MediaNation
Challenge the future

RFM-сегментация: как разделить аудиторию, чтобы продавать больше и чаще?

ПОДГОТОВЛЕНО MEDIANATION

Знакомство и розыгрыш подарков

https://medianation.ru/~MKuMK

MediaNation
Challenge the future

Содержание мастер-класса

1. Основные принципы RFM-сегментации и возможности, которые она открывает маркетингу
2. Какие программные продукты для проведения RFM-анализа использовать
3. Работа с RFM в Яндекс.Аудиториях
4. Стратегии работы с RFM-аудиториями в рекламе
5. Алгоритм проведения RFM анализа

MediaNation
Challenge the future

Что такое RFM анализ и зачем он нужен?

RFM-анализ — это метод анализа данных о покупках клиентов, который позволяет **производить сегментацию базы клиентов по трем критериям:**

частота покупок

сумма покупок

покупки

для дальнейшего анализа и принятия маркетинговых и коммерческих решений

Позволит сэкономить рекламные бюджеты и увеличить рентабельность вложений в маркетинг

MediaNation
Challenge the future

Что такое RFM анализ и зачем он нужен?

Recency
давность
(как давно ваши клиенты что-то у вас покупали)

Frequency
частота
(как часто клиенты у вас покупают)

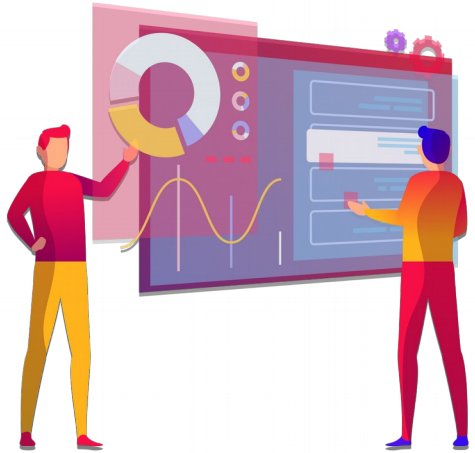
Monetary
деньги
(общая сумма покупок)

MediaNation
Challenge the future

Кому нужно делать RFM-анализ?

Подходит для проектов, продающих свою продукцию довольно часто:

- сферы FMCG
- продавцов подписок (с небольшим изменением R),
- интернет-магазинов с высокой частотой покупок



Мало эффективным RFM анализ является для:

- магазинов дорогой мебели
- эксклюзивных товаров
- предметов искусства

Но помните, что всегда существуют ограничения и исключения

MediaNation
Challenge the future

Условия проведения RFM анализа

1. При RFM мы анализируем данные о покупках за какой-то период. Обычно берем 365 дней

2. В алгоритме RFM необходимо также указать период предсказания. То есть алгоритм может сделать предсказания поведения потребителей на выбранный нами срок. Обычно берем 182 дня

3. В CRM базе должны быть тысячи, миллионы пользователей, которые совершали у вас покупки за выбранный период

4. Списки периодически необходимо обновлять

MediaNation
Challenge the future

Давность, частота, сумма покупки

RFM анализ позволяет разделить клиентов на 27 (64, либо 125) сегментов по сочетанию параметров:

1

Давность покупки:

- недавно (3)
- относительно недавно (2)
- давно (1)

2

Частота покупки:

- часто (3)
- средняя частота (2)
- разовая покупка (1)

3

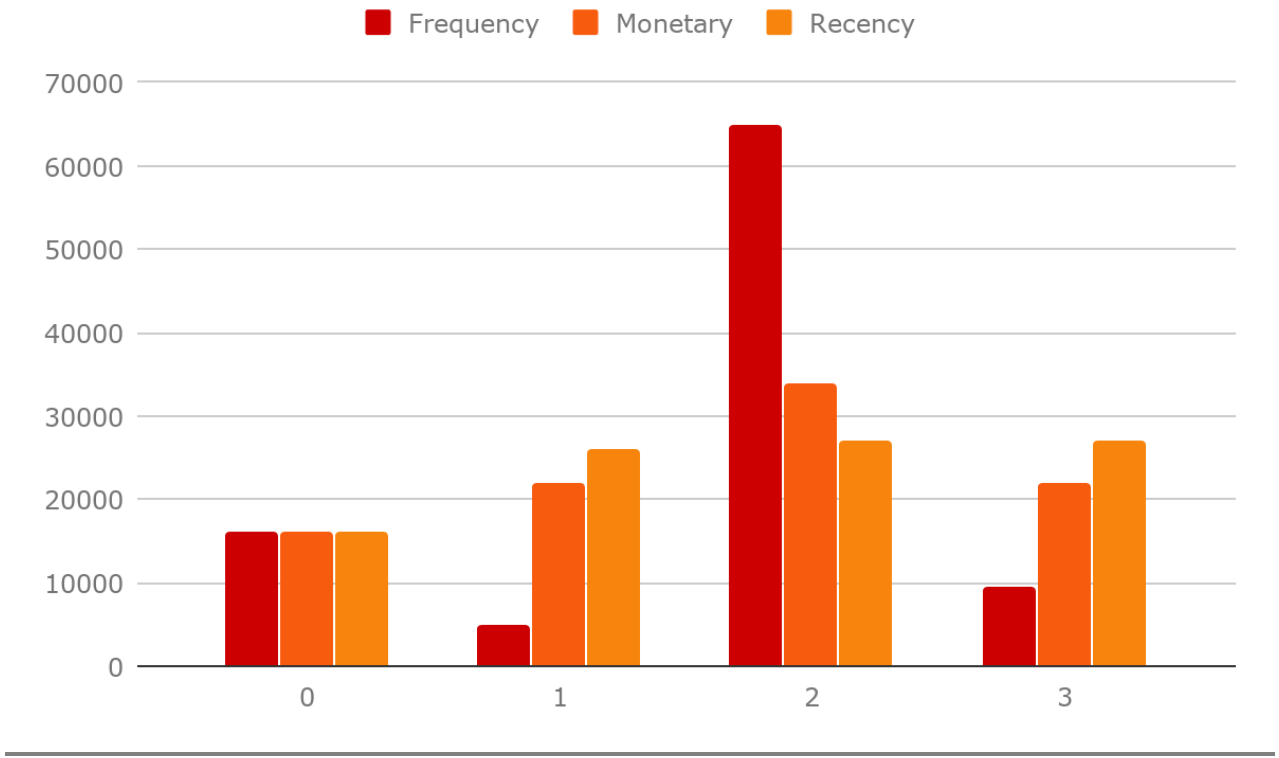
Сумма покупки:

- высокий чек (3)
- средний чек (2)
- низкий чек (1)

```
df_filter = df.groupby('RFM')['user_id'].count()
print (df_filter)

RFM
0-0-0    [redacted]5
1-1-1    5[redacted]
1-1-2    [redacted]06
1-1-3    2[redacted]
1-2-1    [redacted]48
1-2-2    [redacted]3
1-2-3    6[redacted]
1-3-1    [redacted]1
1-3-2    [redacted]5
1-3-3    [redacted]1
2-1-1    4[redacted]
2-1-2    4[redacted]
2-1-3    [redacted]2
2-2-1    [redacted]
2-2-2    [redacted]3
2-2-3    [redacted]
2-3-1    [redacted]1
2-3-2    [redacted]3
2-3-3    [redacted]84
3-1-1    [redacted]54
3-1-2    [redacted]5
3-1-3    [redacted]4
3-2-1    [redacted]3
3-2-2    [redacted]6
3-2-3    [redacted]40
3-3-1    [redacted]3
3-3-2    [redacted]0
3-3-3    [redacted]1
Name: user_id, dtype: int64
```

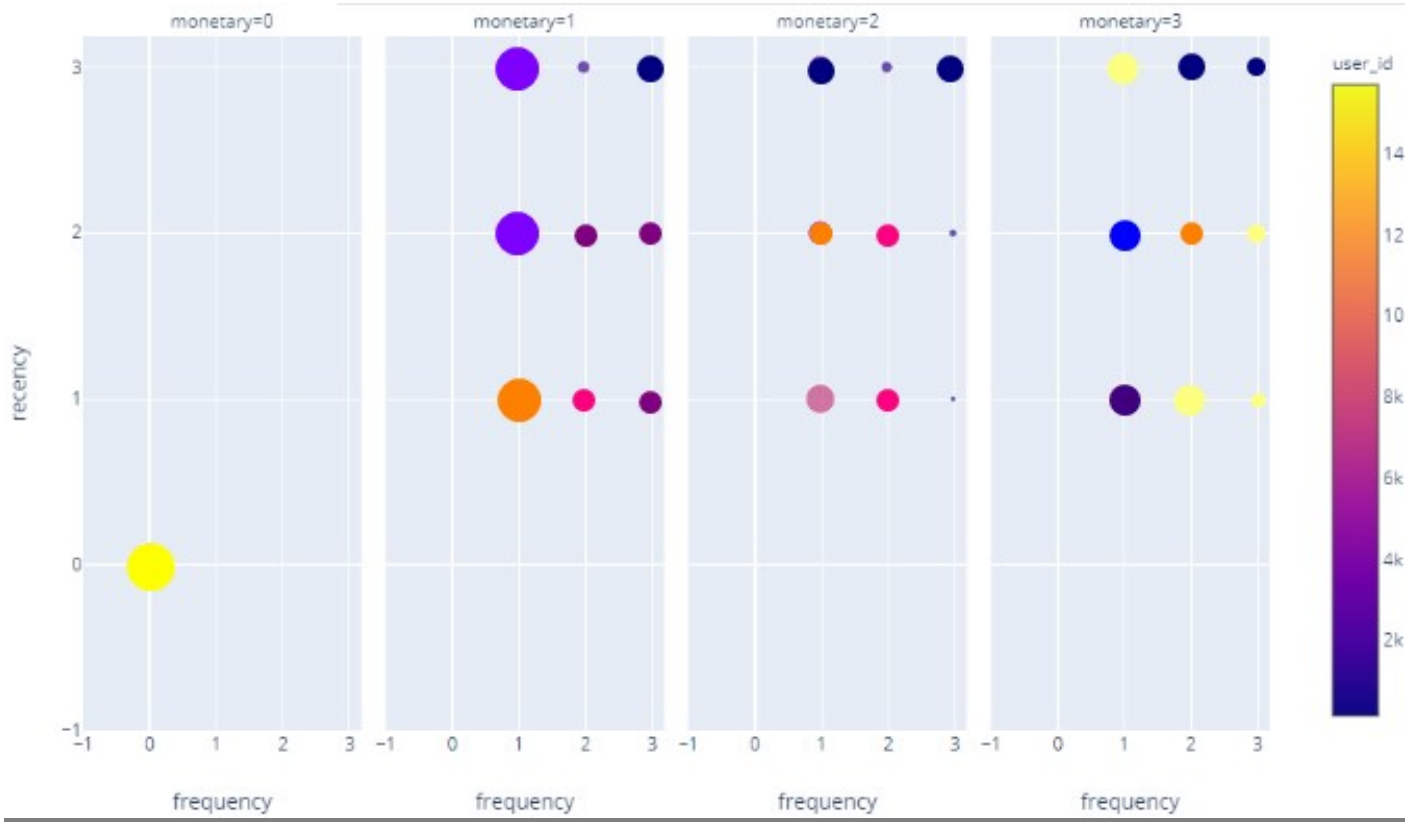
MediaNation
Challenge the future

Визуализация кол-ва пользователей по сегментам 1, 2, 3 каждого параметра R-F-M



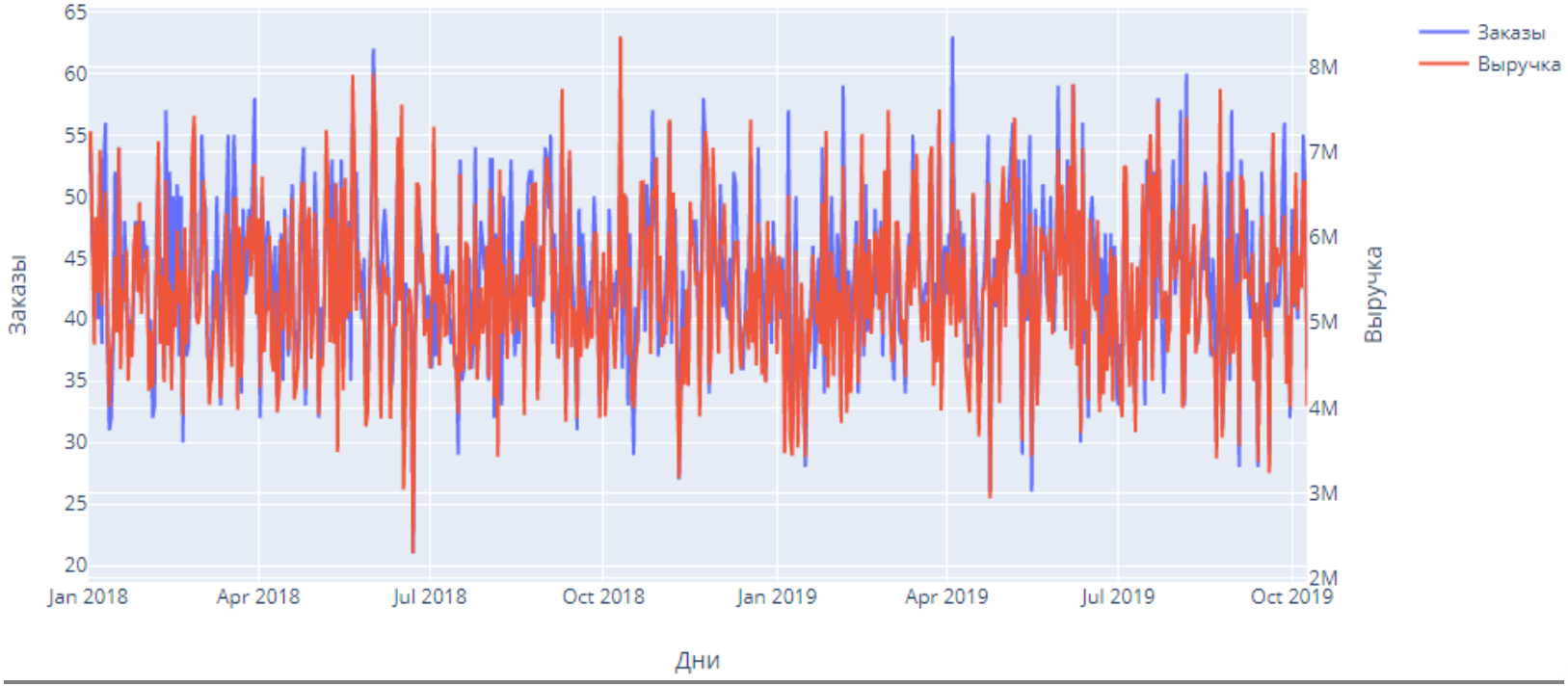
MediaNation
Challenge the future

Визуализация результатов RFM анализа



MediaNation
Challenge the future

Визуализация результатов RFM анализа



MediaNation
Challenge the future

Программные продукты для проведения RFM-анализа

Sublime Text

RFMizer

MediaNation
Challenge the future

CSV файл для RFM анализа

https://medianation.ru/~MEkqk

order_date,user_id,order_value
2018-11-16,49,160418
2018-03-30,49,87424
2019-09-09,67,228445
2018-12-07,67,140733
2019-07-23,67,5143
2019-05-23,67,44158
2019-02-05,67,130638
2019-08-16,67,244325
2018-12-24,67,212708
2018-10-03,67,23277
2018-04-06,67,240145
2018-03-22,67,198593
2018-09-05,67,83526
2018-12-15,67,189733

