



IBM Data Stage Lab Guide#13

Loading Unique Records from Flat Files to SQL Server

Description:

*BISP is committed to provide BEST learning material to the beginners and advance learners. In the same series, we have prepared a list of beginner's guide and FAQs for IBM Data Stage. We have built complete financial Data Model and various data transformation techniques. Download many such learning documents, student guide, Lab Guide and Hands-on practice materials. This document guide you through how to Load sorted data into SQL Server from Flat File. **Join our professional training to learn from Experts.***

History:

Version

0.1
0.1

Description Change

Initial Draft
Review#1

Author

Varun Khare
Amit Sharma

Publish Date

12th Aug 2012
18th Aug 2012

Mapping Sheet

Unique Value Load.

Source- Flat File

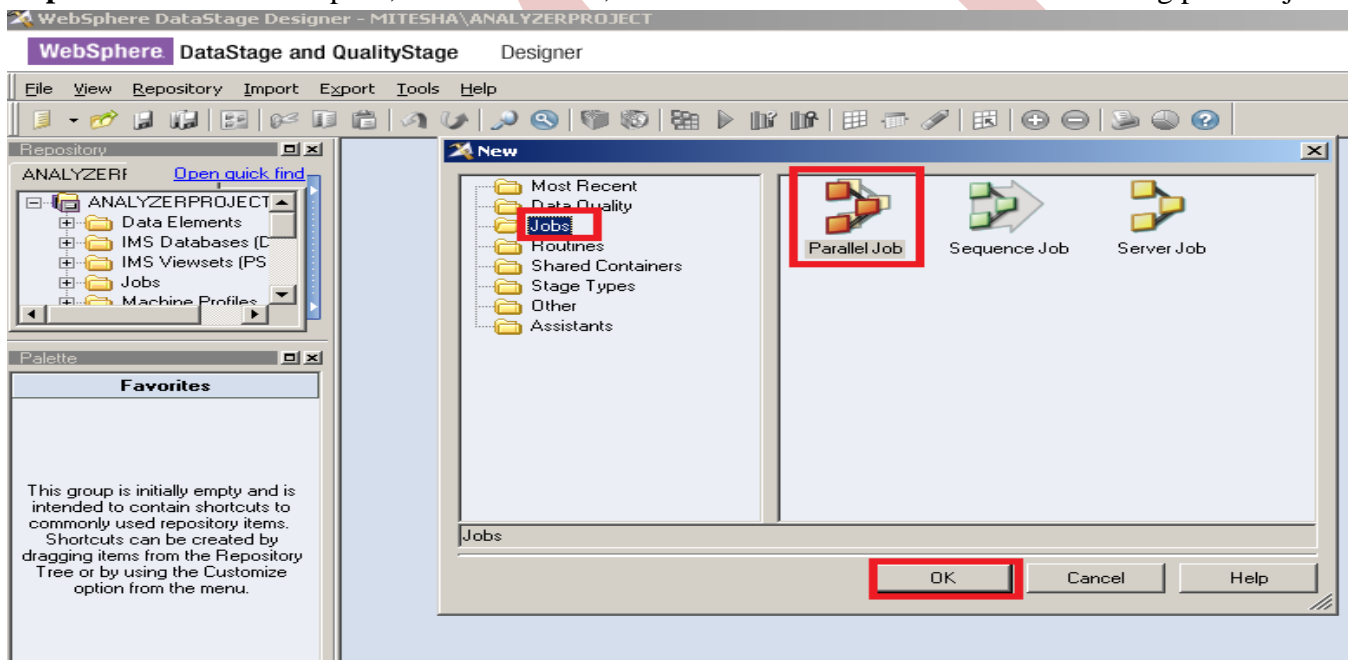
Target- Sql Server

Transformation	Unique Value Load					
Source Name	Flat File					
Target Name	Sql Server					
Flat File Source Details				SQL SERVER Details		
Entity Name	Data Types	Is Null	Expression	Destination Entity Na	DestinationField Name	Is Null
BRANCH_CD	NUMBER(5)	N	Load Unique Value into Sql Server from Flat File	sa	BRANCH_CD	N
CREATED_BY	VARCHAR2(30)	y		sa	CREATED_BY	y
CREATION_DATE	TIMESTAMP(6)	y		sa	CREATION_DATE	y
DEFINITION_LANGUAGE	VARCHAR2(4)	y		sa	DEFINITION_LANGUAGE	y
ENABLED_FLAG	VARCHAR2(1)	y		sa	ENABLED_FLAG	y
LAST_MODIFIED_BY	VARCHAR2(30)	y		sa	LAST_MODIFIED_BY	y
LAST_MODIFIED_DATE	TIMESTAMP(6)	y		sa	LAST_MODIFIED_DATE	y
LEAF_ONLY_FLAG	VARCHAR2(1)	y		sa	LEAF_ONLY_FLAG	y
BRANCH_DISPLAY_CD	VARCHAR2(10)	y		sa	BRANCH_DISPLAY_CD	y

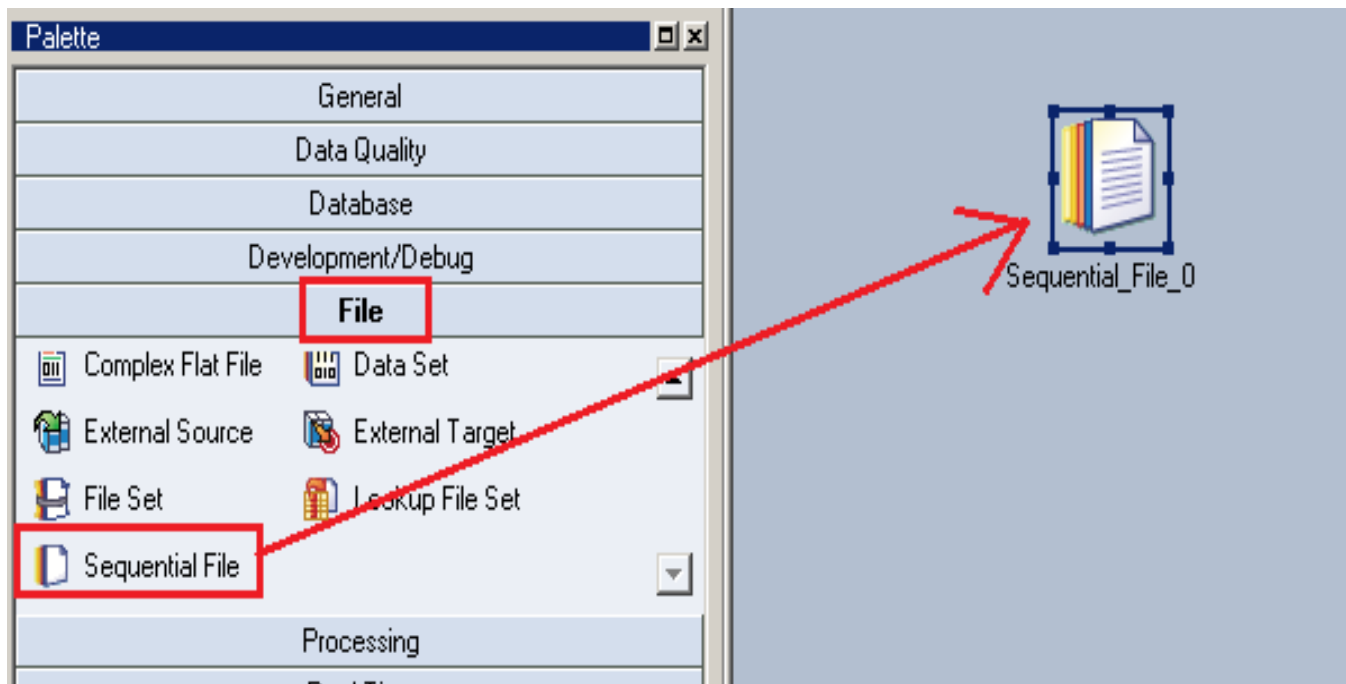
Step-1 :- In Windows Click Designer Client of DataStage



