Задание 1.

Написать SQL-запрос для расчета top10 фильмов по количеству уникальных пользователей в разбивке по месяцам. Посчитать динамику в процентах от месяца к месяцу для фильмов из top10. Для фильмов, которые пришли в top10 первый раз считать, что в предыдущем месяце у них не было показов.  
  
**Пример результата**

Текущий месяц; позиция в top10;название; количество уникальных пользователей; разница с прошлым месяцем

2017-05-01;1;Маша и медведь: Первая встреча;10000;-10.1%

2017-05-01;2;Маша и медведь: До весны не будить;5000;1.3%

**Данные**

сontent - фильмы (контент)  
content\_id integer,  
title varchar  
  
сontent\_watch - показы контента  
watch\_id varchar - уникальный идентификатор показа  
content\_id integer- id контента  
show\_date timestamp - дата и время показа  
user\_id integer - уникальный идентификатор пользователя  
  
сalendar - справочник дат с 1900 по 2100 год по дням  
date timestamp

Задание 2.

Нужно придумать показатель, который характеризует контент (фильм или сериал) с точки зрения привлекательности пользователю, на сколько он нравится ему, как часто он возвращается к просмотру. Сделать запрос который за определенный промежуток времени вычислит этот показатель.   
Задача на размышление, можно рассмотреть разные варианты, их достоинства и недостатки. Если в таблицах, описанной ниже не хватает чего-то или надо добавить новые сущности, то следует изложить свои предложения о загрузке новых данных.

**Данные**

content\_watch - просмотр контента пользователелм

show\_date timestamp - время, когда начался показ  
content\_id int - идентификатор контента (фильма)  
watch\_id varchar(255) - уникальный идентификатор просмотра  
user\_id varchar(255) - уникальный идентификатор пользователя  
show\_duration int - длительность смотрения контента в секундах (на какой секунде контента от начала пользователь закончил просмотр)

content - справочник с контентом

id int - идентификатор контента  
compilation\_title varchar(255) - если контент-сериал, то название сериала (сборника)  
title varchar(255) - название контента  
duration - длительност контента в секундах

Задание 3.

Перед вами стоит задача подготовить рекомендации отделу закупок контента по кинотеатральным релизам за последние 3 месяца. Для этого вы договорились с заказчиком проанализировать данные фонда кино на <https://ekinobilet.fond-kino.ru/>

Необходимо спарсить данные с сайта, подготовить их, провести EDA, визуализировать полученные выводы.

Заказчик ожидает увидеть список топовых кинотеатральных релизов и любые дополнительные выводы, которые вы получите в ходе работы с данными.

В качестве решения мы ожидаем готовый к демонстрации jupyter notebook, который мы будем обсуждать на собеседовании.