

Лабораторна робота 5 (I). Створення і заповнення бази даних MongoDB

Зміст

Лабораторна робота 5(I). Створення і заповнення бази даних MongoDB	1
Зміст	1
Теоретичні положення	1
Термінологія MongoDB	1
Сурогатний ключ	1
Практичні вправи.....	2
Вправа 5(I).1. Встановлення MongoDB	2
Вправа 5(I).2. Навчання початковим командам MongoDB за допомогою навчальної програми	3
Вправа 5(I).3. Інсталяція і використання оболонки Robomongo.....	3
Вправа 5(I).4. Створення бази MongoDB на основі реляційної бази	4
1а. Експорт даних у файли формату csv з бази MsSQL.....	4
1б. Експорт даних у файли формату csv з бази Oracle.....	5
2. Імпорт даних формату csv у БД MongoDB	6
Контрольні питання.....	7
Використані джерела.....	7

Мета: Встановлення MongoDB та інсталяція оболонки Robomongo. Створення бази даних в середовищі MongoDB і заповнення її шляхом експорту таблиць з MsSQL/Oracle та імпорту даних у MongoDB.

Теоретичні положення

Термінологія MongoDB

У порівнянні з реляційними СУБД Microsoft SQL Server та Oracle NoSQL-СУБД MongoDB має іншу термінологію:

Реляційні СУБД	MongoDB
База даних	База даних
Таблиця	Колекція
Рядок таблиці, запис	Документ
Колонка таблиці, поле запису	Поле

Реляційні бази даних визначають «колони» на рівні «таблиці», в той час як документо-орієнтовані бази даних, як-от MongoDB, визначають «поля» на рівні «документа». Це значить, що будь-який документ всередині колекції може мати свій власний унікальний набір полів, а самі колекції безструктурні.

В даній роботі ми на основі таблиць реляційної СУБД ми створюємо колекції, а на основі полів таблиць – поля документів колекцій, фактично не змінюючи схему даних по змісту.

Сурогатний ключ

Якщо це не передбачено явно, для кожного документа кожної колекції автоматично створюється ідентифікуюче поле `_id`, яке містить 24 16-річних символи. Ці поля пронумеровані послідовно. Таким чином, розробнику для ідентифікації документів можна нічого не робити.

Практичні вправи

Вправа 5(І).1. Встановлення MongoDB

Завдання: Встановити MongoDB і перевірити її працездатність.

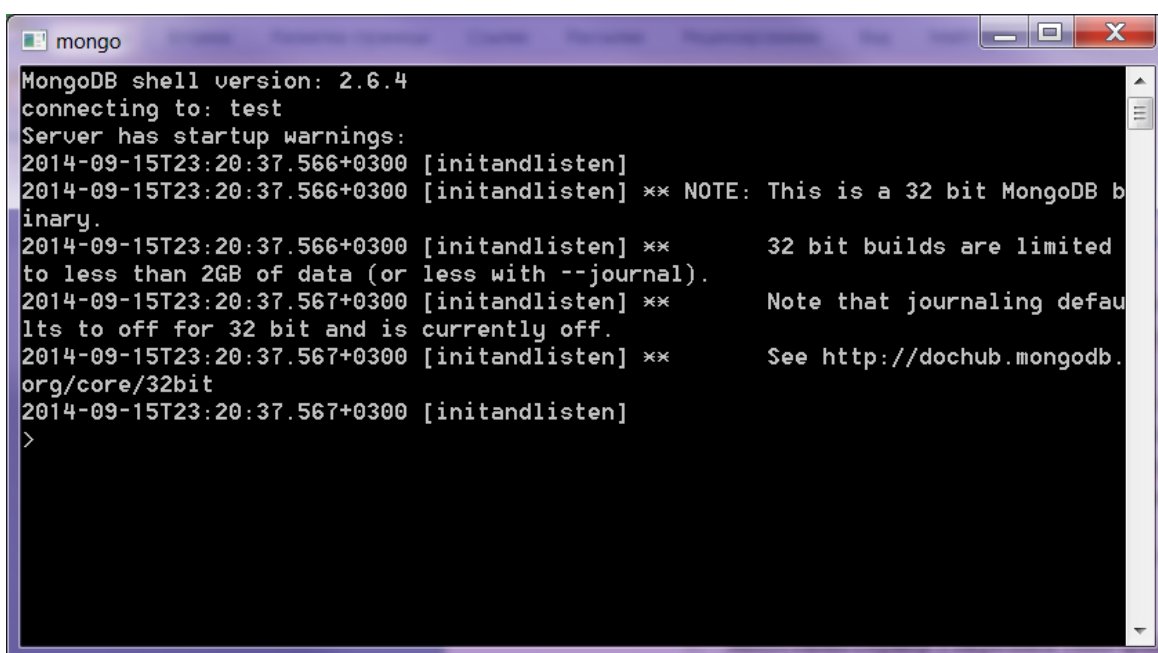
MongoDB розповсюджується вільно. При встановленні вказаного нижче дистрибутиву вона не інсталується, а встановлюється, тобто запускається після копіювання необхідних файлів. Для роботи на Windows XP версія MongoDB не повинна бути старша 2.2.

