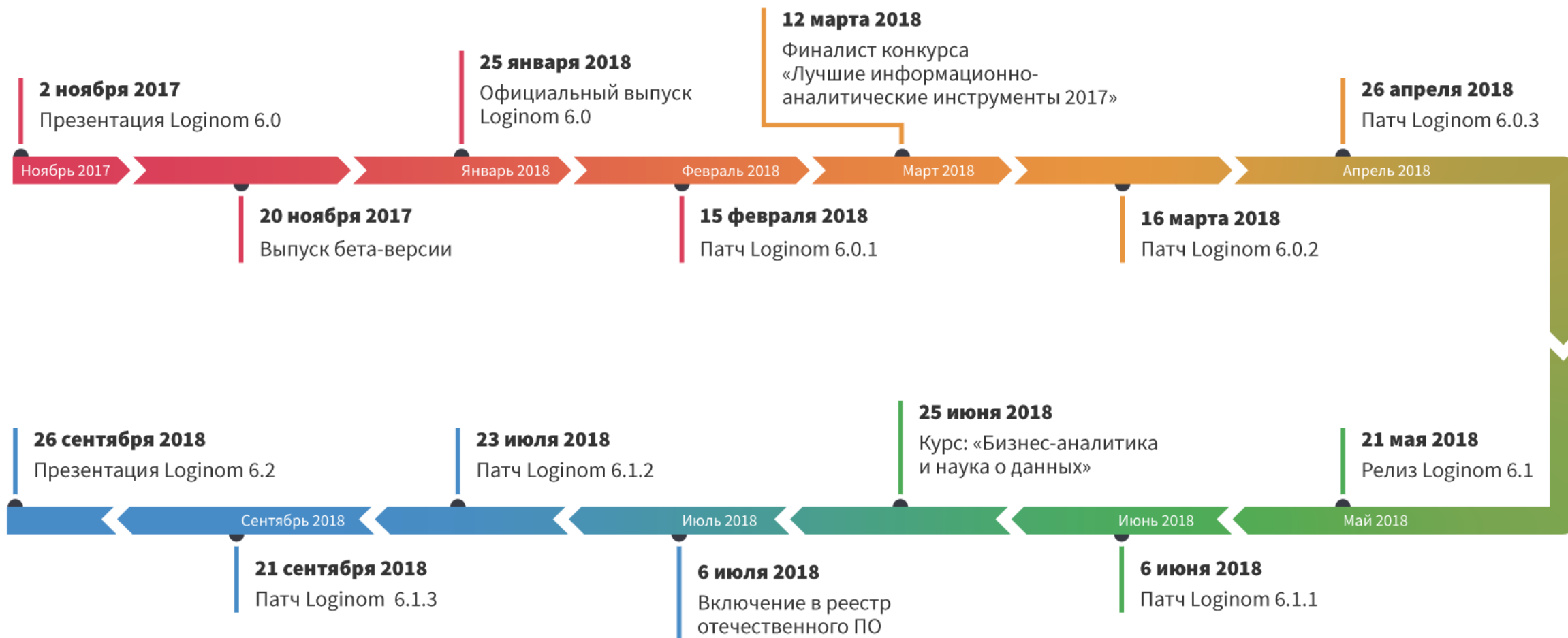



# Logiном 6.2 – ЧТО НОВОГО?

Алексей Арустамов

Logiном Company

# Развитие платформы Loginom






300+

улучшений

- Увеличение стабильности
- Повышение скорости
- Улучшения интерфейса

# Направления:

1. Алгоритмы
  2. Визуализация
  3. Интеграция
  4. Удобство проектирования
  5. Просмотр отчетов
- 

# Алгоритмы анализа



# Проблемы моделирования

1. Подбор структуры модели
2. Отбор значимых факторов
3. Защита от переобучения
4. Интерпретация результатов



### Настройка конечных классов

Состояние входа:  Активировано

Столбец	Индекс F
12 Срок кредита	2,38
90 Сумма кредита	2,26
90 <b>Стоимость кредита</b>	<b>2,26</b>
ab Отрасль предпри...	0,88
12 Время работы пр...	0,74
90 Площадь квартиры	0,41
12 Срок эксплуатац...	0,39
12 Возраст	0,29
ab Основное напра...	0,29
12 Среднемес. расход	0,23

