

Everything You Need to Know about Data Preparation and Modeling with IBM Cognos Analytics — 3682B

think 2019

Michael McGeein, CPA, CA
Senior Offering Manager, Business
Analytics

Please note

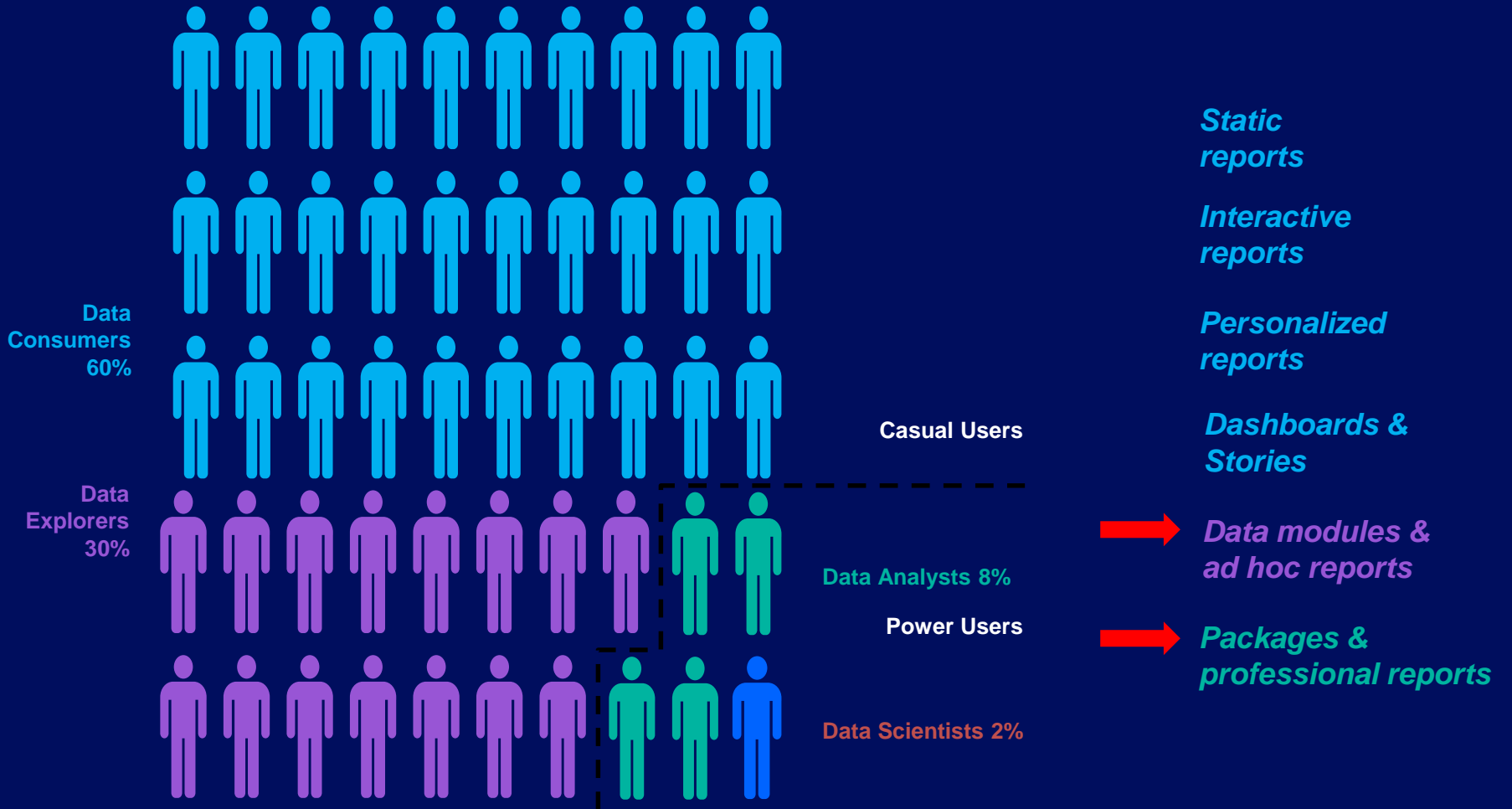
IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice and at IBM's sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

