Final Team Project Fall 2018

- Retail DW & BI Project

Overview

- This project involves a retail company that sells a variety of products to people and to businesses across a variety of sales channels.

- The Retail Company would like to:
  - Create a data warehouse (DW) from its various systems of record (SORs)
  - Create BI applications that enable analysis of business performance

- The DW will store (facts) in the following subject areas:
  - Sales
  - Inventory
  - Sales Quotas
  - Strategy Plans

ATHENA
IT SOLUTIONS

Retail Company

- The categories of products sold:
    - Audio
    - TV and Video
    - Computers
    - Cameras and camcorders
    - Cell phones
    - Music, Movies and Audio Books
    - Games and Toys
    - Home Appliances
- The Retail Company would like to:
    - Create a data warehouse (DW) from its various systems of record (SORs)
    - Create BI applications that enable analysis of business performance

ATHENA
IT SOLUTIONS

Retail Company

- The company has 4 sales channels:
    - Retail
    - Stores
    - Catalog
    - Online
- Three top-level geographic regions:
    - North America
    - Europe
    - Asia
- Customers who purchased products are tracked in:
    - Catalog
    - Online

ATHENA
IT SOLUTIONS

Retail Company - SOR

Database	Business Area	dbms
Retail_SOR_NA	North America	
Retail_SOR_EU	Europe	
Retail_SOR_AS	Asia	
Retail_SOR_CAT	Catalog	
Retail_SOR_PLAN	Sales Quota & Strategy Plan	
Flat Files	Cross-unit reference data	

- All SORs contain:
  - Inventory
  - Sales Quotas
  - Strategy Plans
  - Supporting dimensions

- Continent oriented SORs contain:
  - Sales – Store, Reseller & Online Channels data
- The Catalog SOR contain:
  - Sales – Catalog Channel data from all three continents

ATHENA
IT SOLUTIONS

Retail Company – SOR Table Row Counts

These numbers will change

Table	retail_sor_eu	retail_sor_as	retail_sor_na	retail_sor_cat	retail_sor_plan	Worlwide
DimAccount	24	24	24	24		24
DimChannel						-
DimCustomer_Company	43	67	276			386
DimCustomer_Person	5,505	3,593	9,395	9,395		18,493
DimDate	45,656	45,656	45,656	45,656		45,656
DimEntity	421	421	421	421		421
DimGeography	147	67	303	303		517
DimProduct	2,517	2,517	2,517	2,517		2,517
DimProductCategory	8	8	8			8
DimProductSubcategory	44	44	44			44
DimPromotion	19	10	10	10		39
DimSalesTerritory	53	29	183	1		265
DimScenario						-
DimStore	56	41	209			306
FactInventory	1,745,148	1,410,895	5,912,445	143,512		9,212,000
FactOnlineSalesOrderDetail	4,463,166	4,796,456	5,528,466			14,788,088
FactOnlineSalesOrderHeader	756,126	405,523	748,650			1,910,299
FactSales						
FactCatalogSales				118,537		118,537
factresellersales	175,047	173,574	175,047			523,668
factstoresales	497,523	484,058	1,940,185			2,921,766
FactSalesQuota_All					5,856,453	5,856,453
FactStrategyPlan_All					2,998,090	2,998,090
Total	7,691,503	7,322,983	14,363,839	320,376	8,854,543	38,397,577

ATHENA
IT SOLUTIONS

Systems of Record (SORs)

- SORs have prices, costs & sales in "continent" currency
  - North America & Catalog Operations – US Dollars (USD)
  - Europe - Euro (EUR)
  - Asia - China Yuan (CNY)
  - Quota & Plan has data in local or "continent" currency
- DW has prices, costs & sales:
  - In USD
  - In "local" currency used in SOR
- Prices, costs & sales are in "constant" currency, i.e. recorded using published unit prices & costs
  - Daily currency exchange rate should be used in converting Euro & Yuan to USD

ATHENA
IT SOLUTIONS

Systems of Record (SORs)

- Unit Prices & Costs were changed during 2012-3016

- You need to create SCD dimension for both Unit Price & Unit Cost Dimensions.
  - Step 1 – initial unit prices or costs
  - Step 2-n – prices or costs revised

- Unit Prices & Unit Costs should NOT stores in Sales related tables nor in the Product dimension

ATHENA
IT SOLUTIONS

Systems of Record (SORs)

- DW sales related tables
  - FactSales – contains data for all 4 sales channels
  - FactOnlineSales – contains Online sales channel only and has customer data
- Sales Quotas & Strategy Plan:
  - Have data for three scenarios
    - Actuals
    - Budget – from SORs directly
    - Forecast – from SORs directly
  - Data granularity is monthly
  - For Sales Quota need to aggregate actual sales data to populate Actuals scenario
  - For Strategy Plan need to aggregate data for actuals in 2 accounts:
    - AccountKey = 4 'Sale Revenue'   - sales
    - AccountKey = 5 'Cost of Goods Sold' – total product cost

