Руслан Сафаров

DSPR 14

Задание 4.1

База данных содержит список аэропортов практически всех крупных городов России. В большинстве городов есть только один аэропорт. Исключение составляет:

**SELECT** ap.city

**FROM** dst\_project.airports ap

**GROUP** **BY** 1

**HAVING** **count**(ap.airport\_name) > 1

Задание 4.2

**Вопрос 1.** Таблица рейсов содержит всю информацию о прошлых, текущих и запланированных рейсах. Сколько всего статусов для рейсов определено в таблице?

**SELECT** **count**(**DISTINCT** fl.status)

**FROM** dst\_project.flights fl

**Вопрос 2.** Какое количество самолетов находятся в воздухе на момент среза в базе (статус рейса «самолёт уже вылетел и находится в воздухе»).

**SELECT** **count**(fl.flight\_id)

**FROM** dst\_project.flights fl

**WHERE** fl.status = 'Departed'

**Вопрос 3.** Места определяют схему салона каждой модели. Сколько мест имеет самолет модели  (*Boeing 777-300*)?

**SELECT** **count**(s.seat\_no)

**FROM** dst\_project.seats s

**WHERE** s.aircraft\_code = '773'

**Вопрос 4.** Сколько состоявшихся (фактических) рейсов было совершено между 1 апреля 2017 года и 1 сентября 2017 года?

**SELECT** **count**(\*)

**FROM** dst\_project.flights fl

**WHERE** (fl.actual\_arrival **BETWEEN** '2017-04-01 00:00:00' **AND** '2017-09-01 00:00:00')

**AND** fl.status != 'Cancelled'

Задание 4.3

**Вопрос 1.** Сколько всего рейсов было отменено по данным базы?

**SELECT** **count**(fl.flight\_id)

**FROM** dst\_project.flights fl

**WHERE** fl.status = 'Cancelled'

**Вопрос 2.** Сколько самолетов моделей типа *Boeing, Sukhoi Superjet, Airbus* находится в базе авиаперевозок?

**SELECT** 'Boeing' ac\_name,

**count**(\*)

**FROM** dst\_project.aircrafts ac

**WHERE** ac.model **ILIKE** '%boeing%'

**UNION**

**SELECT** 'Sukhoi Superjet' ac\_name,

**count**(\*)

**FROM** dst\_project.aircrafts ac

**WHERE** ac.model **ILIKE** '%sukhoi superjet%'

**UNION**

**SELECT** 'Airbus' ac\_name,

**count**(\*)

**FROM** dst\_project.aircrafts ac

**WHERE** ac.model **ILIKE** '%airbus%'

**Вопрос 3.** В какой части (частях) света находится больше аэропортов?

**SELECT** 'Asia' POINT,

**count**(\*)

**FROM** dst\_project.airports ap

**WHERE** ap.timezone **ILIKE** '%asia%'

**UNION**

**SELECT** 'Europe' POINT,

**count**(\*)

**FROM** dst\_project.airports ap

**WHERE** ap.timezone **ILIKE** '%europe%'

**Вопрос 4.** У какого рейса была самая большая задержка прибытия за все время сбора данных? Введите *id* рейса (flight\_id).

**SELECT** fl.flight\_id,

(fl.scheduled\_arrival-fl.actual\_arrival) val

**FROM** dst\_project.flights fl

**ORDER** **BY** 2 **ASC**

**LIMIT** 1

Задание 4.4

**Вопрос 1.** Когда был запланирован самый первый вылет, сохраненный в базе данных?

**SELECT** fl.scheduled\_departure

**FROM** dst\_project.flights fl

**ORDER** **BY** 1 **ASC**

**LIMIT** 1

**Вопрос 2.** Сколько минут составляет запланированное время полета в самом длительном рейсе?

**SELECT** (fl.scheduled\_departure-scheduled\_arrival)

**FROM** dst\_project.flights fl

**ORDER** **BY** 1 **ASC**

**LIMIT** 1

**Вопрос 3.** Между какими аэропортами пролегает самый длительный по времени запланированный рейс?

**SELECT** fl.scheduled\_departure-fl.scheduled\_arrival,

fl.departure\_airport,

fl.arrival\_airport

**FROM** dst\_project.flights fl

**ORDER** **BY** 1 **ASC**

**LIMIT** 1

**Вопрос 4.** Сколько составляет средняя дальность полета среди всех самолетов в минутах? Секунды округляются в меньшую сторону (отбрасываются до минут)

