

# РАЗРАБОТКА АЛГОРИТМИЧЕСКИХ СРЕДСТВ ДЛЯ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ В РИТЕЙЛЕ

КИРЕЕВ ВАСИЛИЙ,  
ТИМЛИД КОМАНДЫ DS,  
СТРОИТЕЛЬНЫЙ ДВОР

ТЮМЕНЬ, 2019

# ЗАДАЧИ DS В РИТЕЙЛЕ

- СЕГМЕНТАЦИЯ ПОТРЕБИТЕЛЕЙ
- ПРОГНОЗИРОВАНИЕ СПРОСА
- АНАЛИЗ РЫНОЧНОЙ КОРЗИНЫ
- ЦЕНООБРАЗОВАНИЕ (В ТОМ ЧИСЛЕ ЛОКАЛЬНОЕ)
- АНАЛИЗ ЧУВСТВИТЕЛЬНОСТИ СПРОСА
- ПОВЫШЕНИЕ ОБЪЁМА ПРОДАЖ
- ОПТИМИЗАЦИЯ ТОВАРНОГО АССОРТИМЕНТА
- КОНКУРЕНТНЫЙ АНАЛИЗ
- АНАЛИЗ ТОНАЛЬНОСТИ ОТЗЫВОВ
- ИИ ДЛЯ ЧАТ-БОТОВ

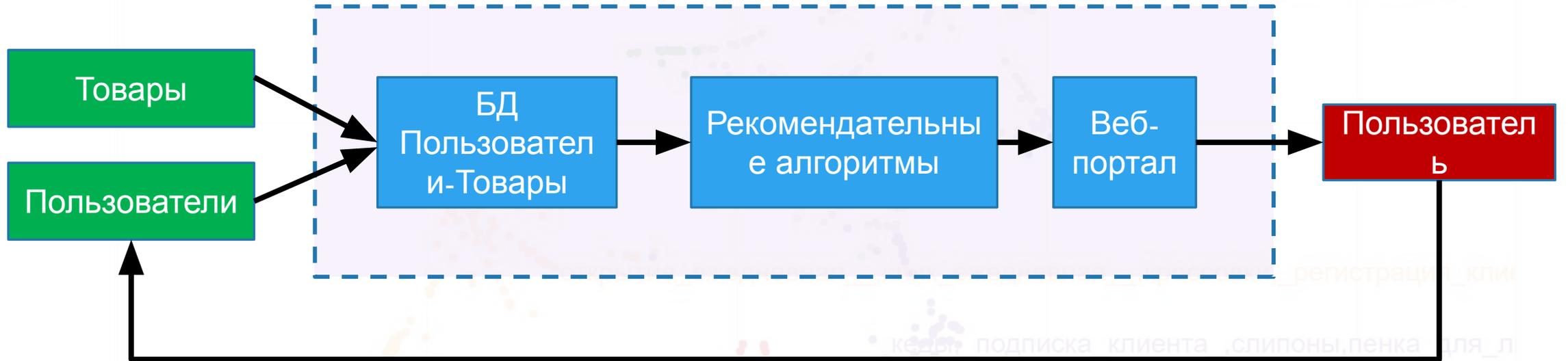
# ВИДЫ ПРОДАЖ, ПОДДЕРЖИВАЕМЫХ РС

- **ДОПРОДАЖИ (UP-SELLING)** – ПОВЫШЕНИЕ УРОВНЯ ПРОДАЖ ОСНОВНЫХ ТОВАРОВ
- **СОПУТСТВУЮЩИЕ ПРОДАЖИ (CROSS-SELLING)** – ПОВЫШЕНИЕ УРОВНЯ ПРОДАЖ ТОВАРОВ, ДОПОЛНЯЮЩИХ ОСНОВНЫЕ, НАПРИМЕР – ЗИП К ОБОРУДОВАНИЮ

