



# **The Data Warehouse Toolkit**

## **Order Management**



# Summary

- Bus Matrix
- Fact Normalization
- Role Playing Dimension
- Junk Dimensions
- Multiple Currencies
- Audit Dimension

# Bus Matrix

“To decide which dimensional models to build, we begin with a top down planning approach we call the Data Warehouse Bus Architecture matrix. The matrix forces us to name all the data marts we could possibly build and to name all the dimensions implied by those data marts”.

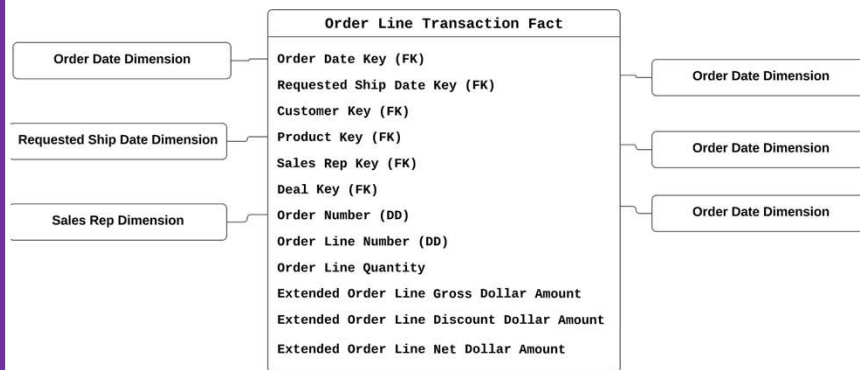
The Data Warehouse Lifecycle Toolkit

	<i>Date</i>	<i>Customer</i>	<i>Product</i>	<i>Sales Rep</i>	<i>Deal</i>	<i>Warehouse</i>	<i>Shipper</i>
Quoting	X	X	X	X	X		
Ordering	X	X	X	X	X		
Shipping to Customer	X	X	X	X	X	X	X
Shipment Invoicing	X	X	X	X	X	X	X
Receiving Payments	X	X		X			
Customer Returns	X	X	X	X	X	X	X

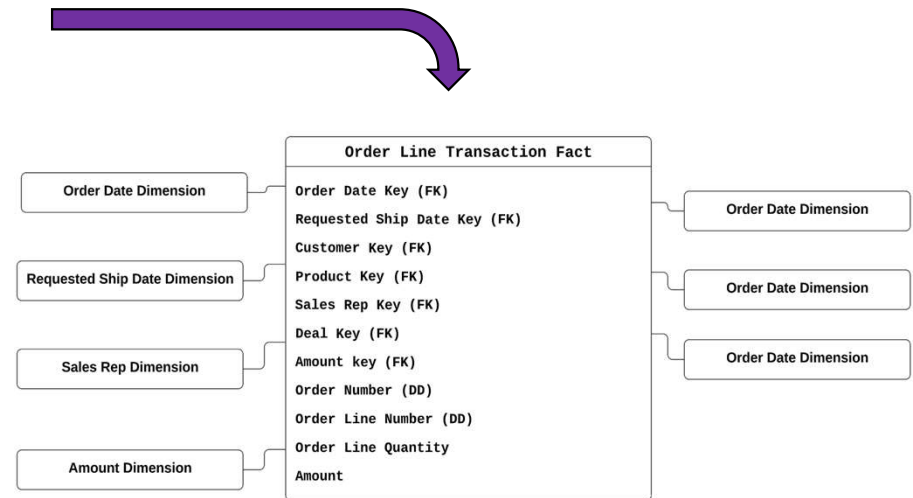
Bus matrix rows for order management processes

# Fact Normalization

- breaking down a fact table into multiple rows, each representing a single measurement and its associated type.
- introducing a new dimension table called "measurement type" to store the different types of measurements.



