

# Untangling Data For Analytics and Reporting Systems

Doug Mack  
[mackd@us.ibm.com](mailto:mackd@us.ibm.com)





All too often, DW/BI teams lose their business focus during the project and concentrate on selecting a BI tool rather than providing full end-to-end solutions.

[www.kimballgroup.com/2013/01/08/design-tip-151-bi-components-for-business-value/](http://www.kimballgroup.com/2013/01/08/design-tip-151-bi-components-for-business-value/)



## Data Transformation

- It's all about the data
- 70% of a BI project can be data “cleansing” and optimization related
- If  $1 + 1 = 3$  in the source data, then getting the wrong answer faster and in graphical form really doesn't buy you much
- If the data is not trusted, the project will fail
- If the data is not understood, the task is very difficult
- BI is intended to REVERSE the 80/20 rule
  - 80% Gathering, 20% Analyzing

# Transformation Examples

## Legacy Date

### OrderDT – P8 Decimal

02132016  
03212017  
06012016



## Date Usage Desired by End Users

Order Date Month  
Order Date Year  
Order Date Quarter  
Order Date Day Name

Current Date  
Current Month  
Current Year  
Today's Date – 30  
Same Date Last Year

Is it a Weekend?  
Is it a Holiday?  
What was the weather?  
Was there a full moon?

# Agenda

- Meta Data as a Foundation
- ETL (Extract, Transformation, and Load) is often the Best Practice to support Analytics Applications

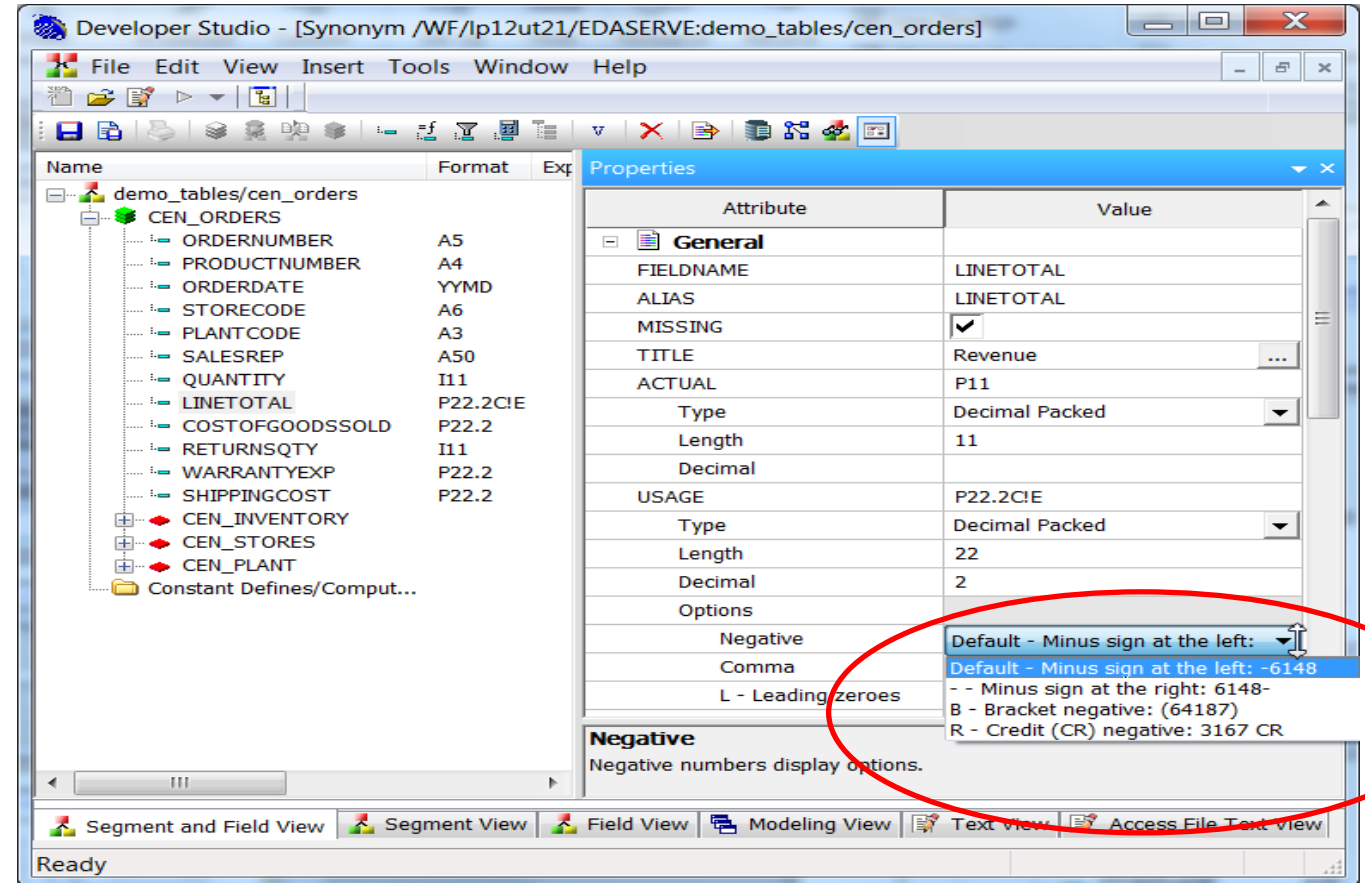


## DB2 Web Query Meta Data

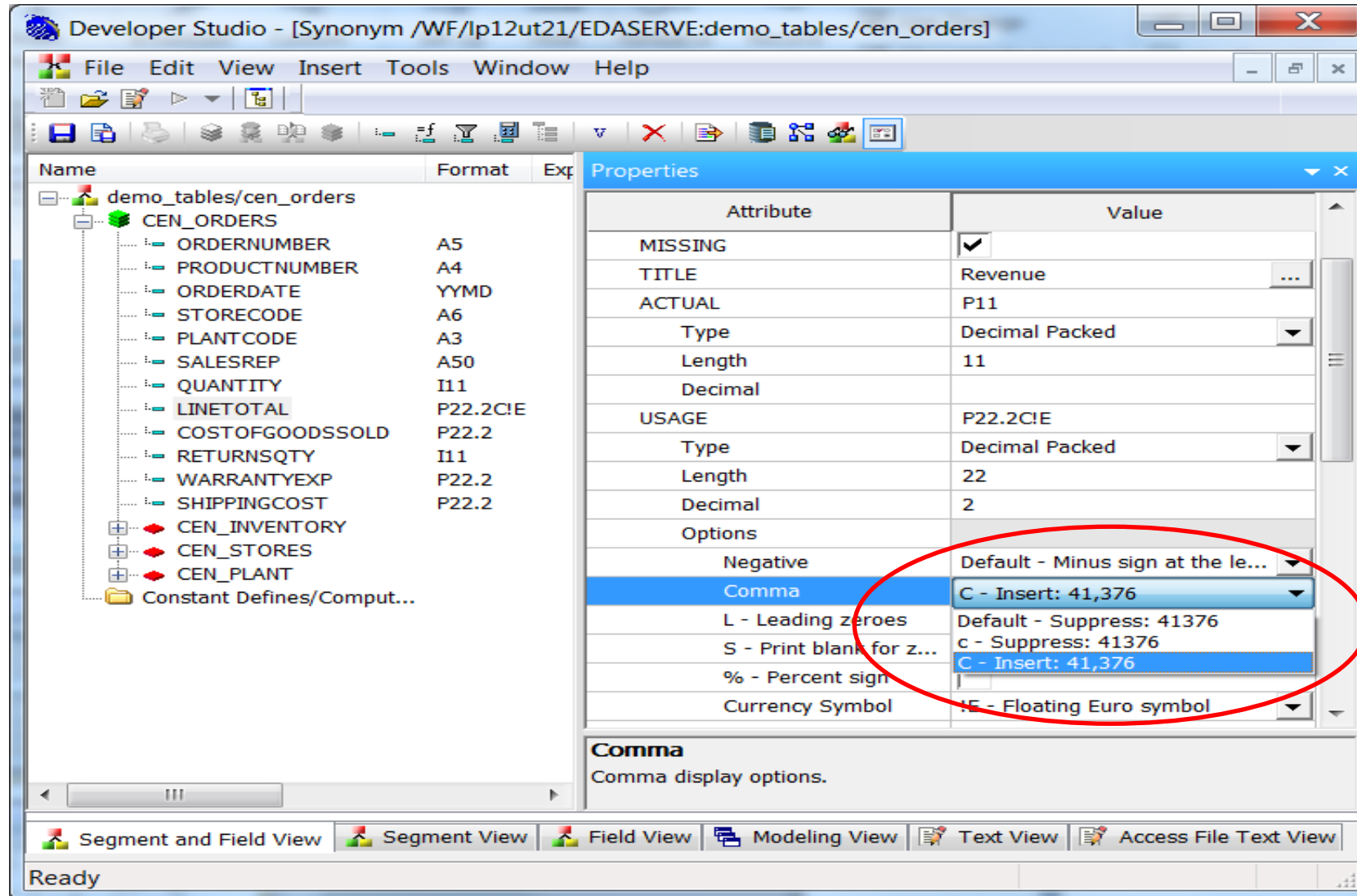
- Called SYNONYMS in DB2 Web Query terms
- Can be used to shield complexity of database from report authors when querying production databases
  - Make the data meaningful to business users through simplification and transformation
  - Add business rules
- Can (and is actually required) to be used for defining source and target data elements in a data warehouse architecture
- Define ONCE, use EVERYWHERE
  - And if you need to change something, you change it ONCE and not in every report

# Examples of Meta Data Transformation Capabilities – Formatting Fields

- Field Formats
  - Set default formats
  - Every Report will use this default format
  - Can be overridden in each report
  
- How do you want negative numbers depicted?



# Inserting thousands separators



The screenshot shows the 'Properties' window for the table 'CEN\_ORDERS'. The 'Options' section is expanded, and the 'Comma' option is selected. A red circle highlights the dropdown menu for the 'Comma' option, which shows the following options:

- Default - Minus sign at the le...
- C - Insert: 41,376
- Default - Suppress: 41376
- c - Suppress: 41376
- C - Insert: 41,376

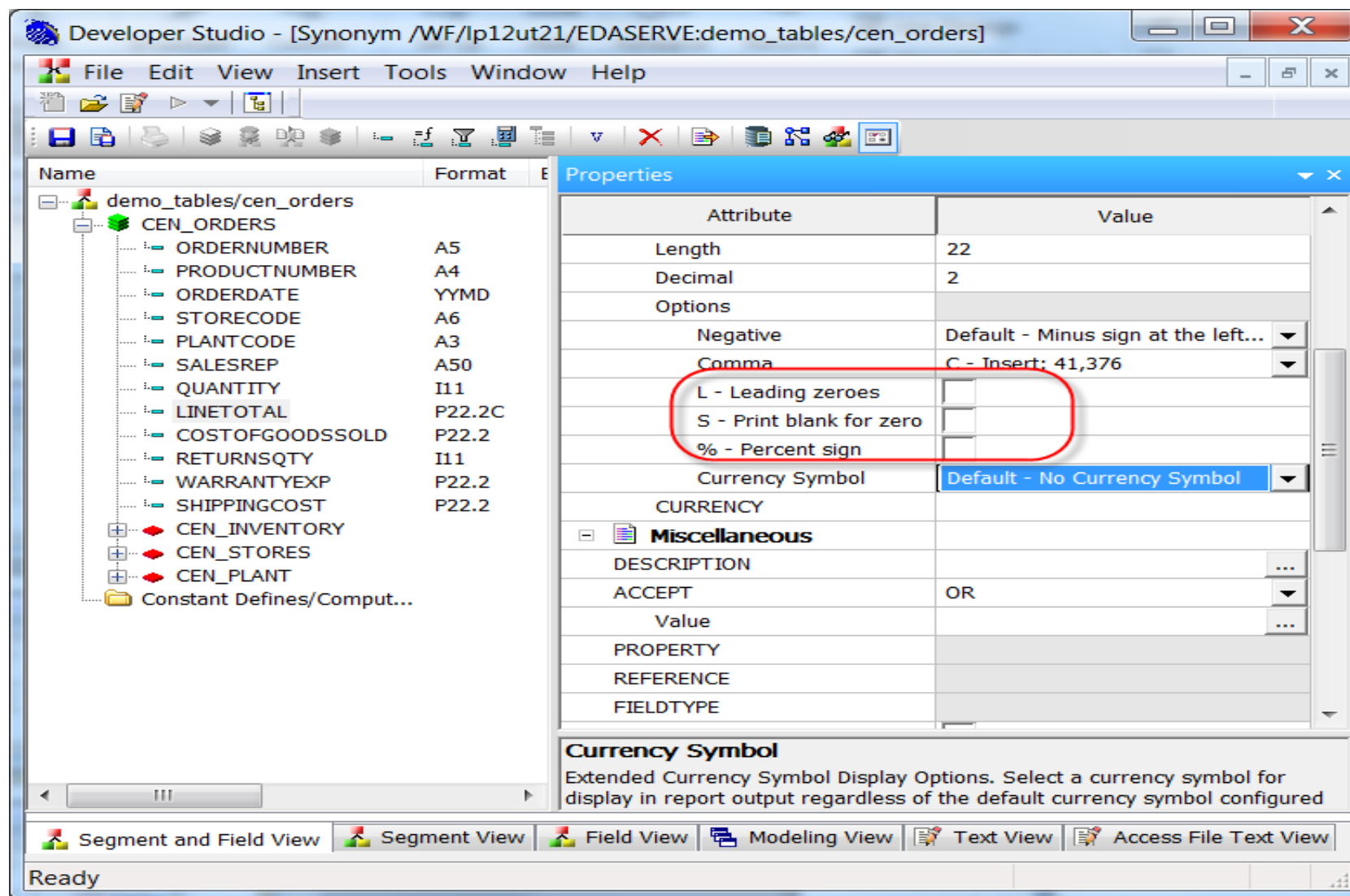
The 'Comma' option is currently selected, and the dropdown menu is open, showing the 'C - Insert: 41,376' option highlighted.

Attribute	Value
MISSING	<input checked="" type="checkbox"/>
TITLE	Revenue
ACTUAL	P11
Type	Decimal Packed
Length	11
Decimal	
USAGE	P22.2C!E
Type	Decimal Packed
Length	22
Decimal	2
Options	
Negative	Default - Minus sign at the le...
Comma	C - Insert: 41,376
L - Leading zeroes	Default - Suppress: 41376
S - Print blank for z...	c - Suppress: 41376
% - Percent sign	C - Insert: 41,376
Currency Symbol	€ - Floating Euro symbol

**Comma**  
Comma display options.



# Leading zeroes, percent sign



Developer Studio - [Synonym /WF/lp12ut21/EDASERVE:demo\_tables/cen\_orders]

File Edit View Insert Tools Window Help

Name	Format	Properties
demo_tables/cen_orders		
CEN_ORDERS		
ORDERNUMBER	A5	
PRODUCTNUMBER	A4	
ORDERDATE	YYMD	
STORECODE	A6	
PLANTCODE	A3	
SALESREP	A50	
QUANTITY	I11	
LINETOTAL	P22.2C	
COSTOFGOODSSOLD	P22.2	
RETURNSQTY	I11	
WARRANTYEXP	P22.2	
SHIPPINGCOST	P22.2	
CEN_INVENTORY		
CEN_STORES		
CEN_PLANT		
Constant Defines/Comput...		

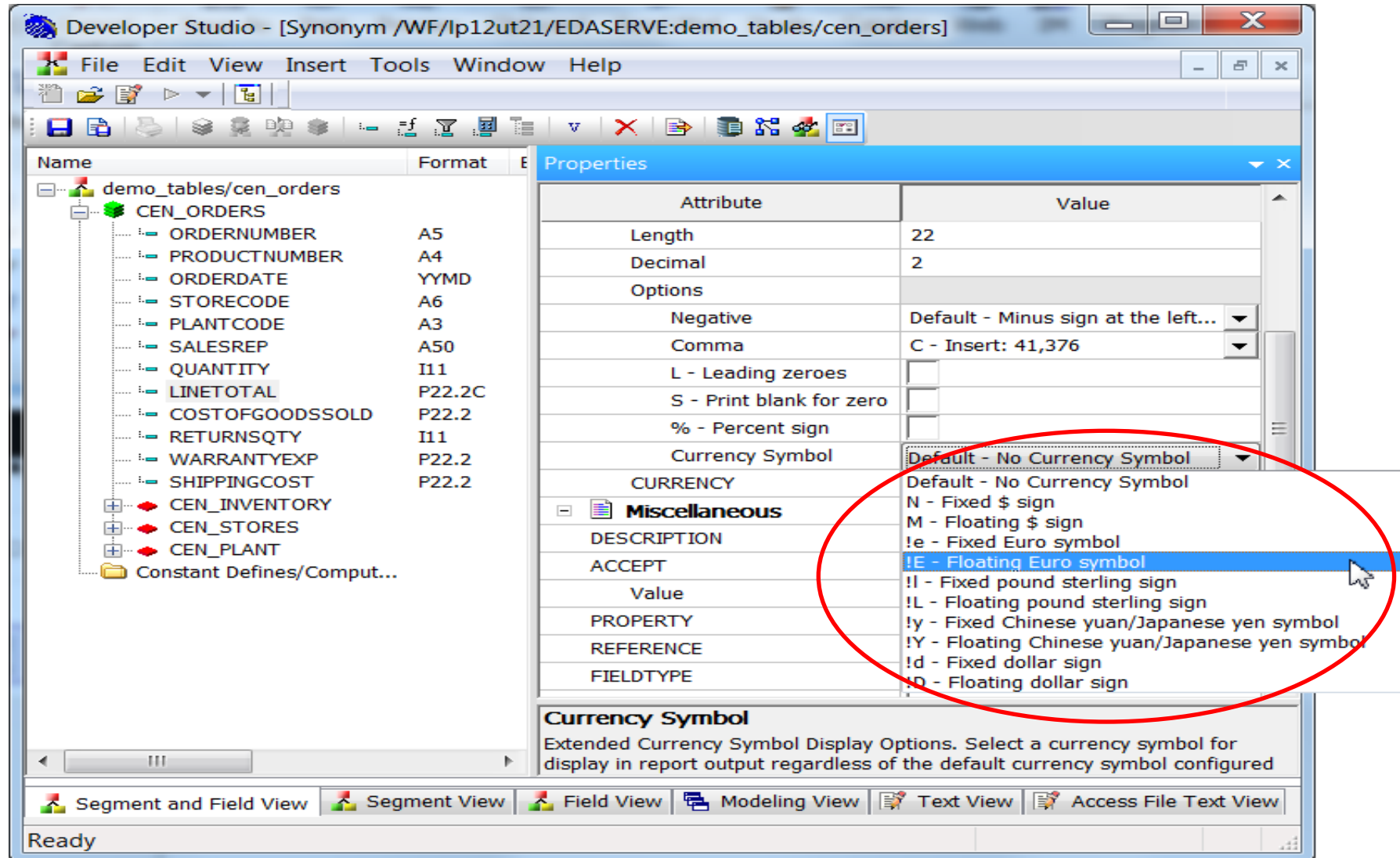
Attribute	Value
Length	22
Decimal	2
Options	
Negative	Default - Minus sign at the left...
Comma	C - Insert: 41,376
L - Leading zeroes	<input type="checkbox"/>
S - Print blank for zero	<input type="checkbox"/>
% - Percent sign	<input type="checkbox"/>
Currency Symbol	Default - No Currency Symbol
CURRENCY	
Miscellaneous	
DESCRIPTION	...
ACCEPT	OR
Value	...
PROPERTY	
REFERENCE	
FIELDTYPE	

**Currency Symbol**  
Extended Currency Symbol Display Options. Select a currency symbol for display in report output regardless of the default currency symbol configured

Segment and Field View Segment View Field View Modeling View Text View Access File Text View

Ready

# Fixed/floating dollar, Euro signs



Developer Studio - [Synonym /WF/lp12ut21/EDASERVE:demo\_tables/cen\_orders]

File Edit View Insert Tools Window Help

Name	Format
demo_tables/cen_orders	
CEN_ORDERS	
ORDERNUMBER	A5
PRODUCTNUMBER	A4
ORDERDATE	YYMD
STORECODE	A6
PLANTCODE	A3
SALESREP	A50
QUANTITY	I11
LINETOTAL	P22.2C
COSTOFGOODSSOLD	P22.2
RETURNSQTY	I11
WARRANTYEXP	P22.2
SHIPPINGCOST	P22.2
CEN_INVENTORY	
CEN_STORES	
CEN_PLANT	
Constant Defines/Comput...	

