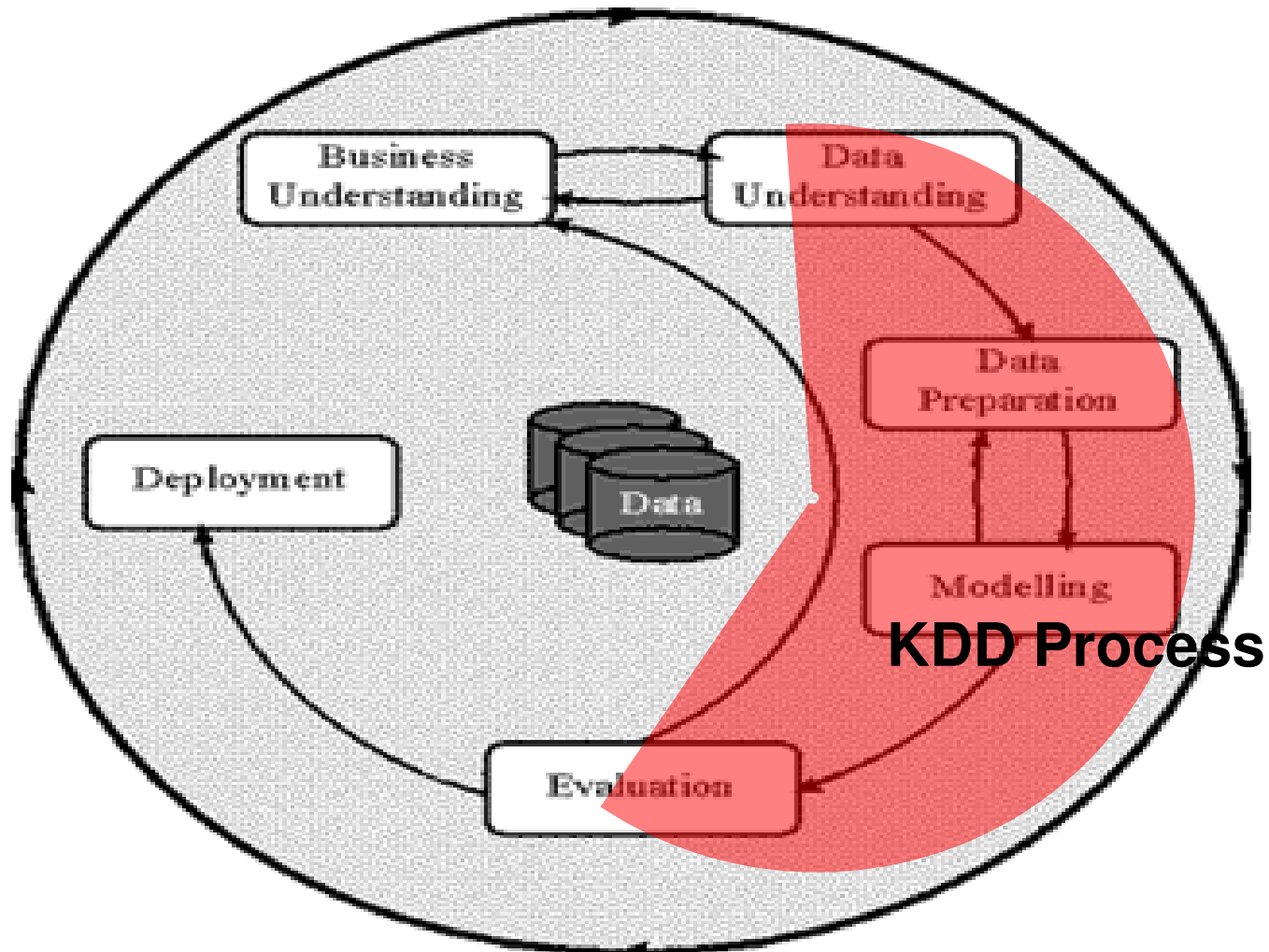


CRISP-DM: The life cycle of a data mining project



Business understanding

- Understanding the project objectives and requirements from a business perspective.
- then converting this knowledge into a data mining problem definition and a preliminary plan.
 - **Determine the Business Objectives**
 - **Determine Data requirements for Business Objectives**
 - **Translate Business questions into Data Mining Objective**



