**PROJECT-5 Рябов Д.В. DARP-1006**

[PA-2. Поведенческая аналитика (проведение ad-hoc исследования)](https://apps.skillfactory.ru/learning/course/course-v1:SkillFactory+DAPR-2.0+MAR2021/block-v1:SkillFactory+DAPR-2.0+MAR2021+type@chapter+block@e958547f9fb04412bdce2a42c294be5c)

**ОПИСАНИЕ КЕЙСА**

**Вы работаете продуктовым аналитиком в компании-сервисе по доставке продуктов на дом. Сервис доступен как в приложении на *iOS*, так и на *Android*. Вы настроили фронтовую аналитику в *AppMetrica*, и в конце квартала маркетинг-менеджер попросил вас проанализировать поведение пользователей, а также оценить эффективность каналов их привлечения.**

**Вы выгрузили данные из *AppMetrica* за период с 1 января по 31 марта 2020 — только по пользователям, зарегистрированным позднее 1 января 2020.**

**ОПИСАНИЕ ДАННЫХ**

**→ date — дата совершения события**

**→ event — событие**

* **app\_install — установка приложения**
* **app\_start — открыть приложения**
* **registration — зарегистрироваться**
* **search — перейти на страницу поиска товаров (каталог)**
* **open\_item — открыть товар**
* **choose\_item — отправить товар в корзину**
* **tap\_basket — перейти в корзину**
* **purchase — подтверждение покупки**

**→ gender — пол пользователя**

**→ os\_name — платформа пользователя**

**→ city — город пользователя**

**→ device\_id — идентификатор устройства пользователя**

**→ utm\_source — канал, с которого пришел пользователь**

* **yandex-direct — Яндекс.Директ**
* **google\_ads — реклама в *Google***
* **vk\_ads — реклама в ВК**
* **instagram\_ads — реклама в Инстаграм\***
* **facebook\_ads — реклама в Фейсбук\***
* **referral — акция «приведи друга»**
* **Если стоит - (прочерк), то канал не определен или это скачивание приложения напрямую или посещение не с рекламы**

**→ purchase\_sum — стоимость покупки (при совершении события ‘purchase’)**

**Обратите внимание!**

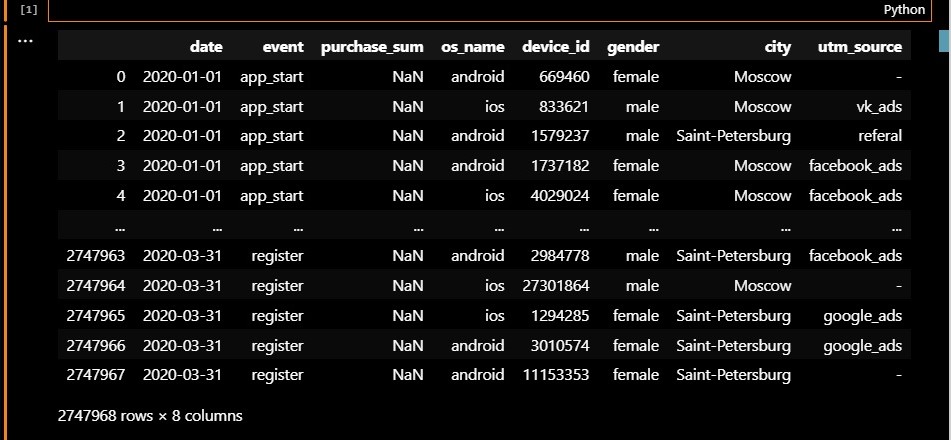
* + **В выгрузке только уникальные действия пользователей за каждый день.**
  + **Можно миновать стадию установки приложения, если оно было установлено ранее.**
  + **Можно миновать стадию регистрации, если пользователь был уже залогинен на момент сессии. Однако незарегистрированные пользователи не могут оформить покупку.**

