**Проектная работа по модулю**

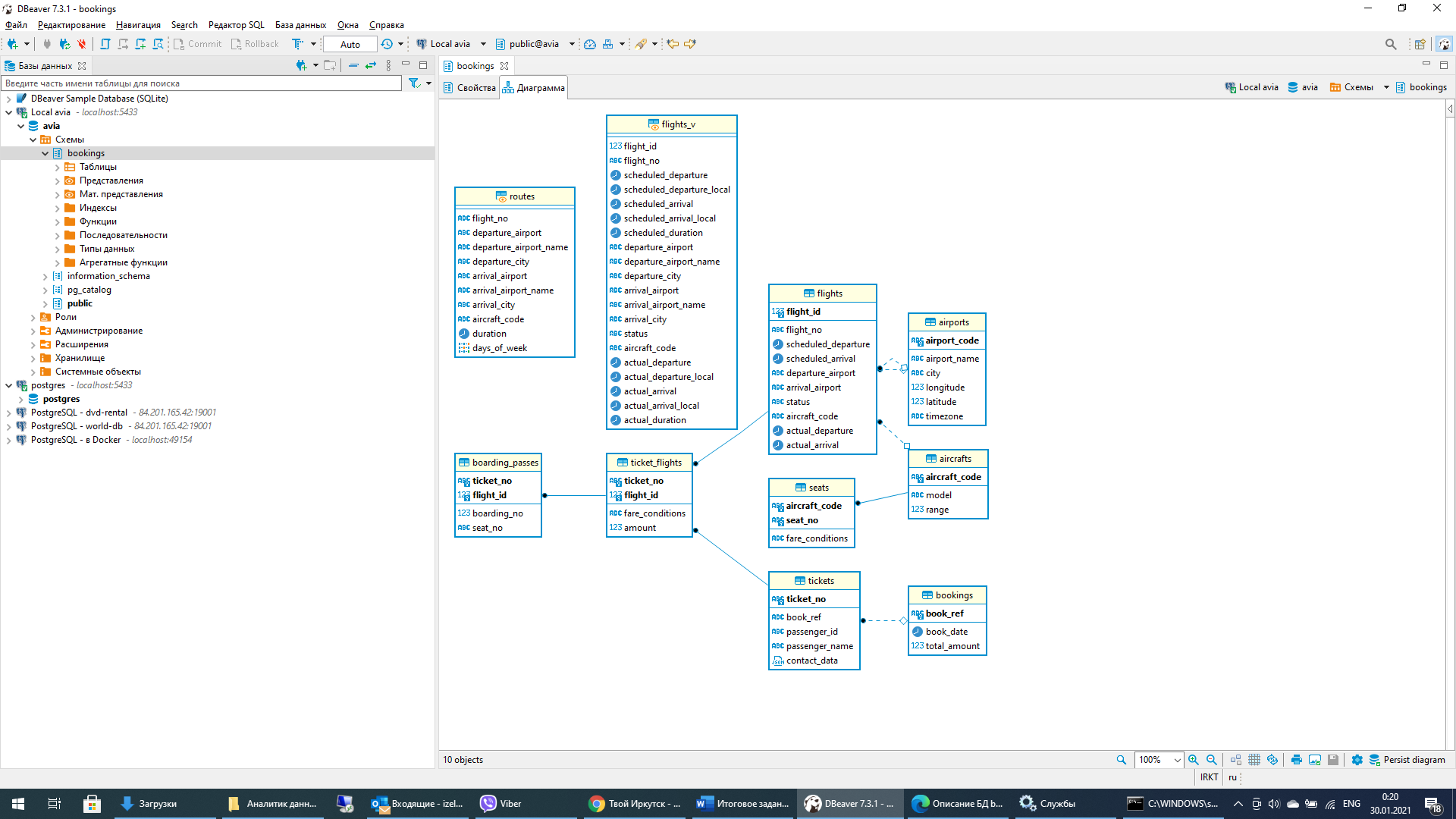
**“SQL и получение данных”**

**Зельберг Ирина,**

**группа SQL-24, обучение профессии «Аналитик данных», группа DA-21**

Приложение №1

1. В работе использовался тип подключения №3: восстановление базы из avia.backup файла.
2. Скриншот ER-диаграммы из DBeaver`a согласно подключения.



1. Краткое описание БД - из каких таблиц и представлений состоит.

