Designing Queries with Power BI Desktop



Agenda

- Deciding What To Measure
- Understanding Queries in Power BI Desktop
- Working with the Query Editor Window
- Designing Advanced Combine Queries
- Importing OLTP Data Into a Star Schema



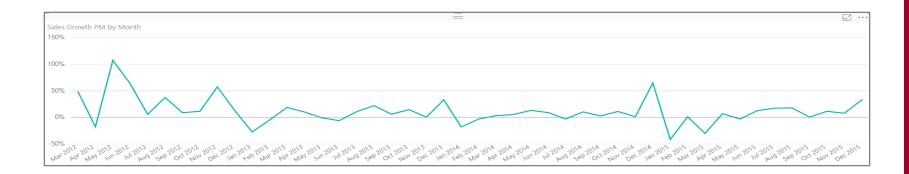
Data Discovery

- Data can live in a variety of sources
 - Files (e.g. CSV file, Excel workbook)
 - OLTP Databases
 - OLAP Databases
 - SharePoint Lists and Document Libraries
 - Azure-based services
 - Online services & SaaS applications



Deciding What To Measure

- You Must Determine Measurable Objectives
 - Financial (revenue, expenses, profit margin, etc.)
 - Business processes efficiency
 - Customer Satisfaction Levels





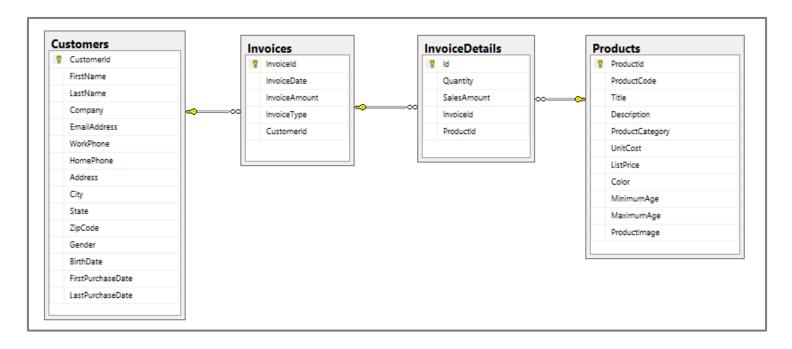
Defining Grain Statements

- Grain statements should be defined in initial design phase
 - Grain statements helps determine requirements for BI queries
 - Grain statements can be created & understood by business users
- Example grain statements for BI project at Wingtip Toys
 - What was the total sales revenue over the last 4 years?
 - What was the sales revenue by year, quarter and month?
 - What was the sales revenue by region, state, city and zip code?
 - What was the sales revenue by category, subcategory and product?
 - What was the growth in sales revenue from month to month in 2013?
 - What was profit margin for each product by year, quarter and month?
 - Have their been any products with significantly decreasing profit margin?



Sample OLTP Database: WingtipSalesDB

- Online Transaction Processing (OLTP) System
 - Used for real-time data access and transaction-based data entry
 - Optimized for faster transactions (e.g. inserts, updates & deletes)
 - Tables normalized to reduce/eliminate redundancies
 - Table schemas can be hard for business users to understand





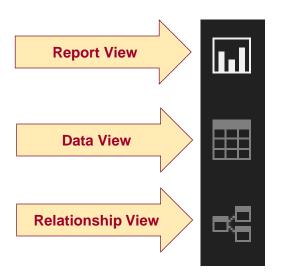
Agenda

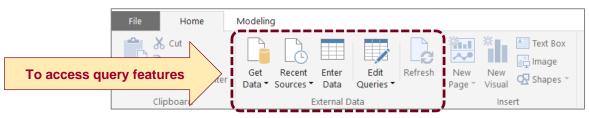
- Deciding What To Measure
- Understanding Queries in Power BI Desktop
- Working with the Query Editor Window
- Designing Advanced Combine Queries
- Importing OLTP Data Into a Star Schema



Getting Around in Power BI Desktop

- What do you need to learn to use Power BI Desktop?
 - Query features for importing data
 - Design features for modeling data
 - Report designer for creating reports
- Navigating between view modes

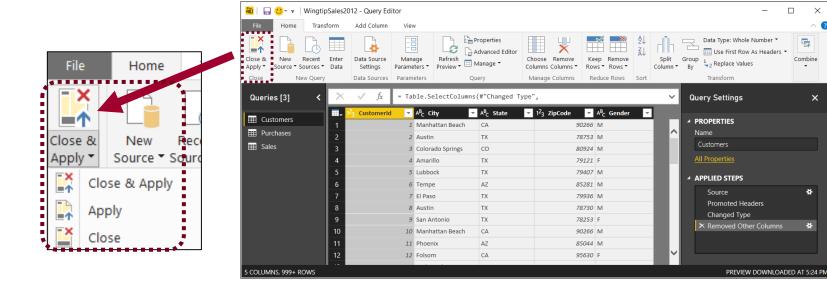






Query Editor Window

- Power BI Desktop provides separate Query Editor window
 - Provides powerful features for designing queries
 - Displays list of all queries in project on the left
 - Displays **Properties** and **Applied Steps** for selected query on right
 - Preview of table generated by query output shown in the middle
 - Query can be executed using Apply or Close & Apply command



