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

Мокрецова Виктория Анатольевна

Тема выпускной квалификационной работы

«Разработка интеллектуальной системы сегментации клиентов и оптимизации программы лояльности на основе многофакторного анализа данных»

**Лабораторная работа 1.1 Установка и настройка ETL-инструмента. Создание конвейеров данных**

Направление подготовки – 38.03.05 «Бизнес-информатика»,

Профиль подготовки – «Аналитика данных и эффективное управление»

(очная форма обучения)

Руководитель:

кандидат технических наук, доцент

Босенко Тимур Муртазович

\_\_\_\_\_

Москва

2025

# Цели

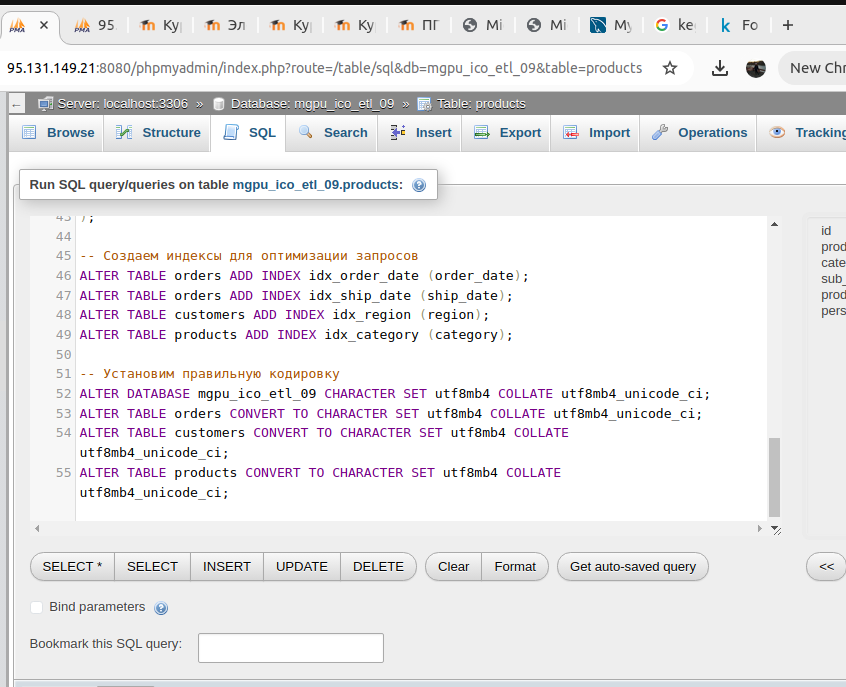
# Научиться открывать пентаго, загружать в него данные, выгружать из него данные в БД.

# Задачи:

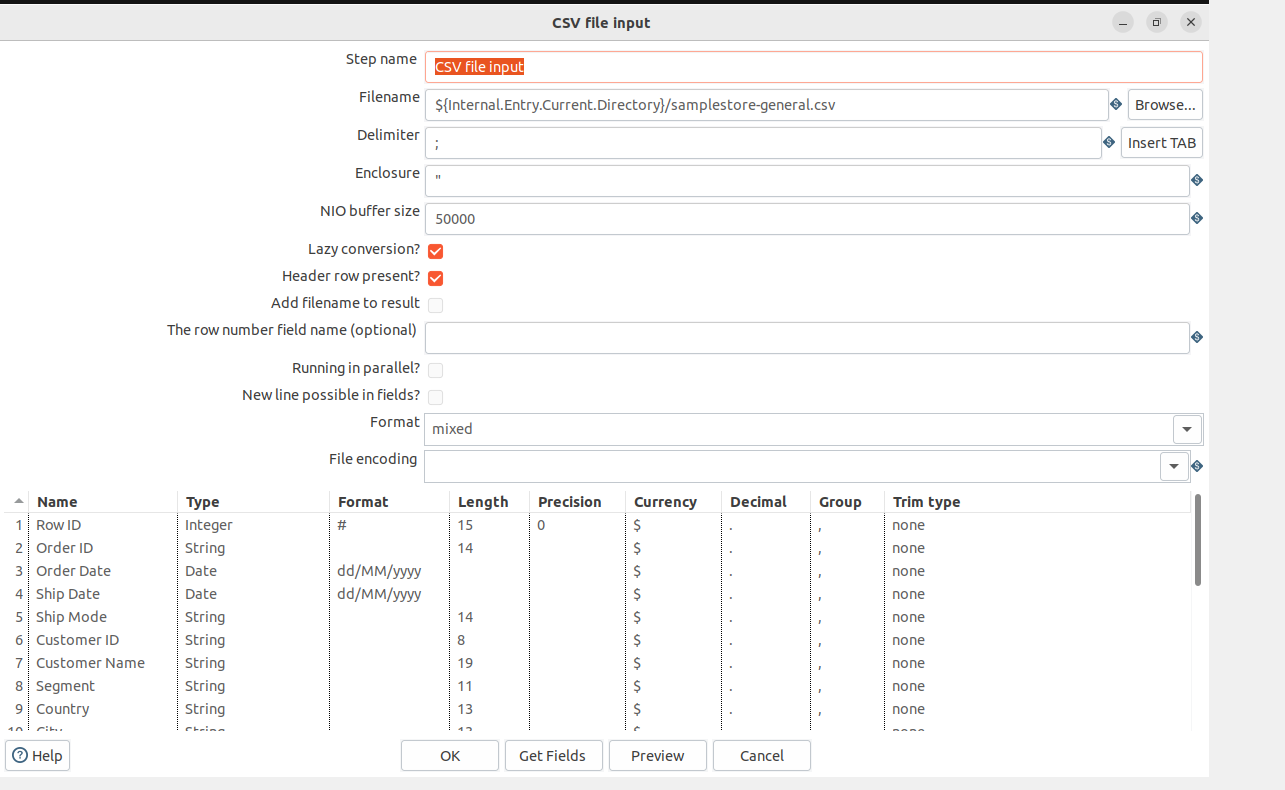
* Скачать и установить необходимые компоненты в виртуальную среду Virtual Box Ubuntu Linux.
* Загрузить данные из CSV в пентаго,
* Загрузить данные в MSSQL.
* Отобразить результаты.