# ПОДХОДЫ К СОЗДАНИЮ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ (РС)

- **ТОВАР-ТОВАР (КОНТЕНТНАЯ ФИЛЬТРАЦИЯ)**
  - ПРЕДЛОЖИТЬ ТОВАРЫ, ПРОДАЮЩИЕСЯ НАИБОЛЕЕ ЧАСТО (ХИТЫ ПРОДАЖ)
  - ПРЕДЛОЖИТЬ ТОВАРЫ, ПОХОЖИЕ НА ДРУГИЕ ТОВАРЫ
  - ПРЕДЛОЖИТЬ ТОВАРЫ, ПРОДАЮЩИЕСЯ С ДРУГИМИ ТОВАРАМИ
- **ПОТРЕБИТЕЛЬ-ТОВАР (КОЛЛАБОРАТИВНАЯ ФИЛЬТРАЦИЯ)**
  - ПРЕДЛОЖИТЬ ТОВАРЫ, КОТОРЫЕ ПОКУПАЮТСЯ ПОЛЬЗОВАТЕЛЯМИ, ПОХОЖИМИ НА ДАННОГО
- **ГИБРИДНЫЕ ПОДХОДЫ**
  - КОМБИНАЦИИ ПЕРЕЧИСЛЕННЫХ ВЫШЕ ВАРИАНТОВ

# ОБЩАЯ АРХИТЕКТУРА РС



# ИЗВЕСТНЫЕ ПРОБЛЕМЫ РС

- **ХОЛОДНЫЙ СТАРТ** – ОТСУТСТВИЕ ИНФОРМАЦИИ О НОВОМ ПОТРЕБИТЕЛЕ ИЛИ НОВОМ ТОВАРЕ, НЕ ПОЗВОЛЯЕТ ИСПОЛЬЗОВАТЬ ИСТОРИЮ ПРОДАЖ ДЛЯ СОЗДАНИЯ КАЧЕСТВЕННЫХ ПЕРСОНАЛЬНЫХ РЕКОМЕНДАЦИЙ
- **ПРОБЛЕМЫ В ОЦЕНКЕ ЭФФЕКТИВНОСТИ РЕКОМЕНДАЦИЙ** – ПОВЫШЕНИЕ УРОВНЯ КОНВЕРСИИ РАЗНЫХ ВИДОВ ПОСЛЕ ВНЕДРЕНИЯ РС, МОЖЕТ БЫТЬ СВЯЗАНО С БОЛЬШИМ КОЛИЧЕСТВОМ ФАКТОРОВ
- **УСТАРЕВАНИЕ ТЕНДЕНЦИЙ** – ВКУСЫ ПОТРЕБИТЕЛЕЙ И ИХ СОСТАВ МЕНЯЮТСЯ СО ВРЕМЕНЕМ, РАЗРАБОТАННЫЕ ОДНАЖДЫ МОДЕЛИ, ТРЕБУЮТ ПОСТОЯННЫХ ОБНОВЛЕНИЙ, ИНОГДА РАДИКАЛЬНЫХ

# ОЦЕНКА КАЧЕСТВА РС

- **ПЕРТИНЕНТНОСТЬ** – ВНУТРЕННЯЯ УДОВЛЕТВОРЕННОСТЬ ПОЛЬЗОВАТЕЛЯ РЕКОМЕНДУЕМОЙ ВЫДАЧЕЙ ТОВАРОВ
  - ПРЯМАЯ ОЦЕНКА ОПРОСНЫМИ МЕТОДАМИ – ОТВЕТЫ В АНКЕТАХ, ВСПЛЫВАЮЩИХ ФОРМАХ НА САЙТЕ
  - КОСВЕННАЯ ОЦЕНКА ПО РЕАКЦИИ ПОЛЬЗОВАТЕЛЕЙ – КЛИКИ, ПЕРЕХОДЫ...
- **РЕЛЕВАНТНОСТЬ** – СООТВЕТСТВИЕ РЕКОМЕНДУЕМОЙ ВЫДАЧИ ТОВАРОВ ЗАПРОСУ ПОЛЬЗОВАТЕЛЯ
  - ПРЯМАЯ ОЦЕНКА - РАСЧЕТ РАССТОЯНИЙ ОТ ВЕКТОРИЗОВАННОГО ЗАПРОСА ПОЛЬЗОВАТЕЛЯ В ПОИСКЕ, ДО ВЕКТОРИЗОВАННОЙ РЕКОМЕНДАЦИОННОЙ ВЫДАЧИ, ТОЧНОСТЬ, ПОЛНОТА, МЕТРИКА ПОЛУРАСПАДА
- **БИЗНЕС-МЕТРИКИ**
  - ИЗМЕНЕНИЯ УРОВНЯ КОНВЕРСИИ, СРЕДНЕЙ СУММЫ ЧЕКА, СРЕДНЕГО КОЛИЧЕСТВА ПОЗИЦИЙ В ЧЕКЕ...

