



Data Warehousing Fundamentals

by Mark Peco



Module 0. About the Course (8 min)

Module 1. Introduction to Data Warehousing (67 min)

- *Fundamental Ideas*
 - Demand for Information
 - Information Demand Segments
 - Information Demand Characteristics
 - Sources of Business Information
 - Information Building Blocks
 - Information Value Chain
 - Classic Definitions of BI
 - Updated Definition of BI
 - Bill Inmon's Definition of the DW
 - Ralph Kimball's Definition of the DW
 - TDWI's Definition of the DW
 - Mark Peco's Definition of the DW
 - Data Warehousing
- *Systems Concepts*
 - Definition of a System
 - System Properties
 - General Structure
 - Example of a System
 - Introduction to Systems Thinking
 - A Classic Example
 - Thinking Style Comparison
- *Architecture Considerations*
 - Definition
 - Purpose and Role of Architecture
 - Properties of Architecture
 - Categories of Architecture
 - Architecture Purposes
 - Dependencies and Relationships
- *Systems View of Data Warehousing*
 - Setting the Context
 - The BI System Components
 - Positioning the Information System
 - Components of the Information System
 - The Data Warehousing System
 - Defining the Data Warehousing System
 - Major Components
 - Detailed Components
- *Data Warehousing System Review*
 - Major Systems
 - Data Acquisition & Refinement
 - Data Provisioning & Retention
 - Information Delivery & Consumption
 - Building & Development
 - Leadership & Control

Module 2. Data Acquisition and Refinement (62 min)



DW-01: Data Warehousing Fundamentals

- *Getting Started*
- *Parts*
- *Exploration & Discovery System*
 - Information Requirements
 - Purpose
 - What Needs to Be Discovered
 - Desired Outputs
 - Components
 - Components Drill Down
 - Inputs and Outputs
 - Input 1 – Business Objectives
 - Input 2 – Ideas and Opportunities
 - Input 3 – Skilled Resources
 - Input 4 – Expectations and Constraints
 - Output 1 – Scope Definition
 - Output 2 – Architecture Decisions
 - Output 3 – Information Requirements
 - Output 4 – Stakeholders
 - Architecture Options
 - Central Hub
 - Hub and Spoke
 - Hub and Spoke with Staging
 - Information Bus
 - Information Bus with Staging
 - Independent Data
 - Variations Using an Operational Data Store
- *Refining & Integration System*
 - Purpose
 - Components
 - Inputs and Outputs
 - Input 1 – Identified Source Databases
 - Input 2 – Raw Data
 - Input 3 – Data Profiling Results
 - Input 4 – Identified Target Database
 - Output 1 – Extracted Data
 - Output 2 – Transformed Data
 - Output 3 – Updated Metadata
- *Transportation System*
 - Purpose
 - Components
 - Inputs and Outputs
 - Input 1 – Extracted Data
 - Input 2 – Transformed Data
 - Input 3 – Updated Metadata
 - Output 1 – Delivered Data
 - Output 2 – Loaded Data
 - Output 3 – Updated Metadata

Module 3. Data Provisioning and Retention (45 min)

- *System Context*
- *Getting Started*



DW-01: Data Warehousing Fundamentals

- *Parts*
- *Storage & Packaging System*
 - Purpose
 - Context
 - Components
 - Design Trade-Off
 - Selecting an Option
 - Inputs and Outputs
 - Input 1 – Loaded Data
 - Input 2 – Loaded Metadata
 - Output 1 – Stored Data
 - Output 2 – Available Data
 - Output 3 – Stored Metadata
 - Output 4 – Available Metadata
- *Technology System*
 - Purpose
 - Components
 - Inputs and Outputs
 - Input 1 – Processing Demands
 - Input 2 – Storage Demand
 - Input 3 – Performance Demand
 - Input 4 – Communications Demand
 - Output 1 – Processed Data
 - Output 2 – Stored Data
 - Output 3 – Measured Performance
 - Output 4 – Communications Delivered
- *Inventory System*
 - Purpose
 - Components
 - Inputs and Outputs
 - Input 1 – Loaded Data
 - Input 2 – Loaded Metadata
 - Input 3 – Expected Usage Patterns
 - Output 1 – Useful and Relevant Data
 - Output 2 – Described and Catalogued Data
 - Output 3 – Profile and Lineage Information

