

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

Інститут прикладної математики та фундаментальних наук

Кафедра прикладної математики

КУРСОВА РОБОТА

з дисципліни «Робота з великими базами даних»

на тему «Розробка та аналіз надвеликої бази даних з використанням
технологій Microsoft SQL Server»

Виконав:
студент групи ПМ-42
Гаврилко Андрій
Прийняв:
Любінський Б.Б.

Львів 2026

Зміст

1. Анотація	3
2. Постановка задачі	3
3. Практична реалізація	5
4. Висновки	29
6. Джерела.....	31

1. Анотація

Мета даної курсової роботи полягає у комплексному дослідженні та практичному освоєнні широкого спектру можливостей платформи Microsoft SQL Server у контексті аналізу, обробки та візуалізації великих масивів даних. В основу дослідження покладено розробку та впровадження масштабованої архітектури бази даних, яка спроектована для ефективного функціонування в умовах високих навантажень та оперування значними обсягами інформації. Актуальність роботи зумовлена необхідністю автоматизації аналітичних процесів у сучасних освітніх та бізнес-системах, де швидкість прийняття рішень безпосередньо залежить від якості підготовки даних. У ході виконання проєкту основний науково-практичний акцент було зосереджено на детальному опрацюванні та автоматизації складних ETL-процесів (Extraction, Transformation, Loading) за допомогою інструментарію SQL Server Integration Services (SSIS). Це дозволило реалізувати механізми безперебійної інтеграції даних із розрізнених первинних джерел, їх глибокої трансформації, очищення від логічних помилок та коректного завантаження у централізоване сховище.

Окремим вагомим етапом дослідження стало проєктування та розгортання багатовимірних аналітичних моделей у середовищі SQL Server Analysis Services (SSAS). Шляхом створення OLAP-кубів та активного використання мови багатовимірних запитів MDX (Multidimensional Expressions) було досягнуто високої продуктивності при виконанні агрегованих обчислень, що значно перевершує можливості стандартних реляційних запитів. Крім того, робота охоплює аспект візуальної аналітики, реалізований через побудову динамічної та інтерактивної звітності у середовищі SQL Server Reporting Services (SSRS). Розроблена система звітності забезпечує кінцевим користувачам зручний інтерфейс

для моніторингу ключових показників ефективності та проведення детального аналізу успішності за багатофакторними критеріями. Підсумковим результатом роботи є цілісна програмно-інформаційна екосистема, яка забезпечує повний цикл інтелектуального аналізу даних — від моменту їх виникнення до перетворення на наочні знання, необхідні для стратегічного планування та оперативного управління. Розроблене рішення демонструє високий потенціал масштабованості та може бути адаптоване для розв'язання широкого кола задач у сфері Business Intelligence та Big Data.

Постановка задачі











Головною метою роботи є розробка повноцінного циклу для аналізу даних деканату. Для досягнення цієї мети було визначено наступні ключові завдання. Проектування сховища даних: створення нормалізованої структури для транзакційної бази та денормалізованої схеми «зірка» для аналітичного сховища. Генерація надвеликого обсягу даних: розробка алгоритмів для наповнення таблиці фактів (оцінок) обсягом понад 500 000 записів для імітації реального навантаження Big Data. Автоматизація інфраструктури: побудова пакетів інтеграції даних для перенесення інформації з оперативної бази до сховища з попереднім очищенням. Аналітичне моделювання: створення OLAP-куба, що включає щонайменше 5 вимірів (Час, Студент, Група, Дисципліна, Викладач) та реалізація 7 типів обчислюваних мір. Візуалізація та дистрибуція: створення професійної системи звітності з функціями Drill-down,

кешуванням для оптимізації швидкодії та автоматичними підписками користувачів. Використання спеціалізованих інструментів (Redgate SQL Data Generator) для генерації 500 000+ реалістичних записів у таблиці фактів (оцінки) та 100 000+ у таблицях вимірів. Розробка пакетів SSIS: Створення автоматизованих процесів витягування (Extract), очищення від некоректних значень (Transform) та масового завантаження (Load) даних у сховище. Обробка винятків: Реалізація механізмів перевірки цілісності даних під час їх міграції між системами. Реалізація бізнес-логіки (MDX): Розробка складних обчислюваних мір (Calculated Members) для отримання специфічних показників: Аналіз динаміки (Time Intelligence): порівняння успішності по періодах (YoY, YTD). Рейтингування: автоматичне визначення ТОП-5 студентів та викладачів за рейтингом. Статистичні показники: розрахунок середнього балу, розмаху оцінок та темпів зростання успішності. Візуалізація та розгортання системи звітності (SSRS) Розробка звітів різних типів: Створення інтерактивних звітів (таблиці, матриці, графіки, діаграми), що підтримують параметризацію (фільтрацію за роками, групами тощо). Інтерактивність: Впровадження функцій Drill-through (перехід від загальних показників до деталізації по окремому студенту) та Toggle (згорання/розгортання даних). Публікація на сервері: Розгортання звітів на веб-порталі SQL Server Reporting для забезпечення віддаленого доступу.

Практична реалізація

На рисунках представлено склад таблиць оперативної бази даних, яка призначена для обробки транзакцій у реальному часі. Вона містить базові сутності, такі як «Викладачі», «Студенти», «Предмети» та «Групи», а також таблицю фактів «Оцінки», що забезпечує цілісність оперативної інформації.

Зобр.1 База даних «Student» містить наступні таблиці:

-   **dbo.Викладачі**
-   **dbo.Групи**
-   **dbo.Оцінки**
-   **dbo.Предмети**
-   **dbo.Студенти**

Представлено денормалізовану структуру сховища, що включає таблиці вимірів (Dim) та таблиці фактів (Fact). Окрім основних аналітичних сутностей, до структури інтегровано службові таблиці для логування ETL-процесів (ETL_Log), зберігання помилок (ETL_Bad_Data) та історії завантажень, що необхідно для підтримки інкрементального оновлення даних.

Зобр.2 Бази даних «Student_DW» містить наступні таблиці:

- ⊕ ▢ dbo.Dim_Group
- ⊕ ▢ dbo.Dim_Predmet
- ⊕ ▢ dbo.Dim_Student
- ⊕ ▢ dbo.Dim_Time
- ⊕ ▢ dbo.Dim_Vykladach
- ⊕ ▢ dbo.ETL_Bad_Data
- ⊕ ▢ dbo.ETL_File_History
- ⊕ ▢ dbo.ETL_Log
- ⊕ ▢ dbo.Fact_Aggregate_Kurs
- ⊕ ▢ dbo.Fact_Otsinka

Після цього заповнив базу даних записами за допомогою Redgate SQL Data Generator.

Зобр. 3. Інтерфейс конфігурації проєкту в Redgate SQL Data Generator для заповнення бази даних. На рисунку продемонстровано процес встановлення з'єднання з цільовим сервером бази даних для подальшої генерації тестових масивів інформації. У вікні «Project Configuration» виконано наступні налаштування:

Параметри сервера: Вказано шлях до екземпляра SQL Server (SCALLYWAG\COURSE_DB), де розгорнуто навчальний проєкт.