# МЕТОДЫ ML ДЛЯ РС

- **ЭВРИСТИЧЕСКИЕ ПОДХОДЫ**

- **RWR...** - РАСЧЕТ ПРЕДПОЧТЕНИЙ ПОЛЬЗОВАТЕЛЕЙ ПО ГРАФУ СОЦИАЛЬНЫХ СВЯЗЕЙ

- **ТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ**

- **WORD2VEC, LDA, LSA...** - ВЫЯВЛЕНИЕ СХОЖИХ ТОВАРОВ ПО ОПИСАНИЮ

- **АССОЦИАТИВНЫЕ ПРАВИЛА**

- **APRIORI, FP-GROWTH** – ВЫЯВЛЕНИЕ СВЯЗАННЫХ ТОВАРОВ ПО ИСТОРИИ ПРОДАЖ

- **КЛАСТЕРИЗАЦИЯ**

- **KMEANS, DBSCAN...** - ВЫЯВЛЕНИЕ КОМПЛЕКТНЫХ ТОВАРОВ, ВЫЯВЛЕНИЕ ПОЛЬЗОВАТЕЛЬСКИХ СЕГМЕНТОВ

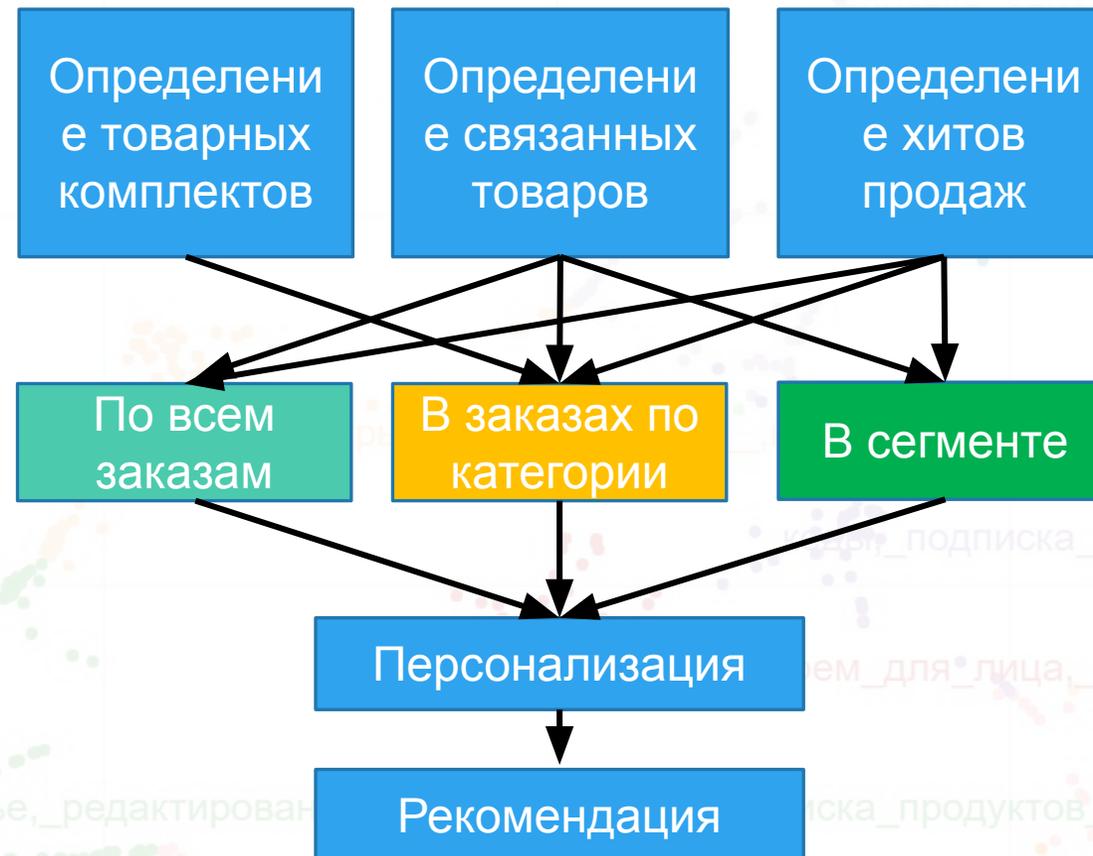
- **КЛАССИФИКАЦИЯ**

- **RF, NB, XGBOOST..** – ВЫЯВЛЕНИЕ ПОЛЬЗОВАТЕЛЕЙ, ПОХОЖИХ НА ДАННОГО, ВЫЯВЛЕНИЕ ТОВАРОВ ПОХОЖИХ НА ДАННЫЙ ПО ИЗОБРАЖЕНИЮ

- **ГЛУБОКОЕ ОБУЧЕНИЕ**

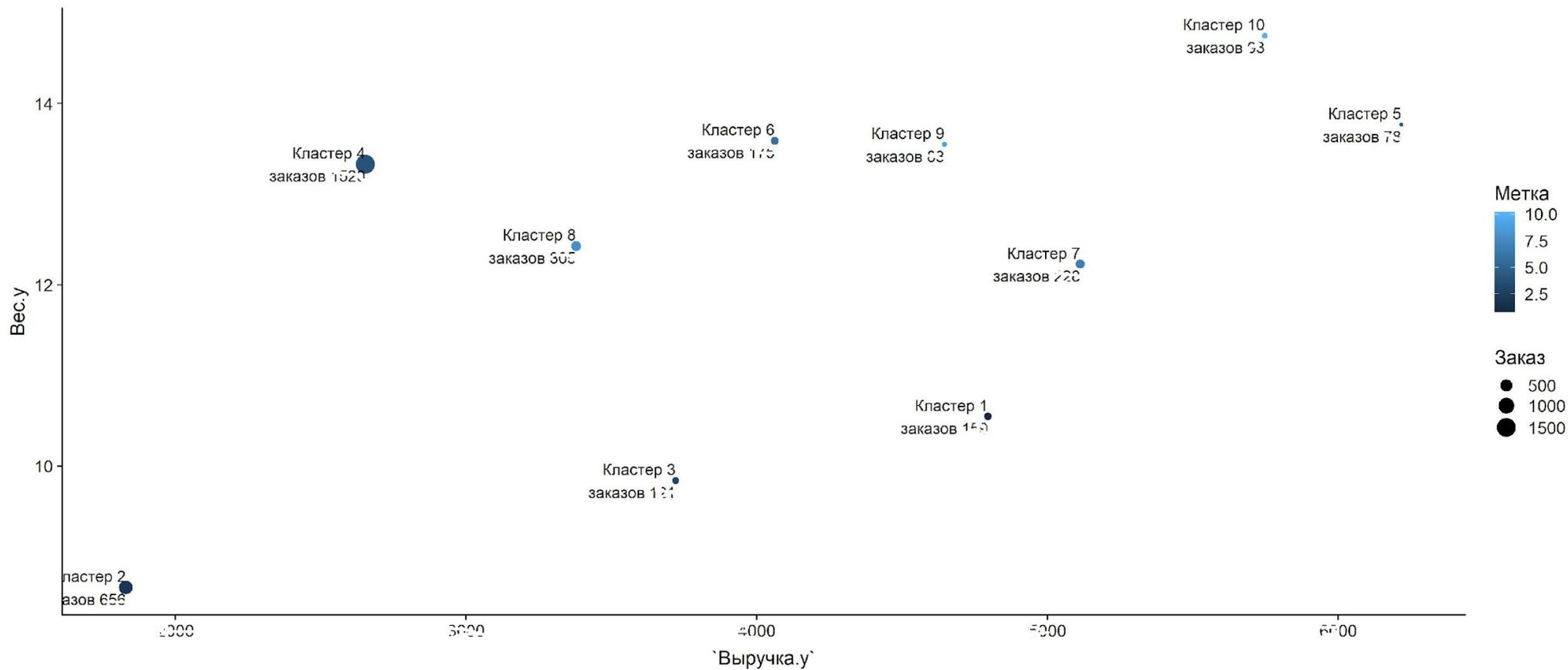
- **GAN, NADE...**-НЕЙРОСЕТЬ, ФОРМИРУЮЩАЯ ОЦЕНКИ ТОВАРОВ

