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:

Проектный практикум по разработке ETL-решений

**Самостоятельная работа**

**Тема:**

«Разработка ETL-процесса для интеграции данных между PostgreSQL и MySQL с использованием Pentaho Data Integration»

Выполнил(а): Морозова Валерия АДЭУ-211

Преподаватель:

Москва

2025

**Задачи:**

− Создать исходные таблицы в PostgreSQL с различными наборами данных.

− Настроить целевые таблицы в MySQL для приема данных.

− Разработать процессы трансформации данных в Pentaho.

− Реализовать механизмы обработки ошибок и валидации данных.

− Создать представления для связанных данных.

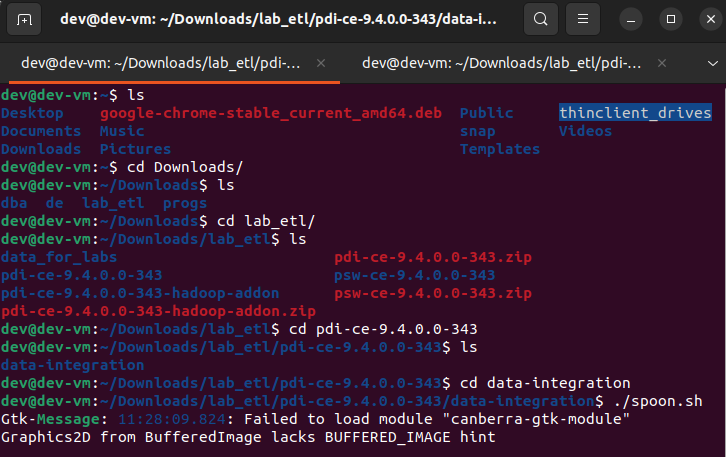


Рисунок 1. Запуск Pentaho Data Integration

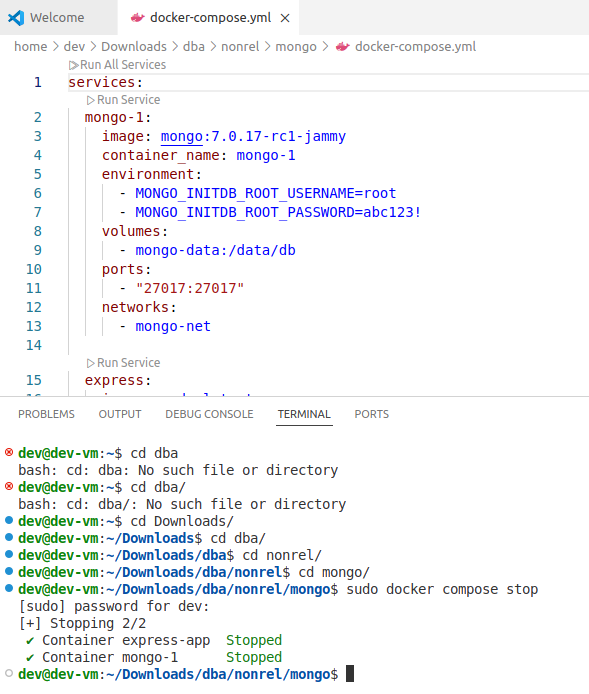


Рисунок 2. Отключение mongo

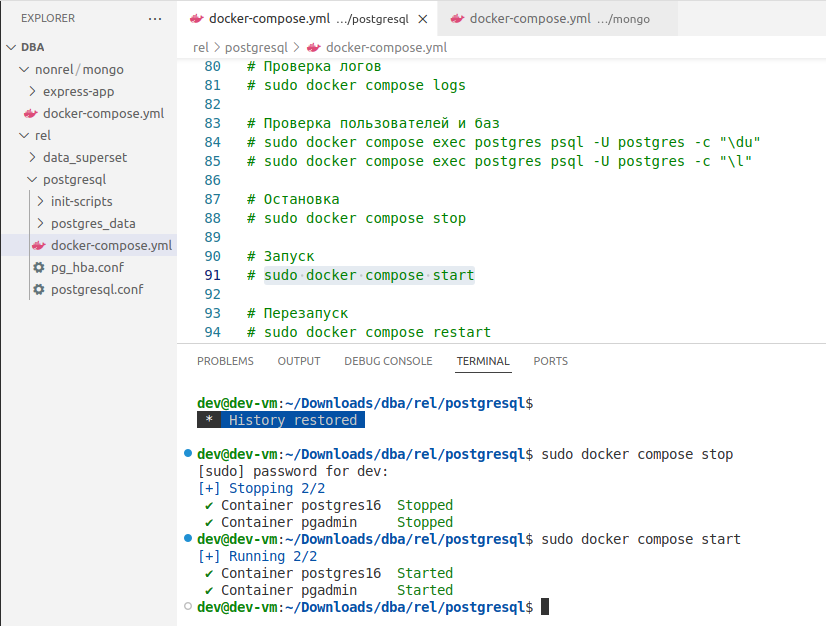


Рисунок 3. Отключение и запуск postgreSQL

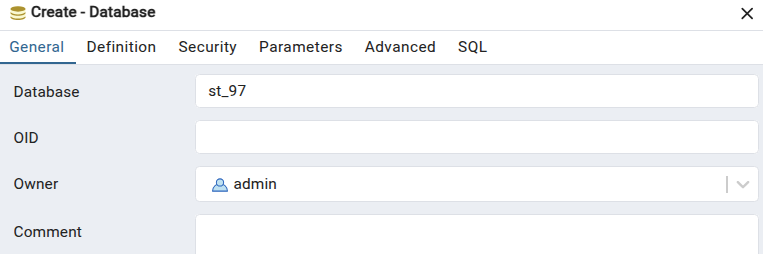


Рисунок 4. Создание собственной базы

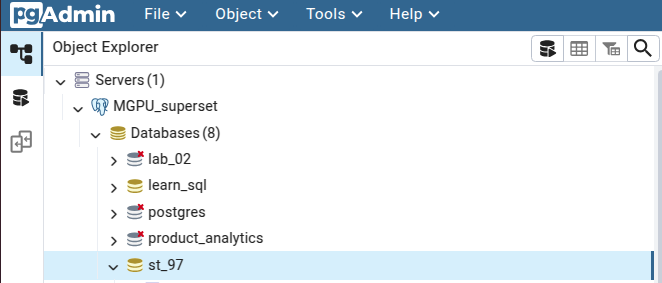


Рисунок 5. База st\_97 успешно создана

Вариант10  
**Задание 1.** Создать таблицу payments (id, customer\_id, amount, date, method)

