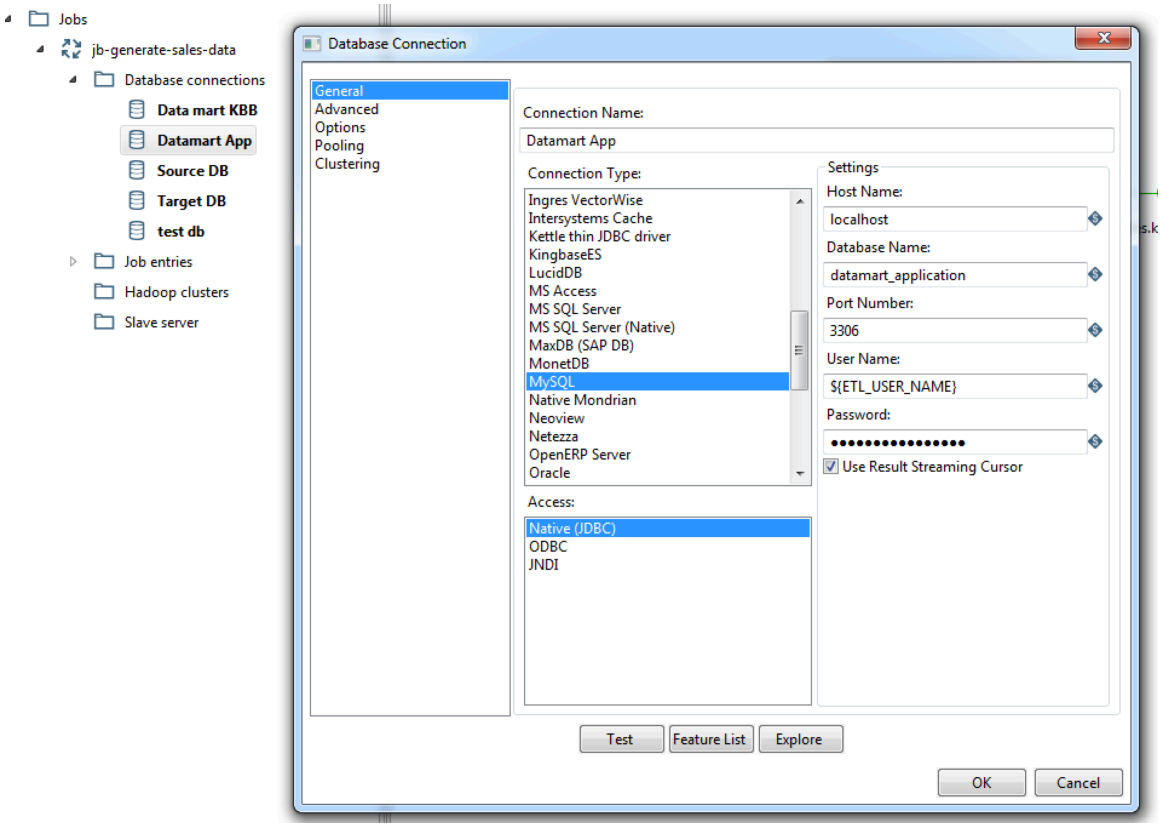
Assignment 02 Specification

Preparation

1. In MySQL, create a database called **datamart_application**
2. Execute the dump file **admissions_info_dump.sql** to import a source table, **admissions_info**.
3. In your Kettle, create a connection called, **Datamart App**, which uses the environment variables, **ETL_USER_NAME** and **ETL_USER_PASS**, as its account.



Step 1. Draw Dimensional Star Schema

Analyze the source table, **admissions_info**, and draw the dimensional schema. The **minimum** requirement for the dimensional start schema is as below:

- **Dimensions**
  - Minimum 1 Role Playing Dimension
  - 1 Junk Dimension (DIM_ETHNICITY)
  - Minimum 1 Bridge Table
  - Minimum Number of Dimensions: 5

- **Measures**
  - GRE percentiles (Quant., Analytics, and Verbal)
  - Numeric flags on admission status

PROG_ACTN	IS_APPLIED	IS_ADMITTED	IS_ACCEPTED
APPL	1	0	0
DENY	1	0	0
WAPP	1	1	0
MATR	1	1	1

Step 2. ETL Implementation

Implement the star schema into physical tables in the **datamart_application** schema of your MySQL server. Not all dimensions in the design of the Step 1 need to be implemented. The minimum implementation requirements are as below:

- **Dimensions**
  - Implement minimum 5 dimensions
  - No bridge table implementation necessary as a BI application of your choice may not support the model
- **Measures**
  - GRE percentiles (Quant., Analytics, and Verbal)
  - Numeric flags on admission status

Step 3. Reporting and Analytics

Deploy your choice of a BI application(s) on top of your star schema implementation.

Add two calculated measures in the BI application layer as below:

- Selectivity = Number of Admitted / Number of Applied
- Yield Ratio = Number of Accepted / Number of Admitted

Explore the star schema via the BI application, and report minimum 5 findings.

Here are the BI application choices you have:

1. **Saiku Analytics (OLAP)**
   a. Saiku Analytics CE (for reporting)
      i. Download: http://community.meteorite.bi/
      ii. Basic Demo: http://www.meteorite.bi/training/videos

   b. Pentaho Schema Workbench (for meta data management)
      i. Download: http://community.pentaho.com/
      ii. Usage Example Demo: https://www.youtube.com/watch?v=Tqw3oOk5jsM&list=PLIS-R80eiu1snl5wW893-BLiE0yDVhQAe

2. **Tableau Desktop (Visualization Tool)**
    a. Download: http://www.tableau.com/academic/students
    b. Free Training Videos: http://www.tableau.com/learn

Assignment 02 Group Deliverables

Each group must submit the followings in one zipped file on BB:

1. Dimensional Star Schema
    a. ERD (Entity-Relation Diagram) either in PDF or MS-DOC format
    b. DDL scripts in pure text format: The filename must be **ddl-scripts.sql**
2. Fully functional Kettle codes
    a. Make sure to use the environment variables, **ETL_UESR_NAME** and **ETL_USER_PASS**, for your DB connection
3. Documentations
    a. Setup instructions if any
    b. ETL execution instructions
    c. Minimum 5 findings from the Reporting & Analytics

Extra Credit

- Voluntary group presentation on April 27
- Must notify the instructor by April 20th.
- The credit can be applied to all but the Final

End.