**Текст для защиты проекта "Yandex Music Analytics"**

**🎵 Слайд 1: Титульный лист**

**Время: 30 секунд**

**Текст:** "Добро пожаловать на защиту моего финального проекта курса SQL Pro. Меня зовут Евгений Зубков, и сегодня я представлю вам аналитическую систему для данных Yandex Music.

Проект называется 'Yandex Music Analytics' и посвящен анализу топ-100 треков популярной музыкальной платформы. В рамках этого проекта я создал полноценную аналитическую базу данных с нормализованной структурой, системой аудита и визуализацией инсайтов."

**Ключевые моменты:**

* Представиться уверенно
* Подчеркнуть, что это финальный проект
* Обозначить масштаб работы

**📊 Слайд 2: Выбранный датасет**

**Время: 1 минута**

**Текст:** "Для анализа я выбрал датасет 'Yandex Music Top-100 Songs' - это актуальные данные о самых популярных треках на платформе.

**Почему именно этот датасет?**

* Во-первых, он содержит **разнообразную структуру данных**: информацию о треках, исполнителях, жанрах и чартах
* Во-вторых, включает **важные метрики**: месячные прослушивания, лайки исполнителей, длительность треков
* В-третьих, позволяет провести **глубокую аналитику** музыкальных трендов и предпочтений

Датасет содержит информацию о 100 треках с 9 ключевыми атрибутами, что дает отличную базу для демонстрации навыков нормализации и аналитики в PostgreSQL."

**Ключевые моменты:**

* Объяснить выбор датасета
* Подчеркнуть разнообразие данных
* Показать понимание аналитического потенциала

**⚡ Слайд 3: Техническое решение**

**Время: 1.5 минуты**

**Текст:** "Теперь перейдем к техническому решению. Я реализовал полноценную аналитическую систему на PostgreSQL с использованием современных подходов.

**Нормализация данных:**

* Создал **4 основные таблицы**: справочники жанров, исполнителей, чартов и основную таблицу треков
* Все таблицы связаны **внешними ключами** для обеспечения целостности данных
* Добавил **оптимизированные индексы**, включая B-tree для быстрого поиска и полнотекстовый индекс для поиска по названиям

**Система аудита:**

* Реализовал **автоматические триггеры** для отслеживания всех изменений в справочных таблицах
* Все операции INSERT, UPDATE, DELETE **автоматически логируются** в формате JSONB
* Создал **удобную функцию** для просмотра истории изменений

**Аналитические представления:**

* Разработал **3 ключевых представления**: анализ популярности жанров, эффективности исполнителей и трендов треков
* Использовал **MATERIALIZED VIEW** для кэширования тяжелых вычислений
* Добавил **автоматическое обновление** аналитики"

**Ключевые моменты:**

* Подчеркнуть использование лучших практик
* Показать понимание производительности
* Выделить автоматизацию процессов

**🎯 Слайд 4: Ключевые находки**

**Время: 1.5 минуты**

**Текст:** "Анализ данных выявил несколько интересных инсайтов о музыкальных предпочтениях:

**Популярность жанров:**

* **Русская поп-музыка** доминирует в топе, что отражает локальные предпочтения аудитории Yandex Music
* Жанры показывают разное распределение по чартам, что говорит о сегментации аудитории

**Оптимальная длительность:**

* **Средняя длительность** успешных треков составляет **3.2 минуты**
* Треки длительностью **3-4 минуты показывают лучшие результаты** по прослушиваниям
* Это подтверждает теорию о 'золотой середине' в музыкальной индустрии

**Explicit контент:**

* **15% треков** содержат explicit контент
* Интересно, что **наличие explicit контента не влияет негативно** на популярность
* Некоторые explicit треки даже показывают выше средних результатов

**Эффективность исполнителей:**

* Выявил **коэффициент эффективности** - соотношение прослушиваний к лайкам
* Некоторые исполнители показывают **высокую конверсию** слушателей в фанатов"

**Ключевые моменты:**

* Представить данные как инсайты, а не просто цифры
* Показать аналитическое мышление
* Связать выводы с музыкальной индустрией

**📈 Слайд 5: Визуализации в SuperSet**

**Время: 1 минута**

**Текст:** "Для наглядной презентации результатов я создал **4 ключевые визуализации** в Apache SuperSet:

**Bar Chart - Популярность жанров:**

* Показывает **топ жанров** по общему количеству прослушиваний
* Позволяет сравнить **относительную популярность** различных музыкальных направлений

**Bubble Chart - Анализ исполнителей:**

* **Трехмерная визуализация**: количество треков, популярность и лайки
* Помогает выявить **наиболее эффективных** исполнителей

**Scatter Plot - Распределение треков:**

* Демонстрирует **корреляцию** между длительностью и популярностью
* Показывает **оптимальные зоны** для успешных треков

**Pie Chart - Explicit контент:**

* Наглядно показывает **долю explicit контента** в топе
* Помогает понять **структуру контента** платформы

Все дашборды **интерактивны** и позволяют детально изучить данные, применять фильтры и получать дополнительные инсайты."

**Ключевые моменты:**

* Подчеркнуть практическую ценность визуализаций
* Показать понимание разных типов графиков
* Выделить интерактивность дашбордов

**🎯 Заключение (дополнительно)**

**Время: 30 секунд**

**Текст:** "В результате работы над проектом я создал **полноценную аналитическую систему**, которая демонстрирует:

* **Глубокое понимание** принципов нормализации данных
* **Практические навыки** создания триггеров и функций в PostgreSQL
* **Аналитическое мышление** при создании инсайтов
* **Навыки визуализации** данных для бизнес-аудитории

Проект показывает, как **технические решения** могут приносить **реальную бизнес-ценность** через понимание пользовательских предпочтений и трендов.

Спасибо за внимание! Готов ответить на ваши вопросы."

**💡 Советы для успешной защиты**

**Подготовка:**

1. **Отрепетируй презентацию 3-4 раза** - засеки время
2. **Подготовь демонстрацию** работающей системы в DBeaver
3. **Изучи свой код** - будут вопросы о деталях реализации

**Во время защиты:**

1. **Говори уверенно** - ты эксперт в своем проекте
2. **Поддерживай зрительный контакт** с аудиторией
3. **Используй технические термины** правильно
4. **Будь готов к вопросам** о производительности и масштабируемости

**Возможные вопросы:**

* **"Почему выбрали именно MATERIALIZED VIEW?"**
* **"Как обеспечиваете производительность при больших объемах данных?"**
* **"Можете показать работу триггеров?"**
* **"Какие еще инсайты можно получить из этих данных?"**

**Демонстрация (если попросят):**

1. **Покажи работу функции аудита**: SELECT \* FROM fnc\_audit('genres');
2. **Продемонстрируй аналитику**: SELECT \* FROM mv\_genre\_popularity LIMIT 5;
3. **Покажи SuperSet дашборды** (если есть доступ)

**Удачи на защите! Твой проект технически сильный и практически ценный.**