Attribute	Value
Length	22
Decimal	2
Options	
Negative	Default - Minus sign at the left...
Comma	C - Insert: 41,376
L - Leading zeroes	<input type="checkbox"/>
S - Print blank for zero	<input type="checkbox"/>
% - Percent sign	<input type="checkbox"/>
Currency Symbol	Default - No Currency Symbol
CURRENCY	
Miscellaneous	
DESCRIPTION	
ACCEPT	
Value	
PROPERTY	
REFERENCE	
FIELDTYPE	

**Currency Symbol**

- Default - No Currency Symbol
- N - Fixed \$ sign
- M - Floating \$ sign
- !e - Fixed Euro symbol
- !E - Floating Euro symbol**
- !! - Fixed pound sterling sign
- !L - Floating pound sterling sign
- !y - Fixed Chinese yuan/Japanese yen symbol
- !Y - Floating Chinese yuan/Japanese yen symbol
- !d - Fixed dollar sign
- !D - Floating dollar sign

**Currency Symbol**  
Extended Currency Symbol Display Options. Select a currency symbol for display in report output regardless of the default currency symbol configured

Segment and Field View Segment View Field View Modeling View Text View Access File Text View

Ready

# Hiding columns (so they do not appear in development tools)

The screenshot shows the IBM Data Studio interface for defining a table. The left pane shows a tree view of the database structure, with the table 'cen\_orders' expanded to show its columns. The 'STORECODE' column is circled in red. The right pane shows the 'Properties' window for the selected column, where the 'INTERNAL' property is checked, also circled in red. Below the properties window, a description for the 'INTERNAL' property is provided.

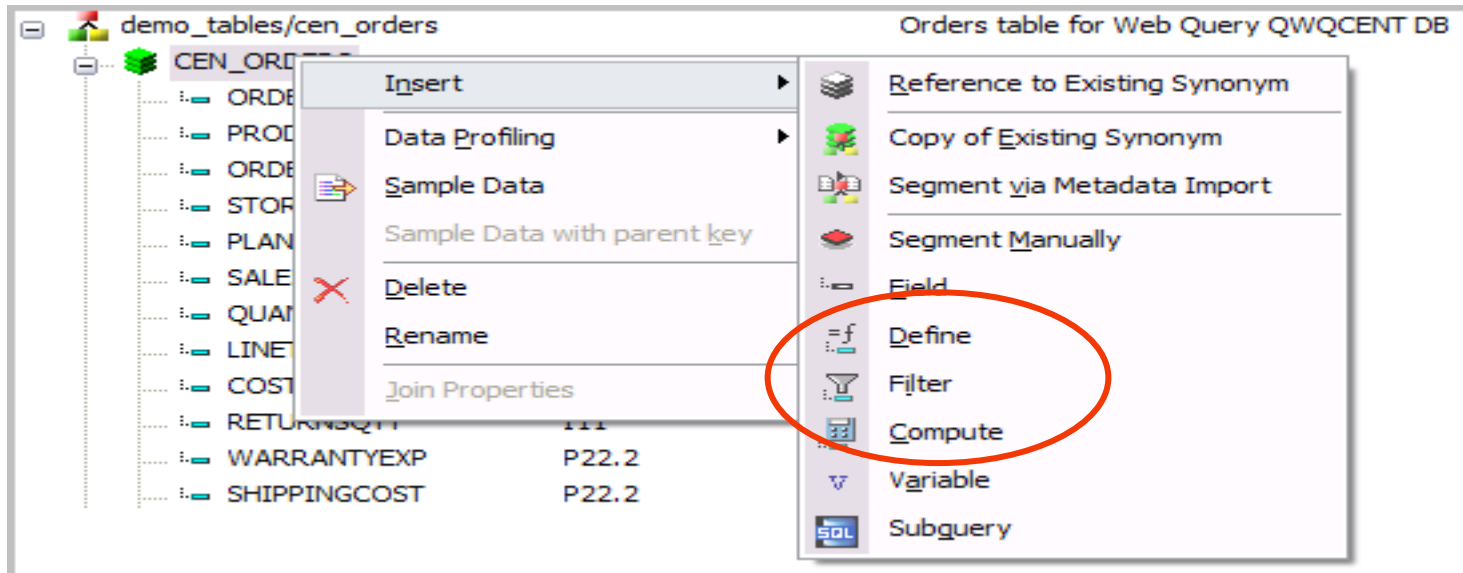
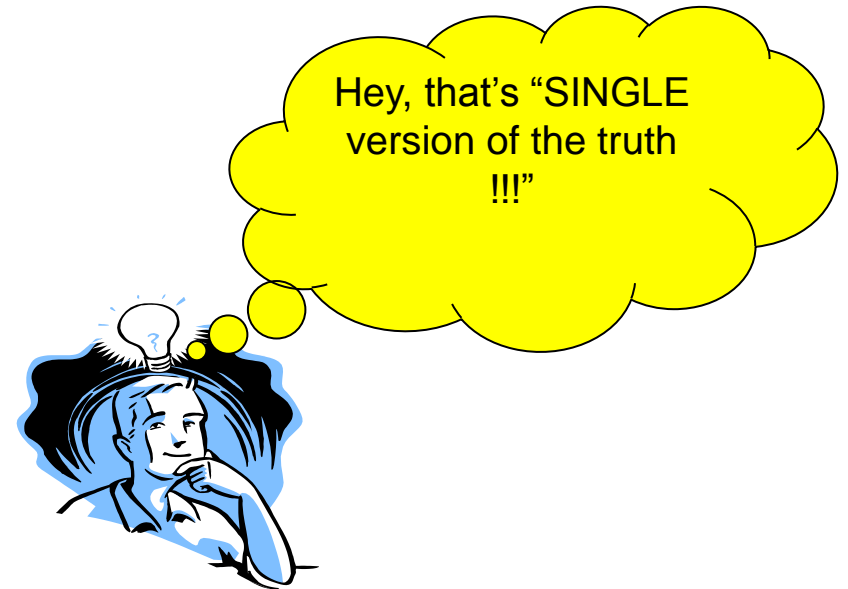
Name	Format
demo_tables/cen_orders	
CEN_ORDERS	
ORDERNUMBER	A5
PRODUCTNUMBER	A4
ORDERDATE	YYMD
STORECODE	A6
PLANTCODE	A3
SALESREP	A50
QUANTITY	I11
LINETOTAL	P22.2
COSTOFGOODSSOLD	P22.2
RETURNSQTY	I11
WARRANTYEXP	P22.2
SHIPPINGCOST	P22.2
CEN_INVENTORY	
PRODUCTNUMBER	A4
PRODUCTTYPE	A15
PRODUCTCATEG...	A30
PRODUCTNAME	A60
MODEL	A10

Attribute	Value
Value	...
PROPERTY	
REFERENCE	
FIELDTYPE	
I - Index	<input type="checkbox"/>
R - Readonly	<input type="checkbox"/>
ACCESS_PROPERTY	
INTERNAL	<input checked="" type="checkbox"/>
NEED_VALUE	<input type="checkbox"/>
Select By	
HELPMESSAGE	
SCD Type	
USE_STYLE	

**INTERNAL**  
Defines a field that will not appear in sample data or in the list of available fields.

# Creating derived (virtual) fields in Meta Data

- Define columns (Global)
  - Sometimes referred to as Virtual fields
  - Also called Define field in authoring tools (local to report)
  - Executed on database read
  - Calculated once for every record
- Compute fields (Global)
  - Also called Compute field in Authoring Tools (local to report)
  - Calculated once for every printed line in report
  - Executed after data is sorted and aggregated





## Editing DB2 Web Query synonyms: Define/Compute Column candidate examples

- Information needed in report but not stored in database in format desired
- Data transformation is necessary
  - Gross Profit Margin
  - Cost Ratio
  - Full customer name (Last\_Name, First\_Name Middle\_Initial)
  - String of codes to represent various values which need to be transformed

Cust Code  
7512469350  
1913003716  
7475386198  
7714825036

## Editing DB2 Web Query synonyms: Define/Compute Column candidate examples

- Information needed in report but not stored in database in format desired
- Data transformation is necessary
  - Gross Profit Margin
  - Cost Ratio
  - Full customer name (Last\_Name, First\_Name Middle\_Initial)
  - String of codes to represent various values which need to be transformed

### Cust Code

751246	935	0
191300	371	6
747538	619	8
771482	503	6

Credit Rating



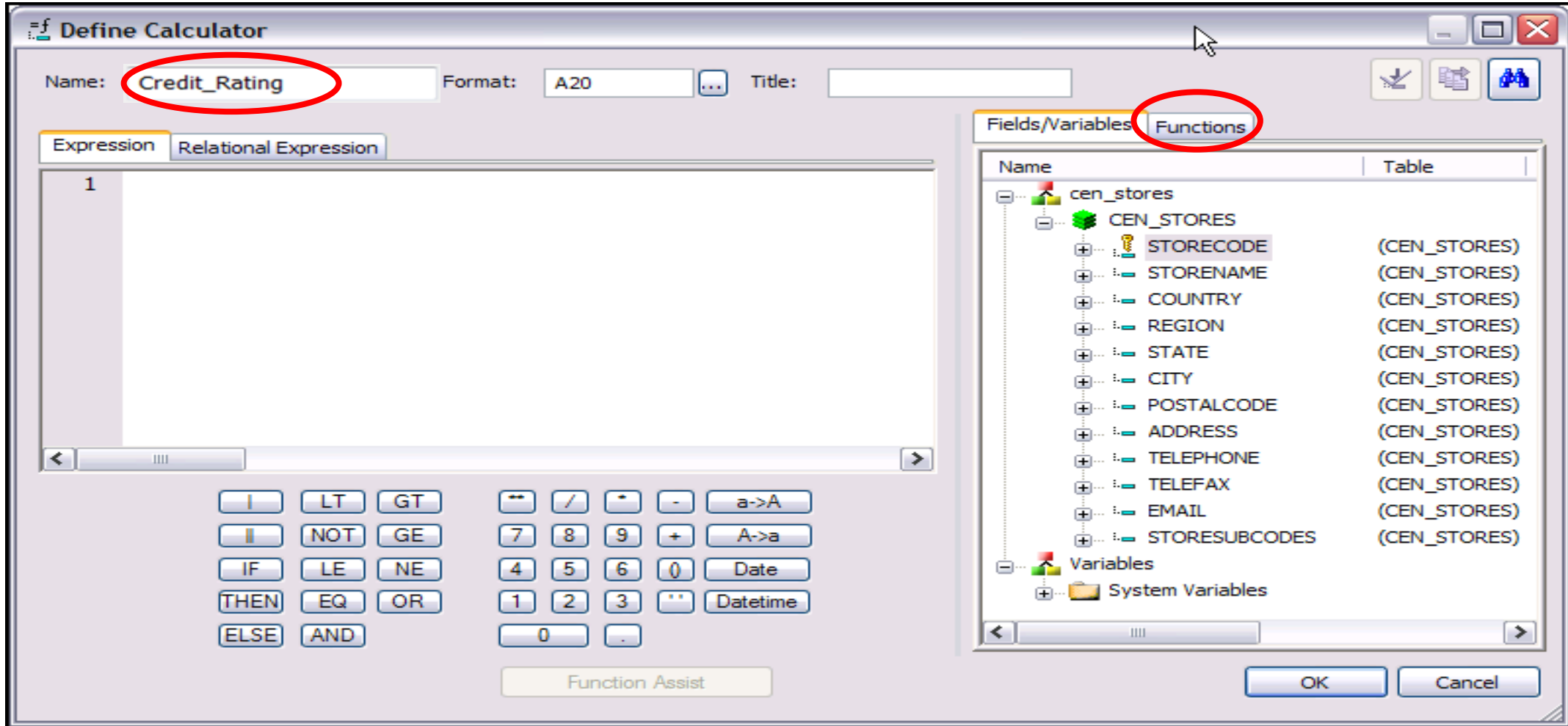
Orders placed in last year



Customer Number



# Creating a derived (virtual) field



**Define Calculator**

Name: **Credit\_Rating** Format: A20 Title:

Expression Relational Expression

1

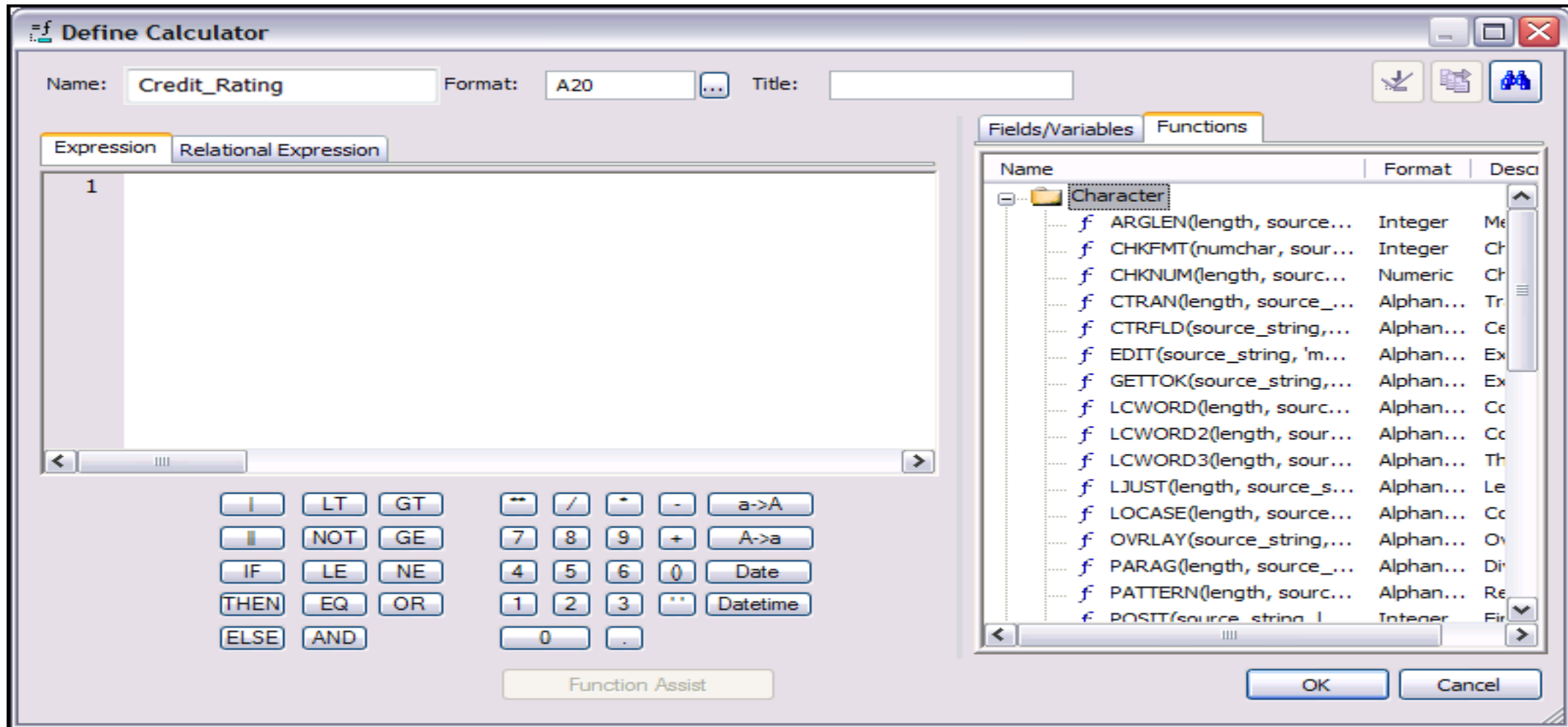
Fields/Variables **Functions**

Name	Table
cen_stores	
CEN_STORES	
STORECODE	(CEN_STORES)
STORENAME	(CEN_STORES)
COUNTRY	(CEN_STORES)
REGION	(CEN_STORES)
STATE	(CEN_STORES)
CITY	(CEN_STORES)
POSTALCODE	(CEN_STORES)
ADDRESS	(CEN_STORES)
TELEPHONE	(CEN_STORES)
TELEFAX	(CEN_STORES)
EMAIL	(CEN_STORES)
STORESUBCODES	(CEN_STORES)
Variables	
System Variables	

Function Assist

OK Cancel

# 100's of Built in Functions (or use DB2 SQL built in functions)



**Define Calculator**

Name:  Format:  Title:

Expression | Relational Expression

1

Function Assist

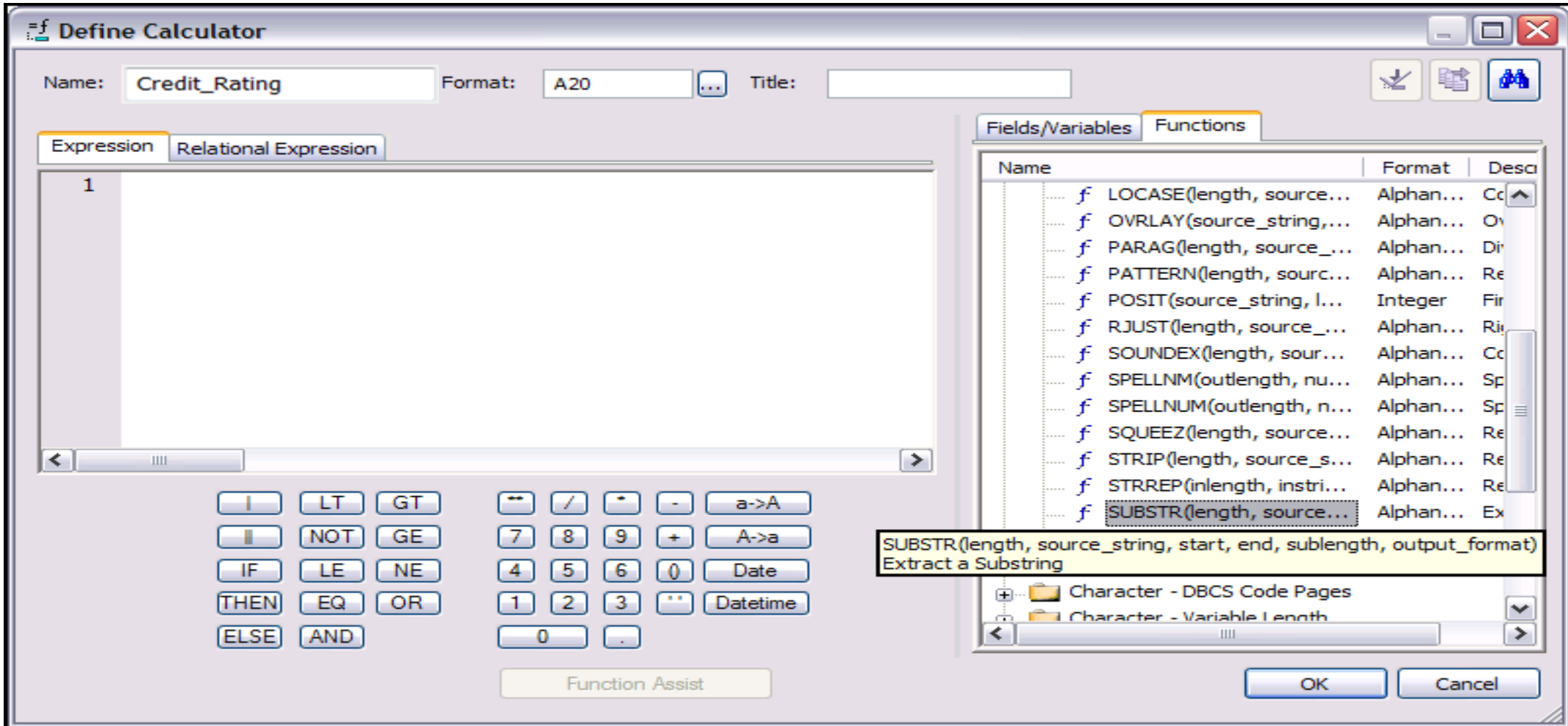
Fields/Variables | Functions

Name	Format	Desc
Character		
f ARGLEN(length, source...	Integer	Me
f CHKFMT(numchar, sour...	Integer	Ch
f CHKNUM(length, sourc...	Numeric	Ch
f CTRAN(length, source_...	Alphan...	Tr
f CTRFLD(source_string,...	Alphan...	Ce
f EDIT(source_string, 'm...	Alphan...	Ex
f GETTOK(source_string,...	Alphan...	Ex
f LCWORD(length, sourc...	Alphan...	Cc
f LCWORD2(length, sour...	Alphan...	Cc
f LCWORD3(length, sour...	Alphan...	Th
f LJUST(length, source_s...	Alphan...	Le
f LOCASE(length, source...	Alphan...	Cc
f OVLAY(source_string,...	Alphan...	Ov
f PARAG(length, source_...	Alphan...	Di
f PATTERN(length, sourc...	Alphan...	Re
f POSIT(source_string	Integer	Fi

