POLY//ATICA

БИЗНЕС-АНАЛИТИКА
КАК НЕ ЗАХЛЕБНУТЬСЯ В ПОТОКЕ ДАННЫХ И ИЗВЛЕЧЬ ИЗ НИХ РЕАЛЬНУЮ ЦЕННОСТЬ

С. Куликовский, Генеральный директор
В. Майоров, Ведущий консультант

НЕМНОГО ЦИФР

- Каждый день накапливается 2,3 млн Гб новых данных (IBM)
- К 2020 году будет создано 43 трлн Гб данных (IBM)
- 90% всех данных накоплено только за последние два года (Forbes)
- 1.7 Мб – новая информация, создаваемая ежесекундно каждым человеком (Grazziti)
- 6 ч 42 м – среднее время, проводимое пользователями в интернете в день (The Next Web)
- 44 трлн Гб – объем глобального ИТ-трафика, прогнозируемый к 2020 году (Grazziti)
- 50 млрд – количество подключенных к сети интернет интеллектуальных устройств в мире к 2020 году, собирающих, анализирующих и обменивающихся данными (Cisco)

Источник: https://financesonline.com/big-data-statistics/

POLY//ATICA

НОВАЯ ЭКОНОМИКА

- Uber & Яндекс.Такси
- Facebook & VK
- Airbnb & HomeAway
- Amazon & Alibaba & Ozon
- eBay & Avito
- DeliveryClub & Яндекс.Еда
- Белка & Яндекс.Драйв

POLY//ATICA

DATA-DRIVEN КОМПАНИИ

- Постоянный поиск новых возможностей
- Повышение скорости принятия решений
- Увеличение вариативности процессов
- Снижение рисков
- Снижение издержек

POLY//ATICA

ЧТО ЕСТЬ BIG DATA?

VOLUME	VARIETY	VELOCITY	VERACITY	VALUE
Объемы сгенерированных и накопленных данных	Разнообразие форматов и представлений данных	Скорость появления новых и устаревания ранее полученных данных	Изменчивость данных во времени	Качество и достоверность данных

POLY//ATICA

ДАННЫЕ – ИНФОРМАЦИЯ - ЗНАНИЕ

POLY//ATICA

ЗАЧЕМ НУЖНА БИЗНЕС-АНАЛИТИКА?

«Я знаю, что половина моего рекламного бюджета расходуется впустую, я только не знаю какая...»

Джон Ванамейкер, создатель первой в мире сети универмагов

Мы можем анализировать весь объем данных о клиентах и товарах, таргетировать аудиторию и эффективно персонализировать предложения, экономя время и деньги компаний

POLY//ATICA

ЗАДАЧИ АНАЛИТИКА

- сбор данных
- подготовка данных к анализу (выборка, очистка, сортировка)
- поиск закономерностей в информационных наборах
- визуализация данных для быстрого понимания имеющихся результатов и будущих тенденций
- формулирование гипотез по улучшению конкретных бизнес-метрик за счет изменения других показателей

POLY//ATICA

ТРИ КИТА АНАЛИТИКИ ДАННЫХ

1

Правильно сформулировать вопрос

100%

успеха

Любая аналитика должна начинаться с правильно поставленного вопроса и моделирования – как будет использован ответ на него / как полученный результат поможет получить бизнес выгоды?

2

Правильно выбрать источники

75%

успеха

Для ответа на вопрос нужно определить на какие данные при ответе на него можно опираться, а на какие нет.

3

Правильно интерпретировать результат

50%

успеха

Получая результат требуется его правильно интерпретировать и что самое главное – превратить в конкретные действия.

