

# Введение в маркетинговую аналитику

# Назначение CRM

Целью создания CRM систем является не только сбор данных о клиентах и выполнение регламентных процедур. Цель – **оптимизация работы с клиентами:**

- Выявление целевой аудитории
  - Определение особенностей потребления
  - Анализ структуры продаж
  - Прогнозирование спроса
  - Оценка эффективности маркетинговых действий
- и многое другое

# Состояние проблемы

Сбор данных без последующего глубокого анализа не позволяет выжать максимум из имеющейся информации, которая лежит «мертвым грузом».

В результате возникает парадоксальная ситуация – **данных много, а пользы от них мало**. Только применение полноценной аналитики позволит трансформировать данные в знания.

# Простые методы анализа

Обычно анализ данных начинается с использования механизмов визуализации:

- Аналитическая отчетность
- Нерегламентированные запросы
- Графики, диаграммы
- OLAP

Это простые и удобные инструменты, но они позволяют изучить только поверхностные и очевидные зависимости.

# Глубокий анализ данных

Реальный бизнес характеризуется сложными зависимостями, большими объемами данных, быстрыми изменениями.

Современные инструменты анализа позволяют выявлять в огромных объемах данных нетривиальные закономерности.

Фактически это единственный способ извлечь пользу из накопленной информации и превратить знания в **конкурентные преимущества**.

# Data Mining

Data Mining – это процесс обнаружения в базах данных нетривиальных и практически полезных закономерностей. Data Mining сводится к решению 5 классов задач:

- Классификация
- Регрессия
- Кластеризация
- Ассоциация
- Последовательность

# Классификация

Классификация – нахождение функциональной зависимости между входными атрибутами и **дискретным** выходным атрибутом.

Классификация позволяет отнести объект к одному из известных классов:

- Оценка перспективности клиентов
- Анализ рисков: давать или нет товарный кредит
- Оценка скидок: какой категории клиентов предоставлять скидки
- Прогнозирование успеха сделки
- Оценка эффективности рекламной компании

# Регрессия

Регрессия – нахождение функциональной зависимости между входными атрибутами и **непрерывным** выходным атрибутом. Позволяет оценивать вероятность возникновения события или его численное значение:

- Прогнозирование спроса
- Оценка ценовой эластичности
- Оценка вероятности повторных продаж
- Расчет загруженности склада, магазина, кассы
- Анализ влияния различных факторов на спрос



# Кластеризация

Кластеризация – разбиение объектов на кластеры, т.е. **группы схожих элементов**. Этот метод позволяет анализировать одни объекты по аналогии с поведением других:

- Кластеризация товаров, выявление товаров со схожей структурой спроса
- Разбиение клиентов на близкие по структуре и особенностям поведения группы
- Анализ спроса в зависимости от комбинации входных показателей
- Обнаружение аномальных отклонений

# Ассоциация

Ассоциация – это анализ транзакций, т.е. событий, **происходящих вместе**. Обнаружение зависимости, что из события А с определенной вероятностью следует событие Б:

- Предсказание поведения клиента и предложение товара, который, скорее всего, его заинтересует
- Размещение товаров на полках, в каталогах
- Кросс-продажи – стимулирование продаж одних товаров за счет продажи других
- Оптимизация складских запасов

# Последовательность

Последовательность – анализ событий, связанных между собой по времени. Обнаружение зависимости, что после события А **спустя определенное время** произойдет событие Б:

- Анализ потребности клиентов в расходных материалах, сопутствующих товарах, ремонте
- Повторные продажи, оценка наиболее вероятного времени потребности в модернизации
- Предсказание наиболее вероятного поведения потребителя

# Бизнес-задачи

Решение большинства задач, связанных со взаимоотношением с клиентами, сводится к применению методов Data Mining:

- Стимулирование продаж
- Прогнозирование спроса
- Анализ предпочтений
- Оценка эффективности действий
- Direct Mail
- Оценка эффективности менеджеров

# Результат

Описанные методы позволяют значительно повысить **эффективность работы** с клиентами и решать те задачи, ради которых внедряются CRM системы:

- Предугадывать потребности
- Предлагать те продукты, которые заинтересуют
- Закупать столько товаров, сколько необходимо
- Использовать наиболее удачные каналы продвижения
- Концентрировать внимание на наиболее перспективных категориях клиентов

# Конкурентная среда



Большинство зрелых рынков, пройдя этапы стандартизации и унификации, переходят к ценовым войнам