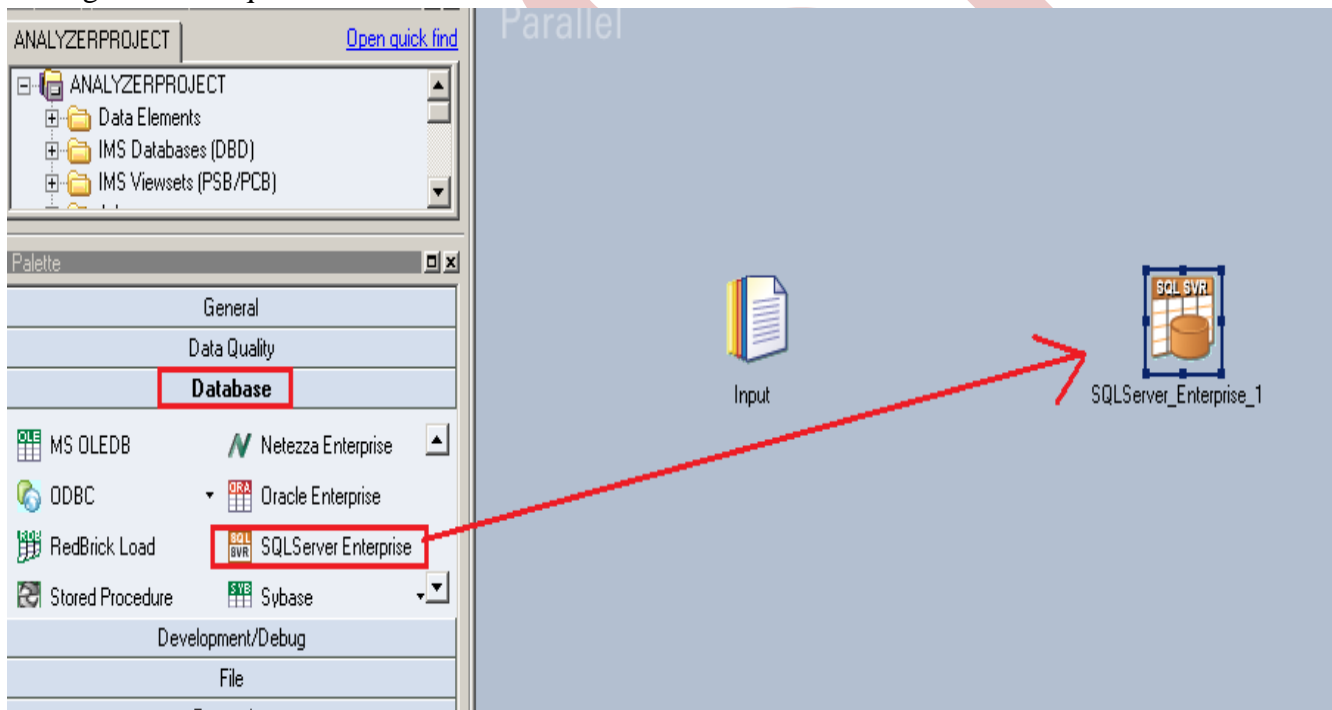
Step-2 :- New Window opens, then click Jobs ,then click Parallel Job because we are using parallel jobs.



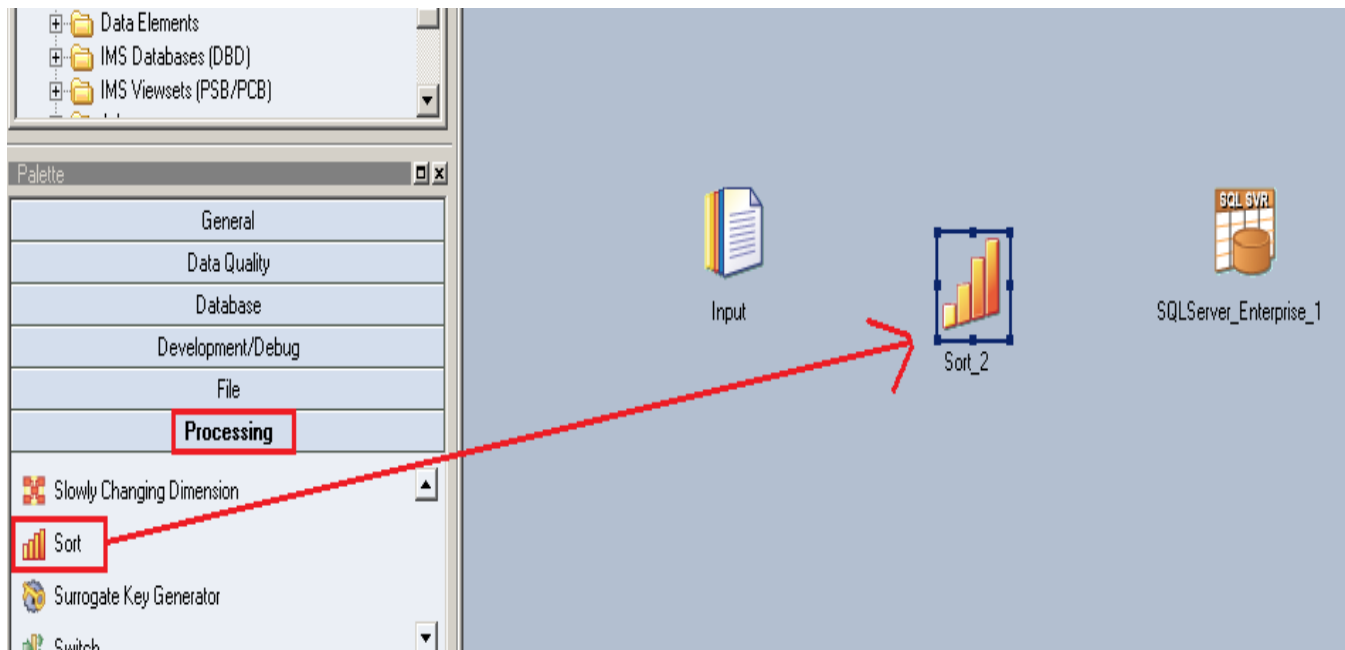
Step-3 :- First go on File Palette then scroll the cursor and choose Sequential File and drag it to Parallel pane and rename it. This file is used for input data.



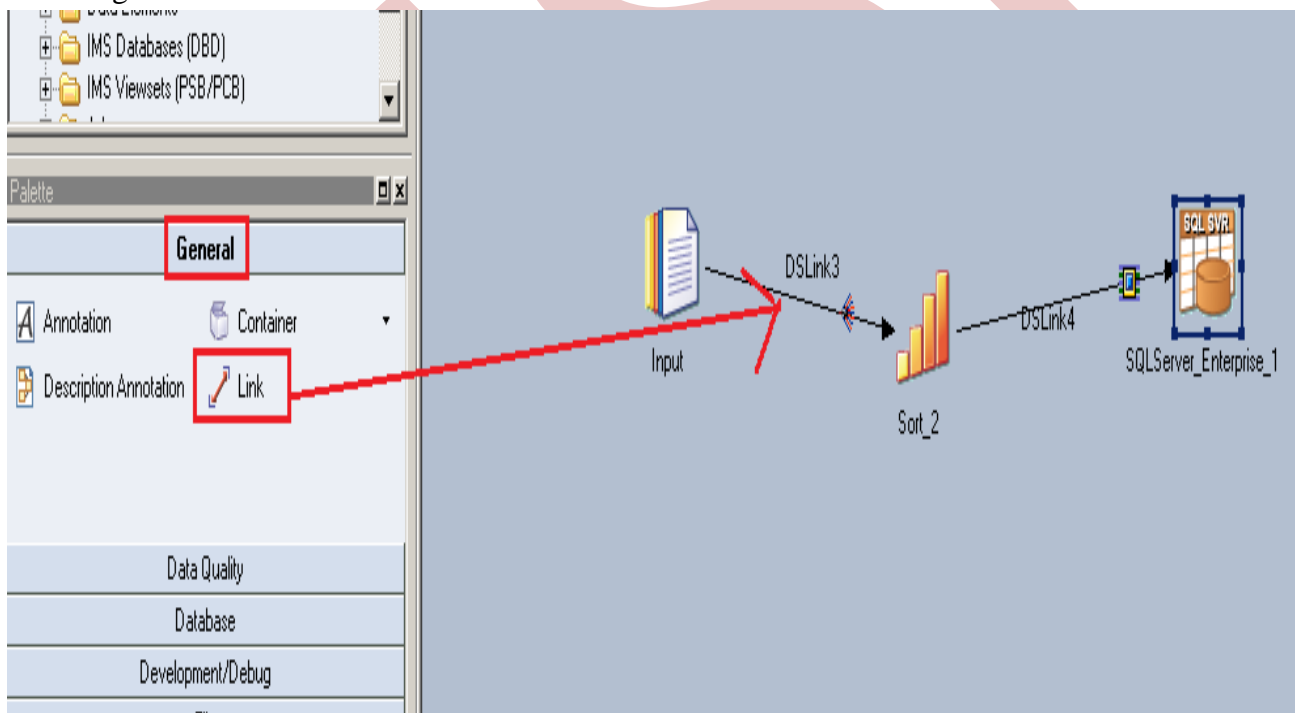
Step-4 :- Choose Sql Server Enterprise from Database Palette and drag it to parallel pane. This is used for loading data into sql server from flat files.



