

PA220: Database systems for data analytics

Power BI

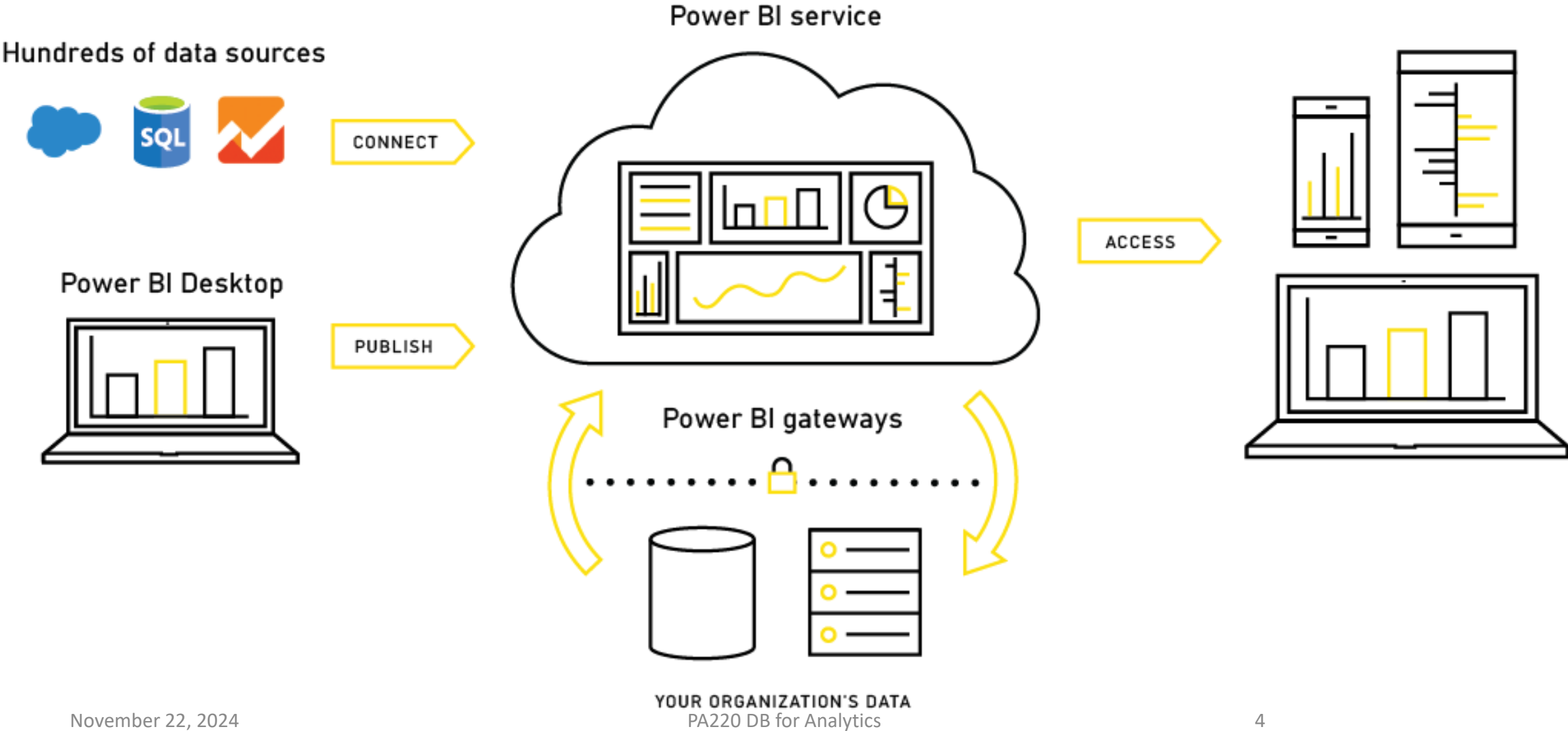
# Microsoft Power BI

- an interactive data visualization software
  - focusing on business intelligence.
  - allows collaborating on and sharing customized dashboards and interactive reports.
- a collection of software services, apps, and data connectors
  - to turn unrelated sources of data into coherent, visually immersive, and interactive insights.
- data input directly from a database, webpage, or structured files
  - e.g., spreadsheets, CSV, XML, and JSON.

# Power BI Ecosystem

- Power BI.com
  - online platform, for O365 tenants
- Desktop
  - free app for dashboarding, reporting, ad-hoc visualizations
- Mobile Apps
  - access to reports, ... from mobile devices
- Visuals Marketplace
  - store of visuals and R-powered visuals
- Gateway
  - sync external data in and out for automated refreshes
- Dataflow
  - implementation of power query for data transformations
- Dataset
  - interface of data sources to reports and visualizations
- Embedded
  - REST API for custom solutions
- Report Server
  - on-premise solution for running dashboard and reports

# Power BI System Architecture



# Power BI announcements

- **Gartner positions Microsoft as a leader in BI and Analytics Platform**
- **February 2016:**
- Gartner has positioned Microsoft as a Leader, for the ninth consecutive year, in the Magic Quadrant for Business Intelligence and Analytics Platforms
- For the first time, Microsoft is placed furthest in vision within the Leaders quadrant

Gartner "Magic Quadrant for Business Intelligence and Analytics Platforms," by Josh Parenteau, Rita L. Sallam, Cindi Howson, Joao Tapadinhas, Kurt Schlegel, Thomas W. Oestreich February 4, 2016

This graphic was published by Gartner, Inc. as part of a larger research document and should be evaluated in the context of the entire document. The Gartner document is available upon request from Microsoft. Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.



# Power BI announcements

- **Gartner positions Microsoft as a leader in BI and Analytics Platform**

- **June 2024:**

- Gartner has positioned Microsoft as a Leader in the Magic Quadrant for Business Intelligence and Analytics Platforms

- Microsoft is placed furthest in vision within the Leaders quadrant

- They added innovations to the Power BI platform:
- introduction of Microsoft Fabric and Copilot in Fabric

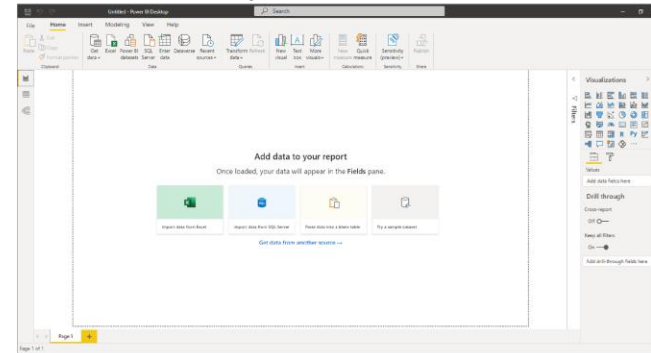


Gartner, Magic Quadrant for Analytics and Business Intelligence Platforms, Kurt Schlegel, Anirudh Ganeshan, David Pidsley, Julian Sun, Georgia O'Callaghan, Christopher Long, Kevin Quinn, Fay Fei, Edgar Macari, Jamie O'Brien [June 20, 2024]

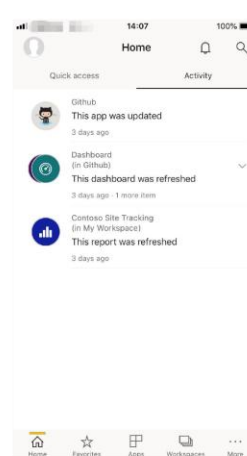
# Parts of Power BI

- Desktop, service, and Mobile apps
  - designed to let people create, share, and consume business insights
  - in the way that serves them, or their role, most effectively.

