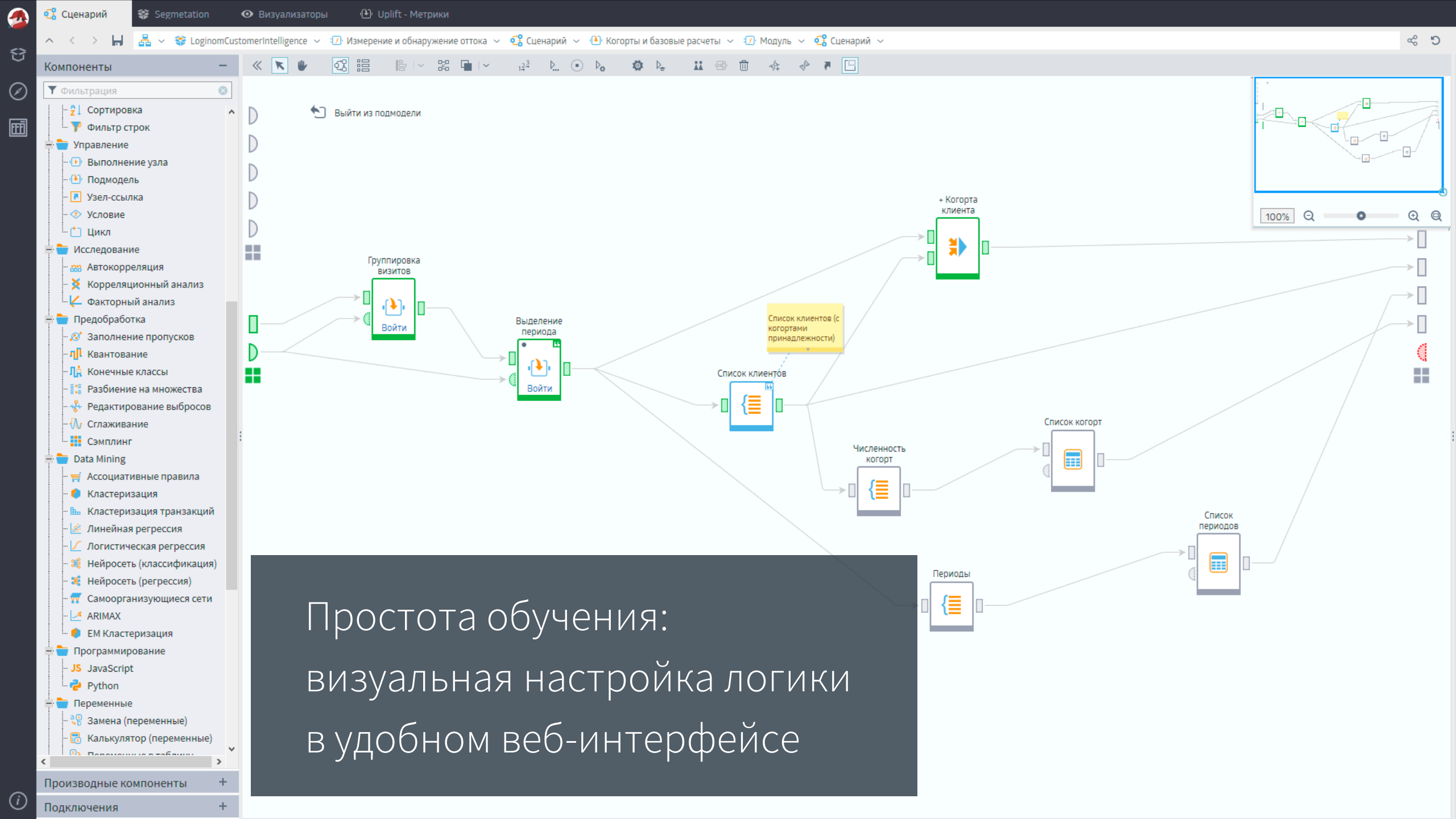


Logiном 6.4: НОВЫЕ ВОЗМОЖНОСТИ



Loginom

Loginom – платформа,
позволяющая бизнес-
экспертам самостоятельно,
без программирования,
проводить сложную
аналитику.





Модуль1

clean_doc

История продаж

ДемоPackage Модуль1 Сценарий История продаж

Компоненты

Производные компоненты

Трансформация

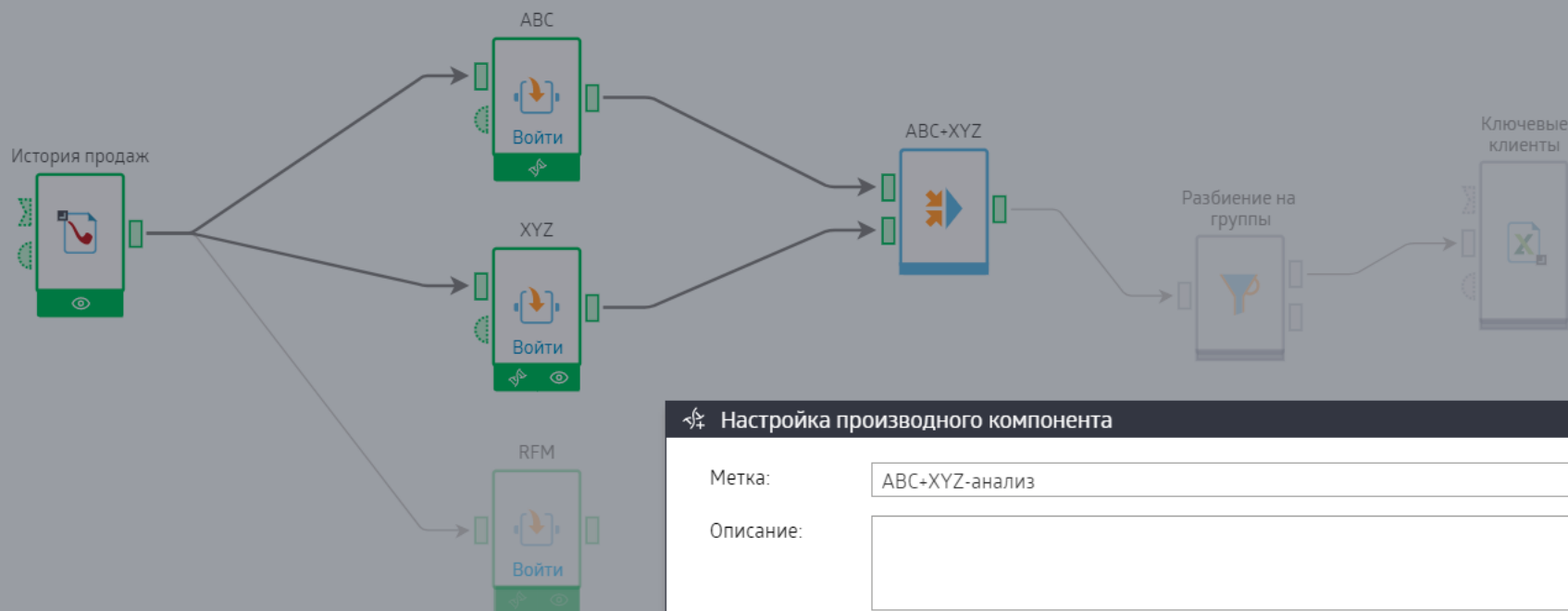
ABC+XYZ #1

Управление

ABC

RFM

XYZ



Настройка производного компонента

Метка: ABC+XYZ-анализ

Описание:

Область видимости

- ☐ Закрытый (доступен только внутри текущего модуля)
- ☐ Частично закрытый (доступен только внутри текущего модуля и его подмоделей)
- ☐ Внутренний (доступен только внутри текущего пакета)
- ☒ Открытый (доступен во всех пакетах)

