--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 ap.city

HAVING count(ap.airport\_code) > 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 f.status)

FROM dst\_project.flights f

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

SELECT count(f.flight\_id)

FROM dst\_project.flights f

WHERE f.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

JOIN dst\_project.aircrafts a ON s.aircraft\_code = a.aircraft\_code

GROUP BY a.model

HAVING a.model = 'Boeing 777-300'

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

SELECT count(f.flight\_id)

FROM dst\_project.flights f

WHERE f.actual\_arrival BETWEEN '2017-04-01 00:00:00'

AND '2017-09-01 00:00:00'

AND f.status != 'Cancelled'

--Задание 4.3

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

SELECT count(f.flight\_id)

FROM dst\_project.flights f

WHERE f.status = 'Cancelled'

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

SELECT count(a.aircraft\_code)

FROM dst\_project.aircrafts a

WHERE a.model like 'Boeing%'

UNION ALL

SELECT count(a.aircraft\_code)

FROM dst\_project.aircrafts a

WHERE a.model like 'Sukhoi Superjet%'

UNION ALL

SELECT count(a.aircraft\_code)

FROM dst\_project.aircrafts a

WHERE a.model like 'Airbus%'

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

SELECT count(ai.airport\_code)

FROM dst\_project.airports ai

WHERE ai.timezone like 'Europe%'

UNION ALL

SELECT count(ai.airport\_code)

FROM dst\_project.airports ai

WHERE ai.timezone like 'Asia%'

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

SELECT f.flight\_id

FROM dst\_project.flights f

WHERE f.actual\_arrival IS NOT NULL

ORDER BY age(f.actual\_arrival, f.scheduled\_arrival) DESC

LIMIT 1

--Задание 4.4

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

SELECT f.scheduled\_departure

FROM dst\_project.flights f

ORDER BY f.scheduled\_departure ASC

LIMIT 1

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

WITH duration AS

(SELECT max(age(f.scheduled\_arrival, f.scheduled\_departure)) t

FROM dst\_project.flights f)

SELECT date\_part('hour', duration.t) \* 60 + date\_part('minute', duration.t)

FROM duration

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

WITH a AS

(SELECT distinct(f.departure\_airport, f.arrival\_airport) airp\_pair,

dense\_rank() OVER (

ORDER BY age(f.scheduled\_arrival, f.scheduled\_departure) DESC) rnk1

FROM dst\_project.flights f)

SELECT a.airp\_pair

FROM a

WHERE a.rnk1 = 1

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

SELECT avg(date\_part('hour', age(f.scheduled\_arrival, f.scheduled\_departure)) \* 60 + date\_part('minute', age(f.scheduled\_arrival, f.scheduled\_departure)))::int

FROM dst\_project.flights AS f

--Задание 4.5

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

SELECT count(s.seat\_no),

s.fare\_conditions

FROM dst\_project.seats s

WHERE s.aircraft\_code = 'SU9'

GROUP BY s.fare\_conditions

ORDER BY 1 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

JOIN dst\_project.tickets t USING(ticket\_no)

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(f.flight\_id)

FROM dst\_project.flights f

JOIN dst\_project.airports a ON f.arrival\_airport = a.airport\_code

WHERE a.city = 'Anapa'

AND date\_part('year', f.actual\_arrival) = 2017

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

SELECT count(f.flight\_id)

FROM dst\_project.flights f

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

WHERE a.city = 'Anapa'

AND date\_part('year', f.actual\_departure) = 2017

AND date\_part('month', f.actual\_departure) in (1,

2,

12)

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

SELECT count(f.flight\_id)

FROM dst\_project.flights f

LEFT JOIN dst\_project.airports a ON f.departure\_airport = a.airport\_code

WHERE a.city = 'Anapa'

AND f.status = 'Cancelled'

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

SELECT count(f.flight\_id)

FROM dst\_project.flights f

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

JOIN dst\_project.airports a1 ON f.arrival\_airport = a1.airport\_code

WHERE a.city = 'Anapa'

AND a1.city != 'Moscow'

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

SELECT DISTINCT f.aircraft\_code,

ac.model,

w.s\_a

FROM dst\_project.flights f

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

JOIN dst\_project.aircrafts ac ON f.aircraft\_code = ac.aircraft\_code

JOIN

(SELECT s.aircraft\_code code,

count(s.seat\_no) s\_a

FROM dst\_project.seats s

GROUP BY s.aircraft\_code) w ON f.aircraft\_code = w.code

WHERE a.city = 'Anapa'

ORDER BY w.s\_a DESC

LIMIT 1

=========================================================

Проект.