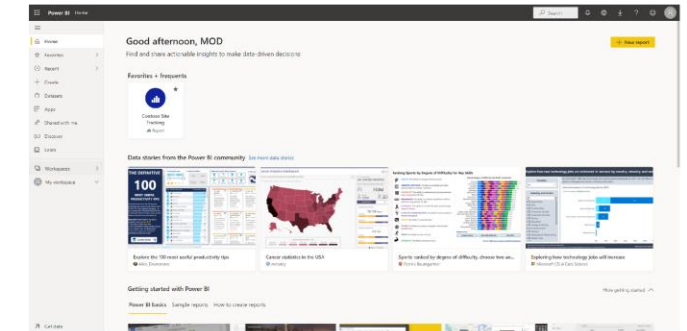
Power BI Desktop



Power BI Mobile



Power BI service



# Power BI Concepts

- Major building blocks: datasets, reports and dashboards
- Organized in workspaces
- Dataset
  - a collection of data that you import or connect to,
  - associated with workspaces,
  - a single dataset can be part of many workspaces.
- Report
  - one or more pages of visualizations (*visuals*) such as line charts, maps, and treemaps
- Dashboard
  - typically, a pinned visual of a report, but not necessarily
  - created in Power BI service



# Power BI Concepts

- Why create dashboards?
  - to see, in one glance, all the information needed to make decisions,
    - to ensure all colleagues are on the same page, viewing and using the same information,
  - to monitor the most important information about your business,
    - to monitor the health of a business or product or business unit or marketing campaign, etc.,
  - to create a personalized view of a larger dashboard and show all the metrics that matter to them.

# Contents

- Introducing Power BI Desktop
  - Creating queries
  - Configuring relationships
  - Enhancing the model
- Designing reports
- Publishing to Power BI
- Describing additional capabilities
- How to enable preview features
- Additional resources

# Power BI Desktop

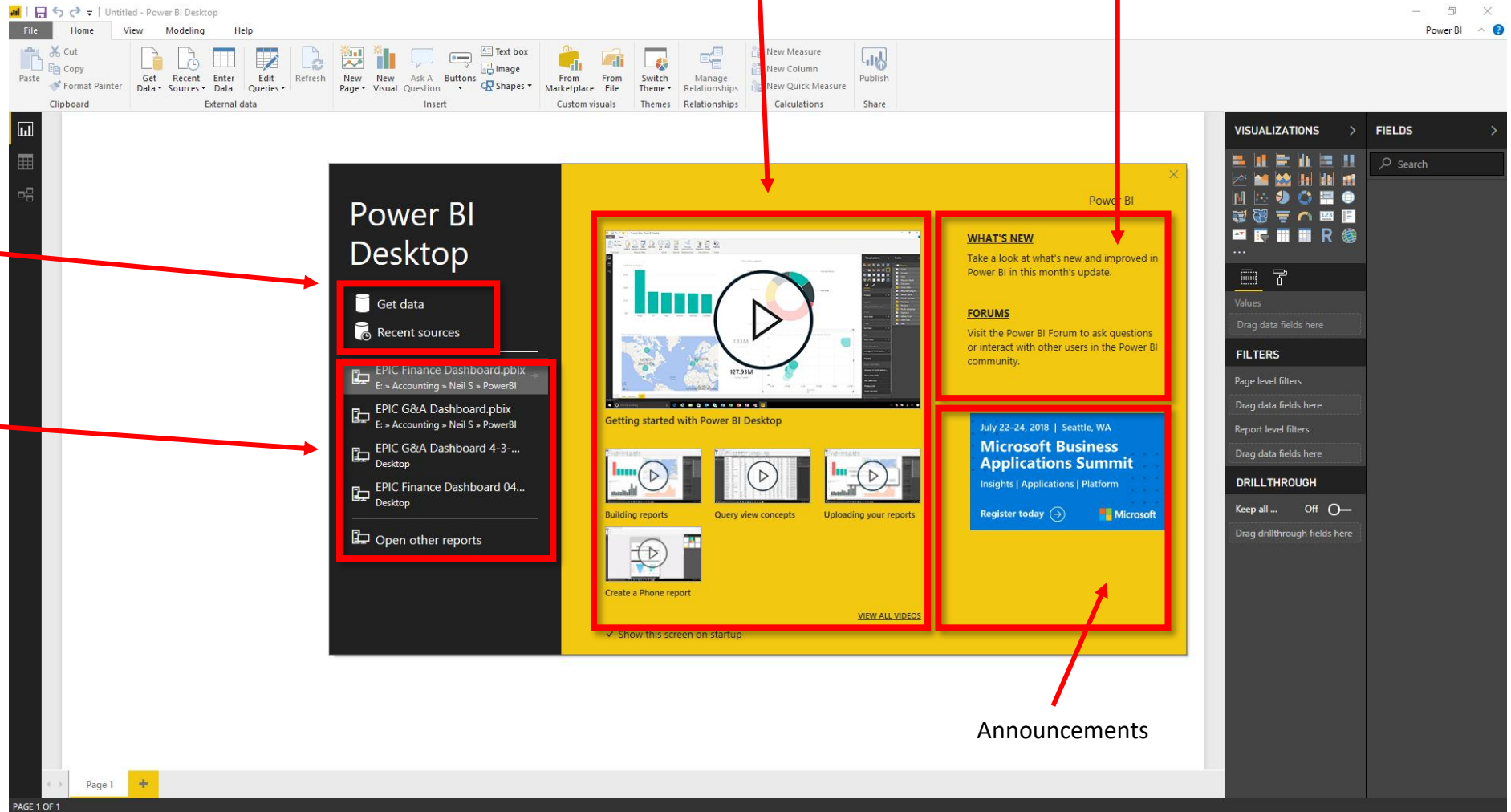
Find out what is new in the latest release or interact on the highly interactive forums, which are a great way to ask questions