Step 5 :- Choose an sort stage from processing palette. This sort process is used when we want to sort any number of values according to ascending or descending order than we apply this on a particular column and output is based on key column in which we apply sorting we get result.



Step-6:- Now make an connection with these stages that's why we have to use this link stage for connecting these stages.



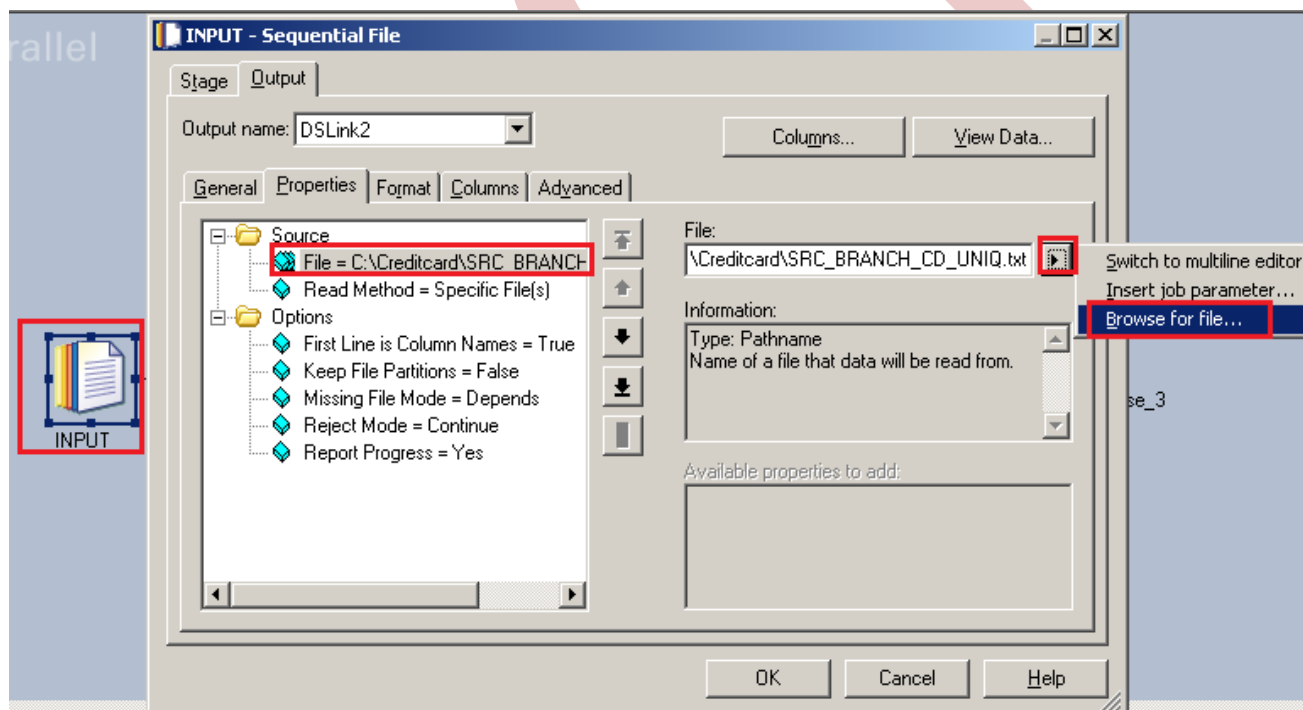
Step-7 : Now first before loading data into output, we have to create a table in sql server. So, we write a create table command and make sure to give same column name as you have in flat file. Goto SQL tab than write query and remember the username and password

```

CREATE TABLE DBO.SRC_BRANCH_CD
(
  BRANCH_CD NUMERIC(5,0) PRIMARY KEY,
  CREATED_BY VARCHAR(30),
  CREATION_DATE TIMESTAMP6),
  DEFINITION_LANGUAGE VARCHAR(4),
  ENABLED_FLAG VARCHAR(1),
  LAST_MODIFIED_BY VARCHAR(30),
  LAST_MODIFIED_DATE TIMESTAMP (6),
  LEAF_ONLY_FLAG VARCHAR(1),
  BRANCH_DISPLAY_CD VARCHAR(10) DEFAULT NULL
);

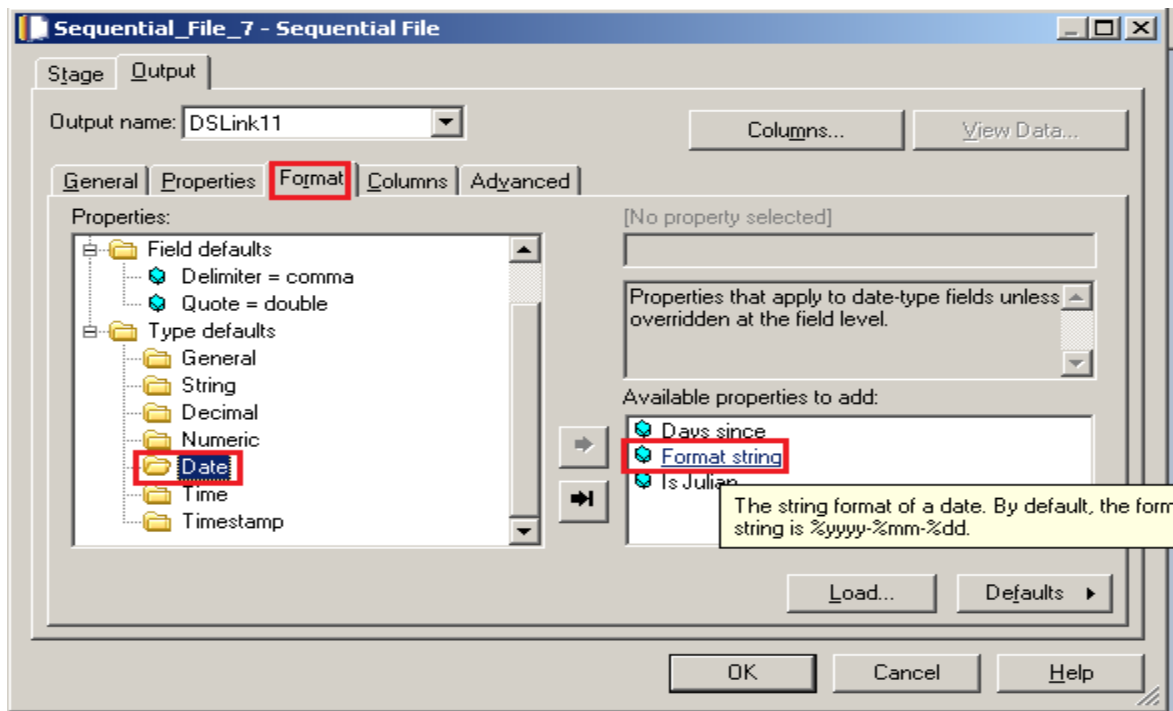
```

Step-8:- Double click on Input file and browse an input file that can be .txt,.csv any test file.

