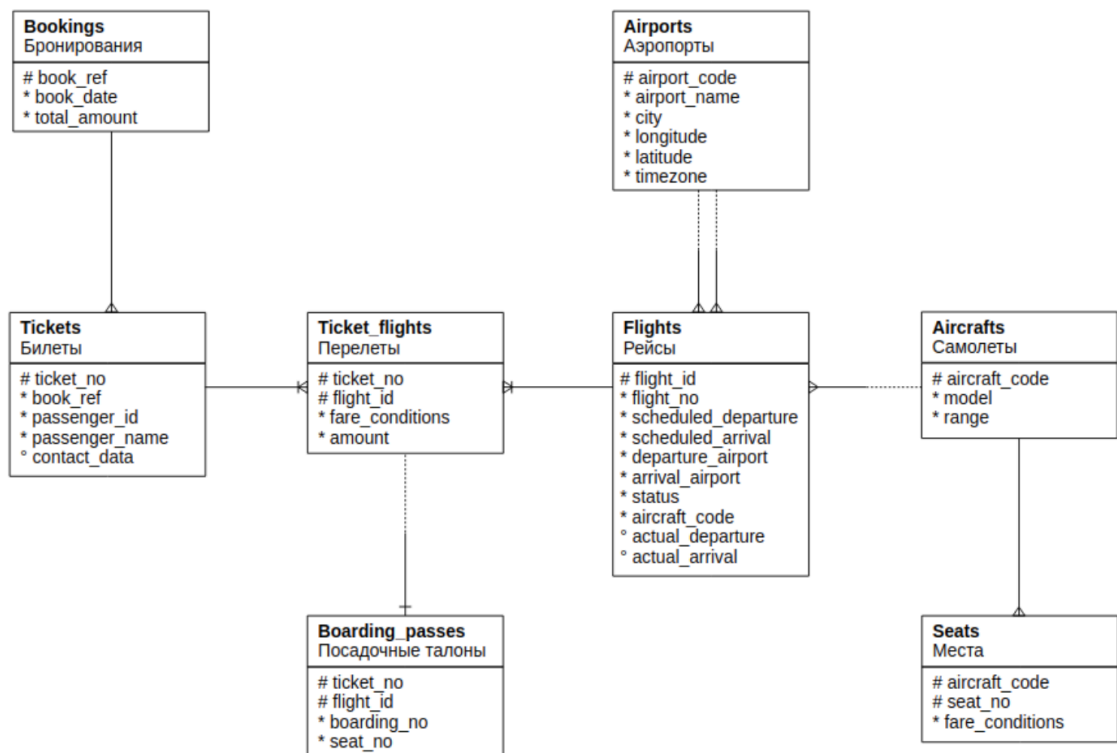


Project 4

SQL query summary



Задание 4.1

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

St. Petersburg, Moscow

Moscow, Ulyanovsk

Moscow, Anapa

Samara, Moscow/*

```
SELECT count(ap.airport_code),  
       ap.city  
FROM dst_project.airports ap  
WHERE ap.city = 'Anapa'  
       OR ap.city = 'Moscow'  
       OR ap.city = 'Ulyanovsk'  
       OR ap.city = 'St. Petersburg'  
       OR ap.city = 'Samara'  
GROUP BY ap.city
```

QUERY RESULTS:

count	city
1	Anapa
2	Ulyanovsk
3	Moscow
1	St. Petersburg
1	Samara

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 status)  
FROM dst_project.flight
```

QUERY RESULTS: 6

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

→ **Departed**

Самолет уже вылетел и находится в воздухе.

статус рейса «самолёт уже вылетел и находится в воздухе»

```
SELECT count(fl.status)
FROM dst_project.flights fl
WHERE fl.status = 'Departed'
```

QUERY RESULTS: 58

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

```
SELECT ac.model,
       ac.aircraft_code,
       COUNT (s.seat_no)
FROM dst_project.aircrafts ac
JOIN dst_project.seats s ON ac.aircraft_code = s.aircraft_code
WHERE ac.model = 'Boeing 777-300'
GROUP BY ac.model,
         ac.aircraft_code
```

QUERY RESULTS:

model	aircraft_code	Количество
Boeing 777-300	773	402

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

- Здесь и далее *состоявшийся рейс* означает, что он не отменён, и самолёт прибыл в пункт назначения.