OK Cancel



# 100's of Built in Functions (or use DB2 SQL built in functions)



**Define Calculator**

Name:  Format:  Title:

Expression:

Fields/Variables | **Functions**

Name	Format	Desc
f LOCASE(length, source...	Alphan...	Cc
f OVERLAY(source_string,...	Alphan...	Ov
f PARAG(length, source_...	Alphan...	Di
f PATTERN(length, sourc...	Alphan...	Re
f POSIT(source_string, l...	Integer	Fi
f RJUST(length, source_...	Alphan...	Ri
f SOUNDEX(length, sour...	Alphan...	Cc
f SPELLNM(outlength, nu...	Alphan...	Sp
f SPELLNUM(outlength, n...	Alphan...	Sp
f SQUEEZ(length, source...	Alphan...	Re
f STRIP(length, source_s...	Alphan...	Re
f STRREP(inlength, instri...	Alphan...	Re
f <b>SUBSTR(length, source...</b>	Alphan...	Ex

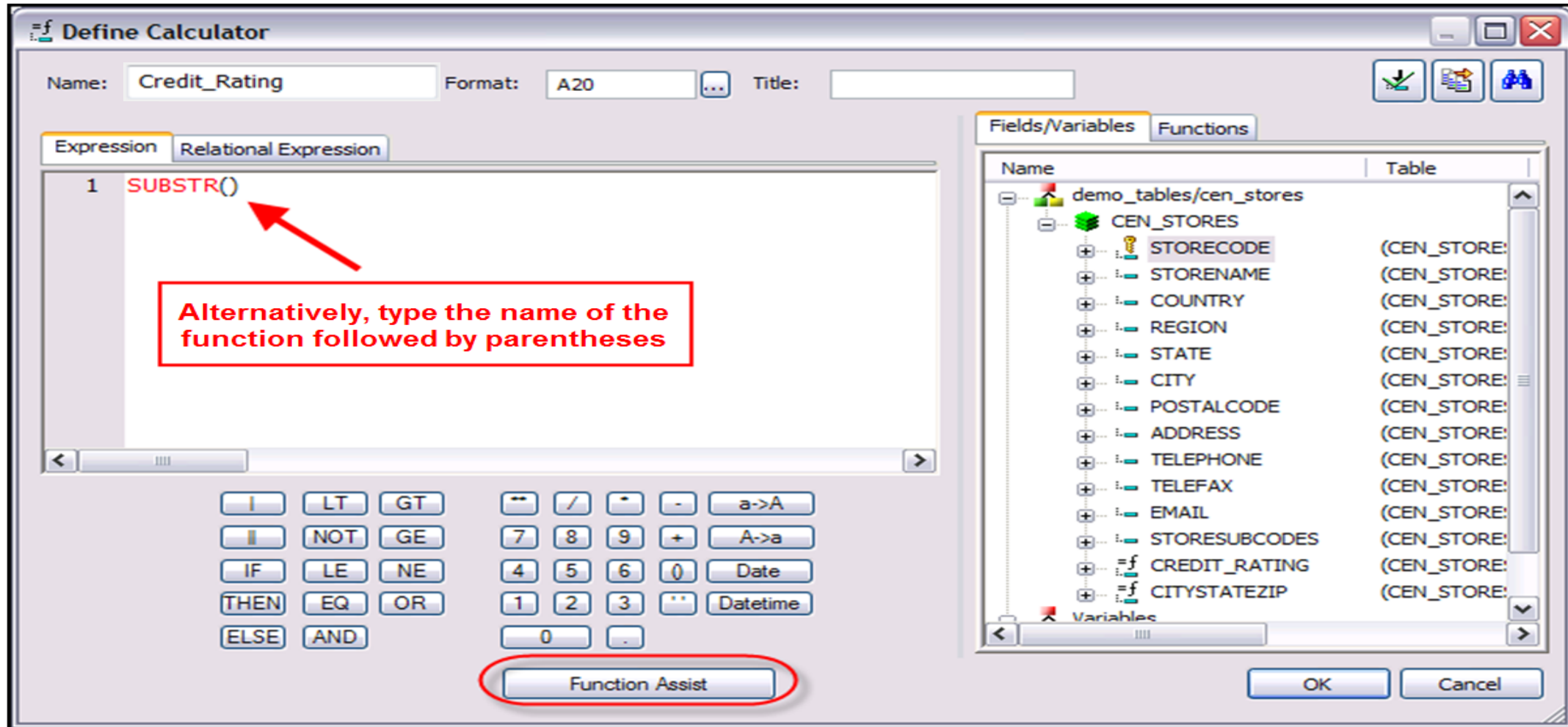
**SUBSTR(length, source\_string, start, end, sublength, output\_format)**  
Extract a Substring

Character - DBCS Code Pages  
Character - Variable Length

Function Assist

OK Cancel

# Expression Builder Makes it Easier to Create and Validate



**Define Calculator**

Name:  Format:  Title:

Expression **Relational Expression**

1 **SUBSTR()**

**Alternatively, type the name of the function followed by parentheses**

**Fields/Variables** Functions

Name	Table
demo_tables/cen_stores	
CEN_STORES	
STORECODE	(CEN_STORE:)
STORENAME	(CEN_STORE:)
COUNTRY	(CEN_STORE:)
REGION	(CEN_STORE:)
STATE	(CEN_STORE:)
CITY	(CEN_STORE:)
POSTALCODE	(CEN_STORE:)
ADDRESS	(CEN_STORE:)
TELEPHONE	(CEN_STORE:)
TELEFAX	(CEN_STORE:)
EMAIL	(CEN_STORE:)
STORESUBCODES	(CEN_STORE:)
CREDIT_RATING	(CEN_STORE:)
CITYSTATEZIP	(CEN_STORE:)

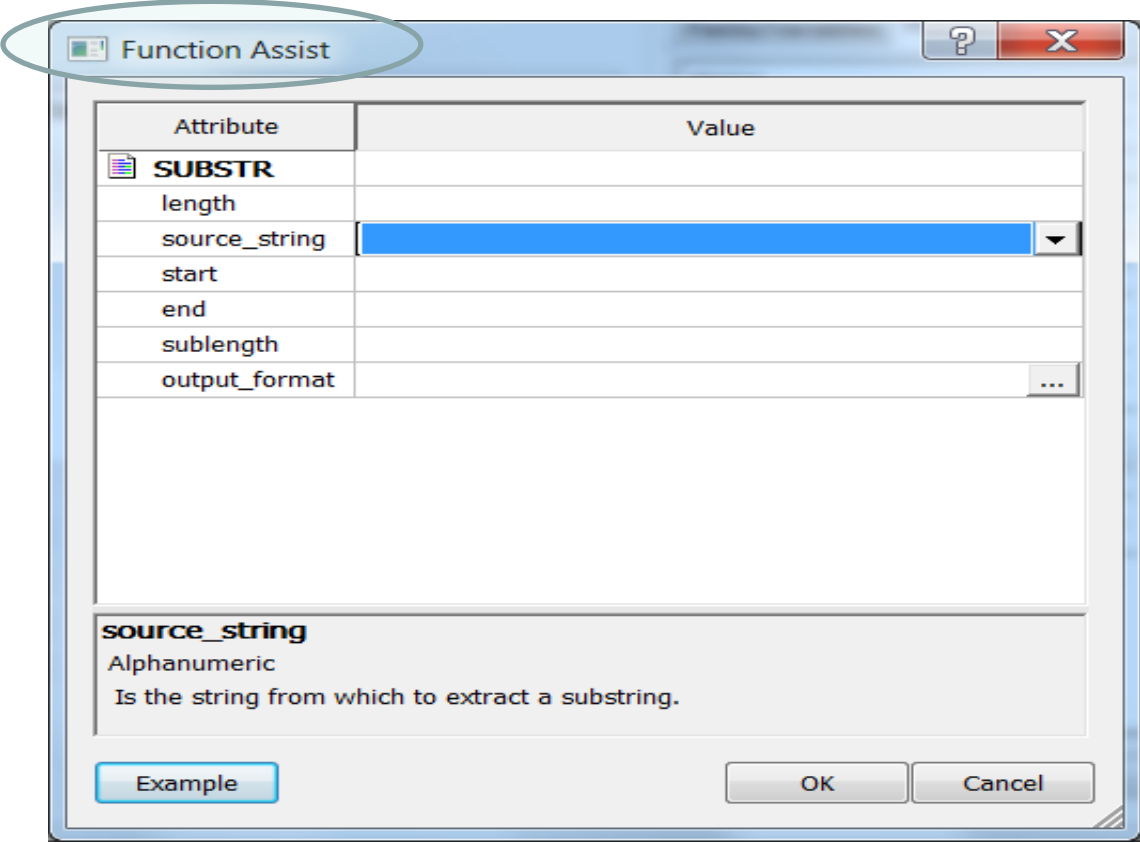
Function Assist

OK Cancel

# Function Assist is Like a Wizard



Function Assist!  
Way cool!!



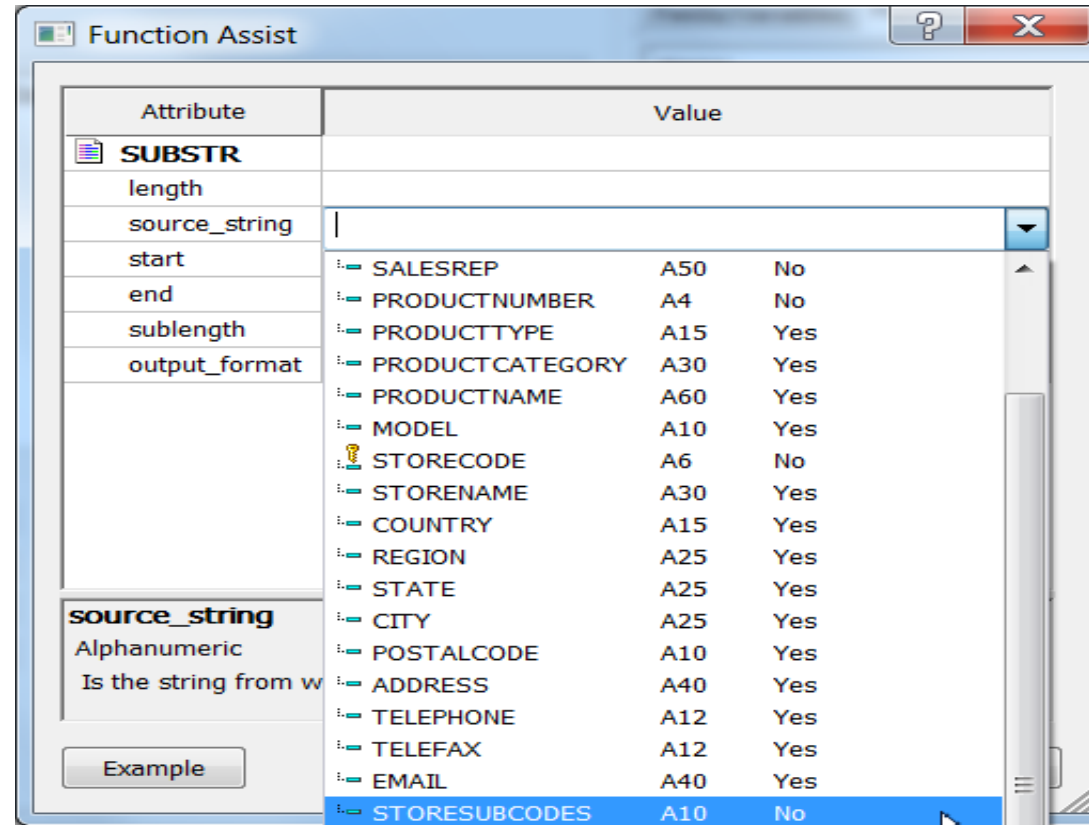
The screenshot shows a window titled "Function Assist" with a table of attributes and values for the SUBSTR function. The "source\_string" attribute is highlighted in blue. Below the table, there is a description for "source\_string" and buttons for "Example", "OK", and "Cancel".

Attribute	Value
<b>SUBSTR</b>	
length	
source_string	
start	
end	
sublength	
output_format	...

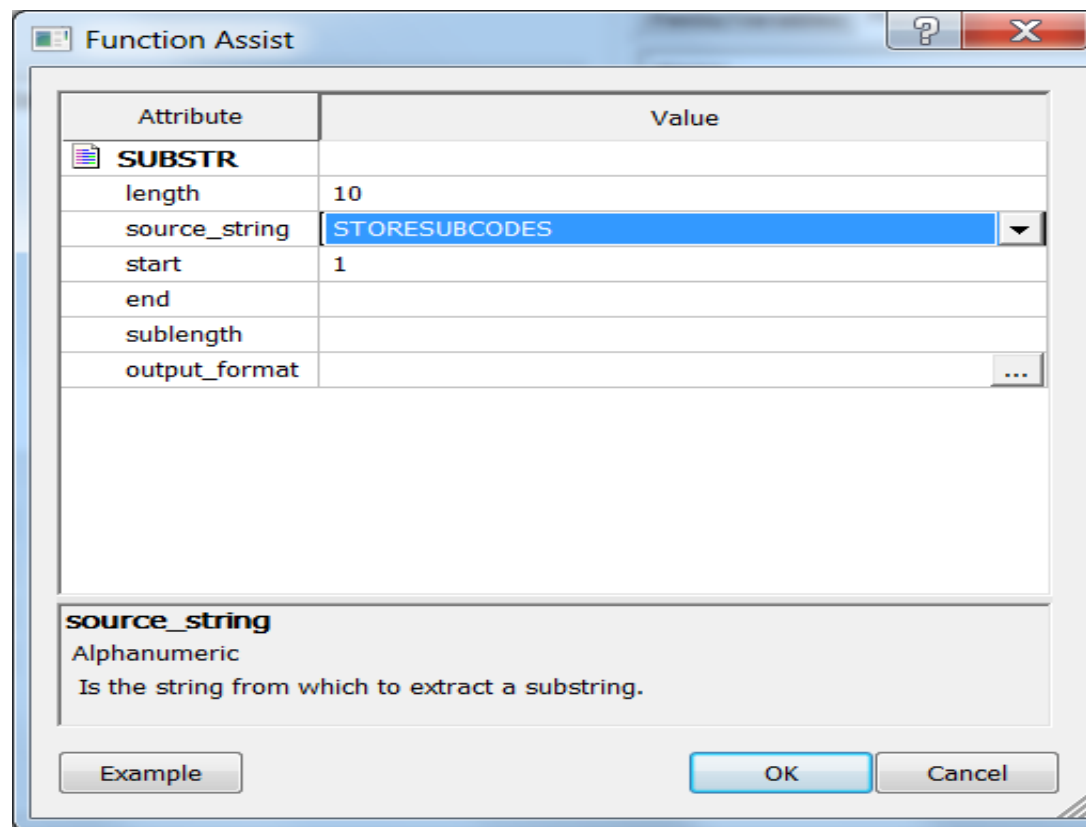
**source\_string**  
Alphanumeric  
Is the string from which to extract a substring.

