

Retrieve Data From A Csv File

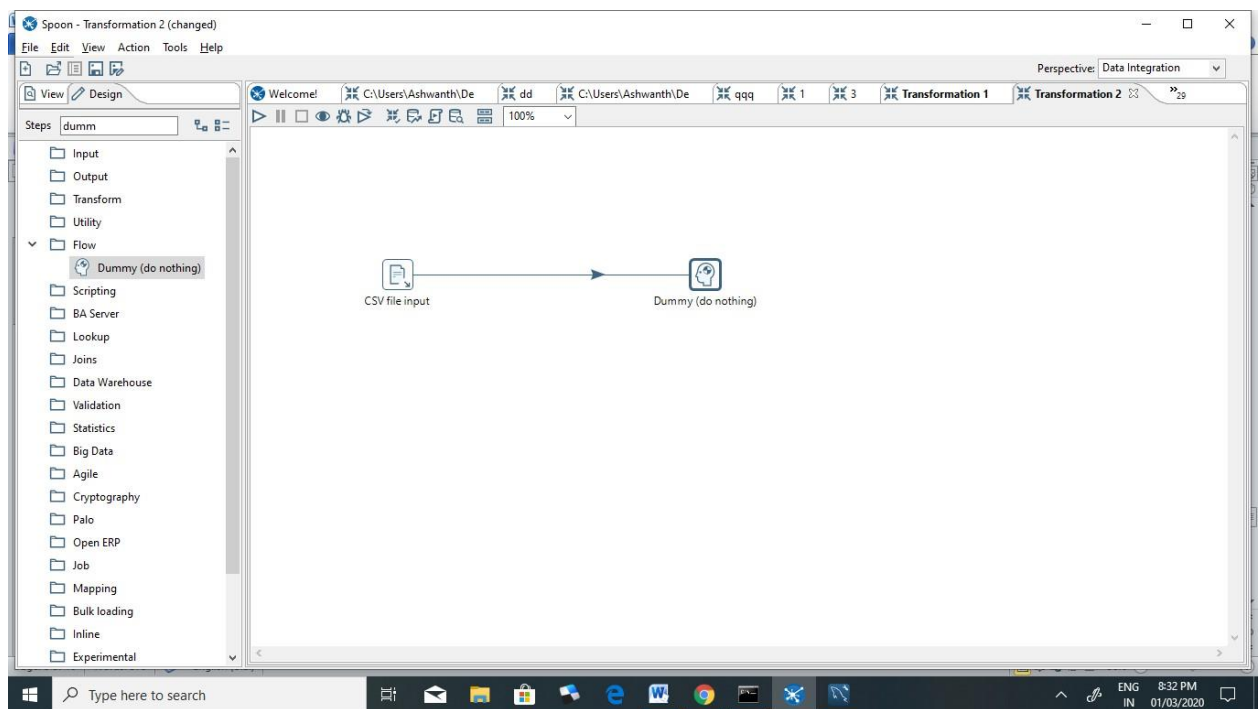
Aim:

To retrieve data from a CSV file and click preview rows to verify the file is being read correctly. (use: sales_data.csv).

Algorithm:

- 1.Start the process.
- 2.Create a Sales database using Excel.
- 3.Save the file in CSV format.
- 4.Open Pentaho Spoon tool for data integration.
- 5.Choose text input and text output for data retrieval.
- 6.Save the file.
- 7.Stop the process.

Screen Shot :



Spoon - Transformation 2 (changed)

File Edit View Action Tools Help

View Design

Steps: dumm

Input

Output

Transform

Utility

Flow

Dummy (do nothing)

Scripting

BA Server

Lookup

Joins

Data Warehouse

Validation

Statistics

Big Data

Agile

Cryptography

Palo

Open ERP

Job

Mapping

Bulk loading

Inline

Experimental

CSV Input

Step name: CSV file input

Filename: E:\BI-II for MBA\Assignment\unit-2\sales_data.csv

Delimiter: ,

Enclosure: "

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result: ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

File encoding:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group
1	ORDERNUMBER	Integer	#	15	0	\$.	
2	QUANTITYORDERED	Integer	#	15	0	\$.	
3	PRICEEACH	Number	##	5	2	\$.	
4	ORDERLINENUMBER	Integer	#	15	0	\$.	
5	SALES	Number	##	7	2	\$.	
6	ORDERDATE	Date	MM/dd/yyyy			\$.	
7	STATUS	String		10		\$.	
8	QTR_ID	Integer	#	15	0	\$.	
9	MONTH_ID	Integer	#	15	0	\$.	
10	YEAR_ID	Integer	#	15	0	\$.	
11	PRODUCTLINE	String		12		\$.	
12	MSRP	Integer	#	15	0	\$.	
13	PRODUCTCODE	String		8		\$.	
14	CUSTOMERNAME	String		30		\$.	
15	PHONE	String		16		\$.	
16	ADDRESSLINE1	String		40		\$.	
17	ADDRESSLINE2	String		9		\$.	

