

# Data Product Engineering

**William McKnight**

**Jake Dolezal**



### Module 0: About the Course (6 min)

### Module 1: Introduction to Data Products (19 min)

- Module Overview
- What is a Data Product and Data Product Engineering?
  - What is a Data Product?
  - Characteristics of a Data Product
  - Examples of Data Product
  - What is Data Product Engineering?
  - Data Product Engineering Focus Areas
  - Database vs. Data Product Engineering Roles
  - Topics Summary
- Dependencies with Database Engineering and Data Pipeline Engineering  
Why Data Products?
  - Database & Data Product Engineering Dependencies
  - Data Pipeline vs. Data Product Engineering Roles
  - Data Pipeline & Data Product Engineering Dependencies
  - Database-Pipeline-Product Engineering Dependency Flow Example
  - Why Data Products?
  - Summary of Data Products Benefits
  - Topics Summary
- Module Summary

### Module 2: The Nature of Data Products (41 min)

- Module Overview
- Data Products Functions
  - Function: Provide Data via APIs
  - Function: Provide Data via Feeds & Streaming Services
  - Function: Provide Data via Self-Service BI & Analytics
  - Function: Provide Data via Data Files & Bulk Exports
  - Function: Provide Data via Embedded Data Services
  - Function: Validate & Assure Quality at Data Ingestion
  - Function: Validate & Assure Quality via Business Rule Validation
  - Function: Validate & Assure Quality via Data Lineage and Provenance Tracking
  - Function: Validate & Assure Quality with Automated Testing and Monitoring
  - Function: Data Transformation
    - Aggregation
    - Normalization
    - Enrichment
  - Topic Summary



## DE-02 Data Product Engineering

- Kinds of Data Products
  - Data Products Classifications
    - Serve Data
    - Validate Data
    - Quality Assurance
    - Transform
    - Aggregate
    - Compose
  - Topic Summary
- Data Products Characteristics
  - Characteristics: Discoverable
  - Characteristics: Addressable
  - Characteristics: Understandable
  - Characteristics: Valuable
  - Characteristics: Trustworthy
  - Characteristics: Self-Describing
  - Characteristics: Interoperable
  - Characteristics: Secure
  - Characteristics: Modular
  - Topic Summary
- Module Summary

### Module 3: Components of Data Products (31 min)

- Module Overview
- Data Product Data and Metadata
  - Data Product Components
  - Core Layers of a Data Product
  - Structured Data Products
  - Semi-Structured Data Products
  - Unstructured Data Products
  - Multi-Structured Data Products
  - Data Products by Structure
  - Raw Data Products
  - Transformed Data Products
  - Aggregated Data Products
  - Derived Data Products
  - Data Products by Granularity
  - Business Domain Data Product Examples: Sales
  - Business Domain Data Product Examples: Marketing
  - Business Domain Data Product Examples: Finance
  - Business Domain Data Product Examples: HR
  - Business Domain Data Product Examples: Operations
  - Business Domain Data Product Examples: Customer Support & IT
  - Business Domain Data Product Examples: Legal
  - Topic Summary



## DE-02 Data Product Engineering

- Data Product Processing and Interface
  - Data Product Key Process: Security and Access Authorization
  - Data Product Key Process: Error Checking
  - Data Product Key Process: Composing Response to Request
  - UI vs. API Interfaces
  - Differences Between UIs & APIs in Data Products – Part 1
  - Differences Between UIs & APIs in Data Products – Part 2
  - When to Use UIs vs. APIs in Data Products
  - Topic Summary
- Module Summary

### Module 4: APIs and Data Contracts (36 min)

- Module Overview
- APIs
  - API Architecture Overview
  - API Architecture Components
  - API Architectural Styles
  - Receiving & Interpreting Requests
  - Responding to Requests
  - API End-to-End Design
  - API End-to-End Code Example 1
  - API End-to-End Code Example 2
  - API End-to-End Code Example 3
  - API End-to-End Code Example 4
  - API Management Lifecycle Design & Development
  - API Management Lifecycle Production Development
  - API Management Lifecycle Security & Access Control
  - API Management Lifecycle Monitoring & Analytics
  - API Management Lifecycle Documentation & Developer Support
  - API Management Lifecycle Versioning & Depreciation
  - API Management Lifecycle Compliance Management
  - API Management Lifecycle Governance & Policy Management
  - API Management Lifecycle Retirement & Decommissioning
  - Topic Summary
- Data Contracts
  - What are Data Contracts?
  - Components of Data Contracts: Definition & Versioning
  - Components of Data Contracts: Quality & SLAs
  - Components of Data Contracts: Roles & Security
  - Benefits of Data Contracts
  - Tools Supporting Data Contracts
  - Example JSON Data Contract
  - Topic Summary
- Module Summary



## DE-02 Data Product Engineering

### Module 5: Developing and Operating Data Products (54 min)

- Module Overview
- Users and Use Cases for Data Products
  - Data Product Users
    - Data Engineers
    - Data Consumers
    - Machine Learning Engineers & Data Scientists
    - Application Developers
    - Data Stewards & Governance Teams
    - External Partners & Customers
    - Product Managers & Business Stakeholders
    - Operation Teams
  - Data Products for Self-Service Analytics
  - Data Products for Products Deployed as Part of a Semantic Layer
  - Topic Summary
- Developments Process
  - Developments Process: Requirements Gathering
  - Developments Process: Functional Design
  - Developments Process: Contract Design
  - Developments Process: Interface Design
  - Developments Process: Process Design
  - Developments Process: Building
  - Developments Process: Testing
  - Developments Process: Deployment
  - Topic Summary
- Operating Data Products
  - Operations: Monitoring & Observability
  - Operations: Security & Access Management
  - Operations: Data Governance
  - Operations: SLA Management
  - Operations: Scalability & Capacity Planning
  - Operations: Error Handling & Recovery
  - Operations: Change Management
  - Operations: Usage & Analytics
  - Operations: Maintenance & Support
  - Operations: Deprecation & Retirement
  - Topic Summary
- Module Summary

### Module 6: Data Products Best Practices (10 min)

- Module Overview
- Design Principles
- Development Practices
- Operational Excellence
- Scalability & Performance



## DE-02 Data Product Engineering

- Documentation & Support
- Governance & Compliance
- Lifecycle Management
- Final Thoughts...
- Module Summary