Pentaho Data Modelling

James O'Reilly

© Hitachi, Ltd. 2018. All Rights Reserved

Agenda

HITACHI
Inspire the Next

- Overview of Pentaho Analytics Platform
- Metadata Editor
- Schema Workbench
- Datasource Wizard

© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

Pentaho Business Analytics Platform

© Hitachi, Ltd. 2018. All Rights Reserved

Overview of Pentaho Platform

HITACHI
Inspire the Next

Server Side Tools and Plugins

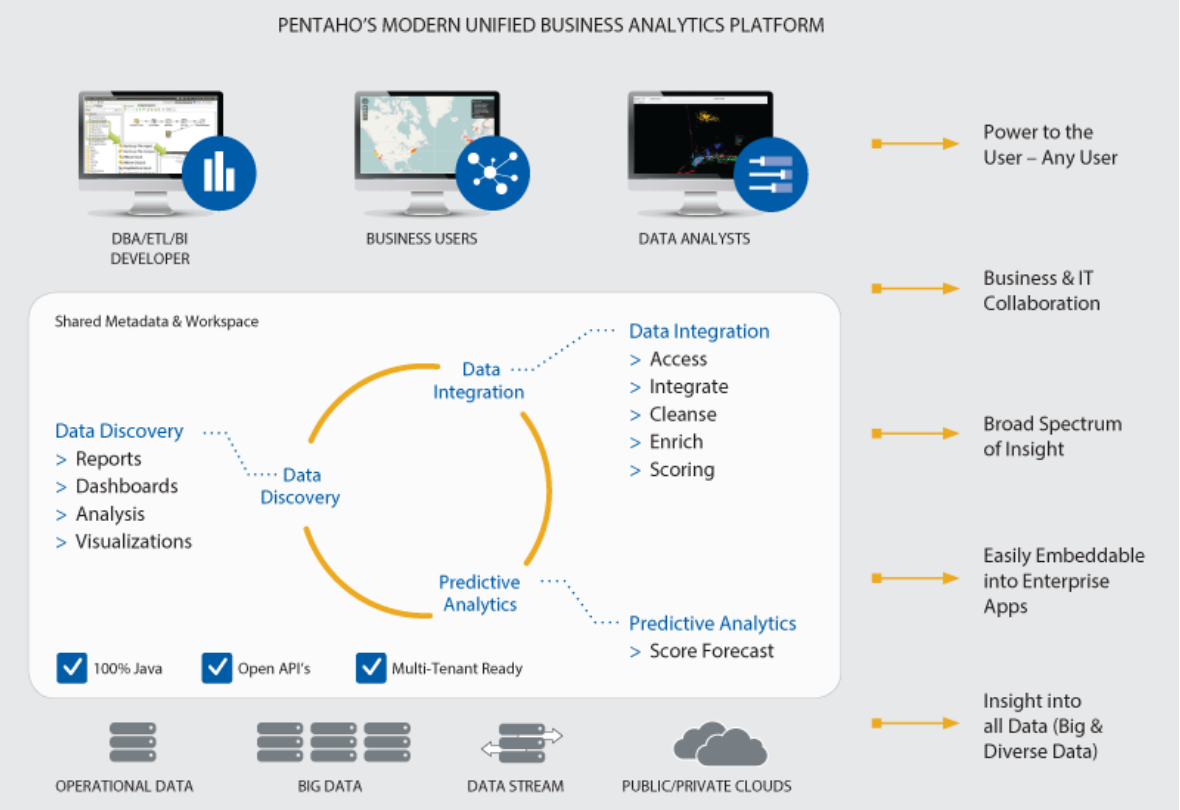
- Analyzer
- Interactive Reporting
- Dashboard Designer

Client side Tools and Plugins

- Data Integration
- Metadata Editor
- Report Designer
- Schema Workbench
  - Aggregation Designer

© Hitachi, Ltd. 2018. All Rights Reserved

Overview of Pentaho Platform



© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

Metadata Editor

© Hitachi, Ltd. 2018. All Rights Reserved

Topics

HITACHI
Inspire the Next

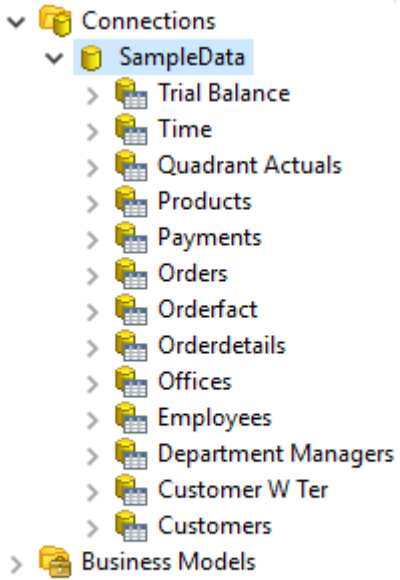
- Overview Metadata Editor
- Physical Layer
- Business Layer
- Key Relationships
- Business Views