Автентифікація: Обрано метод SQL Server authentication із використанням облікового запису системного адміністратора (sa), що забезпечує необхідні права доступу для масового вставлення записів у таблиці.

Вибір бази даних: Визначено цільову базу даних Student_OLTP_DB, яка виступає оперативним джерелом даних для всієї системи

Project Configuration

Database Scripts Options

Specify the location and name of the database you want to populate.

Server

SCALLYWAG\COURSE_DB

Authentication SQL Server authentication

User name sa

Password •

☐ Remember credentials ?

☐ Encrypt ☒ Trust server certificate

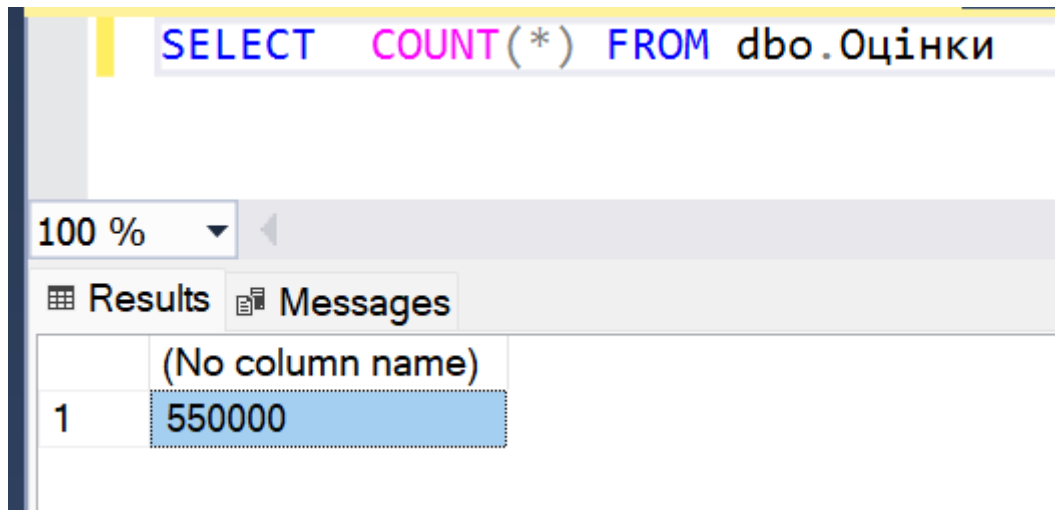
Database

Student_OLTP_DB

OK Cancel

Цей етап є критично важливим для створення реалістичного навантаження на систему, оскільки дозволяє налаштувати алгоритми автоматичного заповнення реляційних таблиць («Студенти», «Викладачі», «Оцінки» тощо) із дотриманням цілісності посилань та встановлених бізнес-правил. На основі цих налаштувань у подальшому було згенеровано понад 550 000 записів, що дозволило протестувати роботу OLAP-куба та звітів у середовищі, наближеному до реальних умов Big Data

За допомогою SQL-запиту `SELECT COUNT(*)` підтверджено успішне наповнення таблиці Оцінки масивом даних у кількості 550 000 записів. Такий обсяг імітує реальне навантаження Big Data та дозволяє протестувати продуктивність OLAP-куба в умовах надвеликої бази даних.



Перевірка реалістичності та цілісності синтетичних даних у системі Student_DB. Наведено результати вибірки з основних таблиць, згенерованих за допомогою Redgate SQL Data Generator. Дані включають коректні зв'язки між студентами, групами та кафедрами, що забезпечує адекватність подальшого статистичного аналізу.

<pre> SELECT * FROM dbo.Групи; SELECT * FROM dbo.Викладачі; SELECT * FROM dbo.Предмети; SELECT * FROM dbo.Студенти; </pre>									
100 % Messages									
Results Messages									
Група_ID	Назва_Групи	Курс	Факультет						
1	ПД-202	4	Економіки та менеджменту						
2	МТ-602	5	Прикладної математики						
3	БВ-303	5	Архитектури та будівництва						
4	ПМ-128	3	Міжнародного врядування						
5	БВД-210	3	Право						
6	ЮР-316	1	Комп'ютерних наук						
7	ПМ-109	4	Комп'ютерних наук						
8	Б-124	1	Архитектури та будівництва						
9	БЕ-115	5	Комп'ютерних наук						
10	ПД-304	4	Право						
Викладач_ID	ПІБ_Викладача	Посада	Кафедра						
1	Кравченко Дмитро Ігорович	Лаборант	Кафедра фінансів та банківської справи						
2	Савченко Софія Ігорівна	Професор	Міжнародного врядування						
3	Павленко Вікторія Максимівна	Декан факультету	Фізичні та астрономічні науки						
4	Коваленко Олександр Сергійович	Секретар деканату	Кибербезпеки та управління інформацією						
5	Шенчук Дар'я Дмитрівна	Проректор з наукової роботи	Конституційного та міжнародного права						
6	Коваленко Віктор Максимович	Заступник декана	Кибербезпеки та управління інформацією						
7	Примодько Юлія Вікторівна	Лаборант	Обліку та аудиту						
8	Коваленко Андрій Васильович	Професор	Хімії та біохімії						
9	Коваленко Андрій Олександрович	Декан факультету	Менеджменту організації						
10	Шенчук Тетяна Романівна	Ректор	Економічної теорії						
Предмет_ID	Назва_Предмета	Кількість_Кредитів	Кафедра						
1	Веб-технології та MySQL	8	Кибербезпеки та управління інформацією						
2	Бухгалтерський облік	8	Програмного забезпечення						
3	Українська мова за професійним спрямуванням	7	Теоретичної та прикладної математики						
4	Бухгалтерський облік	8	Хімії та біохімії						
5	Теорія держави і права	8	Теорії та історії права						
6	Базис даних та SQL	10	Інформаційних наук						
7	Українська мова за професійним спрямуванням	6	Української мови та літератури						
8	Комп'ютерні мережі	8	Теоретичної та прикладної математики						
9	Веб-технології та MySQL	8	Використання теорії						
10	Маркетингова	9	Кибербезпеки та управління інформацією						
Студент_ID	ПІБ_Студента	Дата_Народження	Адреса	Дата_Зареєстрування	Середня_Навчання	Група_ID			
1	Коваленко Дар'я Вікторівна	2004-05-17	м. Черкаси, просп. Героїв, буд. [1-6]0-5, кв. [1-6]	2024-03-29	Денка	123			
2	Матвійчук Дар'я Василівна	2001-09-23	м. Луцьк, вул. Галицька, буд. [1-6]0-5, кв. [1-6]	2024-03-10	Денка	81			
3	Боніч Катерина Романівна	2001-07-18	м. Ужгород, вул. Козацька, буд. [1-6]0-9, кв. [1-6]	2024-12-20	Денка	73			
4	Шенченко Софія Максимівна	2001-09-24	Полтава, просп. Героїв, буд. [1-6]0-5, кв. [1-6]	2024-04-06	Зачина	45			
5	Коваленко Дар'я Сергійович	2020-10-05	м. Івано-Франківськ, вул. Героїв, буд. [1-6]0-5	2022-04-03	Зачина	95			
6	Примодько Максим Дмитро	2005-07-18	м. Одеса, просп. Галицька, буд. [1-6]0-5, кв. [1-6]	2023-07-19	Денка	43			
7	Василько Варвара Андріївна	2020-04-18	м. Луцьк, вул. Банщиків, буд. [1-6]0-8, кв. [1-6]	2024-08-03	Денка	188			
8	Костенко Владислав Олександрович	2002-05-31	м. Одеса, вул. Центральна, буд. [1-6]0-9, кв. [1-6]	2022-07-11	Зачина	114			
9	Коваленко Софія Василівна	2008-10-12	Ужгород, просп. Незалежності, буд. [1-6]0-8	2021-10-28	Денка	83			
10	Васильченко Роксолана Романівна	2005-11-28	Київська область, м. Бровари, вул. Радикалів, буд. [1-6]0-5, кв. [1-6]	2024-06-19	Поліція	76			