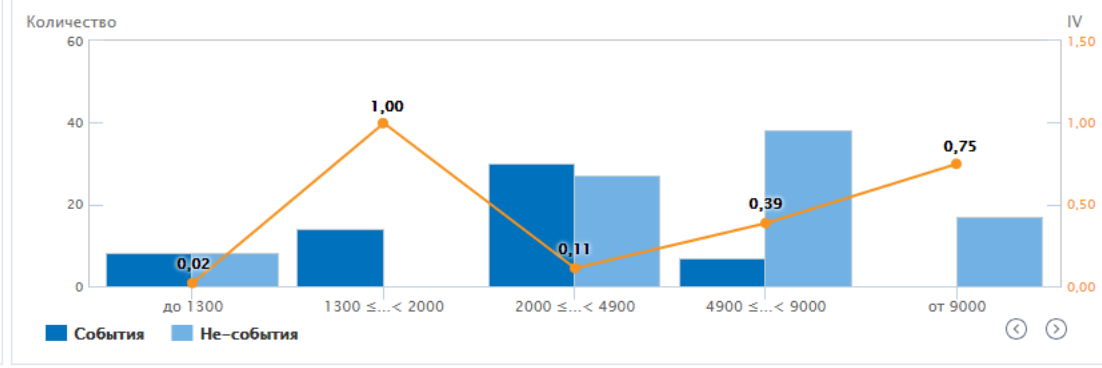
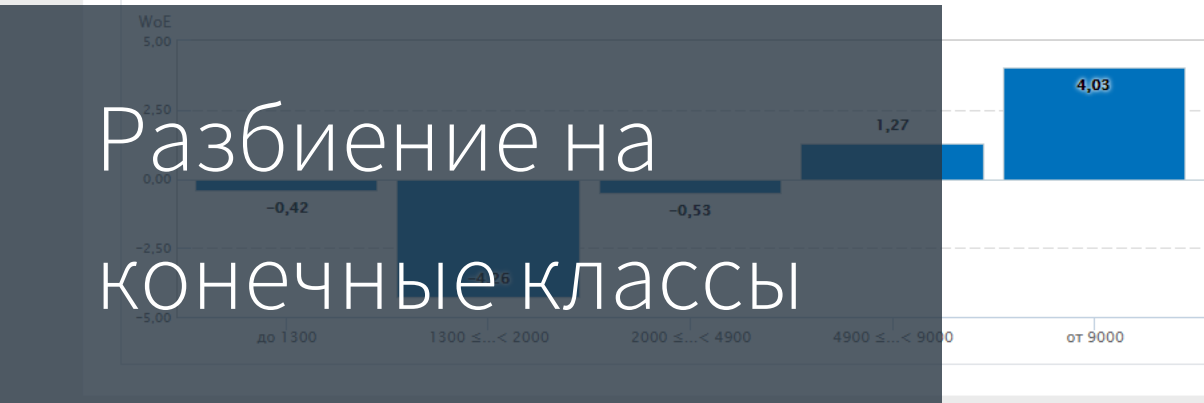
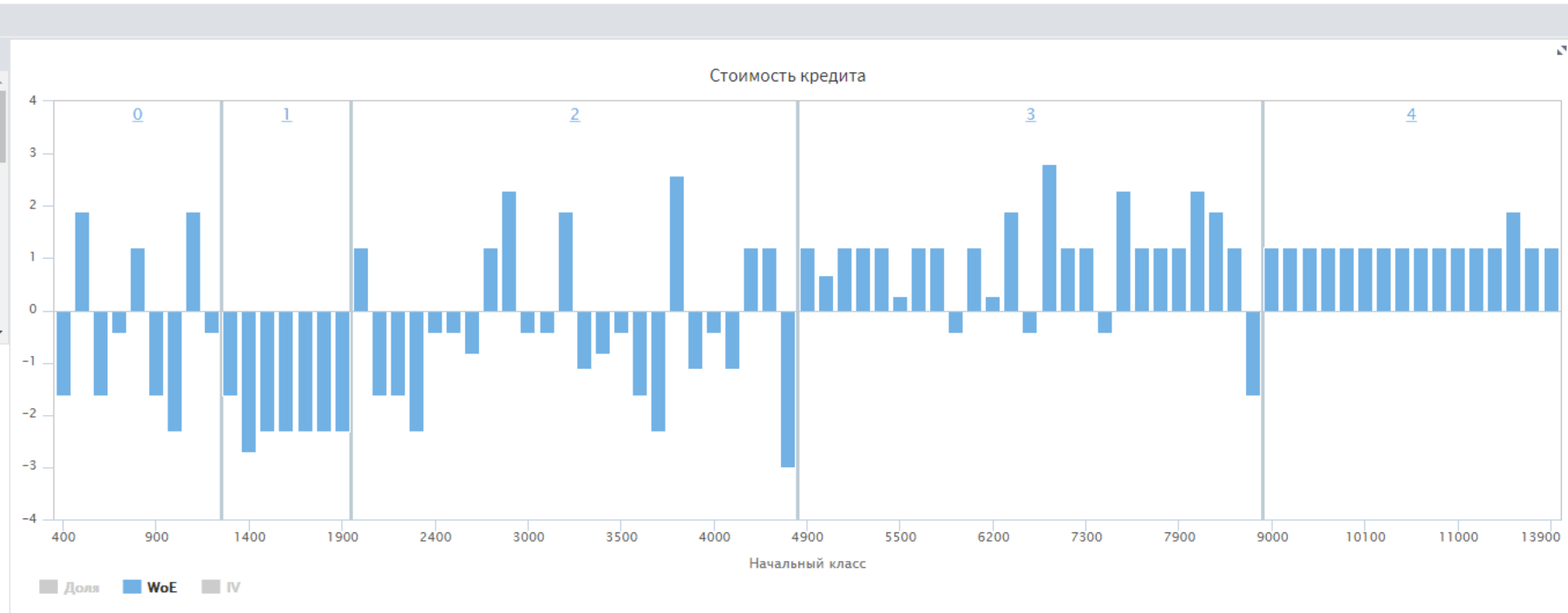
**Конечные классы**

Минимальный вес, %

Максимальное кол-во

**Оптимизация**

Равномерность, %



# Разбиение на конечные классы

Настройка

Пакеты | Скоринг | Модуль1 | Сценарий | Логистическая регрессия | Настройка

### Настройка логистической регрессии

Тип события: Более редкое

Индекс заданного события: 0

Автоматическая настройка:

Приоритет автоматической настройки: [Slider]

Отбор факторов и защита от переобучения: Elastic-Net

Настройки приоритетов

- Приоритет точность/скорость: [Slider]
- Приоритет точные/недоверенные данные: [Slider]
- Приоритет меньше/больше факторов: [Slider]

Денормализовать коэффициенты модели:

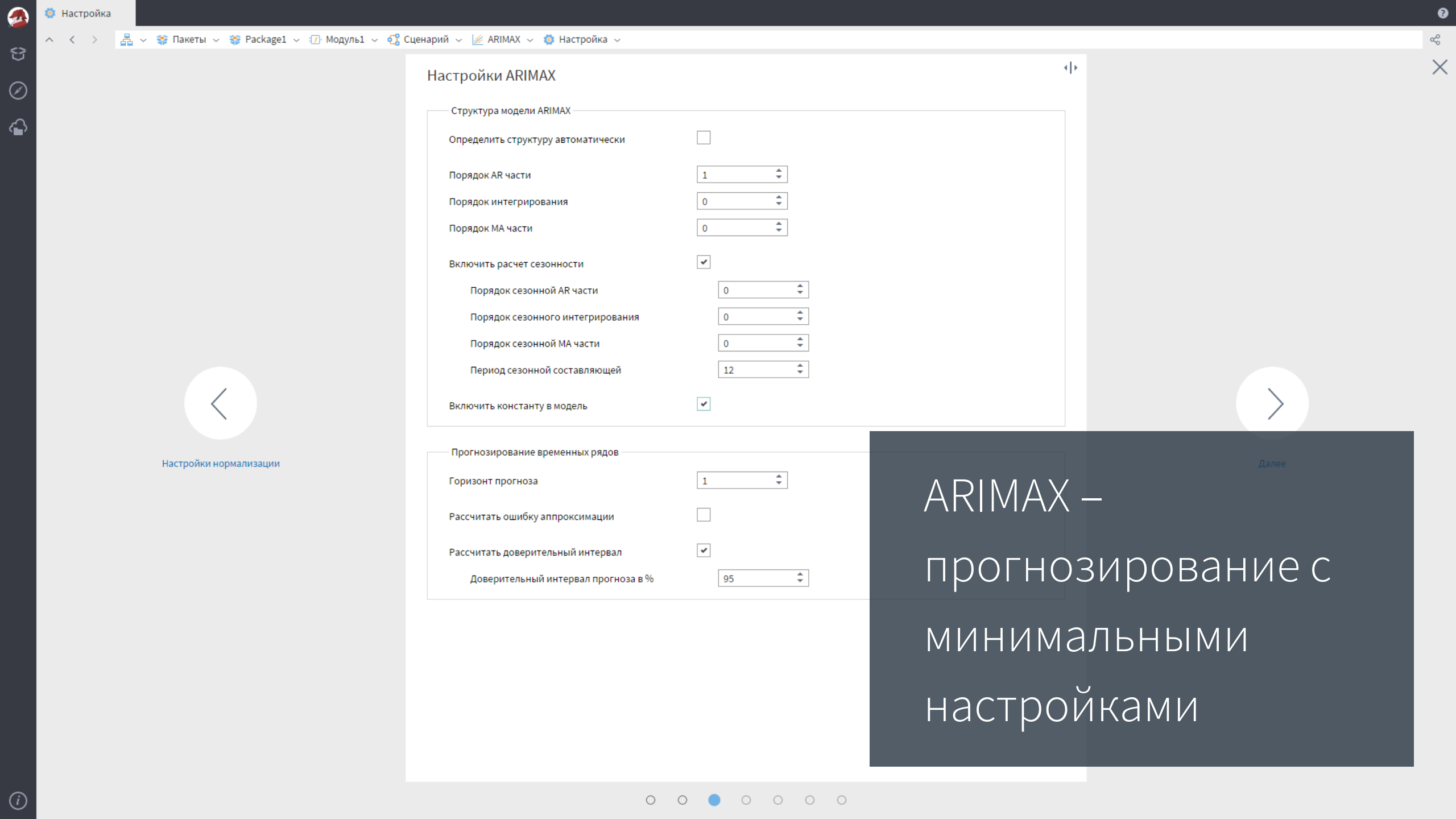
Использовать детальные настройки:

Разбиение на множества

Далее

Линейная и логистическая регрессия с отбором значимых факторов и защитой от переобучения





## Настройки ARIMAX

### Структура модели ARIMAX

Определить структуру автоматически

Порядок AR части

1

Порядок интегрирования

0

Порядок MA части

0

Включить расчет сезонности

Порядок сезонной AR части

0

Порядок сезонного интегрирования

0

Порядок сезонной MA части

0

Период сезонной составляющей

12

Включить константу в модель

### Прогнозирование временных рядов

Горизонт прогноза

1

Рассчитать ошибку аппроксимации

Рассчитать доверительный интервал

Доверительный интервал прогноза в %

95

Настройки нормализации

Далее

ARIMAX –  
прогнозирование с  
минимальными  
настройками

# Нейронные сети

- Автоподбор структуры
- Определение степени регуляризации
- Оценка значимости входных полей для данной модели нейросети

A yellow square containing the letters 'JS' in a bold, black, sans-serif font.

JS

Поддержка JavaScript:

- Безопасный код
- Встраивание в поток обработки данных
- Загрузка внешних модулей

Поддержка Python скоро...

# Запуск внешней программы

1. Любое внешнее приложение
2. Встраивание в поток обработки данных
3. Передача параметров

Потенциально опасная операция, требует явного предоставления прав.

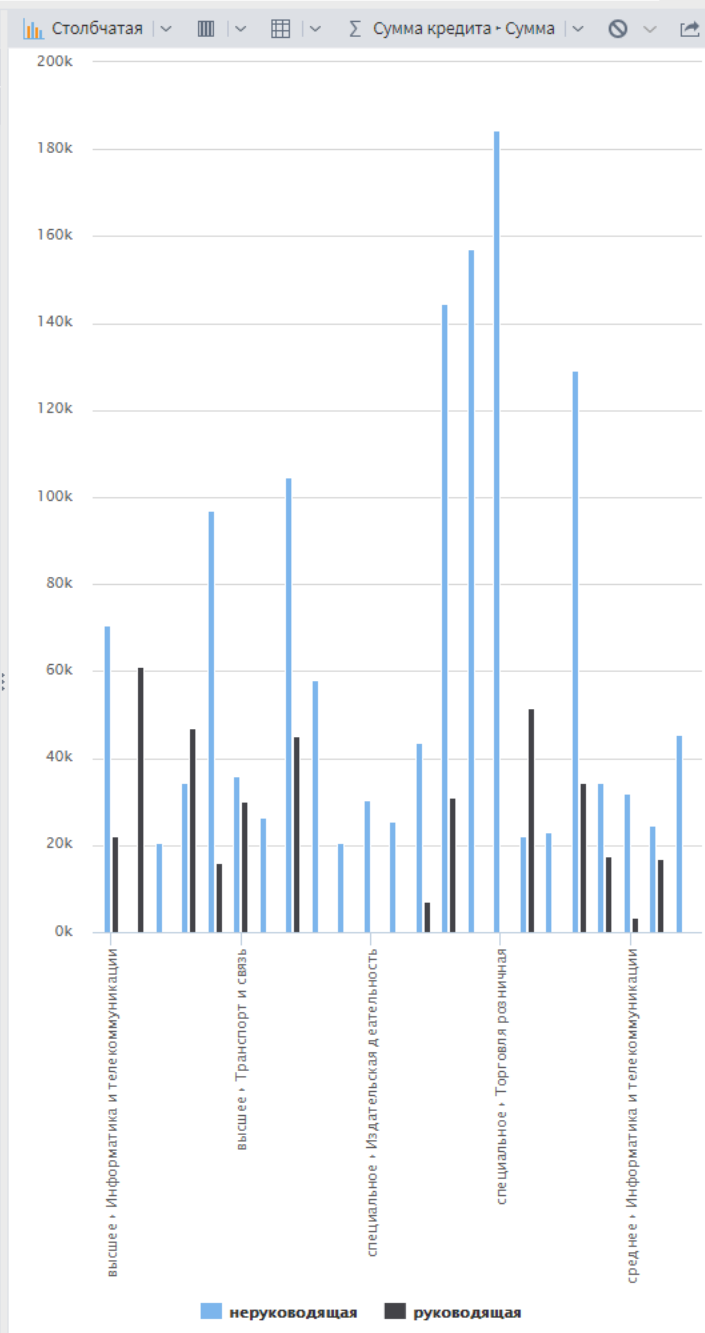
# Визуализация





- Поля
- 9.0 Сумма кредита
- 9.0 Стоимость кредита
- 12 Срок кредита
- ab Цель кредитования
- 12 Количество
- 12 Возраст
- ab Пол
- ab Образование
- ab Частная собственность
- ab Квартира
- 9.0 Площадь квартиры
- ab Способ приобретения собств.
- ab Расположение
- ab Машина
- 12 Срок эксплуатации машины
- ab Загородный дом
- ab Земельный участок
- ab Прописка в данном районе
- ab Гараж
- ab Класс предприятия
- 12 Время работы предприятия
- ab Отрасль предприятия
- ab Специализация
- ab Должность
- 12 Срок работы на предприятии
- 12 Срок работы по специальности
- 9.0 Среднемесячный доход
- 9.0 Среднемесячный расход
- ab Основное направление расходов
- 12 Количество иждивенцев
- ab Гражданское состояние
- ab Занятость супруга
- 9.0 Срок проживания в браке
- ab Давать кредит

