Return of the Variables

David G. Rathbun

BI Solutions Architect, PepsiCo B+IS

Presentation Abstract

- It's time to pull out some variables again as we revisit a popular topic from past conferences. Report variables can provide a more creative, efficient, and effective means of completing the detailed analysis required for your reports. The focus of this presentation is building documents, using report functions, and reviewing variable tricks that can be used to solve reporting challenges. The presentation this year will focus primarily on the SAP BusinessObjects Web Intelligence product and include some of the newest features that have recently become available, but there will be some goodies for users of prior versions as well.

About Dave

- Dedicated to BusinessObjects solutions since 1995
  - Consultant and trainer for fifteen years
  - Currently BI Solutions Architect for PepsiCo
- 14 consecutive years presenting at major BI conferences
  - United States, Europe, Australia
- Charter member of BOB
  - http://busobj.forumtopics.com
- I Blog! Dave's Adventures in Business Intelligence
  - http://www.dagira.com
- Selected to the SAP Mentor program for 2009

Demonstration Platform

- Demonstration universes
  - Island Resorts Marketing
  - Summit Sporting Goods
  - Prestige Motors
  - Demonstration databases were converted to Oracle
- Software configuration
  - BusinessObjects Enterprise XI 3.0
  - Oracle 10g
- BusinessObjects toolset
  - Web Intelligence Rich Client

🖥 Demonstration Icon

Introduction

- Variables and More Variables… the legacy lives on
  - First presented in 1997
  - Many of the techniques are still valid today
    - Calculation context
    - Using functions in creative ways
- Agenda for today
  - Brief review of calculation options
  - New "Cool Tricks"
  - Questions and discussion at the end

Calculation Options

- ETL process

- Universe objects

- Report structure items
    - Variables
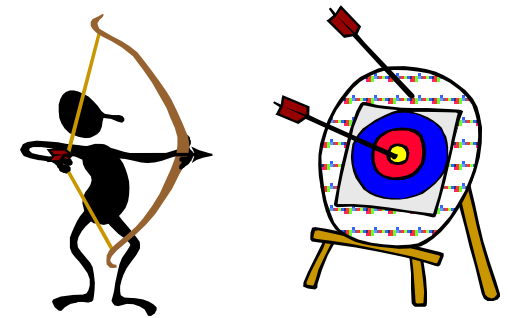   - Formulas
    - Constants

ETL Process

- Advantages
  - Build on core data
  - Can use procedural language
  - Ensure consistency across all data access paths
  - Calculate and store once, retrieve many times
- Disadvantages
  - Change management
  - Complexity
  - Impact analysis
- If a calculation is performed frequently and has high overhead consider pushing it to the ETL

Universe Objects

- Advantages
  - Build once - use many times
  - Use full range of database functions
  - Ensure consistency from report to report
  - Updates automatically propagate
- Disadvantages
  - Limited to information from one universe
  - Maintenance required by universe designer
  - Some aggregation issues can be tricky
    - See the Database Delegated feature from XI 3.0 for more ideas
  - Some functionality might be missing from the database
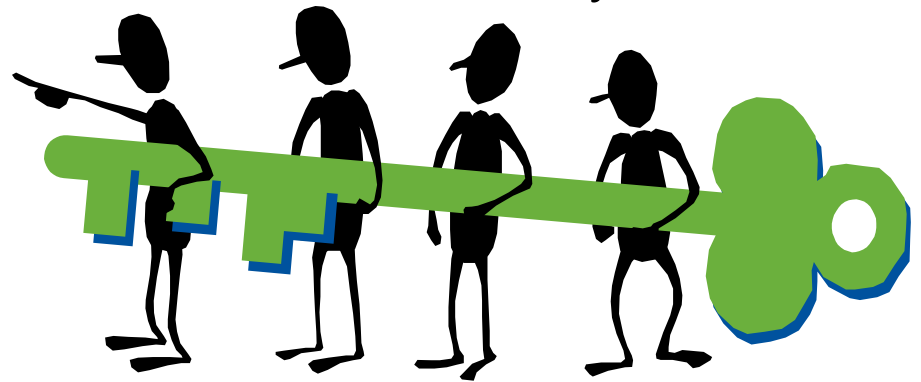   - If – Then – Else logic

Report Structure Items

- Advantages
  - Available on all platforms
  - Independent of SQL restrictions
  - Calculations based on document data
- Disadvantages
  - Stored in a single document
  - Require some level of technical expertise
  - Volume of data could impact performance

Why Variables?