Получение датасета:

SELECT f.flight\_id,

f.flight\_no,

f.departure\_airport,

a.airport\_name departure\_airport\_name,

a.city departure\_city,

a.latitude departure\_latitude,

a.longitude departure\_longitude,

f.arrival\_airport,

a1.airport\_name arrival\_airport\_name,

a1.city arrival\_city,

a1.latitude arrival\_latitude,

a1.longitude arrival\_longitude,

f.scheduled\_departure,

f.scheduled\_arrival,

f.actual\_departure,

f.actual\_arrival,

date\_part('dow', f.scheduled\_departure) dow,

date\_part('hour', age(f.scheduled\_arrival, f.scheduled\_departure)) \* 60 + date\_part('minute', age(f.scheduled\_arrival, f.scheduled\_departure)) scheduled\_len,

date\_part('hour', age(f.actual\_arrival, f.actual\_departure)) \* 60 + date\_part('minute', age(f.actual\_arrival, f.actual\_departure)) actual\_len,

date\_part('hour', age(f.actual\_departure, f.scheduled\_departure)) \* 60 + date\_part('minute', age(f.actual\_departure, f.scheduled\_departure)) departure\_delay,

date\_part('hour', age(f.actual\_arrival, f.scheduled\_arrival)) \* 60 + date\_part('minute', age(f.actual\_arrival, f.scheduled\_arrival)) arrival\_delay,

f.aircraft\_code,

ac.model,

ac.range,

w.seats\_count economy\_seats,

w1.seats\_count business\_seats,

z.seats\_sold economy\_seats\_sold,

z.revenue economy\_revenue,

z1.seats\_sold business\_seats\_sold,

z1.revenue business\_revenue

FROM dst\_project.flights f

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

JOIN dst\_project.airports a1 ON f.arrival\_airport = a1.airport\_code

JOIN dst\_project.aircrafts ac ON f.aircraft\_code = ac.aircraft\_code

JOIN (

SELECT s.aircraft\_code code,

count(s.seat\_no) seats\_count

FROM dst\_project.seats s

GROUP BY s.aircraft\_code, s.fare\_conditions

HAVING s.fare\_conditions = 'Economy') w ON f.aircraft\_code = w.code

JOIN (

SELECT s.aircraft\_code code,

count(s.seat\_no) seats\_count

FROM dst\_project.seats s

GROUP BY s.aircraft\_code, s.fare\_conditions

HAVING s.fare\_conditions = 'Business') w1 ON f.aircraft\_code = w1.code

LEFT JOIN (

SELECT count(tf.ticket\_no) seats\_sold,

sum(tf.amount) revenue,

tf.flight\_id

FROM dst\_project.ticket\_flights tf

GROUP BY tf.flight\_id, tf.fare\_conditions

HAVING tf.fare\_conditions = 'Economy') z ON f.flight\_id = z.flight\_id

LEFT JOIN (

SELECT count(tf.ticket\_no) seats\_sold,

sum(tf.amount) revenue,

tf.flight\_id

FROM dst\_project.ticket\_flights tf

GROUP BY tf.flight\_id, tf.fare\_conditions

HAVING tf.fare\_conditions = 'Business') z1 ON f.flight\_id = z1.flight\_id

WHERE f.departure\_airport = 'AAQ'

AND (date\_trunc('month', f.scheduled\_departure) IN ('2017-01-01','2017-02-01', '2017-12-01'))

AND f.status NOT IN ('Cancelled')