Сохранить

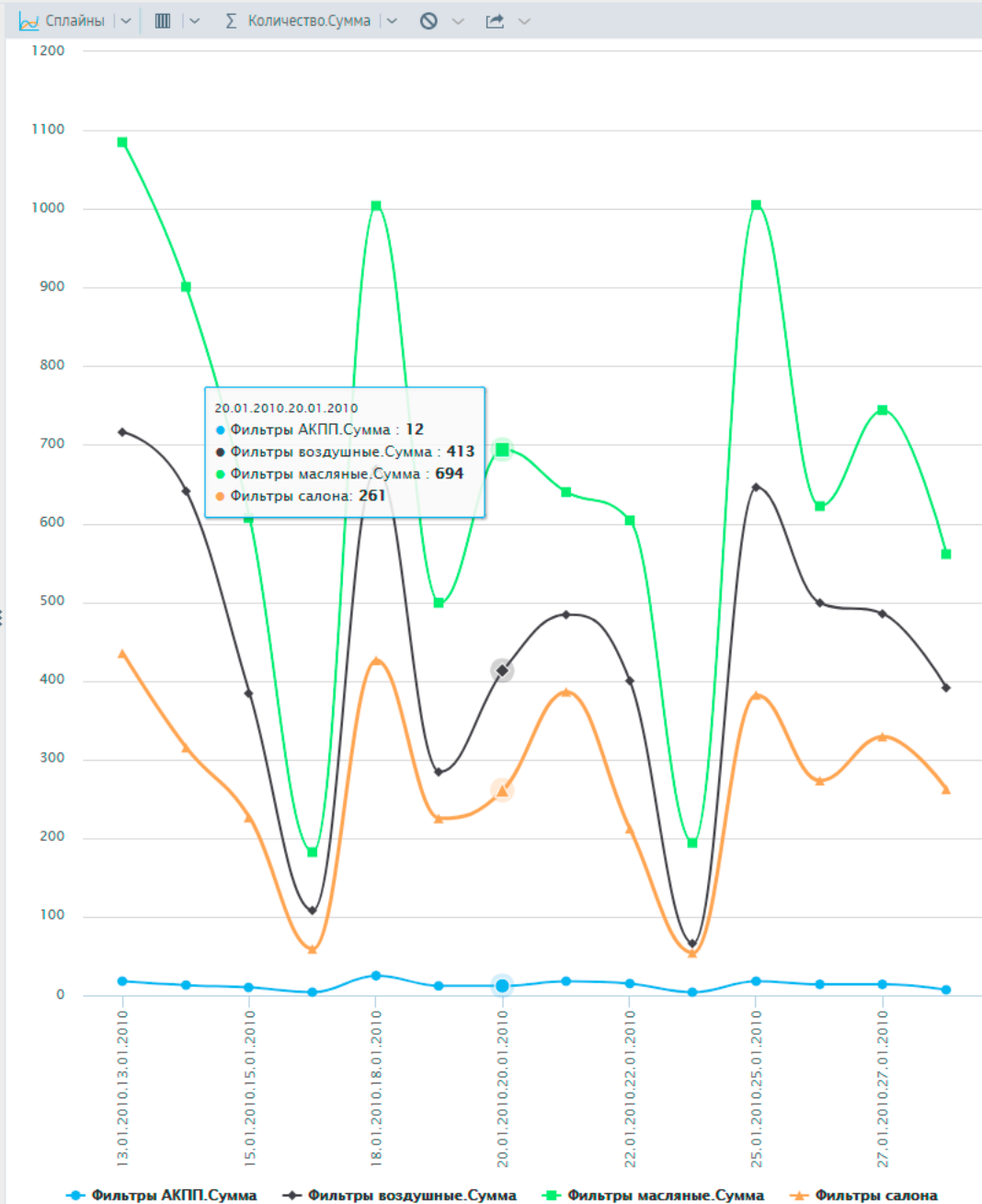
[Отмена](#)

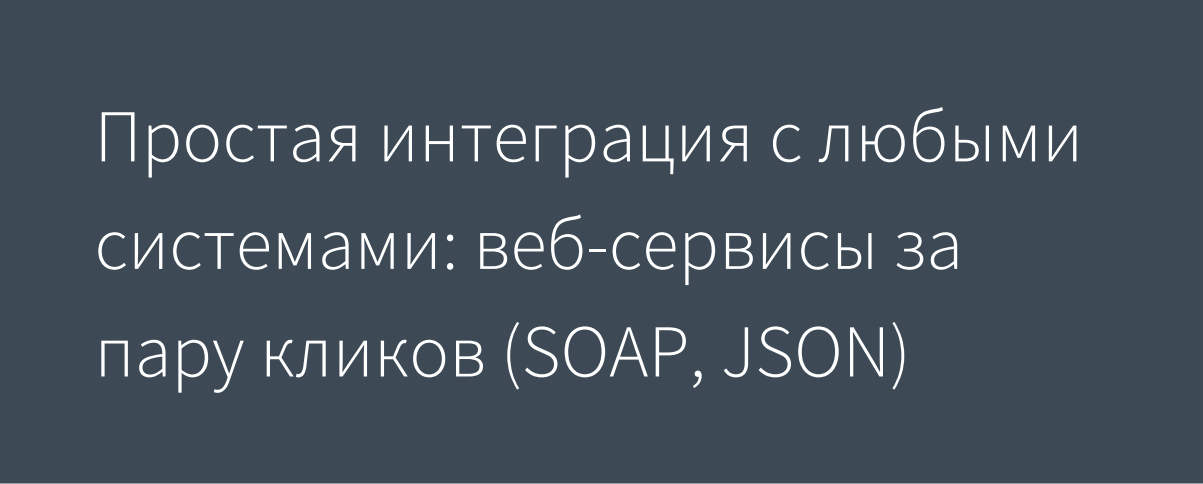
Подключения



Дата	Группа товара	Σ Факты		Фильтры АКПП		Фильтры воздушные		Фильтры масляные		Фильтры салона	
		Сумма	Количес...	Сумма	Количес...	Сумма	Количес...	Сумма	Количес...	Сумма	Количес...
		# Колич...	Σ Сумма	# Колич...	Σ Сумма	# Колич...	Σ Сумма	# Колич...	Σ Сумма	# Колич...	Σ Сумма
> 13.01.2010		9	18,00	214	716,00	179	1 085,00	137	435,00		
> 14.01.2010		9	13,00	191	641,00	157	901,00	116	315,00		
> 15.01.2010		8	10,00	150	384,00	129	607,00	92	226,00		
> 16.01.2010		4	4,00	64	108,00	61	182,00	50	59,00		
> 18.01.2010		15	25,00	251	669,00	188	1 004,00	145	426,00		
> 19.01.2010		9	12,00	138	284,00	119	499,00	111	225,00		
> 20.01.2010		8	12,00	145	413,00	126	694,00	115	261,00		
> 21.01.2010		6	18,00	175	484,00	143	640,00	127	386,00		
> 22.01.2010		10	15,00	122	400,00	131	604,00	84	212,00		
> 23.01.2010		2	4,00	40	66,00	52	194,00	32	54,00		
> 25.01.2010		9	18,00	199	646,00	176	1 005,00	135	382,00		
> 26.01.2010		7	14,00	169	499,00	151	622,00	107	273,00		
> 27.01.2010		8	14,00	174	485,00	164	744,00	115	329,00		
> 28.01.2010		7	7,00	140	391,00	146	561,00	102	262,00		

Удобная визуализация: OLAP и другие способы отображения







70+

НОВЫХ фиЧ

1. Обработчики
2. Визуализация
3. Проектирование сценариев
4. Отладка, мониторинг
5. Хранилища данных
6. Оптимизация работы
7. Языки программирования

Обработчики и визуализаторы





Настройка входных столбцов

Метка	Имя	Вид данных	Назначение
12 Срок кредита	COL3	Непрерывный	Входное
31 Дата кредитования	COL4	Непрерывный	Не задано
ab Цель кредитования	COL5	Дискретный	Входное
12 Количество	COL6	Непрерывный	Входное
12 Возраст	COL7	Непрерывный	Входное
ab Пол	COL8	Дискретный	Входное
ab Образование	COL9	Дискретный	Входное
ab Частная собственность	COL10	Дискретный	Входное
ab Квартира	COL11	Дискретный	Входное
9.0 Площадь квартиры	COL12	Непрерывный	Входное
ab Способ приобретения собств.	COL13	Дискретный	Входное
ab Расположение	COL14	Дискретный	Входное
ab Машина	COL15	Дискретный	Входное
12 Срок эксплуатации машины	COL16	Непрерывный	Входное
ab Загородный дом	COL17	Дискретный	Входное
ab Земельный участок	COL18	Дискретный	Входное
ab Прописка в данном районе	COL19	Дискретный	Входное
ab Гараж	COL20	Дискретный	Входное
ab Класс предприятия	COL21	Дискретный	Входное
12 Время работы предприятия	COL22	Непрерывный	Не задано
ab Отрасль предприятия	COL23	Дискретный	Не задано
ab Специализация	COL24	Дискретный	Не задано
ab Должность	COL25	Дискретный	Не задано
12 Срок работы на предприятии	COL26	Непрерывный	Не задано
12 Срок работы по специальности	COL27	Непрерывный	Не задано
9.0 Среднемес. доход	COL28	Непрерывный	Не задано
9.0 Среднемес. расход	COL29	Непрерывный	Не задано
ab Основное направление расходов	COL30	Дискретный	Не задано
12 Количество иждивенцев	COL31	Непрерывный	Не задано
ab Гражданское состояние	COL32	Дискретный	Не задано
ab Занятость супруга	COL33	Дискретный	Не задано
9.0 Вес записи	COL34	Непрерывный	Вес
ab Давать кредит	COL35	Дискретный	Выходное
12 Давать кредит(число)	COL36	Дискретный	Не задано

Логистическая регрессия:
задание веса записи



Настройка логистической регрессии

Состояние входа

Не активировано

[Активировать](#)

Тип события

Более редкое

Индекс заданного события

0

Автоматическая настройка



Приоритет автоматической настройки

Отбор факторов и защита от переобучения

Ridge

Настройки приоритетов

Приоритет точность/скорость

Приоритет точные/недостоверные данные

Приоритет меньше/больше факторов

Денормализовать коэффициенты модели



Использовать детальные настройки



Показывать коэффициенты опорных категорий



Настройки калибровки

☐ На основе обучающего множества☐ На основе тестового множества☒ Ручная калибровка

Множество	Пропорции события	Пропорции не-события
Генеральное	0,66	0,34
Обучающее	0,40	0,60

Логистическая регрессия:
калибровка модели – поправки на
априорную вероятность





Выбор диаграммы

- ☒ ROC-кривая
- ☐ PR-кривая
- ☐ Базовые показатели
- ☐ Диаграмма точности
- ☐ Диаграмма равновесия
- ☐ % распознанных событий
- ☐ Диаграмма роста
- ☐ Диаграмма отклика
- ☐ Диаграмма выигрыша