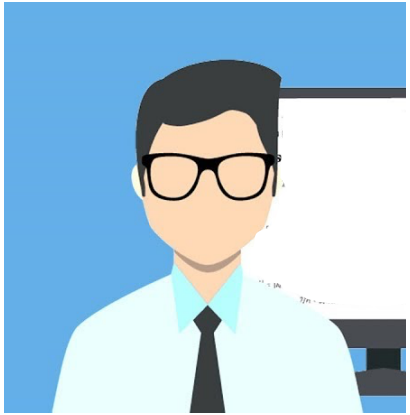
POLY//ATICA

ДВА ТИПА АНАЛИТИКОВ

Бизнес-аналитик

Понимает суть происходящего – владеет контекстом

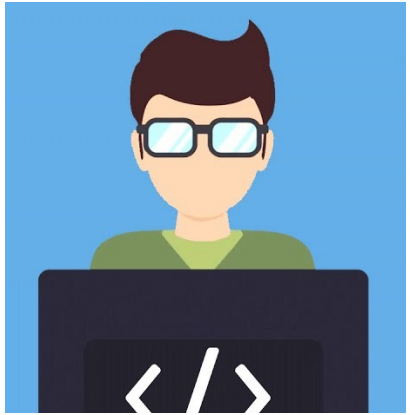
Использует Excel и **BI инструменты**



Data Scientist

Владеет математическим методами работы с данными

Использует R, Python, SQL



POLY//ATICA

ТИПОВЫЕ БАРЬЕРЫ

- Чистота & качество данных
- Интеграция данных
- Пропуски и аномалии в данных
- Сильный разброс значений
- Сложность инструментария для обработки

POLY//ATICA

ПРИМЕР # 1 КАЧЕСТВО ДАННЫХ

Субъект Российской Федерации	Производитель	Наименование продукции	Производственная мощность выпуска продукции в сутки, шт.	Объем фактически произведенной продукции, шт.	Остаток на складе, шт.
Московская область	ИП ААА	маска из СМС	40 000	75 000	289 000
Московская область	ООО "ААА"	маска из СМС	40 000	45 000	225 500
г. Москва	ООО "БББ"	маска из СМС	0* продукция не производится в связи с поиском поставщиков сырья	149 000	407 000
город Москва	ООО "ВВВ"	Маски медицинские SMS	5 200 тыс	5 200 тыс	нет
Московская обл.	ООО "ГГГ"	Маски медицинские SMS, тыс. шт.	1 500	0	6 250
Пензенская обл.	ООО "ДДД"	Маски медицинские (SMS, иные производства), тыс. шт.*	1 200	1 350	3 630
Пермский край	ООО "ЕЕЕ"	Маски медицинские (SMS, производители подгузников), тыс. шт.*	15 000	43 000	0

