

EXCEL INPUT SPREADSHEET PROCESSING:

1. Microsoft Excel Input:

Step name: Excel spreadsheet

Spread sheet type (engine): Excel 97-2003 XLS (XtL)

File or directory: [Empty] Add Browse...

Regular Expression: [Empty]

Exclude Regular Expression: [Empty]

Selected files:

#	File/Directory	Wildcard (RegExp)	Exclude wi
1	C:\Users\Sanreshv\Desktop\coursea\Inventory\ExcelInventory.xls		

Accept filenames from previous steps

Accept filenames from previous step ☐

Step to read filenames from: [Empty]

Field in the input to use as filename: [Empty]

Show filename(s)...

OK Preview rows Cancel

2	03	08/13/2012	17	3	07	22.33	JPY	20.00
3	05	08/20/2012	09	5	01	33.11	USD	30.00
4	07	08/27/2012	07	1	03	33.22	JPY	40.00

Step name: Excel spreadsheet

List of sheets to read

#	Sheet name	Start row	Start column
1	Sheet1	0	0

Get sheetname(s)...

OK Preview rows Cancel

2	03	08/13/2012	17	3	07	22.33	JPY	20.00
---	----	------------	----	---	----	-------	-----	-------

Microsoft Excel input

Step name: Excel spreadsheet

Files | Sheets | Content | Error Handling | **Fields** | Additional output fields

#	Name	Type	Length	Precision	Trim type	Repeat	Format	Currency	Decimal	Grouping
1	BranchPlantKey	Integer	2	-1	none	N				
2	Date	Date	-1	-1	none	N	MM/dd/yyyy			
3	ItemMasterKey	Integer	2	-1	none	N				
4	TransTypeKey	Integer	1	-1	none	N				
5	CustVendorKey	Integer	2	-1	none	N				
6	UnitCost	Number	4	2	none	N				
7	Currency	String	3	-1	none	N				
8	Quantity	Number	2	-1	none	N				

Get fields from header row...

Help OK Preview rows Cancel

2	03	08/13/2012		17	3	07	22.33	JPY	20.00	
3	05	08/20/2012		09	5	01	33.11	USD	30.00	
4	07	08/27/2012		07	1	03	33.22	JPY	40.00	
5	09	09/03/2012		15	1	11	22.01	GBP	50.00	
6	11	09/10/2012		05	3	15	11.03	USD	20.00	
--	--	-----		--	--	--	--	---	--	--

2. Reject a record if any field is null

Filter rows

Step name: Filter rows

Send 'true' data to step: _____

Send 'false' data to step: _____

The condition:

To edit a subcondition, simply click on it

```

(
    BranchPlantKey IS NOT NULL
)
AND
(
    Date IS NOT NULL
)
AND
    ItemMasterKey IS NOT NULL
AND
    TransTypeKey IS NOT NULL
AND
    CustVendorKey IS NOT NULL
AND
    UnitCost IS NOT NULL
AND
    Currency IS NOT NULL
AND
    Quantity IS NOT NULL
  
```

3. Reject a record if any field value does not match its data type

Select a validation to edit :

BranchPlantKey
DateStrChk
ItemMasterKey
TransTypeKey
CustVendorKey
UnitCost
Currency
Quantity

☐ Report all errors, not only the first
☒ Output one row, concatenate errors with separator : |

Validation description BranchPlantKey
 Name of field to validate BranchPlantKey
 Error code
 Error description

Type

Verify data type? ☒
 Data type Integer
 Conversion mask
 Decimal Symbol
 Grouping Symbol

Data

Null allowed? ☐
 Only null values allowed? ☐
 Only numeric data expected? ☐
 Max string length
 Min string length
 Maximum value
 Minimum value
 Expected start string
 Expected end string
 Not allowed start string
 Not allowed end string
 Regular expression expected to match
 Regular expression not allowed to

4. Parse dates in an Excel data source - a Select values step

Step name

Select values

Select & Alter

Remove

Meta-data

Fields to alter the meta-data for :