Fact table before normalization



Fact table after normalization



### Consideration :

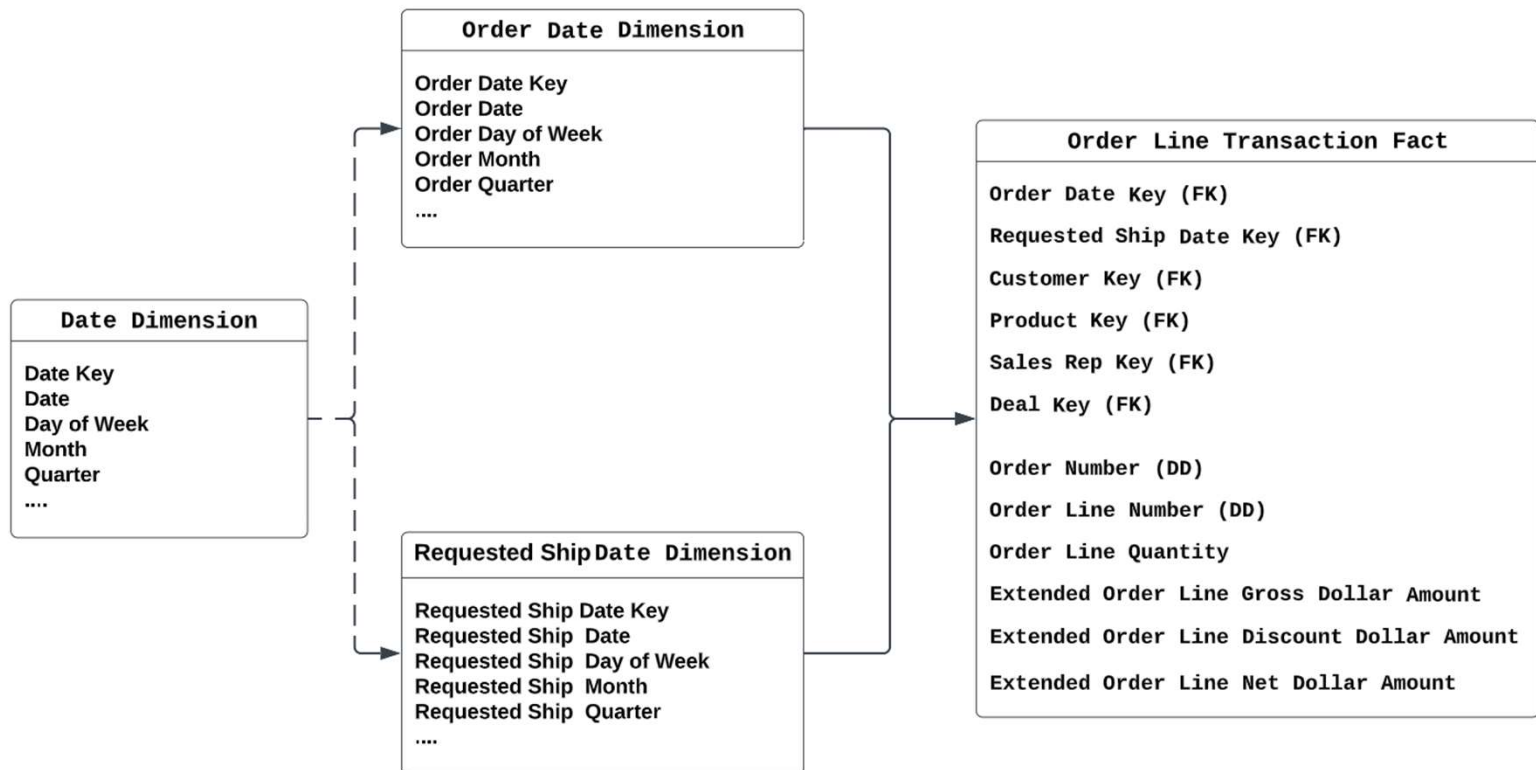
- when there are a large number of sparsely populated facts (many empty values) within a row, and when there's little need for computations between different facts.

### Disadvantage :

- Increase table size

# Role Playing Dimension

- Role playing in a dimensional model, it means that a single dimension can be used in multiple ways within the same fact table.
- Each use of the dimension in the fact table is referred to as a "role," and each role is treated as a separate entity
- This allows analysts to look at the same dimension in different contexts without creating separate physical tables for each role.



One dimension table, multiple roles

# Junk Dimension

- A dimension table that combines multiple low-cardinality attributes (flags and indicators) that typically have a limited number of possible values.
- used to avoid cluttering the fact table with numerous foreign keys and to make it easier to filter and analyze data based on these attributes.

## Why use a junk dimension :

- Reduces fact table size
- Improves query performance
- Enhances data organization



