

Проектная работа по модулю «SQL и получение данных»

Краткое описание таблиц и представлений базы данных.

В данной работе используется демонстрационная база данных (БД) пассажирских авиаперевозок по России, доступная на edu.postgrespro.ru.

База данных состоит из 8 таблиц. На ее основе созданы представления, основные из которых указаны в таблице 1.

Табл. 1. Таблицы и представления БД пассажирских авиаперевозок

Название	Описание	Тип	Размер	
			строк	байтов
aircrafts_data	самолеты	таблица	9	1 kB
airports_data	аэропорты	таблица	104	24 kB
boarding_passes	посадочные талоны	таблица	1 894 295	108 MB
bookings	бронирования	таблица	593 433	29 MB
flights	рейсы	таблица	65 664	6,3 MB
seats	места в салоне	таблица	1 339	64 kB
tickets	билеты	таблица	829 071	108 MB
ticket_flights	перелеты	таблица	2 360 335	153 MB
aircrafts	самолеты	представление	0	0
airports	аэропорты	представление	0	0
flights_v	рейсы	представление	0	0
routes	маршруты	представление	0	0
unique_flights	уникальные перелеты	представление	0	0
rownumber	варианты билетов на рейсах	материализованное представление	103 472	15 MB

БД содержит временной «срез» данных – как будто в некоторый момент была сделана резервная копия. Она сформирована на 18:00:00 15.08.2017. Относительно этой временной точки все рейсы (перелеты) делятся на прошедшие и будущие.

Развернутый анализ базы данных.

БД отражает приближенную к реальности ситуацию. Предполагается, что некая авиакомпания выполняет пассажирские перевозки на территории России. Она обладает собственным парком самолетов различных моделей (см. `aircrafts_data`). Каждая модель самолета имеет код IATA, Международной ассоциации авиаперевозчиков. Аэропорты (`airports_data`) также имеют уникальные коды. Местоположение аэропортов характеризуется географическими координатами – долготой и широтой, а также часовым поясом. В одном городе может быть несколько аэропортов, поэтому для аэропортов указаны названия городов. Между городами сформированы маршруты перелетов. Предполагается, что каждый маршрут имеет аэропорт отправления и аэропорт посадки (промежуточные посадки отсутствуют). Он также идентифицируется отдельным кодом. Перечень маршрутов является основой составления расписания полетов. В расписании указывается номер рейса, время и аэропорт вылета, время и аэропорт прибытия, самолет. При фактическом выполнении рейсов возникает его статус, фактическое время вылета и прилета (`flights`). Для совершения перелета пассажир бронирует авиабилет(ы). В одно бронирование можно включить несколько пассажиров, каждому из которых выписывается отдельный билет (`tickets`). Билет включает один или несколько перелетов (`ticket_flights`). Несколько перелетов могут включаться в билет в двух случаях:

- а) отсутствует прямой рейс, соединяющий пункты отправления и назначения (полет с пересадками);
- б) необходим билет «туда и обратно».

Для каждого перелета указываются: номер рейса, аэропорты отправления и назначения, время вылета и время прибытия, стоимость перелета и класс обслуживания.

Для каждой процедуры бронирования указывается дата бронирования и рассчитывается стоимость оформленных билетов. Все билеты в одном бронировании имеют одинаковый набор перелетов (допущение).

Для билета наряду с его уникальным номером указываются номер бронирования, идентификатор пассажира, его имя и фамилия, контактные данные.

При регистрации на рейс пассажиру выдается посадочный талон (boarding_passes), в котором указано место в самолете. Пассажир может зарегистрироваться только на тот рейс, который есть у него в билете. Комбинация рейса и места в самолете должна быть уникальной, чтобы не допустить выдачу двух посадочных талонов на одно место.

Каждому креслу в салоне самолета соответствует определенный класс обслуживания (seats). Количество мест в самолете и их распределение по классам обслуживания зависит от модели самолета, выполняющего рейс. Предполагается, что каждая модель имеет только одну компоновку салона. Схема данных не контролирует, что места в посадочных талонах соответствуют имеющимся в самолете.

Далее представлено описание таблиц БД и основных представлений, созданных на их основе.

airports_data (Аэропорты)

Аэропорт идентифицируется трехбуквенным кодом (airport_code) и имеет своё имя (airport_name). Для города не предусмотрено отдельной сущности, но введено поле с названием города (city), позволяющее найти аэропорты одного города. Для аэропорта также указаны его географические координаты (coordinates) и часовой пояс (timezone).

Поля airport_name и city этой таблицы содержат названия на двух языках, в формате jsonb. По этой таблице сделано представление airports. В нем значения полей airport_name и city определяются в зависимости от выбранного языка.

База данных содержит список аэропортов практически всех крупных городов России.

Табл. 3.