- Convenience
    - Use with all product features
- Clarity
    - A name defines a purpose
- Complexity
    - Complex items can be built in steps
- Consistency
    - Reuse variables throughout the document and only edit once
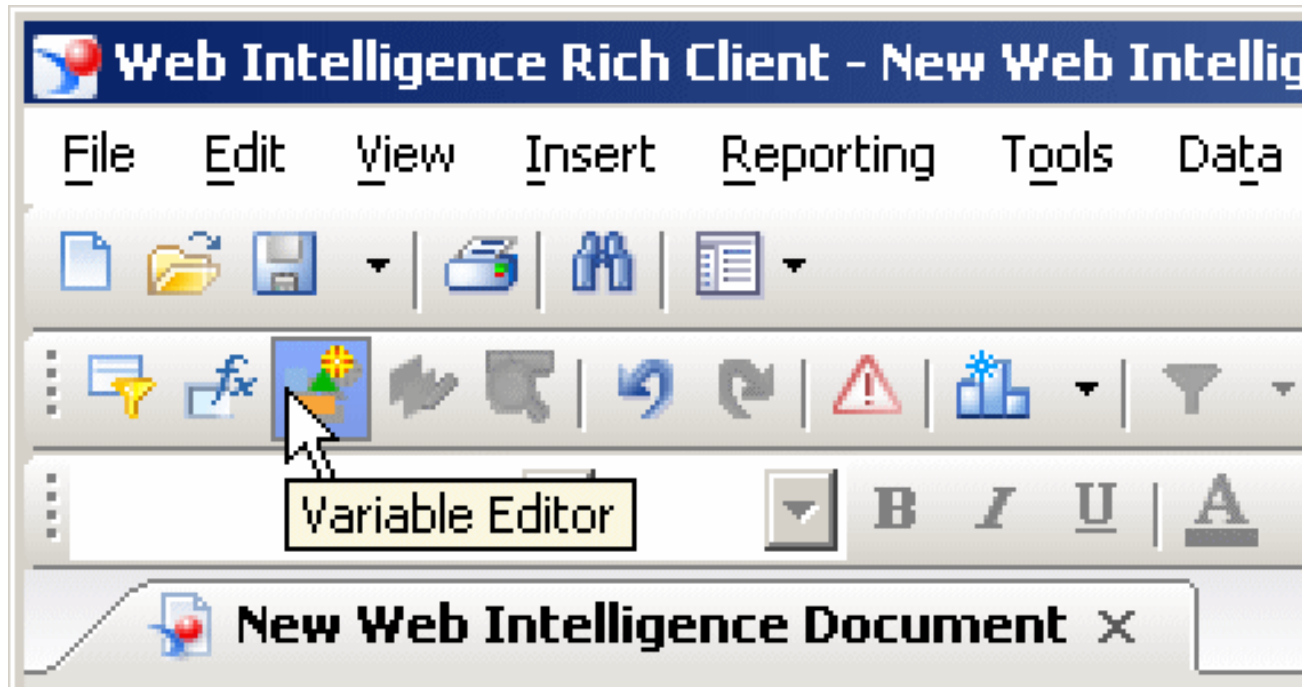
My Wish List

- I want to see "Universe Variables"
- The concept:
  - Pre-build report variables in the universe
  - Allow users to select the variables along with query objects
  - Variables are evaluated by the report engine rather than the database

Building Variables

- Click the Variable Editor

- Alternative: **Data** + **Variables** from menu

🖳 Create a basic formula, copy formula, convert to a variable

Cool Trick #1 Fixing Unbalanced Data Providers

- Web Intelligence allows multiple data providers
    - Common dimension objects are merged on the report
    - Report can contain common data from multiple sources
- What are unbalanced data providers and why are they a problem?

What Are Unbalanced Data Providers?