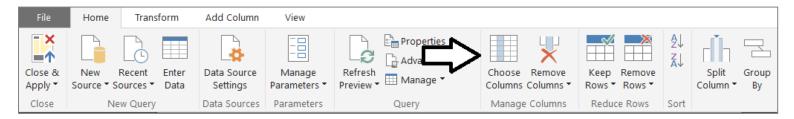


 \times

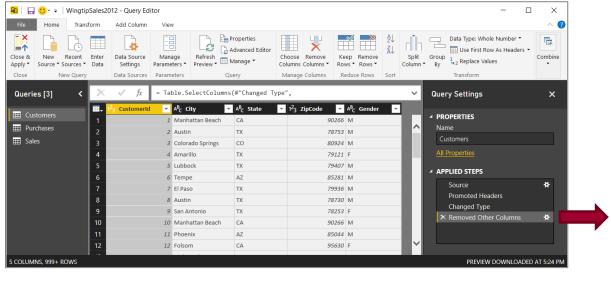
Combine

Queries Are Defined as Sequence of Steps

- Here is an example of adding a query step
 - Click a column header then click Choose Column button in ribbon



Query definition is modified by adding new step

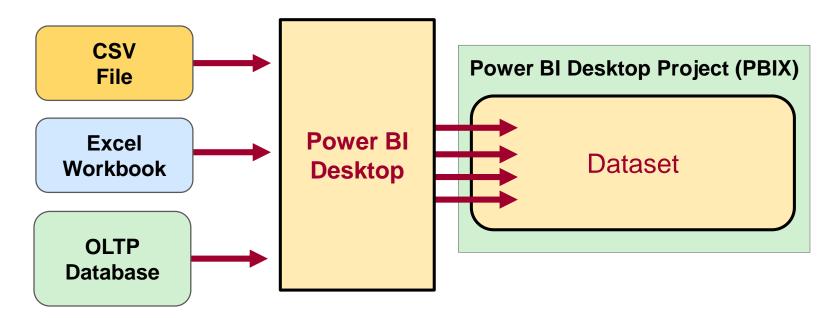






Power BI Desktop is an ETL Tool

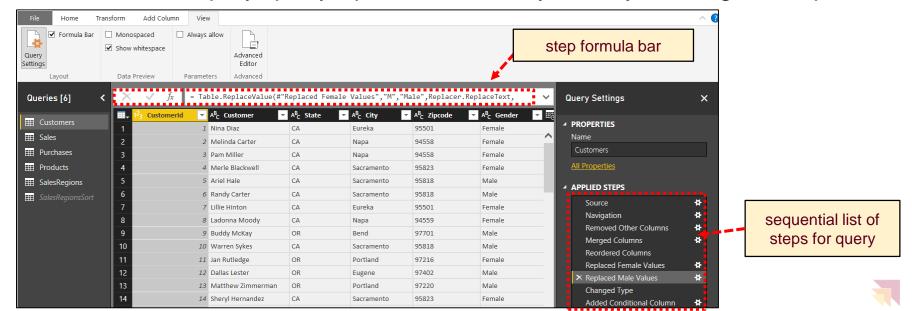
- ETL process is essential part of any BI Project
 - Extract the data from wherever it lives
 - Transform the shape of the data for better analysis
 - Load the data into dataset for analysis and reporting





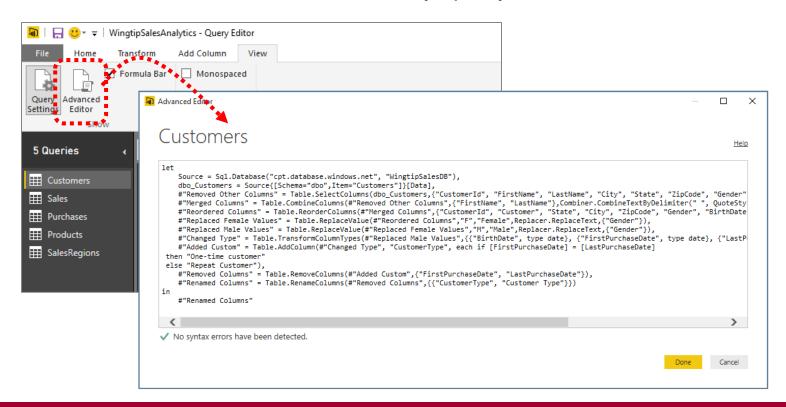
Query Steps

- A query is created as a sequence of steps
 - Each step is a parameterized operation on the data
 - Each step has formula which can be viewed/edited in formula bar
 - Query starts with Source step to extract data from a data source
 - Additional steps added to perform transform operations on data
 - You can replay query operations one by one by clicking on steps



Advanced Editor

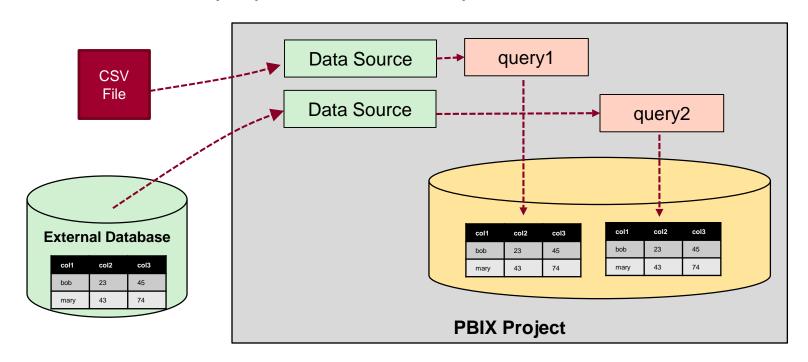
- Power BI Desktop based on "M" functional language
 - Query in Power BI Desktop saved as set of M statements in code
 - Query Editor generates code in M behind the scenes
 - Advanced users can view & modify query code in Advanced Editor





