яка призначена для автоматизації процесів обліку в невеликому магазині. Основна мета роботи полягала у створенні комплексного рішення, що охоплює всі етапи обробки даних: від генерації та завантаження даних до аналізу і візуалізації.

Для досягнення поставленої мети було виконано такі завдання: генерація даних для бази "Облік товарів", реалізація ETL процесу з використанням SSIS, побудова багатовимірного куба за допомогою SSAS та проведення аналізу даних з використанням SSRS

Результатом виконання курсової роботи стала функціональна система для обліку та аналізу товарів у магазині, яка дозволяє автоматизувати процеси обліку, виявляти тенденції продажів, та приймати ефективні управлінські рішення. Інтеграція SSIS, SSAS і SSRS забезпечила повний цикл роботи з даними – від їх збору до аналітики.

Зміст

Анотація	2
1. Теоретична частина.....	4
1.1. ETL-процес.....	4
1.2. OLAP-куб	4
1.3. SSRS	5
2. База даних та сховище даних	6
2.1. База даних	6
2.2. Сховище даних	6
3. ETL-процес (SSIS)	7
4. SSAS	9
5. SSRS	11
Висновок	15
Список використаної літератури:	16

1. Теоретична частина

1.1. ETL-процес

ETL (Extract, Transform, Load) — це процес обробки даних, який включає три основні етапи: витяг даних із джерел (екстракцію), зміну структури або формату даних відповідно до потреб (трансформацію) та завантаження оброблених даних у сховище в нашому випадку Data Warehouse. ETL-процес є важливою частиною побудови сховищ даних, оскільки саме він відповідає за об'єднання, очищення та підготовку інформації з різних джерел для подальшого аналізу.

Для реалізації ETL-процесів розроблено багато інструментів, і найпопулярніший є SQL Server Integration Services (SSIS). SSIS — це інструмент для інтеграції даних, який дозволяє налаштувати автоматичні робочі процеси, з'єднуватися з джерелами даних, виконувати різні трансформації (наприклад, фільтрацію, обчислення або зміну структури даних) і завантажувати дані в сховище.

1.2. OLAP-куб

OLAP (Online Analytical Processing) куб — це багатовимірна структура даних, яка використовується для швидкого виконання складних запитів і аналізу даних з різних точок зору. OLAP-куби дозволяють виконувати аналітичні операції, такі як агрегація, фільтрація, порівняння даних та деталізація.

Структура куба складається з вимірів (dimensions) та фактів (facts):

- Виміри — це "осі аналізу", наприклад, час, регіон, продукт. Вони допомагають класифікувати та групувати дані.
- Факти — це числові дані, які піддаються обчисленням, наприклад, кількість продажів, дохід або обсяг товару.

Для створення OLAP-кубів використовується SQL Server Analysis Services (SSAS) — це частина SQL Server, яка дозволяє будувати багатовимірні моделі даних. SSAS забезпечує швидкий доступ до

великих обсягів інформації та підтримує виконання складних аналітичних запитів. Це дозволяє бізнес-користувачам отримувати необхідну інформацію в режимі реального часу.

1.3. SSRS

SQL Server Reporting Services (SSRS) — це інструмент для створення звітів на основі даних. SSRS є частиною SQL Server і дозволяє створювати звіти різних типів — від простих таблиць до складних графіків та матриць. Завдяки вбудованим компонентам і можливості інтеграції з іншими інструментами, такими як SSAS, SSRS дозволяє створювати інтерактивні та динамічні звіти.

SSRS підтримує експорт звітів у різноманітні формати, включаючи PDF, Excel, Word, що робить їх зручними для подальшого аналізу та обміну. Це ідеальний інструмент для візуалізації даних, оскільки користувачі можуть легко отримувати інформацію у зрозумілому вигляді.

Таким чином, кожен із цих компонентів (ETL-процес, OLAP-куб і SSRS) виконує важливу роль у повному циклі роботи з даними: від їх збору та обробки до аналізу й візуалізації. Усі ці технології інтегруються між собою та створюють цілісну систему для ефективного управління даними.

