

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\_stream\_quality (**В каком качестве?**)

PK: StreamQuality\_SK - суррогатный ключ  
StreamQuality\_ID - естественный ключ  
Format  
Bitrate

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\_stream\_quality

**Метрики:**

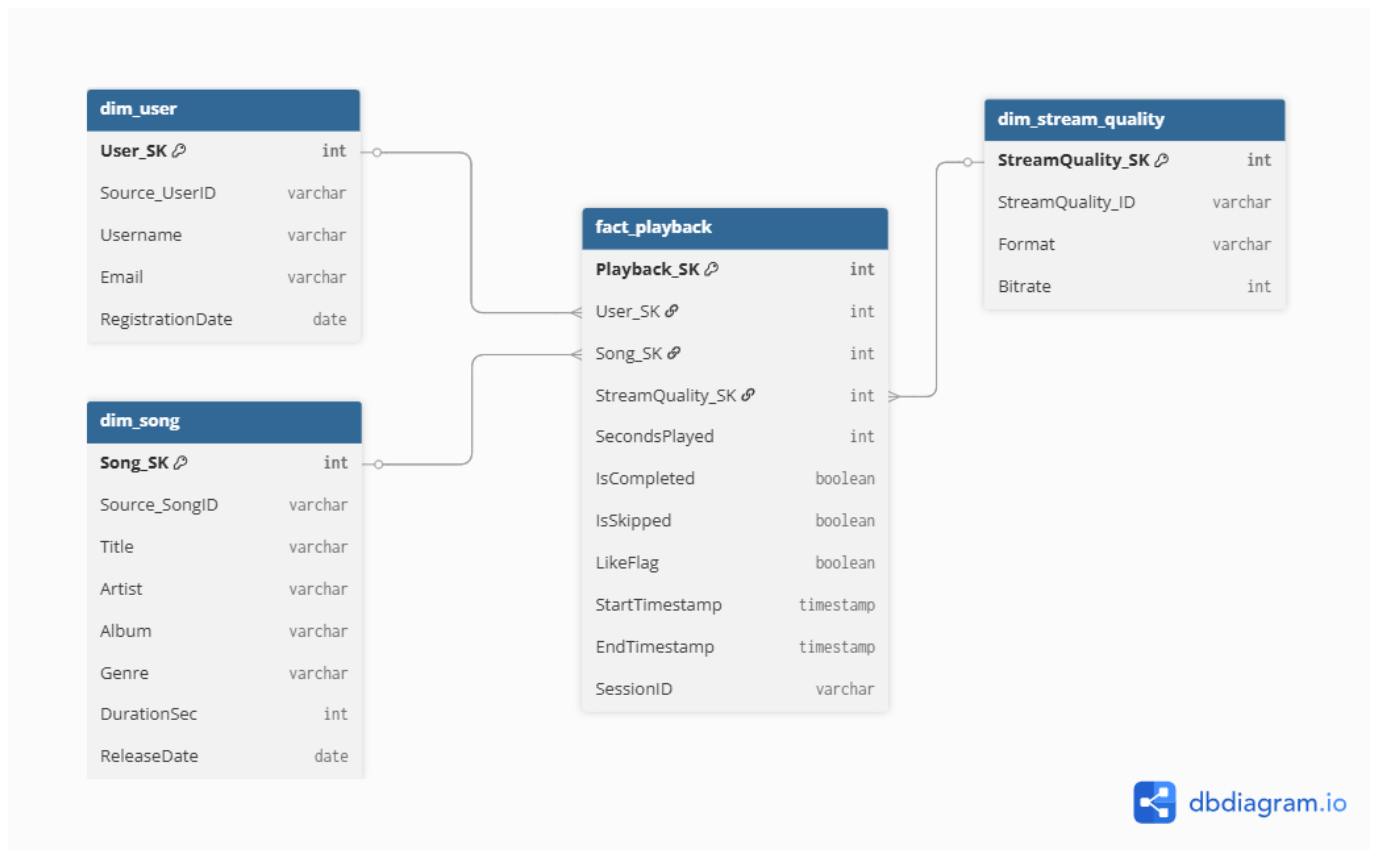
SecondsPlayed - сколько секунд реально прослушано  
IsCompleted - дослушал до конца  
IsSkipped - прервал рано  
LikeFlag — лайк

**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
    EXTRACT(YEAR FROM StartTimestamp AT TIME ZONE 'UTC') AS Year,
    EXTRACT(MONTH FROM StartTimestamp AT TIME ZONE 'UTC') AS Month,
    COUNT(Playback_SK) AS TotalPlaybacks
FROM fact_playback
GROUP BY Year, Month
ORDER BY Year, Month;
```

### 2. Активность пользователей

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

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

```
SELECT
    s.Song_SK,
    s.Title,
    s.Artist,
    COUNT(f.Playback_SK) AS TotalPlaybacks,
    SUM(CASE WHEN f.LikeFlag THEN 1 ELSE 0 END) AS TotalLikes
FROM fact_playback f
JOIN dim_song s ON f.Song_SK = s.Song_SK
GROUP BY s.Song_SK, s.Title, s.Artist
ORDER BY TotalPlaybacks DESC
LIMIT 10;
```

4. Популярные жанры

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

5. Сколько времени пользователь проводит в приложении

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

6. Какой формат предпочитают пользователи

```
SELECT
    sq.Format,
    sq.Bitrate,
    COUNT(f.Playback_SK) AS TotalPlaybacks,
    SUM(f.SecondsPlayed) AS TotalSeconds
FROM fact_playback f
JOIN dim_stream_quality sq
    ON f.StreamQuality_SK = sq.StreamQuality_SK
GROUP BY sq.Format, sq.Bitrate
ORDER BY TotalPlaybacks DESC;
```