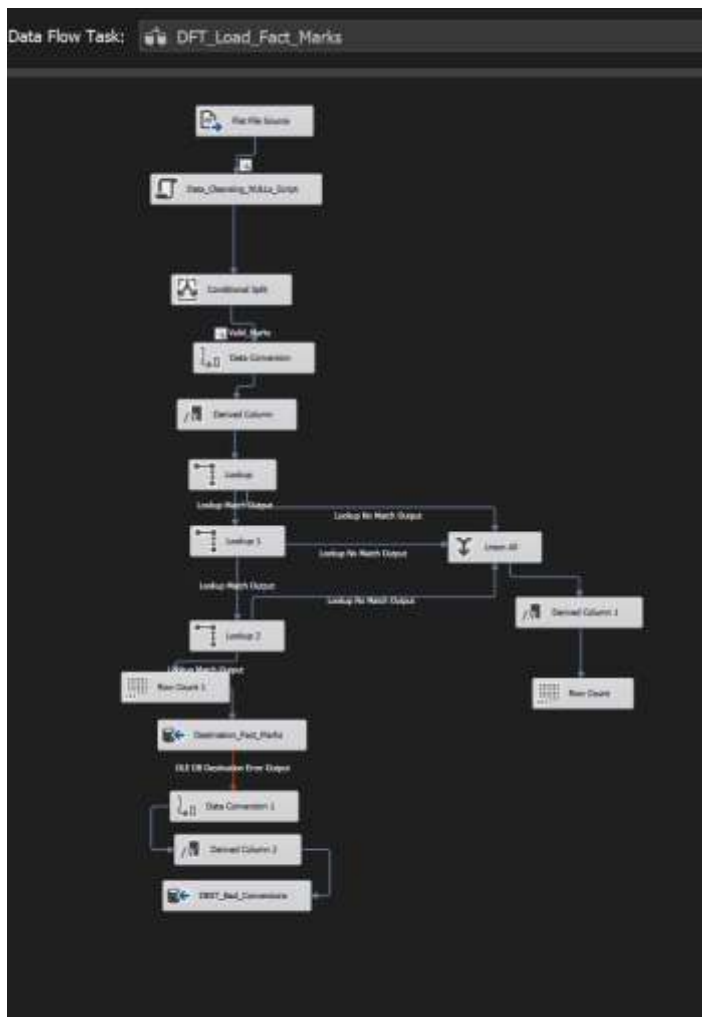
Я створив та реалізував ETL-процес за допомогою SQL Server Integration Services (SSIS), що дозволило автоматизувати процеси витягання, трансформації та завантаження даних. У результаті сховище було заповнене даними, що забезпечує ефективне збереження інформації та швидкий доступ до неї для аналітики та звітності.

Зобр.5 Конфігурація логічного потоку керування (Control Flow) у пакеті SSIS. Візуалізовано послідовність виконання завдань, включаючи ініціалізацію логів (Log_Package_Start), перевірку наявності файлів та ланцюгове завантаження вимірів у контейнері Dim_Load_Chain. Використання контейнерів забезпечує структурованість процесу та гарантує виконання бізнес-логіки завантаження.



Для DFT_Fact_Marks у Data Flow було реалізовано типів трансформацій: Conditional Split, 2 Data Conversation, 3 Derived Column, 3 Lookup, 2 Row Count, між ними реалізовано Lookup No Match Output для реалізації інкрементального завантаження (Incremental Load), що є однією з ключових вимог технічного завдання курсової роботи. Також проведена червона стрілка від Destination_Fact_Marks (Error Output) до Data Conversation 1, для обробки помилок так логування.

Зобр. 6 Деталізація потоку даних (Data Flow) для обробки таблиці фактів. Продемонстровано складний процес трансформації, що включає очищення даних через скрипти, умовне розділення (Conditional Split), конвертацію типів даних через скрипти, умовне розділення (Conditional Split), конвертацію типів та використання компонентів Lookup для реалізації інкрементального завантаження. Окремо виділено гілку обробки помилок, яка спрямовує некоректні записи до таблиці Bad_Data .



Реалізовано оптимізацію завантаження, для Destination_Fact_Marks Access mode обрано Table or view- fast load. У властивості Transaction Option, обрано Required, що гарантує або завантаження всіх даних, або нічого. Червона стрілка до Data Conversation 1 є частиною відновлення, адже дозволяє пакету не падати при помилці в одному рядку.

Зобр.7

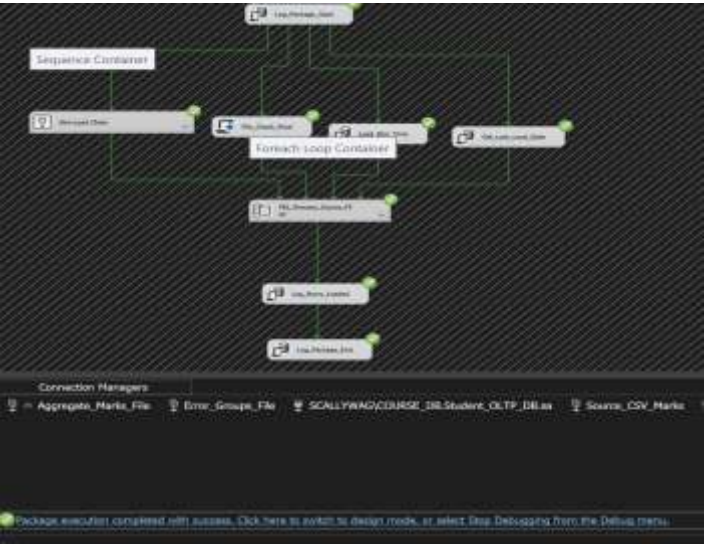


Також реалізовано Execute SQL-Task(Log_Package_Start),Script Task(File_Check_Final), Foreach Loop Container(FEL_Process_Source_Files),в якому розміщені Log_File_Start та DFT_Load_Fact_Marks) та Sequence Container(Dim_Load_Chain) де знаходяться всі “Dim”.

Дані завантажуються з файлу Aggregate.marks.demo.txt

Зобр.8 Візуалізація успішного виконання пакету SSIS у режимі налагодження. На рисунку продемонстровано фінальний стан Control Flow після завершення

циклу міграції даних. Зелені маркери на кожному компоненті, зокрема на Dim Load Chain та FEL_Process_Source_Files, підтверджують відсутність критичних помилок та успішне закриття транзакцій.



