Учреждение образования

«Белорусский государственный университет информатики

и радиоэлектроники»

Инженерно-экономический факультет

Кафедра экономической информатики

Средства и технологии анализа и разработки информационных систем

Лабораторная работа № 1

«Консолидация данных»

Выполнил: студент группы 772302, Шмыров И.Д.

Проверил: Кунцевич А.А.

Минск 2020

1 Ход работы:

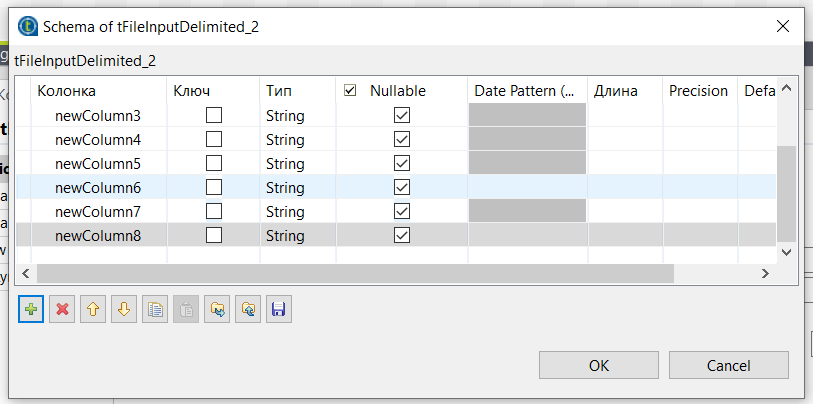


Рисунок 1 – настройка элемента tFileInputDelimited

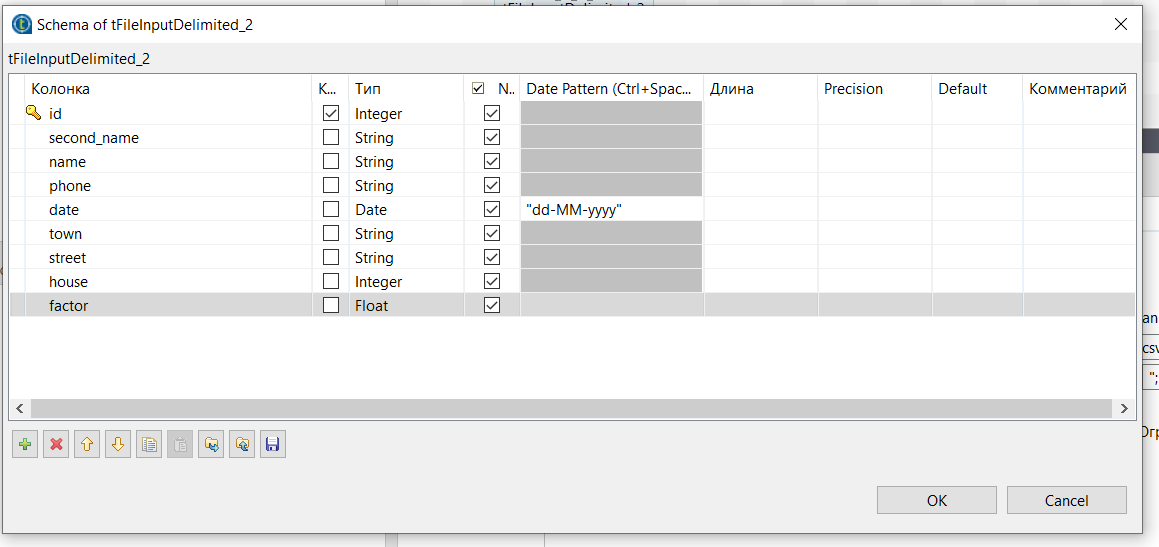


Рисунок 2 – настройка схемы элемента tFileInputDelimited

Схему надо отредактировать в соответствии со своим документом.

