# Министерство науки и высшего образования Российской Федерации ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

#### ОТЧЕТ

# ПО ЛАБОРАТОРНОЙ РАБОТЕ №4

«Запросы на выборку и модификацию данных. Представления. Работа с индексами»

по дисциплине «Проектирование и реализация баз данных»

Обучающийся Клименков Владислав Максимович Факультет прикладной информатики Группа К3241 Направление подготовки 09.03.03 Прикладная информатика Образовательная программа Мобильные и сетевые технологии 2023 Преподаватель Говорова Марина Михайловна

# 1 Цель работы

Овладеть практическими навыками создания представлений и запросов на выборку данных к базе данных PostgreSQL, использования подзапросов при модификации данных и индексов.

# 2 Практическое задание

- Создать запросы и представления на выборку данных к базе данных PostgreSQL (согласно индивидуальному заданию лабораторной работы №2, часть 2 и 3).
- 2. Составить 3 запроса на модификацию данных (INSERT, UPDATE, DELETE) с использованием подзапросов.
- 3. Изучить графическое представление запросов и просмотреть историю запросов.
- 4. Создать простой и составной индексы для двух произвольных запросов и сравнить время выполнения запросов без индексов и с индексами. Для получения плана запроса использовать команду EXPLAIN.

## 3 Схема базы данных (ЛР 3)

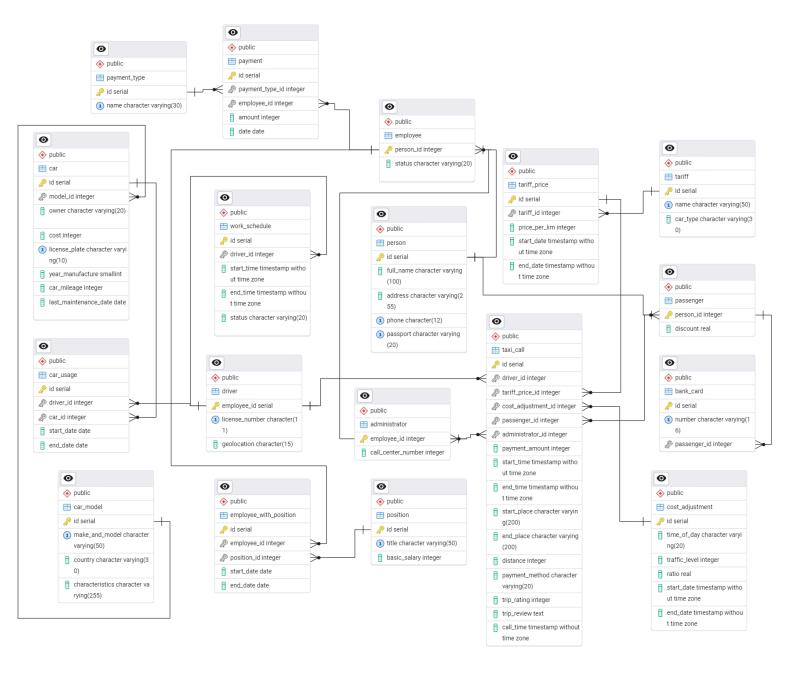


Рисунок 1 – ERD-схема базы данных

## 4 Выполнение

# 4.1 Запросы к базе данных

# Запрос №1

# Описание запроса:

Вывести данные о водителе, который чаще всех доставляет пассажиров на заданную улицу.

# SQL-код запроса:

WITH delivery counts AS (

```
SELECT
       d.employee_id,
       p.full_name,
       COUNT(tc.id) AS delivery count
    FROM
       driver d
    JOIN
       taxi call tc ON d.employee id = tc.driver id
       public.person p ON d.employee id = p.id
   WHERE
       tc.end_place LIKE '%Восточная%'
   GROUP BY
       d.employee_id, p.full_name
)
SELECT
   employee_id,
   full_name,
   delivery_count
FROM
   delivery counts
WHERE
   delivery_count = (SELECT MAX(delivery_count) FROM delivery_counts)
ORDER BY
    employee_id;
```