Understanding Query Input and Output

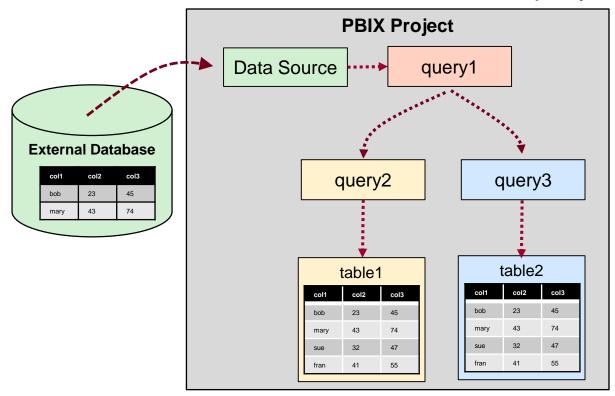
- PBIX project is container for data sources and queries
 - Queries created and saved within scope of Power BI project
 - Queries can pull data from local files
 - Queries can pull data from external content sources
 - Queries main purpose is to load imported data into data model

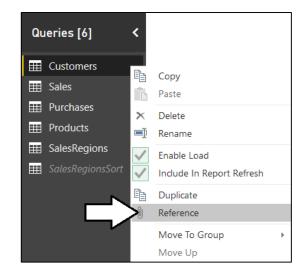




Query Composition

- Query can serve as source for other queries
 - Allows for creation of reusable base queries & query composition
 - Complexity can be hidden in base queries
 - Reference command creates new query based on another query

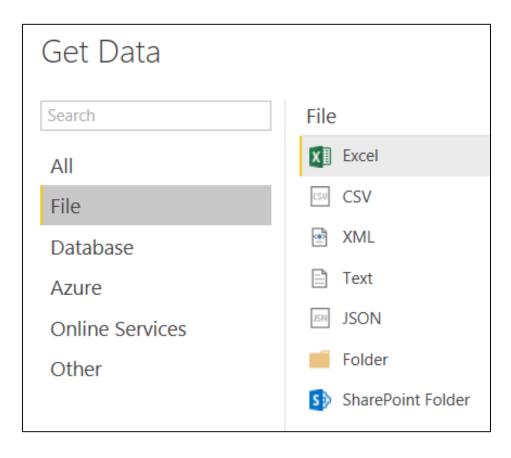






File-based Data Sources

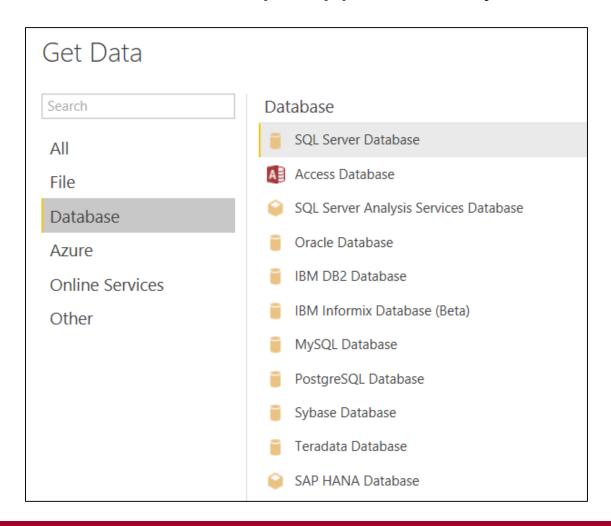
Power BI Desktop supports common file types





Supported Databases

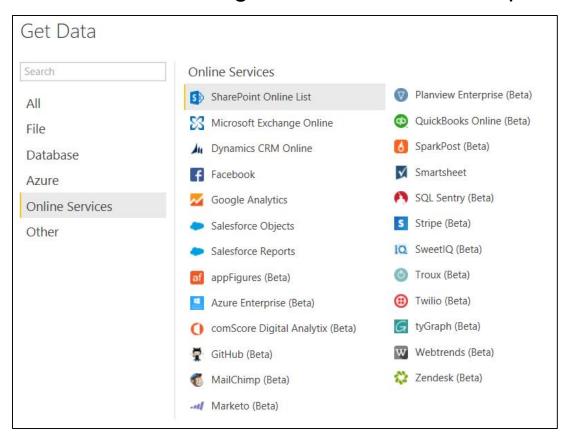
Power BI Desktop supports many database systems





Online Service Data Sources

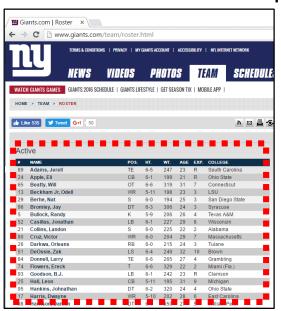
- Power BI Desktop Supports Online Services
 - Includes popular Software-as-a-Service (SaaS) applications
 - Microsoft is working with 3rd vendors to expand this list