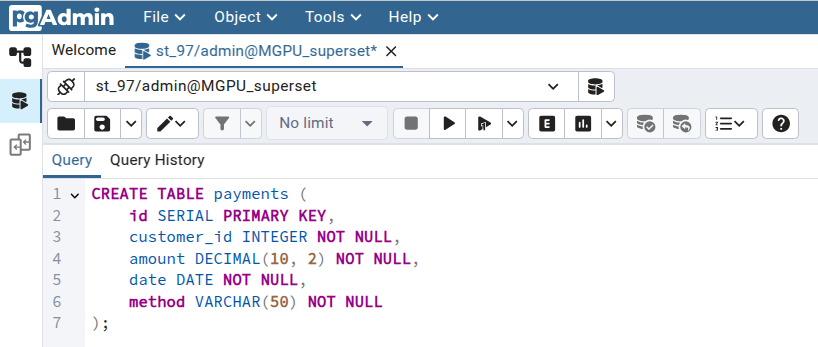


Рисунок 6. Запрос на создание таблицы в postgreSQL

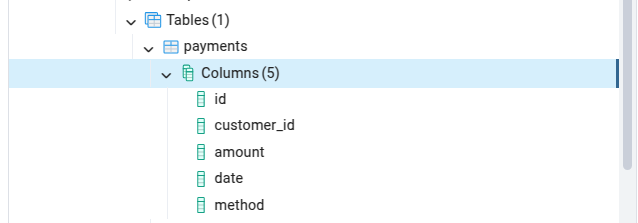


Рисунок 7. Проверка таблицы payments в базе

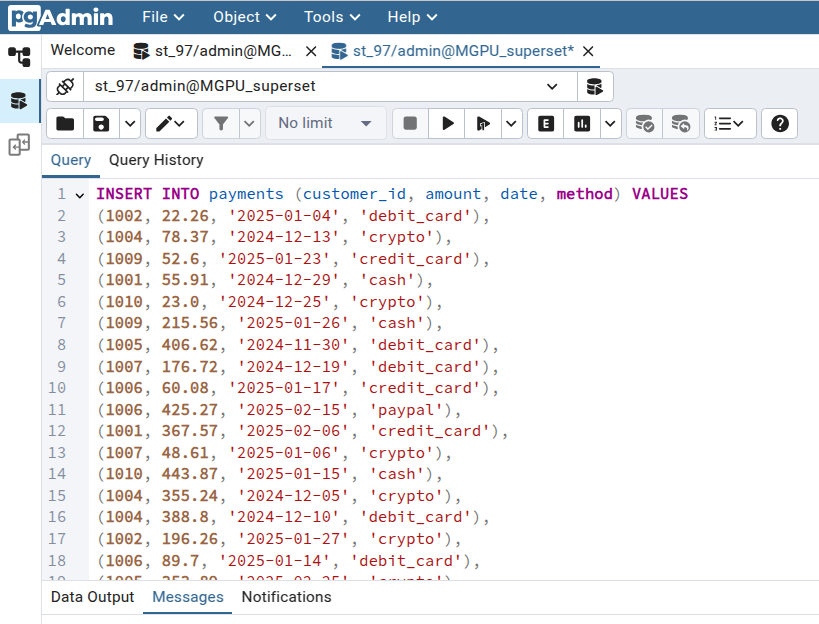


Рисунок 8. Заполнение таблицы сгенерированными данными

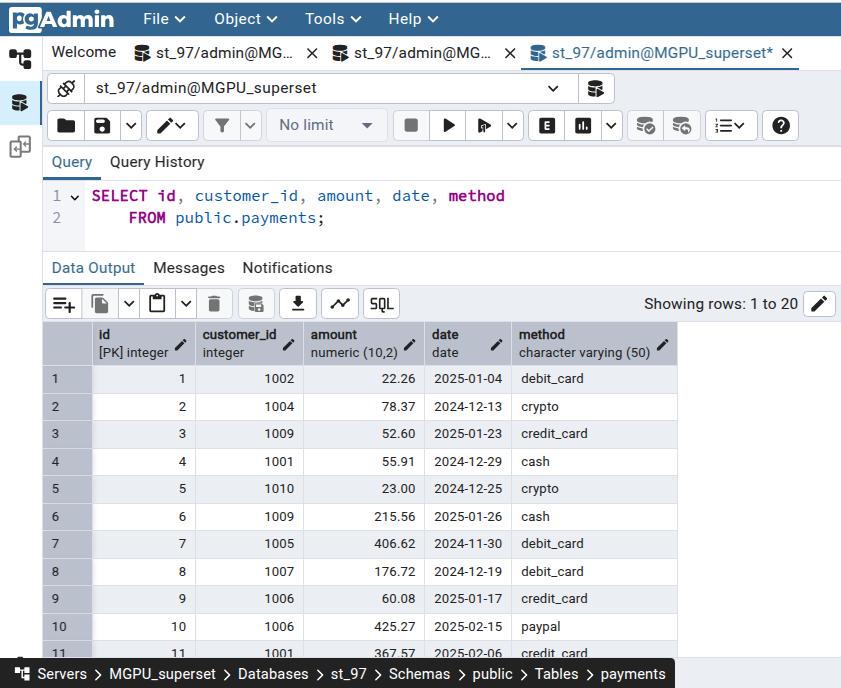


Рисунок 9. Генерация запроса для проверки содержания таблицы