Основна частина

2. База даних та сховище даних

2.1. База даних

Розроблена база даних "Облік товарів" призначена для автоматизації процесів обліку товарів у невеликому магазині. Вона забезпечує збереження, обробку та аналіз інформації про продажі, постачання, магазини, товари та регіони. База даних має такі основні таблиці:

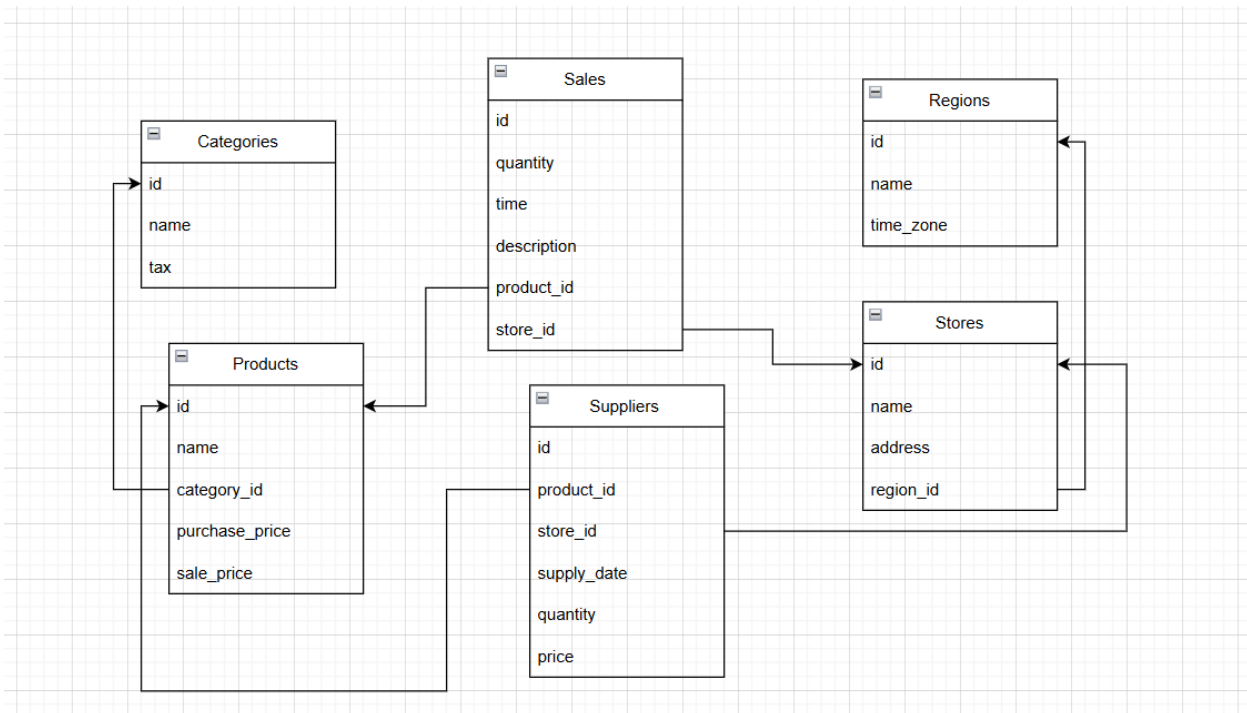


рис. 1

2.2. Сховище даних

Сховище даних, розроблене для завдання має зорієнтовану на факти та виміри структуру, що дозволяє виконувати ефективний багатовимірний аналіз даних.

