Data Modeling with Power BI Desktop



Agenda

- Creating Table Relationships
- Creating Calculated Columns and Measure
- Creating Tables using DAX Expressions
- Configuring Fields for Geographic Mapping
- Creating Dimensional Hierarchies
- Using the DAX Calculate Function
- Calendar Tables and Time Intelligence



Data Modeling with Power BI Desktop

- Steps to create a data model with Power Pivot
 - Create relationships between tables
 - Modify native columns (e.g. set formatting and data category)
 - Create calculated columns
 - Create measures
 - Create dimensional hierarchies
 - Add Calendar table(s)



Table Relationships

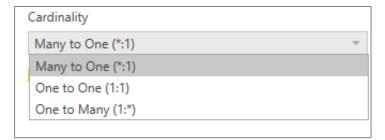
- Tables in data model associated with relationships
 - Relationships based on single columns
 - Tabular model supports [1-to-1] and [1-to-many] relationships



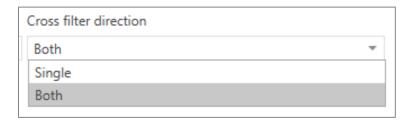


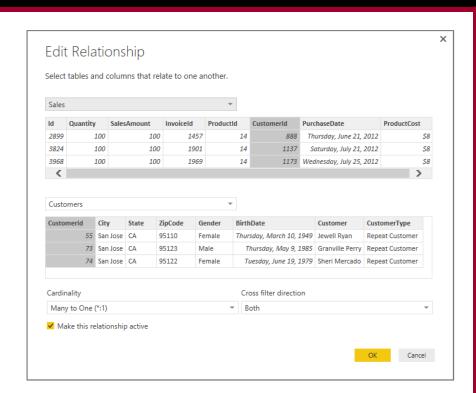
Relationship Properties

Cardinality



Cross filter direction

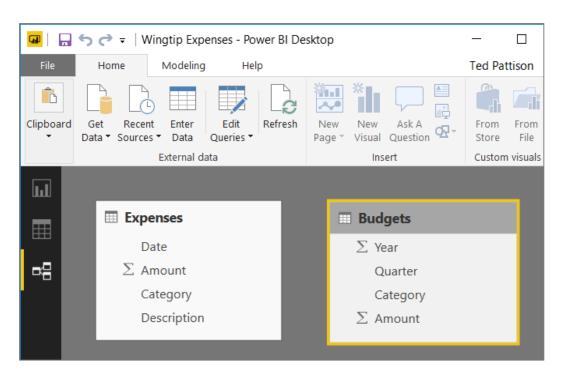






How Do You Create a Relationship Here?

- Two tables don't have fields to create relationship
 - The solution is to create two new calculated columns



