**Описание полей:**

|  |  |
| --- | --- |
| **Наименование поля** | **Описание** |
| **Система источник: oltp\_src\_system.account\_data** | |
| Id | Бизнес ключ |
| Status\_nm | Наименование статуса |
| Create\_dttm | Время создания |
| Update\_dttm | Время загрузки |
|  |  |
| **Система источник: oltp\_cdc\_src\_system. account\_data\_changes** | |
| Id | Бизнес ключ |
| Status\_nm | Наименование статуса |
| Create\_dttm | Время создания |
| Update\_dttm | Время загрузки |
| Operation\_type | Тип операции |
| Updated\_dttm | Время загрузки в **oltp\_cdc\_src\_system** |
|  |  |
| **Система источник: dwh\_stage.account\_data\_src** | |
| Id | Бизнес ключ |
| Status\_nm | Наименование статуса |
| Create\_dttm | Время создания |
| Update\_dttm | Время загрузки |
| Operation\_type | Тип операции |
| Updated\_dttm | Время загрузки в **oltp\_cdc\_src\_system** |
| hash | Хэш данных |
|  | |
| **Система источник: cdc\_from\_stage\_to\_ods.stage\_account\_data\_cdc** | |
| Id | Бизнес ключ |
| Status\_nm | Наименование статуса |
| Create\_dttm | Время создания |
| Update\_dttm | Время загрузки |
| Operation\_type | Тип операции |
| Updated\_dttm | Время загрузки в **oltp\_cdc\_src\_system** |
| hash | Хэш данных |
| Load\_dttm | Время загрузки в слой cdc |
|  |  |
| **Система источник: dwh\_ods.account\_data\_hist** | |
| Id | Бизнес ключ |
| Status\_nm | Наименование статуса |
| Create\_dttm | Время создания |
| Update\_dttm | Время загрузки |
| Operation\_type | Тип операции |
| Updated\_dttm | Время загрузки в **oltp\_cdc\_src\_system** |
| hash | Хэш данных |
| Load\_dttm | Время загрузки в слой cdc |
| Valid\_from\_dttm | Начало версионности |
| Valid\_to\_dttm | Конец версионности |
| Deleted\_flg | Флаг удаленной записи |
| Deleted\_dttm | Время удаления записи |
|  |  |
| **Система источник: dwh\_ods.dim\_date** | |
| Date\_key | Ключ |
| Date\_actual | Полная дата |
| Year | Год |
| Quarter | Квартал |
| Month | Месяц |
| week | Неделя |
| Day\_of\_month | Календарный день |
| Day\_of\_week | День недели |
| Is\_weekday | Флаг рабочего дня |
| Is\_holiday | Флаг выходного дня |
| Fiscal\_year | Финансовый год |
|  |  |

**Описание функций:**

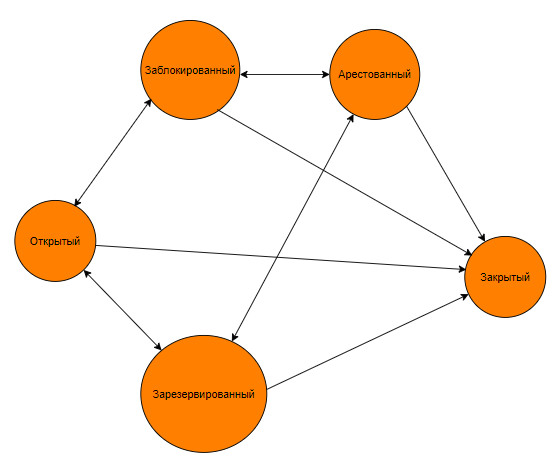
|  |  |  |
| --- | --- | --- |
| **Функция** | **Описание** | **Краткий алгоритм** |
| account\_data\_changes() | Функция для логирования изменений таблицы запускается по триггеру который запускается при изменении таблицы account\_data | Реагирует на функции вставки удаления или замены для каждого условия в запись добавляет значение D, U, I |
| **create\_tasks()** | для создания фейковых задач | Выбирает максимальное значение поля id таблицы account\_date и генерирует несколько новых id |
| **deleted\_existed\_task()** | для случайного удаления фейковых задач | Генерирует число и относительно этого числа удаляет запись из таблицы account\_date |
| **update\_existed\_task()** | для обновления случайно выбранных задач | Генерирует случайный статус отличный от opened и в случайном порядке заменяет его у существующих полей |
| **load\_stage\_from\_oltp\_cdc\_src\_system()** | Загрузка данный в stag слой | Перезаписывает данных из oltp\_cdc\_src\_system в слой stage |
| **stage\_account\_data\_cdc()** | Загрузка в CDC слой | Сравнивает записи слоя stage и слоя dwh\_ods  Добавляет записи с новым id или с новым hash |
| **load\_from\_dwh\_stage\_account\_data\_hist()** | Загрузка в Ods слой  Таблицу account\_data\_hist | Записывает данные из cdc слоя |
| **dim\_date\_uploading()** | Справочник календаря | Относительно данных таблицы account\_data\_hist  создает календарь событий |

