

RETAIL DATA WAREHOUSE DATA INTEGRATION

TALEND JOBS SUMMARY

- Total Jobs
- Dimensions : 10
 - ✓ Dim Promotion
 - ✓ Dim Geography
 - ✓ Dim Product
 - ✓ Dim Sales Territory
 - ✓ Dim Customer Company
 - ✓ Dim Customer Person
 - ✓ Dim Product Cost_SCD
 - ✓ Dim Product Price_SCD
 - ✓ Dim Store
 - ✓ Dim date and many more...

Facts : 6

- ✓ Fact Online Sales
- ✓ Fact sales
- ✓ fact strategy plan
- ✓ fact exchange rate
- ✓ Fact Inventory
- ✓ Fact Sales Quota

CURRENCY CONVERSION IN FACT TABLE

The screenshot displays the Talend Expression Builder window. The main text area contains the following Java code for currency conversion:

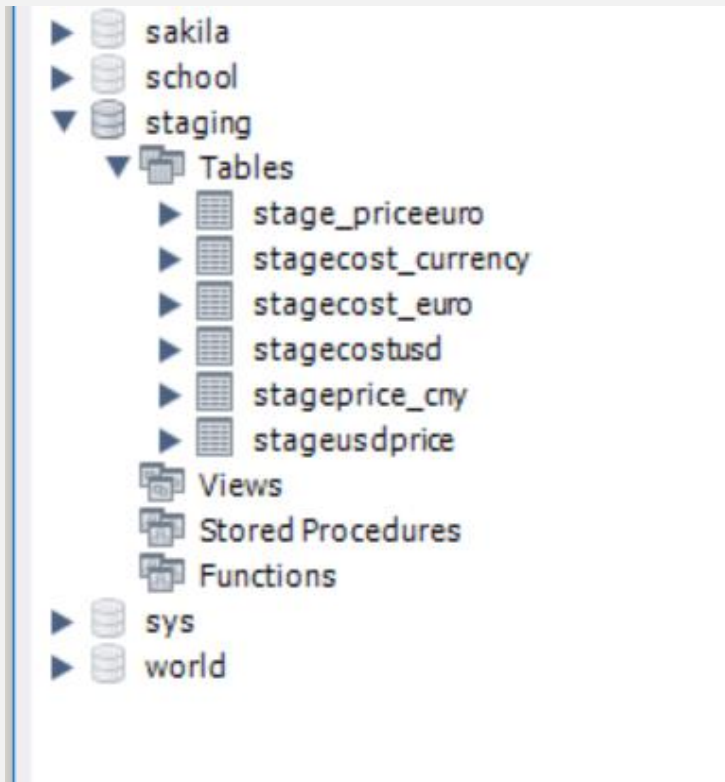
```
factor.EndOfDayRate==null?new BigDecimal(0):  
(fodMysql.SalesAmount).divide(factor.EndOfDayRate,9,BigDecimal.ROUND_HALF_UP)
```

Below the code editor is a toolbar with mathematical operators (+, -, *, /, ==, <, <=, !=, >=, >) and logical operators (and, or, not, parentheses). To the right of the code editor is a 'Test' section with a table showing variable values:

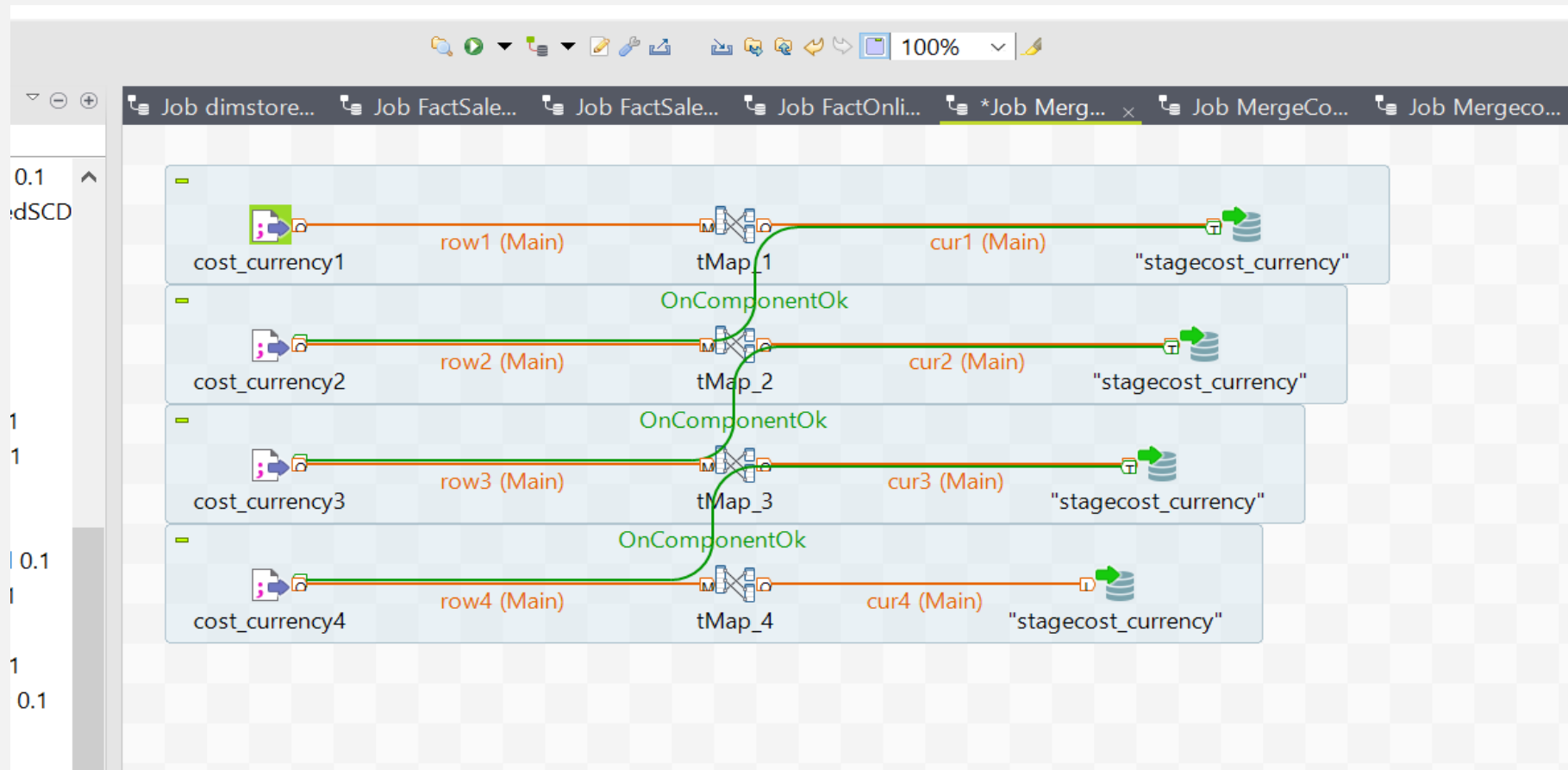
Var	Value
fodMysql...	0
fodMysql...	null
fodMysql...	null
fodMysql...	null
fodMysql...	null
fodMysql...	null
fodMysql...	null
fodMysql...	0

At the bottom of the window, there are three panels: 'Categories' (listing *All, *User Defined, Data Masking, DataOperation, Data Quality, Mathematical), 'Functions' (empty), and 'Help' (containing the text 'Please select a category and function.').