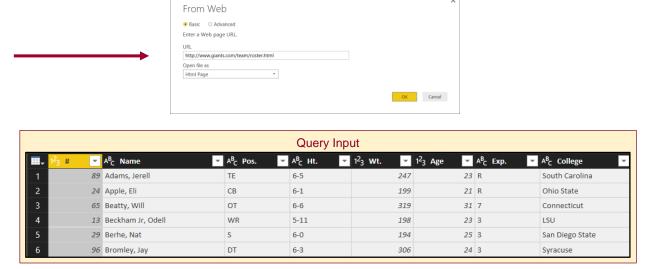




Working with Web Data Sources

- Many public websites publish data using HTML tables
 - Power BI desktop can scrape data from tables in HTML pages



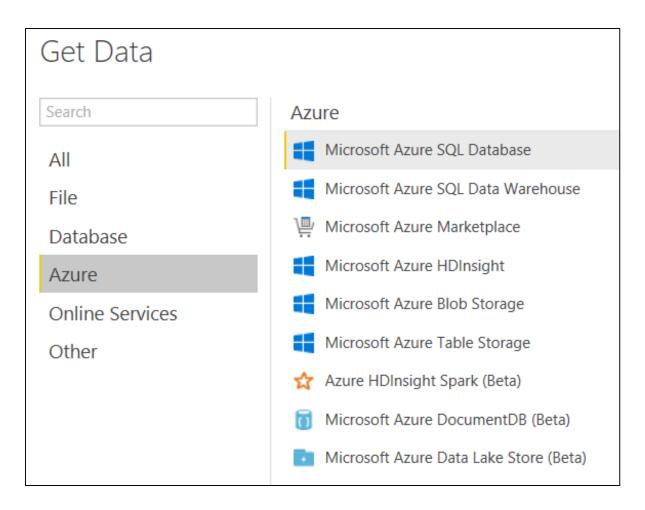






Azure Data Sources

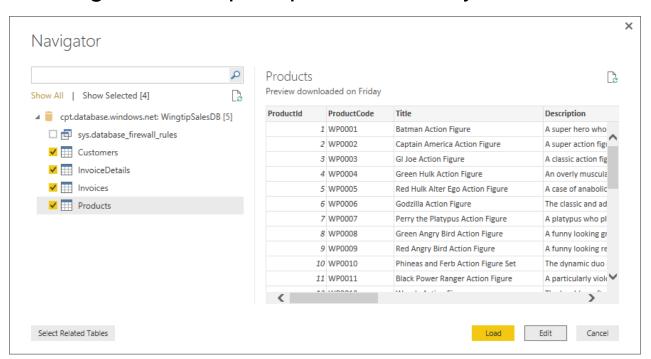
Power BI Desktop supports many Azure data sources





Selecting Tables from a SQL Database

- Power BI Desktop provides Navigator dialog
 - Allows you to select tables
 - Navigator understands existing table relationships
 - Clicking Load will run query and import data
 - Clicking Edit will open queries in Query Editor window







Agenda

- ✓ Deciding What To Measure
- ✓ Understanding Queries in Power BI Desktop
- Working with the Query Editor Window
- Designing Advanced Combine Queries
- Importing OLTP Data Into a Star Schema

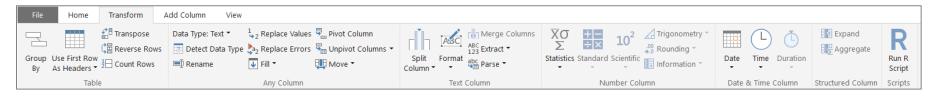


Query Editor Ribbon Tabs

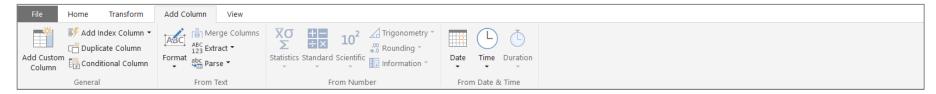
Home tab



Transform tab



Add Column tab



View tab





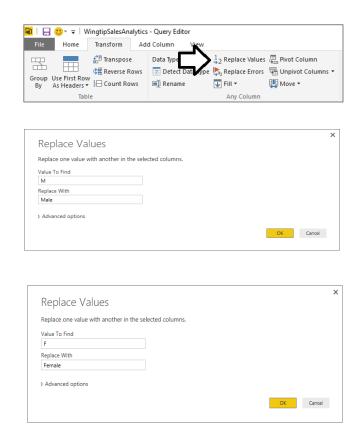
Examples of Basic Power BI Desktop Steps

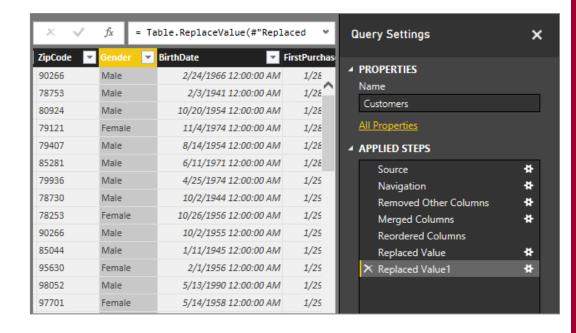
- Rename column
- Convert column type
- Format column values
- Reorder columns
- Replace column values
- Expanding related column
- Merging columns
- Splitting columns



Replacing Values

Used to substitute values during import







Converting Column Types

- Transform data to make it more reliable
 - Convert date-time column to date column
- Transform data to make it more efficient
 - Convert decimal to fixed decimal number for currency