Столбец	Тип	Модификаторы	Описание
airport_code	char(3)	not null	Код аэропорта
airport_name	jsonb	not null	Название аэропорта
city	jsonb	not null	Город
coordinates	point	not null	Координаты аэропорта (долгота и широта)
timezone	text	not null	Часовой пояс аэропорта

Индексы:

PRIMARY KEY, btree (airport_code)

Ссылки извне:

TABLE "flights" FOREIGN KEY (arrival_airport)

REFERENCES airports_data(airport_code)

TABLE "flights" FOREIGN KEY (departure_airport)

REFERENCES airports_data(airport_code)

aircrafts_data (Самолеты)

Таблица содержит данные по модельному ряду самолетов авиаперевозчика. Каждая модель идентифицируется своим трехзначным кодом (aircraft_code). Поле model этой таблицы содержит названия моделей самолетов на двух языках, в формате jsonb.

По этой таблице создано представление aircrafts, где значение поля model определяется в зависимости от выбранного языка.

Табл. 2.

Столбец	Тип	Модификаторы	Описание
aircraft_code	char(3)	not null	Код самолета, IATA
model	jsonb	not null	Модель самолета
range	integer	not null	Максимальная дальность полета, км

Индексы:

PRIMARY KEY, btree (aircraft_code)

Ограничения-проверки:

CHECK (range > 0)

Ссылки извне:

TABLE "flights" FOREIGN KEY (aircraft_code)

REFERENCES aircrafts_data(aircraft_code)

TABLE "seats" FOREIGN KEY (aircraft_code)

boarding_passes (Посадочные талоны)

Посадочный талон выдается пассажиру при регистрации на рейс, которая возможна за сутки до плановой даты вылета. Он идентифицируется также 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).

Табл. 4.

Столбец	Тип	Модификаторы	Описание
ticket_no	char(13)	not null	Номер билета
flight_id	integer	not null	Идентификатор рейса
boarding_no	integer	not null	Номер посадочного талона
seat_no	varchar(4)	not null	Номер места

Индексы:

PRIMARY KEY, btree (ticket_no, flight_id)
 UNIQUE CONSTRAINT, btree (flight_id, boarding_no)
 UNIQUE CONSTRAINT, btree (flight_id, seat_no)

Ограничения внешнего ключа:

FOREIGN KEY (ticket_no, flight_id)
 REFERENCES ticket_flights(ticket_no, flight_id)

bookings (Бронирования)

Таблица содержит данные по бронированию пассажирами билетов. Бронирование, как правило, осуществляется заранее. Для каждой процедуры бронирования указывается дата бронирования (book_date) и рассчитывается стоимость оформленных билетов (total_amount). Бронирование идентифицируется номером (book_ref, представляющим собой шестизначную комбинацию букв и цифр).

Табл. 5.

Столбец	Тип	Модификаторы	Описание
book_ref	char(6)	not null	Номер бронирования
book_date	timestampz	not null	Дата бронирования
total_amount	numeric(10,2)	not null	Полная сумма бронирования

Индексы:

PRIMARY KEY, btree (book_ref)

Ссылки извне:

TABLE "tickets" FOREIGN KEY (book_ref) REFERENCES bookings(book_ref)

flights (Рейсы)

Таблица содержит информацию по всем прямым рейсам. 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 запланированные дата и время вылета (scheduled_departure) и прибытия (scheduled_arrival). Фактические время вылета (actual_departure) и прибытия (actual_arrival) могут отличаться, если рейс задержан.

Прямые рейсы соединяют далеко не все аэропорты, но из любого можно добраться в любой другой с несколькими пересадками.

Табл. 6.

Столбец	Тип	Модификаторы	Описание
flight_id	serial	not null	Идентификатор рейса
flight_no	char(6)	not null	Номер рейса

scheduled_departure	timestampz	not null	Время вылета по расписанию
scheduled_arrival	timestampz	not null	Время прилёта по расписанию
departure_airport	char(3)	not null	Аэропорт отправления
arrival_airport	char(3)	not null	Аэропорт прибытия
status	varchar(20)	not null	Статус рейса
aircraft_code	char(3)	not null	Код самолета, IATA
actual_departure	timestampz		Фактическое время вылета
actual_arrival	timestampz		Фактическое время прилёта

Индексы:

PRIMARY KEY, btree (flight_id)
 UNIQUE CONSTRAINT, btree (flight_no, scheduled_departure)

Ограничения-проверки:

CHECK (scheduled_arrival > scheduled_departure)
 CHECK ((actual_arrival IS NULL)
 OR ((actual_departure IS NOT NULL AND actual_arrival IS NOT NULL)
 AND (actual_arrival > actual_departure)))
 CHECK (status IN ('On Time', 'Delayed', 'Departed',
 'Arrived', 'Scheduled', 'Cancelled'))

Ограничения внешнего ключа:

FOREIGN KEY (aircraft_code)
 REFERENCES aircrafts(aircraft_code)
 FOREIGN KEY (arrival_airport)
 REFERENCES airports(airport_code)
 FOREIGN KEY (departure_airport)
 REFERENCES airports(airport_code)

Ссылки извне:

TABLE "ticket_flights" FOREIGN KEY (flight_id)
 REFERENCES flights(flight_id)

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).