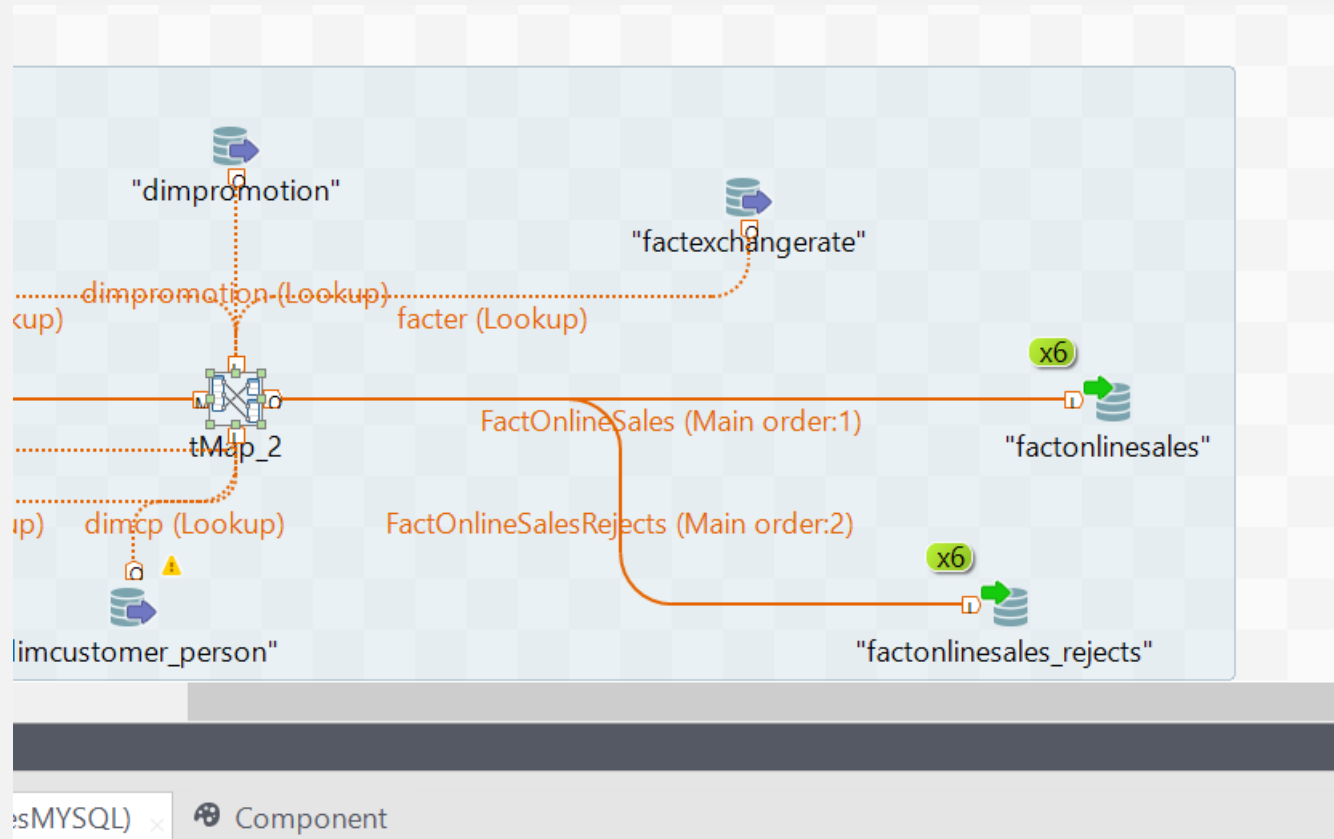
STAGING AREA



COST CURRENCY



TUNING THE JOBS- PARALLEL EXECUTION AND JVM SIZE



The screenshot displays the Talend Studio interface. At the top, a data pipeline is visible in the Designer tab, showing a sequence of components: "factonlinesalesorderdetail", "fodMySQL (Main)", "fodMySQL (Lookup)", "tMap_2", "FactOnlineSales (Main order:1)", and "factonlinesales". The "factonlinesales" component is highlighted, and its configuration is shown in the right-hand pane.

The configuration pane for "factonlinesales" (tMySQLOutput_2) is open, showing the "Advanced settings" tab. The "Additional JDBC Parameters" field is set to "noDatetimeStringSync=true". The "Extend Insert" checkbox is checked, and the "Number of rows per insert" is set to 100. The "Commit every" field is set to 500000.

The "Additional columns" table is empty, with columns for Name, Data type, SQL expression, Position, and Reference column. Below the table are icons for adding, removing, and reordering columns.

The "Use field options" section includes the following checkboxes:

- ☐ Use field options
- ☐ Use Hint Options
- ☒ Enable debug mode
- ☐ Use duplicate key update mode insert
- ☐ tStatCatcher Statistics
- ☒ Enable parallel execution

The "Number of parallel executions" field is set to 6.

RUNNING, TESTING AND DEBUGGING

etailDW_Final (Connection: Local)

The screenshot displays the Talend Studio interface for a job named 'etailDW_Final'. The main workspace shows a job design with components like 'dimstore', 'tMap_1', 'dimgeography', 'tMap_3', and 'dimstore3'. The 'dimstore' component is selected, and its data preview is shown on the right. The preview table has columns: StoreID, StoreKey, SalesTerritorySK, StoreManager, StoreType, StoreName, StoreDescription, Status, and OpenDate. The data row shows: 2, 201, 2, 219, Store, Contoso London Store, Contoso London Store, On, and Sat Apr 28 08:51:37 EDT 2018.

Below the job design, the 'Run (Job dimstore)' button is visible. The 'Component' tab is active, showing the 'Job dimstore' component. The 'Debug' tab is also active, displaying the debug console output.

Debug Console Output:

```
Starting job dimstore at 08:49 28/04/2018.

[statistics] connecting to socket on port 3896
[statistics] connected
[trace] connecting to socket on port 4870
[trace] connected

joblet: CNTL_JobStats_Child --> vJobMessage: [null] --- vJobName: dimstore vJobPID: afkpzC
vChildName: dimstore vChildPID: afkpzC from context
vRootName: null vRootPID: null
```

The right sidebar shows the 'Integration' palette with a search bar 'Find component...'. The 'Favorites' section lists components: tMap, tMSSqlInput, tMSSqlOutput, tMysqlInput, and tMysqlOutput. The 'Recently Used' section lists: tRunJob, tOracleInput, tMSSqlInput, and tPostgresallIno.

REJECT CODES SAMPLE

Expression Builder

Expression

☒ Wrap Undo(Ctrl + Z) Clear

```
\
  ((Relational.ISNULL(dimstore.GeographySK )) || dimstore.GeographySK ==0)
? 80
: (
  (
    ((Relational.ISNULL(fodMysql.SalesOrderNumber)))
    ? 90
    : (
      (
        ((Relational.ISNULL(fodMysql.SalesOrderLineNumber)) || fodMysql.SalesOrderLineNumber==0)
        ? 100
        : (
          (
            ((Relational.ISNULL(fodMysql.SalesAmount.divide(facter.EndOfDayRate,2,
java.math.RoundingMode.HALF_UP) )) || fodMysql.SalesAmount.divide(facter.EndOfDayRate,2,
java.math.RoundingMode.HALF_UP).equals(0))
            ? 130
```

+ - * / == < <= != >= > and or not ()

Categories

- *All
- *User Defined
- Data Masking
- DataOperation
- Data Quality

Functions