

Бизнес-область: сервис потоковой передачи музыки.

1. Бизнес-процесс: **Track Playback** — воспроизведение трека пользователем.
2. Уровень детализации (grain): Одна строка факта = одно воспроизведение трека (playback event) пользователем. Гранулярность user x song x timestamp
Где user — идентификатор пользователя;
song — идентификатор трека;
timestamp — точное время начала воспроизведения;

3. Таблицы измерений (Dimension tables).
Определяем измерения (Dimensions) для факта

dim_user (**Кто?**):

PK: User_SK - суррогатный ключ
Source_UserID - естественный ключ
Username
Email
RegistrationDate

dim_song (**Что?**)

PK: Song_SK - суррогатный ключ
Source_SongID - естественный ключ
Title
Artist
Album
Genre
DurationSec
ReleaseDate

dim_date (**Когда?**)

PK: Date_sk - суррогатный ключ
Source_date_id — естественный ключ
Year
Month
Day

4. Таблица фактов (Fact table) с метриками и атрибутами
Определяем таблицу фактов.

fact_playback

PK: Playback_sk - суррогатный ключ
FK: dim_user
FK: dim_song
FK: dim_date

Метрики:

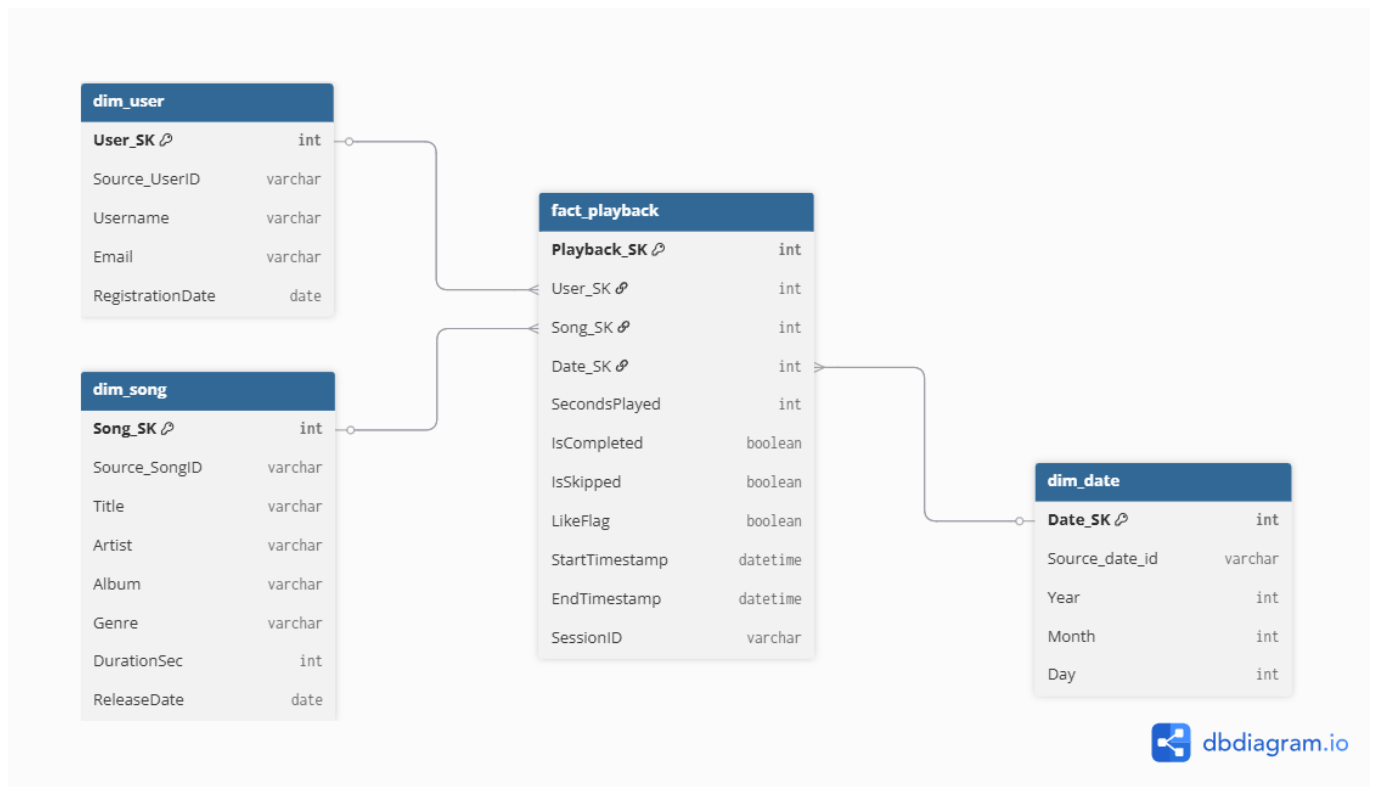
SecondsPlayed - сколько секунд реально прослушано
IsCompleted - дослушал до конца
IsSkipped - прервал рано
LikeFlag — лайк

