Министерство образования Республики Беларусь

Учреждение образования
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Инженерно-экономический факультет
Кафедра экономической информатики

Отчёт

по дисциплине

«Системы и технологии интеллектуальной обработки данных»

**Лабораторная работа №1**

# КОНСОЛИДАЦИЯ ДАННЫХ В ETL

Выполнила: студентка гр.772302
Ковальчук О. Ю.

Проверил: Кунцевич А. А.

Минск 2020

**Задание**:

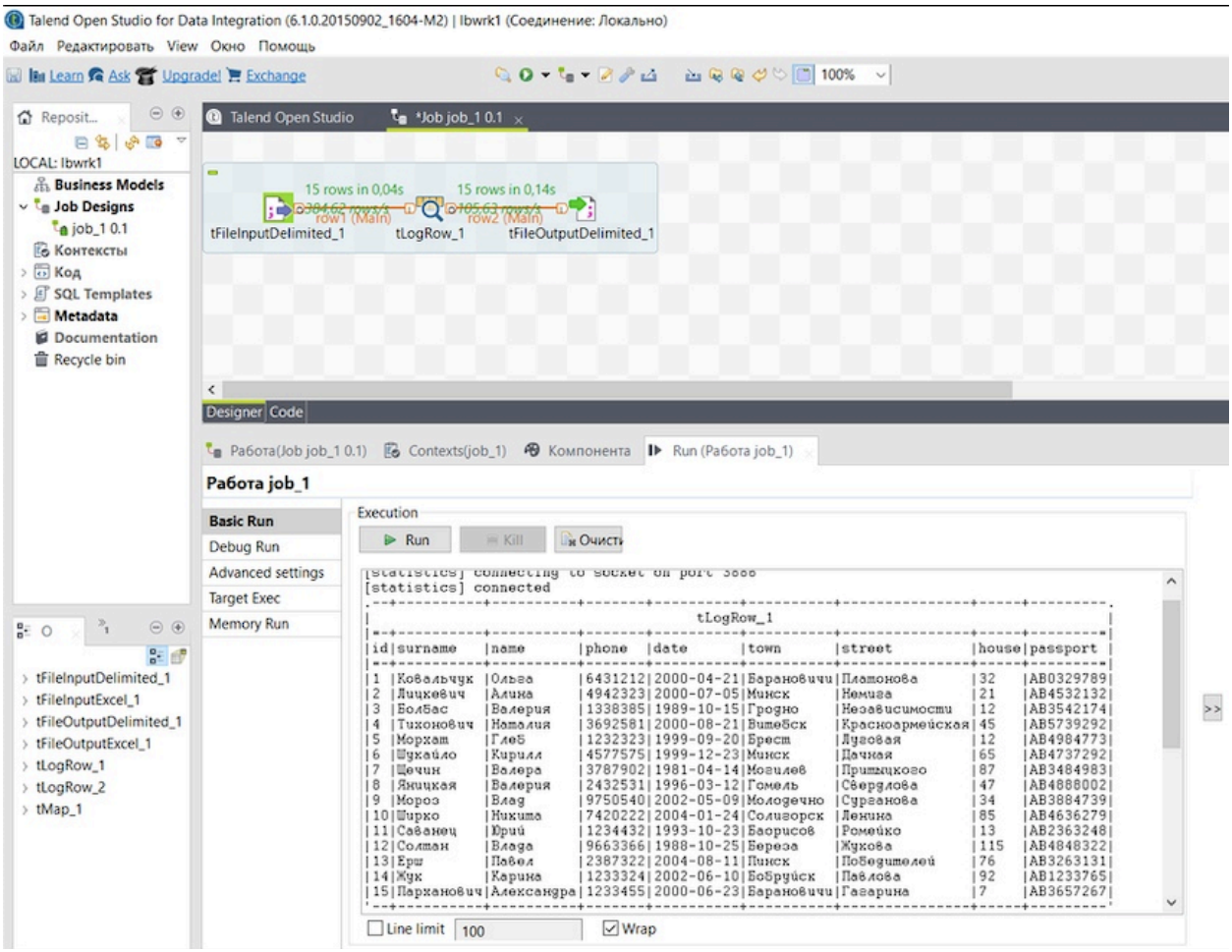
1. Создать CSV файл. Используя Talend, написать сценарий по генерации из этого файла другого CSV файла.
2. Создать Excel файл. Используя Talend, написать сценарий по генерации из этого файла другого Excel файла.
3. Создать Excel файл. Используя элемент tMap Talend(a), написать сценарий по генерации другого Excel файла с заданной структурой полей. Например, некоторые поля одного файла объединены в одно поле другого файла. Таких объединений должно быть 2.