**Задание 2.** Создать таблицу payment\_analytics с полями для анализа транзакций

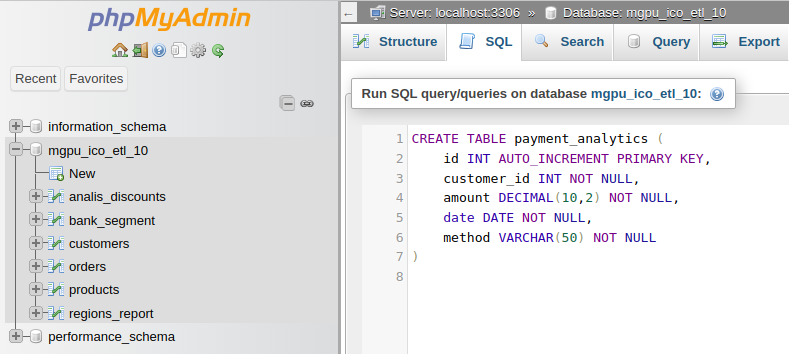


Рисунок 10. Создание таблицы через запрос в MySQL

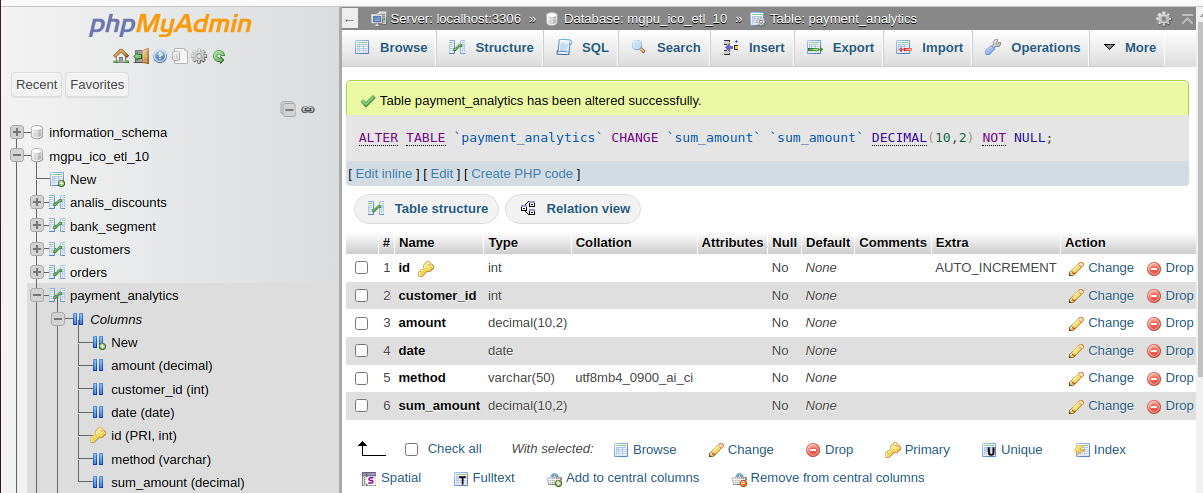


Рисунок 11. Добавлен столбец для группировки по сумме продаж



Рисунок 12. Импортирование данных из PostgreSQL

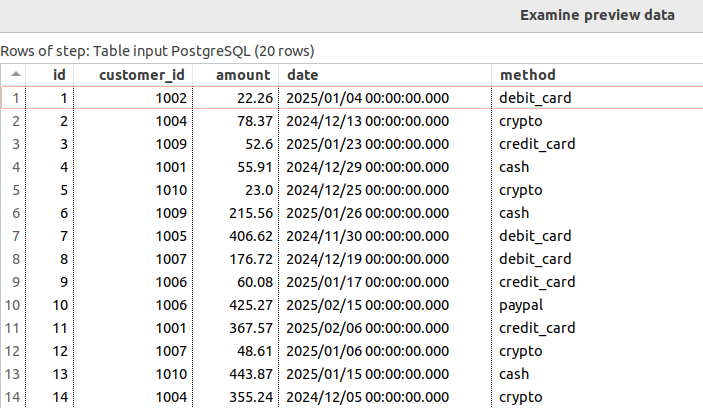


Рисунок 13. Предпросмотр данных

