-- select\*

-- from skyeng\_db.classes

-- select\*

-- from skyeng\_db.payments

-- select\*

-- from skyeng\_db.students

-- select\*

-- from skyeng\_db.teachers

---------------------------------КУРСОВАЯ РАБОТА-----------------------------------

---------ШАГ 1--------------------------

with first\_payment\_date as

(

select

user\_id

, date\_trunc('day', min(transaction\_datetime)) as first\_payment\_date

-- , min(transaction\_datetime) as first\_payment\_date

from skyeng\_db.payments

where status\_name = 'success'

group by user\_id

order by first\_payment\_date

),

-- select \* from first\_payment\_date

---------ШАГ 2---------------------------

all\_dates as

(

select distinct

date\_trunc( 'day', class\_start\_datetime) as dt

from skyeng\_db.classes

where (class\_start\_datetime between '2016-01-01' and'2017-01-01')

),

-- select \* from all\_dates

--------ШАГ 3----------------------------

all\_dates\_by\_user as

(

select

fp.user\_id

, ad.dt

from all\_dates as ad

left join first\_payment\_date as fp

on fp.first\_payment\_date <= ad.dt

order by user\_id, dt

),

-- select \* from all\_dates\_by\_user

--------ШАГ 4-----------------------------

payments\_by\_dates as

(

select

user\_id

,date\_trunc('day', transaction\_datetime) as payment\_date

,sum (classes) as transaction\_balance\_change

from skyeng\_db.payments

where status\_name = 'success'

group by user\_id, payment\_date

order by user\_id, payment\_date

),

-- select \* from payments\_by\_dates

---------ШАГ 5------------------------------

payments\_by\_dates\_cumsum as

(

select

d\_us.user\_id

, d\_us.dt

, p\_d.transaction\_balance\_change

, case when sum(p\_d.transaction\_balance\_change) over (partition by d\_us.user\_id order by d\_us.dt) is null then 0 else

sum(p\_d.transaction\_balance\_change) over (partition by d\_us.user\_id order by d\_us.dt)

end as transaction\_balance\_change\_cs

from all\_dates\_by\_user as d\_us

left join payments\_by\_dates as p\_d

on d\_us.dt = p\_d.payment\_date

and d\_us.user\_id = p\_d.user\_id

order by

d\_us.user\_id,

d\_us.dt

),

-- select \* from payments\_by\_dates\_cumsum

-----------ШАГ 6-------------------------------

classes\_by\_dates as

(

select

user\_id

, date\_trunc('day', class\_start\_datetime) as class\_date

,count(class\_status) \* -1 as classes

from skyeng\_db.classes

where class\_status in ('success', 'failed\_by\_student')

and class\_type <> 'trial'

group by user\_id

, class\_date

),

-- select\* from classes\_by\_dates

------------ШАГ 7-------------------------------

classes\_by\_dates\_dates\_cumsum as

(

select

d\_u.user\_id

,d\_u.dt

,c\_d.classes

,case when sum(c\_d.classes) over (partition by d\_u.user\_id order by d\_u.dt) is null then 0 else

sum(c\_d.classes) over (partition by d\_u.user\_id order by d\_u.dt)

end as classes\_cs

from all\_dates\_by\_user as d\_u

left join classes\_by\_dates as c\_d

on d\_u.user\_id= c\_d.user\_id

and d\_u.dt=c\_d.class\_date

),

-- select \* from classes\_by\_dates\_dates\_cumsum

-------------ШАГ 8--------------------------------

balances as

(

select

pds.user\_id

,pds.dt

,pds.transaction\_balance\_change

,pds.transaction\_balance\_change\_cs

,cds.classes

,cds.classes\_cs

,cds.classes\_cs + pds.transaction\_balance\_change\_cs as balance

from payments\_by\_dates\_cumsum as pds

left join classes\_by\_dates\_dates\_cumsum as cds

on pds.user\_id=cds.user\_id

and pds.dt=cds.dt

group by pds.user\_id

,pds.dt

,pds.transaction\_balance\_change

,pds.transaction\_balance\_change\_cs

,cds.classes

,cds.classes\_cs

,balance

)

-- select \* from balances

--------------------------------------ЗАДАНИЕ 1--------------------------------------------------

-- -- Выберите топ-1000 строк из CTE `balances` с сортировкой по `user\_id` и `dt`.

-- -- Посмотрите на изменения балансов студентов.

-- -- Какие вопросы стоит задавать дата-инженерам и владельцам таблицы `payments`?

-- select \*

-- from balances

-- order by user\_id, dt

-- limit 1000

---------------- Вопросы к данной таблице-------------------

-- Просмотрев и проанализировав таблицу с лимитом = 1000 мне хотелось бы получить информацию по правильности

-- начисления и дальнейшему списания уроков по каждому из представленных user\_id.

-- Нет ли сбоев в загрузке данных в таблицу, своевременно ли это выполняется.

-- Так как наблюдаетсячастое списание уроков, которые не начислены данному user\_id.

-- Например user\_id 1984910 начиная с 02.06.20169г. регулярно уходит в -4 урока, а начисление уроков стоит гораздо позже.

-- User\_id 3335256 начиная с 09.11.2016г ушел в -4 урока и на конец 2016 года баланс у него так и остался отрицательным.

-- Может данные подгружаются в таблицу не корректно и с большой задержкой?

---------------------------------------------------------------------------

-- Для решения шага № 9, я проставлю коментарии к запросу по заданию 1

--------------- ШАГ 9-------------------------

select

dt

, sum(transaction\_balance\_change) as sum\_transaction\_balance\_change

, sum(transaction\_balance\_change\_cs) as sum\_transaction\_balance\_change\_cs

, sum(classes) as sum\_classes

, sum(classes\_cs) as sum\_classes\_cs

, sum(balance) as sum\_balance

from balances

group by dt

order by dt

------------------------------------------ЗАДАНИЕ 2---------------------------------------------------

-- -- Создайте визуализацию (линейную диаграмму) итогового результата.

-- -- Какие выводы можно сделать из получившейся визуализации?

-- Вывод:

-- Изучив линейный график по нашей таблице можно увидеть, что при нулевых показателях на начало года

-- к концу года суммовые показатели баланса студентов только по транзакциям (sum\_transaction\_balance\_change\_cs) уже равнялись 21650,

-- суммовые показатели пройденных уроков (sum\_classes\_cs) составляли 17257.

-- Списание уроков в начале года проходило более плавно, но с мая 2016 года наблюдается более резкие скачки по количеству пройденных уроков,

-- а во второй половине года прохождение уроков по дням имеет большой разрыв по количеству списанных уроков. Отсюда можно сделать вывод о неравномерном прохожнении уроков.

-- Линия кумулятивной сумме по транзакциям (sum\_transaction\_balance\_change\_cs) имеет положительную динамику

-- И при этом линия баланса на каждого студента за каждый день (sum\_balance) не имеет резких изменений. Она плавно повышается.