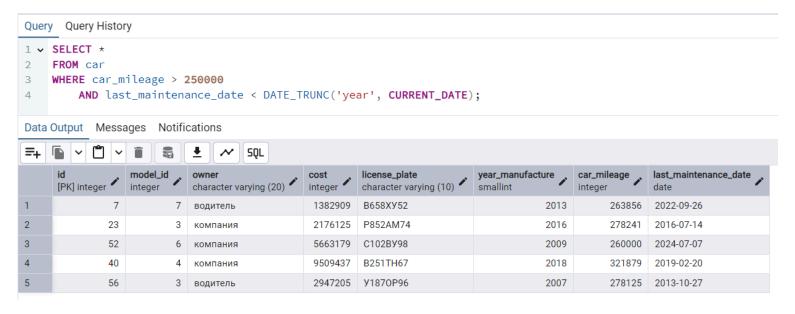
```
Query
      Query History
 5
             COUNT(tc.id) AS delivery_count
6
             driver d
7
         JOIN
8
             taxi_call tc ON d.employee_id = tc.driver_id
9
10
             public.person p ON d.employee_id = p.id
11
12
         WHERE
             tc.end_place LIKE '%Восточная%'
13
         GROUP BY
14
15
             d.employee_id, p.full_name
16
     )
17
     SELECT
18
19
         employee_id,
20
         full_name,
21
         delivery_count
22
     FROM
23
         delivery_counts
24
     WHERE
         delivery_count = (SELECT MAX(delivery_count) FROM delivery_counts)
25
26
     ORDER BY
         employee_id;
27
Data Output Messages Notifications
    SQL
     employee_id
                 full_name
                                         delivery_count
                 character varying (100)
     integer
                                         bigint
             148 Леонтий Викторович Лобанов
```

# Описание запроса:

Вывести данные об автомобилях, которые имеют пробег более 250 тысяч. километров и которые не проходили ТО в текущем году.

# SQL-код запроса:

```
SELECT *
FROM car
WHERE car_mileage > 250000
    AND last maintenance_date < DATE_TRUNC('year', CURRENT_DATE);</pre>
```



### Описание запроса:

Сколько раз каждый пассажир воспользовался услугами таксопарка?

## SQL-код запроса:

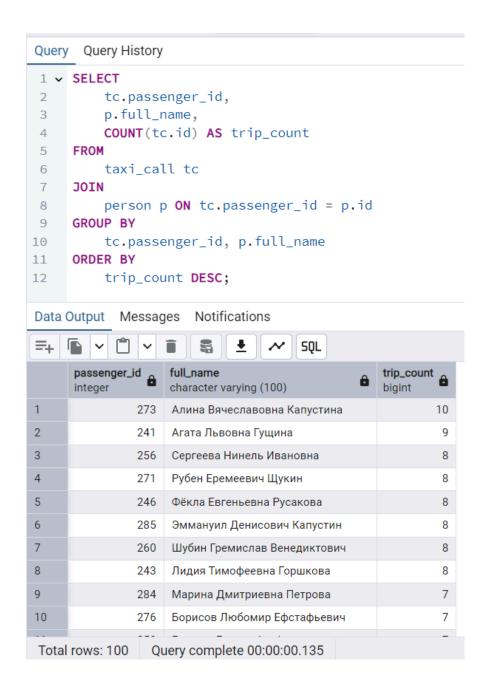
```
tc.passenger_id,
   p.full_name,
   COUNT(tc.id) AS trip_count

FROM
   taxi_call tc

JOIN
   person p ON tc.passenger_id = p.id

GROUP BY
   tc.passenger_id, p.full_name

ORDER BY
   trip count DESC;
```



## Описание запроса:

Вывести данные пассажира, который воспользовался услугами таксопарка максимальное число раз.

# SQL-код запроса:

```
WITH trip_counts AS (

SELECT

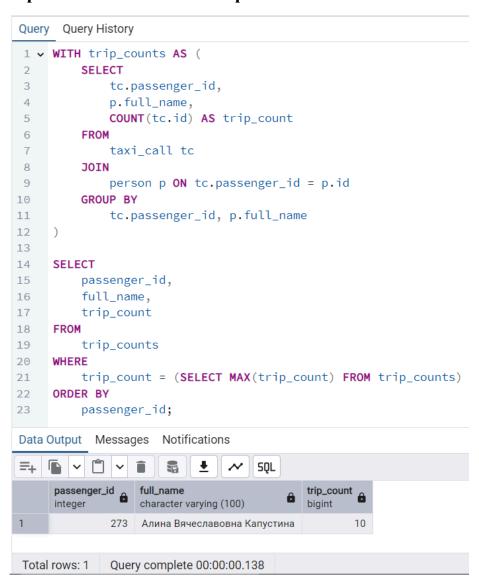
tc.passenger_id,

p.full_name,

COUNT(tc.id) AS trip_count

FROM

taxi call tc
```

