

Домашнее задание. Практическая работа

Дисциплина

Системы хранения данных

Тема

SQL в Hadoop

Форма проверки

Домашнее задание с самопроверкой.

Совет: выполните домашнее задание сразу, как только изучите тему.

Имя преподавателя

Вадим Заигрни

Время выполнения

60 минут.

Цель задания

Научиться работать в Trino

Инструменты для выполнения ДЗ

- Docker
- Trino

Правила приёма работы

- Прикрепите в LMS текстовый файл с результатами работы запросов из пункта 4 задания.

Важно:

- убедитесь, что к файлу есть доступ;
- название файла должно содержать фамилию и имя студента, номер и название ДЗ.

Чек-лист самопроверки

Задание считается выполненным, если:

- Прикреплён файл с выполненным заданием;
- Файл содержит выполненные действия по заданию;
- доступ к материалам открыт.

Задание не выполнено, если:

- файл с заданием не прикреплён или отсутствует доступ к нему;
- файл не содержит выполненные действия по заданию.

Дедлайн

24.11.25

Описание задания

1. Запустите Trino

```
docker run --name trino -d -p 8080:8080 trinodb/trino
```

2. Откройте веб-интерфейс Trino

<http://localhost:8080>

3. Подключитесь к Trino CLI

```
docker exec -it trino trino
SELECT version();
SELECT count(*) FROM tpch.sf1.region;
SELECT * FROM tpch.sf1.region;
SELECT count(*) FROM tpch.sf1.nation;
SELECT * FROM tpch.sf1.nation;
SELECT count(*) FROM tpch.sf1.customer;
SELECT * FROM tpch.sf1.customer LIMIT 20;
```

4. Подключитесь к Trino по JDBC и выполните запросы

host: localhost

port: 8080

user: guest

4.1. Pricing summary report query (Q1)

```
SELECT
returnflag,
linestatus,
sum(quantity) AS sum_qty,
sum(extendedprice) as sum_base_price,
sum(extendedprice * (1 - discount)) AS sum_disc_price,
sum(extendedprice * (1 - discount) * (1 + tax)) AS sum_charge,
avg(quantity) AS avg_qty,
avg(extendedprice) AS avg_price,
avg(discount) AS avg_disc,
count(*) AS count_order
FROM tpch.sf1.lineitem
WHERE shipdate <= date '1998-12-01' - INTERVAL '60' DAY
GROUP BY
returnflag,
linestatus
ORDER BY
returnflag,
linestatus
;
```

4.2. Minimum cost supplier query (Q2)

```
SELECT
    s.acctbal,
    s.name,
    n.name,
    p.partkey,
    p.mfgr,
    s.address,
    s.phone,
    s.comment
FROM
    tpch.sf1.part p,
    tpch.sf1.supplier s,
    tpch.sf1.partsupp ps,
    tpch.sf1.nation n,
    tpch.sf1.region r
WHERE
    p.partkey = ps.partkey
    AND s.suppkey = ps.suppkey
    AND p.size = 15
    AND p.type like '%BRASS'
    AND s.nationkey = n.nationkey
    AND n.regionkey = r.regionkey
    AND r.name = 'EUROPE'
    AND ps.supplycost = (
        SELECT min(ps.supplycost)
        FROM
            tpch.sf1.partsupp ps,
            tpch.sf1.supplier s,
            tpch.sf1.nation n,
            tpch.sf1.region r
        WHERE
            p.partkey = ps.partkey
            AND s.suppkey = ps.suppkey
            AND s.nationkey = n.nationkey
            AND n.regionkey = r.regionkey
            AND r.name = 'EUROPE'
    )
ORDER BY
    s.acctbal DESC,
    n.name,
    s.name,
    p.partkey
;
```

4.3. Прикрепите в ЛМС текстовый файл с результатами работы обоих запросов.