Determine Business Objective

Background

Business Objective

Business Success Criteria

Assess Situation

Inventory of Resources

Requirements Assumptions Constraints

Risk and Contingencies

Terminology

Costs & Benefits

Determine Data Mining Goals

Data Mining Goals

Data Mining Success Criteria

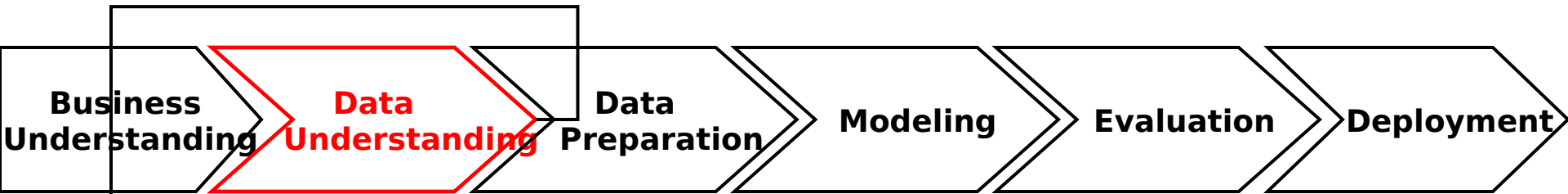
Produce Project Plan

Project Plan

Assessment Of Tools and Techniques

Data understanding

- **Data understanding:** characterize data available for modelling. Provide assessment and verification for data.