MediaNation
Challenge the future

Содержание файла-результата RFM

```
df_filter = df.groupby('RFM')['user_id'].count()
print (df_filter)
```

```
RFM
0-0-0     [redacted]55
1-1-1      5[redacted]
1-1-2     [redacted]106
1-1-3      2[redacted]
1-2-1     [redacted]48
1-2-2     [redacted]3
1-2-3       6[redacted]
1-3-1     [redacted]1
1-3-2     [redacted]5
1-3-3     [redacted]1
2-1-1      4[redacted]
2-1-2      4[redacted]
2-1-3     [redacted]2
2-2-1     [redacted]
2-2-2     [redacted]3
2-2-3     [redacted]
2-3-1     [redacted].1
2-3-2     [redacted]3
2-3-3     [redacted]84
3-1-1     [redacted]54
3-1-2     [redacted]5
3-1-3     [redacted]4
3-2-1     [redacted]3
3-2-2     [redacted]6
3-2-3     [redacted]40
3-3-1     [redacted]3
3-3-2     [redacted]0
3-3-3     [redacted]1
Name: user_id, dtype: int64
```

	A	B	C	D	E
1	user_id,birthdate,email,sex,RFM				
169455	12332,1980-01-01,moscow@mail.ru,N,1-4-5				
169456	10472,1983-09-06,russia@yandex.ru,F,1-2-4				
169457	104836,1970-04-22,europe@yandex.ru,,0-0-0				
169458	1048264,1992-03-22,asia@mail.ru,F,1-1-4				
169459	9462743,1987-05-22,kaliningrad@gmail.com,F,3-2-1				
169460	1234543,1970-01-13,example@mail.ru,N,3-2-2				
169461	983234,1991-07-28,forinstance@mail.ru,F,3-3-5				
169462	3495683,1977-07-06,test@yandex.ru,,0-0-0				
169463	4564234,1980-01-01,rfm@mail.ru,N,1-1-5				
169464	213431,1988-06-01,podbor@mail.ru,F,0-0-0				
169465	2134252,1980-01-01,test43@mail.ru,N,3-3-3				
169466	2134524,1983-09-16,report@yandex.ru,,0-0-0				
169467	3453542,1982-10-15,research@mail.ru,F,4-3-1				
169468	23445329,1993-03-03,search@mail.ru,N,5-5-3				

MediaNation
Challenge the future

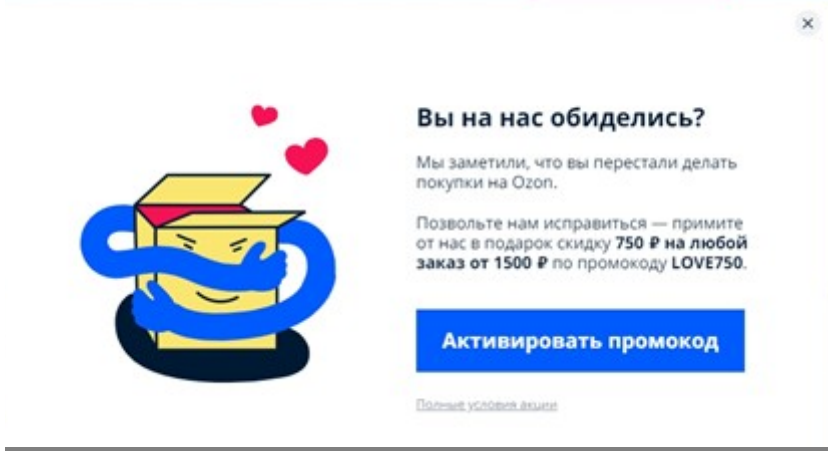
Кастомизированные действия

1. Клиенты с разовыми покупками (111, 112, 113)

Можно предложить:

- специальные бонусы
- скидки
- подарки

Помните:
отток клиентов из базы неизбежен

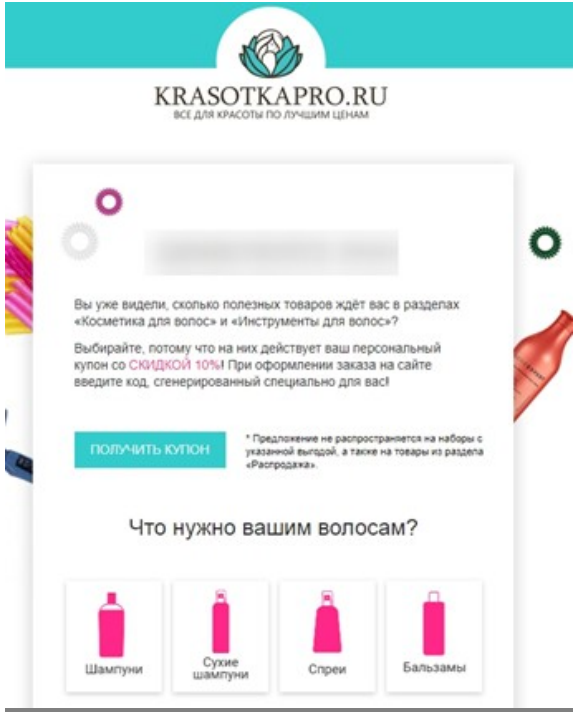


MediaNation
Challenge the future

2. Клиент не совершал покупки некоторое время, но раньше покупал часто и на большие суммы (222, 223, 232, 233)

Предлагаем:

- скидку на заказ (например, «Скидка действует всю неделю»)
- разослать информацию о распродаже
- создать персональную товарную подборку



MediaNation
Challenge the future

3. Клиенты, склонные к покупкам (333)

Предлагаем:

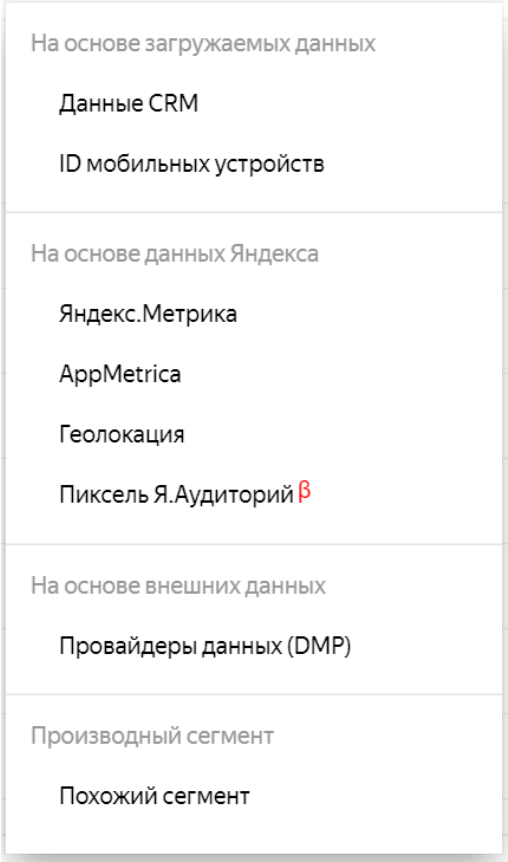
- повышать ставки в динамическом ремаркетинге или в контексте
- программы лояльности для постоянных клиентов
- персональные товарные подборки
- сообщения о важных нововведениях
- полезный контент в рассылках
- попробовать продвигать и свои дорогие продукты

MediaNation
Challenge the future

Яндекс.Аудитории

https://audience.yandex.ru/

Как можно использовать Яндекс.Аудитории?

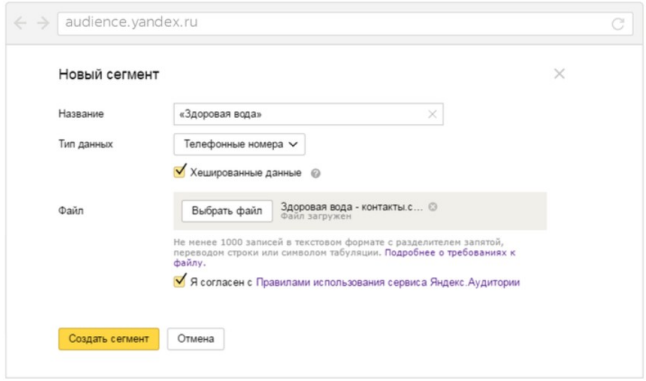


MediaNation
Challenge the future

Подготовили файл CSV и загрузили в Яндекс.Аудитории

phones

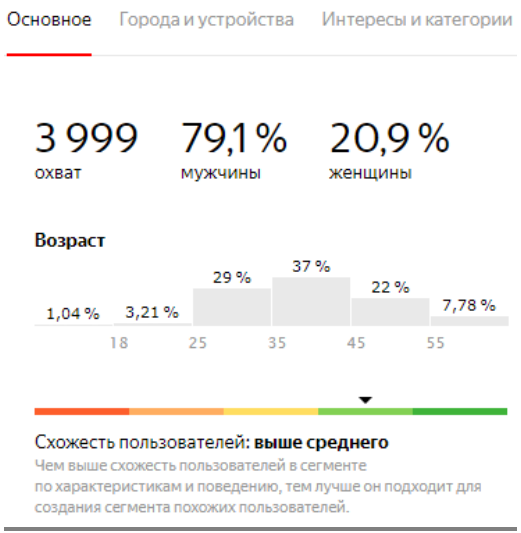
13 500 телефонных контактов выгрузили из CRM-системы, дубликаты удалили и подготовили файлы в формате CSV.

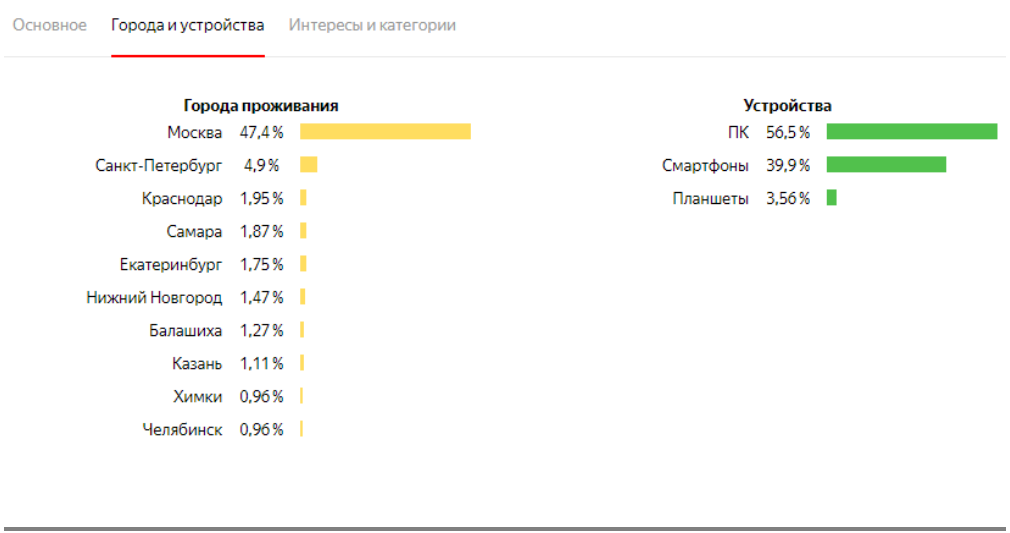


Файл загрузили в сервис Яндекс.Аудитории и создали сегменты.