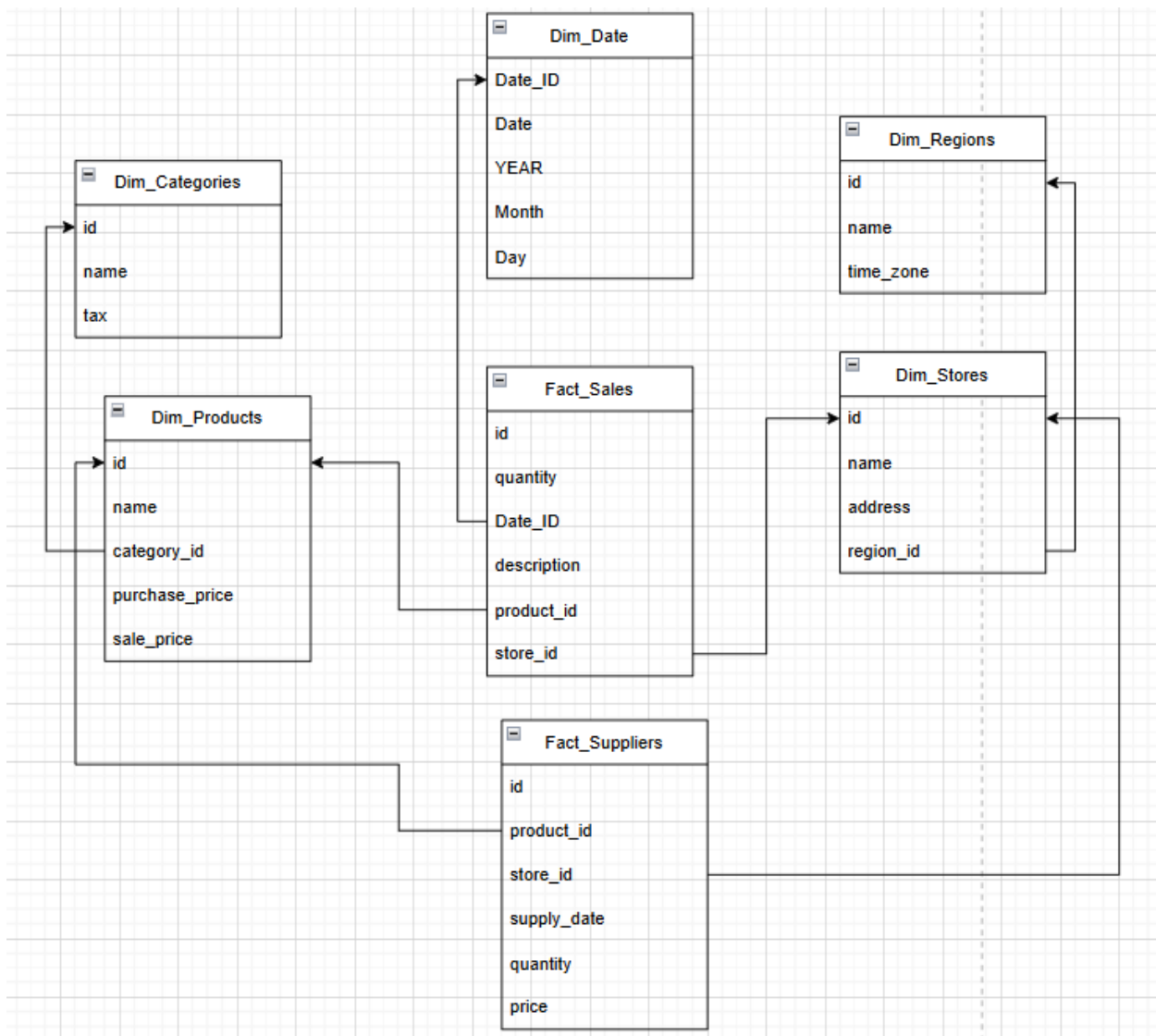


рис. 2

3. ETL-процес (SSIS)

У рамках курсової роботи було розроблено ETL-процес із використанням SQL Server Integration Services (SSIS) для автоматизації завантаження, обробки та перенесення даних у сховище даних. Процес реалізовано для забезпечення інтеграції даних із різних джерел та підготовки їх для подальшого аналізу в багатовимірному кубі.

