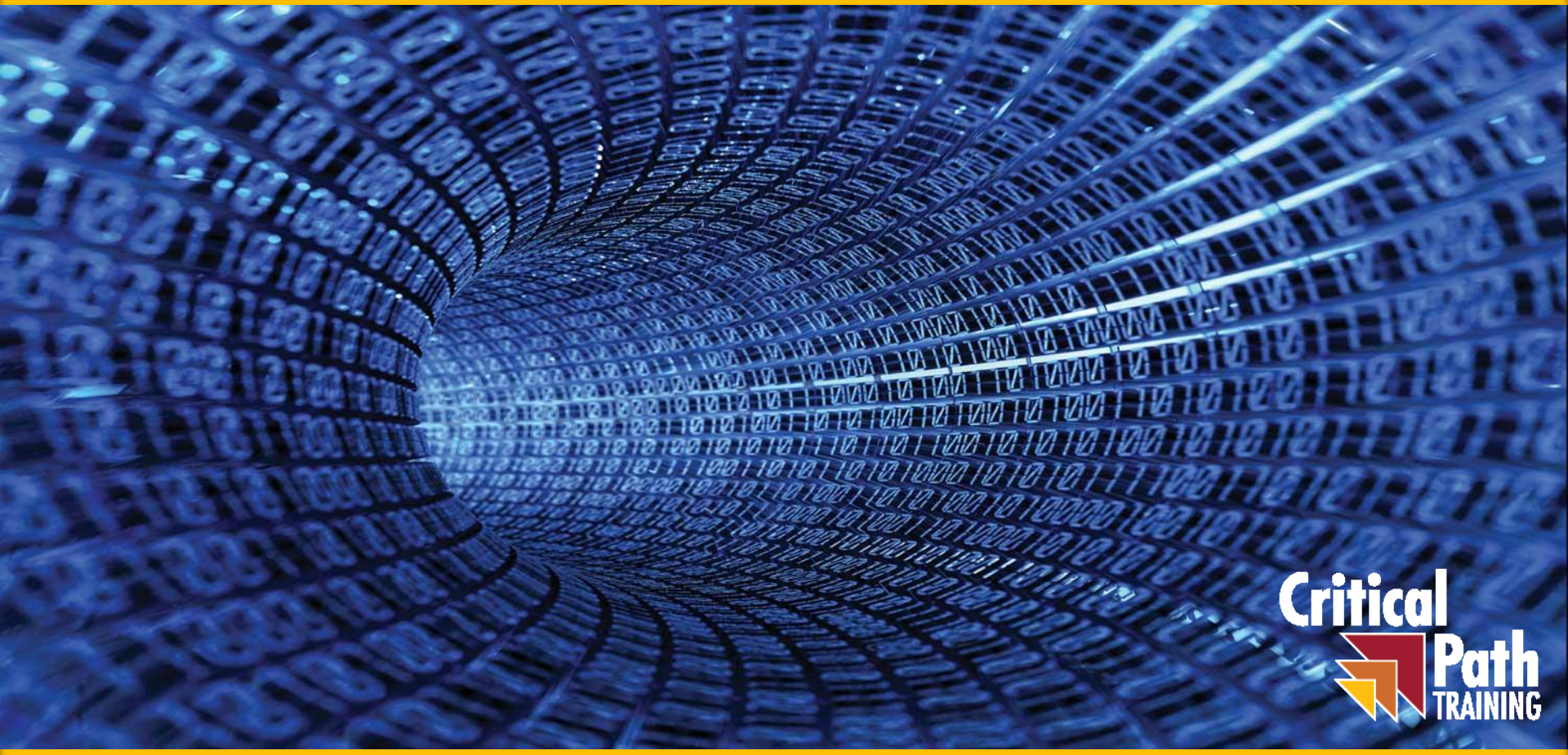
Mastering the Query Features of Power BI Desktop



Agenda

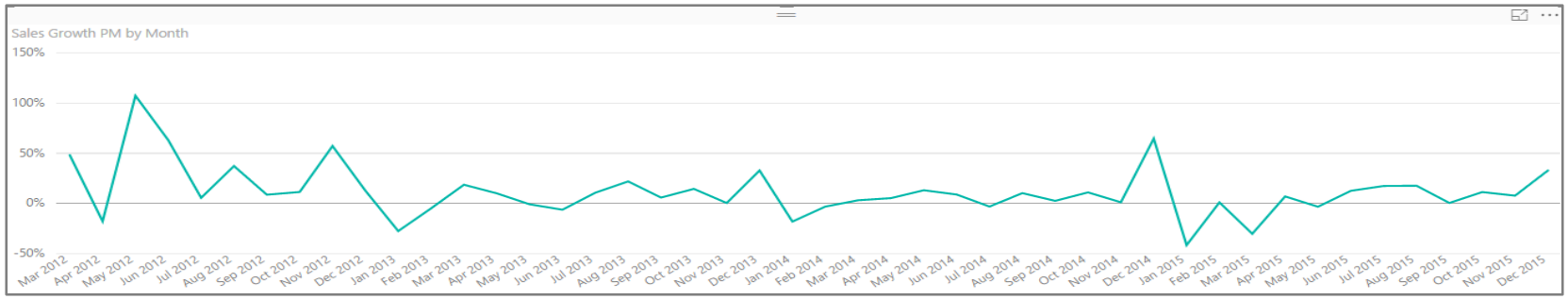
- Deciding What To Measure
- Managing Queries, Datasources and Credentials
- Working with the Query Editor Window
- Designing Queries to Generate a Star Schema
- Merging Columns from Multiple Datasources
- Appending Rows from Multiple Datasources

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

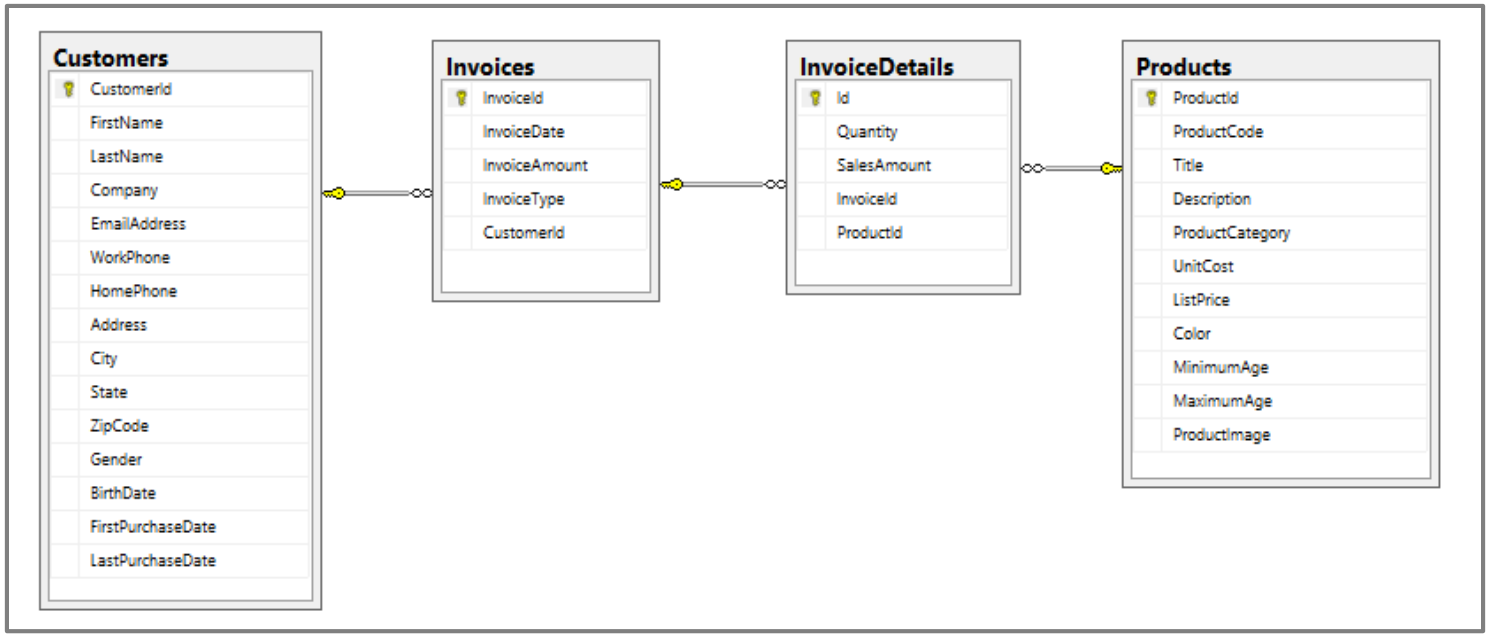


Defining Grain Statements

- 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?

