IBM Data Stage Lab Guide#15
Sort Value and Range Map

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 guides you through Sort Value and Range Map. ***Join our professional training to learn from Experts.***

History:

Version	Description Change	Author	Publish Date
0.1	Initial Draft	Varun Khare	12th Aug 2012
0.1	Review#1	Amit Sharma	18th Aug 2012

Introduction Range Map

The Write Range Map stage takes an input data set produced by sorting a file and writes it to a file in a form usable by the range partitioning method, basically this is debugging stage. The Write Range Map stage is a Development/Debug stage. It allows you to write data to a range map.

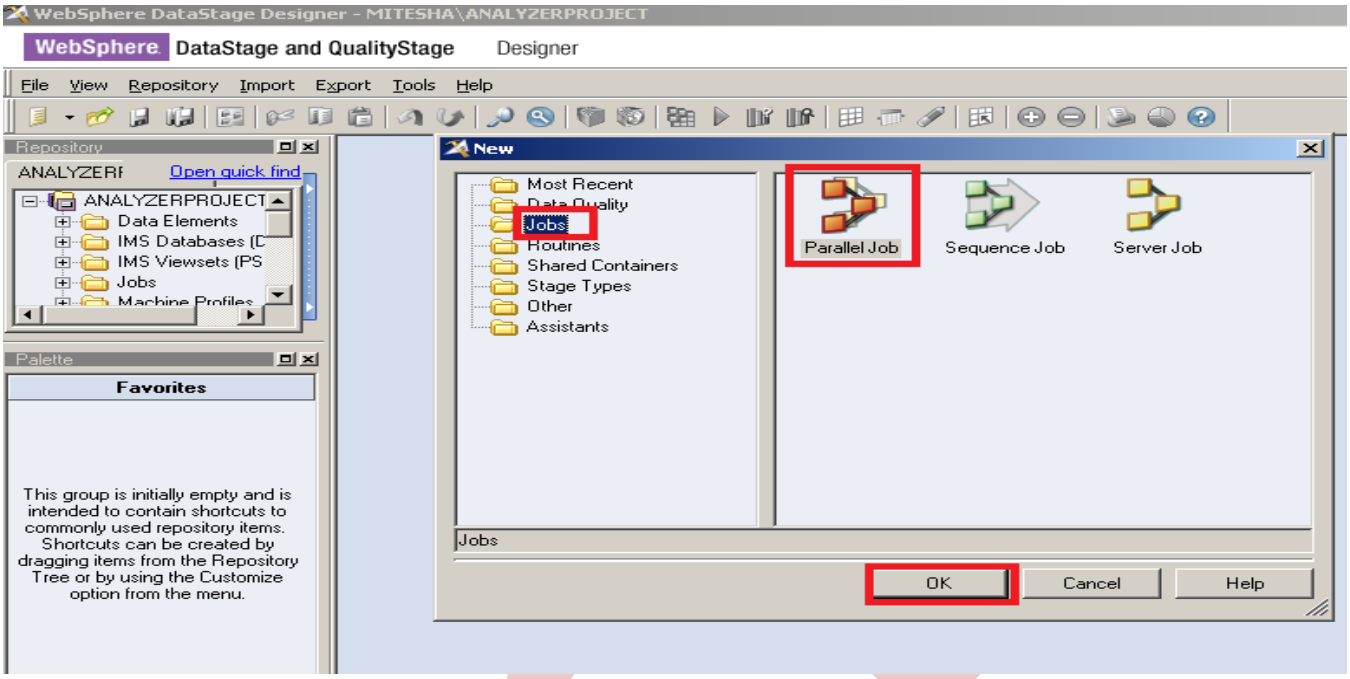
The Write Range Map stage takes an input flat file produced by sorting a flat file and writes it to a file in a form usable by the range partitioning method. The range partitioning method uses the sorted flat file to determine partition boundaries. The stage can have a single input link. It can only run in sequential mode.

A typical use for the Write Range Map stage would be in a job which used the flat file to sort stage, the Sort stage to sort it and the Write Range Map stage to write the range map which can then be used with the range partitioning method to write the original flat file to a file set

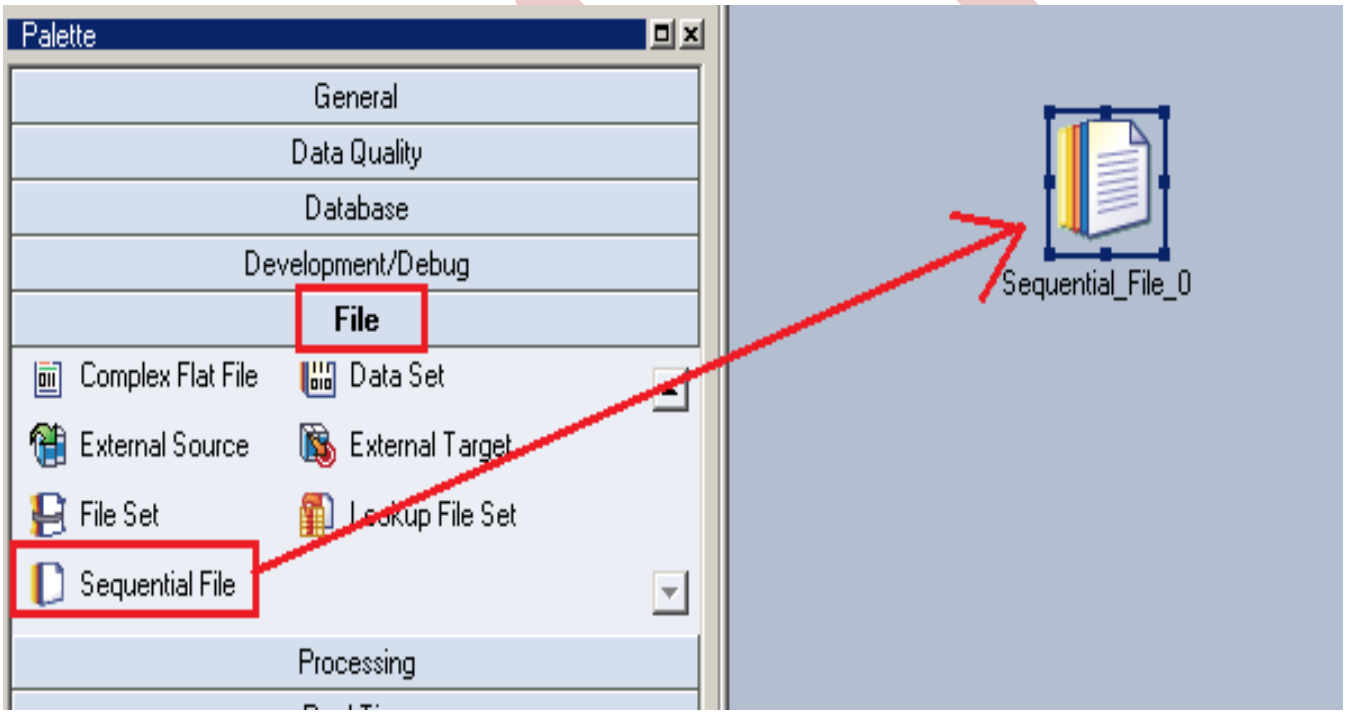
Steps to Load Data Into File Set DATASTAGE

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

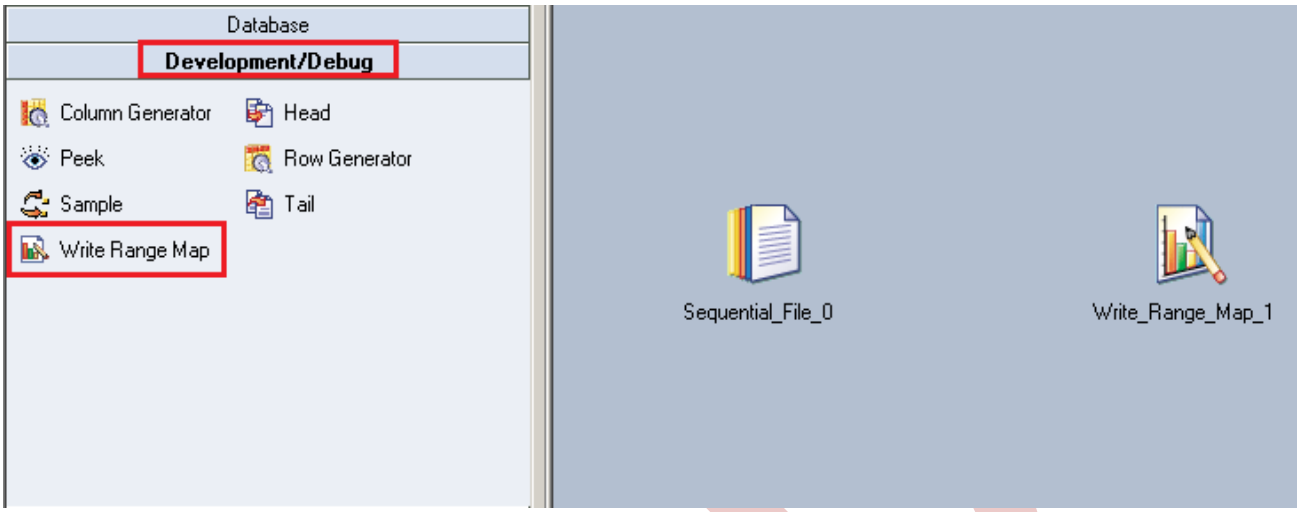
Step-2:- New Window opens, then click Jobs ,then click Parallel Job because we are using parallel jobs.