Start a new import from a new source, or connect to a recently used source

Open a recent file or other report

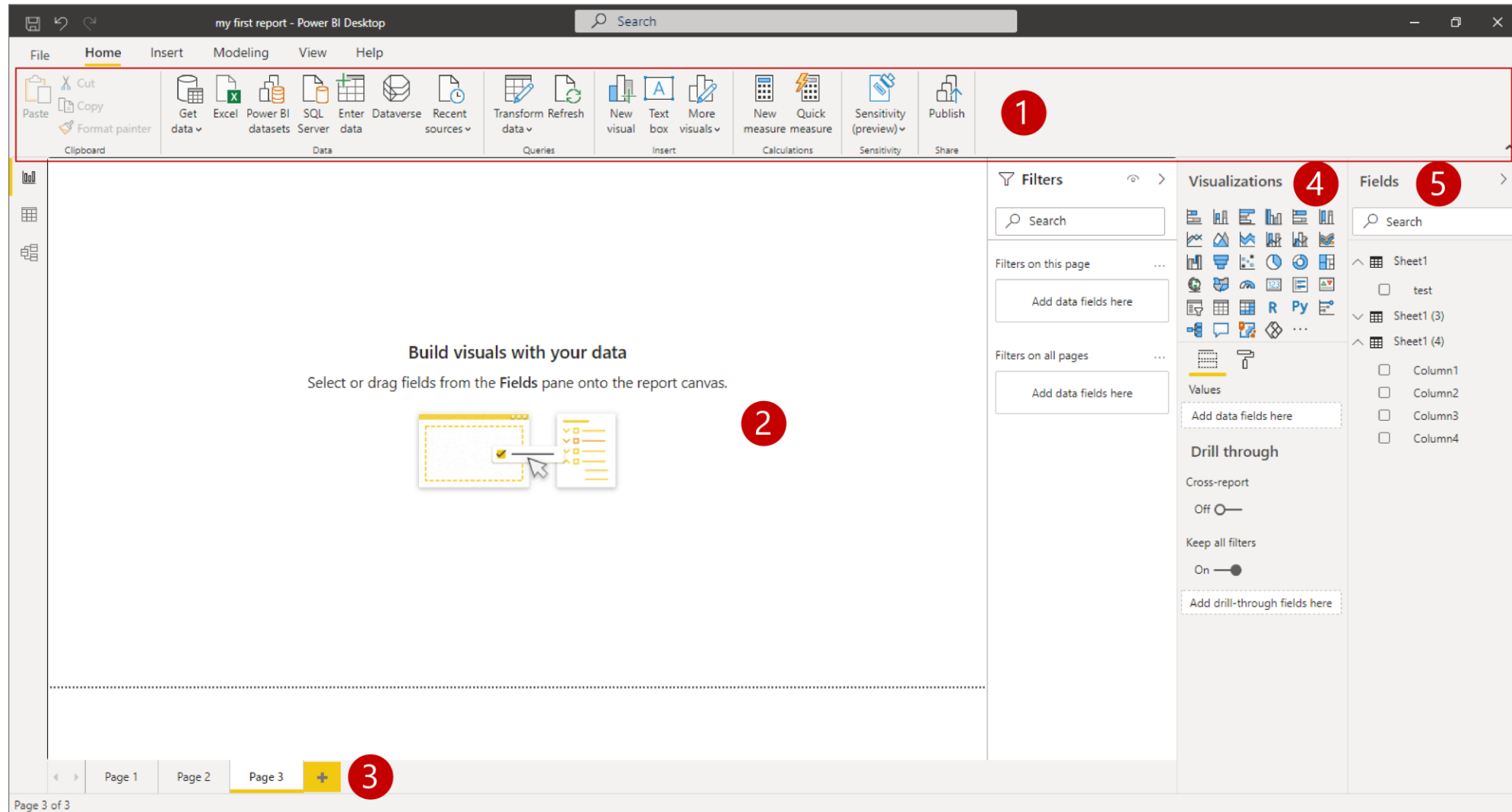
Tutorial Videos

Announcements



# Power BI Desktop

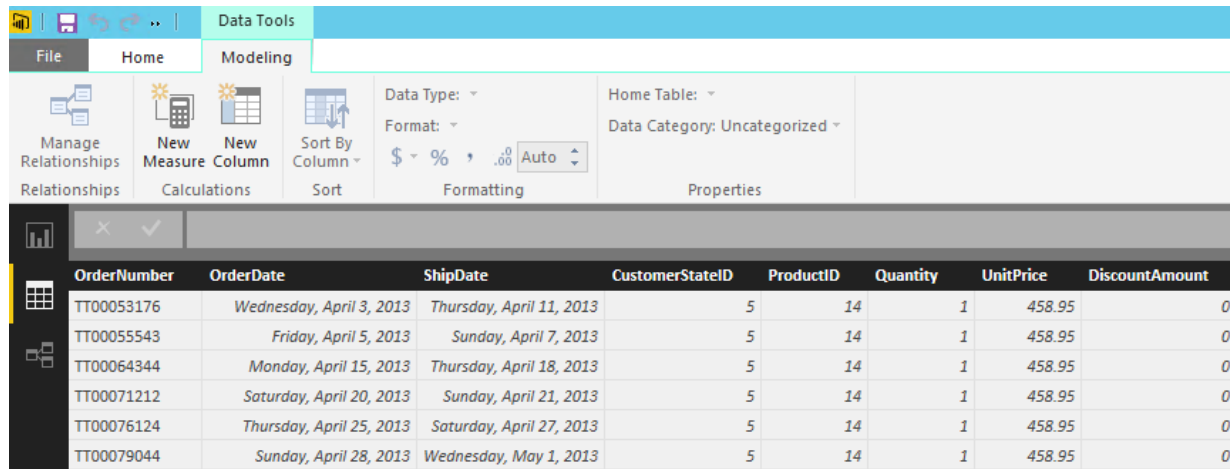
- 1 – ribbon
- 2 – report view or canvas; displays also data and model
- 3 – pages tab (individual report pages)
- 4 – visuals pane
- 5 – fields pane



# Introducing Power BI Desktop

---

- User experience: Data view
- Work in **Data View** to inspect, explore, and understand data in the model
- It is a different experience from how you can view tables, columns, and data in **Query Editor**
- This is a view of the data after it has been loaded into the model

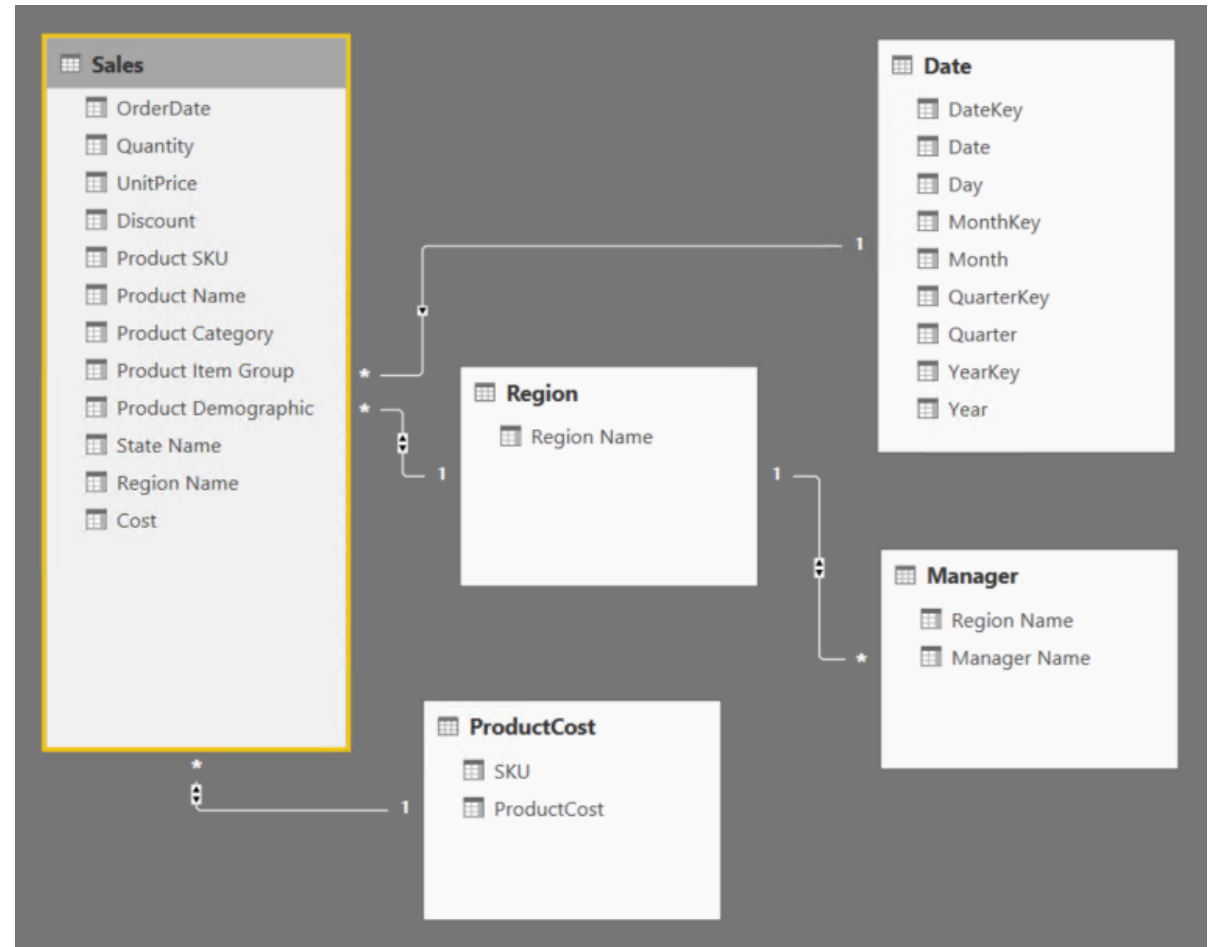


The screenshot shows the Power BI Desktop interface in Data View. The ribbon is set to 'Data Tools' with the 'Modeling' tab selected. The ribbon includes sections for 'Manage Relationships', 'New Measure', 'New Column', 'Sort By Column', 'Data Type', 'Format', 'Home Table', and 'Data Category'. Below the ribbon, a table of order data is displayed with columns for OrderNumber, OrderDate, ShipDate, CustomerStateID, ProductID, Quantity, UnitPrice, and DiscountAmount.

