Slowly Changing Dimension - Transformation

**Pre requisites in Destination table:**

1) Destination table MUST has Surrogate Key.

2) Destination table should either have Start Date and End Data as the fields OR a single column which shows the current and expired rows.

Use Zensar\_DB

go

Drop Table Employee\_Main\_Table

go

Create Table Employee\_Main\_Table

(Empno Varchar(3) Primary Key,

EName Varchar(20),

Location Varchar(20)

)

go

Insert Into Employee\_Main\_Table Values('E1', 'Smith', 'Delhi')

Insert Into Employee\_Main\_Table Values('E2', 'Martin', 'Chennai')

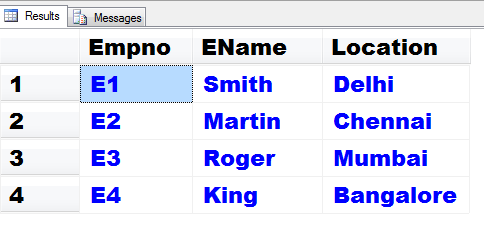
Insert Into Employee\_Main\_Table Values('E3', 'Roger', 'Mumbai')

Insert Into Employee\_Main\_Table Values('E4', 'King', 'Bangalore')

go

Select \* from Employee\_Main\_Table

Go



**Scenario 1: Historical Attribute Options is Use Start and End Dates to identify current and expired row.**

Drop Table Employee\_History\_Table

go

Create Table Employee\_History\_Table

(

Empno Varchar(3),

EName Varchar(20),

Location Varchar(20),

**EmpSK** Integer **Identity**,

**StartDate** Date Default '10-Jun-2010',

**EndDate** Date)

go

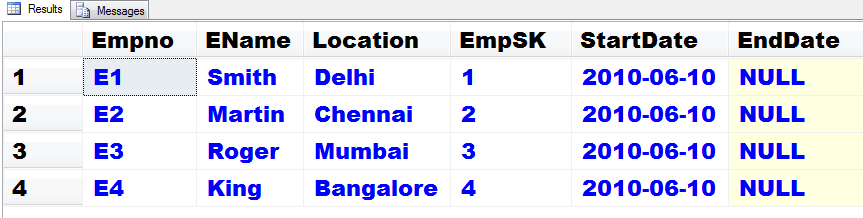
Insert Into Employee\_History\_Table (Empno, EName, Location)

Select \* from Employee\_Main\_Table

Go

Select \* from Employee\_History\_Table

Go

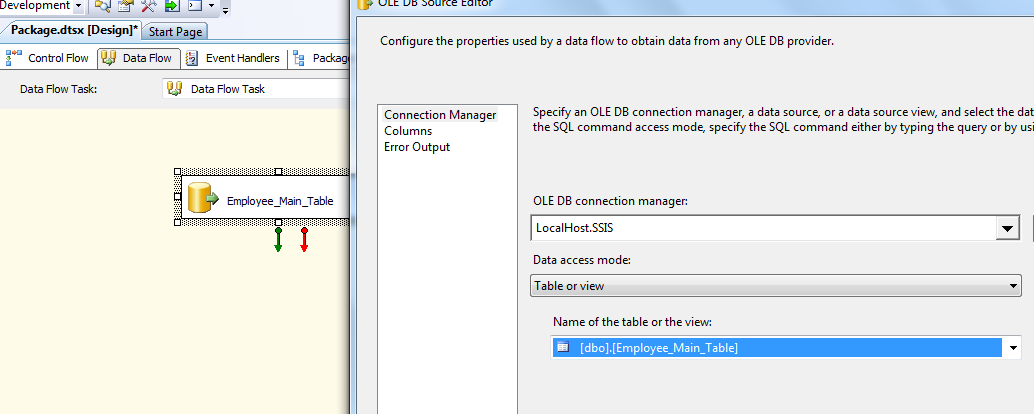


**Employee\_Main\_Table is the Type 1 table (RDBMS table) where in changes will be overridden.**

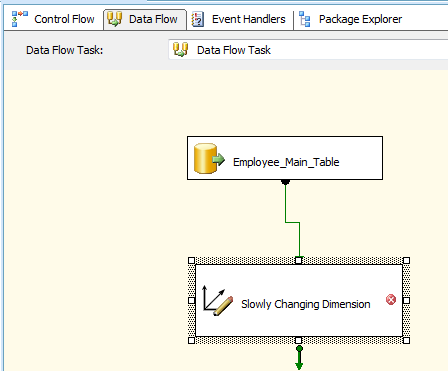
**Employee\_History\_Table will be the Type2 (datawarehouse table) where in changes will be entered as a new record with a unique identifier (Surrogate Key) along with Start and End Dates.**

In SSIS 🡪

Inside the Data Flow OLE DB Source will be Employee\_Main\_Table.