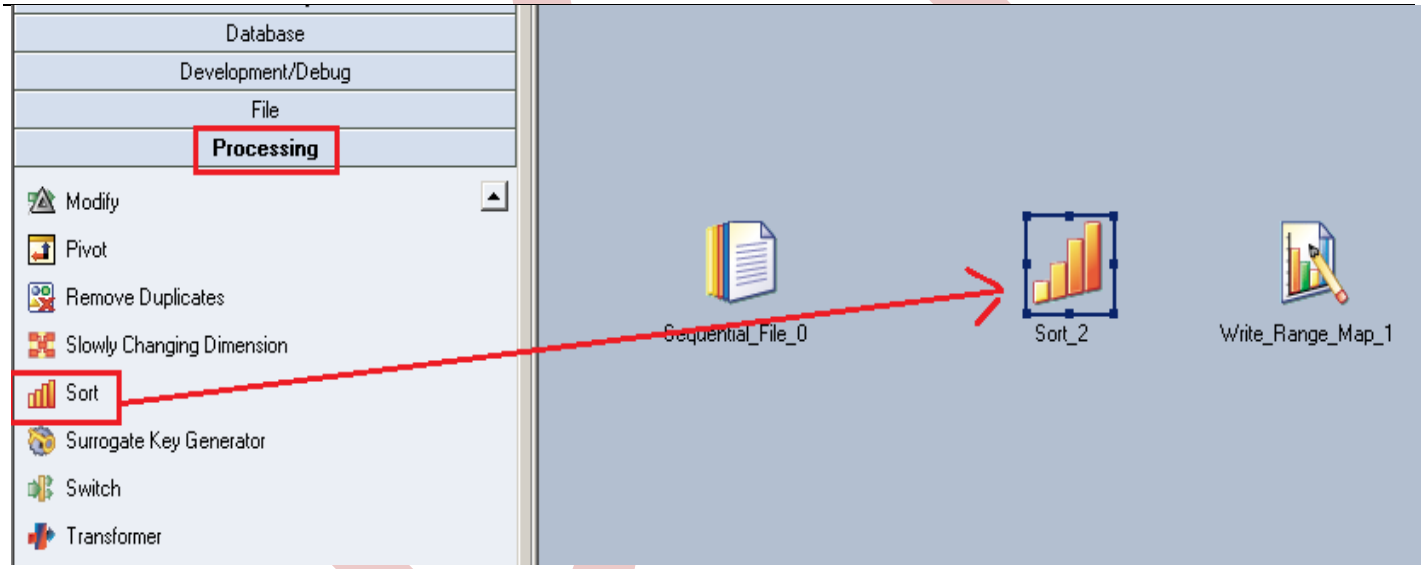
Step-3:- First go on File Palette then scroll the cursor and choose Sequential File and drag it to Parallel pane and rename it. This file is used for input data.



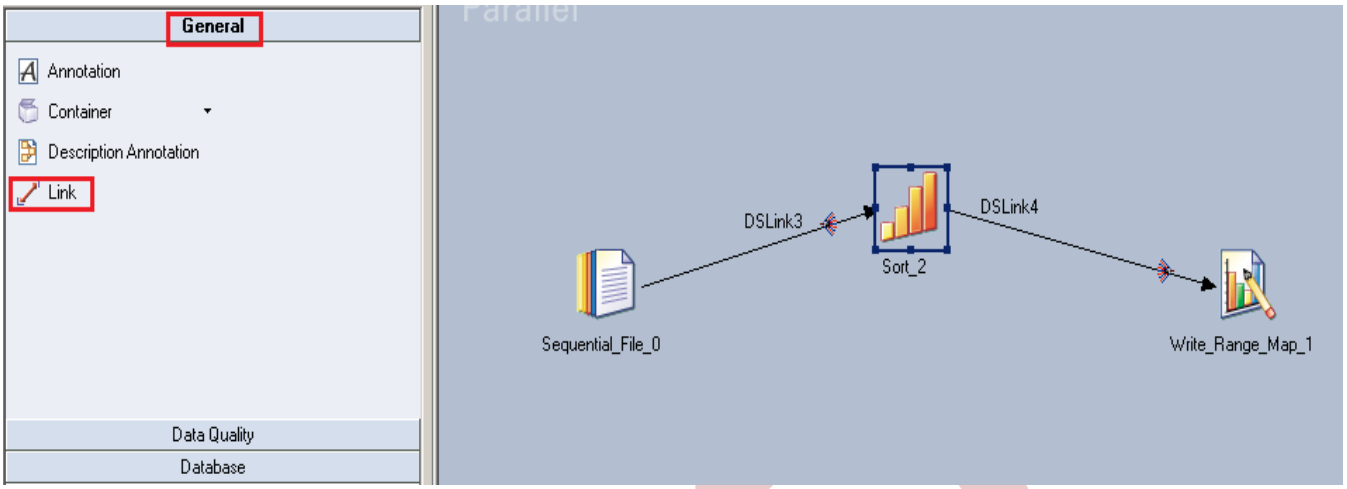
Step-4 :- Choose write range map from development/debug tab this is used for finding range of input file.



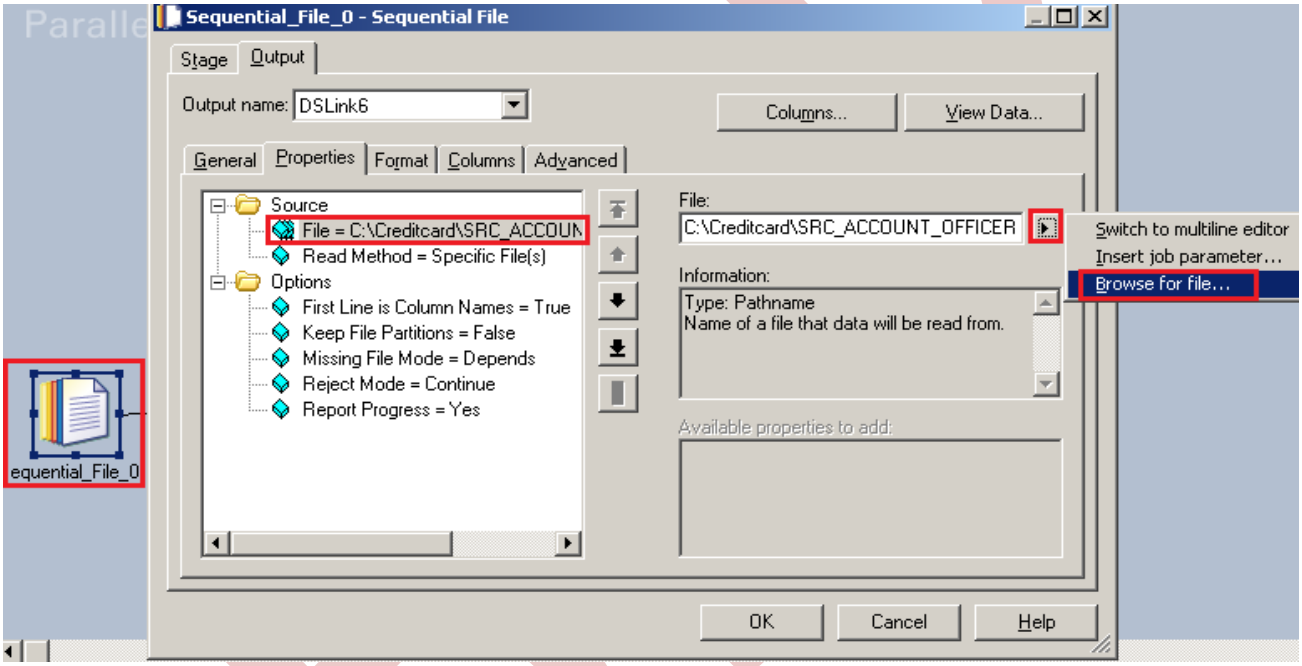
Step-5:- Choose an sort stage from processing palette. This sort process is used when we want to sort any number of values according to ascending or descending order than we apply this on a particular column and output is based on key column in which we apply sorting we get result.



