



LOGINOM
ХАКАТОН 2019

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Библиотека RFM-анализа

Команда TakeFive:
Голов В.А.
Дашпылова М.В.
Шабанова Е.Ю.

Руководитель:
Богданов А.А.



ПОСТАНОВКА ЗАДАЧИ

Дана клиентская база (ID клиента, Дата транзакции, Сумма покупки)

Требуется выполнить сегментирование и анализ базы методами RFM

ВЫБРАННЫЕ МЕТОДЫ

RFM

Дает более полное представление о поведении клиента

RF

Возможность произвести анализ при минимальном объеме исходных данных

LRF
M

Возможность построения матрицы взаимоотношений с клиентами

СРОК ЖИЗНИ КЛИЕНТА (L)



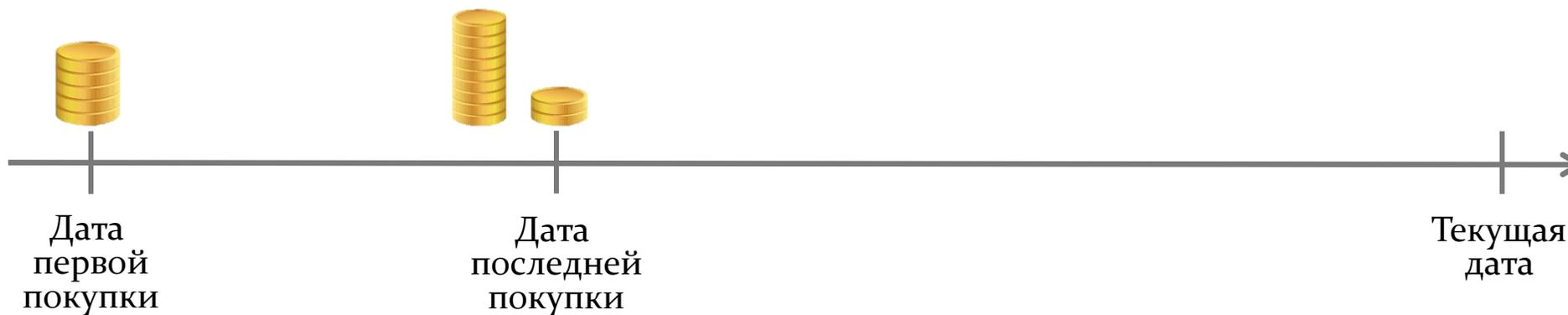
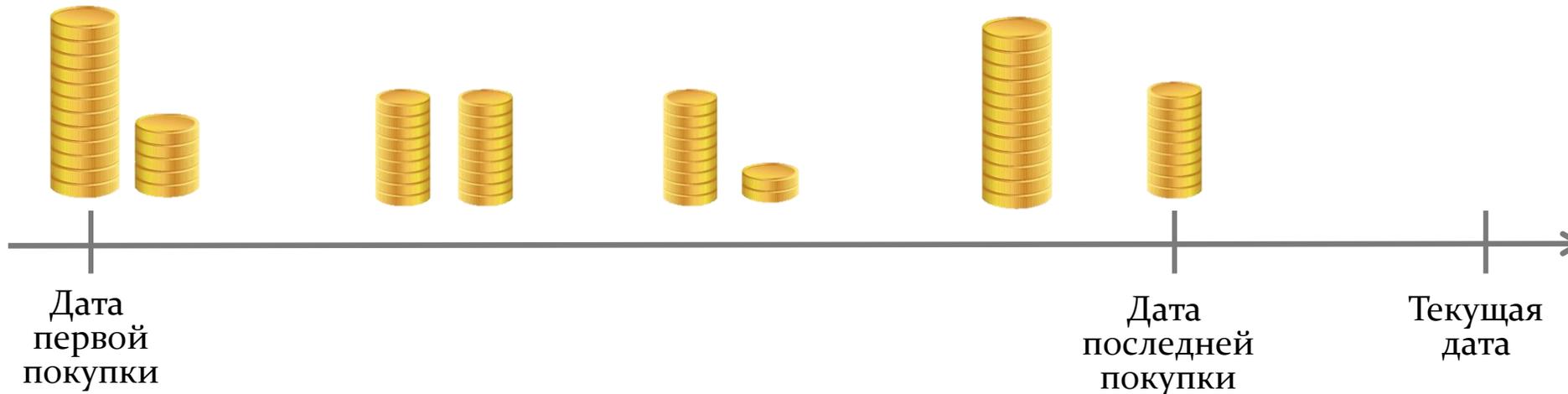
ДАВНОСТЬ ПОСЛЕДНЕЙ ПОКУПКИ (R)



ЧАСТОТА ПОСЕЩЕНИЙ (F)

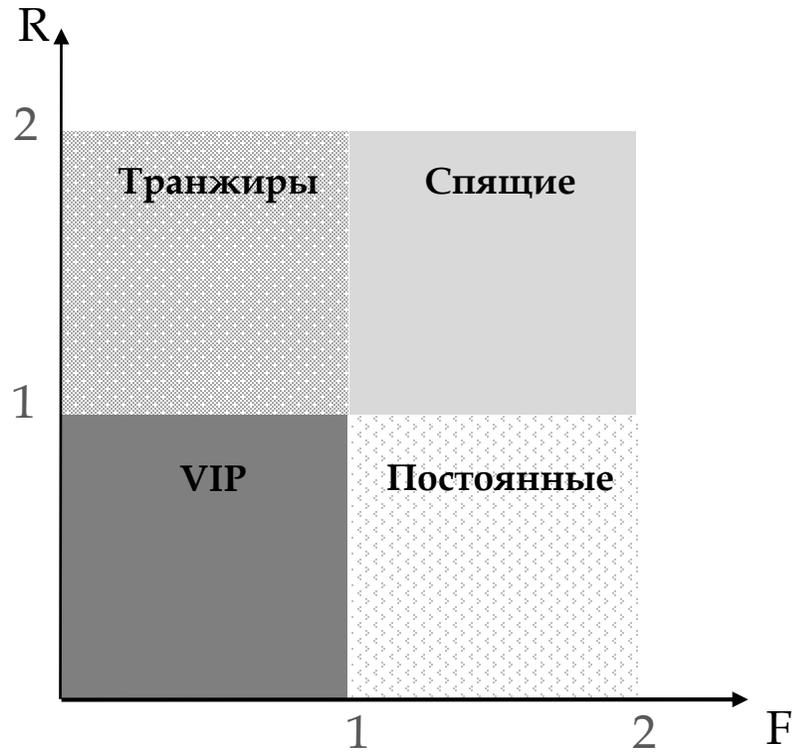


СУММА ПОКУПОК (М)

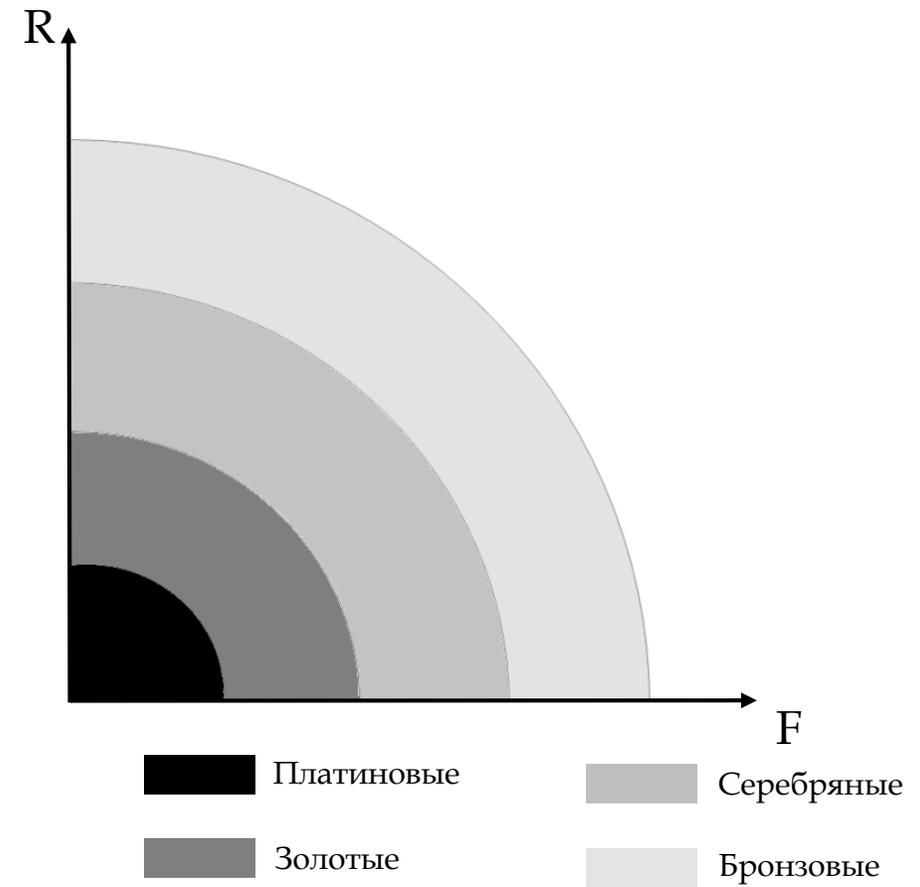


ТЕРМИНОЛОГИЯ

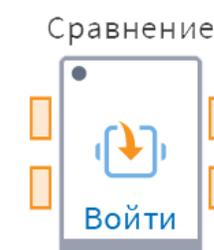
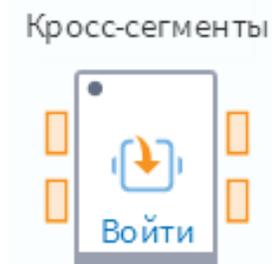
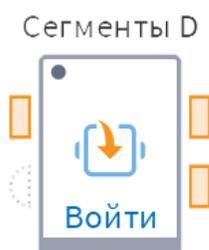
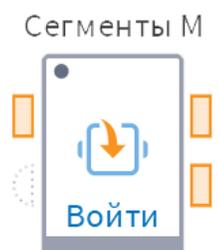
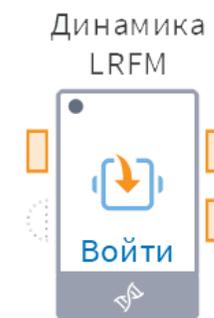
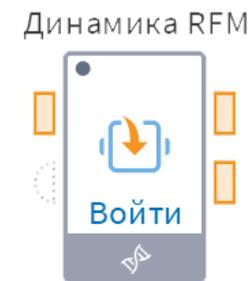
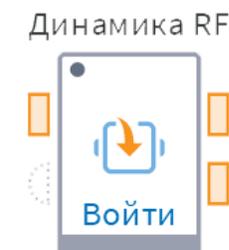
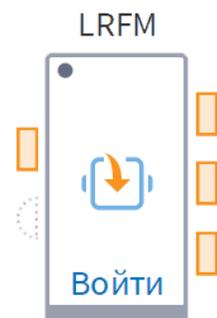
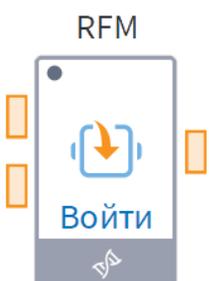
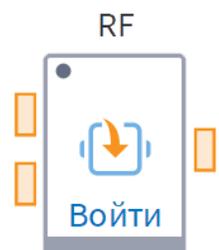
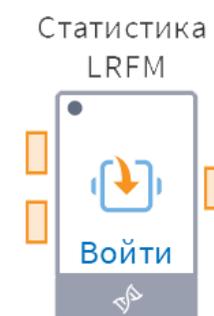
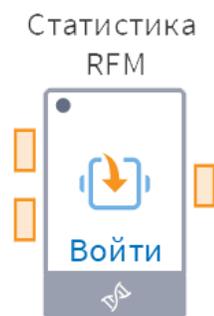
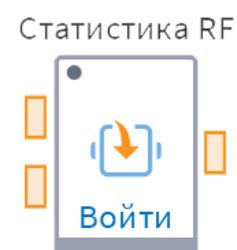
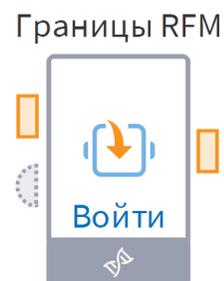
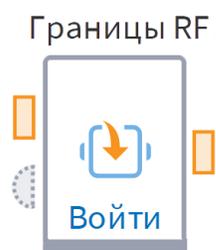
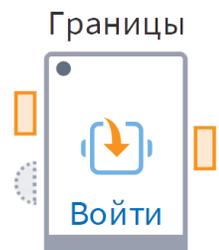
Дискретная сегментация



Метрическая сегментация



БИБЛИОТЕКА КОМПОНЕНТОВ



ВЫЧИСЛЕНИЕ ГРАНИЦ



Вход

Клиентская база:

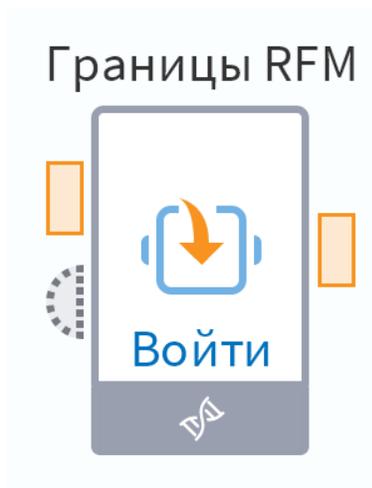
Идентификатор клиента, str

Дата покупки, datetime

Сумма покупки, float

Параметры

Количество интервалов для R, F, M



Выход

Границы:

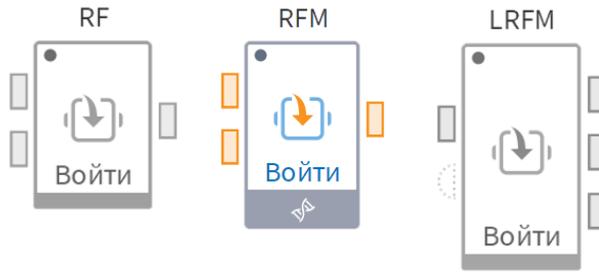
Признак, str

Интервал, int

Нижняя граница, float

Верхняя граница, float

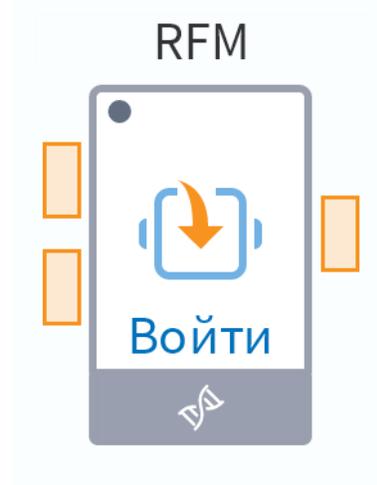
РАСЧЕТ КОДА



Вход

Клиентская база

Границы



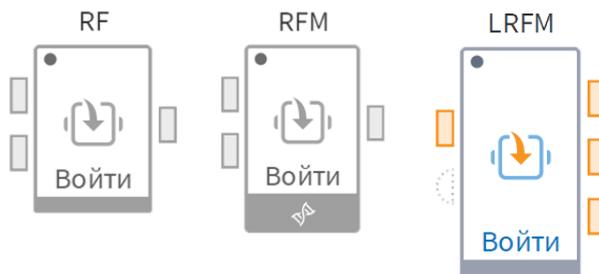
Выход

Клиент-код:

ID клиента, str

RFM-код, str

РАСЧЕТ КОДА



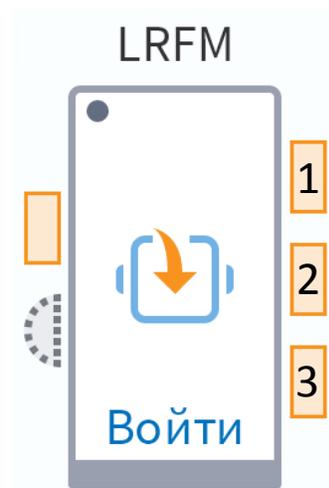
Вход

Клиентская база

Параметры

Файл с описанием сегментов
(Сегмент, Рекомендации)

Файл с распределением сегментов
(Код, Сегмент)



Выход

1 | Клиент-код LRFM

Сегменты:

2 | ID клиента, str
Сегмент, str
Рекомендации, str
Тип отношений, str

Сводка:

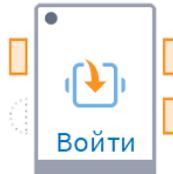
3 | Сегмент, str
Объем сегмента, int
Доля сегмента, float

СЕГМЕНТИРОВАНИЕ

Сегменты M



Сегменты D



Вход

Клиент-код

Параметры

Файл с описанием сегментов

Файл с распределением сегментов

Сегменты D



Выход

1 | Сегменты

2 | Сводка

ПОСТРОЕНИЕ КРОСС-СЕГМЕНТОВ

Кросс-сегменты

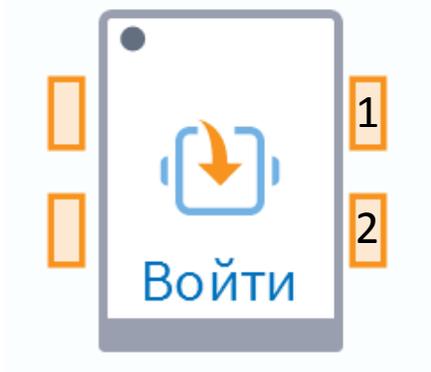


Вход

Сегменты (M)

Сегменты (D)

Кросс-сегменты

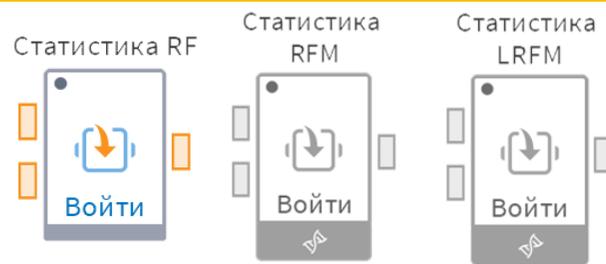


Выход

1 | Сегменты «на пересечении»

2 | Сводка

РАСЧЕТ СТАТИСТИКИ ПО СЕГМЕНТАМ

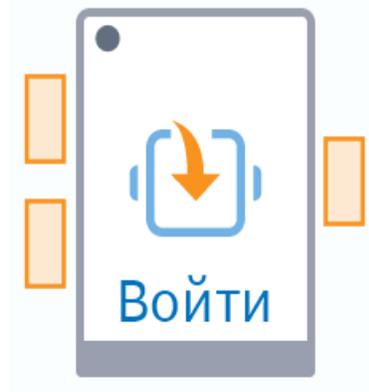


Вход

Клиентская база

Сегменты

Статистика RF



Выход

Сегмент, str

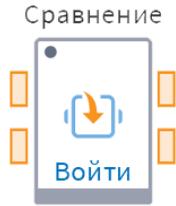
R, F – минимальное, float

R, F – максимальное, float

R, F – среднее, float

R, F – станд. откл., float

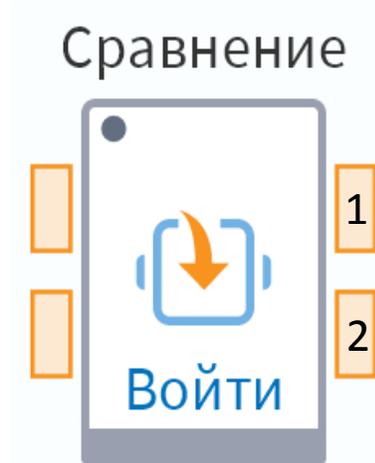
ЭВОЛЮЦИЯ ЦЕННОСТИ КЛИЕНТОВ



Вход

Сегменты для «новой» клиентской базы

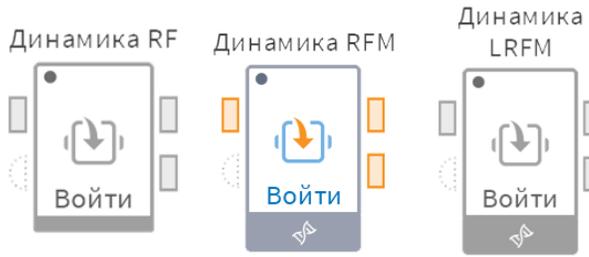
Сегменты для «старой» клиентской базы



Выход

- 1 ID клиента, str
Сравнение (>, <, =, +, -), str
Сегмент «вчера», str
Сегмент «сегодня», str
- 2 Сводка

АНАЛИЗ ДИНАМИКИ СЕГМЕНТОВ



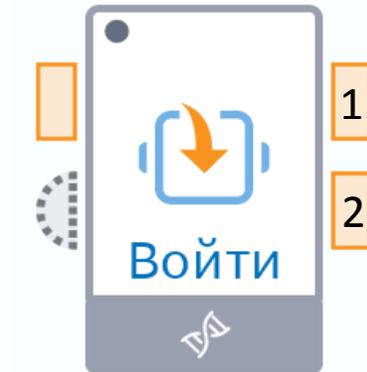
Вход

Клиентская база

Параметры

Файл с границами за месяц

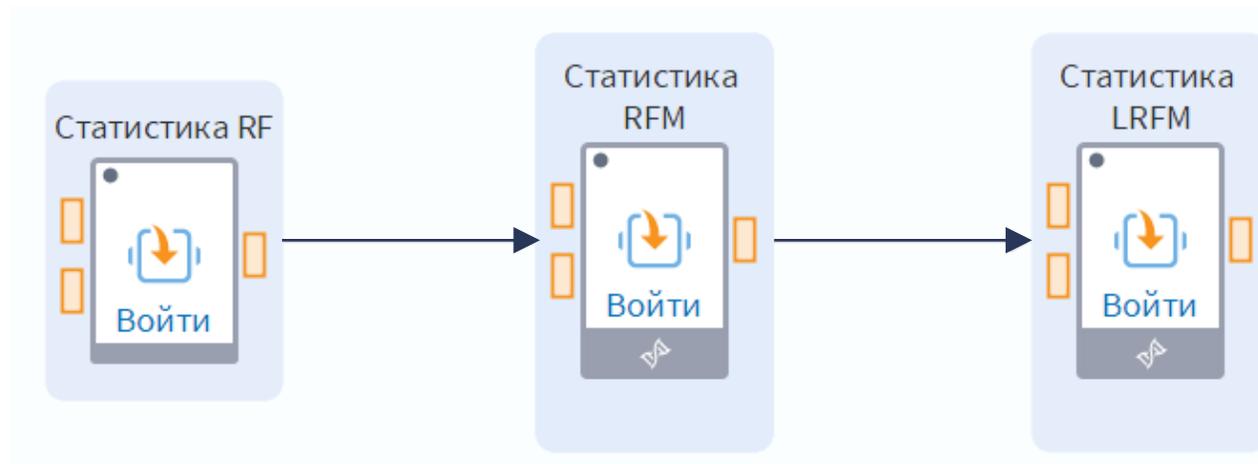
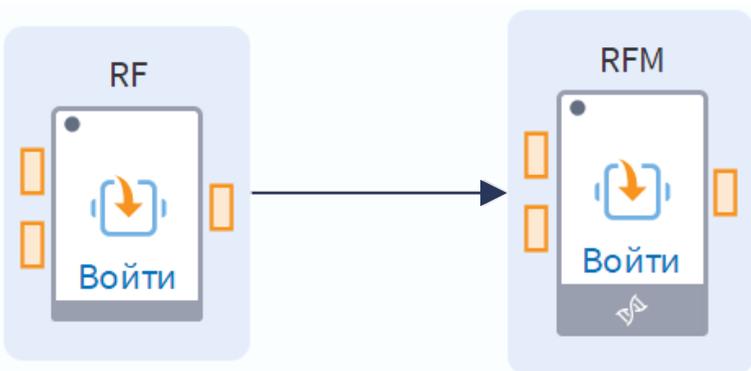
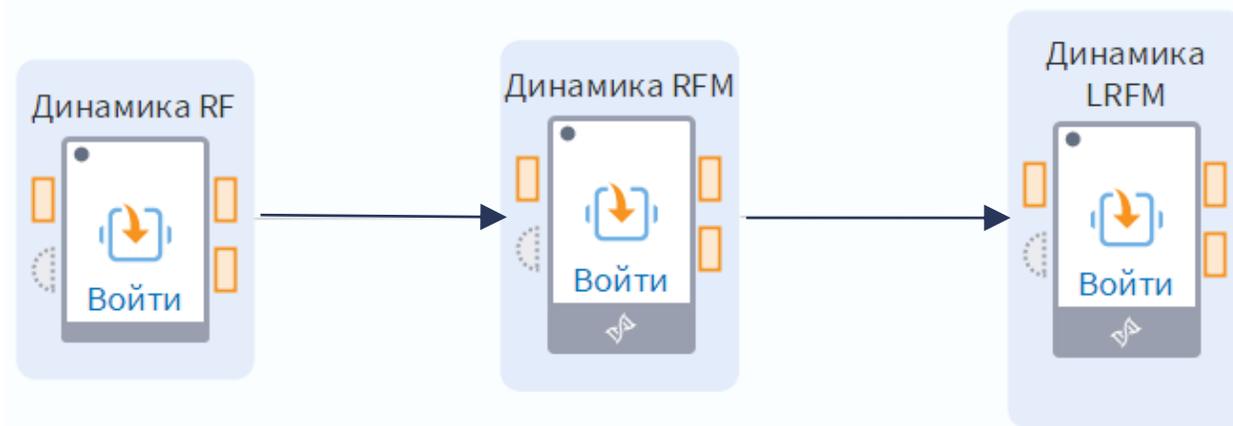
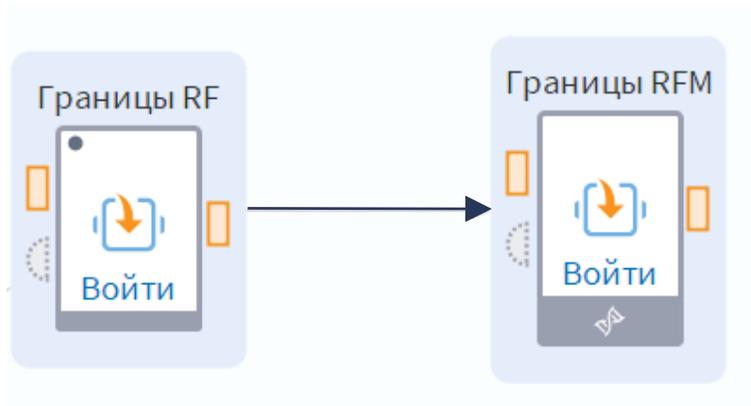
Динамика RFM



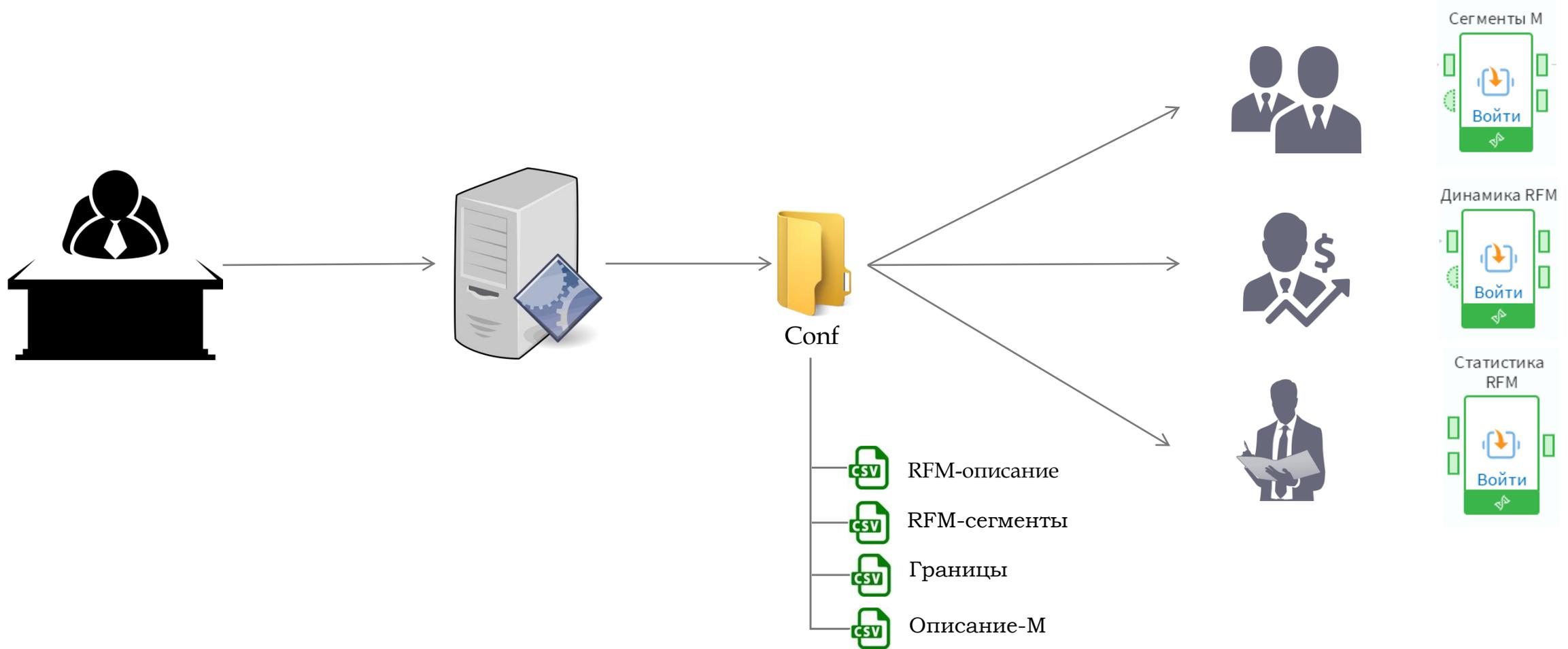
Выход

- 1 Месяц, date/time
Объем i -го сегмента M , $i=1,2,\dots,n$
Статистика i -го сегмента M
- 2 Месяц, date/time
Объем j -го сегмента D , $j=1,2,\dots,m$
Статистика j -го сегмента M