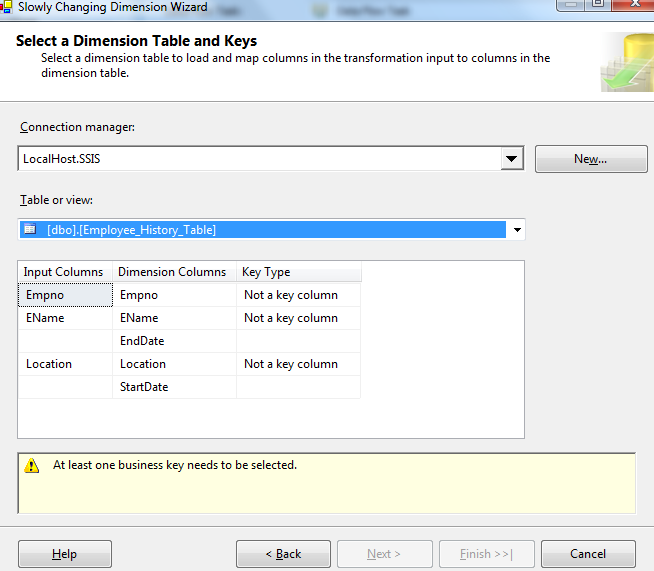


Put **Slowly Changing Dimension** transformation from DFT and link it with the OLE DB Source.

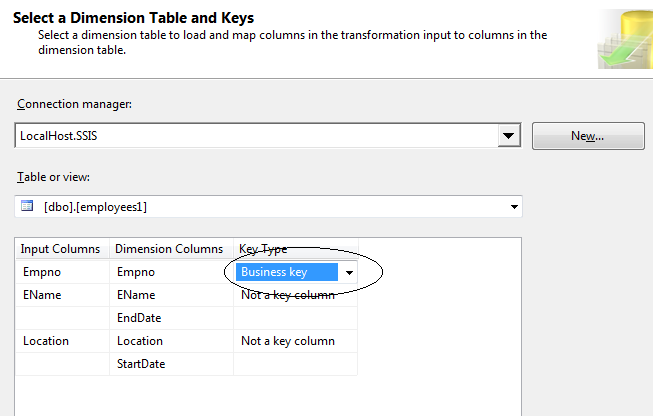


Double click on the SCD, **wizard starts!!!!**

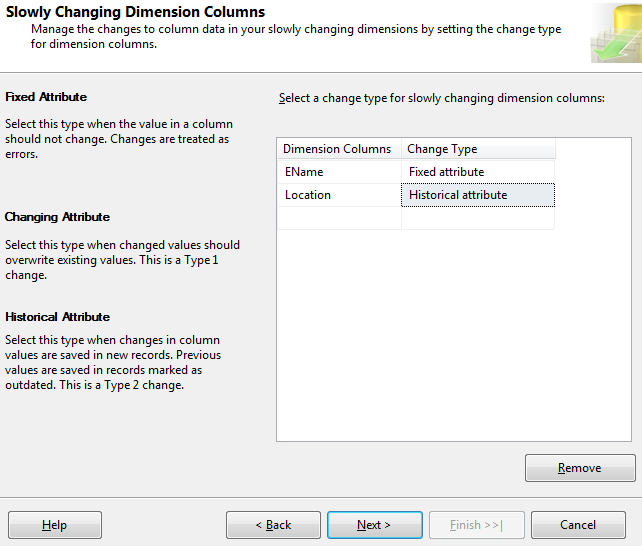
In the **Select a Dimension Table and Keys** step select Employee\_History\_Table as shown next –



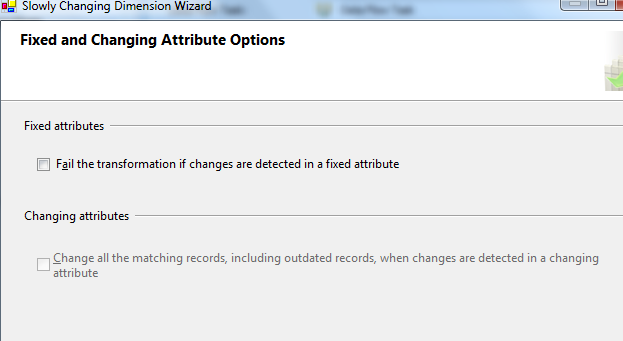
Select **Empno** as the **Business Key.**



