

Learn Datawarehousing Here

This blog remains as a handout for all warehouse & etl professionals

Datastage Scenarios

Datastage interview questions 1

We have a target source table containing 3 columns: Col1, Col2 and Col3. There is only 1 row in the table as follows:

Col1 Col2 Col3

a b c

There is target table contain only 1 column Col. Design a mapping so that the target table contains 3 rows as follows:

Col

a

b

c

Datastage interview questions 2

There is a source table that contains duplicate rows. Designs a mapping to load all the unique rows in 1 target while all the duplicate rows (only 1 occurrence) in another target.

<!--[endif]-->

Datastage interview questions 3

There is a source table containing 2 columns Col1 and Col2 with data as follows:

Col1 Col2

--- ---

a l

b p

a m

a n

b q

x y

Design a mapping to load a target table with following values from the above mentioned source:

Datawarehousing Concepts

► 2014 (35)

▼ 2013 (86)

► November (35)

▼ October (51)

[Types of ETL testing](#)

[How to run datastage Job from Unix](#)

[Datastage Scenarios](#)

[Orchadmin in Datastage](#)

[Datstage Common Errors and Solutions](#)

[Business Intelligence Process](#)

[Datastage 8.5 v certification Dump](#)

[Datastage User Manual](#)

[Order of execution of SQL](#)

[Datawarehousing in 4 stages A conventional metod](#)

[What are Dimensions and Measures?](#)

[Subqueries in SQL](#)

[Few Complex SQL Queries](#)

[Benefits of implementing a perfect DWH](#)

[Codds rules](#)

[Basic Practical Perl Tutorial](#)

[Shell Scripting](#)

[Datastage 8.7 Sample job creation](#)

[Newfeatures in Datastage 8.5](#)

[AWK programming](#)

[Teradata Utilities](#)

[Real time Usage of Grep](#)

[Regular expressions in Unix](#)

[Unix Architecture](#)

[VI Editor in Unix](#)

[Introduction to unix and basics](#)

[Tortoise Subversion](#)

[etl & data integration](#)

[SQL PERFORMANCE TUNING](#)

[Teradata Introduction](#)

[Execution of a sql](#)

[String Functions in SQL](#)

[How do Joins work in SQL?](#)

[Datastage Sample parallel Job creation](#)

[Parallell processing in datastage](#)

Col1 Col2

— — — —

a l, m, n

b p, q

x y

Datastage interview questions 4

Design an Datastage mapping to load first half records to 1 target while other half records to a separate target.

Datastage interview questions 5

A source table contains emp_name and salary columns. Develop an Datastage mapping to load all records with 5th highest salary into the target table.

Datastage interview questions 6

Let's say I have more than have record in source table and I have 3 destination table A,B,C. I have to insert first 1 to 10 records in A then 11 to 20 in B and 21 to 30 in C.

Then again from 31 to 40 in A, 41 to 50 in B and 51 to 60 in C.....So on up to last record.

Datastage interview questions 6

Validation rules for connecting transformations in Datastage ?

Datastage interview questions 7

Source is a flat file and want to load unique and duplicate records separately into two separate targets; right??

Datastage interview questions 8

Input file

10

10

10

20

20

30

output file

1

2

3

Refrential Integrity

Normalisation in a Database

Types of Data

Data Dictionary of a warehouse

HLD OVERVIEW

Incremental loading in a Warehouse

OLTP vs OLAP

Staging area in a Datawarehouse

What is an ETL Tool?

ER MODELLING

Facts & Dimensions

Snow Flake Schema

What is star schema?

Why you need datamart?

Slowly changing dimension

High level Overview of
datawarehousing

About Me



**ASHOK
KUMAR
Chindiri**

Follow 131

[View my complete profile](#)

Followers

Join this site
with Google Friend
Connect

Members (8)

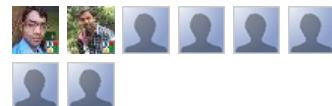


Already a member?
[Sign in](#)

Followers

Join this site
with Google Friend Connect

Members (8)



Already a member? [Sign in](#)

1

2

1

scenario-it will count the no of records for example in this above case first 10 is there so it will count 1,den again 10 is there so it will be 2, when 20 comes it will be 1 again.