- Dimensions are the key values for a data provider

- Just as keys are joined between database tables, dimensions are linked (merged) between data providers

- When the number of dimensions from two data providers are not equal, they are unbalanced

Country	Resort	Revenue
France	French Riviera	$835,420.00
US	Bahamas Beach	$971,444.00
US	Hawaiian Club	$1,479,660.00

Resort	Number of guests
Bahamas Beach	565
French Riviera	446
Hawaiian Club	540

- In this example
  - Resort is common to both data providers and merged
  - Country is unbalanced because it has no match

Unbalanced Data Provider Impact

- Unlinked dimensions cause problems in a report block
    - Measures are out of context for those dimension values
    - Total values are displayed for each dimension value

Country	Resort	Revenue	Number of guests
France	French Riviera	$835,420.00	446
US	Bahamas Beach	$971,444.00	565
US	Hawaiian Club	$1,479,660.00	540

Country	Revenue	Number of guests
France	$835,420.00	1,551
US	$2,451,104.00	1,551

⌨ Create common block from unbalanced data providers

Using ForceMerge()

- The ForceMerge() function can fix unbalanced data issues
  - Requires Web Intelligence to consider all merged dimensions rather than just those participating in the block

```
=ForceMerge([Number of guests])
```

Country	Revenue	Guests
France	$835,420.00	446
US	$2,451,104.00	1,105

⌨ Demonstrate using ForceMerge() to fix unbalanced data providers

ForceMerge() Does Not Fix Everything

- Consider the following two unbalanced data providers

Year	Quarter	Number of guests
FY2005	Q3	36
FY2005	Q4	130
FY2006	Q1	132
FY2006	Q2	139
FY2006	Q3	116
FY2006	Q4	135

Quarter	Number of guests
Q1	394
Q2	410
Q3	357
Q4	390

- Quarters roll up to years…

- … but the data from block #2 has already been summarized by quarter

- Required data is simply not available

🖳 Demonstrate unfixable unbalanced data providers

Cool Trick #1 Fixing Unbalanced Data Providers

- Unbalanced data providers occur when
  - Two or more data providers exist in a document
  - Common dimensions are linked
  - Extra dimensions are present on one (or both) sides
- Impact of unbalanced data providers
  - Measures project to the higher context and repeat for each dimension value
- Fixing unbalanced data providers
  - ForceMerge() or MultiCube() can fix some issues
  - Data must be able to "project up" a hierarchy
  - Data cannot be "broken down" to a lower level of detail

Cool Trick #2 Calculating Business Days

- Question: How can I calculate the number of business days between two dates?

- Answer: Build a calendar table with a business day flag

- Question: What if I can't do that?

- Answer: ...

- At this point the solution often breaks down into a massive "If – Then – Else" process

- Side question: Do we have array functions in Web Intelligence?

Functions To Be Used

- DaysBetween()
  - Returns the number of days between two dates
- DayNumberOfWeek()
  - Returns the number of the day (Monday = 1, Tuesday = 2, …)
- Mod()
  - Returns the remainder of a division
- Substr()
  - Returns a specified portion of an input string
- ToNumber()
  - Converts a character value to a numeric value
- Truncate()
  - Removes the decimal portion from a float value without rounding

Making Complex Simple

- Break a complex calculation down into simple parts
- Business Days between is defined as the number of Monday through Friday days in a date range
- Steps
  - Calculate the number of full weeks * 5 days
  - Calculate the number of remaining days in partial week
  - Determine how many of those days to add

Number of Full Weeks

- No matter what day the start date is on a range of seven days will always include five business days
    - Sunday – Saturday
    - Tuesday – Monday
- Count the number of full weeks and multiply by five business days

October 2009 calendar

Formula For Full Weeks

```
Truncate(DaysBetween([Start Date];
[End Date]) / 7 ; 0) * 5
```

- Calculate total days between two dates
- Divide by 7 to get number of full weeks
- Truncate the result to remove partial week (decimal value)
- Multiply by 5 to get business days per each full week

Number of Days	Divide by 7	Truncate Full Weeks	Business Days
3	0.43	0	0
7	1.00	1	5
13	1.86	1	5
14	2.00	2	10

How Many Days Left?