MediaNation
Challenge the future

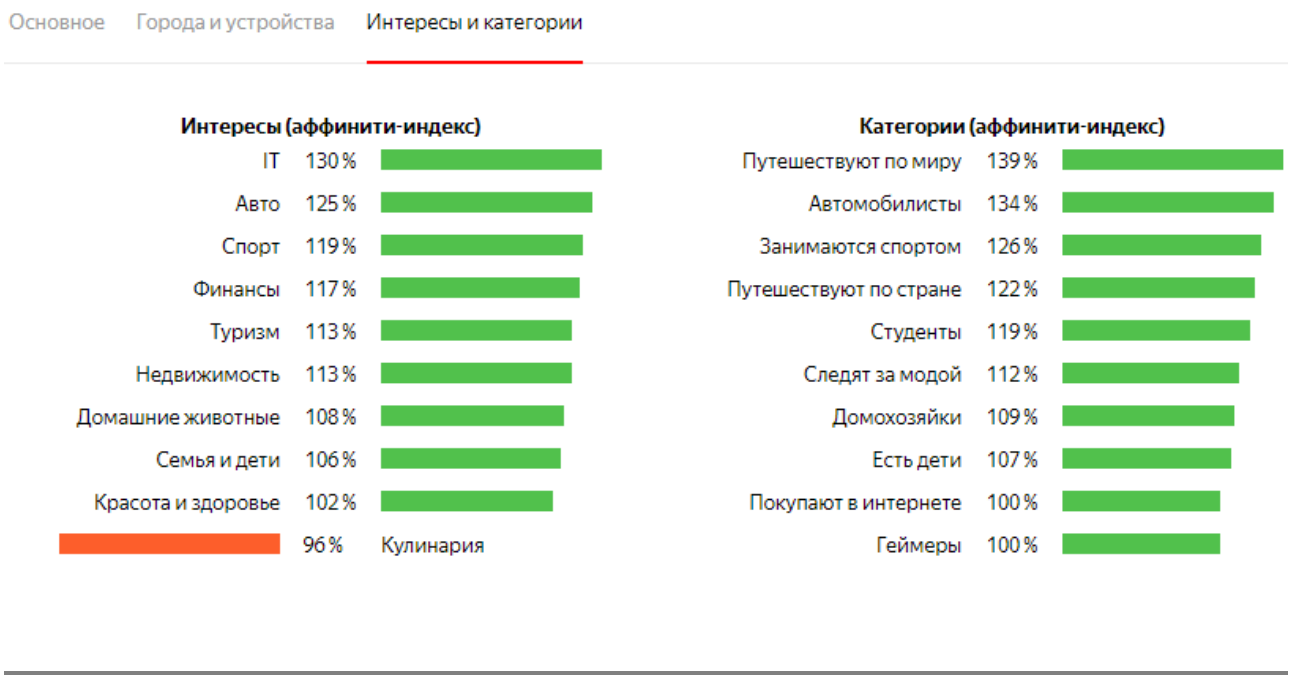
Пример результата обработки данных в Яндекс.Аудиториях





MediaNation
Challenge the future

Пример результата обработки данных в Яндекс.Аудиториях



MediaNation
Challenge the future

Аффинити-индекс

Аффинити-индекс — показатель того, насколько сильно эта тематика интересует посетителей сайтов по сравнению с аудиторией интернета в целом:

- Значение, близкое к 100%, говорит о том, что интерес пользователей к теме ничем не отличается от среднего по интернету.
- Если аффинити-индекс, наоборот, заметно превышает 100%, тематика может быть привлекательной для вашей аудитории. Например, индекс 200% для автомобилей означает, что посетители сайта вдвое чаще интересуются темой авто, чем обычные пользователи в сети.

Анализируя аффинити-индекс, вы можете выявить смежные интересы вашей аудитории. Например, узнать, что ваши целевые посетители интересуются не только авто, но и финансами.

MediaNation
Challenge the future

Сегмент «Арбат 4»

Вы можете создать аудиторию с помощью выбора мест на карте, в которых пользователи находятся прямо сейчас или регулярно их посещают, или настроить более сложное условие посещения.

+ Добавить местоположение

Радиус: 2 км., показан

Аудитория: регулярно посещает

Как аудитория связана с группой местоположения

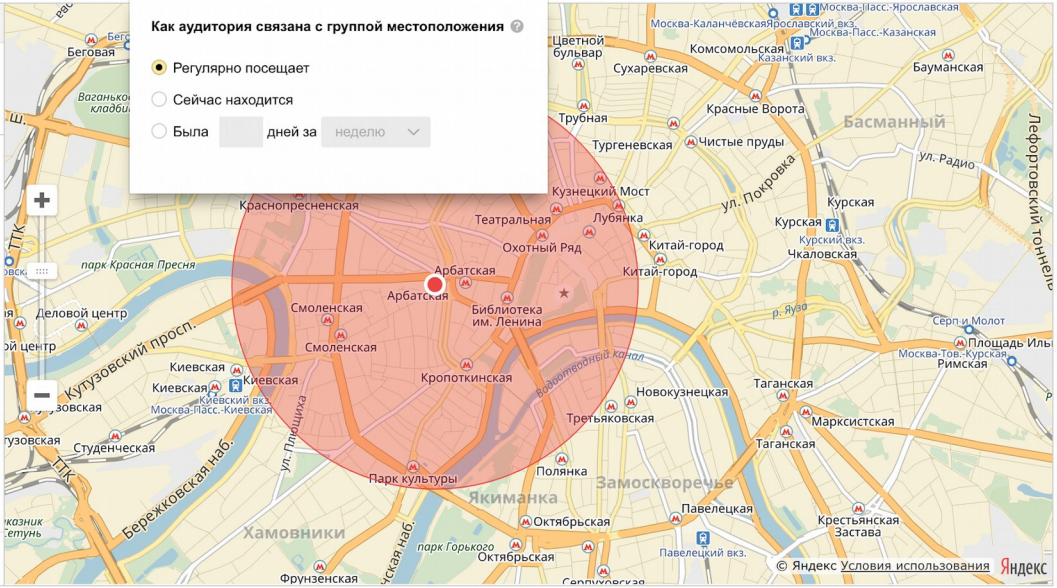
- (•) Регулярно посещает
- ( ) Сейчас находится
- ( ) Была ____ дней за неделю

Местоположения 1

Поиск 1

Арбат 4 | Найти

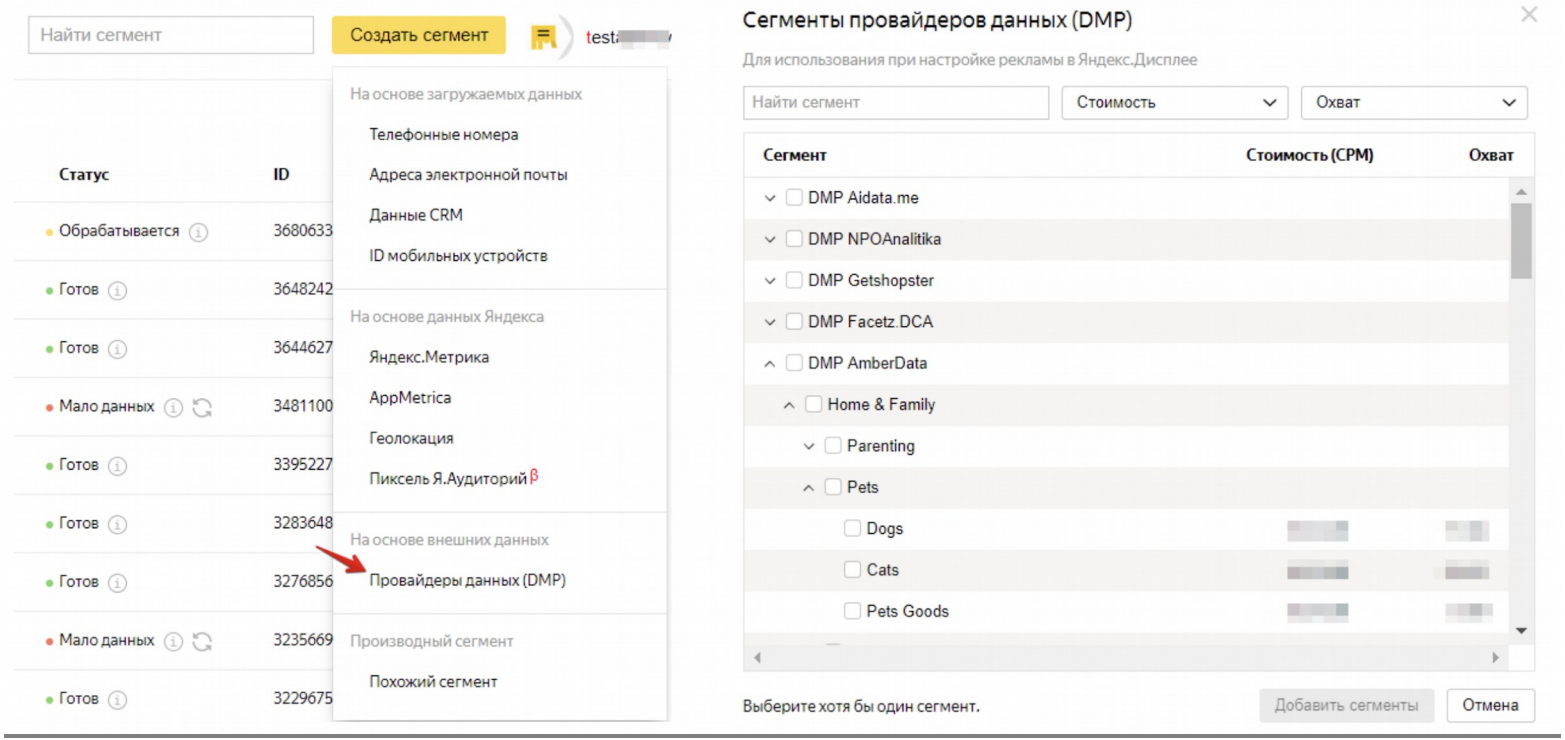
Россия, Москва, улица Арбат, 4с1
55.752035, 37.598492



Добавить все адреса в Местоположения

Создать сегмент | Отмена

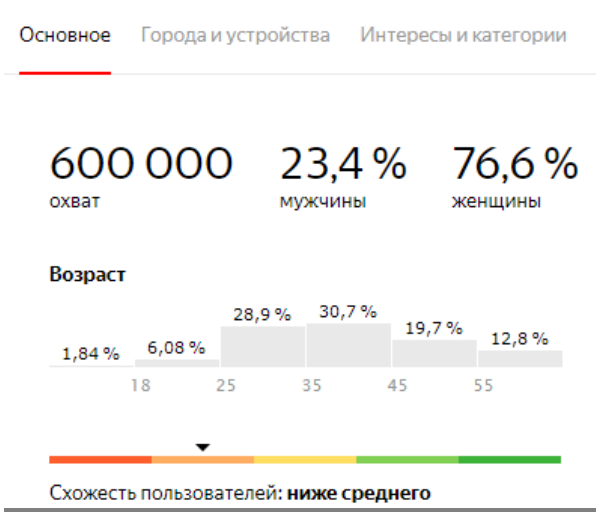
MediaNation
Challenge the future

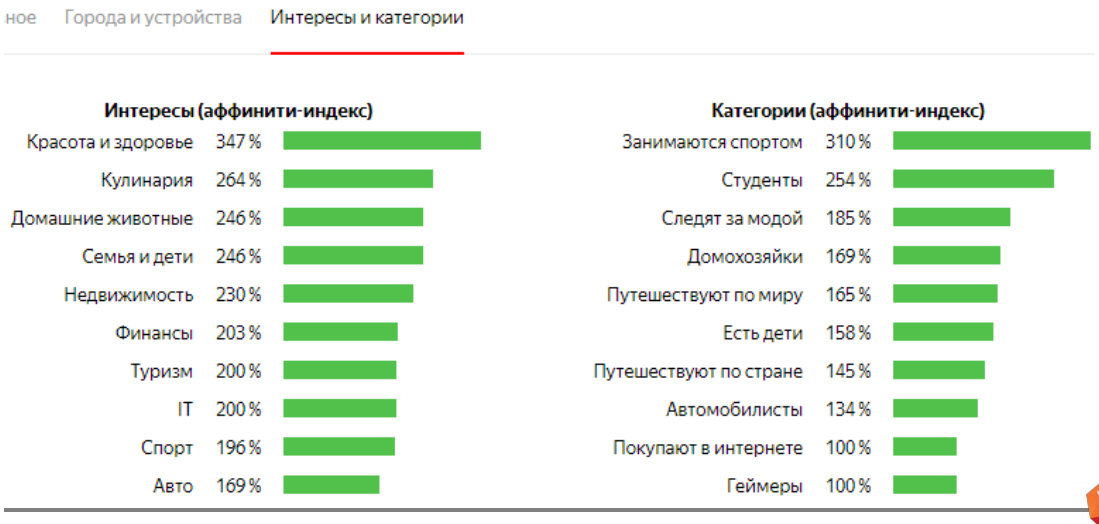

Найти сегмент
Создать сегмент
Статус
ID
Обрабатывается 3680633
Готов 3648242
Готов 3644627
Мало данных 3481100
Готов 3395227
Готов 3283648
Готов 3276856
Мало данных 3235669
Готов 3229675
На основе загружаемых данных
Телефонные номера
Адреса электронной почты
Данные CRM
ID мобильных устройств
На основе данных Яндекса
Яндекс.Метрика
AppMetrica
Геолокация
Пиксель Я.Аудиторий β
На основе внешних данных
Провайдеры данных (DMP)
Производный сегмент
Похожий сегмент
Сегменты провайдеров данных (DMP)
Для использования при настройке рекламы в Яндекс.Дисплее
Найти сегмент
Стоимость
Охват
Сегмент
Стоимость (CPM)
Охват
DMP Aidata.me
DMP NPOAnalitika
DMP Getshopster
DMP Facetz.DCA
DMP AmberData
Home & Family
Parenting
Pets
Dogs
Cats
Pets Goods
Выберите хотя бы один сегмент.
Добавить сегменты
Отмена

Создание Look-a-like

Пример результата Look-a-like

При создании Look-a-like была выбрана максимальная точность против охвата (определяете в зависимости от цели).





