**User Manual**

Instruction to implement ETL.

1. The Folder “etl-datamart-applicant.zip” has been submitted. This should be unzipped and placed in the folder.
2. Open Pentaho kettle spoon.bat which open the Pentaho GUI. .
3. We should also open the SQLYog. Now run the 2 sql scripts provided in the folder “etl-datamart-applicant”.  
   “oltp\_admission\_dump.sql” :- Contains the Dump for the oltp\_admission.  
   “ddl\_script.sql” :- Contains the DDL scripts for Source and Datamart DB.
4. In Pentaho click on File and then Open. Select the job “main-datamart-applicant.kjb”. This Job contains the complete process of Extract, Transform and Load job/transform.
5. Once the main job is opened then click on edit and then select “Set Environment parameter” with below parameters.

ETL\_HOME\_DIR: {Insert the path where etl-datamart-applicant is placed along with the folder name”  
ETL\_USER\_NAME: (Enter the User name which is been used for the source and Database DB in SQL)  
ETL\_USER\_PASS: (Enter the password corresponding to user name which is been used for the source and Database DB in SQL)  
ETL\_DIM\_DATE\_START: 2001-01-01  
ETL\_DIM\_DATE\_END: 2020-12-31

1. Components under each job/ transformation.
   1. EXTRACT DATA FROM OLTP ADMISSION:

Under this Transform we have the input and the output to extract the Data from Table of OLTP\_ADMISSION Database to SOURCE Database.

* 1. LOAD-DIM-ALL

This is a job which contains 4 transformation for the 4 Dimension to be populated.

* + 1. LOAD-DIM-APPLICANT

Here we load the Dimensions table for applicant as the dimension only have the data from the app\_profile and adm\_decision we join the two tables using the stream lookup and load the dim\_applicant in Datamart DB.

* + 1. LOAD-DIM-PROGRAM

Here we load the Dimensions table for program as the dimension only have the data from the adm\_program and adm\_department we join the two tables using the stream lookup and load the dim\_program in Datamart DB.

* + 1. DIM\_ETHNICITY

We load the dim\_ethnicity using the the details from the app\_profile. Here we have generated a junk id by using the checksum.

* + 1. DIM\_DATE

Dim date acts as the role playing dimension. Thus it has the Dates split up as date month and year.

* 1. LOAD\_FACT\_APPLICANT\_TRANSFORM

This transformation is used to load the Fact table.

Here we initially pass the Data from the source table. We calculate the Measure from this table itself. Next we use the steam lookup in order to have the Date surrogate for the Printed date, submitted date, accessed date. Then we use the lookups again for passing the program and the applicant surrogate keys. In case of the ethnicity surrogate key we use the Database lookup and pass it on to the select value in order to pass few parameter needed for the fact tables. AT the end we use the table output which populates the FACT\_APPLICANT.

1. Run the Main job by clicking the  button and launching it.

.