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

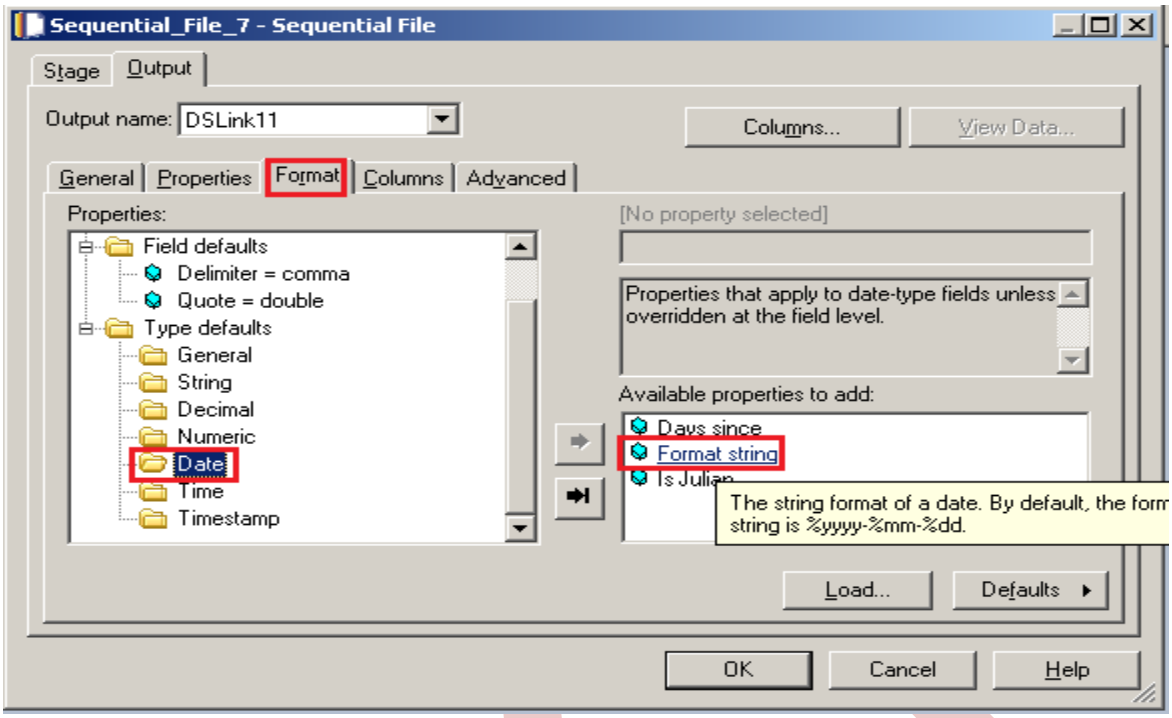


Step-7:- Double click on Input file and browse an input file that can be .txt, .csv any test file.

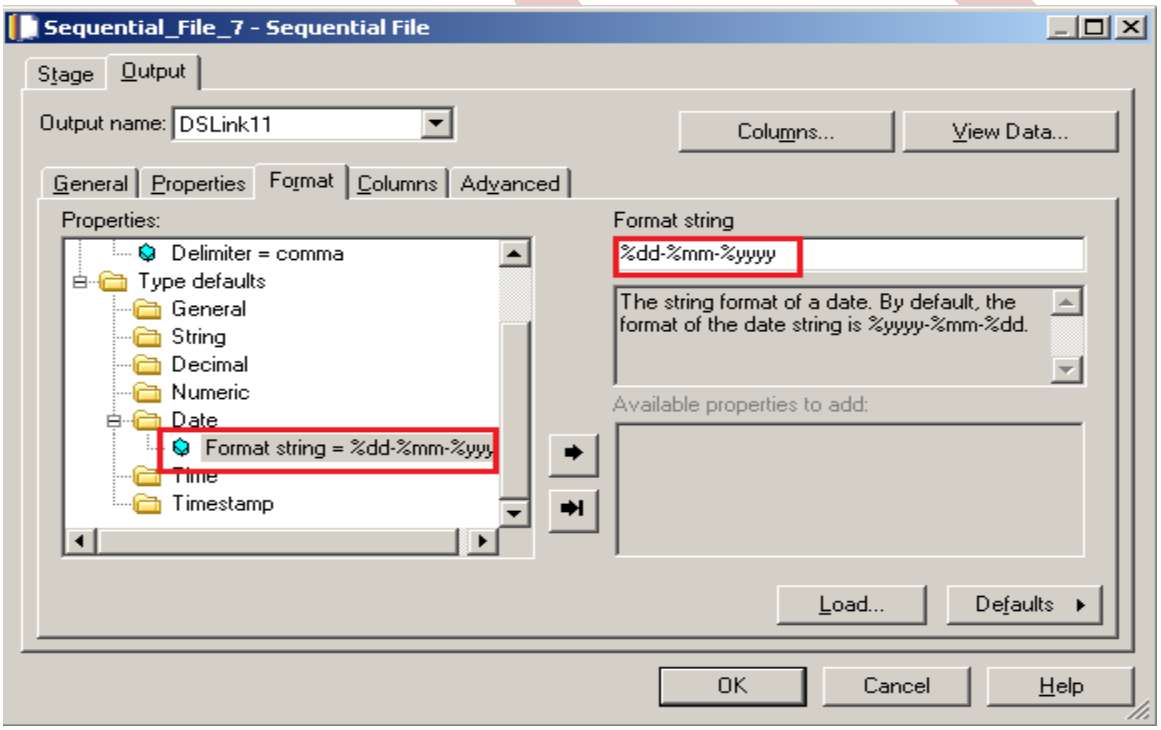


Step-8:- Go to Format tab and Remove double Quote because we don't need them then Click on OK.

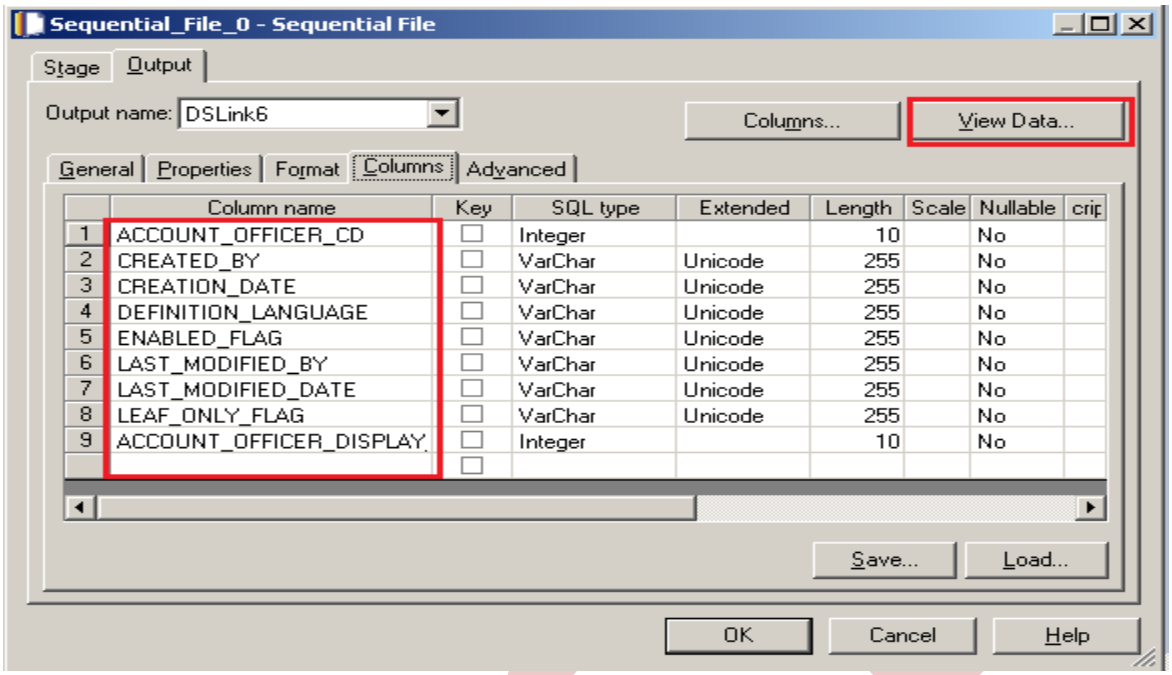
Note:- If you use date as a input than specify here which type of format you used



