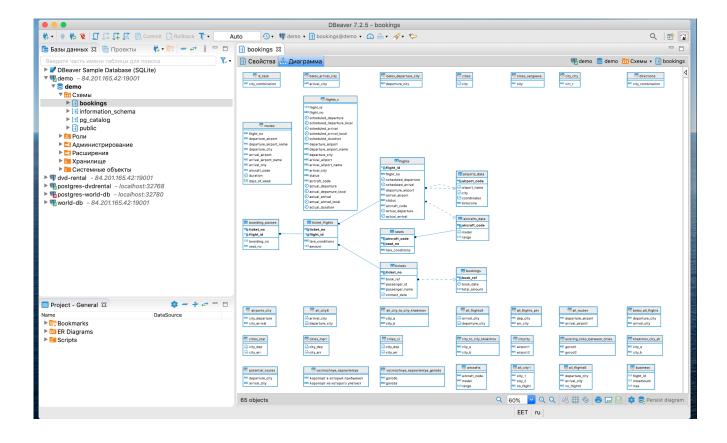
Итоговая работа по курсу "SQL и получение данных"

1 Подключение к базе данных	1
2 Описание и анализ базы данных	2
3 Решение задач на написание SOL-запросов к базе	ланных 11

1 Подключение к базе данных

В работе использовался облачный тип подключения.

Скриншот ER-диаграммы из DBeaver:



2 Описание и анализ базы данных

2.1 Краткое описание базы данных

База данных состоит из восьми таблиц и двух представлений.

Список отношений:

Имя	Тип	Описание
aircrafts_data	таблица	самолеты
airports_data	таблица	аэропорты
boarding_passes	таблица	посадочные талоны
bookings	таблица	бронирования
flights	таблица	рейсы
seats	таблица	места
ticket_flights	таблица	перелеты
tickets	таблица	билеты
flights_v	представление	рейсы
routes	материализованное представление	маршруты

2.2 Развернутый анализ базы данных: описание таблиц, представлений и связей между ними

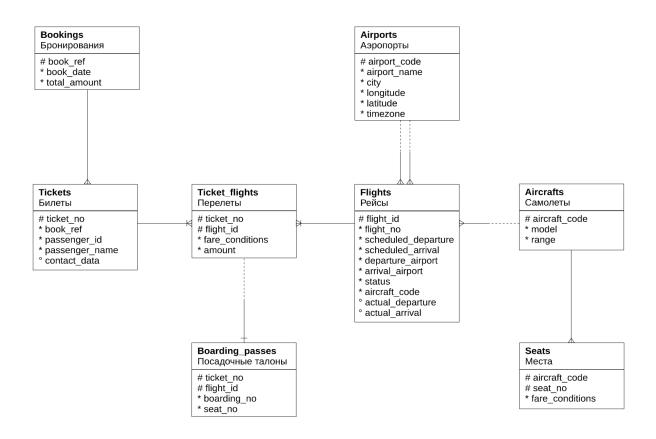


Таблица aircrafts data (самолеты)

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

•	•	Модификаторы	Описание
aircraft_code model	char(3)	NOT NULL NOT NULL	Код самолета, IATA Модель самолета Максимальная дальность полета, км

Ограничение первичного ключа - aircraft code.

Есть проверяющее ограничение для максимальной дальности полета (ее значение должно быть положительным): CHECK (range > 0).

На таблицу есть следующие ссылки извне:

В таблице flights ограничение внешнего ключа для атрибута aircraft code;

В таблице seats ограничение внешнего ключа для атрибута aircraft_code.

Таблица airports data (аэропорты)

Аэропорт идентифицируется трехбуквенным кодом (airport_code) и имеет свое имя (airport_name). Для города не предусмотрено отдельной сущности, но название (city) указывается и может служить для того, чтобы определить аэропорты одного города. Также указывается широта (longitude), долгота (latitude) и часовой пояс (timezone).

Столбец	Тип	Модификаторы	Описание
airport_code airport_name city longitude latitude timezone	char(3) text text float float text	NOT NULL NOT NULL NOT NULL NOT NULL NOT NULL NOT NULL	Код аэропорта Название аэропорта Город Координаты аэропорта: долгота Координаты аэропорта: широта Временная зона аэропорта

Ограничение первичного ключа - airport code.