OrderNumber	OrderDate	ShipDate	CustomerStateID	ProductID	Quantity	UnitPrice	DiscountAmount
TT00053176	Wednesday, April 3, 2013	Thursday, April 11, 2013	5	14	1	458.95	0
TT00055543	Friday, April 5, 2013	Sunday, April 7, 2013	5	14	1	458.95	0
TT00064344	Monday, April 15, 2013	Thursday, April 18, 2013	5	14	1	458.95	0
TT00071212	Saturday, April 20, 2013	Sunday, April 21, 2013	5	14	1	458.95	0
TT00076124	Thursday, April 25, 2013	Saturday, April 27, 2013	5	14	1	458.95	0
TT00079044	Sunday, April 28, 2013	Wednesday, May 1, 2013	5	14	1	458.95	0

# Introducing Power BI Desktop

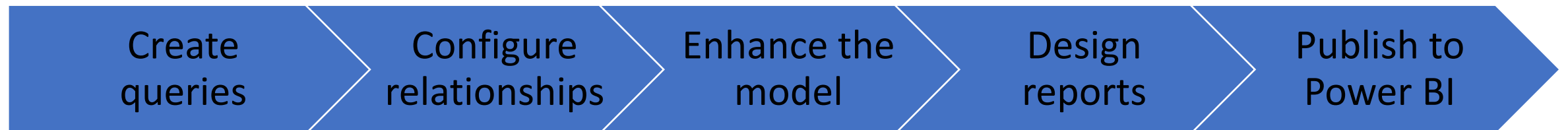
- User experience: Relationships view
- Work in **Relationships View** to view all tables, columns, measures and relationships in the model



# Introducing Power BI Desktop

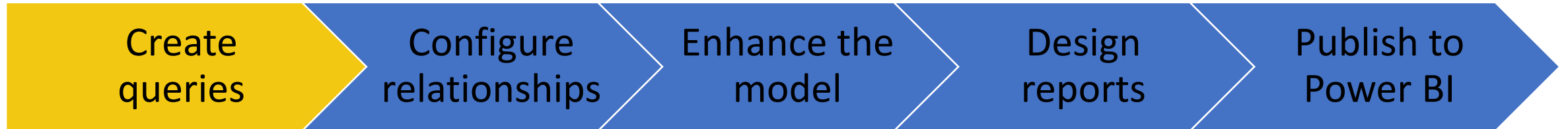
---

- Complex schemas: Development methodology
- In contrast to the simple and quick “get then visualize” approach, more complex schemas may follow this methodology:
  1. Create queries and use the **Query Editor** to filter, cleanse and reshape data
  2. Configure/refine relationships to establish the foundations of a model
  3. Enrich the model with calculation logic and formatting
  4. Design interactive reports with a broad range of modern data visualizations
  5. Publish solutions directly to the Power BI service



# Creating queries

---

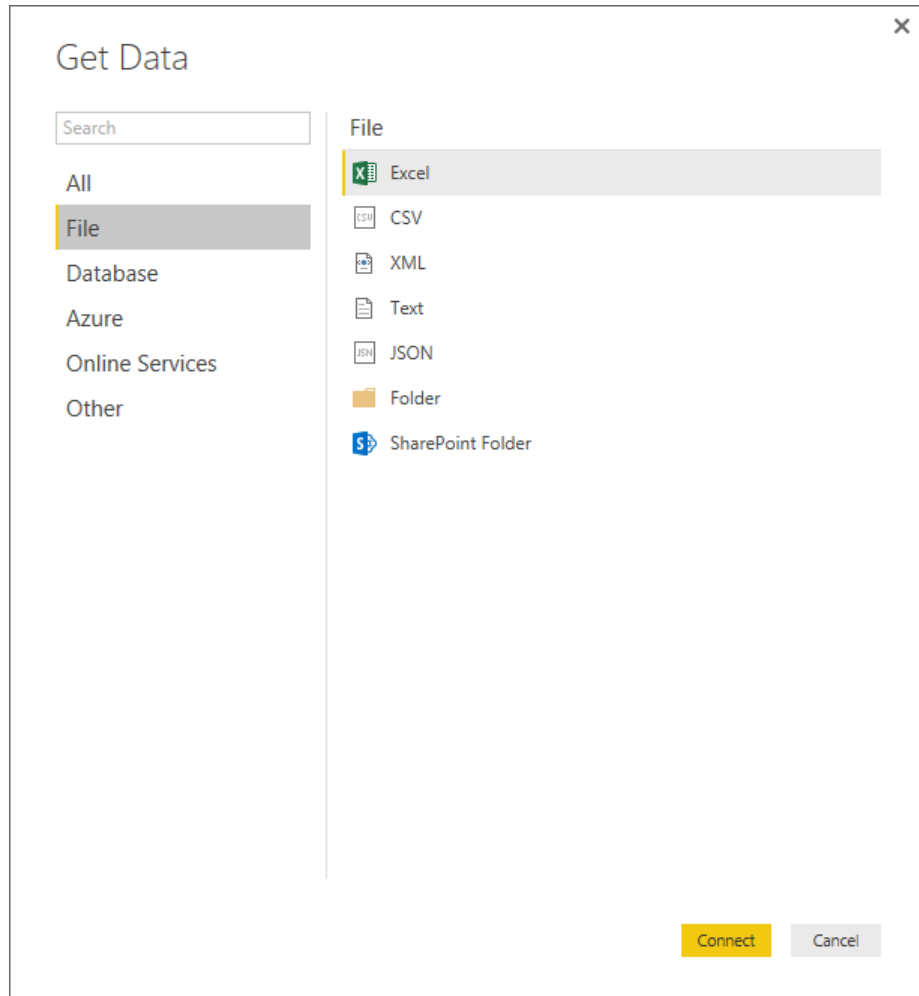


- Queries can be edited before loading into the model
- Once created, each query:
  - Is referenced by a unique name
  - Defines query steps by using the language “M”
    - Most query step logic does not require writing formulas
    - Advanced logic can be written, and even exposed as functions
- Is used to load data into a model table