# Уход от ценовых войн



Монополия



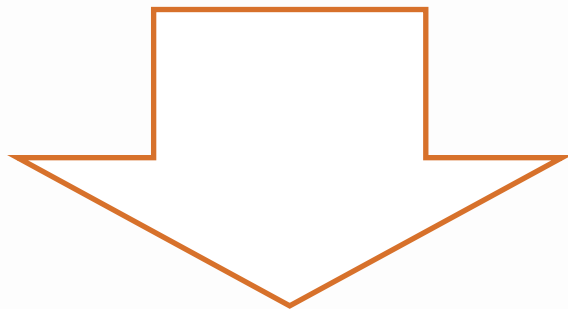
Уникальный  
продукт



Уникальное  
предложение



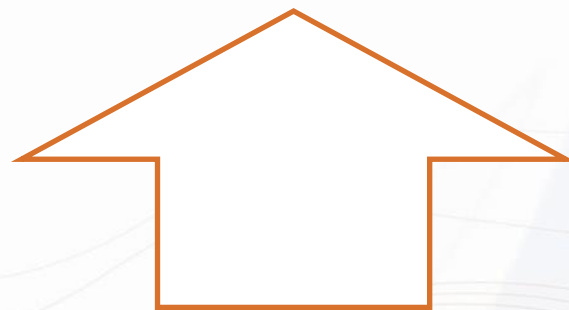
# Формирование лояльности



Учет особенностей  
каждого клиента



Скидки, скидки,  
скидки...





# Типичные ошибки

Действие компании	Взгляд клиента
Постоянная раздача скидок	Дайте еще больше скидок
Массовые рассылки	Не заваливайте меня спамом
Сложные правила участия	Меня обманывают
Одинаковое предложение всем	Мне это не интересно

# Точно в цель

Реальные  
потребности



Правильное  
предложение



Лояльный  
клиент

# Вопросы программы лояльности

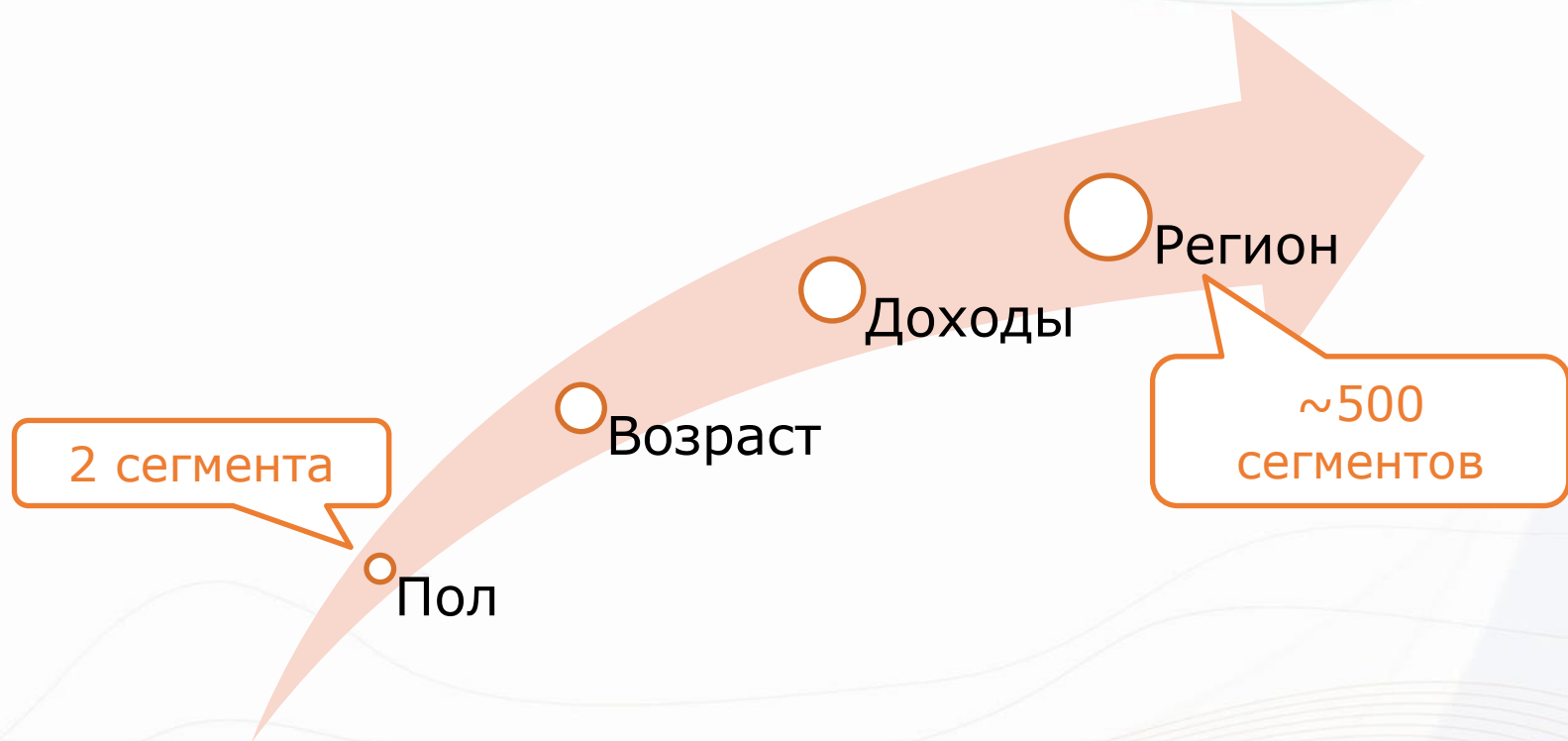


## Идеальный случай

Каждому покупателю предложение с учетом его особенностей: правильному клиенту правильный продукт по правильной цене в правильное время.