MediaNation
Challenge the future

Look-a-like сегмент

Мы можем работать с новыми аудиториями следующим образом:

- Запустить кампании на РСЯ, КМС
- Запустить видео рекламу
- Использовать смарт-баннеры

Аудиторию, не знающую о бренде, может привлечь яркий креатив

MediaNation
Challenge the future

Общая схема взаимодействия с RFM-сегментами

Каждый бизнес требует индивидуального подхода, и он должен проявиться в содержании предложений и акций для покупателей

Много покупок, давно →

- Скидка
- Акция на товары, которые покупал клиент
- Опрос о качестве сервиса

Мало покупок, средняя активность →

- Бонусные баллы, программа лояльности
- Акция на следующую покупку
- Накопительная система скидок

Мало покупок, низкая активность →

- Реактивационная цепочка с провокационной темой
- Промокод на скидку

Дорогие покупки, средняя активность →

- Программы лояльности
- Скидки лояльным клиентам+
- Закрытые распродажи
- Спецпредложения, акционные пакеты товаров

MediaNation
Challenge the future

Плюсы RFM анализа

CRM

Не требуются дополнительные данные. Для такого анализа подойдет уже имеющаяся у вас информация из CRM-системы

Возможность работы с многотысячной клиентской базой

Благодаря автоматизации процесса занимает не очень много времени

Не только не требует дополнительных затрат, но и в последствии помогает оптимизировать бюджет и получить прибыль

MediaNation
Challenge the future

Практика

Проведение RFM анализа

Структура практики

1. Подготовительный этап:
    - Установка необходимых программ
    - Создание файла для RFM анализа
    - Настройка RFMizer-config.yaml
1. Запуск скрипта
2. Подготовка к визуализации
3. Визуализация данных
4. Выбор сегментов для Яндекс.Аудиторий
5. Создание Look-a-like

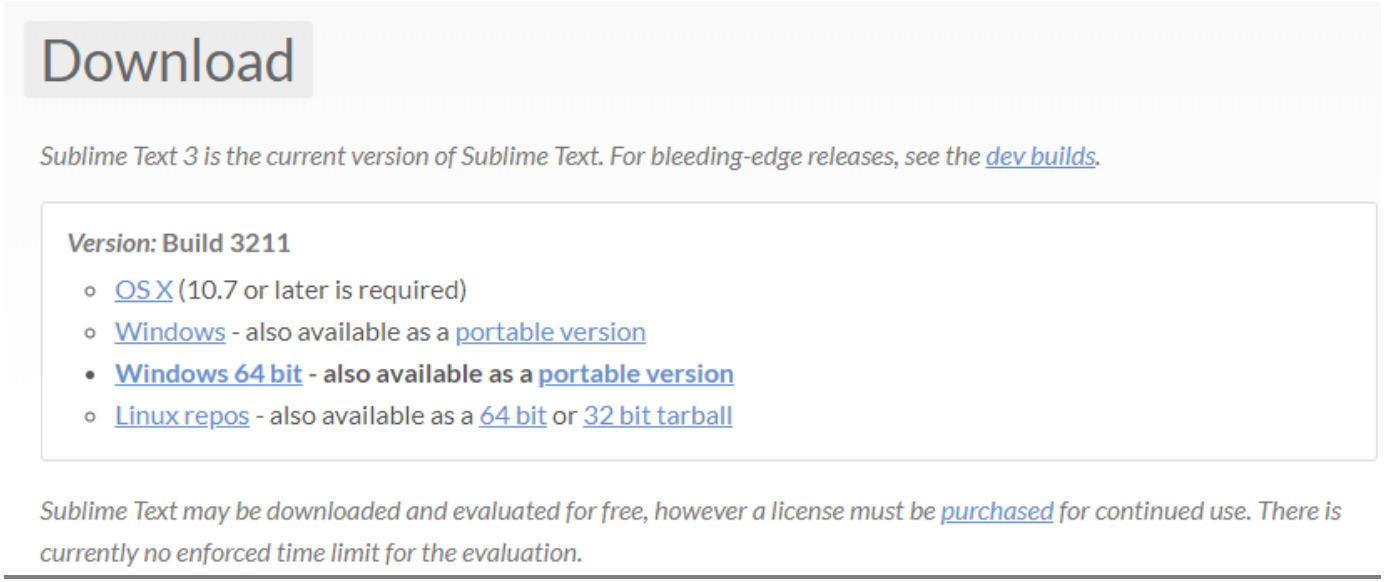
MediaNation
Challenge the future

Подготовительный этап

Установка программ для RFM

Sublime Text

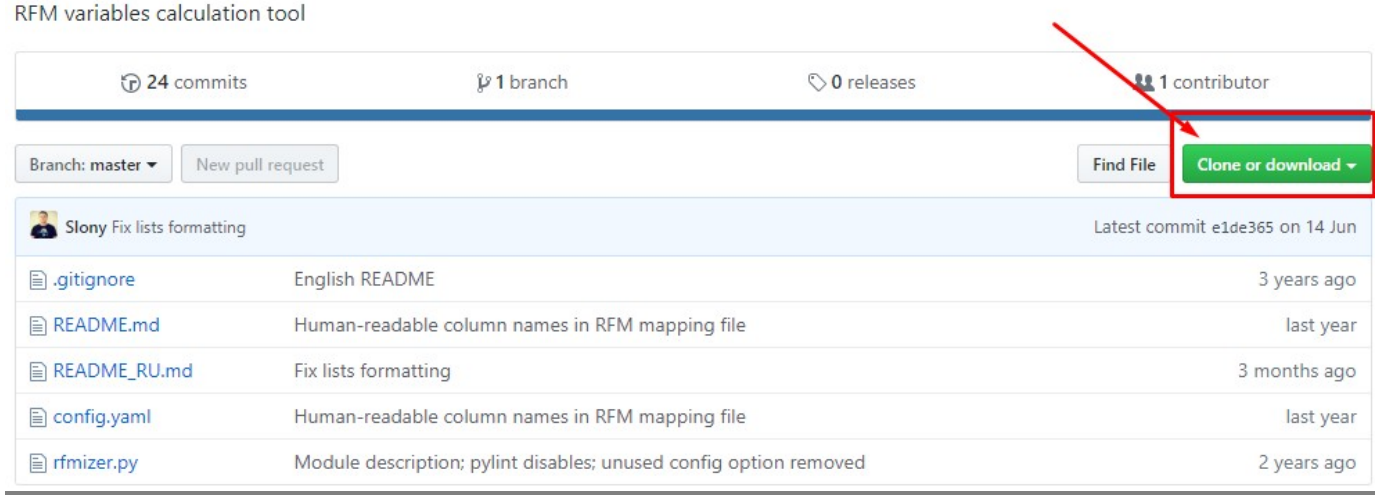
1. Sublime Text - текстовый редактор, ориентированный на серьезную работу с кодом и разметкой https://www.sublimetext.com/3



MediaNation
Challenge the future

Библиотека RFMizer

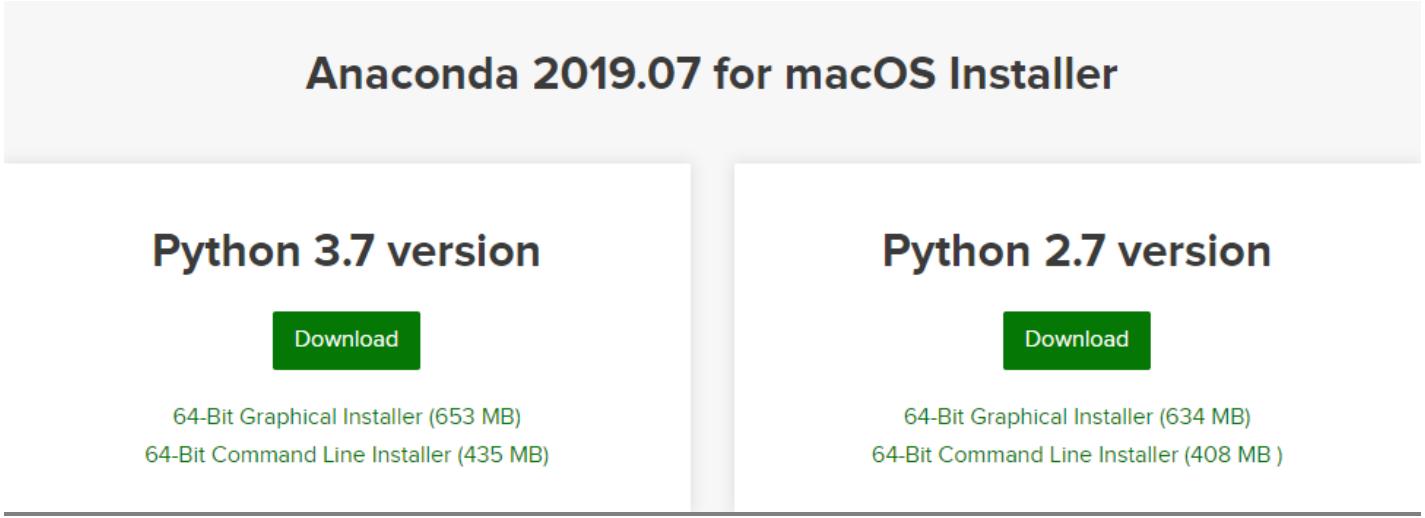
2. Библиотека RFMizer на языке программирования Python
   https://github.com/Slony/rfmizer
   Установите и распакуйте в папку



MediaNation
Challenge the future

Anaconda

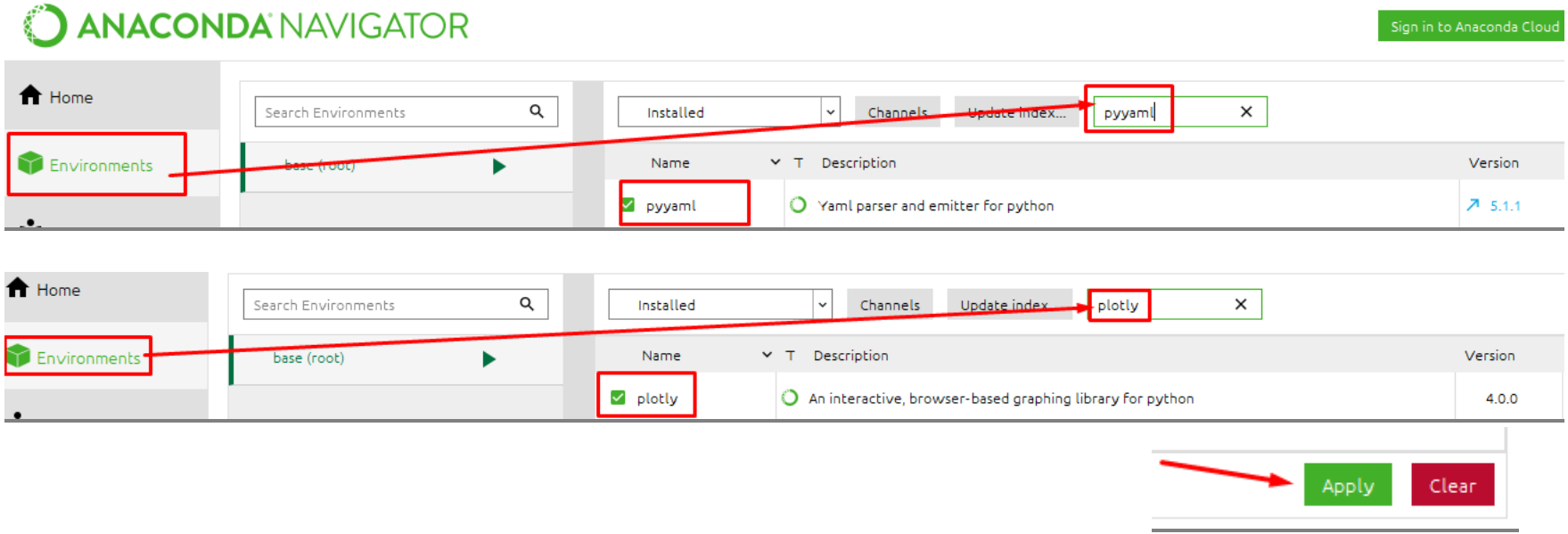
3. Интерпретатор языка Python - Anaconda https://www.anaconda.com/distribution/ (Python 3.7 version)



MediaNation
Challenge the future

Установка pyyaml и plotly

Далее открываем Anaconda и устанавливаем пакеты pyyaml и plotly (через кнопку внизу страницы Apply)



MediaNation
Challenge the future

Подготовительный этап

Создание файла для RFM анализа

Подготовка csv файла для RFM анализа

Для начала Вам необходимо выгрузить из CRM следующие данные:

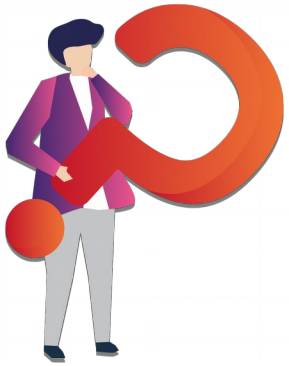
	A	B	C	D	E	F	G
1	**Обязательно**			Дополнительно			
2	**user_id**	**order_date**	**order_value**	phone	email	sex	birthdate
3	**1**	**21.08.2019**	**5399**	79851232288	test@mail.ru	F	01.04.1963
4	**34**	**19.08.2019**	**2400**	79851232471	ekample@yandex.ru	M	03.10.1970

При желании, Вы можете указать еще больше дополнительных данных.