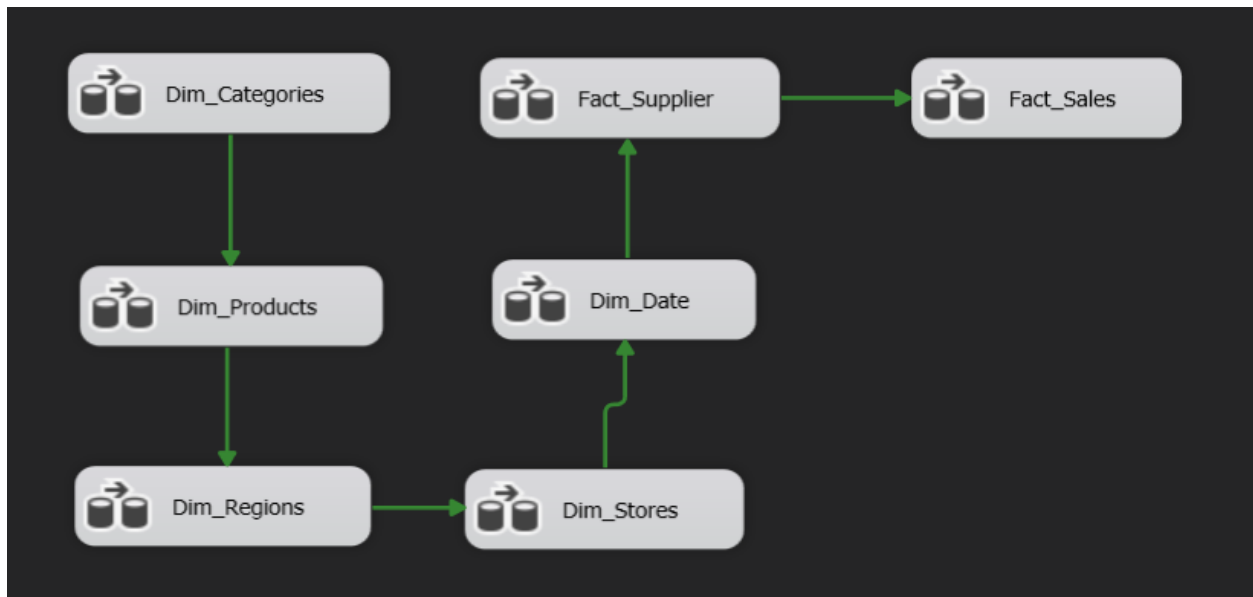


рис. 3

На етапі екстракції дані завантажуються з вихідних джерел у тимчасові таблиці. Дані витягуються з таблиць, таких як Categories, Products, Regions, Stores, Supplier і Sales.

Етап трансформації забезпечує обробку та очищення даних: перевірку цілісності даних, обчислення нових атрибутів Dim_date.

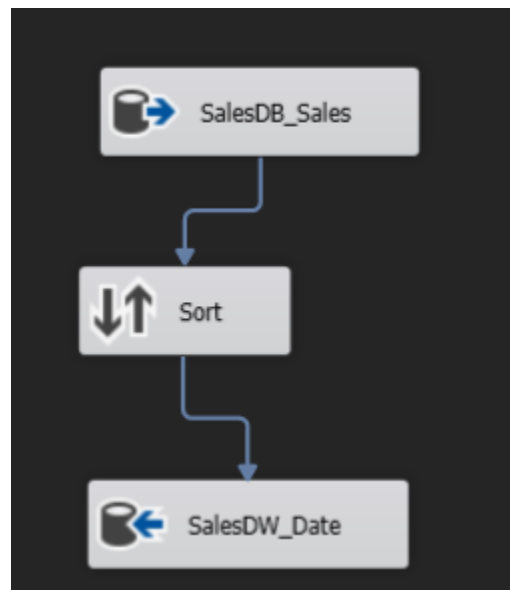


рис. 4

На етапі завантаження дані записуються до сховища даних, що має структуру з таблицями фактів та вимірів

Таблиці фактів:

- Fact_Sales: Завантажуються дані про продажі товарів.
- Fact_Supplier: Завантажуються дані про постачання.

Таблиці вимірів:

- Dim_Categories, Dim_Products, Dim_Regions, Dim_Stores, Dim_Date: Дані з очищенням та трансформацією записуються у відповідні виміри.