**Описание алгоритмов захвата и формирования данных:**

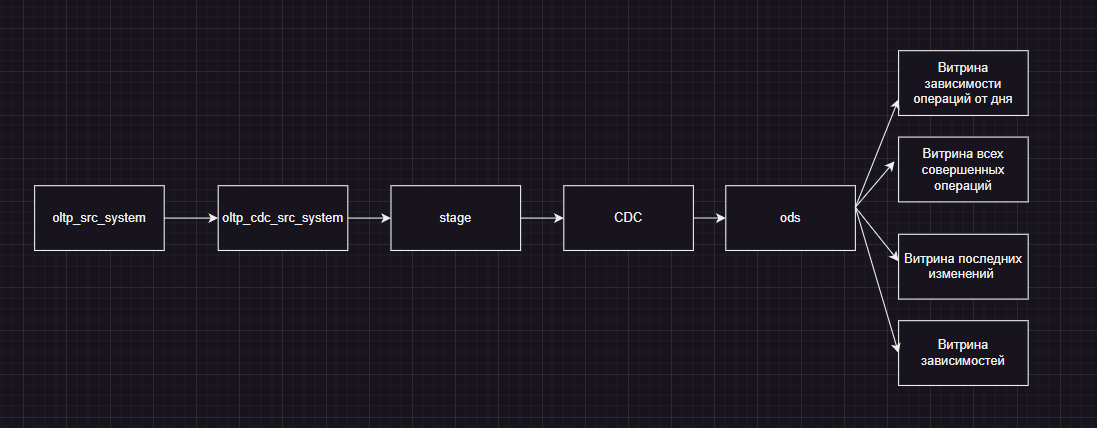
* случайная генерация данных (обновление, удаление)
* загрузка в лог изменений по триггеру после (insert, update, delete)
* дублирование лога изменений в слой stage
* запись разности слоев ods и stage в слой cdc
* запись данных из cdc в ods

слой stage — затирается перед повторной вставки

**Статусная** **модель:**



**Схема решений:**

****

**Описание загрузки данных:**

|  |  |
| --- | --- |
| **Таблицы** | **Функции** |
| **Система источник database: student\_69\_latyshev\_av** | |
| Oltp\_src\_system.account\_data | Oltp\_src\_system.create\_tasks() |
|  | Oltp\_src\_system.update\_task() |
|  | Oltp\_src\_system.deleted\_task() |
|  |  |
| **Слой хранения полного лога изменений Схема oltp\_cdc\_src\_system** | |
| account\_data\_changes | account\_data\_changes() |
|  |  |
| **Первый слой хранения данных Схема dwh\_stage** | |
| account\_data\_src | Load\_stage\_from\_oltp\_cdc\_src\_system() |
|  |  |
| **Слой хранения лога изменений Схема cdc\_from\_stage\_to\_ods** | |
| Stage\_account\_data\_cdc | Load\_stage\_account\_data\_cdc() |
|  |  |
| **Конечный слой хранения истории данных Схема dwh\_stage** | |
| Account\_data\_hist | Load\_to\_dwh\_stage\_account\_data\_hist() |
| Dim\_date | Dim\_date\_uploading() |
|  |  |
| **Витрины данных Схема report** | |
| Account\_data\_recent\_changes | Load\_account\_data\_recent\_changes() |
| Day\_dependence | Load\_Day\_dependence() |
| dependencies | Load\_dependencies() |
| Full\_operations\_report | Load\_ Full\_operations\_report() |

