

Final Team Project

- Retail DW & BI Project



BUSINESS INTELLIGENCE GUIDEBOOK

From Data Integration to Analytics

RICK SHERMAN

MK
MORGAN KAUFMANN

FOREWORD BY **CLAUDIA IMHOFF**
PRESIDENT OF INTELLIGENT SOLUTIONS, INC.

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

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

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

Retail Company - SOR

Database	Business Area	dbms
Retail_SOR_NA	North America	Microsoft SQL Server
Retail_SOR_EU	Europe	MySQL
Retail_SOR_AS	Asia	PostgreSQL
Retail_SOR_CAT	Catalog	Oracle
Flat Files	Cross-unit reference data	csv, text delimited

- 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

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)
- 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

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

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
 - Forecast
 - Data granularity is monthly
 - For Sales Quota need to aggregate actual sales data to populate
 - For Strategy Plan need to aggregate data for actuals in 2 accounts:

Final Team Project

- Sourcing Requirements



Deliverables

- DW has dimensional model to support:
 - Sales analysis
 - Inventory analysis
 - Sales quota & planning analysis
- DI -- Load data sources into DW
 - SOR: flat files, SQL Server, PostgreSQL & MySQL, Oracle
 - DW: MySQL
 - ETL: Talend
- BI -- Develop BI Dashboards, Reports & Visualizations Sales analysis
 - Inventory analysis
 - Sales quota & planning analysis

Deliverables - BI

BI Tools:

- Tableau
- Qlik
 - Qlik Sense
- Microsoft BI
 - PowerBI



Deliverables:

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

Final Team Project

- Deliverables



Deliverables - DI

- Use Talend Enterprise Data Integrator to DW (SQL Server)
- 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

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

Deliverables – Analysis

- Data Subjects:
 - Online Sales (FactOnlineSales)
 - Sales (FactStoreSales)
 - 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

Online Sales Analysis

- Fact: FactOnline Sales
- Dimensions: Customer, Store, Product, Promotion, Sales Territory, Geo – Sales (or Store) & Customer
- Measures: Sales, Cost, Profit, Profit Margin, etc.

Online Sales Analysis

- 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

Overall Sales Analysis

- Fact: FactStoreSales
- Dimensions: Channel, Store, Product, Promotion, Sales Territory, Geo – Sales (or Store)
- Measures: Sales, Cost, Profit, Profit Margin, etc.

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

Overall Sales Analysis

- Fact: FactStoreSales
- Dimensions: Channel, Store, Product, Promotion, Sales Territory, Geo – Sales (or Store)
- Measures: Sales, Cost, Profit, Profit Margin, etc.

Reject Analysis

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

Sales Quota & Strategy Plan Analysis

- Fact: FactSalesQuota, FactStrategyPlan, FactSales
- Scenario: Budget, Forecast, Actual (from FactSales)
- Analysis
 - Compare Budget, Forecast & Plan
 - By product or Product Category
 - By store & geography