MediaNation
Challenge the future

ВОПРОС

Что делать, если у клиента нет user_id пользователей?



MediaNation
Challenge the future

ОТВЕТ

Заменить на email / тел

MediaNation
Challenge the future

Подготовка csv файла для RFM анализа

Далее необходимо обработать данные и создать файл, который должен соответствовать требованиям:

- Формат csv и кодировка UTF-8 (обратите внимание при сохранении документа)
- В начале файла НЕ должно быть маркера последовательности байтов (BOM, byte order marker).
- Поля в строках должны разделяться запятыми: «,».
- данные записываются в одном столбце через запятую и без пробелов
- Номер телефона указывать только через 7.
- Использование пробелов запрещено. При необходимости заменяйте “_”
- В качестве символа десятичной точки в числах должен использоваться символ `.` (точка) или символ `,` (запятая).
- В полях, содержащих двойные кавычки `"`, все символы двойных кавычек должны быть задвоены: везде вместо `"` надо вставить `""`.
- Поля, содержащие запятую `,`, двойные кавычки `"` или переводы строки, должны быть заключены в двойные кавычки `"`.

MediaNation
Challenge the future

Пример файла для RFM анализа

	A	B	C	D	E	F	G
1	user_id,order_date,order_value,phone,email,sex,birthdate						
2	1,2019-08-21,5399,79851232288,test@mail.ru,F,1963-04-01						
3	34,2019-08-19,2400,79851232471,ekample@yandex.ru,M,1970-10-03						

ВАЖНО: сохраните этот файл в распакованную папку RFMizer (в будущем будем называть этот файл “input-file”)

MediaNation
Challenge the future

CSV файл для RFM анализа

https://medianation.ru/~MEkqk

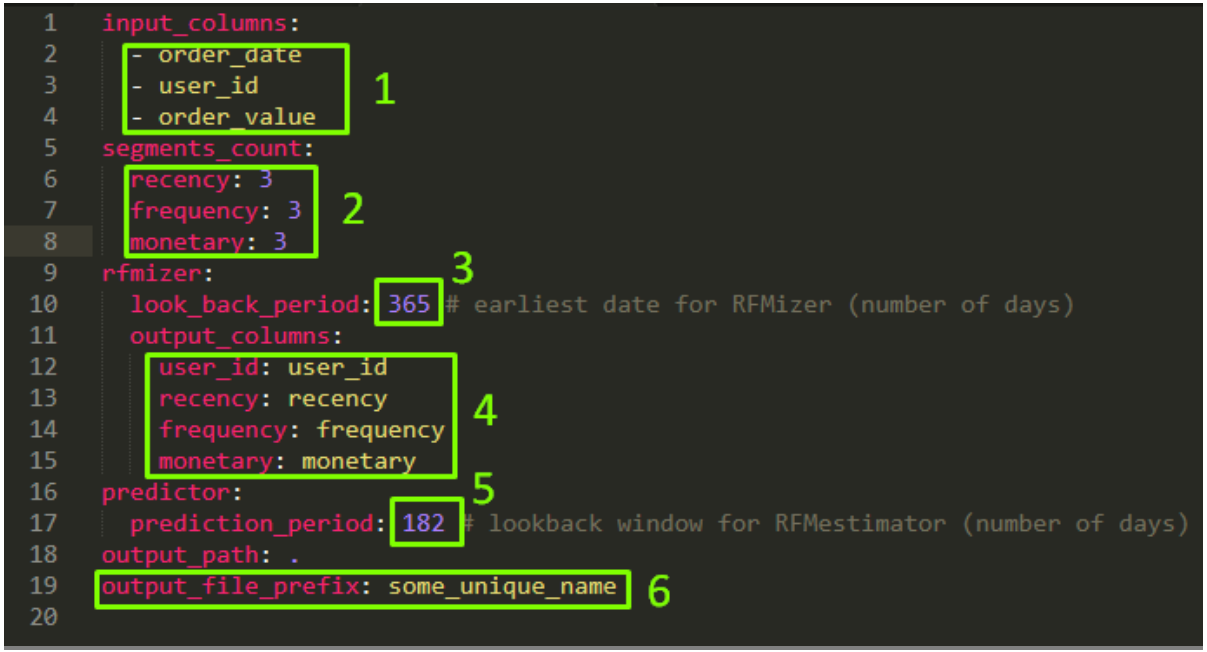
order_date,user_id,order_value
2018-11-16,49,160418
2018-03-30,49,87424
2019-09-09,67,228445
2018-12-07,67,140733
2019-07-23,67,5143
2019-05-23,67,44158
2019-02-05,67,130638
2019-08-16,67,244325
2018-12-24,67,212708
2018-10-03,67,23277
2018-04-06,67,240145
2018-03-22,67,198593
2018-09-05,67,83526
2018-12-15,67,189733

MediaNation
Challenge the future

Настройка RFMizer-config.yaml

Вы установили RFMizer и распаковали файлы в папку (еще раз убедитесь в том, что Вы сохранили csv файл в эту же папку).

Далее необходимо открыть файл “config.yaml” через Sublime Text.



MediaNation
Challenge the future

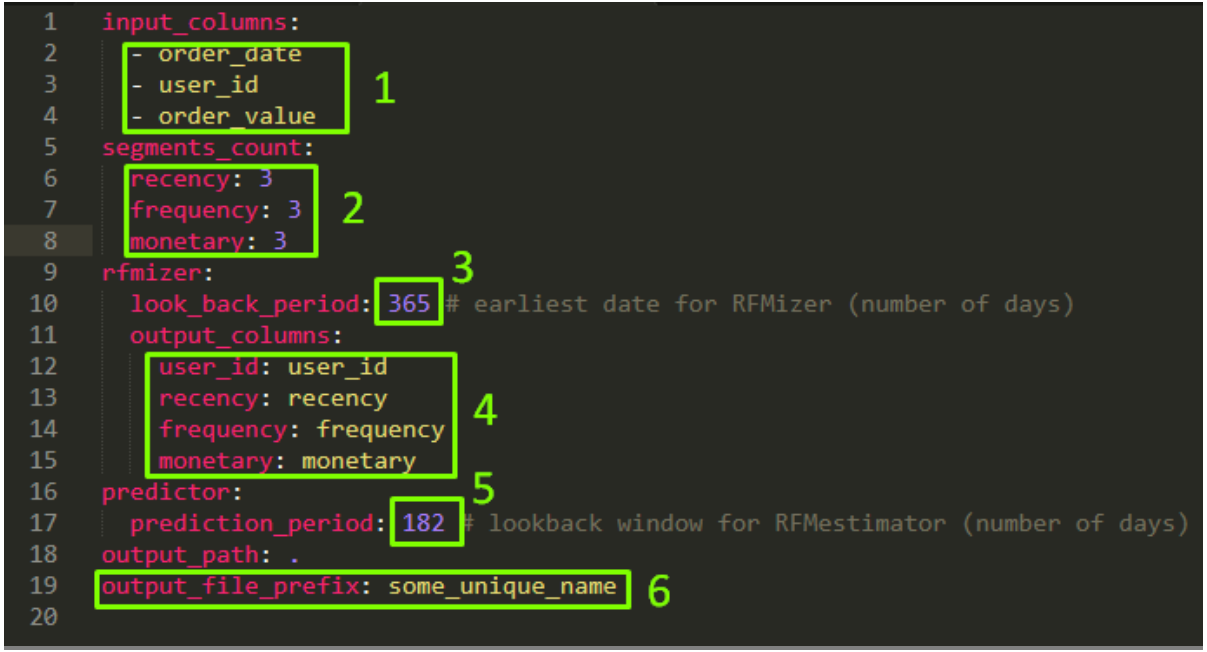
Настройка RFMizer

Блоки №1 и №4:
дополните данными, которые перечислены в csv файле (п: phone, email и т.д.)

Блок №2:
укажите, на какое количество сегментов Вы хотите разделить аудиторию относительно R-F-M (обычно 3/4/5)

Блок №3:
Вы указываете период, данные которого будут использованы в анализе

Блок №5: Вы указываете период предсказания (алгоритм спрогнозирует поведение выбранной аудитории)



Блоки №6:
Укажите уникальное название. Под этим названием будут созданы файлы с результатом RFM анализа

MediaNation
Challenge the future

Запуск скрипта: инструкция

Подготовительные работы закончены. Теперь можно запускать скрипт. Соблюдайте последовательность ниже. Пример будет на следующем слайде.

MAC

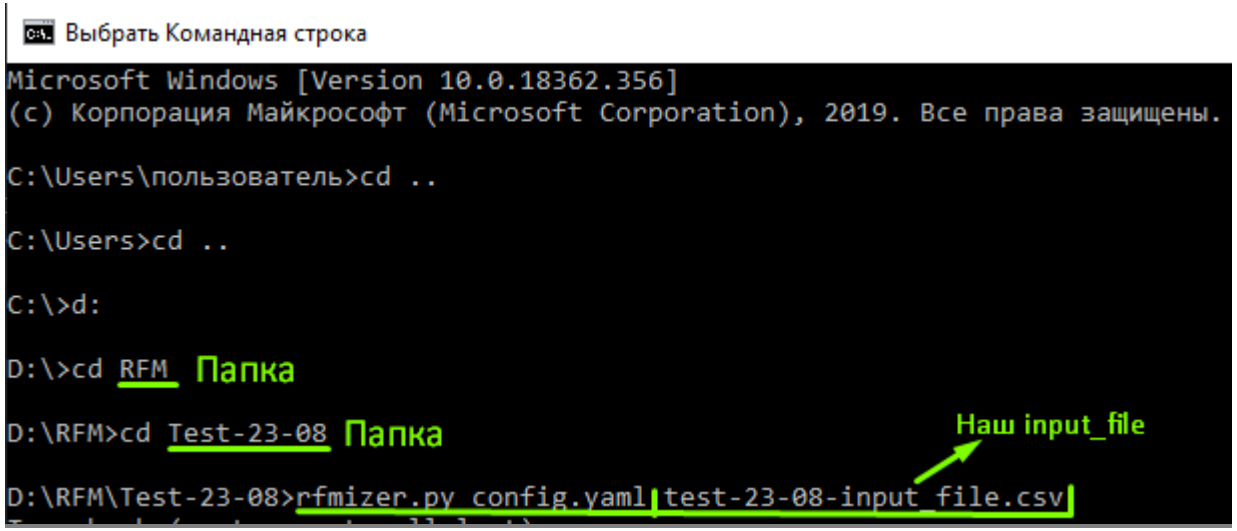
1. Заходим в командную строку
2. Вы увидите свое текущее расположение. Вам необходимо в командной строке попасть в папку, где лежит csv файл и файлы RFMizer.
3. При помощи команды **pwd** переходите в нужную папку.
4. Когда вы окажетесь в нужной директории, необходимо будет ввести:
   python rfmizer.py config.yaml input-file
5. Теперь откройте папку, где лежит csv файл и файлы RFMizer. После выполненных операций внутри появилось 3 новых файла:
   - **input_file_borders.csv**
   - **input_file_mapping.csv**
   - **input_file_ratios.csv**

WINDOWS

1. Заходим в командную строку
2. Вы увидите свое текущее расположение. Вам необходимо в командной строке попасть в папку, где лежит csv файл и файлы RFMizer.
3. При помощи команды **cd ..** переходите в нужную папку.
4. Когда вы окажетесь в нужной директории, необходимо будет ввести:
   python rfmizer.py config.yaml input-file
5. Теперь откройте папку, где лежит csv файл и файлы RFMizer. После выполненных операций внутри появилось 3 новых файла:
   - **input_file_borders.csv**
   - **input_file_mapping.csv**
   - **input_file_ratios.csv**

MediaNation
Challenge the future

Запуск скрипта: пример WINDOWS



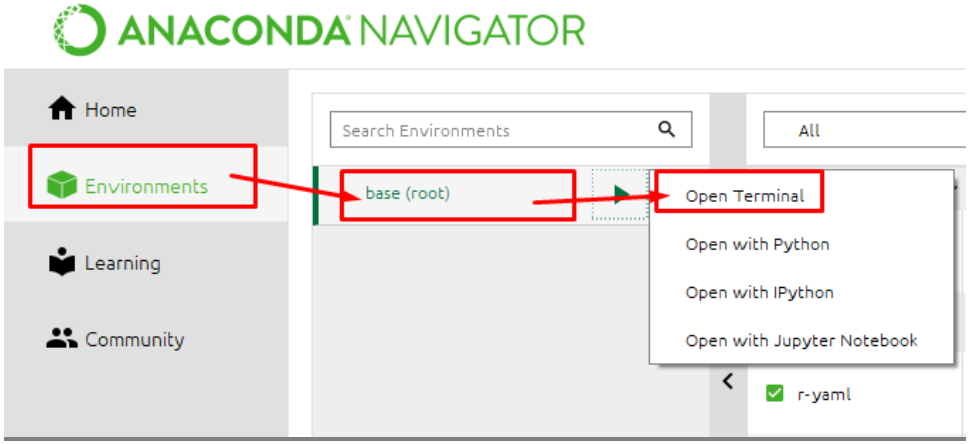
В конце строки нажимаем: ENTER

Результат: в папке Test-23-08 появилось 3 файла

MediaNation
Challenge the future

Дополнительно

Если не получится запустить скрипт, то совершите аналогичные действия, но через командную строку Anaconda.
И как результат в папке появится 3 файла.



MediaNation
Challenge the future

Разбор полученных файлов

1. - файл содержит числовые значения вычисленных границ между сегментами.

