Департамент образования и науки города Москвы

Государственное автономное образовательное учреждение высшего образования города Москвы

«Московский городской педагогический университет»

Институт цифрового образования

Департамент информатики, управления и технологий

ДИСЦИПЛИНА:

Инструменты для хранения и обработки больших данных

**Лабораторная работа №2.1**

**Тема:**

**«**Изучение методов хранения данных на основе NoSQL.**»**

Выполнила: Ванярина Ю. А., группа: АДЭУ-211

Преподаватель: Босенко Т. М.

Москва

2024

## Теоретическая часть

MongoDB: документо-ориентированная NoSQL база данных, где данные хранятся в формате JSON-подобных документов.

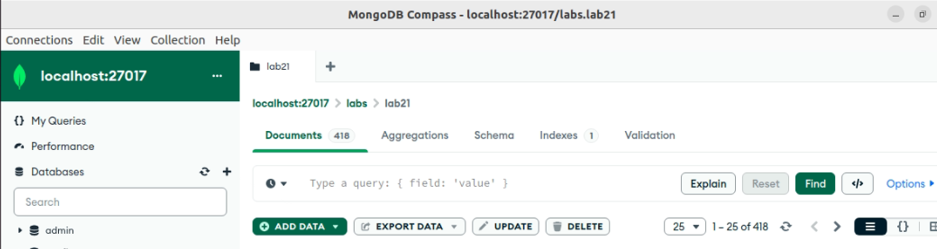
Redis: высокопроизводительная база данных типа "ключ-значение", часто используемая для кеширования и временного хранения данных.

Neo4j: графовая база данных, которая позволяет хранить данные в виде вершин и рёбер графа, что удобно для моделирования сложных взаимосвязей.

## Ход работы

Проверить соединение с Mongo

Запустить Mongo DB Compass



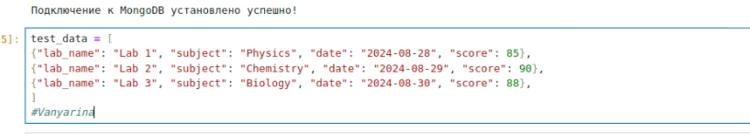
**Шаг 1**: Установка и настройка MongoDB в Jupyter Hub

Установка необходимых библиотек

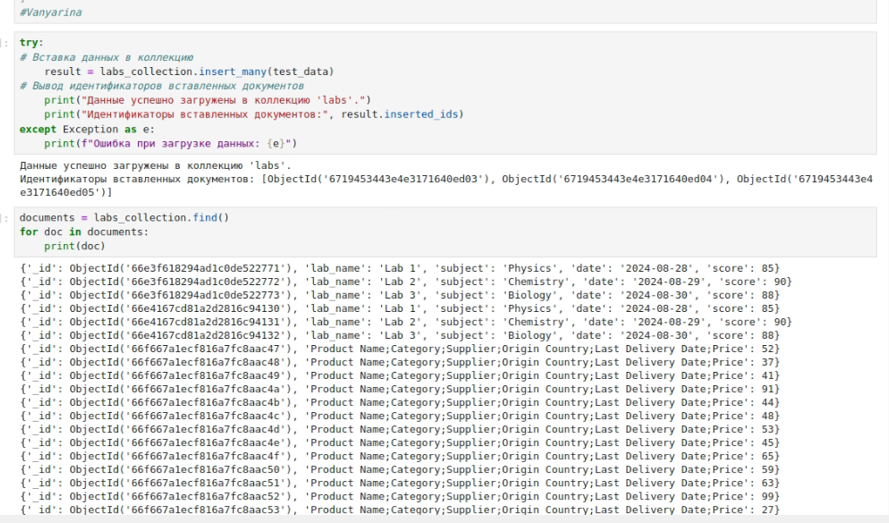
Подключение к MongoDB с аутентификацией



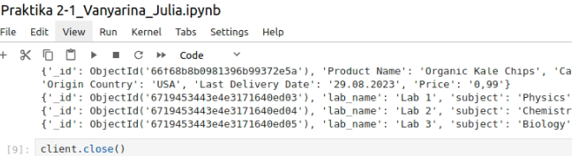
Создание тестовых данных



Загрузка данных в коллекцию labs

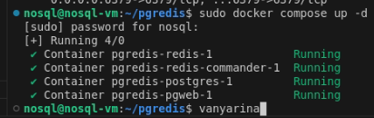


Закрытие подключения



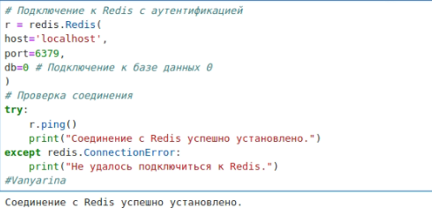
Этот пример демонстрирует, как подключиться к MongoDB с использованием аутентификации, выбрать базу данных и коллекцию, выполнить операции с данными и закрыть соединение.

