DreamCandies File Tool

Feature	CR1888_DRC
Customer	DreamCandies
Release	1.0
Keywords	<keyword>
Author	Daring Star Ltd
Reviewed	No
Document Name	DreamCandies.docx
Document Version	1.0
Disclaimer	The master copy of this document is stored under configuration management in electronic format. Any hardcopies and softcopies not under configuration management are for information only and might be outdated.
Copyright	The copyright in this work is vested in Daring Star and the information contained herein is confidential. This work (either in whole or in part) must not be modified, reproduced, disclosed or disseminated to others or used for purposes other than that for which it is supplied, without the prior written permission of Daring Star. If this work or any part hereof is furnished to a third party by virtue of a contract with that party, use of this work by such party shall be governed by the express contractual terms between the Daring Star company which is a party to that contract and the said party. Copyright of Daring Star, 2010. All rights reserved.

Customer Approval	
Date	
Name	
Signature	

Change History

Version	Date	Author	Description
0.1	27-Oct-2016	Daring Star	Initial Draft

Table of Contents

1 Business Requirements

DreamCandies, a company selling candies has engaged Daring Star to migrate their current billing system to a completely new platform. In order to ensure the final migrated system is bugs free, Daring Star system integration team has decided to implement a test automation process to continuously verify correctness of their implementation throughout the entire project lifecycle.

On a daily basis, the team will receive three FULL database extraction files from DreamCandies. The team will pre-select 1000 customers as test sample to ensure a speedy test automation process.

Based on this input, we should extract all relevant entries from the three full extraction files to produce three smaller files containing all the pre-selected customers' data. The test automation process will then ingest these smaller set of files to perform the test automation execution quickly.

1.1 Extraction File Format

On daily basis, the customer (DreamCandies) will deliver to Daring Star System Integration Team three full extraction files, all extraction files will have the following files specification:

- ASCII file (UNIX format)
- First line of the file is the heading
- Each field is separated by comma character ","
- All fields are double quoted regardless of data type (e.g. "Olivier")
- Thousand separator should not be use in the field with FLOAT data type
- Decimal point "." will be used in field with FLOAT data type as decimal mark

CUSTOMER.CSV (contains approximately 500k customer)

Columns	Data Type
CUSTOMER_CODE	CHAR(30)
FIRSTNAME	CHAR(100)
LASTNAME	CHAR(100)

INVOICE.CSV (contains approximately 1 million invoices)

Columns	Data Type
CUSTOMER_CODE	CHAR(30)
INVOICE_CODE	CHAR(30)
AMOUNT	FLOAT
DATE	DATE

INVOICE_ITEM.CSV (contains approximately 5 million invoice items)

Columns	Data Type
INVOICE_CODE	CHAR(30)
ITEM_CODE	CHAR(30)
AMOUNT	FLOAT
QUANTITY	INTEGER

Example:

The following example show two customers, Maria (CUSTOMER_CODE= CUST0000010231) with two invoices for Jan and Feb 2016, whereas George (CUSTOMER_CODE= CUST0000010235) has only one invoice for Jan 2016.

CUSTOMER.CSV

```
"CUSTOMER_CODE","FIRSTNAME","LASTNAME"
"CUST0000010231","Maria","Alba"
"CUST0000010235","George","Lucas"
```

INVOICE.CSV

```
"CUSTOMER_CODE","INVOICE_CODE","AMOUNT","DATE"
"CUST0000010231","IN0000001","105.50","01-Jan-2016"
"CUST0000010235","IN0000002","186.53","01-Jan-2016"
"CUST0000010231","IN0000003","114.14","01-Feb-2016"
```

INVOICE_ITEM.CSV

```
"INVOICE_CODE","ITEM_CODE","AMOUNT","QUANTITY"
"IN0000001","MEIJI","75.60","100"
"IN0000001","POCKY","10.40","250"
"IN0000001","PUCCHO","19.50","40"
"IN0000002","MEIJI","113.40","150"
"IN0000002","PUCCHO","73.13","150"
"IN0000003","POCKY","16.64","400"
"IN0000003","PUCCHO","97.50","200"
```

1.2 Customer Sample File

System Integration Team will pre-select 1000 customer into an ASCII file, this file be the input of the new tool to be developed. The files will have the following files specification:

- ASCII file (UNIX format)
- First line of the file is the heading

CUSTOMER_SAMPLE.CSV

Columns	Data Type
CUSTOMER_CODE	CHAR(30)

Example:

```
"CUSTOMER_CODE"
"CUST0000010231"
"CUST0000010235"
```

2 New Business Process

Develop a tool with any preferred programming language of choice (e.g. Java, Python, C/C++).

The input of the tool are the customer sample input file (see chapter 1.2) and the full set of Extraction Files (see chapter 1.1).

The expected output from the tool will be a subset of the full extraction files (CUSTOMER.CSV, INVOICE.CSV, INVOICE_ITEM.CSV) only including data for the customers specified within CUSTOMER_SAMPLE.CSV file.

The output should be distributed in three files reflecting the original full extraction file-set structure (see chapter 1.1).