

## **ДОСЛІДЖЕННЯ ШВИДКОСТІ ВИКОНАННЯ ОПЕРАЦІЙ У НОРМАЛІЗОВАНИЙ І ДЕНОРМАЛІЗОВАНИЙ БАЗАХ ДАНИХ**

**Дюков Євген, Пузікова Анна**

*Центральноукраїнський державний педагогічний університет імені*

*Володимира Винниченка, м. Кропивницький, Україна*

*У статті наведені результати дослідження швидкості виконання операцій селекції та вставки в реляційних базах даних, таблиці яких знаходяться у нормалізованій (до нормальної форми Бойса-Кодда) і відповідно у денормалізованій формах. В якості об'єкта дослідження була обрана стандартна база даних World, яка постачається з СУБД MySQL. Для операції вставки у денормалізовану базу даних з метою збереження цілісності даних був написаний тригер. Показано, що проведення денормалізації таблиць реляційних баз даних призводить до пришвидшення виконання операції селекції, але є причиною уповільнення інших операцій, зокрема, операції вставки.*

*Ключові слова: нормалізація, денормалізація, реляційні бази даних, швидкість виконання операцій.*

### **Investigation of the speed of operations in normalized and denormalized databases**

**Dyukov Evgen, Puzikova Anna**

*Volodymyr Vynnychenko Central Ukrainian State Pedagogical University,*

*Kropyvnytsky, Ukraine*

*The article presents the results of the research of the speed of selection and insertion operations in relational databases, the tables of which are in the normalized (to the normal form of Boyce-Codd) and respectively in denormalized forms. The standard database "World" which supplied with MySQL database management system was chosen as the object of study. For the insertion operation in a denormalized database, a trigger was written to maintain of the data integrity. It is shown that the implementation of the denormalization of relational database tables leads to faster execution of the selection operation, but causes the delay of other operations, in particular, insertion operation.*

*Keywords: normalization, denormalization, relational databases, speed of implementation of operations.*

Нормалізація схеми реляційної бази даних (БД) – зведення до відповідних нормальних форм – виконується на етапі логічного проектування [1, 123; 4, 279]. Метою її проведення є звільнення від відомих аномалій (оновлення,

вставки, знищення), які, в свою чергу, виникають внаслідок наявності різних видів обмежень (наприклад, функціональних і багатозначних залежностей). Нормалізація є одним із інструментів розробника схеми реляційної БД для підтримки її цілісності [1, 137; 2, 488; 4, 223]. Але нормалізована БД характеризується уповільненням швидкості виконання запитів, які вимагають проведення вибірки з кількох таблиць, в порівнянні з денормалізованою [3; 5]. Це пояснюється тим, що операція з'єднання двох або більше таблиць є найзатратнішою за часом виконання у порівнянні з іншими операціями (селекцією, проекцією, тощо). Тому практики при роботі з відносно великими реляційними БД з метою підвищення швидкості виконання операції селекції надають перевагу денормалізованим таблицям.

Таким чином, з практичної точки зору [3; 5], класична теорія проектування логічної схеми реляційної БД, яка вимагає здійснення нормалізації, з одного боку забезпечує підтримку цілісності БД, а з іншого – є причиною зниження швидкості виконання запитів в процесі експлуатації цієї БД.

Огляд наукової літератури і IT-ресурсів для спеціалістів показав відсутність статистичних даних з зазначеної проблеми для розповсюдженої в Україні СУБД MySQL, яка і була обрана в якості об'єкта дослідження.

Метою даної роботи є дослідження швидкодії виконання операцій селекції та вставки у нормалізованій і денормалізованій БД MySQL.

За основу була використана стандартна БД World (її схема зображена на рисунку 1), яка постачається з СУБД MySQL.

В якості денормалізованої схеми БД будемо розглядати схему, яка включає усі атрибути (без дублікатів) трьох таблиць БД World. Відповідна таблиця містить 30670 рядків.