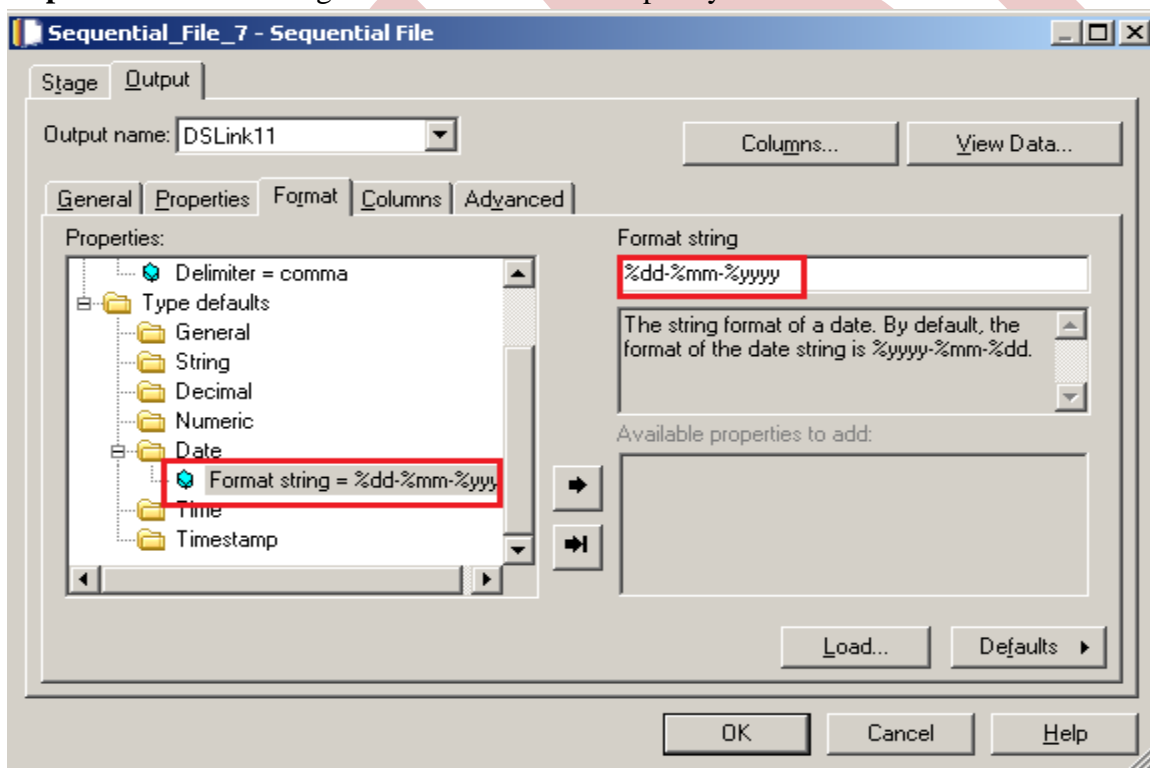


Step-9 Go to Format tab and Remove double Quote because we don't need them then Click on OK.

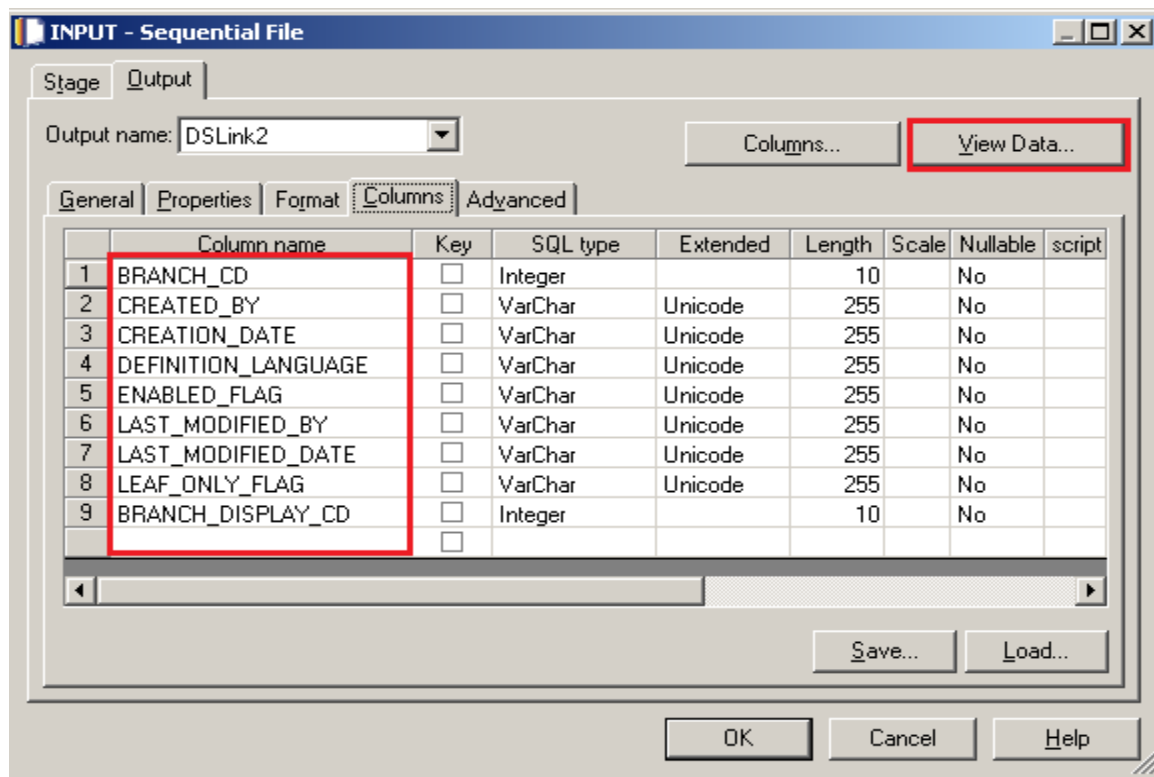
Note:- If you use date as a input than specify here which type of format you used



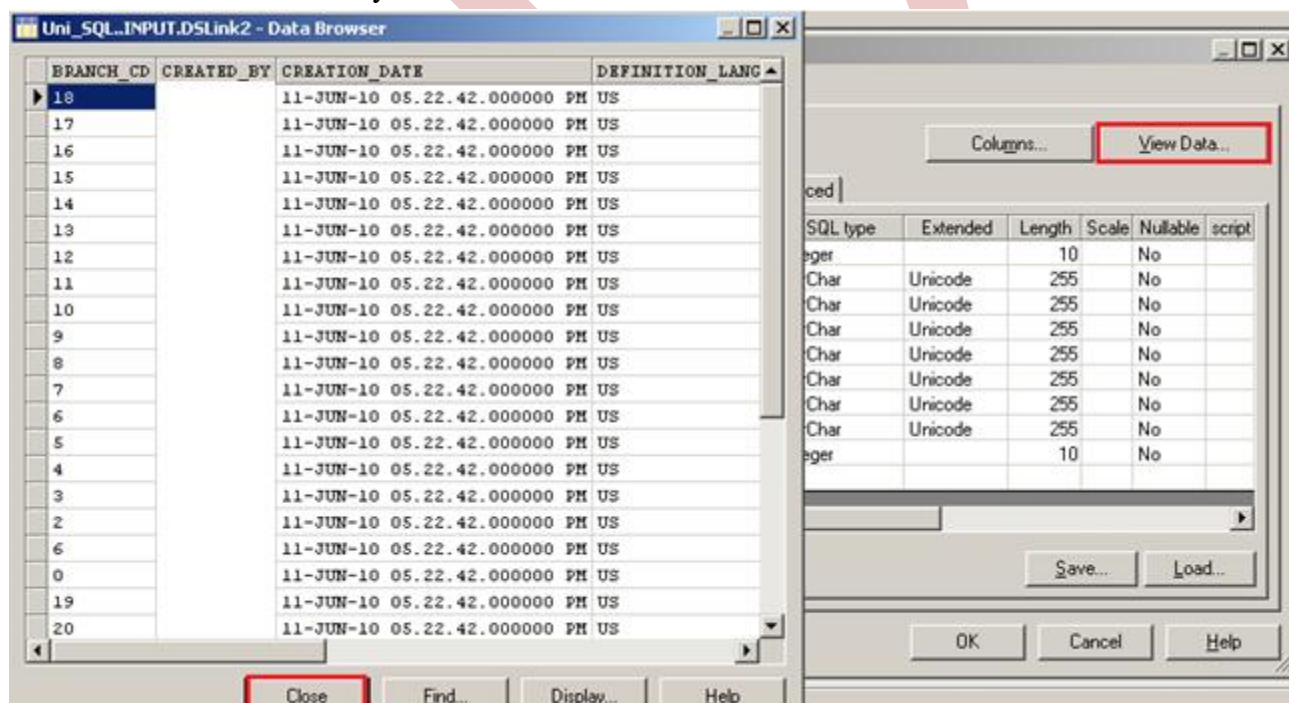
Step-10 :- Here I'm using this date format so we specify here than we able to load the date from flat file.