```
SELECT count(flight_id)
```

```
FROM dst_project.flights fl
WHERE (fl.actual_arrival BETWEEN '2017-04-01 0:00:00' AND '2017-09-01 23:59:59')
AND (fl.status = 'Arrived')
```

QUERY RESULTS: 74227

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'
```

QUERY RESULTS: 437

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

```
SELECT 'Boeing' model,
       COUNT (air.model) count_of_aircrafts
FROM dst_project.aircrafts air
WHERE air.model like '%Boeing%'
UNION ALL
SELECT 'Sukhoi Superjet' model,
       COUNT (air.model) count_of_aircrafts
FROM dst_project.aircrafts air
WHERE air.model like '%Sukhoi Superjet%'
UNION ALL
SELECT 'Airbus' model,
       COUNT (air.model) count_of_aircrafts
FROM dst_project.aircrafts air
WHERE air.model like '%Airbus%'
```

QUERY RESULTS:

Boeing	3
Sukhoi Superjet	1
Airbus	3

Вопрос 3. В какой части (частях) света

находится больше аэропортов?

```
SELECT 'Europe' Continent,
       COUNT (airp.airport_code)
FROM dst_project.airports airp
WHERE airp.timezone like '%Europe%'
UNION ALL
SELECT 'Australia' Continent,
       COUNT (airp.airport_code)
FROM dst_project.airports airp
WHERE airp.timezone like '%Australia%'
UNION ALL
SELECT 'Europe, Asia' Continent,
       COUNT (airp.airport_code)
FROM dst_project.airports airp
WHERE airp.timezone like '%Europe%'
      OR airp.timezone like '%Asia%'
UNION ALL
SELECT 'Asia' Continent,
       COUNT (airp.airport_code)
FROM dst_project.airports airp
WHERE airp.timezone like '%Asia%'
```

QUERY RESULTS:

Europe	52
Australia	0
Europe, Asia	104
Asia	52

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 difference
FROM dst_project.flights fl
```

```
GROUP BY fl.flight_id
ORDER BY 2
LIMIT 1
```

QUERY RESULTS:

flight_id difference

157,571 0 years 0 mons 0 days -5 hours -7 mins -0.00 secs

Задание 4.4

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

```
SELECT min(fl.scheduled_departure)
FROM dst_project.flights fl
```

QUERY RESULTS: август 14, 2016, 11:45 вечера

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

```
SELECT MAX (date_part('hour', (fl.scheduled_arrival -
fl.scheduled_departure))*60 + date_part('minute', (fl.scheduled_arrival -
fl.scheduled_departure)))
FROM dst_project.flights fl
ORDER BY 1 DESC
```

QUERY RESULTS: 530

Вопрос 3. Между какими аэропортами пролегает самый длительный по времени запланированный рейс? DME - UUS or DME - AAQ or DME - PCS or SVO - UUS

```
SELECT fl.departure_airport,
       fl.arrival_airport,
       (fl.scheduled_arrival - fl.scheduled_departure) duration
FROM dst_project.flights fl
ORDER BY 3 DESC
LIMIT 1
```

QUERY RESULTS:

departure_airport

arrival_airport

duration

DME

UUS

0 years 0 mons 0 days 8 hours 50 mins 0.00 secs

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

```
SELECT round(AVG (date_part('hour', (fl.scheduled_arrival -  
fl.scheduled_departure))*60 + date_part('minute', (fl.scheduled_arrival -  
fl.scheduled_departure))))  
FROM dst_project.flights fl
```

QUERY RESULTS:

▼ round

128

Задание 4.5

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

Economy

Business

Standart

Comfort