- I need to know how many days to add for a partial week

Number of Days	Mod(X; 7) = Days to Add
3	3
7	0
13	6
14	0

- And which day my date range starts on

Day	DayNumberOfWeek()
Monday	1
Tuesday	2
Friday	5
Sunday	7

Now It Gets Complicated…

- For each starting day and extra day combination I have to know how many days to add
  - If StartDay = Monday and DaysToAdd = 3
    Then BusinessDays = 3
  - If StartDay = Friday and DaysToAdd = 3
    Then BusinessDays = 1
  - If StartDay = Saturday and DaysToAdd = 1
    Then BusinessDays = 0
- Parameters
  - Seven potential starting days
  - From 0 to 6 additional days or 7 cases to consider
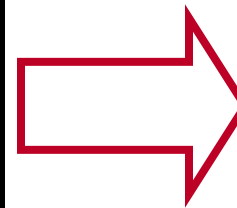 - 7 * 7 = 49 possible results

Building A Matrix

- Table showing all 49 cases and the desired results

Starting Day		Days to Add						
Day Name	Day Number	0	1	2	3	4	5	6
Mon	1	1	2	3	4	5	5	5
Tue	2	1	2	3	4	4	4	5
Wed	3	1	2	3	3	3	4	5
Thu	4	1	2	2	2	3	4	5
Fri	5	1	1	1	2	3	4	5
Sat	6	0	0	1	2	3	4	5
Sun	7	0	1	2	3	4	5	5

Simulating An Array

- Take each row of pre-calculated answers
- Combine them as one long string

Starting Day		Days to Add						
Day Name	Day Number	0	1	2	3	4	5	6
Mon	1	1	2	3	4	5	5	5
Tue	2	1	2	3	4	4	4	5
Wed	3	1	2	3	3	3	4	5
Thu	4	1	2	2	2	3	4	5
Fri	5	1	1	1	2	3	4	5
Sat	6	0	0	1	2	3	4	5
Sun	7	0	1	2	3	4	5	5

1234555
1234445
1233345
1222345
1112345
0012345
0123455

Creating a String Array

- Here is the string from the prior slide again

  1234555123444512333451222345111234500123450123455

- All of the following are true

  - Every entry is a single digit number

  - There are 49 entries in my string

  - Each set of seven characters represents the values to add for a given starting day from a week

- I can treat this string as an array

  - With the Substr() function I can return a value

  - How can I calculate the proper index for my array?

Index / Offset Into The Array

- Consider this formula

```
=DayNumberOfWeek([Start Date])-1)*7)
+Mod(DaysBetween([Start Date];[End
Date]);7)+1
```

- Step through some scenarios

Day Number	(Day Num – 1) * 7	Days to Add	String Index
Monday = 1	0	3	3
Friday = 5	28	3	31
Saturday = 6	35	1	36

- The result of this formula is the index to my string array to the position containing the proper number of business days to add

Some Business Days Examples

- Start Day = Sunday, End Day is Thursday
  - Sunday = 7, 5 days to add (Sunday to Thursday)
  - (7 – 1) * 7 + 5 gives index value of 47

  12345551234445123334512223451112345001234501234 5 5

  ⬆

- Start Day = Tuesday, End Day is Sunday
  - Tuesday = 2, 6 days to add
  - (2 – 1) * 7 + 6 gives index value of 13

  1234555123 4 4 4512333451222345111234500123450123455

  ⬆

Cool Trick #2 Calculating Business Days

- The final formula

```
=(Truncate(DaysBetween([Start Date]; [End
Date]) / 7 ; 0) * 5) +
ToNumber(Substr("123455512344512333451222
345111234500123450123455";
((DayNumberOfWeek([Start Date]) - 1) * 7 )
+ Mod(DaysBetween([Start Date];[End
Date]); 7) + 1; 1))
```

⌨ Demonstrate formula for counting business days

Is This Really The Best Solution?

- Does not require multiple nested "IF" statements
- The array can be adjusted per business needs
  - For example, perhaps Wednesday is not a business day
  - Generate a new string array based on this calendar
  - Formula as shown today will continue to work
- What about leap years?
  - Does not matter because there are no leap weeks

Other Uses For This Technique

- Ever wanted to do HLOOKUP() or VLOOKUP() in Web Intelligence?
- Requirements
  - Constant size for array values
  - Calculation to derive index (offset) into the string array
- Array string data
  - Can be stored in a universe for maximum flexibility
  - Can be generated on the fly
- Do not try to use a delimited string with the Pos() function
  - Pos() only finds the first occurrence
  - Nested Pos() functions get really ugly really fast

Questions and Discussion

- This presentation will be posted on my blog after the conference
  - As time permits I will provide slide narration as blog posts

- Author Information
  - David G. Rathbun
  - BI Solutions Architect, PepsiCo B+IS
  - BI Blog www.dagira.com

Thank you for participating

Please remember to complete and return your evaluation form following this session.