**SELECT** **avg**(fl.scheduled\_departure-fl.scheduled\_arrival)

**FROM** dst\_project.flights fl

date\_part('hour', duration) \* 60 + date\_part('minute', duration) as duration\_mm

Задание 4.5

**Вопрос 1.** Мест какого класса у *SU9* больше всего?

**SELECT** s.aircraft\_code,

**count**(s.seat\_no),

fare\_conditions

**FROM** dst\_project.seats s

**WHERE** s.aircraft\_code = 'SU9'

**GROUP** **BY** 1,

3

**ORDER** **BY** 2 **DESC**

**LIMIT** 1

**Вопрос 2.** Какую самую минимальную стоимость составило бронирование за всю историю?

**SELECT** **min**(b.total\_amount)

**FROM** dst\_project.bookings b

**Вопрос 3.** Какой номер места был у пассажира с id = 4313 788533?

**SELECT** bp.seat\_no

**FROM** dst\_project.boarding\_passes bp

**WHERE** bp.ticket\_no =

(**SELECT** t.ticket\_no tn

**FROM** dst\_project.tickets t

**WHERE** t.passenger\_id = '4313 788533' )

Задание 5.1

**Вопрос 1.** Анапа — курортный город на юге России. Сколько рейсов прибыло в Анапу за 2017 год?

**SELECT** **count**(fl.flight\_id)

**FROM** dst\_project.flights fl

**WHERE** fl.actual\_arrival **BETWEEN** '20170101' **AND** '20171231'

**AND** fl.arrival\_airport =

(**SELECT** a.airport\_code

**FROM** dst\_project.airports a

**WHERE** a.city = 'Anapa')

**Вопрос 2.** Сколько рейсов из Анапы вылетело зимой 2017 года?

**SELECT** **count**(\*)

**FROM** dst\_project.flights fl

**WHERE** fl.departure\_airport = 'AAQ'

**AND** (date\_trunc('month', fl.scheduled\_departure) **in**

('2017-01-01', '2017-02-01', '2017-12-01'))