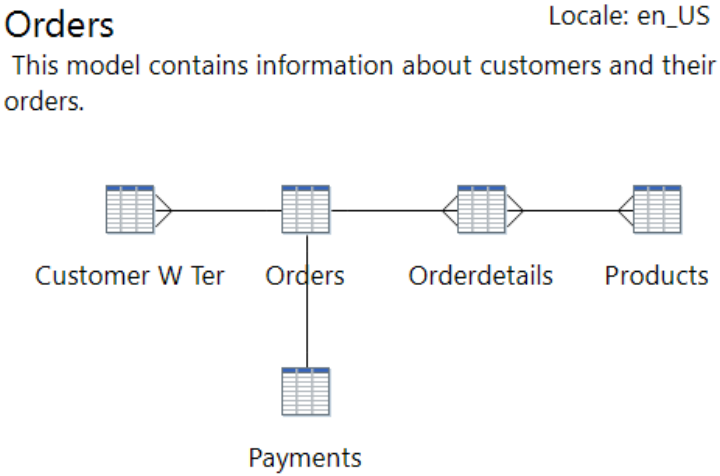
© Hitachi, Ltd. 2018. All Rights Reserved

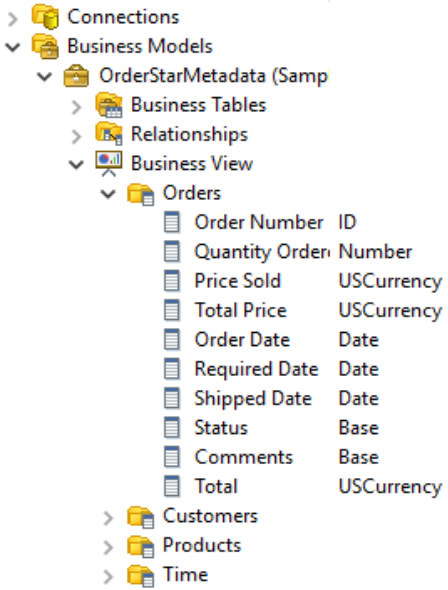
Metadata Editor

HITACHI
Inspire the Next

- Pentaho Metadata Editor (PME) is a tool that builds Pentaho metadata domains and models. A Pentaho metadata model maps the physical structure of your database into a logical business model.







© Hitachi, Ltd. 2018. All Rights Reserved

Metadata Domain

HITACHI
Inspire the Next

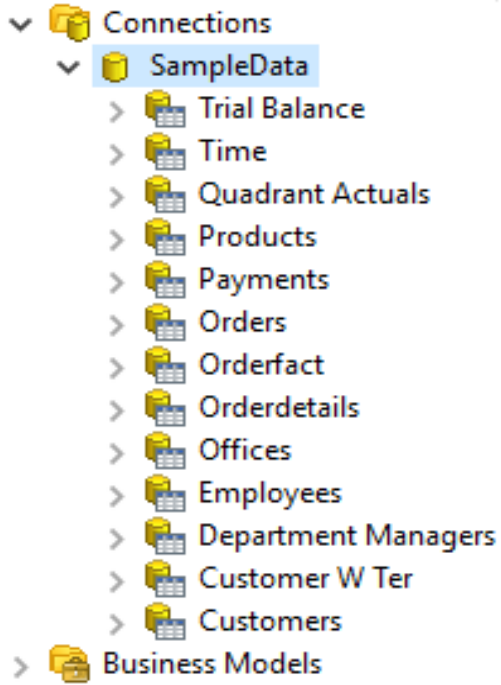
- The metadata domain is a Pentaho term that represents all of the business objects created, stored and used in the metadata layer.

- A domain may consist of one or more connections, one or more models, security information, business tables, business views, categories, columns and concepts. You can create and save multiple metadata domains using the Metadata Editor.

- A metadata domain is accessed by the Pentaho Server by publishing or exporting the domain to an .xmi file, and placing the file in the Pentaho solutions repository.