#	Fieldname	Rename to	Type	Length	Precision	Binary to Normal?	Format	Date Format Lenient?	Date Locale
1	BranchPlantKey		None	2	0	N		N	
2	Date		String			N	MM/dd/yyyy	N	
3	ItemMasterKey		None	2	0	N		N	
4	TransTypeKey		None	1	0	N		N	
5	CustVendorKey		None	2	0	N		N	
6	UnitCost		None	4	2	N		N	
7	Currency		None	3		N		N	
8	Quantity		None	2		N		N	

OK

Cancel

5. Split Date Field into Date-MM, Date-DD and Date-YYYY

Field splitter

Step name:

Field to split:

Delimiter:

Enclosure:

#	New Field	ID	Remove ID?	Type	Length	Precision	Format	Group	Decimal	Currency	Nullif	Default
1	Date-MM		N	Integer	2							
2	Date-DD		N	Integer	2							
3	Date-YYYY		N	Integer	2							

Help

OK Cancel

6. Reject invalid foreign key references : Join Date Dim

Merge Join

Step name: Merge Join

First Step: Sort rows

Second Step: Sort rows 2

Join Type: INNER

Keys for 1st step:

#	Key field
1	Date-DD
2	Date-MM
3	Date-YYYY

Get key fields

Keys for 2nd step:

#	Key field
1	CALDAY
2	CALMONTH
3	CALYEAR

Get key fields

OK Cancel

7. Reject invalid foreign key references : Join Trans Type Dim

Diagram showing data flow: Spreadsheet → Filter rows → Data Validator → Select values → Split Fields → Sort rows → Merge Join → Sort rows 3 → Merge Join 2 → Sort rows 4.

Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

First rows | Last rows | Off

#	BranchPlantKey	Date-MM	Date-DD	Date-YYYY	ItemMasterKey
1	17	10	1	2012	05
2	09	9	3	2012	15
3	01	8	6	2012	19
4	11	9	10	2012	05
5	15	11	12	2012	13
6	03	8	13	2012	17

Merge Join 2

Step name: Merge Join 2

First Step: Sort rows 3

Second Step: Sort rows 4

Join Type: INNER

Keys for 1st step:

#	Key field
1	TransType...

Get key fields

Keys for 2nd step:

#	Key field
1	TRANSTYP...

Get key fields

8. Reject invalid foreign key references : Join Cust Vendor Dim

Merge Join

Step name: Merge Join 3

First Step: Sort rows 5

Second Step: Sort rows 6

Join Type: INNER

Keys for 1st step:

#	Key field
1	CustVendorKey

Get key fields

Keys for 2nd step:

#	Key field
1	CUSTVENDORKEY

Get key fields

9. Reject invalid foreign key references :Join Item Master Dim

spreadsheet Data Validator Split Fields Merge Join Merge Join 2

Merge Join

Step name: Merge Join 4

First Step: Sort rows 7

Second Step: Sort rows 8

Join Type: INNER

Keys for 1st step:

#	Key field
1	ItemMaster...

Keys for 2nd step:

#	Key field
1	ITEMMAST...

Get key fields

Get key fields

Help OK Cancel

10. Reject invalid foreign key references : Join Branch Plant Dim

spreadsheet Data Validator Split Fields Merge Join Merge Join 2 Merge Join 3

Merge Join

Step name: Merge Join 5

First Step: Sort rows 9

Second Step: Sort rows 10

Join Type: INNER

Keys for 1st step:

#	Key field
1	BranchPlantKey

Keys for 2nd step:

#	Key field
1	BRANCHPLANTKEY

Get key fields

Get key fields

Sort rows 6

Dim Item

TransTypeKe

11. Reject invalid foreign key references : Join Currency Dim

spreadsheet Data Validator Split Fields Merge Join Merge Join 2

Merge Join

Step name: Merge Join 6

First Step: Sort rows 11

Second Step: Sort rows 12

Join Type: INNER

Keys for 1st step:

#	Key field
1	Currency

Keys for 2nd step:

#	Key field
1	CURRENCY...

Get key fields

Get key fields

12. Calculator Transformation for Unit Cost and ExtCost:

Calculator

Step name: Calculator

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask	Decimal
1	O_UnitCost	A * B	UnitCost	EXCHANGE_RATE		Number	12	4	N		
2	O_ExtCost	A * B	O_UnitCost	Quantity		Number	14	2	N		

Help OK Cancel

13. Add Sequence

Get Value From Sequence

Step name: Add sequence

Name of value: SEQ_inventory_seq

Use a database to generate the sequence

Use DB to get sequence? ☒

Connection: ExcelConnection Edit... New... Wizard

Schema name: Schema

Sequence name: INVENTORY_SEQ Sequences

Use a transformation counter to generate the sequence

Use counter to calculate sequence? ☐

Counter name (optional):

Start at value: 1

Increment by: 1

Maximum value: 99999999

Help OK Cancel

14. Insert/Update INVENTORY FACT table:

Insert / Update

Step name: Insert / Update

Connection: ExcelConnection [Edit...] [New...] [Wizard...]

Target schema: [Browse...]

Target table: INVENTORY_FACT [Browse...]

Commit size: 100

Don't perform any updates: ☐

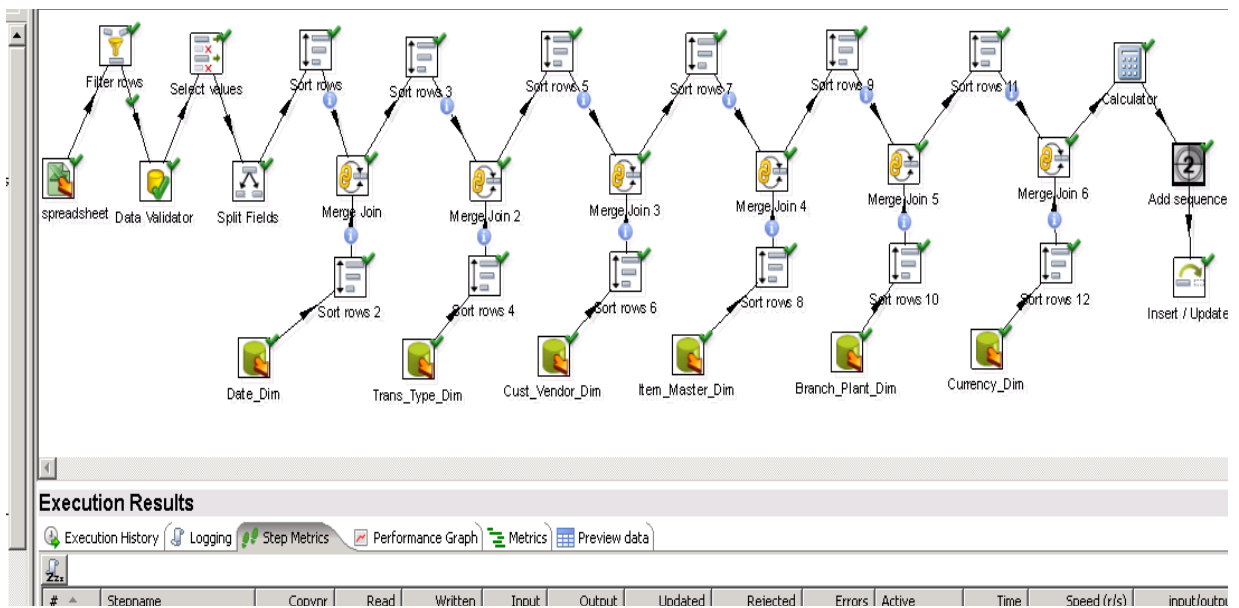
The key(s) to look up the value(s):