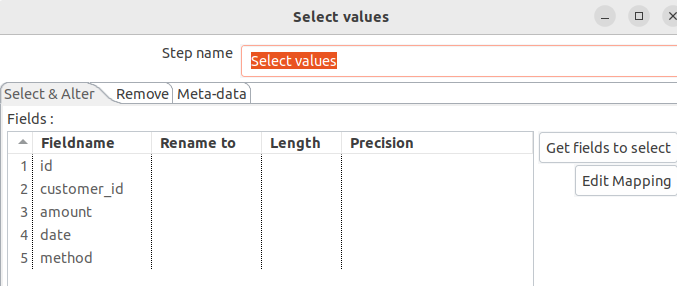


Рисунок 14. Выбор нужных столбцов

**Задание 3.** Фильтр платежей по методу

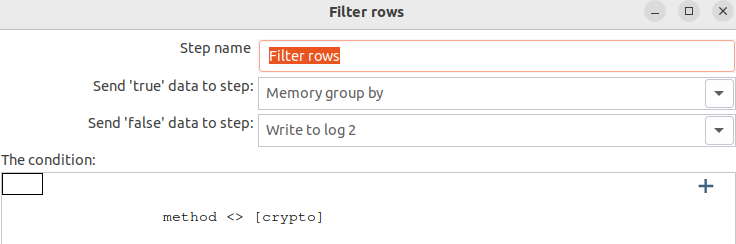


Рисунок 15. Фильтр платежей по методу

Загружаются все методы кроме криптовалютного.

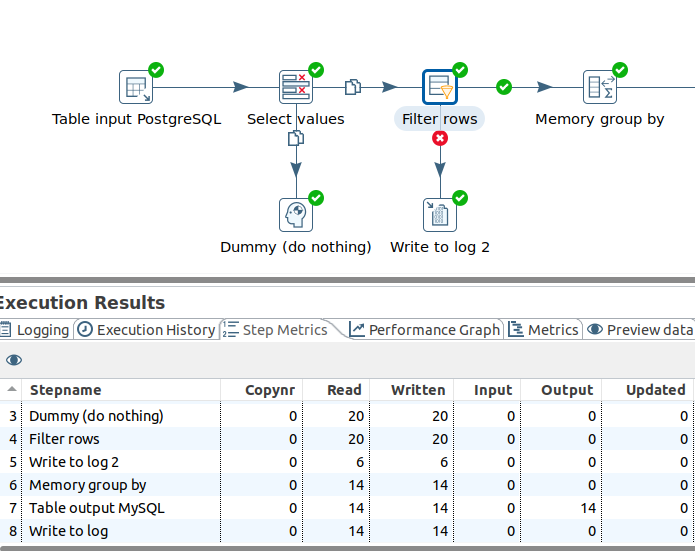


Рисунок 16. Результат фильтрации

Как видно, они отфильтрованы и из 20 в базу выгружены 14.