2. - файл, в котором каждая строка состоит из четырех обязательных полей (user_id,recency,frequency,monetary) и произвольного количества необязательных полей, которыми могут быть, например, дата рождения клиента, email клиента и др. **Именно с ним мы будем работать дальше.**

MediaNation
Challenge the future

Разбор полученных файлов

3. - файл содержит большое количество строк равное количество сочетаний сегментов R, F и M, то есть количество строк получается путем перемножения сегментов R*F*M+1, то есть в задаче, где мы делим сегменты по 3, мы получим 3*3*3+1=28 строк, где последняя 1 имеет значение 0,0,0

*Сочетание 0,0,0 присваивается тем пользователям, которые не вошли в выборку. Например, из CRM у Вас данные за 3 года, а анализ Вы решили производить за 365 дней. Соответственно, те пользователи, действия которых НЕ вошли в 365, будут иметь сочетание 0,0,0.

Пример нескольких строк из файла:

```
1  frequency,monetary,recency,bid ratio
2  0,0,0,0.47034588470828095
3  1,2,1,0.4678753107045928
4  1,2,3,0.17468760997116
5
```

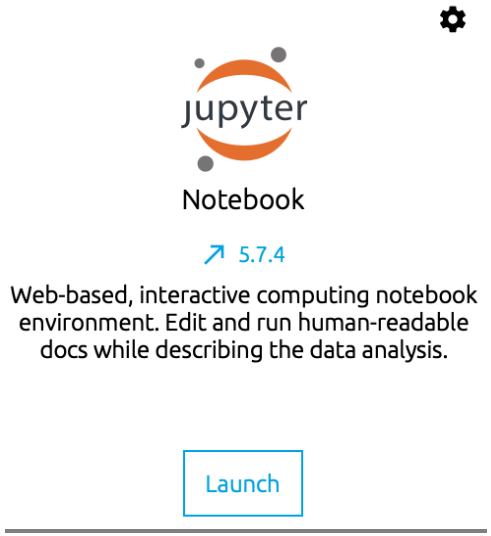
Bid ratio является мультипликатором ставки для каждого из пересечений сегментов пользователей, который создает скрипт RFMizer.
Это значит, что если у вас ранее была ставка для закупки трафика, которая равна 10 рублям, то теперь, после того, как вы стали закупать отдельно сегменты, для сегмента 1,2,1 вы должны умножить ставку 10 рублей на мультипликатор 0,47 (значение округлено до сотых).

MediaNation
Challenge the future

Подготовка к визуализации

Визуализация данных

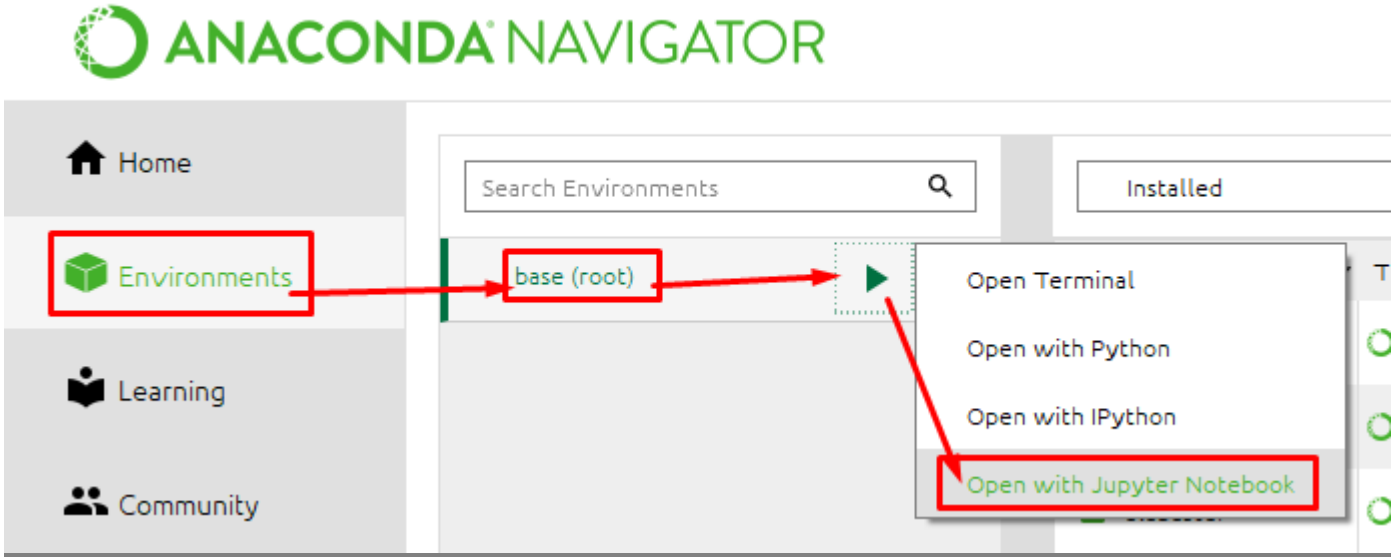
Заходим в Anaconda и устанавливаем (launch) блокнот Юпитер.
Далее либо он сам откроется в браузере, либо Вам придется его открыть.



MediaNation
Challenge the future

Визуализация данных

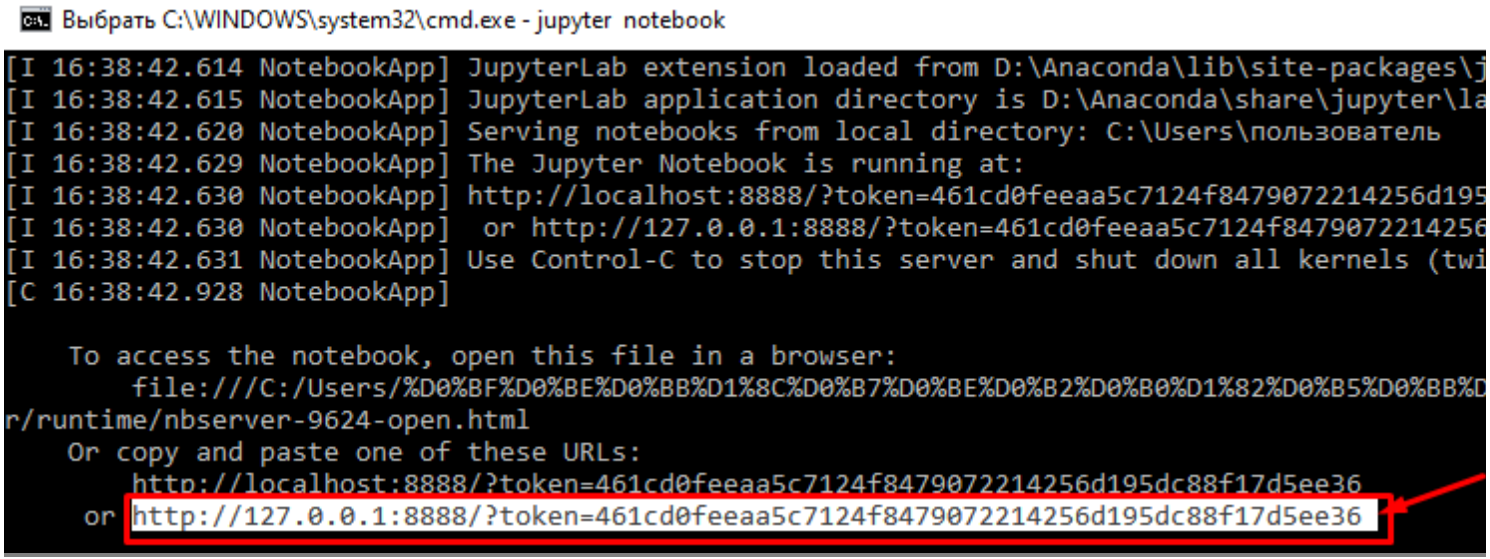
Открываем блокнот Юпитер самостоятельно:



MediaNation
Challenge the future

Визуализация данных

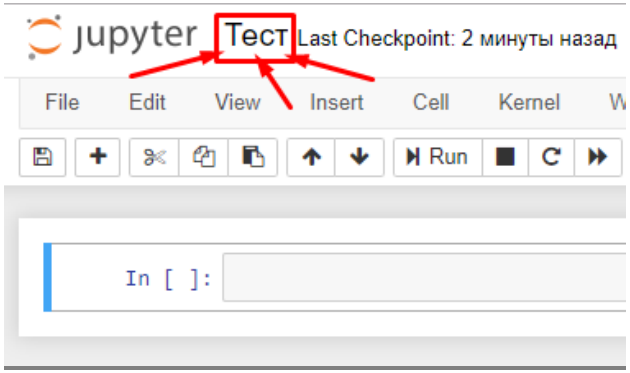
После запуска блокнота откроется командная строка. Вам необходимо вставить в поисковую строку браузера предлагаемую ссылку, чтобы открылся блокнот Юпитер:



MediaNation
Challenge the future

Визуализация данных

В блокноте создаете новый файл и присваиваете ему название:



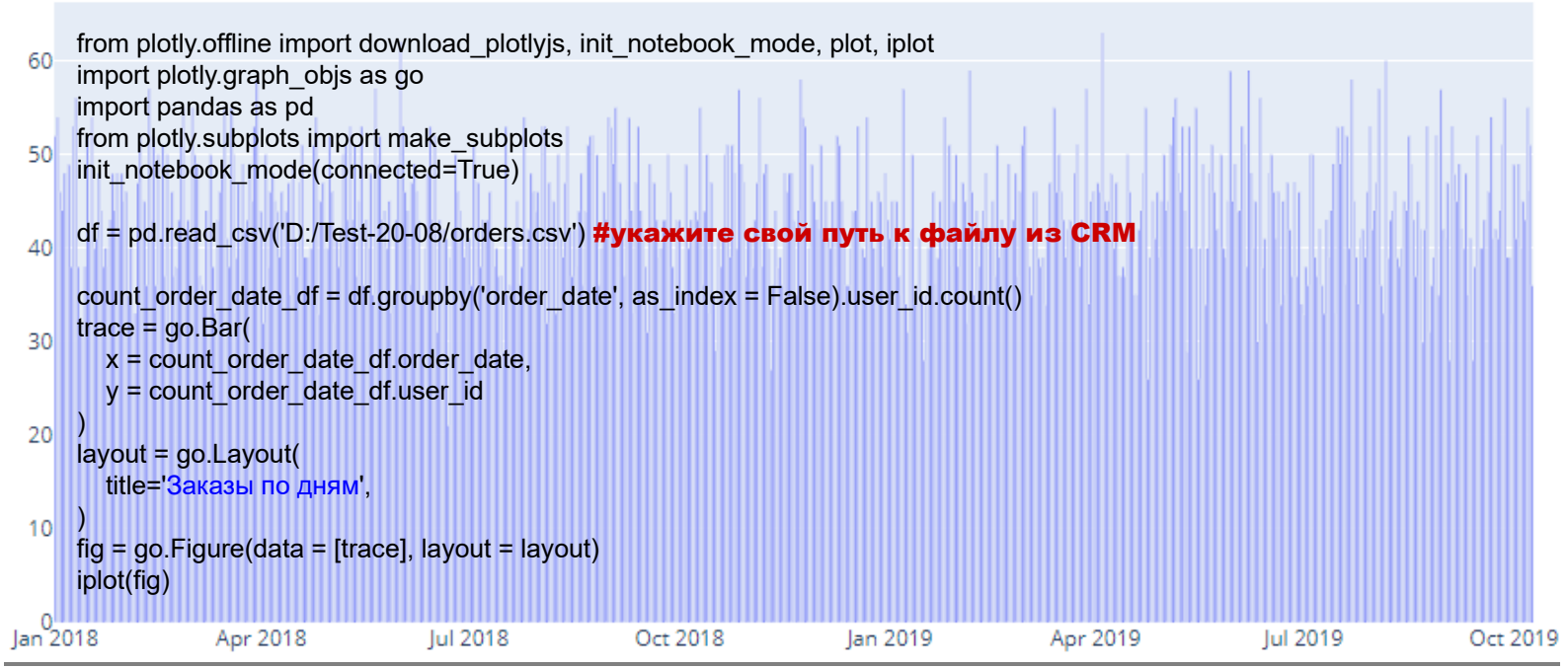
MediaNation
Challenge the future

Визуализация данных

Визуализируем “Заказы по дням”

Копируем код ниже и вставляем в Юпитер. Далее нажимаем ctrl+shift+enter. В качестве результата отобразится график “Заказы по дням”.

```
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot
import plotly.graph_objs as go
import pandas as pd
from plotly.subplots import make_subplots
init_notebook_mode(connected=True)

df = pd.read_csv('D:/Test-20-08/orders.csv') #укажите свой путь к файлу из CRM

count_order_date_df = df.groupby('order_date', as_index = False).user_id.count()
trace = go.Bar(
    x = count_order_date_df.order_date,
    y = count_order_date_df.user_id
)
layout = go.Layout(
    title='Заказы по дням',
)
fig = go.Figure(data = [trace], layout = layout)
iplot(fig)
```