```
SELECT 'Economy' Fare_Conditions,  
COUNT (tf.fare_conditions)  
FROM dst_project.ticket_flights tf  
JOIN dst_project.flights f ON f.flight_id = tf.flight_id  
WHERE f.aircraft_code = 'SU9'  
AND tf.fare_conditions = 'Economy'  
UNION ALL  
SELECT 'Business' Fare_Conditions,  
COUNT (tf.fare_conditions)  
FROM dst_project.ticket_flights tf  
JOIN dst_project.flights f ON f.flight_id = tf.flight_id  
WHERE f.aircraft_code = 'SU9'  
AND tf.fare_conditions = 'Business'  
UNION ALL  
SELECT 'Standart' Fare_Conditions,
```

```

COUNT (tf.fare_conditions)
FROM dst_project.ticket_flights tf
JOIN dst_project.flights f ON f.flight_id = tf.flight_id
WHERE f.aircraft_code = 'SU9'
AND tf.fare_conditions = 'Standart'
UNION ALL
SELECT 'Comfort' Fare_Conditions,
COUNT (tf.fare_conditions)
FROM dst_project.ticket_flights tf
JOIN dst_project.flights f ON f.flight_id = tf.flight_id
WHERE f.aircraft_code = 'SU9'
AND tf.fare_conditions = 'Comfort'
UNION ALL
SELECT 'Total' Fare_Conditions,
COUNT (tf.fare_conditions)
FROM dst_project.ticket_flights tf
JOIN dst_project.flights f ON f.flight_id = tf.flight_id
WHERE f.aircraft_code = 'SU9'

```

QUERY RESULTS:

fare_conditions ▾	Количество ▾
Economy	2,568,738
Business	362,734
Standart	0
Comfort	0
Total	2,931,472

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

```

SELECT t.book_ref,
COUNT(DISTINCT tf.ticket_no),
SUM(tf.amount)
FROM dst_project.tickets t

```



```
JOIN dst_project.ticket_flights tf ON t.ticket_no = tf.ticket_no
GROUP BY t.book_ref
ORDER BY sum(tf.amount)
LIMIT 1
```

QUERY RESULTS:

book_ref ▾	Количество ▾	Сумма ▾
00F39E	1	3,400

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

```
SELECT bs.seat_no
FROM dst_project.tickets t
JOIN dst_project.boarding_passes bs ON t.ticket_no = bs.ticket_no
WHERE t.passenger_id = '4313 788533'
```

QUERY RESULTS:

2A

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
JOIN dst_project.airports a ON fl.arrival_airport = a.airport_code
WHERE a.city = 'Anapa'
AND fl.actual_arrival BETWEEN '2017-01-01 0:00:00' AND '2017-12-31 23:59:59'
```

QUERY RESULTS:

486

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

Зима - 1 января 2017 - 28 - февраля 2017 + 1 декабря 2017 - 31 декабря 2017.

```
SELECT count(fl.flight_id)
FROM dst_project.flights fl
JOIN dst_project.airports a ON fl.departure_airport= a.airport_code
WHERE a.city = 'Anapa'
      AND (fl.actual_departure BETWEEN '2017-01-01 0:00:00' AND '2017-02-28
23:59:59'
      OR fl.actual_departure BETWEEN '2017-12-01 0:00:00' AND '2017-12-31
23:59:59')
```

QUERY RESULTS:

Количество

127

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

Status = 'Cancelled'

```
SELECT count(fl.flight_id)
FROM dst_project.flights fl
JOIN dst_project.airports a ON fl.departure_airport= a.airport_code
WHERE a.city = 'Anapa'
      AND fl.status = 'Cancelled'
```

QUERY RESULTS:

Количество

1

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

```
SELECT count(flight_id) dep_air
FROM dst_project.flights fl
JOIN dst_project.airports a ON fl.departure_airport = a.airport_code
JOIN dst_project.airports a1 ON fl.arrival_airport = a1.airport_code
WHERE a.city = 'Anapa'
      AND a1.city != 'Moscow'
```