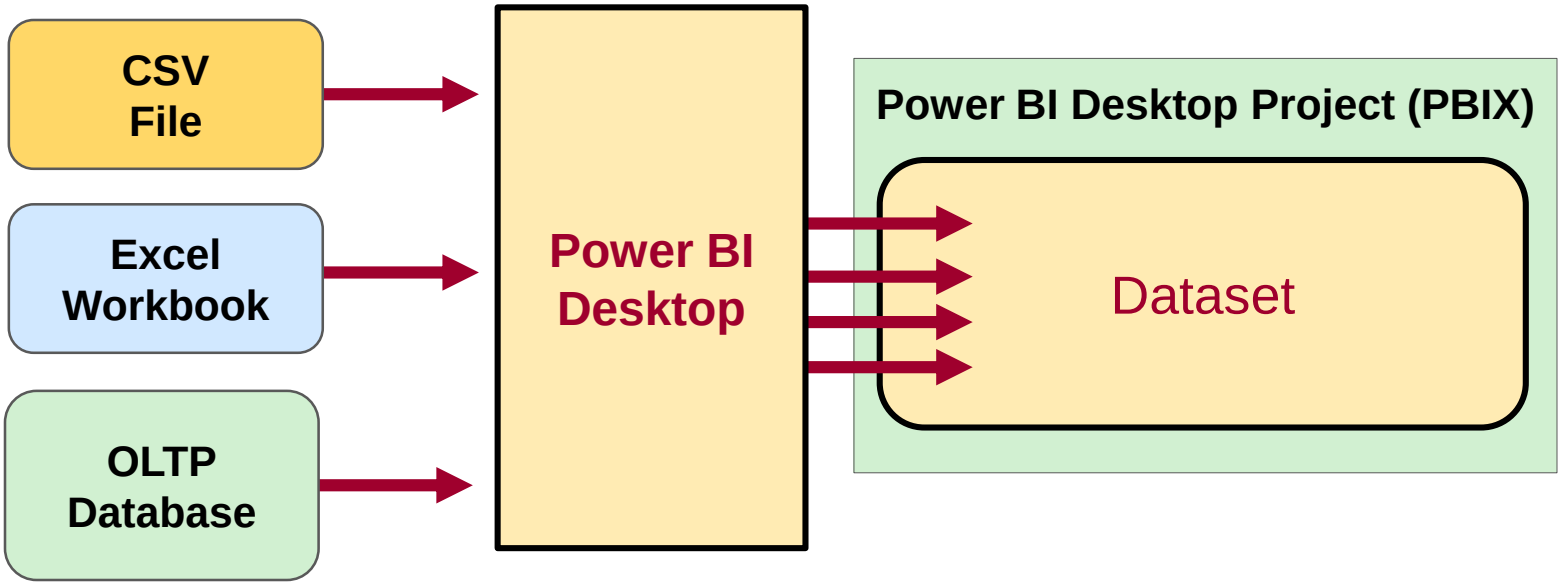
Sample OLTP Database: WingtipSalesDB

- 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



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

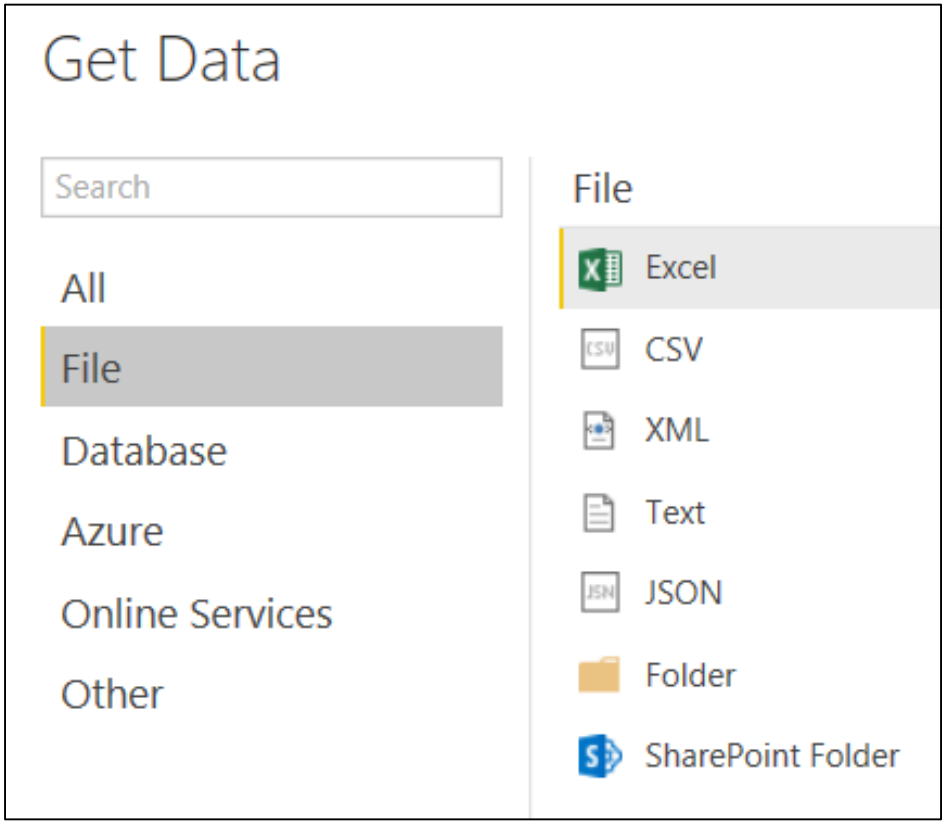


Agenda

- ✓ Deciding What To Measure
- ➢ Managing Queries, Datasources and Credentials
- Working with the Query Editor Window
- Designing Queries to Generate a Star Schema
- Merging Columns from Multiple Datasources
- Appending Rows from Multiple Datasources

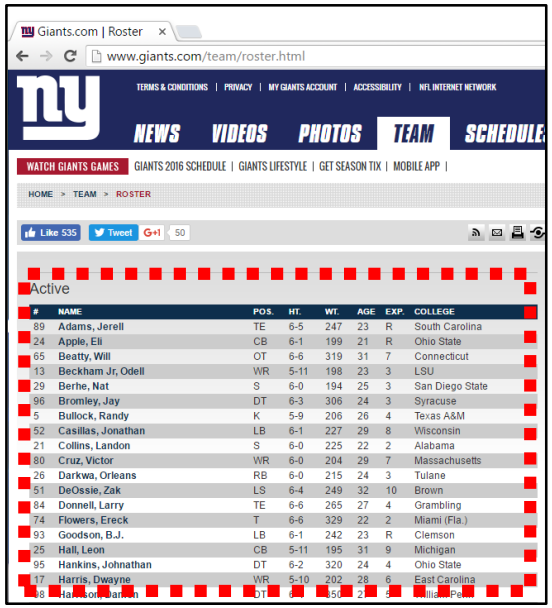
File-based Data Sources

- Power BI Desktop supports common file types



Working with Web Data Sources

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



Query Input

	#	Name	Pos.	Ht.	Wt.	Age	Exp.	College
1	89	Adams, Jerell	TE	6-5	247	23	R	South Carolina
2	24	Apple, Eli	CB	6-1	199	21	R	Ohio State
3	65	Beatty, Will	OT	6-6	319	31	7	Connecticut
4	13	Beckham Jr, Odell	WR	5-11	198	23	3	LSU
5	29	Berhe, Nat	S	6-0	194	25	3	San Diego State
6	96	Bromley, Jay	DT	6-3	306	24	3	Syracuse

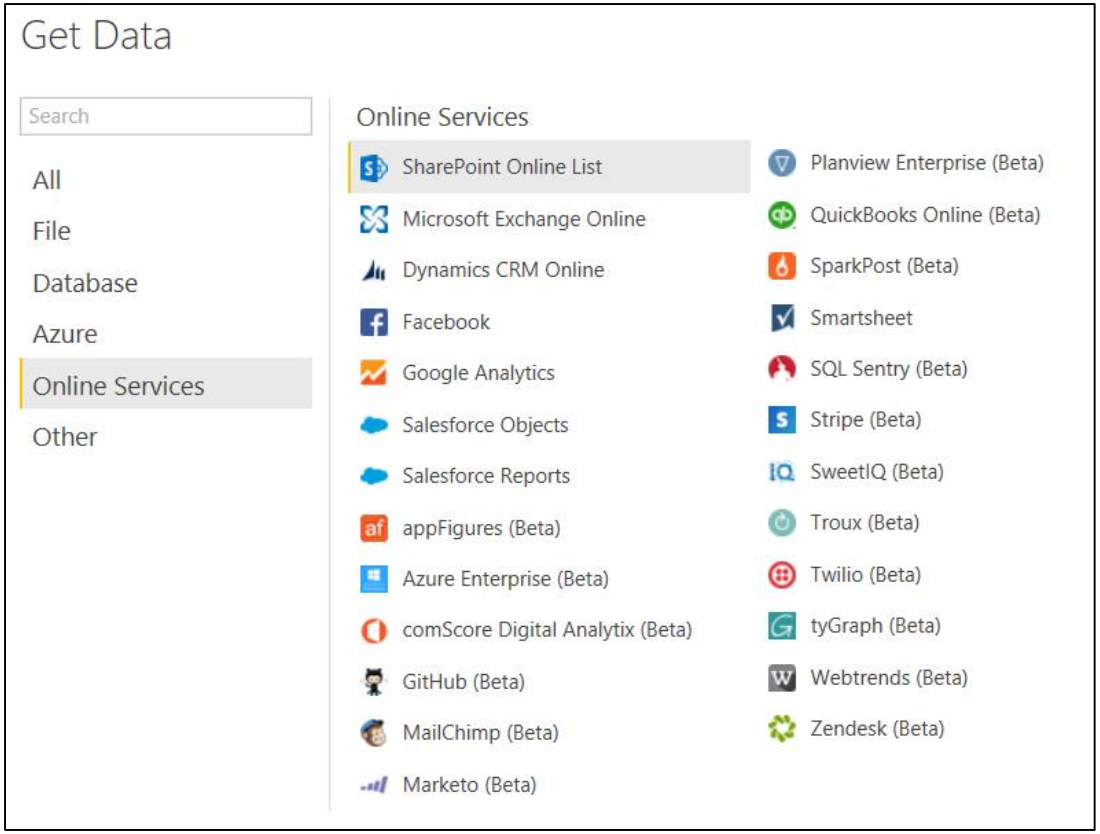
Query Output

	Number	Last Name	First Name	Weight	Height	Age	Experience	Position	Category	Side	College
1	89	Adams	Jerell	247	77	23	0	Tight End	Backs and Receivers	Offense	South Carolina
2	84	Donnell	Larry	265	78	27	4	Tight End	Backs and Receivers	Offense	Grambling
3	45	Tye	Will	262	74	24	1	Tight End	Backs and Receivers	Offense	Stony Brook
4	24	Apple	Eli	199	73	21	0	Cornerback	Defensive Backs	Defense	Ohio State
5	25	Hall	Leon	195	71	31	9	Cornerback	Defensive Backs	Defense	Michigan
6	20	Jenkins	Janoris	198	70	27	5	Cornerback	Defensive Backs	Defense	North Alabama
7	41	Rodgers-Cromartie	Dominique	205	74	30	8	Cornerback	Defensive Backs	Defense	Tennessee State
8	65	Beatty	Will	319	78	31	7	Offensive Tackle	Offensive Line	Offense	Connecticut
9	13	Beckham Jr	Odell	198	71	23	3	Wide Receiver	Backs and Receivers	Offense	LSU
10	80	Cruz	Victor	204	72	29	7	Wide Receiver	Backs and Receivers	Offense	Massachusetts


DEMO
Scraping Data from a Web Page

Online Service Data Sources

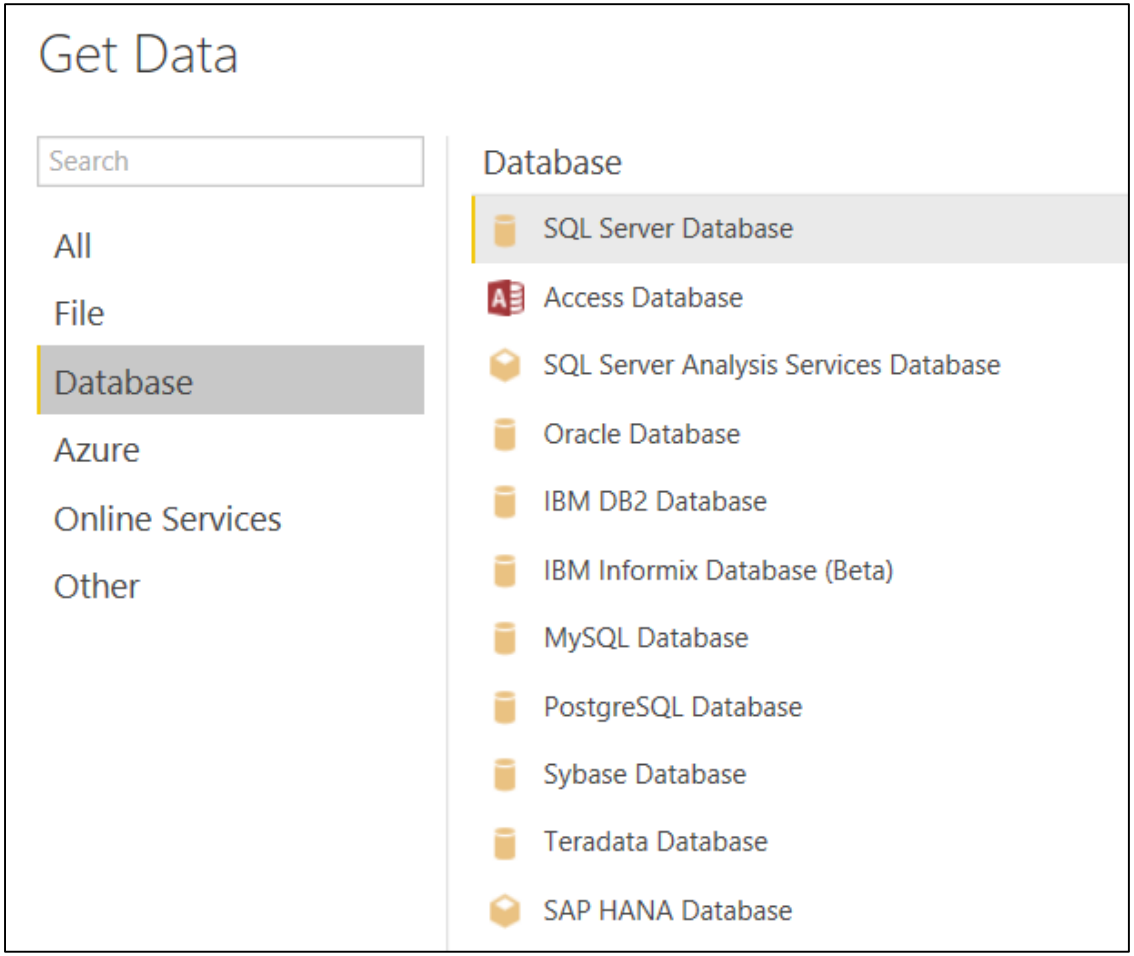
- Power BI Desktop Supports Online Services
  - Includes popular Software-as-a-Service (SaaS) applications
  - Microsoft is working with 3rd vendors to expand this list




