Міністерство освіти і науки України

Черкаський державний технологічний університет

Кафедра програмного забезпечення автоматизованих систем

**ЗВІТ**

з лабораторної роботи №2

з предмету «Аналіз великих даних»

|  |  |
| --- | --- |
| Перевірив:  к. т. н., д. каф. ПЗАС  Рідкокаша А.А.  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  «\_\_\_\_» \_\_\_\_\_\_\_\_\_\_ 2019 р. | Виконав:  студент 1-го курсу  групи МПЗ-1904  Гаврилюк В. Є. |

Черкаси 2019

**Лабораторна робота №2**

**Тема:** Робота з базою даних HBase.

**Мета:** Реалізувати обробку даних в системі управління базами даних HBase.

**Завдання:**

1. Створити базу даних.
2. Наповнити її даними.
3. Відобразити роботу операцій put, delete, get та scan.
4. Навести приклад запиту.

**Короткі теоретичні відомості**

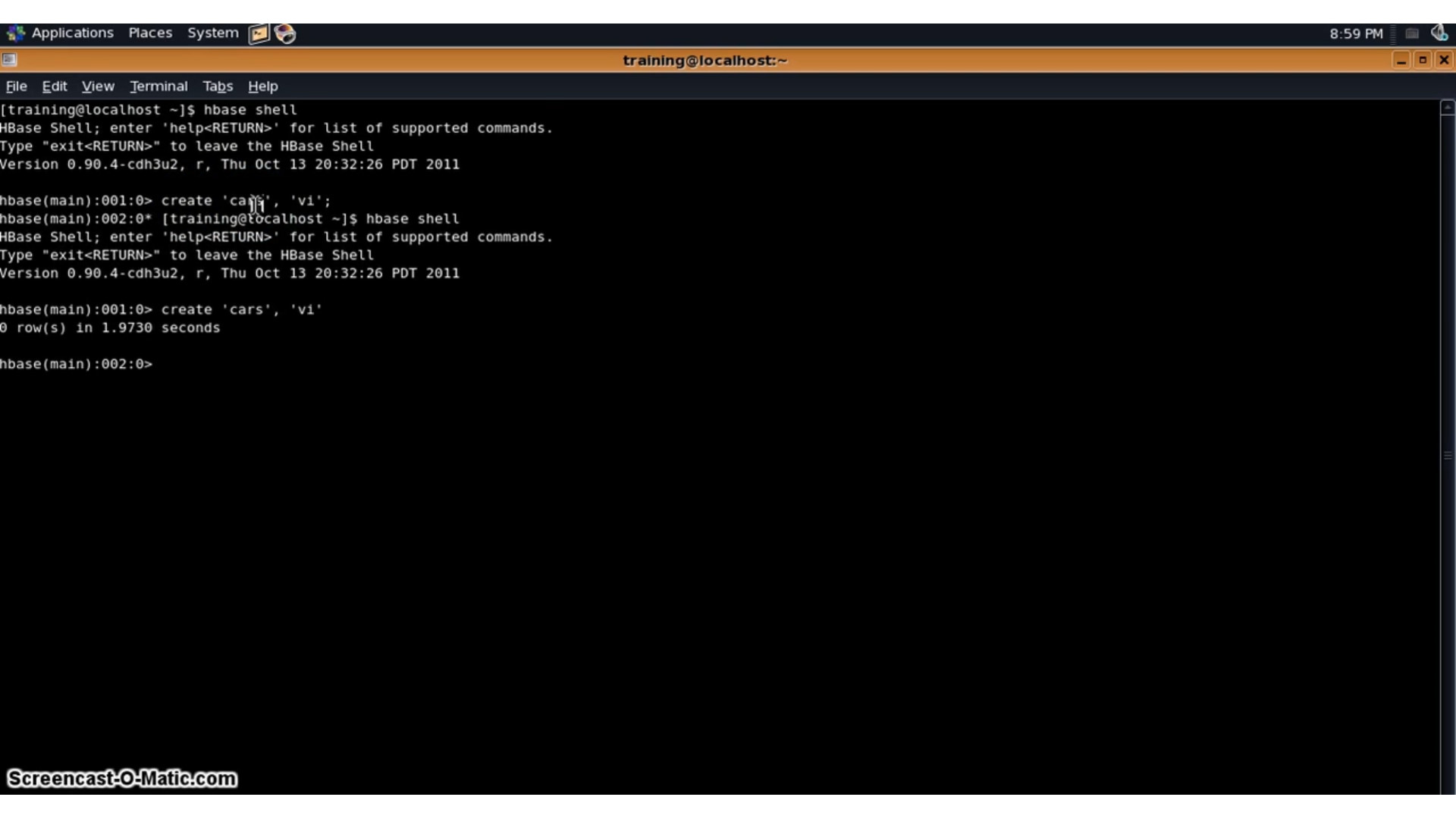
HBase - СУБД класу NoSQL з відкритим вихідним кодом, проект екосистеми Hadoop. Написана на Java та відноситься до категорії «сімейство стовпців» (англ. wide-column store), багато технічних рішень перейняті з Google BigTable. Працює поверх розподіленої файлової системи HDFS і забезпечує BigTable-подібні можливості для Hadoop, тобто забезпечує відмовостійкий спосіб зберігання великих обсягів розріджених даних.