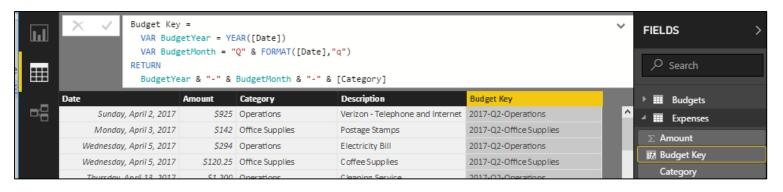


Creating Composite Key Fields

Create composite key column in Budgets

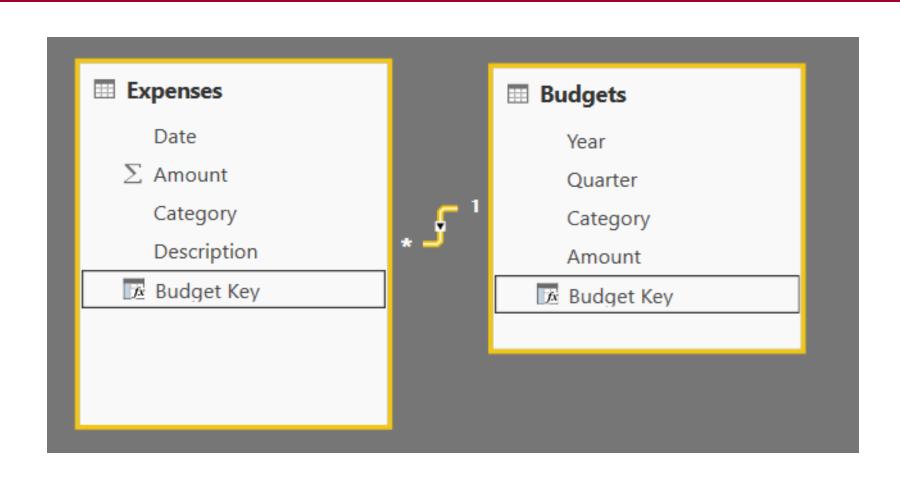


Create composite key column in Expenses





Create Relationship Using Composite Keys





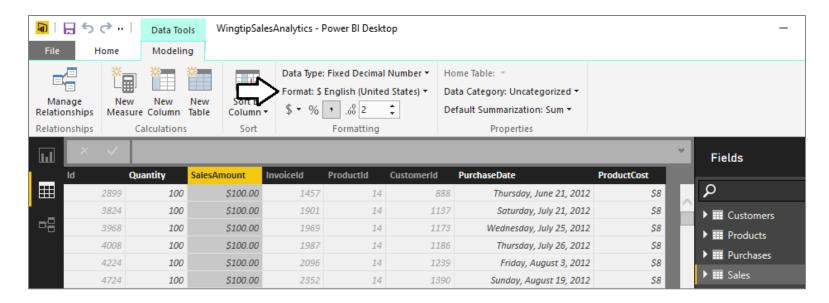
Agenda

- ✓ Creating Table Relationships
- Creating Calculated Columns and Measure
- Creating Tables using DAX Expressions
- Configuring Fields for Geographic Mapping
- Creating Dimensional Hierarchies
- Using the DAX Calculate Function
- Calendar Tables and Time Intelligence



Formatting Columns

- Each column has its own formatting properties
 - Formatting propagated to reports and visuals
 - Makes it easier on data model consumers





Working with DAX

- DAX is the language used to create data models
 - DAX stands for "Data Analysis Expression Language"
- DAX expressions are similar to Excel formulas
 - They always start with an equal sign (=)
 - DAX provides many built-in functions similar to Excel
- DAX Expressions are unlike Excel formulas...
 - DAX expressions cannot reference cells (e.g. A1 or C4)
 - Instead DAX expressions reference columns and tables

```
=SUM('Sales'[SalesAmount])
```



Writing DAX Expressions

Some DAX expressions are simple

```
Sales Revenue = Sum(Sales[SalesAmount])
```

Some DAX expressions are far more complex

```
Sales Growth PM = IF(
  ( ISFILTERED(Calendar[Month]) && ISFILTERED(Calendar[Date]) = FALSE() ),
  DIVIDE(
   SUM(Sales[SalesAmount]) -
   CALCULATE(
      SUM(Sales[SalesAmount]),
      PREVIOUSMONTH(Calendar[Date])
    ),
   CALCULATE(
      SUM(Sales[SalesAmount]),
      PREVIOUSMONTH(Calendar[Date])
  BLANK()
```



Creating Variables in DAX Expressions

- Variables can be added at start of expression
 - Use VAR keyword once for each variable
 - Use RETURN keyword to return expression value

```
Budget Key =
  VAR BudgetYear = YEAR([Date])
  VAR BudgetMonth = "Q" & FORMAT([Date], "q")
RETURN
  BudgetYear & "-" & BudgetMonth & "-" & [Category]
```



Calculated Columns vs Measures

- Calculated Columns (aka Columns)
 - Evaluated based on context of a single row
 - Evaluated when data is loaded into memory

```
Column1 = <DAX expression>
```

- Measures
 - Evaluated at query time based on current filter context
 - Commonly used for aggregations (e.g. SUM, AVG, etc.)
 - Used more frequently than calculated columns

```
Measure1 = <DAX expression>
```



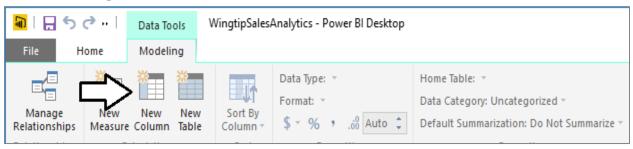
When to Create Calculated Columns

- Measures often better choice than calculate columns
 - Don't create calculated column when you need a measure
 - Prefer to create calculated columns only in specific scenarios
- When should you create calculated columns?
 - To create headers for row labels or column labels.
 - To place calculated results in a slicer for filtering
 - Define an expression strictly bound to current row
 - Categories text or numbers (e.g. customer age groups)



Creating Calculated Columns

- Edited in formula bar of Power Pivot data view
 - Start with name and then equals (=) sign
 - Enter a valid DAX expression
 - Clicking on column adds it into expression



× •	<pre>PurchaseYear = YEAR(Sales[PurchaseDate])</pre>													
ld	Quantity	SalesAmount	Invoiceld	ProductId	CustomerId	PurchaseDate	ProductCost	SalesProfit	PurchaseYear					
2899	100	\$100.00	1457	14	888	6/21/12	\$8.00	\$92.00	2012					
3824	100	\$100.00	1901	14	1137	7/21/12	\$8.00	\$92.00	2012					
3968	100	\$100.00	1969	14	1173	7/25/12	\$8.00	\$92.00	2012					
4008	100	\$100.00	1987	14	1186	7/26/12	\$8.00	\$92.00	2012					
4224	100	\$100.00	2096	14	1239	8/3/12	\$8.00	\$92.00	2012					
4724	100	\$100.00	2352	14	1390	8/19/12	\$8.00	\$92.00	2012					