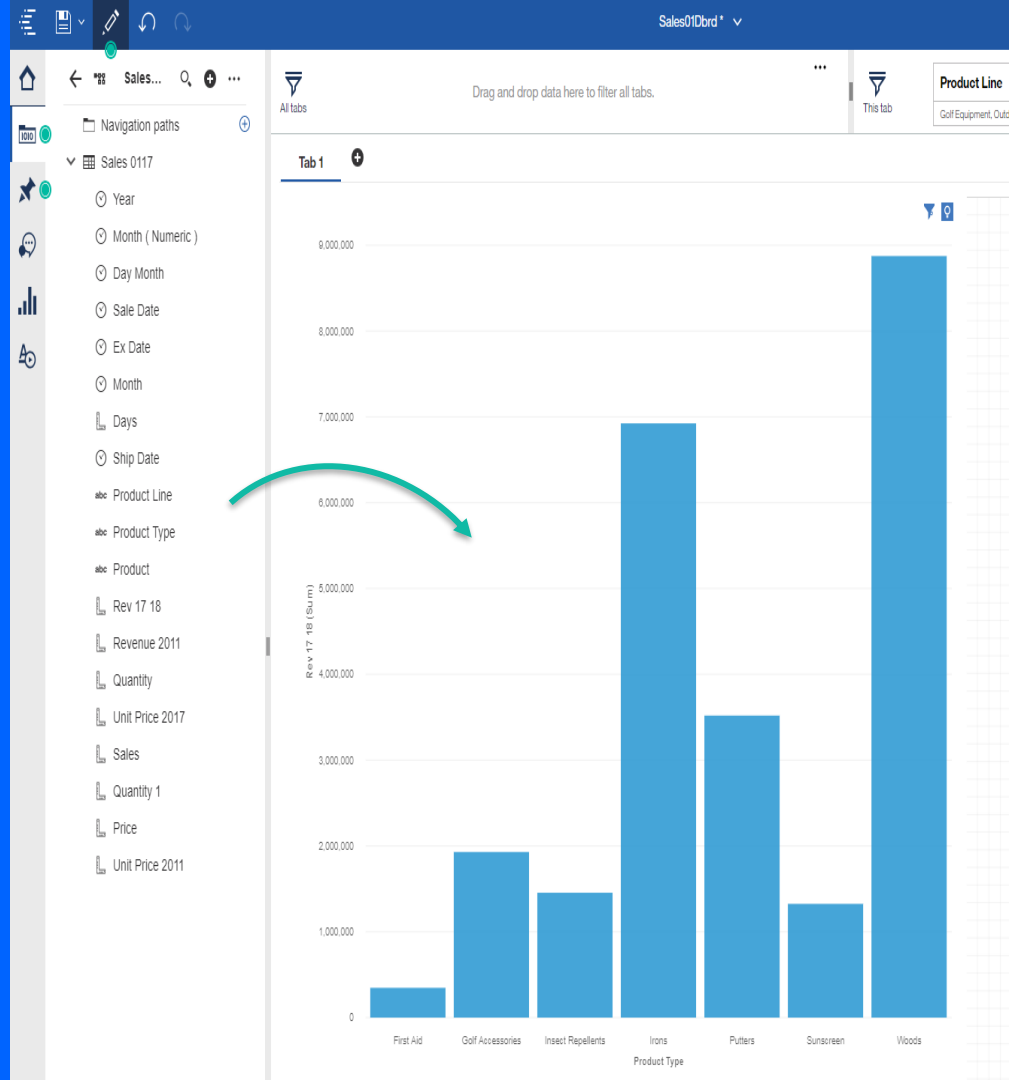
Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.



Data modeling is preparing ...