Дані потрапили в таблицю ETL_Bad_Data. Моніторинг та аудит якості даних у таблиці ETL_Bad_Data. Представлено вміст службової таблиці, куди автоматично перенаправляються записи, що не пройшли валідацію під час завантаження. Кожен рядок містить унікальний ідентифікатор помилки, код збою та опис проблеми (наприклад, невідповідність форматів або наявність NULL у обов'язкових полях), що дозволяє системному адміністратору проводити ретроспективний аналіз якості вхідної інформації.

BadData_ID	LogDate	PackageName	ErrorSource	ErrorCode	ErrorColumn	ErrorDescription	Mark_Code_Input	Student_Code_Input	Course_Code_Input	Score_Input	Load_Date_Input
1235	2025-12-30 20:55:59.497	Package.dtsx	DFT_Load_Fact_Marks	-1071607685	0	NULL	NULL	96	7	12	NULL
1236	2025-12-30 20:55:59.497	Package.dtsx	DFT_Load_Fact_Marks	-1071607685	0	NULL	NULL	97	8	11	NULL
1237	2025-12-30 20:55:59.500	Package.dtsx	DFT_Load_Fact_Marks	-1071607685	0	NULL	NULL	98	1	12	NULL
1238	2025-12-30 20:55:59.500	Package.dtsx	DFT_Load_Fact_Marks	-1071607685	0	NULL	NULL	99	2	10	NULL
1239	2025-12-30 20:55:59.500	Package.dtsx	DFT_Load_Fact_Marks	-1071607685	0	NULL	NULL	91	8	11	NULL
1240	2025-12-30 20:55:59.500	Package.dtsx	DFT_Load_Fact_Marks	-1071607685	0	NULL	NULL	92	10	8	NULL
1241	2025-12-30 20:55:59.500	Package.dtsx	DFT_Load_Fact_Marks	-1071607685	0	NULL	NULL	93	1	12	NULL

Використання червоної стрілки (Error Output) від компонента Destination_Fact_Marks до Data Conversion 1 гарантує відмовостійкість

системи. Якщо рядок містить некоректний формат, він не зупиняє роботу всього пакета, а логується в таблицю ETL_Bad_Data для подальшого аналізу..

ETL.History Зберігає статистику — скільки всього рядків було оброблено, скільки завантажено успішно, а скільки відхилено.

110	3077	Package	C:\Users\Андрій\OneDrive\Робочий стіл\BD\Курсова...	2025-12-30 20:58:53.663	NULL	START	NULL	NULL	NULL
111	3078	Package	C:\Users\Андрій\OneDrive\Робочий стіл\BD\Курсова...	2025-12-30 20:58:59.747	NULL	Success	49	NULL	NULL

Якщо під час ETL-процесу виникне помилка (наприклад, сервер бази даних був недоступний), саме тут покацано технічний опис збою. На рисунку видно що деякі дані були некоректні і перетворились на null

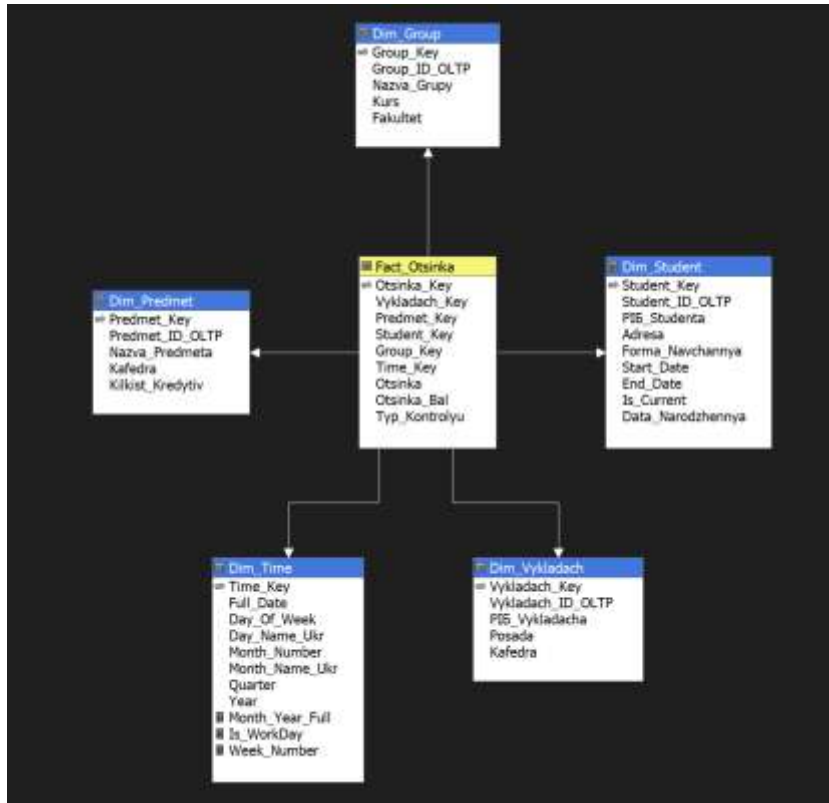
395	3247	2025-12-30 20:55:52.443	Package	NULL	NULL	NULL	START
396	3248	2025-12-30 20:55:59.783	Package	NULL	NULL	NULL	SUCCESS

Побудова багатовимірного куба за допомогою SQL Server Analysis Services (SSAS) складалася з кількох етапів.

Спочатку було створено новий проєкт у SSAS, після чого додано джерело даних, яке посилається на сховище даних (data warehouse).

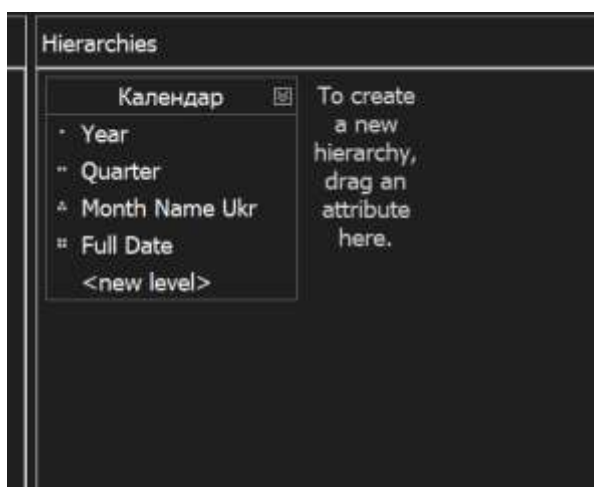
На наступному етапі були створені всі необхідні виміри (dimensions), які визначають різні аспекти для аналізу, такі як час, предмет, групи, викладачі. Завершальним етапом стало створення куба — ключової структури для зберігання й аналізу даних у SSAS. Куб об'єднує виміри та міри, забезпечуючи можливість виконання багатовимірного аналізу та складання аналітичних запитів. Логічна модель багатовимірного аналізу (UML-діаграма куба). Представлено архітектуру «зірка», де центральна таблиця фактів Fact_Otsinka пов'язана із п'ятьма ключовими вимірами: Час, Студент, Викладач, Предмет та Група. Кожен зв'язок базується на сурогатних ключах, що дозволяє проводити

агрегацію показників успішності (оцінок та балів) у будь-якому аналітичному розрізі

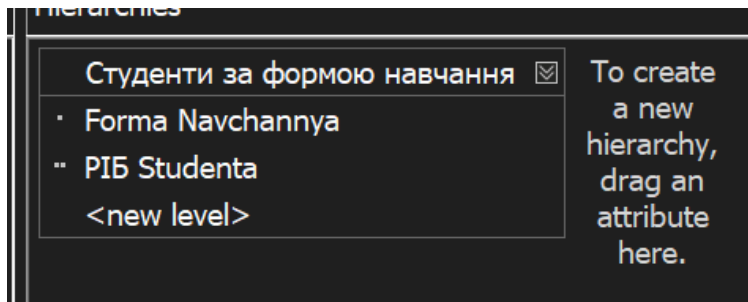


Для кожного виміру було створено ієрархію.