Но есть проблемы...

# Оценка трудоемкости



## Научный подход

Моделирование – единственный способ отойти от кустарных решений к промышленной работе: поставить процесс на поток.

Цель: не удержат конкретное клиента, а запустить механизм **систематического** повышения лояльности.

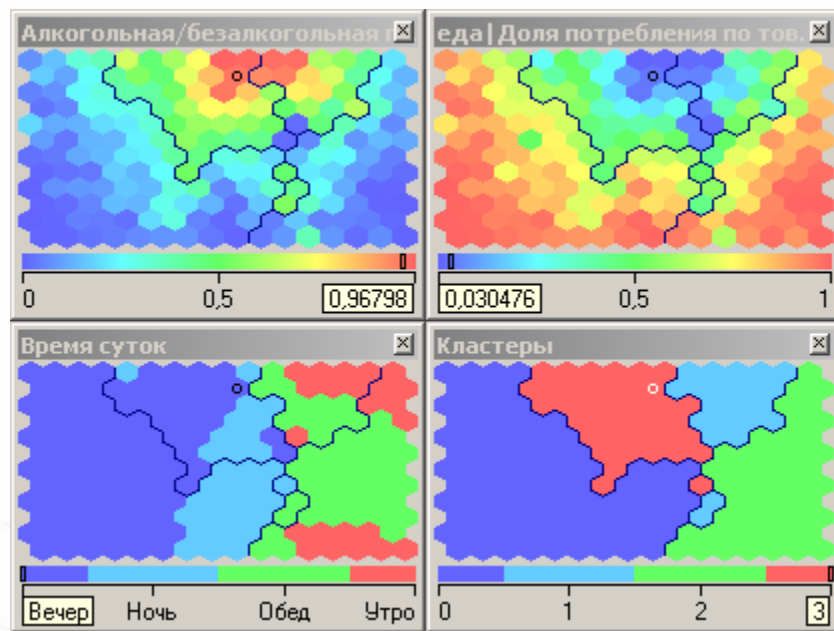
**Пример: Direct Marketing**

# Кейс: Direct Marketing



























# Этап: Сегментация клиентов



# Этап: Выбор целевой аудитории

Кластеры		0	2	3	1	Итого
		 8123 ( 52,2%)	 4242 ( 27,3%)	 2217 ( 14,3%)	 970 ( 6,2%)	   
- + Поля	Показатели					
<div>ab</div> <div>12</div> <div>Время суток</div> <div>Среднее количество позиций в чеке</div>	Значимость	 100,0%	 100,0%	 100,0%	 100,0%	 100,0%
	Значимость	 100,0%	 66,7%	 100,0%	 100,0%	 100,0%
	Доверительный интервал					
	Среднее	9,8714	8,6874	6,0158	6,6845	
	Стандартн. откл.	6,7526	6,7526	4,4161	5,4488	6,5472
	Стандартн. ошиб.	0,074922	0,10368	0,09379	0,17495	0,052501

Характеристики кластера

## Этап: Адресные предложения



















Что уже  
купили

Условие	
Элемент	Поддержка, %
Макаронные изделия	5,77
Мясо животных	8,73
Полуфабрикаты замороженные	6,49
Товары для животных	3,16

Что еще  
предложить

# Этап: Анализ отклика клиентов

Вероятность  
срабатывания

Условие	Следствие	Поддержка	Достоверность
ЕСЛИ		 3040	 2401
Кол-во заказов < 1,5		 314	 184
Сумма за месяц < 11,55		 137	 76
Рассылок всего < 5,5		 85	 46
Сумма за 2 и 3 мес. < 70,4	0	 45	 29
Сумма за 2 и 3 мес. >= 70,4	1	 40	 23
Рассылок всего >= 5,5	1	 52	 37
Сумма за месяц >= 11,55	0	 177	 123
Кол-во заказов >= 1,5		 2726	 2271

Правила

# Этап: Оценка эффективности

ID чека   Номер карты		RFM Score					
		Recency Score					
+ Frequency ...	Recency Monetary	1	2	3	4	5	Итого:
4	4	55	159	330	390	277	1 211
	5	36	103	236	276	209	860
	Итого:	142	456	909	982	679	3 168
5	1					2	2
	2		2	10	11	8	31
	3	6	11	38	74	65	194
	4	4	26	129	305	311	775
	5	10	51	226	766	1 076	2 128
	Итого:	20	90	403	1 158	1 460	3 131
Итого:		3 339	3 189	3 498	3 426	2 805	16 257

1 / 2458

ID чека	Recency	Frequency	Monetary	Recency Score	Frequency Score
1	270	26	6178,98	5	5
100	269	66	5943,32	5	5
1000	271	33	12066,09	5	5
1100067	269	83	11151,25	5	5
1100087	275	16	8487,33	4	5

RFM-  
сегменты

# Результаты

- Учитываем потребности и возможности клиентов
- Делаем интересные предложения каждому
- Предлагаем выгодные продавцу товары
- Не раздражаем бесполезными спамом

Бизнес-кейсы

# Netflix – продажа медиа

The Netflix logo, consisting of the word "NETFLIX" in white, bold, sans-serif capital letters with a slight 3D effect, set against a solid red rectangular background.