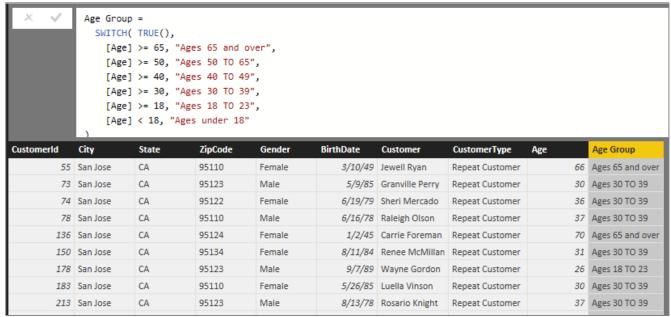


Calculated Column for Customer Age Group

1. Calculate customer age from birthdate

Ш	× ✓	X										
	CustomerId	City	State	ZipCode	Gender	BirthDate	Customer	CustomerType	Age			
	55	San Jose	CA	95110	Female	3/10/49	Jewell Ryan	Repeat Customer		66		
_	73	San Jose	CA	95123	Male	5/9/85	Granville Perry	Repeat Customer		30		
唱	74	San Jose	CA	95122	Female	6/19/79	Sheri Mercado	Repeat Customer		36		
	78	San Jose	CA	95110	Male	6/16/78	Raleigh Olson	Repeat Customer		37		
	136	San Jose	CA	95124	Female	1/2/45	Carrie Foreman	Repeat Customer		70		
	150	San Jose	CA	95134	Female	8/11/84	Renee McMillan	Repeat Customer		31		

2. Calculate age groups using calculated column





Calculated Column used in a Slicer

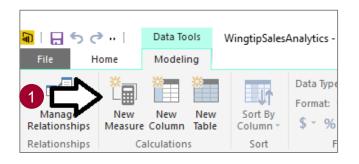
Calculated column can populate slicer values

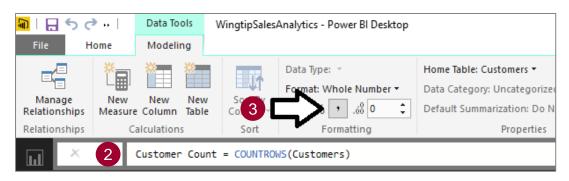




Creating Measures

- Measures have advantage over calculated columns
 - They are evaluated based on the current evaluation context
- Creating a measure with Power BI Desktop
 - Click New Measure button
 - 2. Give measure a name and write DAX expressions
 - 3. Configure formatting

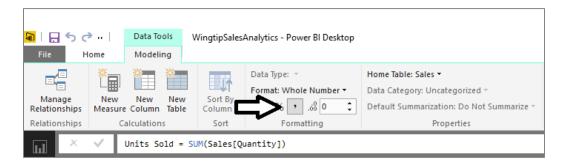




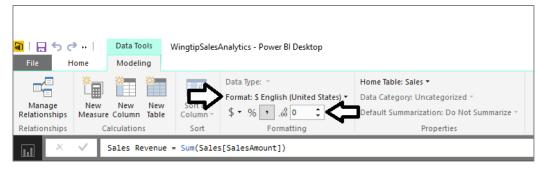


Formatting Measures

Format as whole number



Format as currency





Working with UNICHAR Characters

- Create a measure to return a single UNICHAR character
 - There are many different UNICHAR characters with symbols

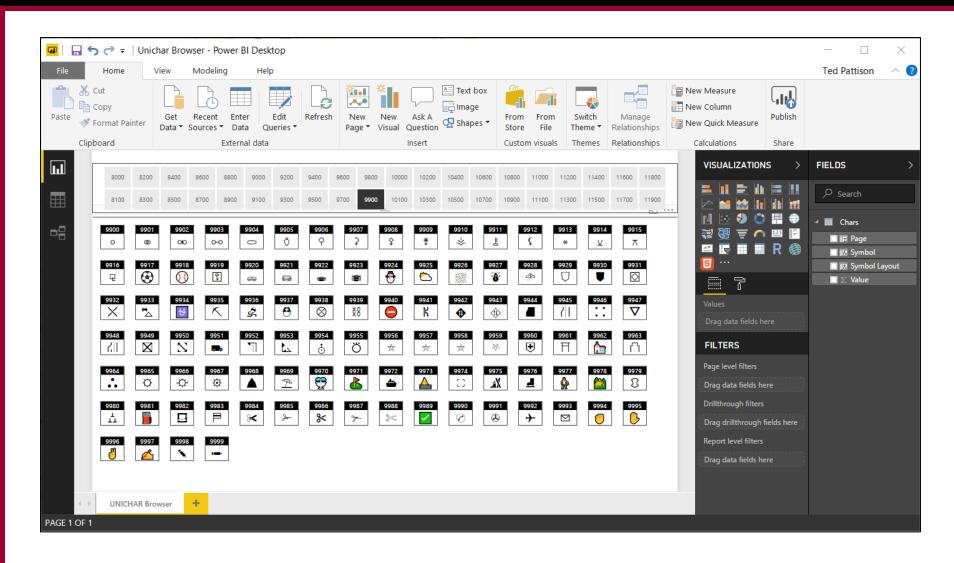
```
Status =
IF(
    [Budget Used] > 1,
    UNICHAR(9940),
    UNICHAR(9989)
)
```

