

# Анализ русских трендов YouTube

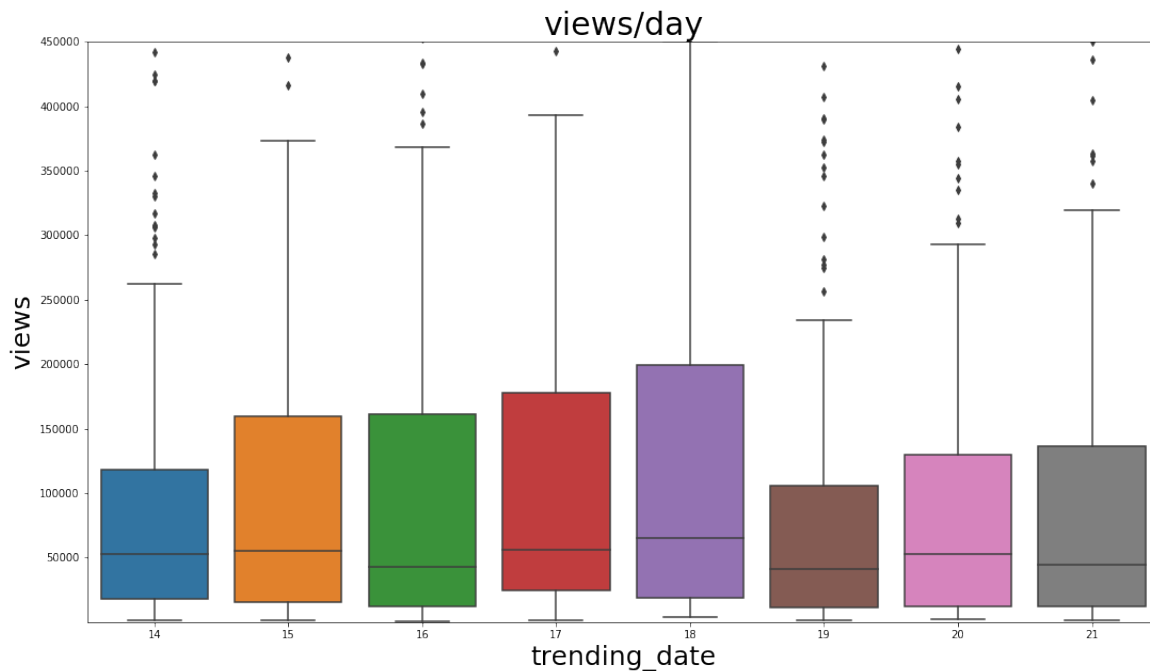
Полев Алексей, Б05-126

## 2. YouTube

### Задача:

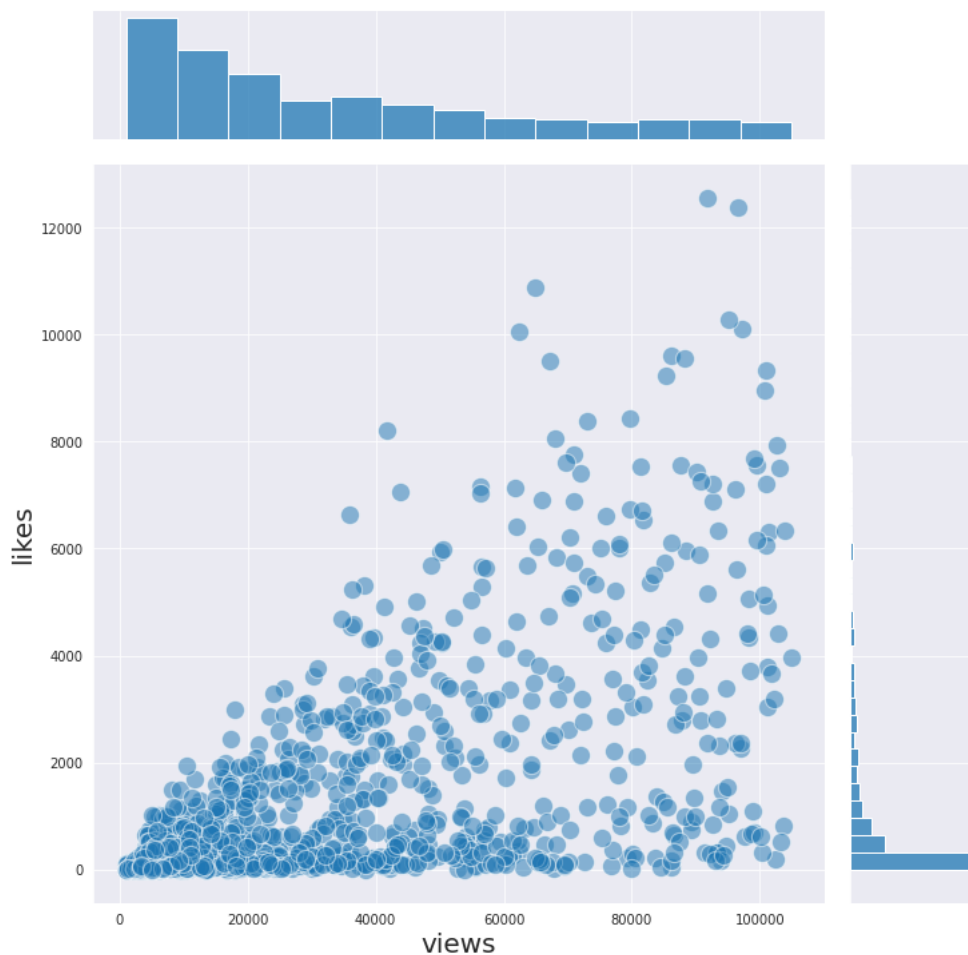
*Найти зависимость между просмотрами, которые получают русскоязычные видео на YouTube и датой, когда они завирусились. Понять как зависит число лайков от числа просмотров на видео.*

Рассмотрим boxplot просмотров видео в день:



Замечаем, что просмотры в день растут по всем параметрам тем сильнее, чем ближе суббота: 18 ноября 2017. В воскресенье заметно резкое падение среднего числа просмотров, что может быть связано с началом рабочей недели. Медианное число просмотров в день - около 50000.

Далее изучим зависимость числа лайков у видео от числа просмотров с помощью joint-plot:



Заметим, что график практически симметричен относительно неглавной диагонали несмотря на то, что единичные отметки у осей отличаются в 50 раз. Значит, что на каждый лайк в среднем приходится около 50 просмотров. Также видно, что большинство видео не набирает больше 250000 просмотров или 600 лайков.