Автоматическое формирование персональных рекомендаций фильмов и сериалов

Результат аналитики:

- Более 80% трафика просмотров приходится на рекомендации пользователям, а не на их поисковую активность



# O'Neill – производство одежды

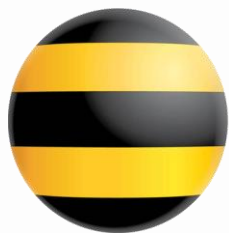


Автоматическая генерация адресных предложений по выбору одежды

Результат аналитики:

- 50% рост отклика на предложение по сравнению с иными вариантами работы

# Билайн - телекоммуникации



Билайн®

Автоматическая формирование предложений на приобретение дополнительных услуг

Результат аналитики:

- Повышение отклика на SMS-предложение в 4 раза по сравнению со стандартной рассылкой

# Tesco – сеть супермаркетов



Выпуск >7 млн. индивидуальных товарных купонов в год с адресными предложениями

Результат аналитики:

- Погашается от 20 до 50% купонов, а по отрасли в среднем – 2%

# Sprint - телекоммуникации



Противодействие оттоку клиентов за счет ненавязчивого наилучшего предложения в критические моменты жизненного цикла

Результат аналитики:

- \$1 млрд. дополнительной прибыли за кросс-продаж + \$500 млн. за счет снижения оттока

# Клиентская аналитика: методы принятия решений

## Вопросы к аналитике

- Кого привлекать?
- Как удержать ценных клиентов?
- Как увеличить прибыльность?
- Как сформировать привлекательные адресные предложения?
- Как увеличить отклик на предложения?
- Как минимизировать отток?
- Как минимизировать негатив?
- Как диагностировать проблемы?

# Варианты анализа



# Простые методы

# ABC-анализ: алгоритм

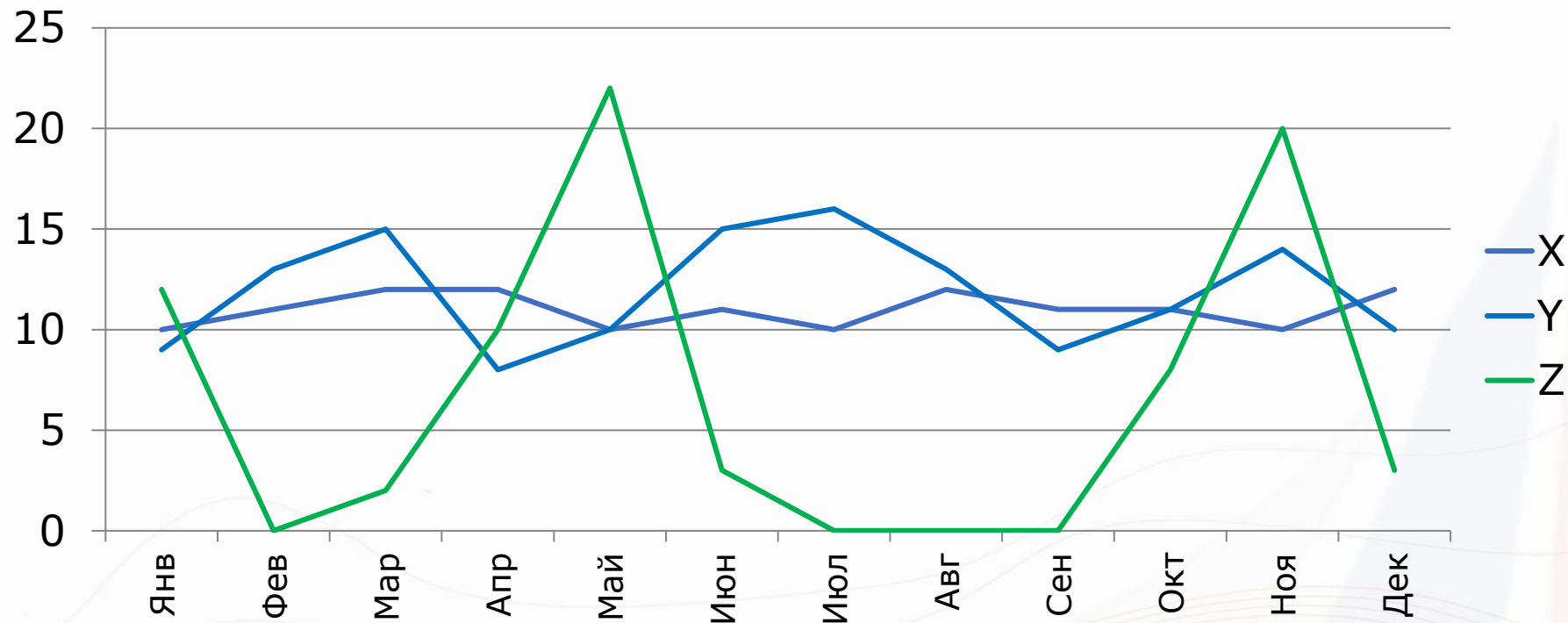
1. Выбрать показатель: выручка, маржа...
2. Агрегировать данные по каждому клиенту за последний период
3. Отсортировать по убыванию показателя, рассчитать накопительную сумму:
  - Первые 20% - самые ценные
  - Следующие 30% - промежуточные
  - Последние 50% - наименее ценные