Example OK Cancel

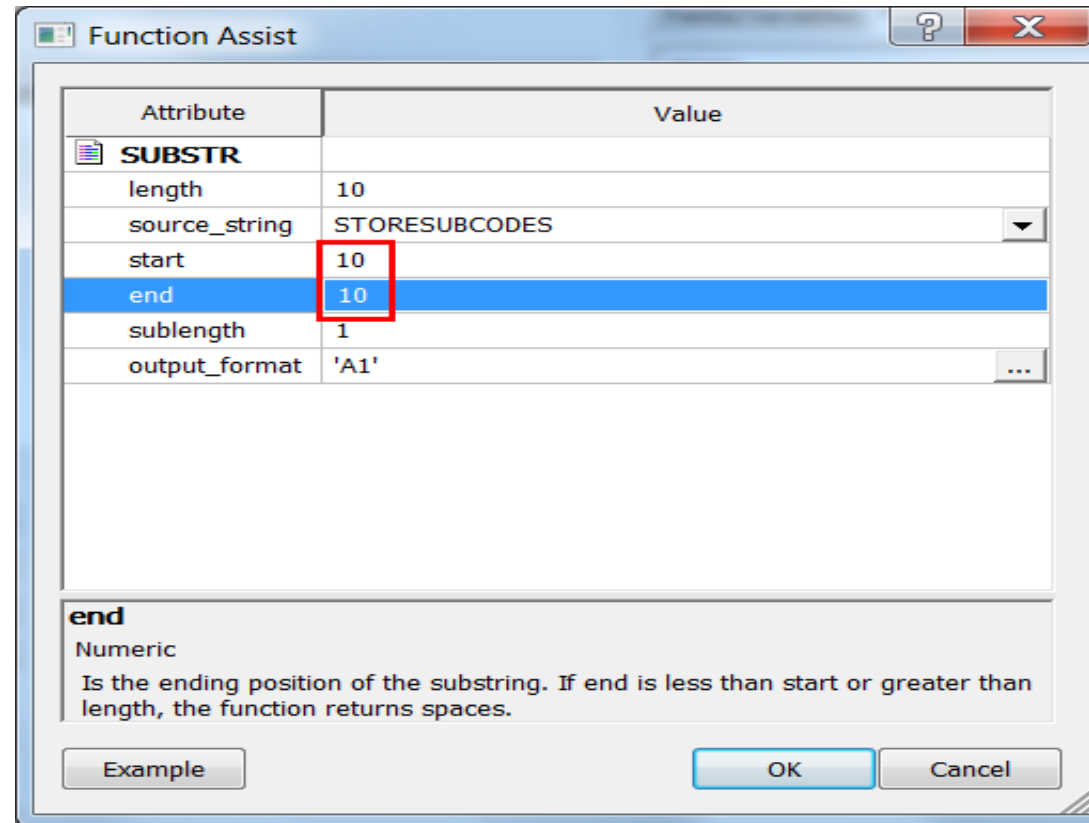
# Function Assist is Like a Wizard



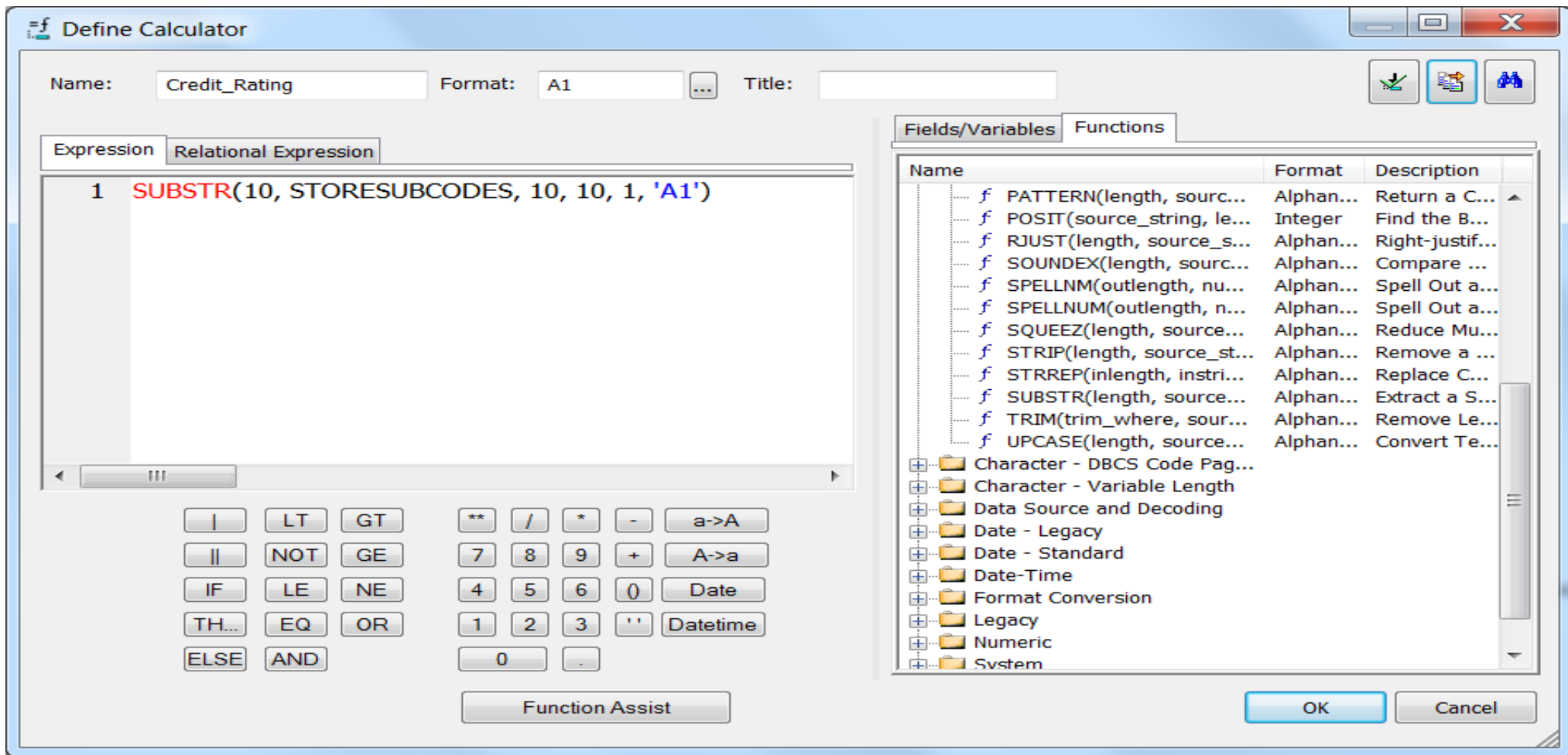
# Function Assist is Like a Wizard



# Function Assist is Like a Wizard



# Expression Builder Makes it Easier to Create and Validate



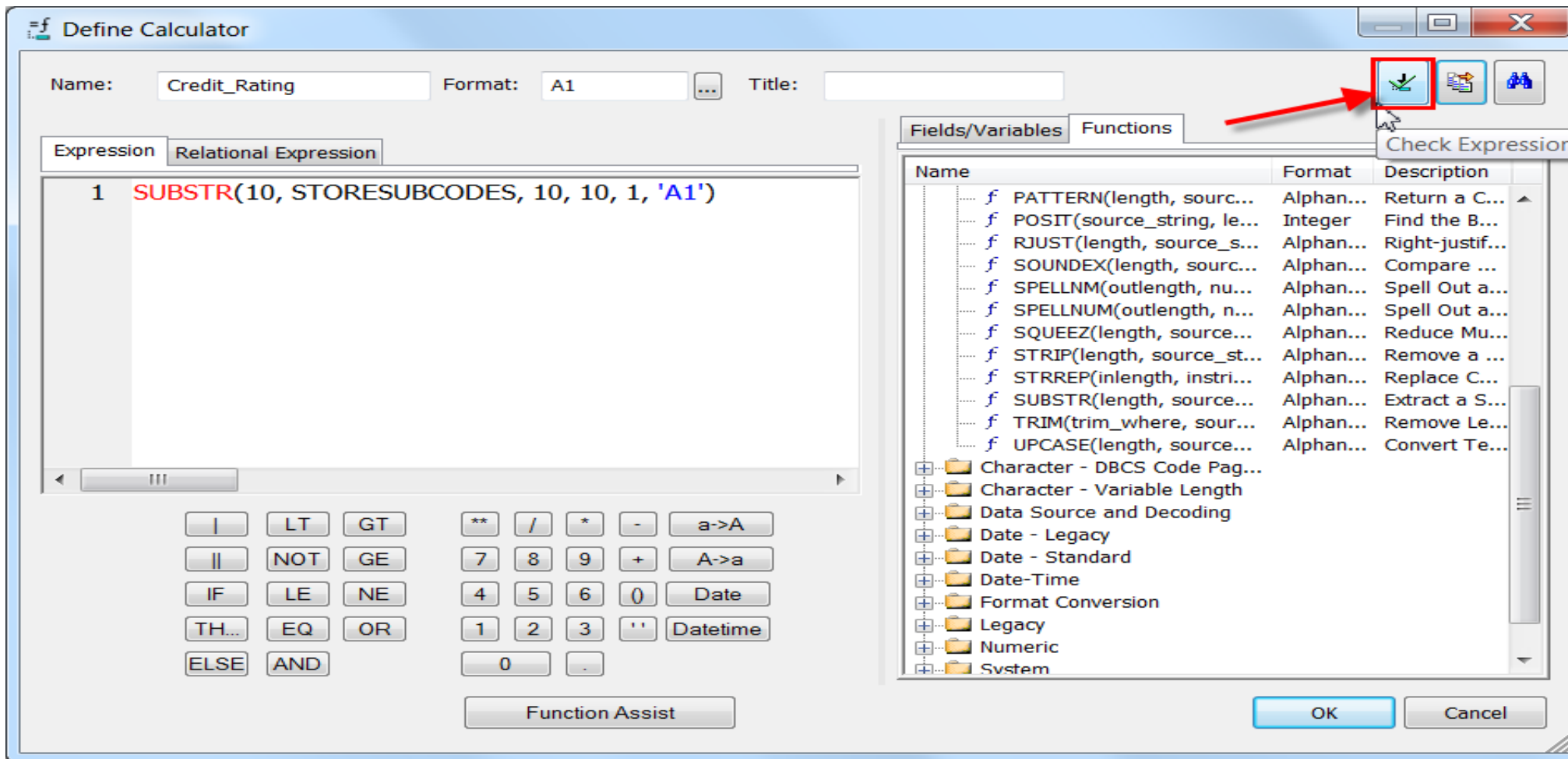
The screenshot shows the 'Define Calculator' dialog box. The 'Name' field is 'Credit\_Rating' and the 'Format' is 'A1'. The 'Expression' field contains the following code:

```
1 SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')
```

The 'Fields/Variables' and 'Functions' tabs are visible. The 'Functions' tab is active, showing a list of functions with columns for Name, Format, and Description. The list includes functions like PATTERN, POSIT, RJUST, SOUNDEX, SPELLNM, SPELLNUM, SQUEEZ, STRIP, STRREP, SUBSTR, TRIM, and UPCASE. There are also folders for Character, Data Source and Decoding, Date, Date-Time, Format Conversion, Legacy, Numeric, and System.

At the bottom of the dialog, there are buttons for 'Function Assist', 'OK', and 'Cancel'.

# Validate your Function Syntax Along the Way



The screenshot shows the 'Define Calculator' dialog box. The 'Name' field contains 'Credit\_Rating' and the 'Format' is set to 'A1'. The 'Expression' field contains the following code:

```
1 SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')
```

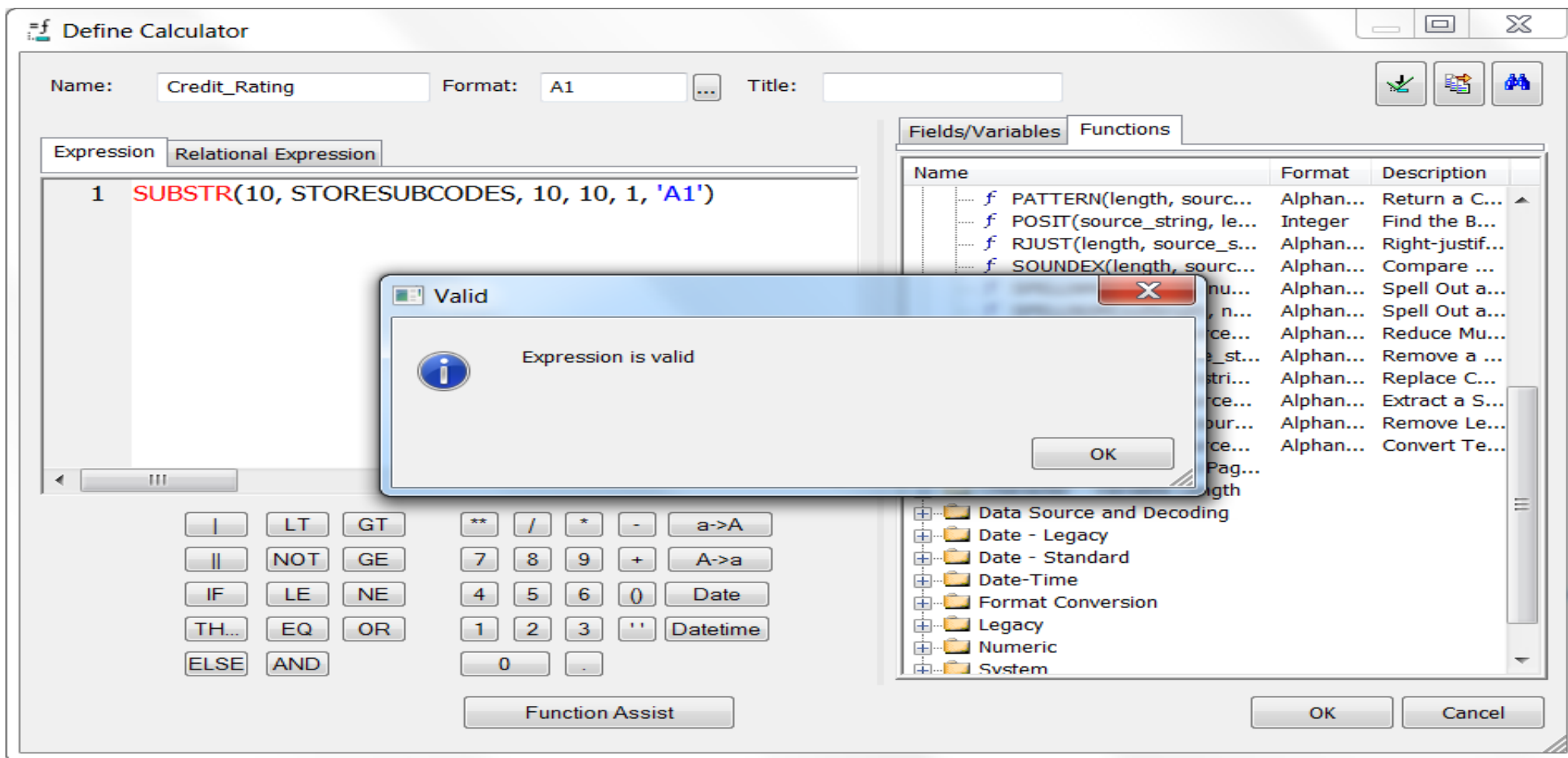
Below the expression field is a keypad with various operators and functions. A red arrow points to the 'Check Expression' button in the top right corner of the dialog box.

The 'Fields/Variables' tab is active, showing a list of functions with columns for Name, Format, and Description.

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
Character - DBCS Code Pag...		
Character - Variable Length		
Data Source and Decoding		
Date - Legacy		
Date - Standard		
Date-Time		
Format Conversion		
Legacy		
Numeric		
System		



# Validate your Function Syntax Along the Way

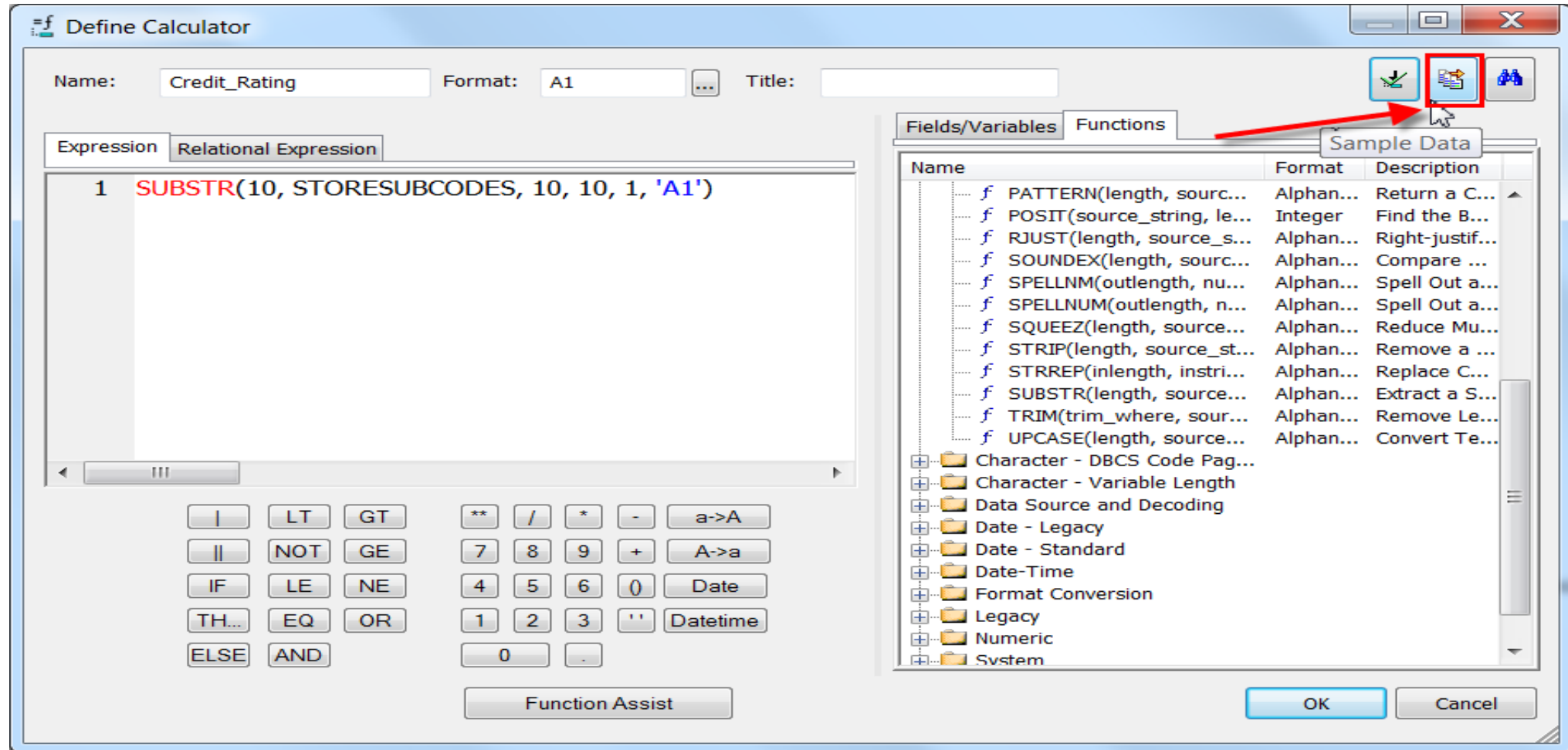


The screenshot shows the 'Define Calculator' dialog box with the following details:

- Name:** Credit\_Rating
- Format:** A1
- Expression:** SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')
- Valid Dialog:** A small 'Valid' dialog box is overlaid on the main window, displaying an information icon and the text 'Expression is valid' with an 'OK' button.
- Functions List:** A list of functions is visible on the right side of the dialog, including:
 

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...

# Sample Data to Validate



**Define Calculator**

Name:  Format:  Title:

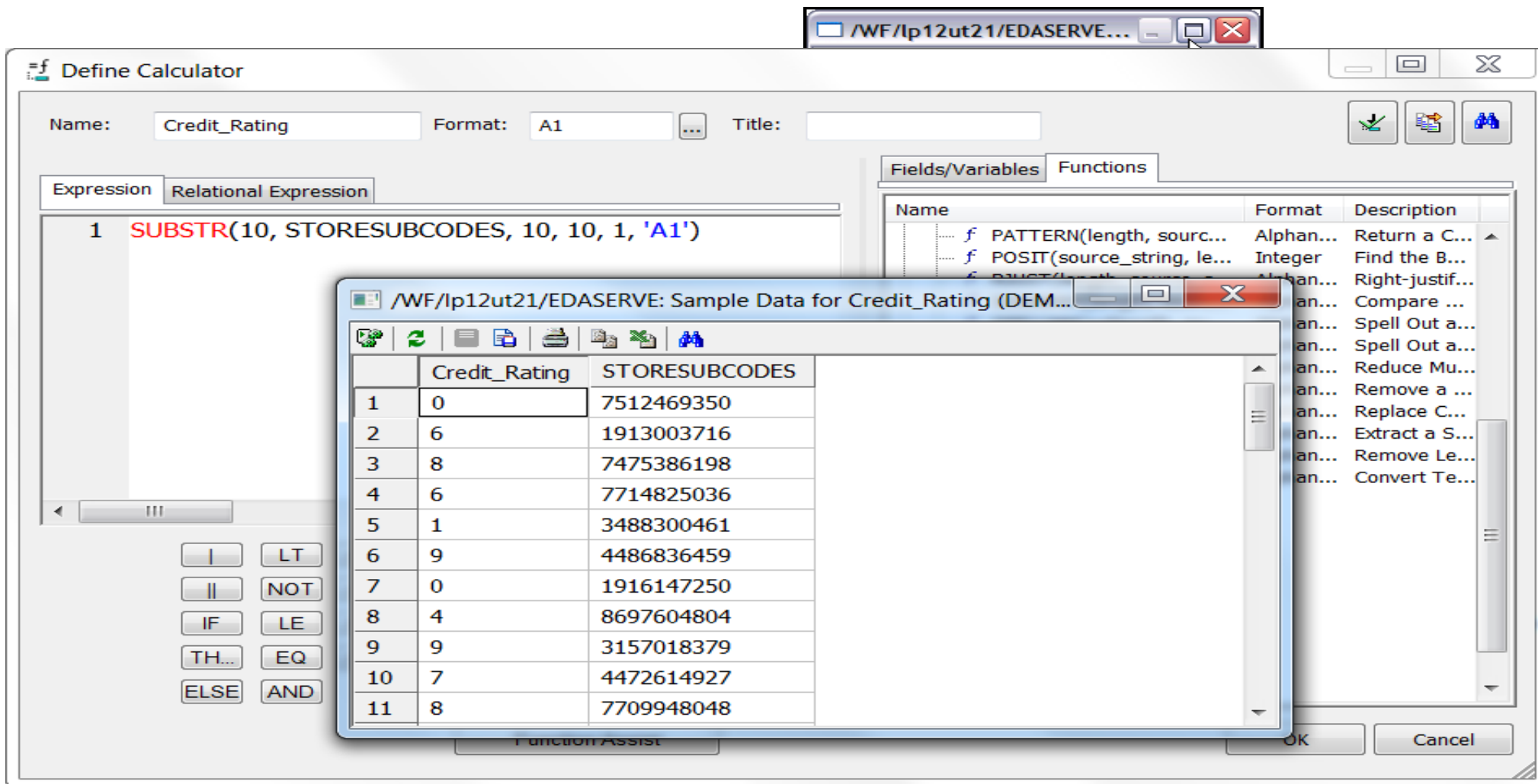
Expression: `1 SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')`

**Fields/Variables** | **Functions** | **Sample Data**

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
Character - DBCS Code Pag...		
Character - Variable Length		
Data Source and Decoding		
Date - Legacy		
Date - Standard		
Date-Time		
Format Conversion		
Legacy		
Numeric		
Svstem		

Function Assist | OK | Cancel

# Sample Data to Validate



The screenshot shows a 'Define Calculator' window with the following details:

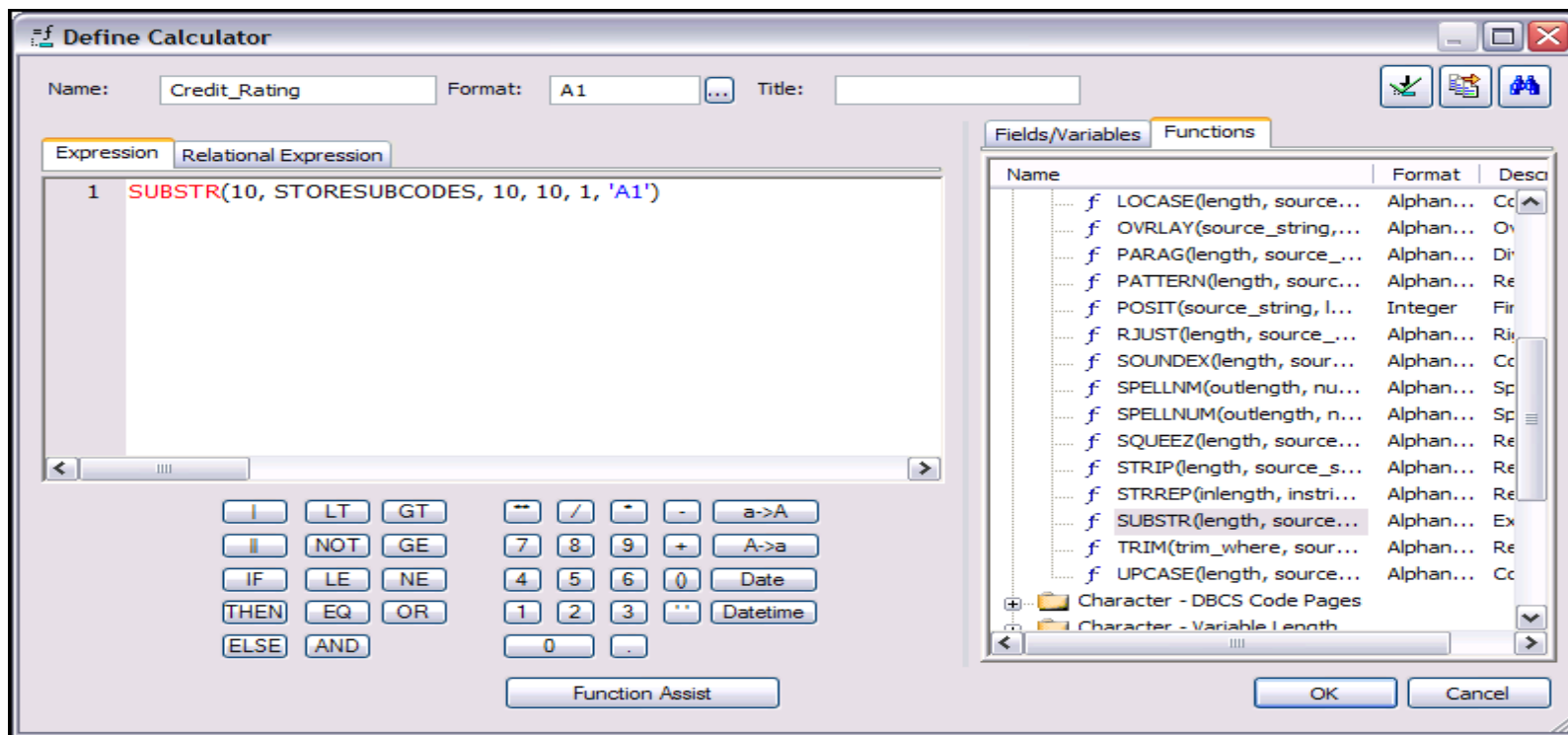
- Name:** Credit\_Rating
- Format:** A1
- Expression:** SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')

A data table window is overlaid on the calculator, displaying the following data:

	Credit_Rating	STORESUBCODES
1	0	7512469350
2	6	1913003716
3	8	7475386198
4	6	7714825036
5	1	3488300461
6	9	4486836459
7	0	1916147250
8	4	8697604804
9	9	3157018379
10	7	4472614927
11	8	7709948048

# Let's Keep Going

That's great, but let's add some transformation logic!



