

SENTURUS

IBM Cognos Analytics Data Modules

Agenda

- Data Modules Overview
- Upload Local Files for analysis
- Create a Data Set from an existing OLAP package
- Blend local files, data sets, and packages with Data Modules
- Create calculations and custom data groupings
- Create a report using a Data Module
- Q & A



SENTURUS

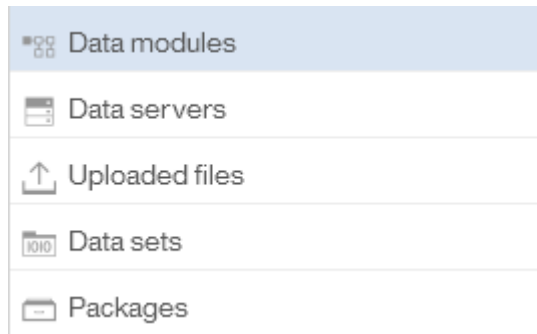
Data Modules Overview

Data Modules Overview

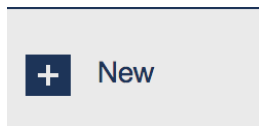
- Data modules are a new Capability in IBM Cognos Analytics that enable users to perform self service analytics
 - Self-service means: not relying on IT or other gatekeepers to analyze data from multiple sources, using the same familiar, powerful, integrated platform as all other BI activities
- Extends our analysis beyond traditional enterprise, or centrally-managed sources of information
- Enables users to combine personal data with other sources
- Can be published and shared so a consistent logical model is available to all users

Data Modules Overview - Sources

- Can be based on:



- Existing Data Modules built in Cognos Analytics
- Pre-configured database connections such as SQL Server, DB2, Redshift, SAP HANA
- Desktop-based files such as Excel Workbooks, CSV
- Extracts, or Snapshots from existing Packages
- Live Connections to existing (relational) Packages




 Report

 Dashboard


 Story

 Data module

 Other

 Upload files

 Manage

 **Data server connections**
Create and manage connections

 **Administration console**

 **New job...**

Data Modules Overview - UI

- Data Modules allow users to conduct light data modeling (integration) in a non-threatening interface, empowering a simple, immediate, and shareable data integration solution for any type of data source, including local (personal) files

The screenshot displays the 'Product Type Data Module' interface. On the left, the 'Sources' pane shows a tree view with 'GO sales (query)', 'Product Type Forecast.xlsx', and 'Product Type Returns Data Set'. The 'Data module' pane on the right shows a hierarchical view of the data model, including 'Product Type Data Module', 'GO sales (query)', 'Product Type Forecast Xlsx', and various fields like 'Product type code', 'Product type', 'Expected volume', 'Forecast revenue', and 'Product type group'. On the right side of the interface, a data model diagram is visible, showing three boxes: 'Product ...Data Set', 'GO sales (query)', and 'Product ...ast Xlsx'. Lines connect 'Product ...Data Set' to 'GO sales (query)', and 'Product ...ast Xlsx' to 'GO sales (query)', indicating data relationships.

Data Modules Overview - Relationships

- Simple Join Relationships can be defined and managed by actual business (end) users

The screenshot shows a configuration interface for defining a relationship between two data tables. On the left, 'Table 1' is 'Products' and on the right, 'Table 2' is 'Product Type Forecast Xlsx'. A blue line with the number '1' at both ends connects the 'Product type code' field in Table 1 to the 'Product type code' field in Table 2. Below the tables, there are search bars and dropdown menus for selecting fields. A 'Match' button with a downward arrow is located below the search bars. On the right side, there are three sections: 'Relationship Type' with radio buttons for 'Inner join' (selected), 'Left outer join', 'Right outer join', and 'Full outer join'; 'Cardinality' with radio buttons for '1-to-1' (selected), '1-to-many', and 'Many-to-1'; and 'Optimization' with radio buttons for 'No filtering' (selected), 'Unique values', and 'Value range'.