UNICHAR character symbols can be displayed in table or matrix

① 44 (A)	=			
Year	Actual Expenses	Budget	Budget Used	Status
2017	\$85,073.74	\$92,000	92.47%	✓
Q1	\$20,788.42	\$26,000	79.96%	~
Marketing Budget for Q1 of 2017	\$3,780.77	\$5,000	75.62%	~
Office Supplies Budget for Q1 of 2017	\$7,191.65	\$8,000	89.90%	✓
Operations Budget for Q1 of 2017	\$7,446.00	\$8,000	93.08%	~
Research & Development Budget for Q1 of 2017	\$2,370.00	\$5,000	47.40%	~
Q2	\$18,283.31	\$22,000	83.11%	~
Marketing Budget for Q2 of 2017	\$5,769.46	\$6,000	96.16%	~
Office Supplies Budget for Q2 of 2017	\$2,602.85	\$4,000	65.07%	~
Operations Budget for Q2 of 2017	\$7,581.00	\$7,000	108.30%	



The UNICHAR Browser Demo



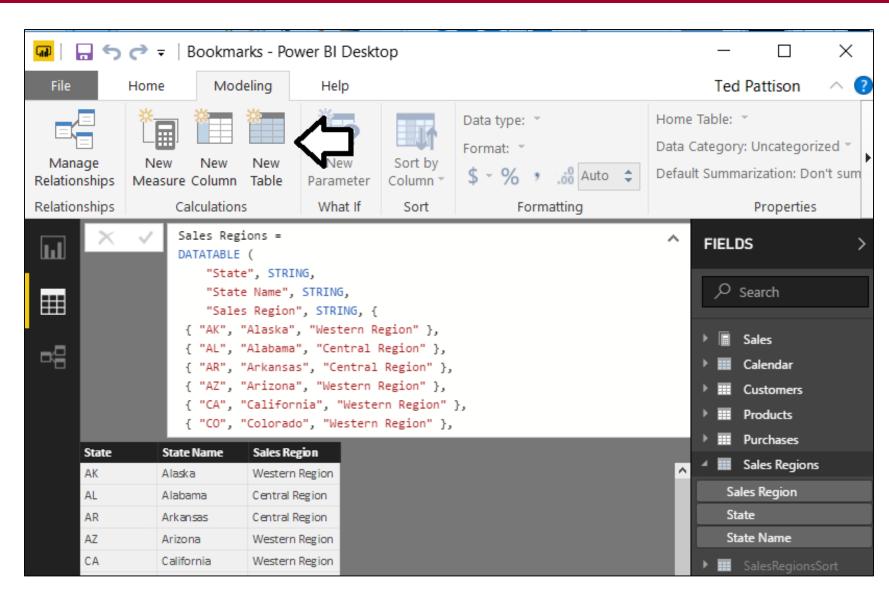


Agenda

- ✓ Creating Table Relationships
- Creating Calculated Columns and Measure
- Creating Tables using DAX Expressions
- Configuring Fields for Geographic Mapping
- Creating Dimensional Hierarchies
- Using the DAX Calculate Function
- Calendar Tables and Time Intelligence



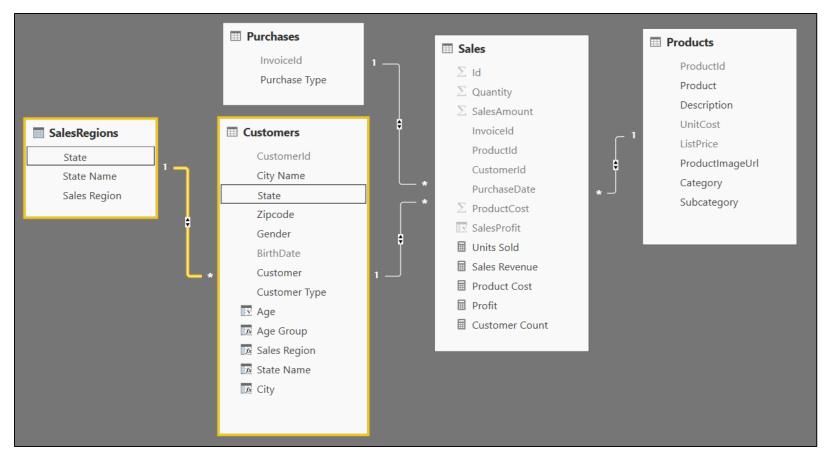
Creating Tables Dynamically using DAX





Integrating the Lookup Table into the Data Model

- Lookup table must be integrated into data model
 - Accomplished by creating relationship to one or more tables