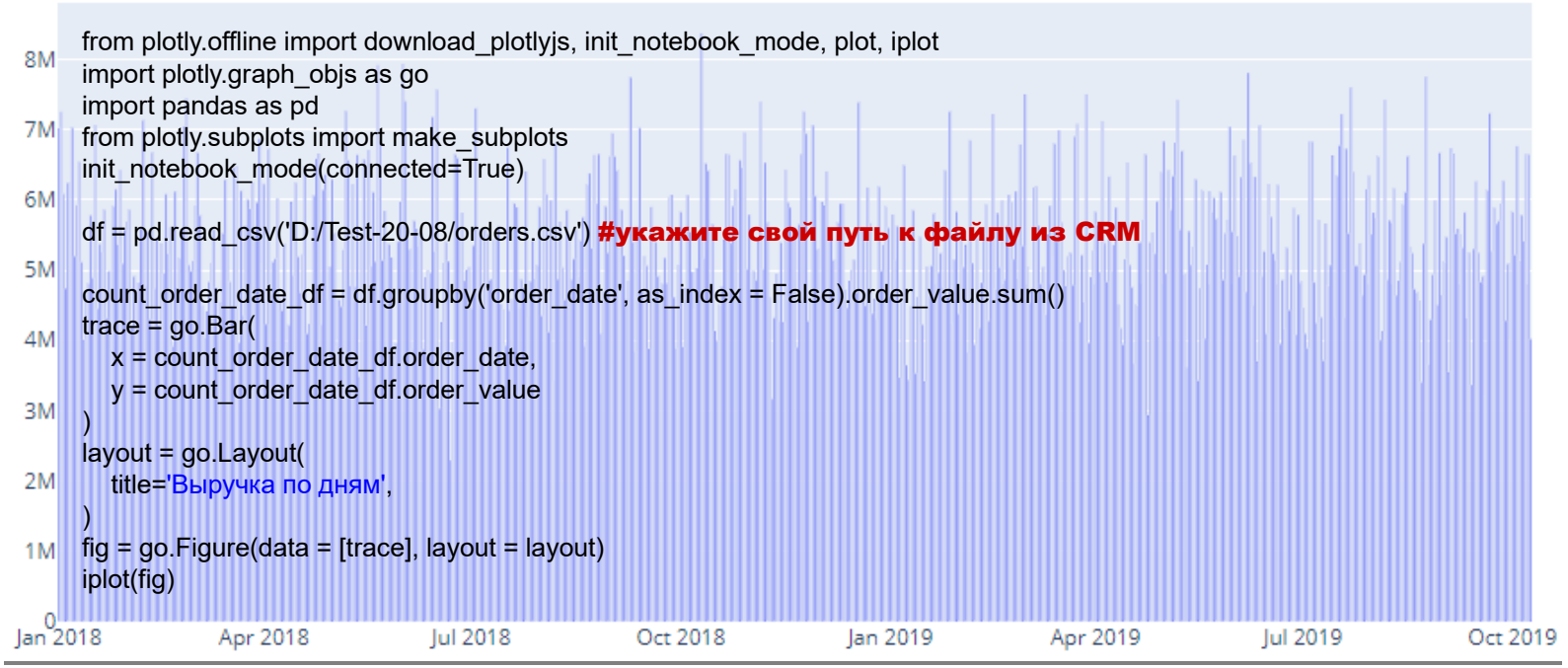


MediaNation
Challenge the future

Визуализируем “Выручка по дням”

В следующую строку копируем код (см.ниже) и вставляем в Юпитер. Далее нажимаем ctrl+shift+enter. В качестве результата отобразится график “Выручка по дням”.

```
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot
import plotly.graph_objs as go
import pandas as pd
from plotly.subplots import make_subplots
init_notebook_mode(connected=True)

df = pd.read_csv('D:/Test-20-08/orders.csv') #укажите свой путь к файлу из CRM

count_order_date_df = df.groupby('order_date', as_index = False).order_value.sum()
trace = go.Bar(
    x = count_order_date_df.order_date,
    y = count_order_date_df.order_value
)
layout = go.Layout(
    title='Выручка по дням',
)
fig = go.Figure(data = [trace], layout = layout)
iplot(fig)
```



MediaNation
Challenge the future

Визуализируем “Заказы и выручка по дням”

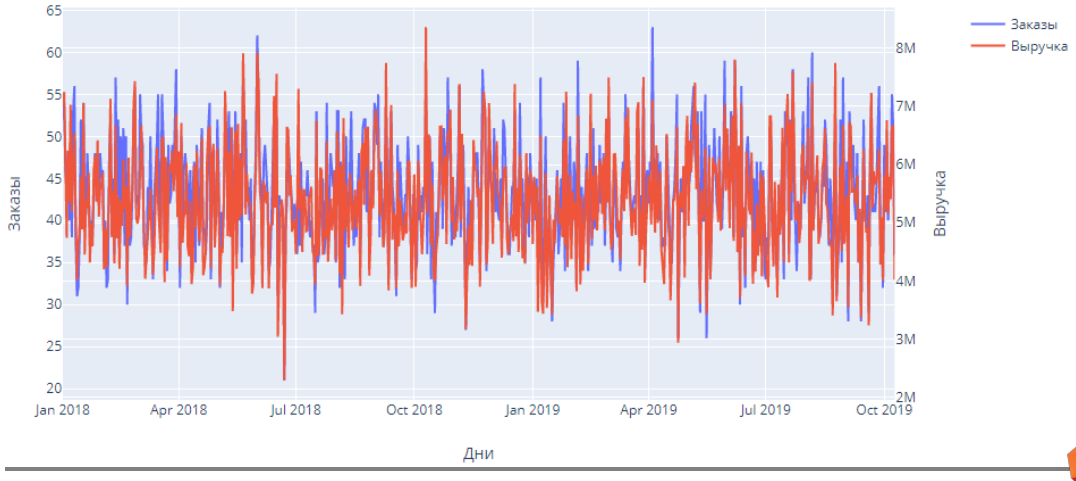
В следующую строку копируем код (см.ниже) и вставляем в Юпитер. Далее нажимаем ctrl+shift+enter. В качестве результата отобразится график “Заказы и выручка по дням”.

#Код

```
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot
import plotly.graph_objs as go
import pandas as pd
from plotly.subplots import make_subplots
init_notebook_mode(connected=True)

df = pd.read_csv('D:/Test-20-08/orders.csv') #укажите свой путь к файлу из CRM

compare_userid_df = df.groupby('order_date', as_index = False).user_id.count()
compare_value_df = df.groupby('order_date', as_index = False).order_value.sum()

fig = make_subplots(specs=[[{"secondary_y": True}]])

fig.add_trace(
    go.Scatter(
    x = compare_userid_df.order_date,
    y = compare_userid_df.user_id,
    name = 'Заказы'
    ),
    secondary_y=False
)

fig.add_trace(
    go.Scatter(
    x = compare_value_df.order_date,
    y = compare_value_df.order_value,
    name = 'Выручка'
    ),
    secondary_y=True
)

fig.update_layout(
    title_text="Заказы и выручка по дням"
)
```

#Продолжение кода

```
# Set x-axis title
fig.update_xaxes(title_text="Дни")

# Set y-axes titles
fig.update_yaxes(title_text="Заказы", secondary_y=False)
fig.update_yaxes(title_text="Выручка", secondary_y=True)

fig.show()
```

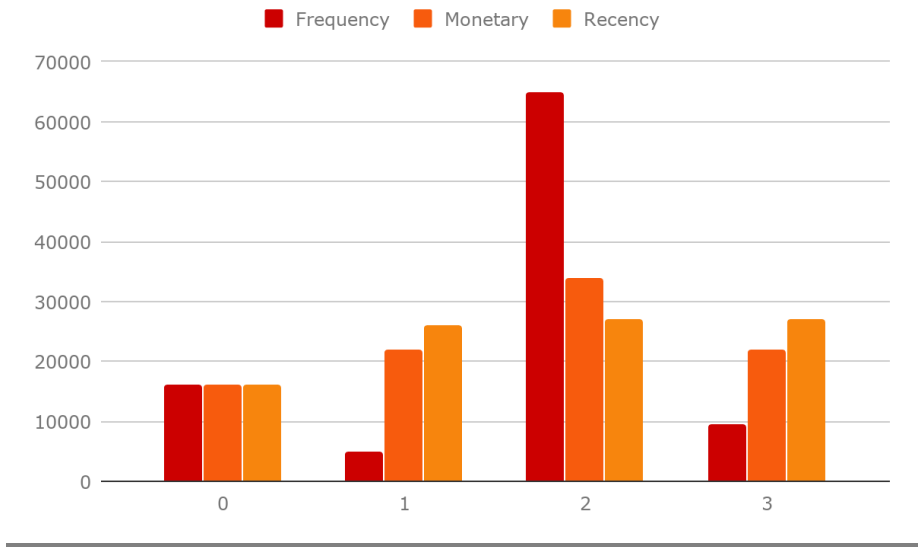


MediaNation
Challenge the future

Визуализируем “Количество пользователей внутри RFM”

#Код

```
df = pd.read_csv('D:/Test-20-08/input_file_mapping.csv') #укажите свой путь к файлу mapping


count_frequency_df = df.groupby('frequency', as_index = False).user_id.count()
count_monetary_df = df.groupby('monetary', as_index = False).user_id.count()
count_recency_df = df.groupby('recency', as_index = False).user_id.count()

trace_frequency = go.Bar(
    x = count_frequency_df.frequency,
    y = count_frequency_df.user_id,
    name = 'frequency'
)
trace_monetary = go.Bar(
    x = count_monetary_df.monetary,
    y = count_monetary_df.user_id,
    name = 'monetary'
)

trace_recency = go.Bar(
    x = count_recency_df.recency,
    y = count_recency_df.user_id,
    name = 'recency'
)

layout = go.Layout(
    title='Количество пользователей внутри RFM',
)

fig = go.Figure(data = [trace_frequency, trace_monetary, trace_recency], layout = layout)
iplot(fig)
```



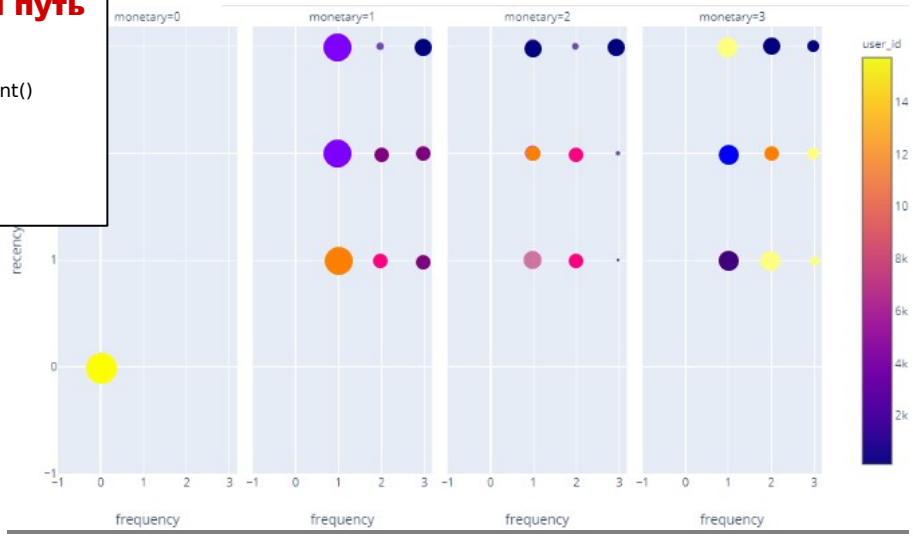
MediaNation
Challenge the future

Визуализируем “объединенный график”

График показывает, сколько пользователей находится в каждом из сегментов RFM. Эти данные нужны для того, чтобы в дальнейшем понять, какие RFM сегменты нужно объединить между собой, чтобы создать загружаемый файл для Яндекс.Аудиторий или любых других сервисов.

#Код

```
df = pd.read_csv('D:/Test-20-08/input_file_mapping.csv') #укажите свой путь к файлу mapping
import plotly.express as px
tips = df.groupby(['frequency', 'recency','monetary'], as_index = False).user_id.count()
fig = px.scatter(tips, x="frequency", y="recency", color="user_id", size='user_id',
facet_col="monetary",
          render_mode="webgl")
fig.update_yaxes(dtick=1)
fig.show()
```



MediaNation
Challenge the future

Считаем количество пользователей

Далее необходимо добавить в файле mapping столбец RFM, благодаря которому сможем узнать кол-во по каждому сегменту.

#Код

```
df = pd.read_csv('D:/adamas/input_file_mapping.csv') #укажите свой путь к файлу mapping

df['RFM'] = df[['recency', 'frequency', 'monetary']].apply(lambda x: '-'.join(map(str, x)), axis=1)
df.head()
```

#Код

```
df_filter = df.groupby('RFM')['user_id'].count()
print (df_filter)
```

MediaNation
Challenge the future

Считаем кол-во пользователей: результат

Полученная таблица позволяет увидеть картину целиком.

Также важно учесть, что в дальнейшем мы будем загружать сегменты в Яндекс.Аудитории, а для этого в загружаемом файле должно быть минимум 1000 строк.

Следовательно, какие-то сегменты можно будет загрузить по-отдельности, а какие-то потребуется объединить.

```
df_filter = df.groupby('RFM')['user_id'].count()
print (df_filter)
```

```
RFM
0-0-0    [illegible]55
1-1-1     5[illegible]
1-1-2    [illegible]06
1-1-3     2[illegible]
1-2-1    [illegible]48
1-2-2    [illegible]3
1-2-3      6[illegible]
1-3-1    [illegible]1
1-3-2    [illegible]5
1-3-3    [illegible]1
2-1-1     4[illegible]
2-1-2     4[illegible]
2-1-3    [illegible]2
2-2-1    [illegible]
2-2-2    [illegible]3
2-2-3    [illegible]
2-3-1    [illegible].1
2-3-2    [illegible]3
2-3-3    [illegible]84
3-1-1    [illegible]54
3-1-2    [illegible]5
3-1-3    [illegible]4
3-2-1    [illegible]3
3-2-2    [illegible]6
3-2-3    [illegible]40
3-3-1    [illegible]3
3-3-2    [illegible]0
3-3-3    [illegible]1
Name: user_id, dtype: int64
```

MediaNation
Challenge the future

Выбор сегментов для Яндекс.Аудиторий

Выбор сегментов для загрузки в Яндекс.Аудитории

Благодаря результатам RFM анализа для дальнейшей работы необходимо выбрать наиболее интересные сегменты, основываясь на специфике бизнеса, на поведении потребителей и т.д.