Атрибуты:

StartTimestamp - точное время начала воспроизведения
EndTimestamp - время окончания
SessionID - идентификатор сессии

5. Физическая модель

Выбрана схема **звезда** самая простая и распространенная



6. Примеры аналитических запросов, которые помогут ответить на ключевые вопросы бизнес-процесса.

1. Количество воспроизведенных треков за период

```
SELECT
    dim_date.Year,
    dim_date.Month,
    COUNT(fact_playback.Playback_SK) AS TotalPlaybacks
FROM fact_playback
JOIN dim_date ON fact_playback.Date_SK = dim_date.Date_SK
GROUP BY dim_date.Year, dim_date.Month
ORDER BY dim_date.Year, dim_date.Month;
```

2. Активность пользователей

```
SELECT
    dim_user.User_SK,
    dim_user.Username,
    COUNT(fact_playback.Playback_SK) AS TotalPlaybacks,
    SUM(fact_playback.SecondsPlayed) AS TotalSeconds
FROM fact_playback
JOIN dim_user ON fact_playback.User_SK = dim_user.User_SK
GROUP BY dim_user.User_SK, dim_user.Username
ORDER BY TotalPlaybacks DESC;
```

3. Популярность треков по числу воспроизведений (Топ 10 треков)

```
SELECT
    dim_song.Song_SK,
    dim_song.title,
    dim_song.artist,
    COUNT(fact_playback.playback_sk) AS totalplaybacks,
    SUM(CASE WHEN fact_playback.LikeFlag THEN 1 ELSE 0 END) AS totallikes
FROM fact_playback
JOIN dim_song ON fact_playback.Song_SK = dim_song.Song_SK
GROUP BY dim_song.Song_SK, dim_song.title, dim_song.artist
ORDER BY totalplaybacks DESC
LIMIT 10;
```

4. Популярные жанры

```
SELECT
    dim_song.Genre,
    COUNT(fact_playback.Playback_SK) AS TotalPlays,
    SUM(fact_playback.SecondsPlayed) AS TotalSeconds
FROM fact_playback
JOIN dim_song ON fact_playback.Song_SK = dim_song.Song_SK
GROUP BY dim_song.Genre
ORDER BY TotalPlays DESC;
```

5. Сколько времени пользователь проводит в приложении

```
SELECT
    fact_playback.User_SK,
    DATE(fact_playback.StartTimestamp) AS playback_date,
    SUM(fact_playback.SecondsPlayed) AS total_seconds_per_day,
    ROUND(SUM(fact_playback.SecondsPlayed) / 60.0, 2) AS total_minutes_per_day
FROM fact_playback
GROUP BY fact_playback.User_SK, DATE(fact_playback.StartTimestamp)
ORDER BY fact_playback.User_SK, playback_date;
```