DEMO
Importing Data from a List in SharePoint Online

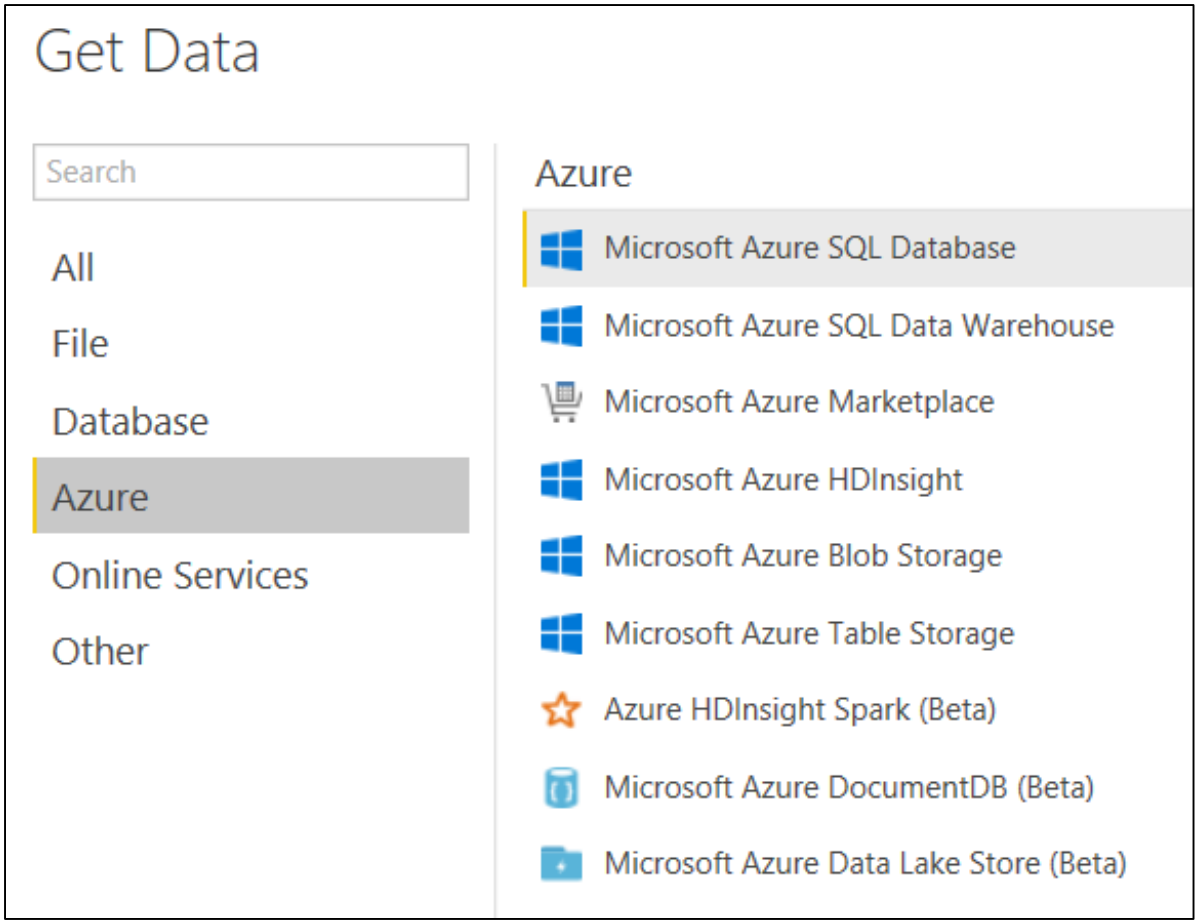
Supported Databases

- Power BI Desktop supports many database systems



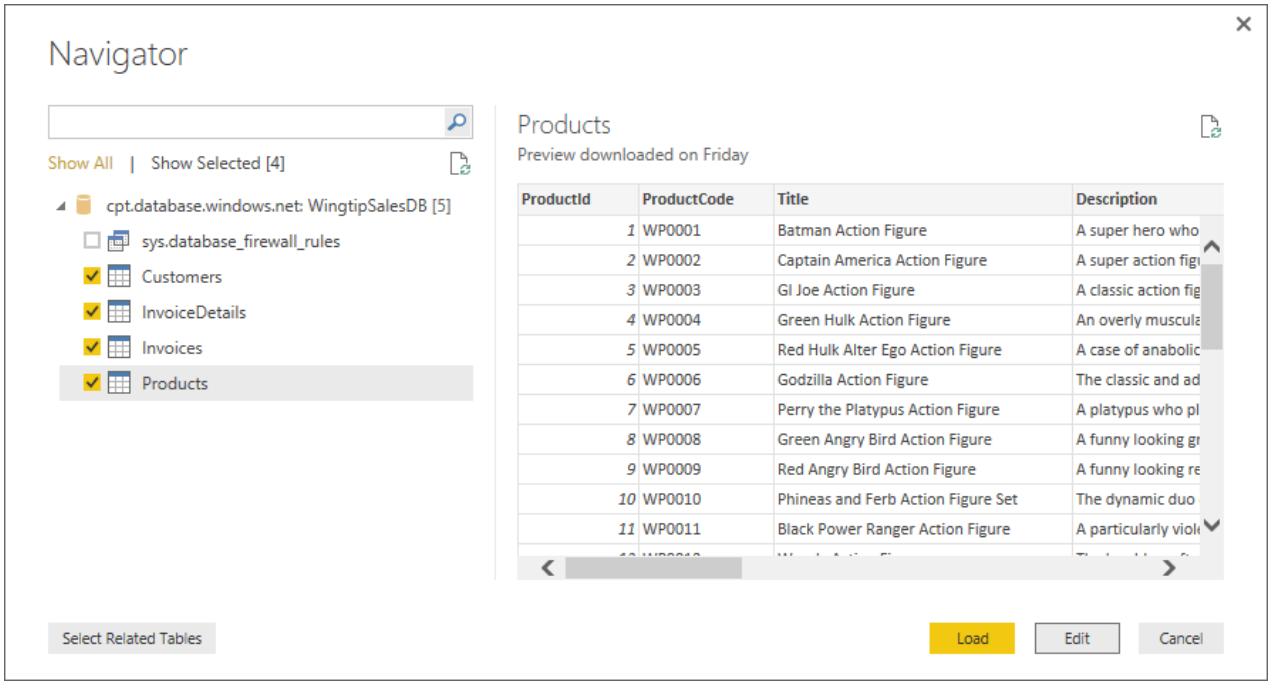
Azure Data Sources

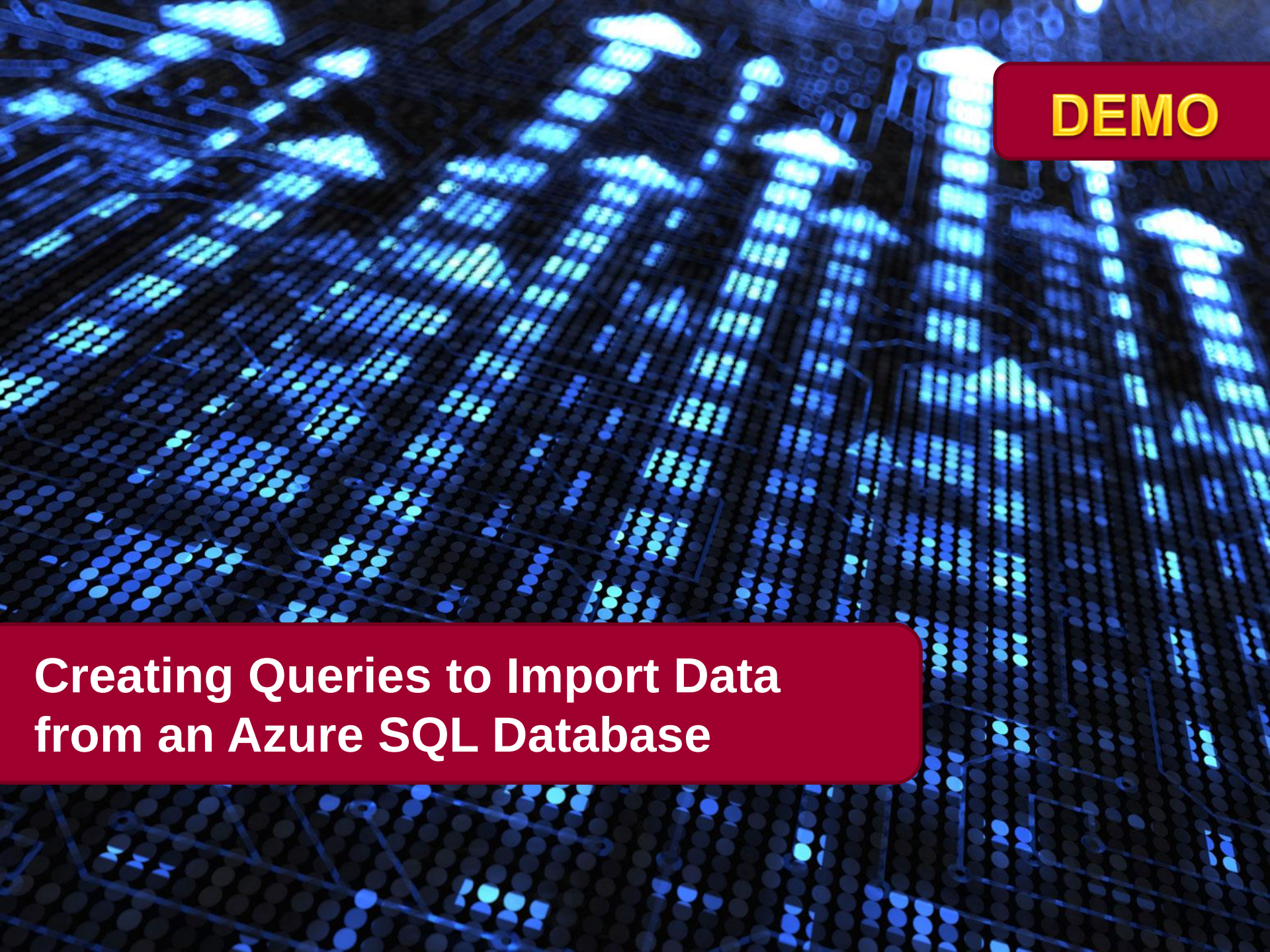
- Power BI Desktop supports many Azure data sources



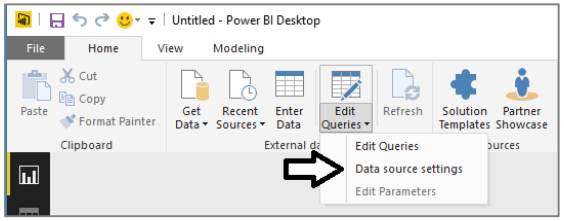
Selecting Tables from a SQL Database

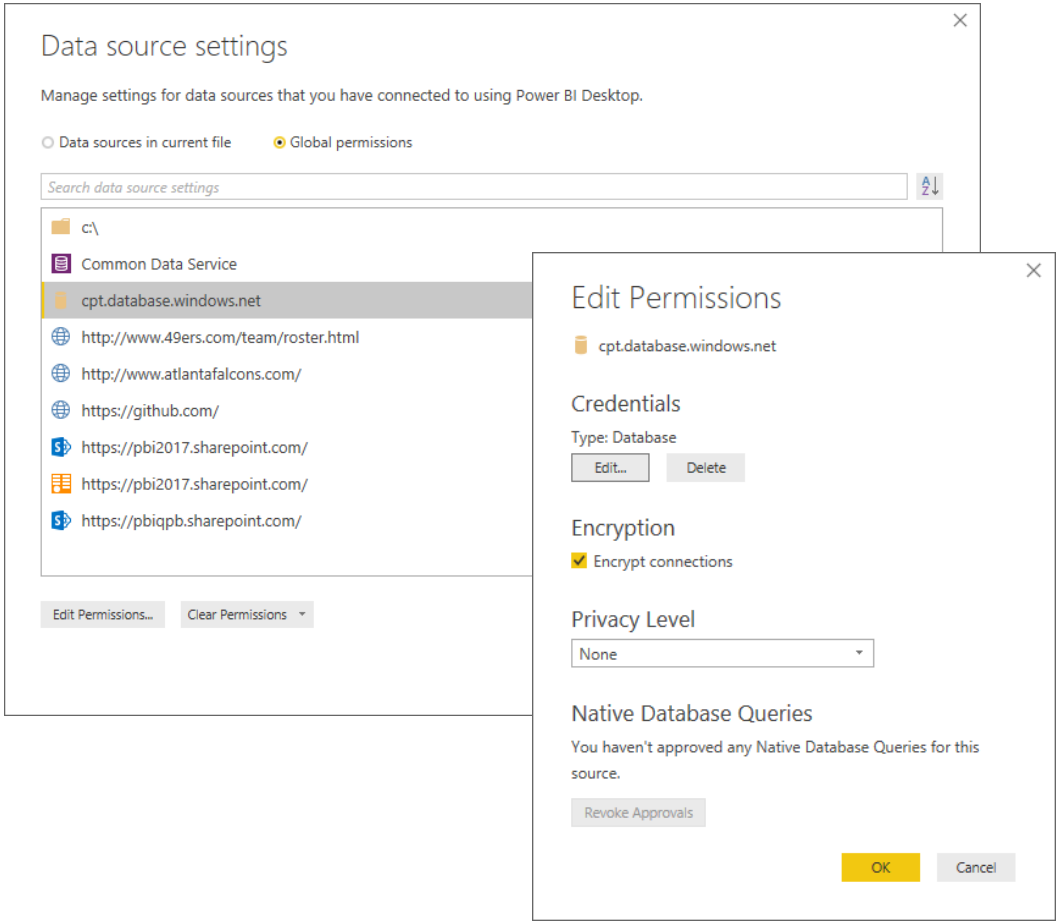
- Power BI Desktop provides Navigator dialog
  - Allows you to select tables
  - Navigator understands existing table relationships
  - Clicking **Load** will run query and import data
  - Clicking **Edit** will open queries in Query Editor window