**Collect
Initial
Data**

**Initial Data
Collection
Report**

**Describe
Data**

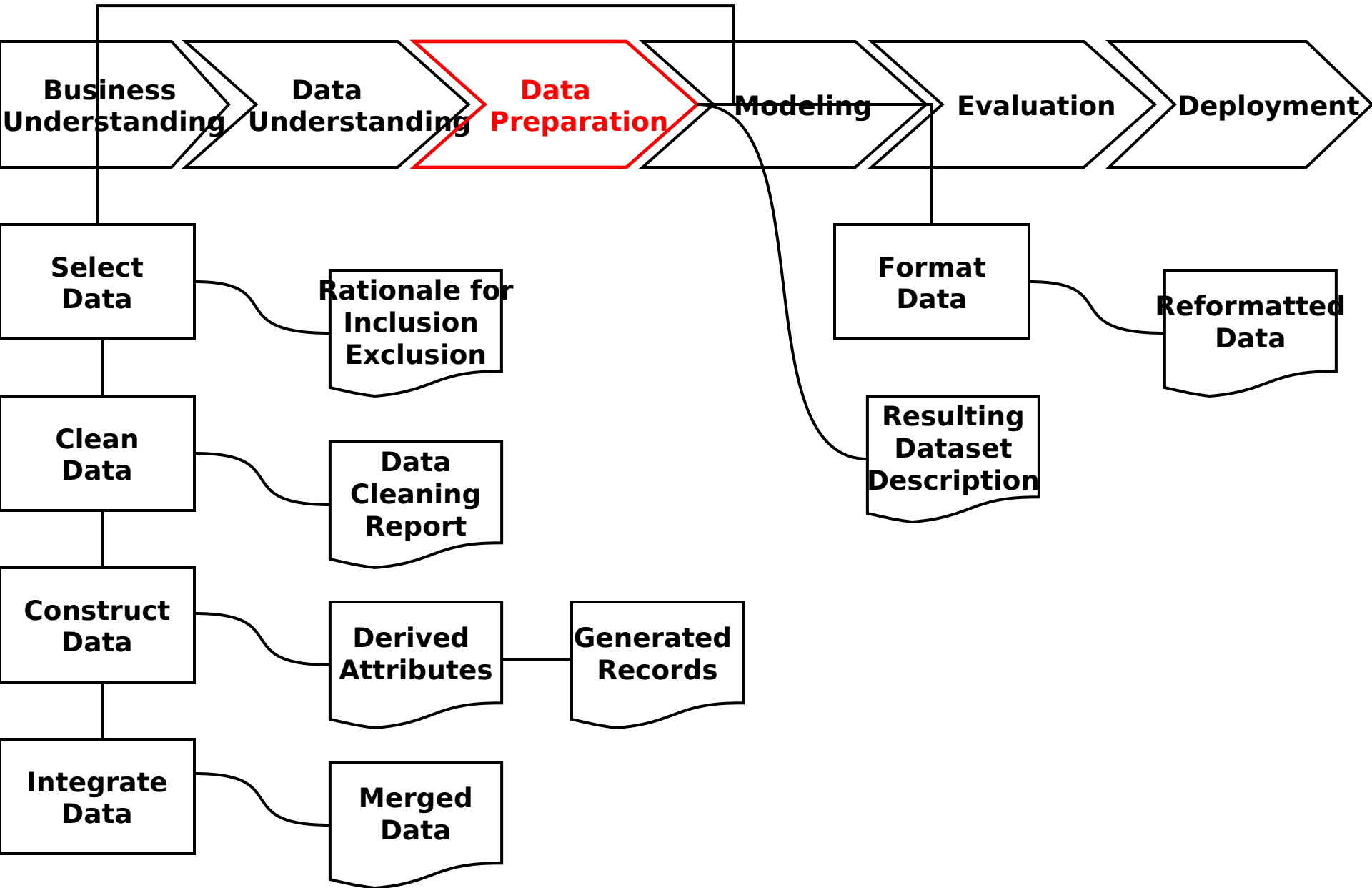
**Data
Description
Report**

