@@@@@ EXP 2- DATA PROFILING @@@@@

1) Open Fateor Microsoft Visual Studio (Administrator) [∞]

2)Create New project

a) file → new project

b) Select Integration type of service

3) Data profiling

a) In SSIS toolbox (left side)

b) Drag and drop to Data profiling task.

c)double click on the box. One page of editor will be appear

d) In destination select new file.

e) dialog dial. box of file connection manager

usage type : Create file

file: filename.xml

f) click ok

4) Go to quick profile (You have to give connecting

a) ADO.NET connection new dialague bar

1) Select one server from it INDIA-POXMSSQLSERVER2.AdventureWorks2012

b) Table Select (Person 7

c) In compute:

Select first 4 catbox only

d) ok (Hill first page opens).

5)Execute (check click)

6) Double click on box Data profile Viewer opened.

7) Select open profile viewer.

@@@@@ EXP 4- SIMPLE ETL TRANSFORMATION @@@@@

(Simple ETL transformation using SQL server Integration)(convert lower case to upper case)

[Open Adventure works 2012 microsoft SQL Server management studio- open dbo.stud table] (If not available then create by right clicking on the Tabes in AdventureWorks2012 and save as dbo.stud (Save the name in right corner) stud_id-int and name-nvarchar(50))

0) Open notepad

1) Create s1.csv as

Stud_Id,Name

1,xyz

2,abc

2) Open SQL Server Data tools (infinity Symbol)

-go to File->New->Project

3) On the left side from installed templates select Integration Services (Select Integration services Project) select ok.

4) From the left side (SSIS Toolbox)

-drag data flow tasks and drop it in control flow and double click on it.
5) From the left side (SSIS Toolbox)
-drag flat file sources and drop it in data flow tab then double click on it
-it opens flat file source editor in that click on new
- it opens flat files connection manager editor in that for file name click on Browse and select s1.csv file click on preview and then click ok.
6) From left side drag and drop data conversion
-connect flat file source to data conversion box.
7) In input column,
-for Stud_Id, Change output Alias as Data_converted_Stud also change datatype to four bite signed integer[Dt_I4].
-for name change output Alis as Data_converted_name and data type as Unicode string[DT_WSTR].
Click OK.
8) Drag and drop derived column from left side.
9) Connect data conversion box to derived column box. It opens up derived column transformation editor.
10) In that derived column name type Derived_Stud_ID .
11) On the upper left side expand the columns and drag Data_Converted_Stud_ID and drop it in Expression. Data type would be same just click on it.
12) In 2nd row type Derived_Stud_Name.
-from upper right side from string functions drag and drop UPPER function in the expression column and change it as UPPER([Data_Converted_Name]). Data type would be same just click on it.
13) Click on Ok.
14) From left side drag and drop OLE DB destination connect Data Conversion box with this Box.
15) double click on it,
-Configure OLE DB Connection manager tab opens click on new
-In data connection there is already Adventure Works 2012.
-click ok.
16) Select the table
-select [dbo].[stud]
-go to mappings and delete existing connections(lines). Connect derived Stud_ID to the Stud_Id and Derived_Stud_Name to the name.
click ok -> save.
17) click on run icon on top (green icon).
18) open dbo.stud
-right click on it click show top 1000 rows.

@@@@@ EXP 5- LOOKUP TRANSFORMATION IN ETL @@@@@

EXP 5
(LOOKUP DATA TRANSFORMATION IN ETL)

1)open Microsoft sql server management studio.
databases->Adventureworks2012->tables->dbo.books
2)delete dbo.books and dbo.bookhistory if already present
3)from left side select adventureworks 2012
-right click on it select new query
-paste the query

```
Create table dbo.books
(
  BookID int primary key clustered,
       [Book Title] NVARCHAR(100),
  [Book Price] money
```

```
            )
      -click on execute(red button)
4) from left side select adventureworks 2012
      -right click on it select new query
      -paste the query
                  Insert into dbo.books values
               (1,'SQL World',560),
               (2,'Oracle overview',780),
               (3,'Explore SSIS',458),
               (4,'Data Visualization',679),
               (5,'SQL Components',353),
               (6,'High Availability',781)

      -click on execute(red button)
5) from left side select adventureworks 2012
      -right click on it select new query
      -paste the query
                  Create table dbo.bookshistory
        (
            BookID int primary key clustered,
            [Book Title] nvarchar(100),
            [Book Price] money
        )


      -click on execute(red button)
6) Open microsoft visual studio (infinity symbol)
7) new->project
8)select type as integration services
      -rename it as Lookuptransformationdemo
9)drag and drop dataflow task in the control flow tab and double click on it.
10)drag and drop OLE DB source and double click on it.
      in OLE DB source editor
            - in OLE DB connection manager
                  -go to new->select AdventureWorks2012
            click ok.
            -in name of the table or view
                  -select [dbo].[books]
            click ok.
11)drag and drop lookup and connect OLE DB source box to lookup box
12)in lookup transformation editor
      -in specify how to handle rows with no matching entities select redirect rows
to match output.
13)on the left side select connection
      -in use a table or view select [dbo].[bookshistory]
14)on left side select column
      -connect BookID to BookID.
   click ok.
15) from SSIS toolbox
      -drag and drop OLE DB destination
      -connect lookup box to OLE DB destination
16) in input output selection tab
      -for output select lookup no match output
    click ok
17)double click on OLE DB destination
      -select OLE DB connection manager as adventureworks2012.
      -name of the table or view select [books].[bookshistory]
      -check for mappings are present
      -click ok
```