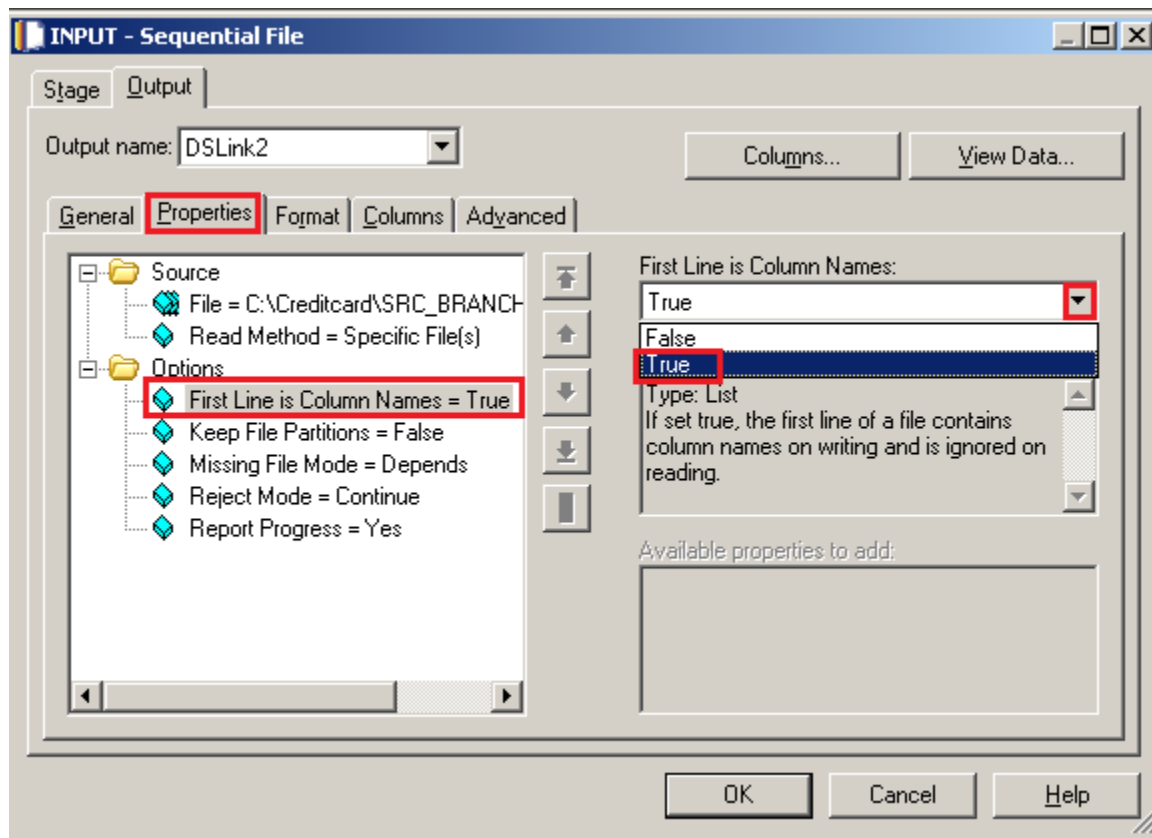
Step-11:- Go to Columns tab and enter same column names as declared in Input file and put datatype, length than click on View Data. This will show your Input Data.



Step-12:- After Clicking on View Data if your data shows like this than our Data is valid and if some error comes than we have to rectify them and check our column names.



Step-13:- After choosing file, select 'True' from First line in column Names dropdown for removing first line from Input file.



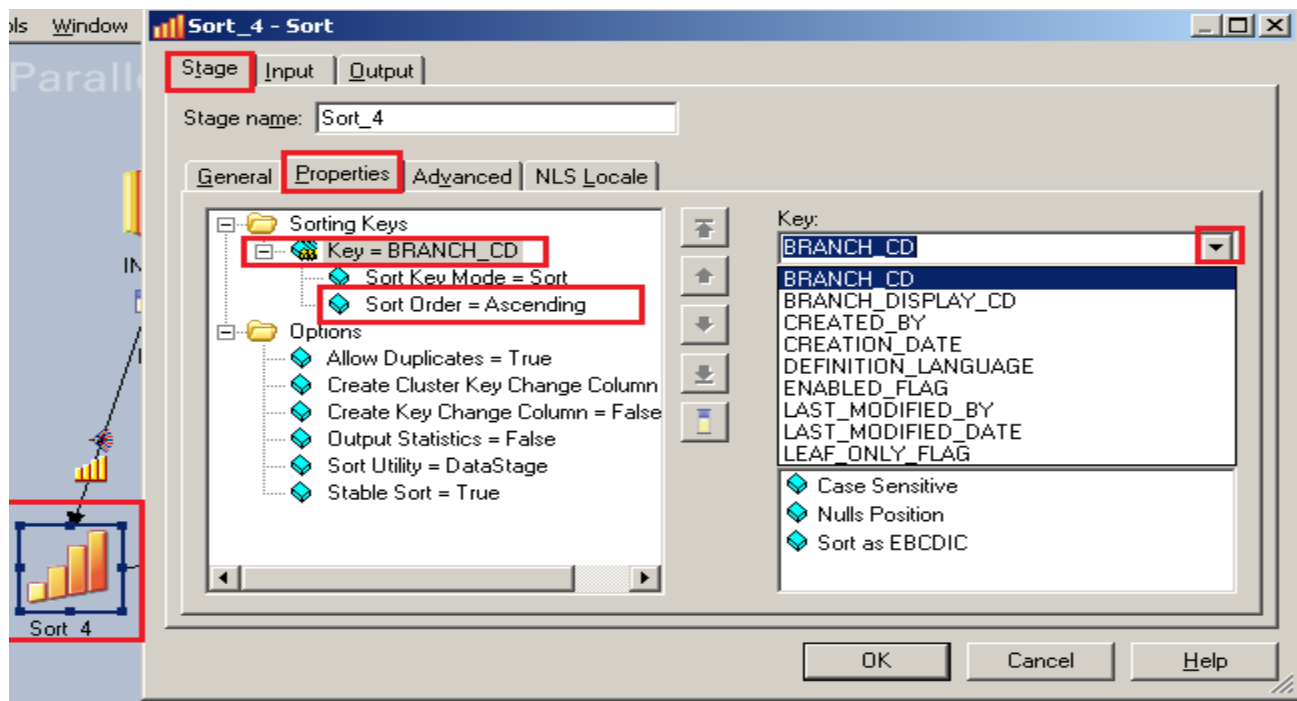
Step-14 :- Make input file and remember column names.

```

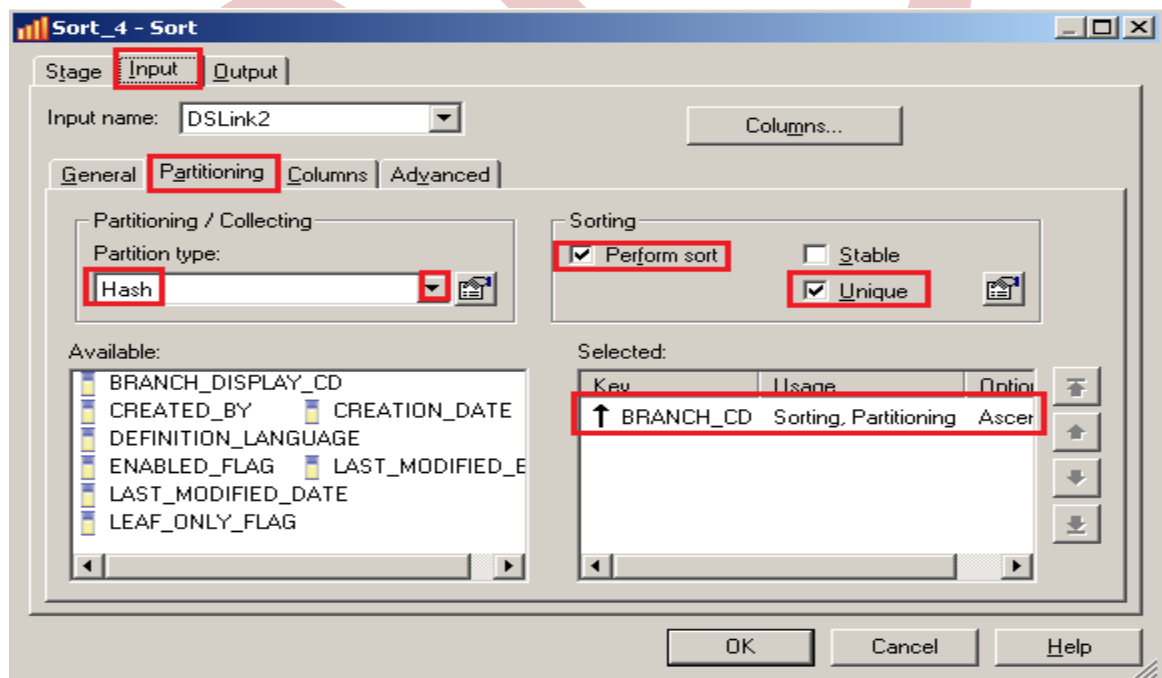
BRANCH_CD,CREATED_BY,CREATION_DATE,DEFINITION_LANGUAGE,ENABLED_FLAG,LAST_MODIFIED_BY,LAST_MODIFIED_DATE,LEAF_ONLY_FLAG,BRANCH_DISPLAY_CD,
18,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,18,
17,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,17,
16,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,16,
15,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,15,
14,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,14,
13,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,13,
12,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,12,
11,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,11,
10,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,10,
9,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,9,
8,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,8,
7,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,7,
6,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,6,
5,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,5,
4,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,4,
3,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,3,
2,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,2,
6,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,1,
0,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,0,
19,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,19,
20,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,20,
20,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,21,
22,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,22,
20,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,23,
24,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,24,
25,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,25,
25,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,26,
27,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,27,
28,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,28,
18,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,29,
30,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,30,
31,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,31,
32,RTUSER,11-JUN-10 05.22.42.000000 PM,US,Y,RTUSER,11-JUN-10 05.22.42.000000 PM,Y,32,