1. Скачаємо файл дистрибутиву з сайту даного продукту <http://www.mongodb.org/downloads> . Оберіть дистрибутив, відповідний до вашої операційної системи.
2. Розпакуємо архів у будь-якій папці, де зручно розташувати базу. Наприклад, у d:\mongodb .
3. [Hotfix for Windows7 для MongoDB](#) можна взяти на сторінці ОБДЗ в Moodle.
4. Створимо папку, де будуть зберігатись дані. Наприклад, d:\mongodb\data .
5. У папці \bin створимо файл mongodb.config і додамо в нього рядок dbpath=ШЛЯХ ДЛЯ ЗБЕРЕЖЕННЯ ФАЙЛІВ БД . В нашому прикладі: dbpath=d:\mongodb\data\
6. В папці \bin\ містяться виконувані файли:
 - mongod.exe – сервер,
 - mongo.exe – інтерактивна консоль,
 - mongoimport.exe – утіліта імпорту,
 - та деякі інші.

Організуємо зручний (напр., з папки на робочому столі) запуск перших двох програм. Сервер запускається з параметром –config ШЛЯХ ДО mongodb.config/mongodb.config . Цей виклик у Windows доцільно помістити в cmd-файл. Приклад змісту cmd-файлу:

```
d:\mongodb\bin\mongod --config d:\mongodb\bin\mongodb.config
pause
```

7. Запустимо сервер і окремим cmd-файлом – консоль. Вигляд консолі:



```

mongo
MongoDB shell version: 2.6.4
connecting to: test
Server has startup warnings:
2014-09-15T23:20:37.566+0300 [initandlisten]
2014-09-15T23:20:37.566+0300 [initandlisten] ** NOTE: This is a 32 bit MongoDB b
inary.
2014-09-15T23:20:37.566+0300 [initandlisten] **      32 bit builds are limited
to less than 2GB of data (or less with --journal).
2014-09-15T23:20:37.567+0300 [initandlisten] **      Note that journaling defau
lts to off for 32 bit and is currently off.
2014-09-15T23:20:37.567+0300 [initandlisten] **      See http://dochub.mongodb.
org/core/32bit
2014-09-15T23:20:37.567+0300 [initandlisten]
>
  
```

Рис.14.1. Інтерактивна консоль MongoDB

8. Для перевірки працездатності сервера виконаємо кілька команд з консолі Регістр має значення):

- `db.version()` – версія СУБД,
- `db` – ім'я підключеної БД,
- `db.help()` – методи об'єкта БД,
- `db.listCommands` – перелік усіх команд БД,
- `use БД` – підключення до іншої БД.

Вправа 5(І).2. Навчання початковим командам MongoDB за допомогою навчальної програми

Завдання: Опанувати базові команди роботи з БД за допомогою навчальної програми.