Ссылки извне:

В таблице flights ограничения внешнего ключа для атрибутов arrival_airport и departure airport (ссылка на airport code).

Таблица boarding passes (посадочные талоны)

При регистрации на рейс, которая возможна за сутки до плановой даты отправления, пассажиру выдается посадочный талон. Он идентифицируется также, как и перелет - номером билета и номером рейса. Посадочным талонам присваиваются последовательные номера (boarding_no) в порядке регистрации пассажиров на рейс (этот номер будет уникальным только в пределах данного рейса). В посадочном талоне указывается номер места (seat no).

Столбец	•	Модификаторы	•
ticket_no flight_id	char(13) integer integer	NOT NULL NOT NULL	Номер билета Идентификатор рейса Номер посадочного талона Номер места

Первичный ключ является составным - ticket no, flight id.

Ограничение типа UNIQUE для атрибутов (flight id, boarding no).

Ограничение типа UNIQUE для атрибутов (flight id, seat no).

Ограничение внешнего ключа: атрибуты (ticket_no, flight_id) ссылаются на таблицу ticket_flights (ticket_no, flight_id).

Таблица bookings (бронирования)

Пассажир заранее (book_date, максимум за месяц до рейса) бронирует билет себе и, возможно, нескольким другим пассажирам.

Бронирование идентифицируется номером (book_ref, шестизначная комбинация букв и цифр). Поле total_amount хранит общую стоимость включенных в бронирование перелетов всех пассажиров.

Столбец		Тип	•	Модификаторы	•	Описание
book_ref book_date total_amount	 	char(6) timestamptz		NOT NULL	İ	Номер бронирования Дата бронирования Полная сумма бронирования

Ограничение первичного ключа - book_ref.

Ссылки извне:

В таблице tickets ограничение внешнего ключа для атрибута book_ref.

Таблица flights (рейсы)

Естественный ключ таблицы рейсов состоит из двух полей - номера рейса (flight_no) и даты отправления (scheduled_departure). Чтобы сделать внешние ключи на эту таблицу компактнее, в качестве первичного используется суррогатный ключ (flight id).

Рейс всегда соединяет две точки - аэропорты вылета (departure_airport) и прибытия (arrival_airport). Такое понятие, как «рейс с пересадками» отсутствует: если из одного аэропорта до другого нет прямого рейса, в билет просто включаются несколько необходимых рейсов.

У каждого рейса есть плановые дата и время вылета (scheduled_departure) и прибытия (scheduled_arrival). Реальные время вылета (actual_departure) и прибытия (actual_arrival) могут отличаться: обычно не сильно, но иногда и на несколько часов, если рейс задержан.

Статус рейса (status) может принимать одно из следующих значений:

- Scheduled рейс доступен для бронирования. Это происходит за месяц до плановой даты вылета до этого запись о рейсе не существует в базе данных.
- On Time рейс доступен для регистрации (за сутки до плановой даты вылета) и не задержан.
- Delayed рейс доступен для регистрации (за сутки до плановой даты вылета), но задержан.
- Departed самолет уже вылетел и находится в воздухе.
- Arrived самолет прибыл в пункт назначения.
- Cancelled рейс отменен.

Столбец	Тип +	Модификаторы +	Описание
flight_id flight_no scheduled_departure scheduled_arrival departure_airport arrival_airport status aircraft_code actual_departure actual_arrival	serial char(6) timestamptz timestamptz char(3) char(3) varchar(20) char(3) timestamptz timestamptz	NOT NULL	Идентификатор рейса Номер рейса Время вылета по расписанию Время прилёта по расписанию Аэропорт отправления Аэропорт прибытия Статус рейса Код самолета, IATA Фактическое время вылета Фактическое время прилёта

Ограничение первичного ключа - flight_id.

Ограничение типа UNIQUE для атрибутов (flight no, scheduled departure).

Проверяющие ограничения:

CHECK (scheduled_arrival > scheduled_departure) - плановое время прибытия должно быть больше планового времени вылета.