Module 4. Information Delivery and Consumption (35 min)

- *System Context*
- *Getting Started*
- *Parts*
- *Usage System*
 - Purpose
 - Components
 - Business Purpose Categories
 - Stakeholders
 - Inputs and Outputs
 - Input 1 – Business Purpose
 - Input 2 – Stakeholder Profile
 - Output 1 – Required Usage Pattern



DW-01: Data Warehousing Fundamentals

- Usage Pattern Examples
- *Delivery System*
 - Purpose
 - Components
 - Inputs and Outputs
 - Input 1 – Required Usage Pattern
 - Input 2 – Useful and Relevant Data
 - Input 3 – Described and Catalogued Data
 - Input 4 – Profile and Lineage Information
 - Output 1 – Delivered Information
 - Output 2 – Delivered Integrated Data
 - Output 3 – Supported Usage Patterns
- *Content Quality System*
 - Purpose
 - Components
 - Inputs and Outputs
 - Input 1 – Loaded Data
 - Input 2 – Loaded Metadata
 - Input 3 – Expected Usage Patterns
 - Output 1 – Usage Suitability Indicators
 - Output 2 – Data Quality Measures

Module 5. Building and Development (58 min)

- *System Context*
- *Getting Started*
- *Parts*
- *Participation System*
 - Purpose
 - Components
 - Inputs and Outputs
 - Input 1 – Understanding
 - Input 2 – Skills
 - Input 3 – Capabilities
 - Input 4 – Motivation
 - Output 1 – Roles
 - Output 2 – Responsibilities
- *Construction System*
 - Purpose
 - Components for Project Management
 - Top-Down Approach
 - Bottom-Up Approach
 - Hybrid Approach
 - Program vs. Project Levels
 - Components for Methodology
 - Waterfall Method
 - Incremental and Iterative Method
 - Components for Asset Portfolios
 - Context
 - Asset Categories
 - Data Provisioning Assets Example
 - Analysis and Design Techniques



DW-01: Data Warehousing Fundamentals

- Data Storage Assets
- Integration Assets
- Access & Delivery Assets
- Analytics Assets
- Inputs and Outputs
 - Input 1 – Architecture Decisions
 - Input 2 – Usage Patterns
 - Input 3 – Skills and Capabilities
 - Input 4 – Approaches and Techniques
 - Output 1 – Data Integration Assets
 - Output 2 – Data Storage Assets
 - Output 3 – Analytics Assets
 - Output 4 – Data Access and Delivery Assets
- *Asset & Process Quality System*
 - Purpose
 - Components
 - Inputs and Outputs
 - Input 1 – Asset Performance
 - Input 2 – Process Performance
 - Input 3 – Stakeholder Expectations
 - Output 1 – Quality Gaps
 - Output 2 – Recommendations

Module 6. Leadership and Control (34 min)

- *System Context*
- *Leadership*
- *Control*
- *Parts*
- *Stakeholder System*
 - Purpose
 - Stakeholder Categories
 - Components
 - Inputs and Outputs
 - Input 1 – Expectations
 - Input 2 – Understanding
 - Input 3 – Priorities
 - Input 4 – Education and Experience
 - Output 1 – Satisfaction
 - Output 2 – Maturity
 - Output 3 – Investment and Sponsorship
- *Governance System*
 - Building Blocks
 - Definition
 - Concepts
 - Purpose
 - Components
 - Ideas and Results
 - Inputs and Outputs
 - Input 1 – Funding
 - Input 2 – Capability Levels
 - Input 3 – Objectives and Constraints



DW-01: Data Warehousing Fundamentals

- Input 4 – Stakeholder Satisfaction
- Output 1 – Vision
- Output 2 – Strategy
- Output 3 – Standards
- Output 4 – Results

Module 7. Putting the Pieces Together (24 min)

- *Perspective & Alignment*
 - Data Warehousing Perspectives
 - Stakeholder and Participant Views
 - Governance Views
- *Data Warehousing System*
 - Review
 - Detailed Components
 - Interactions and Dependencies
 - Putting the Pieces Together
- *Course Summary*
 - Key Concepts
 - Data Warehousing System Defined
 - Data Warehousing Systems
 - The Final Takeaway