

Лабораторная работа 2-1. Изучение методов хранения данных на основе NoSQL в MongoDB, Redis, Neo4j

Цель работы: изучить и освоить методы хранения и работы с данными в NoSQL базах данных MongoDB, Redis и Neo4j. Научиться загружать данные из CSV файлов в указанные СУБД и выполнять базовые операции по работе с данными.

Оборудование и ПО:

- Операционная система Ubuntu.
- Установленные пакеты для работы с NoSQL базами данных: MongoDB, Redis, Neo4j.
- Язык программирования Python (с библиотеками pymongo, redis, neo4j).
- CSV файлы с данными.

Теоретическая часть

1. MongoDB: документо-ориентированная NoSQL база данных, где данные хранятся в формате JSON-подобных документов.
2. Redis: высокопроизводительная база данных типа "ключ-значение", часто используемая для кеширования и временного хранения данных.
3. Neo4j: графовая база данных, которая позволяет хранить данные в виде вершин и рёбер графа, что удобно для моделирования сложных взаимосвязей.

Ход работы

Шаг 1: Установка и настройка MongoDB

1.1. Установка необходимых библиотек

Вначале убедитесь, что у вас установлена библиотека `pymongo`. Если нет, выполните следующую команду:

```
```python
!pip install pymongo
```
```

1.2. Подключение к MongoDB с аутентификацией

Предположим, что у вас есть следующие параметры для подключения:

- URI MongoDB: mongodb://localhost:27017/
- Имя базы данных: labs
- Имя пользователя: mongouser
- Пароль: mongopasswd
- Имя коллекции (таблицы): lab21

Создадим подключение и проверим его успешность.

```
```python
from pymongo import MongoClient
#Задайте параметры подключения
mongo_uri = "mongodb://mongouser: mongopasswd @localhost:27017"

try:
 # Подключение к MongoDB
 client = MongoClient(mongo_uri)

 # Проверка подключения
 client.admin.command('ping')
 print("Подключение к MongoDB установлено успешно!")

 # Выбор базы данных
 db = client[' labs']
```
```

```
# Выбор коллекции
labs_collection = db['lab21']
```

```
except Exception as e:
    print(f"Ошибка подключения: {e}")
... 
```

1.3. Создание тестовых данных

Создадим пример тестовых данных:

```
```python
Пример тестовых данных
test_data = [
 {"lab_name": "Lab 1", "subject": "Physics", "date": "2024-08-28", "score": 85},
 {"lab_name": "Lab 2", "subject": "Chemistry", "date": "2024-08-29", "score": 90},
 {"lab_name": "Lab 3", "subject": "Biology", "date": "2024-08-30", "score": 88},
]
...
```

### 1.4. Загрузка данных в коллекцию labs

Теперь загрузим созданные тестовые данные в коллекцию `labs`:

```
```python
try:
    # Вставка данных в коллекцию
    result = labs_collection.insert_many(test_data)

    # Вывод идентификаторов вставленных документов
    print("Данные успешно загружены в коллекцию 'labs'.")
    print("Идентификаторы вставленных документов:", result.inserted_ids)
except Exception as e:
    print(f"Ошибка при загрузке данных: {e}")
... 
```

```
```python
documents = labs_collection.find()
for doc in documents:
 print(doc)
...
```

### 1.6. Закрытие подключения.

После завершения работы с базой данных закройте соединение:

```
```python
client.close()
... 
```

Этот пример демонстрирует, как подключиться к MongoDB с использованием аутентификации, выбрать базу данных и коллекцию, выполнить операции с данными и закрыть соединение.