БД Авиаперевозки ‘bookings’ состоит из 8 таблиц/отношений, 1 представления, 1 материализованного представления:

|  |  |  |  |
| --- | --- | --- | --- |
| avia | bookings | ticket\_flights | BASE TABLE |
| avia | bookings | boarding\_passes | BASE TABLE |
| avia | bookings | aircrafts | BASE TABLE |
| avia | bookings | flights | BASE TABLE |
| avia | bookings | airports | BASE TABLE |
| avia | bookings | seats | BASE TABLE |
| avia | bookings | tickets | BASE TABLE |
| avia | bookings | bookings | BASE TABLE |
| avia | bookings | flights\_v | VIEW |
| avia | bookings | routes | MATERIALIZED VIEW |

1. Развернутый анализ БД - описание таблиц, логики, связей и бизнес области

База данных может быть использована для запросов по дальности перелетов, востребованности перелетов, заполняемости рейсов, нагруженности аэропортов, интенсивности использования самолетов, анализа стоимости билетов, статистики задержек вылетов, планированию маршрутов перелетов, по клиентской базе пассажиров.

Таблица airports является справочником аэропортов, с указанием международного кода аэропорта, названия аэропорта, города, координат и временной зоны.

Таблица aircrafts является справочником типов самолетов.

Таблица seats является расширением справочника самолетов с указанием номеров мест и класса каждого места.

Таблица tickets – данные по купленным билетам.

Таблица boarding\_passes – данные по выданным посадочным талонам (фактически совершенным вылетам)

Таблица ticket\_flights связывает рейсы, билеты и посадочные талоны

Таблица bookings – данные по бронированиям, может включать несколько билетов.

Таблица flights содержит данные, которые используются для большого количества запросов из БД. Это расписание рейсов со статусом их фактического выполнения: номер рейса, времена вылета и прибытия по расписанию и по факту, текущий статус рейса, коды аэропортов вылета и прибытия, вид самолета.

Имеет наибольшее число связей с другими таблицами БД.

1. Список SQL запросов с описанием логики их выполнения.

Использованы виды функций, операторов, команд:

* Создание подзапросов cte
* Создание оконных функций partition by
* Объединение данных из таблиц join, right join,
* Группировка group by
* Сортировка order by
* Выбор по условиям where, having
* Ограничение вывода строк limit
* Исключение повторяющихся строк distinct
* Подсчет числа записей count
* Сумма sum
* Округление round
* Перевод числовых значений из целых чисел в числа с плавающей точкой decimal
* Перевод даты из timestamp в date
* Создание представления create or replace view

Комментарии по запросам указаны в sql-файле.

Баллы за оформление:

1. 0 - облачная база, 10 - локальная база
2. 5
3. 10
4. 20
5. 15

**Итого: максимум 60 баллов.**

Приложение №2

|  |  |  |
| --- | --- | --- |
| № | Вопрос | В решении обязательно должно быть использовано |
| 1 | В каких городах больше одного аэропорта? |  |
| 2 | В каких аэропортах есть рейсы, выполняемые самолетом с максимальной дальностью перелета? | - Подзапрос |
| 3 | Вывести 10 рейсов с максимальным временем задержки вылета | - Оператор LIMIT |
| 4 | Были ли брони, по которым не были получены посадочные талоны? | - Верный тип JOIN |
| 5 | Найдите свободные места для каждого рейса, их % отношение к общему количеству мест в самолете.  Добавьте столбец с накопительным итогом - суммарное накопление количества вывезенных пассажиров из каждого аэропорта на каждый день. Т.е. в этом столбце должна отражаться накопительная сумма - сколько человек уже вылетело из данного аэропорта на этом или более ранних рейсах за день. | - Оконная функция  - Подзапросы |
| 6 | Найдите процентное соотношение перелетов по типам самолетов от общего количества. | - Подзапрос  - Оператор ROUND |
| 7 | Были ли города, в которые можно добраться бизнес - классом дешевле, чем эконом-классом в рамках перелета? | - CTE |
| 8 | Между какими городами нет прямых рейсов? | - Декартово произведение в предложении FROM - Самостоятельно созданные представления  - Оператор EXCEPT |
| 9 | Вычислите расстояние между аэропортами, связанными прямыми рейсами, сравните с допустимой максимальной дальностью перелетов в самолетах, обслуживающих эти рейсы \* | - Оператор RADIANS или использование sind/cosd |

\* - В облачной базе координаты находятся в столбце airports\_data.coordinates - работаете, как с массивом. В локальной базе координаты находятся в столбцах airports.longitude и airports.latitude.

Кратчайшее расстояние между двумя точками A и B 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·R, где R = 6371 км — средний радиус земного шара.

Баллы за запросы:

1. 10
2. 15
3. 15
4. 15
5. 25
6. 25
7. 25
8. 35
9. 35

**Итого: максимум 200 баллов.**

Для зачета необходимо набрать **минимум 130 баллов**.