|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ **Информатика и системы управления**

КАФЕДРА **Компьютерные системы и сети (ИУ6)**

НАПРАВЛЕНИЕ ПОДГОТОВКИ **09.04.01 Информатика и вычислительная техника**

МАГИСТЕРСКАЯ ПРОГРАММА **09.04.01/07 Интеллектуальные системы анализа,**

**обработки и интерпретации больших данных**

**Отчет**

**по лабораторной работе № 6**

**Название:** Работа с графовой базой данных Neo4j на примере разработки

рекомендательной системы

**Дисциплина:** Технология параллельных систем баз данных

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Студент | ИУ6-12М |  |  | Д.С. Каткова |
|  | (Группа) |  | (Подпись, дата) | (И.О. Фамилия) |
|  |  |  |  |  |
| Преподаватель |  |  |  | А.Д. Пономарев |
|  |  |  | (Подпись, дата) | (И.О. Фамилия) |

Москва, 2023

**Цель лабораторной работы** – изучение работы графовой базы данных Neo4j и ее взаимодействия с документной NoSQL БД Elasticsearch на примере разработки рекомендательной системы.

Строится граф, узлами которого являются рецепты кулинарных блюд и ингредиенты, которые входят в эти блюда (рецепты). Для пользователя определяются предпочтения некоторым блюдам. И на основе анализа пользователю рекомендуются другие блюда, ингредиенты которых входят в предпочтительные блюда пользователя в максимальном количестве.

**Ход работы**

После запуска виртуальной машины файлы recipes.json, ingredients.txt, f1.py, f2.py были переписаны из яндекс-папки в /home/user/lab6 своей ВМ.

recipes.json – это файл json, в котором хранятся описания рецептов блюд (их более 100 тысяч).

ingredients.txt – это текстовый файл, в котором хранится список ингредиентов, которые входят в рецепты (их более 800).

f1.py – это программа, которая вводит рецепты в базу данных Elasticsearch

f2.py – это программа, которая строит граф, связывающий ингредиенты и рецепты, и сохраняет его в базе данных Neo4j

Далее была успешно проведена инсталляция сервиса Neo4j. Результат установки представлен на рисунке 1.



Рисунок 1 – Результат инсталляции Neo4j

Текст программы f1.py представлен на рисунке 2.



Рисунок 2 – Текст программы f1.py

При выполнении программы была произведена запись рецептов в базу данных. Результат выполнения программы f1.py показан на рисунке 3.



Рисунок 3 – Результат выполнения программы f1.py

Текст программы f2.py показан на рисунке 4.