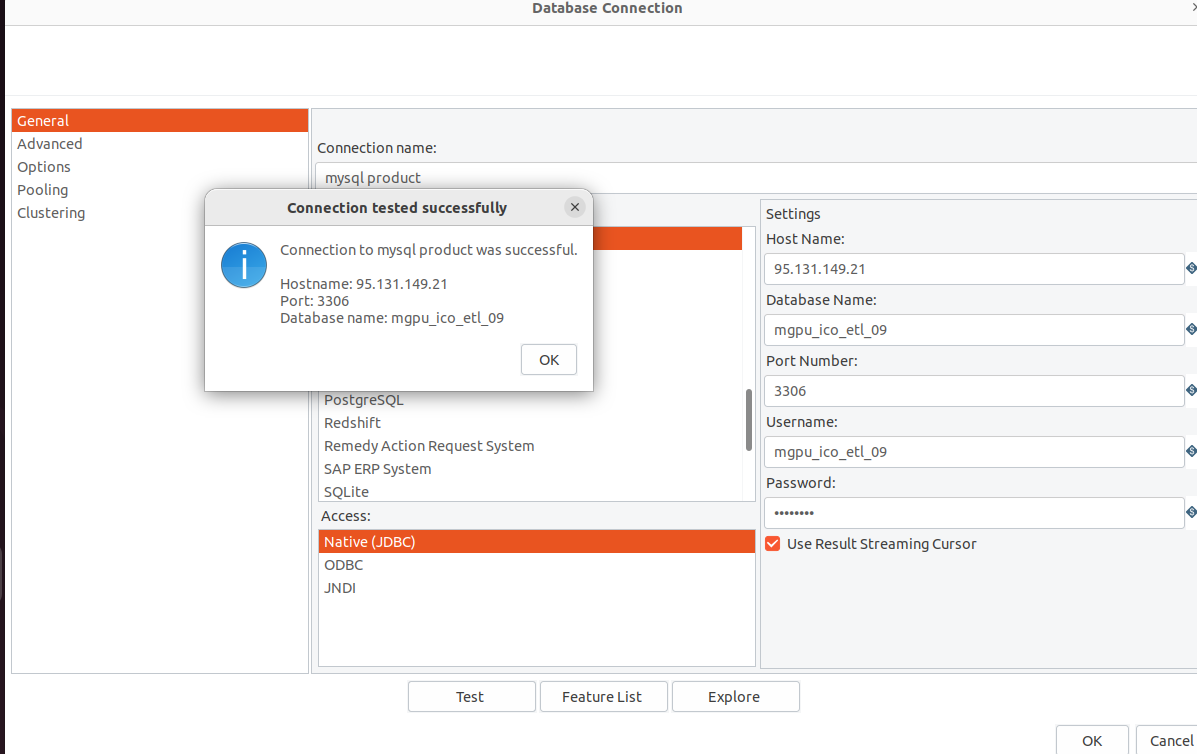
# Установка библиотек.

Создала БД и определила ключи в PhpMyadmin.