- Join type, Cardinality, and Join Filters can be easily managed using non-technical terminology and cues

Data Modules Overview - Joins

- Most of the time a join relationship is defined by simply linking one column in an object to a matching column in the other object
- This is called an equi-join (the values in columnA are = the values in columnB)
- Matching multiple columns is rarely required but available
 - This is called a compound join
 - Necessary when multiple attributes in combination are required to define a unique value

Defined matches

Product type code...roduct type code 

Product type - Product type 

2 Matched columns

Data Modules Overview - Cardinality

- Cardinality describes for any one row in a given table how many rows in the table on the other side of a relationship can be expected to match

1-to-1

Each row in Table 1 is related to exactly one matching row in Table 2

1-to-many

Each row in Table 1 has one or more matching rows in Table 2

Many-to-1

One or more rows in Table 1 match to a single row in Table 2

- We need to provide accurate Cardinality to assist the query engine in its planning
- A basic understanding of data modeling concepts is important

Data Modules Overview - Filtering

- Filter joins can improve performance by limiting the number of rows that are searched (and returned) in one of the tables in a join relationship

No filtering
Join without filtering

Unique values
Filter Table 2 by each unique value in Table 1

Value range
Filter Table 2 by the range of values in Table 1

- For instance, if you have a spreadsheet that only contains the top 10 products in your product family, but your database contains the sales history for thousands of products, you can request that any query to the database include an explicit filter for just the products in the spreadsheet (unique values or a range of values)

Data Modules Overview - Aggregation

- Another fundamental consideration of Data Modeling, or Data Integration, is how measures, or facts, or numbers, are summarized, or aggregated

Product line	Product type	Revenue
Camping Equipment	Cooking Gear	\$272,835,984.18
	Lanterns	\$126,925,660.64
	Packs	\$351,880,402.84
	Sleeping Bags	\$309,172,888.35
	Tents	\$528,221,728.02
Camping Equipment - Summary		\$1,589,036,664.03
Golf Equipment	Golf Accessories	\$51,514,343.88
	Irons	\$254,814,337.99
	Putters	\$106,184,271.37
	Woods	\$313,898,414.65
Golf Equipment - Summary		\$726,411,367.89
Mountaineering Equipment	Climbing Accessories	\$81,096,582.48
	Rope	\$114,426,644.73
	Safety	\$83,236,883.98
	Tools	\$130,900,021.71
Mountaineering Equipment - Summary		\$409,660,132.90
Outdoor Protection	First Aid	\$12,429,699.12
	Insect Repellents	\$36,822,842.52
	Sunscreen	\$26,741,754.61
Outdoor Protection - Summary		\$75,994,296.25

Properties ⓧ

General Navigation paths

Label Return quantity

This item is hidden from users

Expression Return_quantity >

Usage Measure ▾

Aggregate Total ▲

Data type Count

Represents Co...t

Comments M...m

Mi...m

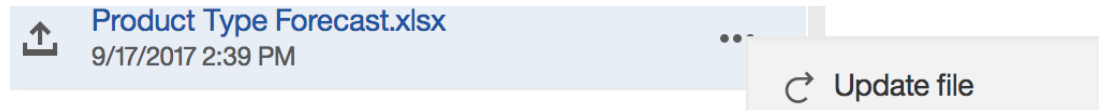
Total

Data Modules Overview - Aggregation

- Measures are almost always additive, meaning they are aggregated using a Total function
 - If I have lots of orders, and each record in the order table records quantity sold, then across any population of orders, I typically want to see the Total of all those individual quantity values
 - By default, query tools will want to summarize data items that are recognized as measures
 - It is our job to control these behaviors
- The key to success is understanding the data and the business requirements

Data Modules Overview - Data Refresh

- When querying a Data Module, the most current values in each source will be returned
- Uploaded files must be manually refreshed



- A Data Set can be defined to create a “snapshot” of select data from a Data Module or a Package



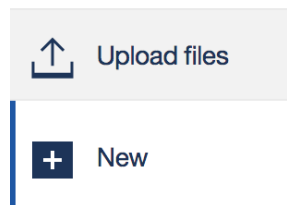
- Data Sets can be refreshed via a Schedule



Upload Local Files for Analysis

Upload Local Files for Analysis

- What types of files can be used for upload?
 - Microsoft Excel (.xlsx and .xls) spreadsheets
 - Only the first sheet in a Microsoft Excel workbook is uploaded. If you want to upload the data from multiple sheets in a workbook, save the sheets as separate workbooks.
 - .CSV, .TXT
 - Text files that contain either comma-separated, tab-separated, semi colon-separated, or pipe-separated values
- Uploaded files are stored on the Cognos server in a columnar format
- Click New and use the Upload tool at the bottom left of the screen





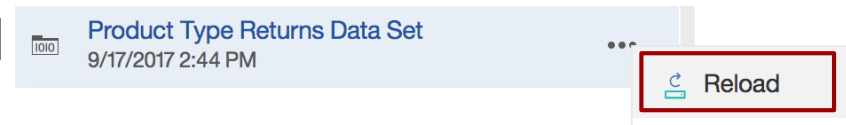
Create a Data Set from an Existing Package