PurchaseDate 🔻	1 ² ₃ Quantity	\$ SalesAmount	\$	ProductCost
1/28/2012	1	2.95	1.2	Decimal Number
1/28/2012	6		\$	Fixed Decimal Number
1/28/2012	1	19.95	1 ² 3	Whole Number
1/28/2012	5	249.75		Date/Time
1/28/2012	1	2.95		Date

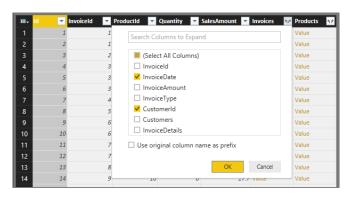
Beware: Conversion can have destructive effect on data



Expanding Related Columns

- Used to pull data from related tables
 - Saves you from performing SQL joins or VLOOKUP



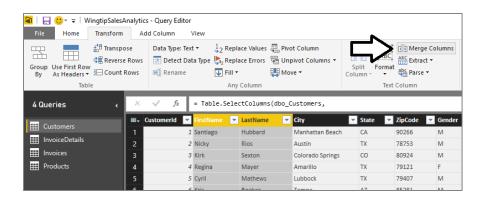


	ld 🔻	InvoiceId 🔻	ProductId 🔻	Quantity 🔻	SalesAmount 🔻	InvoiceDate 🔻	CustomerId 🔻	Products 40
1	1	1	22	4	119.8	1/28/2012 12:00:00 AM	1	Value
2	2	1	22	1	29.95	1/28/2012 12:00:00 AM	1	Value
3	3	2	22	2	59.9	1/28/2012 12:00:00 AM	2	Value
4	4	3	17	8	399.6	1/28/2012 12:00:00 AM	3	Value
5	5	3	18	2	29.9	1/28/2012 12:00:00 AM	3	Value
6	6	3	18	4	59.8	1/28/2012 12:00:00 AM	3	Value
7	7	4	16	1	2.95	1/28/2012 12:00:00 AM	4	Value



Merging Columns

Merge two columns into a single column

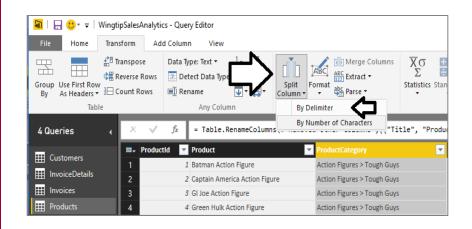


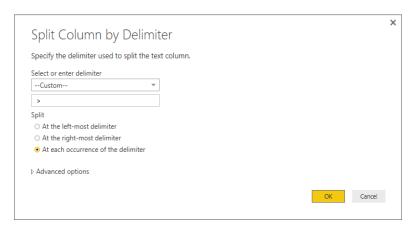


	CustomerId 🔻	Customer
1	1	Santiago Hubbard
2	2	Nicky Rios
3	3	Kirk Sexton
4	4	Regina Mayer
5	5	Cyril Mathews
6	6	Kris Booker
7	7	Tracy Christensen
8	8	Reed Glover

Splitting Columns

Split a single column up into two columns



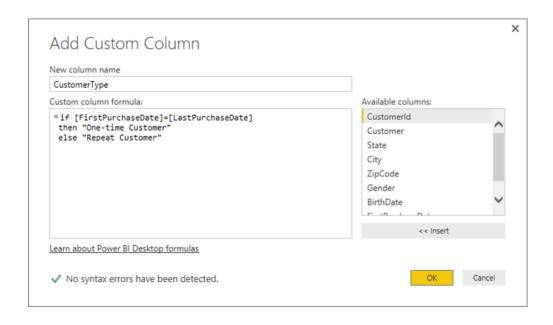


	ProductId 🔻	Product	Category	Subcategory
1	1	Batman Action Figure	Action Figures	Tough Guys
2	2	Captain America Action Figure	Action Figures	Tough Guys
3	3	GI Joe Action Figure	Action Figures	Tough Guys
4	4	Green Hulk Action Figure	Action Figures	Tough Guys
5	5	Red Hulk Alter Ego Action Figure	Action Figures	Tough Guys
6	6	Godzilla Action Figure	Action Figures	Tough Guys
7	7	Perry the Platypus Action Figure	Action Figures	Cute and Huggable
8	8	Green Angry Bird Action Figure	Action Figures	Cute and Huggable

Adding a Custom Column

- Custom column provide custom logic
 - Logic must be written in M programming language



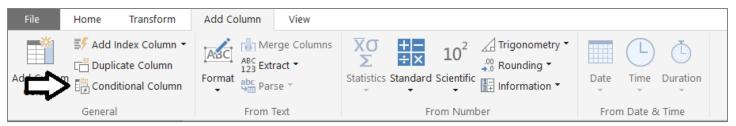


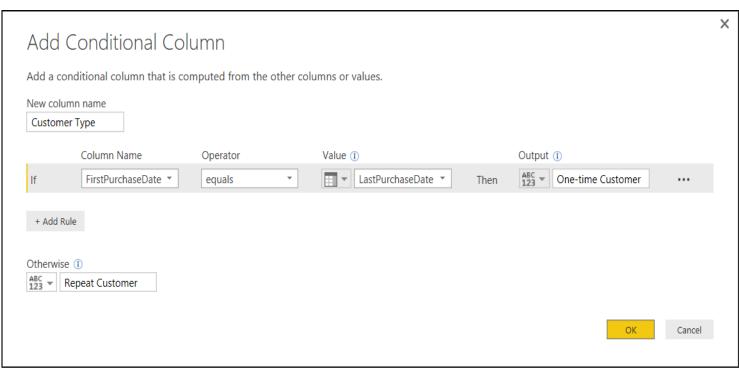
FirstPurchaseDate 🔻	LastPurchaseDate 🔻	CustomerType ~
1/28/2012	1/28/2012	One-time Customer
1/29/2012	11/22/2015	Repeat Customer
1/29/2012	10/2/2015	Repeat Customer
1/29/2012	1/29/2012	One-time Customer
1/29/2012	5/6/2015	Repeat Customer
1/29/2012	1/29/2012	One-time Customer