In the **Slowly Changing Dimension Columns** step select Ename as the Fixed Attribute and Location as the Historical Attribute as shown next –

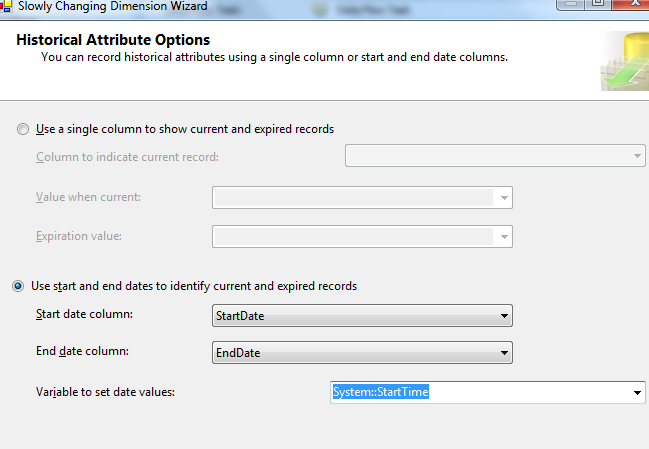


In the Fixed and Changing Attribute Options step remove the selection of Fail the transformation if changes are detected in a fixed attribute.

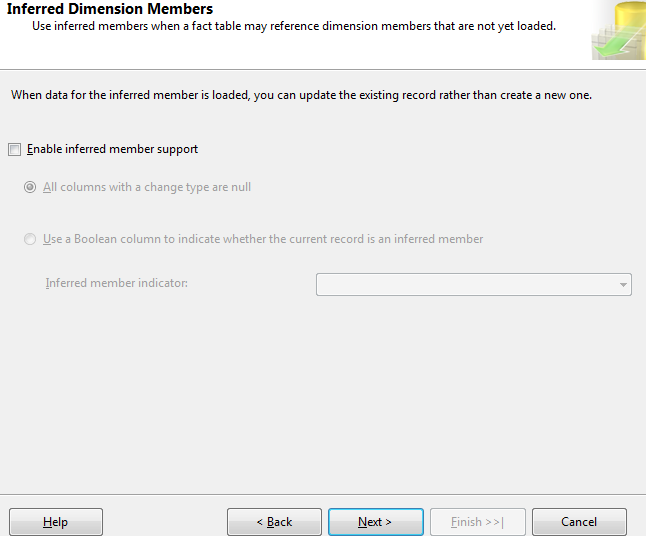


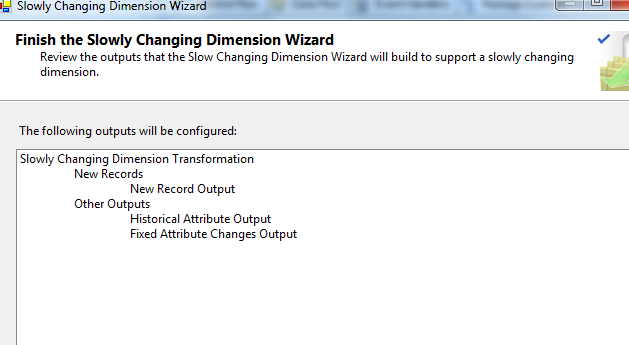
In Historical Attribute Options step select the 2nd radio button of Start and End Date and select Start date column as **StartDate** and End date column as **EndDate**.