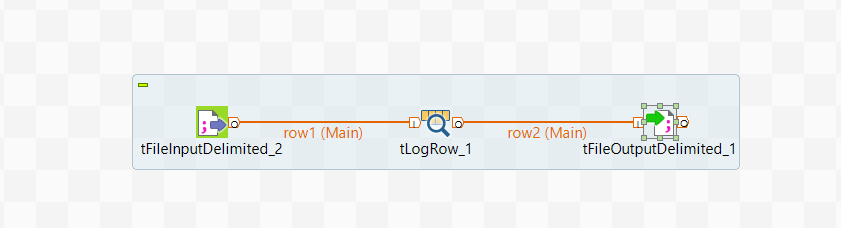


Рисунок 3 – связывание элементов на рабочей области

Теперь переносим на рабочую область еще несколько элементов: tFileOutDelimited и tLogRow. tFileOutDelimited – для вывода файла, tLogRow – для отображения файла. Связываем элементы, щелкнув правой кнопкой мыши по элементу FileInput на рабочей области на пункт Row->Main тянем связь ко второму элементу и повторяем операцию со вторым элементом, протянув связь к третьему.

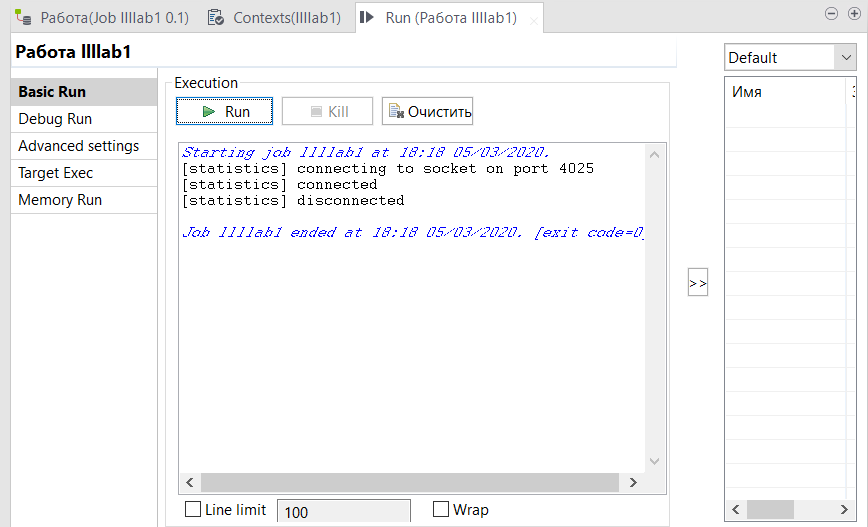


Рисунок 4 – успешный результат выполнения сценария

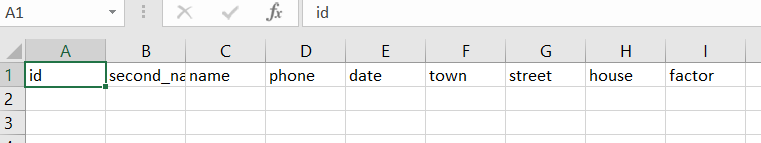


Рисунок 5 – успешно созданный файл вывода

**README**

1. **Что такое Talend OS?**

Talend Open Studio – это программа с открытым исходным кодом которая позволяет производить интеграцию ETL данных в виде дизайнера заданий.

1. **Что такое ETL?**

ETL (Extract, Transform Load*) –э*то процедура извлечения данных из одного источника в систему назначения, которая их преобразует или представляет в другом контексте. Процедура ETL состоит из:

1.Извлечения (Extract) данных из однородных или разнородных источников.

2.Трансформация (Transform) данных и преобразование их в необходимой форме согласно бизнес логике.

3.Загрузка (Load) данных в конечную базу данных или систему назначения.

1. **Консолидация данных.**

Консолидация данных – это процедура получения итоговых данных, при которой данные, расположенные в нескольких различных системах или базах данных, объединяются согласно с определённой функции обработки. В результате консолидации данные обрабатываются и отображаются в одной итоговой таблице, массиве и т.д.