Work with Data

You can refine your Pentaho relational metadata and multidimensional Mondrian data models. You can also learn how to work with big data.

These sections are helpful if you are a data designer, data scientist, or developer.

Data Modeling

Pentaho data modeling include:

- **Multidimensional Modeling**
  Pentaho Analyzer and Report Designer are built on the Mondrian online analytical processing (OLAP) engine, which relies on a multidimensional data model.

- **Relational Modeling**
  A Pentaho relational data model maps the physical structure of your database into a logical business model.

- **Streamlined Data Refinery**
  Streamlined Data Refinery (SDR) augments and blends raw data through a request form, and then publishes it for use in Analyzer.

Big Data

Pentaho supports Hadoop and Spark for the entire big data analytics process from big data aggregation, preparation, and integration to interactive visualization, analysis, and prediction.

- **Hadoop**
  Pentaho Data Integration (PDI) can execute both outside of a Hadoop cluster and within the nodes of a Hadoop cluster.

- **Spark**
  PDI can execute Spark jobs through a Spark Submit entry or the Adaptive Execution Layer (AEL).