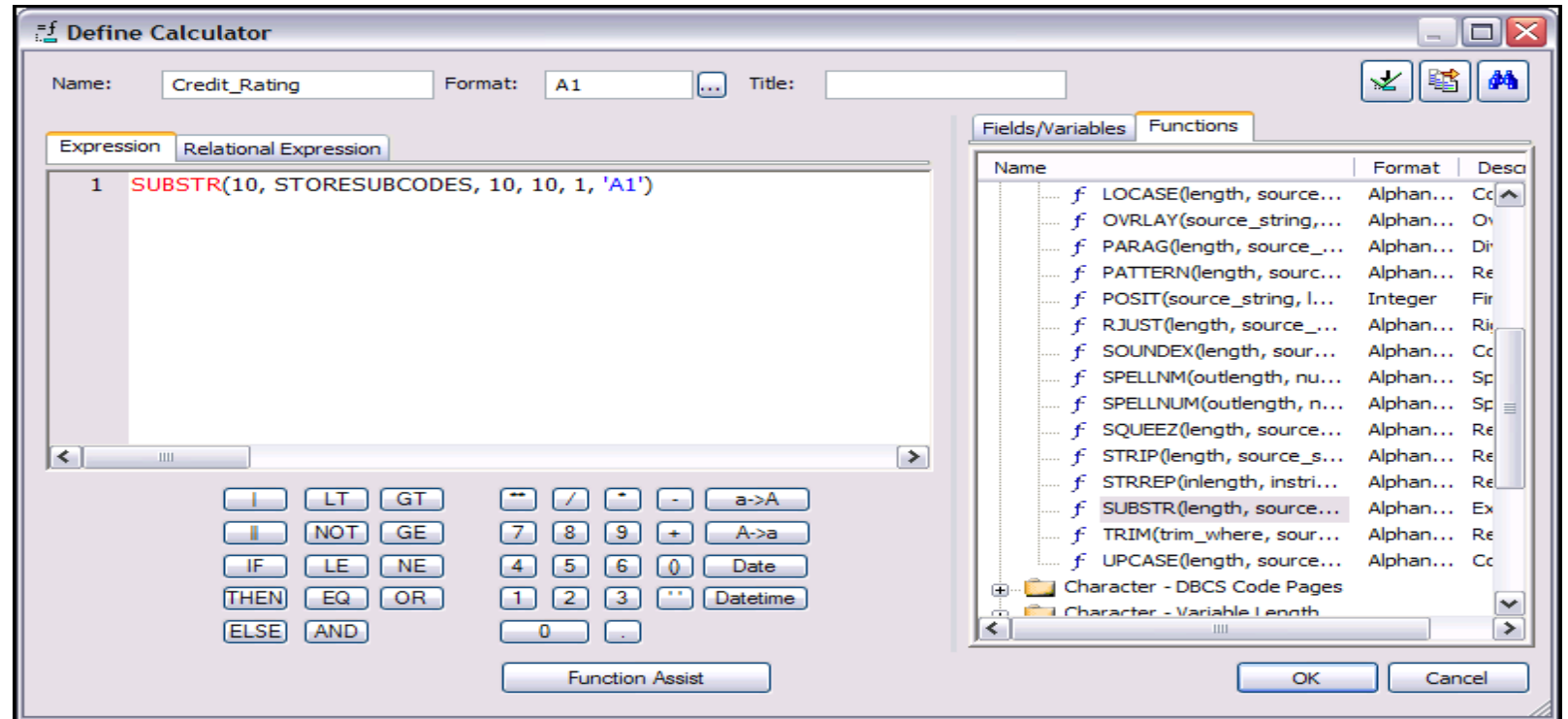
The 'Define Calculator' dialog box is shown with the following details:

- Name: Credit\_Rating
- Format: A1
- Title: (empty)
- Expression: SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')
- Functions list (selected):

Name	Format	Desc
LOCASE(length, source...	Alphan...	Cc
OVRLAY(source_string,...	Alphan...	Ov
PARAG(length, source_...	Alphan...	Di
PATTERN(length, sourc...	Alphan...	Re
POSIT(source_string, l...	Integer	Fi
RJUST(length, source_...	Alphan...	Ri
SOUNDEX(length, sour...	Alphan...	Cc
SPELLNM(outlength, nu...	Alphan...	Sp
SPELLNUM(outlength, n...	Alphan...	Sp
SQUEEZ(length, source...	Alphan...	Re
STRIP(length, source_s...	Alphan...	Re
STRREP(inlength, instri...	Alphan...	Re
SUBSTR(length, source...	Alphan...	Ex
TRIM(trim_where, sour...	Alphan...	Re
UPCASE(length, source...	Alphan...	Cc

# Apply a Business Rule

For credit rating >7 is  
"Good",  
4-6 is "Average",  
<4 is "Bad"

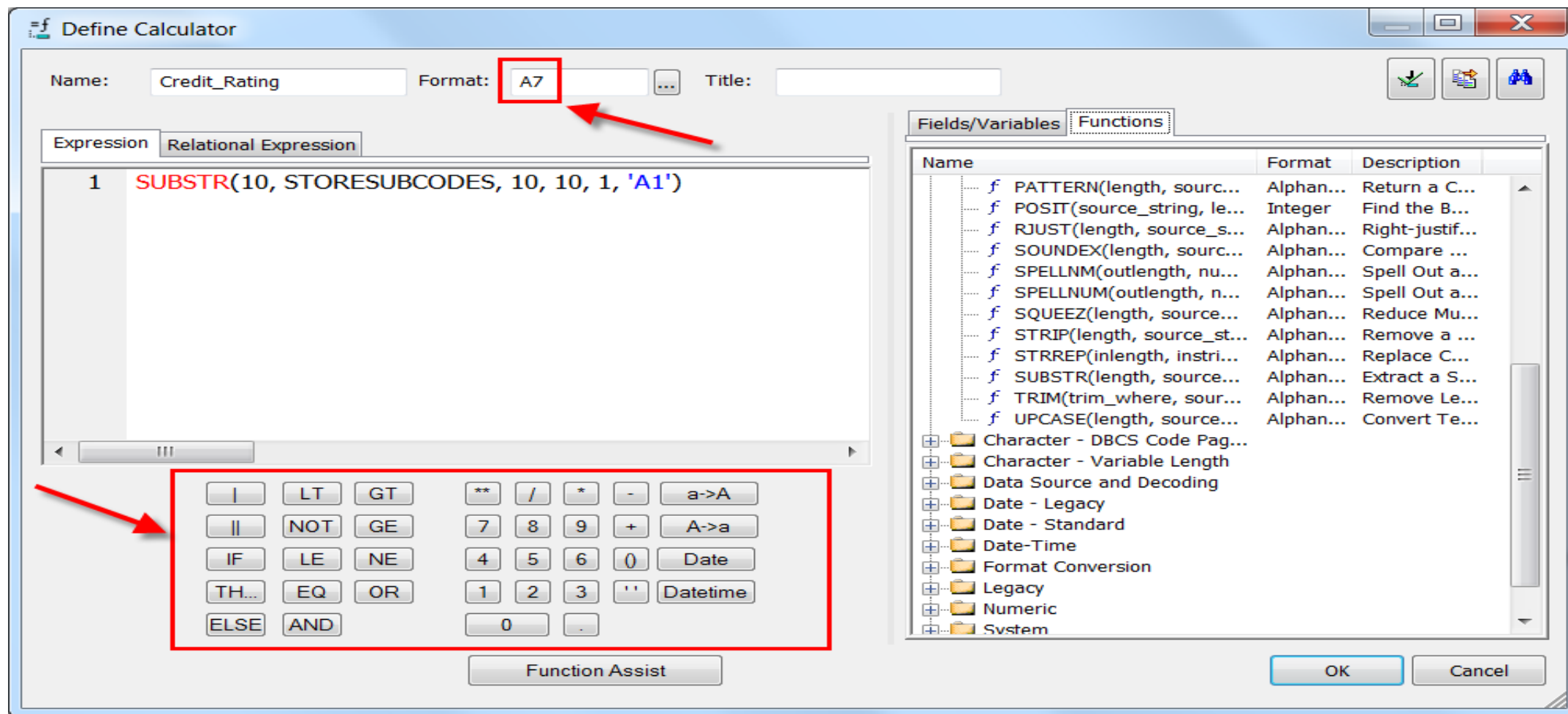


The image shows a 'Define Calculator' dialog box with the following details:

- Name: Credit\_Rating
- Format: A1
- Title: (empty)
- Expression: SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')
- Functions list (selected):

Name	Format	Desc
f LOCASE(length, source...	Alphan...	Cc
f OVLAY(source_string,...	Alphan...	Ov
f PARAG(length, source_...	Alphan...	Di
f PATTERN(length, sourc...	Alphan...	Re
f POSIT(source_string, l...	Integer	Fi
f RJUST(length, source_...	Alphan...	Ri
f SOUNDEX(length, sour...	Alphan...	Cc
f SPELLNM(outlength, nu...	Alphan...	Sp
f SPELLNUM(outlength, n...	Alphan...	Sp
f SQUEEZ(length, source...	Alphan...	Re
f STRIP(length, source_s...	Alphan...	Re
f STRREP(inlength, instri...	Alphan...	Re
f SUBSTR(length, source...	Alphan...	Ex
f TRIM(trim_where, sour...	Alphan...	Re
f UPCASE(length, source...	Alphan...	Cc

# Use Touchpad to Build Logic



**Define Calculator**

Name:  Format:  Title:

Expression: `1 SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')`

Function keypad (highlighted):

	LT	GT	**	/	*	-	a->A
	NOT	GE	7	8	9	+	A->a
IF	LE	NE	4	5	6	()	Date
TH...	EQ	OR	1	2	3	''	Datetime
ELSE	AND		0	.			

Function Assist

Fields/Variables | **Functions**

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
Character - DBCS Code Pag...		
Character - Variable Length		
Data Source and Decoding		
Date - Legacy		
Date - Standard		
Date-Time		
Format Conversion		
Legacy		
Numeric		
Svstem		

OK Cancel

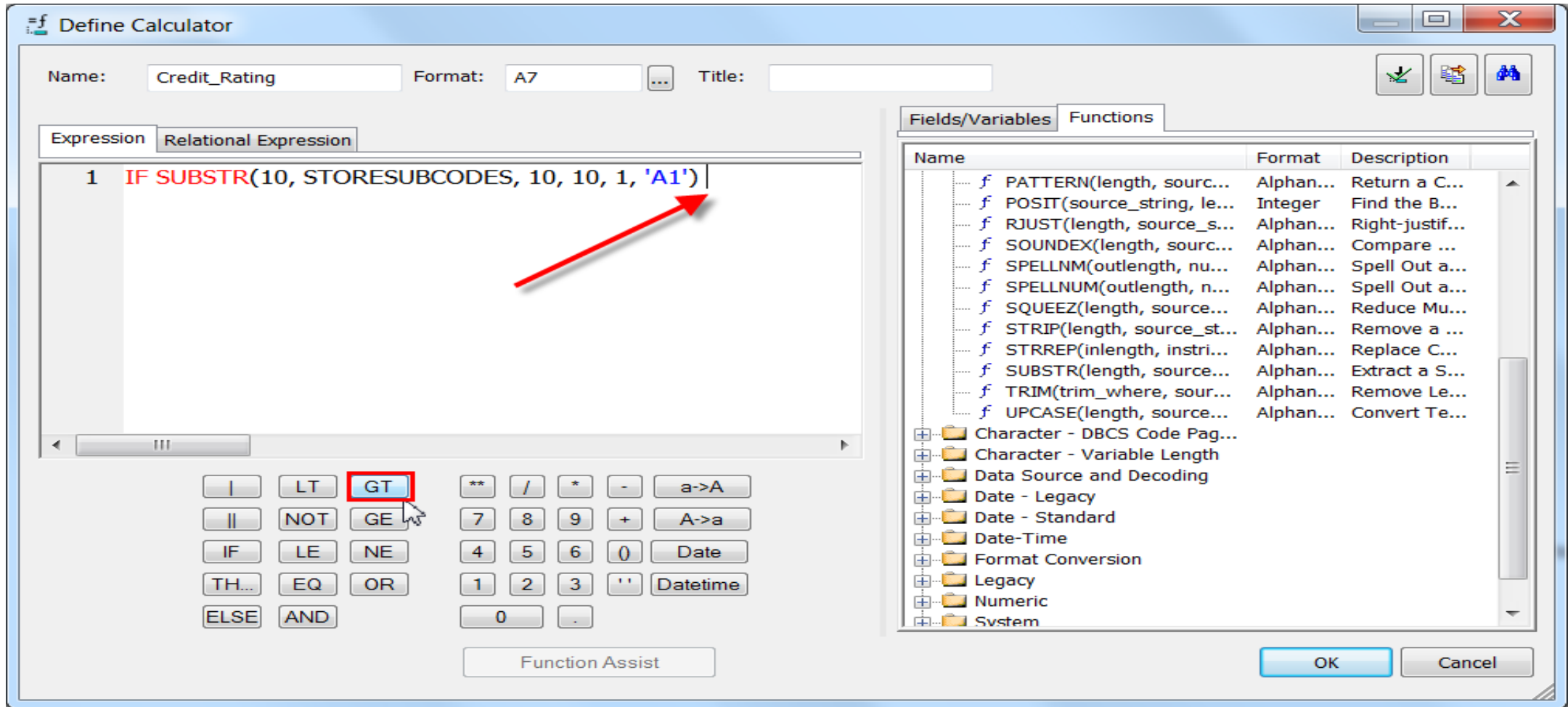
# Use Touchpad to Build Logic

The screenshot shows the 'Define Calculator' dialog box with the following details:

- Name: Credit\_Rating
- Format: A7
- Title: (empty)
- Expression: SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')
- Fields/Variables: (empty)
- Functions: (empty)
- Buttons: IF, LE, NE, EQ, OR, AND, LT, GT, GE, NOT, TH, ELSE, AND, \*\*, /, \*, -, a->A, 7, 8, 9, +, A->a, 4, 5, 6, 0, Date, 1, 2, 3, ', Datetime, 0, ., Function Assist, OK, Cancel

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...

# Use Touchpad to Build Logic



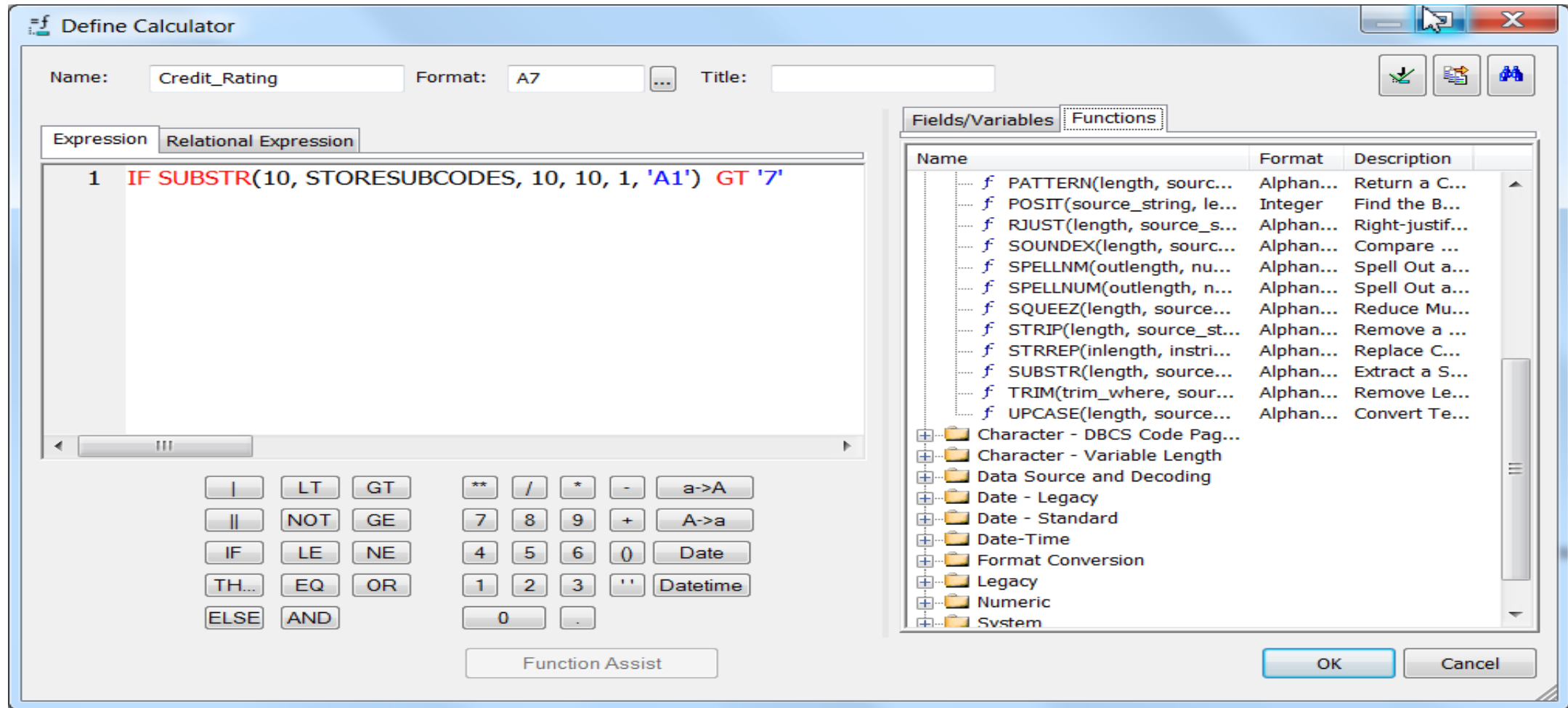
The screenshot shows the 'Define Calculator' dialog box with the following details:

- Name:** Credit\_Rating
- Format:** A7
- Title:** (empty)
- Expression:** 1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1')
- Functions List:**

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
- Buttons:** The 'GT' button is highlighted with a red box and a red arrow points to it from the expression field.