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

Статус	Пояснение
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 отменён.

seats (Места в салоне)

После регистрации на рейс выписывается посадочный талон с указанием места в самолете. 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.

Табл. 7.

Столбец	Тип	Модификаторы	Описание
aircraft_code	char(3)	not null	Код самолета, IATA
seat_no	varchar(4)	not null	Номер места
fare_conditions	varchar(10)	not null	Класс обслуживания

Индексы:

PRIMARY KEY, btree (aircraft_code, seat_no)

Ограничения-проверки:

CHECK (fare_conditions IN ('Economy', 'Comfort', 'Business'))

Ограничения внешнего ключа:

FOREIGN KEY (aircraft_code)

REFERENCES aircrafts(aircraft_code) ON DELETE CASCADE

tickets (Билеты)

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

Табл. 8.

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

Индексы:

PRIMARY KEY, btree (ticket_no)

Ограничения внешнего ключа:

FOREIGN KEY (book_ref) REFERENCES bookings(book_ref)

Ссылки извне:

TABLE "ticket_flights" FOREIGN KEY (ticket_no) REFERENCES tickets(ticket_no)

ticket_flights (Перелеты)

Перелет соединяет билет с рейсом и идентифицируется их номерами.

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

Табл. 9.

Столбец	Тип	Модификаторы	Описание
ticket_no	char(13)	not null	Номер билета
flight_id	integer	not null	Идентификатор рейса
fare_conditions	varchar(10)	not null	Класс обслуживания
amount	numeric(10,2)	not null	Стоимость перелета

Индексы:

PRIMARY KEY, btree (ticket_no, flight_id)

Ограничения-проверки:

CHECK (amount >= 0)

CHECK (fare_conditions IN ('Economy', 'Comfort', 'Business'))

Ограничения внешнего ключа:

FOREIGN KEY (flight_id) REFERENCES flights(flight_id)

FOREIGN KEY (ticket_no) REFERENCES tickets(ticket_no)

Ссылки извне:

TABLE "boarding_passes" FOREIGN KEY (ticket_no, flight_id)

REFERENCES ticket_flights(ticket_no, flight_id)

flights_v (Рейсы)

Это представление содержит дополнительную информацию по рейсам:

- 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`.

Им удобно пользоваться, когда, например, нужно найти ближайший рейс по какому-нибудь маршруту.

Табл. 10.

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

routes (Маршруты)

Это представление содержит информацию обо всех рейсах, которая не зависит от конкретных дат вылетов. Его удобно использовать чтобы узнать, откуда и куда, в какие дни недели можно улететь, какова будет планируемая продолжительность полета.

Табл. 11.

Столбец	Тип	Описание
<code>flight_no</code>	<code>char(6)</code>	Номер рейса
<code>departure_airport</code>	<code>char(3)</code>	Код аэропорта отправления
<code>departure_airport_name</code>	<code>text</code>	Название аэропорта отправления
<code>departure_city</code>	<code>text</code>	Город отправления
<code>arrival_airport</code>	<code>char(3)</code>	Код аэропорта прибытия
<code>arrival_airport_name</code>	<code>text</code>	Название аэропорта прибытия
<code>arrival_city</code>	<code>text</code>	Город прибытия
<code>aircraft_code</code>	<code>char(3)</code>	Код самолета, IATA
<code>duration</code>	<code>interval</code>	Продолжительность полета
<code>days_of_week</code>	<code>interval</code>	Дни недели, когда выполняются рейсы

rownumber (Варианты билетов на рейсах)

Материализованное представление хранит информацию о существующих вариантах билетов на рейсах, которые различаются классом обслуживания и стоимостью (стоимость для одного и того же класса обслуживания на данном рейсе может быть разной).

Табл. 13.

Столбец	Тип	Описание
<code>flight_id</code>	<code>integer</code>	Идентификатор рейса

flight_no	char(6)	Номер рейса
fare_conditions	varchar(10)	Класс обслуживания
amount	numeric(10,2)	Стоимость перелета
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
number_row	integer	Номер строки в группе
count_in_group_	integer	Кол-во строк в группе

unique_flights (Уникальные перелеты)

Это представление служит источником данных для материализованного представления rownumber.

ER – диаграмма

Ссылка на схему: [здесь](#)

Список SQL запросов с их описанием

Все SQL-запросы приведены в файле Script-fw.sql. Он доступен по этой ссылке: [здесь](#)

Самостоятельно развернутая СУБД

Для установки использовалась версия demo_small.zip (21 МБ) демонстрационной базы, содержащая данные по полётам за один месяц (размер БД 265 МБ).

Результаты выполнения SQL-скрипта, создающего базу данных demo представлены скриншотами: [здесь](#) и [здесь](#).