The RELATED Function

- RELATED function performs cross-table lookup
 - Effectively replaces older VLOOKUP function
 - Used in many-side table to look up value from one-side
 - Used to pull data from lookup table into primary table

ш	× 🗸	X											
	CustomerId	City	State	ZipCode	Gender	BirthDate	Customer	CustomerType	Age	Age Group	Sales Region		
	55	San Jose	CA	95110	Female	3/10/49	Jewell Ryan	Repeat Customer	66	Ages 65 and over	Western Region		
_	73	San Jose	CA	95123	Male	5/9/85	Granville Perry	Repeat Customer	30	Ages 30 TO 39	Western Region		
唱	74	San Jose	CA	95122	Female	6/19/79	Sheri Mercado	Repeat Customer	36	Ages 30 TO 39	Western Region		
	78	San Jose	CA	95110	Male	6/16/78	Raleigh Olson	Repeat Customer	37	Ages 30 TO 39	Western Region		
	136	San Jose	CA	95124	Female	1/2/45	Carrie Foreman	Repeat Customer	70	Ages 65 and over	Western Region		
	150	San Jose	CA	95134	Female	8/11/84	Renee McMillan	Repeat Customer	31	Ages 30 TO 39	Western Region		

ш	× ✓	State Nam	State Name = RELATED(SalesRegions[StateFullName])											
	State	ZipCode	Gender	BirthDate	Customer	CustomerType	Age	Age Group	Sales Region	State Name				
	CA	95110	Female	3/10/49	Jewell Ryan	Repeat Customer	66	Ages 65 and over	Western Region	California				
	CA	95123	Male	5/9/85	Granville Perry	Repeat Customer	30	Ages 30 TO 39	Western Region	California				
唱	CA	95122	Female	6/19/79	Sheri Mercado	Repeat Customer	36	Ages 30 TO 39	Western Region	California				
	CA	95110	Male	6/16/78	Raleigh Olson	Repeat Customer	37	Ages 30 TO 39	Western Region	California				
	CA	95124	Female	1/2/45	Carrie Foreman	Repeat Customer	70	Ages 65 and over	Western Region	California				
	CA	95134	Female	8/11/84	Renee McMillan	Repeat Customer	31	Ages 30 TO 39	Western Region	California				



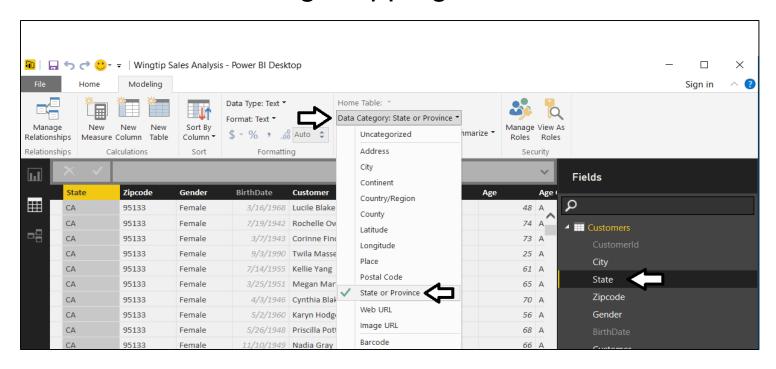
Agenda

- ✓ Creating Table Relationships
- Creating Calculated Columns and Measure
- Creating Tables using DAX Expressions
- Configuring Fields for Geographic Mapping
- Creating Dimensional Hierarchies
- Using the DAX Calculate Function
- Calendar Tables and Time Intelligence



Geographic Field Metadata

- Fields in data model have metadata properties
 - Metadata used by visuals and reporting tools
 - Used as hints to Bing Mapping service





Eliminate Geographic Ambiguity

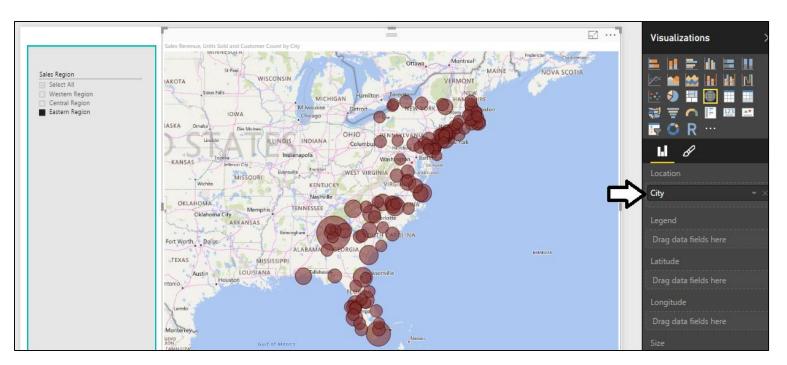
- City name alone is ambiguous
 - "Athens" defaults to Greece not Georgia
 - Concatenate city name with state to disambiguate