DEMO
Creating Queries to Import Data
from an Azure SQL Database

Managing Datasources and Credentials



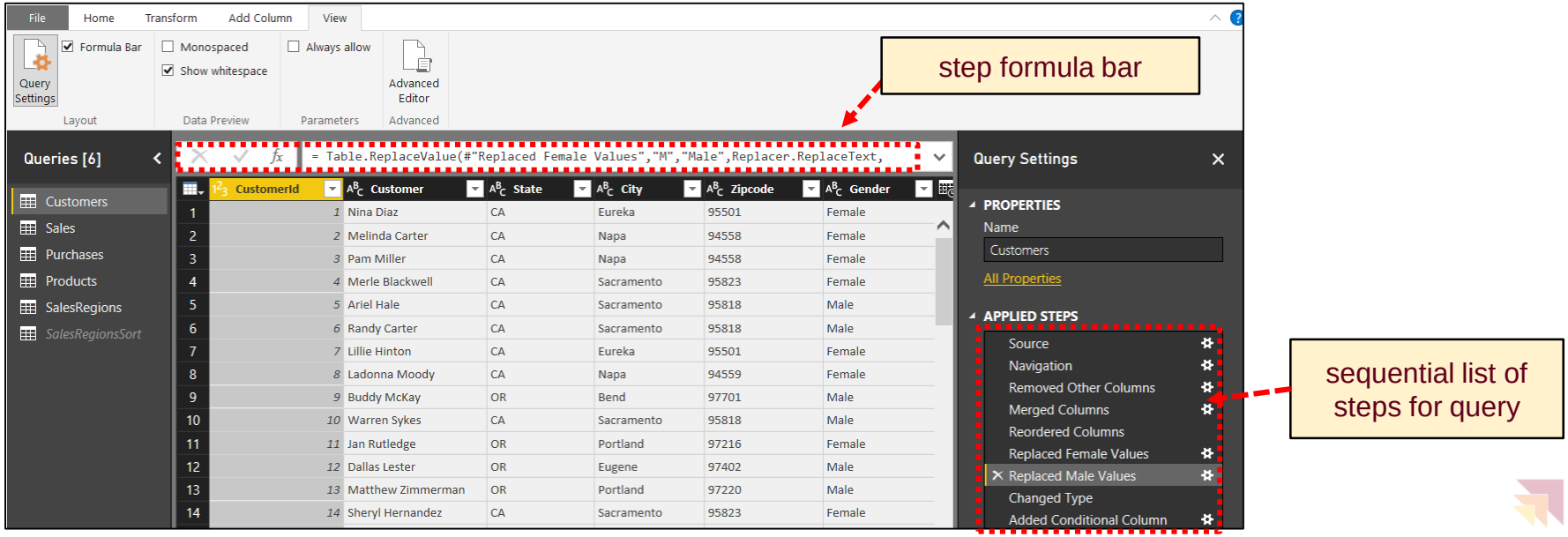


Agenda

- ✓ Deciding What To Measure
- ✓ Managing Queries, Datasources and Credentials
- ➢ Working with the Query Editor Window
- Designing Queries to Generate a Star Schema
- Merging Columns from Multiple Datasources
- Appending Rows from Multiple Datasources

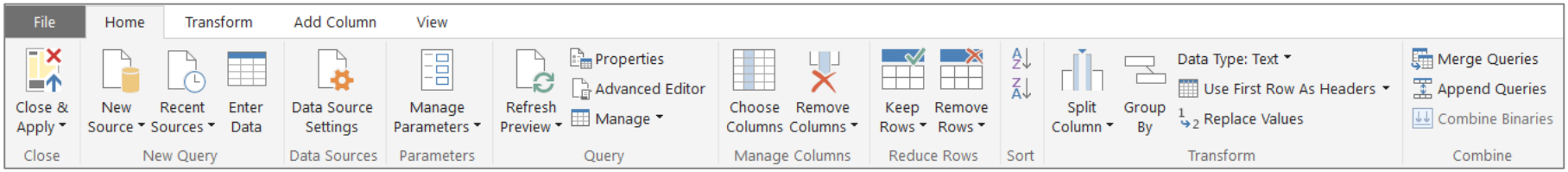
Query Steps

- 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

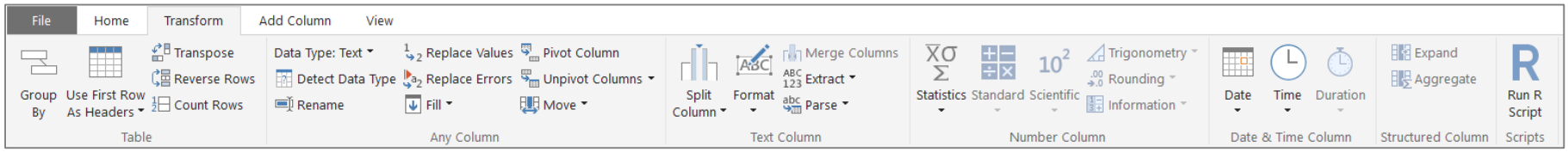


Query Editor Ribbon Tabs

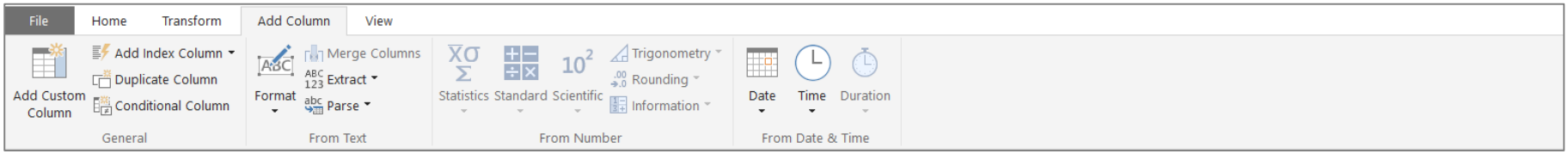
Home tab



Transform tab



Add Column tab



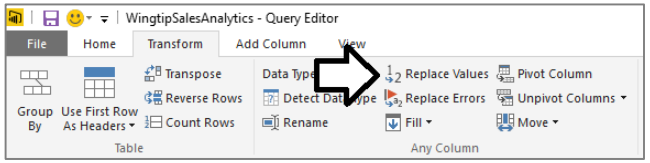
View tab

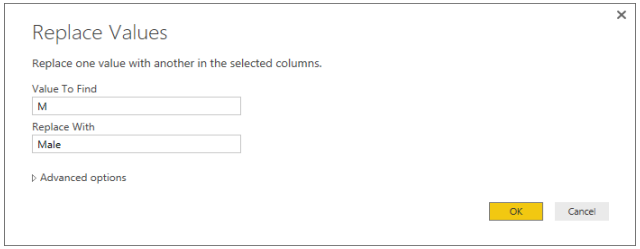
Examples of Basic Power BI Desktop Steps

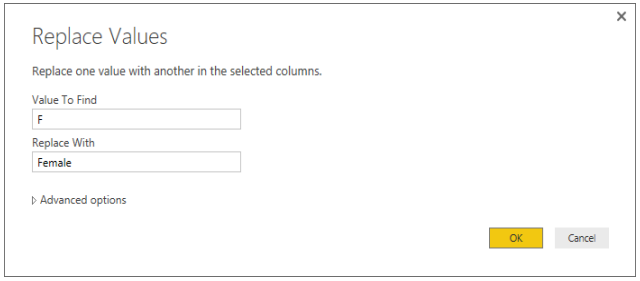
- Rename column
- Convert column type
- Format column values
- Reorder columns
- Replace column values
- Expanding related column
- Merging columns
- Splitting columns