|  |  |
| --- | --- |
| **Система источник: report.account\_data\_recent\_changes** | |
| Id | Бизнес ключ |
| Status\_nm | Наименование статуса |
| Create\_dttm | Время создания |
| Update\_dttm | Время загрузки |
| Operation\_type | Тип операции |
| Updated\_dttm | Время загрузки в **oltp\_cdc\_src\_system** |
| hash | Хэш данных |
| Load\_dttm | Время загрузки в слой cdc |
|  |  |
| **Система источник: report.day\_dependence** | |
| Date\_key | Ключ |
| Date\_actual | Полная дата |
| Is\_holiday | Флаг выходного дня |
| Is\_weekday | Флаг рабочего дня |
| Day\_of\_month | Календарный день |
| Day\_of\_week | День недели |
| Full\_that\_day\_operations | Количество операций, совершенных в этот день |
| insert\_that\_day\_counts | Количество вставок, совершенных в этот день |
| update\_that\_day\_counts | Количество обновлений, совершенных в этот день |
| delete\_that\_day\_counts | Количество удалений, совершенных в этот день |
|  |  |
| **Система источник: report.dependencies** | |
| percentage\_blocked | Процент заблокированных счетов |
| percentage\_arrested | Процент арестованных счетов |
| percentage\_reserved | Процент зарезервированных счетов |
| percentage\_closed | Процент закрытых счетов |
| Reopened\_account | Количесвто переоткрытых счетов |
|  | |
| **Система источник: report.full\_operations\_report** | |
| Insert\_counts | Количество вставок за день |
| Update\_counts | Количество обновлений за день |
| Delete\_counts | Количество удалений за день |

**Описание итоговых таблиц витринного слоя:**

|  |  |  |
| --- | --- | --- |
| **load\_account\_data\_recent\_changes()** | Витрина последних изменений | Записывает данные с valid\_to\_dttm = 2999-12-31 |
| **load\_day\_dependence()** | Витрина зависимости операций от дня | Генерирует витрину зависимости количества операция от дня |
| **load\_dependencies()** | Витрина отношений | Высчитывает в процентах количество различных статусов счетов относительно открытых |
| **load\_full\_opeartions\_report()** | Витрина совершенных операций | Хранит отчет по всем операция за день |

**Описание алгоритмов формирования витринного слоя:**

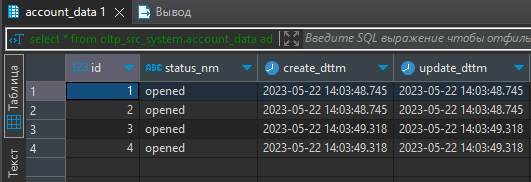
**Описать шаги повторения, которые нужно произвести при добавлении дополнительных таблиц – источников в слое oltp\_src\_system:**

1. Подключить функции генерации данных
2. Добавить триггер для связи со слоем oltp\_cdc\_src\_system
3. Прописать функции загрузки в следующие слои

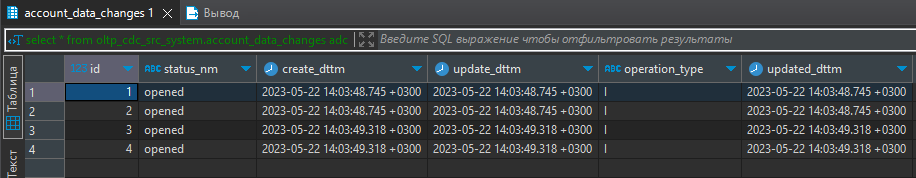
* в слой oltp\_cdc\_src\_system добавить техническое поле **updated\_dttm**
* в слой stage добавить техническое поле **hash**
* в слой cdc добавить техническое поле **load\_dttm**
* в слой dwh\_ods добавить технические поля **valid\_to\_dttm**, **valid\_from\_dttm**