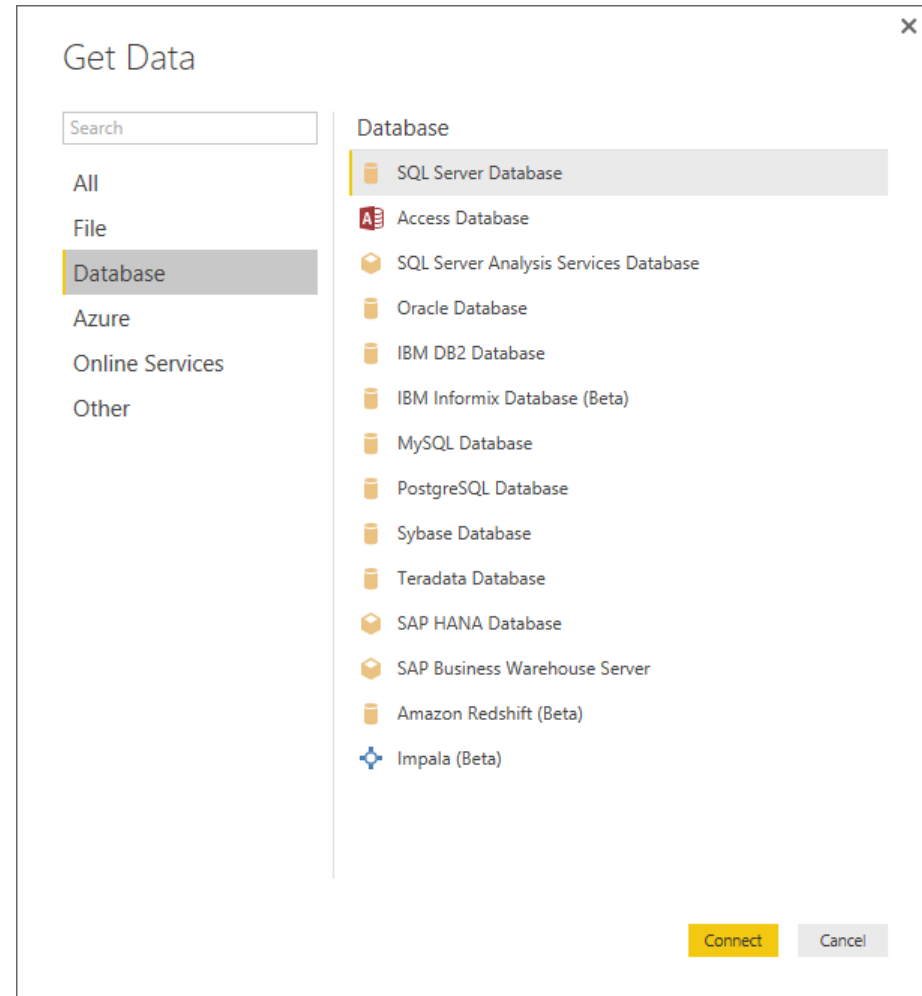


# Creating queries

- Source types: File



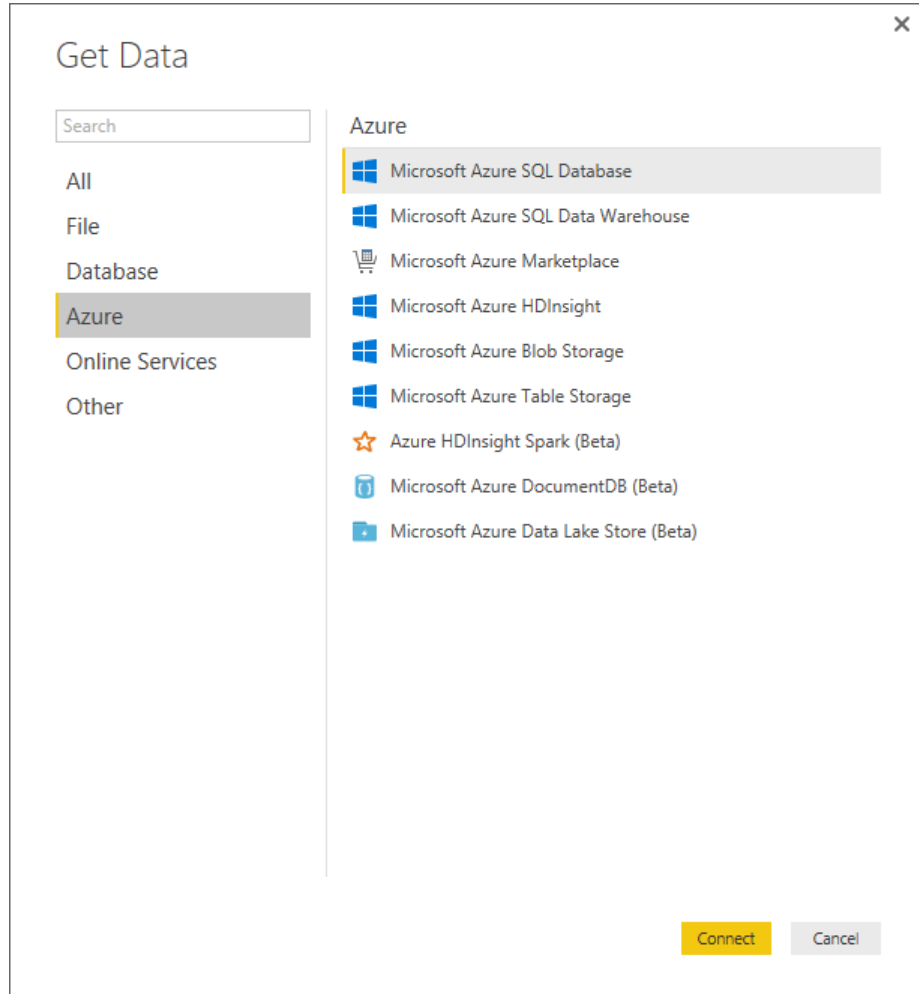
- Source types: Database



# Creating queries

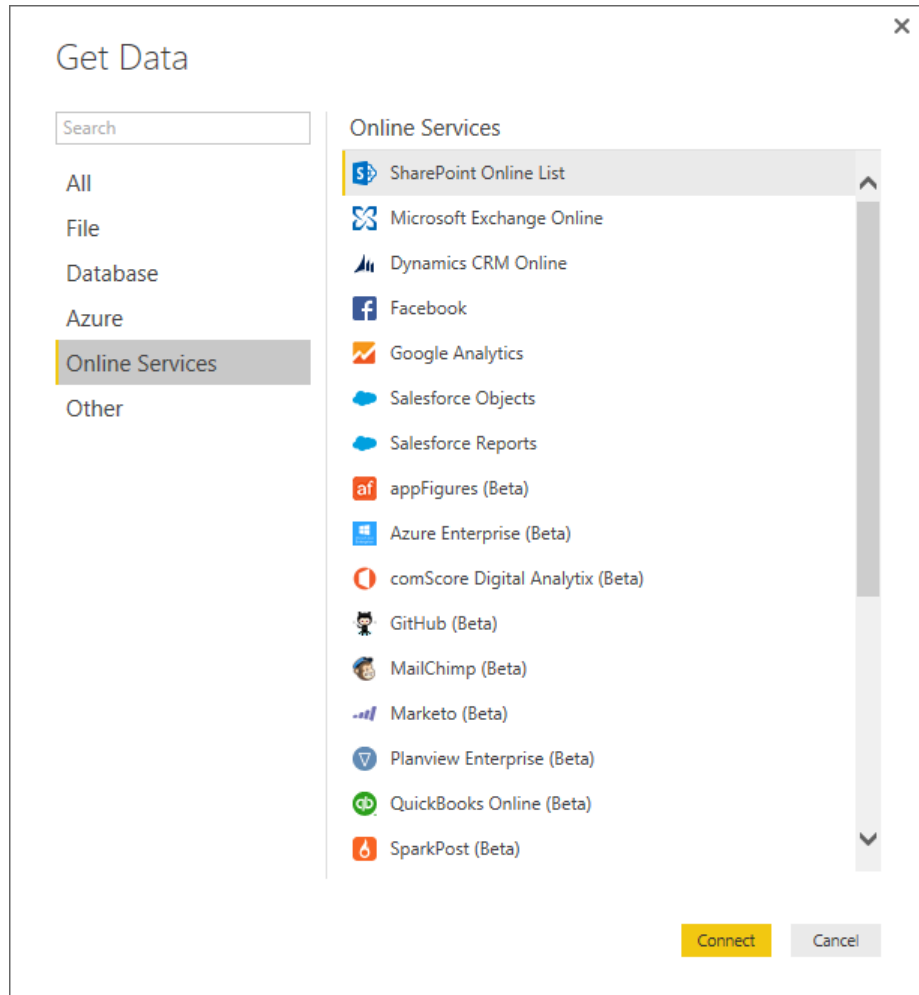
---

- Source types: Azure



# Creating queries

- Source types\*: Online Services

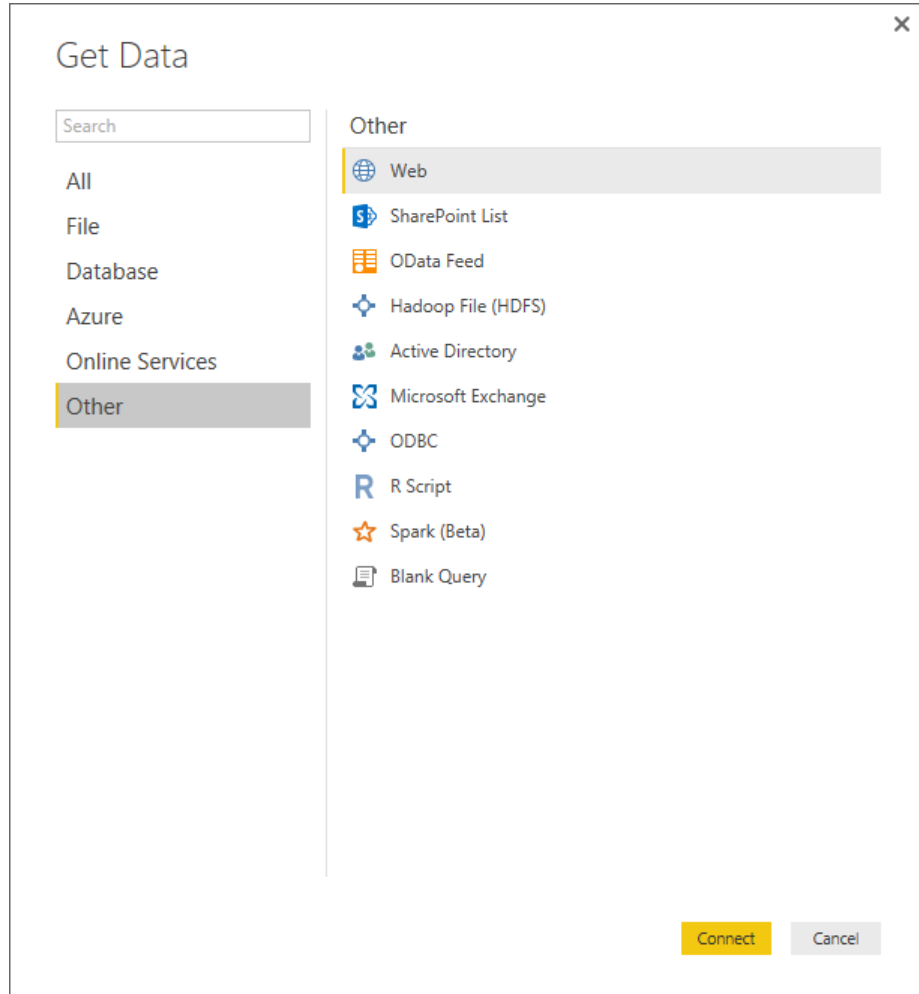


- The list also includes:
- Smartsheet
- SQL Sentry (Beta)
- Stripe (Beta)
- SweetIQ (Beta)
- Troux (Beta)
- Twilio (Beta)
- tyGraph (Beta)
- WebTrends (Beta)
- Zendesk (Beta)

\* As of  
August 25, 2016  
And the list is growing!

# Creating queries

- Source types: Other

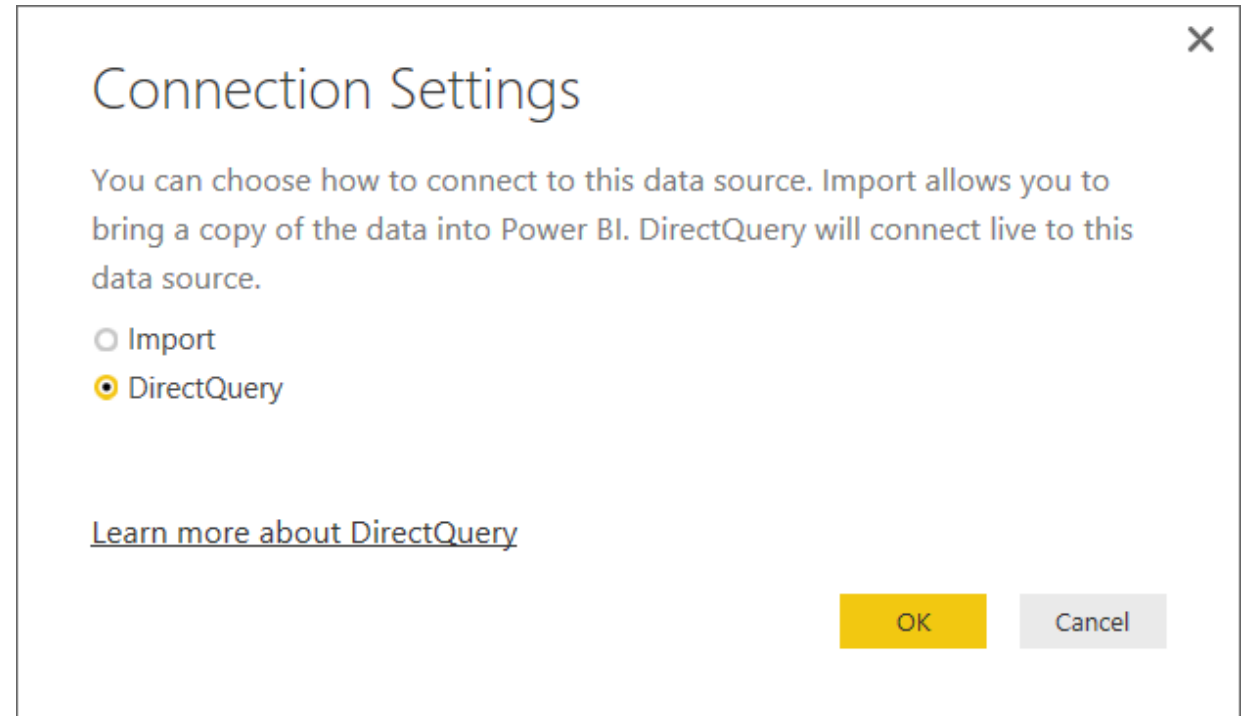


- Note: The following are not supported for on-premises refresh
- Hadoop File (HDFS)
- Microsoft Exchange

# Creating queries

---

- DirectQuery mode
- Connections made to supported data sources can connect directly by using **DirectQuery**:
- SQL Server
- Azure SQL Database
- Azure SQL Data Warehouse
- SAP HANA
- Oracle Database
- Teradata Database
- Amazon Redshift (Preview)



- **Benefits:**
- Allows building visualizations over very large datasets
- Reports always use current data

# Creating queries

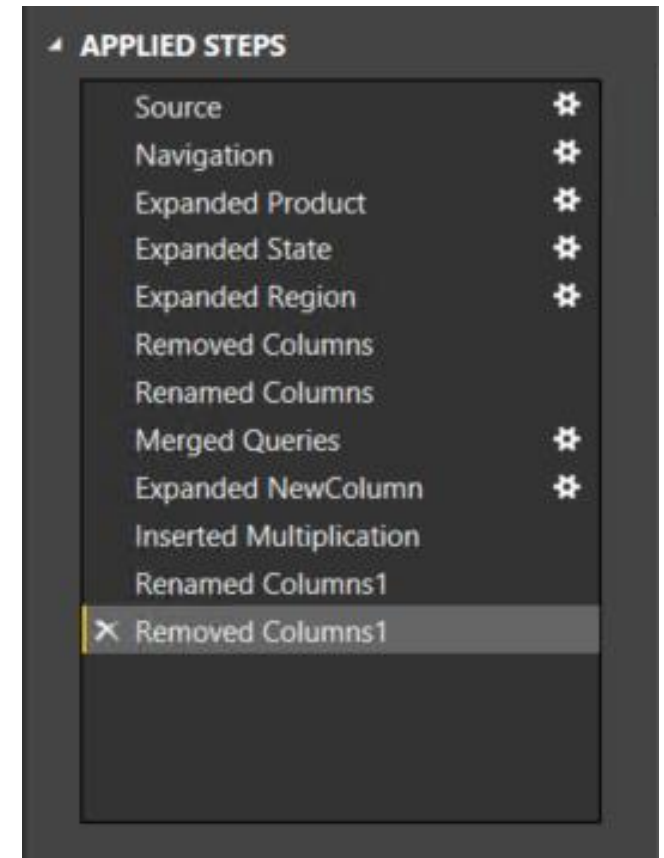
---

- DirectQuery mode: Continued
- Limitations:
  - All tables must come from a single database
  - The **Query Editor** query cannot be overly complex