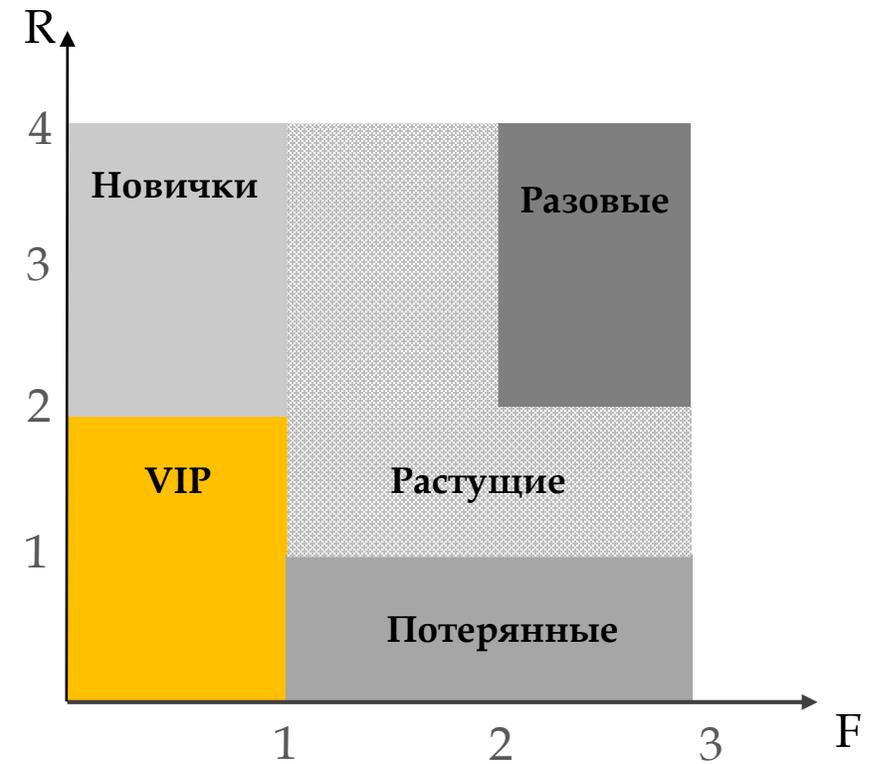
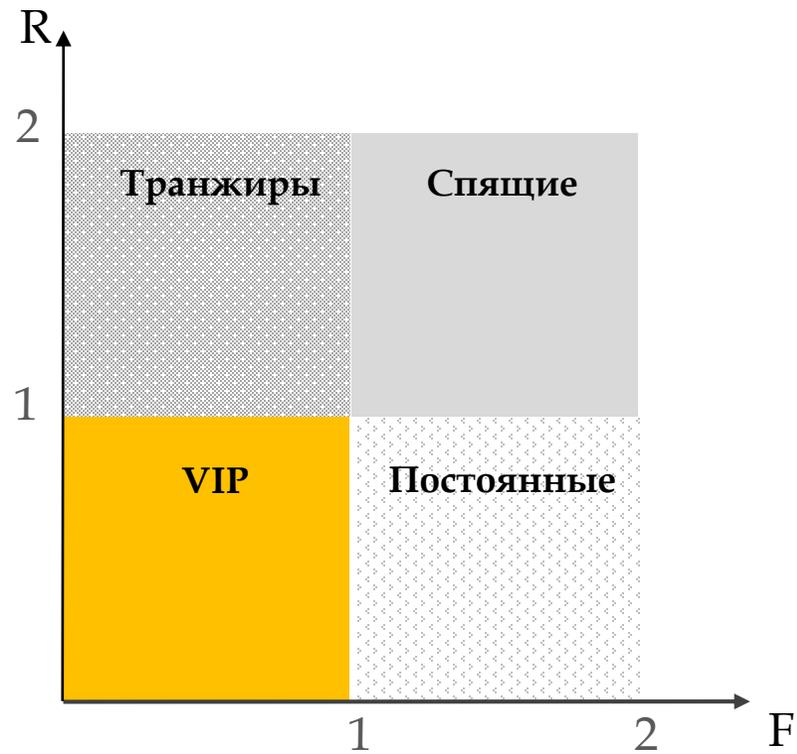
АРХИТЕКТУРА: НАСЛЕДОВАНИЕ



АРХИТЕКТУРА: КОНФИГУРИРОВАНИЕ



АРХИТЕКТУРА: КОНФИГУРИРОВАНИЕ



ДЕМОНСТРАЦИОННЫЙ ПРИМЕР

Исходные данные:

клиентская база сети магазинов формата «Luxury»

Продукция сети:

товары роскоши и услуги класса люкс-премиум

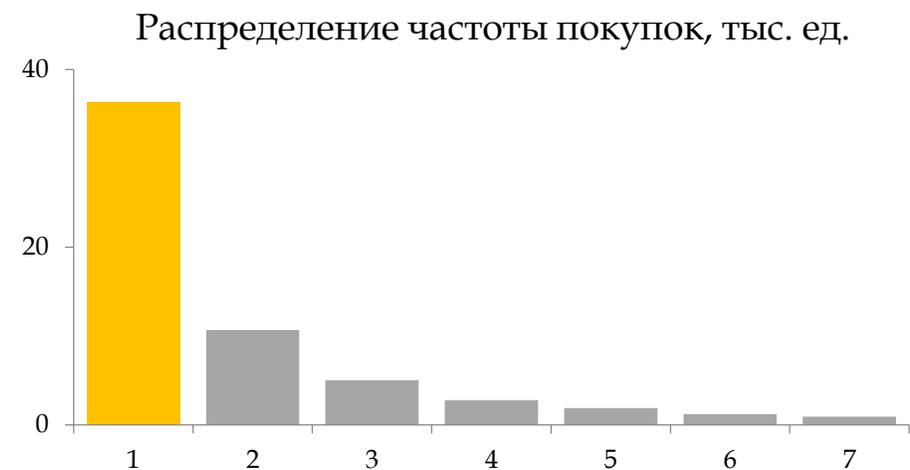
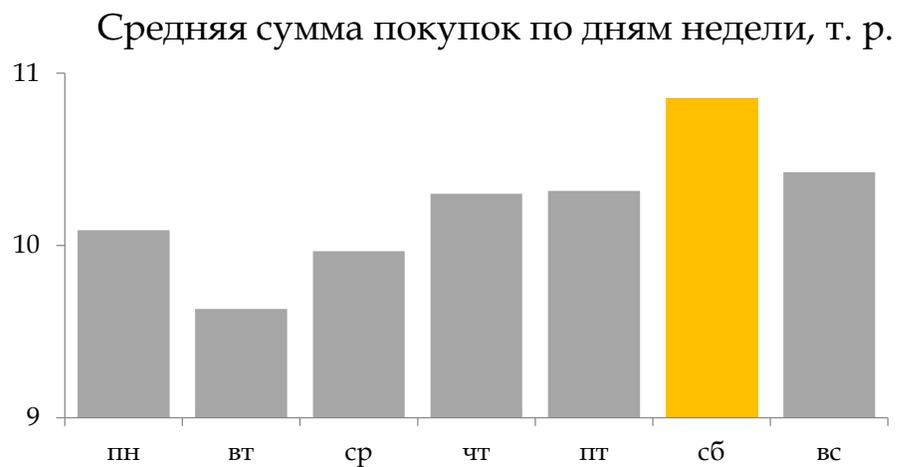
Количество записей – 162 610

Количество клиентов – 62 604

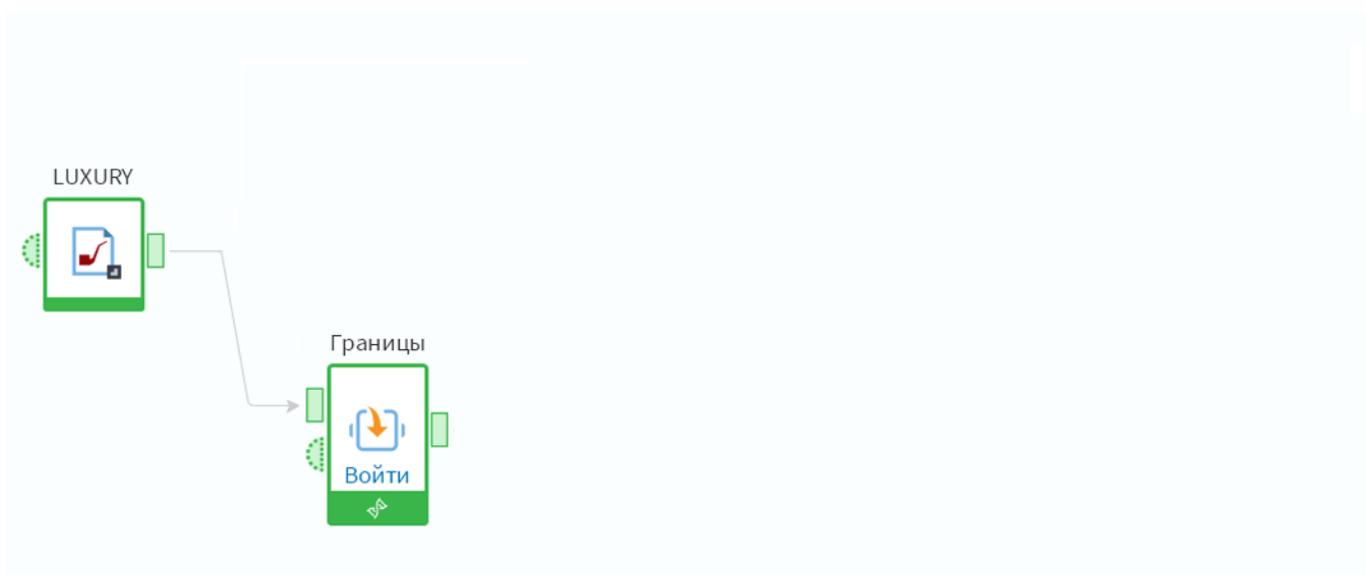
Период транзакций – с 22 апреля 2016г. по 6 декабря 2018г.



КЛИЕНТСКАЯ БАЗА: ОПИСАНИЕ

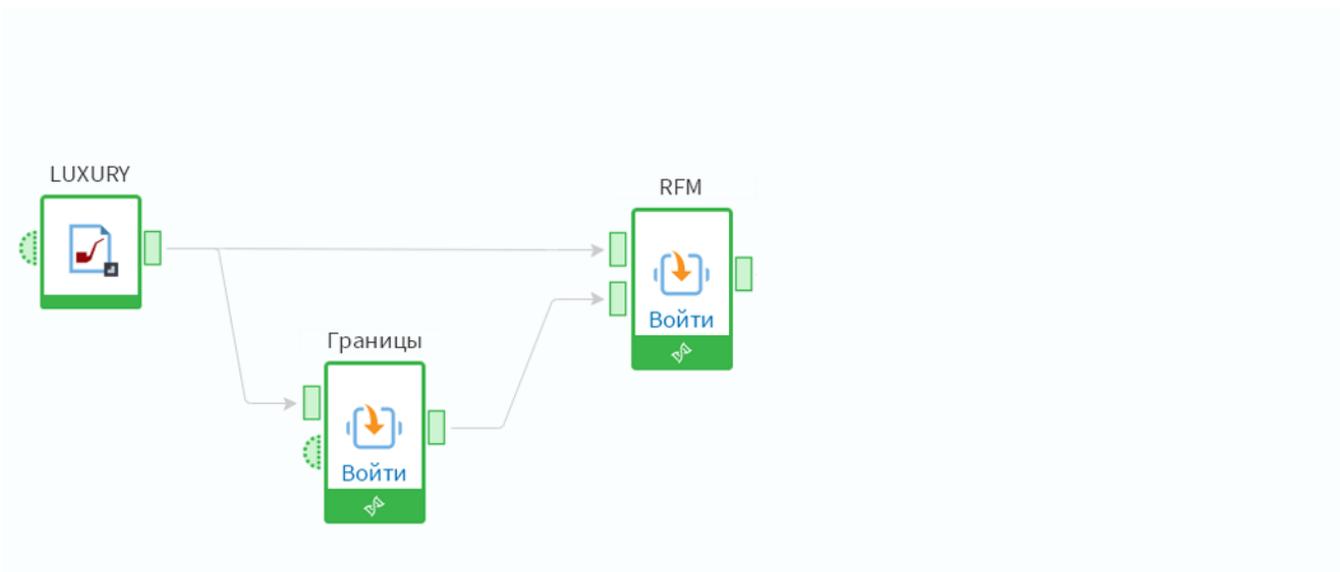


RFM: ВЫЧИСЛЕНИЕ ГРАНИЦ



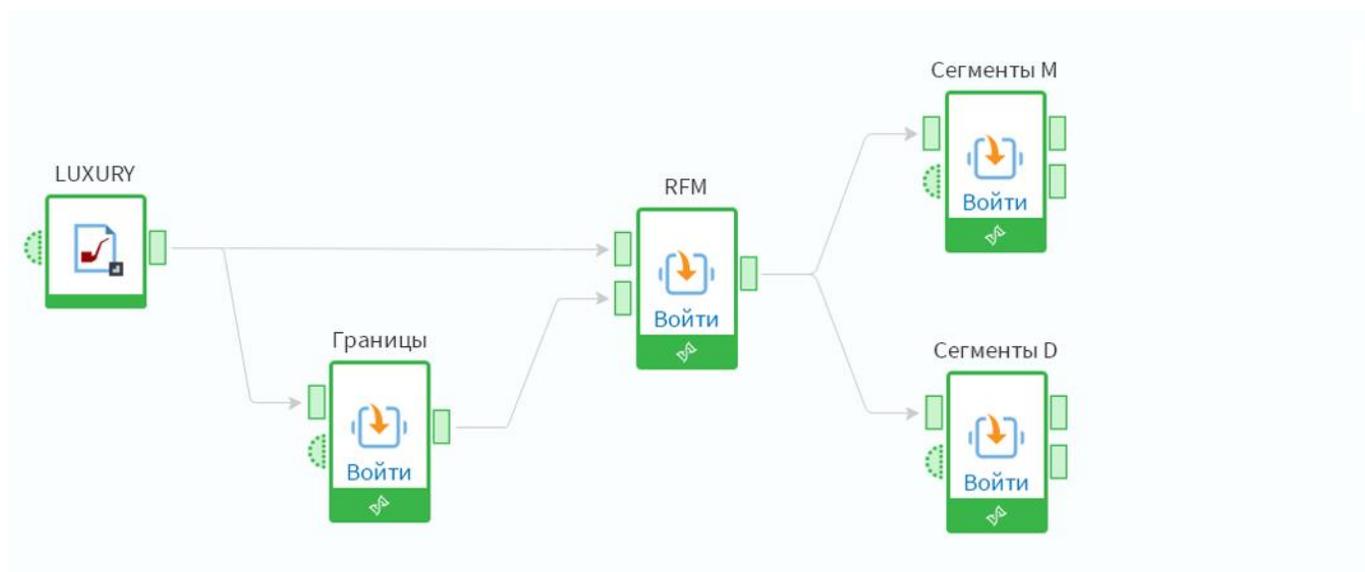
Границы • Границы • Быстрый просмотр данных					
#	ab	Признак	12 Интервал	9.0 Нижняя граница	9.0 Верхняя граница
1	L		1	0,00	0,00
2	L		2	0,00	952,00
3	R		1	174,75	402,75
4	R		2	402,75	1 132,75
5	F		1	1,00	1,00
6	F		2	1,00	131,00
7	M		1	100,00	7 733,00
8	M		2	7 733,00	2 675 857,00