Машина		Специализация		Должность		Факты		
Образование	Отрасль пред...	Цель кредито...	неруководящая	руководящая	Итого:			
			Сумма кре...	Стоимос...	Сумма кре...	Стоимо...		
высшее	> Информатика и телеком...		70 500	14 100	22 000	4 400	92 500	18
	> Легкая и пищевая пром...				61 000	12 200	61 000	12
	> Машиностроение и мета...		20 500	4 100			20 500	4
	> Образование (гос.)		34 500	6 900	47 000	9 400	81 500	16
	> Торговля розничная		97 000	19 400	16 000	3 200	113 000	22
	> Транспорт и связь		36 000	7 200	30 000	6 000	66 000	13
	> Туризм		26 500	5 300			26 500	5
	> Финансы, банковское дело		104 500	20 900	45 000	9 000	149 500	29
	> Юридические и нотари...		58 000	11 600			58 000	11
	Итого:			447 500	89 500	221 000	44 200	668 500
специал...	> Здравоохранение (гос.)		20 500	4 100			20 500	4
	> Издательская деятельно...		30 500	6 100			30 500	6
	> Информатика и телеком...		25 500	5 100			25 500	5
	> Легкая и пи...			8 700	7 000	1 400	50 500	10
	> Машиностроение и мета...		144 500	28 900	31 000	6 200	175 500	35
	> Образование (гос.)		157 000	31 400			157 000	31
	> Торговля розничная		184 500	36 900			184 500	36
	> Транспорт и связь		22 000	4 400	51 500	10 300	73 500	14
	> Туризм		23 000	4 600			23 000	4
	> Юридические и нотари...		34 500	6 900	17 500	3 500	52 000	10
Итого:			814 500	162 900	141 500	28 300	956 000	191
среднее	> Информатика и телеком...		32 000	6 400	3 500	700	35 500	7
	> Химия, парфюмерия, фа...		24 500	4 900	17 000	3 400	41 500	8
	> Химия, парфюмерия, фа...		45 500	9 100			45 500	9
Итого:			102 000	20 400	20 500	4 100	122 500	24
Итого:			1 364 000	272 800	383 000	76 600	1 747 000	349



1. Опциональное автообновление
2. Детализация ячейки
3. Сворачивание/разворачивание



### Выбор диаграммы

- ROC-кривая
  - PR-кривая
  - Базовые показатели
  - Диаграмма точности
  - Диаграмма равновесия
  - % распознанных событий
  - Диаграмма роста
  - Диаграмма отклика
  - Gain-диаграмма
- Кумулятивная
- 10 диапазонов

**Множества**

Обучающее  Тестовое

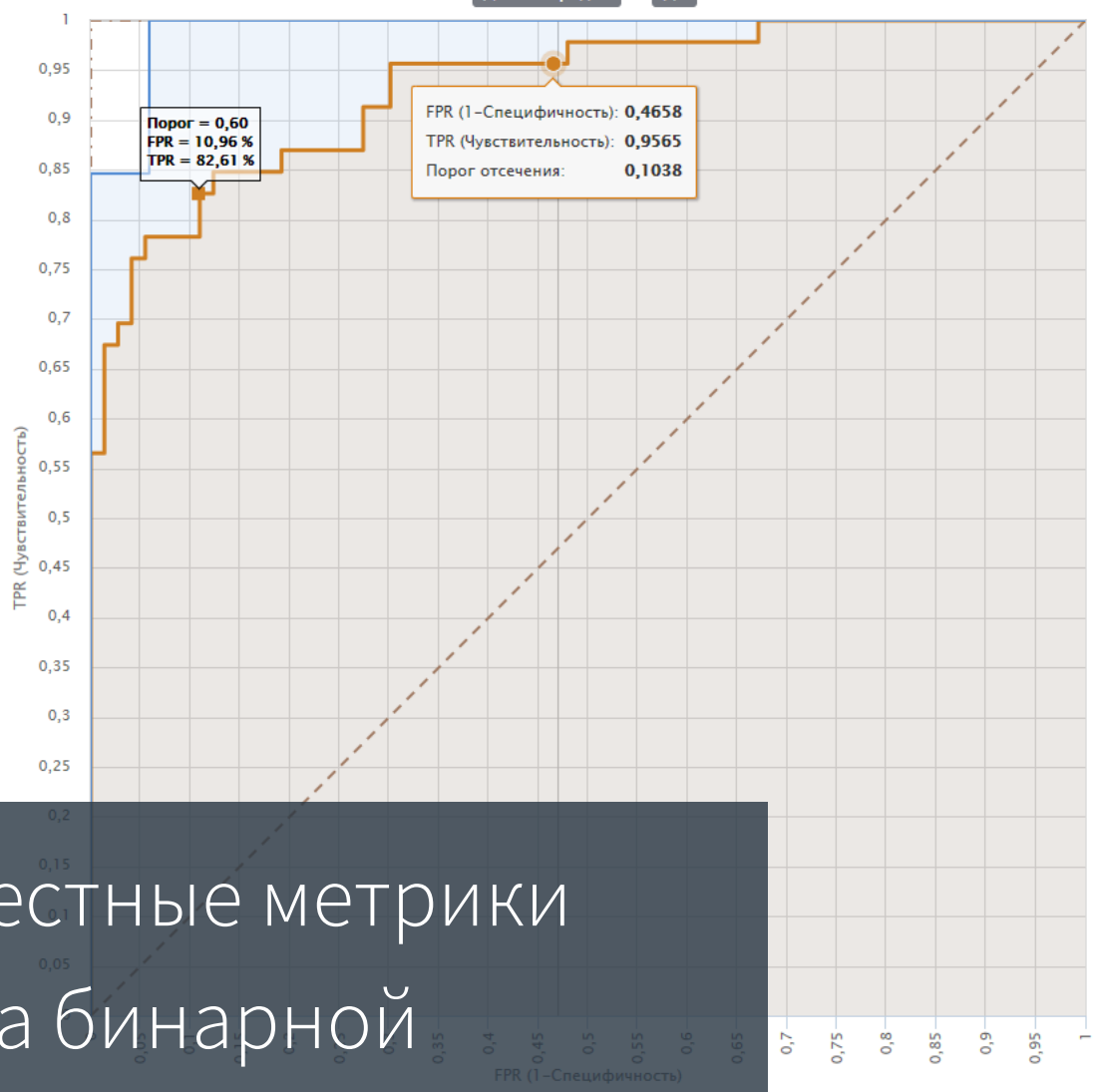
**Порог отсечения**

Заданный

Значение порога: 0,6

### ROC-кривая

Событие: Давать кредит = Да



Все известные метрики качества бинарной классификации

### Оценки классификации

Показатель	Множества	
	Обучающее	Тестовое
Оценки классификатора		
AUC ROC	0,9324	0,9910
AUC PR	0,9201	0,9881
Коэффициент Джини	0,8648	0,9819
KS	72,7814	94,1176
Порог отсечения		
Значение	0,6000	0,6000
TPR (Чувствительность)	0,8261	1,0000
TNR (Специфичность)	0,8904	0,9412
FPR (1-Специфичность)	0,1096	0,0588
PPV	0,8261	0,9286
F1 Score	0,8261	0,9630
MCC	0,7165	0,9349