POLY//ATICA

ПРИМЕР # 2 ИНТЕГРАЦИЯ ДАННЫХ

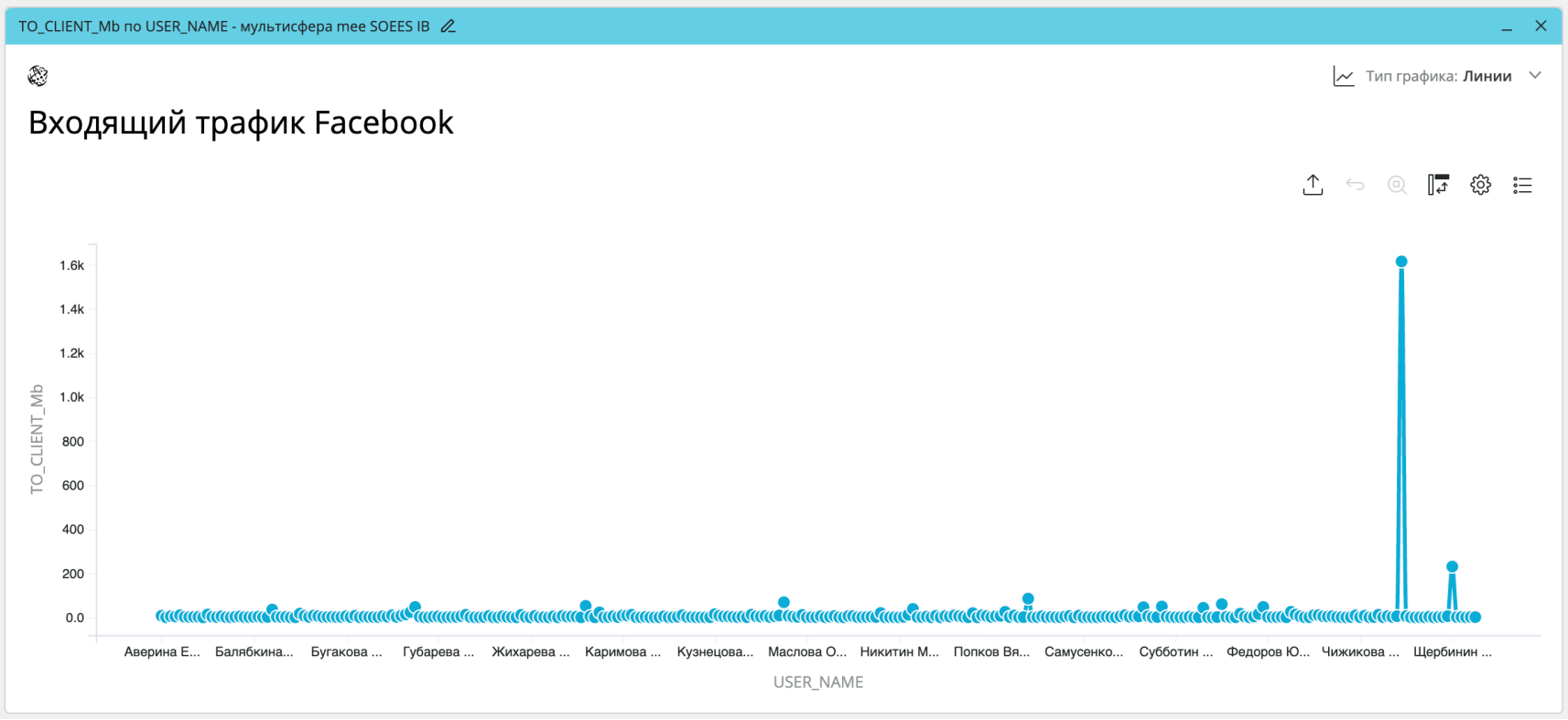
МСС категории

ОКВЭД

Аналитика отраслевых показателей

POLY//ATICA

ПРИМЕР # 3 ПРОПУСКИ И АНОМАЛИИ



POLY//ATICA

ПРИМЕР # 4 РАЗБРОС ЗНАЧЕНИЙ

POLY//ATICA

АНАЛИТИКА И ИНСАЙТЫ

- Многомерная аналитика
- Ad-hoc
- Автоматизация аналитических операций

МИФЫ ПРО АНАЛИТИКУ

- Аналитика требует существенных инвестиций
- Аналитика исключает присущую людям предвзятость
- Лучшие алгоритмы — залог успеха
- Алгоритмы безотказны
- Наука о данных — искусство для избранных
- Аналитика занимает слишком много времени
- Самое сложное — технологии

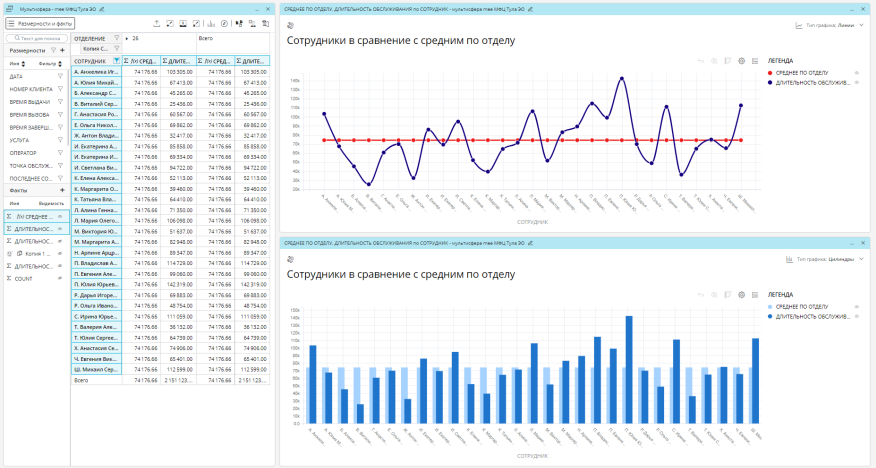
POLY//ATICA

ОТЧЕТЫ ПО ЗАПРОСУ (AD-HOC) В POLYMATICA

Polymatica в режиме реального временя позволяет создать отчет по запросу или нестандартный отчет в интерфейсе, не требующем программирования или написания SQL-запросов