RFM: РАСЧЕТ КОДА



RFM • RFM-код • Быстрый просмотр данных		
#	ab CLIENT_ID	ab RFM
1	900791F9042F669580820F7980F14BFA	211
2	D24FE5BC7A616DD774705E5D54EE2AD0	111
3	11177459EA606EC276715CEEAEF3C02E	211
4	C1A5B87817A0B9622C224A2126A7B468	211
5	BBEF54DEA6DCBA3C2A0B121E34DA3ED8	221
6	50632384AC34FFCBE84A9B8DECBF924C	111
7	1F84EE08E2AF6114523FABF1D6B95303	111
8	D9E940EDC8F29E87A800DDC361D660FA	211
9	EDAAC8D3CB8FA2DB8D272EC9B704D95D	111
10	F7EB23838E7C5C88AD4F39B8D8260BB9	222
62 604	0C201923419FCB9592C85CDF5F575D57	222

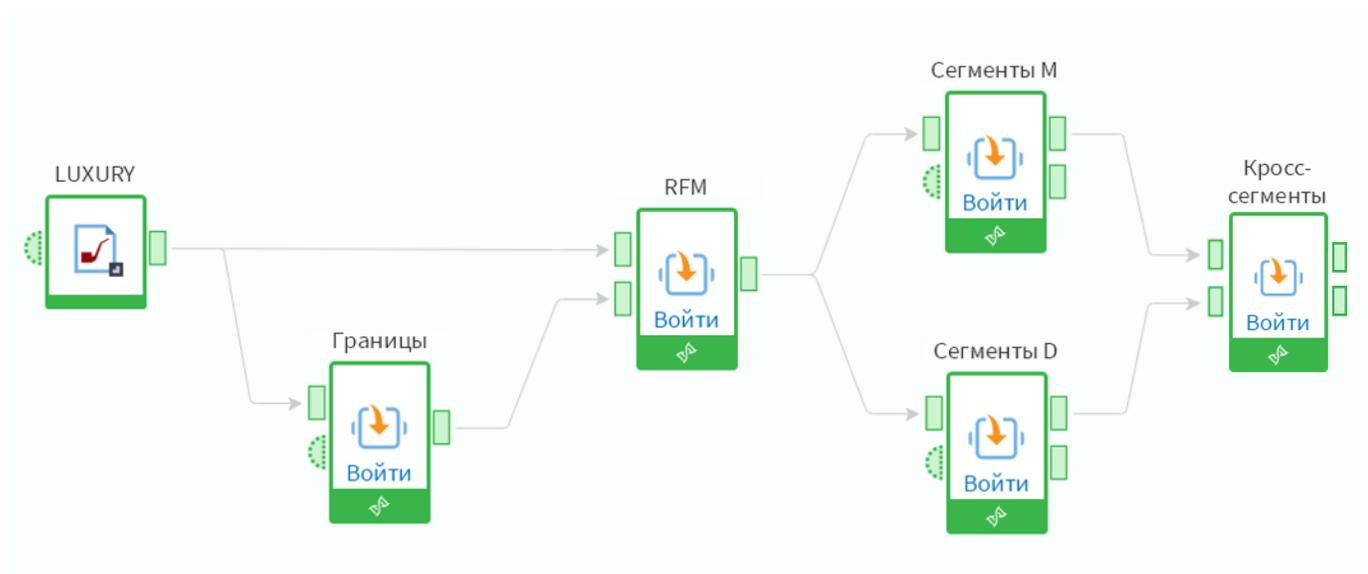
RFM: СЕГМЕНТИРОВАНИЕ



Сегменты M • Сегменты • Быстрый просмотр данных			
#	ab CLIENT_ID	ab Код	ab Сегмент
1	900791F9042F669580820F7980F14BFA	211	Золотые
2	D24FE5BC7A616DD774705E5D54EE2AD0	111	Платиновые
3	11177459EA606EC276715CEEA3C02E	211	Золотые
4	C1A5B87817A0B9622C224A2126A7B468	211	Золотые
5	BBEF54DEA6DCBA3C2A0B121E34DA3ED8	221	Серебряные
6	50632384AC34FFCBE84A9B8DEC924C	111	Платиновые
62 604	1F84EE08E2AF6114523FABF1D6B95303	111	Платиновые

Сегменты D • Сегменты • Быстрый просмотр данных			
#	ab CLIENT_ID	ab Код	ab Сегмент
1	900791F9042F669580820F7980F14BFA	211	Постоянные
2	D24FE5BC7A616DD774705E5D54EE2AD0	111	VIP
3	11177459EA606EC276715CEEA3C02E	211	Постоянные
4	C1A5B87817A0B9622C224A2126A7B468	211	Постоянные
5	BBEF54DEA6DCBA3C2A0B121E34DA3ED8	221	Спящие
6	50632384AC34FFCBE84A9B8DEC924C	111	VIP
62 604	1F84EE08E2AF6114523FABF1D6B95303	111	VIP

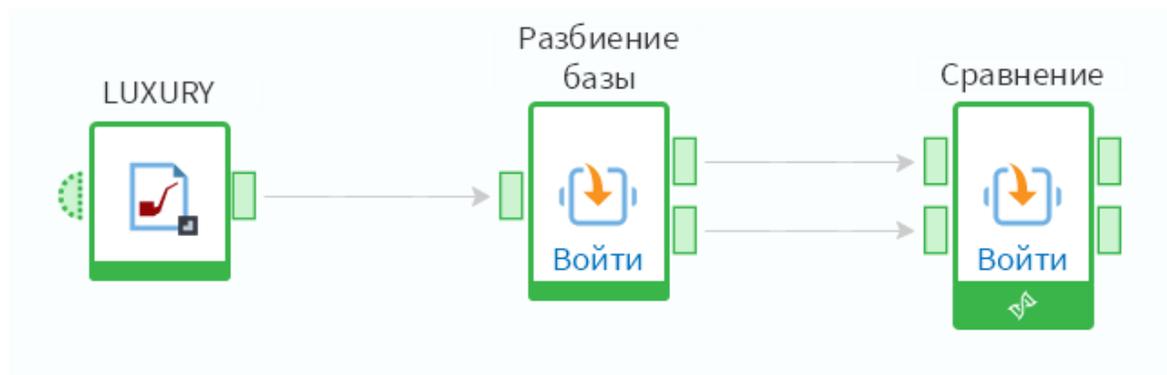
RFM: СЕГМЕНТЫ «НА ПЕРЕСЕЧЕНИИ»



Параметры полей • Выходной набор данных • Быстрый просмотр данных

#	ab CLIENT_ID	ab Кросс-сегменты
1	900791F9042F669580820F7980F14BFA	Золотые-Постоянные
2	D24FE5BC7A616DD774705E5D54EE2AD0	Платиновые-VIP
3	11177459EA606EC276715CEEAEF3C02E	Золотые-Постоянные
4	C1A5B87817A0B9622C224A2126A7B468	Золотые-Постоянные
5	BBEF54DEA6DCBA3C2A0B121E34DA3ED8	Серебряные-Спящие
6	50632384AC34FFCBE84A9B8DECBF924C	Платиновые-VIP
7	1F84EE08E2AF6114523FABF1D6B95303	Платиновые-VIP
8	D9E940EDC8F29E87A800DDC361D660FA	Золотые-Постоянные
9	EDAAC8D3CB8FA2DB8D272EC9B704D95D	Платиновые-VIP
62 604	F7EB23838E7C5C88AD4F39B8D8260BB9	Бронзовые-Спящие

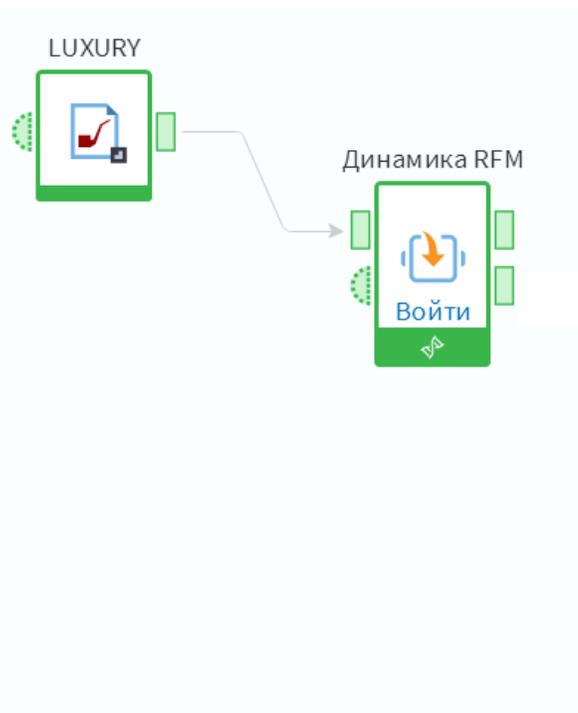
RFM: СРАВНЕНИЕ ЦЕННОСТИ КЛИЕНТОВ



Сравнение • Сравнение • Быстрый просмотр данных

#	ab CLIENT_ID	ab Сравнение	ab Сегмент	ab Сегмент_было
1	900791F9042F669580820F7980F14BFA	!	VIP	<null>
2	D24FE5BC7A616DD774705E5D54EE2AD0	>	VIP	Транжиры
3	11177459EA606EC276715CEEAEF3C02E	!	VIP	<null>
4	C1A5B87817A0B9622C224A2126A7B468	!	Постоянные	<null>
5	BBEF54DEA6DCBA3C2A0B121E34DA3ED8	!	Спящие	<null>
6	50632384AC34FFCBE84A9B8DECBF924C	=	VIP	VIP
7	1F84EE08E2AF6114523FABF1D6B95303	=	Транжиры	Транжиры
8	D9E940EDC8F29E87A800DDC361D660FA	!	VIP	<null>
9	EDAAC8D3CB8FA2DB8D272EC9B704D95D	<	Спящие	VIP
62 604	F7EB23838E7C5C88AD4F39B8D8260BB9	!	Спящие	<null>

RFM: АНАЛИЗ ДИНАМИКИ СЕГМЕНТОВ



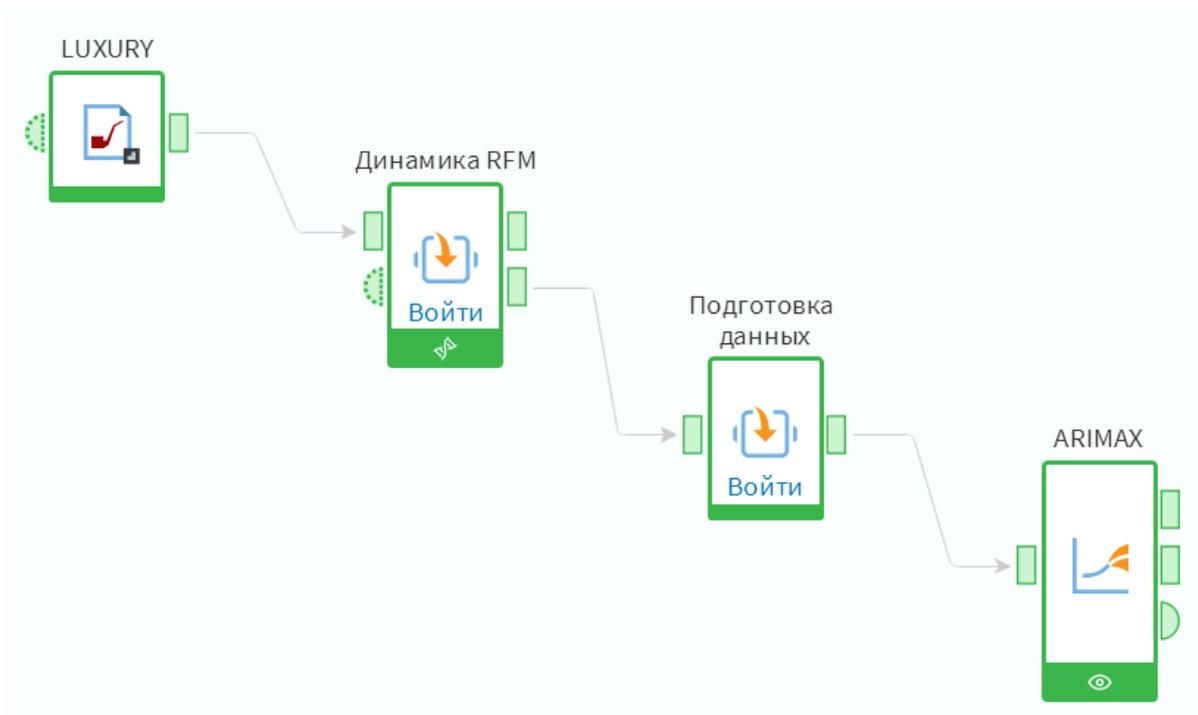
Динамика RFM • Динамика M • Быстрый просмотр данных

#	ab Месяц	9.0 Платиновые Объем	9.0 Золотые Объем	9.0 Серебряные Объем
1	2016-M04	13,00	81,00	202,00
2	2016-M05	65,00	179,00	561,00
3	2016-M06	62,00	187,00	658,00
4	2016-M07	128,00	304,00	948,00
5	2016-M08	77,00	277,00	703,00
6	2016-M09	92,00	293,00	682,00
7	2016-M10	145,00	418,00	816,00
33	2017-M07	287,00	745,00	1.864,00

Динамика RFM • Динамика D • Быстрый просмотр данных

#	ab Месяц	9.0 VIP Объем	9.0 Постоянные Объем	9.0 Транжиры Объем
1	2016-M04	13,00	76,00	207,00
2	2016-M05	65,00	148,00	485,00
3	2016-M06	62,00	158,00	562,00
4	2016-M07	128,00	221,00	860,00
5	2016-M08	77,00	222,00	609,00
6	2016-M09	92,00	244,00	548,00
7	2016-M10	145,00	371,00	604,00
33	2017-M07	287,00	545,00	1.605,00

RFM: ПРОГНОЗИРОВАНИЕ



ARIMAX • Выход модели • Быстрый просмотр данных

#	ab Месяц	9.0 VIP	9.0 VIP Прогноз
1	2016-M04	13,00	
2	2016-M05	65,00	90,89
3	2016-M06	62,00	135,29
4	2016-M07	128,00	132,73
5	2016-M08	77,00	189,09
6	2016-M09	92,00	145,54
7	2016-M10	145,00	158,34
38			