#	▲	Table field	Comparator	Stream field1	Stre. ▲	Get fields
1		INVENTORYKEY	=	SEQ_inventory_seq		
2		BRANCHPLANTKEY	=	BRANCHPLANTKEY_1		
3		DATEKEY	=	DATEKEY		
4		ITEMMASTERKEY	=	ITEMMASTERKEY_1		
5		TRANSTYPEKEY	=	TRANSTYPEKEY_1		
6		CUSTVENDORKEY	=	CUSTVENDORKEY_1		
7		UNITCOST	=	O_UnitCost		
8		QUANTITY	=	Quantity		

Update fields:

#	▲	Table field	Stream field	Update	#	Get update fields	Edit mapping
1		INVENTORYKEY	SEQ_inventory_seq	N			
2		BRANCHPLANTKEY	BRANCHPLANTKEY_1	N			
3		DATEKEY	DATEKEY	N			
4		ITEMMASTERKEY	ITEMMASTERKEY_1	N			
5		TRANSTYPEKEY	TRANSTYPEKEY_1	N			
6		CUSTVENDORKEY	CUSTVENDORKEY_1	N			
7		UNITCOST	O_UnitCost	N			
8		QUANTITY	Quantity	N			
9		EXTCOST	O_ExtCost	N			

Complete Transformation:



Execution: Step Output Stats:

Execution History Logging Step Metrics Performance Graph Metrics Preview data													
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	Input
4	Select values	0	13	13	0	0	0	0	0	Finished	1.6s	8	
5	Split Fields	0	13	13	0	0	0	0	0	Finished	1.6s	8	
6	Sort rows	0	13	13	0	0	0	0	0	Finished	1.6s	8	
7	Date_Dim	0	0	104	104	0	0	0	0	Finished	1.9s	54	
8	Sort rows 2	0	104	104	0	0	0	0	0	Finished	2.1s	51	
9	Merge Join	0	117	11	0	0	0	0	0	Finished	2.4s	50	
10	Sort rows 3	0	11	11	0	0	0	0	0	Finished	2.4s	5	
11	Trans_Type_Dim	0	0	5	5	0	0	0	0	Finished	1.9s	3	
12	Sort rows 4	0	5	5	0	0	0	0	0	Finished	1.9s	3	
13	Merge Join 2	0	16	11	0	0	0	0	0	Finished	2.7s	6	
14	Sort rows 5	0	11	11	0	0	0	0	0	Finished	2.7s	4	
15	Cust_Vendor_Dim	0	0	20	20	0	0	0	0	Finished	1.9s	11	
16	Sort rows 6	0	20	20	0	0	0	0	0	Finished	1.9s	11	
17	Merge Join 3	0	31	10	0	0	0	0	0	Finished	3.0s	10	
18	Sort rows 7	0	10	10	0	0	0	0	0	Finished	3.0s	3	
19	Item_Master_Dim	0	0	20	20	0	0	0	0	Finished	1.9s	11	
20	Sort rows 8	0	20	20	0	0	0	0	0	Finished	1.9s	11	
21	Merge Join 4	0	30	10	0	0	0	0	0	Finished	3.3s	9	
22	Sort rows 9	0	10	10	0	0	0	0	0	Finished	3.3s	3	
23	Branch_Plant_Dim	0	0	20	20	0	0	0	0	Finished	1.8s	11	
24	Sort rows 10	0	20	20	0	0	0	0	0	Finished	1.8s	11	
25	Merge Join 5	0	30	10	0	0	0	0	0	Finished	3.7s	8	
26	Sort rows 11	0	10	10	0	0	0	0	0	Finished	3.7s	3	
27	Merge Join 6	0	14	10	0	0	0	0	0	Finished	4.1s	3	
28	Currency_Dim	0	0	4	4	0	0	0	0	Finished	1.8s	2	
29	Sort rows 12	0	4	4	0	0	0	0	0	Finished	1.8s	2	
30	Calculator	0	10	10	0	0	0	0	0	Finished	4.1s	2	
31	Add sequence	0	10	10	0	0	0	0	0	Finished	4.1s	2	
32	Insert / Update	0	10	10	10	10	0	0	0	Finished	4.1s	2	