18) RUN
19) stop execution
20) repeat
---
21) select dbo.books from microsoft sql server management studio
      -change 560-500 from 1st record
      -save
22) In microsoft visual studio (infinity symbol)
      -drag and drop OLE DB command
      -connect LOOKUP to OLE DB command
      -double click on OLE DB command
      -select the connection manager
      -goto common properties->sql command
      -paste query
      UPDATE [AdventureWorks2012].[dbo].[bookshistory]
               SET
                   [Book Title] = ?,
                   [Book Price] = ?
             WHERE [BookID] = ?;
      -ok
      -goto column mappings and map
                        BookID - param2
                        Book title - param0
                        Book price - param1
      -select ok
      -run
      -stop
23) follow step 3
 - paste the query
 - Insert into dbo.books values
             (7,'The role of DBA',872)
 -run again

      @@@@@ EXP 6 -SLOWLY CHANGING DIMENSIONS USING SSIS @@@@@
1)Go to SQL management server -> Database -> Right click on AdventureWorks2012-> new query

```
Create table Test_Stage
(
Emp_Id int,
Emp_Name Varchar(100),
Country Varchar(100)
)
Create table Test
(
Emp_Id int,
Emp_Name Varchar(100),
Country Varchar(100),
Start_Date Datetime,
End_Date DateTime
)
```

2)Click on Execute and after successful addition of new table close the query.

3)Again right new query to insert the values and do the same.

```
Insert into test_stage Values(1,’Roopesh’,’India’);
```

```
Insert into test_stage Values(2,'Lokesh','UK');
Insert into test_stage Values(3,'Vinay','USA');
Insert into test_stage Values(4,'Rakesh','China');
Insert into test_stage Values(5,'Venki','Japan');


4) go to sql server data tools visual studio
5) go to file --> select project --> Integration services projet(inside the
business Intelligence) --> ok
6) click control flow tab
      i) drag and drop data flow task and double click
      ii) drag and drop OLE DB source and click
            - In OLE DB connection maneger-->create new
            - copy the server name and paste and database sourse is
adventureWork2012-->ok
            -select test stage table-->ok
      iii)drag and drop slowly changing dimension transformation-->connect the OLE
DB arrow to this-->click on slowly changing dimension
            -right click on next
            -select test table-->make emp_id is business key-->next
            -dimensional colums--> historical attribute for contry --> next
            -click use start and end...-->select start_date and end_date-->select
System containerstart time for variable to set date values--click next
            -click next--> finish
7)save it
8)connect server
9)run
10)go to database-->refresh tables-->test_stage-->select top 100 and test table --
>selct top 100(both have same records)
11)go to AdventureWork2012-->new query-->run query(Update test_stage set
country='USA' where emp_id=1)
12)again run the transformation
Note: test table is destination and test_stage is source table


       @@@@@ EXP 7 - EXTRACTING AND UPLOADING DATA @@@@@@
1) Open Microsof Visual Studio

 New → project (Integration service type)

2) Select data flow task from SSIS toolbox

3)Double click on data flow task
      a) Select Flat file Source (left side)
      b) Select new

      (Previous: create 2 csv Files s1.csv

            Stud-id name
            1     xyz

            2     abc

      c) Select sl.csv click ok

      d) Convert data

            a) Select Data conversion (left side)
            b) Drop arrows of FFS to DC base connect them
```

c)Select on DC Select all availaible colm
std_id -> data type change it to four-byte unsigned integer DT_[4]
name: Unicode strin [DT_WSTR]

4) Do same wuth S2.csv

std_id	name
3	pqr
4	abc

a) select flat file source & select S2.CSV

b) Select dates conversion arrange them as men above

c) select all colmns it has.
output Alais:

change: copy of stud-id2

copy of name2

d)datatypes same as previs change

5) We have to take union.

a) Select unian all (left)
b) arrange them
c) Select union All
d)in two fieds <ignore> select
- Stud-1d2
- Copy of name 2.

e)In Output column Name Change

Copy of StudId - Copy of StudId3
Copy of name - Copy of name3

f) Click ok

6)we have to
a) Select OLEDB destination
b) Connect uinion all and DB by arrow.
c) double dick.

1) connection manager (select ok)
2) table name [dbo]. [stud]

* Go to SQL server management Studio * 1) In databases 2) - Adv works 2012 3)tables 4). file tables 5) go to stud 6)Right click & edit this file first 200 6)Select all rows and delete

7)Go back to Admistrator.

a) Select ok
b) Give mapping
copy of stud_id3->studid
copy of name3->name
c)ok & execute

@@@@@ EXP 8 - CUBE ANALYSIS @@@@@

****Start Microsoft Visual Studio(Administrator)****

1) Data Source -select ADVENTUR WORKS DW 2012 5.next Select Fadinternet sales (dbo) & select
-Add related tables
- Delete FactInfan... table from selected list & next finish

2) Data Savere views.

select DW12012

3) Cubes

-Use existing tables.