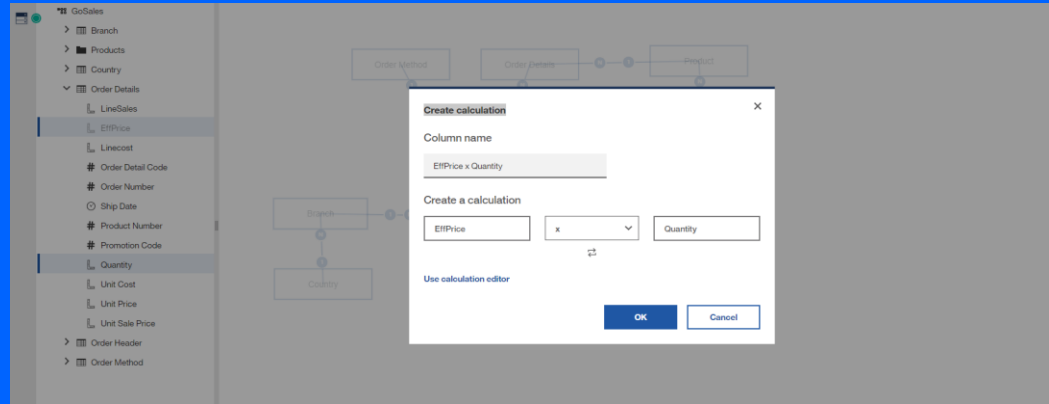
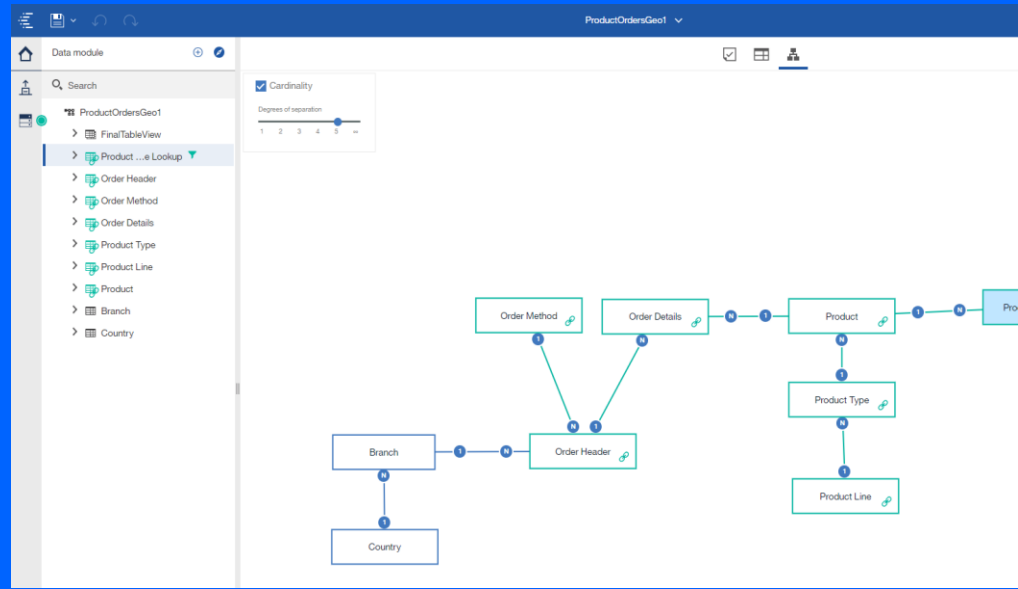
A **business-oriented** presentation of one or more data sources
Easily understood by decision makers

The metadata **building blocks** for assembling reports, dashboards, and stories



Data modeling is shaping...

- **Filtering** data to only what's needed
- Augmenting data with **calculations**
- **Joining** data sets together
- **Cleaning** and **grouping** data
- Setting metadata like **aggregation** and **sorting** to the best defaults for reports, dashboards and stories



Share and reuse

Shaping data in a dashboard only affects that dashboard

Changes to a data model affect all reports, dashboards, and stories based on it

Save a data model into a folder with suitable security permissions so others can use it as appropriate

The screenshot displays the IBM Cognos Analytics interface. The top navigation bar includes the IBM Cognos Analytics logo, a 'Welcome' dropdown, and a breadcrumb path: 'Home > Team content > My Org Only'. The left sidebar contains navigation options: 'Home', 'Search', 'My content', 'Team content', and 'Recent'. The main content area shows a list of data models under 'My Org Only':