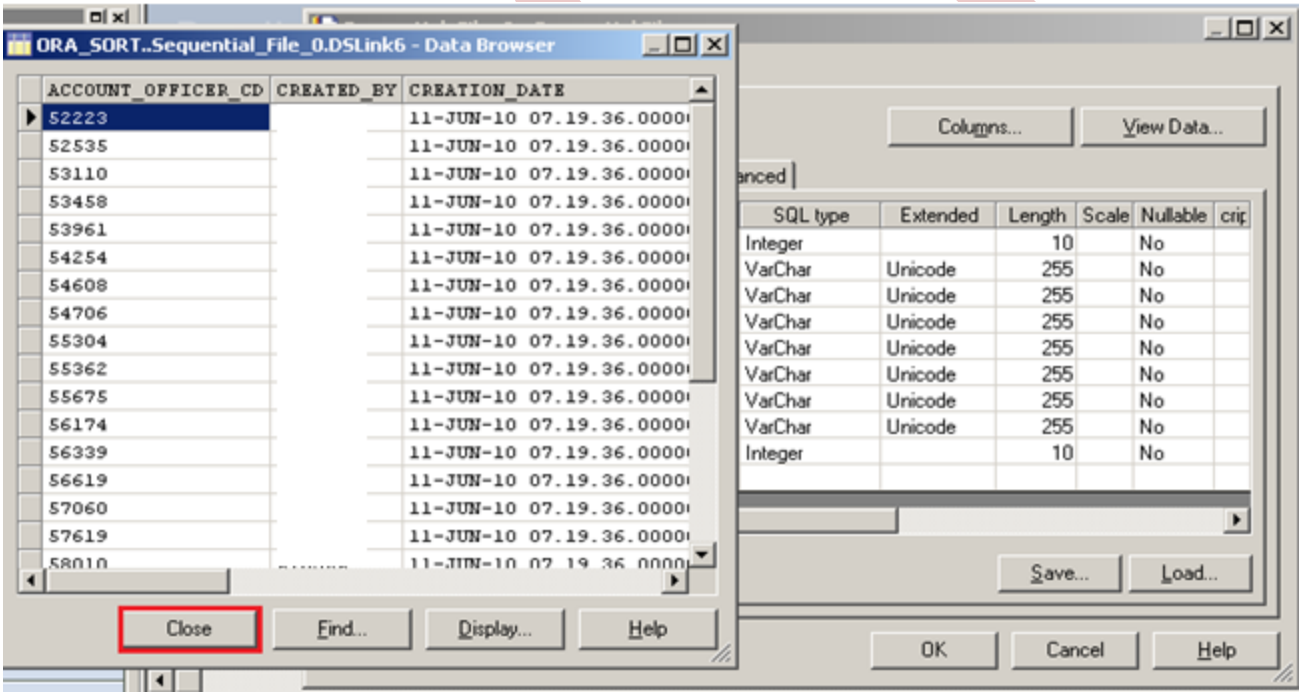
Step-9:- Here I'm using this date format so we specify here than we able to load the date from flat file.



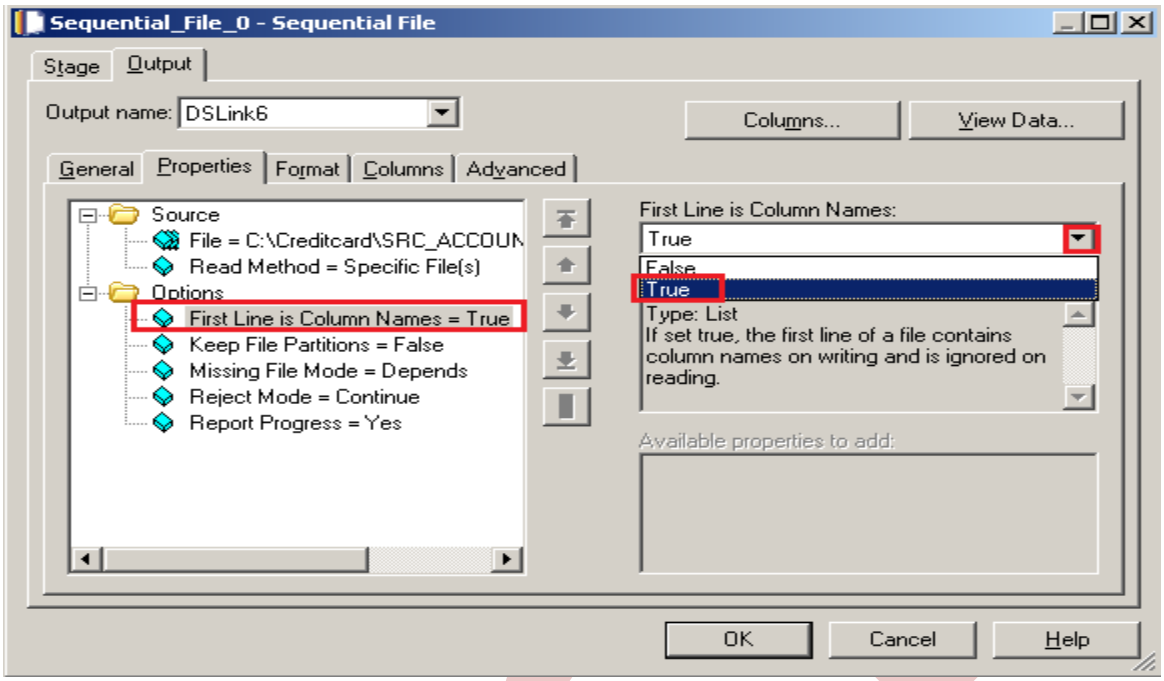
Step-10:- Go to Columns tab and enter same column names as declared in Input file and put data type, length then click on View Data. This will show your Input Data.



Step-11:- After Clicking on View Data if your data shows like this than our Data is valid and if some error comes then we have to rectify them and check our column names.



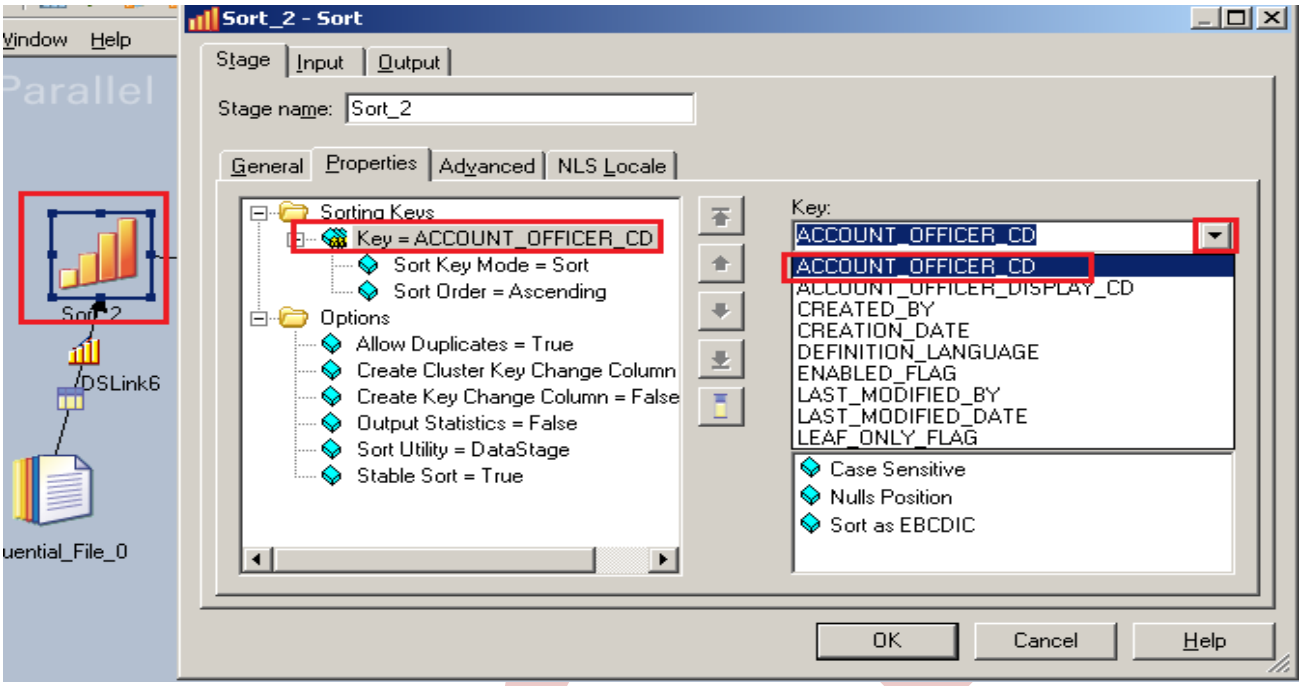
Step-12:- After choosing file, select 'True' from First line in column Names dropdown for removing first line from Input file.