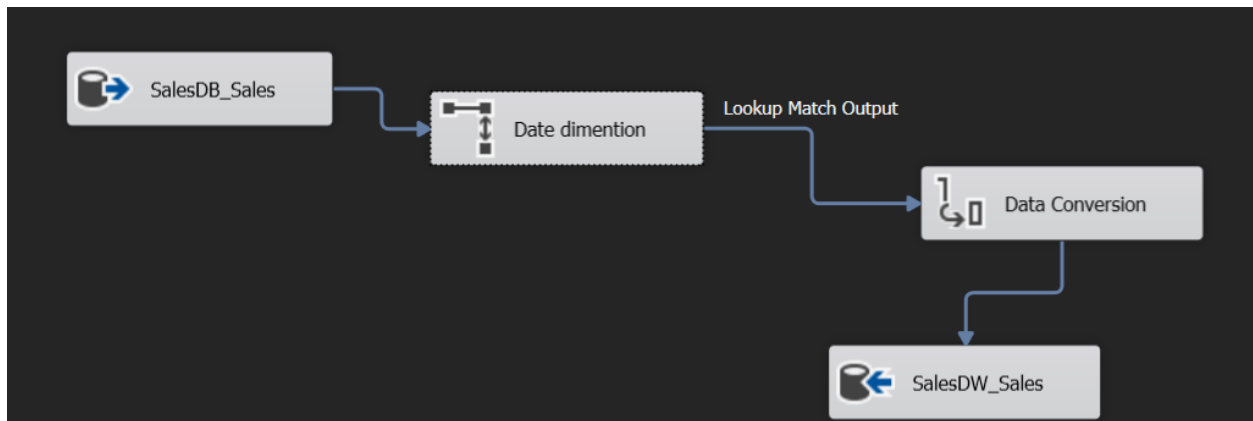


рис. 5

4. SSAS

Також було створено OLAP-куб на основі раніше побудованого сховища даних, що дозволяє здійснити багатовимірний аналіз даних. Для побудови куба використано SQL Server Analysis Services (SSAS).