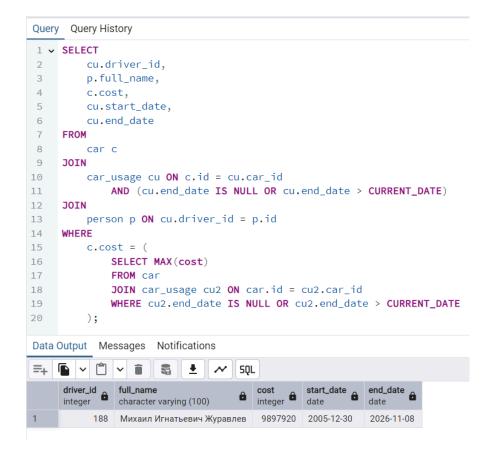


# Описание запроса:

Вывести данные о водителе, который ездит на самом дорогом автомобиле.

# SQL-код запроса:

```
SELECT
  cu.driver_id,
   p.full name,
   c.cost,
   cu.start date,
   cu.end date
FROM
   car c
JOIN
   car_usage cu ON c.id = cu.car_id
       AND (cu.end date IS NULL OR cu.end date > CURRENT DATE)
   person p ON cu.driver_id = p.id
WHERE
   c.cost = (
       SELECT MAX(cost)
       FROM car
       JOIN car_usage cu2 ON car.id = cu2.car_id
       WHERE cu2.end_date IS NULL OR cu2.end_date > CURRENT_DATE
   );
```



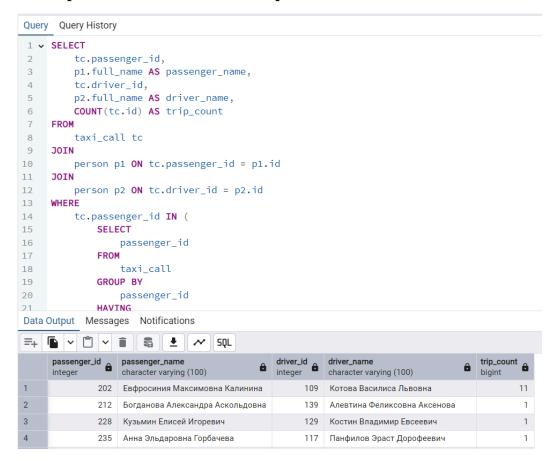
## Описание запроса:

Вывести данные пассажира, который всегда ездит с одним и тем же водителем.

## SQL-код запроса:

```
SELECT
    tc.passenger_id,
    p1.full name AS passenger name,
    tc.driver_id,
    p2.full name AS driver name,
    COUNT(tc.id) AS trip count
FROM
    taxi call tc
JOIN
   person p1 ON tc.passenger id = p1.id
JOIN
   person p2 ON tc.driver id = p2.id
WHERE
    tc.passenger_id IN (
        SELECT
            passenger_id
```

## Скриншот выполнения запроса:



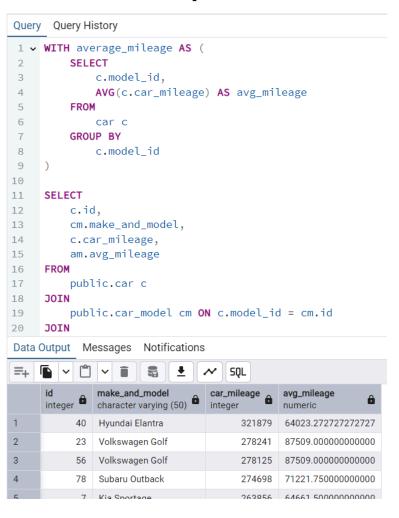
# Запрос №7

## Описание запроса:

Какие автомобили имеют пробег больше среднего пробега для своей марки.