☒ Кумулятивная

10 диапазонов

Множества

☒ Обучающее ☒ Тестовое

Порог отсекаения

Заданный

Значение порога: 0,5

Истинная доля событий

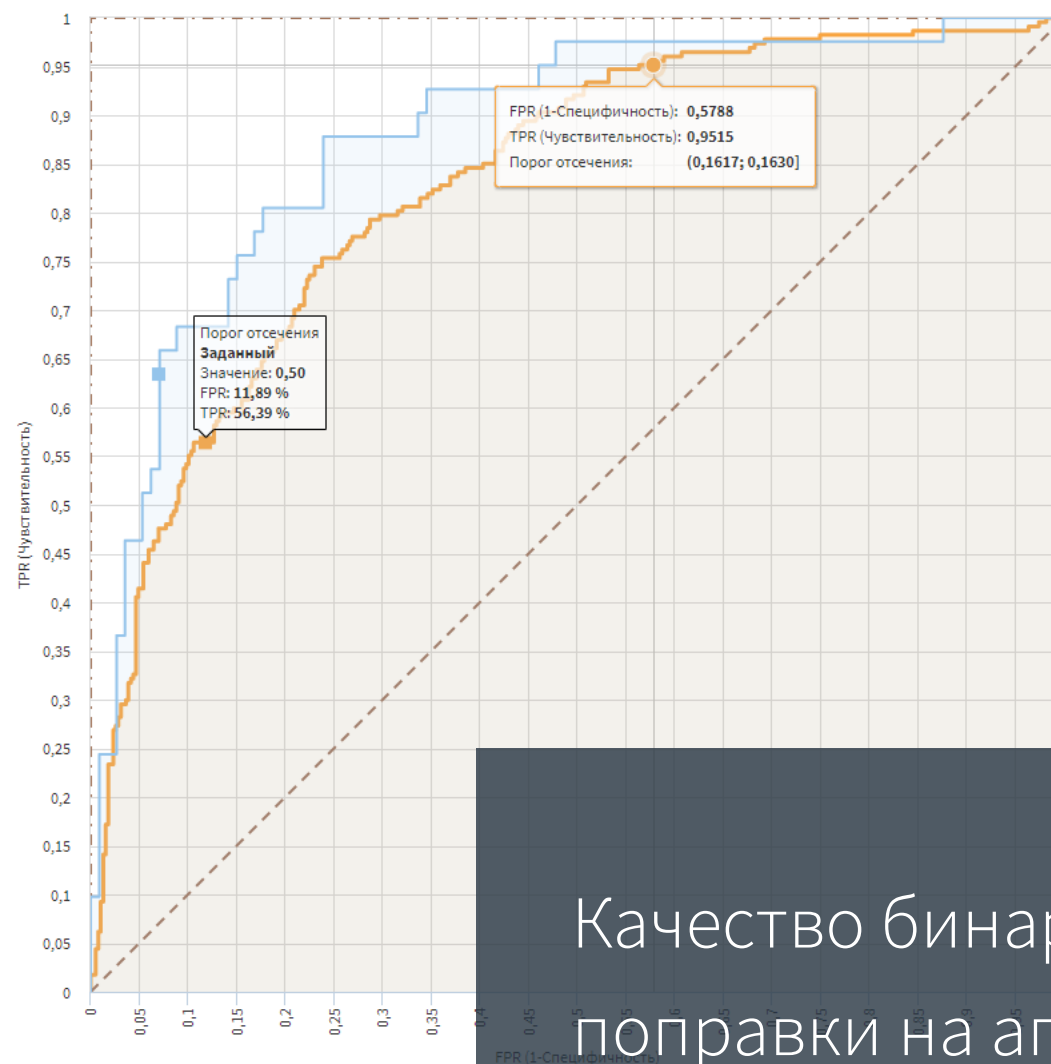
Текущая доля событий: 44 %

Задать вручную

Событий, %: 44

ROC-кривая

Событие: Отток = 1



Обучающее множество
Порог отсекаения
Базовая линия
Идеальная линия
Тестовое множество
Порог отсекаения

Оценки классификации

Показатель	Множества	
	Обучающее	Тестовое
Оценки классификатора		
AUC ROC	0,8280	0,8798
AUC PR	0,7274	0,7374
Коэффициент Джини	0,6561	0,7596
KS	51,5578	63,9111
Порог отсекаения: Заданный		
Значение	0,5000	0,5000
TPR (Чувствительность)	0,6476	0,6829
TNR (Специфичность)	0,8243	0,9115
FPR (1-Специфичность)	0,1757	0,0885
PPV	0,6837	0,7368
F1 Score	0,6652	0,7089
MCC	0,4775	0,6094

Матрицы ошибок

Классифицировано	Фактически		Итого
	Событие	Не-событие	
Обучающее	227	387	
... Событие	147	68	215
... Не-событие	80	319	399
Тестовое	41	113	
... Событие	28	10	38
... Не-событие	13	103	116

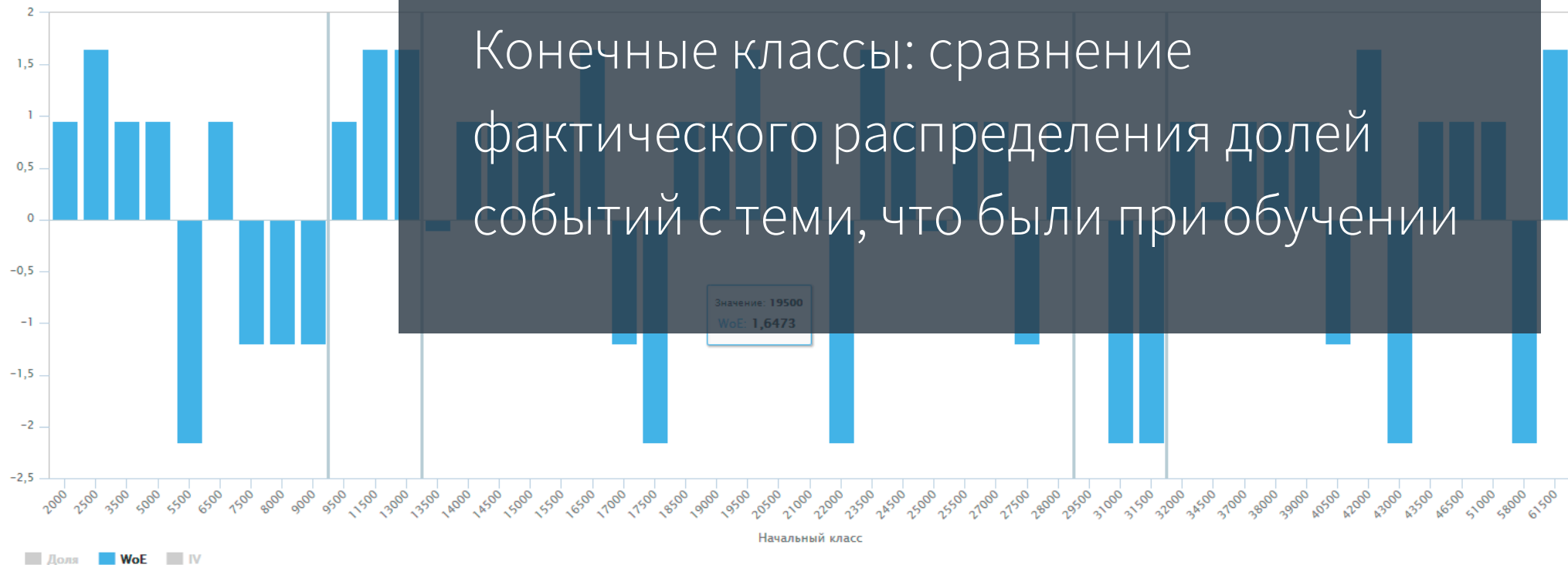
Распознано

Обучающее	466/614
Тестовое	131/154

Качество бинарной классификации:
поправки на априорную вероятность



Столбец	IV	IV (факт)
9.0 Сумма к...	0,63	0,57
9.0 Стоимос...	0,63	0,57
12 Срок кре...	0,07	0,09
12 Количес...	0	0
ab Квартира	0,00	0,00
9.0 Площад...	0,45	0,39
ab Способ п...	0,01	0,01
ab Машина	0,03	0,20
ab Земельн...	0,05	0,08
ab Гараж	0,00	0,00
ab Давать к...	0,09	0,11



№	Метка	Нижняя	Верхняя	Доля событий	Доля событий (факт)	Доля не-событий	Доля не-событий (факт)	Всего	Всего (факт)	Доля	Доля (факт)	Вес док...	Вес доказательства (факт)
0	до 9500	-∞	9 500,00	39,29%	30,77%	60,71%	69,23%	28	13	19%	18%	-0,53	-0,53
1	9500 ≤...< 13500	9 500,00	13 500,00	0,00%	0,00%	100,00%	100,00%	11	5	7%	7%	3,41	3,41
2	13500 ≤...< 29500	13 500,00	29 500,00	21,67%	20,59%	78,33%	79,41%	60	34	40%	46%	0,32	0,32
3	29500 ≤...< 32000	29 500,00	32 000,00	66,67%	100,00%	33,33%	0,00%	9	2	6%	3%	-1,66	-1,66
4	от 32000	32 000,00	∞	26,83%	20,00%	73,17%	80,00%	41	20	28%	27%	0,03	0,03