### Матрицы ошибок

Классифицировано	Фактически		Итого
	Событие	Не-событие	
<b>Обучающее множество</b>	46	73	
Событие	38	8	46
Не-событие	8	65	73
<b>Тестовое множество</b>	13	17	
Событие	13	1	14
Не-событие	0	16	16

### Распознано

Обучающее множество	86,55%
Тестовое множество	96,67%



Модель	Показатель	Изменение поля	Поля
Нулевая модель	0,000000		
1.1	119,811766	90 Сумма кредита	90 Сумма кредита
1.2	100,908726	90 Стоимость кредита	90 Стоимость кредита
1.3	76,914883	12 Срок кредита	12 Срок кредита
1.3.1	95,349324	90 Сумма кредита	12 Срок кредита, 90 Сумма кредита
1.3.2	83,308895	90 Стоимость кредита	12 Срок кредита, 90 Стоимость кредита
1.3.3	70,614898	ab Цель кредитования	12 Срок кредита, ab Цель кредитования
1.3.4	76,914883	12 Количество	12 Срок кредита, 12 Количество
1.3.5	76,273237	12 Возраст	12 Срок кредита, 12 Возраст
1.3.6	75,669926	ab Пол	12 Срок кредита, ab Пол
1.3.7	76,333167	ab Образование	12 Срок кредита, ab Образование
1.3.8	76,720241	ab Частная собственность	12 Срок кредита, ab Частная собственность
1.3.9	76,798565	ab Квартира	12 Срок кредита, ab Квартира

Показатель		Значение
Константа	true	
-2 Log Likelihood		3,317456
R2		0,974947
R2 корр.		0,972954
Chi-квадрат		-129,099363
Число степеней свободы		88
Значимость		0,000000
AIC		21,317456
AICc		23,410479
BIC		44,396589
HQC		30,646418

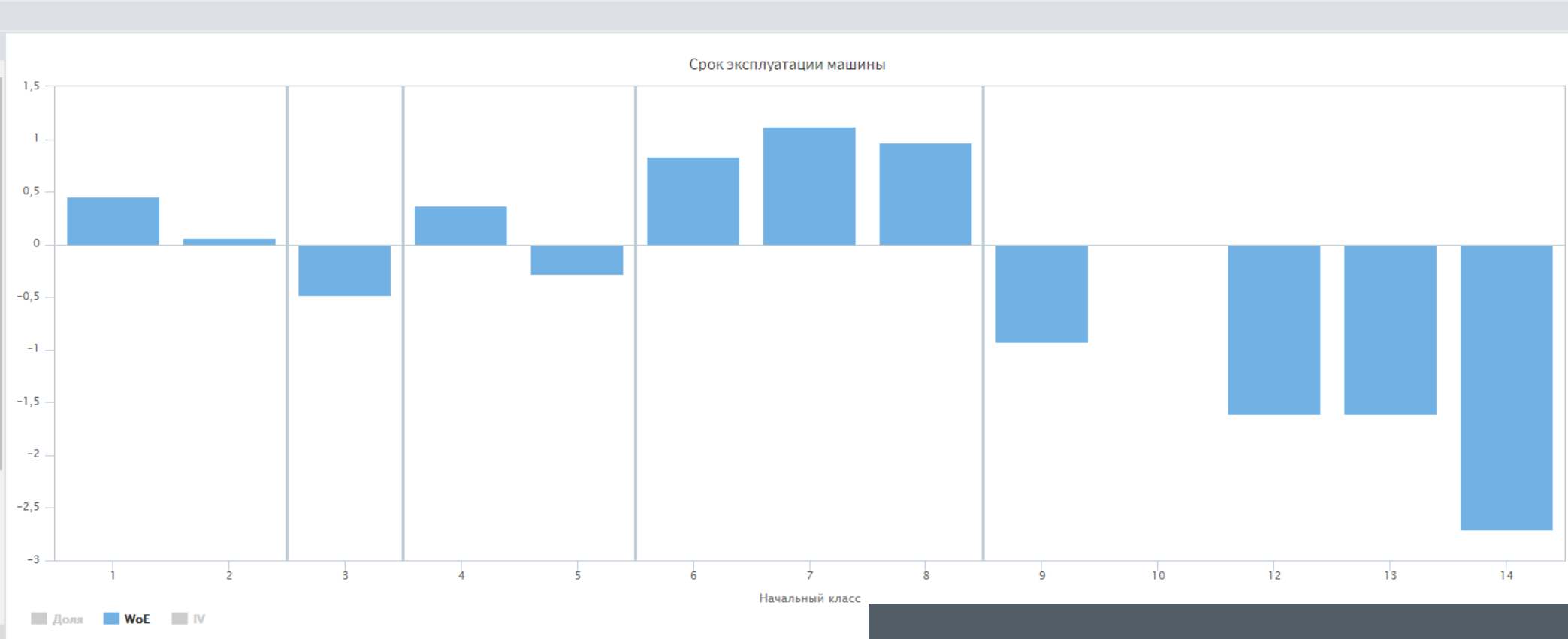
ab Давать кредит	Значение	Кол-во
Событие	Да	59
Не событие	Нет	90

Атрибут	Коэффициент	Стандартная ошибка	Коэффициент Вальда	Значимость	Отношение шансов	Нижняя граница ДИ	Верхняя граница ДИ	Балл
90 Константа	-56,644626	0,857740	4 361,196573	0,000000	0,000000	0,000000	0,000000	0
12 Срок кредита	-3,381617	0,000000	∞	1,000000	0,033992	0,033992	0,033992	0
ab Расположение								
ab область	-11,866497	0,000000	∞	1,000000	0,000007	0,000007	0,000007	0
12 Срок работы по специа...	1,911973	0,179559	113,383948	0,000000	6,766426	9,667869	4,735741	0
90 Среднемес. доход	0,025473	0,001356	353,029997	0,000000	1,025800	1,028567	1,023040	0
90 Среднемес. расход	-0,039792	0,005338	55,576844	0,000000	0,960989	0,971237	0,950849	0
ab Должность								
ab руководящая	-16,294962	0,000000	∞	1,000000	0,000000	0,000000	0,000000	0
12 Срок работы на предпр...	3,534989	0,000000	∞	1,000000	34,294629	34,294629	34,294629	0
ab Частная собственность								
ab Нет	6,258947	1,277368		0,000001	522,668246	6 617,329368	41,262832	0

