## Описание проекта

Вы — маркетинговый аналитик развлекательного приложения Procrastinate Pro+. Несмотря на огромные вложения в рекламу, последние несколько месяцев компания терпит убытки. Ваша задача — разобраться в причинах и помочь компании выйти в плюс.

Есть данные о пользователях, привлечённых с 1 мая по 27 октября 2019 года:

* лог сервера с данными об их посещениях,
* выгрузка их покупок за этот период,
* рекламные расходы.

Вам предстоит изучить:

* откуда приходят пользователи и какими устройствами они пользуются,
* сколько стоит привлечение пользователей из различных рекламных каналов;
* сколько денег приносит каждый клиент,
* когда расходы на привлечение клиента окупаются,
* какие факторы мешают привлечению клиентов.

## Инструкция

Рекомендуем ещё раз просмотреть урок «Разбор кейса» из темы «Юнит-экономика» перед началом работы.

### Шаг 1. Загрузите данные и подготовьте их к анализу

1.1. Положите данные о визитах, заказах и рекламных тратах в переменные.

1.2. Подготовьте данные к анализу. Убедитесь, что тип данных во всех колонках соответствует значениям. Проверьте отсутствие дубликатов.

**Пути к файлам**

* /datasets/visits\_info\_short.csv. [Скачать датасет](https://code.s3.yandex.net/datasets/visits_info_short.csv" \t "_blank);
* /datasets/orders\_info\_short.csv. [Скачать датасет](https://code.s3.yandex.net/datasets/orders_info_short.csv" \t "_blank);
* /datasets/costs\_info\_short.csv. [Скачать датасет](https://code.s3.yandex.net/datasets/costs_info_short.csv" \t "_blank).

### Шаг 2. Задайте функции для расчета и анализа LTV, ROI, удержания и конверсии

Разрешается использовать функции get\_profiles(), get\_retention(), get\_conversion(), get\_ltv(), а также plot\_retention(), plot\_conversion() и plot\_ltv\_roi(), с которыми вы познакомились в теоретических уроках.

### Шаг 3. Проведите исследовательский анализ данных

3.1. Создайте пользовательские профили. Определите минимальную и максимальную даты привлечения пользователей.

3.2. Выясните, из каких стран пользователи приходят в приложение и на какую страну приходится больше всего платящих пользователей. Постройте таблицу, отражающую количество пользователей и долю платящих из каждой страны.

3.3. Узнайте, какими устройствами пользуются клиенты и какие устройства предпочитают платящие пользователи. Постройте таблицу, отражающую количество пользователей и долю платящих для каждого устройства.

3.4. Изучите рекламные источники привлечения и определите каналы, из которых пришло больше всего платящих пользователей. Постройте таблицу, отражающую количество пользователей и долю платящих для каждого канала привлечения.

Напишите промежуточные выводы.

### Шаг 4. Маркетинг

4.1. Посчитайте общую сумму расходов на маркетинг. Выясните, как траты распределены по источникам. Визуализируйте изменения метрик во времени.

4.2. Узнайте, сколько в среднем стоило привлечение одного пользователя из каждого источника. Рассчитайте средний CAC на одного пользователя для всего проекта и для каждого источника трафика. Используйте профили пользователей.

Напишите промежуточные выводы.

### Шаг 5. Оцените окупаемость рекламы

Используя графики LTV, ROI и CAC, проанализируйте окупаемость рекламы. Считайте, что на календаре 1 ноября 2019 года, а в бизнес-плане заложено, что пользователи должны окупаться не позднее чем через две недели после привлечения. Необходимость включения в анализ органических пользователей определите самостоятельно.

5.1. Проанализируйте общую окупаемость рекламы. Постройте графики LTV и ROI, а также графики динамики LTV, CAC и ROI.

5.2. Проанализируйте окупаемость рекламы с разбивкой по рекламным каналам. Постройте графики LTV и ROI, а также графики динамики LTV, CAC и ROI.

5.3. Проанализируйте окупаемость рекламы с разбивкой по странам. Постройте графики LTV и ROI, а также графики динамики LTV, CAC и ROI.

Опишите обнаруженные проблемы и напишите промежуточные выводы.

5.4. Постройте и изучите графики конверсии и удержания с разбивкой по устройствам, странам, рекламным каналам.

Ответьте на такие вопросы:

* Окупается ли реклама в целом?
* Какие устройства, страны и каналы могут снижать окупаемость рекламы?
* Чем могут быть вызваны проблемы окупаемости?

Опишите возможные причины обнаруженных проблем и промежуточные рекомендации для рекламного отдела.

### Шаг 6. Напишите выводы

6.1. Выделите причины неэффективности привлечения пользователей.

6.2. Сформулируйте рекомендации для отдела маркетинга.

## Оформление

Задание выполните в Jupyter Notebook. Код пишите в ячейках типа code, текстовые пояснения — в ячейках типа markdown. Применяйте форматирование и заголовки.

### Описание данных

В вашем распоряжении три датасета. Файл visits\_info\_short.csv хранит лог сервера с информацией о посещениях сайта, orders\_info\_short.csv — информацию о покупках, а costs\_info\_short.csv — информацию о расходах на рекламу.

**Структура visits\_info\_short.csv**

* User Id — уникальный идентификатор пользователя,
* Region — страна пользователя,
* Device — тип устройства пользователя,
* Channel — идентификатор источника перехода,
* Session Start — дата и время начала сессии,
* Session End — дата и время окончания сессии.

**Структура orders\_info\_short.csv**

* User Id — уникальный идентификатор пользователя,
* Event Dt — дата и время покупки,
* Revenue — сумма заказа.

**Структура costs\_info\_short.csv**

* Channel — идентификатор рекламного источника,
* Dt — дата проведения рекламной кампании,
* Costs — расходы на эту кампанию.

## Критерии проверки

Ревьюеры ожидают, что в этом проекте вы:

* подготовили данные к анализу,
* выбрали подходящие типы визуализации,
* верно интерпретировали графики,
* правильно рассчитали и проанализировали метрики,
* аргументировали метриками рекомендации маркетологам,
* сформировали логичную структуру проекта,
* написали «чистый» и эффективный код,
* оставили комментарии,
* сделали корректные выводы.