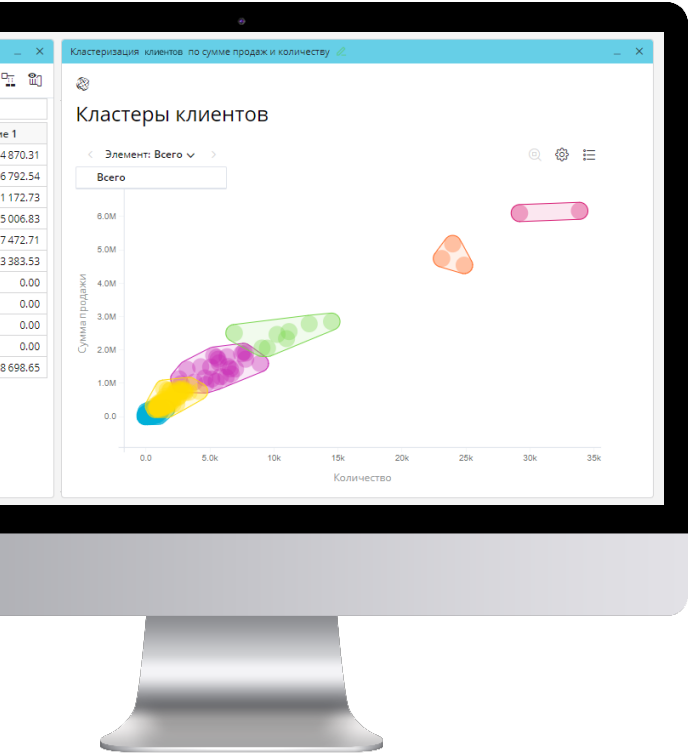
Инструменты:

- Сводные таблицы
- Построение отчетов с возможностью сравнения периодов – день ко дню, месяц к месяцу, и т.п.
- 17 видов графиков
- Поддержка аналитических, математических, статистических функции и логических операций для создания простых и сложных показателей
- Кластеризация и другие методы Data Mining



POLY//ATICA

ИНТЕЛЛЕКТУАЛЬНАЯ КЛАСТЕРИЗАЦИЯ

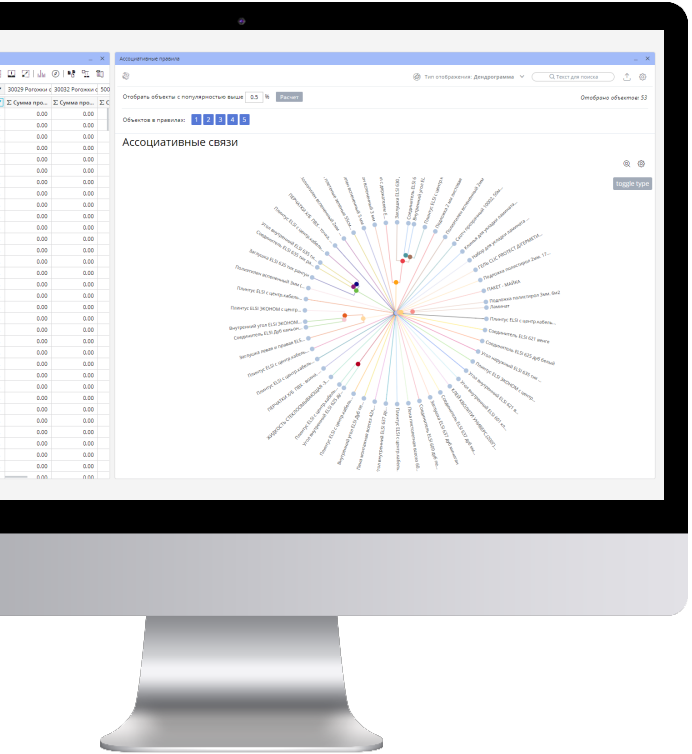


- Три алгоритма кластеризации, запускаемых в фоновом режиме

  От 2 до 100 кластеров, оцениваемых по качеству
- На основе одного или множества параметров
  - Запускается за секунды
  - Результаты воспроизводятся в 100% случаев
- Распределение кластера выстраивается автоматически

POLY//ATICA

НАХОДИТЕ СКРЫТЫЕ ПАТТЕРНЫ ПОВЕДЕНИЯ



Понимание взаимосвязи между событиями

- Классический алгоритм Априори
- Расчеты и фильтрация популярности, лифта и достоверности