**AND** fl.status **not** **in** ('Cancelled')

**Вопрос 3.** Посчитайте количество отмененных рейсов из Анапы за все время.

**SELECT** **count**(fl.flight\_id)

**FROM** dst\_project.flights fl

**WHERE** fl.status = 'Cancelled'

**AND** fl.departure\_airport =

(**SELECT** a.airport\_code

**FROM** dst\_project.airports a

**WHERE** a.city = 'Anapa')

**Вопрос 4.** Сколько рейсов из Анапы не летают в Москву?

**SELECT** **count**(fl.flight\_id)

**FROM** dst\_project.flights fl

**JOIN** dst\_project.airports **ON** fl.arrival\_airport = airports.airport\_code

**WHERE** fl.departure\_airport = 'AAQ'

**AND** airports.city != 'Moscow'

**Вопрос 5.** Какая модель самолета летящего на рейсах из Анапы имеет больше всего мест?

**SELECT** ac.model,

**count**(s.seat\_no)

**FROM** dst\_project.flights fl

**JOIN** dst\_project.seats s **ON** fl.aircraft\_code = s.aircraft\_code

**JOIN** dst\_project.aircrafts ac **ON** fl.aircraft\_code = ac.aircraft\_code

**JOIN** dst\_project.airports a **ON** fl.departure\_airport = a.airport\_code

**WHERE** fl.departure\_airport = 'AAQ'

**GROUP** **BY** 1

**ORDER** **BY** 2 **DESC**

**LIMIT** 1

ФИНАЛЬНОЕ ЗАДАНИЕ

**WITH** table\_1 **AS**

(*-- Информация о полётах и самолётах*

**SELECT** fl.flight\_id,

fl.flight\_no,

fl.aircraft\_code,

fl.arrival\_airport,

fl.departure\_airport,

fl.status,

fl.actual\_arrival,

fl.actual\_departure

**FROM** dst\_project.flights fl

**WHERE** fl.departure\_airport = 'AAQ'

**AND** fl.arrival\_airport != 'NOZ' *--отметаем новокузнецк из датасета сразу*

**AND** (date\_trunc('month', fl.scheduled\_departure) **in** ('2017-01-01',

'2017-02-01',

'2017-12-01'))

**AND** fl.status **not** **in** ('Cancelled') ),

table\_2 **AS**

(*-- Информация о полётах и стоимости проданных билетов (группировка)*

**SELECT** tf.flight\_id,

tf.fare\_conditions fc,

**sum**(tf.amount) summa,

**count**(flight\_id) cnt

**FROM** dst\_project.ticket\_flights tf

**GROUP** **BY** 1,

2),

table\_3 **AS**

(*-- Информация о полётах и кол-во проданных билетов на разные классы*

**SELECT** s.aircraft\_code,

ac.model,

**count**(s.seat\_no) qty,

s.fare\_conditions

**FROM** dst\_project.seats s,

dst\_project.aircrafts ac

**WHERE** s.aircraft\_code = ac.aircraft\_code

**GROUP** **BY** 1,

2,

4) *-- Финальный запрос объединяет все ваши созданные таблицы в одну и выводит нужную информацию*

**SELECT** t1.flight\_id, *-- t1.flight\_no, -- информация полезной смысловой нагрузки не несет*

t2.cnt pass\_qty, *-- число пассажиров конкретного класса*

t3.qty aircraft\_capacity, *-- вместимость судна*

t3.qty - t2.cnt empty\_seats, *-- пустые (невыкупленные) места*

t2.fc, *-- тип места бизнес/эконом*

t2.summa, *-- сумма стоимости купленных билетов на данный класс мест*

(t2.summa / t2.cnt \* (t3.qty - t2.cnt))::int недоп\_прибыль, *-- умножаем пустые места на среднюю стоиомсть места соотв класса и рейса, округляем для наглядности*

arr.city, *-- куда летим*

t3.model aircraft, *-- имя судна*

t1.actual\_arrival дата\_вылета,

(t2.summa / t2.cnt)::int цена\_билета\_средняя, *-- округляем для наглядности*

date\_trunc('day', t1.actual\_arrival),

to\_char(t1.actual\_arrival, 'Day') день\_недели, *-- день недели*

**EXTRACT**(EPOCH

**FROM** to\_char(t1.actual\_arrival - t1.actual\_departure, 'HH24:MI:SS')::INTERVAL)/60, *-- затраченное время в минутах*

**CASE**

**WHEN** t3.aircraft\_code = '733' **THEN** **EXTRACT**(EPOCH

**FROM** to\_char(t1.actual\_arrival - t1.actual\_departure, 'HH24:MI:SS')::INTERVAL)/60 \* 3000 *-- литров топлива на РЕЙС*

**WHEN** t3.aircraft\_code = 'SU9' **THEN** **EXTRACT**(EPOCH

**FROM** to\_char(t1.actual\_arrival - t1.actual\_departure, 'HH24:MI:SS')::INTERVAL)/60 \* 1700 *-- литров топлива на РЕЙС*

**END**

**FROM** table\_1 t1

**LEFT** **JOIN** table\_2 t2 **ON** t1.flight\_id = t2.flight\_id

**LEFT** **JOIN** dst\_project.airports arr **ON** t1.arrival\_airport = arr.airport\_code

**LEFT** **JOIN** table\_3 t3 **ON** t3.aircraft\_code = t1.aircraft\_code

**AND** t3.fare\_conditions = t2.fc

**ORDER** **BY** 7 **DESC** *-- выводим сначала самые пустые рейсы (где много свободных мест и большая недополученная прибыль)*