- Orders (6/16/2017 2:03 PM)
- Returns (11/15/2016 11:55 AM)

The 'Orders' data model is selected, and the right-hand pane shows its details and permissions. The 'Permissions' tab is active, displaying the following information:

- Owner: Jason Tavoularis
- Created: 11/14/2016 3:21 PM
- Modified: 6/16/2017 2:03 PM
- Type: Data module

The 'Permissions' section includes a checked checkbox for 'Override parent permissions'. Below this is a table listing permissions for different user roles:

Name	Permission
Authors	Write
Consumers	Run
Modelers	Write

At the bottom of the permissions pane, there is an unchecked checkbox for 'Apply to all children' and two buttons: 'Apply' and 'Cancel'.

Cognos Analytics 11.1

Data modules

Cognos Analytics 11.1 Data Modules - Focus areas

- Improve First Time Use ('data to insight') experience
 - Improve flat file support (multi sheet, append, replace)
 - Base capability additions (Format, Filtering)
 - Simplify complexity of modeling (data module)
- Experience Improvements
 - Clean / Prep data – Trim, Split, Date/Time
 - Aggregation roll ups – (multi-fact / multi-grain)
 - Built tables – Union, Intersect, Except
 - Usability improvements – Expression Editor, Folders, SQL, Prettify
- Relative Date Support

Data cleaning / preparation

Address 'dirty' data
Data type mismatch
Split data
Combine data
NULL vs 0

Aggregation Roll ups

Address granularity differences
Avoid double counting

Usability Improvements

Create from SQL
Expression guidance / assistance
Set Operations
Format
Filters

Relative Date support

18 pre-configured time periods
User defined Measures
Full period & To-Date

Data cleaning and preparation

- Victim Descent
- # Premise Code
- # Premise Description
- # Weapon Used Code
- # Weapon Description
- Status Code
- Status Description
- # Crime Code 1
- # Crime Code 2
- # Crime Code 3
- # Crime Code 4
- Address
- Cross Street
- Location

Split column - Location

Specify the characters to use as delimiters.

You can either select the delimiters from the template, or type the delimiters.

Select delimiters:

(3 3 . 0 0 3 3 . - 1 1 8 . 2 8 4 2)

Type delimiters:

()

Preview

Location	Location 1	Location 2
(33.9933, -118.2842)	33.9933	-118.2842
(34.1086, -118.2048)	34.1086	-118.2048
(33.9455, -118.3768)	33.9455	-118.3768
(33.998, -118.4354)	33.998	-118.4354
(34.2066, -118.4144)	34.2066	-118.4144
(34.2277, -118.369)	34.2277	-118.369
(34.1814, -118.4263)	34.1814	-118.4263

- Crime Code 1
- Crime Code 2
- Crime Code 3
- Crime Code 4
- Address
- Cross Street
- Location
- Latitude
- Longitude

SalesData1

Data

1/27/2017

- NewDate
- Row Id
- Year
- Month (numeric)
- Day-Month
- Sale Date

Filter...

Edit calculation...

Convert to date...

Split...

Convert to date - NewDate

It appears the data values in this column are of data type date.

1 / 27 / 2017

Month Day Year

Original column: NewDate

Preview: Useable Date

1/27/2017	2017-01-27
1/20/2017	2017-01-20
1/25/2017	2017-01-25
1/26/2017	2017-01-26
1/18/2017	2017-01-18
1/11/2017	2017-01-11
1/14/2017	2017-01-14

Aggregation Roll ups

Multi Fact / Multi Grain

The screenshot shows the Power BI Desktop interface. On the left, the 'Model view' pane displays a table named 'Denormalized.xlsx' with columns for 'CURRENT_YEAR', 'CURRENT_QUARTER', and 'MONTH_EN'. The table contains data for years 2010 and 2011, with rows representing different months and quarters. The 'Specify column dependencies' dialog box is open, showing a list of columns and their dependencies. The columns are grouped into three sections: 'CURRENT_MONTH', 'CURRENT_YEAR', and 'CURRENT_QUARTER'. Each section has a list of columns and their dependencies, with checkboxes for each dependency.