>	City = [City Name] & ", " & [State]										
	Age Group	Sales Region	State Name	SalesRegionSort	City						
48	Ages 40 TO 49	Western Region	California	1	San Jose, CA						
74	Ages 65 and over	Western Region	California	1	San Jose, CA						
73	Ages 65 and over	Western Region	California	1	San Jose, CA						
25	Ages 18 TO 23	Western Region	California	1	San Jose, CA						
61	Ages 50 TO 65	Western Region	California	1	San Jose, CA						
65	Ages 65 and over	Western Region	California	1	San Jose, CA						



Using Map Visual with a Geographic Field

- Map Visual shows distribution over geographic area
 - Visual automatically updates when filtered





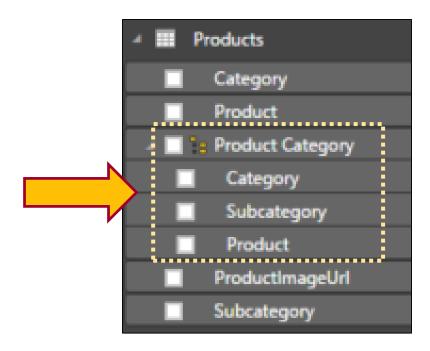
Agenda

- ✓ Creating Table Relationships
- Creating Calculated Columns and Measure
- Creating Tables using DAX Expressions
- Configuring Fields for Geographic Mapping
- Creating Dimensional Hierarchies
- Using the DAX Calculate Function
- Calendar Tables and Time Intelligence



Dimensional Hierarchies

- Hierarchy created from two or more columns
 - All columns in hierarchy must be from the same table
 - Defines parent-child relationship between columns
 - Provides path to navigate through data
 - Provides path to drill down into greater level of detail



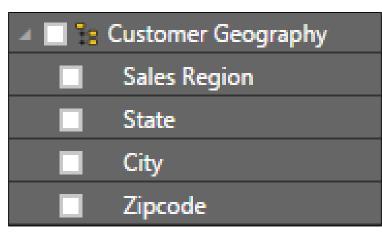


Pulling Columns for Hierarchy into Single Table

- Sometimes hierarchy columns are spread across tables
 - Use RELATED function from DAX to pull columns into single table

× ✓	<pre>Sales Region = RELATED(SalesRegions[SalesRegion])</pre>											
Customer	Customer Type	Age	Age Group	Sales Region	State Name							
Lucile Blake	One-time Customer	48	Ages 40 TO 49	Western Region	California							
Rochelle Owen	One-time Customer	74	Ages 65 and over	Western Region	California							
Corinne Finch	One-time Customer	73	Ages 65 and over	Western Region	California							
Twila Massey	One-time Customer	25	Ages 18 TO 23	Western Region	California							

Then create hierarchy in the table with all the columns





Agenda

- ✓ Creating Table Relationships
- Creating Calculated Columns and Measure
- Creating Tables using DAX Expressions
- Configuring Fields for Geographic Mapping
- Creating Dimensional Hierarchies
- Using the DAX Calculate Function
- Calendar Tables and Time Intelligence



A Tale of Two Evaluation Contexts

Row Context

- Context includes all columns in iteration of current row
- Used to evaluate DAX expression in calculated column
- Only available in measures with iterator function (e.g. SUMX)

Filter Context

- Context includes filter(s) defining current set of rows
- Used by default to evaluate DAX expressions in measures
- Can be fully ignored or partially ignored using DAX code
- Not used to evaluate DAX in calculated columns



Understanding Row Context

Row context used to evaluate calculated columns

>	City = [City Name] & ", " & [State]								
	Age Group	Sales Region	State Name	SalesRegionSort	City				
48	Ages 40 TO 49	Western Region	California	1	San Jose, CA				
74	Ages 65 and over	Western Region	California	1	San Jose, CA				
73	Ages 65 and over	Western Region	California	1	San Jose, CA				
25	Ages 18 TO 23	Western Region	California	1	San Jose, CA				
61	Ages 50 TO 65	Western Region	California	1	San Jose, CA				
65	Ages 65 and over	Western Region	California	1	San Jose, CA				

× ✓	Ag	Age = Floor((TODAY()-Customers[BirthDate])/365, 1)										
Customer		Customer Type	Age	Age Group	Sales Region	State Name						
Lucile Blake		One-time Customer	48	Ages 40 TO 49	Western Region	California						
Rochelle Owen		One-time Customer	74	Ages 65 and over	Western Region	California						
Corinne Finch		One-time Customer	73	Ages 65 and over	Western Region	California						



Understanding Iterators Like SUMX

- Standard aggregation functions (e.g. sum) have no row context
 - You can use SUM to sum values of a single column
 - You cannot use SUM to sum results of an expressions

```
Gross Margin = SUM( Sales[SalesAmount]-Sales[ProductCost] )

The SUM function only accepts a column reference as an argument.
```