# XYZ-анализ: алгоритм

1. Выбрать квант: день, неделя, месяц
2. Агрегировать данные по продажам каждому клиенту и получить временные ряды
3. Рассчитать коэффициент вариации для каждого клиента:
  - 0-10% - стабильное потребление
  - 10-25% - регулярное потребление
  - >25% - хаотическое потребление

# XYZ-анализ: пример



# RFM-анализ: идея

Оценка лояльности клиента на основе показателей:

- **R**ecency (новизна) – давность последней покупки.
- **F**requency (частота) – частота покупок.
- **M**onetary (деньги) – потраченные суммы

# RFM-анализ: алгоритм

Каждый из показателей делится на 5 квантилей, например:

R	Давность покупки
5	<1 мес.
4	2-3 мес.
3	4-8 мес.
2	9-18 мес.
1	>18 мес.

F	Частота покупок
5	>1 в 1 мес.
4	1 в 2-3 мес.
3	1 в 4-8 мес.
2	1 в 9-18 мес.
1	1 в >18 мес.

M	Потраченные суммы
5	>100 000 руб.
4	50-100 000 руб.
3	10-50 000 руб.
2	3-10 000 руб.
1	<3 000 руб.

Считается RFM-код каждого клиента.

# RFM-анализ: визуализация

IDчека   Номер карты		RFM Score					
		Recency Score					
+ Frequency ...	Recency Monetary	1	2	3	4	5	Итого:
4	4	55	159	330	390	277	1 211
	5	36	103	236	276	209	860
	Итого:	142	456	909	982	679	3 168
5	1					2	2
	2		2	10	11	8	31
	3	6	11	38	74	65	194
	4	4	26	129	305	311	775
	5	10	51	226	766	1 076	2 129
	Итого:	20	90	403	1 158	1 460	3 131
Итого:		3 339	3 189	3 498	3 426	2 805	16 257

1 / 2458
----------

IDчека	Recency	Frequency	Monetary	Recency Score	Frequency Score
Номер карты					
1	270	26	6178,98	5	5
100	269	66	5943,32	5	5
1000	271	33	12066,09	5	5
1100067	269	83	11151,25	5	5
1100087	275	16	8487,33	4	5

# RFM-анализ: применение

1. Оценка привлекательности клиента
2. Контроль изменений в поведении
3. Оценка миграции между группами
4. Оценка эффективности программы лояльности
5. Выбор стратегии работы с каждым сегментом