© Hitachi, Ltd. 2018. All Rights Reserved

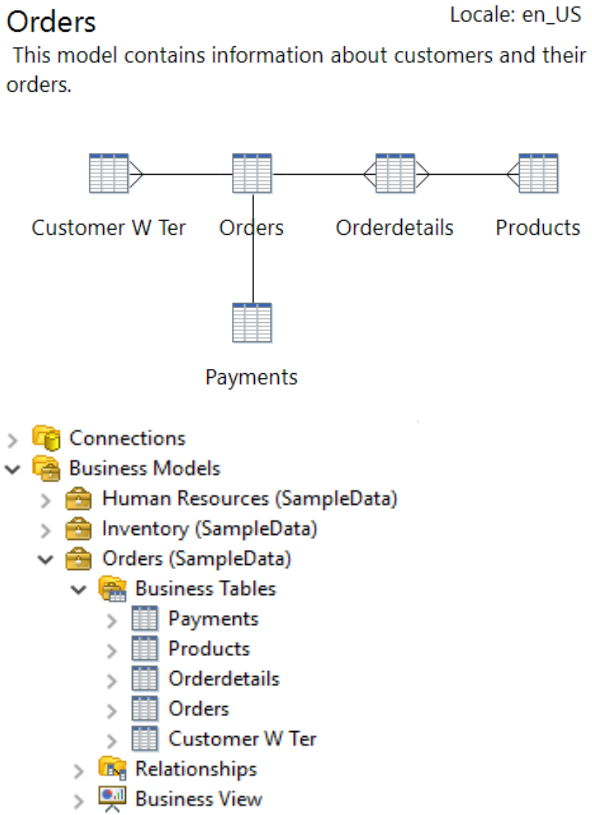
Physical Layer



- The Physical Layer of a Pentaho domain encompasses connections, physical tables and physical columns. These objects represent the database(s) you are trying to model and enrich with metadata.
- The Physical Layer is not considered part of the business model, because not all connections defined in the Physical Layer will be used in every business model.
- Only one connection per domain. (You can get around this limitation by editing the xml file - consultancy).

© Hitachi, Ltd. 2018. All Rights Reserved

Abstract Business Layer

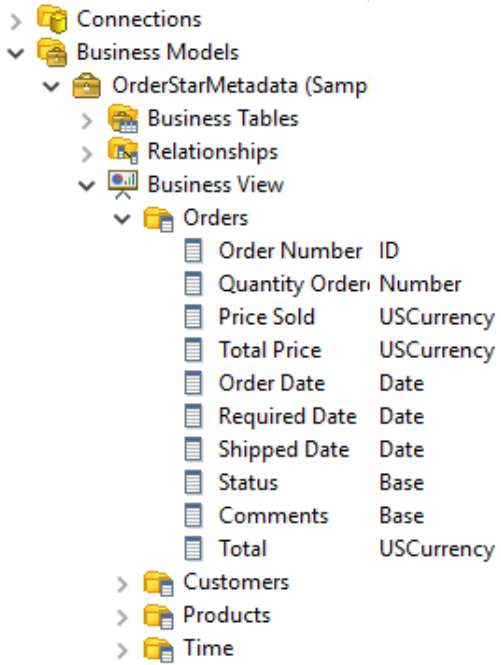


- In the Abstract Business Layer, you have business tables, business columns, and business relationships.
- You can create business tables for any physical table you have defined in the Physical Layer. You can also create more than one business table to reference the same physical table.
- The business table keeps a reference to the physical table that it models, and this allows a metadata inheritance relationship between physical tables and business tables.
- If you define metadata on a physical table, the business table will inherit that metadata, unless and until the business table itself has overridden the inherited metadata.
- This concept operates on a metadata property to property basis, meaning that each property can be overridden individually.

Hitachi Vantara Confidential. All rights Reserved

Business Views

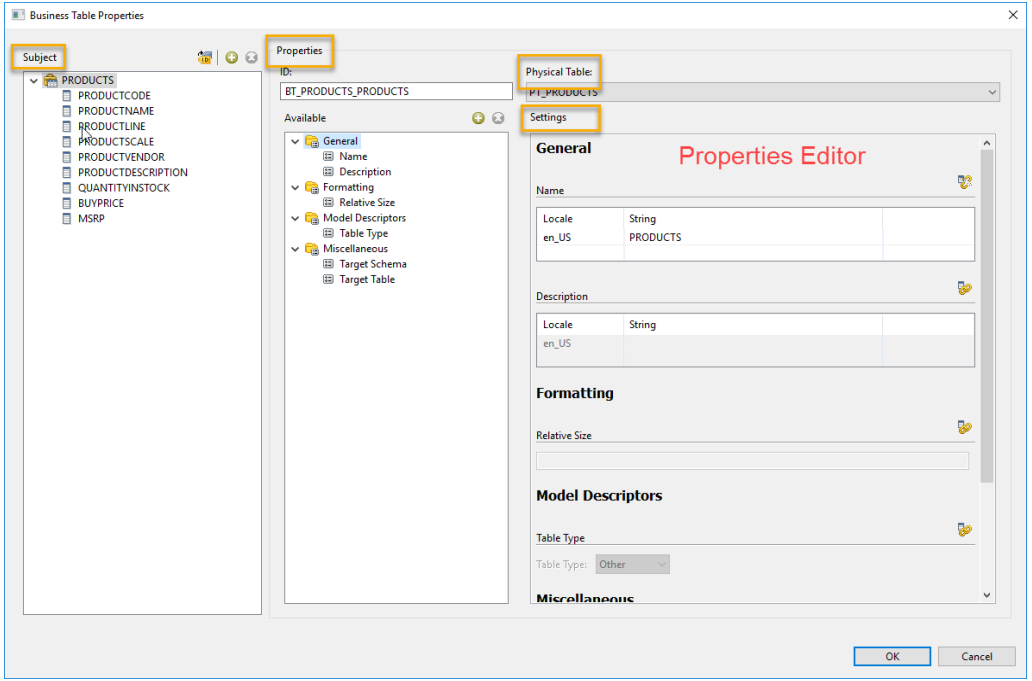
HITACHI
Inspire the Next



- The Business View is the part of the business model that applications will operate against, and end-users will see.
- The Business View is nothing more than "buckets" (called categories) for you to re-arrange and re-organize your business columns in a fashion that makes sense to the consumers of the data.
- In the Business View, you can create any number of categories and arrange your business columns in those categories however it best suits your business, mixing and matching columns that derived from different business tables, even adding business columns more than once to different categories.

