# Структура данных

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Таблица просмотров контента – content\_watch. Один просмотр контента - одна запись** | | | | | | | | | |
| Идентификатор просмотра, ключ. | Дата просмотра | | Длительность  в секундах | Идентификатор площадки, на которой был просмотр | | Юзер | Источник трафика | | Идентификатор контента. |
| **watch\_id** | **show\_date** | | **show\_duration** | **platform** | | **user\_id** | **utm\_medium** | | **content\_id** |
| 10971121570 | 07.01.2018 14:37 | | 1340 | 583 | | 1553139 | organic | | 314472 |
| 4458319751 | 12.01.2018 15:00 | | 12432 | 353 | | 1554866 | organic | | 314472 |
| 31382550 | 08.02.2018 14:39 | | 1800 | 10 | | 5255577 | organic | | 314472 |
| 11254336994 | 07.07.2017 17:56 | | 210 | 11 | | 1554866 | organic | | 314472 |
| 1231646730 | 01.01.2016 12:48 | | 4685 | 11 | | 1554866 | organic | | 132271 |
| 4212172051 | 12.08.2018 10:52 | | 472 | 11 | | 1554866 | organic | | 314480 |
| 8909218338 | 09.05.2017 0:55 | | 297 | 583 | | 9462609 | direct | | 127399 |
| 1904761857 | 24.09.2018 19:31 | | 1635 | 9 | | 320756 | Organic | | 127399 |
| 17947987 | 30.10.2018 4:45 | | 854 | 353 | | 1547421 | Referral | | 184673 |
| 6077839073 | 07.12.2017 23:58 | | 4571 | 353 | | 4066590 | Organic | | 222161 |
| **Таблица контента – content** | | | | | | | | | | |
| Идентификатор контента, ключ | | Идентификатор сериала. NULL - контент единичный, то есть не является серией сериала | | | Номер эпизода в сериале | | | Монетизация контента | | |
| **content\_id** | | **compilation\_id** | | | **episode** | | | **paid\_type** | | |
| 314472 | | 9570 | | | 1 | | | AVOD | | |
| 132271 | | NULL | | | NULL | | | SVOD | | |
| 314480 | | 9570 | | | 2 | | | AVOD | | |
| 127399 | | 9570 | | | 6 | | | TVOD | | |
| 184673 | | 7608 | | | 16 | | | AVOD | | |
| 222161 | | NULL | | | NULL | | | AVOD | | |

## Задание 1

**Составьте SQL-запросы. Укажите, какой диалект SQL вы используете**

1. **На каждый день количество просмотров отдельно по монетизациям SVOD и AVOD на платформах 10 и 11 за последние 30 дней.**
2. **Ежемесячный ТОП-5 сериалов и ТОП-5 единичного контента по количеству смотрящих людей.**
3. **Список пользователей, у которых вчера был сначала просмотр с organic, а сразу следом за ним - просмотр с referral**

## Задание 2

Составьте SQL-запросы. Придумайте, как оценить показатели, и напишите запросы для расчёта придуманных метрик. Представьте, что в вашем распоряжении есть все ресурсы по сбору статистики. Если вам необходимы дополнительный данные, то опишите, чего не хватает в тестовой базе, чтобы посчитать нужные метрики.

1. **"Цепляемость" и "крутость" сериала. Нужна какая-то метрика, которая при наличии трёх-четырёх серий сериала позволит сравнить этот сериал по "крутости" с другими сериалами.**
2. **Ретеншн всех пользователей сервиса. Нужно просегментировать аудиторию, рассчитать её ретеншн по сегментам и дать рекомендации по тому, как увеличить ретеншн каждого из сегментов.**

## Задание 3

Есть А/Б тест направленный на конверсию в покупку (подписки и/или платных фильмов). % что либо купивших в обоих группах одинаковый. Средний чек разный. Каким способом можно проверить, что разница не случайна (с использованием sql и/или python).