Мгновенные результаты

- Ассоциативные правила для сотни миллионов объектов
- Возможность выбирать любое число объектов
- Представление правил в виде графики: графов и дендрограмм

POLY//ATICA

ДЕЛАЙТЕ ПРОГНОЗЫ НА БУДУЩЕЕ

- Множество моделей прогнозирования запускаются одновременно

  ARIMA, ARIMA-T, фильтр Калмана, нелинейная регрессия, логистическая регрессия

- Независимые испытания

  Каждая модель тестируется независимо на последних 10% наборов данных

POLY//ATICA

ВИЗУАЛИЗАЦИЯ В POLYMATICA

- Интерактивные графики и многомерная аналитика
- Локализация результатов на интерактивных географических картах

POLY//ATICA

БИЗНЕС-ЗАДАЧА: СОКРАЩЕНИЕ ОТТОКА

STARBUCKS

1 чашка кофе
3,5 $

LTV клиента
14 099 $

ПОДПИСКА ИЛИ РАЗОВЫЕ ПРОДАЖИ

- Ритейл и HORECA
- Банки
- SaaS
- Медиа и игры
- Операторы связи

POLY//ATICA

ВИДЫ ОТТОКА

Естественный:

- смена места жительства/временное использование
- смерть
- изменение потребностей (детские товары теряют актуальность по мере взросления ребенка)

Мотивированный (осознанный):

- недовольство уровнем оказываемых услуг или качества продукции
- найдена компания, где стоимость аналогичных товаров или услуг ниже при том же качестве
- найдена компания, где качество услуг выше, при той же стоимости
- «усталость» от бренда
- клиент вообще отказался от продукции или услуг, предоставляемых компанией (например, бросил курить)

Скрытый — постепенное снижение объёма и регулярности потребления продуктов компании из-за:

- параллельного пользования услугами конкурентов
- использования продуктов-заменителей

POLY//ATICA

TO DO LIST

- Дать определение оттоку
- Создать метрику, определить периодичность оценки и границы естественного оттока
- Утвердить с руководством

POLY//ATICA

БЕНЧМАРКИ ПО ОТРАСЛЯМ

- Кредитные карты – около 20%
- Операторы мобильной связи – 20-38%
- Провайдеры Software-as-a-Service (SaaS) обычно говорят о 5-7% оттока
- Подписка на развлекательные сервисы – 30%

Источник: https://www.wordstream.com/blog/ws/2014/05/12/customer-churn

POLY//ATICA

ДОПОЛНИТЕЛЬНЫЕ ТРЕБОВАНИЯ

- Горизонт прогнозирования
- Параметры, сигнализирующие о том, что клиент находится в зоне риска по оттоку
- Дизайн эксперимента – подтверждения эффективности

POLY//ATICA

ГОРИЗОНТ ПРОГНОЗИРОВАНИЯ

- Как быстро мы можем связаться с пользователем и какие каналы доступны?
- Какие методики удержания мы используем?
- Сколько времени занимает процесс удержания?

POLY//ATICA

ПРИМЕР

- Отток – разрыв договора подключения к сервису
- Горизонт прогнозирования – 2 недели
- Дизайн эксперимента – А/Б тестирование на 10%-м сегменте случайных пользователей

POLY//ATICA

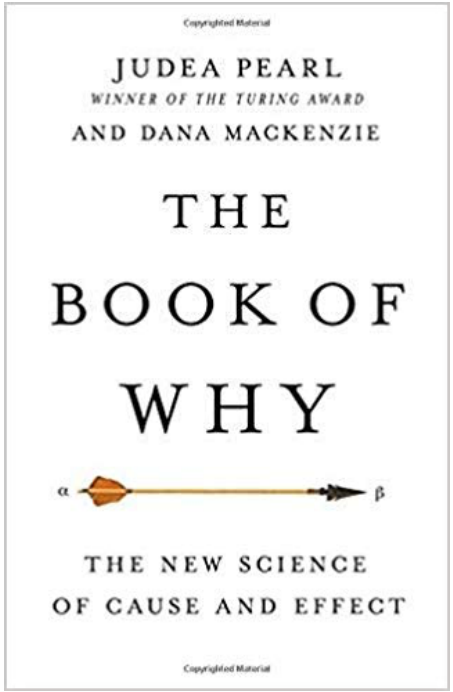
ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ, СИГНАЛИЗИРУЮЩИХ ОБ ОТТОКЕ