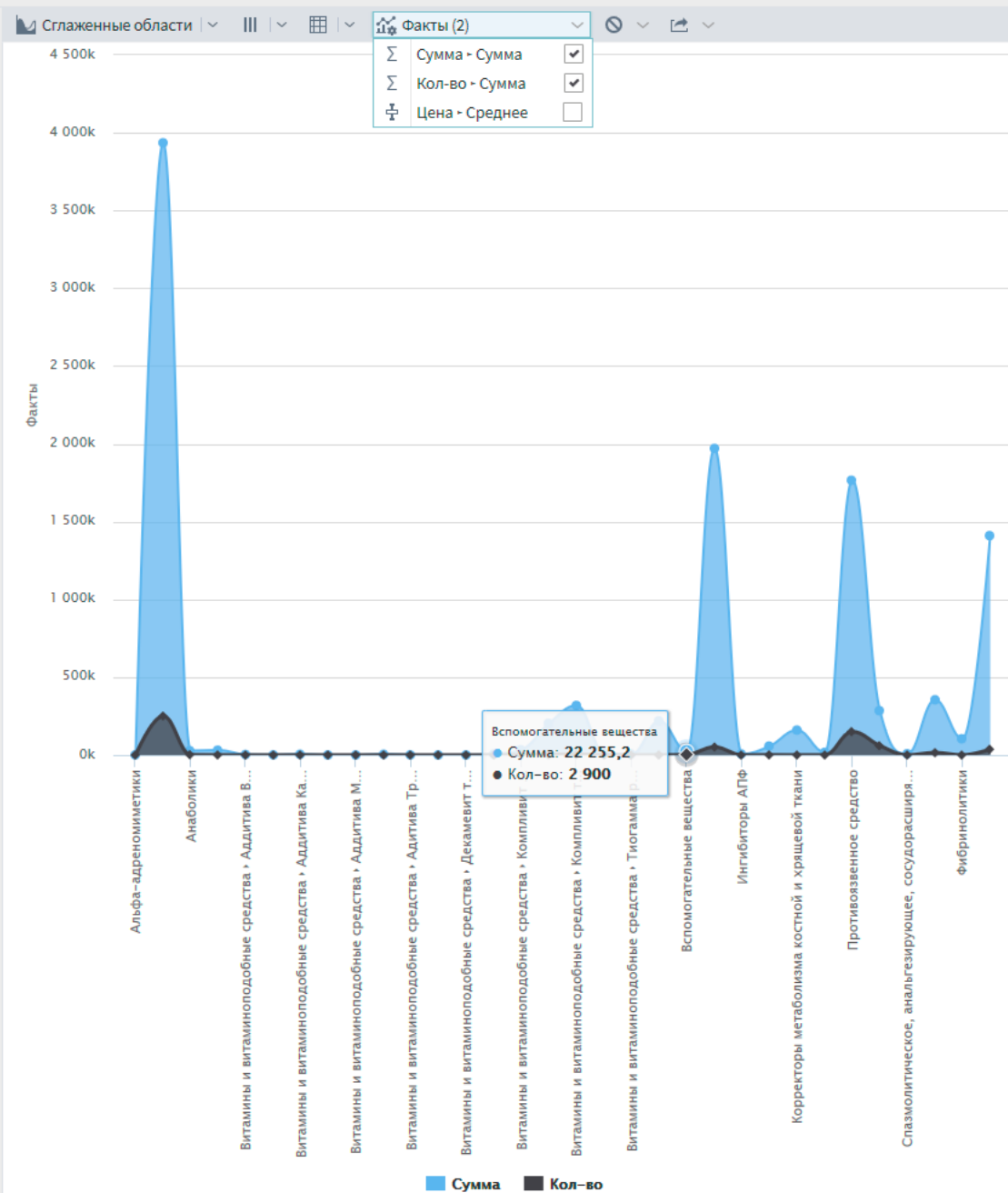
Всего				27,52%	27,52%	72,48%	72,48%	149	74				
-------	--	--	--	--------	--------	--------	--------	-----	----	--	--	--	--



#	12 Группа дубликата	12 Группа противоречия	12 Срок кредита	Пол	Машина	Площадь квартиры	Давать кредит	Образование	ab Способ приобретения собств.	ab Земельный участок	9.0 Сумма кредита
85	13		12	Муж	отечественная	24	Нет	специальное	другое	Нет	19000
86	7		12	Жен	отечественная	35	Нет	специальное	покупка	Да	17000
87			12	Муж	отечественная	21	Нет	высшее	другое	Да	22000
88			18	Муж	импортная	54	Да	специальное	другое	Нет	31000
89			6	Муж	импортная	44	Да	специальное	другое	Нет	12000
90			12	Муж	импортная	21	Нет	специальное	покупка	Да	13500
91			12	Жен	отечественная	15	Нет	специальное	покупка	Нет	16500
92		10	24	Муж	отечественная	20	Нет	специальное	покупка	Нет	34500
93	14		6	Муж	импортная	25	Да	специальное	покупка	Нет	6000
94	20		30	Муж	импортная	38	Нет	высшее	другое	Да	51000
95			12	Жен	отечественная	12	Нет	специальное	нет жилья в собственности	Да	16000
96		6	18	Жен	отечественная	34	Нет	специальное	покупка	Нет	19000
97			36	Жен	импортная	18	Нет	высшее	другое	Да	63000
98	16	14	6	Муж	отечественная	29	Да	специальное	покупка	Да	3000
99			12	Жен	отечественная	39	Нет	специальное	другое	Нет	30000
100			18	Жен	отечественная	61	Нет	высшее	другое	Нет	38500
101	8		6	Жен	отечественная	36	Да	высшее	другое	Да	13000
102		3	18	Жен	импортная	23	Нет	высшее	нет жилья в собственности	Нет	27500
103		14	24	Муж	отечественная	29	Нет	высшее	другое	Нет	50000
104		3	6	Жен	импортная	23	Да	среднее	другое	Нет	9000
105	2	2	36	Жен	отечественная	21	Нет	специальное	другое	Нет	43000
106		11	6	Муж	отечественная	22	Нет	среднее	нет жилья в собственности	Нет	2500
107			24	Жен	отечественная	47	Нет	специальное	другое	Нет	48500
108		1	30	Жен	отечественная	17	Нет	среднее	нет жилья в собственности	Да	45500
109			12	Жен	Нет автомобиля	15	Нет	среднее	другое	Нет	32000
110			6	Муж	отечественная	40	Нет	высшее	другое	Нет	2500
111			6	Муж	отечественная	62	Да	специальное	другое	Нет	6500
112			20	Да	Нет автомобиля	20	Да	специальное	другое	Да	13500
113			24	Муж	импортная	43	Нет	специальное	покупка	Нет	31500
114		8	6	Муж	отечественная	15	Нет	специальное	нет жилья в собственности	Нет	6000
115			18	Муж	отечественная	19	Нет	специальное	покупка	Нет	28000
116			45	Да	отечественная	45	Да	специальное	другое	Да	13500
117			12	Муж	Нет автомобиля	38	Да	специальное	покупка	Нет	19500
118	7		6	Жен	отечественная	35	Да	специальное	другое	Нет	18000
119			21	Нет	отечественная	21	Нет	высшее	другое	Нет	45000
120			31	Да	отечественная	31	Да	специальное	другое	Нет	20000
121	22		30	Муж	отечественная	44	Нет	специальное	покупка	Да	38000
122	17		12	Муж	отечественная	31	Нет	специальное	другое	Да	14500

Обработчик для поиска
дубликатов и противоречий и
специальный визуализатор

+ Σ Факты ▾					
Группа товара ▾		Сумма	Кол-во	Цена	
Медикамент ▾ +	> Альфа-адреномиметики	764,40	9,00	116,40	
	> Аминогликозиды	3 935 504,89	250 546,00	15,70	
	> Анаболики	28 222,34	1 669,00	16,84	
	> Ветрогонное средство	31 130,35	440,00	66,12	
	▾ Витамины и витаминоподо... средства	Аддитива Витамин С таб. шип. №10	3 469,26	49,00	71,23
		Аддитива Железо таб. шип. №10	1 140,39	32,00	35,35
		Аддитива Кальций таб. шип. №10	4 564,02	102,00	40,67
		Аддитива Магний таб. шип. №10	282,60	8,00	35,33
		Аддитива Мультивит. с минералами таб. шип. №10	256,11	7,00	36,59
		Аддитива Мультивит. тропик. таб. шип. №10	4 660,05	105,00	38,50
		Адитива Тропик таб. шип. №10	821,44	26,00	31,41
		Аекол р-р д/местн. и наружн.примен. фл. 100мл	214,00	5,00	42,80
		Декамевит таб. п/о x10	244,80	12,00	20,40
		Карнитен р-р для приема внутрь 1 г/10 мл, фл.№10	6 295,80	14,00	449,70
		Компливит Актив таб. п/о №30	33 804,06	555,00	64,11
		Компливит Мама д/берем. и корм. женщин таб. п...	203 598,80	3 099,00	64,33
		Компливит таб. п/о №60	318 803,24	5 159,00	61,80
		Тиогамма 600 мг таб. п/о № 30	13 071,00	15,00	871,40
		Тиогамма р-р д/инф. 1,2 % фл.50 мл №10	6 669,20	4,00	1 667,30
	Церневит лиоф. д/р-ра для в/в введ. 747 мг 1 доз;...	219 546,00	150,00	1 463,64	
	> Вспомогательные вещества	817 440,77	9 342,00	70,08	
	> Гипогликемические синтетические и другие средства	22 255,20	2 900,00	7,91	
	> Гипотензивные средства	1 971 097,28	50 938,00	135,66	
> Гормональные средства	1 025,70	41,00	243,81		
> Гормональные препараты	3 335,40	849,00	76,17		
> Корректоры метаболизма костной и хрящевой ткани	160 517,33	319,00	760,84		
> Корректоры нарушений мозгового кровообращения	17 537,27	417,00	40,91		
> Кортикостероиды	1 766 372,09	149 739,00	42,63		
> Пульмонология	285 340,12	59 970,00	73,77		
> Спазмолитическое, анальгезирующее, сосудорасширяющее средство	8 046,06	22,00	365,73		
> Средства, влияющие на обмен мочевой кислоты	355 218,72	14 589,00	41,19		