# SQL-код запроса:

```
AVG(c.car mileage) AS avg mileage
    FROM
        car c
    GROUP BY
        c.model id
)
SELECT
    c.id,
    cm.make and model,
    c.car_mileage,
    am.avg_mileage
FROM
    public.car c
JOIN
    public.car model cm ON c.model id = cm.id
JOIN
    average mileage am ON c.model id = am.model id
WHERE
    c.car mileage > am.avg mileage;
```



# 4.2 Представления

# Представление №1

## Описание представления:

Создать представление, содержащее сведения о незанятых на данный момент водителях.

# SQL-код представления:

```
CREATE OR REPLACE VIEW available_drivers AS

SELECT

p.id,
p.full_name,
ws.start_time,
ws.end_time

FROM
work_schedule ws

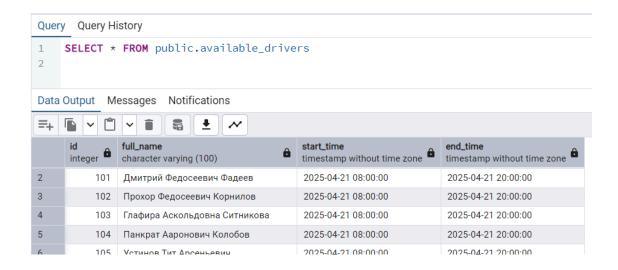
JOIN
person p ON ws.driver_id = p.id

WHERE

NOW() BETWEEN ws.start_time AND ws.end_time
AND ws.status = 'вышел';
```

## Скриншоты создания и отображения представления:

```
Query Query History
1 ➤ CREATE OR REPLACE VIEW available_drivers AS
2 SELECT
3
        p.id,
4
       p.full_name,
5
       ws.start_time,
        ws.end_time
6
7
    FROM
8
        work_schedule ws
9 JOIN
        person p ON ws.driver_id = p.id
10
11 WHERE
       NOW() BETWEEN ws.start_time AND ws.end_time
12
        AND ws.status = 'вышел';
Data Output Messages Notifications
CREATE VIEW
Query returned successfully in 563 msec.
```



## Представление №2

#### Описание представления:

Создать представление с зарплатой всех водителей за вчерашний день.

## SQL-код представления:

```
CREATE OR REPLACE VIEW driver_salary_yesterday AS
SELECT
    pay.employee id,
    per.full name,
    payt.name AS payment type name,
    pay.date,
    pay.amount
FROM
    payment pay
JOIN
    driver d ON pay.employee id = d.employee id
JOIN
    person per ON d.employee id = per.id
JOIN
    payment type payt ON pay.payment type id = payt.id
WHERE
    payt.name = 'Зарплата'
    AND pay.date = CURRENT DATE - INTERVAL '1 day';
```

#### Скриншоты создания и отображения представления:



Query Ustory Query History					
1 SELECT * FROM public.driver_salary_yesterday 2					
Data Output Messages Notifications					
	employee_id integer	full_name character varying (100)  €	payment_type_name character varying (30)	date date	amount integer
1	101	Дмитрий Федосеевич Фадеев	Зарплата	2025-04-20	1904
2	102	Прохор Федосеевич Корнилов	Зарплата	2025-04-20	4434
3	103	Глафира Аскольдовна Ситникова	Зарплата	2025-04-20	9781
4	104	Панкрат Ааронович Колобов	Зарплата	2025-04-20	2607
5	105	Устинов Тит Арсеньевич	Зарплата	2025-04-20	7112
6	106	Никодим Харлампьевич Гаврилов	Зарплата	2025-04-20	2059