- Разрыв договора
- Отсутствие платных транзакций 10/90 дней
- Отсутствие на сервисе более 14/28 дней

POLY//ATICA

КАК ОТБИРАТЬ И ВАЛИДИРОВАТЬ ПАРАМЕТРЫ

- Корреляция
- Здравый смысл
- Модель «what if»



POLY//ATICA

КАКАЯ ДОЛЯ ПОПАДАЕТ ПОД ЭТО ОПРЕДЕЛЕНИЕ ОТТОКА

1. Разрыв договора – 0,1%
2. Отсутствие платных транзакций > 10 дней – 22%
3. Отсутствие платных транзакций > 90 дней – 5%
4. Отсутствие на сервисе более 14 дней – 16%
5. Отсутствие на сервисе более 28 дней – 9%

POLY//ATICA

КАКАЯ ДОЛЯ ВОЗВРАЩАЕТСЯ

1. Разрыв договора – 0%
2. Отсутствие платных транзакций > 10 дней – 80%
3. Отсутствие платных транзакций > 90 дней – 12%
4. Отсутствие на сервисе более 14 дней – 90%
5. Отсутствие на сервисе более 28 дней – 78%

POLY//ATICA

ДАННЫЕ

- Какие данные доступны?
- За какой исторический период?
- Как объединять данные?
- Есть ли в данных сигнал?
- Как данные следует обрабатывать?
- Как рассчитать признаки на основе данных?

POLY//ATICA

ОПИСАТЕЛЬНАЯ АНАЛИТИКА

- Ключевые характеристики пользователей
- Различаются ли пользователи из групп отток/не отток?
- Можем ли мы решать задачу для всех сегментов (новички или старые)?
- Одинаково ли важно решать задачу для всех сегментов?

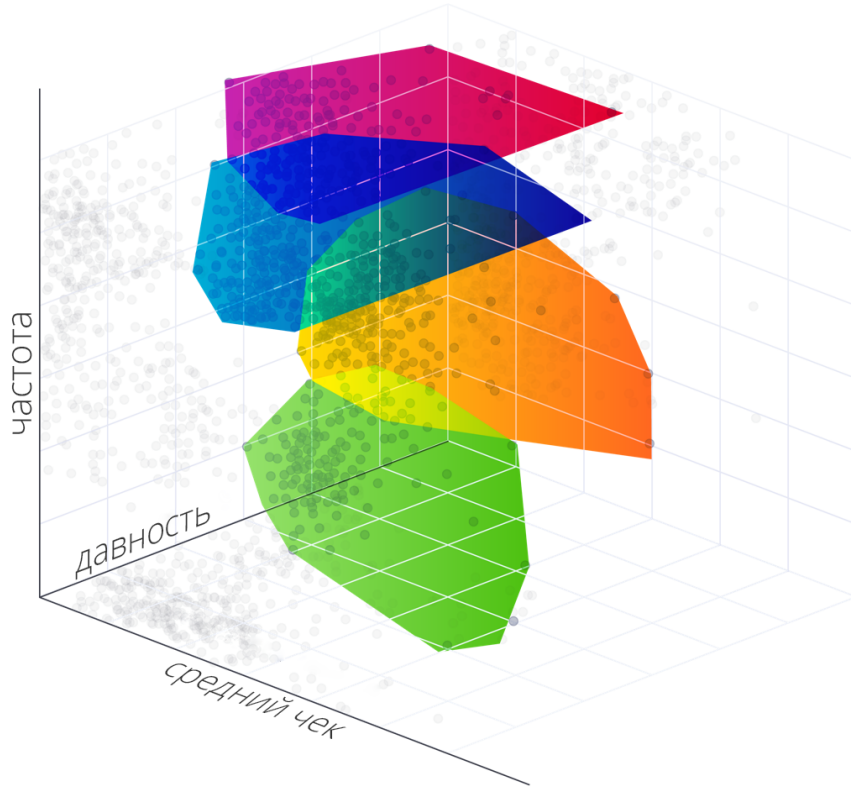
ПРЕДСКАЗАНИЕ ОТТОКА В РИТЕЙЛЕ

ПАРАМЕТРЫ ДЛЯ ОЦЕНКИ

- Периодичность покупок
- Размер чека/количество купленных товаров
- Давность покупок
- Участие в программе лояльности
- Обращение в службу поддержки
- Участие в сообществах
- Возврат товара
- Другие параметры (соц-дем)