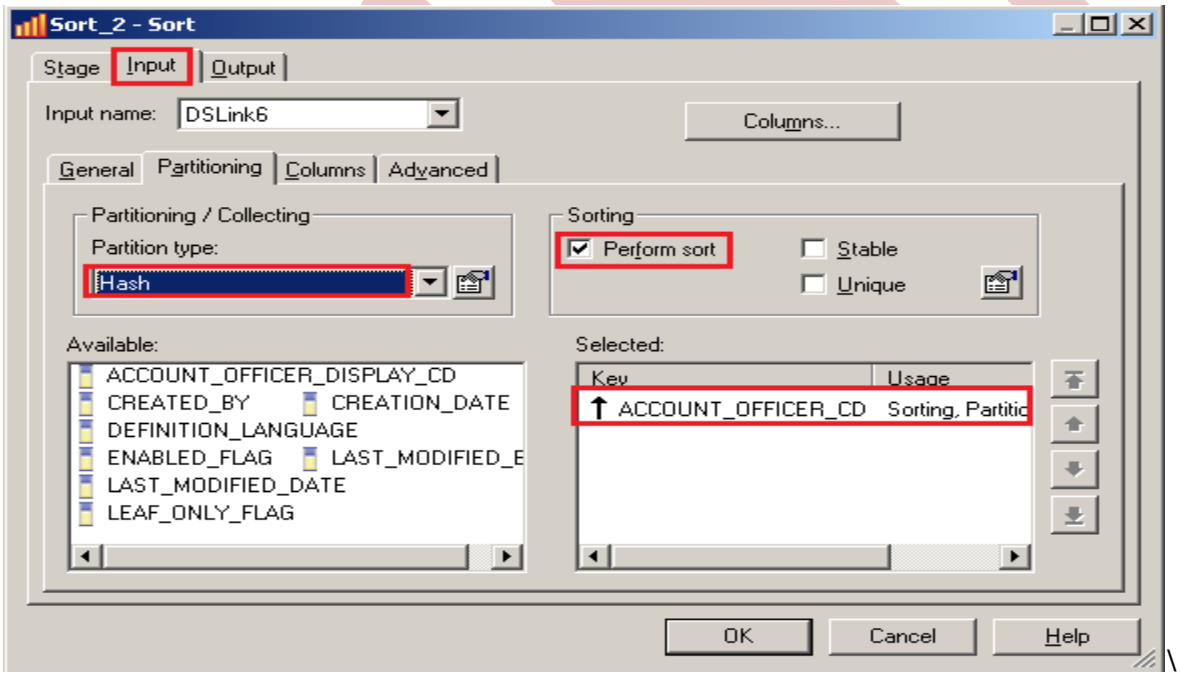
Step-13:- Make input file and remember column names.

```
ACCOUNT_OFFICER_CD,CREATED_BY,CREATION_DATE,DEFINITION_LANGUAGE,ENABLED_FLAG,LAST_MODIFIED_BY,LAST_MODIFIED_DATE,LEAF_ONLY_FLAG,ACCOUNT_OFFICER_DISPLAY_CD,
52223,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,52223.36,
52535,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,52534.89,
53110,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,53110.24,
53458,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,53457.6,,
53961,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,53960.7,,
54254,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,54253.88,
54608,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,54607.99,
54706,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,54705.91,
55304,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,55303.81,
55362,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,55362.01,
55675,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,55674.97,
56174,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,56173.79,
56339,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,56338.61,
56619,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,56618.92,
57060,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,57059.61,
57619,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,57619.17,
58010,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,58010.18,
58235,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,58235.35,
58616,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,58616.48,
58900,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,58899.61,
59154,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,59153.7,,
59567,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,59566.99,
60254,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,60254.29,
60740,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,60740.15,
60951,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,60951.47,
61134,      11-JUN-10 07.19.36.000000 PM,US,Y,      11-JUN-10 07.19.36.000000 PM,Y,61133.62,
```

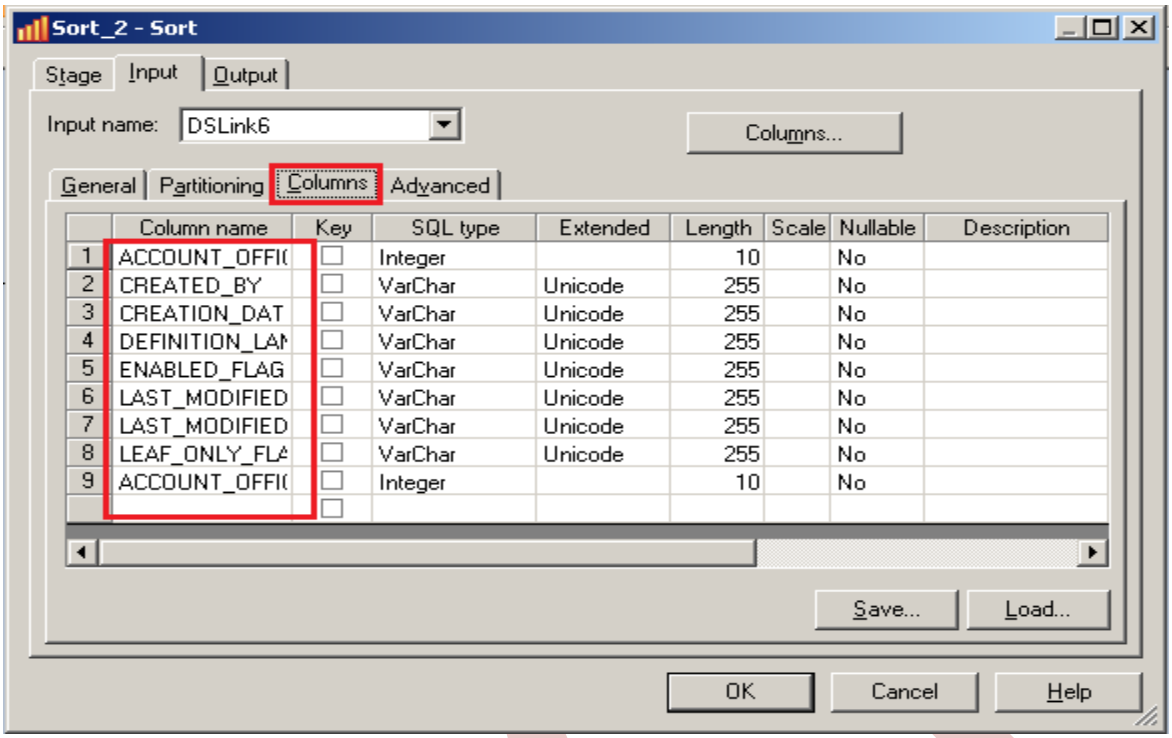
Step-14:- Now double click on sort stage and see here there are some field showing like key field here we have to specify in which column we want to apply sorting so we select here and we also specify in which order we want our target data.



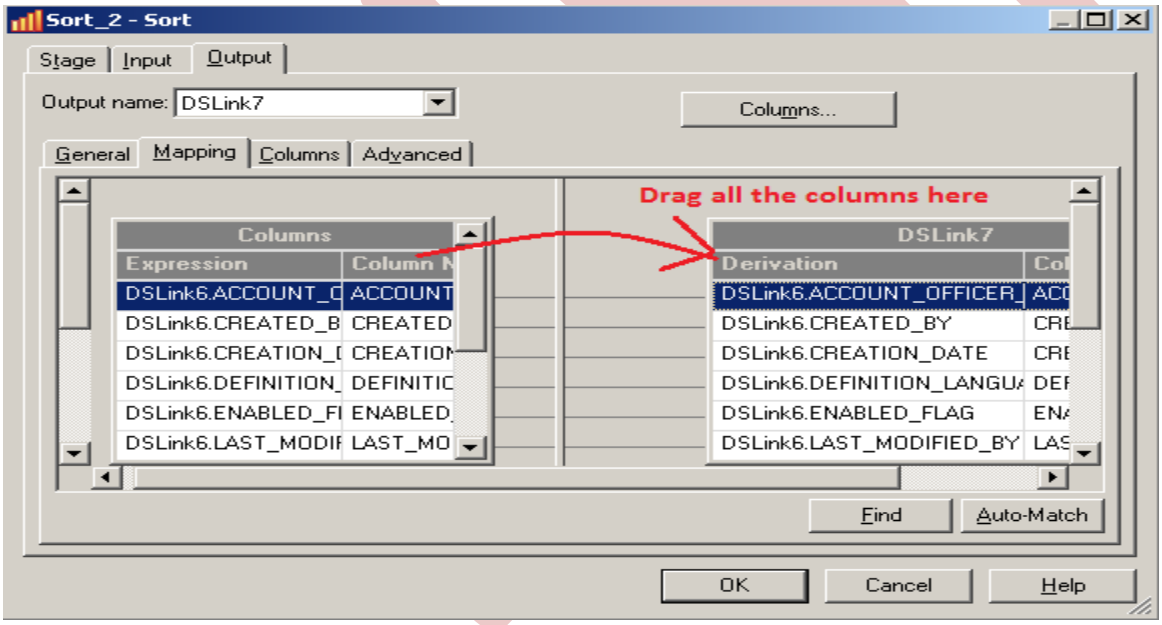
Step-15:- Now click to input and choose hash key partition because these are key based stage so we have to choose hash and check on perform sort for sorting.