Adding a Conditional Column

Abstracts away need to write M code









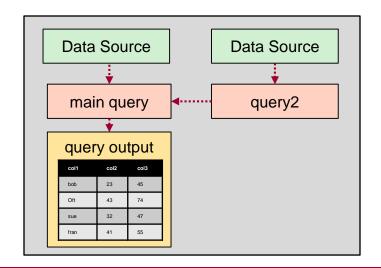
Agenda

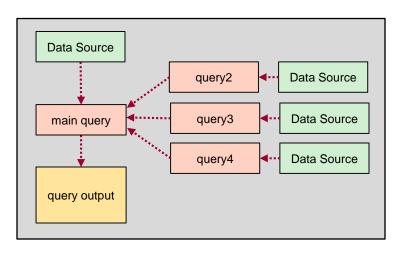
- ✓ Deciding What To Measure
- ✓ Understanding Queries in Power BI Desktop
- ✓ Working with the Query Editor Window
- Designing Advanced Combine Queries
- Importing OLTP Data Into a Star Schema



Combining Queries

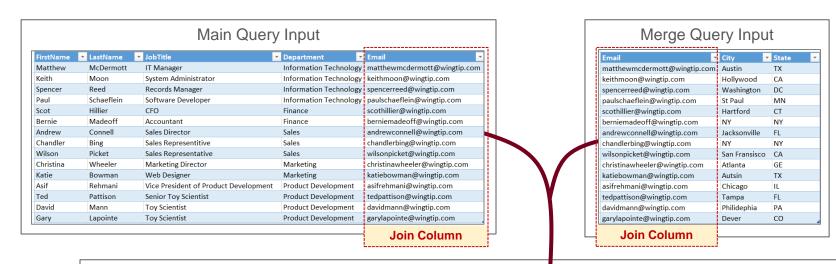
- Query can be merged or appended with another query
 - Merge operation allows you combine columns from two tables
 - Append operation allows you to combine rows from two tables
- Two queries are combined into single output for loading
 - Load settings of main query determines where output is loaded
 - Secondary query acts as source for main query
 - Secondary query be can created with connection-only load setting







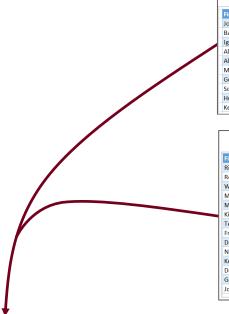
Merging Columns



			Query Output V			
First Name	Last Name	Email	Job Title	Department 🔻	City	State
Matthew	McDermott	matthewmcdermott@wingtip.com	IT Manager	Information Technology	Austin	TX
Keith	Moon	keithmoon@wingtip.com	System Administrator	Information Technology	Hollywood	CA
Spencer	Reed	spencerreed@wingtip.com	Records Manager	Information Technology	Washington	DC
Paul	Schaeflein	paulschaeflein@wingtip.com	Software Developer	Information Technology	St Paul	MN
Scot	Hillier	scothillier@wingtip.com	CFO	Finance	Hartford	CT
Bernie	Madeoff	berniemadeoff@wingtip.com	Accountant	Finance	NY	NY
Andrew	Connell	andrewconnell@wingtip.com	Sales Director	Sales	Jacksonville	FL
Chandler	Bing	chandlerbing@wingtip.com	Sales Representitive	Sales	NY	NY
Wilson	Picket	wilsonpicket@wingtip.com	Sales Representative	Sales	San Fransisco	CA
Christina	Wheeler	christinawheeler@wingtip.com	Marketing Director	Marketing	Atlanta	GE
Katie	Bowman	katiebowman@wingtip.com	Web Designer	Marketing	Autsin	TX
Asif	Rehmani	asifrehmani@wingtip.com	Vice President of Product Development	Product Development	Chicago	IL
Ted	Pattison	tedpattison@wingtip.com	Senior Toy Scientist	Product Development	Tampa	FL
David	Mann	davidmann@wingtip.com	Toy Scientist	Product Development	Philidephia	PA
Gary	Lapointe	garylapointe@wingtip.com	Toy Scientist	Product Development	Dever	со