# ПРЕДЛАГАЕМЫЙ ПОДХОД ДЛЯ СОЗДАНИЯ РЕКОМЕНДАЦИЙ

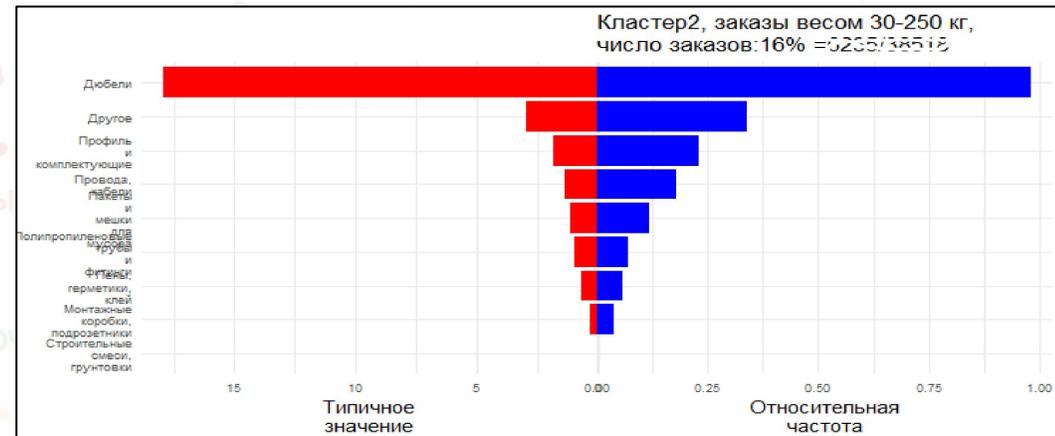
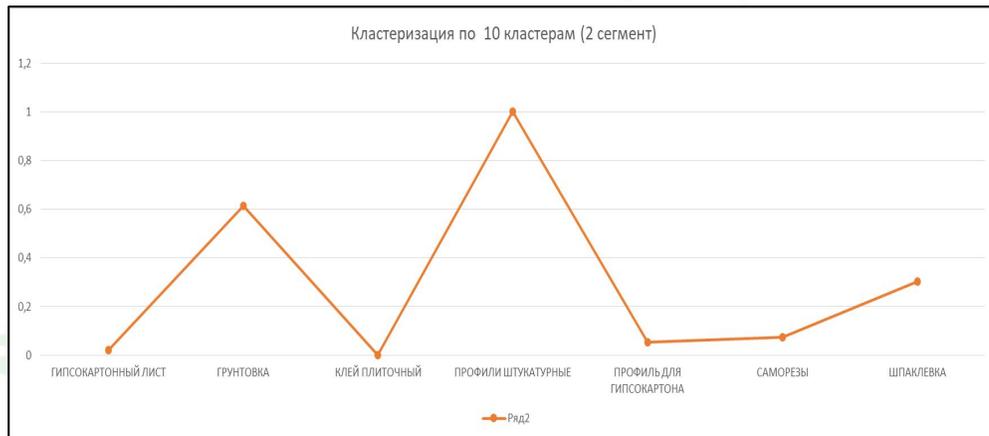
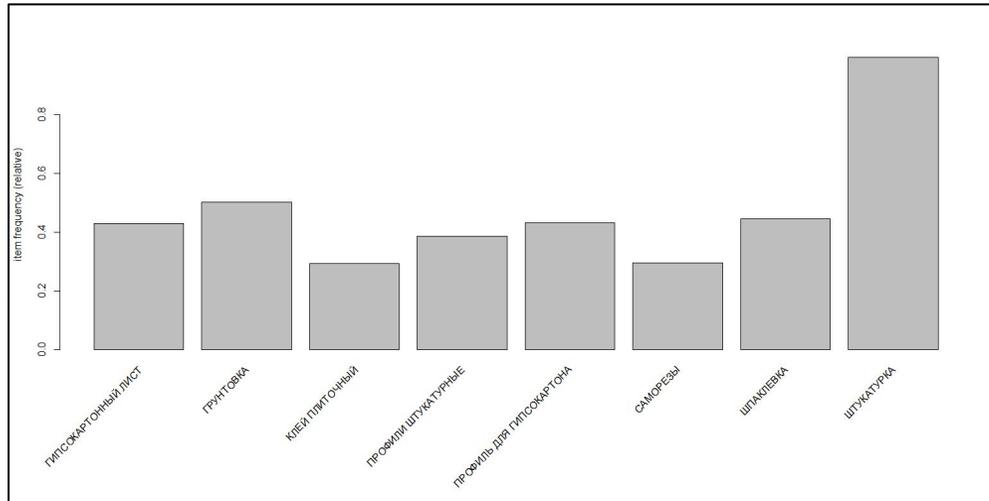