Шаг 2: Установка и настройка Redis

2.1. Установите Redis.

```
```bash
sudo apt-get install redis-server
...
```

### 2.2. Запустите Redis:

```
```bash
sudo systemctl start redis
... 
```

2.3. Загрузка данных из CSV в Redis:

```

python
import csv
import redis

r = redis.Redis(host='localhost', port=6379, db=0)

with open('data.csv', mode='r') as file:
    reader = csv.reader(file)
    header = next(reader) # Считываем заголовок
    for row in reader:
        r.hmset(f"row:{row[0]}", {header[i]: row[i] for i in range(1, len(row))})

```

Шаг 3: Выполнение операций с данными

3.1. MongoDB:

- Найдите документ по определённому критерию:

```

python
result = collection.find_one({"field_name": "value"})
print(result)

```

- Обновите документ:

```

python
collection.update_one({"field_name": "value"}, {"$set": {"field_name": "new_value"}})

```

- Удалите документ:

```

python
collection.delete_one({"field_name": "value"})

```

3.2. Redis:

- Получите данные по ключу:

```

python
data = r.hgetall('row:1')
print(data)

```

- Обновите данные:

```

python
r.hset('row:1', 'field_name', 'new_value')

```

- Удалите ключ:

```

python
r.delete('row:1')

```

Варианты заданий

Каждому студенту предоставляется CSV файл с уникальными данными, который он должен загрузить в MongoDB, Redis. Студент должен продемонстрировать выполнение базовых операций (вставка, выборка, обновление, удаление) для каждой из СУБД. Ниже приводятся примеры тем, с которыми должны быть связаны данные в CSV (не менее 100 записей).

1. Сотрудники компании и их контактные данные.
2. Товары интернет-магазина и их описание.
3. Книги и авторы в библиотеке.
4. Курсы и преподаватели в университете.

5. Пациенты и их медицинские карты.
6. Заказы и клиенты интернет-магазина.
7. Фильмы и режиссеры в кинобазе.
8. Услуги и клиенты сервисного центра.
9. Транспортные средства и их владельцы.
10. Продукты и их поставщики.
11. Сотрудники и их департаменты в организации.
12. Студенты и их курсы в образовательном учреждении.
13. Проекты и участники в IT компании.
14. Клиенты и их заказы в ресторане.
15. Маршруты и станции в транспортной сети.
16. Оборудование и его спецификации в производстве.
17. Музыкальные альбомы и артисты.
18. Туристические маршруты и гиды.
19. Программные продукты и их версии.
20. Клиенты и их счета в банке.
21. Бренды и их товары на рынке.
22. Квартиры и арендаторы в жилом комплексе.
23. Ивенты и участники в системе мероприятий.
24. Лекции и студенты в образовательной платформе.
25. Контракты и их условия для компании.

Следующие задания могут быть выполнены с помощью “песочницы”:
<http://console.neo4j.org> в контексте использования базы «Учебные курсы», которая на языке Cypher описана следующим образом:

```
CREATE (C:course{name:'Discrete Mathematics'}),
(S:course{name:'Databases'}),
(I:course{name:'Data Processing'}),
(E:person{name:'Elena'}),
(N:person{name:'Natalia'}),
(V:person{name:'Victoria'}),
(O:person{name:'Olga'}),
(St:person{name:'Stas'}),
(A:person{name:'Andrey'}),
(D:person{name:'Dan'}),
(Al:person{name:'Alex'}),
(Dm:person{name:'Dmitry'}),
(K:person{name:'Katarina'}),
(Elena:student{name:'Teresa'}),
(Nina:student{name:'Nina'}),
(Victor:student{name:'Victor'}),
(Olga:student{name:'Olga'}),
(Stas:student{name:'Stas'}),
(Anna:student{name:'Elen'}),
(Denis:student{name:'Denis'}),
(Alexandr:student{name:'Alexandr'}),
(Dmitry:student{name:'Dmitry'}),
(Konstantin:student{name:'Konstantin'}),
(Roman:student{name:'Roman'}),
(ilia:student{name:'Ilia'}),
(E)-[:author]->(C),
(N)-[:author]->(C),
(A)-[:author]->(S),
(A)-[:author]->(I),
(Al)-[:author]->(I),
```

```

(Dm)-[:author]->(I),
(O)-[:speaker]->(C),
(St)-[:editor]->(C),
(V)-[:designer]->(C),
(D)-[:speaker]->(S),
(St)-[:editor]->(S),
(V)-[:designer]->(S),
(D)-[:speaker]->(I),
(St)-[:editor]->(I),
(K)-[:designer]->(I),
(Elena)-[:learn]->(I),
(Elena)-[:learn]->(C),
(Nina)-[:learn]->(S),
(Nina)-[:learn]->(C),
(Victor)-[:learn]->(S),
(Victor)-[:learn]->(C),
(Olga)-[:learn]->(S),
(Olga)-[:learn]->(C),
(Stas)-[:learn]->(I),
(Stas)-[:learn]->(C),
(Stas)-[:learn]->(S),
(Anna)-[:learn]->(I),
(Anna)-[:learn]->(C),
(Anna)-[:learn]->(S),
(Denis)-[:learn]->(I),
(Alexandr)-[:learn]->(C),
(Alexandr)-[:learn]->(S),
(Dmitry)-[:learn]->(I),
(Dmitry)-[:learn]->(C),
(Konstantin)-[:learn]->(S),
(Roman)-[:learn]->(C),
(Ilia)-[:learn]->(C)