CURRENT_YEAR	CURRENT_QUARTER	MONTH_EN	QUANTITY	UNIT_COST	UNIT_PRICE	GROSS_MARGIN	SALE_TOTAL	GROSS_PROFIT	PRODUCT_NUMBER	PRODUCT_NAME	PRODUCT_LINE_CODE	PRODUCT_LINE_EN	PRODUCT_TYPE_CODE	PRODUCT_TYPE_EN	Year_Quantity	Quarter_Quantity	Month_Quantity	ProductLine_Quantity	ProductType_Quantity	YearProductLine_Quantity	QuarterProductType_Quantity
2010	1	1																			
2010	1	2																			
2010	1	3																			
2010	2	1																			
2010	2	2																			
2010	2	3																			
2010	3	1																			
2010	3	2																			
2010	3	3																			
2011	1	1																			
2011	1	2																			
2011	1	3																			
2011	2	1																			
2011	2	2																			
2011	2	3																			
2011	3	1																			
2011	3	2																			
2011	3	3																			

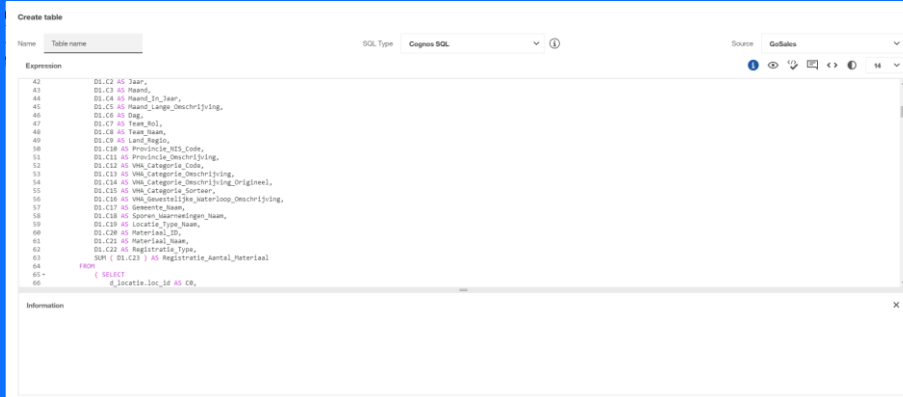
The screenshot shows a Power BI report with a table displaying aggregation roll-ups. The table has columns for 'CURRENT_YEAR', 'CURRENT_QUARTER', 'CURRENT_MONTH', 'Month_SALE_TOTAL', 'Quarter_SALE_TOTAL', and 'Year_SALE_TOTAL'. The data is grouped by year (2010 and 2011) and quarter (1, 2, 3). The 'Month_SALE_TOTAL' column shows the total sales for each month, while the 'Quarter_SALE_TOTAL' and 'Year_SALE_TOTAL' columns show the total sales for each quarter and year, respectively.

CURRENT_YEAR	CURRENT_QUARTER	CURRENT_MONTH	Month_SALE_TOTAL	Quarter_SALE_TOTAL	Year_SALE_TOTAL
2010	1	1	\$72,741,622.65	\$221,704,705.31	\$914,352,803.72
2010	1	2	\$73,242,843.99	\$221,704,705.31	\$914,352,803.72
2010	1	3	\$75,720,238.67	\$221,704,705.31	\$914,352,803.72
2010	2	4	\$65,278,358.76	\$222,143,384.57	\$914,352,803.72
2010	2	5	\$74,695,218.83	\$222,143,384.57	\$914,352,803.72
2010	2	6	\$82,169,806.98	\$222,143,384.57	\$914,352,803.72
2010	3	7	\$80,723,457.30	\$235,750,316.25	\$914,352,803.72
2010	3	8	\$80,174,484.77	\$235,750,316.25	\$914,352,803.72
2010	3	9	\$74,852,374.18	\$235,750,316.25	\$914,352,803.72
2010	4	10	\$75,086,688.05	\$234,754,397.59	\$914,352,803.72
2010	4	11	\$75,470,246.45	\$234,754,397.59	\$914,352,803.72
2010	4	12	\$84,197,463.09	\$234,754,397.59	\$914,352,803.72
2011	1	1	\$94,998,202.61	\$293,228,460.53	\$1,159,195,590.16
2011	1	2	\$99,488,042.66	\$293,228,460.53	\$1,159,195,590.16
2011	1	3	\$98,742,215.26	\$293,228,460.53	\$1,159,195,590.16
2011	2	4	\$88,093,161.37	\$278,180,759.96	\$1,159,195,590.16
2011	2	5	\$90,020,791.42	\$278,180,759.96	\$1,159,195,590.16
2011	2	6	\$100,066,807.17	\$278,180,759.96	\$1,159,195,590.16
2011	3	7	\$90,409,616.73	\$281,079,666.95	\$1,159,195,590.16
2011	3	8	\$91,983,016.90	\$281,079,666.95	\$1,159,195,590.16