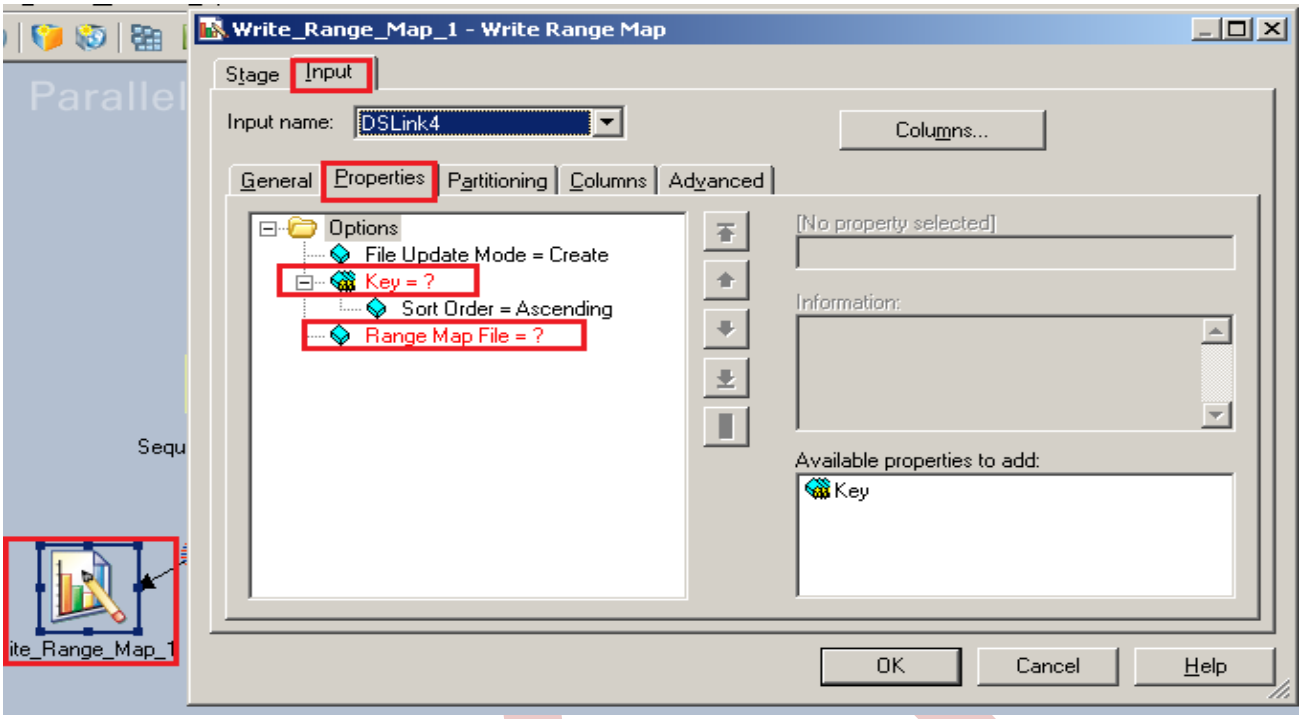
Step-16:- Click on Input and see all columns are available or not



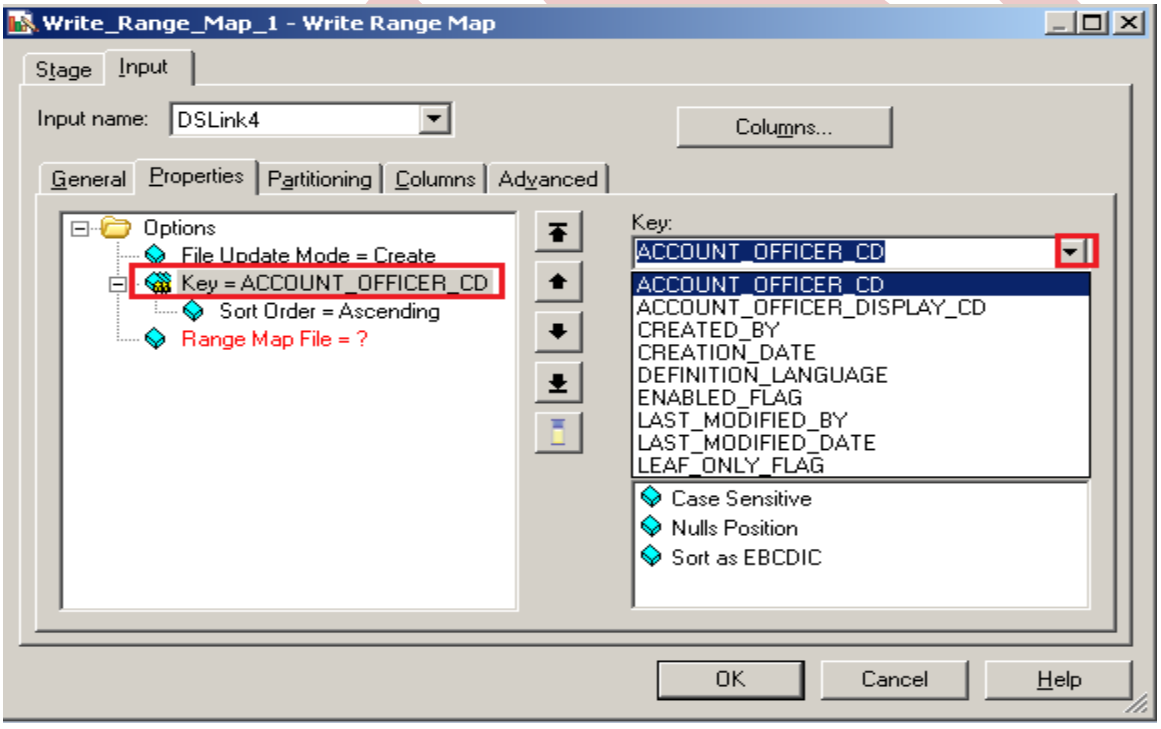
Step-17 Now go to output tab than select all links and drag them into Output Link for sending data which satisfies filter condition into output links..



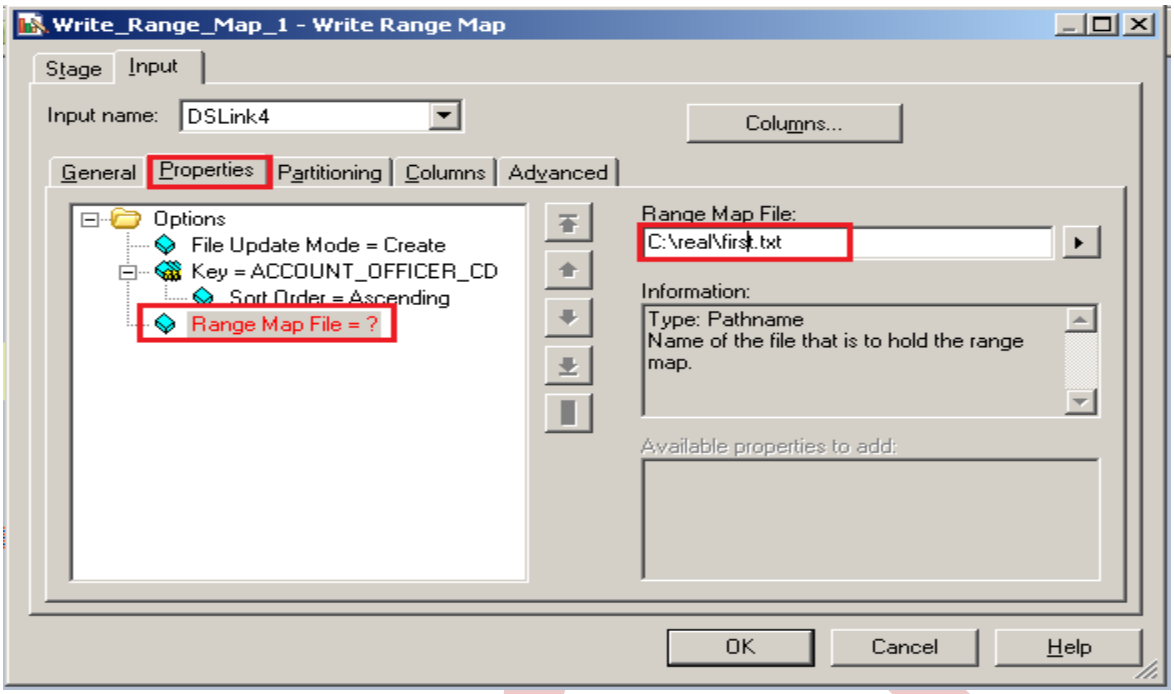
Step-18:- Now see there are multiple options available there now select key what we use in sort stage same key use and also declare output file set file. A typical use for the Write Range Map stage would be in a job which used the flat file to sort stage, the Sort stage to sort it and the Write Range Map stage to write the range map which can then be used with the range partitioning method to write the original flat file to a file set.



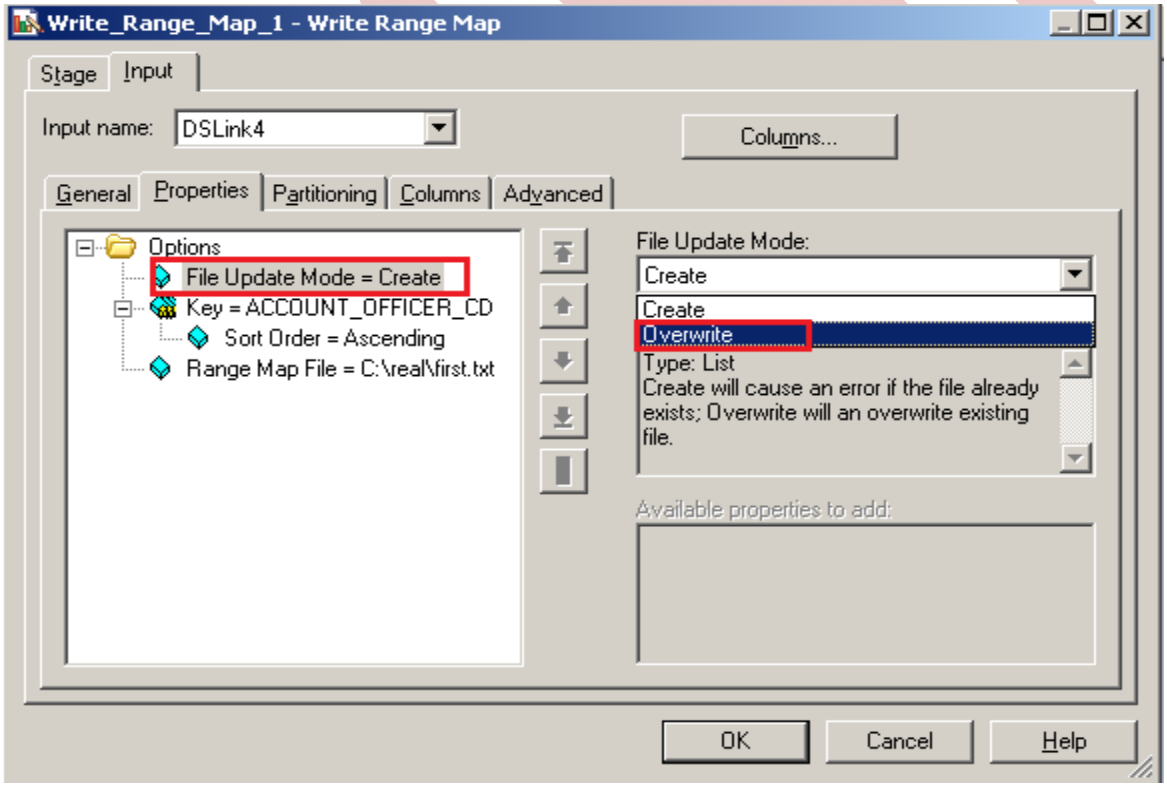
Step-19:- Now select an key value as select in sort stage and select range map file for sending data into target file.