Iterator functions (e.g. SUMX) iterate through rows in target table

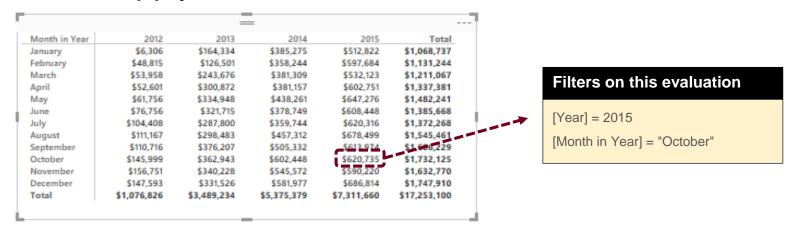
```
X ✓ Gross Margin = SUMX(Sales, Sales[SalesAmount]-Sales[ProductCost] )
```

- First argument accepts expressions that evaluates to table of rows
- Second argument accepts expression that is evaluated for each row



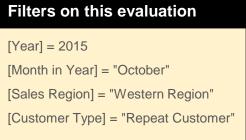
Understanding Filter Context

Visuals apply various filters in different evaluation contexts



Filter context also affected by slicers and other filters







Using the CALCULATE Function

- CALCULATE function provides greatest amount of control
 - First argument defines expression to evaluate
 - Second argument defines table on which to evaluate expression
 - You can evaluate expressions with or without current filter context

```
Pct of All Products =
DIVIDE(
    SUM( Sales[SalesAmount] ),
    CALCULATE(
        Sum (Sales[SalesAmount] ),
        ALL(Products[Category], Products[Subcategory], Products[Product])
    )
)
```

```
Pct of Product Category =
DIVIDE(
    SUM( Sales[SalesAmount] ),
    CALCULATE(
        Sum (Sales[SalesAmount] ),
        ALL( Products[Subcategory], Products[Product] )
    )
)
```



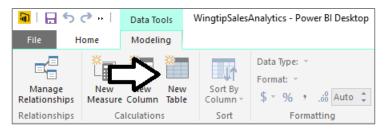
Agenda

- ✓ Creating Table Relationships
- Creating Calculated Columns and Measure
- Creating Tables using DAX Expressions
- Configuring Fields for Geographic Mapping
- Creating Dimensional Hierarchies
- Using the DAX Calculate Function
- Calendar Tables and Time Intelligence

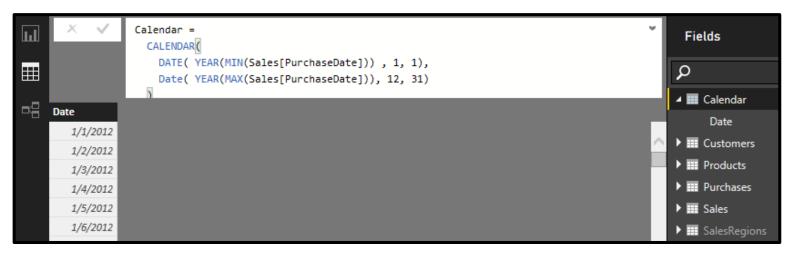


Creating Calendar Table as Calculated Table

Use New Table command in ribbon



Create calendar table using DAX CALENDAR function





Adding Columns to Calendar Table

Creating the Year column



Creating the Quarter column



Creating the Month column



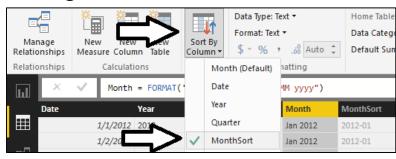


Configuring Sort Columns

- Month column will not sort in desired fashion by default
 - For example, April will sort before January, February and March
- Creating a sort column for the Month column
 - MonthSort sorts alphabetically & chronologically at same time

× •	MonthSort = FORMAT('Calendar'[Date], "yyyy-MM")					
Date	Year	Quarter	Month	MonthSort		
1/1/2012	2012	2012-Q1	Jan 2012	2012-01		
1/2/2012	2012	2012-Q1	Jan 2012	2012-01		

