### Homework 10

[**The Kimball Methodology**](#_z079rn9skn2z) **1**

[Advantages of the Kimball Methodology](#_tf1d2ld699ab) 2

[Disadvantages of the Kimball Methodology](#_q3up50xfjw6o) 3

[Criteria for choosing Kimball approach](#_4ucaczcrtnfx) 3

[**The Inmon Method**](#_8mpwet2gw5ex) **5**

[Advantages of the Inmon Method](#_tviyrglj1ro0) 5

[Disadvantages of the Inmon Method](#_c0bd3wdk2u94) 6

[Criteria for choosing Inmon approach](#_p2vucrsdmz16) 6

### The Kimball Methodology

Initiated by Ralph Kimball, the Kimball data model follows a bottom-up approach to data warehouse architecture design in which data marts are first formed based on the business requirements.

The primary data sources are then evaluated, and an Extract, Transform and Load (ETL) tool is used to fetch data from several sources and load it into a staging area of the relational database server. Once data is uploaded in the data warehouse staging area, the next phase includes loading data into a dimensional data warehouse model that’s denormalized by nature. This model partitions data into the fact table, which is numeric transactional data or dimension table, which is the reference information that supports facts.

Star schema is the fundamental element of the dimensional data warehouse model. The combination of a fact table with several dimensional tables is often called the star schema. Kimball dimensional modeling allows users to construct several star schemas to fulfill various reporting needs. The advantage of star schema is that small dimensional-table queries run instantaneously.

To integrate data, Kimball approach to DW lifecycle suggests the idea of conformed data dimensions. It exists as a basic dimension table shared across different fact tables (such as customer and product) within a data warehouse or as the same dimension tables in various Kimball data marts. This guarantees that a single data item is used in a similar manner across all the facts.

An important design tool in Ralph Kimball’s data warehouse methodology is the enterprise bus matrix or Kimball bus architecture that vertically records the facts and horizontally records the conformed dimensions. The Kimball matrix, which is a part of bus architecture, displays how star schemas are constructed. It is used by business management teams as an input to prioritize which row of the Kimball matrix should be implemented first.

The Kimball approach to data warehouse lifecycle is also based on conformed facts, i.e. data marts that are separately implemented together with a robust architecture.

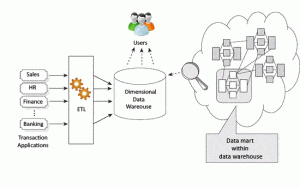
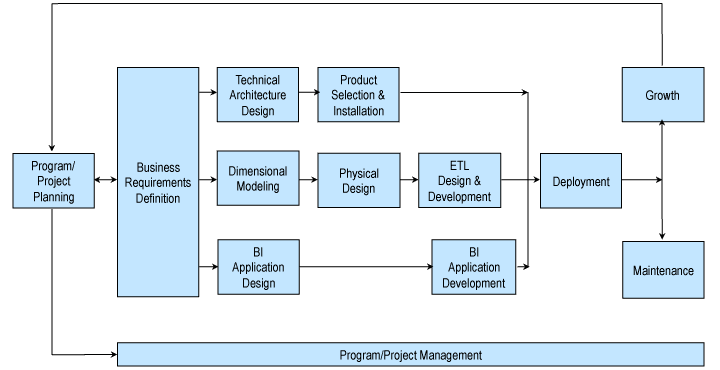


Figure 2. Basic Kimball Data Warehouse architecture explained (Source: Zentut)

#### Advantages of the Kimball Methodology

Some of the main benefits of the Kimball methodology include:

* Kimball dimensional modeling is fast to construct as no normalization is involved, which means swift execution of the initial phase of the data warehousing design process.
* An advantage of star schema is that most data operators can easily comprehend it because of its denormalized structure, which simplifies querying and analysis.
* Data warehouse system footprint is trivial because it focuses on individual business areas and processes rather than the whole enterprise. So, it takes less space in the database, simplifying system management.
* It enables fast data retrieval from the data warehouse, as data is segregated into fact tables and dimensions. For example, the fact and dimension table for the insurance industry would include policy transactions and claims transactions.
* A smaller team of designers and planners is sufficient for data warehouse management because data source systems are stable, and the data warehouse is process-oriented. Also, query optimization is straightforward, predictable, and controllable.
* Conformed dimensional structure for data quality framework. The Kimball approach to data warehouse lifecycle is also referred to as the business dimensional lifestyle approach because it allows business intelligence tools to deeper across several star schemas and generates reliable insights.