Проверить соединение с Redis

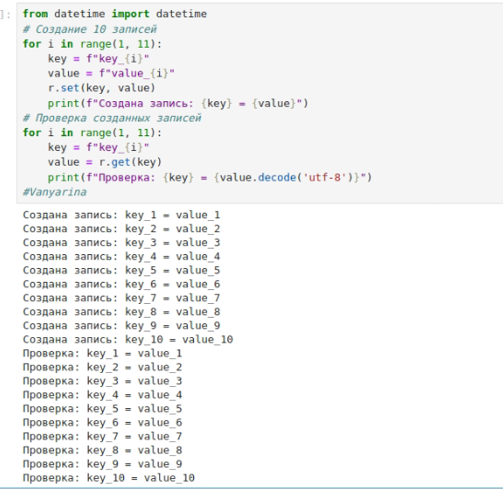


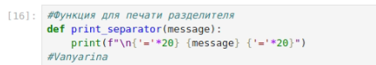


2.4 После установки библиотеки можно подключиться к Redis и выполнить необходимые операции



Генерация и загрузка данных в Redis. Сгенерируем десять записей и загрузим их в базу данных Redis.

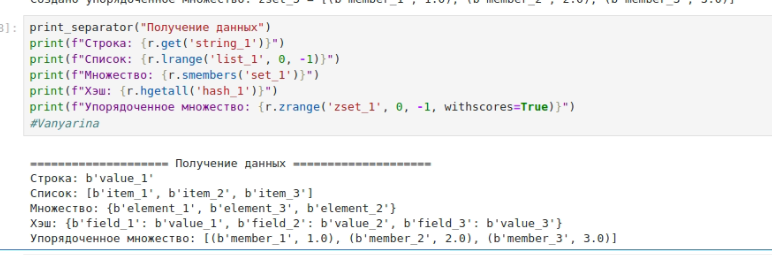




Создание по 5 записей различных типов данных.



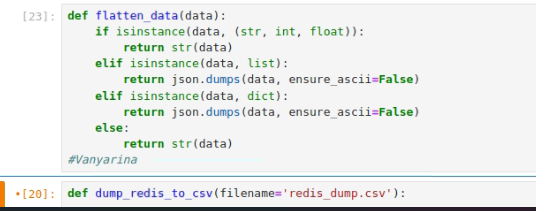
Получение данных по ключу.



2.9Удаления записей



Выгрузить все данные из Redis в csv.





