


# Data modeling in Cognos Analytics

Last Updated: 2024-01-29


IBM® Cognos® Analytics with Watson provides web-based, self-service data modeling capabilities.

You can use data modeling in Cognos Analytics to fuse together many sources of data, including relational databases, Hadoop-based technologies, Microsoft Excel spreadsheets, text files, and so on. Using these sources, a data module  is created that can be **12.13** used in reports, dashboards, or explorations.

Star schemas are the ideal database structure for data modules, but transactional schemas are equally supported.

You can enhance a data module by creating calculations, defining filters and navigation paths, and more.

After you save a data module, other users can access it. Save the data module in a folder that users, groups, and roles have appropriate permissions to access. This process is the same as saving a report or dashboard.

 **Tip:** Data modeling in Cognos Analytics does not replace IBM Cognos Framework Manager, IBM Cognos Cube Designer, or IBM Cognos Transformer, which remain available for maintaining upgraded projects.

## Intent modeling

You can use intent modeling to create a data module. Intent modeling proposes tables to include in the module, based on matches between the terms that you supply and metadata in the underlying sources.

Intent modeling recognizes the difference between fact tables and dimension tables by the number of rows, data types, and distribution of values within columns. When possible, the intent modeling proposal is a star or snowflake of tables. If an appropriate star or snowflake cannot be determined, a single table or a collection of tables is proposed.

For more information, see [Discovering related tables](#).

## Automatic joins

Cognos Analytics automatically creates joins between tables in a data module. The autojoin (automatic join) algorithm adopts a diagnostic scoring approach when deciding which columns to use to join two tables. The algorithm uses a set of rules that are applied when choosing the column combinations

between the two tables. Each rule produces a score. The score could be negative. The total score of all rules decides if a column combination qualifies to be a join column.

The autojoin algorithm uses the following rules:

- The similarity of two column names must exceed a minimum threshold.  
For example, the names SalesCountryCode and CountryCode are highly similar, and can be considered a match.
- Both columns belong to the same semantic category.  
For example, the Employee or Product category.
- Both columns have the same semantic attribute.  
For example, both are IDs.
- None of the columns is a common row identifier.  
The row ID column could be in every table.
- The data in two numeric columns overlaps.
- The relationship between two columns can't be many-to-many.

A join relationship is created if any column combinations between two tables satisfy a minimum qualification score. The collected statistics is used to ensure that cardinality is properly set when building the relationship. The joins created by the autojoin algorithm are saved as the inferred relationships in the data module.

For more information, see [Relationships](#).

- **Modeling user interface**  
Use the web modeling user interface to view, create, enhance, and edit data modules.

**Parent topic:**

→ [Data Modules](#)