ARIMAX • Выход модели • Быстрый просмотр данных

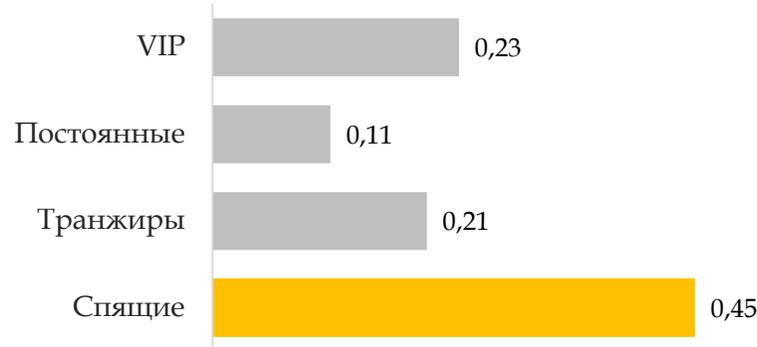
#	ab Месяц	9.0 VIP M_avg	9.0 Прогноз M_avg
1	2016-M04	29 663,46	
2	2016-M05	39 870,78	36 646,46
3	2016-M06	35 135,61	41 598,21
4	2016-M07	32 281,69	39 301,09
5	2016-M08	61 527,30	37 916,60
6	2016-M09	52 603,64	52 104,15
7	2016-M10	65 966,27	47 775,13
38			

RFM-АНАЛИЗ КЛИЕНТСКОЙ БАЗЫ

Доли клиентских сегментов (M)



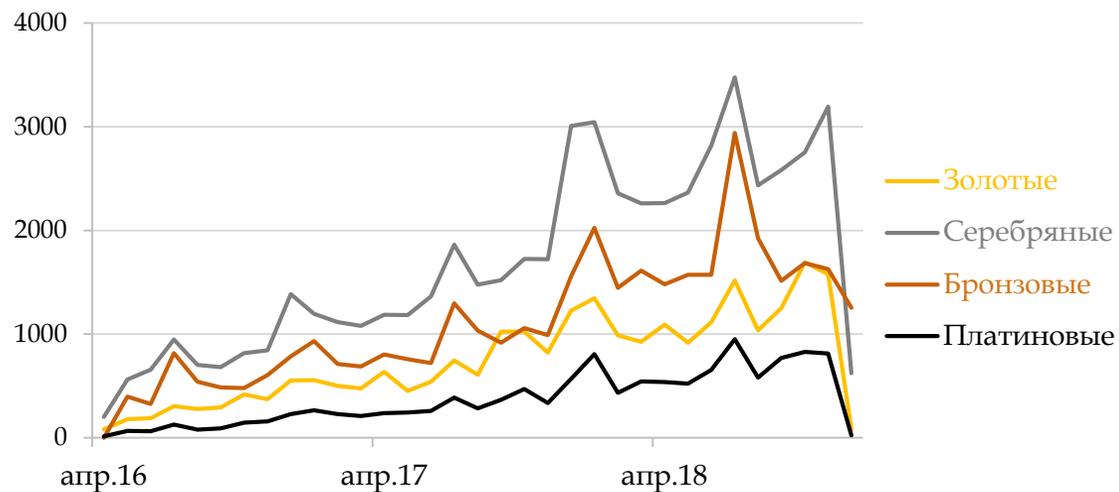
Доли клиентских сегментов (D)



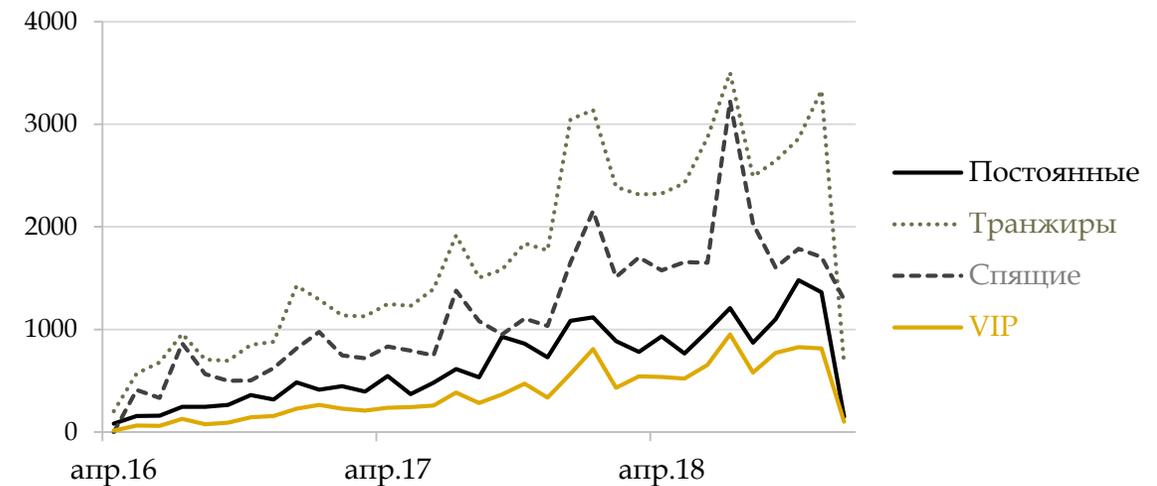
Доли сегментов «на пересечении»

Платиновые	0,23			
Золотые		0,17	0,05	
Серебряные			0,19	0,10
Бронзовые				0,26
	VIP	Постоянные	Транжиры	Спящие

Динамика объема сегментов (M)

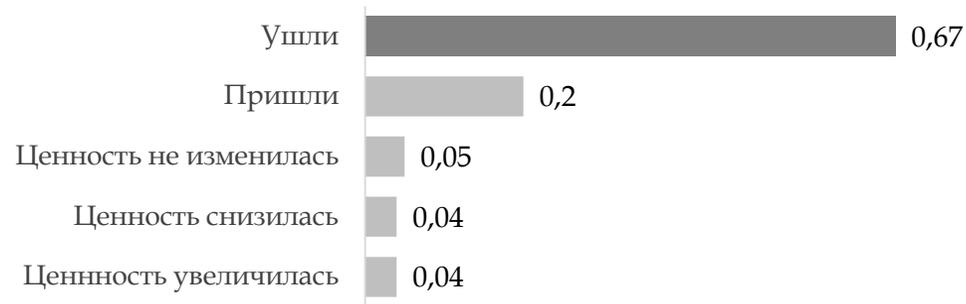


Динамика объема сегментов (D)

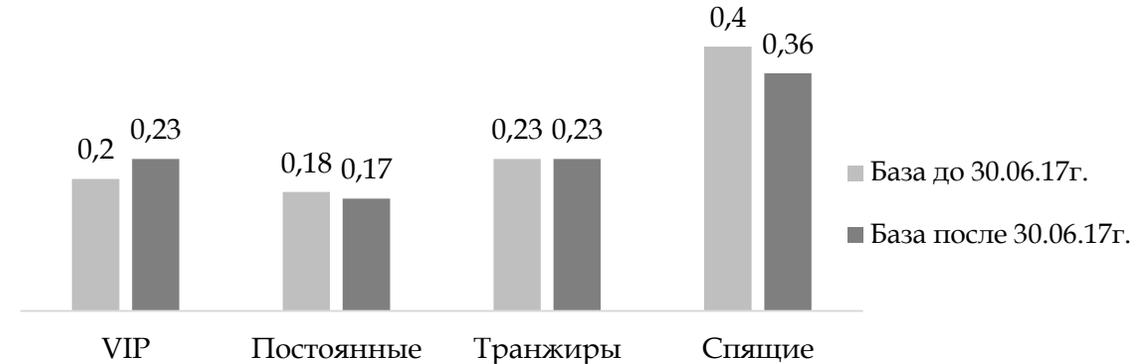


RFM-АНАЛИЗ КЛИЕНТСКОЙ БАЗЫ

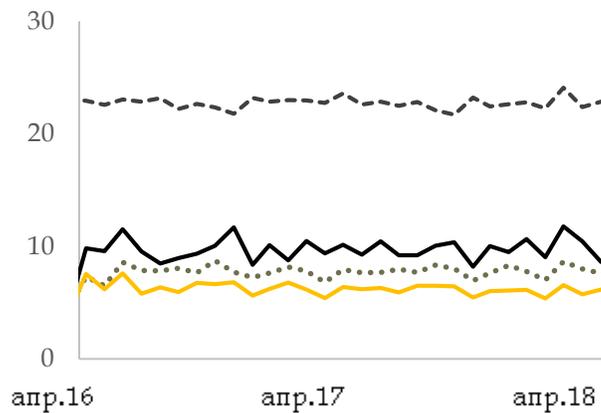
Статистика изменения ценности клиентов



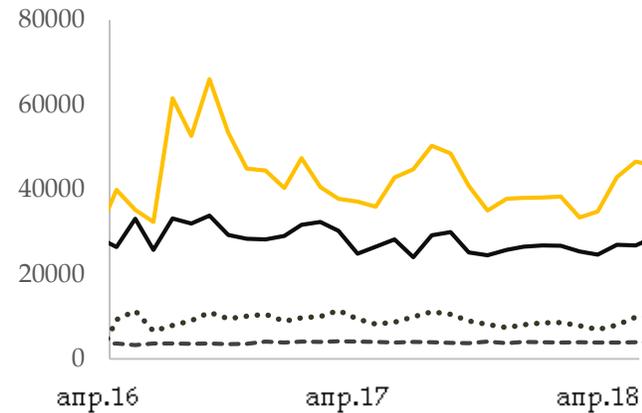
Изменение долей сегментов для двух периодов



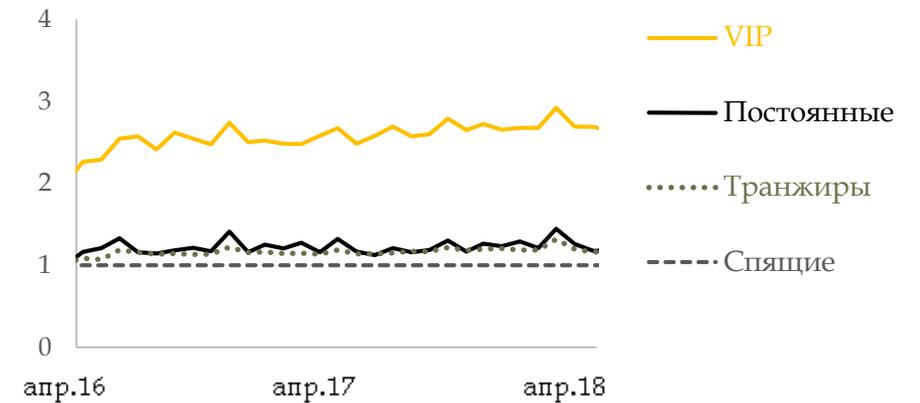
Динамика средней давности



Динамика среднего чека

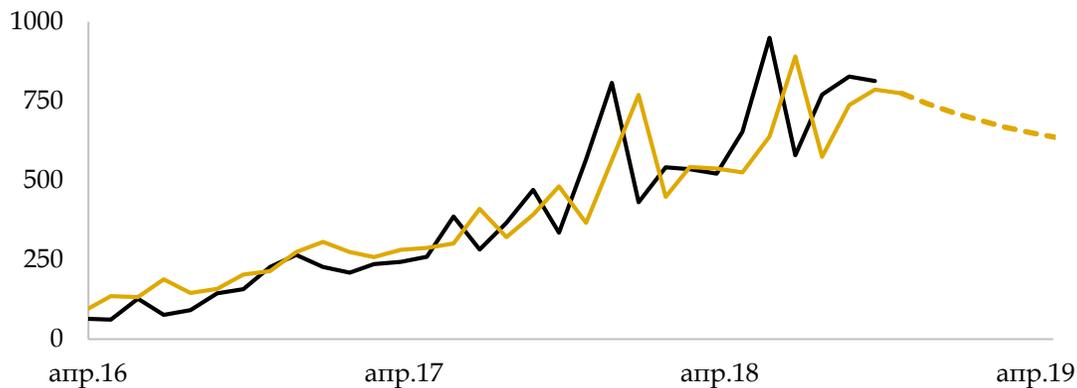


Динамика средней частоты

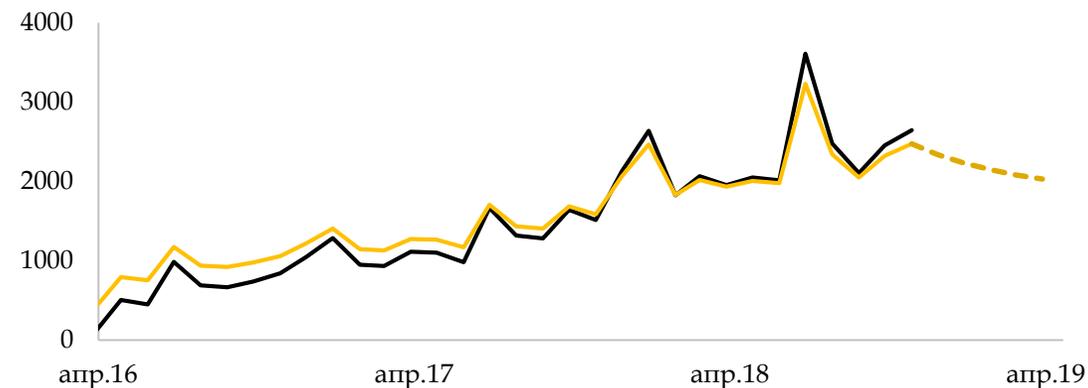


RFM-АНАЛИЗ КЛИЕНТСКОЙ БАЗЫ

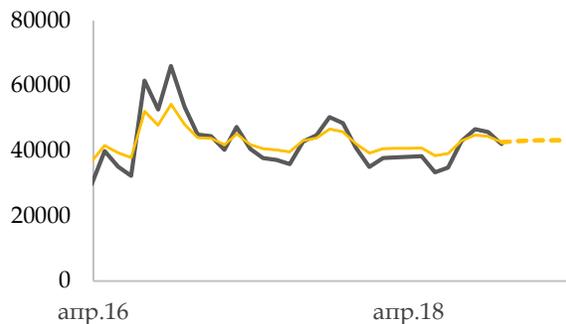
Объем сегмента *VIP* на полгода вперед (декабрь 2016-май 2019)



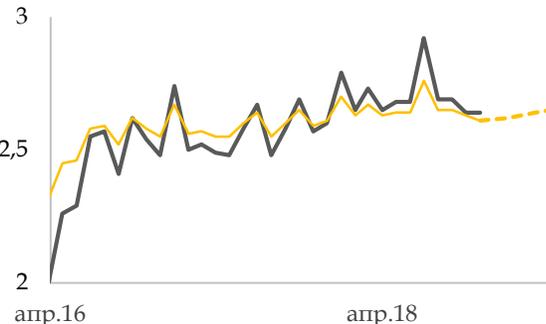
Объем сегмента *Спящие* на полгода вперед (декабрь 2016-май 2019)



