**Пошаговый план выполнения финального задания SQL Pro**

**Шаг 1: Подготовка данных**

1. **Подключение к базе данных:**
   * Используй DBeaver для подключения к PostgreSQL
   * Host: 158.160.76.203, Port: 5432, Database: course, User: course, Password: course
2. **Проверка исходных данных:**
3. SELECT \* FROM evgeniy\_zubkov.yandex\_music\_raw LIMIT 10;
4. DESC evgeniy\_zubkov.yandex\_music\_raw; -- или \d в psql

**Шаг 2: Создание нормализованной структуры**

1. **Выполни первый артефакт** - создание таблиц и индексов
2. **Создай триггер для аудита** на таблице genres
3. **Создай функцию для просмотра аудита** fnc\_audit()

**Шаг 3: Миграция данных**

1. **Выполни скрипт миграции** из третьего артефакта
2. **Проверь качество данных** после миграции
3. **Протестируй триггеры** - добавь/измени/удали тестовый жанр

**Шаг 4: Создание аналитических представлений**

1. **Создай MATERIALIZED VIEW** для популярности жанров
2. **Создай VIEW** для анализа исполнителей
3. **Создай MATERIALIZED VIEW** для трендов треков
4. **Обнови MATERIALIZED VIEW** командой REFRESH

**Шаг 5: Визуализация в SuperSet**

1. **Подключись к SuperSet:**
   * URL: http://158.160.74.56:8088/superset/welcome/
   * User: admin, Password: admin
2. **Создай датасет:**
   * Добавь подключение к PostgreSQL
   * Создай датасеты на базе твоих VIEW и MATERIALIZED VIEW
3. **Создай 3 основных дашборда:**
   * **Дашборд 1:** Популярность жанров (Bar Chart)
   * **Дашборд 2:** Анализ исполнителей (Bubble Chart)
   * **Дашборд 3:** Распределение треков (Scatter Plot)

**Шаг 6: Создание отчета**

Создай DOCX файл со следующей структурой:

**1. Титульный лист**

* Название: "Анализ музыкальных данных Yandex Music"
* Твое имя
* Дата

**2. Схема нормализации**

🗂️ Справочники:

├── genres (жанры)

├── artists (исполнители)

└── charts (чарты)

📊 Оперативная таблица:

└── tracks (треки)

🔍 Аналитические представления:

├── mv\_genre\_popularity (популярность жанров)

├── v\_artists\_analysis (анализ исполнителей)

└── mv\_track\_trends (тренды треков)

**3. Код SQL (скриншоты + код)**

* Создание таблиц
* Триггеры для аудита
* Функция аудита
* Аналитические представления
* Миграция данных

**4. Скриншоты SuperSet**

* Дашборд популярности жанров
* Анализ исполнителей
* Распределение треков по длительности

**5. Интересные находки**

* Самый популярный жанр
* Корреляция между длительностью и популярностью
* Влияние Explicit контента на прослушивания

**Шаг 7: Подготовка презентации (5 слайдов)**

**Слайд 1: Титульный**

* Тема: "Анализ данных Yandex Music"
* Твое имя и презентация

**Слайд 2: Выбранный датасет**

* Описание: "Топ-100 треков Yandex Music"
* Структура данных: треки, исполнители, жанры, чарты
* Объем: ~100 записей, 9 полей

**Слайд 3: Техническое решение**

* Нормализация: 4 основные таблицы + аудит
* Триггеры для отслеживания изменений
* 3 аналитических представления
* Индексы для производительности

**Слайд 4: Ключевые находки**

* Самый популярный жанр: [будет видно после анализа]
* Средняя длительность трека: ~3.5 минуты
* Корреляция explicit контента и популярности
* Топ исполнители по эффективности

**Слайд 5: Визуализации SuperSet**

* Скриншот дашборда популярности жанров
* Скриншот анализа исполнителей
* Скриншот распределения треков

**Полезные команды для проверки:**

-- Проверка аудита

SELECT \* FROM evgeniy\_zubkov.fnc\_audit('genres');

-- Обновление аналитики

SELECT evgeniy\_zubkov.refresh\_analytics();

-- Топ жанров

SELECT \* FROM evgeniy\_zubkov.mv\_genre\_popularity LIMIT 5;

-- Топ исполнители

SELECT artist\_name, tracks\_count, total\_monthly\_listens

FROM evgeniy\_zubkov.v\_artists\_analysis

ORDER BY total\_monthly\_listens DESC LIMIT 10;

**Критерии успеха:**

✅ Нормализованная структура данных  
✅ Работающие триггеры аудита  
✅ Функция просмотра аудита  
✅ 3 аналитических представления  
✅ Визуализации в SuperSet  
✅ Документация с кодом и скриншотами  
✅ Презентация на 5 слайдов

**Возможные проблемы и решения:**

* **Проблема:** Не подключается к базе → Проверь VPN/сетевые настройки
* **Проблема:** Ошибки в SQL → Проверь схему evgeniy\_zubkov
* **Проблема:** SuperSet не отображает данные → Проверь SQL в датасете
* **Проблема:** Медленные запросы → Проверь индексы и REFRESH MATERIALIZED VIEW