**import pandas as pd**

**df = pd.read\_csv('SF\_case\_data.csv')**

**df = pd.DataFrame(df) #создаем дата фрейм**

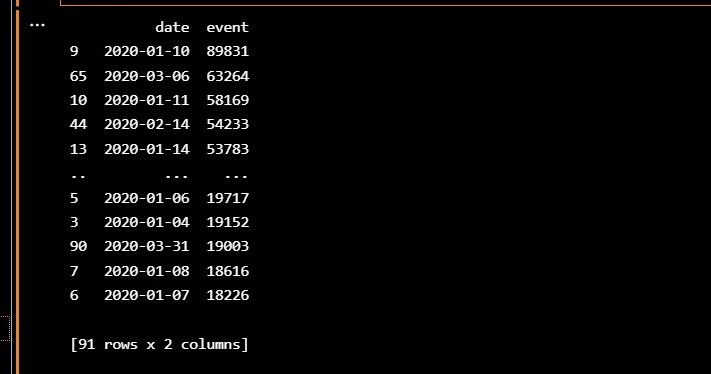
**display(df) #проверяем данные**

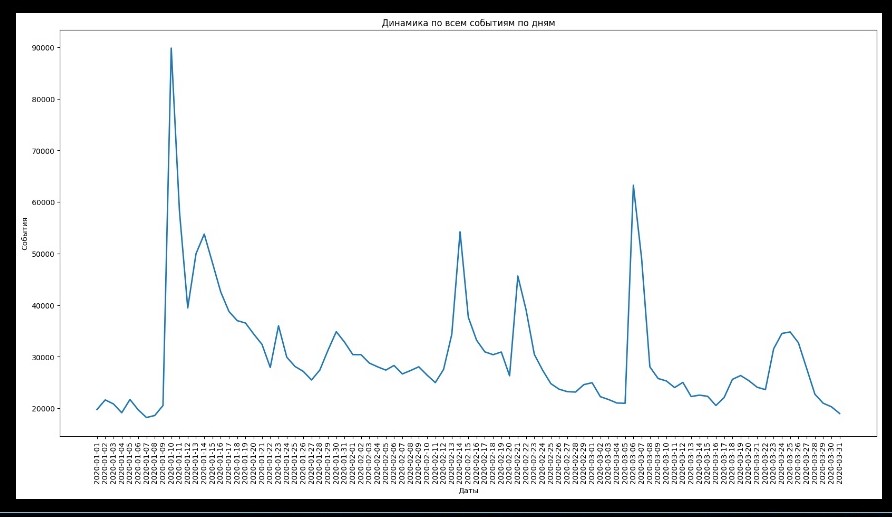
****

**1. Эффективность каналов привлечения и воронка продаж**

### Задание 1.1.1

**Постройте график с динамикой по всем событиям по дням. В какой день было больше всех событий? Ответ впишите в формате: ДД.ММ.ГГГГ.**

****

****

**10.01.2020 было больше всего событий**

### 

### Задание 1.1.2

**Посмотрите на то, как распределен трафик. Постройте график с динамикой открытий приложений для групп «Установили в тот же день» и «Остальные».**

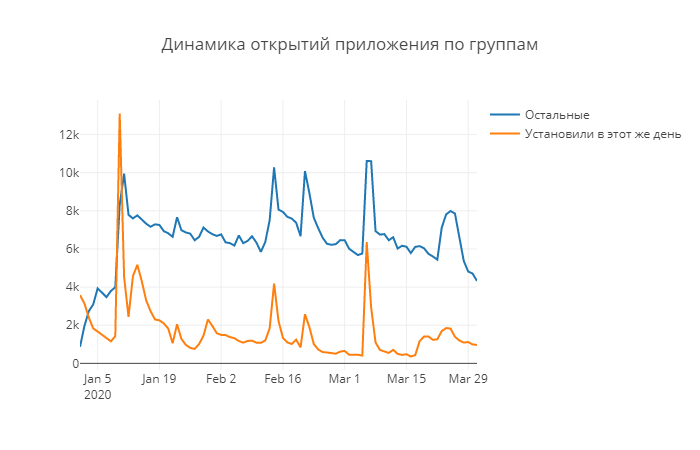
**Вопрос 1. Сколько людей установили приложение 31 марта?**

**968 человек установили приложение 31 марта.**

****

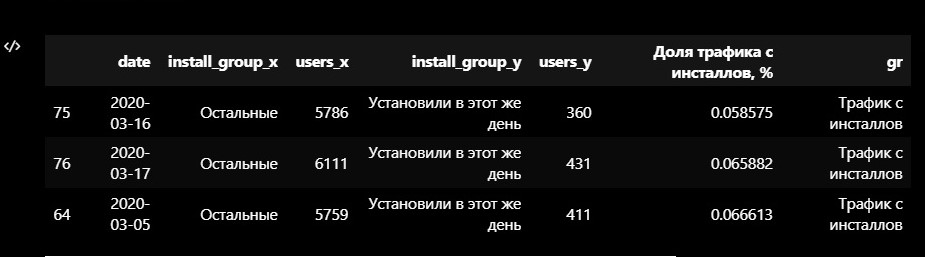
**Вопрос 2. Какая доля открытий приложения 14 февраля пришлась на пользователей, у которых приложение уже было установлено (ответ впишите целым числом, в процентном виде, без знака процента).**

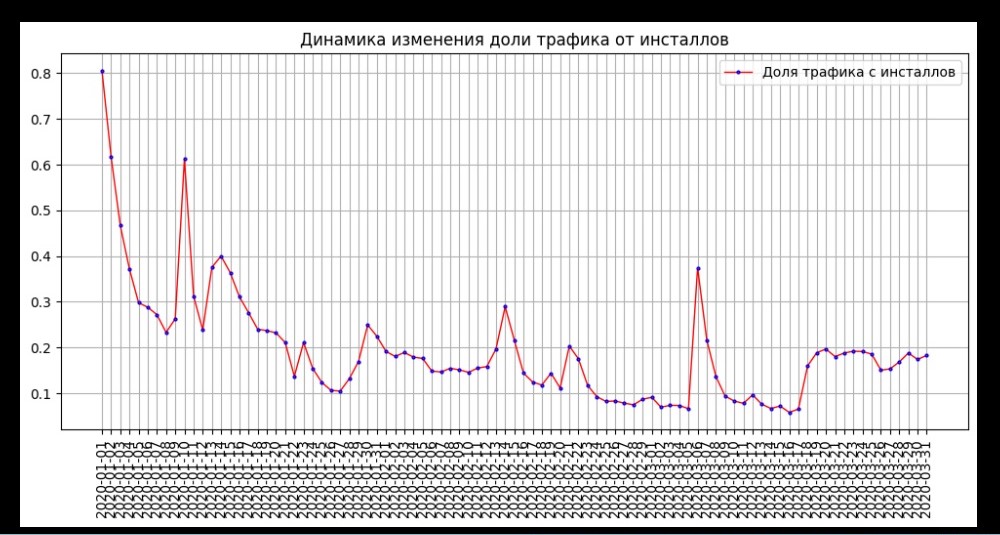
**71% доля открытий приложения 14 февраля пришлась на пользователей, у которых приложение уже было установлено.**

****

### Задание 1.1.3

**В какой день доля трафика от инсталлов была наибольшей? Ответ впишите в формате: ДД.ММ.ГГГГ.**

****

****

**День с максимальной долей трафика от инсталлов — 1 января 2020г. , хотя в абсолютном выражении количество инсталлов в этот день не максимальное (3579 установок).**