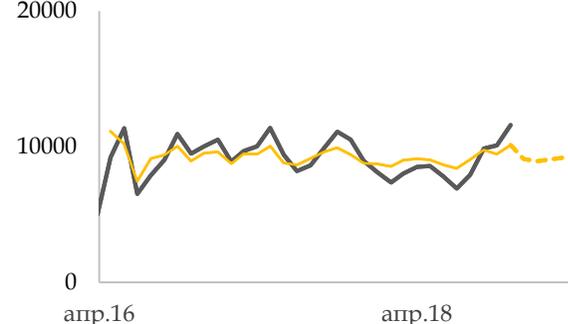
VIP: динамика среднего чека



VIP: динамика посещений



Спящие: динамика среднего чека

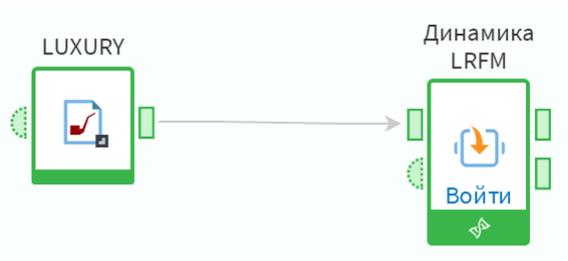
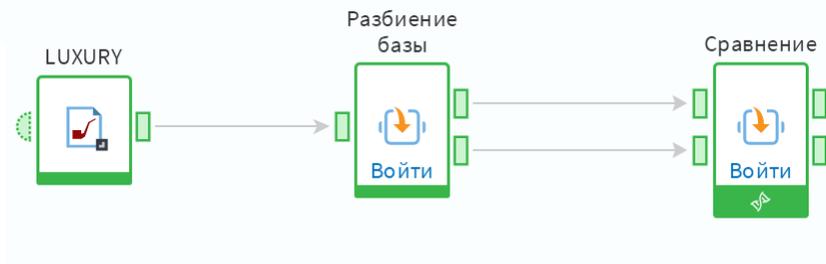
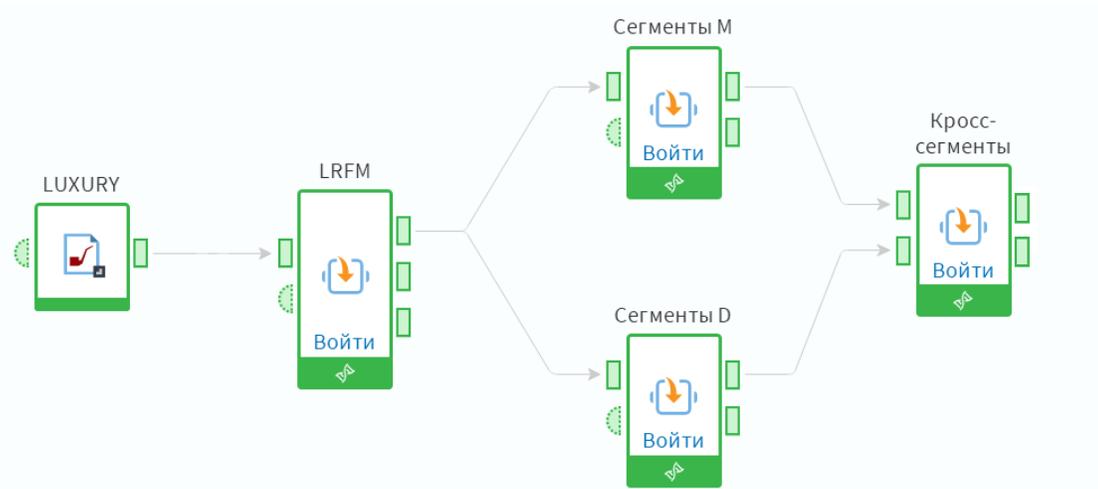


Спящие: динамика посещений



— Факт — Модель - - - Прогноз

LRFM-АНАЛИЗ КЛИЕНТСКОЙ БАЗЫ



Сравнение • Сравнение • Быстрый просмотр данных

#	ab CLIENT_ID	ab Сравнение	ab Сегмент	ab Сегмент_было
1	900791F9042F669580820F7980F14BFA	!	Лучшие	<null>
2	D24FE5BC7A616DD774705E5D54EE2AD0	>	Лучшие	Постоянные
3	11177459EA606EC276715CEEAFF3C02E	!	Лучшие	<null>

LRFM • Сегменты • Быстрый просмотр данных

#	ab CLIENT_ID	ab LRFM	ab Сегмент	ab Тип отношений
1	900791F9042F669580820F7980F14BFA	1211	Лучшие	Потенциальная связь
2	D24FE5BC7A616DD774705E5D54EE2AD0	1111	Лучшие	Тесная связь
3	11177459EA606EC276715CEEAFF3C02E	1211	Лучшие	Потенциальная связь

Сегменты D • Сегменты • Быстрый просмотр данных

#	ab CLIENT_ID	ab Код	ab Сегмент
1	900791F9042F669580820F7980F14BFA	1211	Потенциальные
2	D24FE5BC7A616DD774705E5D54EE2AD0	1111	Основные
3	11177459EA606EC276715CEEAFF3C02E	1211	Потенциальные

LRFM • LRFM-код • Быстрый просмотр данных

#	ab CLIENT_ID	ab LRFM
1	900791F9042F669580820F7980F14BFA	1211
2	D24FE5BC7A616DD774705E5D54EE2AD0	1111
3	11177459EA606EC276715CEEAFF3C02E	1211
4	C1A5B87817A0B9622C224A2126A7B468	2221
5	BBEF54DEA6DCBA3C2A0B121E34DA3ED8	2222
6	50632384AC34FFCBE84A9B8DEC924C	1111
7	1F84EE08E2AF6114523FABF1D6B95303	1112
62 604	D9E940EDC8F29E87A800DDC361D660FA	1222

LRFM-АНАЛИЗ КЛИЕНТСКОЙ БАЗЫ

Доли клиентских сегментов (M)



Доли клиентских сегментов (D)



Доли сегментов «на пересечении»

Платиновые	0,15	0,06	0,01	
Золотые	0,01	0,03	0,03	0,05
Серебряные		0,01	0,02	0,32
Бронзовые				0,33
	Лучшие	Постоянные	Транжиры	Неопред.



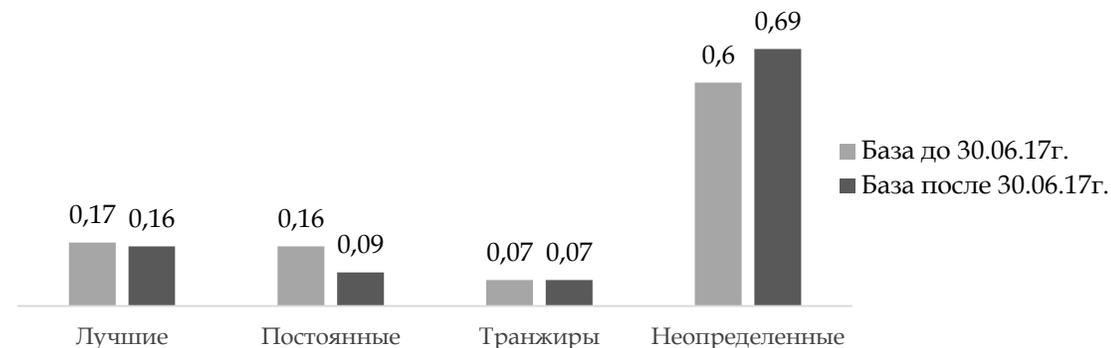
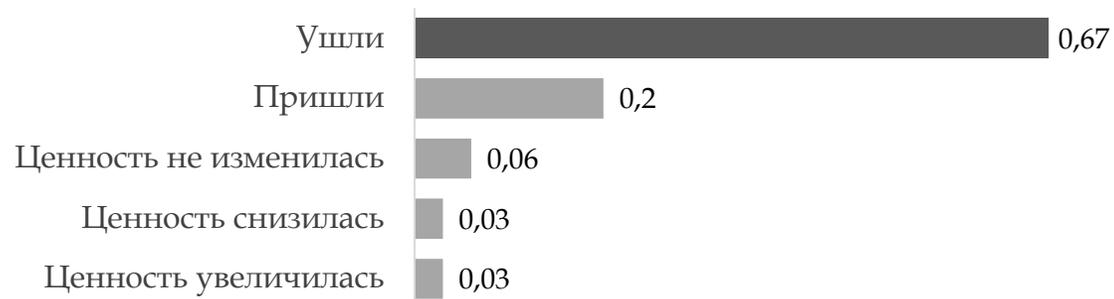
Матрица взаимоотношений с клиентами



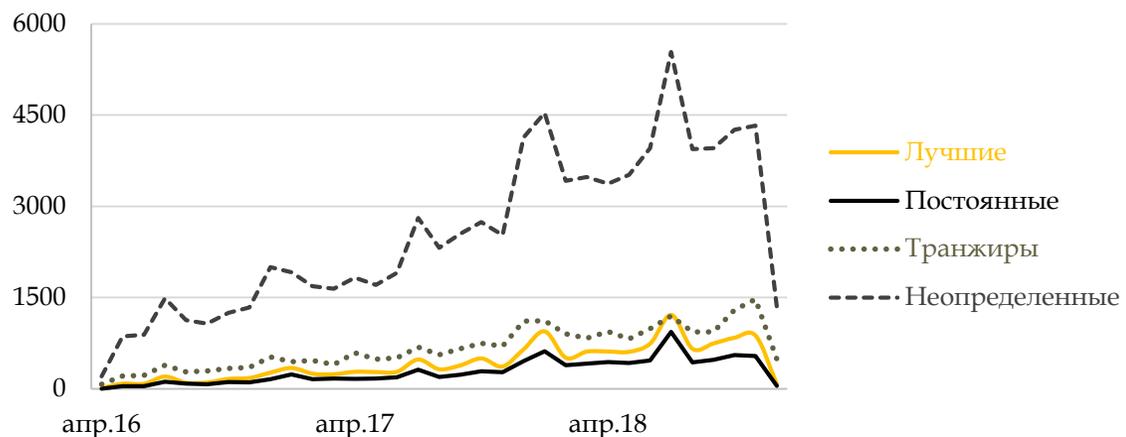
Матрица ценности клиентов

LRFM-АНАЛИЗ КЛИЕНТСКОЙ БАЗЫ

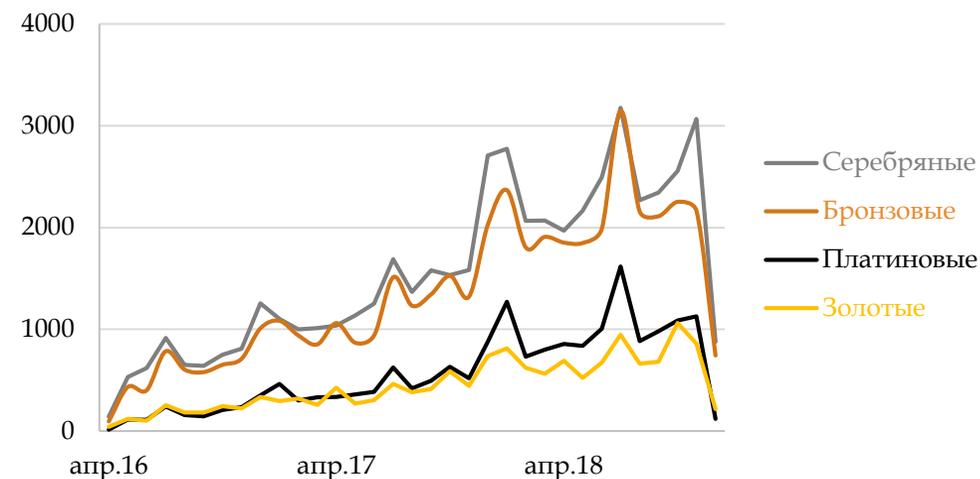
Статистика изменения ценности клиентов



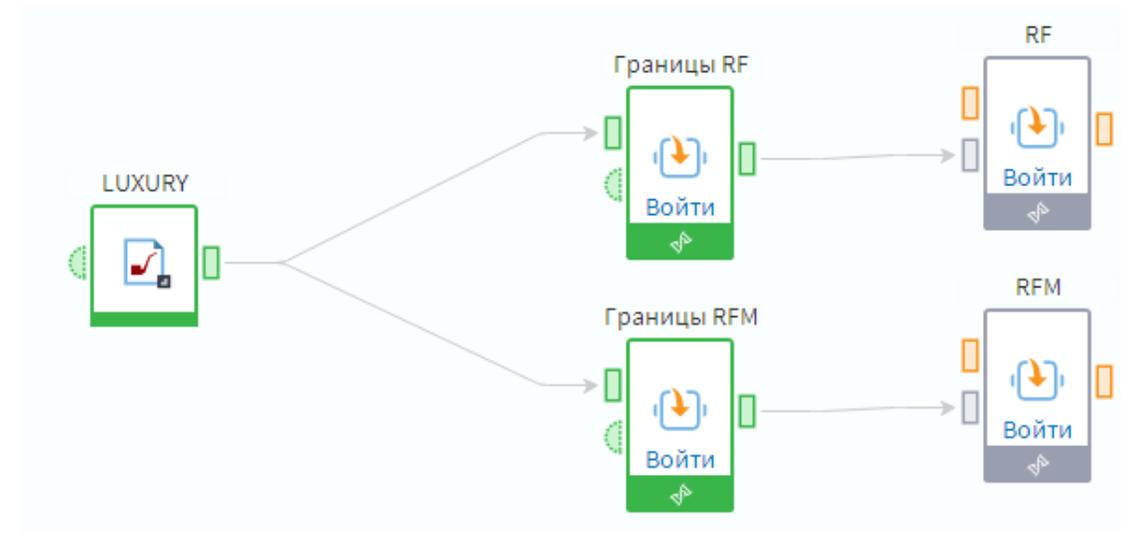
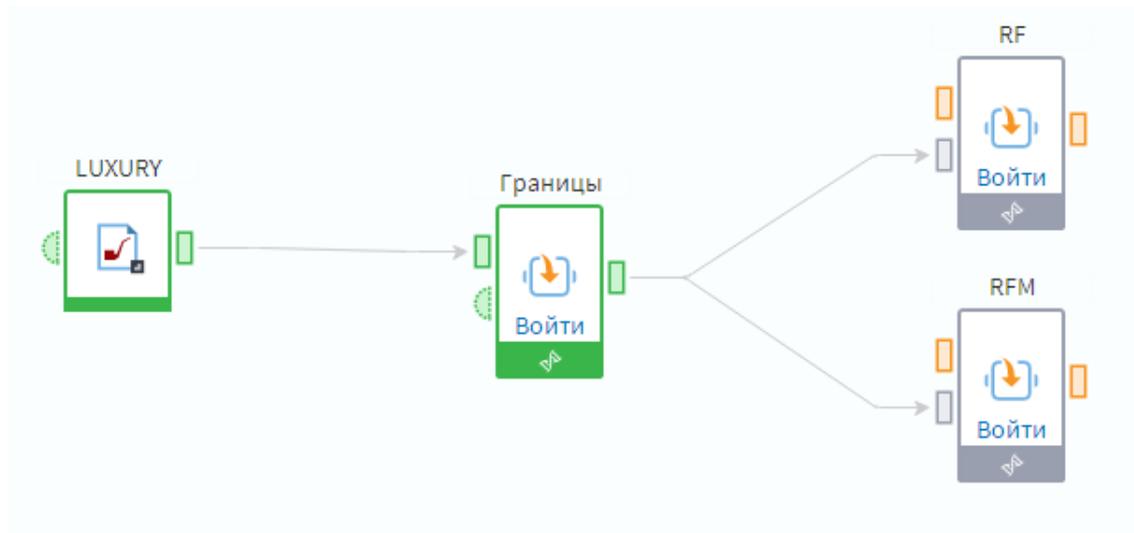
Динамика объемов сегментов (D)



Динамика объемов сегментов (M)



