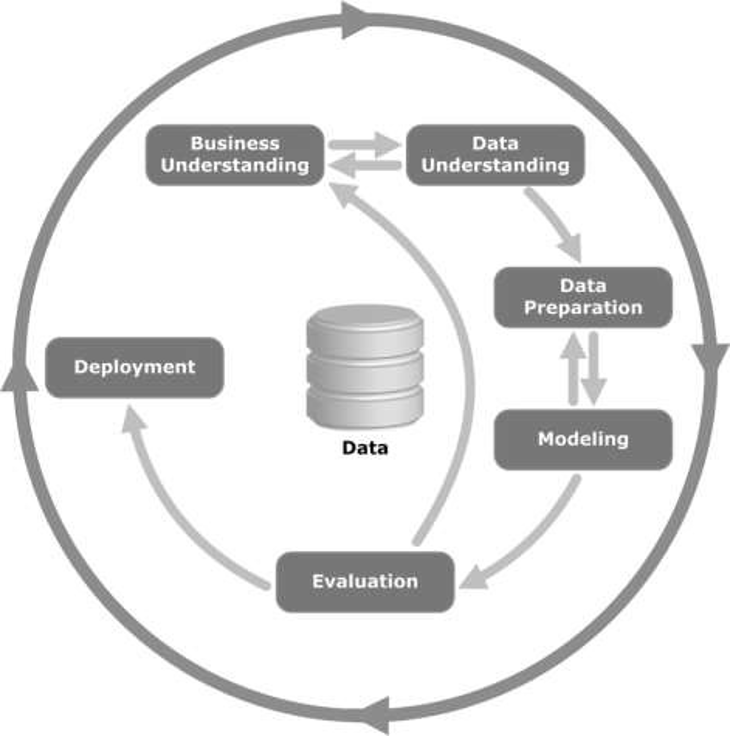
### CRISP-DM (Cross Industry Standard Process for Data Mining) Methodology

The CRISP-DM methodology consists of a cyclical sequence of phases commonly used by data mining experts for traditional business intelligence (BI) data mining and data strategy development. Figure 3 (below) summarizes this cycle, which is dynamic, features bidirectional movement between the six phases, and entails multiple iterations. In the diagram:

* The Iterative CRISP-DM process is shown in the outer circle.
* The most significant dependencies between phases are shown.
* Subsequent phases depend on results from the preceding phases.
* Returning to earlier phases is possible before moving forward.



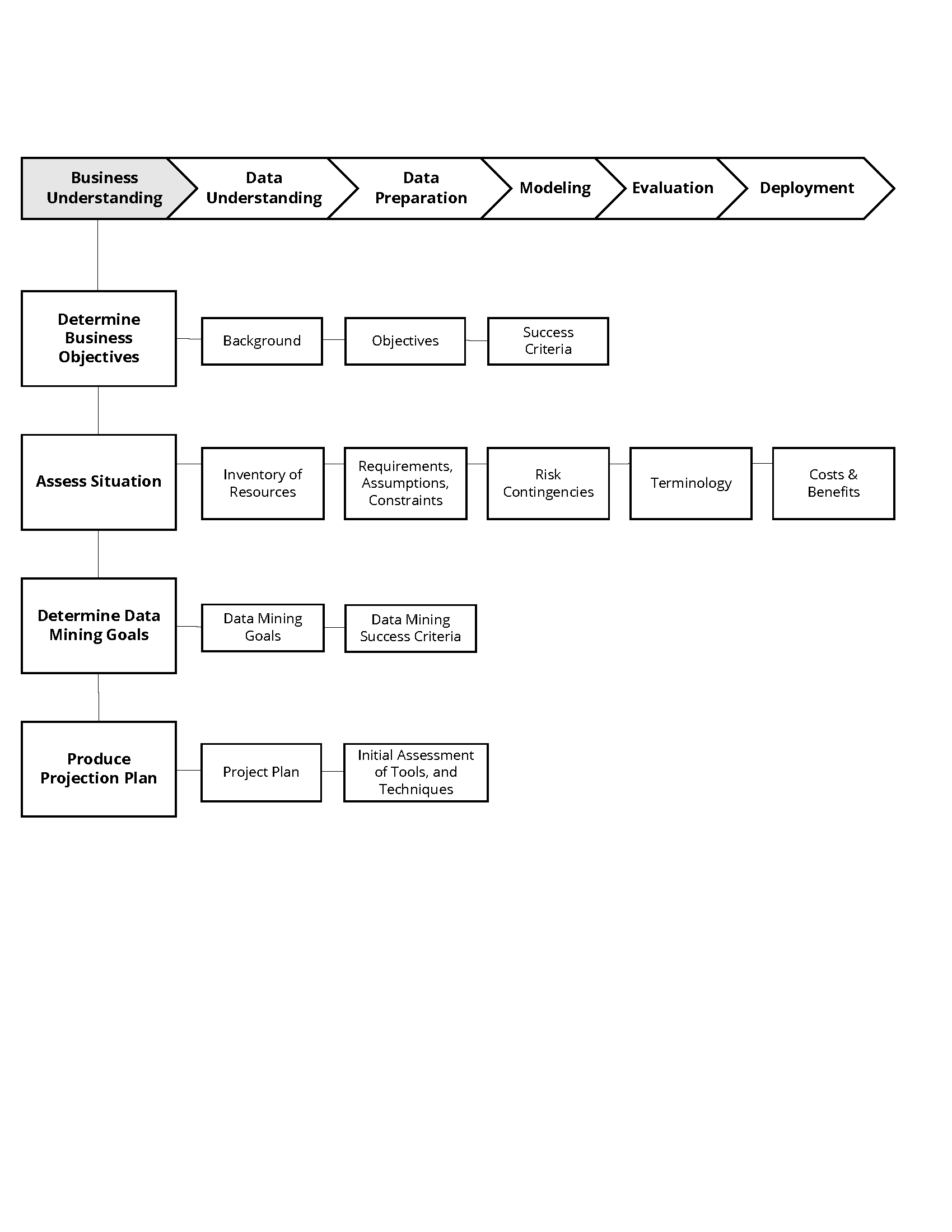
*Image source:* [*https://commons.wikimedia.org/wiki/File:CRISP-DM\_Process\_Diagram.png*](https://commons.wikimedia.org/wiki/File:CRISP-DM_Process_Diagram.png)

*Figure 3: The CRISP-DM process*

Specific activities for each phase are highlighted below:

### Business/Research Understanding Phase (Figure 4)

* Define project requirements and objectives
* Translate objectives into a data mining problem definition
* Prepare a preliminary strategy to meet the objectives



*Figure 4: The business understanding phase*

### Data Understanding Phase (Figure 5)

* Collect data
* Perform exploratory data analysis (EDA)
* Assess data quality
* Select

Ethnic

Variance <-- > bias

Validity <-🡪 reliability

Target prediction

\*\* -🡪 &&&&: validity

&

& \*\* & : reliability

&

Table A customer Key

Customer ID

Customer number

Cusomter age

Customer salary

Xxx purchase history

Table B product

Customer ID

Product ID

Spend

Datetime

Return

Profit

Table C customer per production

Custoemer ID

Number of product

Spend on monthly

Sale by day