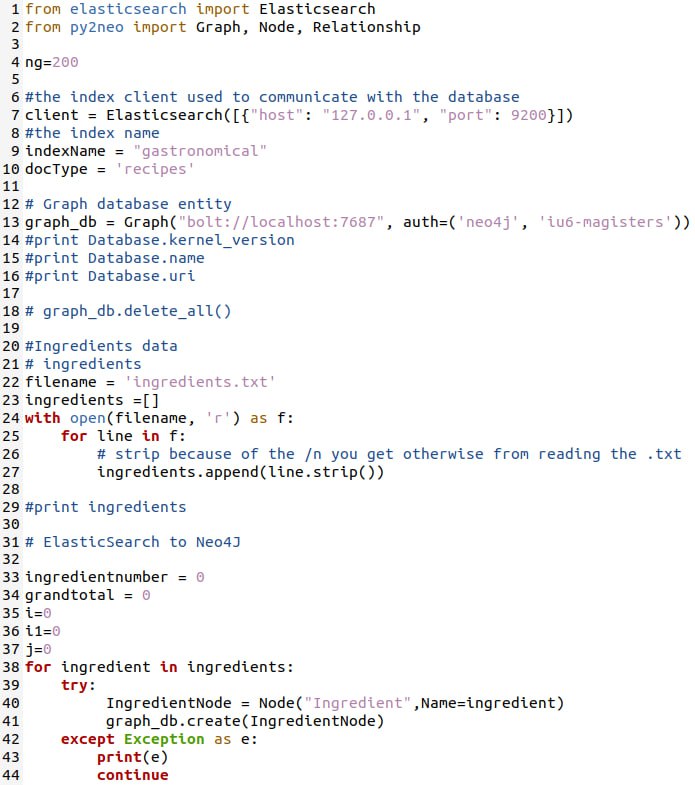




Рисунок 4 – Текст программы f2.py

По выполнению программы была произведена запись ингредиентов в базу данных. Результат выполнения программы f2.py показан на рисунке 5.

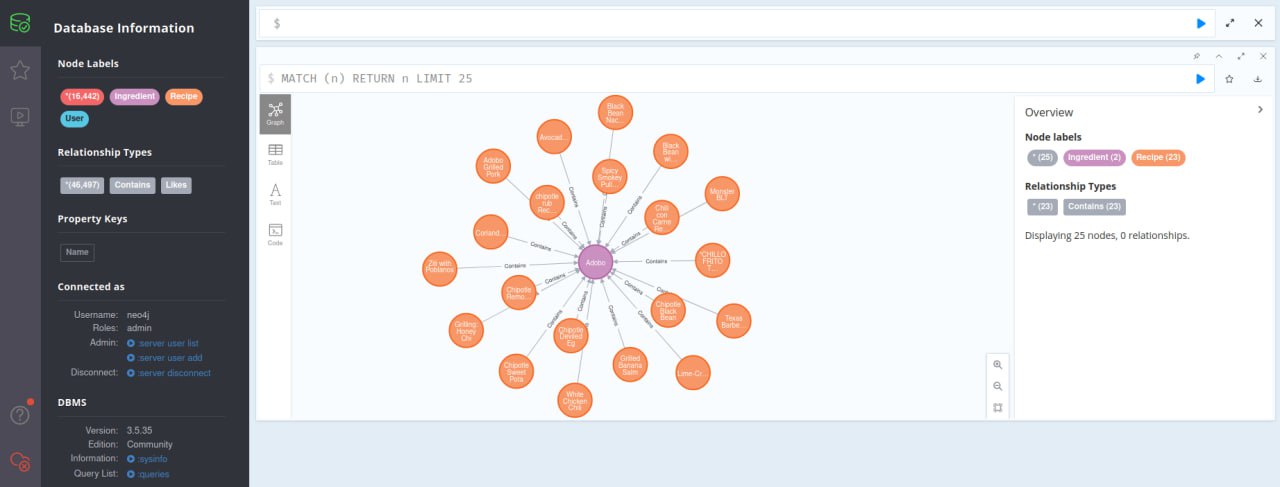


Рисунок 5 – Результат выполнения программы f2.py

**Запрос 1**

Текст запроса 1 записан в файл f3.py, содержимое которого представлено на рисунке 6.

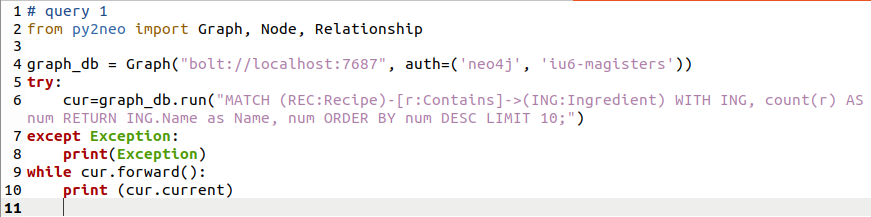


Рисунок 6 – Текст 1 запроса

Результат выполнения показан на рисунке 7.