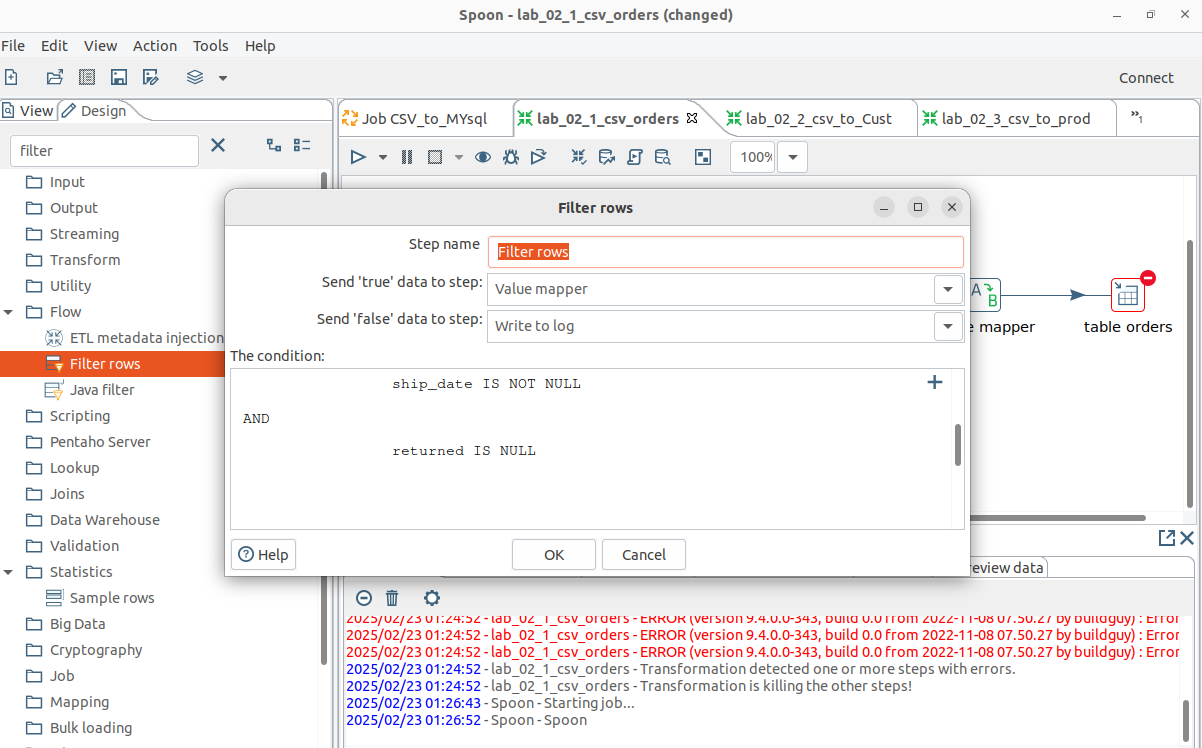
Скачала все файлы, поменяла во всех файлах подключения на свои.