© Hitachi, Ltd. 2018. All Rights Reserved

Metadata Concepts



- Each business object in the domain can have metadata associated with it, with the exception of categories. In Pentaho terminology, a collection of metadata properties is called a concept.
- Each business object can have three levels of concepts:
  - Self-concepts – are applied to the property directly in the editor
  - Parent concepts – are applied to the property by setting the concept in the business view
  - Inherited concepts – are inherited from other concepts in the hierarchy

© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

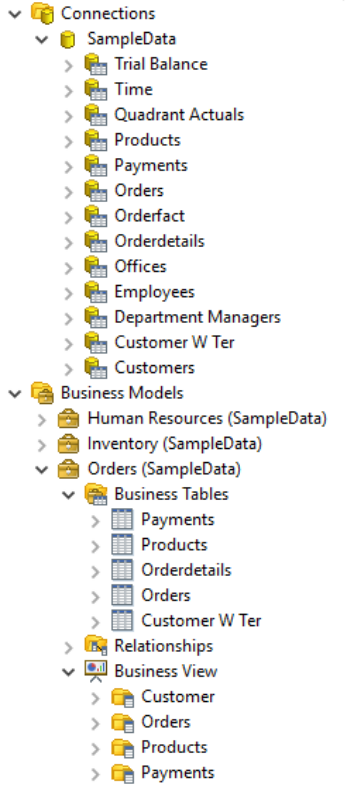
Guided Demonstration

steel-wheels.xmi

© Hitachi, Ltd. 2018. All Rights Reserved

Guided Demo: steel-wheels.xmi

HITACHI
Inspire the Next



- In this demonstration we will open Metadata Editor and view the steel-wheels.xmi domain.
  - To import (open) an existing metadata domain:
  - From the menu, select File > Import from XMI File.
  - Navigate to \pentaho\design-tools\metadata-editor\samples.
  - Double-click steel-wheels.xmi.
  - In the Save Model dialog, type SteelWheelsInc.
  - Click OK.

© Hitachi, Ltd. 2018. All Rights Reserved

Summary

HITACHI
Inspire the Next

- Overview of Metadata Editor
  - Physical Layer
  - Business Layer
  - Key relationships
  - Business View
- Guided Demo: View an Existing Metadata Domain (steel-wheels.xmi)

© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

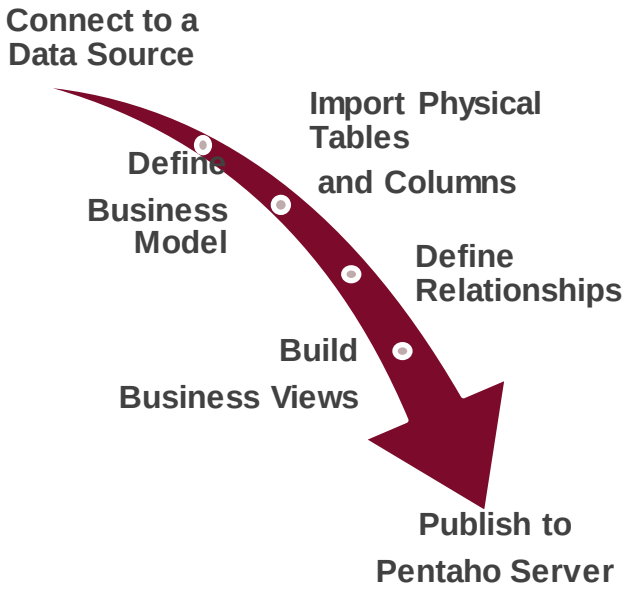
Guided Demonstration

Define a Domain: OrdersME

© Hitachi, Ltd. 2018. All Rights Reserved

Guided Demo: OrdersME Domain

HITACHI
Inspire the Next



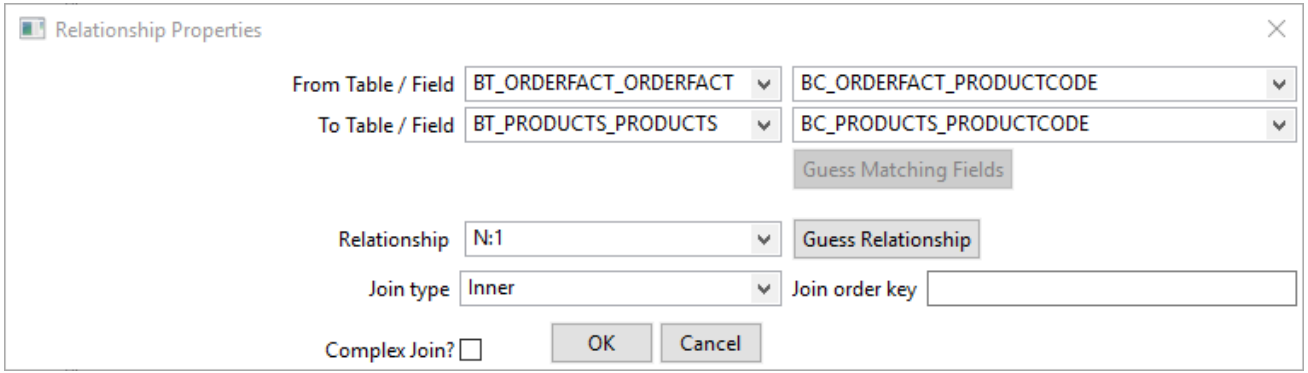
- Define a Domain
  - Connect to Data Source
  - Import Physical Tables and Columns
  - Define a Business Model
  - Define Relationships between tables
  - Build a Business View
  - Publish to BA (Pentaho) server
  - Create an Interactive Report
- Guided Demo: Creating a Metadata Domain
- Exercise: Create an OrdersStarMetadata Domain

© Hitachi, Ltd. 2018. All Rights Reserved

Relationship: OrdersME Domain

HITACHI
Inspire the Next

- Once you have all of your business tables created, you will need to define the relationships between the tables, so that the query generators and SQL generators that work with Pentaho metadata can create the data queries correctly.



© Hitachi, Ltd. 2018. All Rights Reserved

Relationships

HITACHI
Inspire the Next

Relationship	Description
1:N	A **one-to-many** mandatory relationship is the most common relationship in databases. The primary key table contains only one record that relates to none, one, or many records in the related table. This relationship is similar to the one between you and one of your parents. You have one mother, but your mother may have several children.
N:1	A **many-to-one** is opposite of one to many (1:N) relationship.
1:1	In a **one-to-one** relationship, both tables are limited to one record only on either side of the relationship. Each primary key value relates to a single record, or no record, in the associated table. They are like spouses — you may be married, or not; however, if you are married, both you and your spouse can have only one partner. Most one-to-one relationships are forced by business rules. If you do not have a business rule, you can, in most cases, combine both tables into one table without breaking normalization rules.
0:N	A **zero to many** optional relationship indicates that a person may have no phone, one phone, or many phones, and that the phone may not be "owned," but can only be owned by a maximum of one person.
N:0	Opposite of a zero to many relationship
0:1	A **zero to one** relationship might indicate that a person may be a programmer, but a programmer must be a person. It is assumed that the mandatory side of the relationship is the dominant.
1:0	Opposite of a zero to one relationship
N:N	In a many to many relationship each record in both tables can relate to an unlimited number of records (or no records) in the other table. For example, if you have many siblings, your siblings also have many siblings. Many-to-many relationships must have a third table, referred to as an associate or linking table, because relational systems cannot accommodate the relationship directly.
0:0	A **zero to zero** optional relationship indicates that a person may occupy one parking space, but that a person is not necessary to have a space and a space does not need to have a person.

© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

Exercise

OrdersStarMetadata

© Hitachi, Ltd. 2018. All Rights Reserved

Exercise: OrdersStarMetadata

HITACHI
Inspire the Next

- Edit Column Properties - override and edit text
- Apply Aggregation - PRICEEACH change to Price Sold set aggregation to:
  - Average
  - Minimum
  - Maximum
- Override Data Type - Set Orderdate to Date
- Add Formulas - Total = [QUANTITYORDERED]*[PRICEEACH]
- Add Column - PC_CONTACT = [CONTACTLASTNAME]+", "+[CONTACTFIRSTNAME]
- Add Categories

© Hitachi, Ltd. 2018. All Rights Reserved

Summary

HITACHI
Inspire the Next

- Define a Domain
  - Connect to Data Source
  - Import Physical Tables and Columns
  - Define a Business Model
  - Define Relationships between tables
  - Build a Business View
  - Publish to BA (Pentaho) server
  - Create an Interactive Report
- Guided Demo: Creating a Metadata Domain
- Exercise: Create an OrdersStarMetadata Domain

© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

Enriching the Data

OrdersStarMetadata

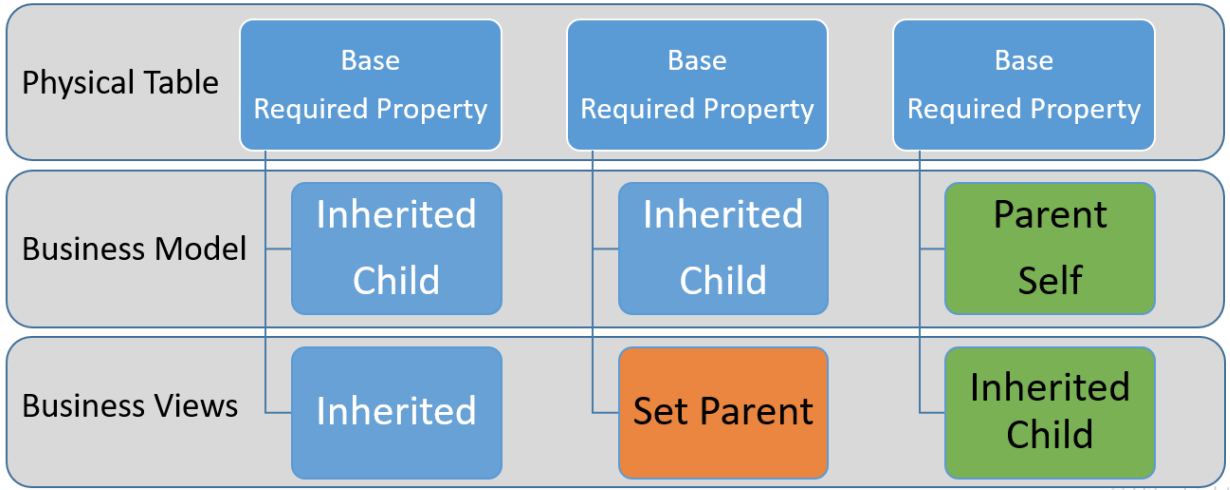
© Hitachi, Ltd. 2018. All Rights Reserved

Concepts

HITACHI
Inspire the Next

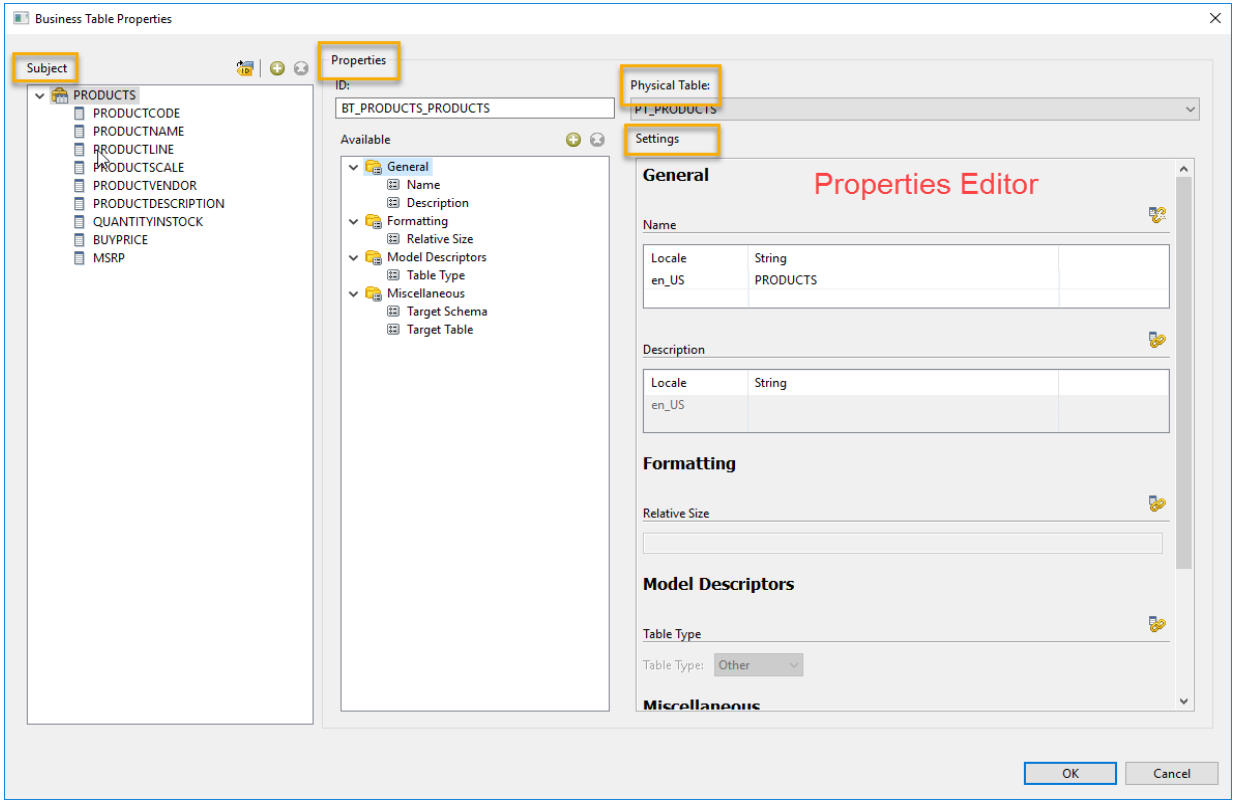
- Metadata concepts are collections of properties that can be applied to business objects, such as currency, number, and text formats. They can be defined on the business object or independently of the object.

- You can create a hierarchy of concepts. In a hierarchy, child concepts inherit from parent concepts. New concepts generally inherit from the pre-configured base concept.