# Use Touchpad to Build Logic



**Define Calculator**

Name:  Format:  Title:

Expression Relational Expression

1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '7'

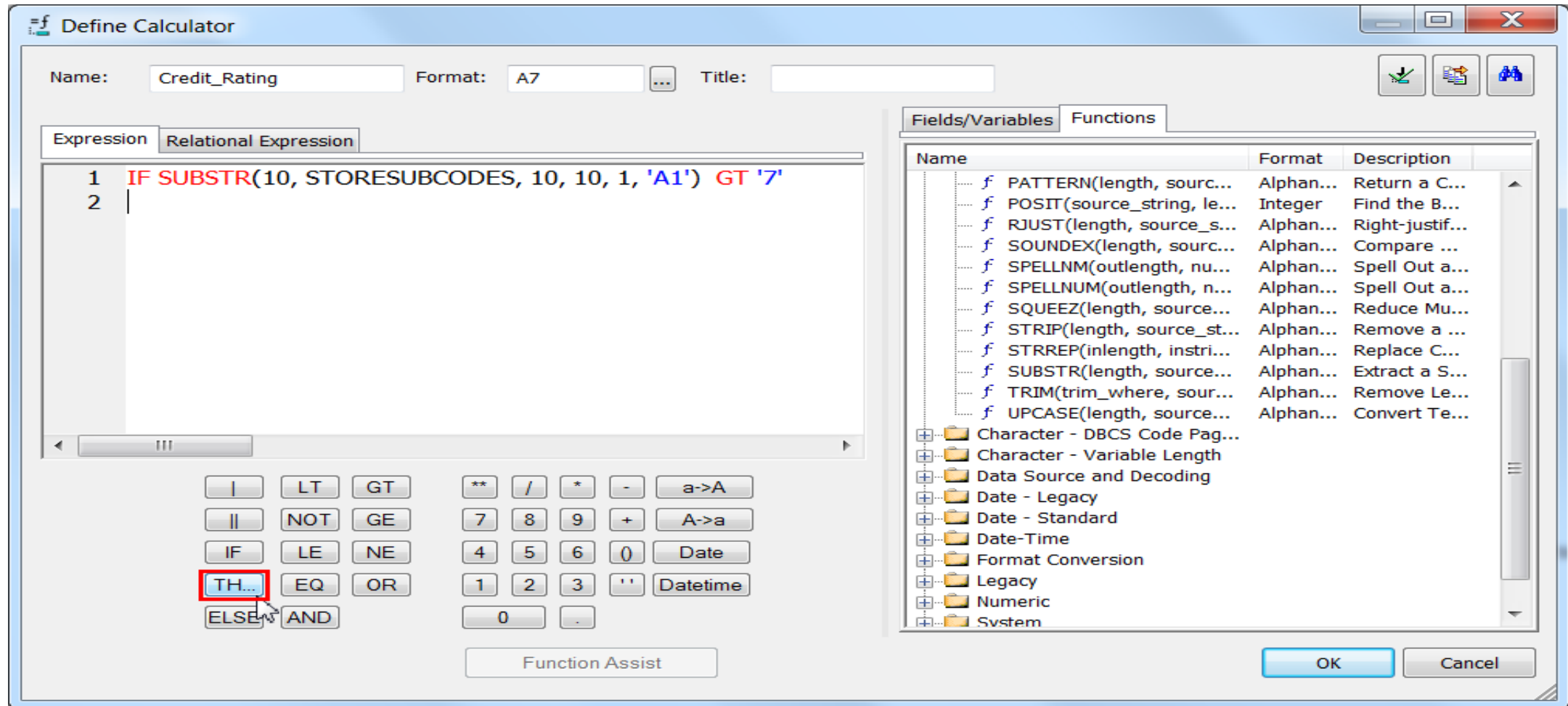
Fields/Variables Functions

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
Character - DBCS Code Pag...		
Character - Variable Length		
Data Source and Decoding		
Date - Legacy		
Date - Standard		
Date-Time		
Format Conversion		
Legacy		
Numeric		
Svstem		

Function Assist

OK Cancel

# Use Touchpad to Build Logic



**Define Calculator**

Name:  Format:  Title:

Expression | Relational Expression

```

1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '7'
2 |
    
```

Function List:

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
Character - DBCS Code Pag...		
Character - Variable Length		
Data Source and Decoding		
Date - Legacy		
Date - Standard		
Date-Time		
Format Conversion		
Legacy		
Numeric		
System		

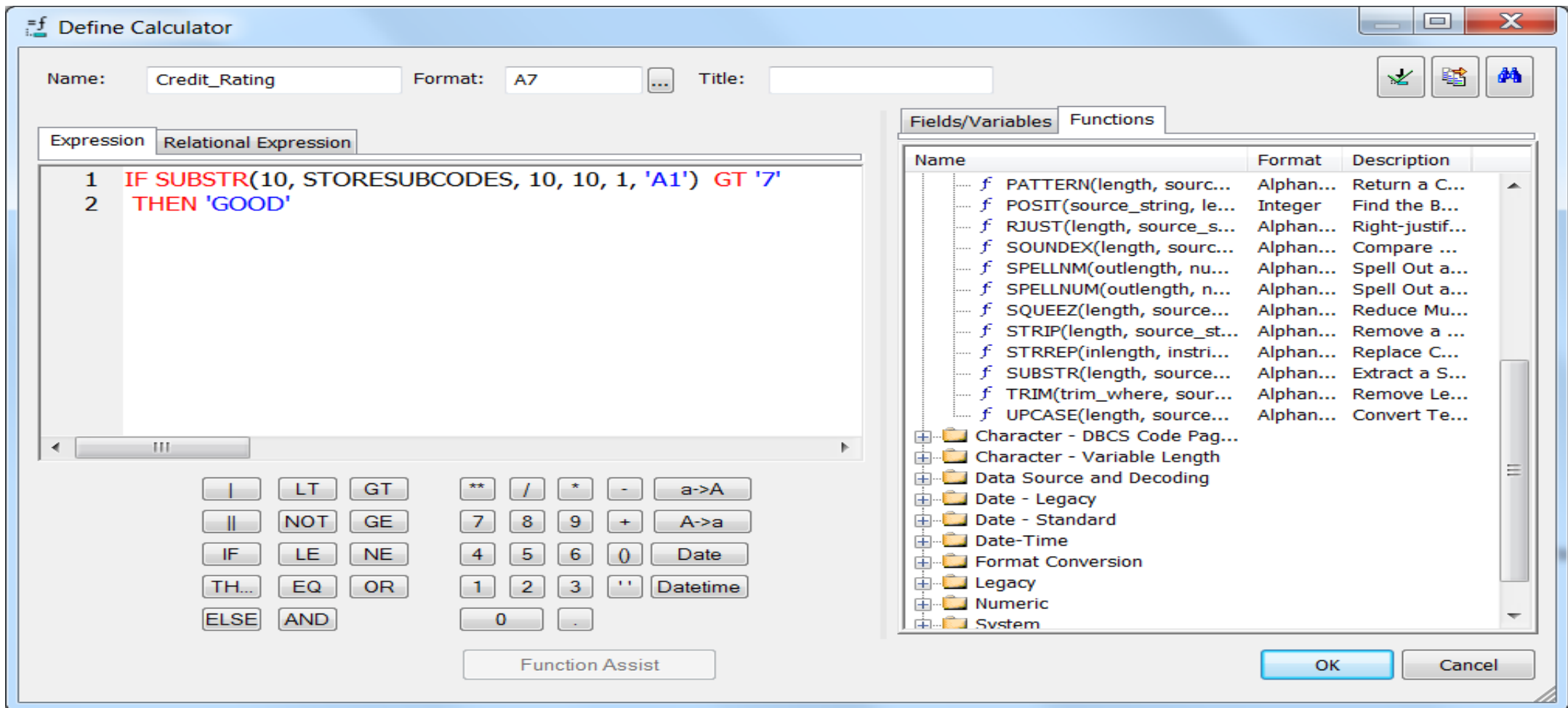
Logic Keypad:

- Buttons: |, LT, GT, \*\*, /, \*, -, a->A
- Buttons: ||, NOT, GE, 7, 8, 9, +, A->a
- Buttons: IF, LE, NE, 4, 5, 6, (), Date
- Buttons: TH..., EQ, OR, 1, 2, 3, "", Datetime
- Buttons: ELSE, AND, 0, .

Function Assist

OK Cancel

# Use Touchpad to Build Logic



The screenshot shows the 'Define Calculator' dialog box with the following details:

- Name:** Credit\_Rating
- Format:** A7
- Title:** (empty)
- Expression:**

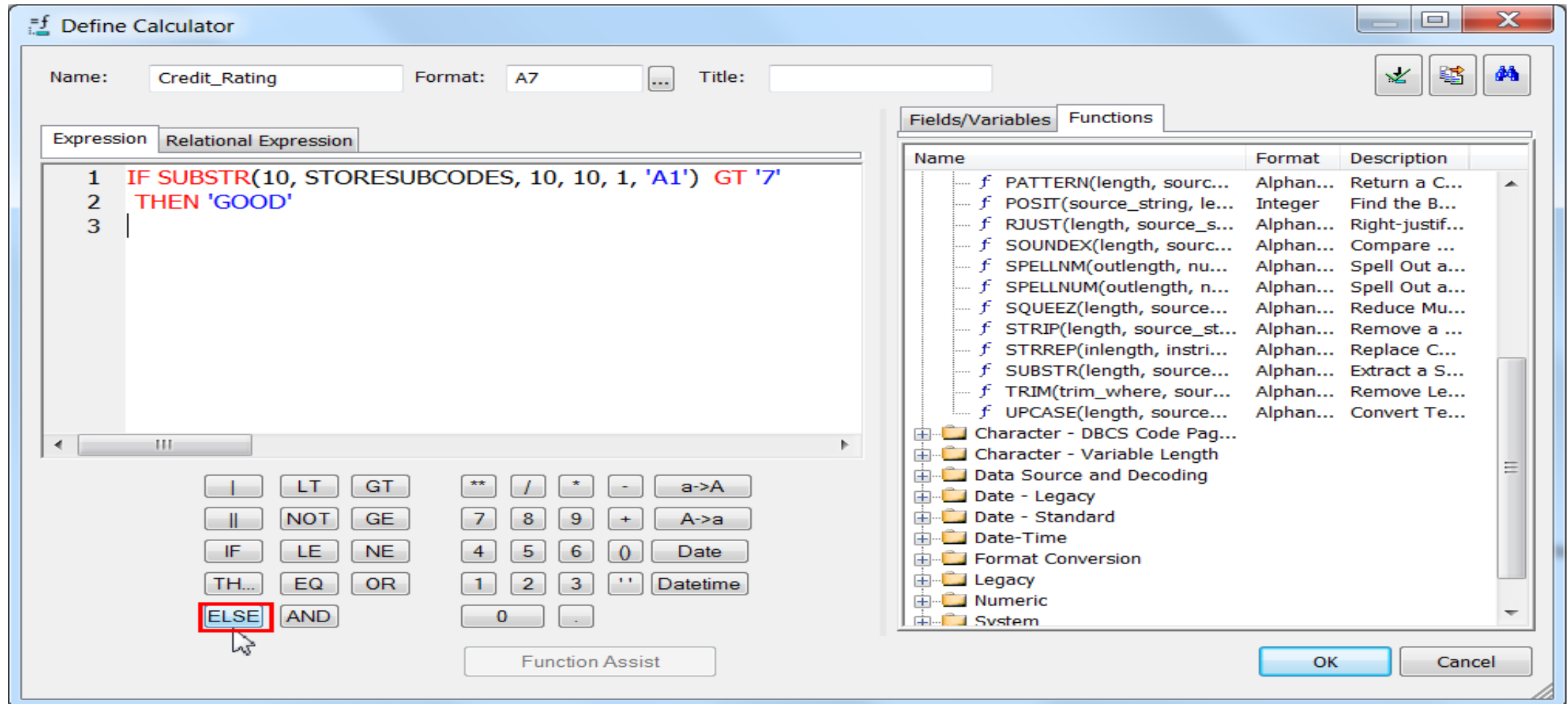
```

1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '7'
2 THEN 'GOOD'

```
- Fields/Variables:** (empty)
- Functions:**

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
Character - DBCS Code Pag...		
Character - Variable Length		
Data Source and Decoding		
Date - Legacy		
Date - Standard		
Date-Time		
Format Conversion		
Legacy		
Numeric		
Svstem		
- Buttons:** Function Assist, OK, Cancel

# Use Touchpad to Build Logic



**Define Calculator**

Name:  Format:  Title:

Expression Relational Expression

```
1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '7'  
2 THEN 'GOOD'  
3
```

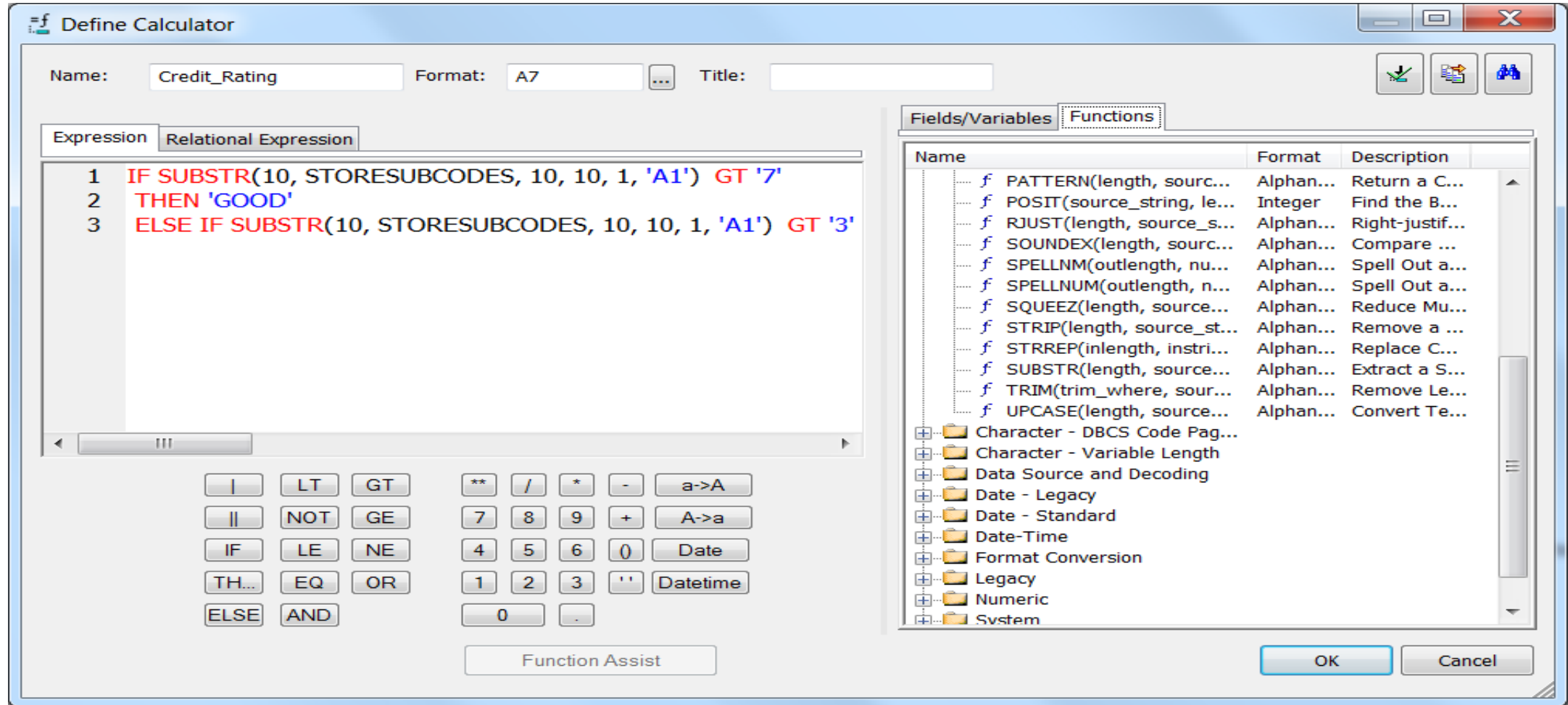
**Fields/Variables** **Functions**

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
Character - DBCS Code Pag...		
Character - Variable Length		
Data Source and Decoding		
Date - Legacy		
Date - Standard		
Date-Time		
Format Conversion		
Legacy		
Numeric		
System		

Function Assist

OK Cancel

# Use Touchpad to Build Logic



**Define Calculator**

Name:  Format:  Title:

Expression Relational Expression

```

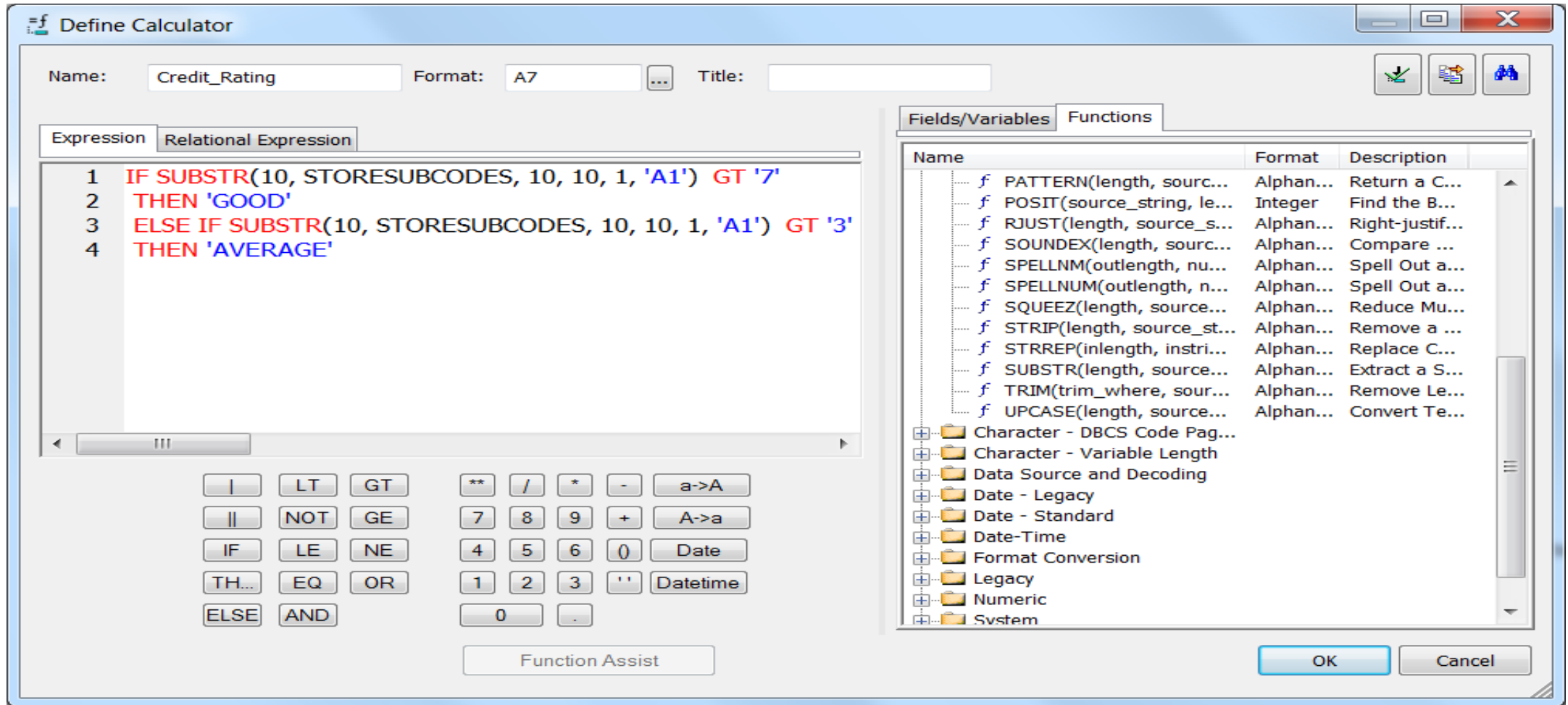
1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '7'
2 THEN 'GOOD'
3 ELSE IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '3'
  
```

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...

