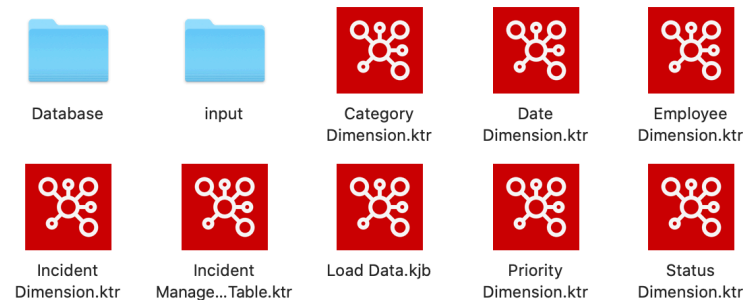# IMS ETL Technical Instructions

This is a manual to establish the technical environment in order to complete the ETL process, details will be explained as following, Downloaded Files, Technical Requirements, and Steps.

1. Downloaded files

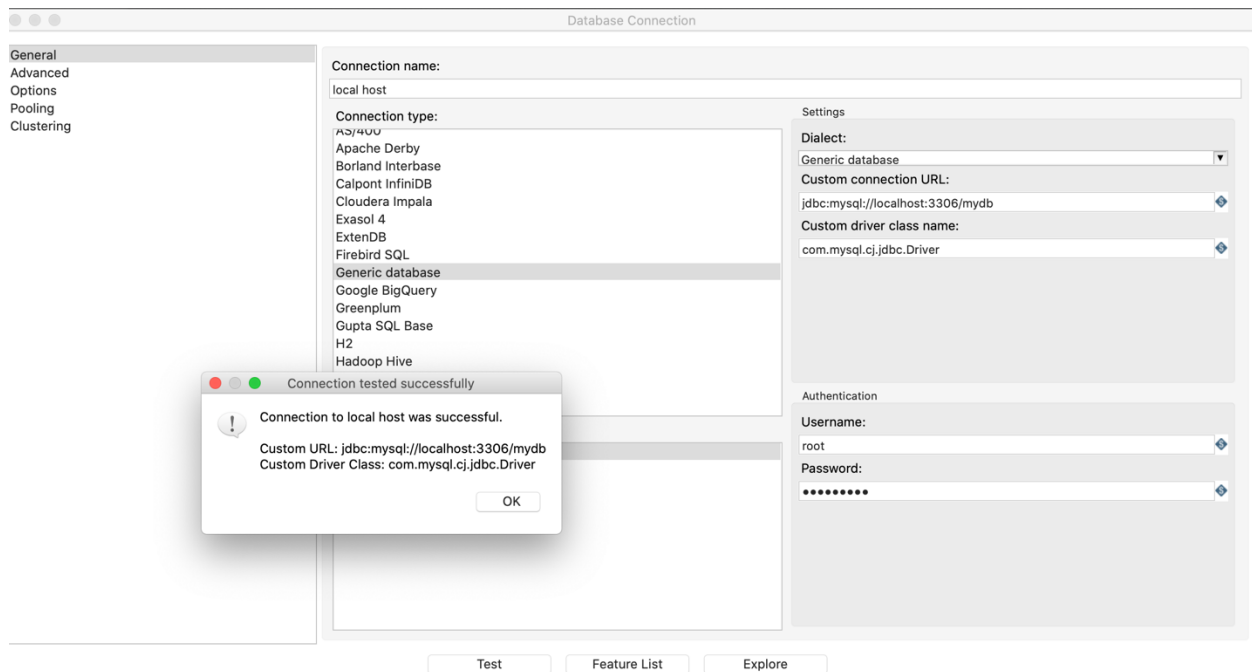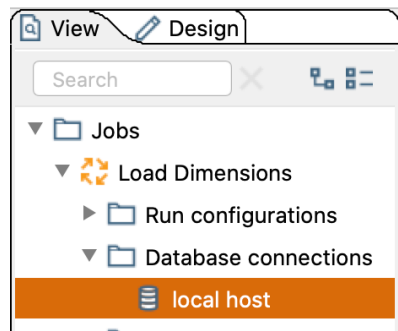- All the files mentioned are provided in a zip file, shown as below:

  o *__Database Folder__* :
    - IMS Star Schema Group E.mwb,
    - IMS_LOADdata.sql (backup of our database)
  o *__Input Folder__* (Original Dataset): incident_event_log.csv;
  o *__The set of transformation__* includes 7 transformations and 1 job linking the entire ETL process.

2. Technical requirements:

For a MacBookAir 1.6Ghz Intel Core i5 with 8GB RAM, the ETL for 6 dimensions takes an average of ***7 min - 8 min***, and ETL for the fact table takes ***11 min 50 Sec.*** In total, the ETL process takes ***19 min -20 min***.

3. To run the ETL process with Schema

1. **To create mydb in local database**: Open IMS Star Schema Group E.mwb with MYSQL WORKBENCH.
    o Database→ Forward Engineer
2. **Open Load Data.kjb in Pentaho**
    o Double check whether the your MYSQL database local host is connected with Pentaho.
        - Custom connection URL: jdbc:mysql://localhost:3306/mydb
        - Custom Driver Class Name: com.mysql.cj.jdbc.Driver
        - Your username and password.

3. Run Load Data.kjb in **Pentaho** to run the **entire ETL process**.