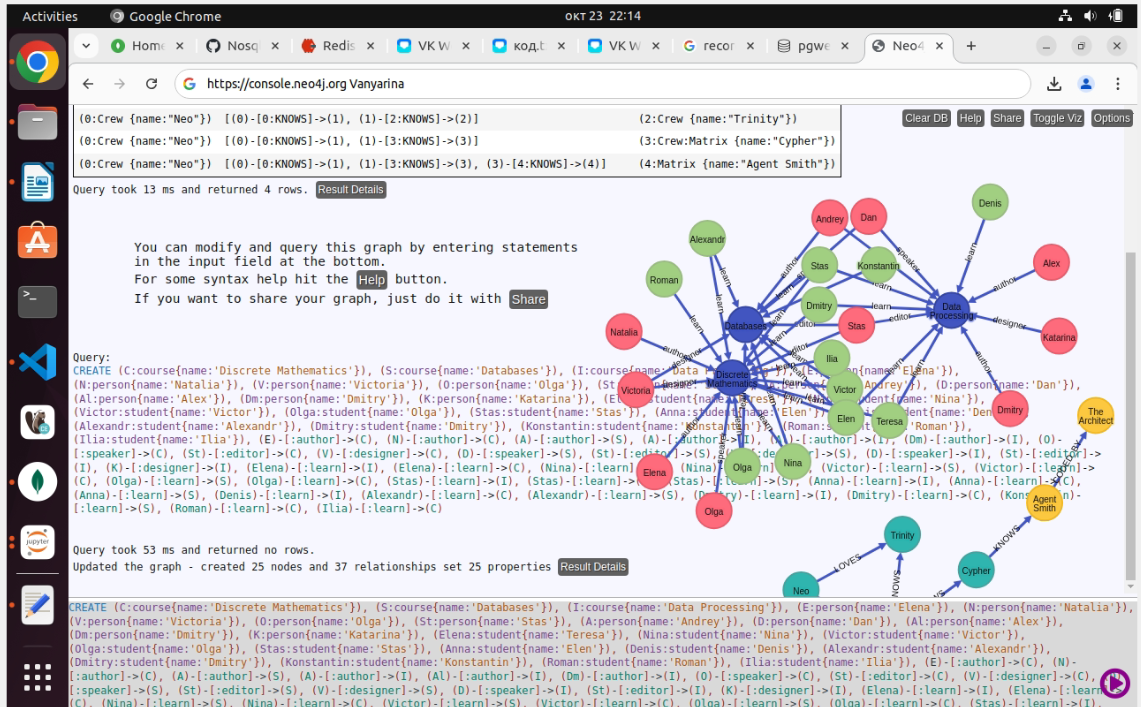


## Neo4j

Neo4j — это графовая система управления базами данных с открытым исходным кодом, реализованная на Java.

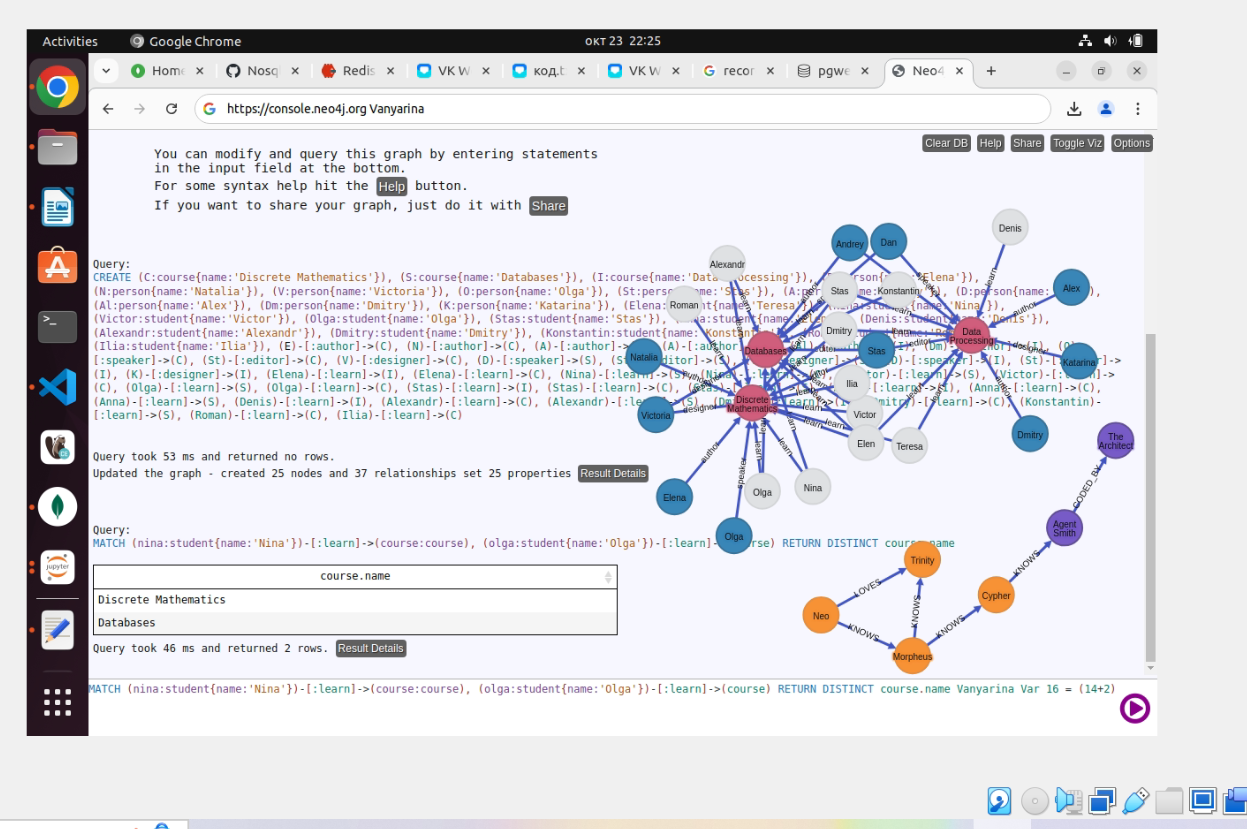
Модель Neo4j в основном состоит из следующих основных компонентов:

* Узлы (эквивалентны вершинам в теории графов): основные элементы данных, которые связаны отношениями. Узлы могут иметь метки и свойства.
* Отношения (эквивалентно ребрам в теории графов): описывает соединения между узлами и соединяет их вместе. Отношения могут иметь одно или несколько свойств.
* Метки: представляют роль узлов. Метки используются для группировки узлов. У каждого узла может быть несколько меток. Метки также индексируются для ускорения процесса поиска узлов в графе.
* Свойства: атрибуты узлов и отношений, включающих пары имен или значений.



Вариант 2. Self: Составьте запрос, который вернет список курсов, на которые записаны студенты с именем "Nina" и "Olga"

MATCH(nina:student{name:'Nina'})-[:learn]->(course:course), (olga:student{name:'Olga'})-[:learn]->(course) RETURN DISTINCT course.name



## Индивидуальное задание.Вариант 16.

Оборудование и его спецификации в производстве:

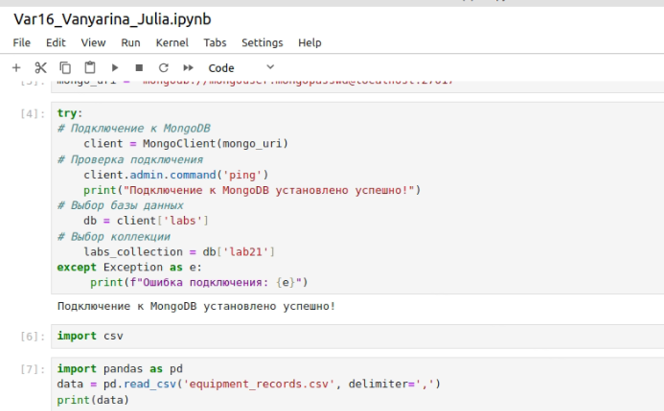
* ID оборудования
* Название оборудования
* Тип
* Производитель
* Дата установки
* Мощность
* Габариты (ДxШxВ)
* Вес (кг)
* Срок службы (лет)
* Статус (работает/на обслуживании)

Проверяем соединение в Монго и импортируем данные

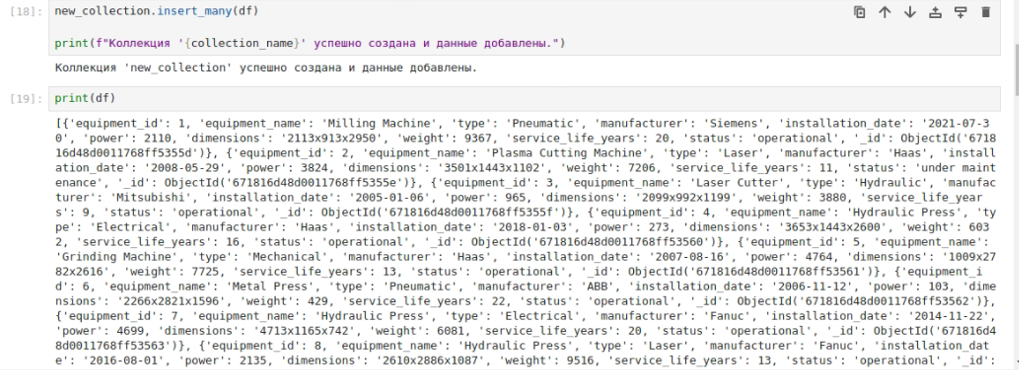
```

*ID,name,type,manufactor,date\_instalation,power,size,weight,service\_life,status 1,Robotic arm,Mechanical,Haas,2013-11-04,1168,4455x2860x1144,8725,24,works 2,Milling machine,Laser,Mazak,2024-08-15,4946,4541x1194x1151,9434,16,in service 3,Grinding machine,Laser,Yaskawa,2007-12-14,2759,1085x2733x2035,6832,25,works 4.3D Printer,Robotic,Fanuc,2012-12-05,3036,2291x1253x1203,2424,15,works 5,Milling machine,Mechanical,ABB,2017-08-23,3389,3213x2321x895,6384,11,works 6.3D printer,Laser,Siemens,2013-10-08,4228,2650x2618x2033,682,14,works*

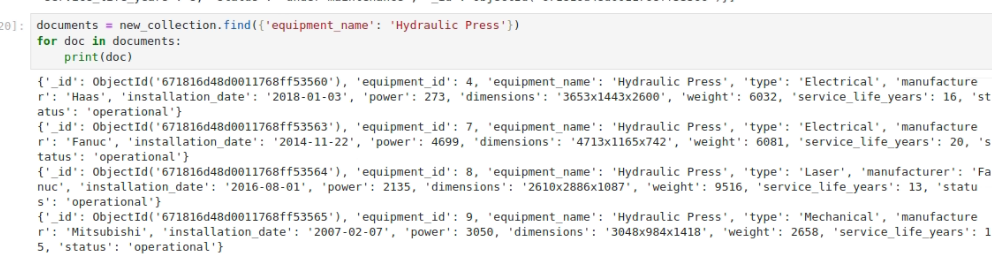
```



