Практическая работа №3

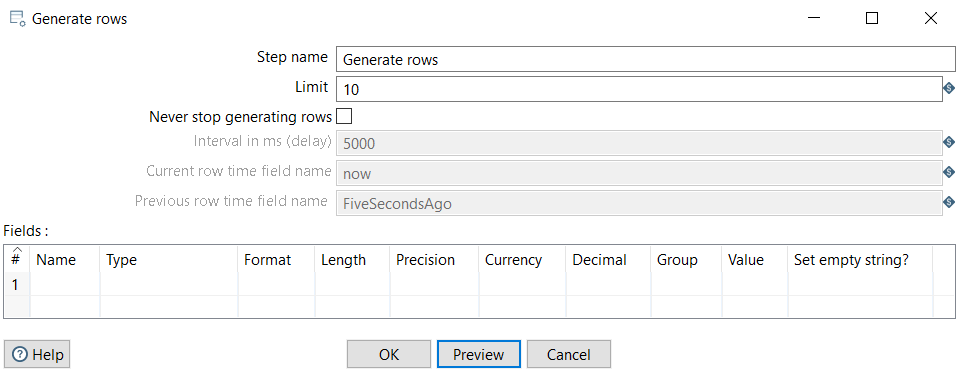
Pentaho ETL подсистемы

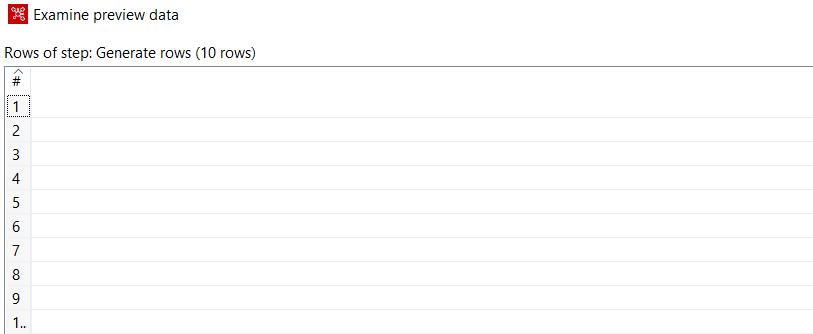
1. Input- добавление файла. Можно добавить csv,json, excel и другие расширения файлов.

Изображение выглядит как текст

Автоматически созданное описание

Generate rows- создание строк





Data Grid- компонент для создания статистических значений

Изображение выглядит как квадрат

Автоматически созданное описание

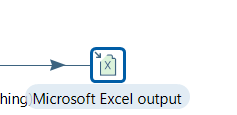
2. OutPut используется для создания описательной документации для одного или нескольких заданий. Он принимает в качестве входных данных список типов имен файлов (преобразование или задание) и генерирует соответствующее состояние файлов документации, содержащих различные сведения о них.

Изображение выглядит как текст

Автоматически созданное описание

Microsoft Excel output

Этот шаг экспортирует данные в файл электронной таблицы Microsoft Excel



Update- обновление данных по ключ поиска. Если ключ не может быть найден, вставляется строка. Если его можно найти и поля для обновления совпадают, ничего не меняется. Если они не все одинаковы, строка в таблице обновляется.



3.Job

Get Rows From Result- Этот компонент возвращает строки, которые ранее были сгенерированы другим преобразованием в задании. Строки были переданы на этот шаг с помощью шага Копирования строк в результат.

Изображение выглядит как текст

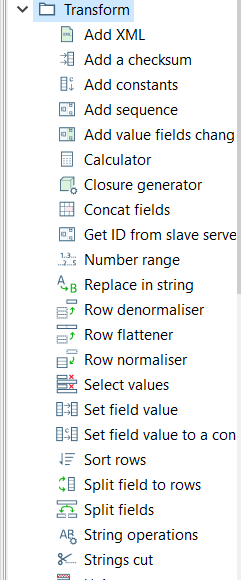
Автоматически созданное описание

Get Variables- позволяет вам получить значение переменной. Этот шаг может возвращать строки или добавлять значения к входным строкам.

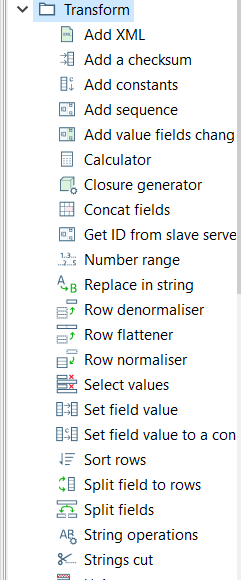
Изображение выглядит как текст

Автоматически созданное описание

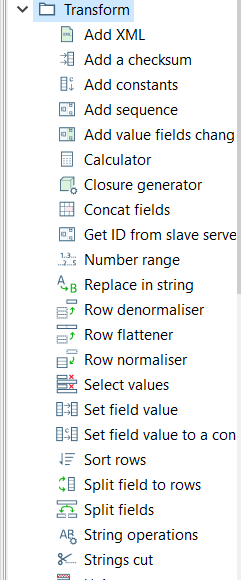
4. Transform- объект, состоящий из шагов, связанных переходами. Эти шаги и переходы создают пути, по которым передаются данные: данные вводятся или создаются на шаге, шаг применяет к ним какое-либо преобразование и, наконец, данные покидают этот шаг. Г



Add a Checksum- вычисляет контрольные суммы для одного или нескольких полей во входном потоке и добавляет их к выходным данным в качестве нового поля.



Calculator-предоставляет вам предопределенные функции, которые могут быть выполнены для значений полей ввода.



5.Utility-обработка данных

Изображение выглядит как текст

Автоматически созданное описание

Change File Encoding-изменение кодировки

Изображение выглядит как текст

Автоматически созданное описание

Email Message- определяет содержимое электронного письма

Изображение выглядит как текст

Автоматически созданное описание

Clone Row- клонирование поля

Изображение выглядит как текст

Автоматически созданное описание

6. Flow

Изображение выглядит как текст

Автоматически созданное описание

Abort -позволяет прервать трансформацию при просмотре входных данных. Его основное применение заключается в обработке ошибок.

Изображение выглядит как текст

Автоматически созданное описание

Dummy - компонент для завершения операции

Изображение выглядит как текст

Автоматически созданное описание

7. Look up

Получение обратной информации, вызвав процедуру БД



Проверяет существует ли файл



8. Statistics – статистика и анализ

Изображение выглядит как текст

Автоматически созданное описание

Analytic Query позволяет выполнять аналитические запросы по отсортированному набору данных.

Изображение выглядит как текст

Автоматически созданное описание

Group by позволяет вычислить значения по определенной группе полей.

Изображение выглядит как текст

Автоматически созданное описание

Talend ETL подсистемы

1.Считывает файл по строкам с простым разделением полей



2. Позваоляет выполнять соединения, фильтрацию столбцов или строк, преобразование и множественные выходные данные



3. Подключает и загружает данные из файла в google bigquery tbigquerysqlrowИзображение выглядит как текст

Автоматически созданное описание

4. Подключить и запустить запрос в google bigquery

Изображение выглядит как текст

Автоматически созданное описание

5-6 Выводит отчёт и создаёт отчет в расширенных форматах, используя iReport от JaspersoftИзображение выглядит как текст

Автоматически созданное описание

7-8. Отправляет неблокирующее триггерное сообщение в tassertcatcher и прослушивает команды tassert или другие команды die в задании. извлекает информацию журнала из сообщения и отправляет ее на стандартный вывод talend

Изображение выглядит как текст

Автоматически созданное описание