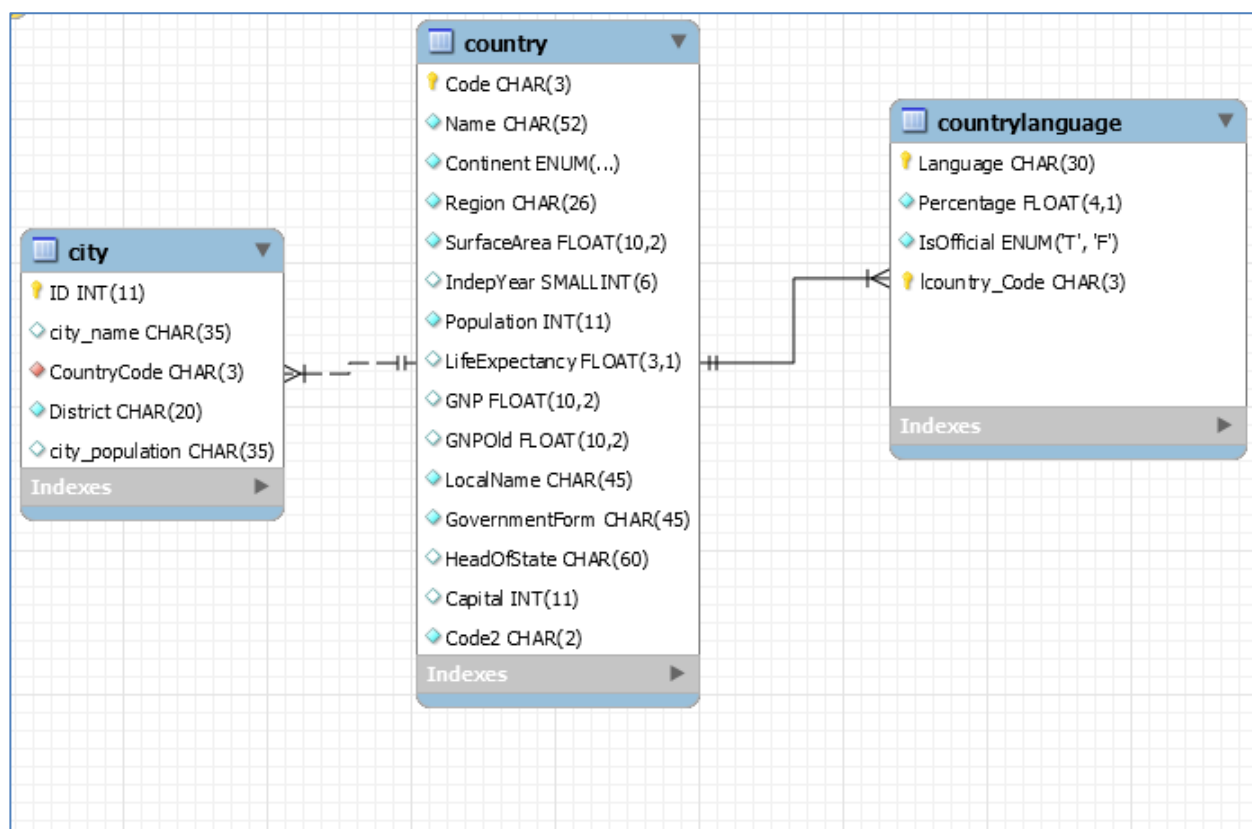


Рис. 1. Логічна схема нормалізованої БД World

Для дослідження швидкості виконання запитів, було розроблено набір з 40 тестів, що містять такі групи запитів: з умовою на різні типи полів, з агрегатними функціями без операції групування, з операцією групування та агрегатними функціями, запити з підзапитами. П'ять результатів виконання запитів (з різних тестових груп) з таблиць нормалізованої БД (НБД) і відповідно денормалізованої БД (ДНБД) наведені у таблиці 1.

Таблиця 1.

### Результати виконання операції селекції у НБД і ДНБД

№	Кількість рядків у результаті		Час виконання, мс		Кількість рядків за 1 мс	
	ДНБД	НБД	ДНБД	НБД	ДНБД	НБД
1	30670	30670	80	200	383	153
2	823	823	40	220	20	3
3	15169	15169	50	70	303	216
4	523	523	50	160	10	3
5	1	1	40	130	0,025	0,007
...	...	...	...	...	...	...
Середнє значення					115,019	43,906

Таким чином, відповідно до отриманих результатів, в денормалізованій БД швидкість виконання запитів в середньому збільшилась приблизно в 2,6 рази.

Діаграма, яка відображає швидкість виконання запитів у нормалізованій і денормалізованій БД, наведена на рисунку 2.