Навчальна програма запускається за адресою <http://try.mongodb.org/>.

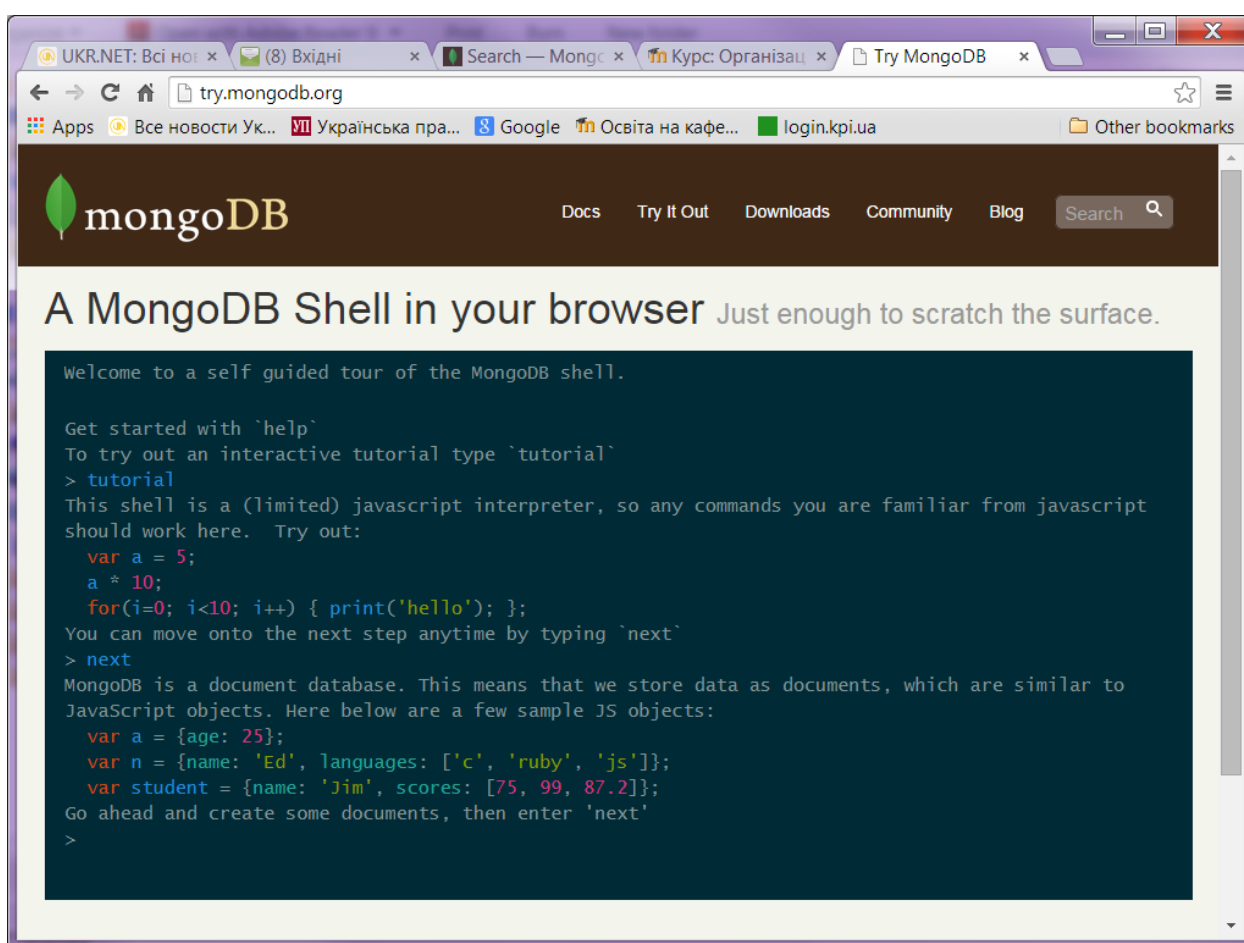


Рис.14.2. Вікно навчальної програми

Пройдіть усі кроки навчальної програми. Це може зайняти 20-30 хвилин.