```

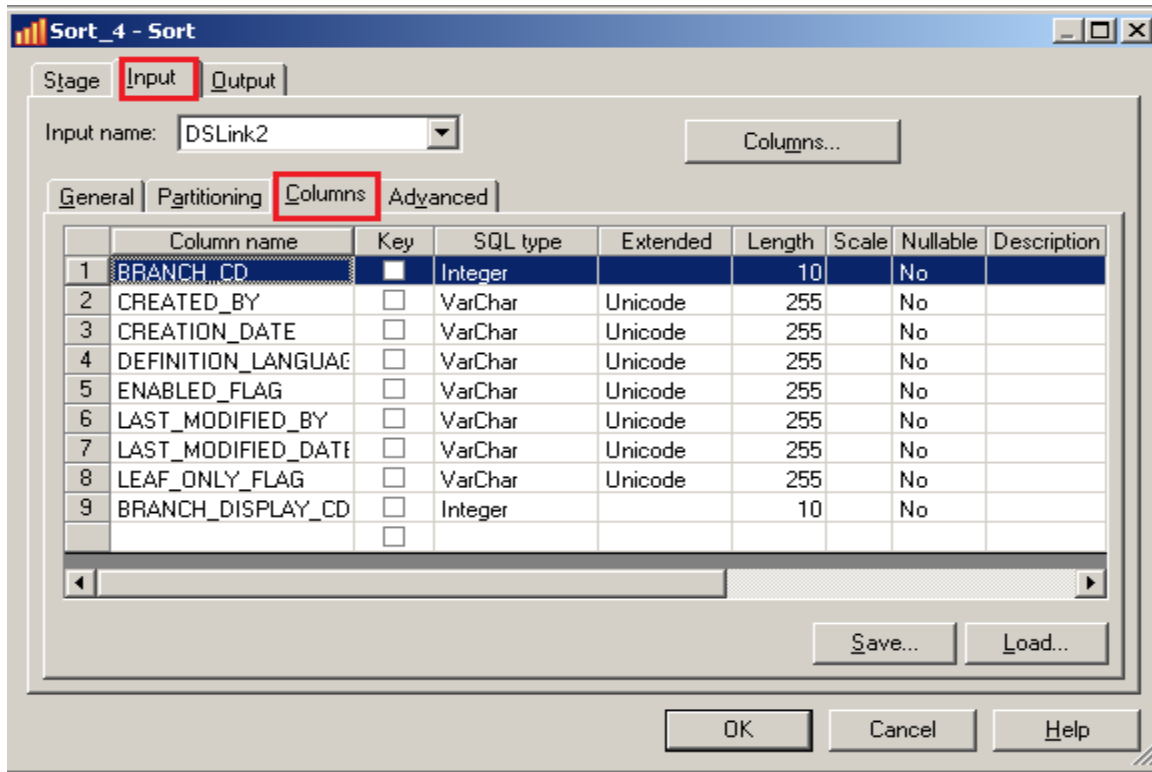
Step-15:- Now double click on sort stage and see here there are some field showing like key field here we have to specify in which column we want to apply sorting so we select here and we also specify in which order we want our target data.



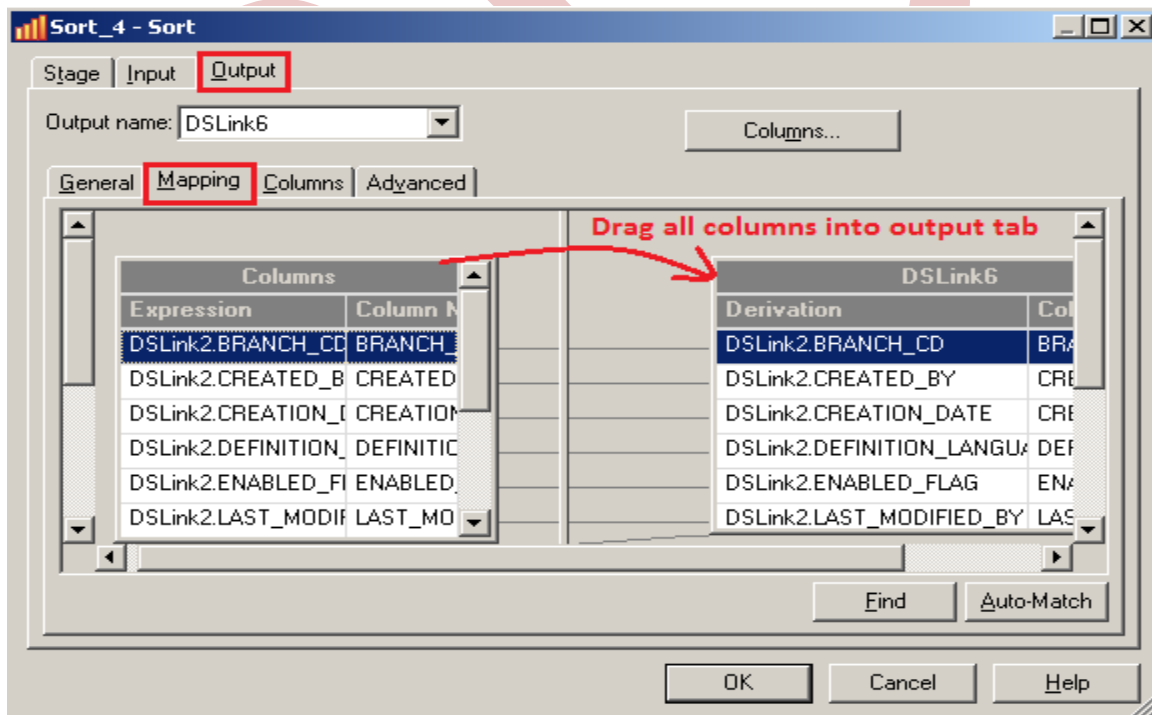
Step-16:- Now click to input and choose hash key partition because these are key based stage so we have to choose hash and check on unique key with the help of this only unique values will store in sql server.



Step-17:- Click on Input and see all columns are available or not

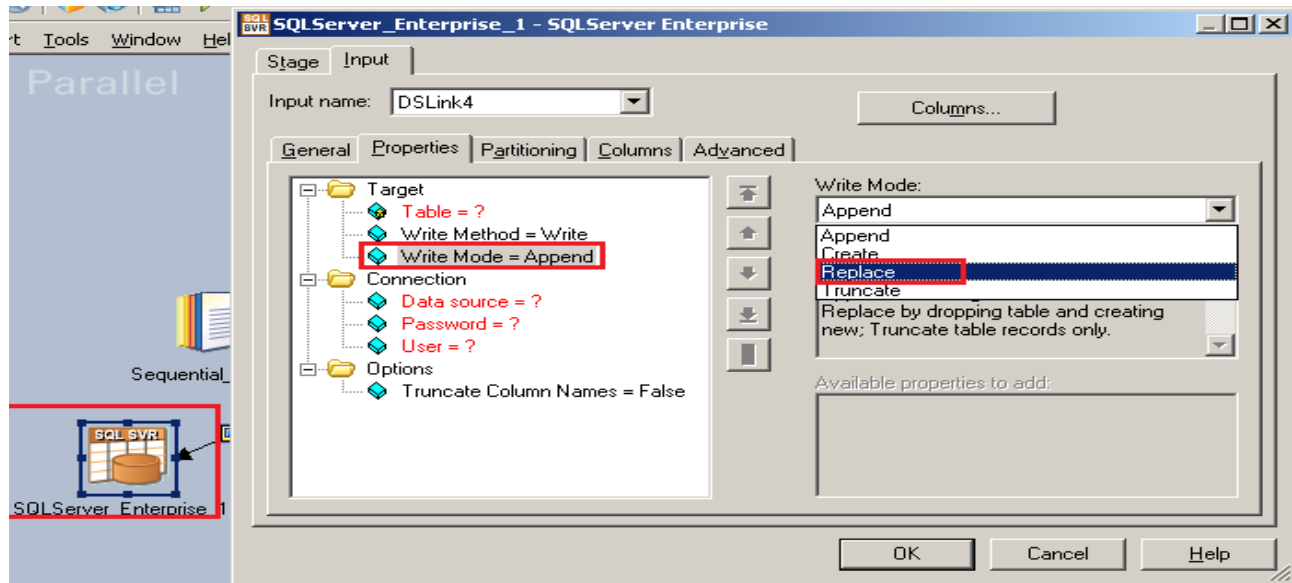


Step-18 Now go to output tab than select all links and drag them into OutLink for sending data which satisfies filter condition into output links..