Настройка

Пакеты

dwh (только чтение)

Модуль1

Сценарий

Хранилище данных-2

Настройка

Импорт из хранилища данных

Фильтрация

Параметр	Тип
Продажи	
Дата	
31 Дата	
Неделя	
31 Неделя	
12 Год	
Месяц	
31 Месяц	
Квартал	
31 Квартал	
12 Год	
Покупатель	
ab Организация	
ab Категория	
Ответственный сотрудник	
ab Сотрудник	
Должность	
Получатель	
ab Организация	
ab Категория	
Ответственный сотрудник	
Поставщик	
ab Организация	
ab Категория	
Ответственный сотрудник	
ab Сотрудник	
Должность	
Товар	
ab Товар	
ab Группа	
9.0 Сумма	

Выводить	Поле	Тип	Условие	Значение	
☒	31 Дата Месяц				
☒	ab Покупатель				
☒	ab Покупатель Категория				
☒	ab Получатель		начинается на	Аптека	
☒	ab Получатель Категория				
☒	ab Товар				
☒	ab Товар Группа		в списке	Косметика,Детские товары	
☒	9.0 Сумма			(Сумма)	

Импорт из хранилища данных:
отображение структуры объектов
в виде дерева



Экспорт в хранилище данных

Связать все автоматически Удалить все связи

Фильтрация

Столбцы исходного набора

31 Дата Месяц
31 Дата Месяц Квартал
12 Дата Месяц Квартал Год
ab Покупатель
ab Покупатель Категория
ab Покупатель Ответственный сотрудник
ab Покупатель Ответственный сотрудник Должность
ab Покупатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Покупатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Покупатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Покупатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Покупатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Получатель
ab Получатель Категория
ab Получатель Ответственный сотрудник
ab Получатель Ответственный сотрудник Должность
ab Получатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Получатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Получатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Получатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Получатель Ответственный сотрудник Должность Вышестоящая Дол...
ab Поставщик
ab Поставщик Категория
ab Поставщик Ответственный сотрудник
ab Поставщик Ответственный сотрудник Должность
ab Поставщик Ответственный сотрудник Должность Вышестоящая Дол...
ab Поставщик Ответственный сотрудник Должность Вышестоящая Дол...
ab Поставщик Ответственный сотрудник Должность Вышестоящая Дол...
ab Поставщик Ответственный сотрудник Должность Вышестоящая Дол...
ab Поставщик Ответственный сотрудник Должность Вышестоящая Дол...
ab Товар
ab Товар Группа
9.0 Сумма Сумма

Фильтрация

Параметр	Тип	Способ удаления	Способ агрегации
Продажи			
31 Дата	↕		
ab Покупатель	↕		
ab Получатель	↕		
ab Поставщик	↕		
ab Товар	↕		
9.0 Сумма	↕		Сумма

Экспорт в хранилище данных

Оптимизация и улучшения

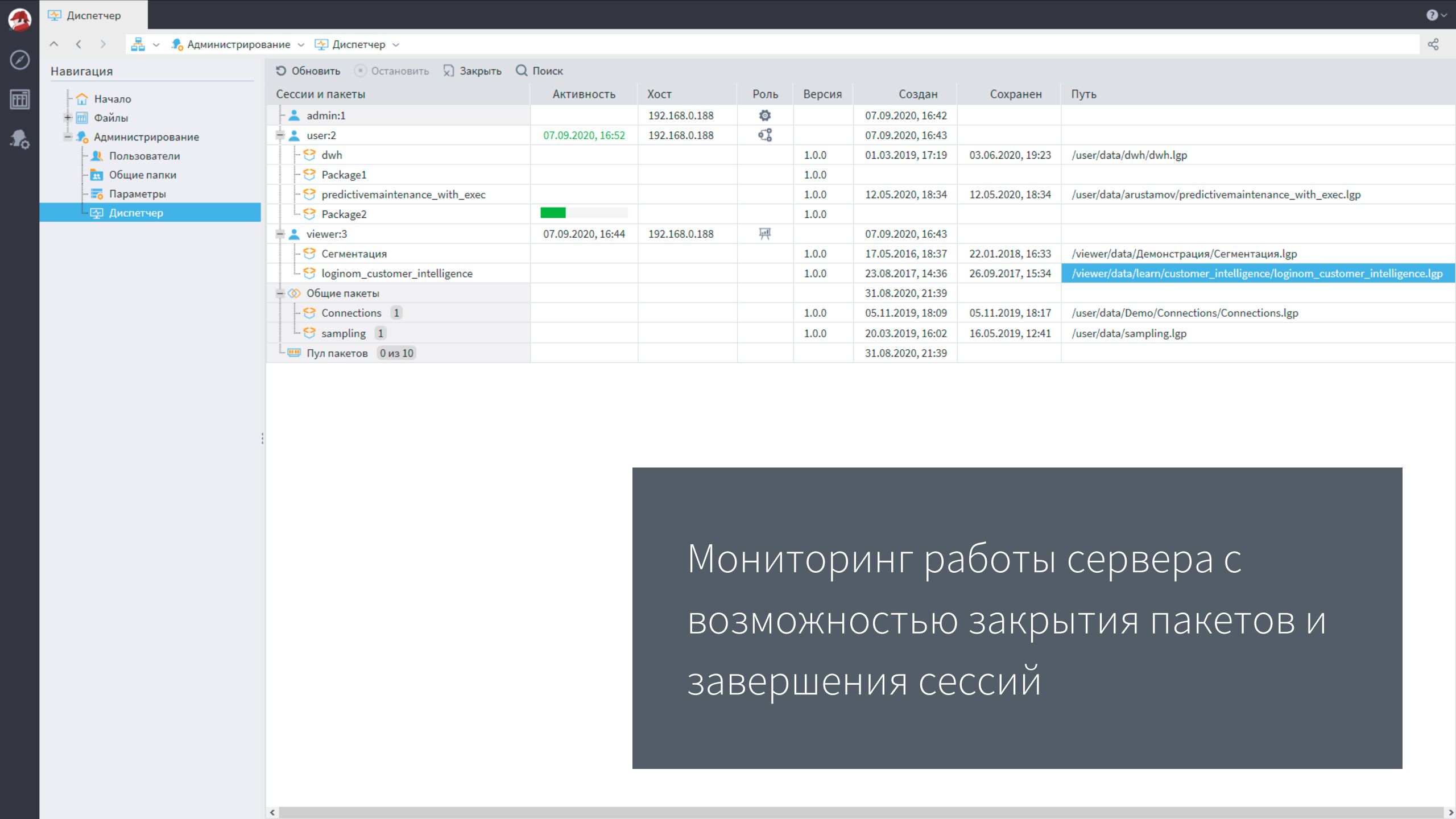
Категория	Название
Обработчик	Заполнение пропусков
Обработчик	Корреляционный анализ
Обработчик	Фильтрация
Визуализатор	Профили кластеров
Визуализатор	Таблица

Администрирование



Логирование

1. Включение и выключение отладочного режима без перезапуска сервера
2. Улучшенная диагностика и более информативные сообщения в логах



Мониторинг работы сервера с
возможностью закрытия пакетов и
завершения сессий

Языки программирования



Расширение API узла JavaScript – возможность проектировать произвольную структуру выходного набора из кода JavaScript. Расширение API объекта OutputTable.

Метод	Назначение
AssignColumns	Создание столбцов выходного набора из коллекции имен/описаний столбцов
AddColumn	Добавление столбца в конец списка столбцов выходного набора
InsertColumn	Вставка столбца по заданному индексу в выходной набор
DeleteColumn	Удаление столбца по имени или индексу
ClearColumns	Очистка списка столбцов



СPython 3.4 и выше

1. API аналогичный JS
2. Подключаемые пакеты
3. Поддержка DataFrame

Основной кейс применения
Python – встраивание расчетов
в сценарий обработки



Ограничения компонента Python

1. Запрещено работать с GUI
2. Обработка выполняется в один поток
3. Не поддерживаются модули пакета multiprocessing
4. В случае критических ошибок в Python приложение может аварийно завершиться
5. При конвертации в DataFrame нужно передавать все данные

Безопасность Python

1. Надо дать права на использование компонента Python
2. Код на Python выполняется с теми же правами, что есть у пользователя под которым запущен сервер Loginom

Подготовка к использованию

1. Установить Python версии 3.4 или выше
2. Установить необходимые пакеты
3. Установить IDE для Python

Пример с использованием структур DataFrame



О программе

Сценарий

Параметры

Администрирование

Параметры

Навигация

Начало

Пакеты

Classification_2

- Переменные
- Ссылки
- Отчеты
- Модель в файле
 - Сценарий
 - Переменные сценария
 - Разбиение на множес
 - Loginom Data файл
 - Сэмплинг
 - Python
 - Подключения
 - Компоненты
 - Модель в строке

