

Team Assignment

Databases:

- Source db:
 - AdventureWorks2016_us - sql server - backup AdventureWorks2016_us.bak
 - AdventureWorks2016_eu - mysql
 - AdventureWorks2016_au - postgresql
 - AdventureWorks2016_ca - oracle
 - Source files
- Target db:
 - AdventureWorksDW_neu: MySQL for Talend Target
 - AdventureWorksDW_neu: SQL Server for SSIS Target
- SQL Scripts
 - AdventureWorksDW2014 Purchasing DW - 2018-03-01 sql server.sql
 - AdventureWorksDW2014 Purchasing DW - 2018-03-01 mysql.sql

Team Assignment

Updated: April 14, 2018

Target DW

TABLE_CATALOG	TABLE_NAME	DW	Comments
AdventureWorksDW_neu	DimCurrency	X	
AdventureWorksDW_neu	DimCustomerPerson	X	
AdventureWorksDW_neu	DimDate	X	
AdventureWorksDW_neu	DimEmployee	X	
AdventureWorksDW_neu	DimGeography	X	
AdventureWorksDW_neu	DimLocation	X	
AdventureWorksDW_neu	DimEmployeeSalaryHistory	X	SCD 2, Replaces DimPayHistory
AdventureWorksDW_neu	DimProducts_Purchased	X	only products that are purchased from vendors
AdventureWorksDW_neu	DimProductsAll	X	
AdventureWorksDW_neu	DimProductVendor	X	
AdventureWorksDW_neu	DimPromotion	X	
AdventureWorksDW_neu	DimSalesTerritory	X	
AdventureWorksDW_neu	DimScrapReason	X	
AdventureWorksDW_neu	DimShipMethod	X	
AdventureWorksDW_neu	DimStore	X	
AdventureWorksDW_neu	DimVendorContacts	X	
AdventureWorksDW_neu	DimVendors	X	
AdventureWorksDW_neu	Fact_WorkOrder	X	
AdventureWorksDW_neu	Fact_WorkOrder_Rejects	X	
AdventureWorksDW_neu	Fact_WorkOrderRouting	X	
AdventureWorksDW_neu	Fact_WorkOrderRouting_Rejects	X	
AdventureWorksDW_neu	FactInternetSales	X	Sources are in native currency & need to be converted to USD
AdventureWorksDW_neu	FactInternetSales_Rejects	X	
AdventureWorksDW_neu	FactPurchases	X	Purchase Orders
AdventureWorksDW_neu	FactPurchases_Rejects	X	
AdventureWorksDW_neu	FactStoreSales	X	Sources are in native currency & need to be converted to USD
AdventureWorksDW_neu	FactStoreSales_Rejects		

Team Assignment

Updated: April 14, 2018

TABLE_CATALOG	TABLE_NAME	DW	Comments
AdventureWorksDW_neu	Dim_SORSystem	X	Track SOR_ID, Note: Team manually inputs
AdventureWorksDW_neu	Dim_RejectCodes	X	Track Rejection Codes, Note: Team manually inputs

Data Sources

FileName	File Type	Data Subject
adventureworks2016_au.sql	PostgreSQL export	Australia "AU"
AdventureWorks2016_CA.zip	Oracle export	Canada "CA"
AdventureWorks2016_us.bak	SQL Server backup	United States "US"
adventureworksdw_eu.sql	MySQL export	Europe "EU" - 3 countries Germany "DE", France "FR" & Great Britain "GB"
BillofMaterials.csv	CSV file	BillofMaterials
DimCurrency.csv	CSV file	DimCurrency
EmployeePayHistory.csv	CSV file	DimEmployeeSalaryHistory
Employee_Manager.csv	CSV file	Employee_Manager
FactCurrencyRateConversion.csv	CSV file	FactCurrencyRateConversion
Location.csv	CSV file	Location
SalesReason.csv	CSV file	SalesReason
SalesTaxRate.csv	CSV file	SalesTaxRate
SalesTerritoryHistory.csv	CSV file	SalesTerritoryHistory
ScrapReason.csv	CSV file	ScrapReason
Shift.csv	CSV file	Shift

Team Assignment

Updated: April 14, 2018

Output:

Need to load target DW:

- AdventureWorksDW_neu: MySQL for Talend Target
- AdventureWorksDW_neu: SQL Server for SSIS Target

Note:

- Assume SCD Type 2 for DimEmployeeSalaryHistory and SCD 1 (only get current or active rows) for DimEmployee
- Cannot use AdventureWorksDW2016 for any data except DimDate
- Cannot use any views, stored procedures or functions from AdventureWorks2016

Submit:

- Schema, List of Tables, Row Counts – in spreadsheet or table
- Screen capture of success run of Talend & SSIS jobs with row counts
- Upload Talend & SSIS jobs
- BI Queries Purchasing related data - PowerBI
 - Ranked order of Vendors by purchase amount \$
 - Ranked order of products by amount \$
 - By category
 - By subcategory
 - By product model (top 20)
 - By product (top 20)
 - Ranked list of employees purchasing products amount \$
 - List of employees who purchased products with phone, email & address
 - List of purchasing contacts with vendor name, phone, email & address
- BI Queries Sales-related Data - Tableau
 - Ranked order of products by amount \$
 - By category
 - By subcategory
 - By product model (top 20)
 - By product (top 20)

Team Assignment

Updated: April 14, 2018

- Internet Sales by Geo – Qty & \$ (USD)
- Store (or Reseller) Sales by Geo – Qty & \$ (USD)

Grading criteria:

- Loading all tables
- Handling errors in loading Fact tables
- Correct target table row counts – total rows, valid rows and invalid rows.
- Split of invalid rows by type and reason for error.
- Structure of SSIS code
- Structure of Talend code
- Completing SSIS job in one pass & job time
- Completing Talend job in one pass & job time