**День с минимальной долей трафика от инсталлов — 16 марта (360 установок).**

### Задание 1.1.4

**С какого канала в Q1 пришло больше всего пользователей? Впишите в поля ниже один ответ из возможных вариантов: *Яндекс, Гугл, ВК, Фейсбук, Инстаграм, Реферальная программа*.**

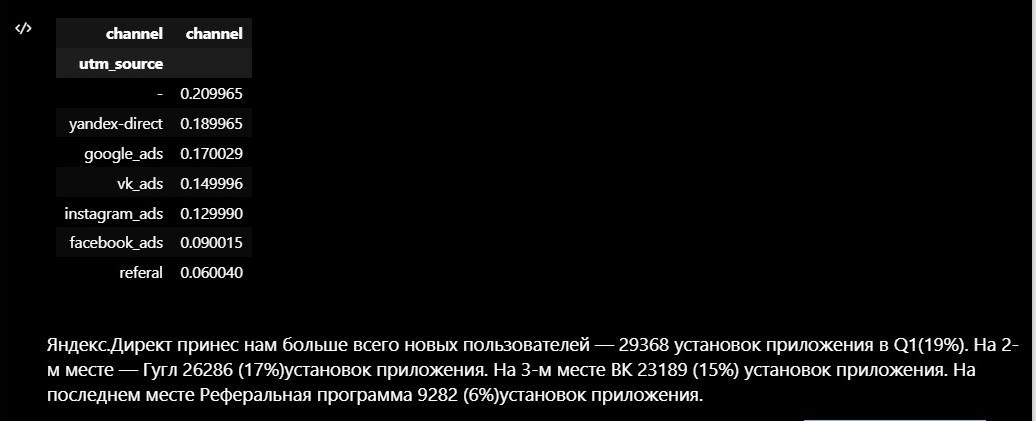
**Вопрос 1. С какого канала было больше всего инсталлов?**

**Вопрос 2. С какого канала было больше всего первых открытий приложения?**

****

****

****

****

**Топ-3 канала, по которым мы получили больше всего инсталлов в Q1: Яндекс-директ, Гугл, ВК.**

**По первым посещениям топ-3 канала в Q1 такие же, как и для инсталлов: Яндекс.Директ, Гугл, ВК.**

**Яндекс.Директ принес нам больше всего новых пользователей — 29368 установок приложения в Q1(19%). На 2-м месте — Гугл 26286 (17%)установок приложения. На 3-м месте ВК 23189 (15%) установок приложения. На последнем месте Реферальная программа 9282 (6%)установок приложения.**