**Тест-кейс:**

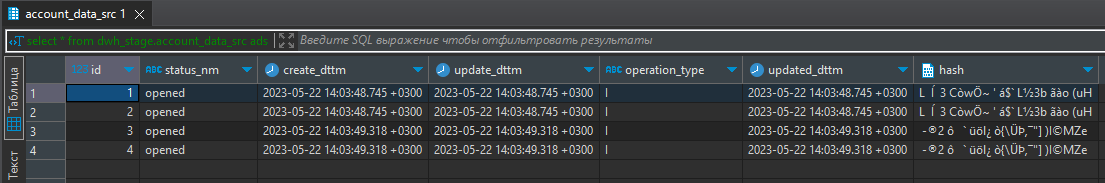
* Генерируем данные в таблице account\_data

****

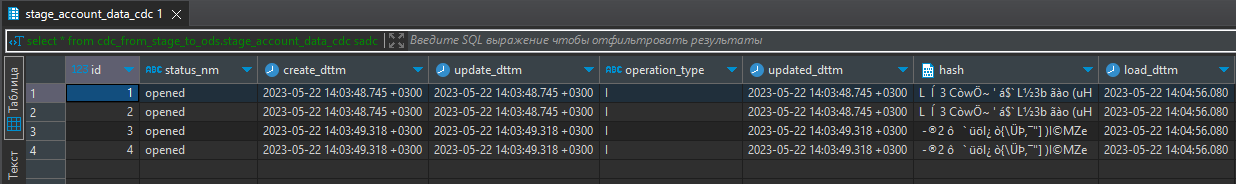
* Данные автоматически появляются в логе изменений account\_data\_changes слоя oltp\_cdc\_src\_system