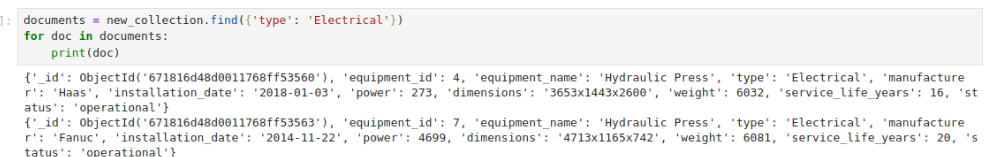
Создание новой коллекции



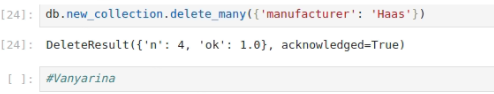
Поиск записей с наименованием гидравлический пресс



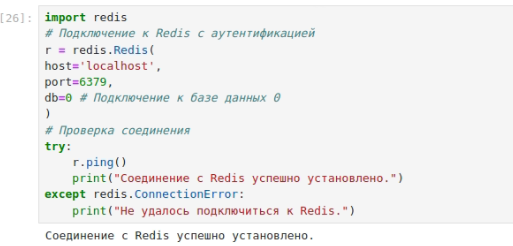
Выборка по типу оборудования Электрический

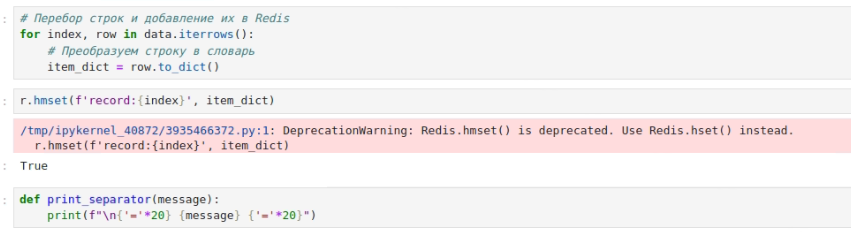


Удаление производителя Haas

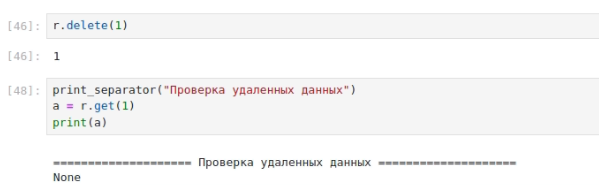


Подключение к Redis









Выводы по использованию инструментов по хранению данных nosql

1. MongoDB:

MongoDB является документо-ориентированной NoSQL базой данных, которая хранит данные в формате документов JSON.

+MongoDB хорошо подходит для хранения больших объемов текстовых данных.

-Было сложно разобраться с установкой монго,так как последняя версия не подходила, были перепробованы все версии до 2016 года. Много раз не запускался контейнер.При подключении в монго компас отсутствовала бд student.Получилоcь выполнить задание после 8 попыток установки на разные конфигурации вм.

2. Redis:

Redis - это высокопроизводительная база данных типа ключ-значение, которая часто используется для сеансов хранения кеширования и обработки очередей сообщений.

+Его скорость и способность к хранению данных в оперативной памяти делают его отличным выбором для обработки высоконагруженных приложений.

3. Neo4j:

Neo4j - графовая база данных, которая хранит данные в виде узлов, связей и свойств. Он отлично подходит для хранения и обработки сложных структур данных, таких как социальные сети.

+Neo4j обеспечивает эффективную работу с запросами, а также обеспечивает глубокие аналитические возможности.

Понятнее и удобнее всего работа проходила с Redis.