CHECK ((actual_arrival is null) or ((actual_departure is not null and actual_arrival is not null) and (actual_arrival > actual_departure))) - реальное время прилета должно быть null либо реальное время вылета не null и реальное время прилета не null (реальное время прилета при этом больше реального времени вылета).

CHECK (status in ('On Time', 'Delayed', 'Departed', 'Arrived', 'Scheduled', 'Cancelled')) - статус рейса может иметь одно из перечисленных значений.

Ограничения внешнего ключа:

Атрибут aircraft_code ссылается на таблицу aircrafts_data (aircraft_code);

Атрибут arrival airport ссылается на таблицу airports data (airport code);

Атрибут departure airport ссылается на таблицу airports data (airport code).

Ссылки извне:

В таблице ticket_flights атрибут flight_id имеет ограничение внешнего ключа - ссылка на атрибут flight_id.

Таблица seats (места)

Места определяют схему салона каждой модели. Каждое место определяется своим номером (seat_no) и имеет закрепленный за ним класс обслуживания (fare conditions) - Economy, Comfort или Business.

Столбец	•	Модификаторы +	Описание
	char(3) varchar(4)	NOT NULL NOT NULL	Код самолета, IATA Номер места Класс обслуживания

Первичный ключ таблицы является составным - aircraft_code, seat_no.

Проверяющие ограничения:

CHECK (fare_conditions in ('Economy', 'Comfort', 'Business')) - класс обслуживания должен иметь одно из приведенных значений.

Ограничения внешнего ключа:

Атрибут aircraft code ссылается на таблицу aircrafts data (aircraft code).

Таблица ticket flights (перелеты)

Перелет соединяет билет с рейсом и идентифицируется их номерами.

Для каждого перелета указываются его стоимость (amount) и класс обслуживания (fare_conditions).

Столбец	Тип +	Модификаторы	Описание
ticket_no flight_id fare_conditions amount	varchar(10)	NOT NULL	Номер билета Идентификатор рейса Класс обслуживания Стоимость перелета

Первичный ключ таблицы является составным - ticket no, flight id.

Ограничения-проверки:

Стоимость перелета должна быть неотрицательна: CHECK (amount >= 0);

Класс обслуживания должен иметь одно из приведенных значений: CHECK (fare conditions in ('Economy', 'Comfort', 'Business')).

Ограничения внешнего ключа:

Атрибут flight_id ссылается на таблицу flights (flight_id);

Атрибут ticket_no ссылается на таблицу tickets (ticket_no).

Ссылки извне:

В таблице boarding_passes есть ограничение внешнего ключа для атрибутов (ticket_no, flight_id).

Таблица tickets (билеты)

Билет имеет уникальный номер (ticket_no), состоящий из 13 цифр.

Билет содержит идентификатор пассажира (passenger_id) - номер документа, удостоверяющего личность, его фамилию и имя (passenger_name) и контактную информацию (contact_data).

Ни идентификатор пассажира, ни имя не являются постоянными (можно поменять паспорт, можно сменить фамилию), поэтому однозначно найти все билеты одного и того же пассажира невозможно.

Столбец	Тип +	Модификаторы	Описание
ticket_no book_ref passenger_id passenger_name contact_data	char(13) char(6) varchar(20) text jsonb	NOT NULL NOT NULL NOT NULL NOT NULL	Номер билета Номер бронирования Идентификатор пассажира Имя пассажира Контактные данные пассажира

Ограничение первичного ключа - ticket no.

Ограничение внешнего ключа:

Атрибут book ref ссылается на таблицу bookings (book ref).

Ссылки извне:

В таблице ticket_flights ограничение внешнего ключа для атрибута ticket no.

