

ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

Харківський національний університет радіоелектроніки
Кафедра ЕОМ

Модель OLAP-гіперкуба системи аналізу банківських операцій

Кваліфікаційна робота
Другий (магістерський рівень)

Виконав:
ст. гр. СПМ-21-1
Лифар Д.С.

Керівник:
Мартовицький В.О.
доц. каф. ЕОМ

Мета

□ Модель системи аналізу банківських операцій:

Вимоги:

- обробка та валідація вхідних даних;
- трансформація даних у аналітичну форму;
- масштабованість системи;
- розширення даних додатковою інформацією.

□ Реалізація описаної моделі:

- розробка структури БД;
- розробка об'єктів OLAP-гіперкуба;
- проектування та створення процесу потоку даних (ETL).

□ Тестування продуктивності:

- генерація та завантаження тестових даних;
- вимірювання часу виконання запитів даних у різних умовах.

Концепція багатовимірного аналізу (OLAP – On-Line Analytical Processing)

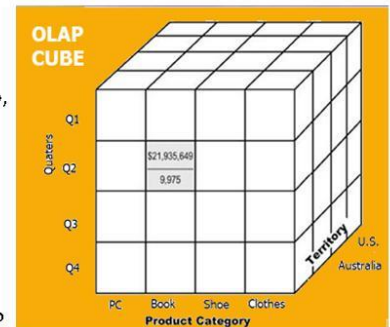
Куб – це сукупність об'єктів багатовимірного аналізу;
структура даних, яка оптимізована для швидкого аналізу даних.

Вимір – це розріз даних, наприклад «Квартал», «Категорія продукту», «Країна», тощо.

Перетин вимірів – це агреговані значення, метрики.

Приклади питань, які може вирішити багатовимірний аналіз:

- Яка сумарна вартість замовлень, зроблених у Австралії?
- Яка сумарна вартість замовлень, зроблених у США в 2020 році?
- Яка середня вартість замовлень товару «Кавун» в Україні в 2019 році?



3

Інструменти та принципи

Microsoft SQL Server

Система управління базами даних від Microsoft.
Має Transact-SQL – процедурне розширення мови SQL.

Microsoft SQL Server Analysis Services (SSAS)

Набір інструментів для роботи з OLAP-технологіями.
Інтеграція з іншими продуктами Microsoft (SQL Server не виключення).

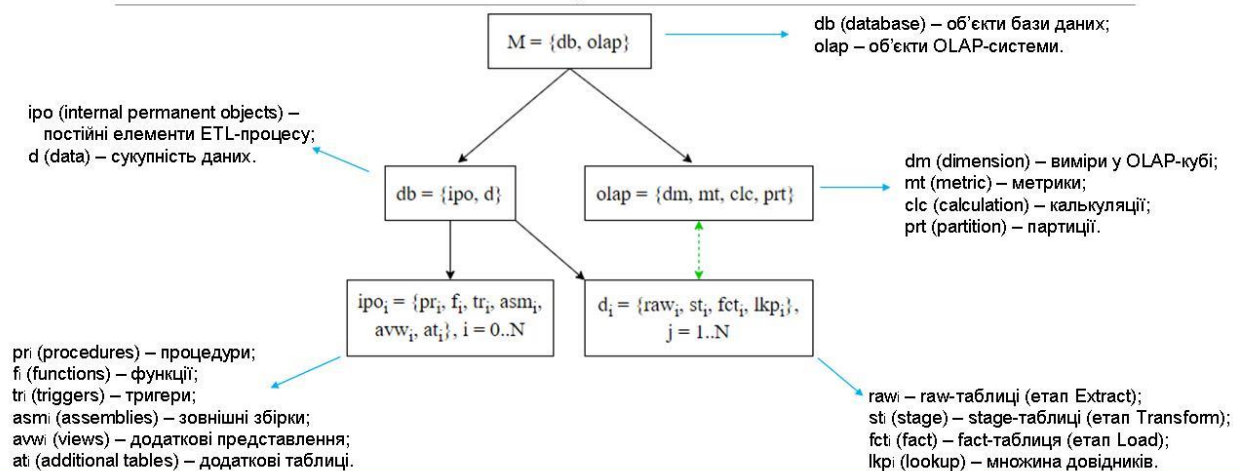
ETL-процес

Принцип завантаження даних:

- Extract (випущення);
- Transform (трансформація);
- Load (завантаження).

4

Модель системи аналізу банківських операцій



5

«іро» складова – ETL-процес

Extract:

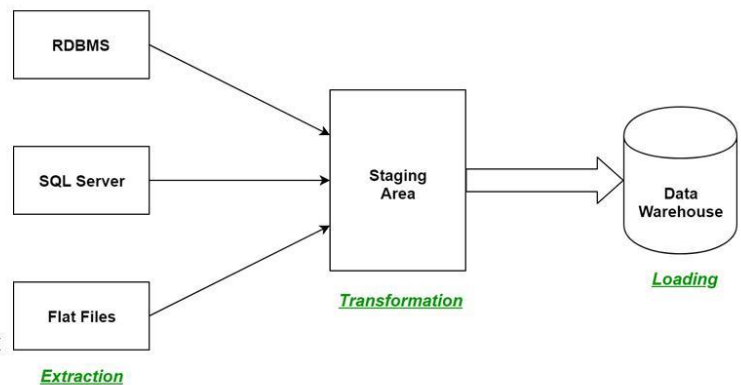
- завантаження файлів;
- перевірка рядків та чек-сум;
- логування вхідних файлів.