In Variables to set date values select **System::StartTime**

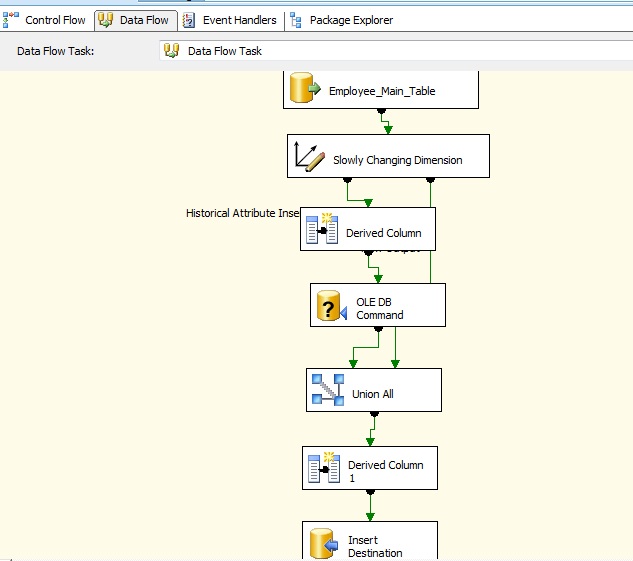


In the Inferred Dimension Members step remove the selection of Enable inferred member support.





After finishing the wizard the Data Flow gets further populated!



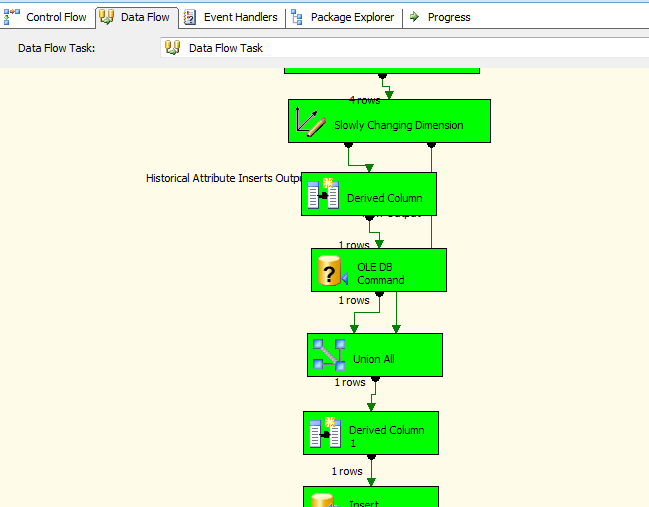
Now in **SSMS** change the location of Smith from Delhi to Mumbai in the Employee\_Main\_Table.

Update Employee\_Main\_Table

Set Location = 'Mumbai'

Where Empno = 'E1'

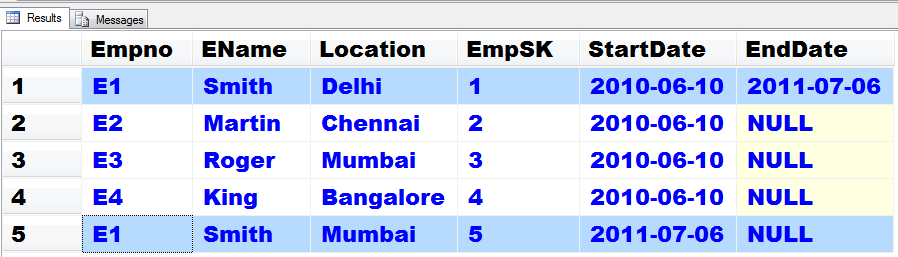
**In SSIS run the package.**



Check the Employee\_history\_table in SSMS.

Select \* from Employee\_History\_Table

Go



It means that a new record for Smith has been added with a new surrogate key value as 5, keeping start date as the date of update. End date of the first Smith record is entered as the date of update.

**The null end date for the new Smith record indicates that the record is active.**

**Scenario 2: Historical Attribute Options is Use Single Column Demo to show current and expired rows.**

Use Zensar\_DB

go

Drop Table Employee\_Main\_Table

go

Create Table Employee\_Main\_Table

(Empno Varchar(3) Primary Key,

EName Varchar(20),

Location Varchar(20)

)

go

Insert Into Employee\_Main\_Table Values('E1', 'Smith', 'Delhi')

Insert Into Employee\_Main\_Table Values('E2', 'Martin', 'Chennai')

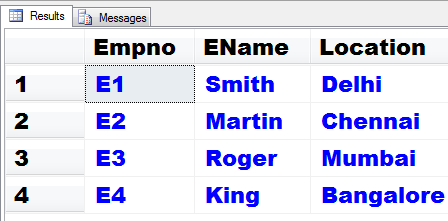
Insert Into Employee\_Main\_Table Values('E3', 'Roger', 'Mumbai')

Insert Into Employee\_Main\_Table Values('E4', 'King', 'Bangalore')