Представление "flights v" (рейсы)

Столбец	Тип	Описание
flight_id	integer	Идентификатор рейса
flight_no	char(6)	Номер рейса
scheduled_departure	timestamptz	Время вылета по расписанию
scheduled_departure_local	timestamp	Время вылета по расписанию,
		местное время в пункте отправления
scheduled_arrival	timestamptz	Время прилёта по расписанию
scheduled_arrival_local	timestamp	Время прилёта по расписанию,
		местное время в пункте прибытия
scheduled_duration	interval	Планируемая продолжительность полета
departure_airport	char(3)	Код аэропорта отправления
departure_airport_name	text	Название аэропорта отправления
departure_city	text	Город отправления
arrival_airport	char(3)	Код аэропорта прибытия
arrival_airport_name	text	Название аэропорта прибытия
arrival_city	text	Город прибытия
status	varchar(20)	Статус рейса
aircraft_code	char(3)	Код самолета, ІАТА
actual_departure	timestamptz	Фактическое время вылета
actual_departure_local	timestamp	Фактическое время вылета,
		местное время в пункте отправления
actual_arrival	timestamptz	Фактическое время прилёта
actual_arrival_local	timestamp	Фактическое время прилёта,
		местное время в пункте прибытия
actual_duration	interval	Фактическая продолжительность полета

Над таблицей flights создано представление flights_v.

Представление содержит дополнительную информацию:

- расшифровку данных об аэропорте вылета (departure_airport, departure_airport_name, departure_city),
- расшифровку данных об аэропорте прибытия (arrival_airport, arrival_airport_name, arrival_city),
- местное время вылета (scheduled_departure_local, actual_departure_local),
- местное время прибытия (scheduled_arrival_local, actual_arrival_local),
- продолжительность полета (scheduled duration, actual duration).

Материализованное представление routes (маршруты)

Таблица рейсов содержит избыточность: из нее можно было бы выделить информацию о маршруте (номер рейса, аэропорты отправления и назначения), которая не зависит от конкретных дат рейсов. Именно такая информация и составляет материализованное представление routes.

Столбец	Тип	Описание
flight_no departure_airport departure_airport_name departure_city arrival_airport arrival_airport_name arrival_city aircraft_code duration days_of_week	char(6) char(3) text text char(3) text text char(3) text char(3) interval integer[]	Номер рейса Код аэропорта отправления Название аэропорта отправления Город отправления Код аэропорта прибытия Название аэропорта прибытия Город прибытия Код самолета, IATA Продолжительность полета

2.3 Описание логики

Основной сущностью является бронирование (bookings). В одно бронирование можно включить несколько пассажиров, каждому из которых выписывается отдельный билет (tickets), соответственно, в одно бронирование может входить несколько билетов.

Как таковой пассажир не является отдельной сущностью. Как имя, так и номер документа пассажира могут меняться с течением времени, так что невозможно однозначно найти все билеты одного человека; для простоты можно считать, что все пассажиры уникальны.

Каждый билет включает один или несколько перелетов (ticket_flights). При этом, все билеты в одном бронировании имеют одинаковый набор перелетов.

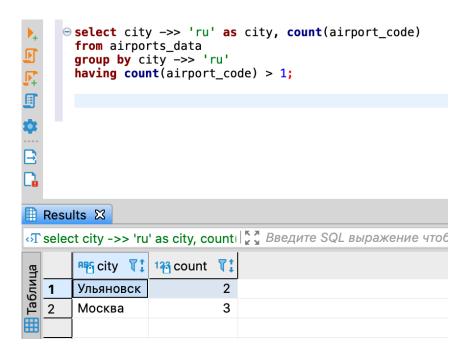
Каждый рейс (flights) - перелет следует из одного аэропорта (airports) в другой. Рейсы с одним номером имеют одинаковые пункты вылета и назначения, но будут отличаться датой отправления.

При регистрации на рейс пассажиру выдается посадочный талон (boarding_passes), в котором указано место в самолете. Комбинация рейса и места в самолете должна быть уникальной, чтобы не допустить выдачу двух посадочных талонов на одно место.