OK Get Fields Preview Cancel

Perspective: Data Integration

Transformation 2

ENG 8:32 PM 01/03/2020

Spoon - Transformation 2 (changed)

File Edit View Action Tools Help

View Design

Examine preview data

Rows of step: CSV file input (1000 rows)

#	ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	ORDERDATE	STATUS	QTR_ID	MONTH_ID	YEAR_ID	PRODUCTLINE	MSRP	PRODUCTCODE	CUSTOMERNAME
1	10107	30	95.7	2	2871	02/24/2003	Shipped	1	5	2003	Motorcycles	95	S10_1678	Lan
2	10121	34	81.4	5	2765.9	05/07/2003	Shipped	2	7	2003	Motorcycles	95	S10_1678	Reir
3	10134	41	94.7	2	3884.3	07/01/2003	Shipped	3	8	2003	Motorcycles	95	S10_1678	Lyo
4	10145	45	83.3	6	3746.7	08/25/2003	Shipped	3	10	2003	Motorcycles	95	S10_1678	Toy
5	10159	49	100	14	5205.3	10/10/2003	Shipped	4	10	2003	Motorcycles	95	S10_1678	Cor
6	10168	36	96.7	1	3479.8	10/28/2003	Shipped	4	10	2003	Motorcycles	95	S10_1678	Tecl
7	10180	29	86.1	9	2497.8	11/11/2003	Shipped	4	11	2003	Motorcycles	95	S10_1678	Dae
8	10188	48	100	1	5512.3	11/18/2003	Shipped	4	11	2003	Motorcycles	95	S10_1678	Her
9	10201	22	98.6	2	2168.5	12/01/2003	Shipped	4	12	2003	Motorcycles	95	S10_1678	Min
10	10211	41	100	14	4708.4	01/15/2004	Shipped	1	1	2004	Motorcycles	95	S10_1678	Aut
11	10223	37	100	1	3965.7	02/20/2004	Shipped	1	2	2004	Motorcycles	95	S10_1678	Aus
12	10237	23	100	7	2333.1	04/05/2004	Shipped	2	4	2004	Motorcycles	95	S10_1678	Vita
13	10251	28	100	2	3188.6	05/18/2004	Shipped	2	5	2004	Motorcycles	95	S10_1678	Teku
14	10263	34	100	2	3676.8	06/28/2004	Shipped	2	6	2004	Motorcycles	95	S10_1678	Gift
15	10275	45	92.8	1	4177.4	07/23/2004	Shipped	3	7	2004	Motorcycles	95	S10_1678	La F
16	10285	36	100	6	4099.7	08/27/2004	Shipped	3	8	2004	Motorcycles	95	S10_1678	Mar
17	10299	23	100	9	2597.4	09/30/2004	Shipped	3	9	2004	Motorcycles	95	S10_1678	Toy
18	10309	41	100	5	4394.4	10/15/2004	Shipped	4	10	2004	Motorcycles	95	S10_1678	Baa
19	10319	46	94.7	1	4758	11/02/2004	Shipped	4	11	2004	Motorcycles	95	S10_1678	Pha