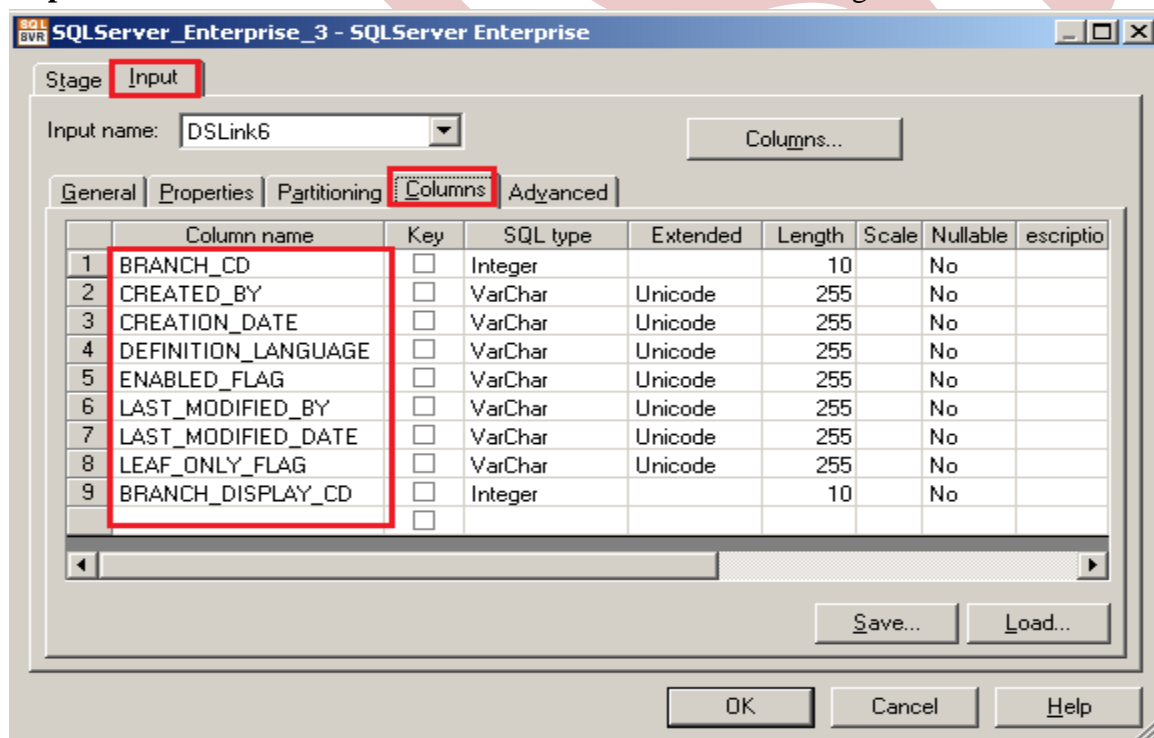


Step-20:- Now double click on Sql Server enterprise stage then it looks like below screenshot. Here we have to enter our table name where we want to insert our data. Multiple options are there for write mode dropdown. If we directly want to load data than we simple used write method as load otherwise we manually

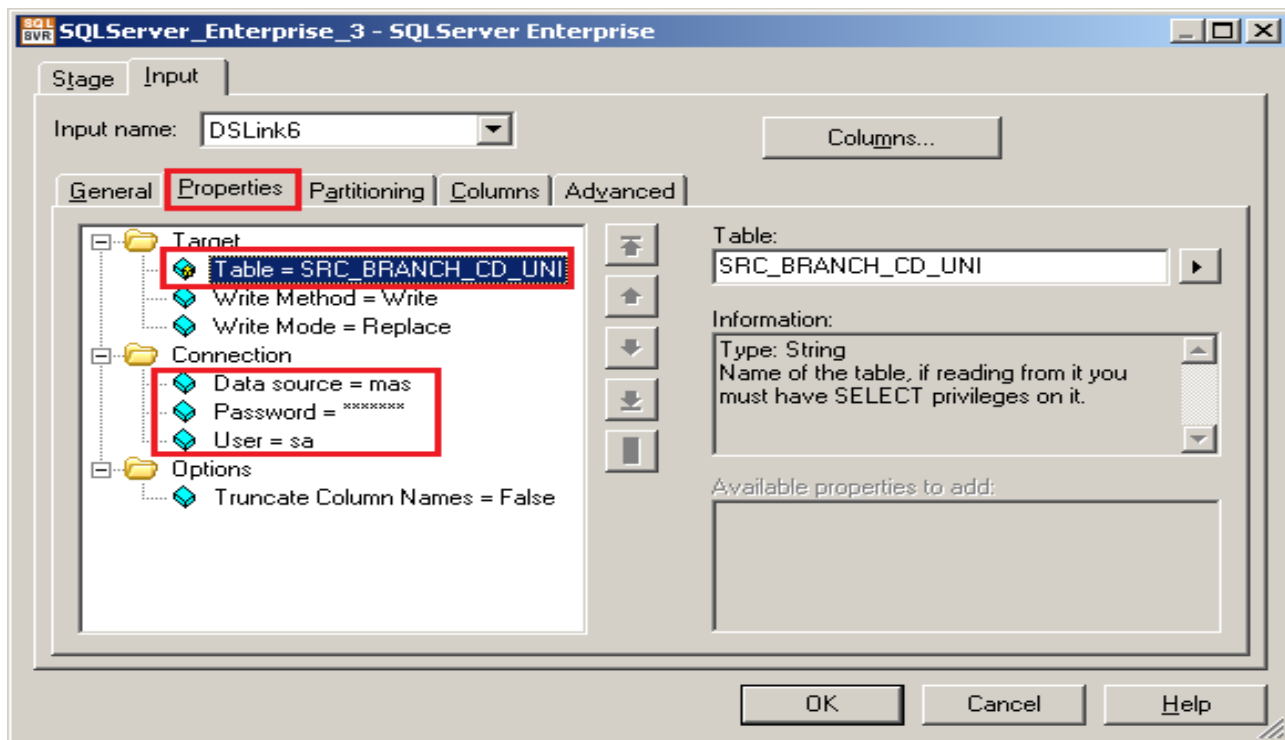
put queries on it. Also we have to specify username and password and this password should match with oracle username and password. Below we specify Replace as write Mode that means if already data is present in table than replace the data with new data.