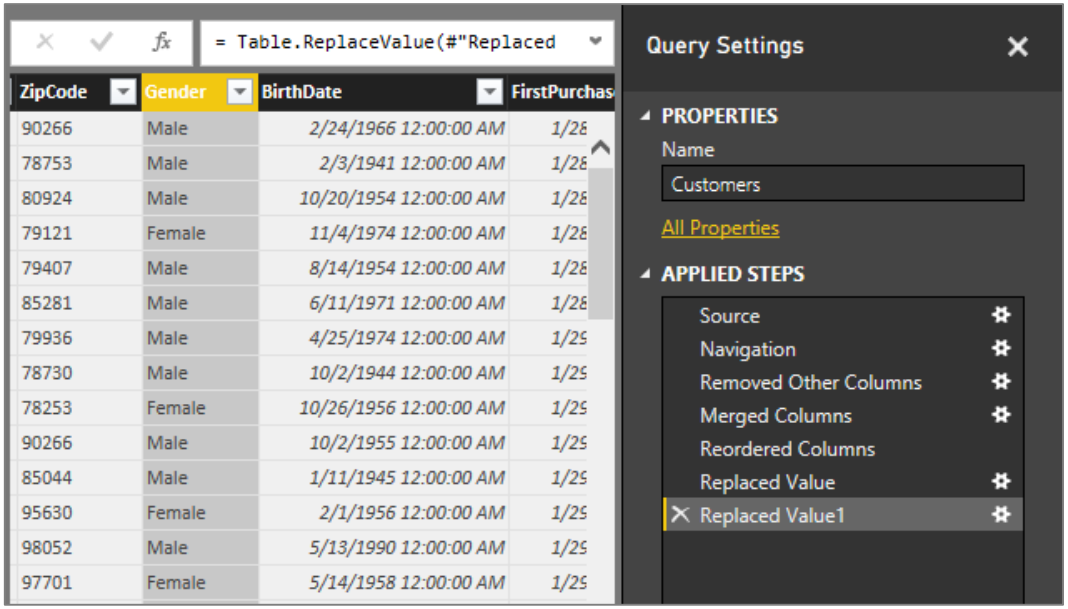
Replacing Values

- Used to substitute values during import



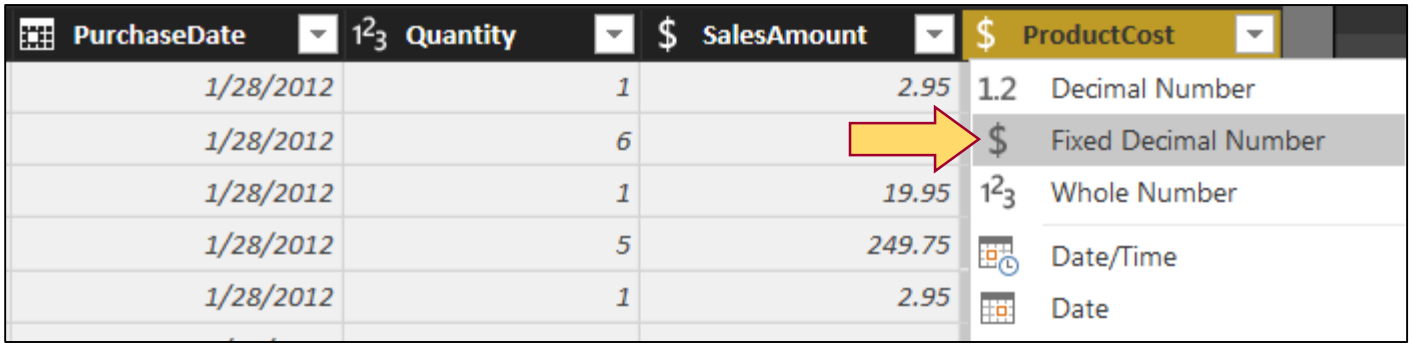






Converting Column Types

- 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

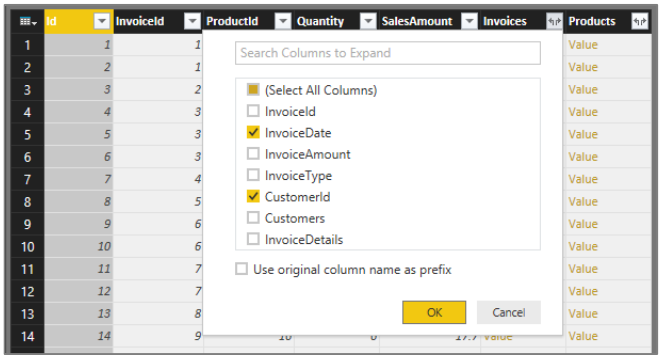


- Beware: Conversion can have destructive effect on data

Expanding Related Columns

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

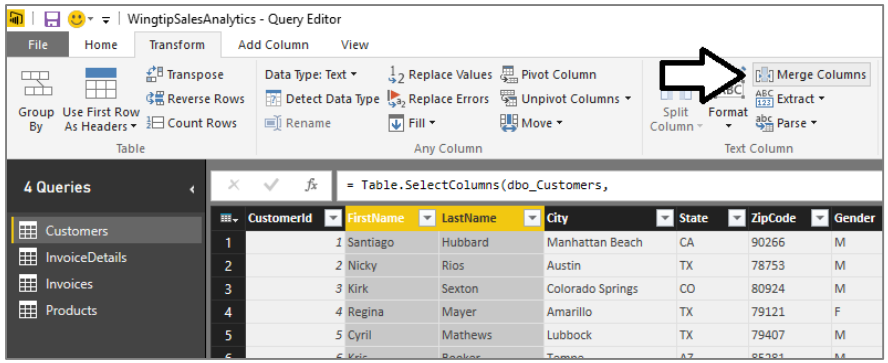
SalesAmount	Invoices	
119.8	Value	Value
29.95	Value	Value
59.9	Value	Value
399.6	Value	Value
29.9	Value	Value
59.8	Value	Value

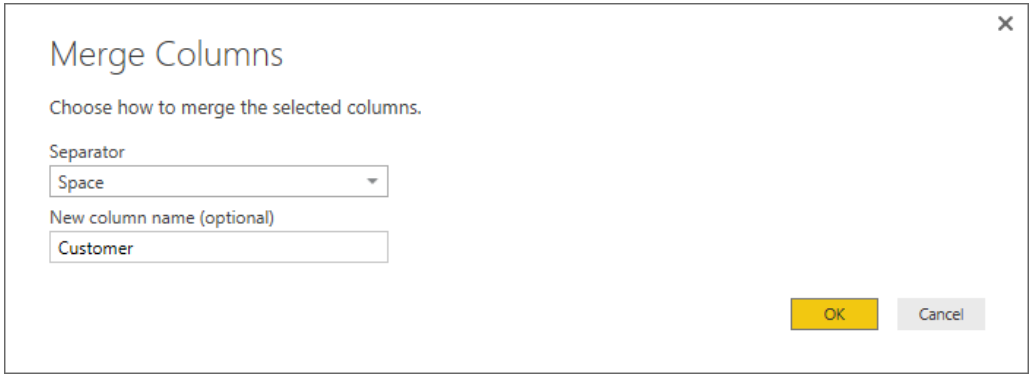


Id	InvoiceId	ProductId	Quantity	SalesAmount	InvoiceDate	CustomerId	Products
1	1	22	4	119.8	1/28/2012 12:00:00 AM	1	Value
2	1	22	1	29.95	1/28/2012 12:00:00 AM	1	Value
3	2	22	2	59.9	1/28/2012 12:00:00 AM	2	Value
4	3	17	8	399.6	1/28/2012 12:00:00 AM	3	Value
5	3	18	2	29.9	1/28/2012 12:00:00 AM	3	Value
6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

Merging Columns

- Merge two columns into a single column

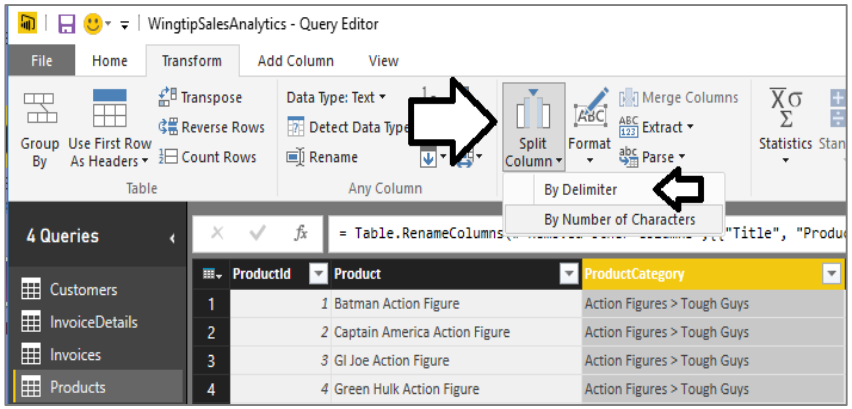


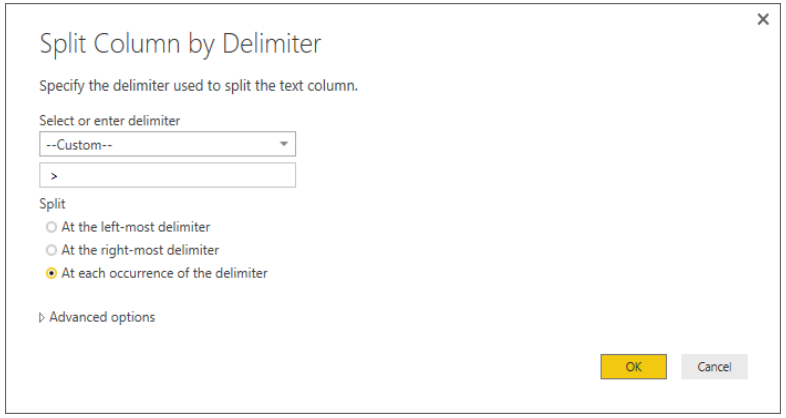


	CustomerId	Customer
1	1	Santiago Hubbard
2	2	Nicky Rios
3	3	Kirk Sexton
4	4	Regina Mayer
5	5	Cyril Mathews
6	6	Kris Booker
7	7	Tracy Christensen
8	8	Reed Glover

Splitting Columns

- Split a single column up into two columns

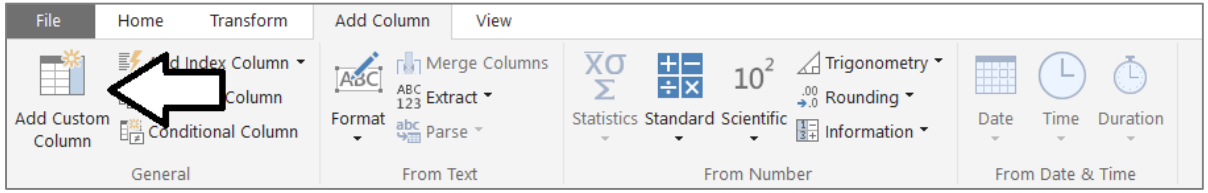


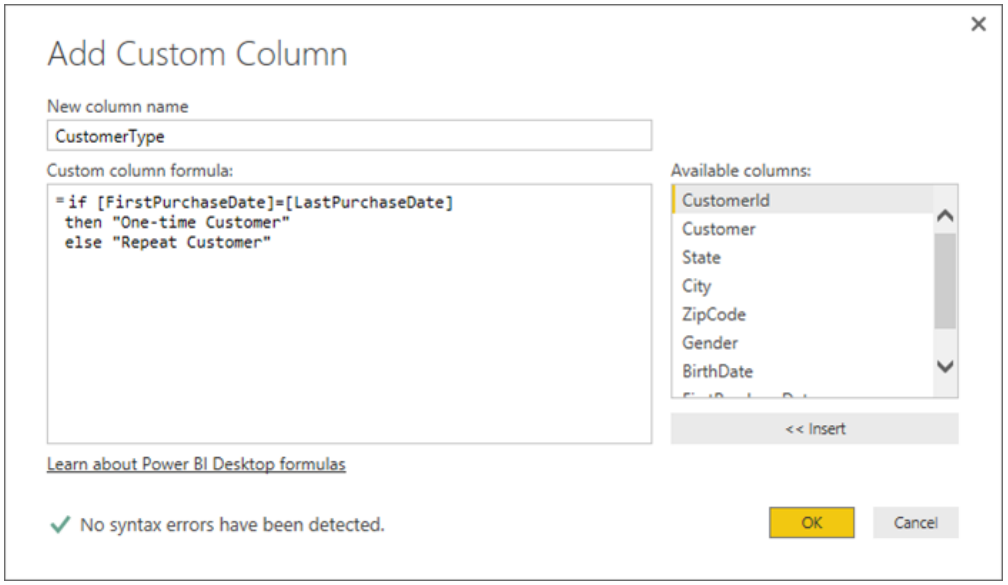


	ProductId	Product	Category	Subcategory
1	1	Batman Action Figure	Action Figures	Tough Guys
2	2	Captain America Action Figure	Action Figures	Tough Guys
3	3	GI Joe Action Figure	Action Figures	Tough Guys
4	4	Green Hulk Action Figure	Action Figures	Tough Guys
5	5	Red Hulk Alter Ego Action Figure	Action Figures	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