Отчет по регрессии



Столбец	Индекс
90 Сумма кредита	2,26
90 Стоимость кредита	2,26
12 Срок кредита	2,38
ab Цель кредитования	0,13
12 Количество	0
12 Возраст	0,29
ab Пол	0,02
ab Образование	0,10
ab Частная собствен...	0,00
ab Квартира	0,00
90 Площадь квартиры	0,41
ab Способ приобрет...	0,17
ab Расположение	0,07
ab Машина	0,01
12 Срок эксплуатац...	0,39
ab Загородный дом	0,01
ab Земельный участок	0,04
ab Прописка в данн...	0,01
ab Гараж	0,01
ab Класс предприятия	0,02
12 Время работы пр...	0,74
ab Отрасль предпри...	0,88
ab Специализация	0,21
ab Должность	0,02
ab Срок работы на п...	0,05

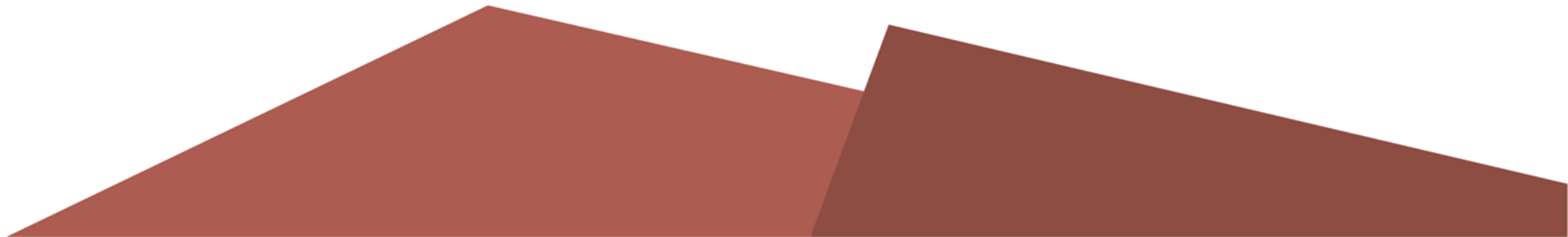


№	Метка	Нижняя	Верхняя	События	Не-события	Всего	Доля	Вес доказательства	Инф.индекс
0	до 3	-∞	3,00	13	25	38	26%	0,23	0,01
1	3 ≤...< 4	3,00	4,00	16	15	31	21%	-0,49	0,05
2	4 ≤...< 6	4,00	6,00	12	19	31	21%	0,04	0,00
3	6 ≤...< 9	6,00	9,00	6	25	31	21%	1,00	0,18
4	от 9	9,00	∞	12	6	18	12%	-1,12	0,15

Конечные классы – интерпретация логики разбиения на классы

Всего			59	90	149			0.39391448642...
-------	--	--	----	----	-----	--	--	------------------

# Интеграция

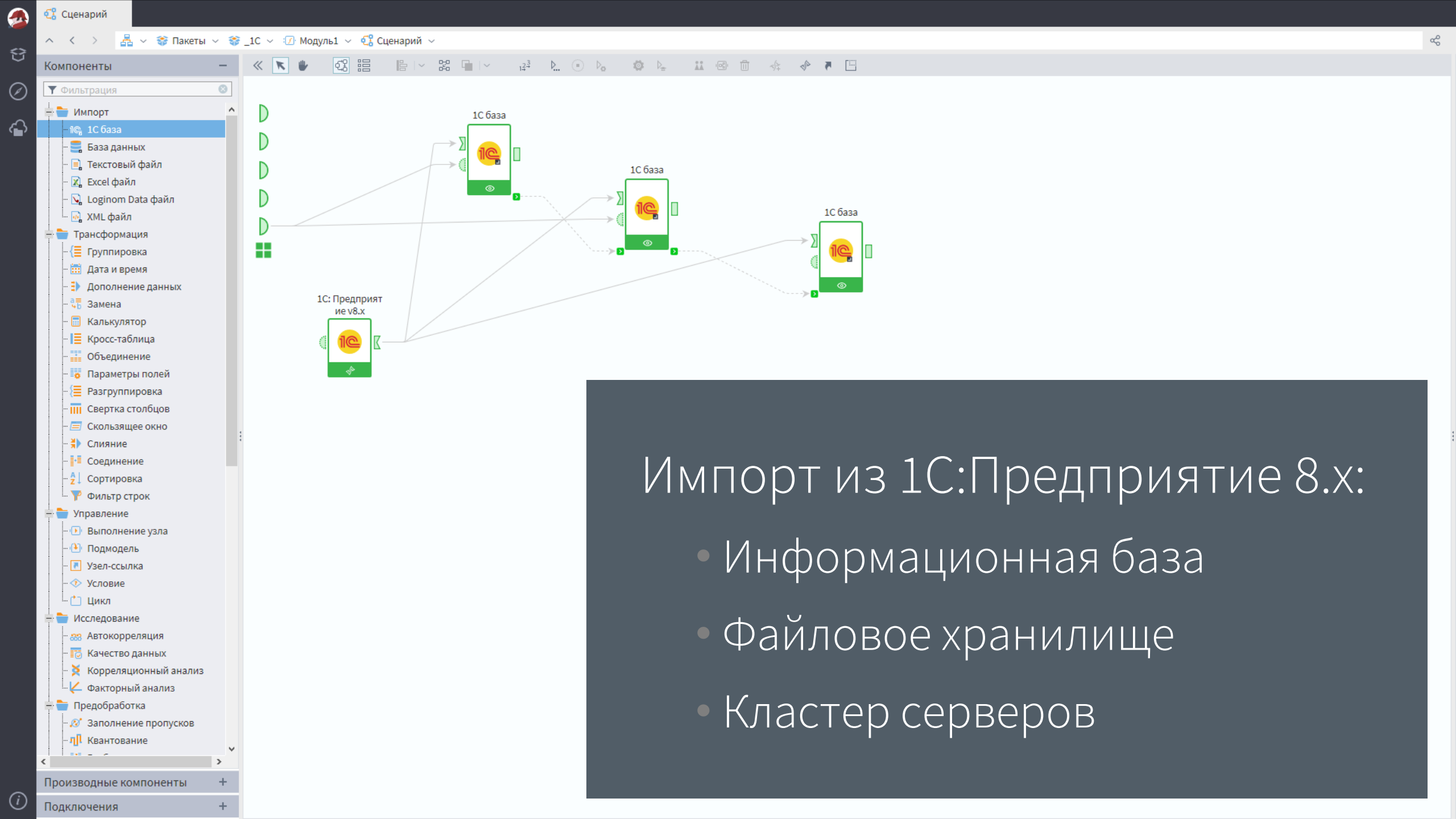


### Выбор объекта Хранилища данных

Подключение:  Активировано

Объекты хранилища данных	
Кубы	
Продажи за последние 12 месяцев	
Продажи за последние 30 дней	
Потоварные продажи за последн...	За 3 полных месяца
Недельные продажи	Недельные продажи в количестве
Процессы	
Дневные продажи	Информация по продажам
Поставки товаров	Товар, который был доставлен
Товары в пути	
Остатки	
Движение товара	
Заявки	Заявки формируемые поставщику
Изменения продажных цен	
Заявка от покупателя	
Доставка	Стоимость доставки поставляемого товара
Прогноз продаж для планирования	
Прогноз по группам	
План продаж	
Исключение аномальных продаж	
Восстановленные продажи	
Прогноз по товарам	Результаты прогноза по товарам
Прогноз продаж	Скорректированный прогноз продаж по товарам
Оценка качества прогноза по гру...	
Лучший ретро-прогноз	
=====Неиспользуе...	