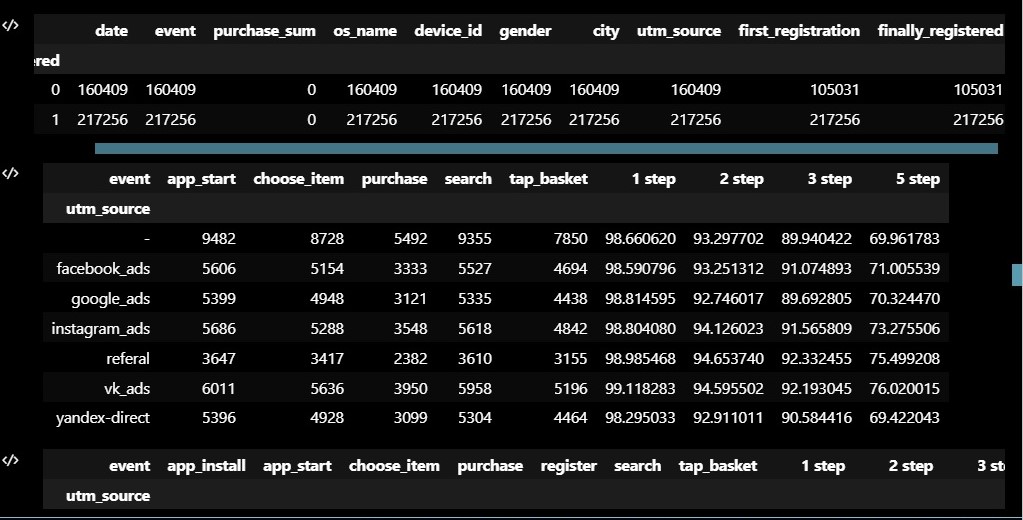
### 

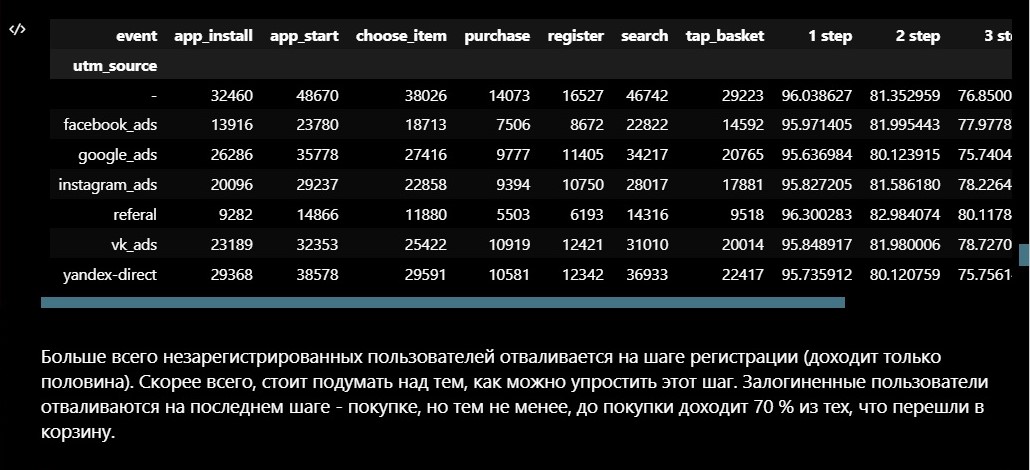
### Задание 1.1.5

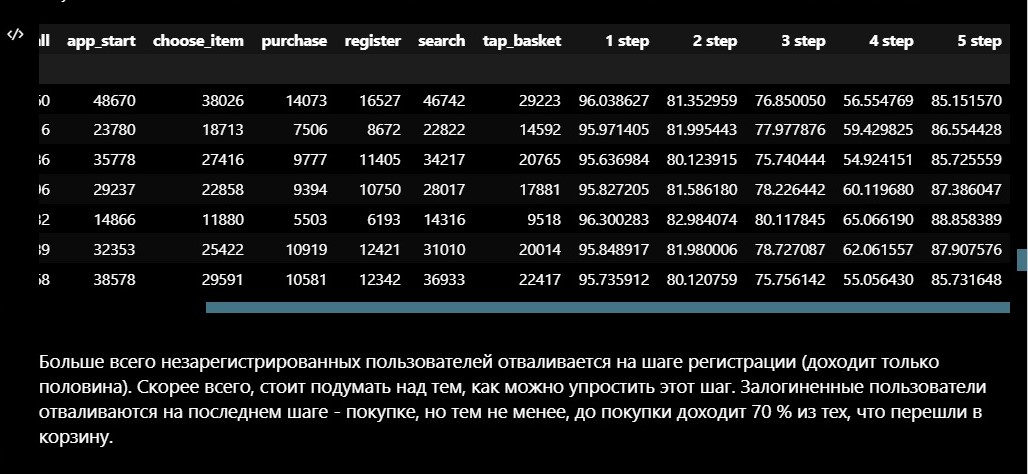
**Проанализируйте, на каком этапе воронки отваливается б'ольшая часть клиентов. Посмотрите отдельно сценарии для зарегистрированных и для незарегистрированных пользователей. Подумайте, можно ли что-то оптимизировать. Впишите в поля ниже один из шагов, конверсия в который (из предыдущего шага) самая низкая. Возможные варианты ответа: *Поиск, Добавление товара, Переход в корзину, Регистрация, Покупка*.**

**Вопрос 1. На каком шаге отваливается больше всего зарегистрированных пользователей?**

**Вопрос 2. На каком шаге отваливается больше всего незарегистрированных пользователей?**

****

****

****

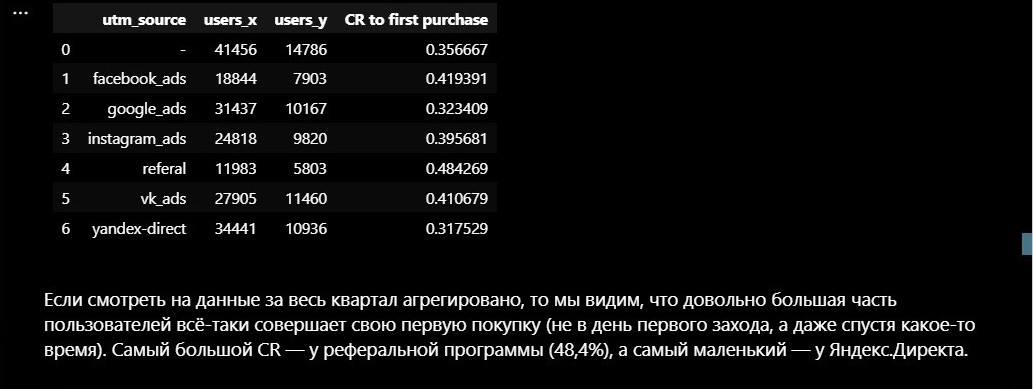
**Больше всего незарегистрированных пользователей отваливается на шаге регистрации (доходит только половина). Скорее всего, стоит подумать над тем, как можно упростить этот шаг. Залогиненные пользователи отваливаются на последнем шаге, но тем не менее, до покупки доходит 70 % из тех, что перешли в корзину.**

### Задание 1.1.6

**Впишите в поля ниже один ответ из возможных вариантов: *Яндекс, Гугл, ВК, Фейсбук, Инстаграм, Реферальная программа*.**

**Вопрос 1. Пользователи, пришедшие с каких каналов, показали (за весь Q1) самую высокую конверсию в первую покупку?**

**Вопрос 2. Пользователи, пришедшие с каких каналов, показали (за весь Q1) самую низкую конверсию в первую покупку?**

****

**Пользователи, пришедшие с канала Реферальной программы, показали (за весь Q1) самую высокую конверсию в первую покупку CR=48.43%.**

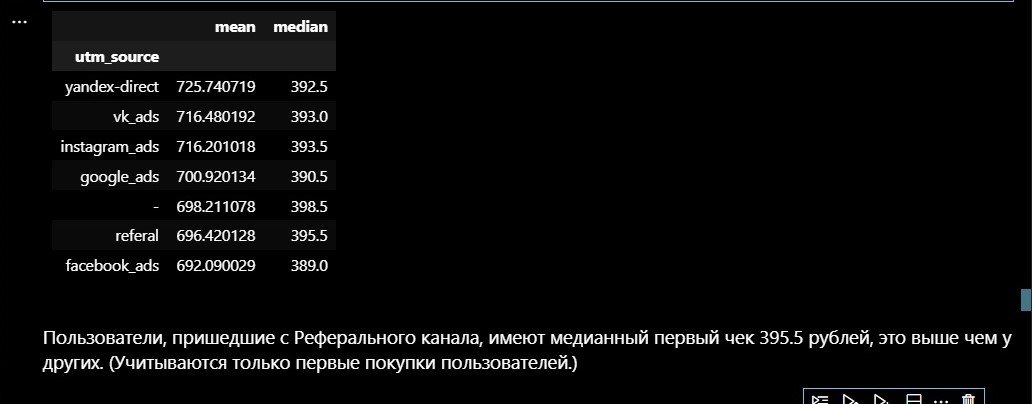
**Пользователи, пришедшие с канала Яндекс, показали (за весь Q1) самую низкую конверсию в первую покупку CR=31.75%.**