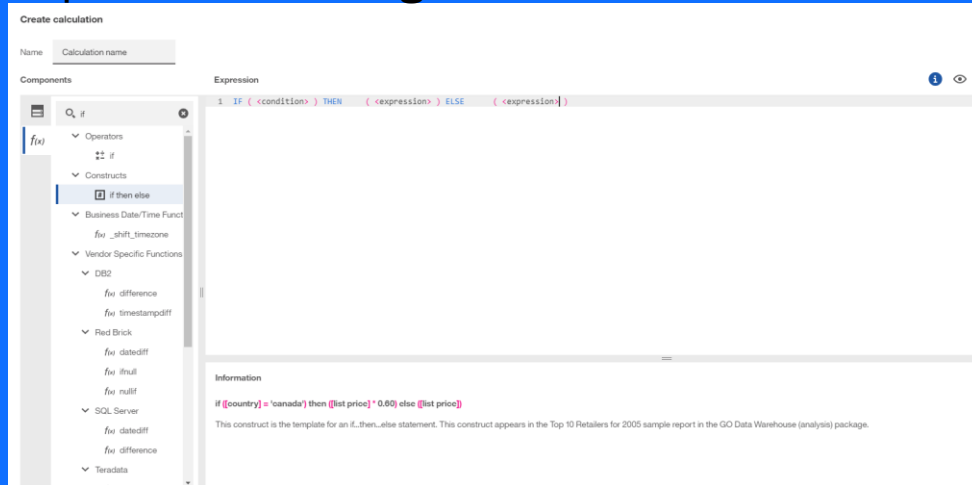
Usability Improvements

Tables and Set Operations

Create table from SQL



Expression Editor guidance



Relative Dates

- Calendars shipping with 11.1
 - Gregorian Calendar – 1/1 to 12/31
 - Samples now includes all 12 months
- Include Calendar in Data module
 - It will be hidden by default after addition
 - Relate Date field(s) to calendar (Lookup Reference)
 - Link measure to date field
 - Instructions available to generate another calendar

Note: Currently Months MUST begin on the first day of the month

Split Columns

Aggregation Roll up

Relative Dates

A hand is shown from the bottom right, reaching up to touch a large, white-bordered button with the word "Demo" written in white. The background is a vibrant blue, featuring a faint world map and several white-outlined rounded rectangles of varying sizes, some containing network-like patterns with glowing nodes. The overall aesthetic is modern and tech-oriented.

Demo

Data modelling improvements

- **Flat file - multi tab support & append**
- **Comparative dates - month to date, quarter to date(QTD), prior QTD**
- **Data preparation – split column, trim, case**
- **Recommendation on relationships – join, blend**
- **Aggregation across grains –Multi grain analysis**
- **Ease of use – expression editor, SQL based tables, format, filter**
- **Set operations – Union, Intersect, Except**
- **Security Filters**

Notices and disclaimers

© 2018 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights – use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed “as is” without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply.”

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer’s responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer’s business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

Notices and disclaimers continued

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.**

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com and [names of other referenced IBM products and services used in the presentation] are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at: www.ibm.com/legal/copytrade.shtml.

Wrap Up





**THANK YOU
FOR YOUR
TIME**