Вправа 5(І).3. Інсталяція і використання оболонки Robomongo

Завдання: Встановити оболонку Robomongo і опанувати її інтерфейс.

Robomongo – вільно розповсюджуваний оболонко-центричний кросплатформений інструмент для управління MongoDB (Shell-centric cross-platform MongoDB management tool). Скачаємо дистрибутив, відповідний до вашої операційної системи, з сайту

<http://robomongo.org/> . Вивантажений архів потребує розархівування (напр., у \ Program Files) та інсталювання. Варіанти встановлення для Windows 7 і наступних версій:

- Windows installer – пакет, що потребує інсталяції,
- Application archive – не потребує.

Для Windows XP версія Robomongo з Application archive не працює.

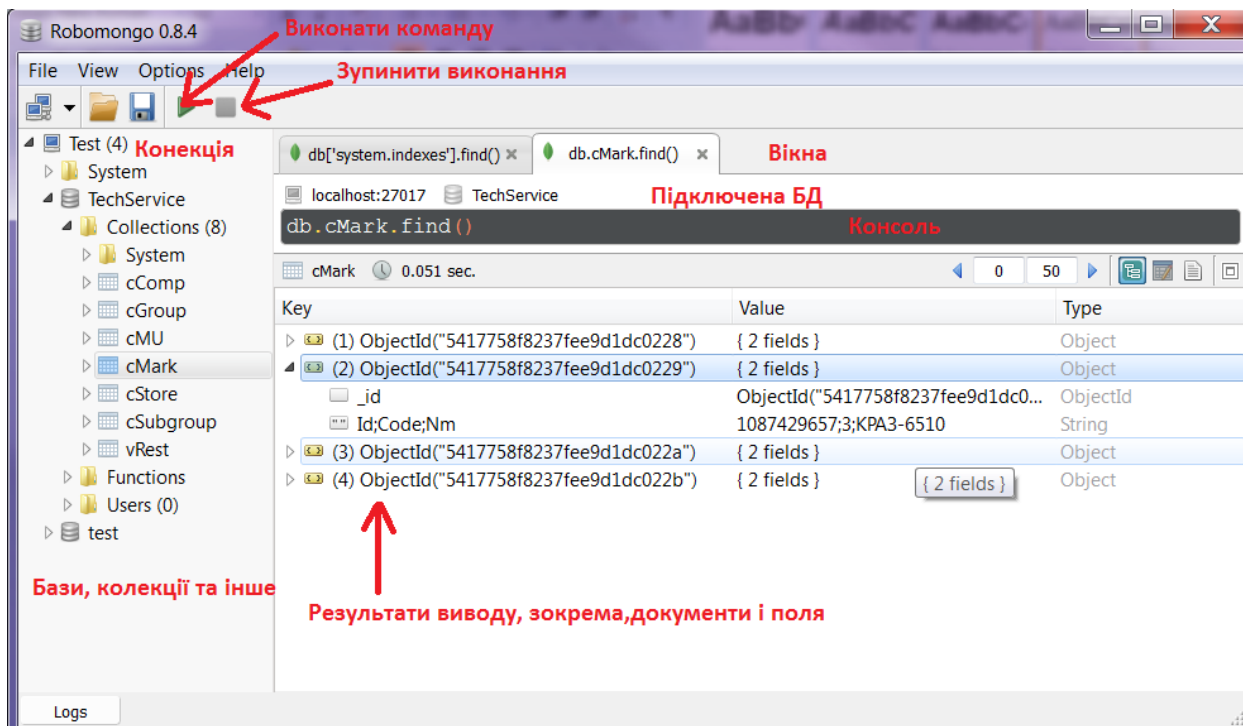


Рис.14.3. Вікно Robomongo

Вправа 5(І).4. Створення бази MongoDB на основі реляційної бази

Завдання: Експортувати дані з реляційної БД вашої предметної області у проміжний формат та імпортувати їх у MongoDB без зміни схеми даних. Тобто колекції мають відповідати таблицям, а документи – записам таблиць. У звіті треба відобразити кінцевий зміст вашої бази MongoDB.