Хранилище данных:  
оперирование бизнес-терминами для  
извлечения данных



## Импорт из 1С:Предприятие 8.x:

- Информационная база
- Файловое хранилище
- Кластер серверов

# Вызов REST-сервиса

## Функции:

- Настройка мастером
- GET и POST запросы
- Разбор ответа в сценарии

## Применение:

- Передача документов
- Обмен JSON

# odant



Российская  
технологическая  
платформа:

1. Распределенная
2. Объектная

Viewer

Пользователи

Администрирование Пользователи

Навигация

- Начало
- Пакеты
- Файлы
- Администрирование
  - Пользователи
  - Общие папки
  - Параметры

Добавить Редактировать Удалить Обновить

Логин	Полное имя	Администратор	Просмотр отчетов	Доступ ко всем файлам	Интерактивный режим	Режим службы	Заблокирован
admin		•			•		
user					•		
service				•		•	
viewer			•		•		

Редактирование пользователя

Логин: viewer

Полное имя: Иванов И.И. (менеджер по продажам)

Пароль: .....

Администратор

Просмотр отчетов

Полный доступ к файловому хранилищу

Вход в интерактивном режиме

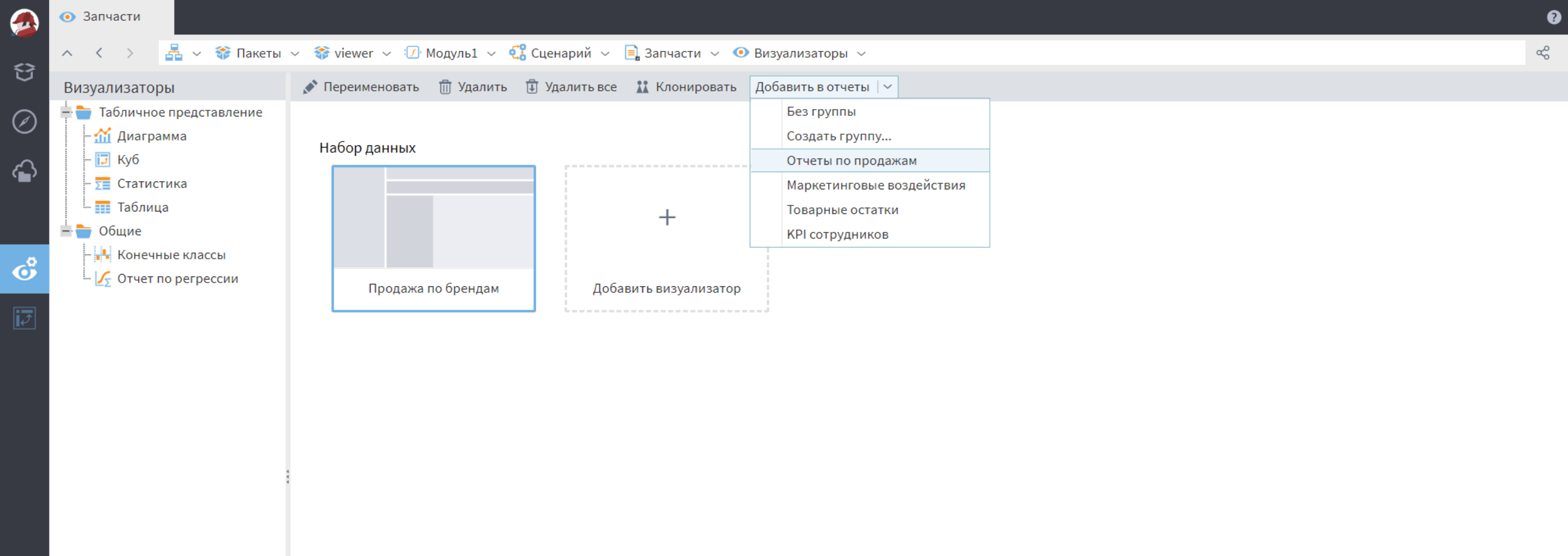
Вход в режиме службы

Заблокировать учётную запись

Сохранить [Отменить](#)

Новая роль – только  
для просмотра отчетов



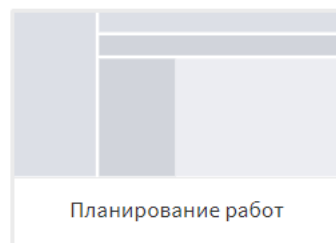
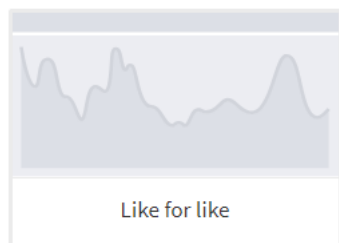
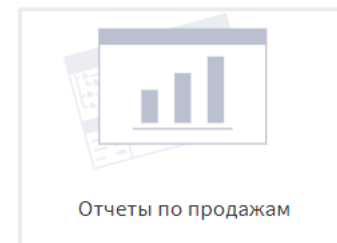
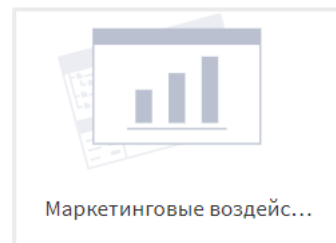
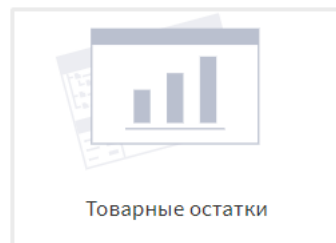
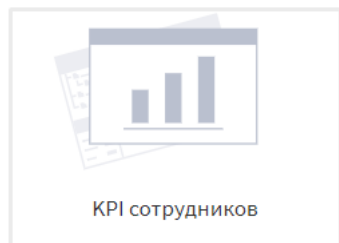


Вынесение отчетов на  
отдельную панель



Навигация

- Начало
- Пакеты
  - viewer
    - Отчеты**
      - Отчеты по продажам
        - Региональные продажи
        - Продажи в разрезе регионов
        - Статистика по SKU
        - Возврат продукции
        - Продажа по брендам
      - Маркетинговые воздействия
        - Адресные рассылки
        - Анализ каналов
        - Сводка по воздействиям
        - Отклик на рассылки
        - Конверсия
      - Товарные остатки
        - Остатки по складам
        - Оборачиваемость товаров
        - Неликвиды
        - ABC анализ
      - KPI сотрудников
        - Выполнение планов продаж
        - Табель
      - Like for like
      - Планирование работ



