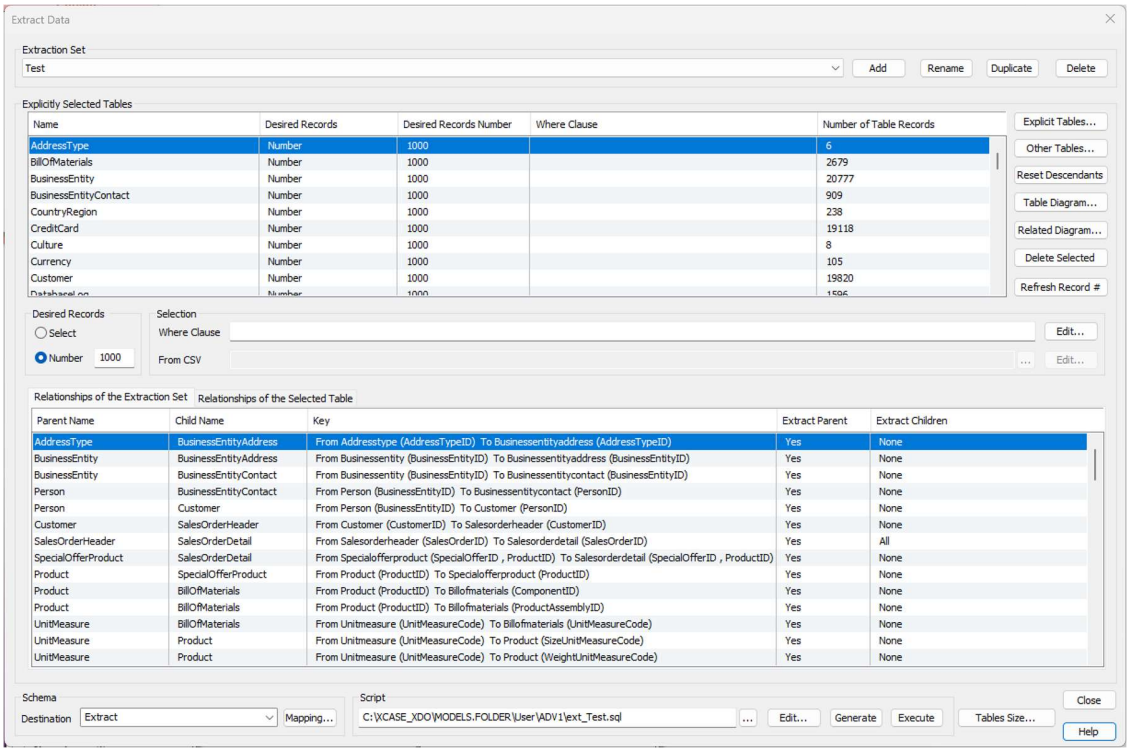
Extracting Data from Prod Databases



The Extract Data Dialog is launched by selecting "Database" and then "Extract Data" from the main Xcase menu.

This Dialog generates an SQL script that can be run to relationally extract data from the database represented in the model.

The extraction helps update Dev Databases using Prod Data. It can be used to refresh the Dev data regularly and to quickly extract compact yet coherent sets of data from Prod when debugging requires specific data.

Extraction Set Group

Multiple extraction sets can be defined and saved by using the different Buttons in the Extraction Set group.

- The drop-down allows you to select an existing Extraction Set.
- The Add Button allows you to create a new Extraction Set.
- The Rename Button allows you to rename it.
- The Duplicate Button allows you to duplicate the currently selected Set and to name it uniquely.

- The Delete Button allows you to delete the currently selected Set.

Explicitly Selected Tables list

This list displays the tables that you have explicitly selected. All tables that are recursively related (in the model, not necessarily in the database) to the table will be automatically implicitly selected.

For example, let's say that your model describes the table Invoice with Customer as its parent and Invoice_Line as its child.

If you select the table invoice explicitly using the Explicit Tables Button, Customer and Invoice_Line, which are related to it, will be implicitly selected. If they are related to other tables, they will also be implicitly selected recursively.

