SQL и получение данных

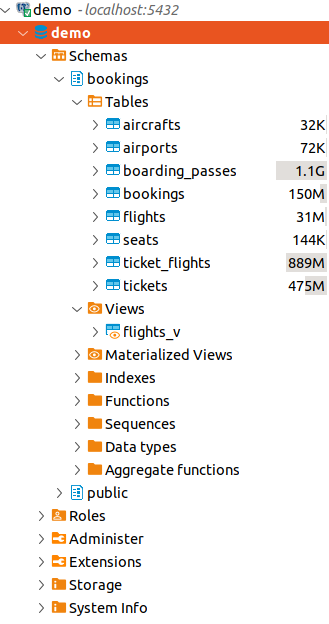
Итоговая работа

Котик для привлечения внимания

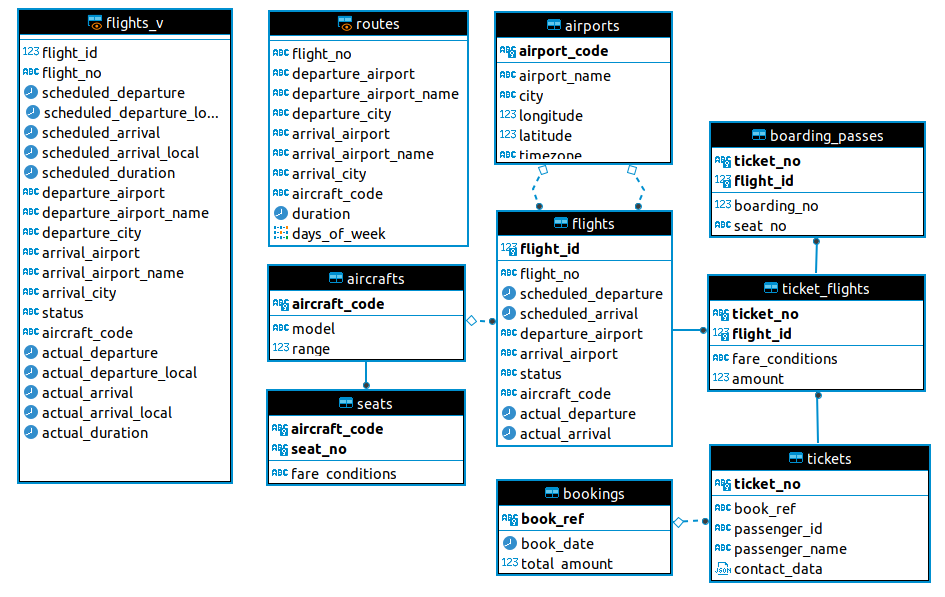


# 1. Тип подключения:

В работе использовался локальный тип подключения.



# 2. ER-диаграмма



# **3**. Описание БД

### **Т**аблицы

**aircrafts** — самолеты

**airports** — аэропорты

**boarding\_passes** — информация о посадке на самолет

**bookings** — бронирования билетов

**flights** — рейсы

**seats** — информация о местах в самолете

**ticket\_flight** — связь между билетами и перелетами

**tickets** — билеты

### **Материализованные представления**

**flights\_v** — информация о перелетах, обогащённая информацией об аэропортах, собирается из таблиц airports и flights

### **Материализованные представления**

**Routes** — Информация о существующих маршрутах, собирается из таблиц bookins, airports, flights

# **4**.1 **Развернутый анализ БД: таблицы**

#### **aircrafts**

Каждая модель воздушного судна идентифицируется своим трехзначным кодом (aircraft\_code). Указывается также название модели (model) и максимальная дальность полета в километрах (range).

**Первичный ключ:** **aircraft\_code**

#### **airports**

Аэропорт идентифицируется трехбуквенным кодом (airport\_code) и имеет свое имя (airport\_name). Также указывается название города (city), широта (longitude), долгота (latitude) и часовой пояс (timezone).

**Первичный ключ:** **airport\_code**

#### **boarding\_passes**

При регистрации на рейс, пассажиру выдается посадочный талон. Он идентифицируется номером билета (seat\_no) и номером рейса (boarding\_no). Посадочным талонам присваиваются последовательные номера в порядке регистрации пассажиров. Также указывается номер места (seat\_no).

**Первичный ключ:** **составной** **flight\_id + seat\_no**

**Внешний ключ:** по первичному ключу, ссылается на таблицу **ticket\_flights**

#### **bookings**

Информация о бронировании билетов, может быть забронировано больше 1 билета за раз. Присваивается уникальный идентификатор ((book\_ref). Также сохраняется дата бронирования (book\_date) и сумма (total\_amount )

**Первичный ключ:** **book\_ref**

#### ***seats***

Каждое место определяется своим номером (seat\_no) и имеет закрепленный за ним класс обслуживания (fare\_conditions) — Economy, Comfort или Business.