© Hitachi, Ltd. 2018. All Rights Reserved

Concepts

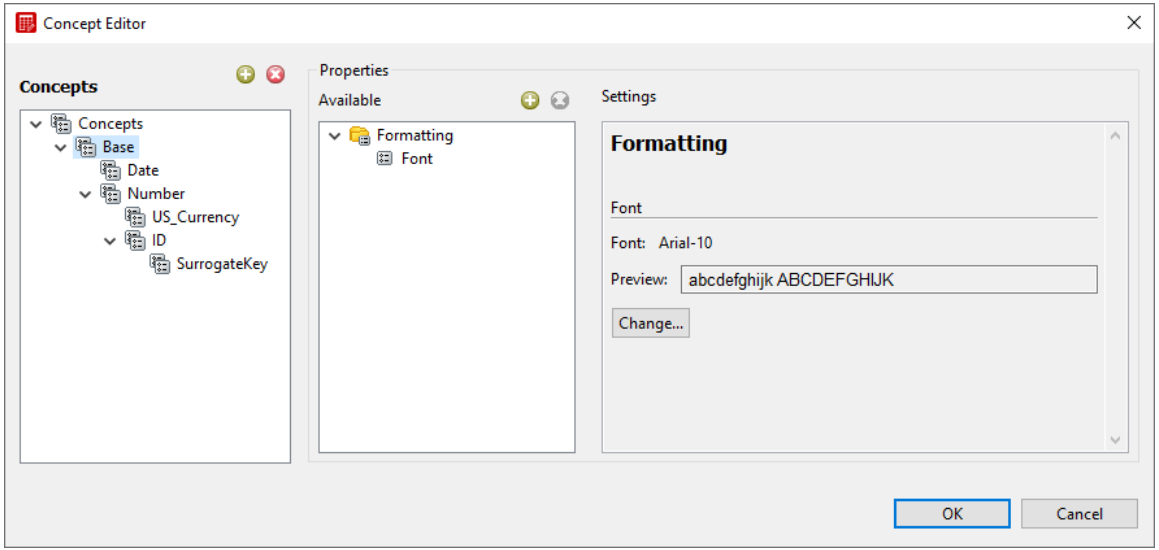


© Hitachi, Ltd. 2018. All Rights Reserved

Concept Editor

HITACHI
Inspire the Next

- Parent concepts are independent hierarchies of concepts that can be assigned to one or more business objects through the navigation tree. Before you can assign a parent concept you must first learn how to create one.



© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

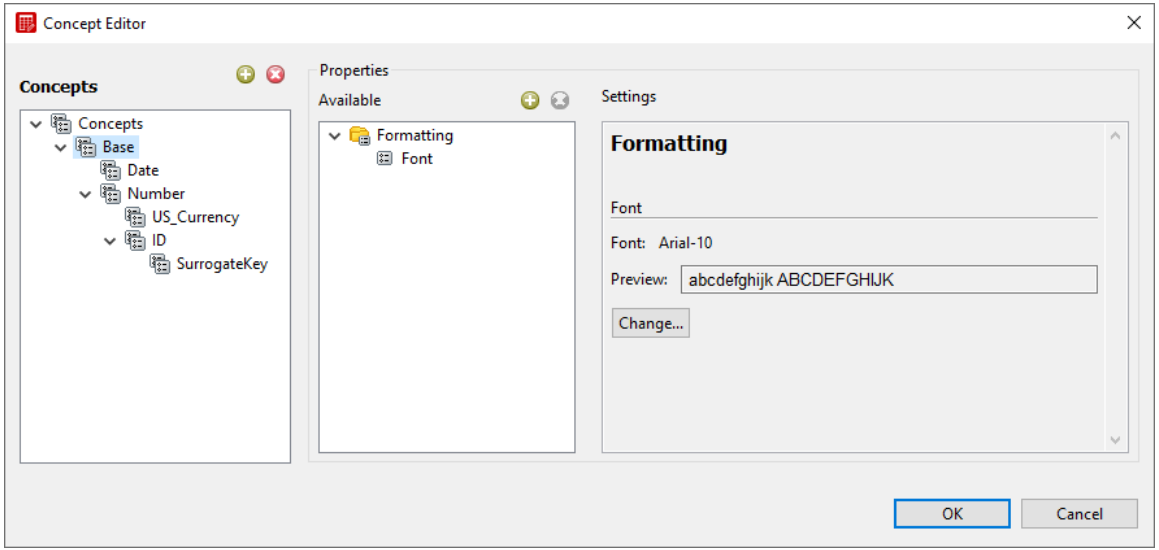
Guided Demonstration

Concept Editor

© Hitachi, Ltd. 2018. All Rights Reserved

Guided Demo: Concept Editor

HITACHI
Inspire the Next



Using the Concept Editor

- Numeric Concept
- Currency Concept
- ID Concept
- Date Concept

- Apply the concepts
- Publish Metadata Domain
- Create Interactive Report

© Hitachi, Ltd. 2018. All Rights Reserved

Summary

HITACHI
Inspire the Next

© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

Security
MQL Constraints

© Hitachi, Ltd. 2018. All Rights Reserved

Security

HITACHI
Inspire the Next

- Adding Security
  - Configure the Security service
  - Add Column Level Security
  - Add Row Level Security
    - Global Constraints
    - User / Role Constraints
- Guided Demo: Adding Security to Metadata Objects