POLY//ATICA

КЛАСТЕРИЗАЦИЯ



Кластеризация позволяет группировать объекты по схожим характеристикам

POLY//ATICA

RFM

Recency	Frequency	Monetory		
		Крупные суммы 3 балла	Средние суммы 2 балла	Малые суммы 1 балл
Недавно 3 балла	Часто 3 балла			
	Периодически 2 балла			
	Редко 1 балл			
Давно 2 балла	Часто 3 балла			
	Периодически 2 балла			
	Редко 1 балл			
Очень давно 1 балл	Часто 3 балла			
	Периодически 2 балла			
	Редко 1 балл			

- Лучшие клиенты
- С потенциалом
- Новые клиенты
- Хорошие клиенты в зоне риска
- Плохие в зоне риска
- Бывшие лучшие клиенты
- Бывшие клиенты

POLY//ATICA

ПРЕДСКАЗАНИЕ ОТТОКА ДЛЯ ОПЕРАТОРА СВЯЗИ

ЧТО НУЖНО ЗНАТЬ, ЧТОБЫ РАБОТАТЬ С ОТТОКОМ

Склонность к оттоку в первые 2-3 месяца после смены провайдера на 30% выше, чем у абонентов, пользующихся услугой год и более

Отток абонентов, которые посещают сайты конкурентов или звонят на КЦ конкурента в 5 раз выше оттока абонентов, обратившегося на техническую поддержку

Склонность к оттоку абонентов архивных тарифных планов в 3 раза выше чем абонента текущей линейки

Деградация скорости приводит к негативному пользовательскому опыту и как следствие к желанию сменить оператора

Работа на этапе проактивного сохранения в 3 раза эффективнее, чем на этапе реактивного сохранения

POLY//ATICA

ПАРАМЕТРЫ

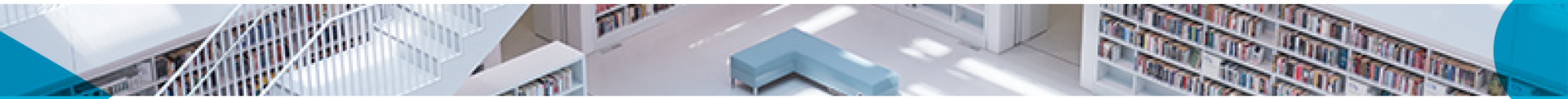
- Блокировки
- Участие в программе лояльности
- Продолжительность звонков
- Интернет-трафик
- Заход на сайты конкурентов
- Соц-дем
- Тарифные планы
- Обращение в службу поддержки

POLY//ATICA

ПРАКТИЧЕСКИЙ ПРИМЕР

О КОМПАНИИ ПОЛИМАТИКА

РОССИЙСКИЙ РАЗРАБОТЧИК УНИКАЛЬНОЙ АНАЛИТИЧЕСКОЙ ПЛАТФОРМЫ POLYMATICA, ПОЗВОЛЯЮЩЕЙ ОПЕРАТИВНО ОБРАБАТЫВАТЬ ЛЮБОЙ ОБЪЁМ ДАННЫХ ДЛЯ ПОИСКА ТОЧЕК РОСТА И ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ



Более **4 000 пользователей** в России и Европе

Аналитика по всем юридическим и физическим лицам РФ – более **150 млн объектов**

Используется в **3 критически-важных** приложениях в Правительстве РФ

5 отраслей: госсектор, ритейл, энергетика, банки и телеком

29 проектов **в России**
7 проектов **в Европе**

78 серверных инсталляций

POLY//ATICA

POLY//ATICA

E-MAIL
sales@polymatica.ru

ТЕЛ.
+7 495 748 84 84

САЙТ:
WWW.POLYMATICA.RU