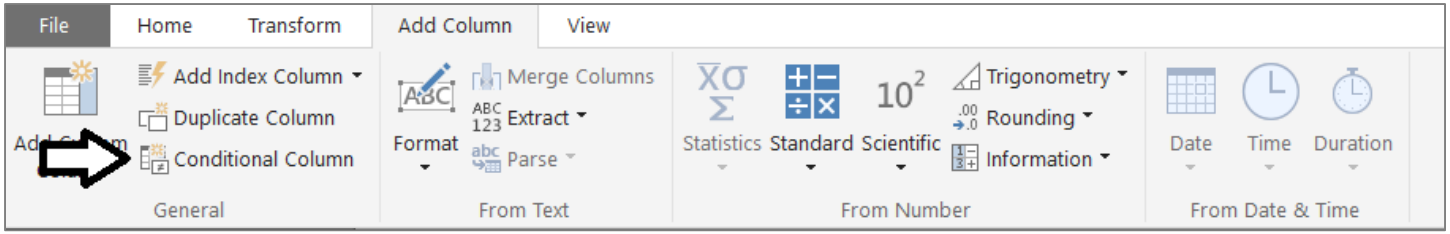


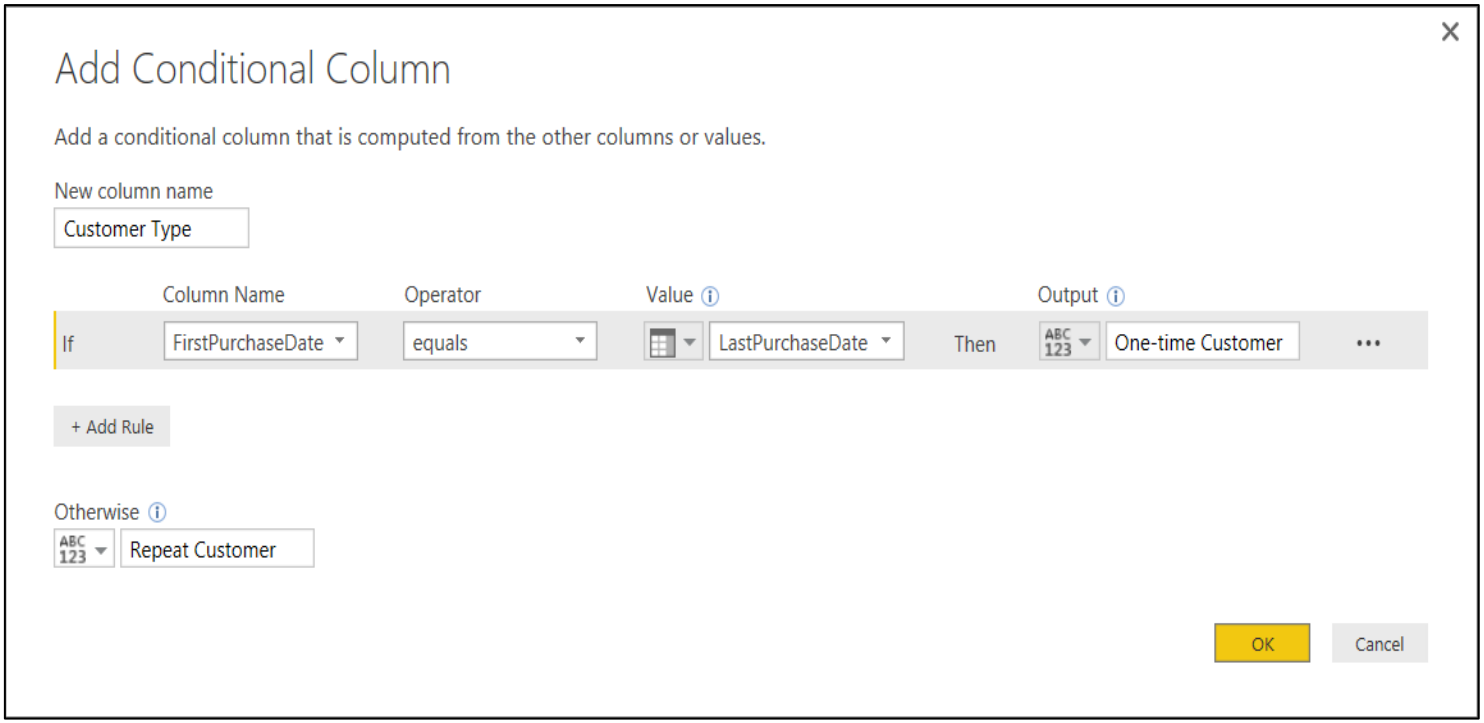


FirstPurchaseDate	LastPurchaseDate	CustomerType
1/28/2012	1/28/2012	One-time Customer
1/28/2012	1/28/2012	One-time Customer
1/28/2012	1/28/2012	One-time Customer
1/28/2012	1/28/2012	One-time Customer
1/28/2012	1/28/2012	One-time Customer
1/28/2012	1/28/2012	One-time Customer
1/29/2012	11/22/2015	Repeat 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

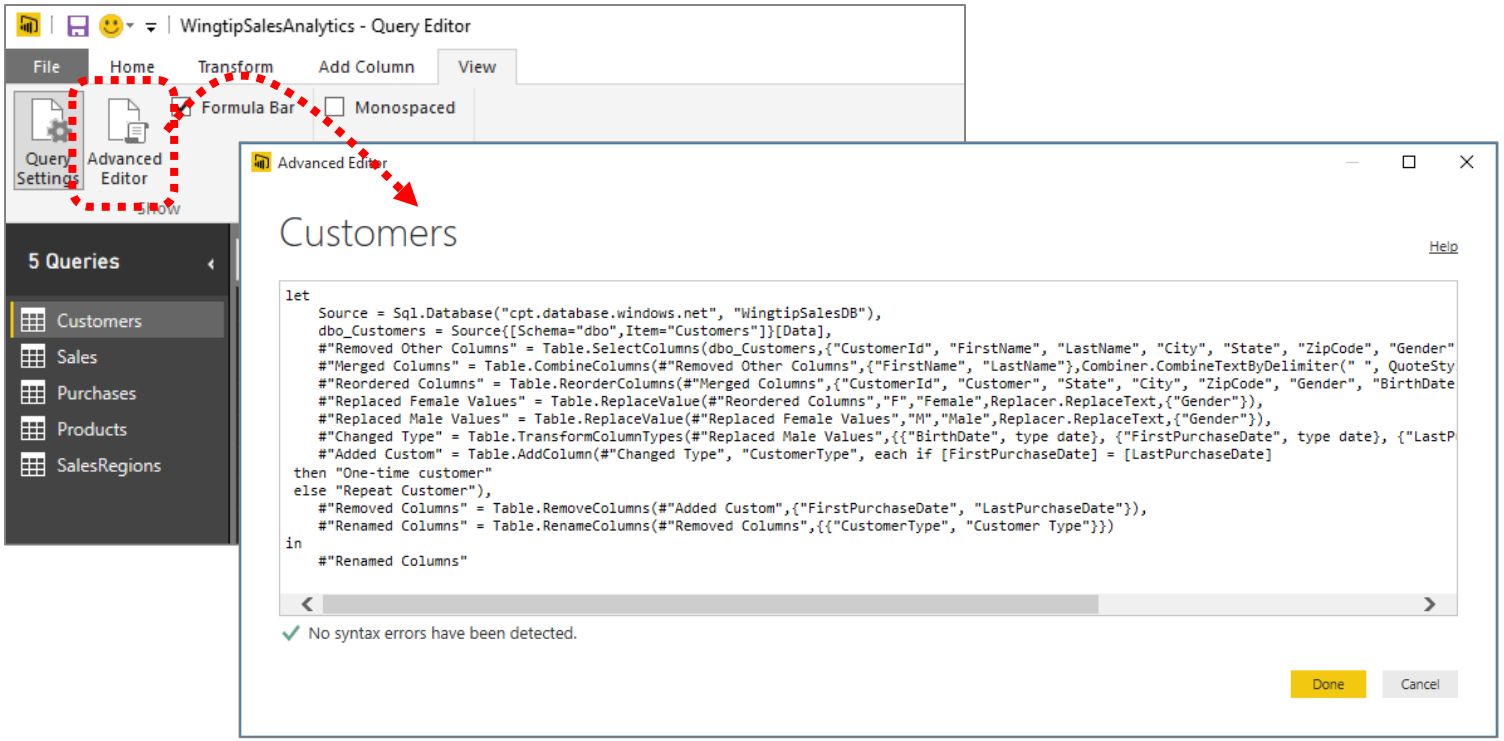
- Abstracts away need to write M code





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

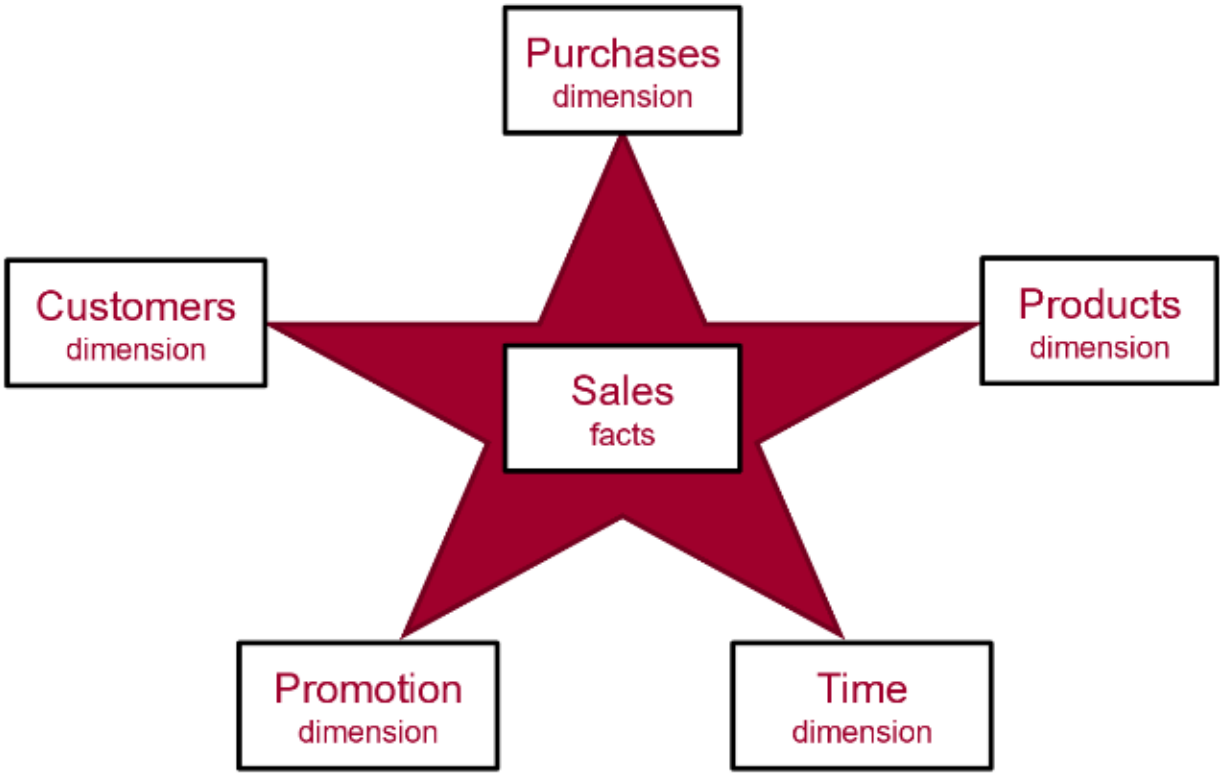


Agenda

- ✓ Deciding What To Measure
- ✓ Managing Queries, Datasources and Credentials
- ✓ Working with the Query Editor Window
- ➢ Designing Queries to Generate a Star Schema
- Merging Columns from Multiple Datasources
- Appending Rows from Multiple Datasources