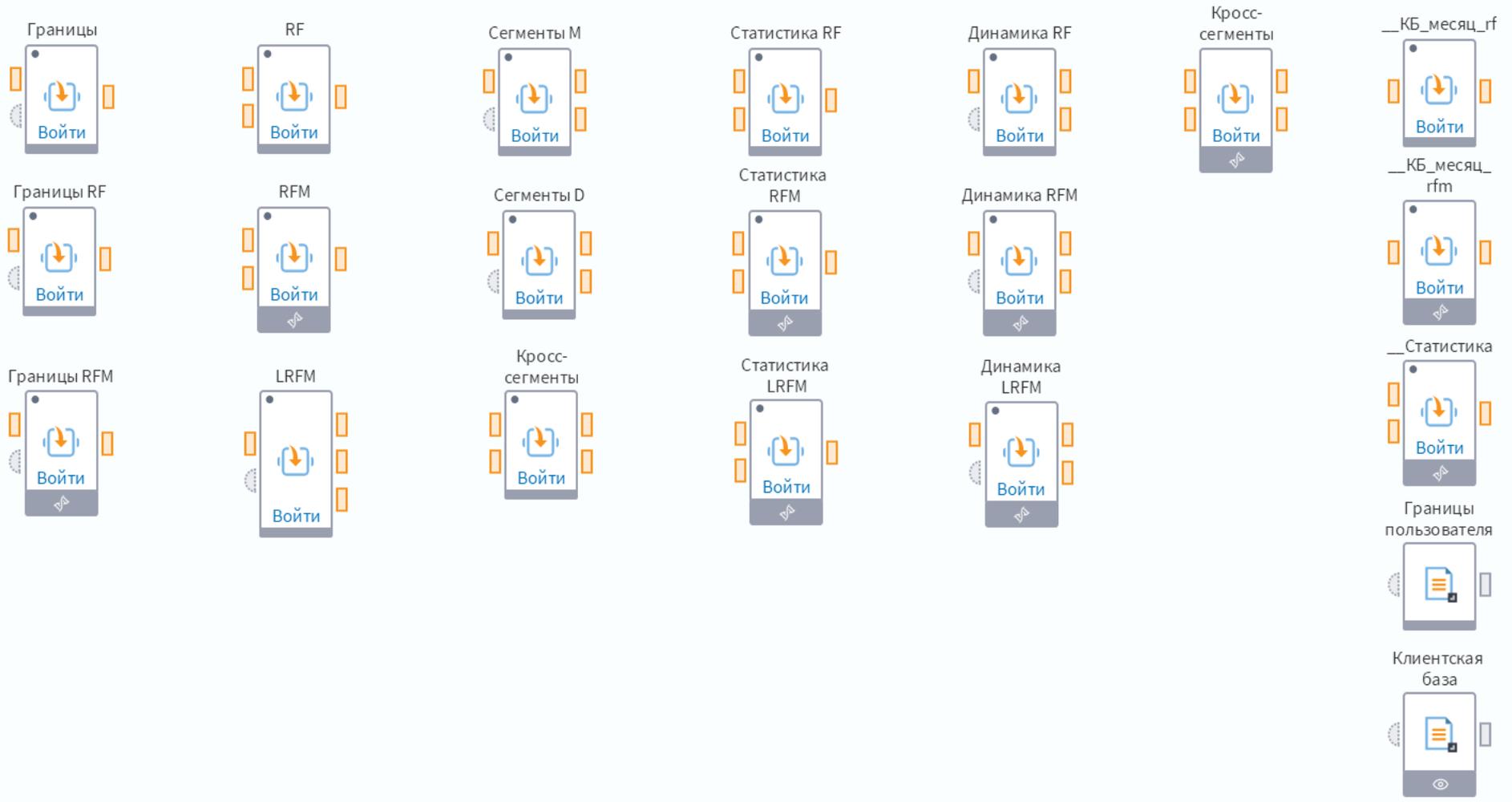
СПОРНЫЙ ВОПРОС



Компоненты

Фильтрация

- Импорт
 - Текстовый файл
 - Loginom Data файл
- Трансформация
 - Группировка
 - Дата и время
 - Дополнение данных
 - Замена
 - Калькулятор
 - Кросс-таблица
 - Объединение
 - Параметры полей
 - Разгруппировка
 - Свертка столбцов
 - Скользящее окно
 - Слияние
 - Соединение
 - Сортировка
 - Фильтр строк
 - JavaScript
- Управление
 - Выполнение узла
 - Подмодель
 - Узел-ссылка
 - Условие
 - Цикл
- Исследование
 - Автокорреляция
 - Корреляционный анализ
 - Факторный анализ
- Предобработка
 - Заполнение пропусков
 - Квантование
 - Конечные классы
 - Разбиение на множества



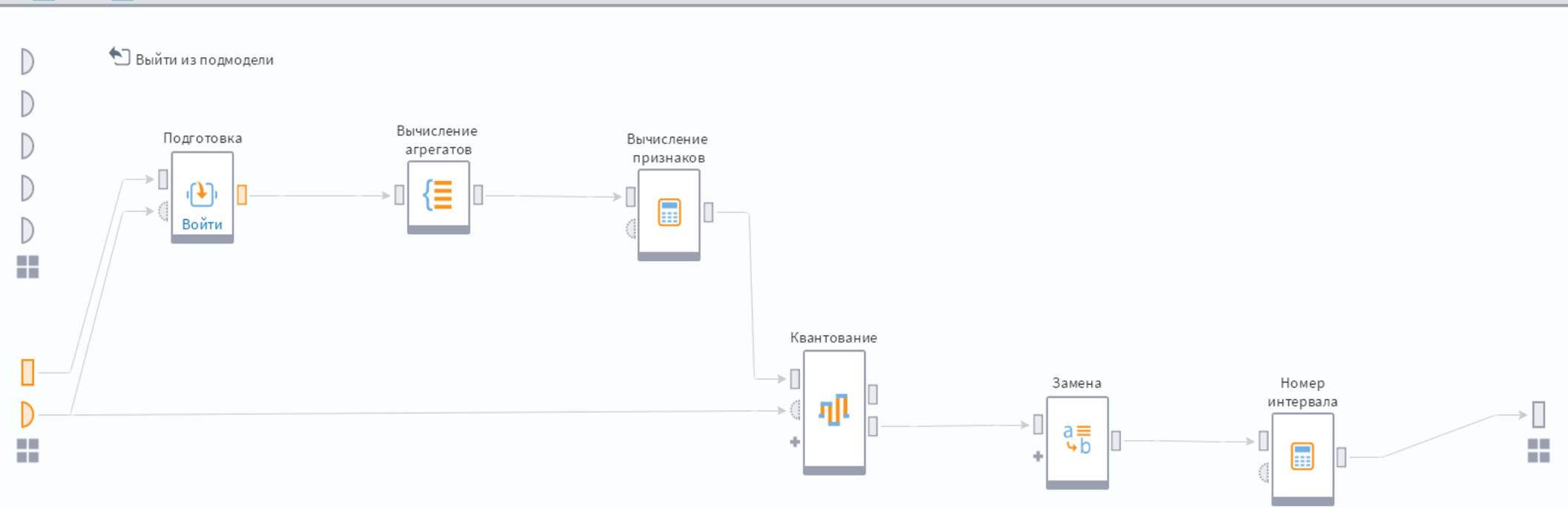
Производные компоненты +

Подключения +

Компоненты

Фильтрация

- Импорт
 - Текстовый файл
 - Loginom Data файл
- Трансформация
 - Группировка
 - Дата и время
 - Дополнение данных
 - Замена
 - Калькулятор
 - Кросс-таблица
 - Объединение
 - Параметры полей
 - Разгруппировка
 - Свертка столбцов
 - Скользящее окно
 - Слияние
 - Соединение
 - Сортировка
 - Фильтр строк
 - JavaScript
- Управление
 - Выполнение узла
 - Подмодель
 - Узел-ссылка
 - Условие
 - Цикл
- Исследование
 - Автокорреляция
 - Корреляционный анализ
 - Факторный анализ
- Предобработка
 - Заполнение пропусков
 - Квантование
 - Конечные классы
 - Разбиение на множества
 - Редактирование выбросов

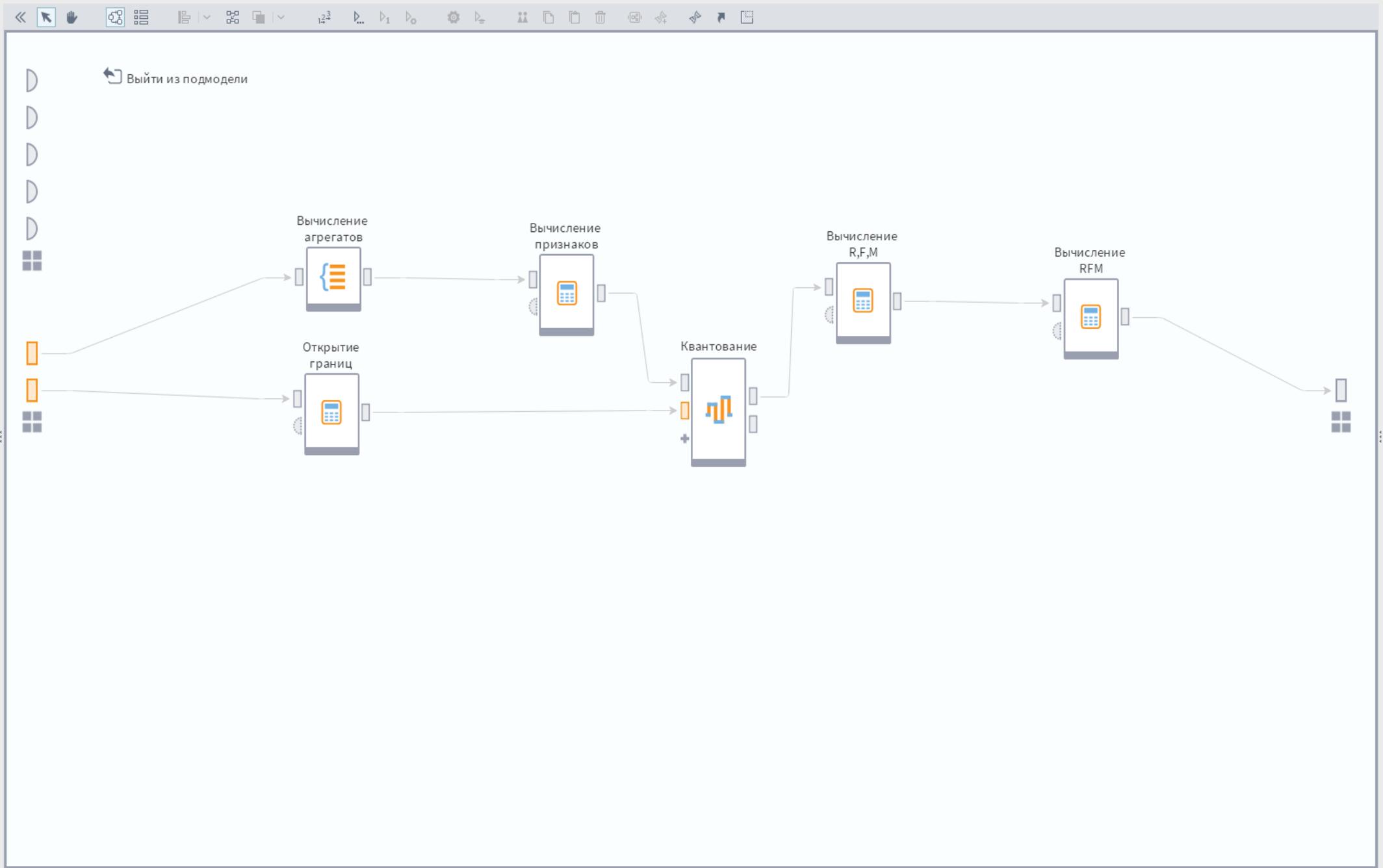


Производные компоненты +

Подключения +

Компоненты

- Фильтрация
- Импорт
 - Текстовый файл
 - Loginom Data файл
 - Трансформация
 - Группировка
 - Дата и время
 - Дополнение данных
 - Замена
 - Калькулятор
 - Кросс-таблица
 - Объединение
 - Параметры полей
 - Разгруппировка
 - Свертка столбцов
 - Скользящее окно
 - Слияние
 - Соединение
 - Сортировка
 - Фильтр строк
 - JavaScript
 - Управление
 - Выполнение узла
 - Подмодель
 - Узел-ссылка
 - Условие
 - Цикл
 - Исследование
 - Автокорреляция
 - Корреляционный анализ
 - Факторный анализ
 - Предобработка
 - Заполнение пропусков
 - Квантование
 - Конечные классы
 - Разбиение на множества



Производные компоненты +

Подключения +

Компоненты

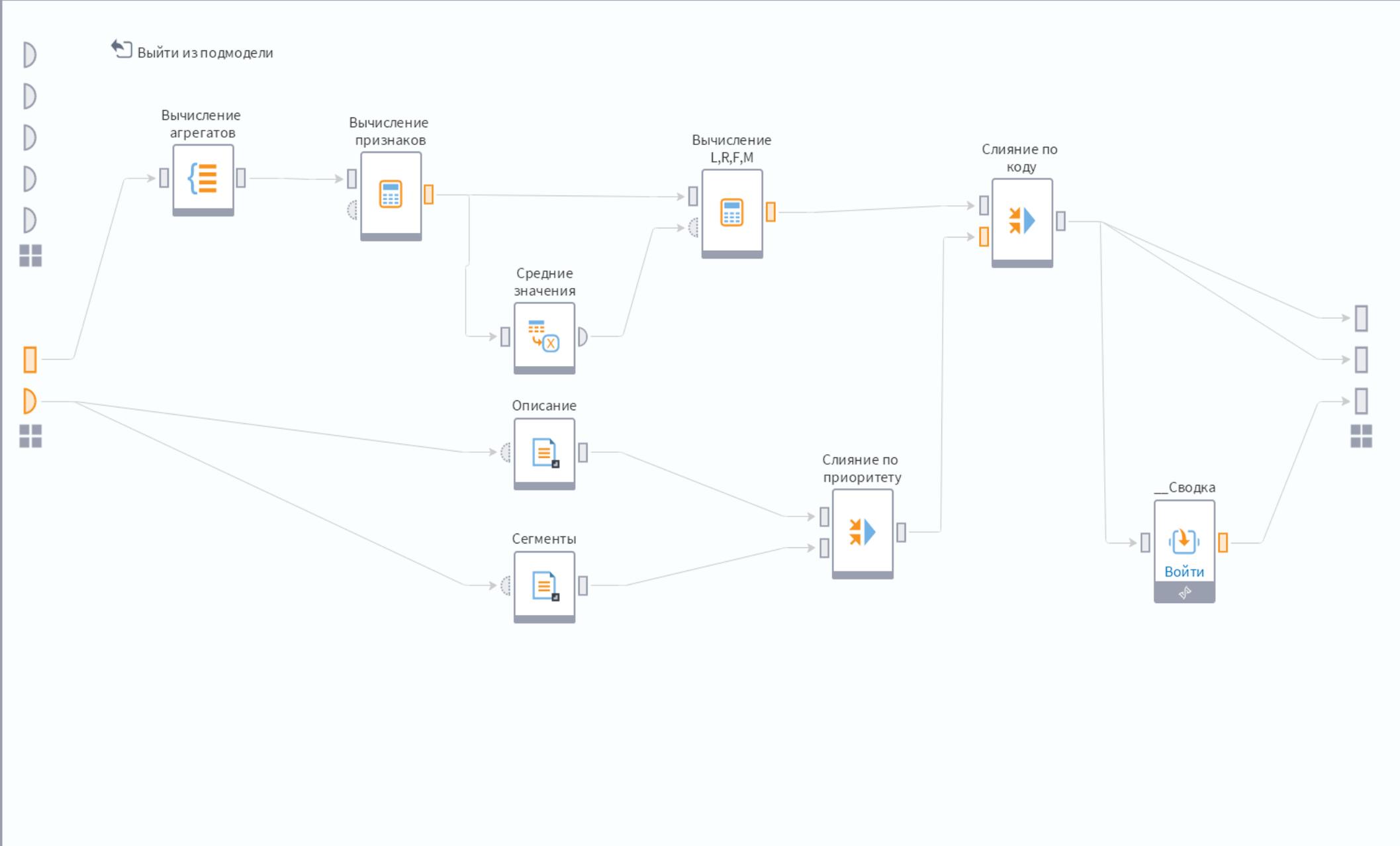
Фильтрация

- Импорт
 - Текстовый файл
 - Loginom Data файл
- Трансформация
 - Группировка
 - Дата и время
 - Дополнение данных
 - Замена
 - Калькулятор
 - Кросс-таблица
 - Объединение
 - Параметры полей
 - Разгруппировка
 - Свертка столбцов
 - Скользящее окно
 - Слияние
 - Соединение
 - Сортировка
 - Фильтр строк
 - JavaScript
- Управление
 - Выполнение узла
 - Подмодель
 - Узел-ссылка
 - Условие
 - Цикл
- Исследование
 - Автокорреляция
 - Корреляционный анализ
 - Факторный анализ
- Предобработка
 - Заполнение пропусков
 - Квантование
 - Конечные классы
 - Разбиение на множества