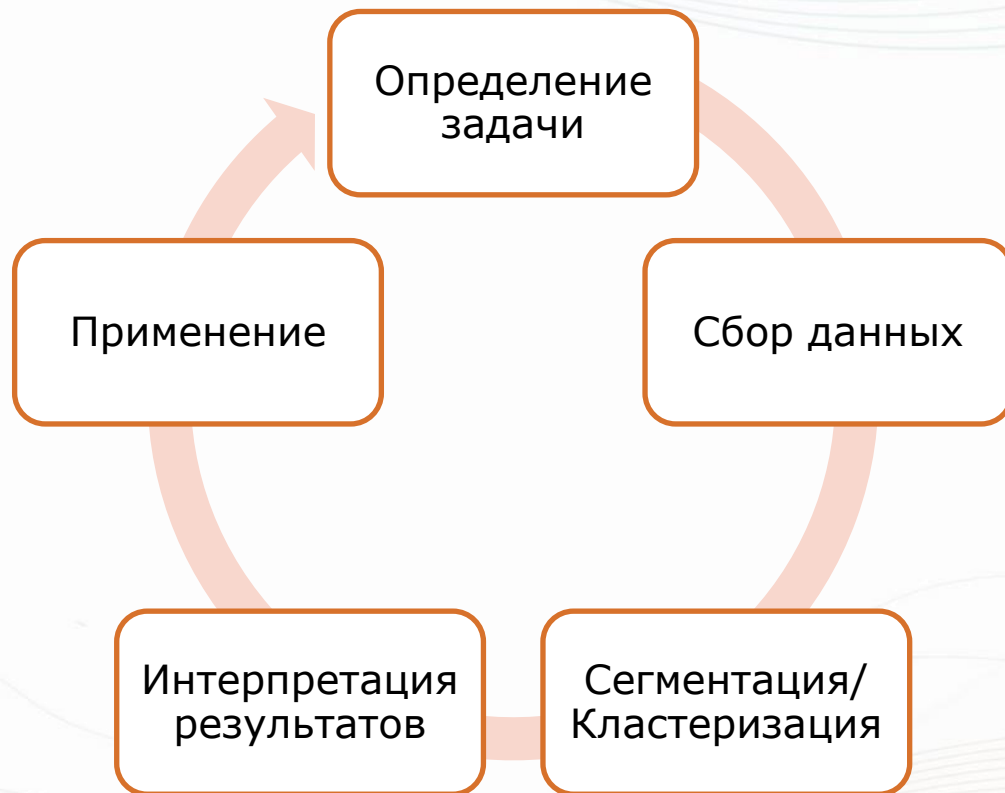
# RFM-анализ: выводы

Клиенты	R	F	M	Количество клиентов
Лояльные				
Промежуточные				
Нелояльные				
Потерянные				

# Продвинутые методы



# Цикл работ



# Математический аппарат

Разработано множество алгоритмов кластеризации. Некоторые наиболее популярные:

- Карты Кохонена
- К-средних
- EM-кластеризация
- Иерархическая

## Требования к данным

1. Оцифрованы и систематизированы
2. Содержат значимые факторы
3. Очищены от ошибок
4. Рассчитаны агрегаты
5. Сведены в одну таблицу