Всегда самый лучший сегмент - это 3-3-3 (клиенты, которые покупают часто, на большие суммы и покупали недавно). Но не стоит ограничиваться. Существуют и другие интересные сегменты.

Например, если Вы занимаетесь продажей ювелирных украшений, то для Вас имеют значения сегменты потребителей и их объединения, представленные на рисунке справа.

№	Сегмент R-F-M	Описание сегмента
1	3-3-3	Покупали недавно Высокая частота покупок Высокий чек
2	3-2-3	Покупали недавно Средняя частота покупок Высокий чек
3	3-2-2	Покупали недавно Средняя частота покупок Средний чек
4	1-3-3 2-3-3 3-3-3	Покупали недавно + относительно недавно + давно Высокая частота покупок Высокий чек
5	1-2-3 2-2-3 3-2-3	Покупали недавно + относительно недавно + давно Средняя частота покупок Высокий чек
6	1-3-2 2-3-2 3-3-2	Покупали недавно + относительно недавно + давно Высокая частота покупок Средний чек

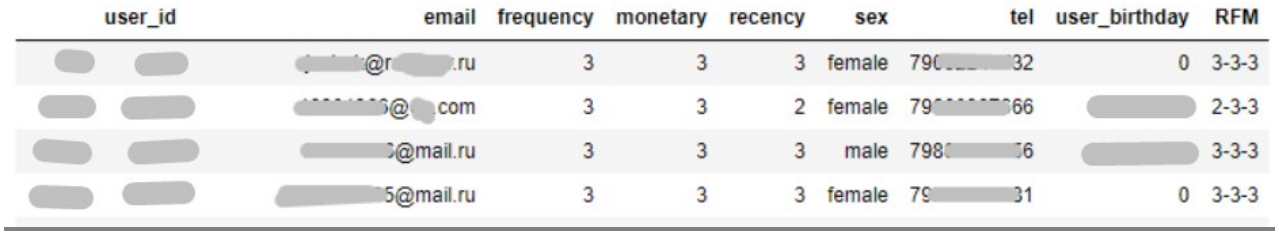
MediaNation
Challenge the future

Реализуем фильтр по выбранным сегментам

Делаем фильтрацию строк, которые содержат значения из приведенного ниже списка в конкретном столбце.

#Код

```
df = df[df['RFM'].isin(['1-3-3', '2-3-3', '3-3-3']) ] #выбрали и объединили 3 сегмента
df.head(100) #отображаем для просмотра первые 100 строк
```



user_id	email	frequency	monetary	recency	sex	tel	user_birthday	RFM
[redacted]	[redacted].ru	3	3	3	female	[redacted]32	0	3-3-3
[redacted]	[redacted].com	3	3	2	female	[redacted]66	[redacted]	2-3-3
[redacted]	[redacted]@mail.ru	3	3	3	male	[redacted]6	[redacted]	3-3-3
[redacted]	[redacted]5@mail.ru	3	3	3	female	[redacted]31	0	3-3-3

MediaNation
Challenge the future

Загрузка сегментов в Яндекс.Аудитории

Для сохранения отдельных сегментов, либо их совокупностей необходимо в новой строке написать код:

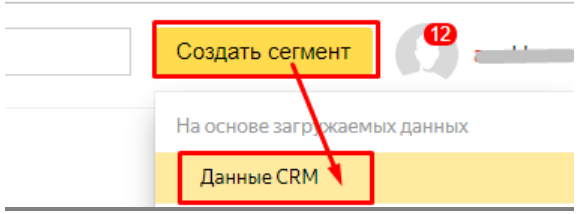
df.to_csv('D:/adamas/final-rfm.csv') **#Вы указываете место сохранения и название файла.**

Далее повторяем фильтрацию и сохранение файла по нужным сегментам и их объединениям. В результате, у Вас должно получиться от 3 и более файлов.

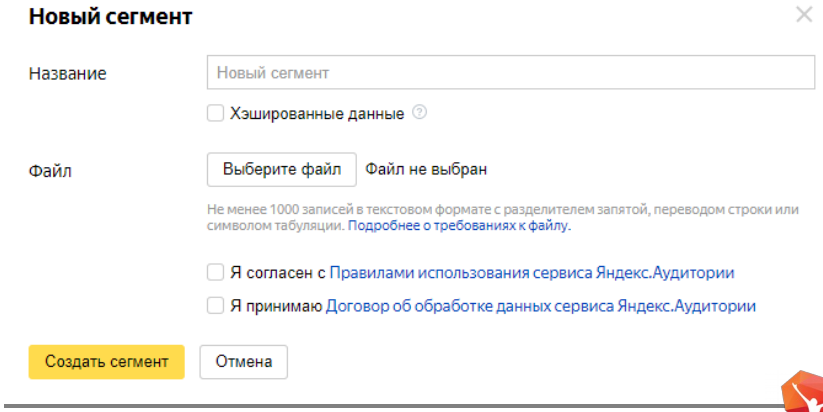
MediaNation
Challenge the future

Загрузка файлов в Яндекс.Аудитории

1. Открываем https://audience.yandex.ru/ (проходим регистрацию)
2. Создаем сегмент на базе данных из CRM:

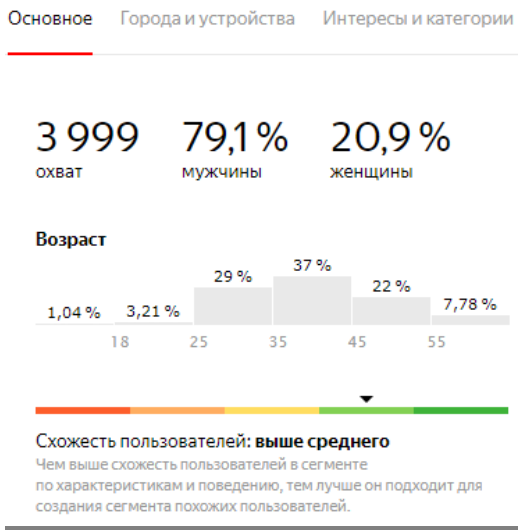


3. При загрузке данных, ознакомьтесь с требованиями к файлам.
4. Поменяйте в своих файлах название user_id на external_id.
5. Создайте требуемое Вам кол-во сегментов.



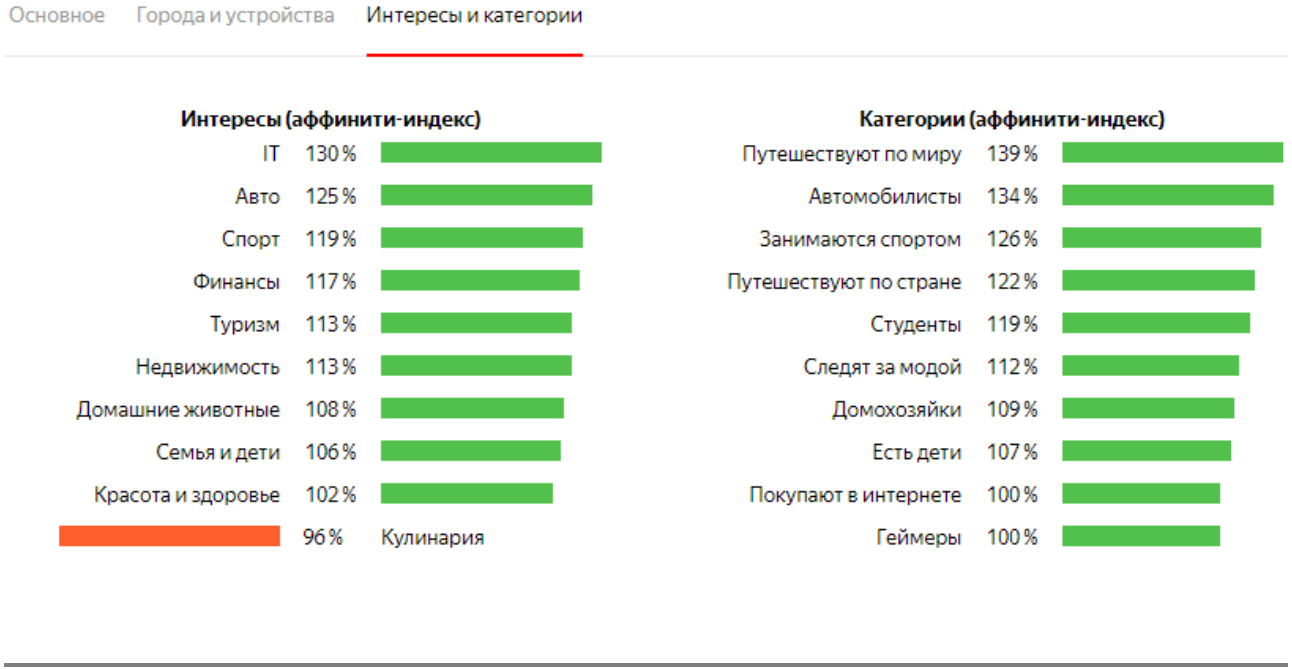
MediaNation
Challenge the future

Пример результата обработки данных



MediaNation
Challenge the future

Пример результата обработки данных

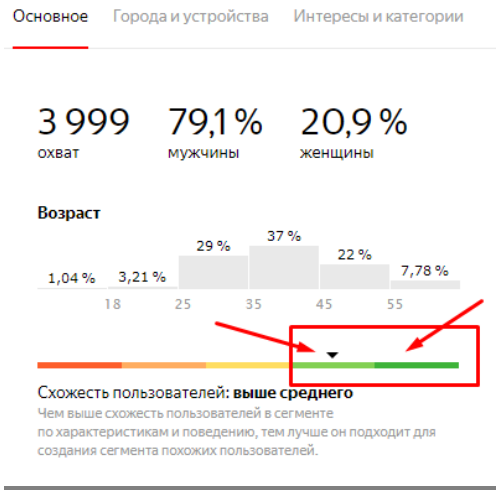


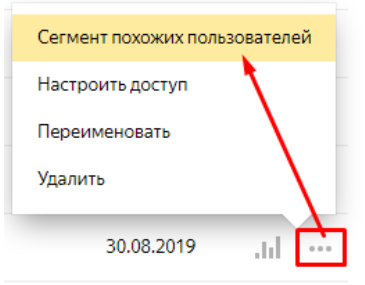
MediaNation
Challenge the future

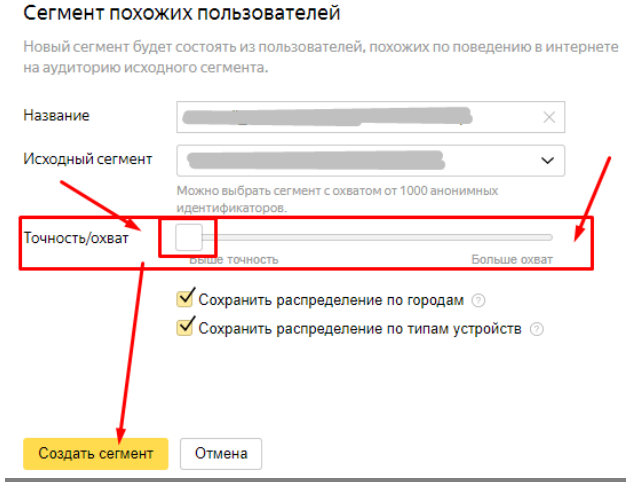
Создание Look-a-like

Создание Look-a-like аудиторий

Если схожесть пользователей загруженного сегмента выше среднего/высокая (т.е. пользователи сегмента гомогенны), тогда на базе этой аудитории имеет смысл создать Look-a-like аудиторию.







MediaNation
Challenge the future

Создание Look-a-like аудиторий

При создании Look-a-like была выбрана максимальная точность против охвата (определяете в зависимости от цели).

*Вы сами решаете, насколько пользователи в новом сегменте должны быть похожи на пользователей из исходного сегмента. Помните: чем точнее будет соответствие между пользователями в новом и в исходном сегменте, тем меньше анонимных идентификаторов пользователей будет добавлено в сформированный сегмент.

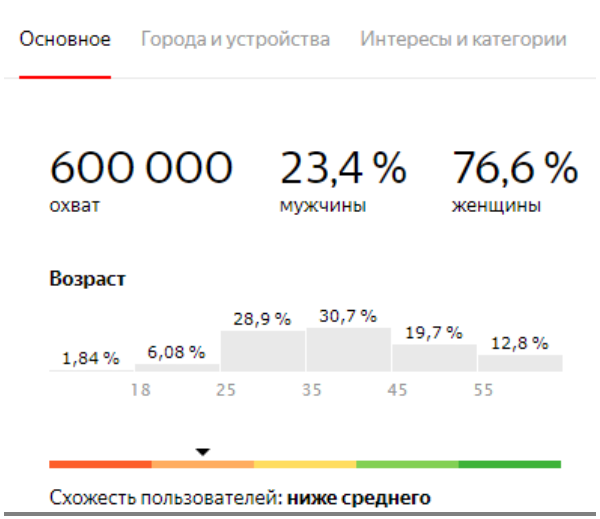
В итоге Вы создаете несколько Look-a-like аудиторий исходя из результатов гомогенности базовых аудиторий.

Подробную информацию можете прочитать в справке Яндекса

MediaNation
Challenge the future

Пример результата Look-a-like

При создании Look-a-like была выбрана максимальная точность против охвата (определяете в зависимости от цели).



MediaNation
Challenge the future

MediaNation
Challenge the future

Спасибо за внимание!

Телефон: +7 (495) 215-10-97
E-mail: comeon@medianation.ru
Адрес: Москва, Малая Полянка, д. 2, 119180