

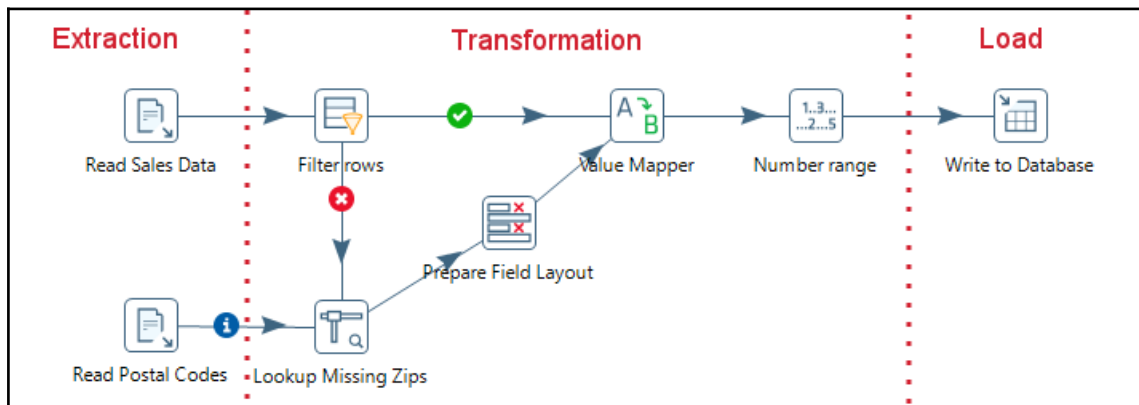
Loading data warehouses or data marts

The loading of a data warehouse or a data mart involves many steps, and there are many variants depending on business area or business rules.

However, in every case, with no exception, the process involves the following steps:

1. Extracting information from one or more databases, text files, XML files, and other sources. The extract process may include the task of validating and discarding data that doesn't match expected patterns or rules.
2. Transforming the obtained data to meet the business and technical needs required on the target. Transforming includes such tasks such as converting data types, doing some calculations, filtering irrelevant data, and summarizing.
3. Loading the transformed data into the target database or file store. Depending on the requirements, the loading may overwrite the existing information or may add new information each time it is executed.

Kettle comes ready to do every stage of this loading process. The following screenshot shows a simple ETL designed with the tool:



ETL process