```

Узлы базы: `person` — сотрудники, `student` — студенты. Отношение `learn` связывает студентов с курсами, на которые они записались, а отношения `author`, `speaker` и `editor` связывают авторов, дикторов и монтажеров с курсами, которые они создавали. Именам сотрудников, студентов и курсов соответствуют атрибуты `name`.

Вариантов заданий в контексте базы данных «Учебные курсы» на Neo4j

1. Напишите Cypher-запрос, который вернет список всех студентов, записанных на курс "Discrete Mathematics".
2. Составьте запрос, который вернет список курсов, на которые записаны студенты с именем "Nina" и "Olga".
3. Найдите всех авторов курса "Data Processing" и верните их имена.
4. Напишите запрос, который вернет список курсов, созданных сотрудником с именем "Andrey".
5. Составьте запрос для получения списка всех курсов, где "Stas" является редактором.
6. Напишите запрос, который вернет всех студентов, записанных на курс "Databases", и укажите, какие сотрудники связаны с этим курсом как авторы, дикторы и редакторы.
7. Найдите всех сотрудников, которые имеют отношение к созданию курса "Discrete Mathematics", и определите их роль.
8. Напишите запрос, чтобы найти всех студентов, которые учатся на всех трех курсах ("Discrete Mathematics", "Databases", "Data Processing").

9. Составьте запрос, который вернет список студентов, обучающихся на курсе "Data Processing", и перечислите авторов этого курса.
10. Напишите запрос, который вернет количество студентов, записанных на каждый из курсов.
11. Найдите всех студентов, которые учатся у "Elena", и верните список курсов, на которые они записаны.
12. Напишите запрос, чтобы получить список всех сотрудников, которые участвуют в создании хотя бы одного курса в роли автора, диктора или редактора.
13. Составьте запрос, который вернет список студентов, записанных на курсы, где "Katarina" является дизайнером.
14. Напишите запрос, чтобы найти все курсы, которые созданы командой сотрудников (автор, диктор, редактор), включающей "Dmitry".
15. Найдите всех студентов, которые не записаны ни на один курс.
16. Напишите запрос, который вернет список студентов и количество курсов, на которые они записаны.
17. Составьте запрос, чтобы получить список всех курсов, где один и тот же человек выполняет несколько ролей (например, автор и редактор).
18. Напишите запрос, который вернет список всех студентов, записанных на курс "Discrete Mathematics", и их наставников.
19. Найдите всех сотрудников, которые являются одновременно авторами и дикторами какого-либо курса.
20. Напишите запрос, чтобы найти все курсы, на которые записаны студенты "Elena" и "Stas".
21. Составьте запрос, чтобы получить список студентов, которые записаны на курс "Databases" и имеют того же наставника, что и "Nina".
22. Напишите запрос, который вернет список курсов, на которых "Roman" является студентом, а "Olga" - сотрудником (в любой роли).
23. Найдите всех сотрудников, которые принимали участие в создании курсов, на которых учится студент "Konstantin".
24. Напишите запрос, чтобы найти студентов, которые записаны на курсы, на которых "Victoria" является дизайнером.
25. Составьте запрос, чтобы получить список студентов, которые учатся на более чем двух курсах, и перечислите эти курсы.

Отчет по работе

Отчет должен включать:

- Описание загрузки данных в каждую из NoSQL СУБД.
- Пример кода для работы с MongoDB, Redis и Neo4j.
- Результаты выполнения операций с данными.
- Выводы по работе с различными типами NoSQL баз данных.