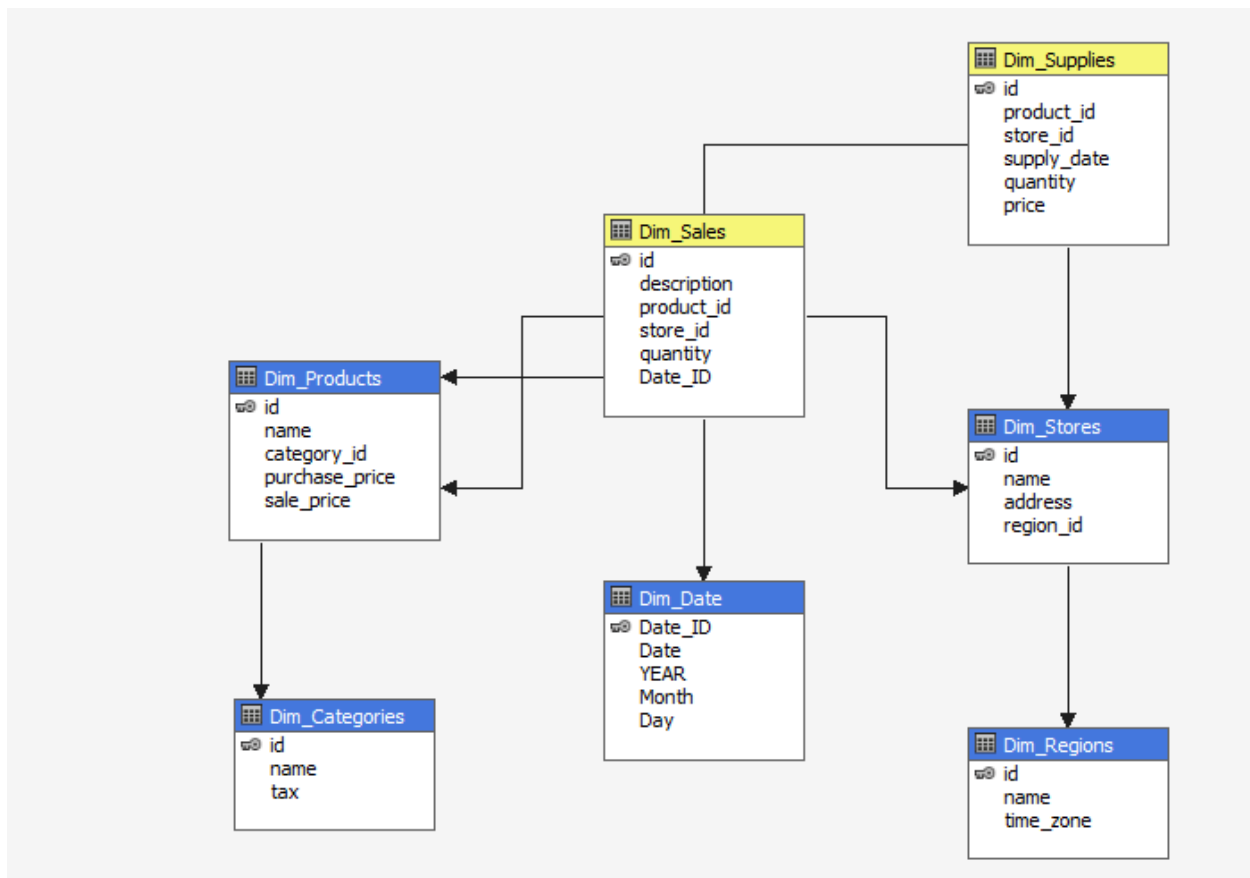


рис. 6

OLAP-куб містить такі виміри: Products, Date, Category, Regions та Stores

Для виміру Date була розроблена ієрархія. Ієрархія складається з кількох рівнів: Date, Year, Month та Day, що робить аналіз даних більш зручним і ефективним.

5. SSRS

У рамках курсової роботи було реалізовано систему аналітичних звітів за допомогою SQL Server Reporting Services (SSRS), яка дозволяє користувачам отримувати інформацію про ключові показники бізнесу, пов'язані з постачаннями та продажами товарів у магазині. Звіти надають можливість аналізувати дані, виявляти тенденції та приймати управлінські рішення на основі даних.

Stores by 10 regions

Region name	Id	Name	Address
Aruba	121	Monsipegentor WorldWide	374 White First Street
	99625	Endtinplax Holdings	445 Rocky Clarendon Parkway
Bhutan	3251	Tipfropicator Holdings	988 Old Boulevard
	99892	Qwivenopover	20 South Oak Street
Cuba	1363	Redimegentor WorldWide Group	737 West Green Clarendon Way
	99855	Revenantor Group	48 Cowley Boulevard
Greece	244	Surwerepar Company	33 Green Milton Road
	1195	Frogliban International	97 White Cowley Avenue
Isle of Man	99896	Bardudar Holdings Corp.	23 Green Nobel Road
	575	Froerilar Direct	96 South Green Nobel Avenue
Mongolia	99132	Hapmunantor Group	918 West Rocky Cowley Boulevard
	2114	Cipsaponistor Group	92 Hague Drive
Palau	99115	Surdudupower Holdings	18 Green Second Avenue
	836	Inbanan Holdings Group	96 Milton Street
Panama	99096	Recador WorldWide	24 East First Road
	2186	Enddimegar WorldWide	826 West New Road
Russia	964	Klitinplentor	27 Oak Boulevard
Timor-Leste			

табл. 1 Магазины за 10 першими регіонами

Name	Sales Count	Quantity
Lomquestar	8	42169
Thruerin	4	10220
Doppickower	6	18940
Inweropazz	1	4287
Kliwerpaman	1	6280
Emtumax	5	14569
Ciptumover	8	44268
Endquestadazz	1	6384
Thrufropantor	11	42861
Tupvenonazz	1	4031
Rapfropex	6	27034
Doprobefor	1	8647
Tippebistor	5	33724
Winquestentor	6	28681
Lomhupentor	8	38951
Hapdimonin	2	9166
Parwerpopex	3	6202
Zeepebax	7	35481
Surbanadistor	1	4167
Parwerplover	1	4722

рис. 7 Кількість продажів за Продуктом

Categories

Id	Name	Tax
1	Dairy	38.56
2	Beverages	20.21
3	Cereals	8.3
4	Grain	49.26
5	Snails	5.45
6	Confections	15.33
7	Seafood	40.11
8	Poultry	22.28
9	Meat	11.25
10	Shell fish	0.56