**Задание 4.** Суммы платежей по периодам

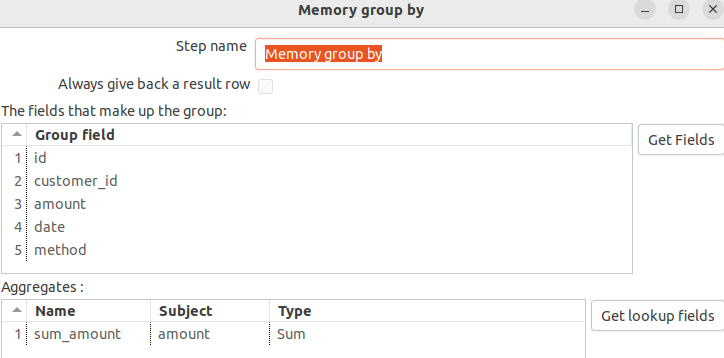


Рисунок 17. Расчет суммы платежей

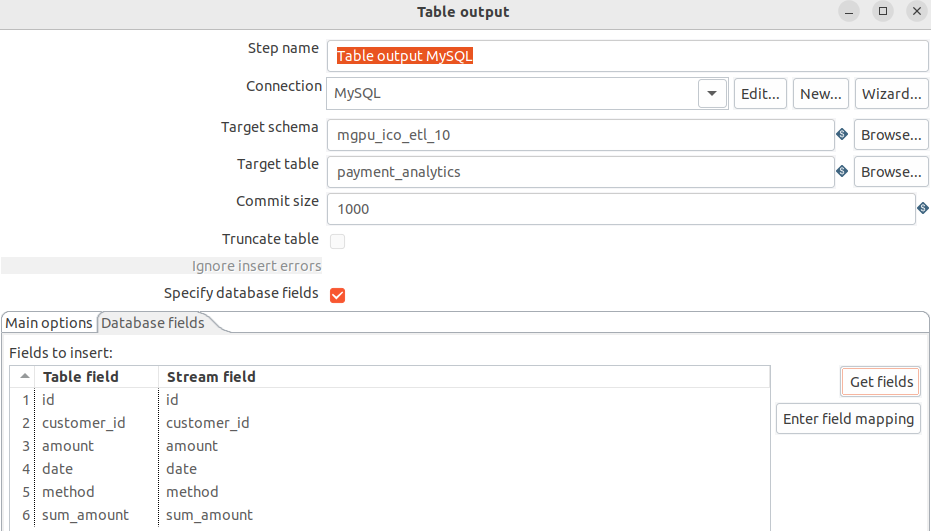


Рисунок 18. Настройка коннектора для выгрузки данных в таблицу MySQL

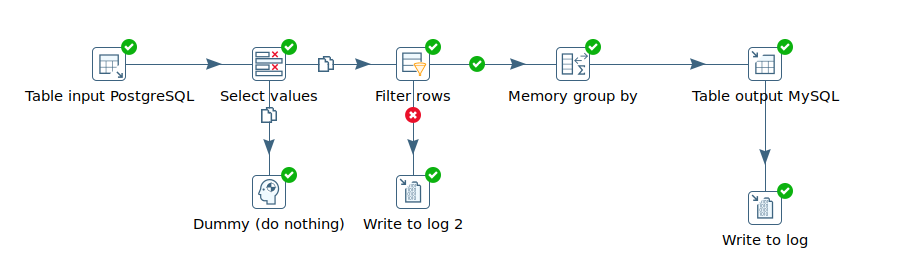


Рисунок 19. Схема итоговой транмформации

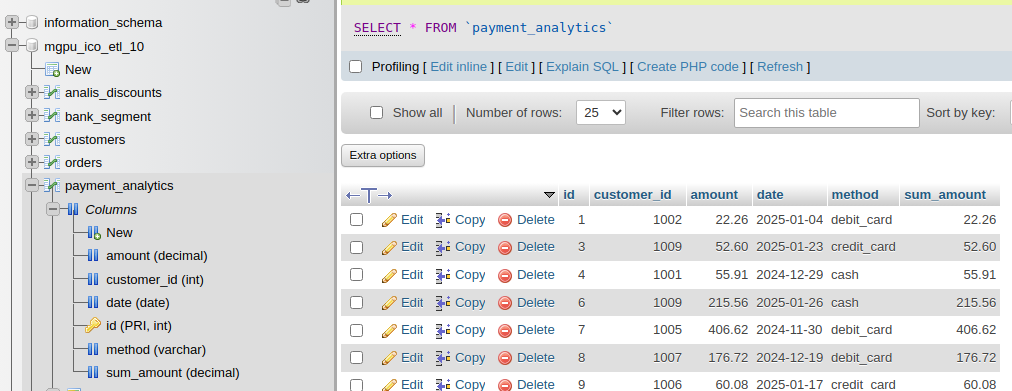


Рисунок 20. Данные успешно выгружены и загружены в базу

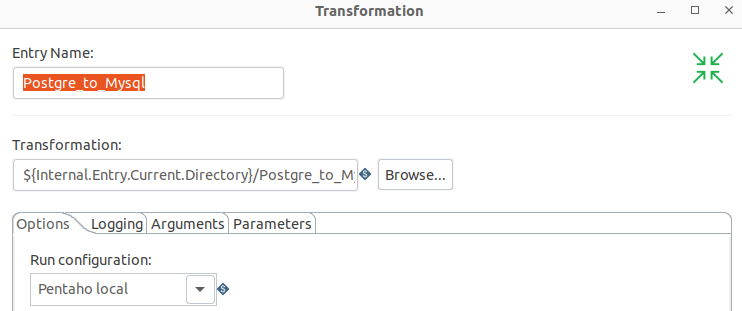


Рисунок 21. Подключение трансформации в job

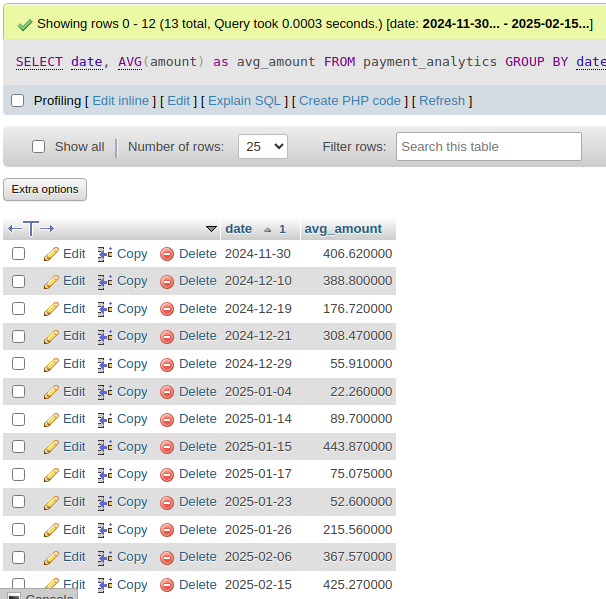