****

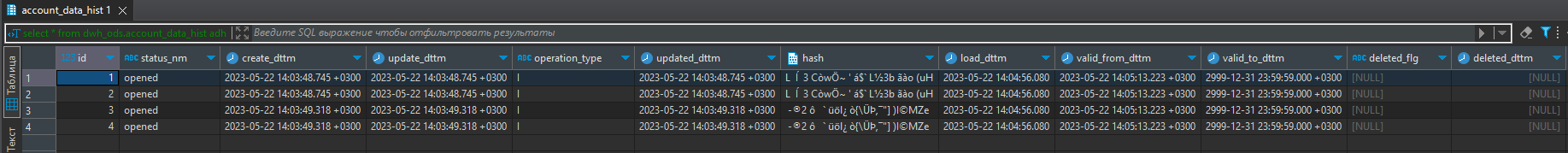
* Загрузка данных в слой stage

****

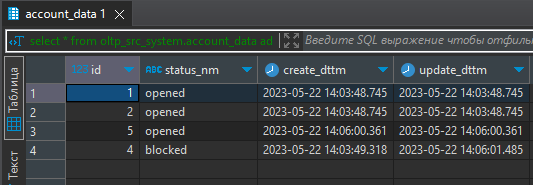
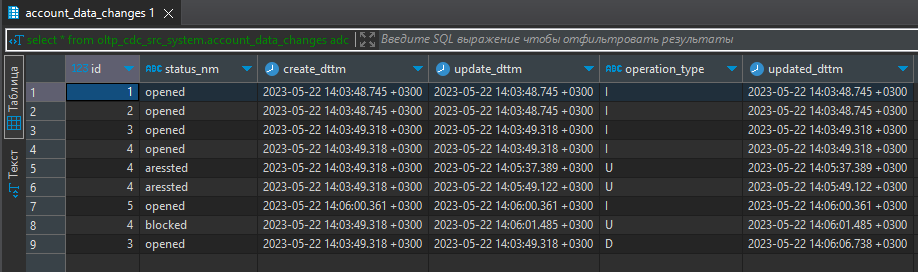
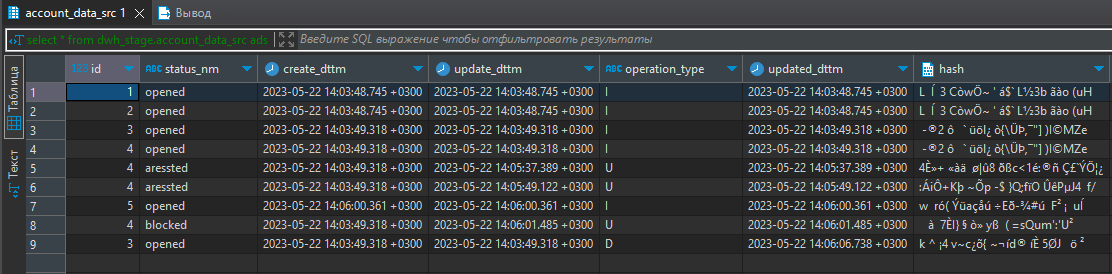
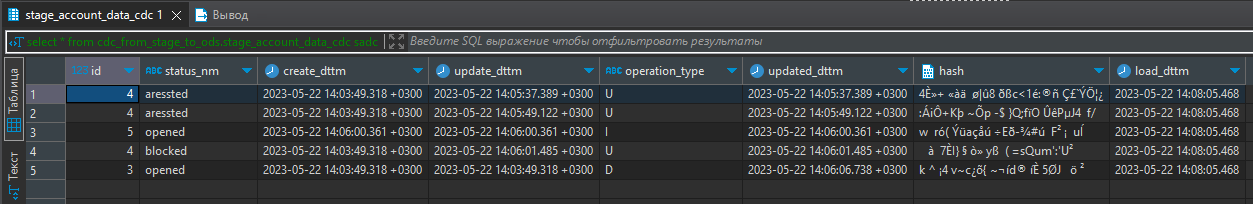
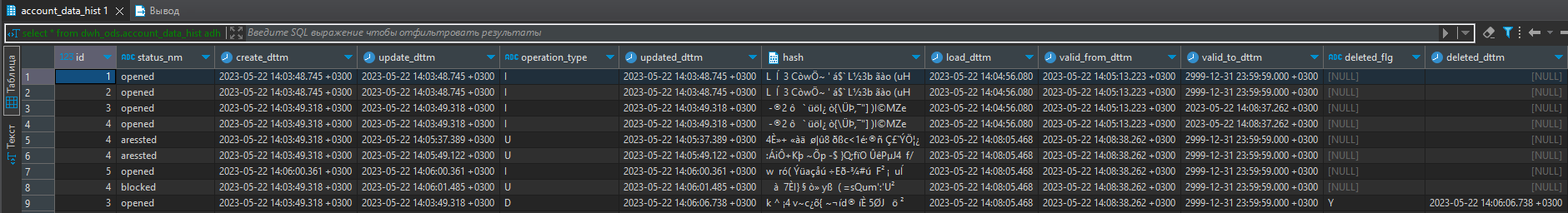
* Загрузка данных в слой cdc

****

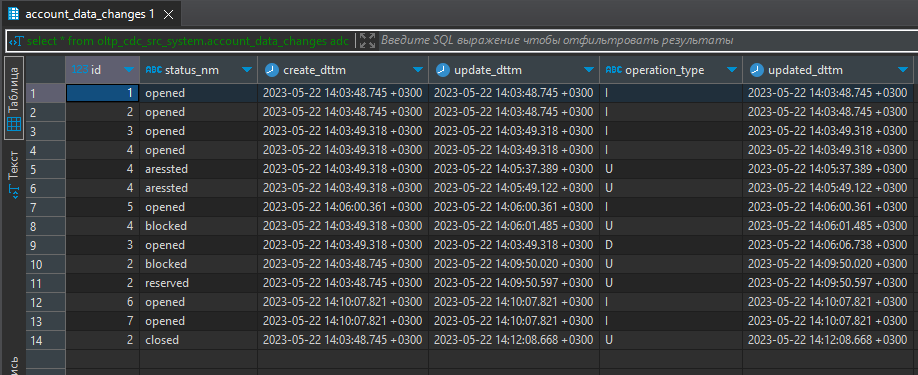
* Запись данных из cdc в ods

****

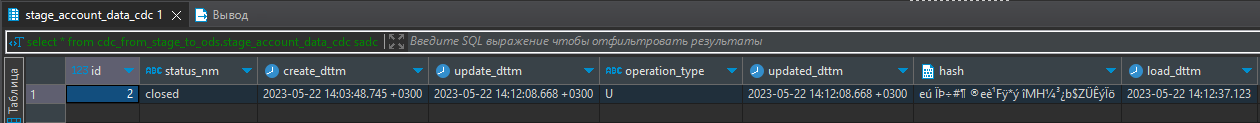
**Далее выполняем операции удаления и обновления данных**