Виконаємо цю вправу на прикладі бази Технічної служби. Експортуємо таблиці з реляційної БД у файли формату .csv та імпортуємо їх як колекції в MongoDB.

Файл формату .csv містить табличні дані. В кожному рядку запис; поля розділені розподільником «;» або «;». В першому рядку містяться імена полів. Крім блокноту файл може бути відчинений Excel.

1а. Експорт даних у файли формату .csv з бази MsSQL

1а.1. Зайдемо у SQL Server Management Studio та відчинимо таблицю для проглядання.

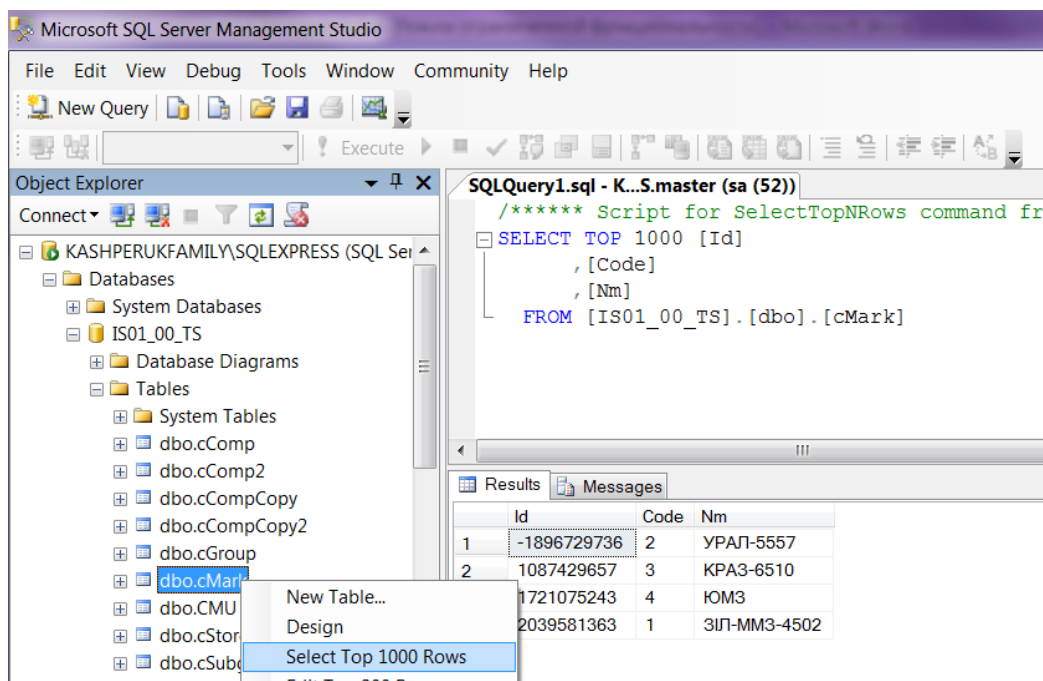


Рис.14.4. Відчинення таблиці для поглядання в SQL Server Management Studio

1а.2. Виділимо всі записи таблиці та в контекстному меню оберемо Save Results As... . У вікні діалогу обираємо папку збереження, ім'я файлу відповідно до імені таблиці і тип: CSV (Comma delimited). Виконуємо Save.

1а.3. Попередні пункти виконуємо для кожної таблиці, яку переносимо в MongoDB.

16. Експорт даних у файли формату .csv з бази Oracle

16.1. Скористаємось оболонкою Oracle Database Express Edition. Оберемо пункти меню Utilities \ Data Load/Unload \ Unload .