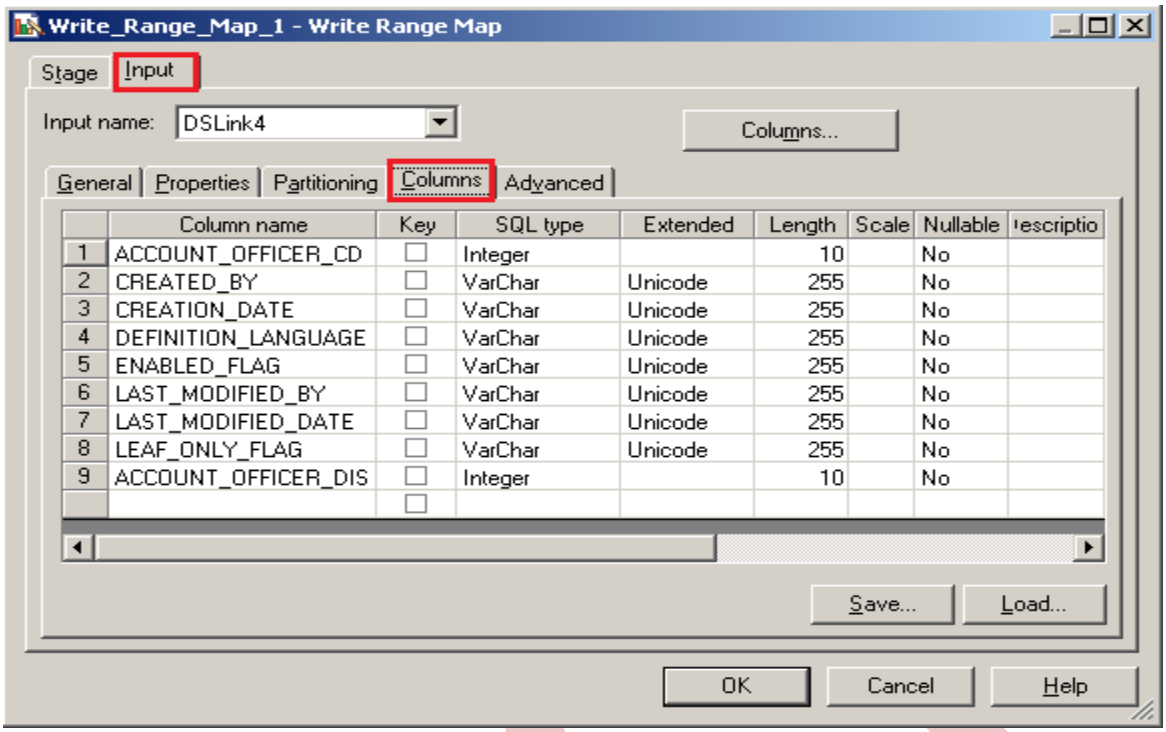
Step-20:- As I provide path for target file we have to specify target file name than we able to load data into target file.



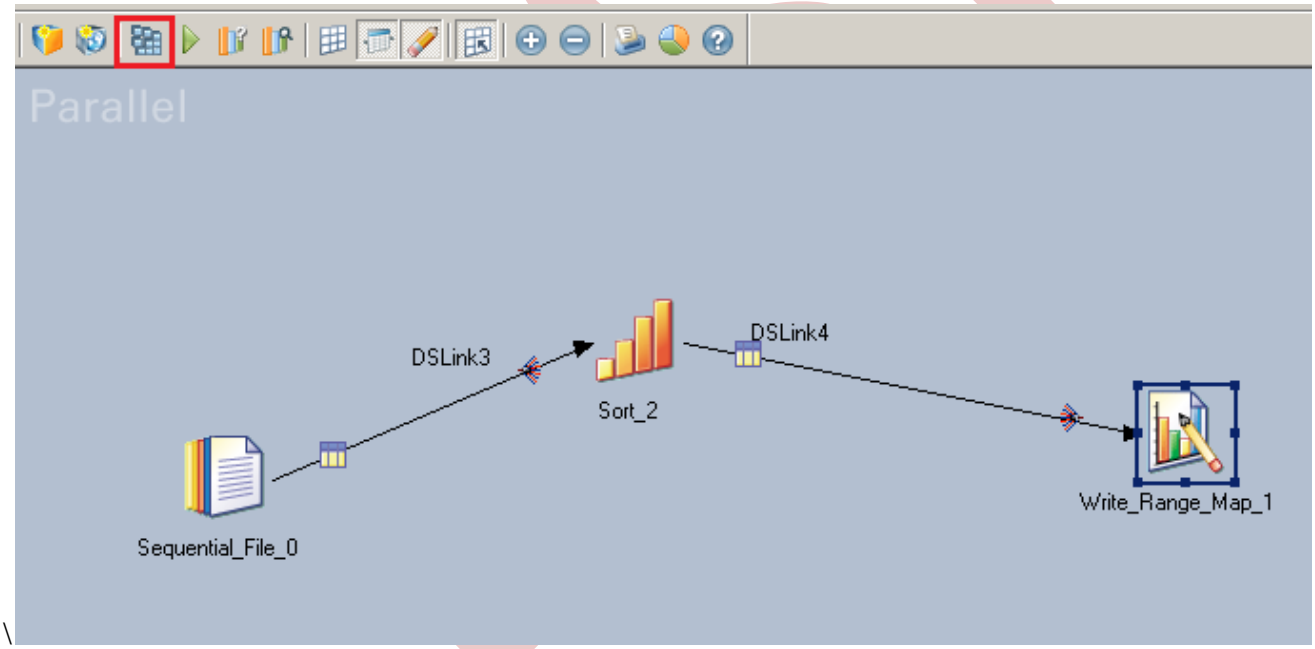
Step-21:- See here there are two option available here create or overwrite that means if create new file than we choose create otherwise if we have existing file and we want to overload than we use this file.



Step-22:- Click on column tab and see all the columns are showing or not.

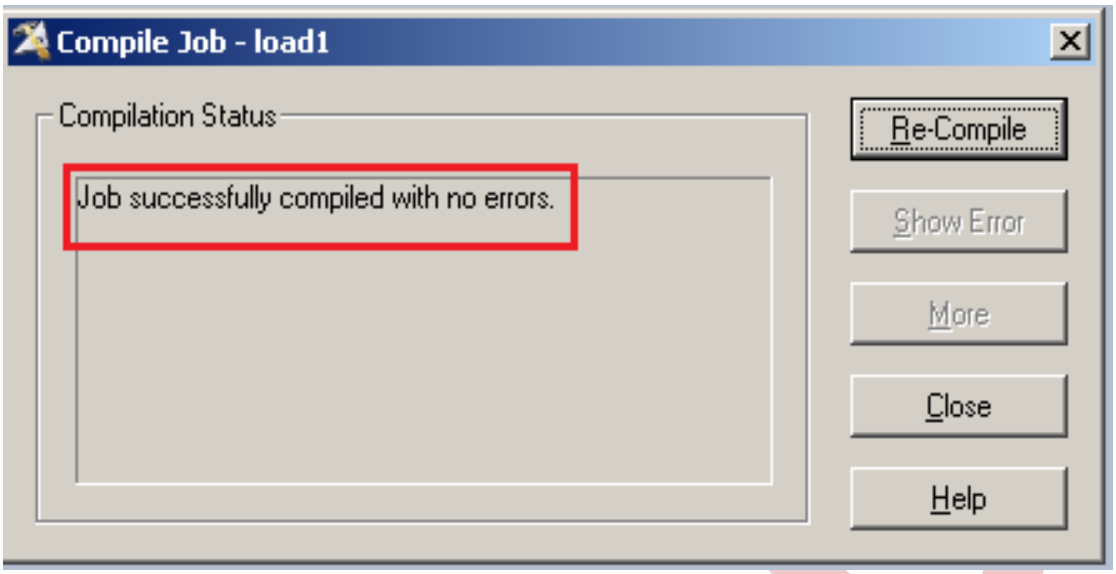


Step-23:- Now compiling the job, click on compile button.