You can define whether you want to extract all the records from the Explicit Table or a subset consisting of a randomly selected number by clicking the Number Radio Button. The selection can also be restricted by specifying a Where Clause in the Selection Group or by naming and locating a CSV file containing the values of the Primary Key for the records to be selected.

The records of the implicitly related tables will be selected based on the settings you establish at the relationship level.

For example, you could decide to extract the parents of the invoice (Customer) and the children of the Invoice (Invoice_Line).

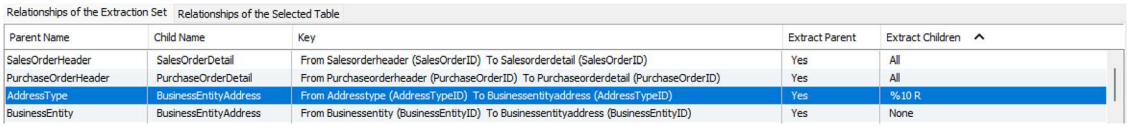
Only the customers and the invoice lines of the selected invoices will be selected.

By default, the system selects all parents and ancestors, as doing otherwise would lead to potential integrity errors. Also, by default, only the direct children and descendants of the explicit table will be selected.

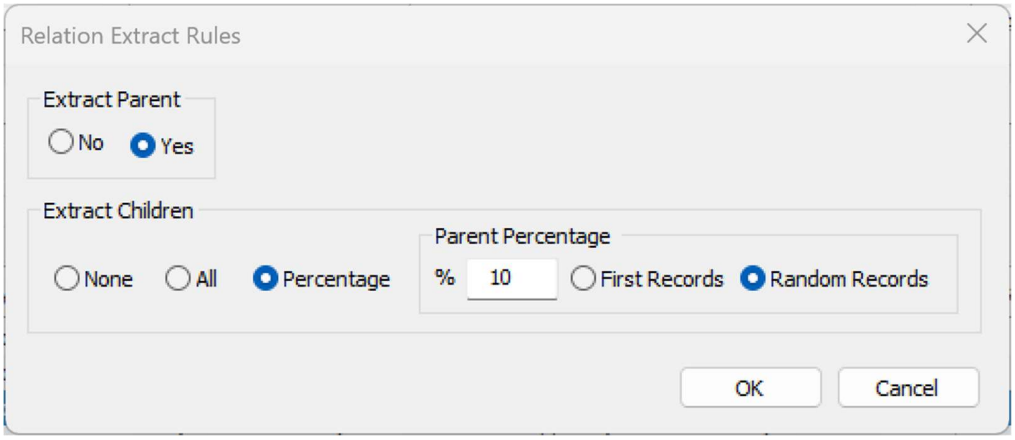
You can modify these default decisions in two ways: using the Relationships List or an Entity Relationship Diagram (ERD) that displays the selected and related tables.

Relationships List

Using the list displaying the list of tables explicitly selected and their related implicit tables:



This list displays the recursive relationships derived from the explicitly selected tables. You can multi-select relationships (CTRL + A selects them all) and right-click on them to produce a dialog that allows you to set whether you want to retrieve the parent and/or the child records for the selected relationships.