**Если смотреть на данные за весь квартал агрегировано, то мы видим, что довольно большая часть пользователей всё-таки совершает свою первую покупку (не в день первого захода, а даже спустя какое-то время). Самый большой CR — у реферальной программы (48,4%), а самый маленький — у Яндекс.Директа(31.75).**

### Задание 1.1.7

**Впишите в поле ниже один ответ из возможных вариантов: *Яндекс, Гугл, ВК, Фейсбук, Инстаграм, Реферальная программа*.**

**Вопрос. Пользователи, пришедшие с какого канала, имеют медианный первый чек выше? (Учитываются только первые покупки пользователей.)**

****

**Пользователи, пришедшие с канала Реферальной программы, имеют медианный первый чек 395.5 рублей, это выше, чем у других. Пользователи, пришедшие с канала Фэйсбук, имеют самый низкий медианный первый чек 389 рублей. Разница между максимальным и минимальным медианным чеком по каналам 1.017% (Учитываются только первые покупки пользователей.)**

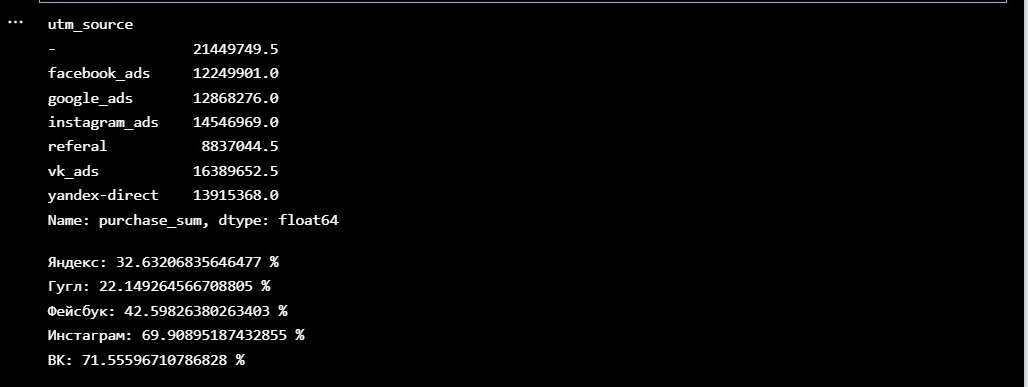
### Задание 1.1.8

**Расходы на реферальную программу: если пользователь приведет друга и последний совершит первую покупку, то оба получат по 100 рублей.**

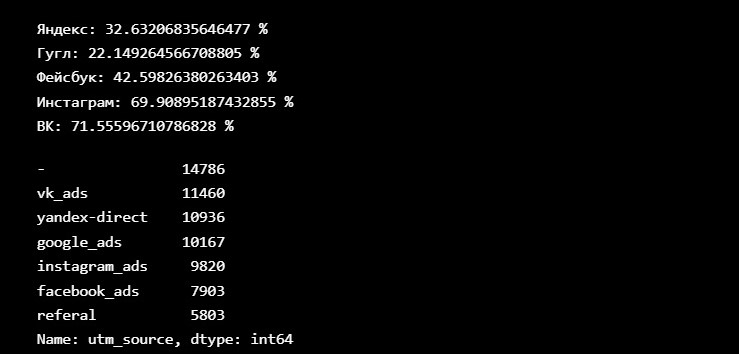
**Впишите в поля ниже один ответ из возможных вариантов: *Яндекс, Гугл, ВК, Фейсбук, Инстаграм*.**

**Вопрос 1. Какой платный канал привлечения имеет самый высокий ROMI?**

**Вопрос 2. Какой платный канал привлечения имеет самый низкий ROMI?**

****

**есть подозрение, что в коде ошибка и данные последней таблицы неверные**

****

**Самый высокий ROMI — у ВК (71.56 %). Скорее всего, это было достигнуто за счёт лояльности аудитории, пришедшей оттуда, а также их высокого LTV. Однако самая выгодная опция — реферальная программа, в рамках которой за 3 месяца 5803(тут вероятно должно получиться 5874) пользователей принесло 8837044.5 рублей.**