# Анализируемые данные

## Соц-дем

- Пол
- Возраст
- Семья
- Образование
- Собственность

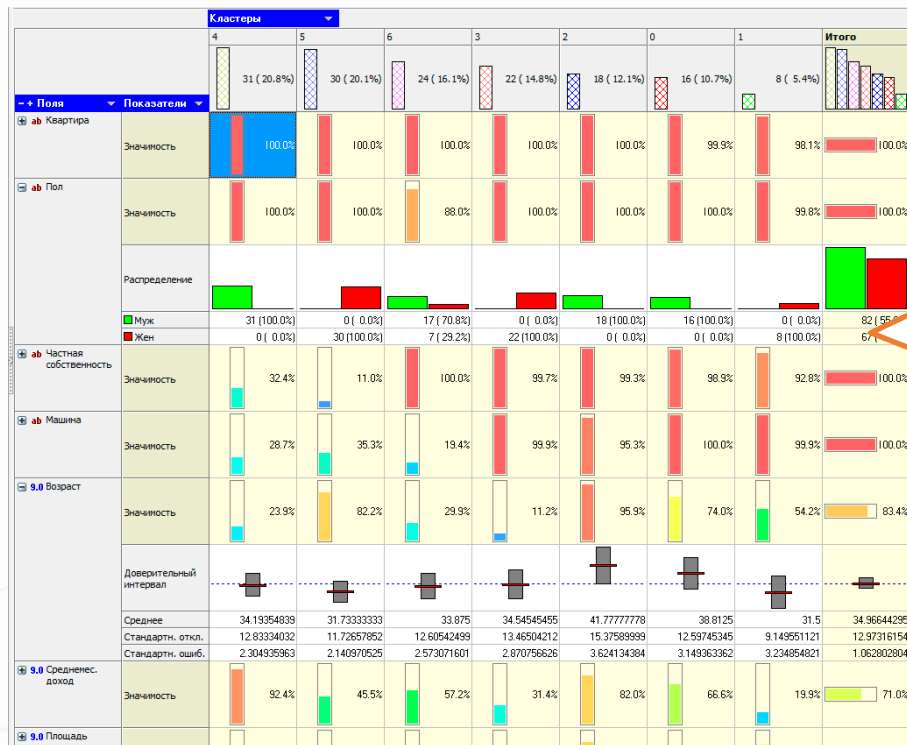
## Поведение

- Частота покупок
- Средний чек
- Последняя покупка
- Покупки товаров по акциям

## География

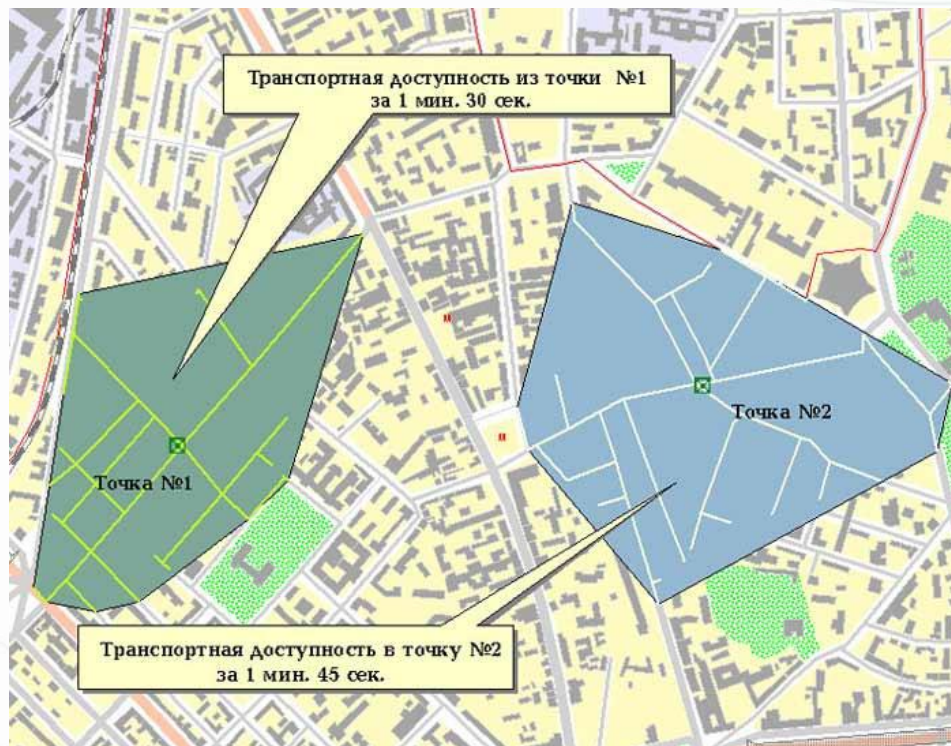
- Место жительства
- Место работы
- Место покупки
- Перемещения