Appending Rows



	Main Query Input							
FirstName	LastName ×	Company	Email Email	▼ City ▼	State 💌	ZipCode 💌		
Joe	Stephens	Initech	Joe.Stephens@Initech.com	Sacramento	CA	95823		
Basil	Frazier	Nordyne Defense Dynamics	Basil.Frazier@NordyneDefenseDynamics.com	Sacramento	CA	95823		
Ignacio	Duran	Yoyodyne Propulsion Systems	Ignacio.Duran@YoyodynePropulsionSystems.com	Beaverton	OR	97005		
Alvaro	Brock	W.C. Boggs & Co.	Alvaro.Brock@W.C.Boggs&Cocom	Beaverton	OR	97005		
Alec	Miller	Krusty Burger	Alec.Miller@KrustyBurger.com	Vancouver	WA	98662		
Maureen	Griffin	VersaLife Corporation	Maureen.Griffin@VersaLifeCorporation.com	Vancouver	WA	98662		
Guillermo	Sykes	Soar Airlines	Guillermo.Sykes@SoarAirlines.com	Sacramento	CA	95818		
Solomon	Warner	Wayne Enterprises	Solomon.Warner@WayneEnterprises.com	Seattle	WA	98125		
Humberto	Petersen	W.C. Boggs & Co.	Humberto.Petersen@W.C.Boggs&Cocom	San Francisco	CA	94118		
Keven	Griffith	Fabrikam	Keven.Griffith@Fabrikam.com	Bend	OR	97701		

	Append Query Input								
FirstName -	LastName 💌	Company	Email	City	State	ZipCode 🔽			
Rich	Pierce	The Regal Beagle	Rich.Pierce@TheRegalBeagle.com	Portland	OR	97216			
Ronnie	Donaldson	Union Aerospace Corporation	Ronnie.Donaldson@UnionAerospaceCorporation.com	Vancouver	WA	98684			
Willard	Frazier	Benthic Petroleum	Willard.Frazier@BenthicPetroleum.com	Portland	OR	97220			
Marina	Caldwell	Ewing Oil	Marina.Caldwell@EwingOil.com	Portland	OR	97205			
Milagros	Mercer	Hishii Industries	Milagros.Mercer@HishiiIndustries.com	Salem	OR	97301			
Kirsten	Little	Binford	Kirsten.Little@Binford.com	Ventura	CA	93003			
Terri	Ferrell	Black Mesa Research Facility	Terri.Ferrell@BlackMesaResearchFacility.com	Portland	OR	97205			
Francine	Doyle	Peach Pit	Francine.Doyle@PeachPit.com	Vancouver	WA	98662			
Dannie	Powers	Deon International	Dannie.Powers@DeonInternational.com	Issaquah	WA	98027			
Noah	Best	Shinra Electric Power Company	Noah.Best@ShinraElectricPowerCompany.com	Portland	OR	97217			
Keenan	Holmes	Biffco	Keenan.Holmes@Biffco.com	Portland	OR	97220			
Douglas	Maynard	Cyberdyne Systems	Douglas.Maynard@CyberdyneSystems.com	San Francisco	CA	94118			
Gerald	Harrington	Yoyodyne Propulsion Systems	Gerald.Harrington@YoyodynePropulsionSystems.com	Portland	OR	97217			
Josue	Robinson	North Western Railway	Josue.Robinson@NorthWesternRailway.com	Beaverton	OR	97005			

			Query Output			
FirstName	▼ LastName	Company	Email	City	State -	ZipCode
Joe	Stephens	Initech	Joe.Stephens@Initech.com	Sacramento	CA	95823
Basil	Frazier	Nordyne Defense Dynamics	Basil.Frazier@NordyneDefenseDynamics.com	Sacramento	CA	95823
Ignacio	Duran	Yoyodyne Propulsion Systems	Ignacio.Duran@YoyodynePropulsionSystems.com	Beaverton	OR	97005
Alvaro	Brock	W.C. Boggs & Co.	Alvaro.Brock@W.C.Boggs&Cocom	Beaverton	OR	97005
Alec	Miller	Krusty Burger	Alec.Miller@KrustyBurger.com	Vancouver	WA	98662
Maureen	Griffin	VersaLife Corporation	Maureen.Griffin@VersaLifeCorporation.com	Vancouver	WA	98662
Guillermo	Sykes	Soar Airlines	Guillermo.Sykes@SoarAirlines.com	Sacramento	CA	95818
Solomon	Warner	Wayne Enterprises	Solomon.Warner@WayneEnterprises.com	Seattle	WA	98125
Humberto	Petersen	W.C. Boggs & Co.	Humberto.Petersen@W.C.Boggs&Cocom	San Francisco	CA	94118
Keven	Griffith	Fabrikam	Keven.Griffith@Fabrikam.com	Bend	OR	97701
Rich	Pierce	The Regal Beagle	Rich.Pierce@TheRegalBeagle.com	Portland	OR	97216
Ronnie	Donaldson	Union Aerospace Corporation	Ronnie.Donaldson@UnionAerospaceCorporation.com	Vancouver	WA	98684
Willard	Frazier	Benthic Petroleum	Willard.Frazier@BenthicPetroleum.com	Portland	OR	97220
Marina	Caldwell	Ewing Oil	Marina.Caldwell@EwingOil.com	Portland	OR	97205
Milagros	Mercer	Hishii Industries	Milagros.Mercer@HishiiIndustries.com	Salem	OR	97301
Kirsten	Little	Binford	Kirsten.Little@Binford.com	Ventura	CA	93003
Terri	Ferrell	Black Mesa Research Facility	Terri.Ferrell@BlackMesaResearchFacility.com	Portland	OR	97205
Francine	Doyle	Peach Pit	Francine.Doyle@PeachPit.com	Vancouver	WA	98662
Dannie	Powers	Deon International	Dannie.Powers@DeonInternational.com	Issaquah	WA	98027
Noah	Best	Shinra Electric Power Company	Noah.Best@ShinraElectricPowerCompany.com	Portland	OR	97217
Keenan	Holmes	Biffco	Keenan.Holmes@Biffco.com	Portland	OR	97220
Douglas	Maynard	Cyberdyne Systems	Douglas.Maynard@CyberdyneSystems.com	San Francisco	CA	94118
Gerald	Harrington	Yoyodyne Propulsion Systems	Gerald.Harrington@YoyodynePropulsionSystems.com	Portland	OR	97217
Josue	Robinson	North Western Railway	Josue.Robinson@NorthWesternRailway.com	Beaverton	OR	97005