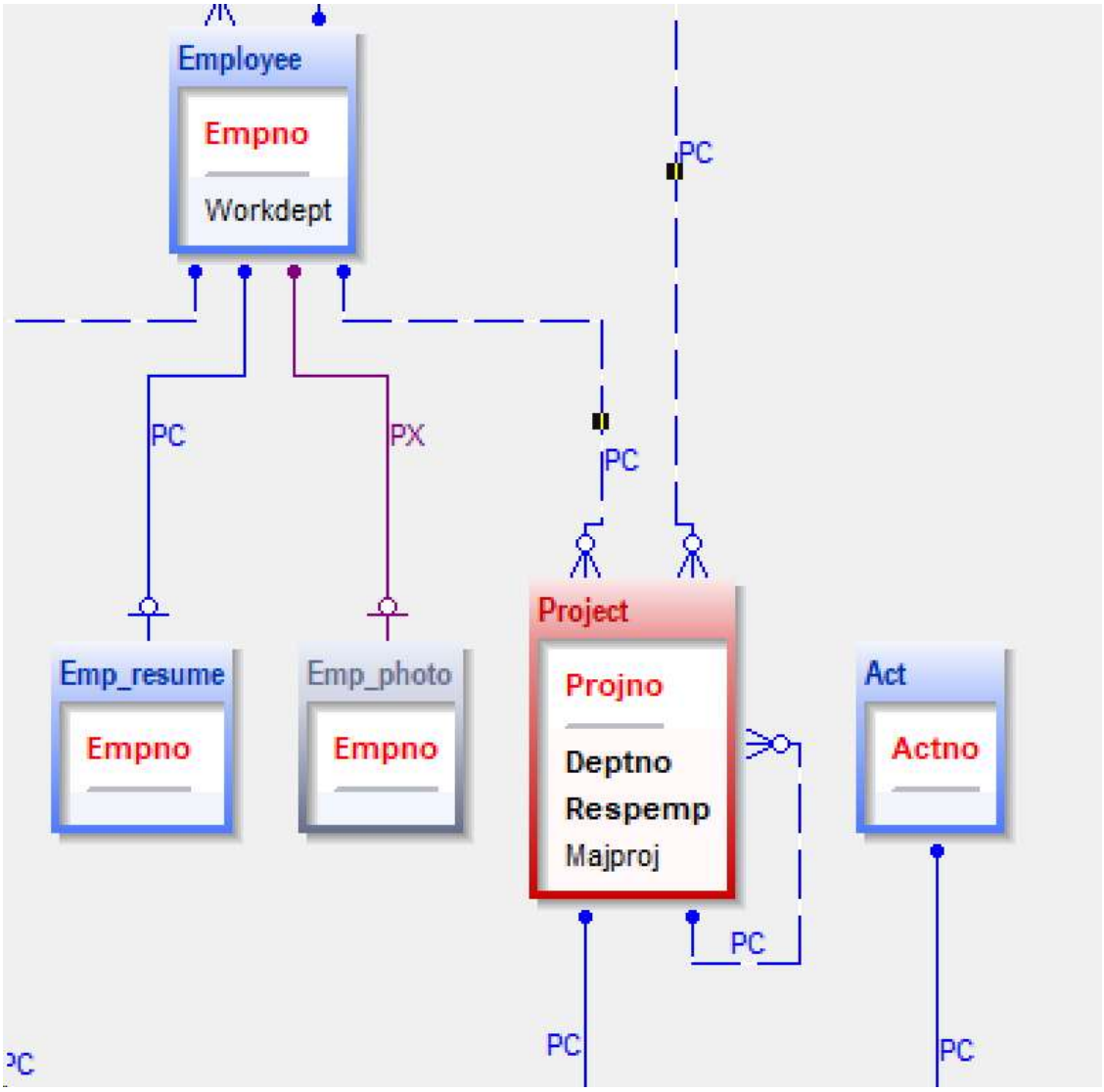


For the parent, you can choose whether to include it. It is recommended to set this to Yes to maintain relational integrity. For the children, you can select None or All; for example, you might want all the children of a specific parent (such as the detail lines of an invoice). You can also choose to extract children only for a percentage of the parents' children, with those parents being either the first or randomly selected.

Related Diagram Button

An alternative way to set up the relationships between parents and children is by using an Entity-Relationship Diagram, which graphically presents the explicitly selected tables and their recursively related relationships.

To produce the diagram, click the Related Diagram Button.



Explicitly selected tables appear in Red. Implicitly selected tables in Blue, except for those for which extraction will not occur, given the relationship settings, which will appear in Grey.

