

Data Science & Business Intelligence Methodologies

1. Introduction

Data plays a central role in decision-making. Methodologies provide structured approaches to ensure quality, efficiency, and success in Data Science and Business Intelligence (BI) projects.

We'll explore some of the most recognized methodologies:

- CRISP-DM
- IBM Master Plan
- Ralph Kimball's Approach
- GIMSI Method

2. Data Science Methodology

CRISP-DM (Cross Industry Standard Process for Data Mining)

The most widely used methodology for data mining and data science projects.

Phases:

1. Business Understanding - define objectives and requirements.
2. Data Understanding - collect, describe, and explore the data.
3. Data Preparation - cleaning, transforming, integrating datasets.
4. Modeling - applying machine learning/statistical models.
5. Evaluation - assess models against business goals.
6. Deployment - implement in production, monitor results.

Strengths: Iterative, flexible, domain-independent.

Use case: Predictive analytics, fraud detection, customer segmentation.

IBM Master Plan

Developed by **IBM** for guiding data science projects.

Provides a **structured and iterative framework** similar to CRISP-DM but adapted to IBM's ecosystem.

Phases:

1. **Business Understanding** - clarify business problem.
2. **Analytic Approach** - define type of analytics to use (descriptive, predictive, prescriptive).

3. **Data Requirements** - identify data needed.
4. **Data Collection** - gather and integrate data sources.
5. **Data Understanding & Preparation** - explore, clean, transform data.
6. **Modeling & Evaluation** - build, test, and evaluate models.
7. **Deployment** - implement model into business processes.
8. **Feedback & Refinement** - monitor and improve over time.

🔑 **Strengths:** Business-focused, iterative, integrates well with enterprise environments.

💡 **Use case:** Enterprise-level data science projects with strong alignment between IT and business.

3. Business Intelligence Methodologies

Ralph Kimball Methodology (Dimensional Modeling)

A bottom-up approach for data warehouses. Based on star schemas and dimensional modeling.

Steps:

1. Identify business process.
2. Choose grain (level of detail).
3. Identify dimensions (who, what, where, when).
4. Identify facts (measures/metrics).
5. Build star schema.

Strengths: Fast delivery, user-friendly, strong for OLAP & reporting.

Use case: Retail sales analysis, financial reporting.

GISI Method (Generalized Information Management System Integration)

Methodology for developing BI systems. Emphasizes user needs and iterative development.

Phases:

1. Preliminary study (identify BI needs).
2. Conceptual modeling (define business concepts).
3. Logical modeling (map to database structures).
4. Physical implementation (DW + ETL + reporting tools).
5. Deployment & evaluation.

Strengths: Structured, iterative, clear link between business needs & technical design.

Use case: Medium to large BI projects with strong data modeling needs.

4. Comparison

Methodology	Domain	Approach	Key Strength
CRISP-DM	Data Science	Iterative	Flexible, domain-independent
IBM Master Plan	Data Science	Top-down	Strategic alignment
Ralph Kimball	Data Warehousing	Bottom-up	User-friendly, fast delivery
GIMSI	BI/DW	Iterative	Strong modeling, business-driven

5. Conclusion

CRISP-DM is best for predictive modeling and analytics.

IBM Master Plan is ideal for enterprise-wide BI strategy.

Ralph Kimball focuses on practical dimensional modeling for warehouses.

GIMSI bridges business requirements and technical implementation.

Choosing the right methodology depends on project scope, goals, and organizational context.