**2. Пользовательская аналитика**

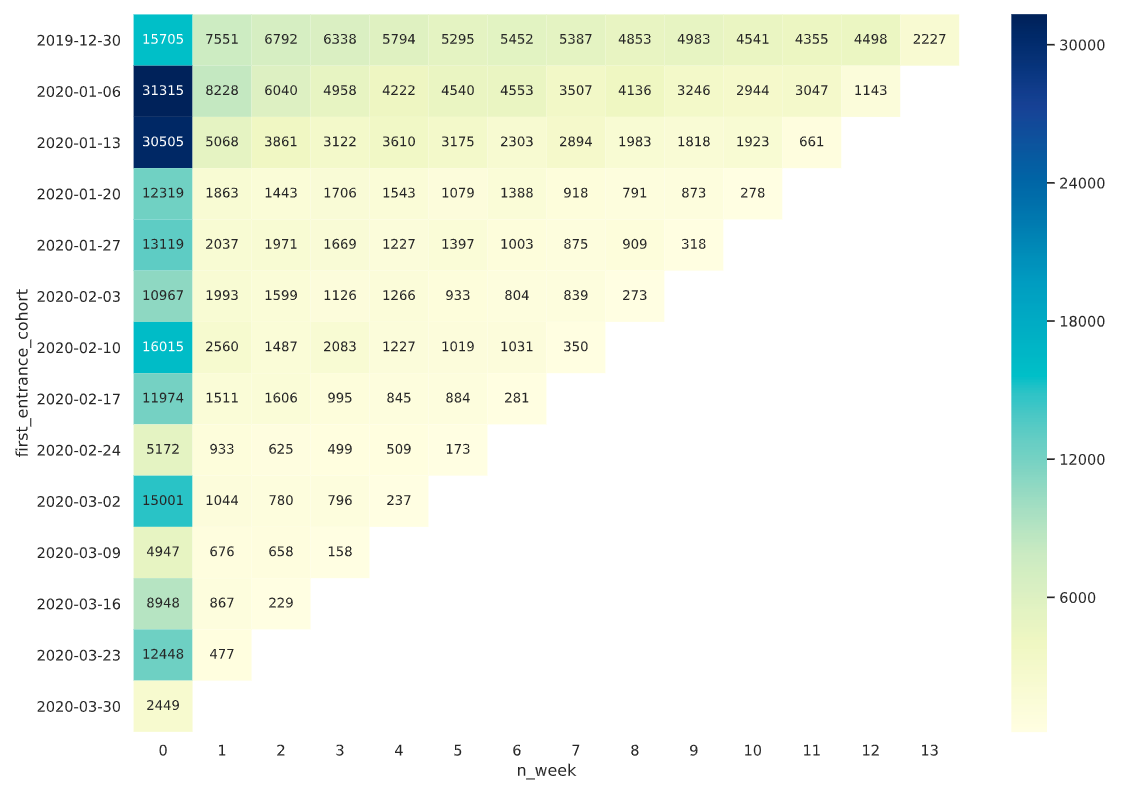
**Задачи**

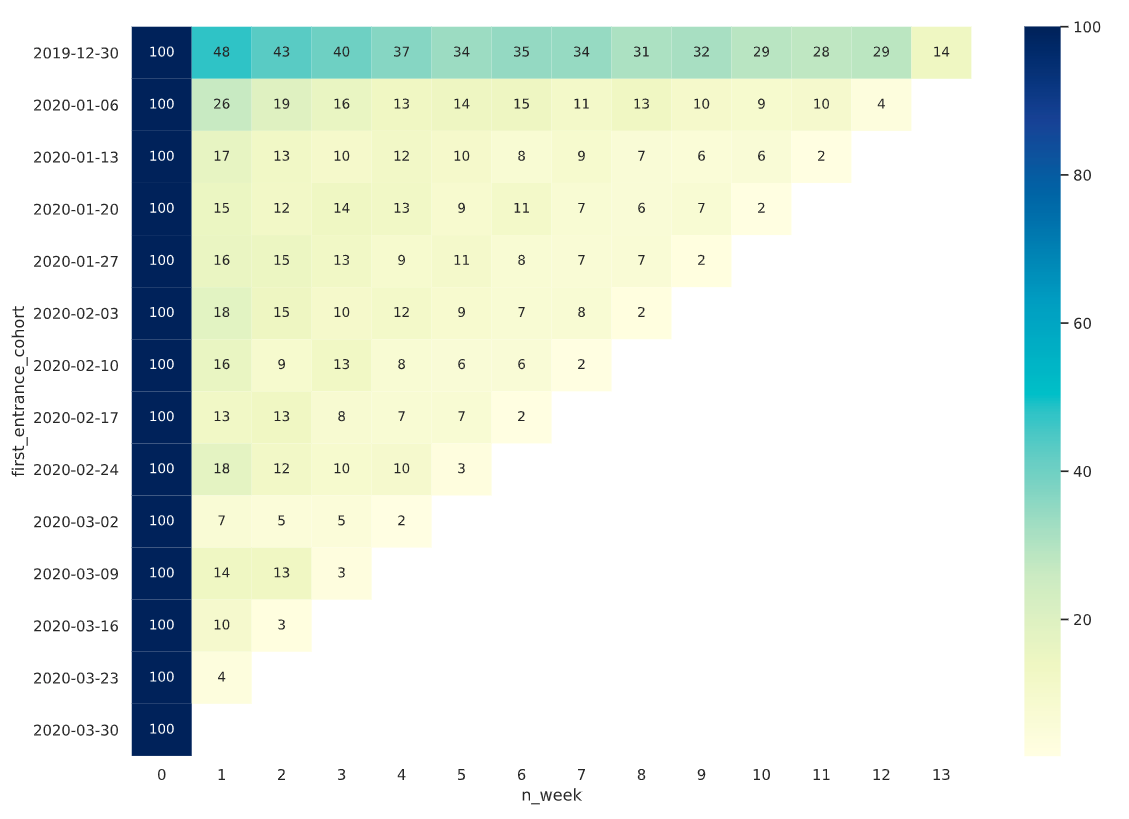
**Проведите когортный анализ по:**

* **неделе первого посещения;**
* **неделе регистрации;**
* **неделе первой покупки.**