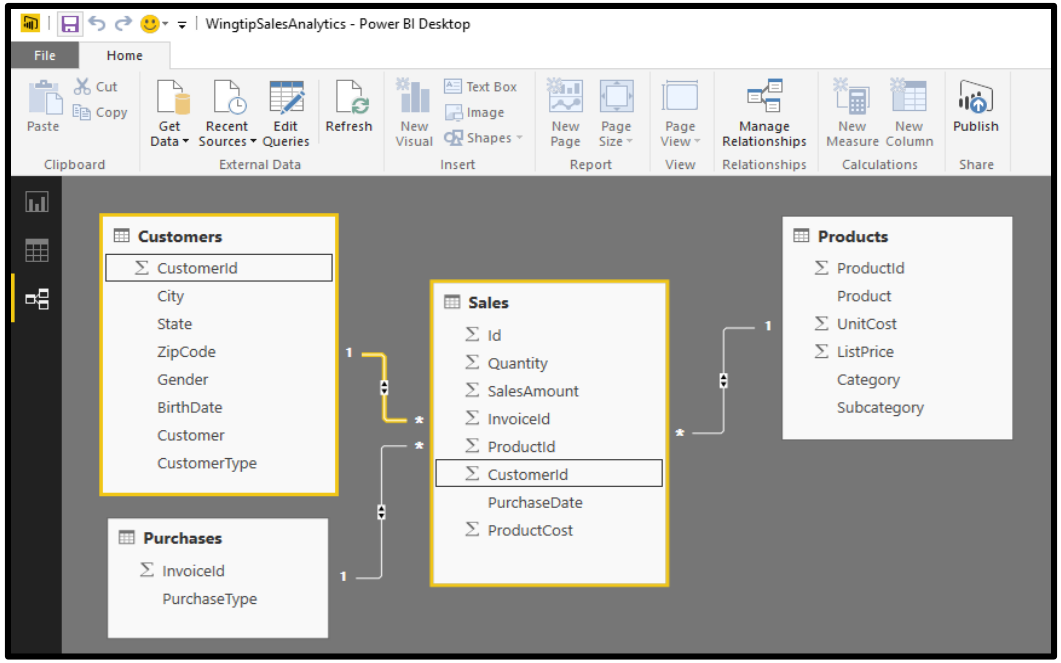
Data Modeling using a Star Schema

- 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



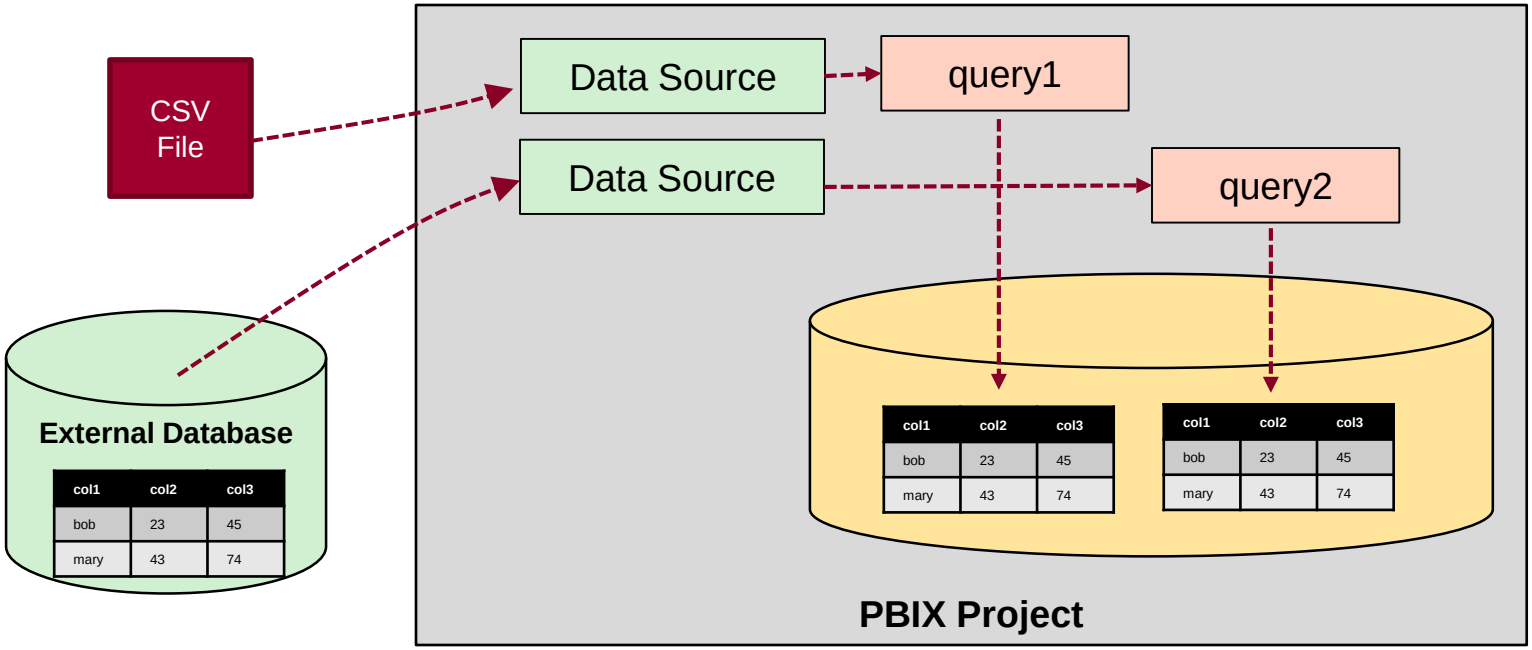

DEMO
Using Power BI Desktop to Import Data into a Star Schema

Agenda

- ✓ Deciding What To Measure
- ✓ Working with the Query Editor Window
- ✓ Managing Queries, Datasources and Credentials
- ✓ Designing Queries to Generate a Star Schema
- ➢ Merging Columns from Multiple Datasources
- Appending Rows from Multiple Datasources

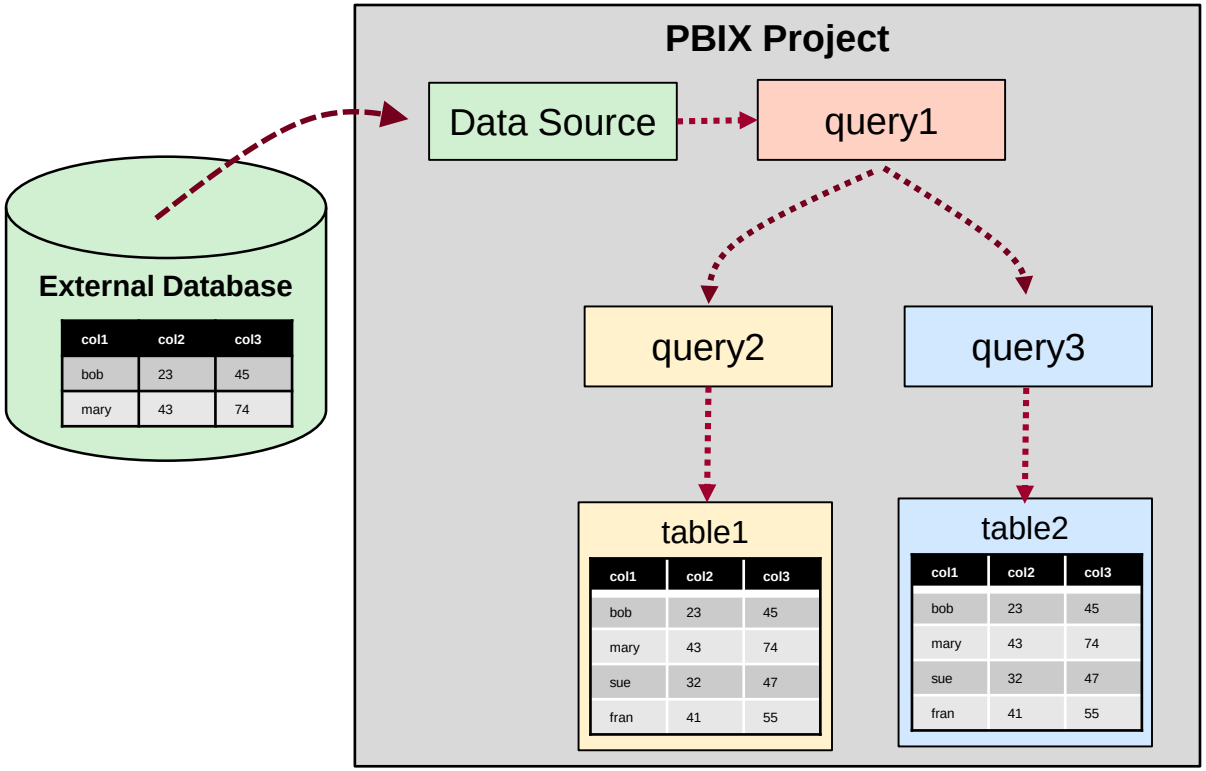
Understanding Query Input and Output

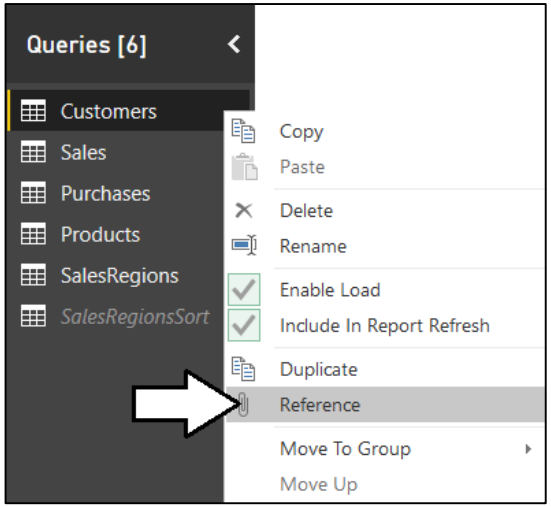
- PBIX project is container for data sources and queries
  - Queries created and saved within scope of Power BI project
  - Queries can pull data from local files
  - Queries can pull data from external content sources
  - Queries main purpose is to load imported data into data model



Query Composition

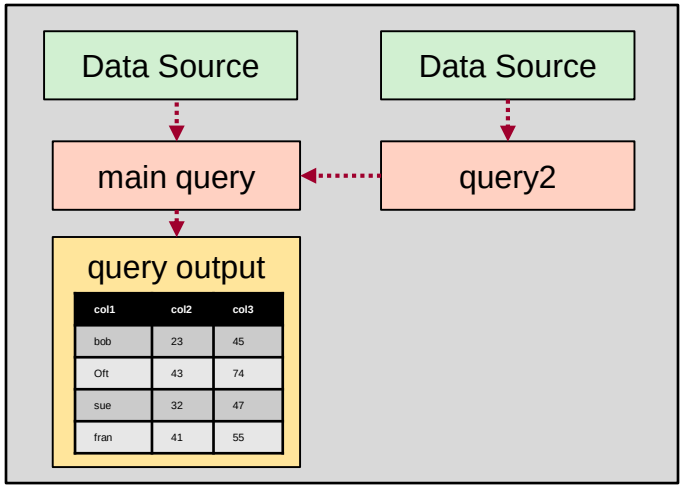
- Query can serve as source for other queries
  - Allows for creation of reusable base queries & query composition
  - Complexity can be hidden in base queries
  - **Reference** command creates new query based on another query