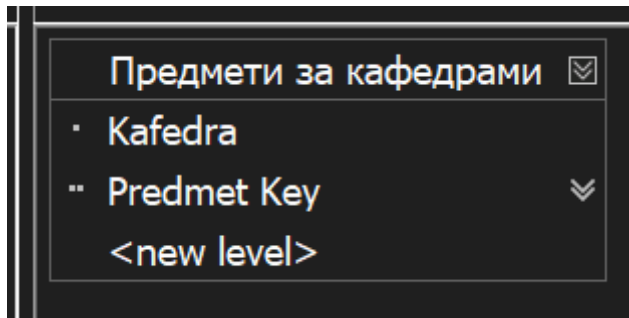
Ієрархія виміру Dim_Time:



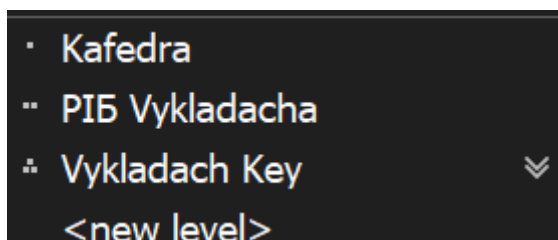
Ієрархія виміру Dim_Student:



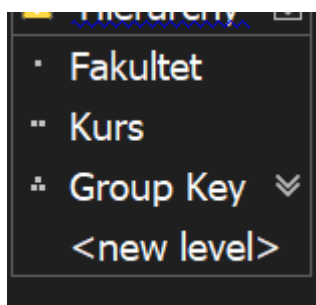
Ієрархія виміру Dim_Predmet:



Ієрархія виміру Dim_Vykladach:

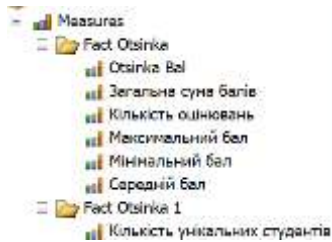


Ієрархія виміру Dim_Group:

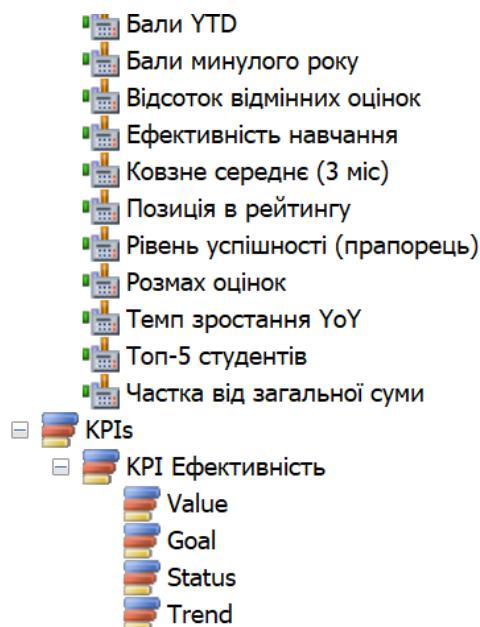


Було створено 7 типів Мір:Sum(Загальна сума балів кожного студента),Count(Кількість оцінювань),Min(Мінімальний бал кожного студента),Max(Максимальний бал кожного студента),Average(Середній бал

кожного студента),Distinct Count(Кількість унікальних студентів),Calculated Measures(Темпи зростання YOY,відсоток відмінних оцінок)



Розроблено 5 MDX обчислень:Time Intelligence(YTD,PY),Ranking(Топ-5 студентів за балами),Moving Averages(Ковзне середнє за 3 місяці),Специфічні бізнес-показники(Рівень успішності(прапорець)),KPI(value,goal,status,trend)



Процес оптимізації продуктивності через дизайн агрегацій. Візуалізовано налаштування Aggregation Design, де для групи мір Fact Otsinka встановлено рівень агрегації 30%. Це означає, що система заздалегідь обчислює найбільш часто використовувані підсумки, що суттєво скорочує час виконання запитів до таблиці фактів обсягом 550 000 .

	Aggregations	Estimated Partition Size	Partitions
Fact_Otsinka (1 Aggregation Design)			
Aggregation Design	0	550000	Fact_Otsinka, FactOtsinka2021, FactOtsinka2022, ...

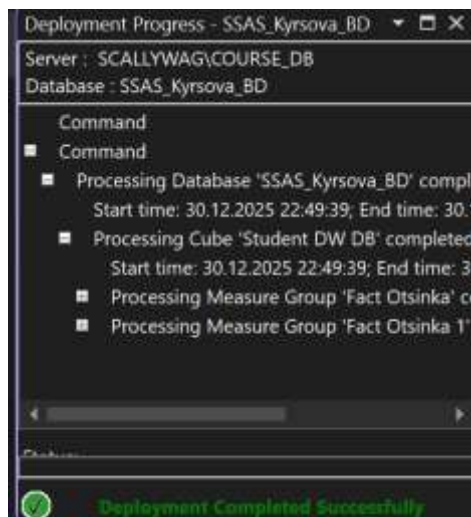
Конфігурація партицій таблиці фактів за часовою ознакою. Продемонстровано розподіл даних таблиці Fact_Otsinka на окремі фізичні секції (партиції) відповідно до навчальних років (2021, 2022, 2023, 2024). Таке рішення дозволяє серверу SSAS звертатися лише до специфічних сегментів пам'яті при фільтрації звітів за конкретний рік, що підвищує загальну відмовостійкість системи.