Data Sets

- Data sets are created from packages or data modules
- You define a data set by choosing one or more items (columns) from a package or data module
- You can also apply filters to reduce the data
- The data is extracted and stored within the Cognos Analytics system
- Because the data is cached, data sets can improve query performance and reduce the workload on your database(s)
- Some reasons to use a data set:
 - improve query performance if your database is slow
 - retain a version of the data at a specific time

Data Sets

- For data sets created from *relational* packages or data modules, you have the option to Summarize detailed values, suppressing duplicates
 - When you use this option, measure values will be aggregated to the lowest grain that is explicitly included in the data set
- Refreshing your Data Set
 - Updates can occur on-demand
 - Updates can be scheduled from the Properties menu to occur automatically including weekly, daily, hourly or every X minutes



Data Sets

- Use the More action next the entry

GO Sales Cube
4/25/2017 2:29 AM

- To Filter the data set click a value or a column header in the Data Set Preview

Product type code	Product type	Return quantity
951	Cooking Gear	141,731
955	Lanterns	65,590

Create report

Create dashboard

Create data set

Properties

- Or, right-click a data item in the Source pane
- (relational only)

Time

- Year
- Qu
- Qu
- Qu
- Qu
- Mo
- Mo
- Mo

Insert
Collapse
Collapse All
Cancel
Filter for report...
Properties

Filter condition - Year

Condition:

Show only the following values:

Values

Comparison

Operator:

>

Value:

2015

2 filters applied



SENTURUS

Blend Local Files, Data Sets, and Packages with Data Modules

Data Modules - Select Sources

- Select the Data Module Sources

Selected sources	Sources	Packages
<p>🔍 Search</p> <p>📁 GO sales (query) Team content > Samples > Models</p> <p>📁 Product Type Forecast.xlsx My content > Presentation Files</p> <p>📁 Product Type Returns Data Set My content > Presentation Files</p> <p>Done</p>	<p>📁 Data modules</p> <p>📁 Data servers</p> <p>📁 Uploaded files</p> <p>📁 Data sets</p> <p>📁 Packages</p>	<p>🔍 Search</p> <p>📁 Census Cube Team content > Samples > Models <input type="checkbox"/></p> <p>📁 GO data warehouse (analysis) Team content > Samples > Models <input type="checkbox"/></p> <p>📁 GO data warehouse (query) Team content > Samples > Models <input type="checkbox"/></p> <p>📁 GO sales (analysis) Team content > Samples > Models <input type="checkbox"/></p> <p>📁 GO sales (query) <input checked="" type="checkbox"/> Team content > Samples > Models</p> <p>📁 GO Sales Cube Team content > Samples > Models <input type="checkbox"/></p>

- Select the entire source or only a subset of tables for the Data Module (i.e. connecting directly to a database)



SENTURUS

Create Calculations, and Custom Data Groupings

Calculations and Custom Data Groupings

- There are properties to be considered when defining a Data module
 - Calculations
 - Custom Groups
 - Measure behavior

The screenshot shows a data module configuration interface. On the left, a list of items is displayed under the heading "Product Type...rns Data Set". The items are:

- # Product type code (highlighted)
- abc Product type
- Return quantity

A context menu is open over the "Product type code" item, listing the following actions:

- Filter...
- Create calculation...
- Clean...
- Create data group...
- Create navigation path...
- Rename
- Hide
- Remove
- Properties

The screenshot shows the "Properties" dialog box for a data module. The dialog has two tabs: "General" and "Navigation paths". The "General" tab is active.

The "Label" field is set to "Product type".

The "This item is hidden from users" checkbox is unchecked.

The "Expression" field is set to "Product_type" with a right-pointing arrow.

The "Usage" dropdown is set to "Identifier".

The "Aggregate" dropdown is set to "None".

The "Data type" is set to "Text".

The "Represents" dropdown is set to an empty value.

The "Comments" field is empty.

The "Sorting" section is expanded, showing:

- The "Sort" checkbox is checked.
- The "Sort by" dropdown is set to "Product type".
- The "Order" radio buttons are set to "Ascending".

Calculations and Custom Data Groupings

- We can also do the following:
 - Rename objects
 - Define Filters
 - Clean up columns

Clean - Product type ×

Whitespace
 Trim leading and trailing whitespace

Convert case to
 UPPERCASE lowercase Do not change

Return a substring of characters
Start Length

Preview
this is a preview placeholder

NULL values
 Replace this value with NULL
 Replace NULL values with



SENTURUS

Data Modules Q & A