табл. 2 Категорії продуктів та їх податок

Supplies_by_product

Name	Quantity	Price
Lomquestar	4682	246692.62
Thruerin	2926	122055.55
Doppickower	4955	361455.58
Inweropazz	526	71132.22
Kliwerpaman	257	42296.6
Emtumax	2062	86560.44
Ciptumover	5405	387027.37
Endquestadazz	631	81942.91
Thrufropantor	6869	590631.84
Tupvenonazz	137	70525.62
Rapfropex	2817	368386.87
Doprobefor	306	17338.72
Tippebistor	3870	187283.11
Winquestentor	2086	311421.49
Lomhupentor	3600	455330.86
Hapdimonin	767	49761.88
Parwerpopex	1115	128850.97
Zeepebax	3299	385807.67

рис. 8 Поставки за продуктами

Products report

Id	Name	Purchase Price	Sale Price	Category Name	Tax
65536	Froquestexantor	32732.46	52847.19	Beverages	20.21
65537	Dopwerpaquower	65471.81	42648.63	Confections	15.33
65538	Vartinistor	51525.99	73814.39	Poultry	22.28
65539	Zeequestollentor	39078.90	62426.57	Cereals	8.30
65540	Grovenopicator	44642.07	373.57	Dairy	38.56
65541	Upbanefor	136.14	50215.14	Cereals	8.30
65542	Doptumistor	74203.16	68460.23	Beverages	20.21
65543	Parcadar	26559.28	32243.04	Poultry	22.28
65544	Hapdimonin	78197.78	34971.59	Shell fish	.56
65545	Unbanollantor	50476.34	57353.62	Dairy	38.56
65546	Surpickadentor	7935.01	78438.61	Cereals	8.30
65547	Incadupower	99768.92	7841.19	Seafood	40.11
65548	Enddudax	63020.17	23346.67	Shell fish	.56
65549	Grocadicazz	61299.82	33696.85	Cereals	8.30
65550	Dopniponex	67493.28	29231.76	Meat	11.25
65551	Tupsapicor	74143.39	84913.29	Confections	15.33
65552	Emsipor	89253.06	58971.44	Grain	49.26

рис. 9 Звіт про Продукти, їх категорію та податок

Висновок

У межах виконання курсової роботи було реалізовано комплексне рішення для зберігання, обробки та аналізу даних. Основні результати включають:

Розробка бази даних та сховища даних:

- Створено реляційну базу даних для структурованого зберігання інформації про проекти, працівників та фінансові операції.
- Побудоване сховище даних забезпечує централізоване та оптимізоване зберігання інтегрованих даних для подальшого аналізу.

Реалізація ETL-процесу:

- Впроваджено процес, який автоматизує екстракцію даних із бази даних, їх трансформацію та завантаження у сховище. Це дозволило підготувати дані до подальшого аналізу, забезпечивши їхню цілісність та якість.

Створення OLAP-куба:

- На основі даних зі сховища побудовано OLAP-куб, який організовує дані у багатовимірній структурі. Це спростило виконання складних аналітичних запитів, таких як аналіз за часом, проектами чи працівниками.
- Реалізація OLAP-куба дала змогу швидко виконувати агрегацію метрик та аналізувати інформацію з різних перспектив.

Таким чином, розроблена система дозволяє ефективно зберігати, обробляти та аналізувати дані, підтримуючи виконання управлінських рішень та бізнес-аналітики. Впроваджені компоненти демонструють можливості інтеграції бази даних, сховища, ETL-процесів та OLAP-технологій у єдиній інформаційній системі.

Список використаної літератури:

1. SQL Server Integration Services

<https://learn.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver16>

2. OLAP Cube Definition

<https://olap.com/learn-bi-olap/olap-bi-definitions/olap-cube>

3. Business Intelligence:

<https://www.tableau.com/business-intelligence/what-is-business-intelligence>

4. Multidimensional Model Databases (SSAS)

<https://learn.microsoft.com/en-us/analysis-services/multidimensional-models/multidimensional-model-databases-ssas?view=asallproducts-allversions>

5. SQL Server Reporting Services (SSRS)

<https://learn.microsoft.com/en-us/sql/reporting-services/create-deploy-andmanage-mobile-and-paginated-reports?view=sql-server-ver16>