# 4.3 Запросы на модификацию данных

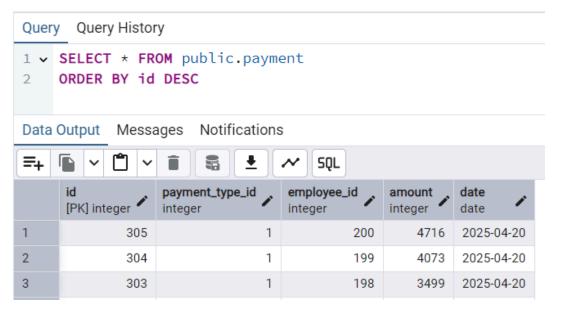
# INSERT-запрос

# Описание запроса:

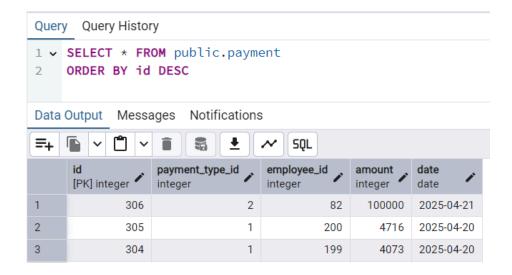
Выплатить премию в 100000 рублей самым долго работающим сотрудникам.

# SQL-код запроса:

INSERT INTO payment (payment\_type\_id, employee\_id, amount, date)



```
Query Query History
1 ➤ INSERT INTO payment (payment_type_id, employee_id, amount, date)
2
     SELECT
3
         employee_id,
4
         100000,
5
         CURRENT_DATE
6
 7
     FROM
         employee_with_position
8
9
     WHERE
         end_date IS NULL OR end_date > CURRENT_DATE
10
11
         AND start_date = (SELECT MIN(start_date)
12
                            FROM employee_with_position
                            WHERE end_date IS NULL OR end_date > CURRENT_DATE);
13
Data Output Messages Notifications
INSERT 0 1
Query returned successfully in 116 msec.
```



## UPDATE-запрос

#### Описание запроса:

Обновить на текущую все даты технического обслуживания для всех автомобилей компании с заданной маркой, которые используются хотя бы одним водителем.

## SQL-код запроса:

```
UPDATE car

SET last_maintenance_date = CURRENT_DATE

WHERE owner = 'компания'

AND model_id = 2

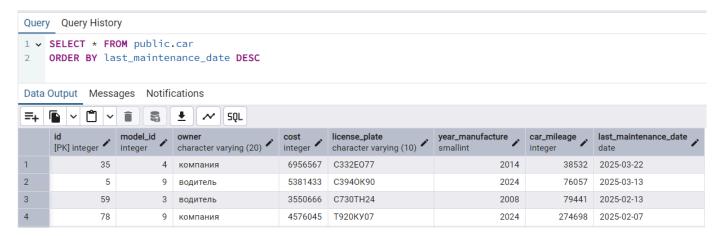
AND id IN (

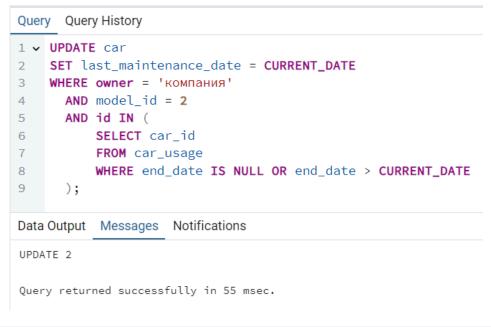
SELECT car_id

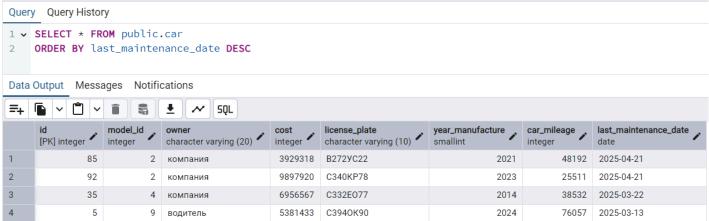
FROM car_usage

WHERE end_date IS NULL OR end_date > CURRENT_DATE

);
```







## DELETE-запрос

## Описание запроса:

Удалить всех пассажиров, у которых нет привязанной банковской карты и которые не совершили ни одного вызова такси.