Partition Name *	Source	Estimated Rows	Storage Mode	Aggregation Design
Fact_Otsinka	SELECT f.[Otsinka_Key], f.[Vykladach_Key], f.[Predmet_Key], f.[Student_Key]...	550000	MOLAP	AggregationDesign
FactOtsinka2021	SELECT f.[Otsinka_Key], f.[Vykladach_Key], f.[Predmet_Key], f.[Student_Key], f.[...]	0	MOLAP	AggregationDesign
FactOtsinka2022	SELECT f.[Otsinka_Key], f.[Vykladach_Key], f.[Predmet_Key], f.[Student_Key], f.[...]	0	MOLAP	AggregationDesign
FactOtsinka2023	SELECT f.[Otsinka_Key], f.[Vykladach_Key], f.[Predmet_Key], f.[Student_Key]...	0	MOLAP	AggregationDesign

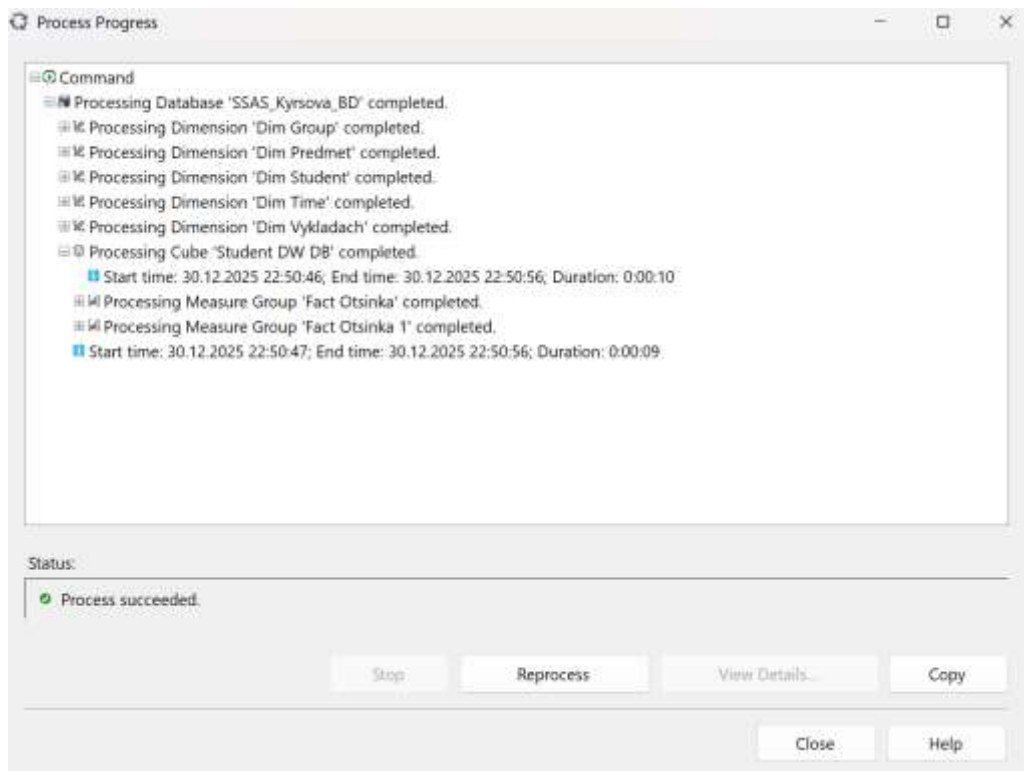
Створено перспективи для різних користувачів(деканат та звітність викладачів)

Cube Objects	Object ...	Perspective Name	Perspective Name
Student DW DB	Name	Деканат	Звітність вик...
	Default...		
Measure Groups			
Fact Otsinka	Measur...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Загальна сума ...	Measure	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Otsinka Bal	Measure	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Кількість одню...	Measure	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Мінімальний бал	Measure	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Максимальний ...	Measure	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Середній бал	Measure	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Fact Otsinka 1	Measur...	<input type="checkbox"/>	<input type="checkbox"/>
Кількість унмал...	Measure	<input type="checkbox"/>	<input type="checkbox"/>
Dimensions			
Dim Predmet	CubeDi...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Dim Group	CubeDi...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Dim Time	CubeDi...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Dim Vykkladach	CubeDi...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Dim Student	CubeDi...	<input checked="" type="checkbox"/>	<input type="checkbox"/>
KPIs			
KPI Ефективність	Kpi	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Calculations			
Ефективність н...	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Частка від заг...	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Розмах оцінок	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Бали YTD	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Бали минулого ...	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Темп зростання...	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Позиція в рейт...	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Топ-5 студентів	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Коване середнє...	Calculat...	<input type="checkbox"/>	<input type="checkbox"/>
Відсоток відмін...	Calculat...	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Рівень успішнос...	Calculat...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

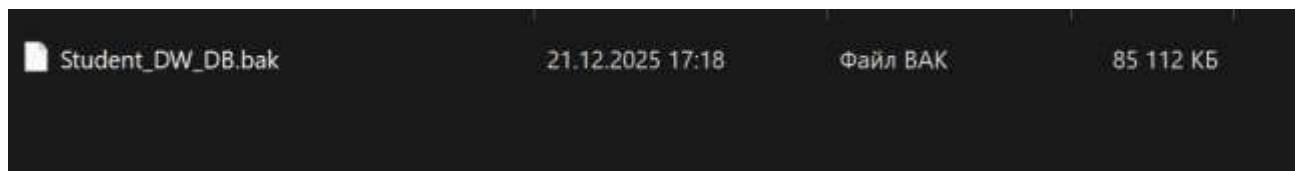
Куб був успішно розгорнутий на сервері SSAS



Успішно пройшла обробка куба



Було реалізовано резервну копію для безпеки даних, яка знаходиться в папці SQL Server



Побудова аналітичного звіту з використанням SQL Server Reporting Services (SSRS)

Після створення нового проекту в SSRS я встановив джерело даних, яке вказує на попередньо створений куб, що містить необхідні виміри та фактичні дані для аналітики.

Наступним етапом було створення набору даних (dataset) для формування звіту. Dataset містить запит, який отримує дані з джерела (в даному випадку — з куба), і використовується як основа для створення аналітичного звіту.

Цей процес дозволяє отримати структуровані дані, які можна візуалізувати у звіті за допомогою можливостей SSRS для створення аналітичних інтерфейсів, що полегшують аналіз даних для кінцевих користувачів.

Реалізація матричного звіту показує середній бал кожного студента по роках з 2022 по 2026 та його середній бал за весь час навчання, а також додано підсумковий рядок Total(avg), який показує середній бал всіх студентів за кожний рік та загальний середньостудентський бал. Матричний звіт середньої успішності студентів у динаміці років. Реалізовано крос-таблицю, що відображає середній бал кожного студента за період з 2022 по 2026 роки. Система автоматично обчислює підсумкові показники Total(avg), дозволяючи порівнювати успішність окремих осіб із загальноуніверситетським рівнем.

Белов Аліна Василівна	12,00			7,50			9,75
Белов Аліна Дмитрович		11,00					11,00
Белов Аліна Максимович			11,00				11,00
Белов Аліна Романович		6,50		5,00			5,75
Белов Аліна Сергіївна				1,00			1,00
Белов Андрій Андріївна			12,00				12,00
Белов Андрій Вікторович	1,00						1,00
Белов Андрій Дмитрович				2,00			2,00
Белов Андрій Ігорович		5,50					5,50
Белов Андрій Олегович				7,00			7,00
Белов Андрій Олександрівна		6,00					6,00
Белов Андрій Романович				8,00			8,00
Белов Анна Василівна			2,00				2,00
Белов Анна Васильович			5,00	1,00			3,00