**Первичный ключ: seat\_no**

**Внешний ключ:** колонка **aircraft\_code** ссылается на таблицу **aircrafts**

#### ***flights***

Информация номера рейса (flight\_no), самолете (aircraft\_code), датах отправления (scheduled\_departure) и прибытия (scheduled\_arrival). Рейс всегда соединяет две точки — аэропорты вылета (departure\_airport) и прибытия (arrival\_airport). После рейса сохраняются фактические времена вылета(actual\_departure) и прибытия (actual\_arrival). Так же хранится текущий статус рейса:

Scheduled — запланирован

Departed — вылетел

Arrived — прибыл с задержкой

On Time — прибыл по расписанию

Cancelled — отменен

Delayed — задерживается

**Первичный ключ: flight\_no**

**Внешнии ключи:** колонка **aircraft\_code** ссылается на таблицу **aircrafts**, колонки **departure\_airport** и **arrival\_airport** ссылаются на таблицу **airports**

#### **ticket\_flights**

Связь билета с рейсом, хранятся данные о стоимости (amount) и классе обслуживания (fare\_conditions)

**Первичный ключ: составной flight\_id + ticket\_no**

**Внешний ключ:** по первичному ключу, ссылается на таблицу **ticket\_flights**

#### **tickets**

Билет имеет уникальный номер (ticket\_no), состоящий из 13 цифр. Хранится информация о документе, удостоверяющем личность (passenger\_id), фамилии и имени (passenger\_name) и контактную информацию (contact\_date).

**Первичный ключ: ticket\_no**

**Внешний ключ:** колонка **book\_ref** ссылается на таблицу **bookings**

# **4.2 Развернутый анализ БД: проблемы в архитектуре**

1) Нет таблицы с информацией о маршрутах, представление не позволят хранить дополнительную информацию именно о маршруте.

2) Не используется тип данных ENUM, там, где он напрашивается

flights.status

flights.dair\_condition

seats.fait\_condition

3) Мои предложения по оптимизации:

3а) Создать справочник с городами ( с указанием часового пояса, что позволило бы убрать эти данные из таблицы с самолетами.

3б) Создать справочник с классами обслуживания, и связать билеты с посадочными билетами, а посадочные билеты с местами в самолете, что позволит бы удалить таблицу tickets\_flights и упростить некоторые запросы. Возможно, стоит подумать о

3в) Создать таблицу с пассажирами, что позволит нормализовать данные и решать отдельный класс бизнес-задач — хранение паспортных данных пассажира, внедрение программы лояльности и т.п.

# **4.3 Развернутый анализ БД: решаемые бизнес задачи**

1) По каждому самолету можно посмотреть средний налет за определенное время. Это позволит сбалансировать нагрузку между самолетами, запланировать техническое обслуживание и обновление парка.

2) Можно проанализовать совершающиеся рейсы на предмет сезонности, и оптимизировать частоту перелетов в зависимости от времени года.

3) Можно проанализировать востребованность новых прямых рейсов, отслеживая стыковочные перелеты (по пассажиру)

4) Анализировать средний процент посадки от проданных билетов, что даст возможность использовать меньший по размеру самолет. (возможно, используя овербукинг)

5) Анализировать статистику по планируемым и фактическим временам перелета, есть ли связь с конкретными аэропортами, самолетами, пассажирами

6) Рассчитать выручку за определённый период

7) Найти наиболее лояльных пассажиров

8) Влияет ли задержка рейса на вероятность возвращения пассажира к авиакомпании

# **5. Результаты выполнения SQL-запросов:**

**В каких городах больше одного аэропорта?**

- Ульяновск, Москва

**В каких аэропортах есть рейсы, выполняемые самолетом с максимальной дальностью перелета?**

- Внуково, Домодедово, Кольцово, Пермь, Сочи, Толмачёво, Шереметьево

**Вывести 10 рейсов с максимальным временем задержки вылета**

- PG0073, PG0040, PG0533, PG0132, PG0531, PG0589, PG0560, PG0590, PG0201, PG0114

**Были ли брони, по которым не были получены посадочные талоны?**

- Да

**Найдите свободные места для каждого рейса, их % отношение к общему количеству мест в самолете и накопительную сумму - сколько человек уже вылетело из данного аэропорта на этом или более ранних рейсах за день.**

- Результаты в приложении

**Найдите процентное соотношение перелетов по типам самолетов от общего количества.**

Airbus A319-100 3.738%

Airbus A321-200 5.898%

Boeing 737-300 3.846%

Boeing 767-300 3.686%

Boeing 777-300 1.843%

Bombardier CRJ-200 27.278%

Cessna 208 Caravan 28.015%

Sukhoi SuperJet-100 25.696%

**Были ли города, в которые можно добраться бизнес - классом дешевле, чем эконом-классом в рамках перелета?**

- Нет

**Между какими городами нет прямых рейсов?**

- Есть 9584 возможных прямых рейса, которых нет (или 4792, если считать туда/обратно)

**Вычислите расстояние между аэропортами, связанными прямыми рейсами, сравните с допустимой максимальной дальностью перелетов в самолетах, обслуживающих эти рейсы**

- Результаты в приложении