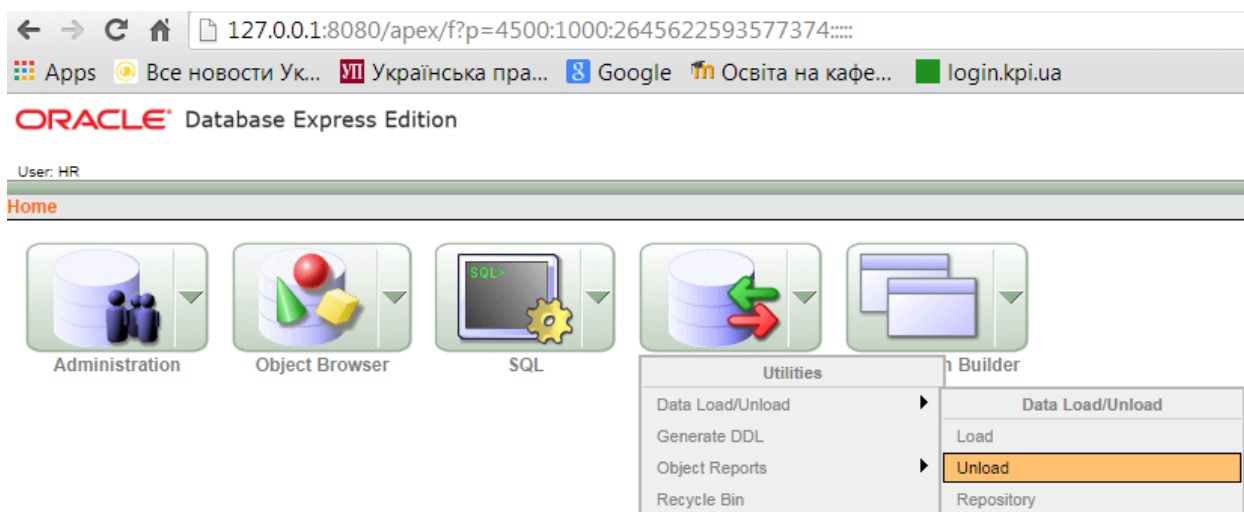


Рис.14.5. Пункт Unload в меню Oracle Database Express Edition

16.2. В подальших формах обираємо схему, таблицю, та додатково кілька опцій:

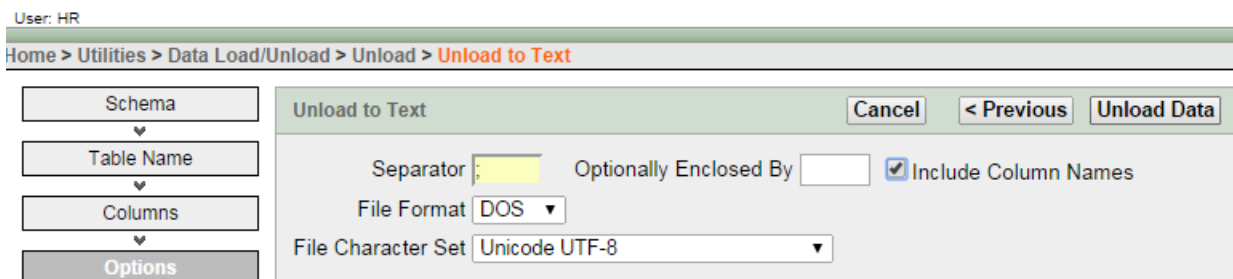


Рис.14.6. Опції для дії Unload to Text в Oracle Database Express Edition

16.3. Тиснемо кнопку Unload Data. Текстовий файл (.txt) створюється в папці Downloads. Скопіюємо його в потрібну папку і змінимо розширення на .csv .

2. Імпорт даних формату csv у БД MongoDB

Імпорт даних виконується утилітою mongoimport, яка знаходиться в папці \bin\ .