**Explore
Data**

**Data
Exploration
Report**

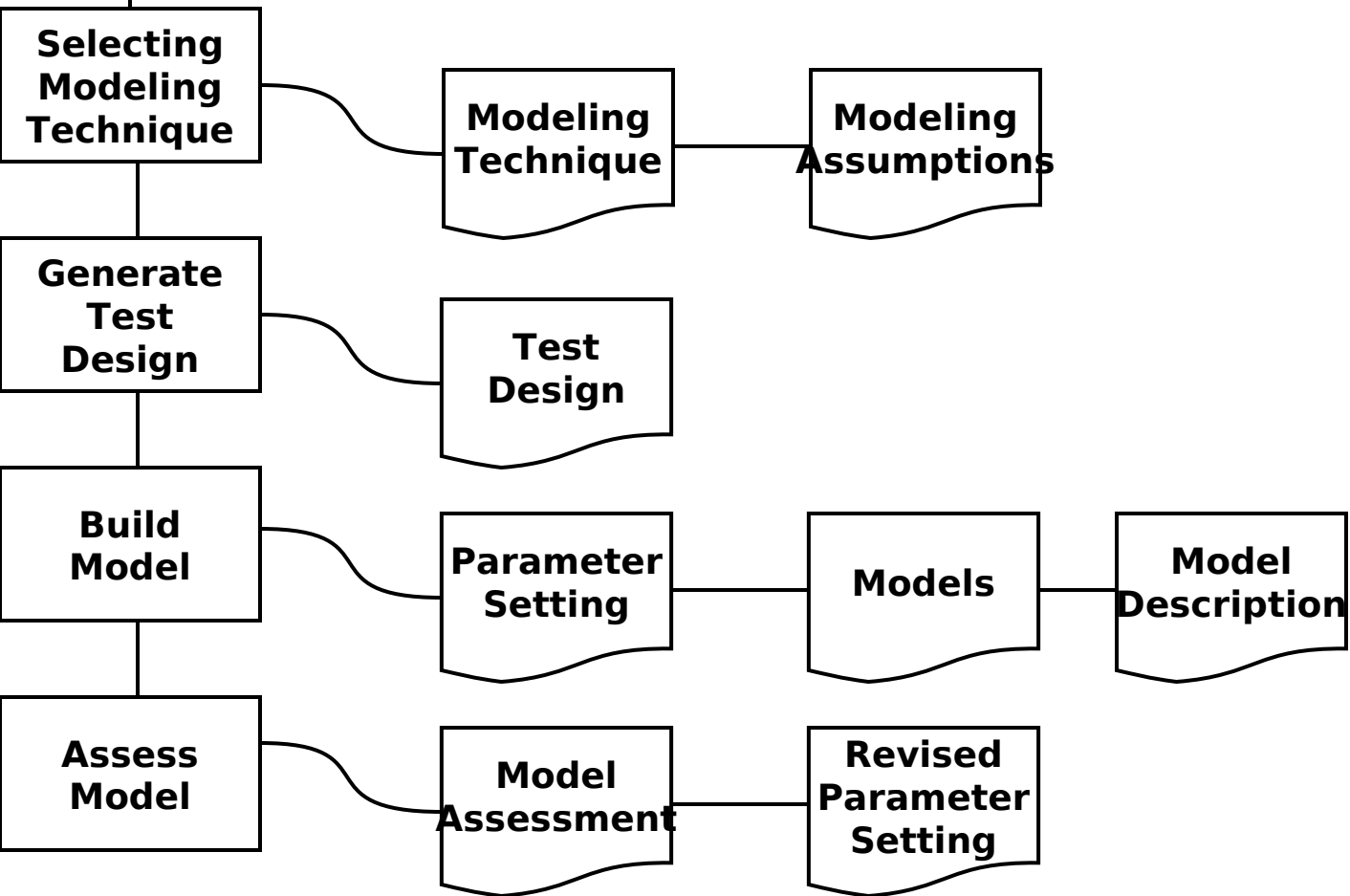
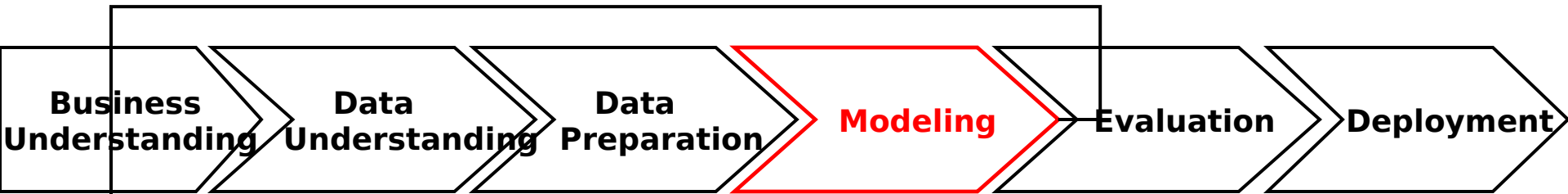
**Verify
Data
Quality**