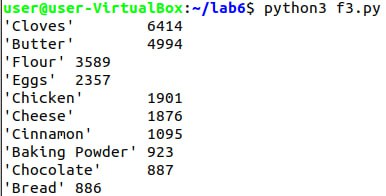


Рисунок 7 – Результат выполнения 1 запроса

В результате выполнения запроса были выведены 10 ингредиентов, наиболее часто встречающиеся в рецептах (число упоминаний также выведено, порядок вывода – по убыванию).

**Запрос 2**

Текст запроса 2 записан в файл f4.py, содержимое которого представлено на рисунке 8.

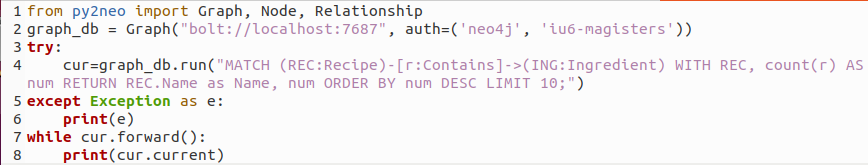


Рисунок 8 – Текст 2 запроса

Результат выполнения показан на рисунке 9.

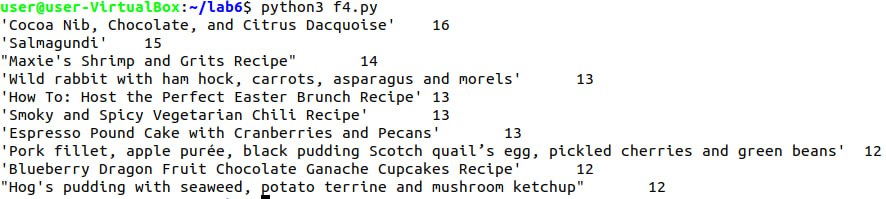


Рисунок 9 – Результат выполнения 2 запроса

В результате выполнения запроса были выведены 10 рецептов, требующих наибольшее количество ингредиентов для приготовления (число ингредиентов также выведено, порядок вывода – по убыванию).

**Запрос 3**

Текст запроса 10 записан в файл f5.py, содержимое которого представлено на рисунке 10.

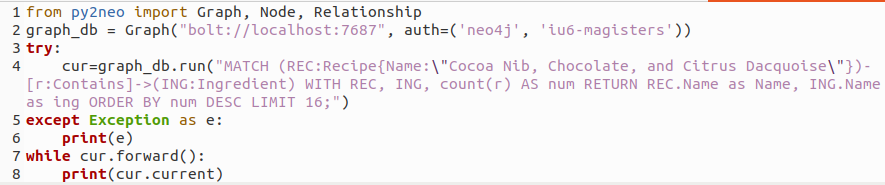


Рисунок 10 – Текст 3 запроса

Результат выполнения показан на рисунке 11.

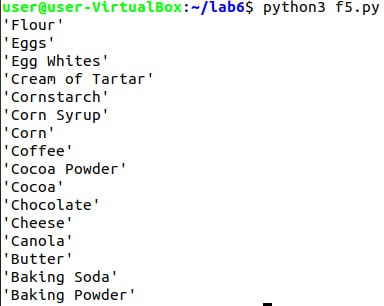


Рисунок 11 – Результат выполнения 3 запроса

В результате выполнения запроса были выведены 16 ингредиентов, используемых в рецепте, требующем наибольшее число ингредиентов для приготовления (рецепт взят из предыдущего запроса).

**Задача 1**

Текст выполнения 1 задачи записан в файл f6.py, содержимое которого представлено на рисунке 12.