Припустимо, ця папка розташована у d:\mongodb\ , а файли .csv – у

D:\Proj\KPI\OBDZ\Lab\CSV\ . Тоді виклик утиліти імпорту з консолі операційної системи Windows матиме вигляд

```
D:\>d:\mongodb\bin\mongoimport --db TechService --collection cMark --type csv
--headerline --file D:\Proj\KPI\OBDZ\Lab\CSV\cMark.csv
connected to: 127.0.0.1
2014-09-16T02:26:07.714+0300 imported 4 objects
```

D:\>

Опції утиліти:

- db – база даних;
- collection - колекція, в яку імпортуються дані. Якщо колекція не існує, вона створюється;
- type – формат файлу-джерела: json, .csv або .tsv .
- headerline – ознака того, що перший рядок файлу-джерела містить імена полів;
- file – шлях та ім'я до файлу-джерела.

Як бачимо, утиліта повідомляє про конекцію до сервера та кількість імпортованих об'єктів - рядків (документів).

Такий імпорт треба виконати для кожного файлу .csv, що містить дані таблиці реляційної БД. Після цього треба перевірити наявність даних у базі за допомогою оболонки, як-от Robomongo (рис.14.3).

Існує імовірність при імпорті документів з .csv в колекцію MongoDB несприйняття символу «;» як розподільника полів. У цьому разі рядок файлу .csv з кількома полями імпортується як одне поле.

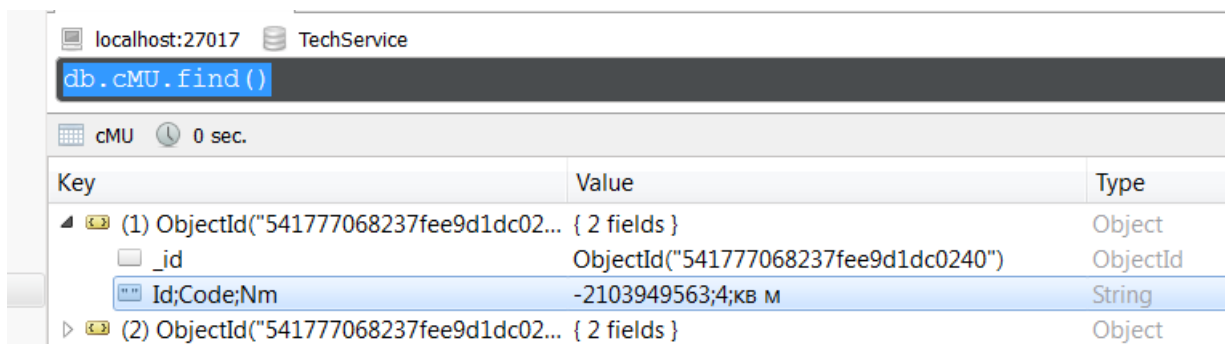


Рис.14.7. Невірно імпортовані дані в результаті несприйняття утилітою MongoImport «;» як розподільника полів

Тоді треба:

- У файлі .csv контекстною заміною в Notepad замінити десятковий розподільник «;» на «.»
- Там же замінити розподільник полів «;» на «.»
- Видалити всі документи з колекції в MongoDB
- Імпортувати колекцію ще раз.

Може трапитись, що з бази Oracle інформація в .csv експортується взагалі без розподільника полів. В цьому разі при налаштуванні експорту з Oracle треба вказати розподільник полів «.» і потурбуватись, аби у файлі .csv в числах десятковим розподільником була «.».

Контрольні питання

1. Як називаються таблиці, рядки таблиць (записи) і колонки таблиць у MongoDB?
2. В чому відмінність колекцій і документів у MongoDB від таблиць і записів у реляційній СУБД?
3. Які основні компоненти БД можна побачити на лівій панелі вікна Robomongo?
4. У чому переваги використання Robomongo перед інтерактивною консоллю MongoDB?
5. Яка команда виводить зміст колекції?
6. Як представляється таблиця в форматі .csv?
7. З якого вікна запускається утиліта mongoimport?

Використані джерела

1. Karl Seguin. The Little MongoDB Book (Маленькая книга о MongoDB). <http://openmymind.net/mongodb.pdf>
2. The MongoDB 2.6 Manual. <http://docs.mongodb.org/manual/>
3. mongoimport. <http://docs.mongodb.org/manual/reference/program/mongoimport/>