Администрирование

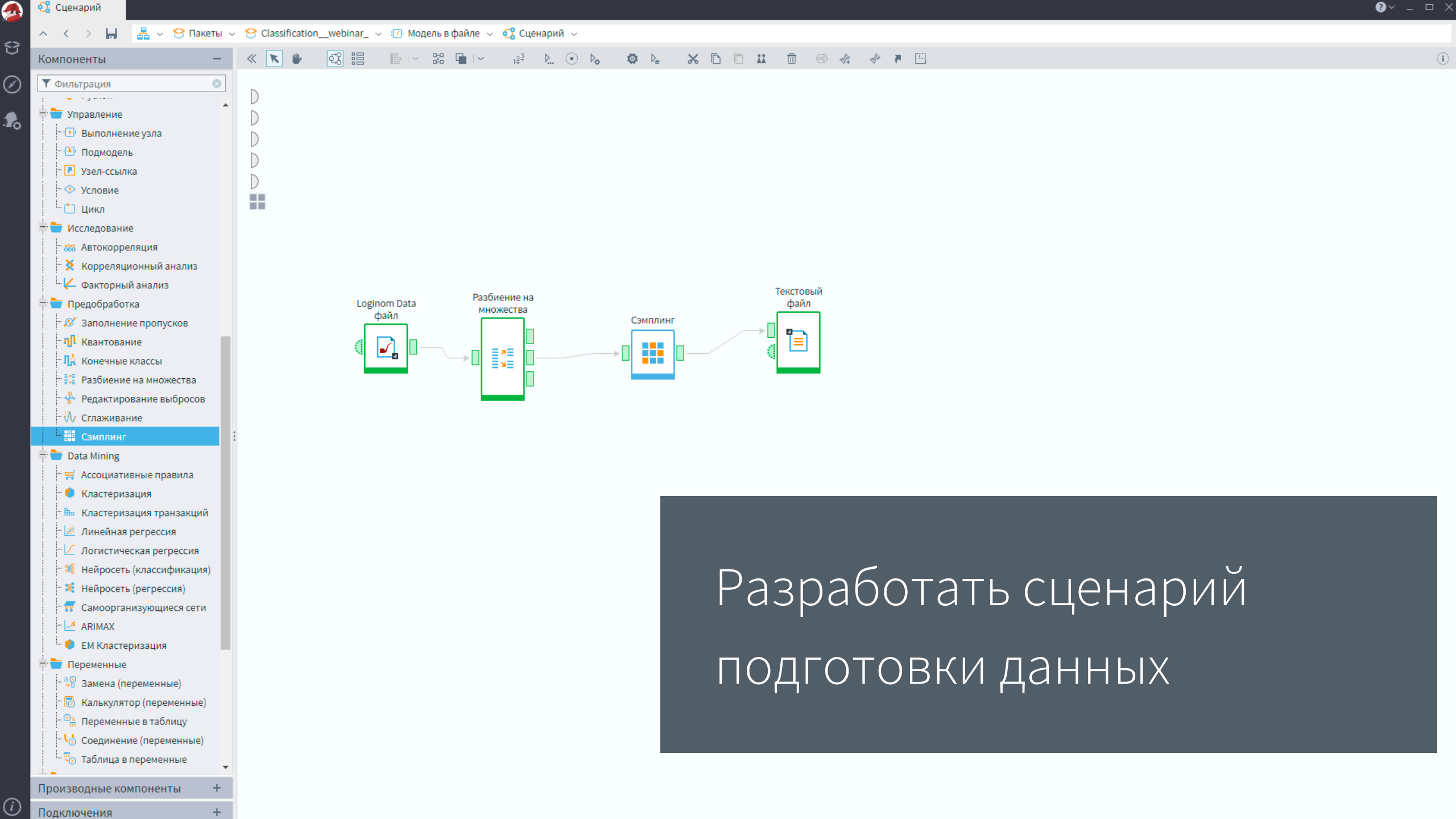
Параметры

Сохранить

Обновить

Параметр	Значение
Параметры логирования	
Уровень логирования	Информация
Имя файла	app.log
Перезапись файла	false
Максимальный размер файла	10485760
Количество резервных файлов	10
Параметры приложения	
Размер пула потоков	По умолчанию
Параметры компонента: Python	
Выполнение запрещено	false
Путь библиотеки	D:\Users\alex\AppData\Local\Programs\Python\Python38\python38.dll
Параметры компонента: Выполнение программы	
Выполнение запрещено	false

Разрешить выполнение
Python и указать путь к
библиотеке python3x.dll



In [1]: `import pandas as pd`

In [2]: `data = pd.read_csv('covtype_test.csv')`
`data.head()`

Out[2]:

	Elevation	Aspect	Slope	Horizontal_Distance_To_Hydrology	Vertical_Distance_To_Hydrology	Horizontal_Distance_To_Roadways	Hillshade_9am	Hillshade_Noon
0	2562	354	12	67	9	1057	200	259
1	2886	22	8	285	31	3849	215	279
2	3201	331	5	0	0	4002	208	259
3	2952	62	8	192	30	4487	227	259
4	2791	41	17	458	-18	1740	220	259

5 rows x 9 columns

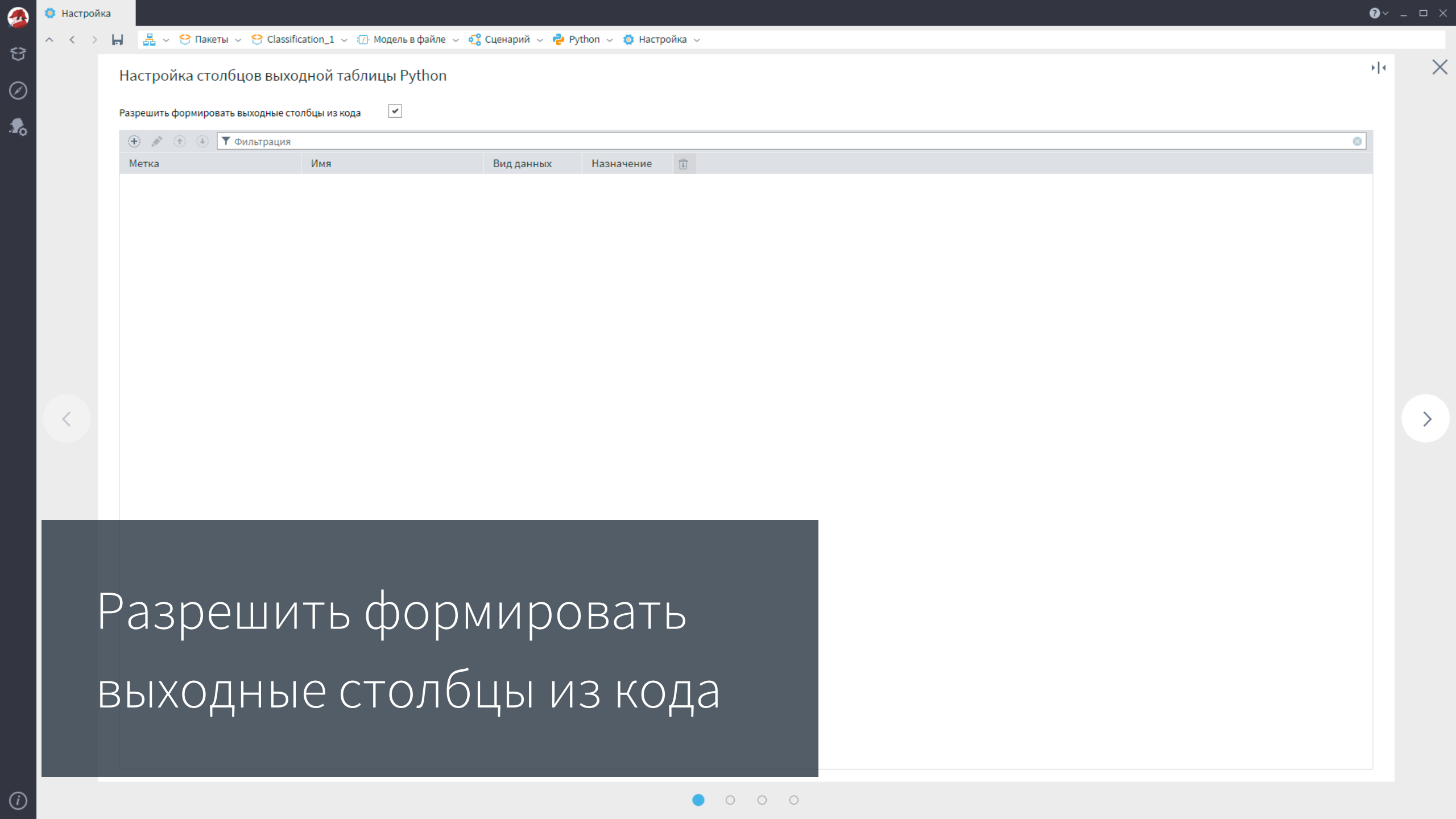
In [3]: `from sklearn.cluster import Birch`
`from sklearn.metrics import calinski_harabasz_score`

In [4]: `clustering = Birch(n_clusters=7).fit(data)`
`data["cluster"] = clustering.labels_`

In [5]: `print("Calinski and Harabasz score =", calinski_harabasz_score(data, clustering.labels_))`
Calinski and Harabasz score = 6788.11193881867

In [3]: `data.to_csv('covtype_result.csv')`

Написать и отладить
код программы



Настройка столбцов выходной таблицы Python

Разрешить формировать выходные столбцы из кода ☒

Фильтрация				
Метка	Имя	Вид данных	Назначение	

Разрешить формировать
выходные столбцы из кода



Python

```
1 import builtin_data
2 from builtin_data import InputTable, InputTables, InputVariables, OutputTable, DataType, DataKind, UsageType
3 '''
4 import numpy as np, pandas as pd
5 from builtin_pandas_utils import to_data_frame, prepare_compatible_table, fill_table
6
7 # Входной порт необязательный
8 if InputTable:
9     # Создать pd.DataFrame по входному набору №1
10     input_frame = to_data_frame(InputTable)
11
12 # Здесь может быть код работы с данными
13
14 # Полученный выходной pd.DataFrame
15 output_frame = pd.DataFrame({"COL1": [1, 2], "COL2": ["a", "b"], "COL3": [np.datetime64("2020"), np.datetime64("nat")]})
16
17 # Если включена опция "Разрешить формировать выходные столбцы из кода", структуру выходного набора можно подготовить по pd.DataFrame
18 if isinstance(OutputTable, builtin_data.ConfigurableOutputTableClass):
19     prepare_compatible_table(OutputTable, output_frame, with_index=False)
20 fill_table(OutputTable, output_frame, with_index=False)
21 '''
```

Исправить шаблон
написания кода с DataFrame

Тайм-аут ожидания (мс)

∞

Предпросмотр...



