

Delivering a Relational Data Warehouse

Week 2 – Designing a Relational Data Warehouse Schema

Module 05

Designing Fact Tables



Module Outline

05 | Designing Fact Tables

| | Topic |
|---|---|
| ▶ | Fact Table Fundamentals |
| ▶ | Fact Table Design Concepts |
| ▶ | Demo: Exploring the AdventureWorksDW Fact Tables |
| | |
| | |
| | |



©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

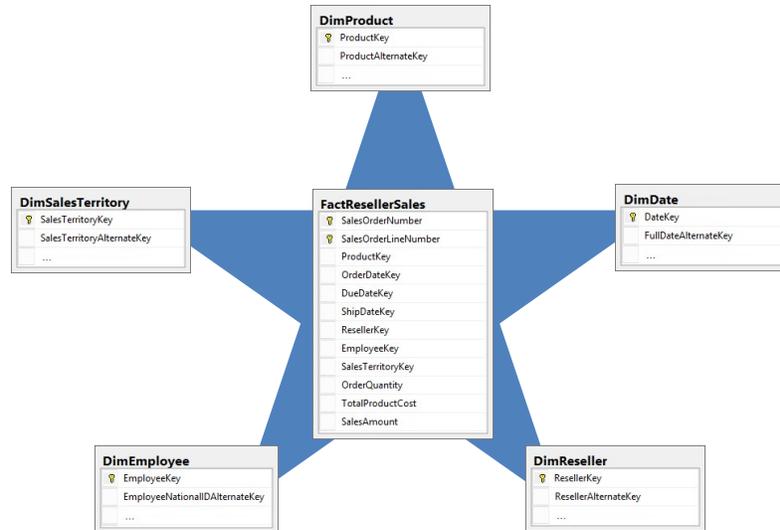
Module Outline

05 | Designing Fact Tables

| Topic |
|---|
| Fact Table Fundamentals |
| Fact Table Design Concepts |
| Demo: Exploring the AdventureWorksDW Fact Tables |
| |
| |
| |

Fact Table Fundamentals

Star Schema Review



Fact Table Fundamentals

- Fact tables are a collection of measurements associated with a specific business process
 - Table names are typically prefixed with **Fact**
- Specific column types:
 - Dimension keys (dimensionality)
 - Measures—numeric and typically additive
- Define a consistent granularity—the most atomic level by which the facts can be defined
- A primary key may serve no useful purpose

Fact Table Fundamentals

Example ► AW Fact Tables

- AdventureWorksDW2016
 - Database Diagrams
 - Tables (filtered)
 - System Tables
 - FileTables
 - External Tables
 - dbo.FactAdditionalInternationalProductDescription
 - dbo.FactCallCenter
 - dbo.FactCurrencyRate
 - dbo.FactFinance
 - dbo.FactInternetSales
 - dbo.FactInternetSalesReason
 - dbo.FactProductInventory
 - dbo.FactResellerSales
 - dbo.FactResellerSalesXL_CCI
 - dbo.FactResellerSalesXL_PageCompressed
 - dbo.FactSalesQuota
 - dbo.FactSurveyResponse
 - dbo.NewFactCurrencyRate

Fact Table Fundamentals

Example ► AW Fact Tables (Continued)

| FactResellerSales | |
|----------------------|--|
| SalesOrderNumber | |
| SalesOrderLineNumber | |
| ProductKey | |
| OrderDateKey | |
| DueDateKey | |
| ShipDateKey | |
| ResellerKey | |
| EmployeeKey | |
| SalesTerritoryKey | |
| OrderQuantity | |
| TotalProductCost | |
| SalesAmount | |

- Reseller Sales, by:
 - Product
 - Order/Due/Ship date
 - Reseller
 - Employee
 - Sales Territory
- Measures:
 - Order Quantity
 - Total Product Cost
 - Sales Amount, etc.
- Grain:
 - Sales Order Line

| FactSalesQuota | |
|------------------|--|
| SalesQuotaKey | |
| EmployeeKey | |
| DateKey | |
| CalendarYear | |
| CalendarQuarter | |
| SalesAmountQuota | |
| Date | |

- Reseller Sales Quota, by:
 - Employee
 - Date
- Measure:
 - Sales Amount Quota
- Grain:
 - Employee
 - Calendar Quarter



©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Module Outline

05 | Designing Fact Tables

| Topic |
|---|
| Fact Table Fundamentals |
| Fact Table Design Concepts |
| Demo: Exploring the AdventureWorksDW Fact Tables |
| |
| |
| |

Fact Table Design Concepts

- Grain
- Dimensionality
- Fact table type
- Measures
- Primary keys
- Foreign keys

Fact Table Design Concepts

Grain

- The fact table grain is the business definition of what a single record represents, at the most atomic level
- Typically, design tables to store the lowest grain possible
 - For example, sales order line, or quarterly salesperson quota
- While it is possible to store facts at higher levels, it is not (easily) possible to split out measure values beneath the grain
 - Sheer volumes, together with analytic requirements, may provide valid reason to store at a higher grain
- A single fact table should not mix different grains