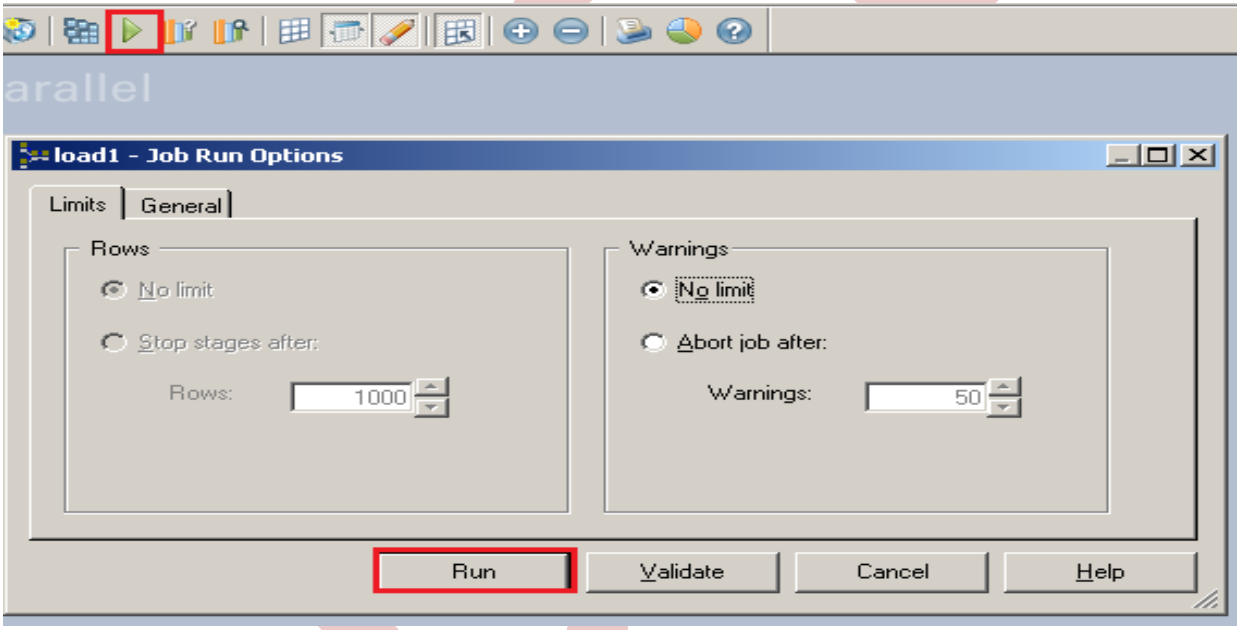


\

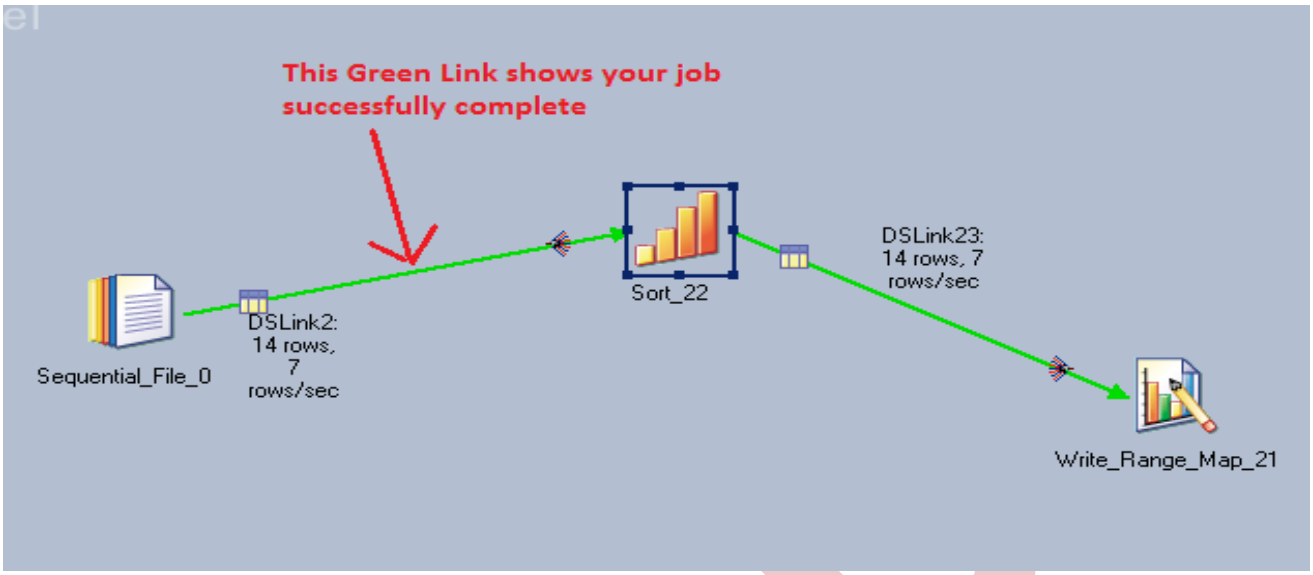
Step-24:- If this is shown that means your compilation is done otherwise it shows error.Now click on run button or (CTRL+F5).



Step-24:- Now click on Run tab and see here there are some options available here means see here there's are warning so we have select no limit of warning like this. and we can validate before run the job.



Step-25:- After click on run wait for a while than it shows GREEN line that means your tranformation is successfully done otherwise if it shows RED Line that means not Done and BLUE Line means Under Process.



Step-26:- Now Go to SQL Console and connect with the same login credentials as I mentioned in oracle enterprise stage src_creditcard/password 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:

>Occurred	>On date	Type	Event
2:46:04 PM	3/7/2013	Info	Parallel job initiated
2:46:04 PM	3/7/2013	Info	Parallel job default NLS map ASCL_MS1252, default locale OFF
2:46:05 PM	3/7/2013	Info	main_program: IBM WebSphere DataStage Enterprise Edition 8.0.1.4458 (...)
2:46:05 PM	3/7/2013	Info	main_program: orchgeneral: loaded (...)
2:46:05 PM	3/7/2013	Info	main_program: APT configuration file: C:/IBM/InformationServer/Server/Confi...
2:46:05 PM	3/7/2013	Warning	Write_Range_Map_21: When checking operator: A sequential operator cann...
2:46:05 PM	3/7/2013	Warning	Sequential_File_0,0: Delimiter for field "empno" not found; input: <empty>, at ...
2:46:05 PM	3/7/2013	Warning	Sequential_File_0,0: Import warning at record 14.
2:46:05 PM	3/7/2013	Warning	Sequential_File_0,0: Import unsuccessful at record 14.
2:46:05 PM	3/7/2013	Info	Sequential_File_0,0: Import complete; 14 records imported successfully, 1 reje...
2:46:05 PM	3/7/2013	Info	main_program: Step execution finished with status = OK.
2:46:05 PM	3/7/2013	Info	main_program: Startup time, 0:01; production run time, 0:00.
2:46:05 PM	3/7/2013	Info	Parallel job reports successful completion
2:46:05 PM	3/7/2013	Control	Finished Job write.
4:44:02 PM	3/7/2013	Control	Starting Job write.
4:44:03 PM	3/7/2013	Info	Environment variable settings: (...)
4:44:03 PM	3/7/2013	Info	Parallel job initiated
4:44:03 PM	3/7/2013	Info	Parallel job default NLS map ASCL_MS1252, default locale OFF
4:44:04 PM	3/7/2013	Info	main_program: IBM WebSphere DataStage Enterprise Edition 8.0.1.4458 (...)
4:44:04 PM	3/7/2013	Info	main_program: orchgeneral: loaded (...)
4:44:04 PM	3/7/2013	Info	main_program: APT configuration file: C:/IBM/InformationServer/Server/Confi...
4:44:06 PM	3/7/2013	Warning	Write_Range_Map_21: When checking operator: A sequential operator cann...
4:44:06 PM	3/7/2013	Warning	Sequential_File_0,0: Delimiter for field "empno" not found; input: <empty>, at ...
4:44:06 PM	3/7/2013	Warning	Sequential_File_0,0: Import warning at record 14.
4:44:06 PM	3/7/2013	Warning	Sequential_File_0,0: Import unsuccessful at record 14.
4:44:06 PM	3/7/2013	Info	Sequential_File_0,0: Import complete; 14 records imported successfully, 1 reje...
4:44:06 PM	3/7/2013	Info	main_program: Step execution finished with status = OK.
4:44:06 PM	3/7/2013	Info	main_program: Startup time, 0:02; production run time, 0:00.
4:44:06 PM	3/7/2013	Info	Parallel job reports successful completion
4:44:06 PM	3/7/2013	Control	Finished Job write.

Log for job: write — 100 entries