Pivoting Columns

- Pivot column adds its values are new columns
 - Create table layout like PivotTable

City ▼	Year ▼	Population <a>
New York, NY	2010	8175133
New York, NY	2000	8008278
New York, NY	1990	7322564
Los Angeles, CA	2010	3792621
Los Angeles, CA	2000	3694820
Los Angeles, CA	1990	3485398
Chicago, IL	2010	2695598
Chicago, IL	2000	2896016
Chicago, IL	1990	2783726
Houston, TX	2010	2100263
Houston, TX	2000	1953631
Houston, TX	1990	1630553
Philadelphia, PA	2010	1526006
Philadelphia, PA	2000	1517550
Philadelphia, PA	1990	1585577
Phoenix, AZ	2010	1445632
Phoenix, AZ	2000	1321045
Phoenix, AZ	1990	983403





Unpivoting Columns

- Unpivot columns to collapse them into single column
 - Removes PivotTable layout
 - Can be useful to prepare data for charting and analysis

City	•	2010	•	2000	₩	1990	~
New York, NY		8175	133	80	08278	73	22564
Los Angeles, CA		3792	621	36.	94820	34	85398
Chicago, IL		2695	598	28.	96016	27	83726
Houston, TX		2100	263	19	53631	16	30553
Philadelphia, PA		1526	006	15	17550	15	85577
Phoenix, AZ		1445	632	13.	21045	9	83403
San Antonio, TX		1327	407	11	44646	9	35933
San Diego, TX		1307	402	12	23400	11	10549
Dallas, TX		1197	816	11	88580	10	06877
San Jose, CA		945	942	8.	94943	7	82248



City	▼ Year	₩	Population	₩
New York, NY	2010		817	5133
New York, NY	2000		800	8278
New York, NY	1990		732	2564
Los Angeles, CA	2010		379	2621
Los Angeles, CA	2000		369	4820
Los Angeles, CA	1990		348	5398
Chicago, IL	2010		269	5598
Chicago, IL	2000		289	6016
Chicago, IL	1990		278	3726
Houston, TX	2010		210	0263
Houston, TX	2000		195	3631
Houston, TX	1990		163	0553
Philadelphia, PA	2010		152	6006
Philadelphia, PA	2000		151	7550
Philadelphia, PA	1990		158	5577
Phoenix, AZ	2010		144	5632
Phoenix, AZ	2000		132	1045
Phoenix, AZ	1990		98	3403



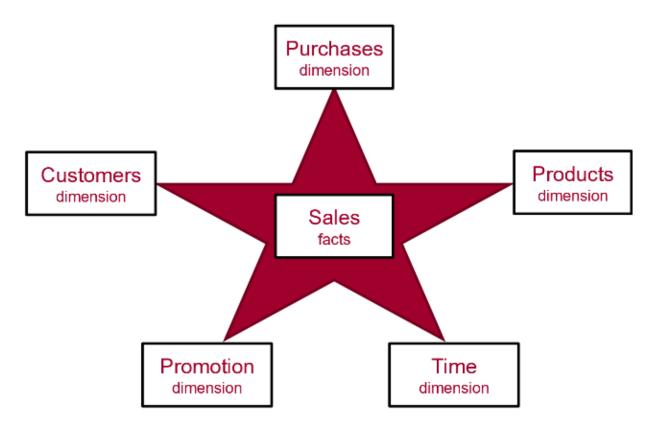
Agenda

- Deciding What To Measure
- ✓ Understanding Queries in Power BI Desktop
- ✓ Working with the Query Editor Window
- ✓ Designing Advanced Combine Queries
- Importing OLTP Data Into a Star Schema



Data Modeling using a Star Schema

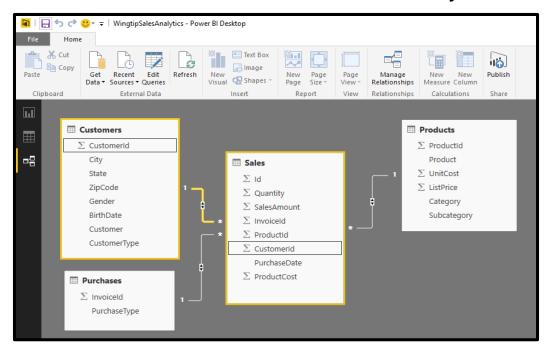
- OLAP Modeling often based on Star Schema
 - Tables defined as fact tables or dimension tables
 - Fact tables related to dimension table using 1-to-many relationships





Designing Queries to Build a Star Schema

- Converts OLTP Data Model to OLAP Data Model
 - Sales table is modeled as a OLAP Fact Table
 - Other tables are modeled as OLAP Dimension tables
 - Requires pulling CustomerId column into Sales table
 - All dimension tables should be directly related to fact table







Summary

- Deciding What To Measure
- ✓ Understanding Queries in Power BI Desktop
- ✓ Working with the Query Editor Window
- ✓ Designing Advanced Combine Queries
- ✓ Importing OLTP Data Into a Star Schema