### Задание 1.2.1

**Какая когорта самая активная? Ответ впишите в формате: ДД.ММ.ГГГГ. Дата здесь — это дата начала недели когорты.**

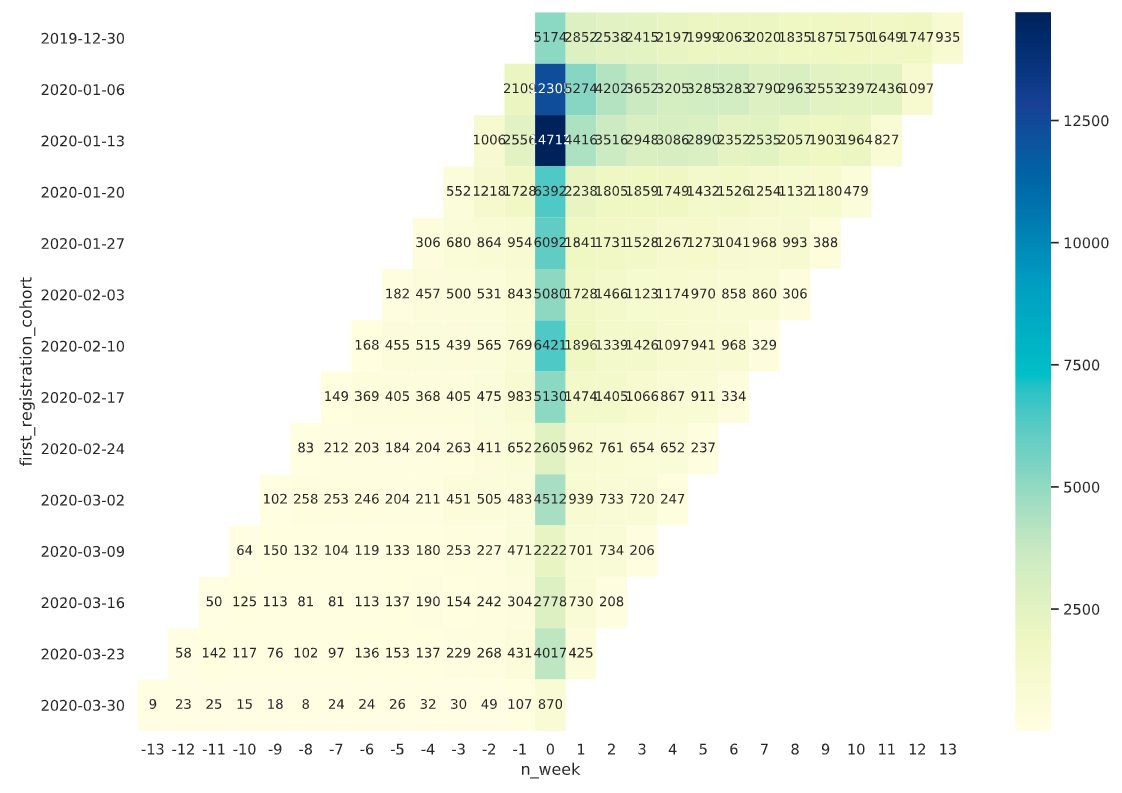
****

****

**Самая активная когорта — когорта, пришедшая на первой неделе Q1 30.12.2019 г.. Активность этой когорты в несколько раз выше, чем у остальных. Эта же когорта приносит больше всего денег. Стоит стремиться к тому, чтобы все остальные когорты были похожи на неё.**

### Задание 1.2.2

**Регистрируются ли все пользователи в день первого посещения?**

****

**Нетрадиционная форма heat map показывает нам, что многие пользователи несколько недель подряд заходят в приложение, прежде чем зарегистрироваться.**

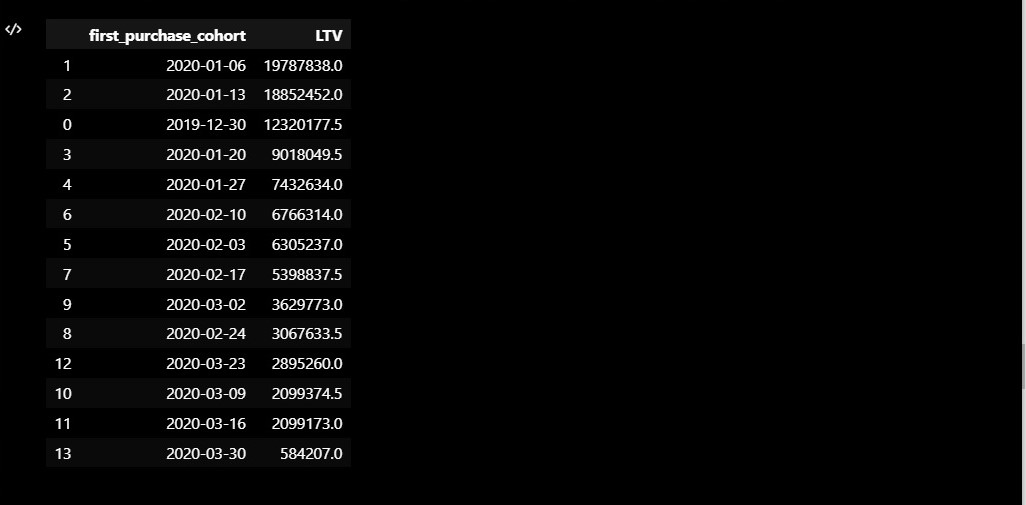
### 

### 

### 

### Задание 1.2.3

**Какая когорта принесла больше всего денег? Ответ впишите в формате: ДД.ММ.ГГГГ. Дата здесь — это дата начала недели когорты.**

****

**Когорта от 06.01.2020 принесла больше всего денег 19787838 рублей. Когорта от 30.03.2020 принесла меньше всего денег 584207 рублей.**