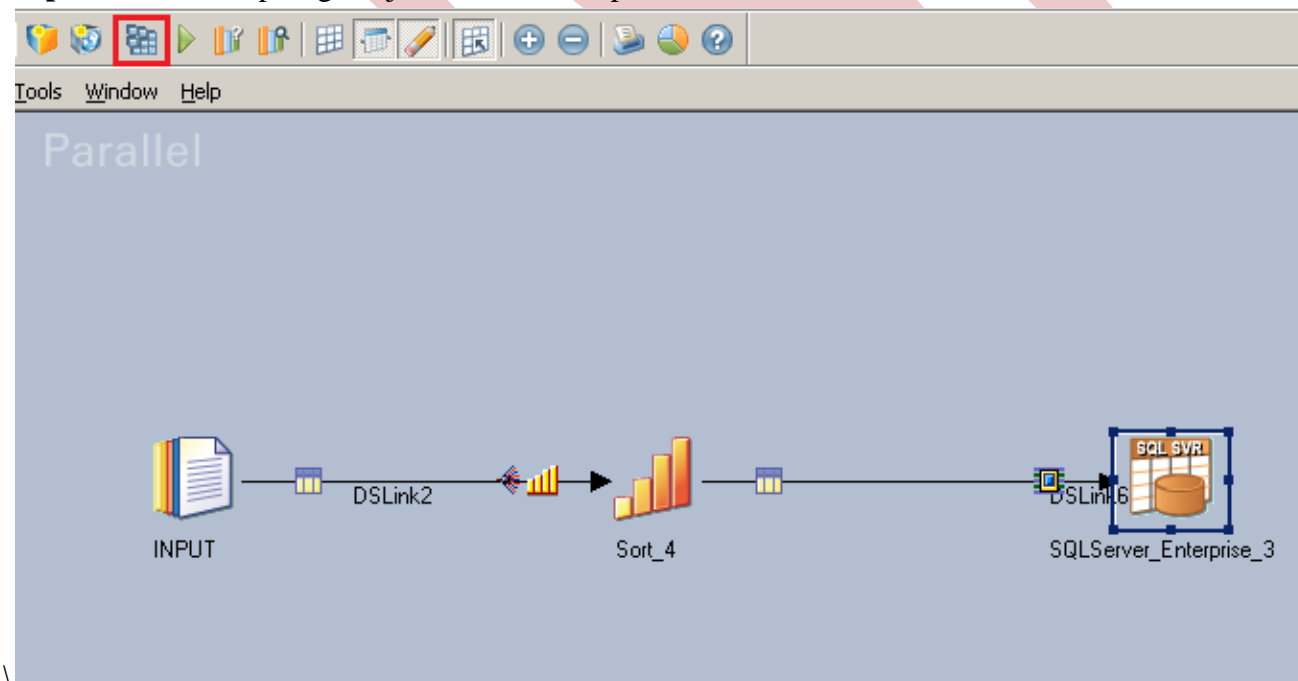
Step-21:- Click on column tab and see all the columns are showing or not.



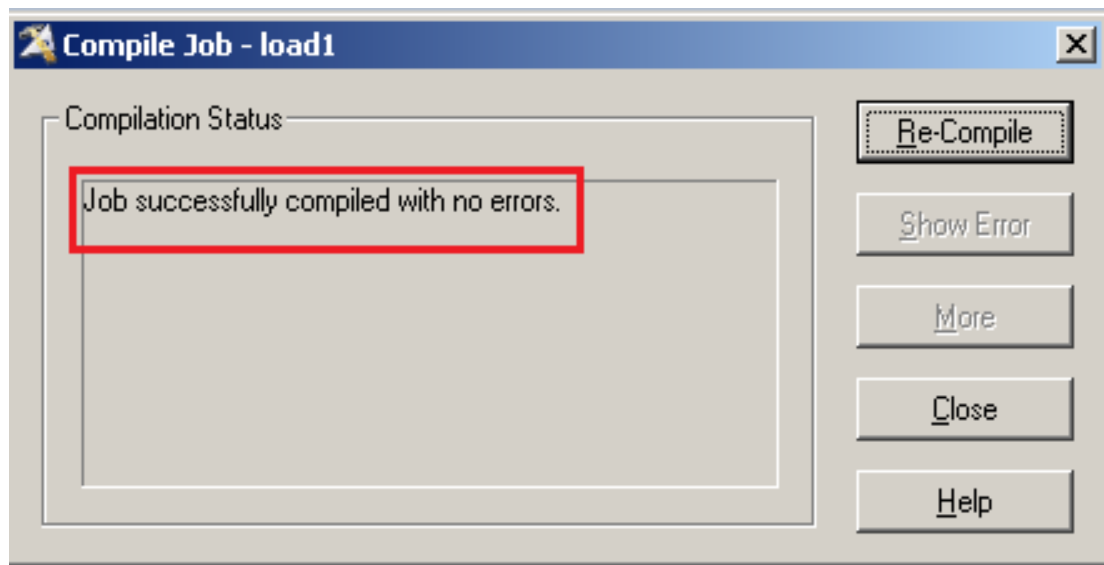
Step-22:- Now here we have to specify the table name and schema name and make an ODBC connection. System DSN and all the other entries should be same as where your table is situated and whatever you specify in System DSN.



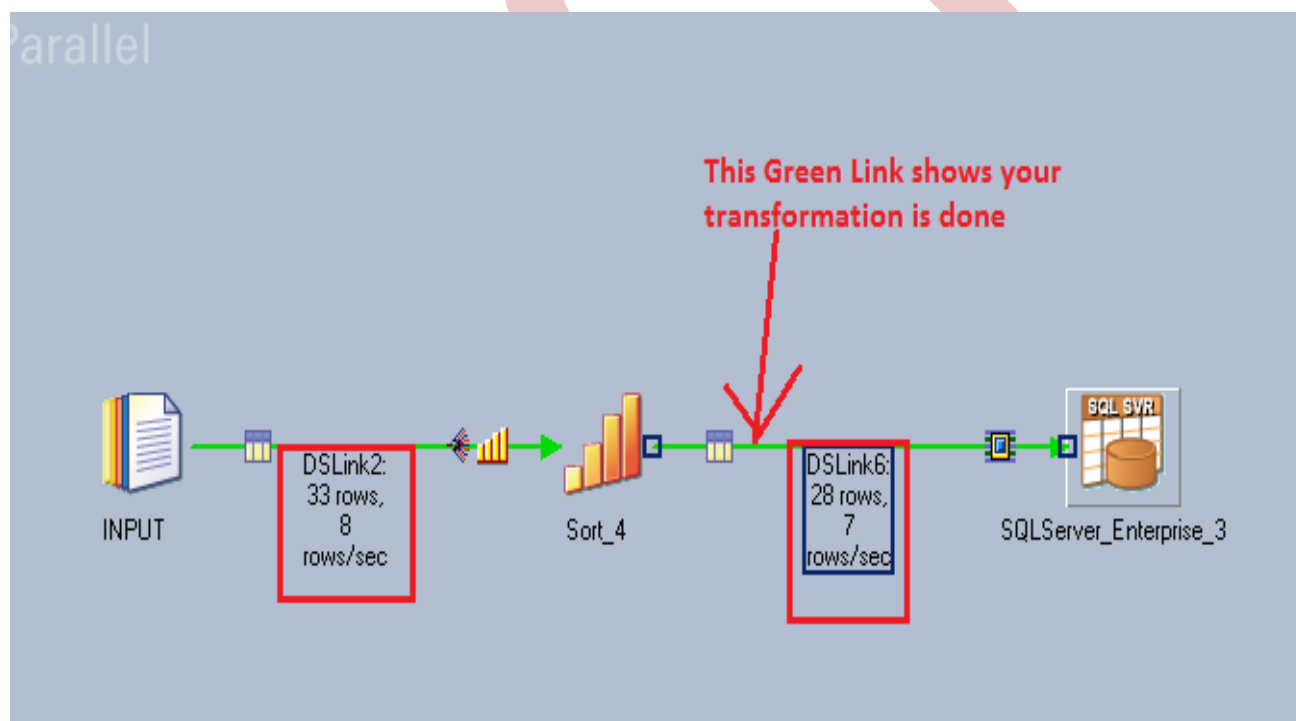
Step-23:- Now compiling the job, click on compile button.



Step-24:- If this is shown that means your compilation is done otherwise it shows error. Now click on run button or (CTRL+F5).



Step-26:- After click on run wait for a while than it shows GREEN line that means your transformation is successfully done otherwise if it shows RED Line that means not Done and BLUE Line means Under Process.



Step-27:- Now Goto SQL Console and connect with the same login credentials as I mentioned in sql server enterprise stage sa/password123 so you can choose your own username and password. Before loading, remember to check structure of this file is available and then you should be able to load the data, then simple query:

```
<SELECT * FROM SRC_BRANCH_CD_UNI;>
```

MITESHA.mast...QLQuery1.sql* Summary									
SELECT * FROM SRC_BRANCH_CD_UNI;									
Results Messages									
	BRANCH_CD	CREATED_BY	CREATION_DATE	DEFINITION_LANGUAGE	ENABLED_FLAG	LAST_MODIFIED_BY	LAST_MODIFIED_DATE	LEAF_ONLY_FLAG	BRANCH_CD
1	0		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	0
2	2		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	2
3	3		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	3
4	4		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	4
5	5		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	5
6	6		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	6
7	7		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	7
8	8		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	8
9	9		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	9
10	10		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	10
11	11		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	11
12	12		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	12
13	13		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	13
14	14		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	14
15	15		11-JUN-10 05:22:42.000000 PM	US	Y	RTUSER	11-JUN-10 05:22:42.000000 PM	Y	15