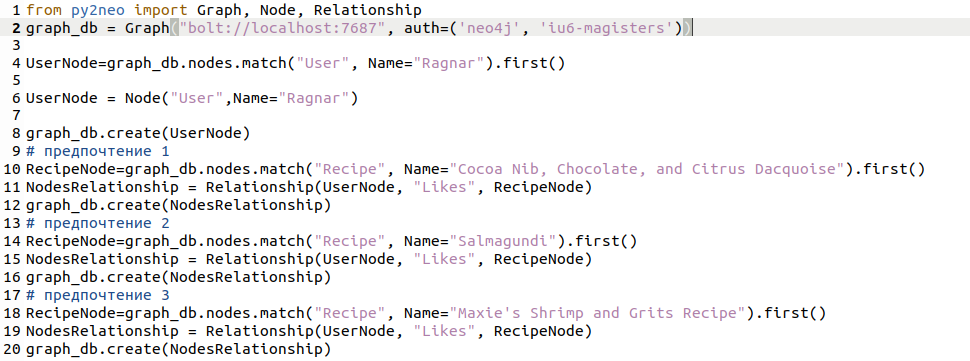


Рисунок 12 – Текст 1 задачи

В результате выполнения задачи 1 был создан пользователь Ragnar и описаны его предпочтения (рецепты "Cocoa Nib, Chocolate, and Citrus Dacquoise", "Salmagundi", "Maxie's Shrimp and Grits Recipe").

**Запрос 4**

Текст запроса 4 записан в файл f7.py, содержимое которого представлено на рисунке 13.

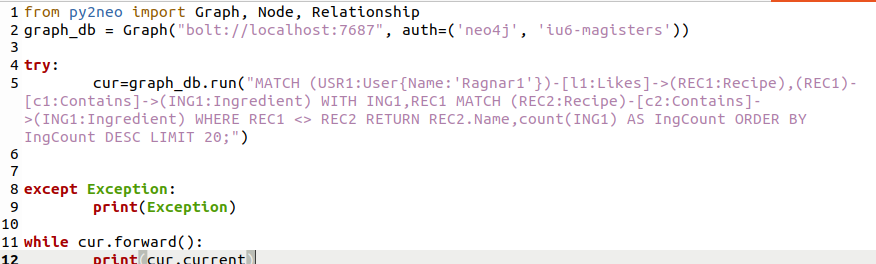


Рисунок 13 – Текст 4 запроса

Результат выполнения показан на рисунке 14.

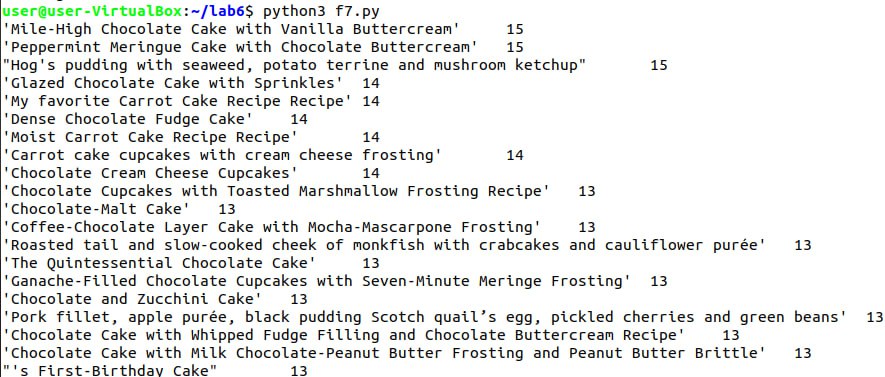


Рисунок 14 – Результат выполнения 4 запроса

В результате выполнения запроса 4 были выведены 20 рецептов, которые могут быть рекомендованы пользователю Ragnar на базе его предпочтений (число ингредиентов также выведено, порядок вывода – по убыванию).

**Вывод:** при выполнении лабораторной работы была изучена работа графовой базы данных Neo4j и её взаимодействие с документной NoSQL БД Elasticsearch на примере разработки рекомендательной системы. Был построен граф, узлами которого являются рецепты кулинарных блюд и ингредиенты, которые входят в эти блюда (рецепты). Для пользователя Ragnar были определены предпочтения некоторым блюдам, а также были были рекомендованы другие блюда, ингредиенты которых входят в предпочтительные блюда пользователя в максимальном количестве.