Close Show Log

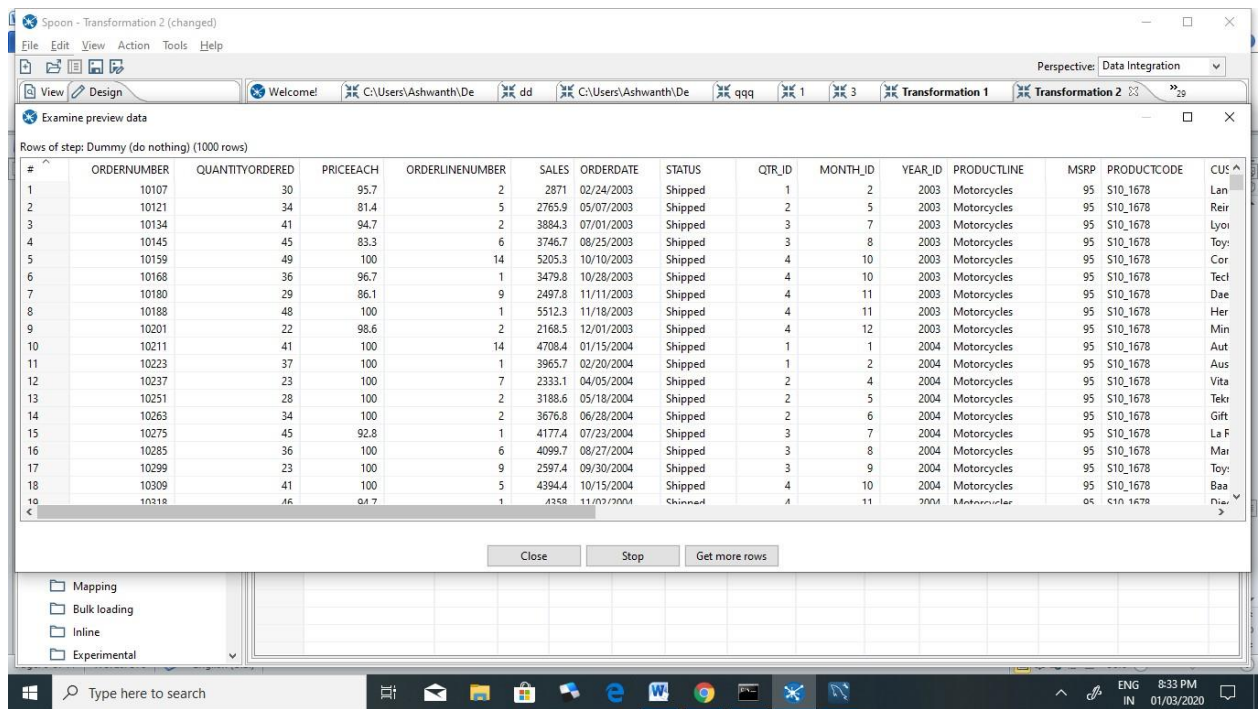
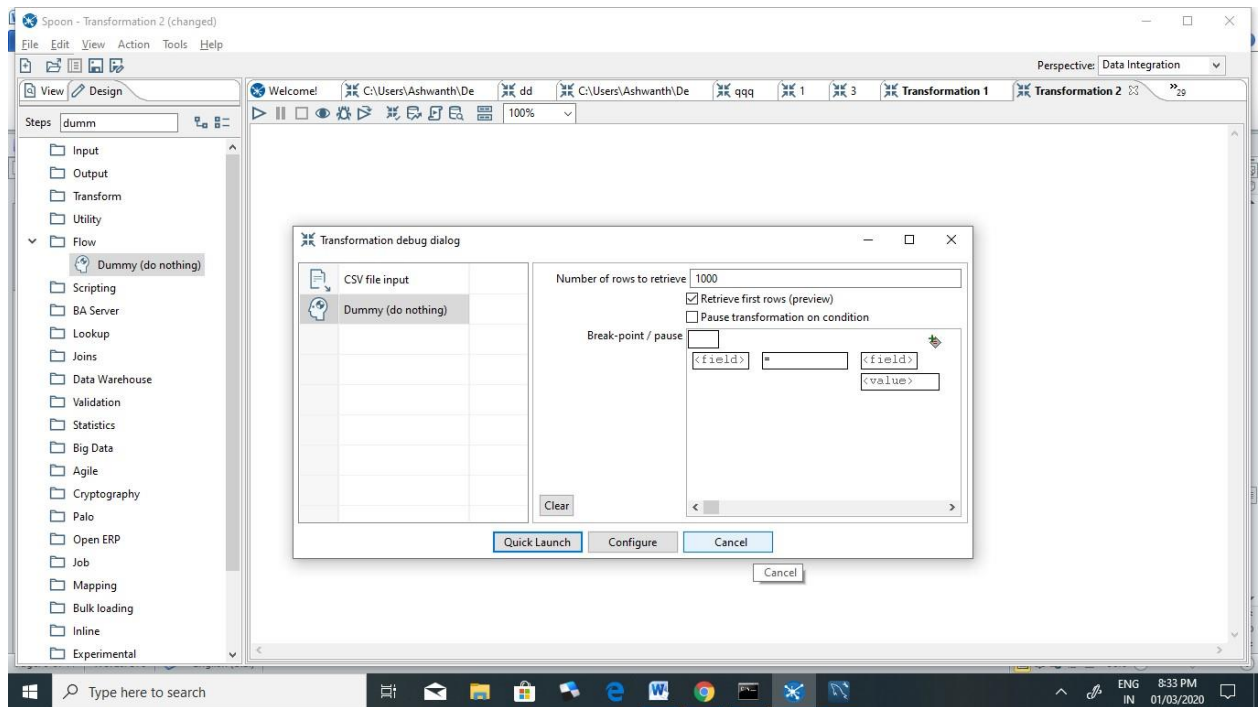
Mapping

Bulk loading

Inline

Experimental

ENG 8:33 PM 01/03/2020



Result :

Thus the expected output is achieved.