Производные компоненты +

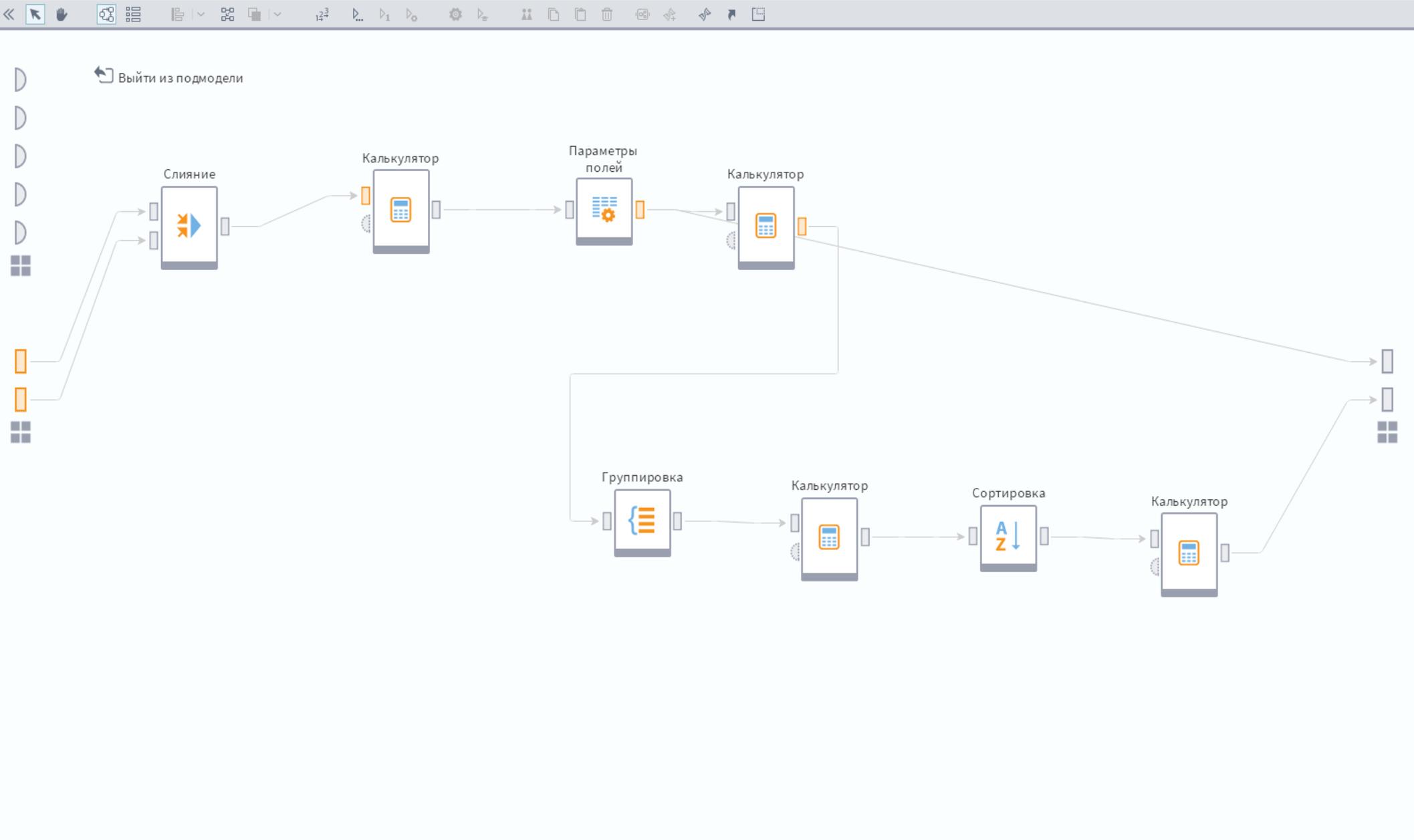
Подключения +

Выйти из подмодели



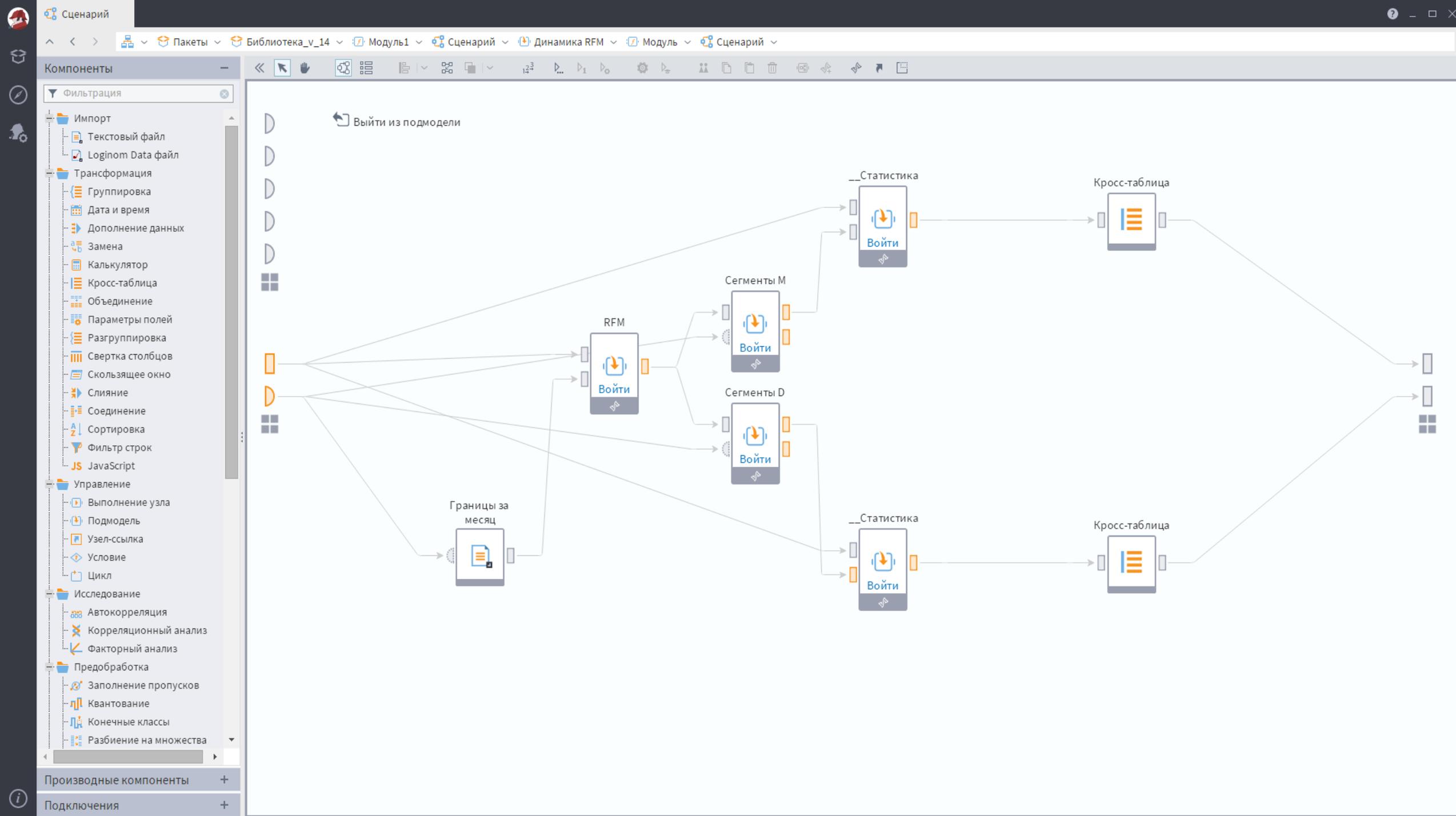
Компоненты

- Фильтрация
- Импорт
 - Текстовый файл
 - Logiном Data файл
 - Трансформация
 - Группировка
 - Дата и время
 - Дополнение данных
 - Замена
 - Калькулятор
 - Кросс-таблица
 - Объединение
 - Параметры полей
 - Разгруппировка
 - Свертка столбцов
 - Скользящее окно
 - Слияние
 - Соединение
 - Сортировка
 - Фильтр строк
 - JS JavaScript
 - Управление
 - Выполнение узла
 - Подмодель
 - Узел-ссылка
 - Условие
 - Цикл
 - Исследование
 - Автокорреляция
 - Корреляционный анализ
 - Факторный анализ
 - Предобработка
 - Заполнение пропусков
 - Квантование
 - Конечные классы
 - Разбиение на множества
 - Редактирование выбросов



Производные компоненты +

Подключения +



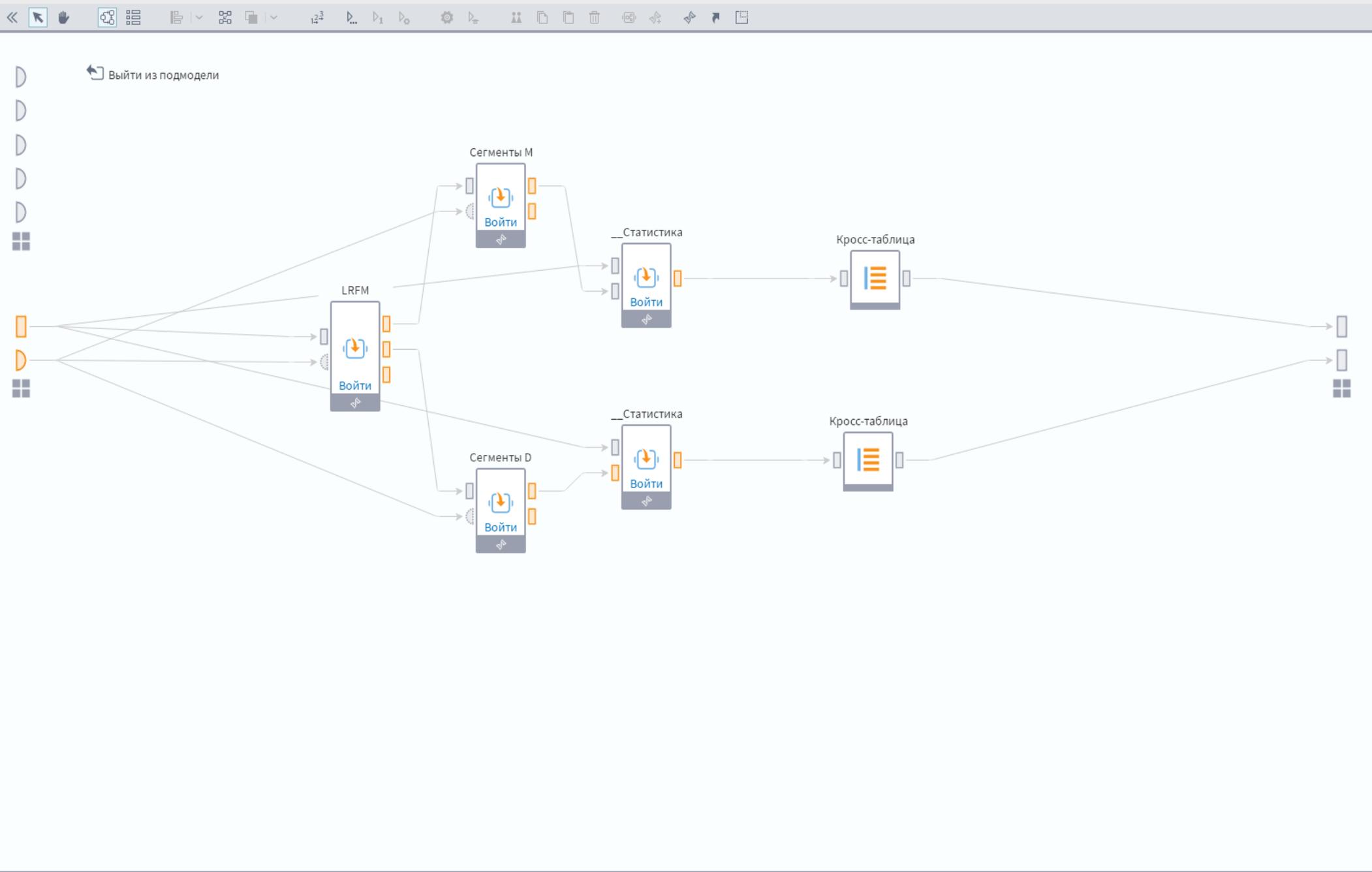
Компоненты

Фильтрация

- Импорт
 - Текстовый файл
 - Loginom Data файл
- Трансформация
 - Группировка
 - Дата и время
 - Дополнение данных
 - Замена
 - Калькулятор
 - Кросс-таблица
 - Объединение
 - Параметры полей
 - Разгруппировка
 - Свертка столбцов
 - Скользящее окно
 - Слияние
 - Соединение
 - Сортировка
 - Фильтр строк
 - JavaScript
- Управление
 - Выполнение узла
 - Подмодель
 - Узел-ссылка
 - Условие
 - Цикл
- Исследование
 - Автокорреляция
 - Корреляционный анализ
 - Факторный анализ
- Предобработка
 - Заполнение пропусков
 - Квантование
 - Конечные классы
 - Разбиение на множества

Производные компоненты +

Подключения +





LOGINOM
ХАКАТОН 2019

СПАСИБО ЗА ВНИМАНИЕ