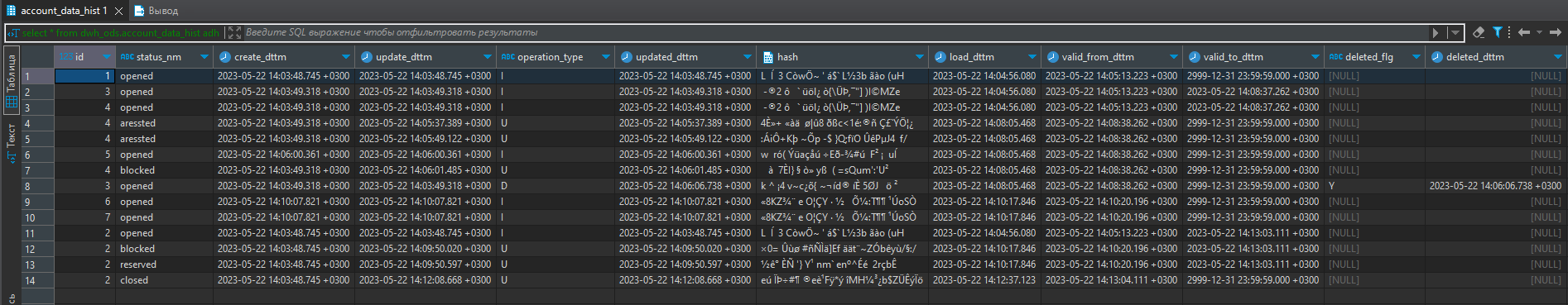
* исходная таблица принимает вид  
  ****
* лог изменений принимает форму  
  ****
* загружаем данные в stage  
  ****
* находим разность в слое cdc  
  ****
* Записываем разность в ods