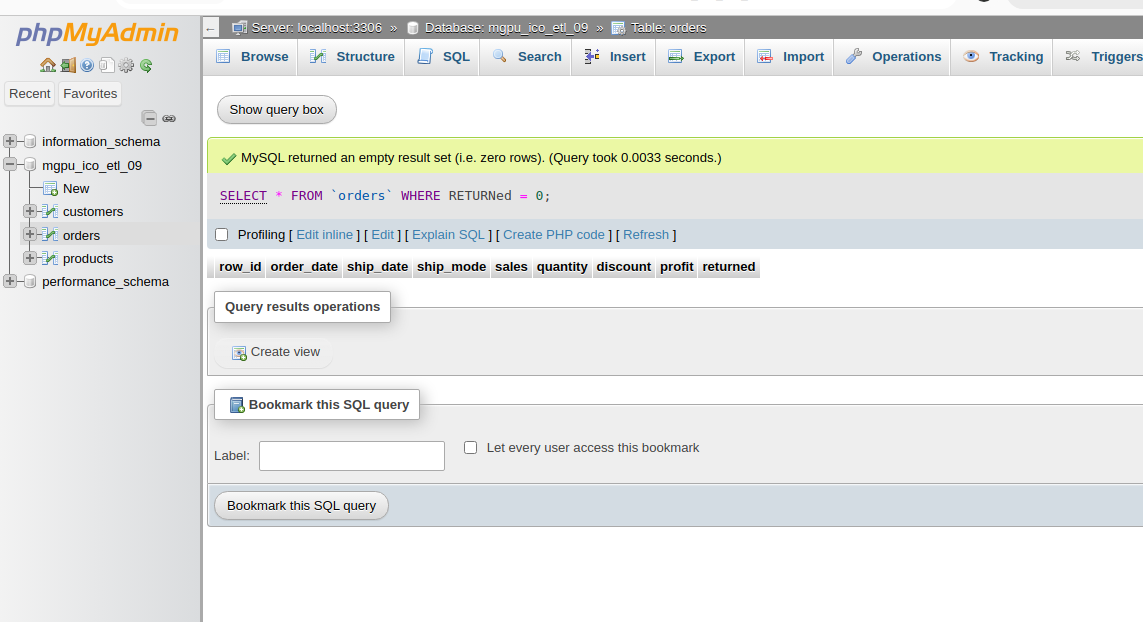


# Фильтр: только заказы без возвратов

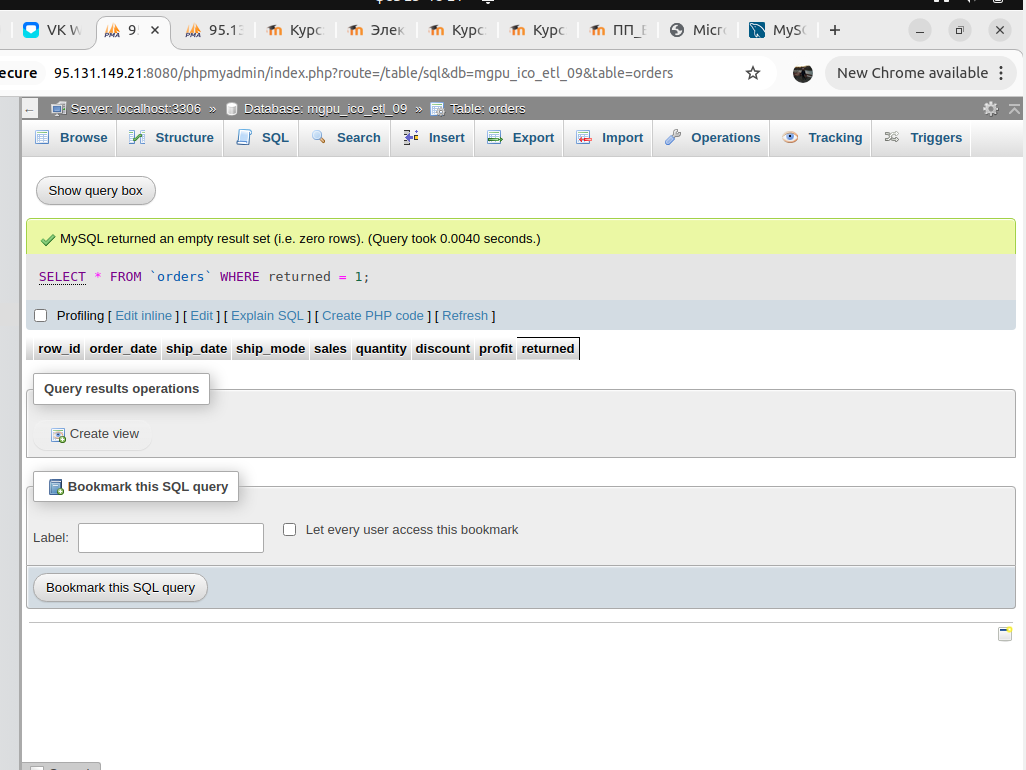
Добавила фильтр в потоке обработки заказов на поле returned is null, так как возврат обозначается единицей.



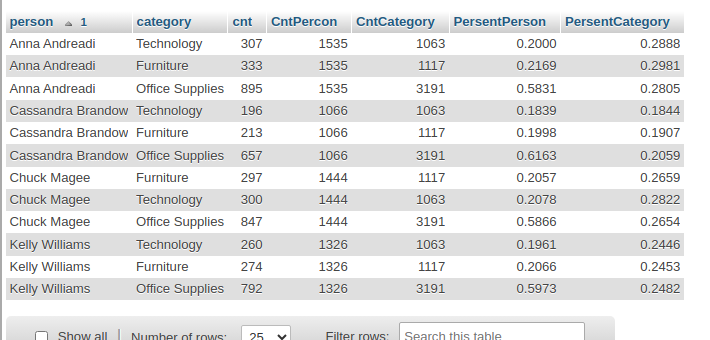
Проверила существуют ли записи с returned = 0, таких записей нет.



С фильтром убрали записи с returned = 1



# Отчет по менеджерам и категориям

Менеджеры по продажам расположены в таблице продуктов, так как каждый менеджер ответственен за свои продукты, выведем какие продукты продаёт каждый менеджер: 

Всего 4 менеджера, каждый из которых продаёт товары из всех категорий, у менеджеров есть небольшие различия в количестве проданных товаров, наибольшим количеством номенклатур владеет Анна Андреади – 1535, наименьшим Кассандра Брендоу – 1066. В категориях больший разброс, Офисные принадлежности – 3191 номенклатура, технологии – 1063 номенклатур. Процент распределения номенклатур соответствует количеству, в среднем 60% номенклатур приходится на каждого менеджера по категории Офисные принадлежности и по 20% на остальные категории. В остальных категориях та же ситуация.