Підтримка стиснення, операції в пам'яті і фільтр Блума для кожного базового стовпчика реалізовані відповідно до документації BigTable. Таблиці в HBase можуть служити входом і виходом для роботи реалізації MapReduce в проекті Hadoop, і можуть бути отримані не тільки через Java API, але і через API (REST, Avro, Thrift). Проекти Phoenix і Trafodion забезпечують SQL-доступ до даних під керуванням HBase.

Використовується для управління даними в ряді великих проектів, зокрема, Facebook в період 2010-2018 років використовував HBase для платформи повідомлень (в 2018 році платформа переведена на MyRocks), до числа постійних користувачів також відносяться Adobe, StumbleUpon, Twitter, Yahoo! (Експлуатує HBase-кластер з 3 тис. вузлів).

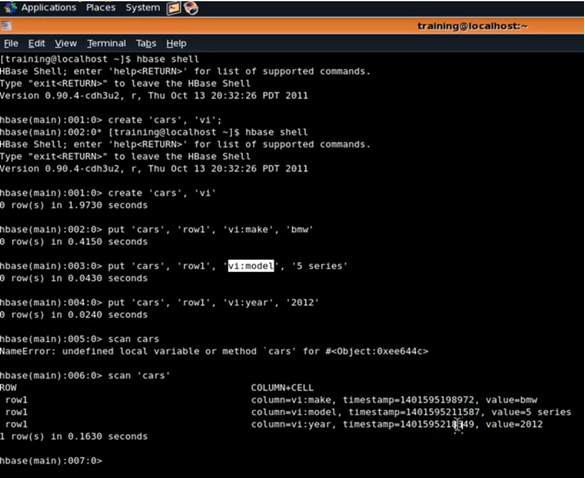
**Хід роботи**

1. Створення бази даних.

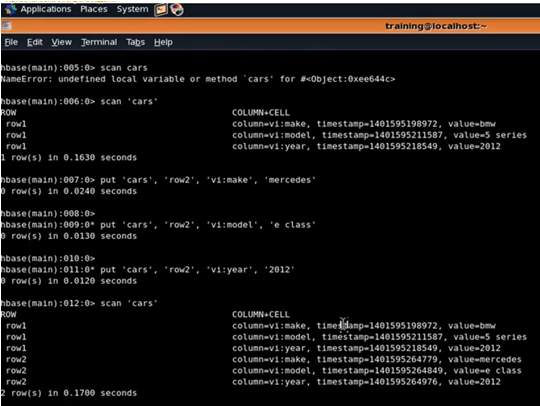


*Рис. 1 – Створення таблиці «cars».*

1. Заповнення таблиці даними за допомогою «put», та перевірка введених даниз за допомогою «scan».