The label of the Relationships contains two letters. The first refers to the action on the Parent (P: Extract the Parent, X: Do not Extract the Parent). The second refers to the child (C: Extract the Children, X: Do not Extract the Children, %: Only a percentage of the parents will have their children extracted)

Relationships can be multi-selected, and by right-clicking on one of them, you can set the desired behavior in the same way as in the Relationship List.

Right-clicking a white space on the Entity Relationship Diagram produces a menu with the following options:

- Fit to Window (F can be used instead)
- Zoom In (+ can be used instead)
- Zoom Out (- can be used instead)
- Do Hierarchical Layout
- Do Orthogonal Layout
- Refresh

It is important to note that you can explicitly select (and set conditions for) multiple tables.

The diagram window contains three resizable sections.

- The diagram on the right shows where relationship rules can be set.
- The list of tables in the diagram on the top left. Double-clicking a table will position the mouse cursor on that table in the diagram.
- The Zoom window on the bottom left allows you to view in full size the portion of the diagram where the mouse cursor is positioned.

Table Diagram Button

If more practical, instead of clicking the Related Diagram Button, you can click the Table Diagram Button, which will produce the diagram of the recursively related tables of only the currently selected explicit table.

Reset Descendants Button

This Button allows you to set the children and descendants of the explicitly selected tables to be set as “Get Children”. Note that other relationship settings are not affected.

Delete Selected Button

This button removes from the list of explicitly selected tables those marked as disabled.

Refresh Record #

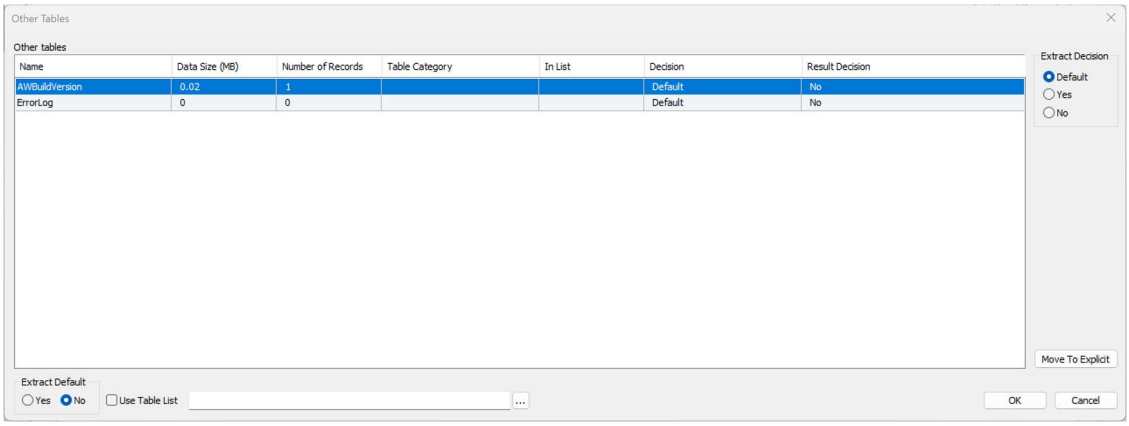
Clicking this button will refresh the column that displays the number of table records in the explicitly selected tables list.

Other Tables Button

Model tables that are not explicitly or implicitly selected are displayed in the Dialog produced when clicking the Other Tables Button.

Contrary to explicit and implicit tables, which use inserts recursively to comply with integrity rules and the conditions set (Number, Where, CSV), these tables should be either copied entirely or not at all if they are not needed in the Dev/Test environment.

The Other Tables Dialog presents them and allows you to decide if they should be part of the extraction process.



The Move to Explicit Button will move the selected tables to the Explicitly Selected Tables list.

The Extract Default group enables you to specify whether the tables in the list should be extracted by default.

This default can be overridden in multiple ways:

- If the Table Category column value is “Temporary” or “Obsolete”, the Extract Decision is set to “No”. This table attribute can be edited in the Entities Data Dictionary.
- If the name of the table is present as a separate line in the specified Table List text file and the “Use Table List” Checkbox is checked, the Extract decision is set to “Yes”.
- If the table is used by an IBM i program provided in the Program List and the “Use Program List” Checkbox is checked, the Extract decision is set to “Yes”.
- The Resulting Decision can always be overridden, despite the above rules, by setting it to “Yes” or “No” instead of “Default” using the Radio Buttons in the Extract Decision Group.