# Сегментация: профили кластеров



Социально-демографические данные + поведение

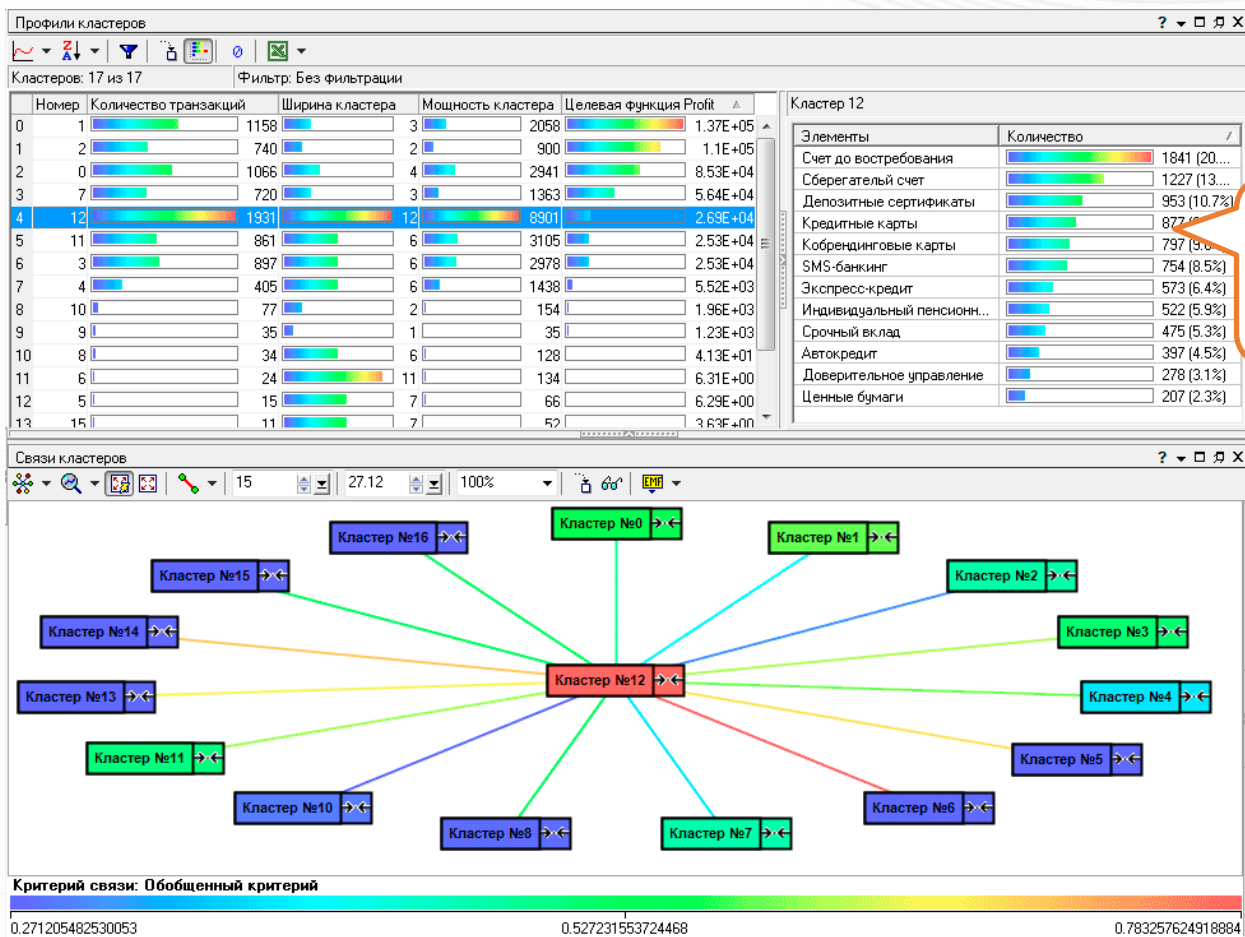
# Сегментация: география



## Предпочтения: алгоритм

1. Консолидировать данные о транзакциях (чеках)
2. Определить товары-маркеры
3. Кластеризовать транзакции
4. Рассчитать предпочтения по товарам-маркерам

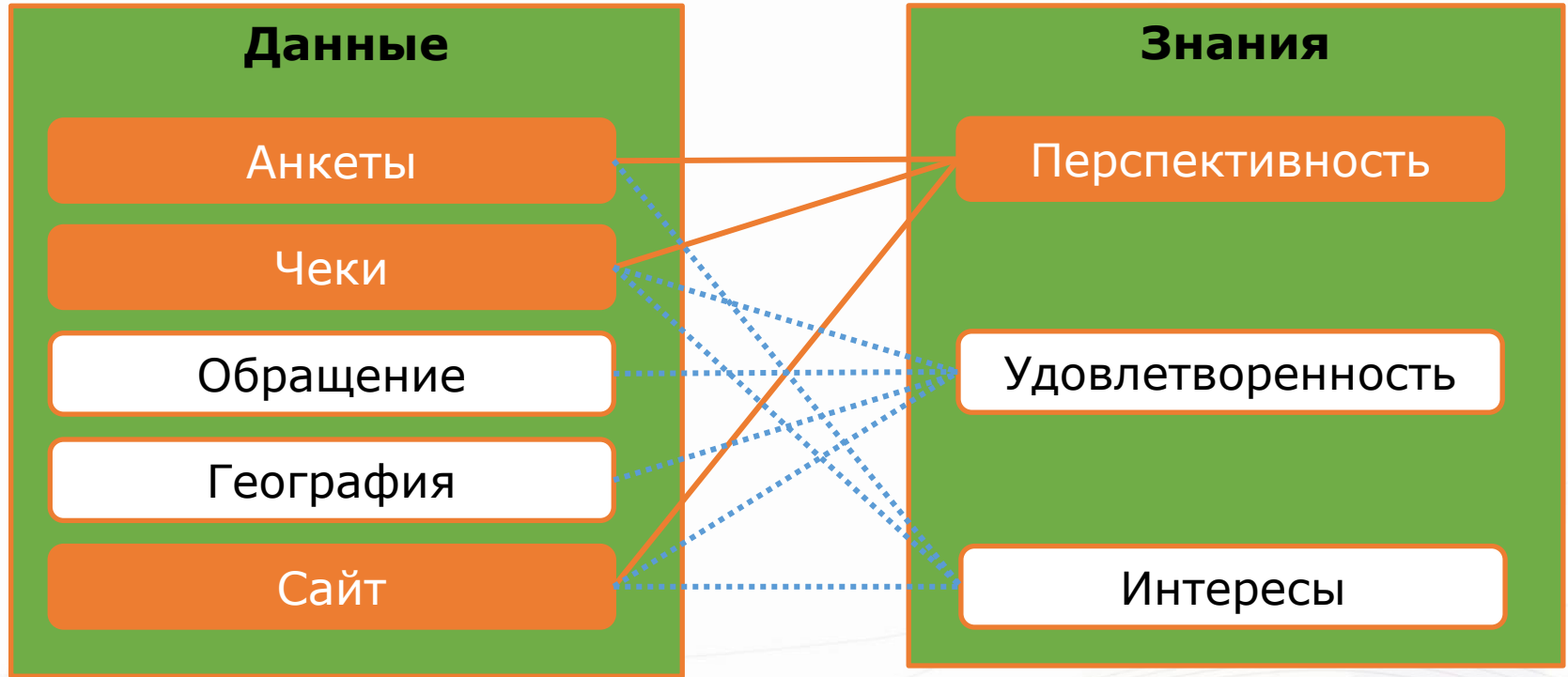
# Предпочтения: визуализация



Типичная  
потребительская  
корзина



# Комплексная аналитика



# Оценка сложности

Метод	Сложность	Специалист
ABC	Простой	Маркетолог
XYZ	Простой	Маркетолог
RFM	Простой	Маркетолог
Соц-дем.	Средний	Аналитик
География	Средний	Аналитик
Поведение	Сложный	Data Mining аналитик
Предпочтения	Сложный	Data Mining аналитик
Комплексная	Сложный	Команда аналитиков

# Технологии анализа