Datastage interview questions 9

Input file

10

10

10

20

20

30

output file

1

2

3

4

5

6

Datastage interview questions 10

Input file

10

10

10

20

20

30

output file

----->

1

1

1

2

2

3

Datastage interview questions 11

There are 2 tables(input table)

table aa table bb

id name id name

-- ---- --

101 ramesh 106 harish

102 shyam 103 hari

103 --- 104 ram

104 ---

output file

id name

-- --

101 ramesh

102 shyam

103 hari

104 ram

Datastage interview questions 12

There are 2 tables(input table)

table aa table bb

id name id name

-- ---- --

101 ramesh 106 harish

102 shyam 103 hari

103 --- 104 ram

104 ---

output file

id name

-- --

101 ramesh

102 shyam

103 hari

104 ram

Datastage interview questions 12

table aa(input file)

id name

-- --

10 aa

10 bb

10 cc

20 aa

20 bb

30 aa

Output

id name1 name2 name3

- - - - -

10 aa bb cc

20 aa bb -

30 aa - -

Datastage interview questions 14

table aa(input file)

id name

--

10 a

10 b

10 c

20 d

20 e

output

id name

--

10 abc

20 de

Datastage interview questions 15

In the below scenario how can I split the row into multiple depending on date range?

The source rows are as

ID Value from _date(mm/dd) To _date(mm/dd)

1 \$10 1/2 1/3

2 \$5 1/5 1/8

3 \$20 1/9 1/11

The target should be

ID Value Date

1 \$10 1/2

1 \$10 1/3

2 \$5 1/5

2 \$5 1/6

2 \$5 1/7

2 \$5 1/8

3 \$20 1/9

3 \$20 1/10

3 \$20 1/11

What is the Datastage solution?

Datastage interview questions 16

How is the following be achieved with single Datastage Mapping.

* If the Header table has error value or no value (NULL) then

those records and their corresponding child records in the

SUBHEADER and DETAIL tables should be rejected from the target
(TARGET1,TARGET 2 or TARGET3).

* If the HEADER table record is valid, but the SUBHEADER or

DETAIL table record has an error value (NULL) then the no

data should be loaded into either of the target TARGET1,TARGET 2 or TARGET3.

* If the HEADER table record is valid and the SUBHEADER or DETAIL table record also
has valid records only then the

data should be loaded into the target TARGET1,TARGET 2 and TARGET3.

HEADER

C1 C2 C3 C4 C5 C6

1 ABC null null C1

2 ECI 756 CENTRAL TUBE C2

3 GTH 567 PINCDE C3

SUBHEADER

C1 C2 C3 C4 C5 C6

1 01001 VALUE3 748 543

1 01002 VALUE4 33 22

1 01003 VALUE6 23 11

2 02001 AAP1 334 443

2 02002 AAP2 44 22

3 03001 RADAR2 null 33

3 03002 RADAR3 null 234

3 03003 RADAR4 83 31

DETAIL

C1 C2 C3 C4 C5 C6

1 D01 TXXD2 748 543

1 D02 TXXD3 33 22

1 D03 TXXD4 23 11

2 D01 PXXD2 56 224

2 D02 PXXD3 666 332

TARGET1

2 XYZ 756 CENTRALTUBE CITY2

TARGET2

2 02001 AAP1 334 443

2 02002 AAP2 44 22

TARGET3

2 D01 PXXD2 56 224

2 D02 PXXD3 666 332

Datastage interview questions 17

If i had source like unique & duplicate records like 1,1,2,3,3,4 then i want load unique records in one target like 2,4 and i want load duplicate records like 1,1,3,3

Datastage interview questions 18

I Have 100 Records in a relational table and i want to load the record in 3 targets , first records goes to target 1 and second to target 2 and third to target 3 and so on ,what are the tx used in this.

Datastage interview questions 19

There are three columns empid, salmonth, sal contains the values

101,january,1000

101 february 1000 ...

like twelve rows are there then my required out put is like contains 13 columns empid jan feb march dec and the values are 101 1000, 1000, 1000 etc

Datastage interview questions 20

I have a source either file or db table

Eno ename sal dept

0101 Max 100 1

0102 steve 200 2

0103 Alex 300 3

0104 Sean 76 1

0105 swaroop 120 2

Want to run a session 3 times.

First time: it should populate dept 1

Second time: dept 2 only

Third time: dept 3 only

How can we do this?

[Newer Post](#)

[Home](#)

[Older Post](#)