Код всех решений нужно сохранить в отдельном закрытом репозитории в bitbucket.
И выдать доступ пользователю nazarenkr@gmail.com(Назаренко Роман).
Задача 1
Написать простую программу/скрипт(желательно на питоне, но можно и на java) для трансформации "битого" текстового файла (data.tsv, должен быть во вложении к письму) в правильно-отформатированный tsv файл. Результирующий файл должен : иметь в каждой строке одинаковое количество полей (колонок) значения в строке должны быть разделены знаком табуляции "\t" поля которые содержат специальные символы (\t, \r, \n) должны быть экранированы. кодировка UTF-8 (файл источник имеет кодировку UTF-16LE) Главное написать алгоритм для исправления конкретного набора данных, а не делать общее решение для всевозможных кейсов.

Задача 2
Написать ETL-процедуру, которая преобразует файл из TSV-формата, подготовленного в задаче 1, в ORC-файл. Использовать Hadoop MapReduce Framework. Язык Java.
Требования:

1. На вход подается TSV файл
2. Необходимо логировать все ошибки во входящих данных
3. На выходе ORC-файл
4. Написать декларацию create table ... чтобы можно было из ORC-файла сделать external table в Hive

Задача 3
Привести таблицу вида:

id	name
1	name1
2	name2
3	name3

в таблицу вида:

name1	name2	name3
1	2	3

Описание
Задание должно быть выполнено одним sql-запросом без промежуточных / временных таблиц. Использовать либо диалект postgree либо sql server от MS.
Результаты протестировать здесь: https://rextester.com
И сохранить ссылки на результаты в readme.md в репо.