**Выполнение**:

– Создадим исходный файл in.csv:

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | id | surname | name | phone | date | town | street | house | passport |
| 2 | 1 | Ковальчук | Ольга | 6431212 | 2000-04-21 | Баранови | Платонова | 32 | AB0329789 |
| 3 | 2 | Лицкевич | Алина | 4942323 | 2000-07-05 | Минск | Немига | 21 | AB4532132 |
| 4 | 3 | Болбас | Валерия | 1338385 | 1989-10-15 | Гродно | Независимости | 12 | AB3542174 |
| 5 | 4 | Тихонович | Наталия | 3692581 | 2000-08-21 | Витебск | Красноармейская | 45 | AB5739292 |
| 6 | 5 | Морхат | Глеб | 1232323 | 1999-09-20 | Брест | Луговая | 12 | AB4984773 |
| 7 | 6 | Шукайло | Кирилл | 4577575 | 1999-12-23 | Минск | Дачная | 65 | AB4737292 |
| 8 | 7 | Щечин | Валера | 3787902 | 1981-04-14 | Могилев | Притыцкого | 87 | AB3484983 |
| 9 | 8 | Яницкая | Валерия | 2432531 | 1996-03-12 | Гомель | Свердлова | 47 | AB4888002 |
| 10 | 9 | Мороз | Влад | 9750540 | 2002-05-09 | Молодечн | Сурганова | 34 | AB3884739 |
| 11 | 10 | Ширко | Никита | 7420222 | 2004-01-24 | Солигорск | Ленина | 85 | AB4636279 |
| 12 | 11 | Саванец | Юрий | 1234432 | 1993-10-23 | Баорисов | Ромейко | 13 | AB2363248 |
| 13 | 12 | Солтан | Влада | 9663366 | 1988-10-25 | Береза | Жукова | 115 | AB4848322 |
| 14 | 13 | Ерш | Павел | 2387322 | 2004-08-11 | Пинск | Победителей | 76 | AB3263131 |
| 15 | 14 | Жук | Карина | 1233324 | 2002-06-10 | Бобруйск | Павлова | 92 | AB1233765 |
| 16 | 15 | Парханович | Александра | 1233455 | 2000-06-23 | Баранови | Гагарина | 7 | AB3657267 |

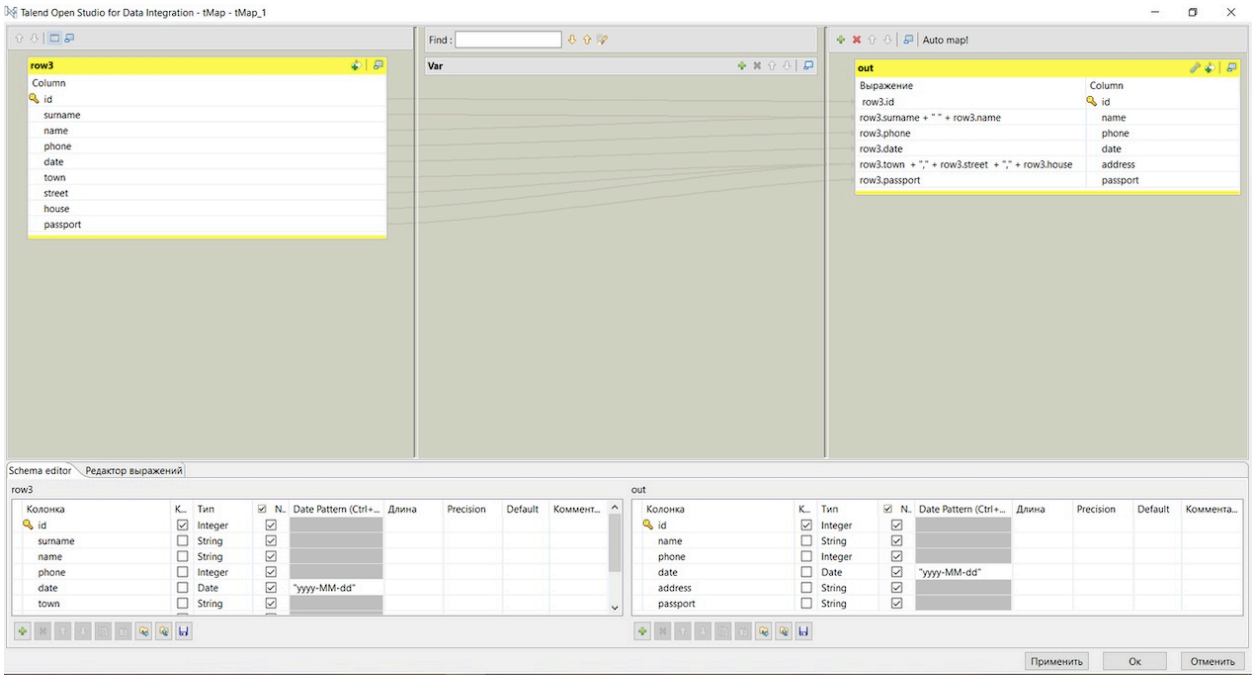
– Создадим сценарий преобразования in.csv в out.csv и выполним работу:



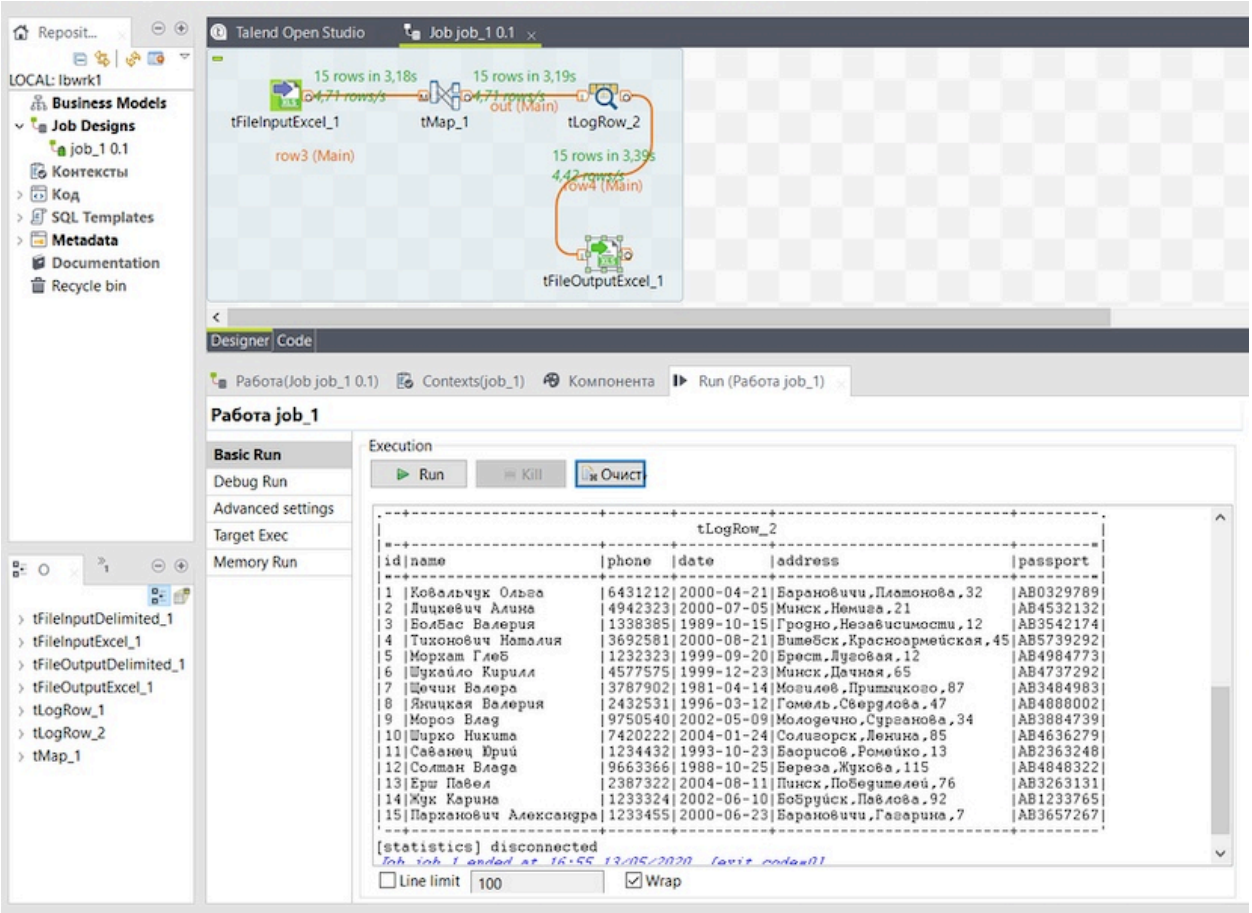
– Создадим исходный файл in.xls:

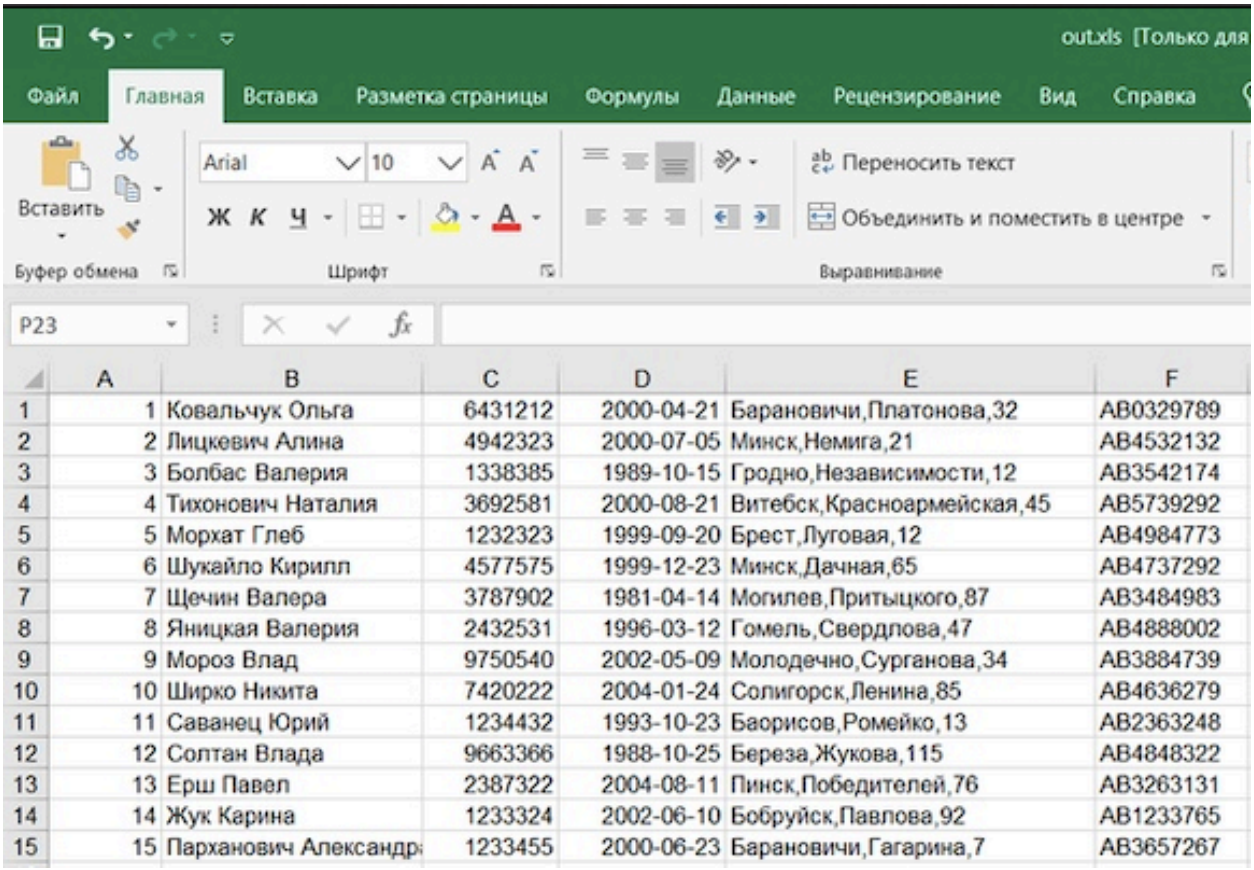
| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | id | surname | name | phone | date | town | street | house | passport |
| 2 | 1 | Ковальчук | Ольга | 6431212 | 2000-04-21 | Баранови | Платонова | 32 | AB0329789 |
| 3 | 2 | Лицкевич | Алина | 4942323 | 2000-07-05 | Минск | Немига | 21 | AB4532132 |
| 4 | 3 | Болбас | Валерия | 1338385 | 1989-10-15 | Гродно | Независимости | 12 | AB3542174 |
| 5 | 4 | Тихонович | Наталия | 3692581 | 2000-08-21 | Витебск | Красноармейская | 45 | AB5739292 |
| 6 | 5 | Морхат | Глеб | 1232323 | 1999-09-20 | Брест | Луговая | 12 | AB4984773 |
| 7 | 6 | Шукайло | Кирилл | 4577575 | 1999-12-23 | Минск | Дачная | 65 | AB4737292 |
| 8 | 7 | Щечин | Валера | 3787902 | 1981-04-14 | Могилев | Притыцкого | 87 | AB3484983 |
| 9 | 8 | Яницкая | Валерия | 2432531 | 1996-03-12 | Гомель | Свердлова | 47 | AB4888002 |
| 10 | 9 | Мороз | Влад | 9750540 | 2002-05-09 | Молодечн | Сурганова | 34 | AB3884739 |
| 11 | 10 | Ширко | Никита | 7420222 | 2004-01-24 | Солигорск | Ленина | 85 | AB4636279 |
| 12 | 11 | Саванец | Юрий | 1234432 | 1993-10-23 | Баорисов | Ромейко | 13 | AB2363248 |
| 13 | 12 | Солтан | Влада | 9663366 | 1988-10-25 | Береза | Жукова | 115 | AB4848322 |
| 14 | 13 | Ерш | Павел | 2387322 | 2004-08-11 | Пинск | Победителей | 76 | AB3263131 |
| 15 | 14 | Жук | Карина | 1233324 | 2002-06-10 | Бобруйск | Павлова | 92 | AB1233765 |
| 16 | 15 | Парханович | Александра | 1233455 | 2000-06-23 | Баранови | Гагарина | 7 | AB3657267 |