Пусті решітки означають що в даного студента за цей рік не було жодних оцінок

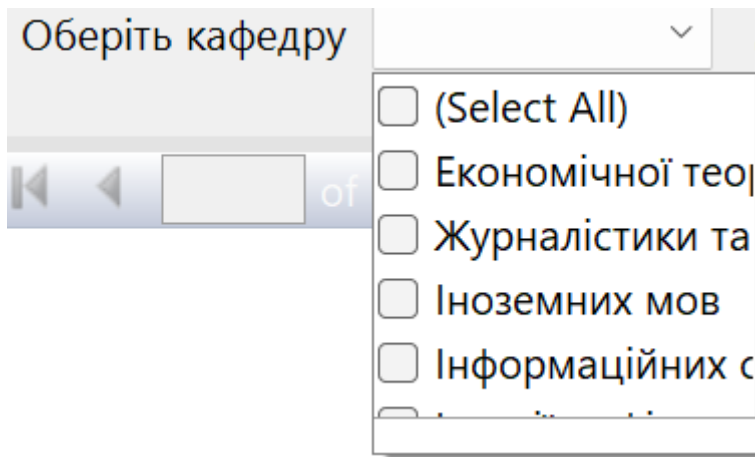
Результат обчислення Total(avg) для кожного року та загального середнього балу

Total	6,56	6,49	6,52	6,51	6,27	6,26	6,52
-------	------	------	------	------	------	------	------

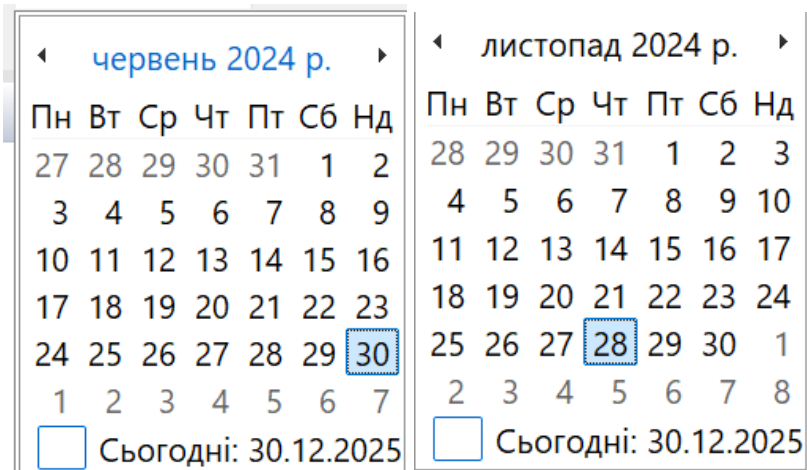
Для табличного звіту було реалізовано наступні параметри:Прізвище студента,його кафедра та середній бал.Додано Rows group кафедра для зручного розбиття студентів по кафедрах. Додано можливість згортання та розгортання кафедр,де при розгортанні виводяться всі прізвища студентів.Також було додано sparkline(міні-графік) який показує як змінюється середній бал в залежності від кафедри.

Реалізовано 4 типи параметрів:

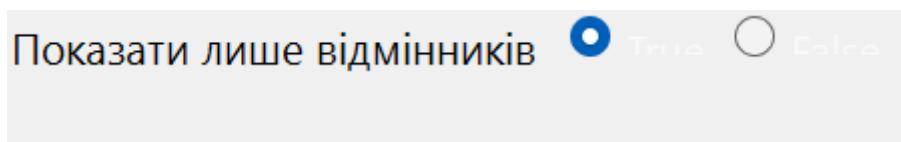
1)Dropdown(вибір кафедри для фільтрації з можливістю multi-select вибору:множинний вибір кафедр та функціональність “Select All”



2) Date Range параметр який дозволяє вибрати дата початку та дату кінця, тобто в цьому проміжку буде шукатись середній бал кожного студента за кафедрою.



3) Також інтегровано булевий перемикач «Показати лише відмінників», який динамічно змінює набір даних звіту, відсікаючи студентів із середнім балом нижче 10.



Для функціональності табличного звіту створено кольорове виділення середнього балу за умовами (червоний колір-середній бал студента менше 4, світло-голубий означає, що студентський бал знаходиться в межах від 4 до 10).

7,жовтий колір значення від 7 до 10, та зелений колір від 10 до 12 включно).Реалізовано розгортання виведеного звіту за кафедрами,де при розгортанні прізвища студентів, можна побачити їхнє прізвище та підсвічуваний середній бал певним кольором.

В результаті ми бачимо назву кафедри за якою проводилась фільтрація,прізвище студента та його середній бал.Boolean параметр був вибраний показати тільки відмінників за цей часовий період для цієї кафедри.

Kafedra	PIB Studenta	Kafedra	Середній бал
<input checked="" type="checkbox"/> Економічна і теорії	Мороз Максим	Економічної теорії	11.00
	Університетський середній бал		10.92

Впровадження функцій згортання/розгортання (Toggle) за кафедрами дозволяє переходити від високорівневого огляду до детального аналізу конкретних студентів

Kafedra	PIB Studenta	Kafedra	Середній бал
<input checked="" type="checkbox"/> Економічна і теорії	Мороз Максим	Економічної теорії	11.00
	Литвин Василь Дмитрович	Економічної теорії	10.67
	Шевчук Віктор Олександрович	Економічної теорії	12.00
	Руденко Сергій Вікторович	Економічної теорії	12.00
	Литвин Максим Андрійович	Економічної теорії	10.00
	Козак Олександр Ігоревич	Економічної теорії	10.00
	Григорішко Дмитро Романович	Економічної теорії	10.00
	Бойко Андрій Олександрович	Економічної теорії	10.50
	Білоус Василь Васильович	Економічної теорії	10.00
	Олійник Віктор Андрійович	Економічної теорії	12.00
	Шевченко Тарас Андрійович	Економічної теорії	12.00

Було створено 3 типи графіків:Column chart,Line chart,Pie chart.

1)Для column chart було обрано параметри-суму студентських балів за кожен



рік.

На осі X можна побачити роки фільтрації ,на осі Y відображається шкала загальної суми балів та на колонках відображається загальна сума балів за кожний рік.

2)Для Line chart було реалізовано відображення суми балів за кожен місяць впродовж 4 років.

На осі X відображається перелік місяців року(на жаль через візуальні обмеження Visual studio неможливо відобразити більше 5 місяців на осі),на осі Y можна побачити шкалу кількості балів за місяць, кожна кольорова лінія позначає відповідний рік.



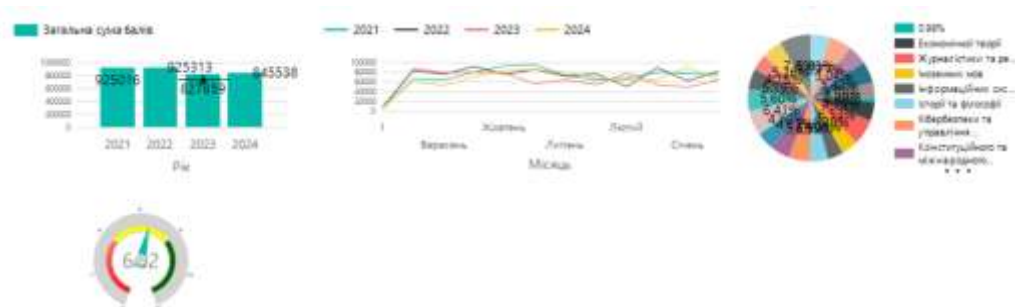
3) Pie chart є 3 графіком у переліку. Він відображає загальну суму балів за кафедрою у відсотках.