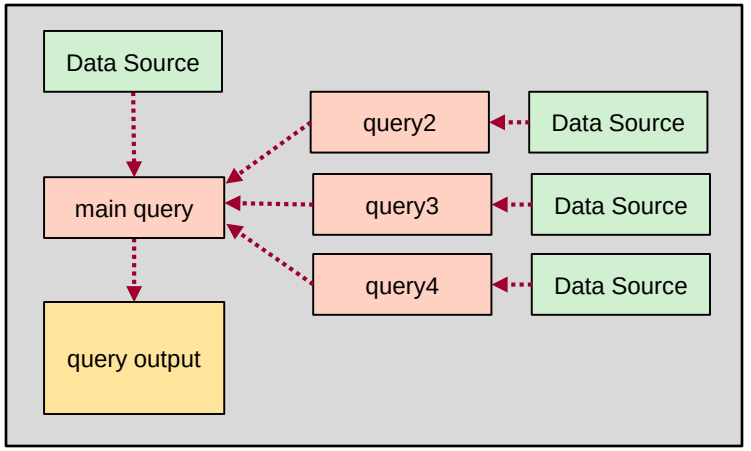




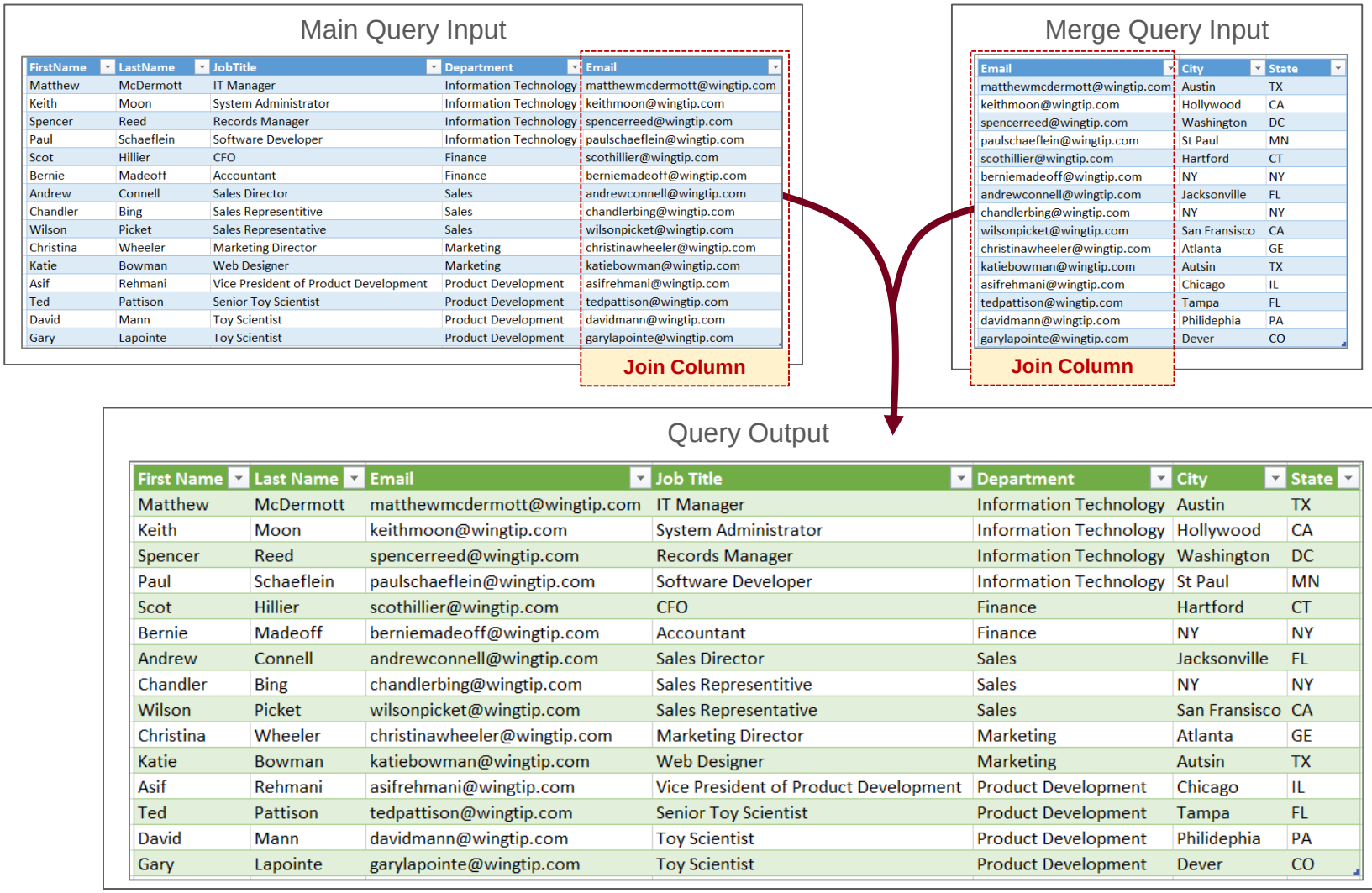
Combining Queries

- 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



Main Query Input

FirstName	LastName	JobTitle	Department	Email
Matthew	McDermott	IT Manager	Information Technology	matthewmcdermott@wingtip.com
Keith	Moon	System Administrator	Information Technology	keithmoon@wingtip.com
Spencer	Reed	Records Manager	Information Technology	spencerreed@wingtip.com
Paul	Schaeflein	Software Developer	Information Technology	paulschaeflein@wingtip.com
Scot	Hillier	CFO	Finance	scothillier@wingtip.com
Bernie	Madeoff	Accountant	Finance	berniemadeoff@wingtip.com
Andrew	Connell	Sales Director	Sales	andrewconnell@wingtip.com
Chandler	Bing	Sales Representitive	Sales	chandlerbing@wingtip.com
Wilson	Picket	Sales Representative	Sales	wilsonpicket@wingtip.com
Christina	Wheeler	Marketing Director	Marketing	christinawheeler@wingtip.com
Katie	Bowman	Web Designer	Marketing	katiebowman@wingtip.com
Asif	Rehmani	Vice President of Product Development	Product Development	asifrehmani@wingtip.com
Ted	Pattison	Senior Toy Scientist	Product Development	tedpattison@wingtip.com
David	Mann	Toy Scientist	Product Development	davidmann@wingtip.com
Gary	Lapointe	Toy Scientist	Product Development	garylapointe@wingtip.com

Merge Query Input

Email	City	State
matthewmcdermott@wingtip.com	Austin	TX
keithmoon@wingtip.com	Hollywood	CA
spencerreed@wingtip.com	Washington	DC
paulschaeflein@wingtip.com	St Paul	MN
scothillier@wingtip.com	Hartford	CT
berniemadeoff@wingtip.com	NY	NY
andrewconnell@wingtip.com	Jacksonville	FL
chandlerbing@wingtip.com	NY	NY
wilsonpicket@wingtip.com	San Fransisco	CA
christinawheeler@wingtip.com	Atlanta	GE
katiebowman@wingtip.com	Autsin	TX
asifrehmani@wingtip.com	Chicago	IL
tedpattison@wingtip.com	Tampa	FL
davidmann@wingtip.com	Philidephia	PA
garylapointe@wingtip.com	Dever	CO

Query Output

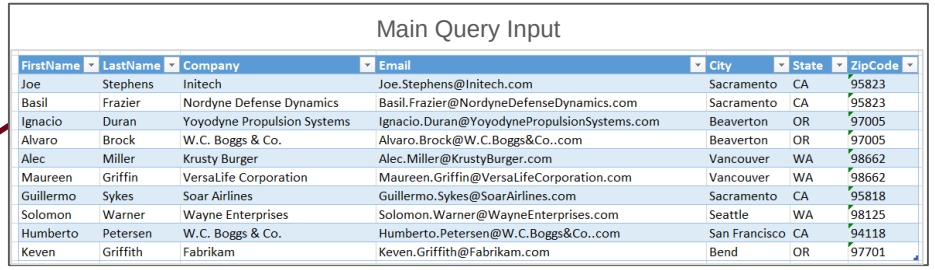
First Name	Last Name	Email	Job Title	Department	City	State
Matthew	McDermott	matthewmcdermott@wingtip.com	IT Manager	Information Technology	Austin	TX
Keith	Moon	keithmoon@wingtip.com	System Administrator	Information Technology	Hollywood	CA
Spencer	Reed	spencerreed@wingtip.com	Records Manager	Information 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	CO


DEMO
Merging Columns from Multiple Datasources

Agenda

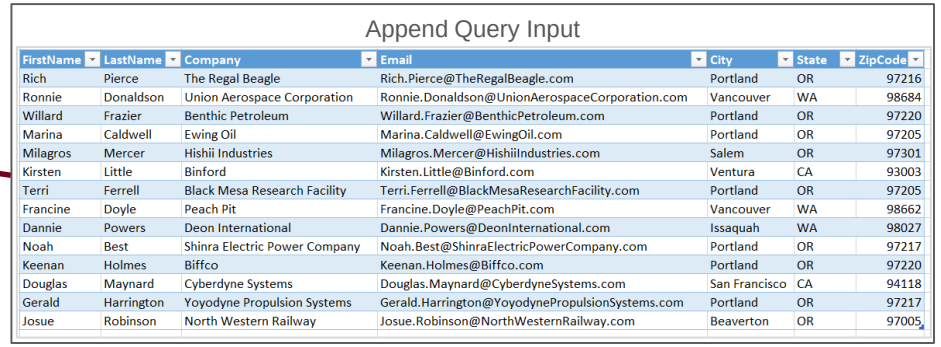
- ✓ Deciding What To Measure
- ✓ Working with the Query Editor Window
- ✓ Managing Queries, Datasources and Credentials
- ✓ Designing Queries to Generate a Star Schema
- ✓ Merging Columns from Multiple Datasources
- ➢ Appending Rows from Multiple Datasources

Appending Rows



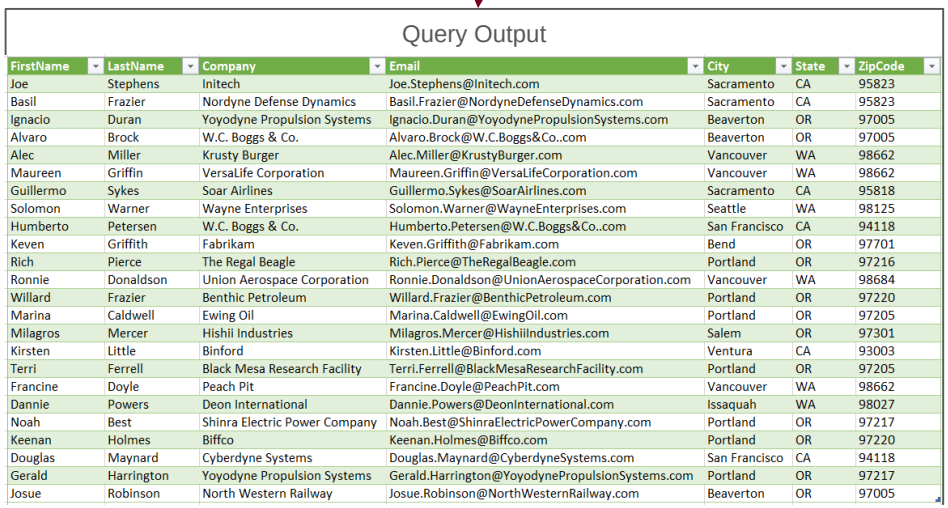
Main Query Input

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&Co..com	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&Co..com	San Francisco	CA	94118
Keven	Griffith	Fabrikam	Keven.Griffith@Fabrikam.com	Bend	OR	97701



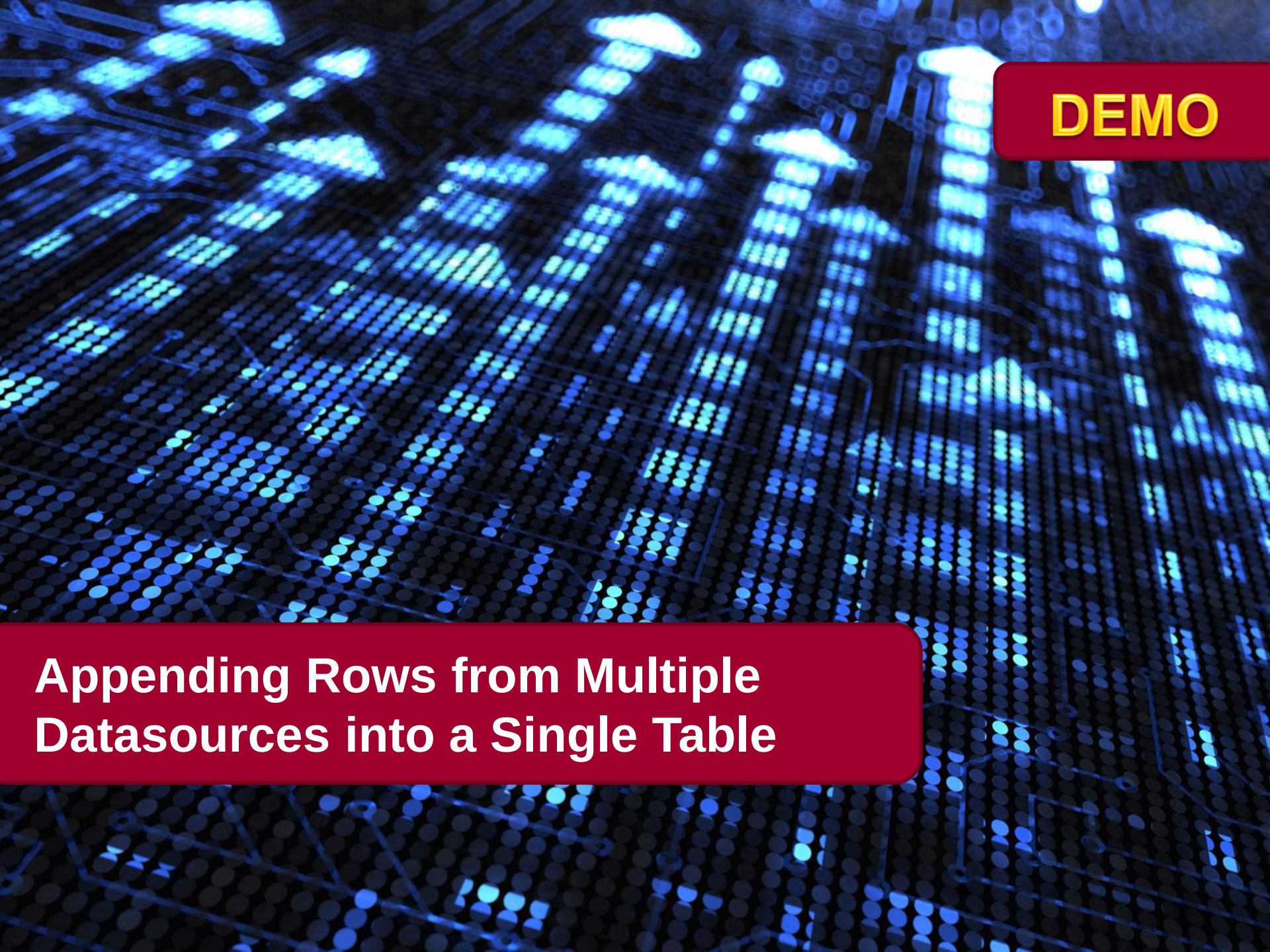
Append Query Input

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&Co..com	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&Co..com	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


DEMO
Appending Rows from Multiple Datasources into a Single Table

Summary

- ✓ Deciding What To Measure
- ✓ Working with the Query Editor Window
- ✓ Managing Queries, Datasources and Credentials
- ✓ Designing Queries to Generate a Star Schema
- ✓ Merging Columns from Multiple Datasources
- ✓ Appending Rows from Multiple Datasources