Function Assist

OK Cancel

# Use Touchpad to Build Logic



The screenshot shows the 'Define Calculator' dialog box with the following details:

- Name:** Credit\_Rating
- Format:** A7
- Title:** (empty)
- Expression:**

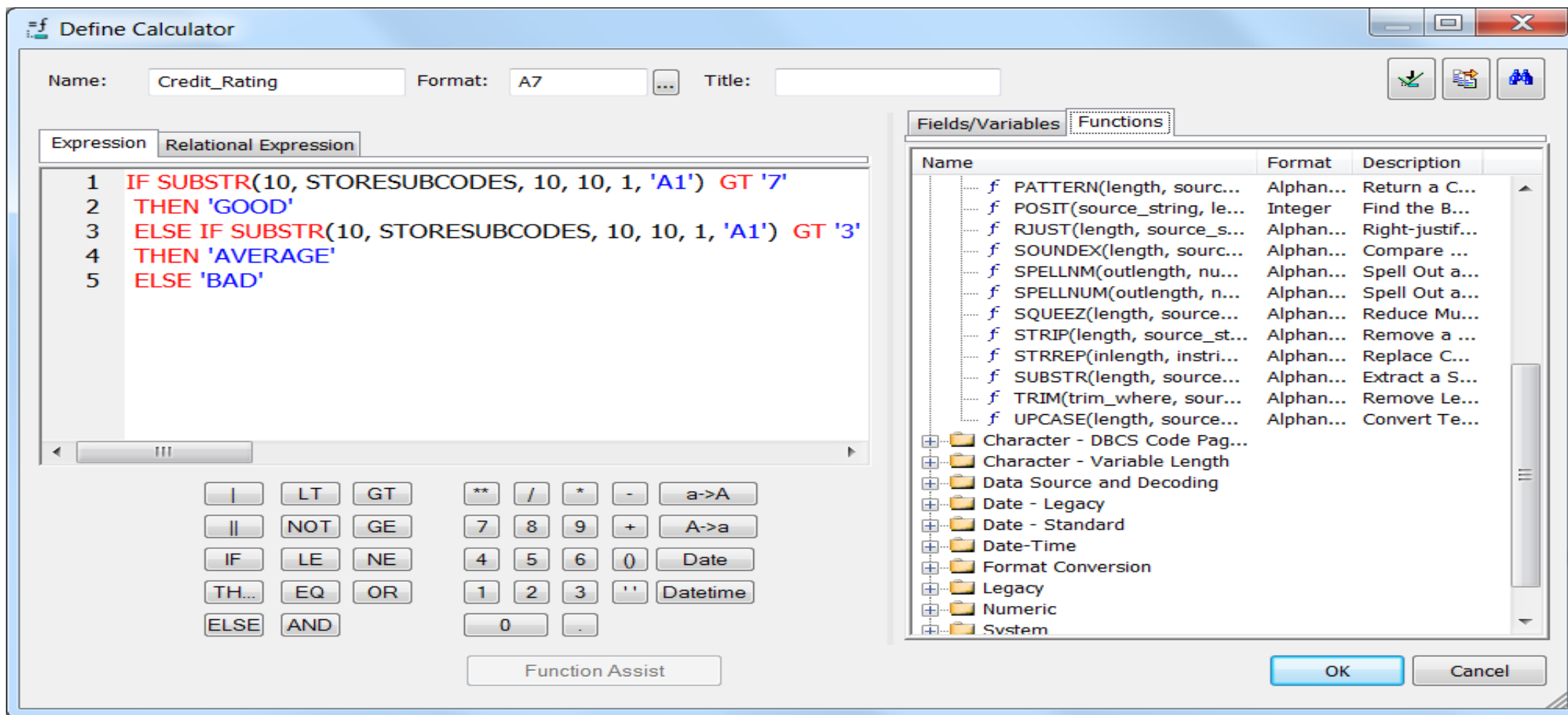
```

1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '7'
2 THEN 'GOOD'
3 ELSE IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '3'
4 THEN 'AVERAGE'

```
- Fields/Variables:** (empty)
- Functions:**

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
- Keyboard:**
  - Row 1: |, LT, GT, \*\*, /, \*, -, a->A
  - Row 2: ||, NOT, GE, 7, 8, 9, +, A->a
  - Row 3: IF, LE, NE, 4, 5, 6, (), Date
  - Row 4: TH..., EQ, OR, 1, 2, 3, '', Datetime
  - Row 5: ELSE, AND, 0, .
- Buttons:** Function Assist, OK, Cancel

# Use Touchpad to Build Logic



The screenshot shows the 'Define Calculator' dialog box with the following details:

- Name:** Credit\_Rating
- Format:** A7
- Title:** (empty)
- Expression:**

```

1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '7'
2 THEN 'GOOD'
3 ELSE IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '3'
4 THEN 'AVERAGE'
5 ELSE 'BAD'

```
- Fields/Variables:** (empty)
- Functions:**

Name	Format	Description
f PATTERN(length, sourc...	Alphan...	Return a C...
f POSIT(source_string, le...	Integer	Find the B...
f RJUST(length, source_s...	Alphan...	Right-justif...
f SOUNDEX(length, sourc...	Alphan...	Compare ...
f SPELLNM(outlength, nu...	Alphan...	Spell Out a...
f SPELLNUM(outlength, n...	Alphan...	Spell Out a...
f SQUEEZ(length, source...	Alphan...	Reduce Mu...
f STRIP(length, source_st...	Alphan...	Remove a ...
f STRREP(inlength, instri...	Alphan...	Replace C...
f SUBSTR(length, source...	Alphan...	Extract a S...
f TRIM(trim_where, sour...	Alphan...	Remove Le...
f UPCASE(length, source...	Alphan...	Convert Te...
- Calculator Interface:** Includes a keypad with logical operators (|, LT, GT, NOT, GE, LE, NE, EQ, OR, AND), arithmetic operators (\*\*, /, \*, -, +, 0, .), and conversion functions (a->A, A->a, Date, Datetime). A 'Function Assist' button is also present.
- Buttons:** OK and Cancel buttons are located at the bottom right.

# Sample Data

**Define Calculator**

Name:  Format:  Title:

Expression Relational Expression

```

1 IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') GT '7'
2 THEN 'GOOD'
3 ELSE IF SUBSTR(10, STORESUBCODES, 10, 10, 1, 'A1') G
4 THEN 'AVERAGE'
5 ELSE 'BAD'
    
```

Function Assist

	Credit_Rating	STORESUBCODES
1	BAD	7512469350
2	AVERAGE	1913003716
3	GOOD	7475386198
4	AVERAGE	7714825036
5	BAD	3488300461
6	GOOD	4486836459
7	BAD	1916147250
8	AVERAGE	8697604804
9	GOOD	3157018379
10	AVERAGE	4472614927
11	GOOD	7709948048
12	BAD	3359468361
13	GOOD	5804993948
14	AVERAGE	3379383827
15	AVERAGE	1150593726
16	BAD	2836049271
17	AVERAGE	7577415605
18	AVERAGE	2614825015
19	AVERAGE	6591592014
20	AVERAGE	7993792105
21	AVERAGE	7683604814
22	GOOD	3159388148
23	AVERAGE	0713735837
24	BAD	4363389371
25	BAD	0359271593

at Description

- n... Return a C...
- er Find the B...
- n... Right-justif...
- n... Compare ...
- n... Spell Out a...
- n... Spell Out a...
- n... Reduce Mu...
- n... Remove a ...
- n... Replace C...
- n... Extract a S...
- n... Remove Le...
- n... Convert Te...

OK Cancel





# Dealing with Dates

One of the most common reporting issues is dealing with dates

- Dates can be stored in various formats in the underlying database
- Dates should be displayed in various formats
- Dates are often rolled up to create buckets based on week, month, quarter, year, etc.
- Dates are often used for prompts in reports (Start\_date, End\_date)
- Date calculations based on today's date are often required (today – 30)
- Period to date calculations and comparisons to that same time frame in a previous period (QTD compared to QTD Last Year)

Meta Data is a great way to re-use date manipulation definitions

- Option 1: Create true date fields out of date fields being used in non date data types
  - Can then apply date functions to the virtual true date field to get date attributes such as “quarter” or “day of the week”, etc.
- Option 2: Use a Date Dimension Table that contains date attributes
  - Probably a better approach as it moves date processing to DB2 which improves performance



## Advanced Date Processing with Date Dimension Table

- Sample Date Dimension table NOW included in sample database (QWQCENT) that comes with DB2 Web Query
- Table can be joined with others within synonym to provide advanced date processing
  - What are the profit margins on days before and after holidays?
  - How many bags of corn chips are sold the week before the Super Bowl?
  - What is the rate of product returns on the day after Christmas as compared to any other day of the year?
  - Are more galoshes sold in the spring or the fall?
  - How many boxes of diapers are sold on days when there is a full moon?

# Date Dimension Table

IBM DB2® for i

ORDDTL legacy file



ORDER	CUST	ORDDAT	SHPDAT	SHPVIA	ORDSTS	ORDAMT	TOTLIN
1715810	H4541	2152005	6302005	UPS	1	933346.39	2
1563685	R1948	2202005	8042005	Mule Train	2	0.00	0
7195900	Q7881	2232005	5022005	Pick Up	2	0.00	0
8854635	S1511	2232005	9102005	Train	1	118086.88	2
6694902	X8863	2242005	7192005	Camel	2	0.00	0
8054679	F4327	2272005	9112005	Pnuematic Tube	2	0.00	0
527879	C3233	2282005	5022005	US Mail	1	897524.71	3
4011038	G1496	2282005	10162005	Train	1	865114.60	2
5417918	J3825	2282005	10022005	FedEx	1	84608.50	2
9456994	N2796	2282005	7012005	Train	2	0.00	0
3526155	O7881	3012005	8062005	UPS	1	194660.67	2

Date dimension table

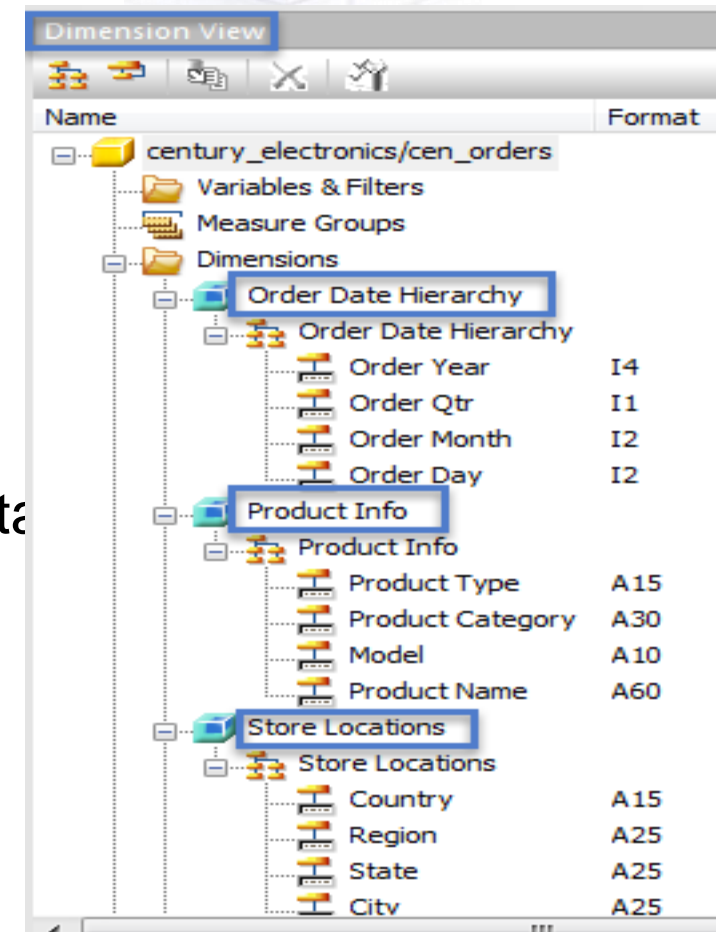


Join (RI constraints, SQL view, synonym join, report join)

Date	MDYY	Year	Month	Day	Month Name	Day of Week	Day Of Year	Day Name	Weekend Flag	Season	Fiscal Year	Fiscal Qtr	Week Starting Date	Same Day Last Year
2005/02/15	2152005	2005	2	15	February	2	46	Tuesday	N	Winter	2005	2	2005/02/12	2004/02/17
2005/02/16	2162005	2005	2	16	February	3	47	Wednesday	N	Winter	2005	2	2005/02/12	2004/02/18
2005/02/17	2172005	2005	2	17	February	4	48	Thursday	N	Winter	2005	2	2005/02/12	2004/02/19
2005/02/18	2182005	2005	2	18	February	5	49	Friday	N	Winter	2005	2	2005/02/12	2004/02/20
2005/02/19	2192005	2005	2	19	February	6	50	Saturday	Y	Winter	2005	2	2005/02/19	2004/02/21
2005/02/20	2202005	2005	2	20	February	7	51	Sunday	Y	Winter	2005	2	2005/02/19	2004/02/22
2005/02/21	2212005	2005	2	21	February	1	52	Monday	N	Winter	2005	2	2005/02/19	2004/02/23
2005/02/22	2222005	2005	2	22	February	2	53	Tuesday	N	Winter	2005	2	2005/02/19	2004/02/24
2005/02/23	2232005	2005	2	23	February	3	54	Wednesday	N	Winter	2005	2	2005/02/19	2004/02/25
2005/02/24	2242005	2005	2	24	February	4	55	Thursday	N	Winter	2005	2	2005/02/19	2004/02/26
2005/02/25	2252005	2005	2	25	February	5	56	Friday	N	Winter	2005	2	2005/02/19	2004/02/27
2005/02/26	2262005	2005	2	26	February	6	57	Saturday	Y	Winter	2005	2	2005/02/26	2004/02/28
2005/02/27	2272005	2005	2	27	February	7	58	Sunday	Y	Winter	2005	2	2005/02/26	2004/02/29
2005/02/28	2282005	2005	2	28	February	1	59	Monday	N	Winter	2005	2	2005/02/26	2004/03/01

## Multi Dimensional Analysis

- Business Analysts want to slice and dice data across dimensions
  - Product
  - Geography
  - Time
  - Sales Channel
- Each Dimension contains a hierarchy
  - Product -> Product Category -> Product Type -> Class
- DB2 Web Query doesn't care where the data resides – the metadata that defines the relationships will point to the source
  - Enables Auto Drill Down Reports
  - Enables chained (smart) parameters
  - Data can be production database or isolated datamart





# Agenda

- Meta Data as a Foundation

- ETL (Extract, Transformation, and Load) is often the Best Practice to support Analytics Application



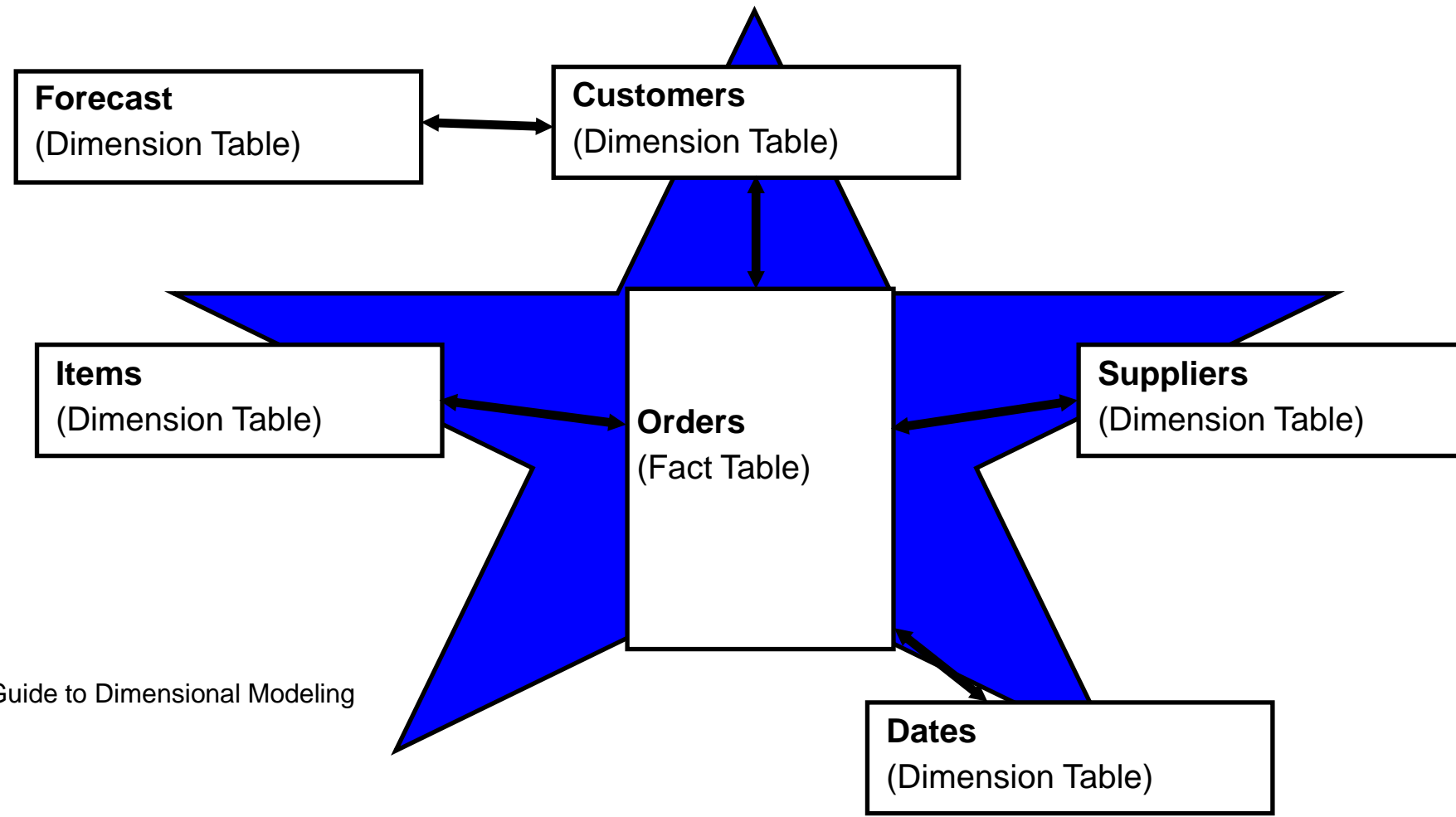
## Performance

- Many BI queries are complex and **summary in nature**
  - “Show me revenue vs. forecast across these 3 product lines, these 2 regions, and the trend over the last 3 quarters.”
- These kinds of queries against databases designed for transaction processing on systems tuned for transaction processing....can be problematic
  - Data not in multidimensional formats
  - Data in detail level, not summary
  - Data transformation work is being done at report run time
  - Data consolidation done at run time
  - Data is point in time, not historical (usually)



## Multi Dimensional Analysis with Star Schema

- Many people choose to isolate the data being queried into its own LPAR or Server and.....
- Re shape that data into a multi-dimensional model such as a star schema or snowflake design



The Data Warehouse Toolkit – The Definitive Guide to Dimensional Modeling  
Ralph Kimball, Margy Ross

# Difficult Data Issues

Customer File - US	
CUSTNO	CUSTNAME
1001	John Smith
1002	Mary Jones
1003	Chris Anderson
1004	David Perry

Customer File - Canada	
CUSTNO	CUSTNAME
1001	Harry Potter
1002	Jeremy Carr
1003	Penny Hayes
1004	Debbie Thornton



Surrogate key is a sequential number with no correlation to replaced value(s)

Customer File - Data Warehouse			
CUSTNUMBER	CUSTNAME	REGION	OLDNUM
1	John Smith	US	1001
2	Mary Jones	US	1002
3	Chris Anderson	US	1003
4	David Perry	US	1004
5	Harry Potter	CANADA	1001
6	Jeremy Carr	CANADA	1002
7	Penny Hayes	CANADA	1003
8	Debbie Thornton	CANADA	1004

PK

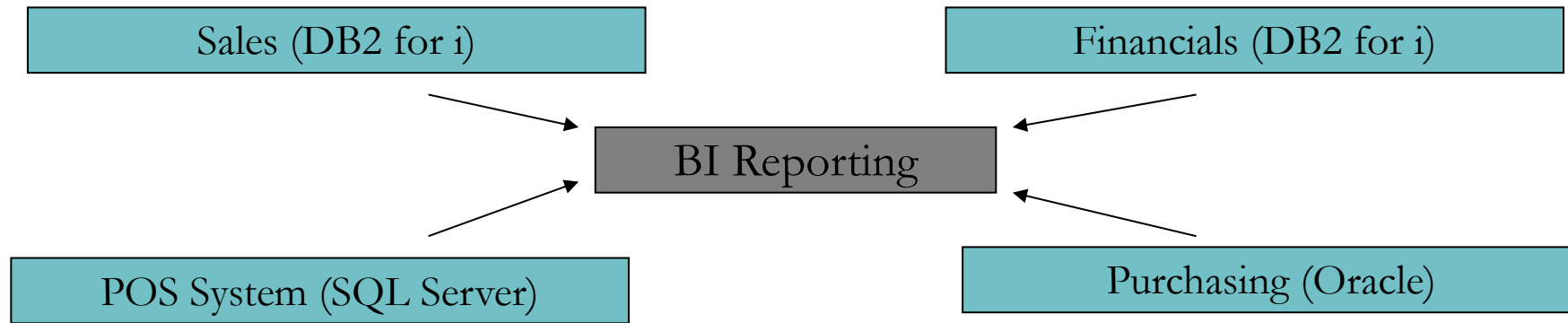
Secondary Index





# Difficult Data Issues

- Multiple application databases
  - May even be different types (DB2, Oracle, SQL Server)
  - Data stored in totally different structures – but related information



- Very difficult, if not impossible to join tables across databases
- Different security, availability etc
- Adds significant complexity



# Difficult Data Issues

- Keeping History

- Customer attributes (group, territory) are periodically changed

2010	100	Smith & Jones Electrical	Small Retailer	Jenny Brown
2013	100	Smith & Jones Electrical	Major Retailer	Jenny Brown
2017	100	Smith & Jones Electrical	Major Retailer	Rob McAdam

What happens if :

- I need to run a report to get status of account and who covered them in 2012?

You cannot get the same report that was run in 2012! The original groupings (customer group, sales rep etc) are no longer available in the operational database.