ATHENA
IT SOLUTIONS

Final Team Project

- Sourcing Requirements

ATHENA
IT SOLUTIONS

Deliverables

- DW has dimensional model to support:
  - Sales analysis
  - Inventory analysis
  - Sales quota & planning analysis
- DI -- Load data sources into DW
  - SOR: TBD
  - DW: TBD with ETL: Talend
  - DW: TBD with ETL: SSIS

ATHENA
IT SOLUTIONS

Deliverables

- BI -- Develop BI Dashboards, Reports & Visualizations analysis
  - Sales analysis
  - Inventory analysis
  - Sales quota & planning analysis

ATHENA
IT SOLUTIONS

Deliverables - BI

BI Tools:

- Tableau

- Microsoft BI
    - PowerBI

Deliverables:

- Dashboards with visualization for each analysis & for each BI tool
- Comparison of each tool – key differences – strengths & weaknesses

ATHENA
IT SOLUTIONS

Final Team Project

- Deliverables

Deliverables - DI

This will be updated

- Use Talend Enterprise Data Integrator to DW (TBD)

- Use SSIS to DW (TBD)

- Load DW from data sources
  - Document all jobs
  - Provide load statistics
  - Provide analysis of load jobs using one of BI tools used in this project

- Handle data quality & error messages
  - Document error handling results

- Load rejection records
  - Track reasons for rejections
  - Provide analysis of rejections & reasons using one of BI tools used in this project

- Note: Follow project standards

ATHENA
IT SOLUTIONS

Deliverables – Data Integration

- Table Row Counts
  - Total Rows - Dimensions, Facts & Reject Tables
  - Row counts by region (SOR source) - Dimensions, Facts & Reject Tables
  - Row counts by reject reason – Fact Tables
- Sales Amounts (USD)
  - By Region
  - By Channel (Internet versus Store)
- Time to load DW
  - Note: Need one root job that loads all sources – databases and files

ATHENA
IT SOLUTIONS

Deliverables – Analysis

This will be updated

- ## Data Subjects:
    - Online Sales (FactOnlineSales)
        - Online & Catalog both have customers
    - Sales (FactStoreSales)
        - All four sales channels
    - Inventory (FactInventory)
    - Sales Quota
    - Strategy Plan

- ## Types of analysis:
    - Trending
    - Ranking
    - Comparison
    - Period over Period
    - Geo Map
    - Contribution

- ## Measures:
    - Sales $, Profit, Profit Margin, Avg Order Size,…

- ## Dimensions:
    - Customers: Company & Person, demographics
    - Product: Product Hierarchy (Category, Subcategory, Product), Brand, other attributes
    - Store: Type, other attributes
    - Dates
    - Geography

ATHENA
IT SOLUTIONS

Online Sales Analysis

This will be updated

- Fact: FactOnline Sales

- Dimensions: Customer, Store, Product, Promotion, Sales Territory, Geo – Sales (or Store) & Customer

- Measures: Sales, Cost, Profit, Profit Margin, etc.

ATHENA
IT SOLUTIONS

Online Sales Analysis

This will be updated

- Product sales analysis
  - Contribution, Ranked (Top x)
    - Product Hierarchy, Brand
- Sales by time
  - Period over Period analysis
  - Trends
- Customer sales analysis
  - B2C - demographics such as education, income, etc. & geo
  - B2B – by annual sales, geo
- Sales Analysis by Geography
- Promotion Analysis - Impact of sales & margin with promotions

ATHENA
IT SOLUTIONS

Overall Sales Analysis

- Fact: FactStoreSales

- Dimensions: Channel, Store, Product, Promotion, Sales Territory, Geo – Sales (or Store)

- Measures: Sales, Cost, Profit, Profit Margin, etc.

ATHENA
IT SOLUTIONS

Sales Analysis

- Channel sales analysis
  - Contribution
  - Trends
- Product sales analysis
  - Contribution, Ranked (Top x)
    - Product Hierarchy, Brand
- Sales by time
  - Period over Period analysis
  - Trends
- Sales Analysis by Geography
- Promotion Analysis - Impact of sales & margin with promotions

ATHENA
IT SOLUTIONS

Overall Sales Analysis

This will be updated

- Fact: FactStoreSales

- Dimensions: Channel, Store, Product, Promotion, Sales Territory, Geo – Sales (or Store)

- Measures: Sales, Cost, Profit, Profit Margin, etc.

ATHENA
IT SOLUTIONS

Reject Analysis

- Fact: FactSales, FactOnlineSales, FactInventory
- Analysis by Region SOR and/or by Source Table
  - Row counts & type of rejects

ATHENA
IT SOLUTIONS

Sales Quota & Strategy Plan Analysis

This will be updated

- Fact: FactSalesQuota, FactStrategyPlan, FactSales

- Scenario: Budget, Forecast, Actual (from FactSales)

- Analysis

  - Compare Budget, Forecast & Plan
    - By product or Product Category
    - By store & geography

ATHENA
IT SOLUTIONS