- Select factInternetSales is next Select all & next & next

- Select Dim Date dim from Right side select financial Year

- Select Sales Teratory and Select Region

- Select product dim Select English Product Name

4)Select project name.

- Go to properpties.

- Go to deployment Change servername to our servername

- right click & build. After success deploy

5)Select cube & brows

-Goto measures & select sales amount

@@@@@ EXP 9-DASHBOARD CREATION @@@@@

Experiment 9 : Dashboard creation

1) go to sql server data tools
2)start sql server management studio
3)search servies-->click services (ensure the sql server started)
4)file-->new project
5)selct report server project wizard from reporting service--> ok-->minimize the window
6)selct shared date source(right side)-->add-->new item
i)click data source-->add
ii)select mysql data server for type-->edit
iii)copy the server name and paste and select the database-->AdventureWorkDW2012-->test connection-->ok-->ok-->ok
7)select shared datasets(right side)-->add-->new iteam
i)select dataset-->add
ii)type query--> SELECT Cat.[EnglishProductCategoryName] AS ProductCategory,
SubCat.[EnglishProductSubcategoryName] AS

```
ProductSubCategory,
                        Prod.EnglishProductName AS ProductName,
                        Prod.Color,
                        Fact.OrderQuantity,
                        Fact.TotalProductCost,
                        Fact.SalesAmount,
                        Fact.TaxAmt,
                        Fact.[Freight]
                  FROM FactInternetSales AS Fact
                        INNER JOIN DimProduct AS Prod
                              ON Fact.ProductKey=Prod.ProductKey
                        INNER JOIN DimProductSubCategory AS SubCat
                              ON
Prod.ProductSubCategoryKey=SubCat.ProductSubCategoryKey
                        INNER JOIN DimProductCategory AS Cat
                              ON SubCat.ProductCategoryKey=Cat.ProductCategoryKey
      iii)click ok
8)select report (right side) -->add-->new item
      i)select report-->add
      ii)in design tab-->right click-->insert-->chart-->piechart-->ok
      iii)select dataset-->ok
      iv)double click chart-->drag and drop salesAmount in values(chart data)
      v)in category group-->details-->productSubCategory
      vi)(for formatting) right click on chart-->show data lables
      vii)save and preview
9)click design tab
10)right click dataset1-->properties-->filters-->add
      i)productCategory for expression, = for operator and Bikes for value-->ok
      ii) click preview
11)in design tab-->right click-->insert-->chart-->column chart-->ok
      i)double click chart-->drag and drop salesAmount in values(chart data)
      ii)drag and drop totalProductCost in values(chart data)
      iii)drag and drop color in category group(chart data)
      vii)save and preview


      @@@@@ EXP 10 -PARAMETERIZED REPORT CREATION @@@@@
1) go to sql server data tools
2)start sql server management studio
3)search servies-->click services (ensure the sql server started)
4)file-->new project
5)selct report server project wizard from reporting service--> ok-->minimize the
window
6)selct shared date source(right side)-->add-->new item
      i)click data source-->add
      ii)select mysql data server for type-->edit
      iii)copy the server name and paste and select the database--
>AdventureWorkDW2012-->test connection-->ok-->ok-->ok
7)select shared datasets(right side)-->add-->new iteam
      i)select dataset-->add
      ii)type query->
                  SELECT Geo.[EnglishCountryRegionName] AS [Country] ,Geo.
[StateProvinceName] AS [State] ,Geo.[City] ,Cust.FirstName +' '+ Cust.LastName AS
[Full Name] ,Cust.EnglishEducation AS Education ,Cust.EnglishOccupation AS
Occupation ,SUM(Cust.YearlyIncome) AS YearlyIncome ,SUM(Fact.SalesAmount)AS
SalesAmount
                  FROM DimCustomer AS Cust INNER JOIN
                  FactInternetSales AS Fact ON
                  Cust.CustomerKey = Fact.CustomerKey INNER JOIN [DimGeography] AS
```

```
Geo ON
                  Cust.GeographyKey = Geo.GeographyKey
                  GROUP BY Geo.[EnglishCountryRegionName] ,Geo.
[StateProvinceName] ,Geo.[City] ,Cust.FirstName
                  ,Cust.LastName ,Cust.EnglishEducation ,Cust.EnglishOccupation
                  ORDER BY [Country]
```

8)Reports -> add -> new item -> Report
9)Right click -> insert -> TAble -> Dataset1 -> ok
10)country,full name, occupation, yearly income, sales
11)preview
12)Parameters -> add parameter -> name -> countryparameter -> prompt -> "enter country name" -> data type -> text ->(visible)
13)preview
14)Dataset -> dataset properties -> filters -> expressoin : country -> operator : like -> value : [@countryparameter]
15)preview