**Повторяем операцию обновления —>**

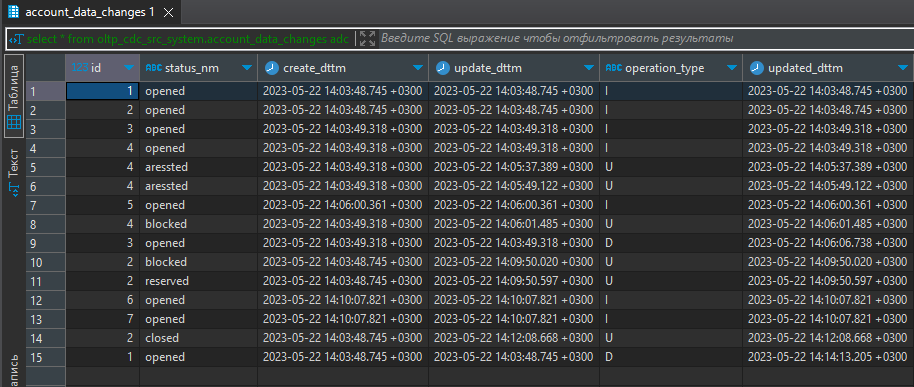
****

**загружаем в stage и находим разность —>**

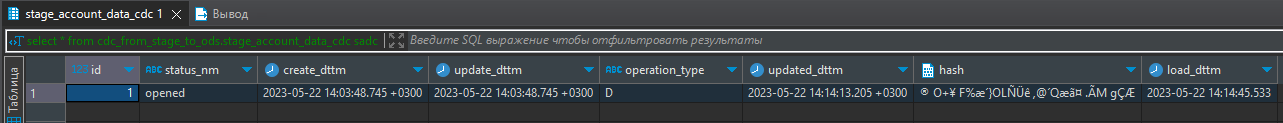
****

**слой ods выглядит следующим образом —>  
**

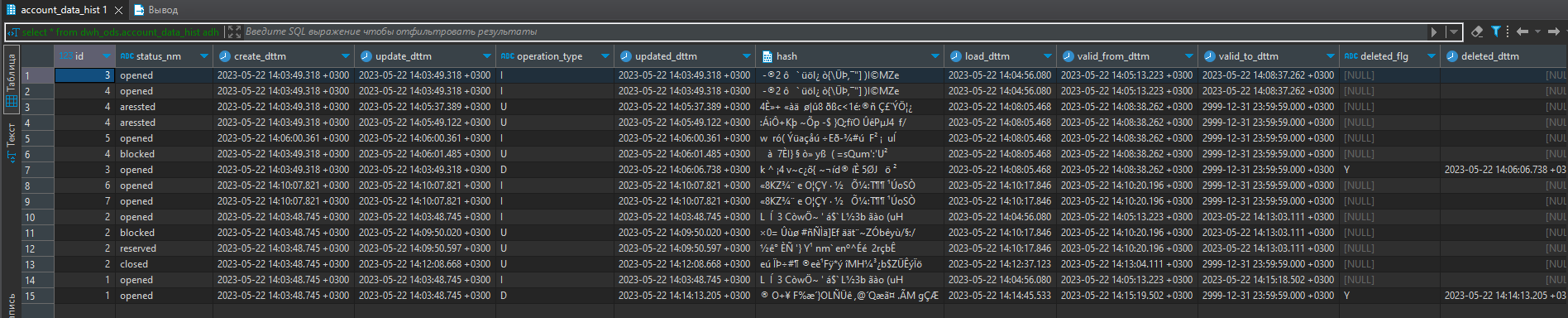
**Повторяем операцию удаления —>**

****

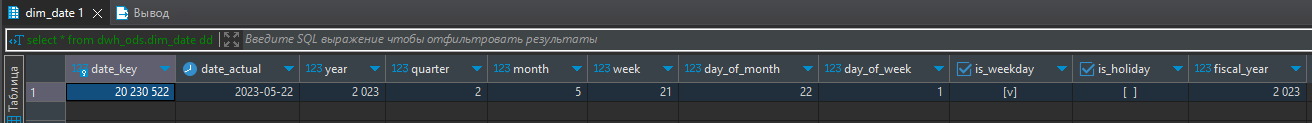
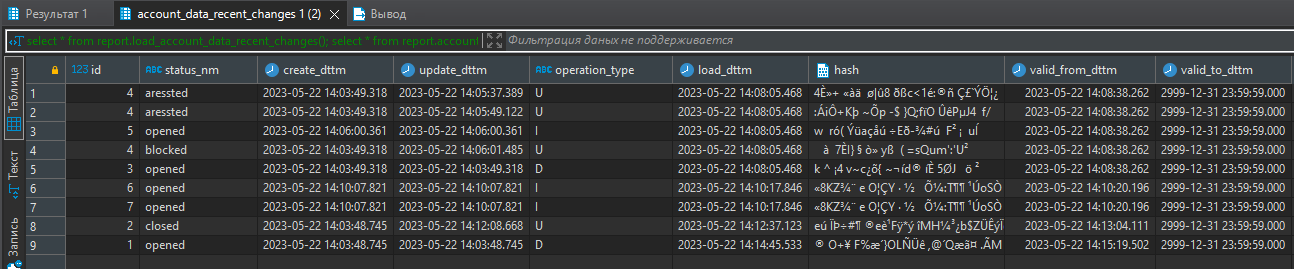
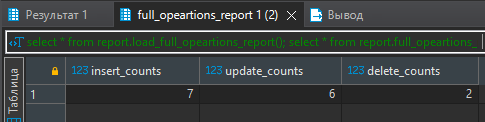
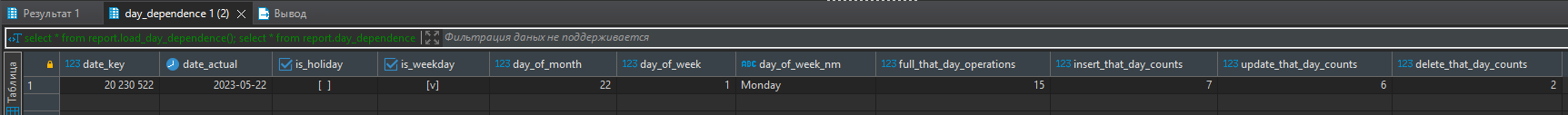
**загружаем в stage и находим разность —>**

****

**слой ods выглядит следующим образом —>**

****

**все отрабатывает успешно.**

* генерация календаря   
  
* генерация витрины последних изменений  
  
* генерация витрин отчета по операциям   
  
* генерация витрин зависимости операций от дня
* генерация витрин отношений  
  