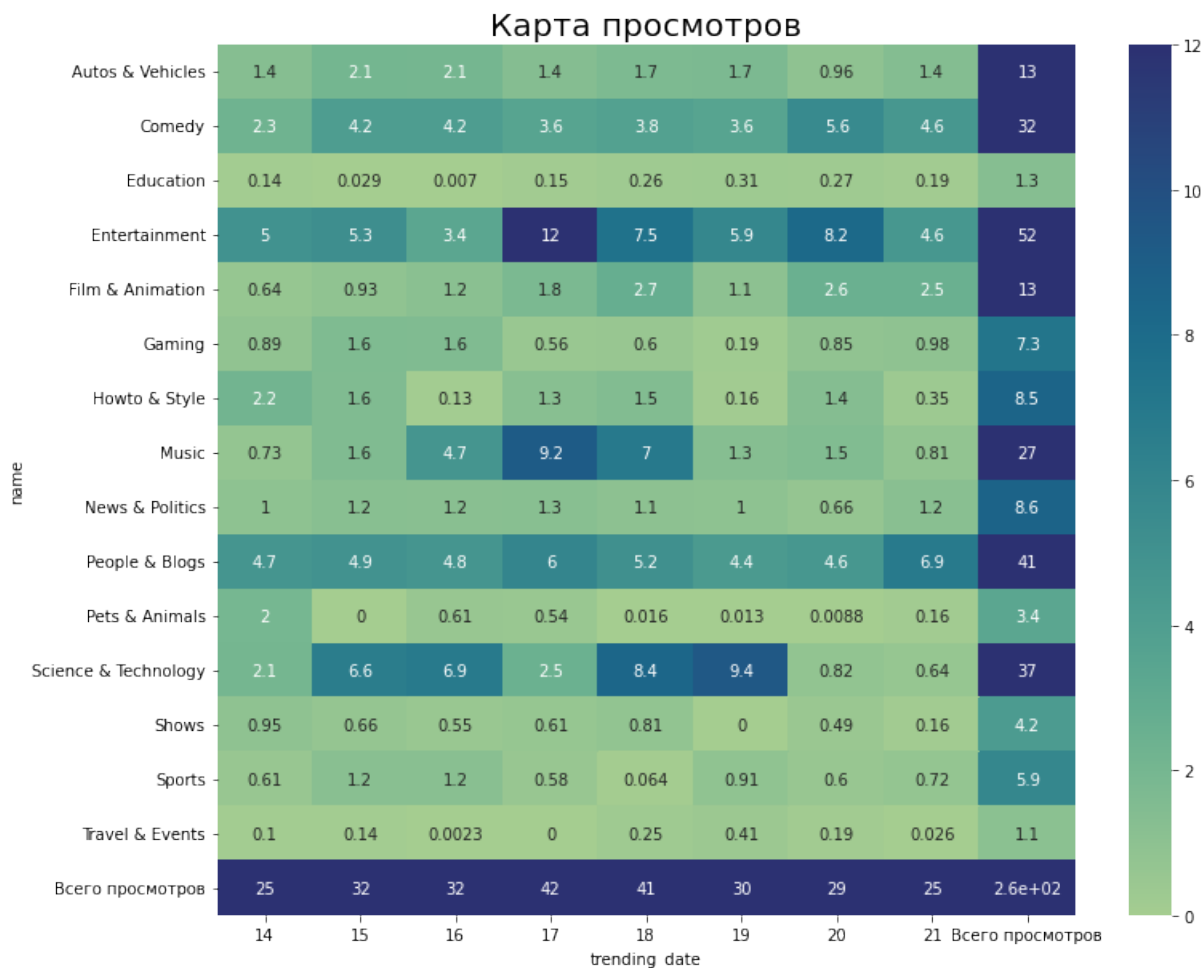
### 3. YouTube 2

**Задача:**

*Для ярких категорий найти периоды, в которые им наиболее вероятно стать трендовыми.*

Чтобы разобраться с трендовостью категорий, изобразим heatmap со следующими значениями в ячейках: суммарное количество просмотров в день у видео из этой категории.

Так как ячейки с суммой будут иметь значения на порядок больше, чем в обычных, только они окрасятся в самый синий цвет и потеряется статистика у обычных ячеек, соответственно ограничиваем окрашиваемость ячейки максимальным значением 12.



Тут снова можем заметить возрастающие просмотры во всех категориях при приближении даты к 17 – 18му числу и падение сразу после него (ряд "Всего просмотров").

Также в ряду "Всего просмотров" сильно выделяется общее число за 8 дней у категории "Развлечения" -  $52 \times 10^6$ , после него "Блоги" -  $41 \times 10^6$ .

Самые же не востребоваанные категории видео в России в этом периоде: "Путешествия" и "Образование".

Пиковая активность за день зафиксирована у категории "Развлечения" 17 ноября.

**Вывод:**

*Мы поняли, что на русском YouTube набирать лайки гораздо сложнее, чем просмотры, однако оба эти параметра зависят от дня недели трендовости видео (Самый загруженный - пятница, свободный - воскресенье). Россияне предпочитают развлечения и блоги путешествиям и образованию.*