### 

### 

### 

### 

### 

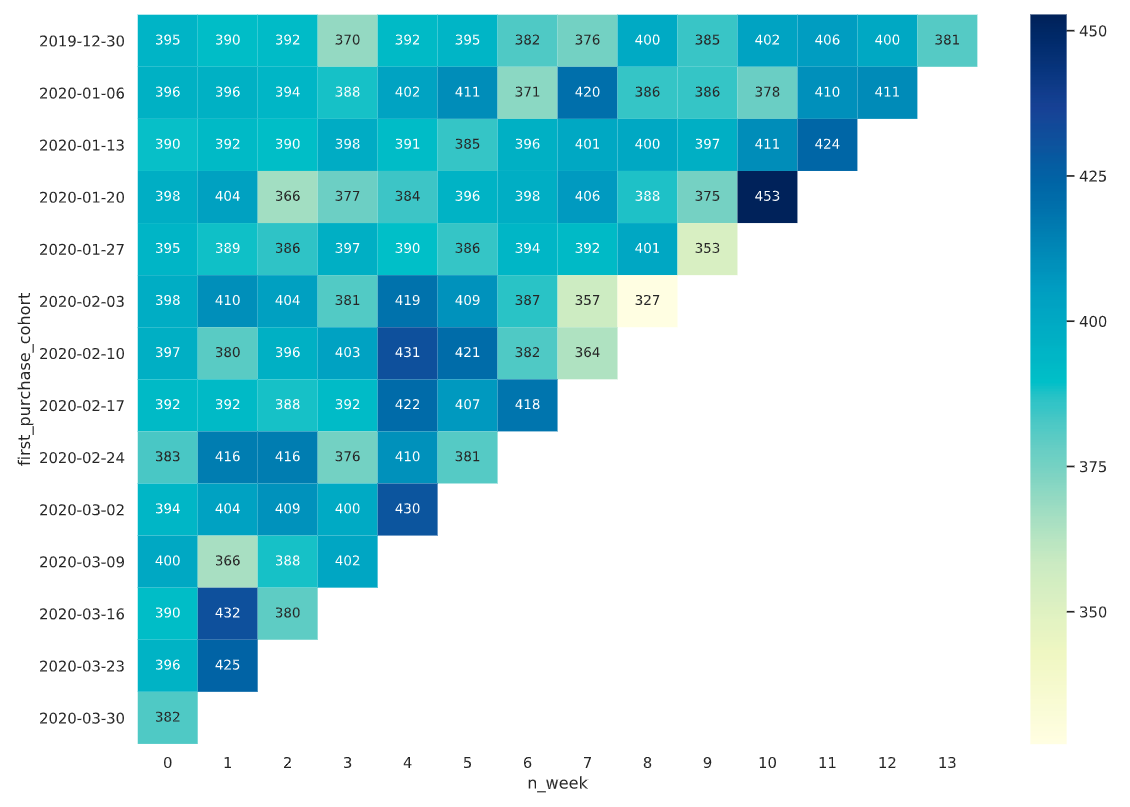
### 

### 

### 

### Задание 1.2.4

**У какой когорты выше медианный чек на неделе первой покупки? Ответ впишите в формате: ДД.ММ.ГГГГ. Дата здесь — это дата начала недели когорты.**

**Медианный чек на неделе первой покупки выше у когорты от 9 марта 2020 г.(=400 рублей). Самый маленький Медианный чек на неделе первой покупки у когорты от 30 марта 2020 г.(=382 рубля). Диапазон медианных чеков за три месяца 327 - 453 рублей.**

### 

### 

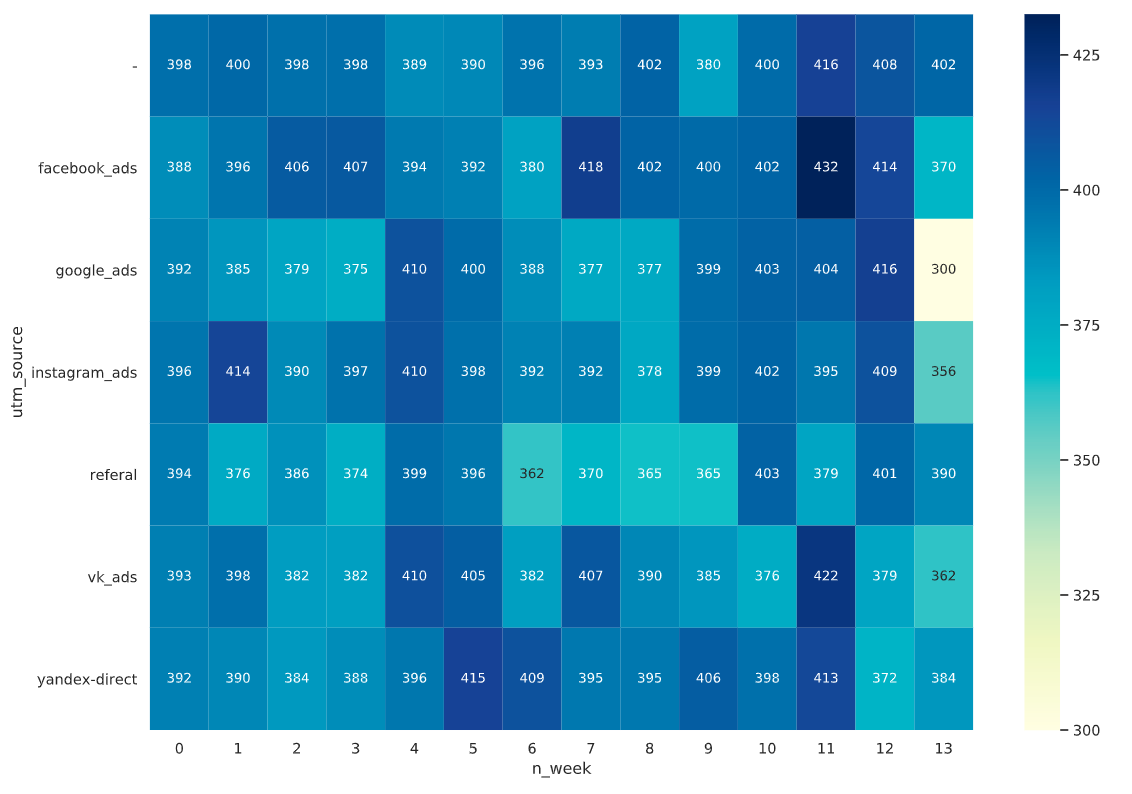
### 

### 

### 

### Задание 1.2.5

**Сравните медианный чек на второй неделе с момента первой покупки для всех каналов привлечения. У какого канала он выше?**

****

**Медианный чек на второй неделе с момента первой покупки на канале привлечения Инстаграм 414 рублей выше, чем у остальных.**