

УДК 004.657

АНАЛІЗ ПІДХОДІВ ДО ІНТЕГРАЦІЇ SQL-ЗАПИТІВ З ВІЗУАЛІЗАЦІЄЮ ДАНИХ

Руденко А.В. – гр. БІП-21, студент, rudenko.a1337@gmail.com

Лебеденко Ю.О. – к.т.н., доц., lebedenko.yo@knutd.edu.ua

Київський національний університет технологій та дизайну

Аналіз великих баз даних є одним із ключових напрямків сучасних інформаційних технологій, що потребує застосування ефективних підходів до вилучення, обробки та візуалізації інформації. Попри значні досягнення у розвитку інструментів для роботи з великими обсягами даних і інтеграції SQL-запитів із засобами візуалізації, залишаються нерозв'язаними питання забезпечення максимальної ефективності їх взаємодії [1]. Серед основних викликів можна виділити необхідність досягнення високої продуктивності, універсальності на різних платформах та зручності використання для кінцевих користувачів. Таким чином, актуальною є задача аналізу підходів до інтеграції SQL-запитів із інструментами візуалізації даних, що дозволить підвищити ефективність їх обробки та аналізу для прийняття обґрунтованих рішень.

Метою дослідження є створення та оцінка оптимальних методів інтеграції SQL-запитів із інструментами візуалізації даних, які забезпечують зручність використання, адаптивність та ефективність аналізу великих баз даних.

Інтеграція SQL-запитів із засобами візуалізації, такими як Power BI [2], забезпечує синергію можливостей реляційних баз даних із сучасними інструментами аналітики, що дозволяє приймати обґрунтовані управлінські рішення (рис. 1,а). Ключовою перевагою такого підходу є точне вилучення даних за допомогою SQL та створення інтерактивних дашбордів і звітів у Power BI. Додаткове застосування Python сприяє автоматизації процесів та розширенню можливостей аналітики.

Процес обробки даних складається з кількох етапів: розробка SQL-запитів для отримання даних, їх трансформація за допомогою Python-скриптів, імпорт у Power BI та створення візуальних презентацій (рис. 1,б) [3 – 5]. Для підвищення ефективності застосовуються методи, зокрема індексація бази даних, розпаралелювання обчислювальних задач та кешування результатів. Процес також включає проектування SQL-запитів для вилучення релевантних даних, їх обробку через Python, завантаження в Power BI для створення візуалізацій та звітів.

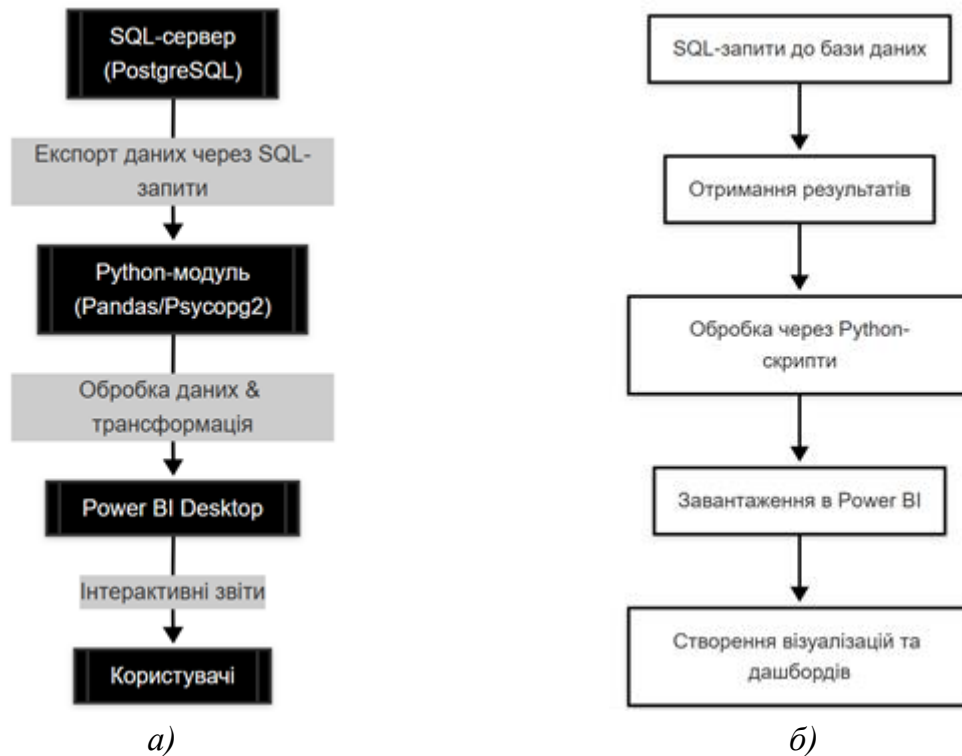


Рисунок 1 – Архітектура (а) та потік даних для інтеграції SQL, Python і Power BI у процесі візуалізації (б)

Для покращення ефективності використовуються методи індексації, розпаралелювання та кешування результатів. Система також передбачає автоматизацію оновлення звітів через планувальник задач, що значно знижує час обробки даних. Тестування на наборі даних обсягом понад 10 ГБ показало зменшення часу формування звітів на 78% завдяки автоматизації процесів.

Висновок. Таким чином, інтеграція SQL, Python та Power BI є ефективним підходом для роботи з великими даними, що дозволяє поєднати точність SQL з інтуїтивною візуалізацією даних. Система може бути адаптована для різних галузей дата-аналітики.

Л і т е р а т у р а

1. Beaulieu A. Learning SQL: Generate, Manipulate, and Retrieve Data. O'Reilly Media, 2022.
2. Ferrari A., Russo M. Introducing Microsoft Power BI. Microsoft Press, 2023.
3. McKinney W. Python for Data Analysis. O'Reilly Media, 2022.
4. Larson B. Data Analysis with Microsoft Power BI. Packt Publishing, 2023.
5. Nield T. Essential SQL: A Practical Guide to SQL in PostgreSQL and MySQL. Manning Publications, 2022.