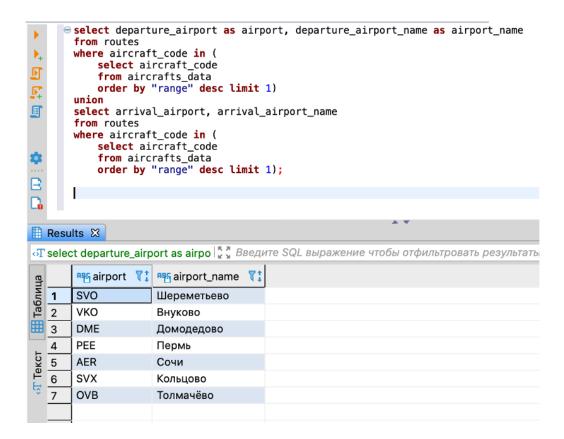
Количество мест (seats) в самолете и их распределение по классам обслуживания зависит от модели самолета (aircrafts), выполняющего рейс. Предполагается, что каждая модель самолета имеет только одну компоновку салона. Схема данных не контролирует, что места в посадочных талонах соответствуют имеющимся в самолете (такая проверка может быть сделана с использованием табличных триггеров или в приложении).

3 Решение задач на написание SQL-запросов к базе данных

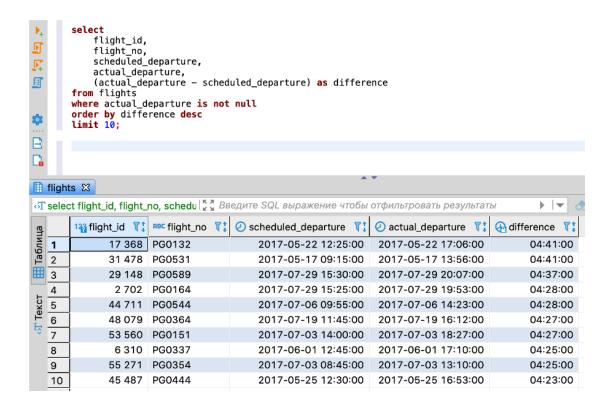
1. В каких городах больше одного аэропорта?



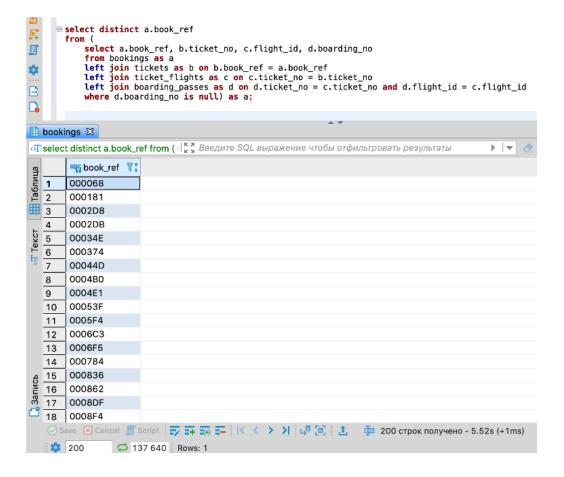
2. В каких аэропортах есть рейсы, выполняемые самолетом с максимальной дальностью перелета?



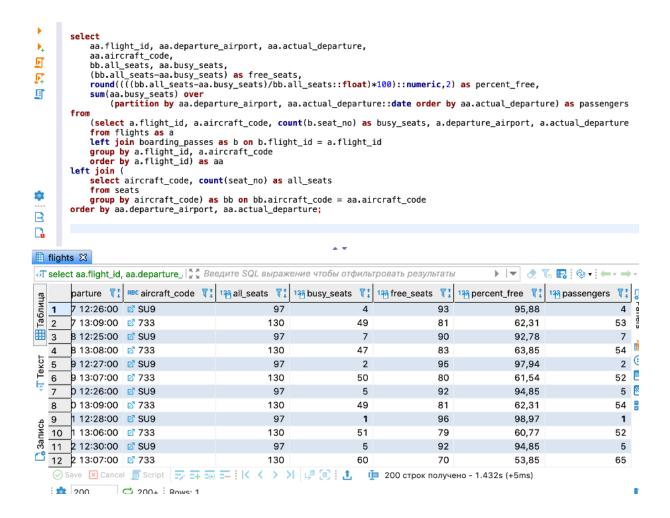
3. Вывести 10 рейсов с максимальным временем задержки вылета.