Fact Table Design Concepts

Dimensionality

- After declaring the grain, it is usually straight forward to determine the dimensional keys that exist at that grain
 - By declaring the grain first, the discussion of keys remains grounded and precise
 - Dimensionality can be validated against requirements by referring to all of the “by” mentions
- Fact tables will always include a date dimension
- It is possible to reference a dimension table more than once
 - This is known as a **role playing** dimension

Fact Table Design Concepts

Fact Table Type

- There are three types of fact tables:
 - Transaction grain
 - Periodic snapshot grain
 - Accumulating snapshot grain

Fact Table Design Concepts

Fact Table Type (Continued)

- Transaction grain
 - Measurements are taken at a single instant, for example a sales order
- Periodic snapshot grain
 - Measurements are taken at a pre-defined time, or period of time
 - Often used for financial, stock or weather observation recording
- Accumulating snapshot grain
 - Measurements accumulate across a well-defined period or workflow
 - Often events are recorded by progression in the workflow, or by a status update

Fact Table Design Concepts

Measures

- Measure columns are typically numeric, and commonly additive
 - For example, **OrderQuantity**, and **SalesAmount**
- However, some measures cannot always be added, and these are categorized as either:
 - Non-additive, or
 - Semi-additive

| FactResellerSales | |
|-------------------|----------------------|
| 🔑 | SalesOrderNumber |
| 🔑 | SalesOrderLineNumber |
| | ProductKey |
| | OrderDateKey |
| | DueDateKey |
| | ShipDateKey |
| | ResellerKey |
| | EmployeeKey |
| | SalesTerritoryKey |
| | OrderQuantity |
| | TotalProductCost |
| | SalesAmount |

Fact Table Design Concepts

Measures ► Non-Additive

- Non-additive measures can include rates and ratios
- Requirements to report non-additive measures should be translated to measure columns that store the required parts
 - For example, **UnitPrice** is actually a rate, so consider storing $(\text{UnitPrice} * \text{OrderQuantity})$ as **SalesAmount**
- Note that this does not mean that **UnitPrice** is not a valid measure
 - Values can still be aggregated in other meaningful ways:
 - Average, Median, Min, Max, Count, Distinct count, etc.—just not Sum

Fact Table Design Concepts

Measures ► Semi-Additive

- Semi-additive measures can only be added *sometimes*
 - For example, stock balances can be added only within a time period, and are stored in snapshot fact tables
 - Another example, foreign currency sales amounts can only be added within their currency, and can be stored in transactional fact tables

| FactProductInventory | |
|----------------------|--|
| ProductKey | |
| DateKey | |
| MovementDate | |
| UnitCost | |
| UnitsIn | |
| UnitsOut | |
| UnitsBalance | |

| FactFinance | |
|--------------------|--|
| FinanceKey | |
| DateKey | |
| OrganizationKey | |
| DepartmentGroupKey | |
| ScenarioKey | |
| AccountKey | |
| Amount | |
| Date | |

Fact Table Design Concepts

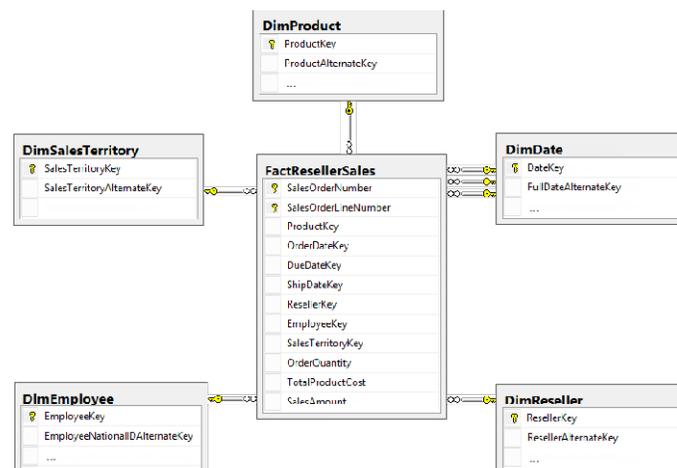
Primary Keys

- A fact table does not usually define a primary key, as it is often implied by the set of dimension keys
- There is typically no advantage to assign a surrogate key to the fact records
 - However, certain ETL methods for updating, or logging, fact rows may only be feasible if a surrogate key is assigned to each fact row

Fact Table Design Concepts

Foreign Keys

- Foreign keys can be used to enforce referential integrity between the dimension and fact tables
 - However, they may not be defined, or enabled, in order to optimize load performance





©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

Module Outline

05 | Designing Fact Tables

| Topic |
|---|
| Fact Table Fundamentals |
| Fact Table Design Concepts |
| Demo: Exploring the AdventureWorksDW Fact Tables |
| |
| |
| |

Demo

Exploring the AdventureWorksDW Fact Tables

Demo objectives:

1. Introduce the AdventureWorksDW fact tables
2. Explore the FactResellerSales table
3. Explore the FactSalesQuota table
4. Explore the FactProductInventory table



©2016 Microsoft Corporation. All rights reserved. Microsoft, Windows, Office, Azure, System Center, Dynamics and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.