QUERY RESULTS: ▼ dep_air

453

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

```
SELECT DISTINCT ac.model,
                count(s.seat_no),
                fl.flight_id
FROM dst_project.seats s
JOIN dst_project.aircrafts ac ON ac.aircraft_code = s.aircraft_code
JOIN dst_project.flights fl ON fl.aircraft_code = s.aircraft_code
JOIN dst_project.airports ap ON fl.departure_airport = ap.airport_code
WHERE ap.city = 'Anapa'
GROUP BY ac.model,
         fl.flight_id
ORDER BY count(s.seat_no) DESC
LIMIT 1
```

QUERY RESULTS:

model ▼	▼ Количество	▼ flight_id
Boeing 737-300	130	136,305

Задача

Определить от каких самых **малоприбыльных рейсов** из Анапы можно отказаться в зимнее время.

прибыльность рейса = доход от продаж билетов - расход на полет

Простой способ определить прибыльность рейса: стоимость билетов - стоимость топлива на рейс.

Обязательно в дотает:

- 1) id рейса
- 2) Город вылета (Анапа)
- 3) Город прилета

Что входит в расходы:

- 1) Топливо
- 2) Затраты на ТО самолета
- 3) Затраты на зарплату сотрудников
- 4) Налоги и сборы
- 5) Затраты на владение авиапарком

Что входит в income:

- 1) Оплата за проезд (усреднённое по всем пассажирам)



```
SELECT fl.flight_id,  
       fl.flight_no,  
       fl.departure_airport,  
       fl.aircraft_code,  
       ac.model aircraft_model,  
       ac.range,
```

```

        a_arrival.city arrival_city,
        fl.arrival_airport,
        a_arrival.longitude arrival_longitude,
        a_arrival.latitude arrival_latitude,
        COUNT (DISTINCT s.seat_no) total_num_of_seats,
        COUNT (DISTINCT tf.ticket_no) total_num_of_tickets,
        SUM (tf.amount) income,
        count(DISTINCT tf.ticket_no)*100/count(DISTINCT
s.seat_no) filled_pers,
        fl.actual_departure,
        fl.actual_arrival,
        date_part('hour', (fl.actual_arrival -
fl.actual_departure))*60 + date_part('minute', (fl.actual_arrival -
fl.actual_departure)) duration
FROM dst_project.flights fl
JOIN dst_project.airports a_arrival ON fl.arrival_airport =
a_arrival.airport_code
JOIN dst_project.aircrafts ac ON fl.aircraft_code = ac.aircraft_code
JOIN dst_project.seats s ON fl.aircraft_code=s.aircraft_code
JOIN dst_project.ticket_flights tf ON tf.flight_id = fl.flight_id
WHERE departure_airport = 'AAQ'
        AND (date_trunc('month', scheduled_departure) in ('2017-01-01',
                                                             '2017-02-01',
                                                             '2017-12-01'))

        AND status not in ('Cancelled')
GROUP BY fl.flight_id,
        ac.model,
        a_arrival.city,
        a_arrival.longitude,
        a_arrival.latitude,
        ac.range

```

QUERY RESULTS:

<div>▼ flight_id</div>	<div>flight_no</div>	<div>departure_airport</div> ▼	<div>aircraft_code</div> ▼	<div>aircraft_model</div> ▼	<div>range</div>	<div>arrival_city</div> ▼	<div>arrival_airport</div> ▼	<div>arrival_longitude</div> ▼	<div>arrival_latitude</div> ▼	<div>total_num_of_seats</div> ▼
136,119	PG0252	AAQ	733	Boeing 737-300	4,200	Moscow	SVO	37.41	55.97	120
136,120	PG0252	AAQ	733	Boeing 737-300	4,200	Moscow	SVO	37.41	55.97	120
136,121	PG0252	AAQ	733	Boeing 737-300	4,200	Moscow	SVO	37.41	55.97	120

Датасет загружен в csv file.