Go

Select \* from Employee\_Main\_Table

Go



Drop Table Employee\_History\_Table

go

Create Table Employee\_History\_Table

(

Empno Varchar(3),

EName Varchar(20),

Location Varchar(20),

EmpSK Integer Identity,

**Status Varchar(15)** **Default 'Current'**

)

go

Insert Into Employee\_History\_Table (Empno, EName, Location)

Select \* from Employee\_Main\_Table

Go

Select \* from Employee\_History\_Table

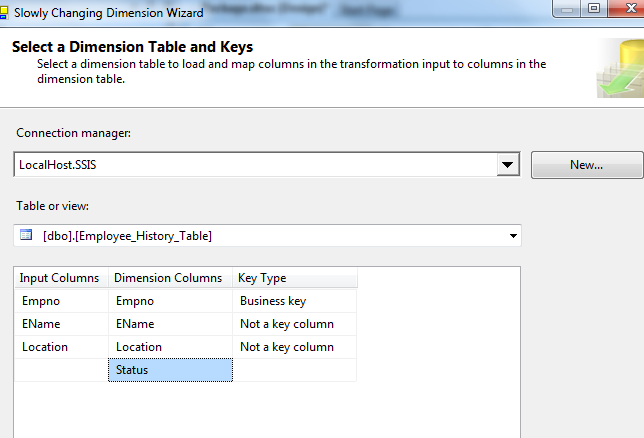
Go



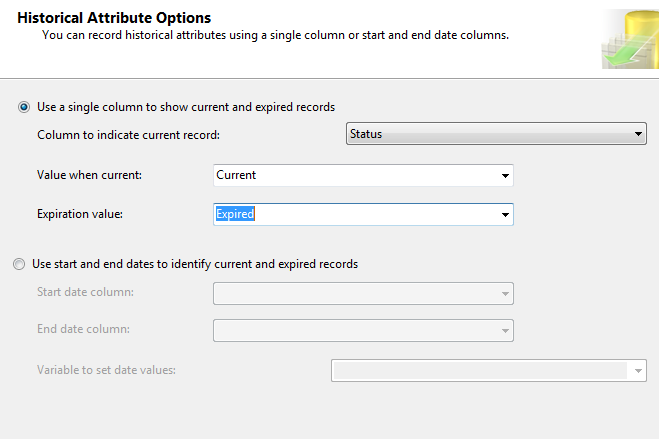
So there are no Start Date and End Date columns. Instead of that a single column status is used.

In SSIS the initial steps are same.

OLE DB source, starting steps wizard of SCD are same.



**Important Step** 🡪



After finishing the wizard, in SSMS update Smith’s location as done before.

Update Employee\_Main\_Table

Set Location = 'Mumbai'

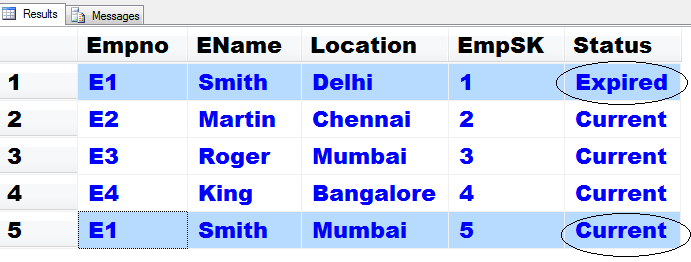
Where Empno = 'E1'

Run the package.

In SSMS,

Select \* from Employee\_History\_Table

Go



**Slowly Changing Dimension Columns – Location as a Changing Column instead of Historical Column**

Use Zensar\_DB

go

Drop Table Employee\_Main\_Table

go

Create Table Employee\_Main\_Table

(Empno Varchar(3) Primary Key,

EName Varchar(20),

Location Varchar(20)

)

go

Insert Into Employee\_Main\_Table Values('E1', 'Smith', 'Delhi')

Insert Into Employee\_Main\_Table Values('E2', 'Martin', 'Chennai')

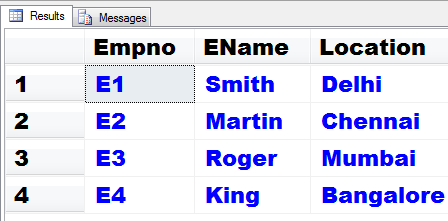
Insert Into Employee\_Main\_Table Values('E3', 'Roger', 'Mumbai')

Insert Into Employee\_Main\_Table Values('E4', 'King', 'Bangalore')