On IBM i, the tables set to be extracted in this Dialog will be efficiently saved on one or more SAVF files together with their indexes, logical files, triggers, and constraints. Please note that trigger programs are not currently supported and must be manually handled. The SAVF files are created in the Destination Schema set in the Extract Data Dialog.

Schema Group

The Destination Dropdown allows you to select the schema where the tables with the extracted data (and the DB2 for i SAVF files) will be stored.

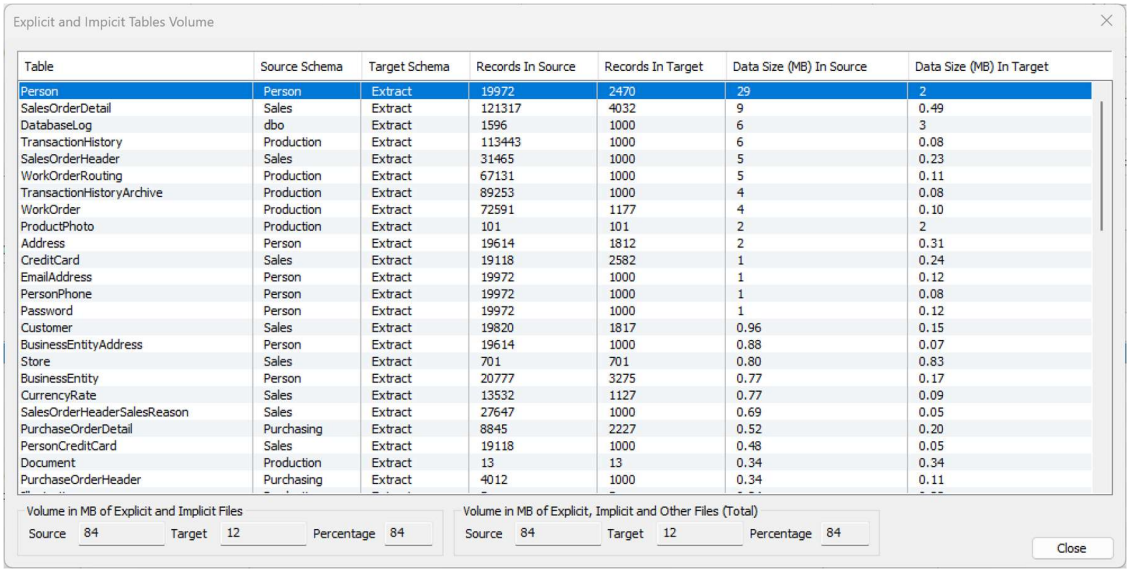
The Mapping Button produces a Dialog that allows you to map the existing schemas to other values in the resulting script.

Script Group

In this Group, you can set the name of the extraction script, generate it, modify it, and execute it.

Table Size Button

Once the extraction script has been executed, click this button to produce the Tables Volume dialog.



In this dialog, you can see how the volume of the extracted tables has been reduced.

Extraction strategies

Depending on your goal, you can select various extraction strategies to produce compact yet coherent and representative extraction sets. If you want to obtain a non-production database representing the entire Production, the following settings are recommended:

Select all relevant tables for your non-production database into the Explicitly Selected Tables List and set them to, for example, a random selection of 1000 records each. Note that sometimes you might want to extract the full content of a table. For example, all products, regardless of whether they are referenced in the extraction, are included. In that case, do not specify any condition.

Groups of tables that are not related to others can be moved to “Other Tables,” and they will be restored in their entirety more efficiently.

In the relationship list, set Extract Parent to Yes, and Extract Children to None. If a child table requires all Children of a Parent (for example, the detail lines of an invoice), then set Extract Children to All.