  - DAX formula are restricted to a subset of the DAX functions
  - Relationship filtering is limited to a single direction, rather than both directions
    - If enabled, a preview feature allows cross filtering in both directions
  - The data type of a column cannot be changed

# Creating queries

---

- Defining query steps
- The **Query Editor** can be used to filter, cleanse and reshape data into the desired result
- Steps can easily be produced by applying column filters, and by using the commands available from the ribbon, or the query and column context menus
- It is possible to select a step and preview the data at that step
- Steps can be removed – but take care not to remove a step that downstream steps depend on
- Step formulas can be viewed or edited in the formula bar



# Creating queries

---

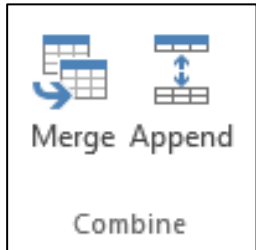
- Defining query steps: Commands
- Numerous commands are available on the **Query Editor** ribbons and context menus to:
  - Manage columns
  - Reduce rows, including removing errors
  - Transform
  - Split
  - Add columns, by using formulas
- As the editor experience is WYSIWYG, users can easily experiment and roll-back changes



# Creating queries

---

- Defining query steps: Combining queries



- New queries can be created by:
- Merging two queries (joining on a common column)
- Appending two queries (union)
- Merge join options include:

## Join Kind

Left Outer (all from first, matching from second)

Right Outer (all from second, matching from first)

Full Outer (all rows from both)

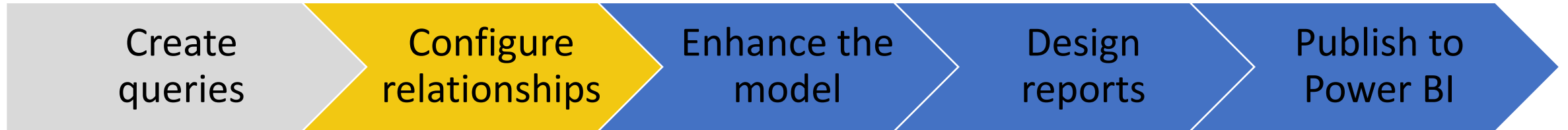
Inner (only matching rows)

Left Anti (rows only in first)

Right Anti (rows only in second)

# Configuring relationships

---



- Relationships enable filters on one table to appropriately propagate filter to other tables
- Note: Relationships are not foreign key (integrity) constraints
- They can be defined between any two tables, regardless of their data connection or source type
- However, they must be based on single columns with matching data types
- There is no support for self-referencing relationships

# Configuring relationships

---

- Advanced options: Continued
- Configure relationship advanced options to set:
  - Cardinality
  - Cross filter direction
    - The **Both** option enables many-to-many analysis
  - Active/inactive

## Advanced options

### Cardinality

Many to One (\*:1)

One to One (1:1)

One to Many (1:\*)

### Cross filter direction

Single

Both

Make this relationship active

Tip: All options can be interpreted from symbols used in the relationship diagram

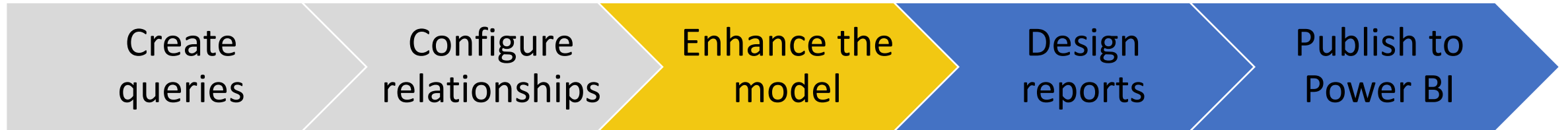
# Configuring relationships

---

- Advanced options
- It is possible to define multiple relationship paths between two tables
- One relationship must be **active**, and will be used by default during model browsing
- The remaining relationships are **inactive**, and can be used explicitly in model calculations

# Enhancing the model

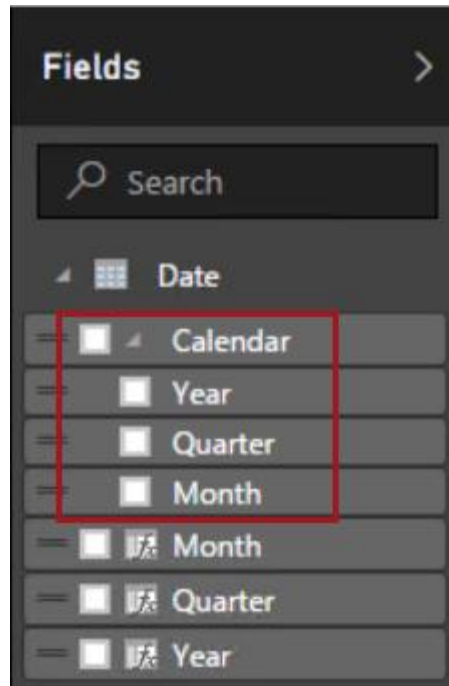
---



- Model tables can be extended with hierarchies, calculations, and may be hidden (i.e., not for reporting)
- Model column properties can be set:
  - Data type
  - Format
  - Categorization
    - Spatial types (GPS), or web URLs
  - Sort order, based on another column from the same table
  - Default summarization (numeric, text and date columns)
  - Visibility

# Enhancing the model

- Hierarchies
- Hierarchies can be added to a table, enabling navigation between columns (in the one-to-many direction)



The image shows a table view with columns for 'Year', 'Quarter', and 'Month'. The data is organized into a hierarchy, with rows for each combination of Year, Quarter, and Month. The table is scrollable and has a search bar at the top.

Year	Quarter	Month
CY2013	CY2013 Q1	CY2013 Jan
CY2013	CY2013 Q1	CY2013 Feb
CY2013	CY2013 Q1	CY2013 Mar
CY2013	CY2013 Q2	CY2013 Apr
CY2013	CY2013 Q2	CY2013 May
CY2013	CY2013 Q2	CY2013 Jun
CY2013	CY2013 Q3	CY2013 Jul
CY2013	CY2013 Q3	CY2013 Aug
CY2013	CY2013 Q3	CY2013 Sep
CY2013	CY2013 Q4	CY2013 Oct
CY2013	CY2013 Q4	CY2013 Nov
CY2013	CY2013 Q4	CY2013 Dec
CY2014	CY2014 Q1	CY2014 Jan
CY2014	CY2014 Q1	CY2014 Feb
CY2014	CY2014 Q1	CY2014 Mar
CY2014	CY2014 Q2	CY2014 Apr

# Enhancing the model

---

- Calculations
- There are three different types of calculations, and each are defined by using DAX:
  - Calculated columns
  - Calculated tables
  - Measures
- DAX = Data Analysis Expressions (DAX), consisting of:
  - Excel functions (~80 functions)
  - Table functions
  - Aggregate functions
  - Relationship navigation functions
  - Context modification functions
  - Time Intelligence functions

# Enhancing the model

---

- Calculations: Calculated columns
- Define **Calculated Columns** to add new columns to tables
- Column values for each row are materialized and stored in the model
- If possible, to reduce the model size, avoid creating them
- Column values are recalculated when:
  - The table is refreshed
  - Formula dependencies are refreshed



# Enhancing the model

---

- Calculations: Calculated tables
- Define **Calculated Tables** to add new tables to the model, based on existing data from other tables
- Calculated tables are generally best for intermediate calculations of data stored in the model, rather than calculated on the fly
- Calculated tables appear in Relationship View, and relationships can be defined with other tables
- If possible, to reduce the model size, avoid creating them
- Calculated tables are recalculated in the same circumstances as calculated columns

Note: This is an advanced design concept

# Enhancing the model

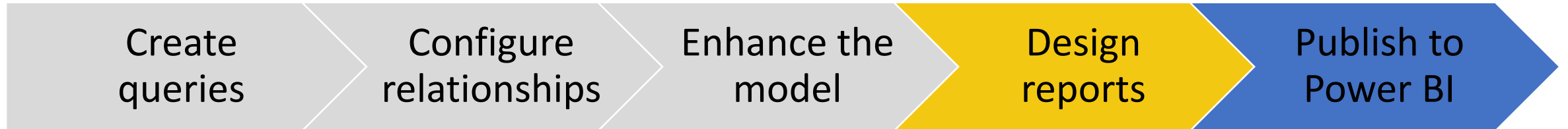
---

- Calculations: Measures
- Define **Measures** to add aggregation logic to the model
- Values are not materialized in the model
- Formulas are evaluated at query time
- Common aggregate functions:
  - Sum
  - Count, Distinct Count
  - Average
  - Minimum, Maximum

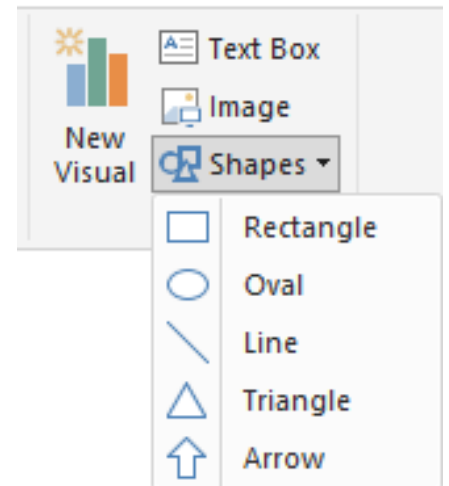
Note: Summarizable columns are aggregated by reports—so defining measures is not necessary for simple aggregation requirements