Configure Month column with MonthSort as sort column





Columns for Month in Year and Day in week

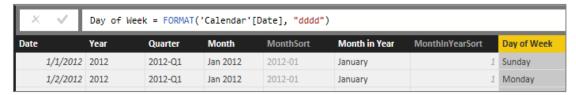
Creating the Month in Year column



Creating the MonthlnYearSort column



Creating the Day of Week column



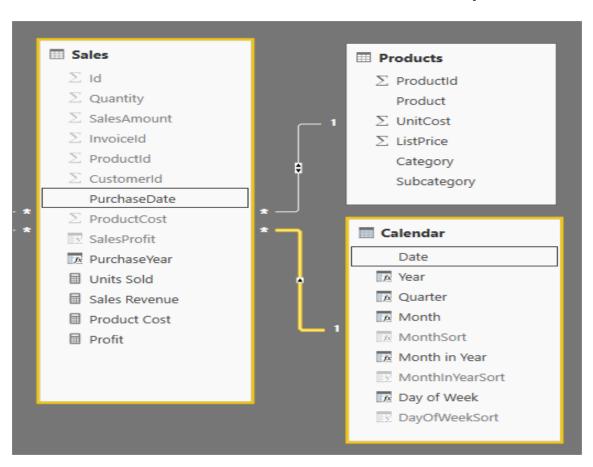
Creating the DayOfWeekSort column





Integrating Calendar Table into Data Model

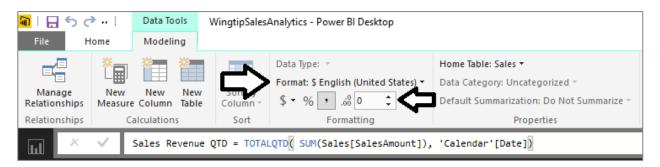
Calendar table needs relationship to one or more tables



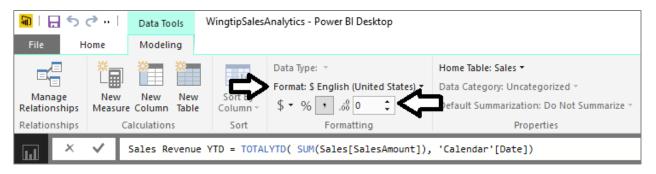


Calculated Fields for QTD and YTD Sales

TOTALQTD function calculates quarter-to-date totals



TOTALYTD function calculates year-to-date totals





Creating Running Total using CALCULATE

- Calculate a running total of sales revenue across years
 - This must be done using CALCULATE function

```
Data Type: "
                                                  Format: $ English (United States) *
  Manage
                New
                               New
                                        Sort By
              Measure Column Table
Relationships
                                       Column ▼
Relationships
                    Calculations
                                         Sort
                                                            Formatting
                     Sales Revenue RT =
1.1
                     CALCULATE(
                          SUM(Sales[SalesAmount]),
FILTER(
                              ALL('Calendar'),
唱
                              'Calendar'[Date] <= MAX('Calendar'[Date])
```



Sales Growth PM Measure - First Attempt

Create a measure named Sales Growth PM

```
Sales Growth PM =
DIVIDE(
    SUM(Sales[SalesAmount]) -
    CALCULATE(
    SUM(Sales[SalesAmount]),
    PREVIOUSMONTH(Calendar[Date])
),
    CALCULATE(
    SUM(Sales[SalesAmount]),
    PREVIOUSMONTH(Calendar[Date])
)
)
```

- Use measure in matrix evaluating month and quarter
 - Measure returns correct value when filtered by Month
 - Measure returns large, erroneous value when filtered by Quarter

Г					=	
	Year	Quarter	Month	Sales Revenue	Sales Growth PM	
	2014	2014-Q1	Jan 2014	\$629,969	-18.13 %	
			Feb 2014	\$609,637	-3.23 %	
			Mar 2014	\$628,618	3.11 %	4
			Total	\$1,868,225	142.79 %	$\langle \Box$
		2014-Q2	Apr 2014	\$661,588	5.24 %	√
			May 2014	\$748,193	13.09 %	
			Jun 2014	\$814,333	8.84 %	4
			Total	\$2,224,114	253.81 %	$\langle \neg$
		2014-Q3	Jul 2014	\$788,469	-3.18 %	7



Using the ISFILTERED Function

ISFILTERED function used to determine when perform evaluation

```
Sales Growth PM =
IF(
    ( ISFILTERED(Calendar[Month]) && NOT(ISFILTERED(Calendar[Date])) ),
    DIVIDE(
        SUM(Sales[SalesAmount]) -
        CALCULATE(
            SUM(Sales[SalesAmount]),
            PREVIOUSMONTH(Calendar[Date])
        ),
        CALCULATE(
            SUM(Sales[SalesAmount]),
            PREVIOUSMONTH(Calendar[Date])
        )
            PREVIOUSMONTH(Calendar[Date])
        )
            BLANK()
        )
```

Expression returns Blank value when evaluation context is invalid

_				=	
Year	Quarter	Month	Sales Revenue	Sales Growth PM	
2014	2014-Q1	Jan 2014	\$629,969	-18.13 %	
		Feb 2014	\$609,637	-3.23 %	
		Mar 2014	\$628,618	3.11 %	4
		Total	\$1,868,225		$< \square$
	2014-Q2	Apr 2014	\$661,588	5.24 %	7
		May 2014	\$748,193	13.09 %	
		Jun 2014	\$814,333	8.84 %	4
		Total	\$2,224,114	,	$\leq \Box$
	2014-Q3	Jul 2014	\$788,469	-3.18 %	7
		Aug 2014	\$869,143	10.23 %	



Summary

- Creating Table Relationships
- Creating Calculated Columns and Measure
- Creating Tables using DAX Expressions
- Configuring Fields for Geographic Mapping
- Creating Dimensional Hierarchies
- Using the DAX Calculate Function
- ✓ Calendar Tables and Time Intelligence