Kimball Approach to Data Warehouse Lifecycle (Source: Kimball Group)

#### Disadvantages of the Kimball Methodology

Some of the drawbacks of the Kimball design approach include:

* Data isn’t entirely integrated before reporting; the idea of a ‘single source of truth is lost.’
* Irregularities can occur when data is updated in Kimball DW architecture. This is because in denormalization technique, redundant data is added to database tables.
* In the Kimball DW architecture, performance issues may occur due to the addition of columns in the fact table, as these tables are quite in-depth. The addition of new columns can expand the fact table dimensions, affecting its performance. Also, the dimensional data warehouse model becomes difficult to alter with any change in the business needs.
* As the Kimball model is business process-oriented, instead of focusing on the enterprise as a whole, it cannot handle all the BI reporting requirements.
* The process of incorporating large amounts of legacy data into the data warehouse is complex.

#### Criteria for choosing Kimball approach

| Business decision support requirements | Tactical |
| --- | --- |
| Data integration requirements | Individual business requirements |
| The structure of data | KPI, business performance measures, scorecards… |
| Persistence of data in source systems | Source systems are quite stable |
| Skill sets | A small team of generalists |
| Time constraint | Urgent needs for the first data warehouse |
| Cost to build | Low start-up cost |

### The Inmon Method

Bill Inmon, the father of data warehousing, came up with the concept to develop a data warehouse which identifies the main subject areas and entities the enterprise works with, such as customers, product, vendor, and so on. Bill Inmon’s definition of a data warehouse is that it is a “subject-oriented, nonvolatile, integrated, time-variant collection of data in support of management’s decisions.”

The model then creates a thorough, logical model for every primary entity. For instance, a logical model is constructed for products with all the attributes associated with that entity. This logical model could include ten diverse entities under product, including all the details, such as business drivers, aspects, relationships, dependencies, and affiliations.

The Bill Inmon design approach uses the normalized form for building entity structure, avoiding data redundancy as much as possible. This results in clearly identifying business requirements and preventing any data update irregularities. Moreover, the advantage of this top-down approach in database design is that it is robust to business changes and contains a dimensional perspective of data across data mart.

Next, the physical model is constructed, which follows the normalized structure. This Bill Inmon model creates a single source of truth for the whole business. Data loading becomes less complex due to the normalized structure of the model. However, using this arrangement for querying is challenging as it includes numerous tables and links.

This Inmon data warehouse methodology proposes constructing data marts separately for each division, such as finance, marketing sales, etc. All the data entering the data warehouse is integrated. The data warehouse acts as a single data source for various data marts to ensure integrity and consistency across the enterprise.

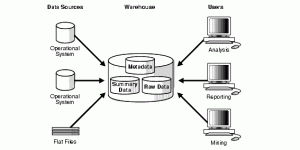


Figure 3. Basic Bill Inmon data warehousing architecture explained (Source: Stanford University)

#### Advantages of the Inmon Method

The Bill Inmon design approach offers the following benefits :

* Data warehouse acts as a unified source of truth for the entire business, where all data is integrated.
* This approach has very low data redundancy. So, there’s less possibility of data update irregularities, making the ETL-concept based data warehouse process more straightforward and less susceptible to failure.
* It simplifies business processes, as the logical model represents detailed business objects.
* This approach offers greater flexibility, as it’s easier to update the data warehouse in case there’s any change in the business requirements or source data.
* It can handle diverse enterprise-wide reporting requirements.

#### Disadvantages of the Inmon Method

The possible drawbacks of this approach are as follows:

* Complexity increases as multiple tables are added to the data model with time.
* Resources skilled in data warehouse data modeling are required, which can be expensive and challenging to find.
* The preliminary setup and delivery are time-consuming.
* Additional ETL process operation is required since data marts are created after the creation of the data warehouse.
* This approach requires experts to manage a data warehouse effectively.

#### Criteria for choosing Inmon approach

| Business decision support requirements | Strategic |
| --- | --- |
| Data integration requirements | Enterprise-wide integration |
| The structure of data | Data that meet multiple and varied information needs and non-metric data |
| Persistence of data in source systems | Source systems have a high rate of change |
| Skill sets | Bigger team of specialists |
| Time constraint | Longer time is allowed to meet business needs. |
| Cost to build | High start-up costs |