# ПРИМЕР СЕГМЕНТАЦИИ

Заказы до 30 кг, доставка, всего заказов 1521, заказов с 15 самыми частыми РТ: Пистолеты, Пены, герметики, клей,  
15 самых частых РТ: Пистолеты, Пены, герметики, клей,  
Канализация из ПВХ, Хомуты, Средства герметизации, Трубы и фитинги,  
Шаровые краны, Резьбовые фитинги, Полипропиленовая канализация, Полипропиленовые трубы и фитинги,

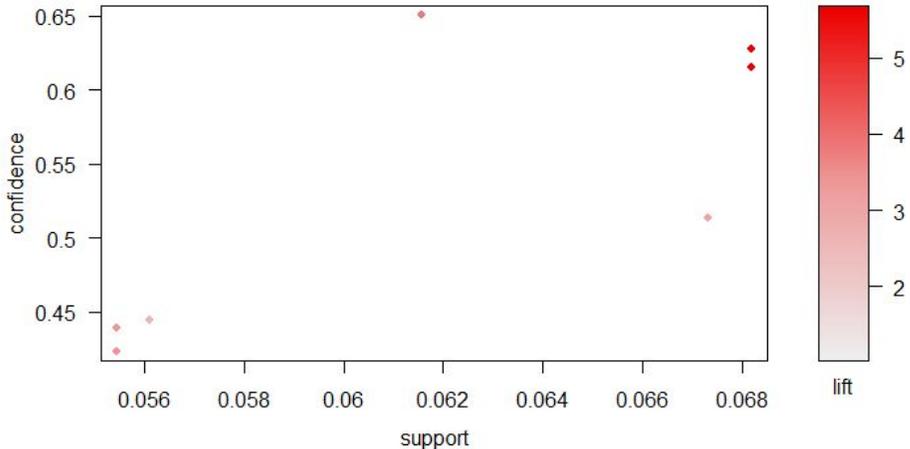


# ПРОФИЛЬ СЕГМЕНТОВ



# ПРИМЕР СВЯЗОК ТОВАРОВ, ПОЛУЧЕННЫХ АССОЦИАТИВНЫМИ ПРАВИЛАМИ

Scatter plot for 7 rules



	rules	support	confidence	lift	count
1	{ГИПСОКАРТОННЫЙ ЛИСТ} => {ПРОФИЛЬ ДЛЯ ГИПСОКАРТОНА}	0,07	0,63	5,66	938
2	{ПРОФИЛЬ ДЛЯ ГИПСОКАРТОНА} => {ГИПСОКАРТОННЫЙ ЛИСТ}	0,07	0,62	5,66	938
3	{ПРОФИЛИ ШТУКАТУРНЫЕ} => {ШТУКАТУРКА}	0,06	0,65	3,77	847
4	{ШПАКЛЕВКА} => {ГРУНТОВКА}	0,06	0,44	3,35	763
5	{ГРУНТОВКА} => {ШПАКЛЕВКА}	0,06	0,42	3,35	763
6	{ШПАКЛЕВКА} => {ШТУКАТУРКА}	0,06	0,44	2,57	772
7	{ГРУНТОВКА} => {ШТУКАТУРКА}	0,07	0,51	2,97	926

Rule:  $X \Rightarrow Y$

$$Support = \frac{freq(X, Y)}{N}$$

$$Confidence = \frac{freq(X, Y)}{freq(X)}$$

$$Lift = \frac{Support}{Supp(X) \times Supp(Y)}$$