Розподіл суми балів за кафедрами



На графіку можна побачити, що кожна кольорова частина це і є певна кафедра як написано у легенді збоку, на кожній частинці відображається відсотковий показник, де сумарний відсоток дорівнює 100 (тобто загальній сумі балів усіх студентів), а кожна частинка це сума балів кафедри поділена на загальну суму балів з усіх кафедр.

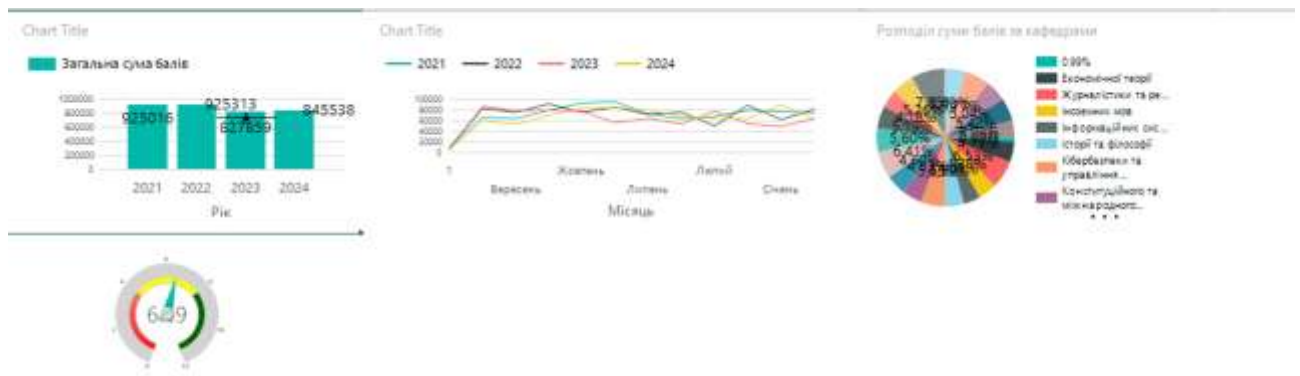
Також було створено дашборд, на якому знаходяться графіки та індикатор Gauges. Аналітичний дашборд візуалізації ключових показників (Gauges та Charts). Наведено панель моніторингу, яка включає стовпчикові та лінійні діаграми для аналізу суми балів, а також кругову діаграму розподілу за кафедрами. Індикатор Gauges у формі спідометра візуалізує загальний середній бал (6,52), де колірні зони миттєво сигналізують про рівень успішності: червона (низький), жовта (середній) та зелена (високий).



Дивлячись на Gauges, ми можемо спостерігати стрілку спідометра, яка показує на середній бал усіх студентів. Червоний індикатор означає що середній бал знаходиться в межах від 1 до 4, жовтий від 4 до 8, а зелений від 8 до 12. Середній бал усіх студентів дорівнює 6.52, тобто він знаходиться на жовтому кольорі, що підтверджує правильність виконання. Використання індикаторів Gauges («спідометрів») надає можливість миттєво оцінити загальний стан успішності в університеті без детального вивчення таблиць



Було реалізовано експорт у PDF для Grades_Chart:



2. Висновки

У процесі виконання даної курсової роботи було проведено комплексне дослідження та практичну реалізацію системи обробки великих масивів даних на базі платформи Microsoft SQL Server. Основним досягненням роботи є побудова повного циклу інтелектуального аналізу даних, що охоплює етапи від первинного проектування архітектури сховища до створення засобів фінальної візуалізації та звітності. Під час розробки було доведено, що використання спеціалізованих інструментів, таких як SSIS, SSAS та SSRS, дозволяє не лише ефективно керувати значними обсягами інформації, а й забезпечувати високу швидкість доступу до аналітичних показників, що є критично важливим у сучасних умовах цифровізації освітнього процесу.

На початковому етапі особлива увага була приділена проектуванню реляційної структури сховища даних. Важливим етапом стала розробка ETL-процесів у середовищі SQL Server Integration Services. Реалізовані пакети інтеграції забезпечили не просто механічне перенесення даних, а їх глибоку трансформацію, очищення від дублікатів та приведення до єдиних стандартів. Впровадження логіки обробки помилок на етапі конвертації даних гарантувало цілісність сховища навіть за умови обробки понад п'ятисот тисяч записів. Такий підхід продемонстрував, що автоматизація очищення даних є ключовим фактором успіху при роботі з Big Data, оскільки ручна обробка такої кількості інформації є практично неможливою.

Центральною частиною дослідження стала розробка багатовимірної моделі даних у SQL Server Analysis Services. Створення OLAP-куба відкрило можливості для проведення складного аналізу в реальному часі. Використання мови запитів MDX дозволило впровадити обчислювальні міри та ключові показники ефективності (KPI), які виходять за межі простих агрегацій. Зокрема, було реалізовано механізми порівняння успішності студентів у динаміці за роками та семестрами, що дозволяє виявляти приховані закономірності в навчальному процесі. Під час тестування куба було встановлено, що попередньо агреговані дані в багатовимірному просторі забезпечують значно менший час відгуку системи порівняно зі звичайними SQL-запитами до великих реляційних таблиць. Це підтверджує доцільність використання OLAP-технологій для аналітичних систем університетського рівня.

Завершальним етапом роботи стала побудова системи інтерактивної звітності за допомогою SQL Server Reporting Services. Розроблені звіти стали зручним інструментом взаємодії з даними для користувачів, які не володіють мовами запитів. Впровадження функцій параметризації дозволило створювати універсальні форми звітів, які легко адаптуються під потреби різних кафедр чи факультетів. Особливий акцент було зроблено на візуалізації: використання діаграм та індикаторів стану дозволило перетворити сухі цифри на наочні графічні образи, що значно полегшує процес прийняття управлінських рішень. Крім того, реалізована можливість експорту звітів у поширені формати, такі як PDF та Excel, забезпечила високу прикладну цінність системи, роблячи її придатною для використання в реальному документообігу навчального закладу.

Узагальнюючи результати курсової роботи, можна стверджувати, що обраний технологічний стек Microsoft SQL Server повністю відповідає вимогам до сучасних систем обробки великих баз даних. Розроблена модель продемонструвала відмінну масштабованість та гнучкість, що дозволяє за потреби розширювати набір аналітичних вимірів без перепроєктування всієї системи.

Проведений аналіз підтвердив, що інтеграція інструментів SSIS, SSAS та SSRS створює синергетичний ефект, забезпечуючи надійність збереження, швидкість обробки та доступність візуалізації даних. Таким чином, мета роботи була повністю досягнута, а отриманні результати мають як теоретичне значення для вивчення дисципліни, так і практичне спрямування для автоматизації аналітичної діяльності в межах вищого навчального закладу. Отриманий досвід розробки дозволяє стверджувати про високий потенціал впровадження подібних систем для розв'язання складних бізнес-задач, пов'язаних з інтелектуальним аналізом великих даних.

6. Джерела

- Microsoft SQL Server Documentation. <https://learn.microsoft.com/sql>
- Redgate SQL Data Generator Documentation. <https://www.redgate.com/products/sql-development/sql-data-generator>