Соккрытие информации о сценариях, просмотр только подготовленных отчетов

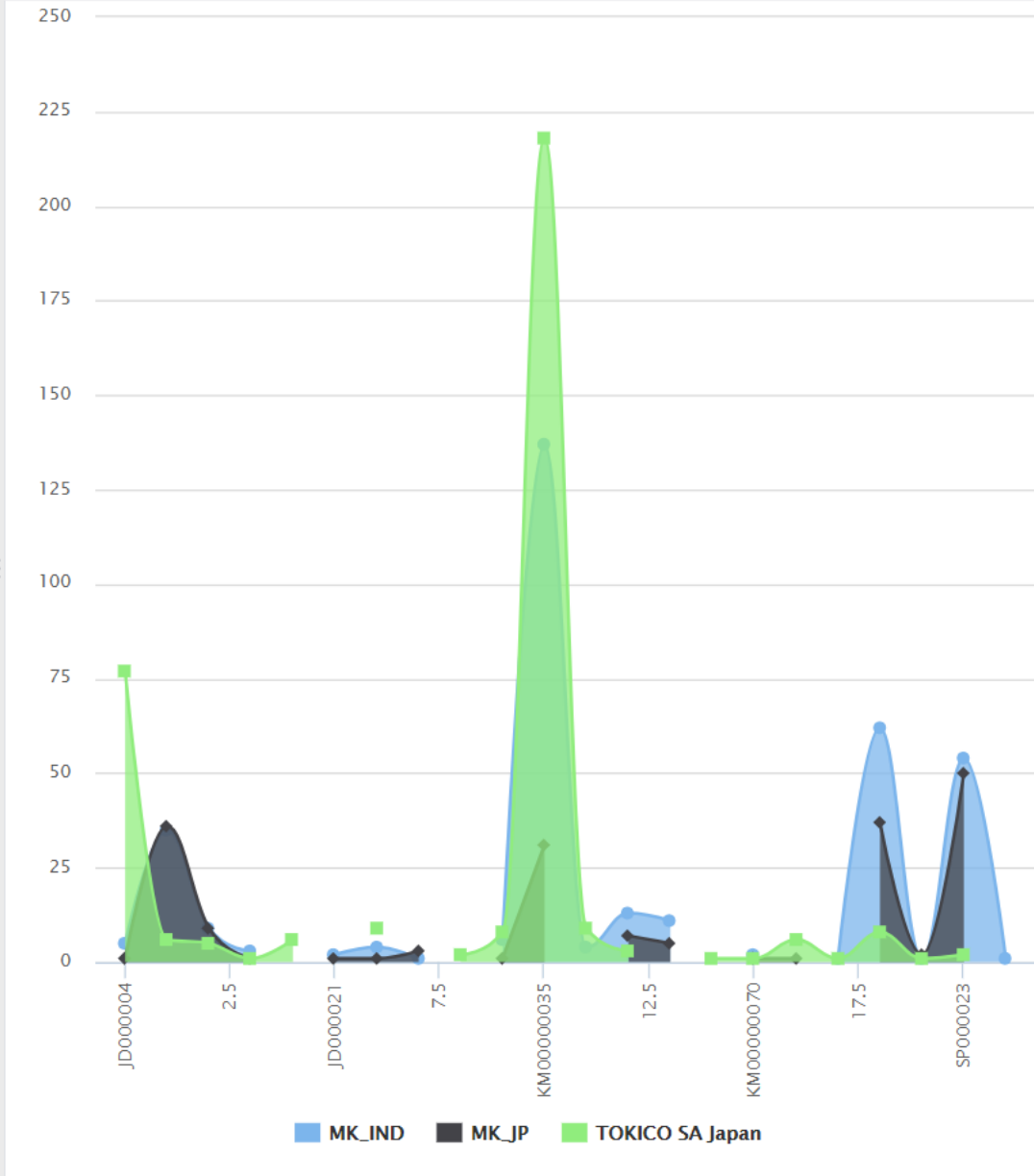


Продажа по брендам



Покупатель [...]	Товар   Бренд	MK_IND	MK_JP	TOKICO SA Ja...	Итого:
	JD000004	5	1	77	83
	JD000005	35	36	6	77
	JD000006	9	9	5	23
	JD000007	3	1	1	5
	JD000010			6	6
	JD000021	2	1		3
	KM00000...	4	1	9	14
	KM00000...	1	3		4
	KM00000...			2	2
	KM00000...	6	1	8	15
	KM00000...	137	31	218	386
	KM00000...	4		9	13
	KM00000...	13	7	3	23
	KM00000...	11	5		16
	KM00000...			1	1
	KM00000...	2	1	1	4
	KM00000...			1	1
	KM00000...	1		1	2
	SP000002	62	37	8	107
	SP000023	54	50	2	106
	SP000025	1			1
Итого:		251	187	264	902

Просмотр результатов анализа с возможностью манипулирования



Учебный центр



# Учебные модули

- Бизнес-аналитика и наука о данных
- Основы работы в Logiном
- Стандартизация и очистка НСИ
- Обнаружение и измерение оттока
- Введение в кредитный скоринг

#### ▼ Лекция 4. Подготовка данных к анализу

Особенности бизнес-данных. Формализация данных. Методы сбора данных. Информативность данных. Требования к данным.

#### ▼ Лекция 5. Технологии бизнес-аналитики

Обнаружение данных - Knowledge Discovery. Схема шагов KDD. Data Mining. Значение и перевод термина. Модели и задачи Data Mining. Ассоциация. Кластеризация. Классификация и регрессия. Описательные и предсказательные модели. Связь Data Mining с другими областями. Машинное обучение, глубокое обучение (определения). Причины популярности Data Mining

#### ▼ Лекция 6. Прикладные задачи бизнес-аналитики

Где применяется бизнес-аналитика? Традиционный BI. Описательное моделирование. Предсказательное моделирование. Задачи бизнес-аналитики в банках. Консолидированная отчетность. Кредитный скоринг. Скоринговая модель, карта. Виды кредитного скоринга. Управление инкассациями. Кредитный конвейер. Сегментация заемщиков. Задачи бизнес-аналитики в торговле. Шаблоны покупок. Скоринги откликов. Управление маркетинговыми кампаниями. Оптимизация запасов. Задачи бизнес-аналитики в телекоме. Скоринг удержания клиентов. Кривые выживания.

# Курс «Бизнес-аналитика и наука о данных»

- 21 лекция
- 20 практических занятий
- Мастер-классы
- Тренажеры
- Контрольные задания
- Итоговое тестирование

[loginom.ru](http://loginom.ru)