Go

Select \* from Employee\_Main\_Table

Go



Drop Table Employee\_History\_Table

go

Create Table Employee\_History\_Table

(

Empno Varchar(3),

EName Varchar(20),

Location Varchar(20),

EmpSK Integer Identity,

Status Varchar(15) **Default 'Current'**

)

go

Insert Into Employee\_History\_Table (Empno, EName, Location)

Select \* from Employee\_Main\_Table

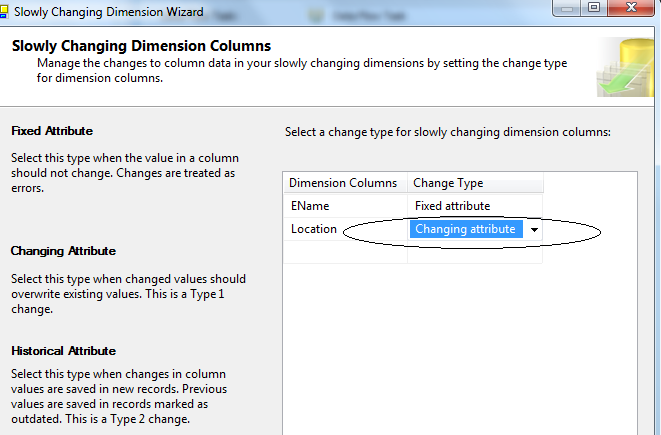
Go

Select \* from Employee\_History\_Table

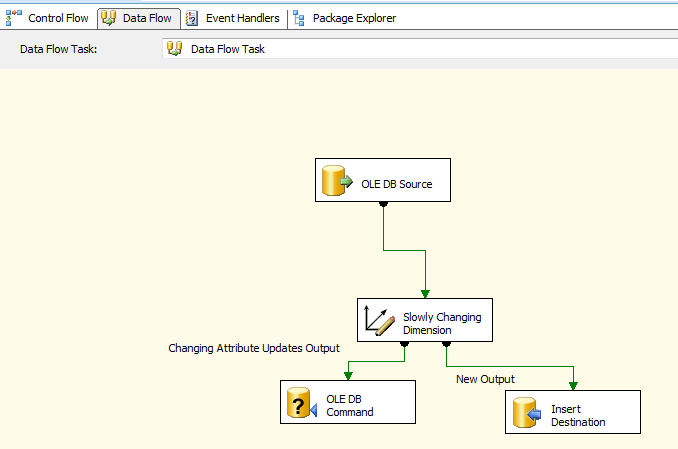
Go



In The SCD wizard all steps are same except the following changed step.



The Data Flow Task gets a different look; a simpler model comes up.



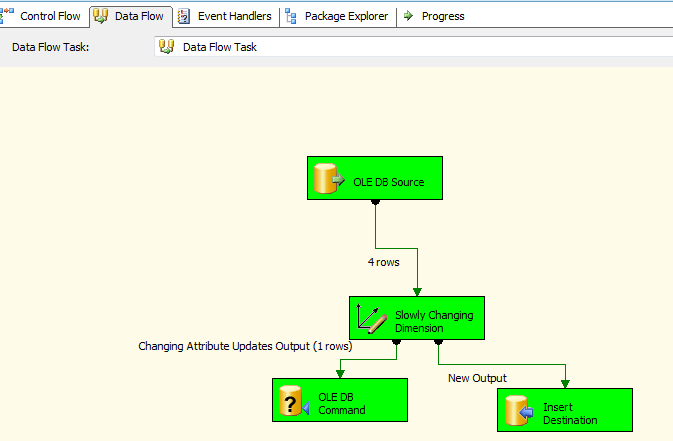
After finishing the wizard, in SSMS update Smith’s location as done before.

Update Employee\_Main\_Table

Set Location = 'Mumbai'

Where Empno = 'E1'

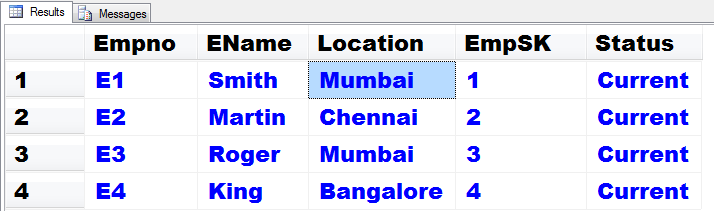
Run the package.



In SSMS,

Select \* from Employee\_History\_Table

Go



The location of Smith got changes there itself. Here surrogate key is meaningless.