- I need to compare ThisYear sales against LastYear for all sales reps?

Same problem – You’ve lost visibility to last year’s coverage.

## Answer: The Data Warehouse

- Automate Data Consolidation
  - Multiple Applications or Databases; Bring in Spreadsheet data(?)
- Move Data Transformations to the ETL process
  - Automate this process to run in the background
  - Validate data as it enters the warehouse
- Transform the data into an analytical data model
  - E.g., Star Schema or Snowflake
- Automate Capturing and Maintaining Historical Data
  - Slowly changing dimensions
- Optimize Environment for Reporting
  - E.g., Create Summaries





## Transformation Examples

- Pull from files with character or numeric based 'dates' and transform them into true Date/Time/Timestamp fields
- Turn conditional based processing into database centric e.g. convert `when FLAGFIELD = 'A' then use field1 else use field2` into a single column in the target table
- Pull a subset of fields from a source table and put into a target table
- Pull data from a legacy data model and convert into a star schema
- Aggregate data from a legacy database and put into a resulting table
  - E.g. pull data from legacy database that has all individual entries in customer orders, aggregate the multiple records of a given customer order into one row (with sums and counts) and put into the target table



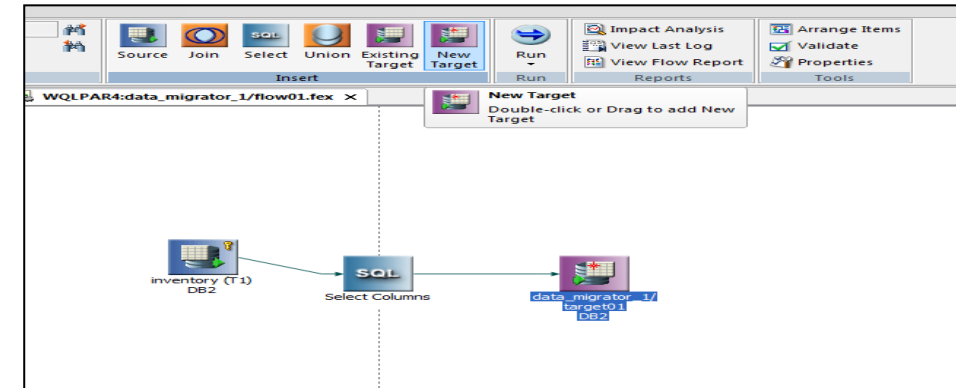
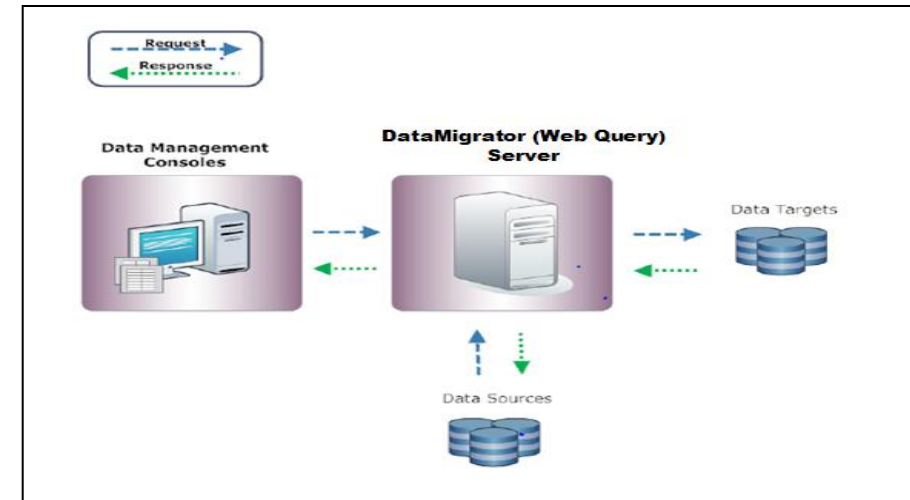
# Load Types

- Multi Table Load
  - Extracts data using a single pass of the data and loads multiple targets
- Slowly Changing Dimensions
  - Choose how you want to handle history
  - Insert Begin/End Timestamp and activation flag columns (similar to DB2 temporal)
- Consolidate related data and generate surrogate key
  - Similar to identity column in DB2
- Clear - then Load
- Change-Data-Capture
  - Log (journal receiver) based extraction
  - Can be from remote journals\*



# DataMigrator ETL (Extract, Transformation, Load)

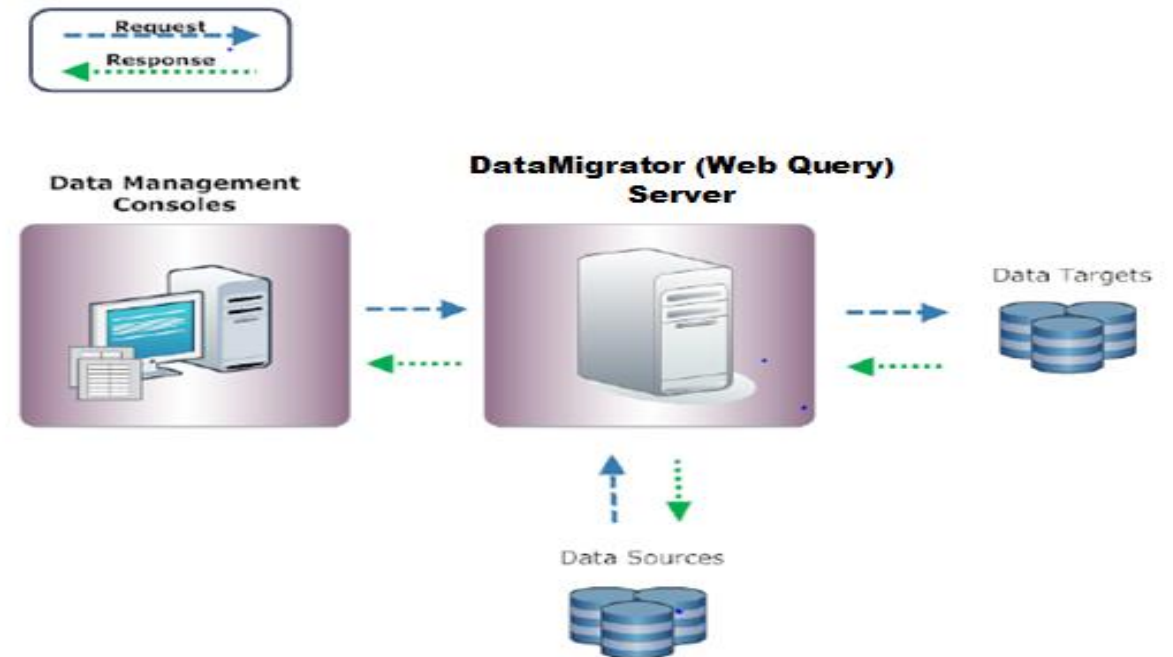
- New Product ID 5733-WQM
  - Requires either DB2 Web Query Express or Standard Edition V2.1 with Group PTF Level 11
    - Core based. Must have same number of cores as DB2 Web Query core licenses
  - Standard Edition required for non DB2 relational databases as a data source
  - One license of DB2 Web Query Developer Workbench is required to set up data and process flows
  
- ETL for IBM i
  - Like DB2 Web Query, all server code and ETL objects run in and/or are stored in IBM i
  
- INTEGRATED with DB2 Web Query
  - Can leverage meta data created with DB2 Web Query or vice versa
  - Same look and feel
  - Shared services and administration



# DataMigrator Architecture

## DataMigrator for i has two main components:

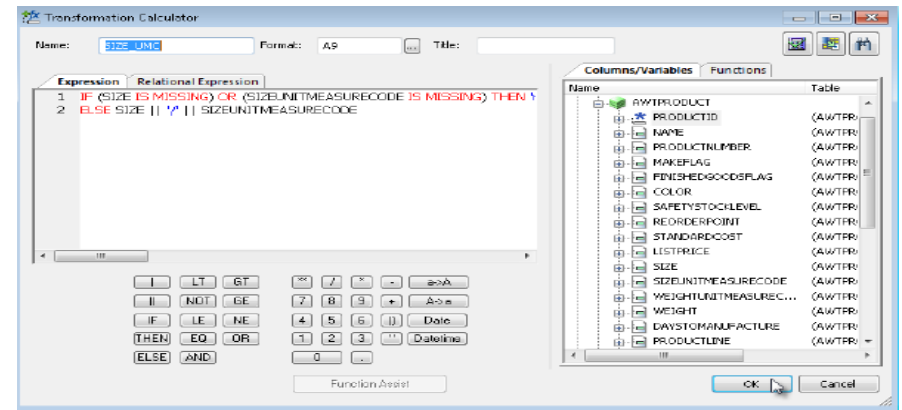
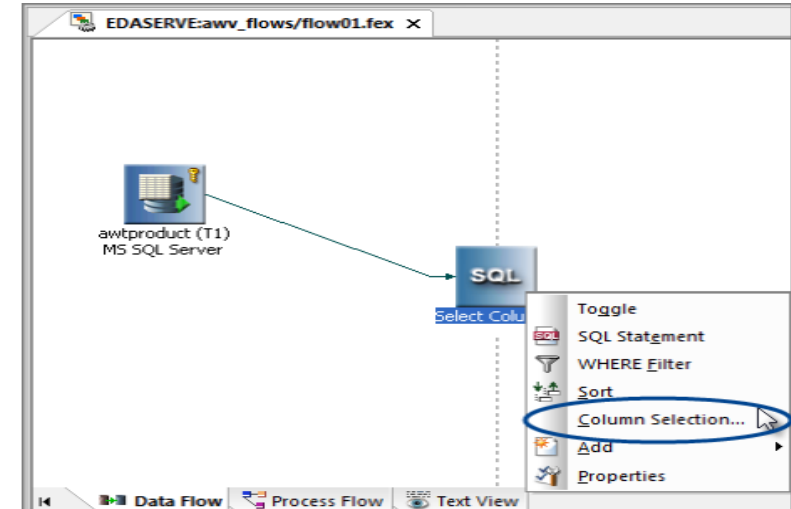
- DataMigrator Server
  - Engine that stores and runs ETL 'Flows'
  - Same reporting server as Web Query
- Data Management Console (DMC)
  - Graphical user interface
    - Build and manage ETL 'Flows'
    - Set up Schedules to Automate
  - Included in Developer Workbench
    - Installed on PC



# What is a Data Flow?

A Data Flow is where you define the data sources, transformation, and map to the target tables.

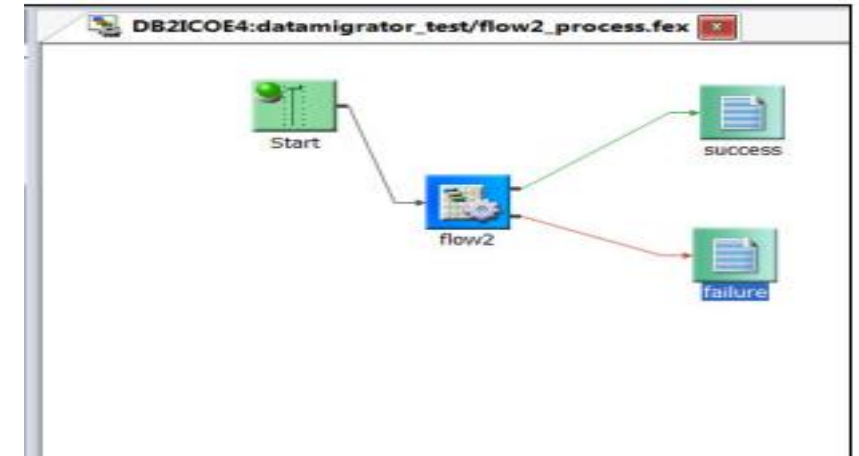
- Leverage existing DB2 Web Query meta data or create as part of the process
- Select columns, add where clause, edit SQL, or join files as part of the source transformations
- Create virtual (derived) columns
- Choose or create new target table(s)
- Determine load type (default is insert/update)
- Test your transformations, edit if necessary





## What is a Process Flow?

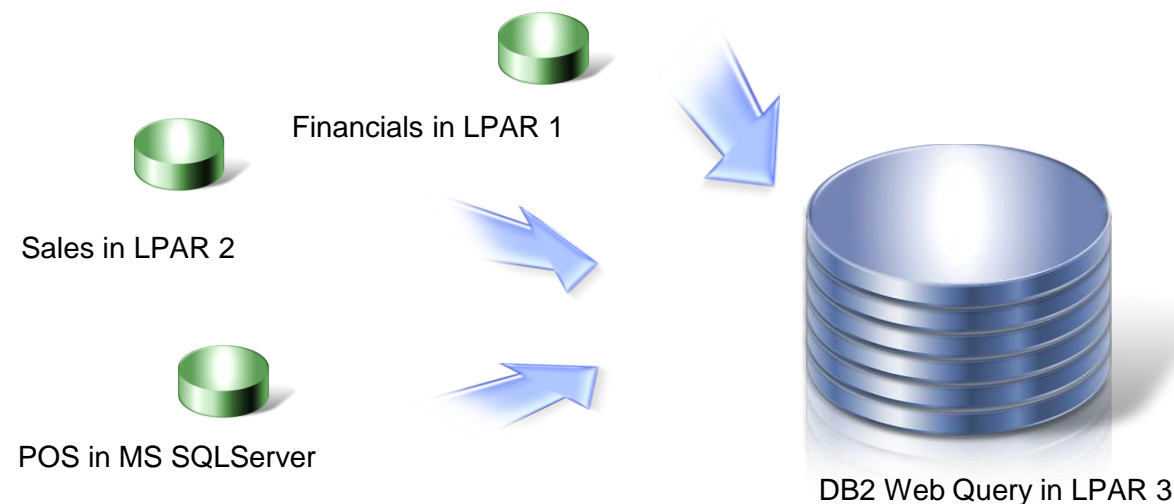
A Process Flow defines a set of steps that execute a series of components or objects, including the data flows that define your extractions, transformation, and load logic. You can define:



- **Data Flow** objects that indicate what data to extract and move from data sources to data targets.
- **Email** objects that notify users about the status of the process at specified points.
- **Stored Procedure** objects that perform some task before or after a step in the process flow
  - Example: Make a note in the log about success or failure of a load step
- **Connector** objects, represented as arrows, that specify execution logic for the other objects included in the process flow
  - Example: If step 1 is successful, take this path. If unsuccessful, take this other path
- **Group** objects that specify the flow of a subset of objects within the total flow
  - Example: Group two steps on the process flow and don't send notification email until both steps are completed

## DataMigrator Example

- Dedicated LPAR or Server for DB2 Web Query and Data Warehouse/Mart
  - DB2 Web Query and DataMigrator ONLY installed in LPAR3
  - DataMigrator automates process of pulling data from sources into LPAR3
    - Pull directly from operational databases (full refresh or based on some selection criteria, example where date=current\_date)
    - Pull from a Journal Receiver (Change Data Capture)
    - All reports run against data in LPAR 3 database

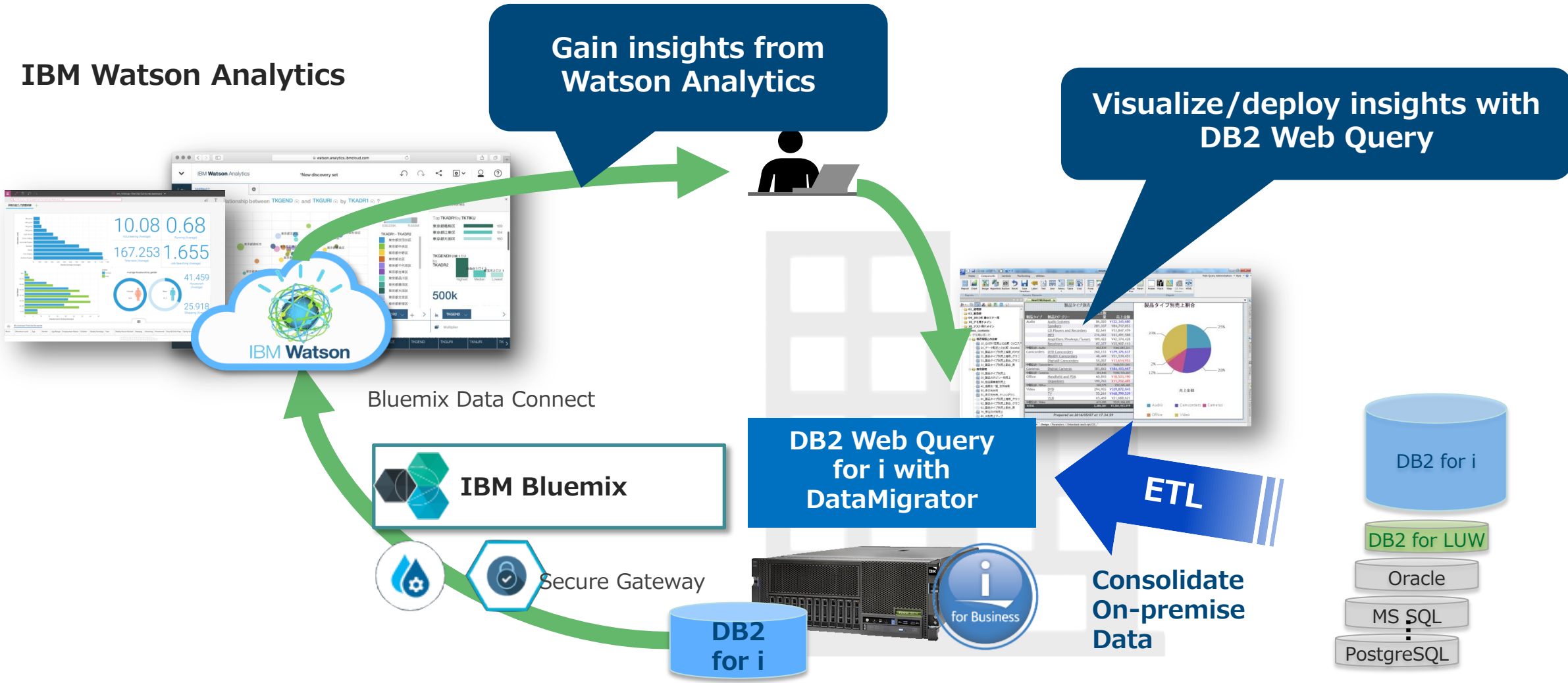




## DataMigrator Use Cases

- Data Replication
  - Ongoing Log (Journal Receiver based) in near real time
- Data Migration
  - One time or on going data synchronization across different databases
- Data Preparation
  - Consolidate and/or transform data in preparation for feeding into predictive analytics or Watson
- Data Warehousing or Data Mart Automation
  - Timed ETL
- Spreadsheet Consolidation
  - Automate upload of spreadsheet of flat file data to offer reporting combining this data with DB2 data

# Preparing Data For Watson; Visualizing Insights with DB2 Web Query



Power Systems(IBM i)

**Any Questions?**



## For Assistance: DB2 for i Lab Services Offerings

- DB2 Web Query Getting Started Service
  - 3 Day skills transfer and workshops to build prototype
  - Goes beyond tutorials with Best Practices and hand-holding
  - Best way to ensure successful implementation
  - Should include with EVERY DB2 Web Query proposal (proposal insert available)
- Query/400 Modernization
  - Discovery, Analysis, and Conversion
  - Addresses the question of how a customer can move from 100's or 1000's of Query/400 definitions to modernized environment that leverages DB2 query optimization!
  - Out with the (very) OLD, in with the NEW – improve performance, productivity, perception of IBM i, and move into “analytics”
- DB2 Web Query V1 to V2 Migration Service
  - Remote assistance to help customers move to Version 2
  - While there exists migration tooling, migrations can become more complex depending on the customer's DB2 Web Query Environment - inexpensive service to hand hold the migration
- DB2 and SQL Performance Assessments
  - Really critical assessment that should be conducted regularly by the customer
    - But lack of DBAs in our i community means no one is paying attention to the database
  - Assessment collects database (and system) performance data
  - Recommendations made in report

## To Learn More

- DB2 Web Query for i Website
  - [Ibm.biz/db2webqueryi](http://ibm.biz/db2webqueryi)
- DB2 Web Query for i Wiki
  - [Ibm.co/db2wqwiki](http://ibm.co/db2wqwiki)
- DB2 Web Query Getting Started Enablement:
  - <https://ibm.biz/db2wqconsulting>
- Demonstrations:
  - Wizard Analytics: <https://ibm.biz/DB2WQWizards>
  - End User Demos: <https://ibm.biz/db2wqreportingdemos>
- Follow DB2 Web Query guy Doug Mack on twitter at @mckdrmoly or check out his blog at <http://db2webqueryi.blogspot.com/> for all the latest