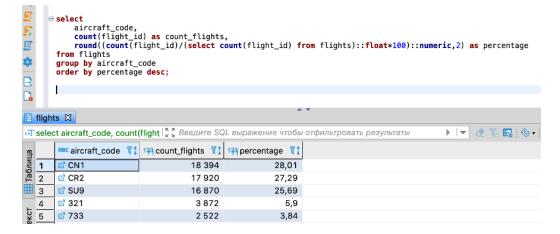
4. Были ли брони, по которым не были получены посадочные талоны?



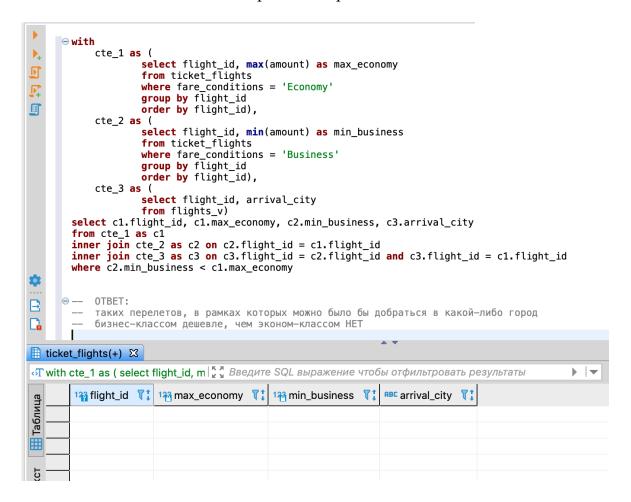
5. Найдите свободные места для каждого рейса, их % отношение к общему количеству мест в самолете. Добавьте столбец с накопительным итогом - суммарное количество вывезенных пассажиров из аэропорта за день. Т.е. в этом столбце должна отражаться сумма - сколько человек уже вылетело из данного аэропорта на этом или более ранних рейсах за день.



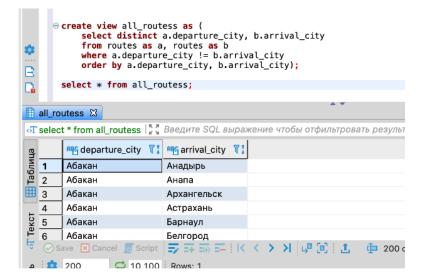
6. Найдите процентное соотношение перелетов по типам самолетов от общего количества.



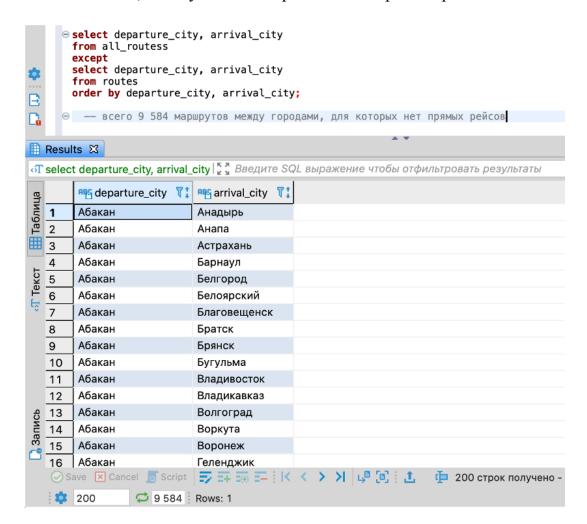
7. Были ли города, в которые можно добраться бизнес - классом дешевле, чем эконом-классом в рамках перелета?



- 8. Между какими городами нет прямых рейсов?
- Создали представление для всех возможных прямых рейсов между городами:



- Нашли, между какими городами нет прямых рейсов:



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

Кратчайшее расстояние между двумя точками А и В на земной поверхности (если принять ее за сферу) определяется зависимостью:

```
d = arccos {sin(latitude_a)·sin(latitude_b) + cos(latitude_a)·cos(latitude_b)·cos(longitude_a - longitude_b)}, где latitude_a и latitude_b - широты, longitude_a, longitude_b - долготы данных пунктов,
```

d - расстояние между пунктами, измеряется в радианах длиной дуги большого круга земного шара.

Расстояние между пунктами, измеряемое в километрах, определяется по формуле: $L = d \cdot R$, где R = 6371 км - средний радиус земного шара.