# SQL-код запроса:

```
DELETE FROM passenger
WHERE person_id NOT IN (
        SELECT passenger_id FROM bank_card
)
AND person_id NOT IN (
        SELECT passenger_id FROM taxi_call
);
```



#### 4.4 Создание индексов

# Простой индекс

#### SQL-код запроса:

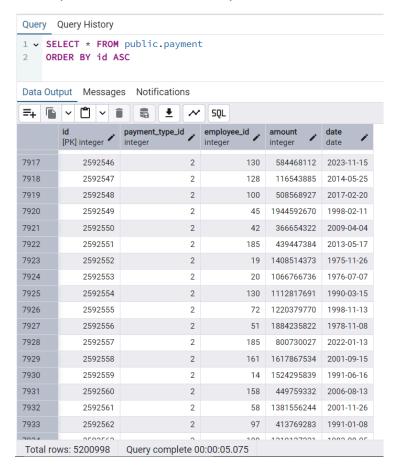
```
SELECT * FROM payment WHERE amount > 1999999000 ORDER BY amount DESC;
```

#### SQL-код индекса:

CREATE INDEX idx amount ON payment(amount);

## Скриншоты работы с индексом:

Чтобы нагляднее увидеть влияние индекса заполним таблицу payment большим количеством записей (больше 5 000 000) с различными значениями в столбце amount (от 1 до 2 000 000 000):



Проверим скорость выполнения запроса без индекса:

```
Query Query History

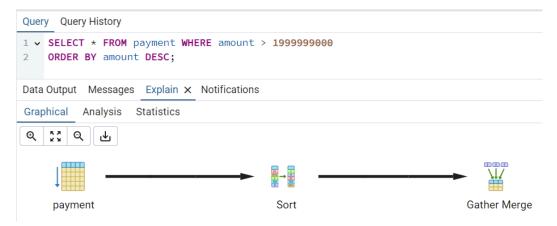
1  SELECT * FROM payment WHERE amount > 1999999000
2  ORDER BY amount DESC;

Data Output Messages Explain × Notifications

Successfully run. Total query runtime: 447 msec.
1 rows affected.
```

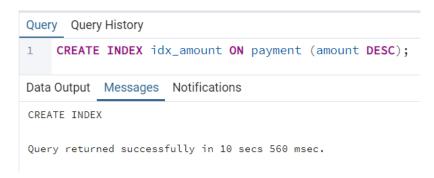
Среднее время выполнения запроса 400-500 мс.

Теперь рассмотрим раздел Explain:



Как можно заметить, использовался механизм "Gather Merge", позволяющий разбить вычисления на несколько параллельных процессов, а потом объединить их результаты. Это позволяет увеличить скорость выполнения запроса.

Теперь создадим индекс для столбца "amount" (используем DESC, так как в оптимизированном запросе используется "ORDER BY ... DESC"):



Выполним тот же запрос, но на этот раз при созданном индексе:

```
Query Query History

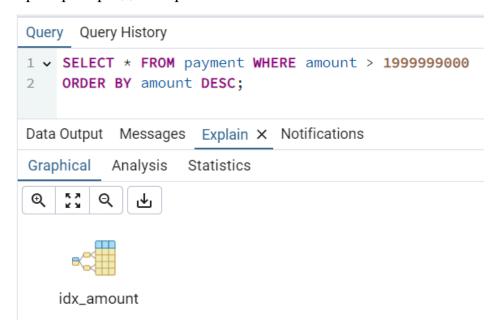
1 V SELECT * FROM payment WHERE amount > 1999999000
2 ORDER BY amount DESC;

Data Output Messages Notifications

Successfully run. Total query runtime: 117 msec.
1 rows affected.
```

Среднее время выполнение запроса снизилось до 100-200 мс. Разница заметна, но не настолько ощутима, так как и без индекса выполнение запроса происходит достаточно быстро (возможно, в том числе и за счёт параллельных вычислений).

## Проверим раздел Explain:



Как можно заметить, в этот раз в качестве плана использовался созданный нами индекс, что и позволило увеличить скорость выполнения.

#### Теперь удалим индекс:

```
Query Query History

1    DROP INDEX idx_amount;

Data Output    Messages    Notifications

DROP INDEX

Query returned successfully in 277 msec.
```

#### Итоговое время:

Без индекса: 400-500 мс в среднем.

С индексом: 100-200 мс в среднем.