Python

```
1 from builtin_data import InputTable, OutputTable # Удаляем неиспользуемые объекты
2 from builtin_pandas_utils import to_data_frame, prepare_compatible_table, fill_table # Оставляем имеющийся в шаблоне блок работы с DataFrame
3
4 data = to_data_frame(InputTable) # Преобразовываем входную таблицу в pandas.DataFrame
5
6 # *****
7 # Фрагмент кода по обработке данных
8 # *****
9 from sklearn.cluster import Birch
10 from sklearn.metrics import calinski_harabasz_score
11
12 clustering = Birch(n_clusters=7).fit(data) # Создаем и обучаем модель кластеризации
13 data["Cluster"] = clustering.labels_ # Полученные номера кластеров добавляем в DataFrame
14
15 print("Calinski and Harabasz score =", calinski_harabasz_score(data, clustering.labels_)) # Метрика качества
16 # *****
17
18 prepare_compatible_table(OutputTable, data, with_index=False) # Формируем структуру выходной таблицы по DataFrame
19 fill_table(OutputTable, data, with_index=False) # Заполняем выходную таблицу из DataFrame
20
21
```

Оставить нужный код из
шаблона и добавить код для
встраивания в Loginom

Тайм-аут ожидания (мс)

∞

[Предпросмотр...](#)



Python

```
1 from builtin_data import InputTable, OutputTable # Удаляем неиспользуемые объекты
2 from builtin_pandas_utils import to_data_frame, prepare_compatible_table, fill_table # Ос
3
4 data = to_data_frame(InputTable) # Преобразовываем входную таблицу в pandas.DataFrame
5
6 # *****
7 # Фрагмент кода по обработке данных
8 # *****
9 from sklearn.cluster import Birch
10 from sklearn.metrics import calinski_harabasz_score
11
12 clustering = Birch(n_clusters=7).fit(data) # Создаем и обучаем модель кластеризации
13 data["Cluster"] = clustering.labels_ # Полученные номера кластеров добавляем в DataFrame
14
15 print("Calinski and Harabasz score =", calinski_harabasz_score(data, clustering.labels_))
16 # *****
17
18 prepare_compatible_table(OutputTable, data, with_index=False) # Формируем структуру выход
19 fill_table(OutputTable, data, with_index=False) # Заполняем выходную таблицу из DataFrame
20
21
```

Предпросмотр

#	9.0 Elevation	9.0 Aspect	9.0 Slope	9.0 Horizontal_Distance_To_Hydrology	9.0 Vertical_Distance_To_Hydrology	9.0 Horizontal_Dis
1	2 713,00	117,00	30,00	60,00	17,00	
2	2 795,00	18,00	3,00	218,00	17,00	
3	2 499,00	42,00	13,00	497,00	-7,00	
4	3 231,00	45,00	4,00	648,00	59,00	
5	3 162,00	43,00	15,00	484,00	144,00	
6	2 916,00	138,00	21,00	127,00	49,00	
7	3 237,00	27,00	14,00	997,00	36,00	
8	2 978,00	27,00	4,00	192,00	13,00	
9	3 149,00	188,00	7,00	134,00	17,00	
10	3 044,00	23,00	5,00	210,00	-25,00	
11	2 611,00	99,00	14,00	30,00	2,00	
12	2 939,00	288,00	5,00	379,00	39,00	
13	2 955,00	235,00	21,00	451,00	55,00	
14	2 826,00	11,00	7,00	366,00	73,00	
15	2 699,00	165,00	3,00	390,00	2,00	
16	2 645,00	45,00	4,00	134,00	-3,00	
100	2 764,00	61,00	15,00	42,00	7,00	

Вывод

Bce

Calinski and Harabasz score = 6187.754706725524

> stdout

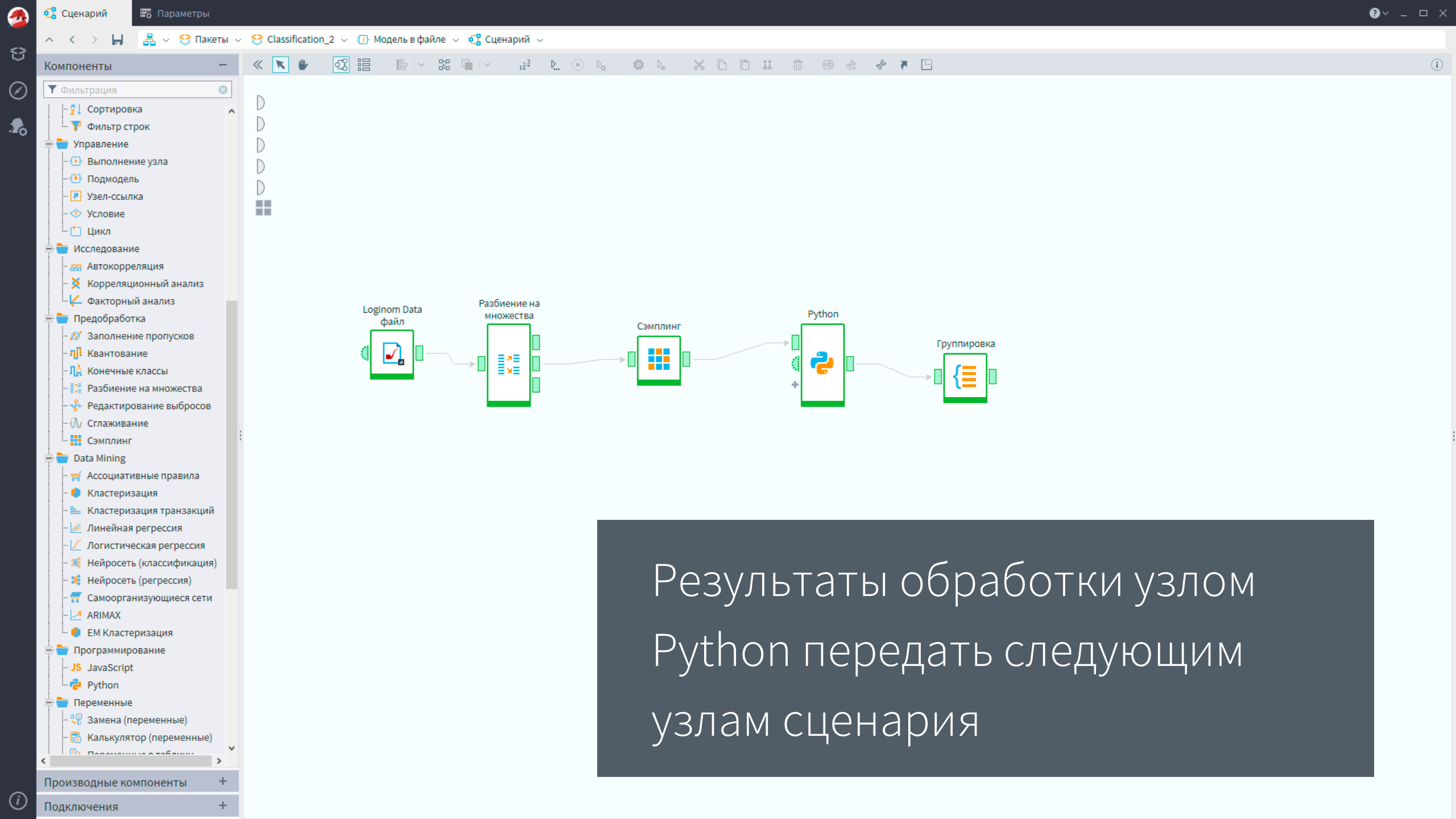
stderr

Предварительный просмотр результатов

Тайм-аут ожидания (мс)

∞

Предпросмотр...



Пример без использования структур DataFrame





Python

```
1 import apache_log_parser, geoip2.database
2 from geoip2.errors import AddressNotFoundError
3 from builtin_data import InputVariables, OutputTable, DataType
4 from datetime import datetime
5 from typing import TextIO
6
7 def prepare_output_table(log_line_data: dict) -> None:
8     for key, value in log_line_data.items():
9         datatype = DataType.String
10         if isinstance(value, bool):
11             datatype = DataType.Boolean
12         elif isinstance(value, datetime):
13             datatype = DataType.DateTime
14         OutputTable.AddColumn(Name=key, DataType=datatype)
15     OutputTable.InsertColumn(1, Name="country", DataType=DataType.String)
16     OutputTable.InsertColumn(2, Name="city", DataType=DataType.String)
17     OutputTable.InsertColumn(3, Name="latitude", DataType=DataType.Float)
18     OutputTable.InsertColumn(4, Name="longitude", DataType=DataType.Float)
19
20 def fill_output_table(log_file: TextIO, geoip_reader: geoip2.database.Reader) -> None:
21     for line in file:
22         OutputTable.Append()
23         log_line_data = line_parser(line)
24         for key, value in log_line_data.items():
25             OutputTable.Set(key, value)
26         try:
27             location_data = geoip_reader.city(log_line_data["remote_host"])
28         except AddressNotFoundError:
29             continue
30         OutputTable.Set("country", location_data.country.name)
31         OutputTable.Set("city", location_data.city.name)
32         OutputTable.Set("latitude", location_data.location.latitude)
33         OutputTable.Set("longitude", location_data.location.longitude)
34
35 line_parser = apache_log_parser.make_parser("%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\"")
36 with open(filestoragepath(InputVariables.Items["log_file"].Value)) as file, \
37     geoip2.database.Reader(filestoragepath(InputVariables.Items["geoip_db_path"].Value)) as reader:
```

Код Python может через
API работать напрямую со
структурами Loginom

Тайм-аут ожидания (мс)

Предпросмотр...