© Hitachi, Ltd. 2018. All Rights Reserved

MQL – Metadata Query Language

HITACHI
Inspire the Next

- MQL is the syntax Pentaho Metadata uses for generating SQL queries based off of metadata. Normally a user would generate MQL via Pentaho's Metadata Editor.
  - A constraint function references business table columns and uses various comparison operators to determine which subset of data the business user is interested in.

```
BT_CUSTOMERS.BC_CUSTOMERS_CUSTOMERNAME] = "EuroCars";
(([BT_CUSTOMERS.BC_CUSTOMERS_CREDITLIMIT] * 2) / 3 > 1000))
```

  - Here is an example of a Physical Table Column Formula:

```
[QUANTITYORDERED]*[PRICEEACH]
```


© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

Guided Demonstration

Applying Security Constraints

© Hitachi, Ltd. 2018. All Rights Reserved

Guided Demo: Security

HITACHI
Inspire the Next

- Pentaho metadata provides a Security Information property that allows you to define table or column level security that the Pentaho BI Server can make use of.
- Before you can use this property, you need to tell the Pentaho Metadata Editor about your Pentaho BI Server, so that the program can retrieve the list of Users, Roles and Access Control Lists needed.
- Configure security services
- Column Level security constraints
- Row Level security constraints
  - Global – all users
  - User or Role based

© Hitachi, Ltd. 2018. All Rights Reserved

Summary

HITACHI
Inspire the Next

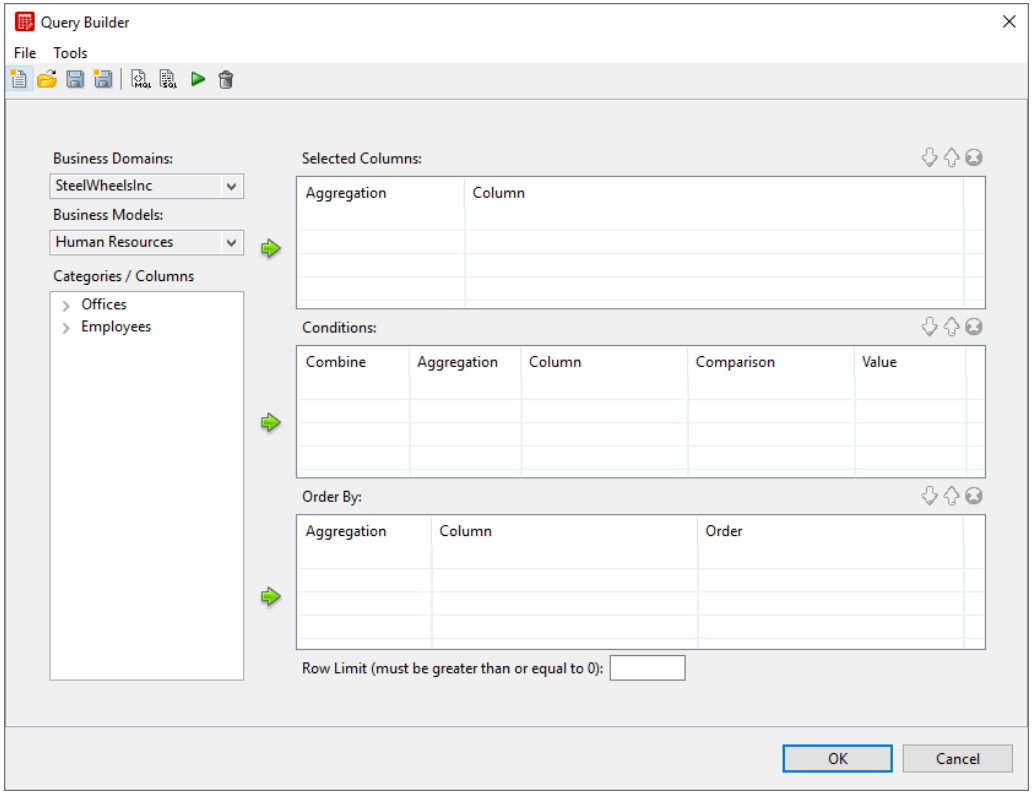
© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

Query Builder

© Hitachi, Ltd. 2018. All Rights Reserved

Query Builder



- Once all the metadata data source options have all been defined, you can use the Query Editor to build and execute a query to test the data source.
- The results do not incorporate formatting, so this is not a thorough test of the data source; however, it is a good tool for testing relationships and calculated values.

© Hitachi, Ltd. 2018. All Rights Reserved

HITACHI
Inspire the Next

Summary

© Hitachi, Ltd. 2018. All Rights Reserved

Summary

HITACHI
Inspire the Next

- Metadata Editor
- Schema Workbench
- Datasource Wizard

© Hitachi, Ltd. 2018. All Rights Reserved

© Hitachi, Ltd. 2018. All Rights Reserved