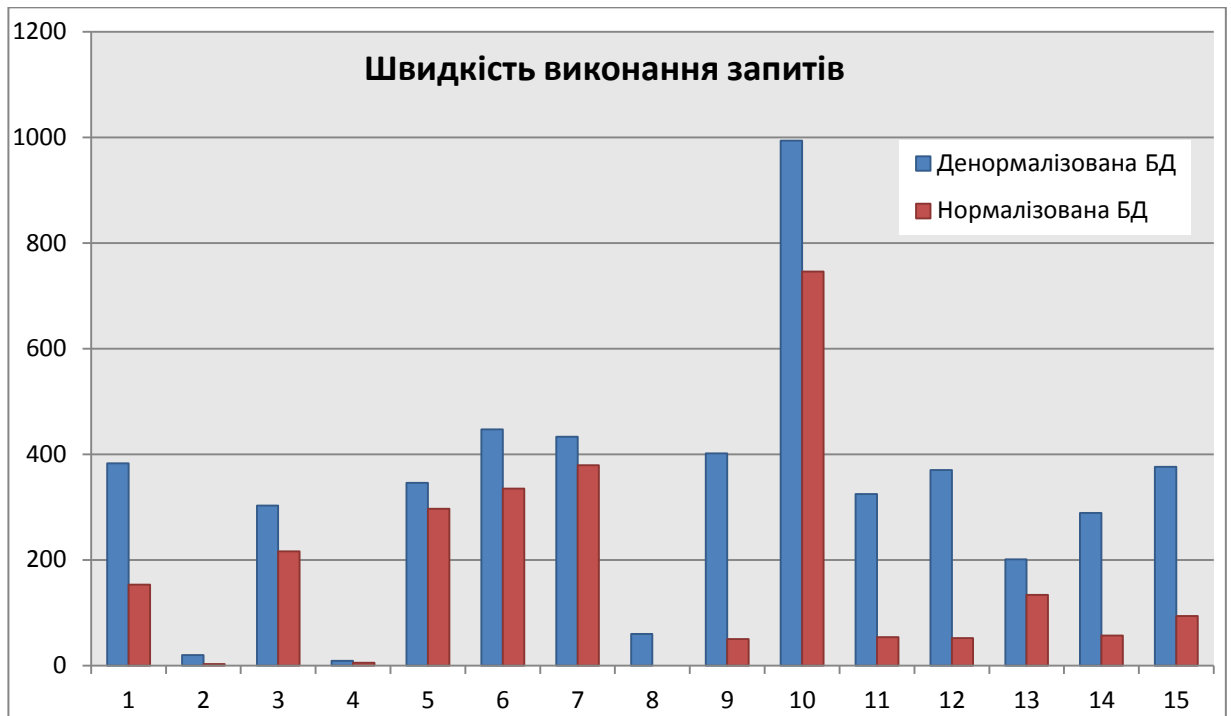


Рис. 2. Швидкість виконання запитів у нормалізованій і денормалізованій БД

Наступним кроком стало дослідження швидкості виконання операцій вставки у нормалізованій і відповідно денормалізованій БД. Для автоматизації даного процесу для кожної БД були написані процедури. Оскільки в процесі денормалізації за рахунок надлишковості даних виникає проблема появи загальновідомих аномалій, для денормалізованої БД з метою підтримки наявних в її таблиці функціональних залежностей додатково був створений тригер.

Зведені результати щодо виконання вставки різної кількості рядків у нормалізовану та денормалізовану БД наведені в таблиці 2.

Як можна бачити, в денормалізованій БД швидкість виконання операцій вставки в середньому зменшилась приблизно в 2,3 рази.

Діаграма, яка відображає швидкість виконання операцій вставки у нормалізованій і денормалізованій БД, наведена на рисунку 3.

## Результати виконання операції вставки у НБД і ДНБД

№	Кількість рядків у результаті	Час виконання, мс		Кількість рядків за 1 мс	
		ДНБД	НБД	ДНБД	НБД
1	10	700	2190	0,014	0,004
2	20	1460	5050	0,013	0,003
3	50	3460	10750	0,014	0,004
4	100	6900	22060	0,014	0,004
5	250	14480	71500	0,017	0,003
...	...	...	...	...	...
Середнє значення				0,0016	0,0037