Сценарий

Параметры

Таблица

Пакеты

ApacheLog_mini

Модуль1

Сценарий

Разбор access.log

Визуализаторы

Таблица

Формат

Сортировка

Фильтр

Найти

XLS

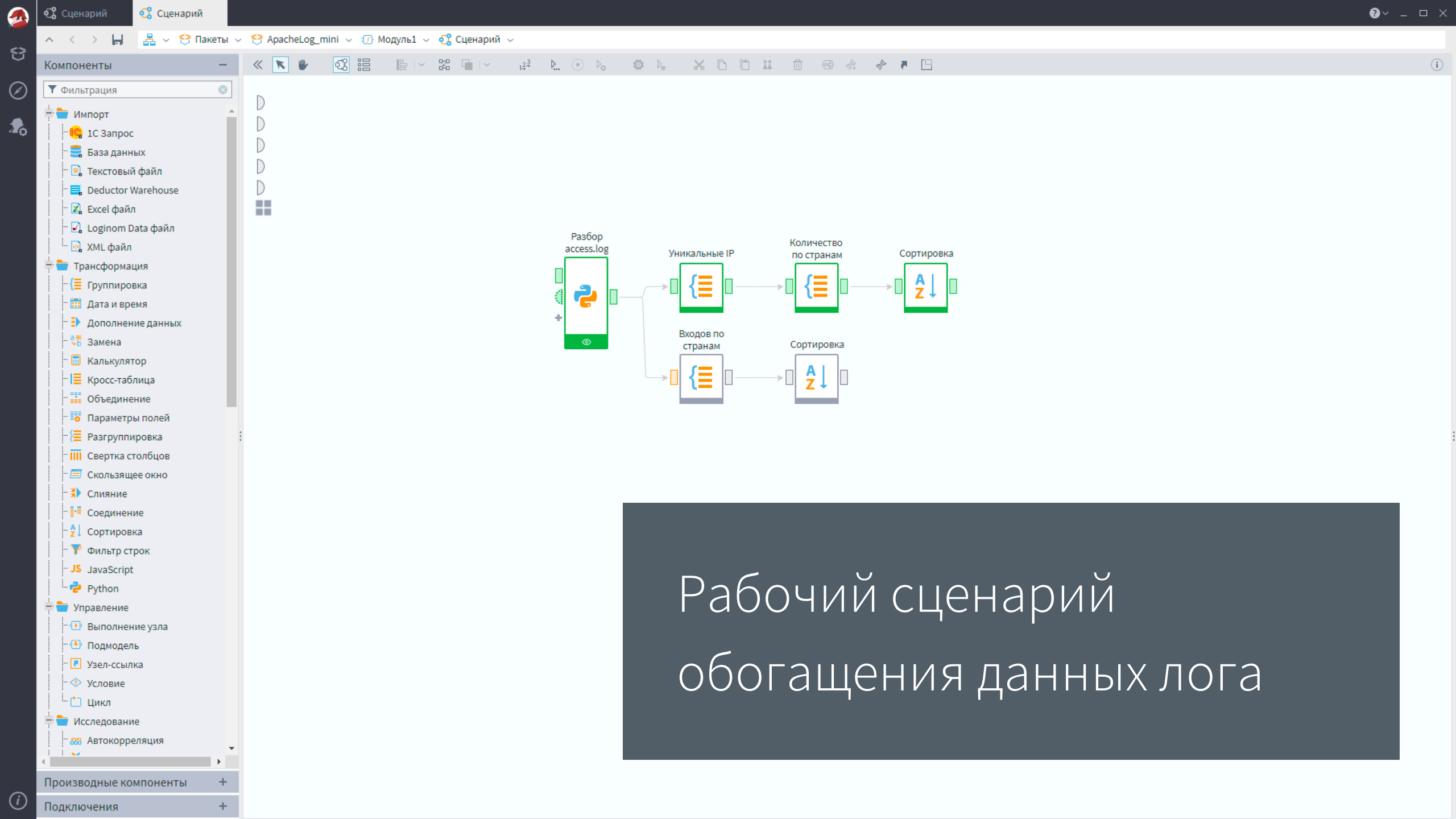
#	ab remote_host	ab country	ab city	latitude	longitude	ab remote_logname	ab remote_user	ab time_received	time_received_date
1	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:03 +0000]	17.05.20
2	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:43 +0000]	17.05.20
3	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:47 +0000]	17.05.20
4	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:12 +0000]	17.05.20
5	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:07 +0000]	17.05.20
6	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:34 +0000]	17.05.20
7	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:57 +0000]	17.05.20
8	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:50 +0000]	17.05.20
9	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:24 +0000]	17.05.20
10	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:50 +0000]	17.05.20
11	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:46 +0000]	17.05.20
12	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:11 +0000]	17.05.20
13	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:19 +0000]	17.05.20
14	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:33 +0000]	17.05.20
15	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:00 +0000]	17.05.20
16	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:25 +0000]	17.05.20
17	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:59 +0000]	17.05.20
18	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:30 +0000]	17.05.20
19	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:53 +0000]	17.05.20
20	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:24 +0000]	17.05.20
21	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:54 +0000]	17.05.20
22	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:33 +0000]	17.05.20
23	83.149.9.216	Russia	Moscow	55.7522	37.6156	-	-	[17/May/2015:10:05:56 +0000]	17.05.20
24	24.236.252.67	United States	Holland	42.8123	-86.1421	-	-	[17/May/2015:10:05:40 +0000]	17.05.20
25	93.114.45.13	Romania	Bucharest	44.4333	26.1	-	-	[17/May/2015:10:05:14 +0000]	17.05.20
26	93.114.45.13	Romania	Bucharest	44.4333	26.1	-	-	[17/May/2015:10:05:04 +0000]	17.05.20
27	93.114.45.13	Romania	Bucharest	44.4333	26.1	-	-	[17/May/2015:10:05:45 +0000]	17.05.20
28	93.114.45.13	Romania	Bucharest	44.4333	26.1	-	-	[17/May/2015:10:05:14 +0000]	17.05.20
29	93.114.45.13	Romania	Bucharest	44.4333	26.1	-	-	[17/May/2015:10:05:17 +0000]	17.05.20
30	93.114.45.13	Romania	Bucharest	44.4333	26.1	-	-	[17/May/2015:10:05:21 +0000]	17.05.20
31	66.249.73.185	United States	Ashtabula	41.7141	-81.8228	-	-	[17/May/2015:10:05:40 +0000]	17.05.20
32	50.16.19.13	United States	Ashtabula	39.0481	-77.1128	-	-	[17/May/2015:10:05:10 +0000]	17.05.20
33	66.249.73.185	United States	Ashtabula	37.751	-97.822	-	-	[17/May/2015:10:05:37 +0000]	17.05.20
34	110.136.166.128	Indonesia	Batang	-6.9078	109.7303	-	-	[17/May/2015:10:05:35 +0000]	17.05.20
35	46.105.14.53	France	Batang	48.8582	2.3387	-	-	[17/May/2015:10:05:03 +0000]	17.05.20
36	110.136.166.128	Indonesia	Batang	-6.9078	109.7303	-	-	[17/May/2015:10:05:06 +0000]	17.05.20
37	110.136.166.128	Indonesia	Batang	-6.9078	109.7303	-	-	[17/May/2015:10:05:03 +0000]	17.05.20
10 000	110.136.166.128	Indonesia	Batang	-6.9078	109.7303	-	-	[17/May/2015:10:05:41 +0000]	17.05.20

Детализация

Метка	Значение
#	1
ab remote_host	83.149.9.216
ab country	Russia
ab city	Moscow
latitude	55.7522
longitude	37.6156
ab remote_logname	-
ab remote_user	-
ab time_received	[17/May/2015:10:05:03 +0000]
time_received_datetimeobj	17.05.2015 10:05:03
time_received_isoformat	2015-05-17T10:05:03
time_received_tz_datetimeobj	17.05.2015 10:05:03
time_received_tz_isoformat	2015-05-17T10:05:03+00:00
time_received_utc_datetimeobj	17.05.2015 10:05:03
time_received_utc_isoformat	2015-05-17T10:05:03+00:00
ab request_first_line	GET /presentations/logstash-monitorama-2013/im...
ab request_method	GET
ab request_url	/presentations/logstash-monitorama-2013/image...
ab request_http_ver	1.1
ab request_url_scheme	
ab request_url_netloc	
ab request_url_path	/presentations/logstash-monitorama-2013/image...
ab request_url_query	
ab request_url_fragment	
ab request_url_username	
ab request_url_password	
ab request_url_hostname	
ab request_url_port	
ab request_url_query_dict	{}
ab request_url_query_list	[]
ab request_url_query_simple_dict	{}
ab status	200
ab response_bytes_clf	203023
ab request_header_referer	http://semicomplete.com/presentations/logstash...
ab request_header_user_agent	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_1) App...
ab request_header_user_agent_browser_family	Chrome
ab request_header_user_agent_browser_version...	32.0.1700
ab request_header_user_agent_os_family	Mac OS X

Результаты обработки

Страница 1 из 1



Рабочий сценарий обогащения данных лога

Запрос бета-версии
help@loginom.ru