#### Составной индекс

## SQL-код запроса:

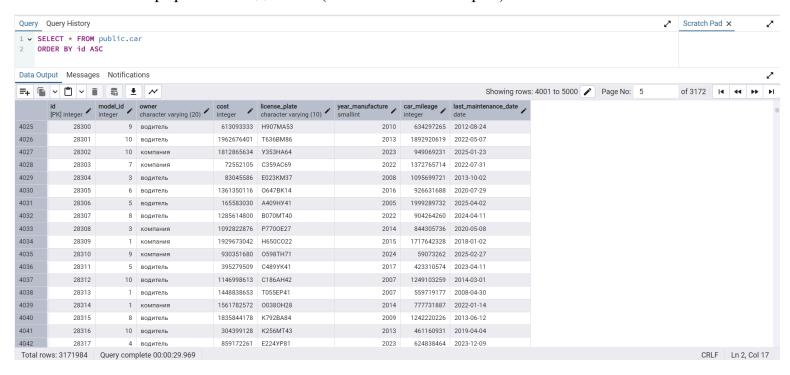
```
SELECT \*
FROM car
WHERE year_manufacture >= 2024 AND car_mileage <= 1000
ORDER BY cost DESC;</pre>
```

## SQL-код индекса:

CREATE INDEX idx\_car\_year\_cost\_mileage ON car (year\_manufacture, cost DESC,
car mileage);

## Скриншоты работы с индексом:

Заполним таблицу саг большим количеством случайно сгенерированных данных (более 3 000 000 строк):



Проверим скорость выполнения запроса без индекса:

```
Query Query History

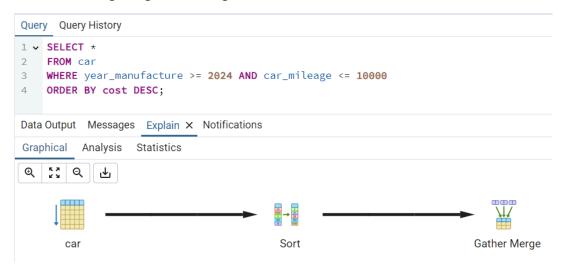
1  SELECT *
FROM car
WHERE year_manufacture >= 2024 AND car_mileage <= 10000
ORDER BY cost DESC;

Data Output Messages Notifications

Successfully run. Total query runtime: 435 msec.
1 rows affected.
```

Среднее время выполнения 400-500 мс.

Рассмотрим раздел Explain:

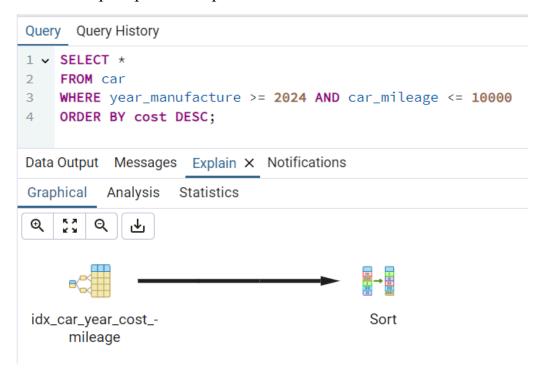


Теперь создадим составной индекс для столбцов "year\_manufacture", "cost DESC" и "car mileage":

Проверим скорость работы запроса при созданном составном индексе:

Среднее время выполнения снизилось до 100-200 мс.

Рассмотрим раздел Explain:



Как можно заметить, созданный индекс применяется при выполнении запроса.

Теперь удалим индекс:

```
Query Query History

1    DROP INDEX idx_car_year_cost_mileage;

Data Output    Messages    Notifications

DROP INDEX

Query returned successfully in 55 msec.
```

#### Итоговое время:

Без индекса: 400-500 мс в среднем.

С индексом: 100-200 мс в среднем.

#### 5 Выводы

В рамках данной работы удалось выполнить все поставленные практические задачи:

- Создать запросы и представления на выборку данных к базе данных PostgreSQL (согласно индивидуальному заданию лабораторной работы №2, часть 2 и 3).
- Составить 3 запроса на модификацию данных (INSERT, UPDATE, DELETE) с использованием подзапросов.
- Создать простой и составной индексы для двух произвольных запросов и сравнить время выполнения запросов без индексов и с индексами. Для получения плана запроса использовать команду EXPLAIN.

При работе с индексами удалось выяснить, что при большом количестве обрабатываемых данных простые и составные индексы способны ускорить выполнение определённых SQL-команд. В рамках работы удалось добиться ускорения с примерно 400-500 мс до 100-200 мс на запрос как для простого, так и для составного индексов.