– Создадим сценарий преобразования in.xls в out.xls и настроим свойства элемента map:



– Выполним работу:




| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | 1 | Ковальчук Ольга | 6431212 | 2000-04-21 | Барановичи,Платонова,32 | AB0329789 |
| 2 | 2 | Лицкевич Алина | 4942323 | 2000-07-05 | Минск,Немига,21 | AB4532132 |
| 3 | 3 | Болбас Валерия | 1338385 | 1989-10-15 | Гродно,Независимости,12 | AB3542174 |
| 4 | 4 | Тихонович Наталия | 3692581 | 2000-08-21 | Витебск,Красноармейская,45 | AB5739292 |
| 5 | 5 | Морхат Глеб | 1232323 | 1999-09-20 | Брест,Луговая,12 | AB4984773 |
| 6 | 6 | Шукайло Кирилл | 4577575 | 1999-12-23 | Минск,Дачная,65 | AB4737292 |
| 7 | 7 | Щечин Валера | 3787902 | 1981-04-14 | Могилев,Притыцкого,87 | AB3484983 |
| 8 | 8 | Яницкая Валерия | 2432531 | 1996-03-12 | Гомель,Свердлова,47 | AB4888002 |
| 9 | 9 | Мороз Влад | 9750540 | 2002-05-09 | Молодечно,Сурганова,34 | AB3884739 |
| 10 | 10 | Ширко Никита | 7420222 | 2004-01-24 | Солигорск,Ленина,85 | AB4636279 |
| 11 | 11 | Саванец Юрий | 1234432 | 1993-10-23 | Баорисов,Ромейко,13 | AB2363248 |
| 12 | 12 | Солтан Влада | 9663366 | 1988-10-25 | Береза,Жукова,115 | AB4848322 |
| 13 | 13 | Ерш Павел | 2387322 | 2004-08-11 | Пинск,Победителей,76 | AB3263131 |
| 14 | 14 | Жук Карина | 1233324 | 2002-06-10 | Бобруйск,Павлова,92 | AB1233765 |
| 15 | 15 | Парханович Александр | 1233455 | 2000-06-23 | Барановичи,Гагарина,7 | AB3657267 |