Рисунок 22. Анализ платежей по дням

**Задание 5.** Анализ популярных методов

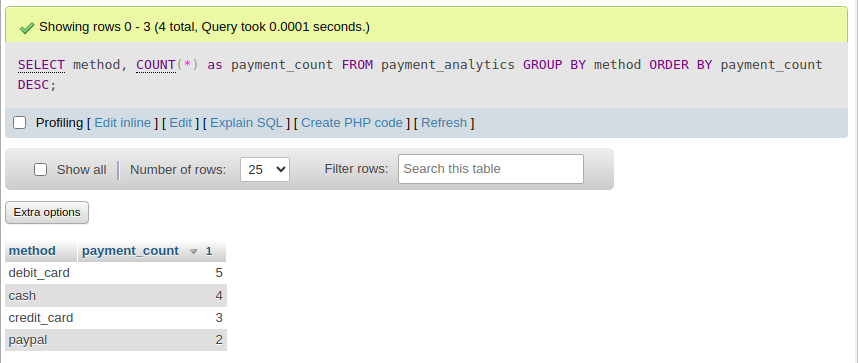


Рисунок 23. Количество платежей по методам оплаты

**Вывод:**

Созданы исходные таблицы в PostgreSQL с различными наборами данных

Настроены целевые таблицы в MySQL для приема данных.

Разработаны процессы трансформации данных в Pentaho.

Реализованы механизмы обработки ошибок и валидации данных.

Созданы представления для связанных данных.

По результату анализа видно, что больше всего платежей совершают дебетовыми картами.