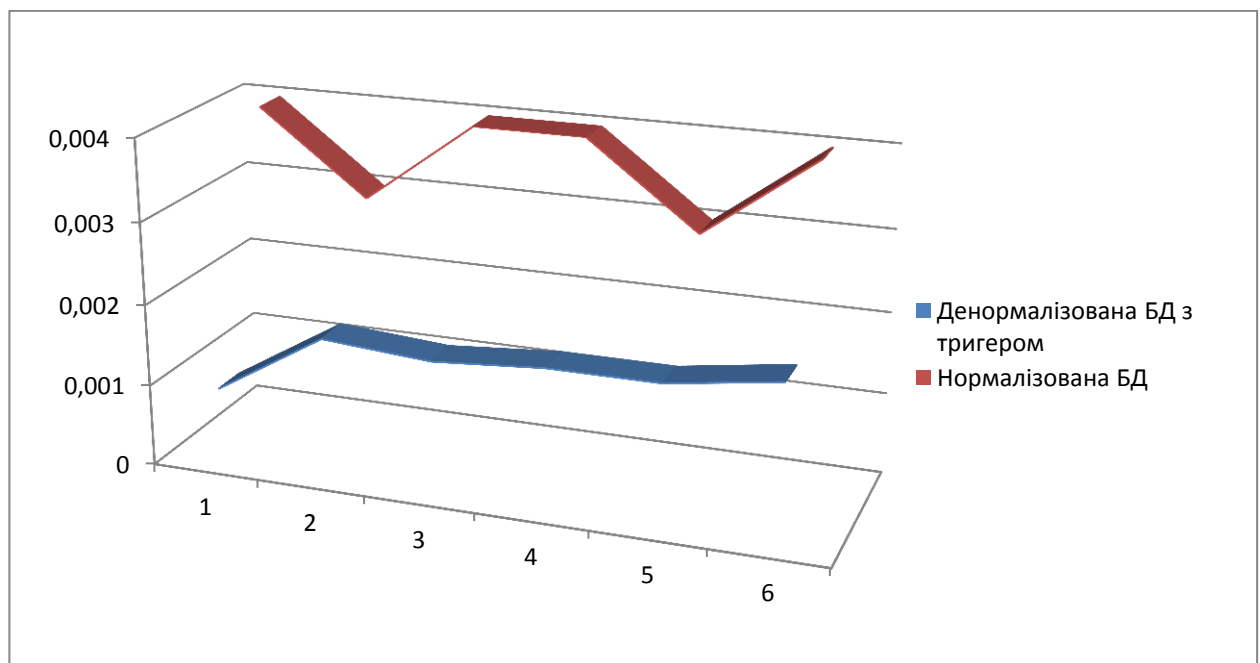


Рис. 3. Швидкість виконання операцій вставки у нормалізованій і денормалізованій БД

Слід зауважити, що у даній роботі не наведенні результати дослідження щодо операцій оновлення та знищення рядків, які є близькими до результатів стосовно операції вставки.

У випадку денормалізованих таблиць, які містять надлишкові дані і потребують додаткового місця на диску, з метою забезпечення цілісності даних, операції оновлення та знищення також вимагають написання тригерів,

транзакцій і збережених процедур. Їх наявність в свою чергу теж зменшує швидкість виконання зазначених вище операцій.

Таким чином, проведення денормалізації таблиць реляційної БД призводить до пришвидшення виконання операції селекції, але, в свою чергу, є причиною уповільнення інших операцій. Отже, приймаючи рішення щодо підвищення ефективності виконання операцій в конкретній реляційній БД, потрібно враховувати співвідношення кількості операцій селекції та операцій вставки, оновлення і знищення рядків.

### Список літератури

1. Гарсия-Молина Г., Ульман Д., Уидом Д. Системы баз данных. Полный курс. М.: Вильямс, 2003. 1088 с.
2. Дейт К. Дж. Введение в системы баз данных. М.: Вильямс, 2005. 1328 с.
3. Зачем нужна денормализация баз данных, и когда ее использовать (Блог компании Латера Софтвр, Разработка веб-сайтов) URL: <https://habr.com/ru/company/latera/blog/281262/> (дата звернення: 18.04.2019).
4. Коннолли Т., Бегг К., Страчан А. Базы данных: проектирование, реализация, сопровождение. Теория и практика: уч. пос. М.: Вильямс, 2000. 1120 с.
5. Databases: Normalization or Denormalization. Which is the better technique? URL: <http://www.ovaistariq.net/199/databases-normalization-or-denormalization-which-is-the-better-technique/#.XLroWegzaUI> (дата звернення: 18.04.2019).