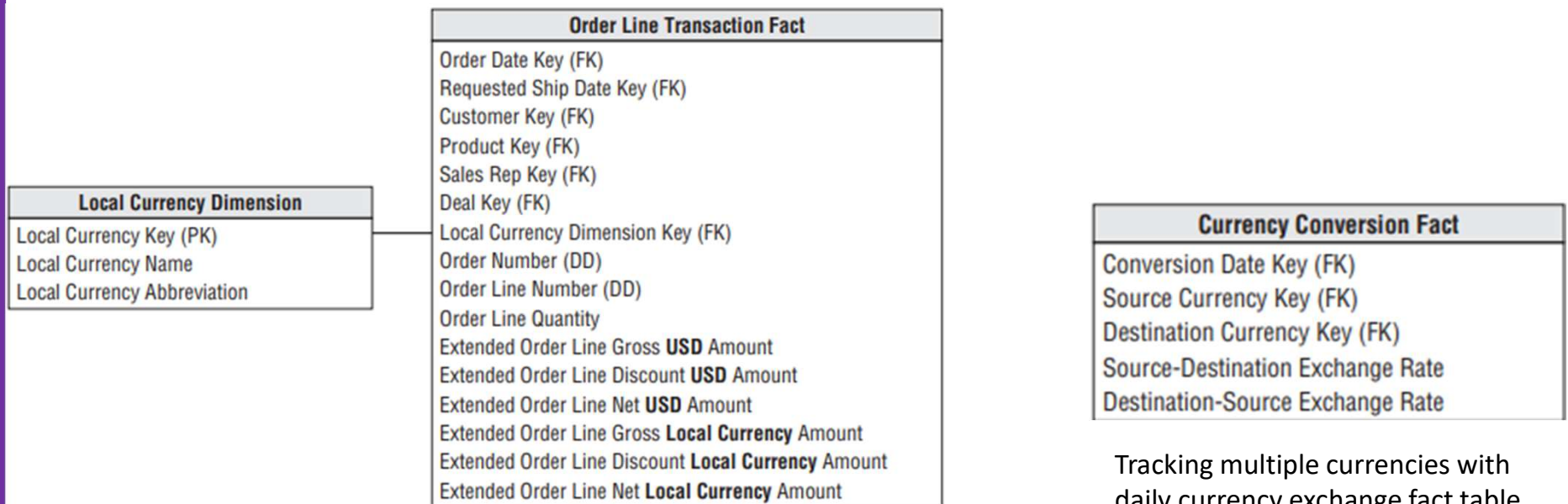
Customer_id	Employee_id	Product_id	Confirmed	Received	Shipped	Delivered
1	100	169	TRUE	TRUE	TRUE	TRUE
2	100	150	FALSE	FALSE	FALSE	FALSE
3	200	100	TRUE	TRUE	FALSE	FALSE
4	205	99	TRUE	TRUE	TRUE	FALSE

Customer_id	Employee_id	Product_id	Order_status
1	100	169	10
2	100	150	20
3	200	100	30
4	205	99	40

DWH_id	Confirmed	Received	Shipped	Delivered
10	TRUE	TRUE	TRUE	TRUE
20	FALSE	FALSE	FALSE	FALSE
30	TRUE	TRUE	FALSE	FALSE
40	TRUE	TRUE	TRUE	FALSE

# Multiple Currencies

- Dual facts for local and corporate currencies:
- Currency conversion fact table:



Metrics in multiple currencies within the fact table

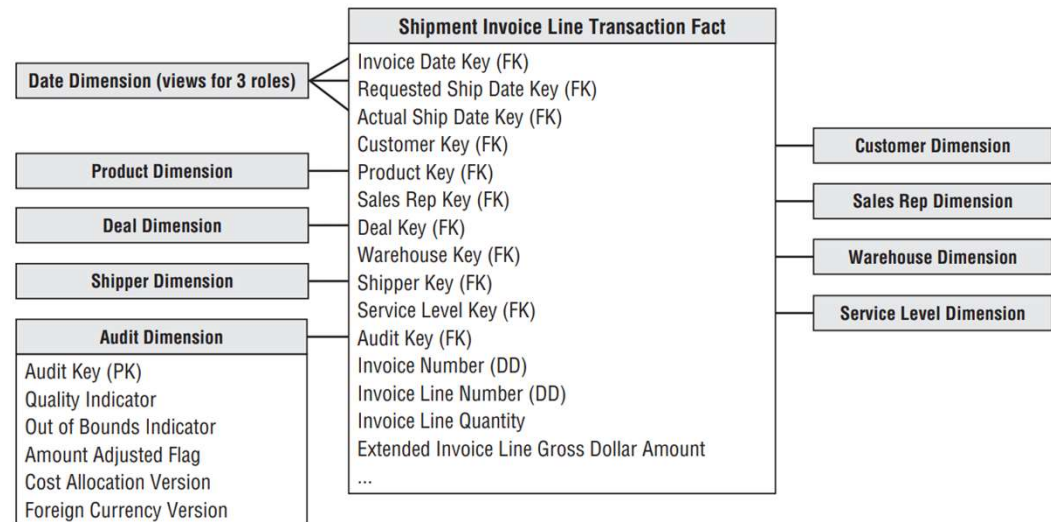
Tracking multiple currencies with  
daily currency exchange fact table

# Audit Dimension

## How to answer the following questions :

- What is my confidence in these reported numbers?
- Were there any anomalous values encountered while processing this source data?
- What version of the cost allocation logic was used when calculating the costs?
- What version of the foreign currency conversion rules was used when calculating the revenues?

**By including an audit dimension with any fact table to expose the metadata context that was true when the fact table rows were built**



Two vertical purple bars of different widths are positioned on the left side of the slide.

**Thank you for your attention**



# **The Data Warehouse Toolkit**

## **Order Management**