Transform:

- фільтрування та перевірки даних;
- очищення даних;
- стандартизація даних;
- встановлення ключів;
- поповнення довідниковими даними.

Load:

- синхронізація рядків / визначення нових даних;
- завантаження порції даних.



6

Приклад роботи з OLAP у Microsoft Excel

Названия строк	Count of Accounts	Count of Accounts as %	Current Balance	Current Balance as %
Austria	22 414	3,54%	556 439 996,10	3,52%
Bulgaria	14 873	2,35%	373 285 446,53	2,36%
Czech Republic	14 594	2,31%	364 065 016,25	2,30%
France	33 525	5,30%	841 908 382,17	5,32%
Germany	91 045	14,38%	2 281 405 353,82	14,41%
Hungary	19 469	3,08%	488 507 647,94	3,09%
Italy	66 829	10,56%	1 677 477 458,69	10,60%
Netherlands	10 170	1,61%	256 627 669,24	1,62%
Poland	29 997	4,74%	750 418 572,90	4,74%
Romania	24 610	3,89%	617 570 033,94	3,90%
Spain	56 215	8,88%	1 398 852 010,82	8,84%
Sweden	11 078	1,75%	276 118 077,24	1,74%
Ukraine	238 123	37,62%	5 946 319 382,21	37,57%
Общий итог	632 942	100,00%	15 828 995 047,85	100,00%

Поля сводной табл...

Показать поля: Accounts

- Registration Duration as %
- Tax Amount
- Tax Amount as %
- Total Txn Amount
- Total Txn Amount as %
- Total TXNS Count
- Total TXNS Count as %
- Transaction Duration
- Account Closing Date
- Account Closing Date
- Account Suffix
- Account Suffix
- Alpha Split First Name
- Alpha Split First Name.Alpha...

Перетащите поля в нужную область:

ФИЛЬТРЫ: Currency

КОЛОННЫ: Значения

СТРОКИ: Location

ЗНАЧЕНИЯ: Count of Acc..., Current Balance, Current Balan...

Отложить обновления... ОБНОВИТЬ

Тестування продуктивності моделі

1. Запит до даних у SQL Server (мова SQL).

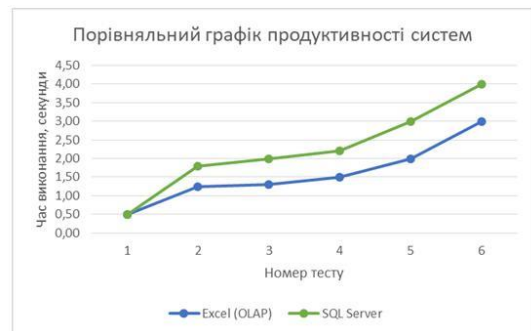
```

1 SELECT COUNT(*), SUM(CurrentBalance),
2 SUM(CreditAmount), AVG(Age), AVG(TotalTXNSCount)
3 FROM vwFactAccounts f
4 JOIN vw_dicCity c ON f.City_ID = c.ID
5 JOIN vw_dicAccountSuffix a ON f.AccountSuffix_ID = a.ID
6 WHERE ARStatus = 'Active'
7 AND a.Code = 'Private'
8 AND c.City = 'Kyiv'
    
```

(No column name)	(No column name)	(No column name)	(No column name)	(No column name)
11226	282372761,24	1417132431,00	36	68

2. Запит до даних OLAP-системи (Немає необхідності у знанні мови).

Account Suffix	Private	.T		
Location	Kyiv	.T		
Count of Accounts	Current Balance	Credit Amount	Avg Age	Avg Total TXNS Count
11 226	282 372 761,24	1 417 132 431,00	36,92	67,75



#	Вид тесту	#	Вид тесту
1	Запит метрик	4	Запит виміру з великою кількістю
2	Запит з фільтрами	5	Запит ієрархічного виміру
3	Запит з фільтрами за додатковими атрибутами	6	Запит вкладених вимірів

Висновки

- ✓ Проведено аналіз існуючих систем аналітики.
- ✓ Проаналізовано інструменти та принципи роботи з потоками даних.
- ✓ Спроектвано модель системи аналізу банківських операцій з її складовими та вимогами.
- ✓ Розроблено структуру бази даних, низку об'єктів OLAP-гіперкуба та повний ETL-процес.
- ✓ Розгорнуто модель на локальному сервері.
- ✓ Згенеровано та завантажено тестові дані.
- ✓ Проаналізовано продуктивність розробленої моделі.