

Adventure Works Purchasing DW

Databases:

- Source databases:
 - AdventureWorks2017 - SQL Server all schemas except for below
 - aw2017neu_hr - Oracle (HR Schema)
 - aw2017neu_person - Oracle (Person Schema)
 - aw2017neu_purchasing – MySQL (Purchasing Schema)
 - dimdate.csv
- Target databases: AdventureWorks_PurchasingDW (student creates). Implement complete DW schema on:
 - MySQL
 - Oracle
- Load one of DWs with Alteryx & other Talend

Objectives:

Create a DW containing data from AdventureWorks2017 that can be used by the business the answer questions on:

- All product purchases made by AdventureWorks from vendors, i.e. purchase orders
- List of products that might be purchased from Vendors, in other words, all the products that are not made by AdventureWorks
 - Need to include product attributes such as list price, standard cost, model name, subcategory name and category name.
 - Note: if hierarchy applicable then this needs to be a star and NOT a snowflake design
- A slowly changing dimension (SCD) for list prices for products above
 - SCD 2 or higher
- A slowly changing dimension (SCD) for standards costs for products above
 - SCD 2 or higher
- Vendors that AdventureWorks might purchase products from
- Products that AdventureWorks might purchase from vendors and associated with the vendors that sell them.
- Vendors' employees who interact with AdventureWorks and include attributes about those people.
- AdventureWorks' employees who have bought products from vendors and include attributes about those people.
- A slowly changing dimension (SCD) for AdventureWorks employee pay (only employees listed above)
 - SCD 2 or higher
- AdventureWorks' sales territories
- AdventureWorks' ship methods
- A geography dimension that AdventureWorks could use as outrigger for dimensions above.

- A calendar dimension that AdventureWorks could use for facts and dimensions
- AdventureWorks product inventory for purchase products and locations stored in

The tables created and data in each is based on your design. There may need to be additional dimensional tables based on how you design the target database

Tasks:

Week 1: 1/30/20 Perform data profiling using:

- SQL
- Microsoft Power BI
- List of tables from source system that is needed to answer queries and build DW

Week 2: 2/6/20 Create DW in target database.

- Create data model in E/R studio
- Create DDL
- Create tables in your DW

Week 3: 2/13

- Load data into Purchasing DW using Alteryx
- Perform data profiling & creating BI answering questions from Week 1 with new Purchasing DW

Week 4: 2/20/20 Load data into Purchasing DW using Talend

DW Tables:

Target DW - AdventureWorks_PurchasingDW

TABLE_NAME	DW	Comments
DimEmployee (Purchasing only)	X	Only purchasing employees
DimGeography	X	Only US locations
DimLocation	X	
DimProducts_Purchased	X	only products purchased from vendors
DimProductVendor	X	
DimProductCostHistory	X	SCD 2
DimProductPriceHistory	X	SCD 2
DimScrapReason	X	
DimShipMethod	X	
DimVendorContacts	X	
DimVendors	X	
Fact_WorkOrder	X	
Fact_WorkOrder_Rejects	X	
Fact_WorkOrderRouting	X	
Fact_WorkOrderRouting_Rejects	X	
FactPurchases	X	Purchase Orders
FactPurchases_Rejects	X	
Dim_SORSystem	X	Track SOR_ID, Note: Team manually inputs
Dim_RejectCodes	X	Track Rejection Codes, Note: Team manually inputs
DimDate		

Output:

Submit:

- Schema, List of Tables, Row Counts – in spreadsheet or table
- Screen capture of success run of Talend & Alteryx jobs with row counts
- Upload Talend & Alteryx 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

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 Talend code
- Completing Talend job in one pass & job time

Business questions for data profiling & BI visualizations:

- Queries (SQL queries): Answer using existing AdventureWorks2017, not new schema
 - Ranked order of Vendors by purchase amount \$
 - Ranked order of products purchased by amount \$
 - By category
 - By subcategory
 - By product model (top 20)
 - By product (top 20)
 - List of employees who purchased products with phone, email & address
 - List of employees who purchased products with pay rate & raises (SCD)
 - List of purchasing vendor contacts with vendor name, phone, email & address
 - List of product prices by product order by product and SCD effective ascending
 - List of standard costs by product order by product and SCD effective ascending
- Microsoft Power BI: Answer using existing AdventureWorks2017, not new schema
 - Ranked order of Vendors by purchase amount \$
 - Ranked order of products purchased by amount \$
 - By category
 - By subcategory
 - By product model (top 20)
 - By product (top 20)
 - List of employees who purchased products with phone, email & address
 - List of employees who purchased products with pay rate & raises
 - List of purchasing vendor contacts with vendor name, phone, email & address