**Data
Quality
Report**



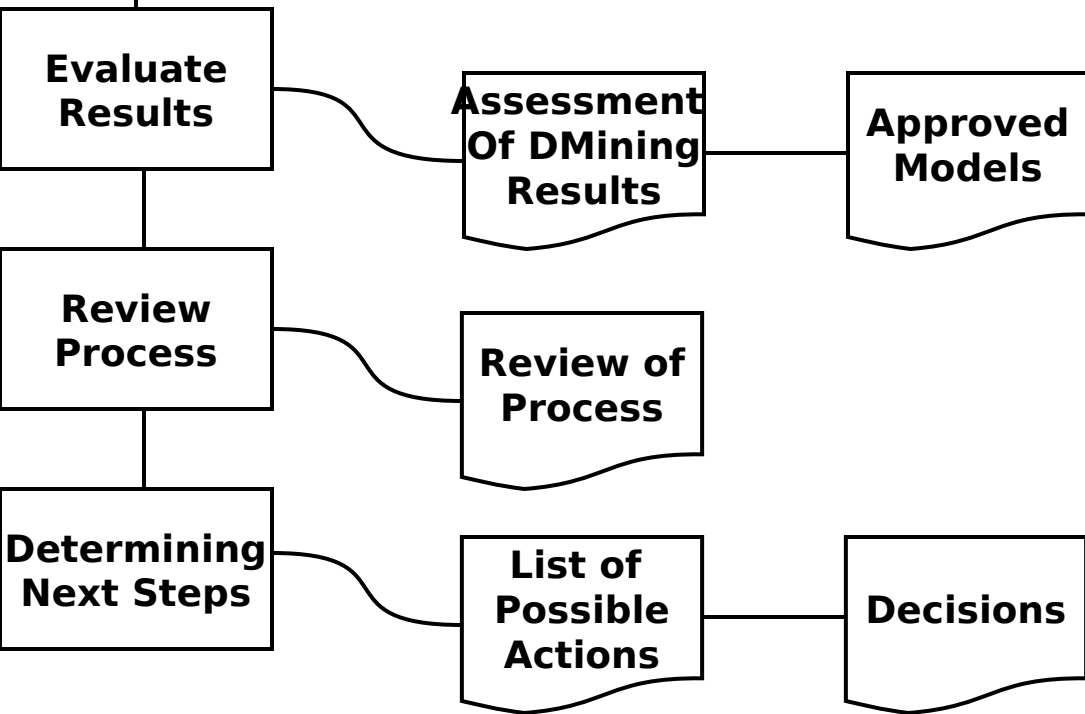
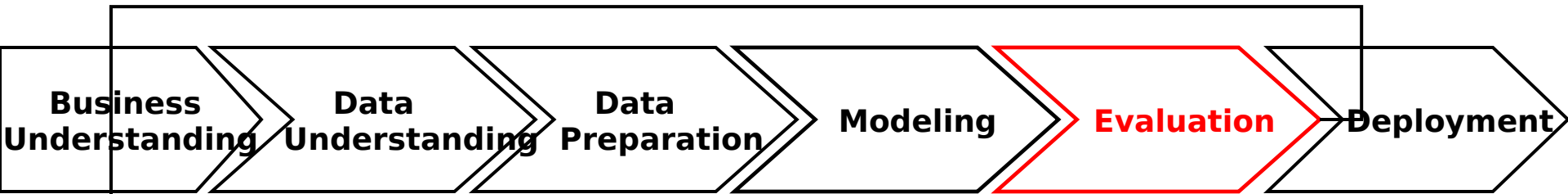
Modeling:

- In this phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values.
- Typically, there are several techniques for the same data mining problem type. Some techniques have specific requirements on the form of data.
- Therefore, stepping back to the data preparation phase is often necessary.



Evaluation

- At this stage in the project you have built a model (or models) that appears to have high quality from a data analysis perspective.
- Evaluate the model and review the steps executed to construct the model to be certain it properly achieves the business objectives.
- A key objective is to determine if there is some important business issue that has not been sufficiently considered.

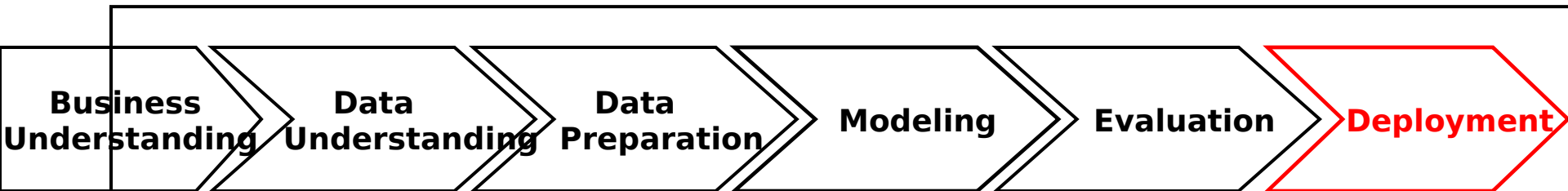


Deployment:

- The knowledge gained will need to be organized and presented in a way that the customer can use it.
- It often involves applying “live” models within an organization’s decision making processes, for example in real-time personalization of Web pages or repeated scoring of marketing databases.

Deployment:

- It can be as simple as generating a report or as complex as implementing a repeatable data mining process across the enterprise.
- In many cases it is the customer, not the data analyst, who carries out the deployment steps.



**Plan
Deployment**

**Deployment
Plan**

**Plan
Monitoring and
Maintenance**

**Monitoring
and
Maintenance
Plan**

**Produce
Final
Report**

**Final
Report**

**Final
Presentation**

**Review
Project**

**Experience
Documentation**