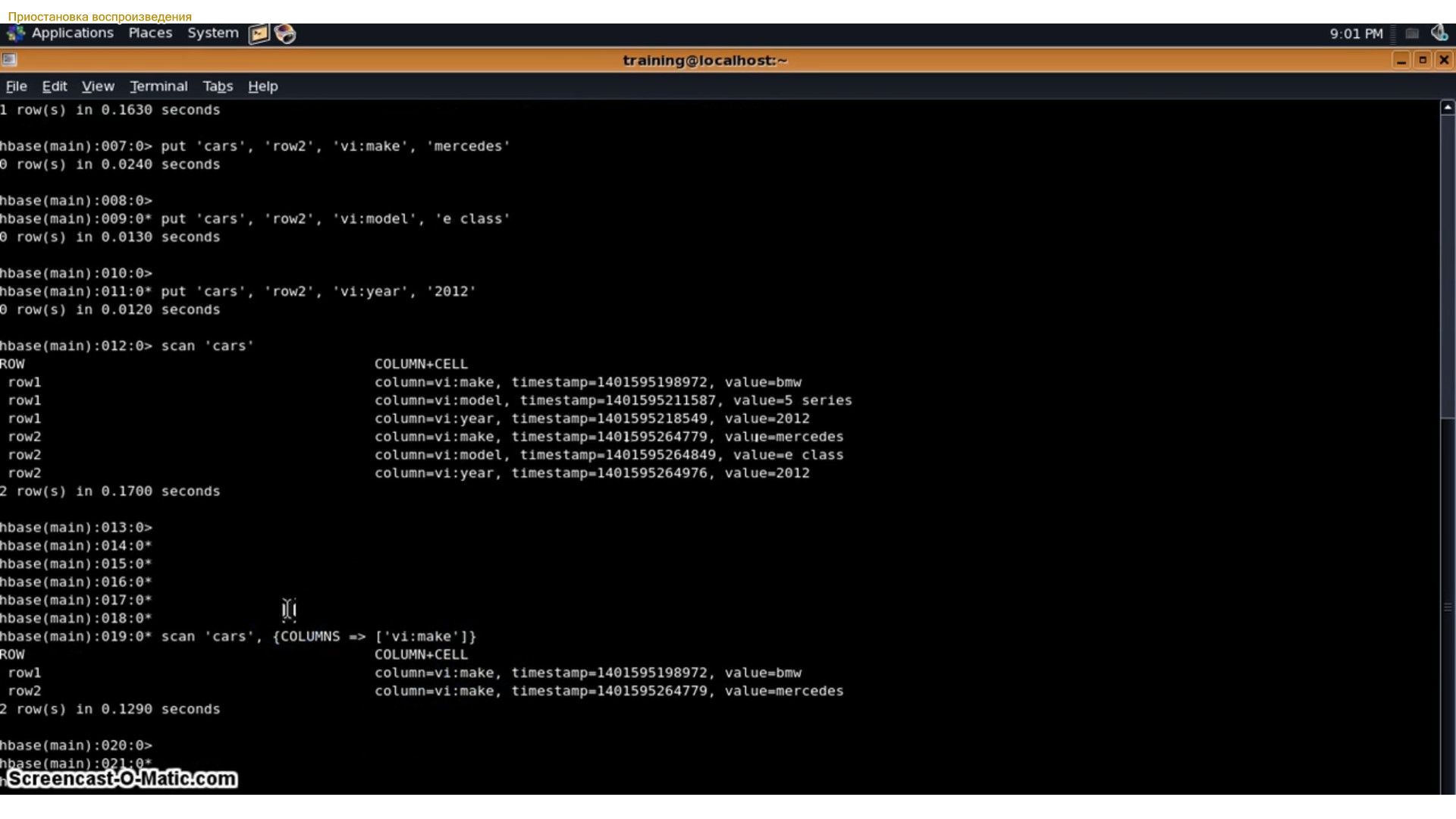


*Рис. 2 – Додавання першого елементу та перевірка введеного.*

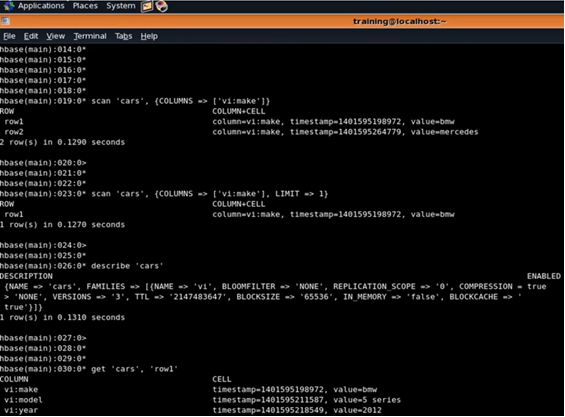


*Рис. 3 – Додавання другого елементу та перевірка введеного.*

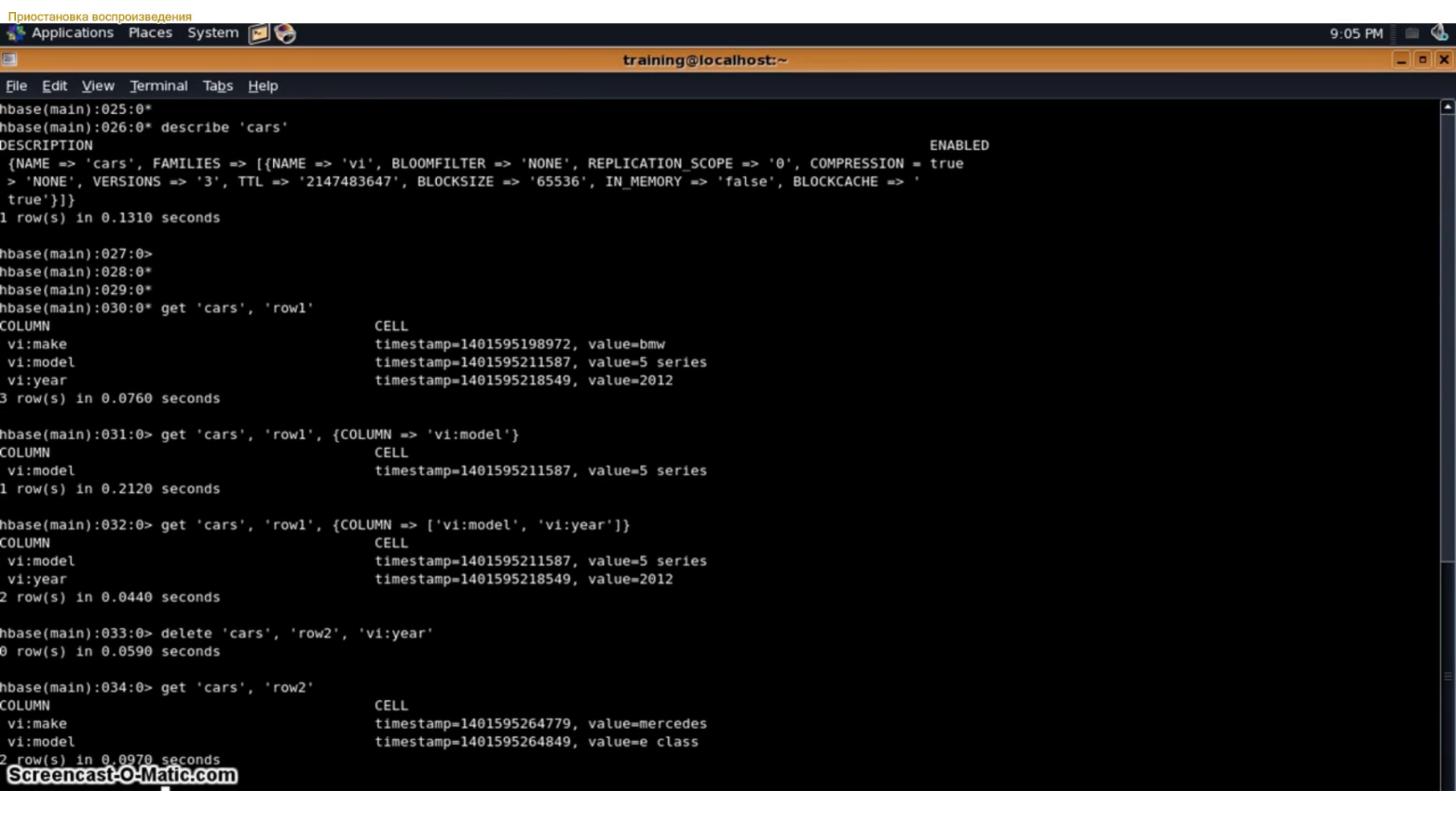
1. Презентувати операції put, delete, get, scan, навести запит.



*Рис. 4 – Перевірка введеного за певною колонкою.*



*Рис. 5 – Отримання першого рядка з таблиці за допомогою «get».*



*Рис. 6 – Видалення рядка за допомогою «delete».*

**Висновок:** Під час виконання даної лабораторної роботи я реалізував обробку даних в системі управління базами даних HBase. Навчився створювати таблиці, заповнювати їх даними. Навчився проводити перевірку введених даних за певною колонкою. Навчився використовувати команди для взаємодії з таблицями.