# Designing reports

---



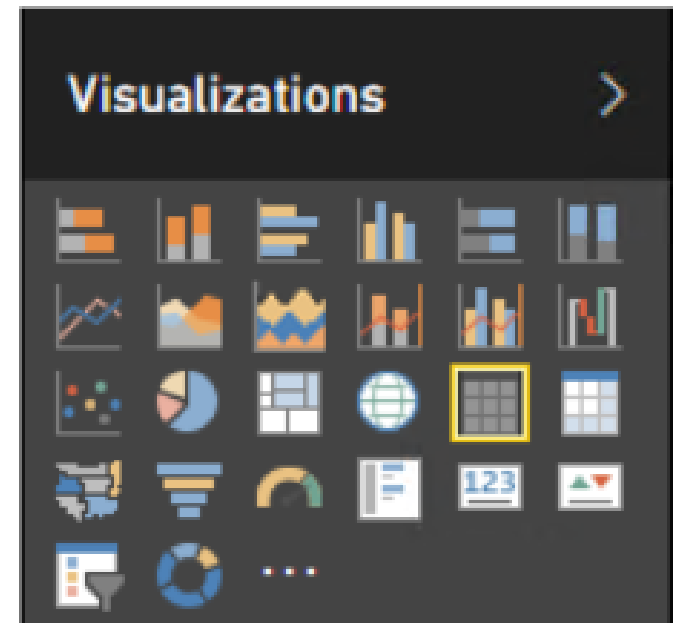
- Reports can be designed based on the visible model interface
- The design experience is almost the same as is available in the Power BI online service
- Text boxes, images and shapes can be added
- Text boxes can contain hyperlinks



# Designing reports

---

- Custom visuals
- Custom visuals can be imported to extend beyond the out-of-the-box visualizations
- A gallery of visuals created by the Power BI community is available at <https://app.powerbi.com/visuals>
- Browse through the visuals or submit one of your own for others to use
- The list of available visuals is growing each month
- Custom visuals will render in the Power BI service



# Designing reports

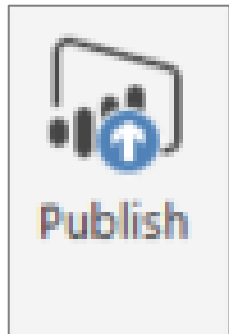
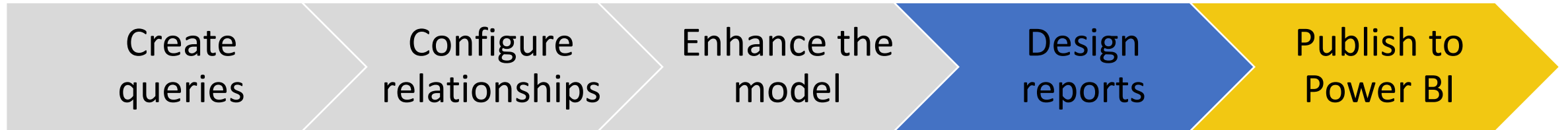
- Custom visuals: Gallery (subset)



\* And the list is growing!

# Publishing to Power BI

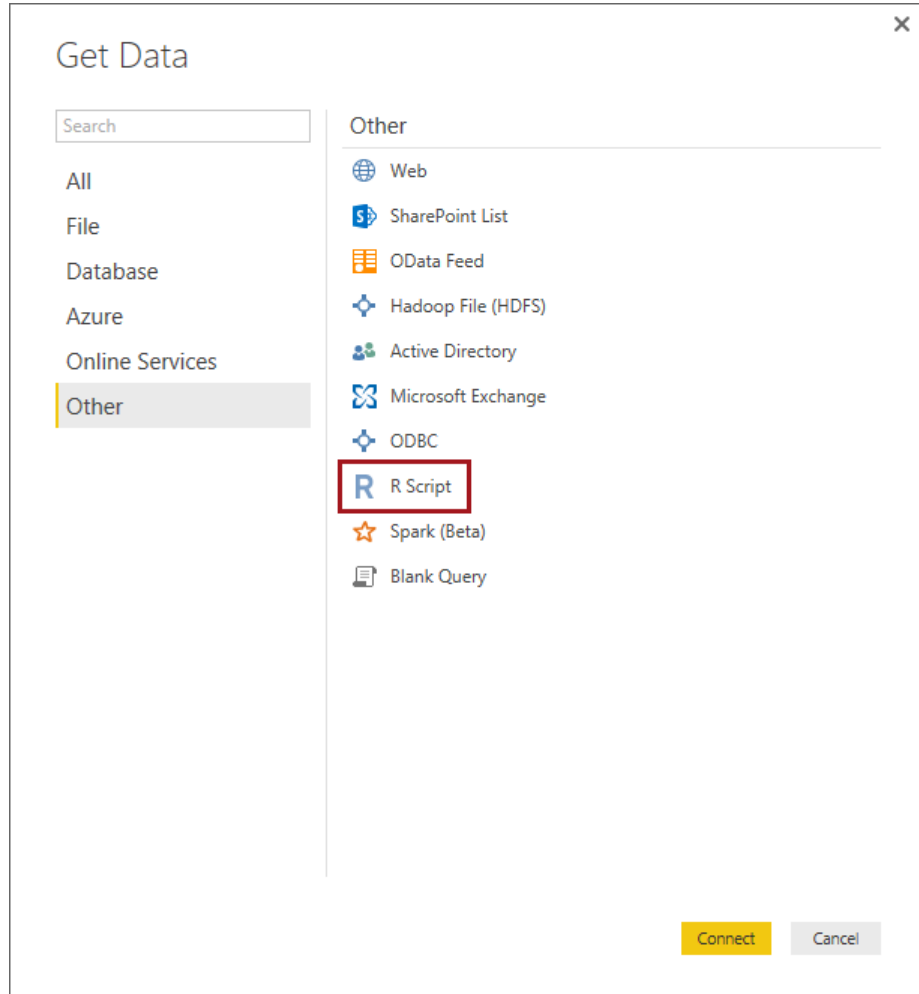
---



- The Power BI Desktop file can be uploaded to the Power BI service, or published directly
- Publish directly to your workspace or a group
- Note, if overwriting an existing dataset:
  - If there are two or more datasets with the same name, remove one, or rename the Power BI Desktop file
  - Renaming columns or measures can break existing reports or dashboard tiles

# Describing additional capabilities

- Running R scripts

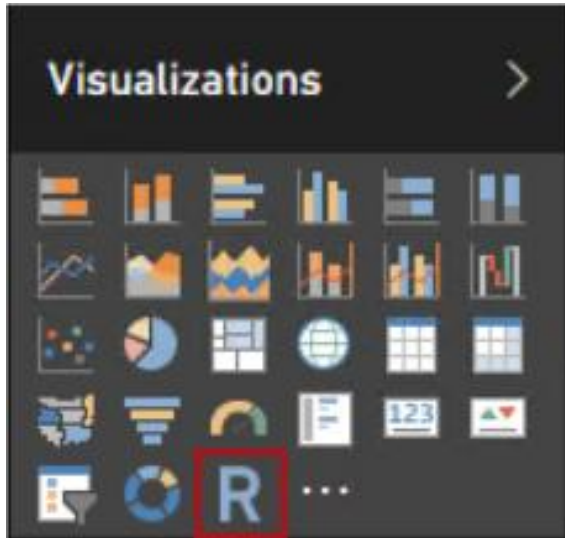


- R scripts can be run directly in Power BI Desktop, and resulting datasets imported into a Power BI Desktop data model
- R must be installed on the local machine
- Only data frames are imported
- Columns typed as Complex and Vector are not imported
- Can be refreshed with a gateway

# Describing additional capabilities

---

- Generating R visuals



- R visuals render from R script, accepting input fields
- Benefits:
  - Leverage the voluminous and growing number of out-of-the-box plots available in R
  - Easily customize R visuals by developing the script
  - Combine advanced analytics in visuals
  - Interact with R visuals in Power BI Desktop (filter, and cross-filter are supported)



# Describing additional capabilities

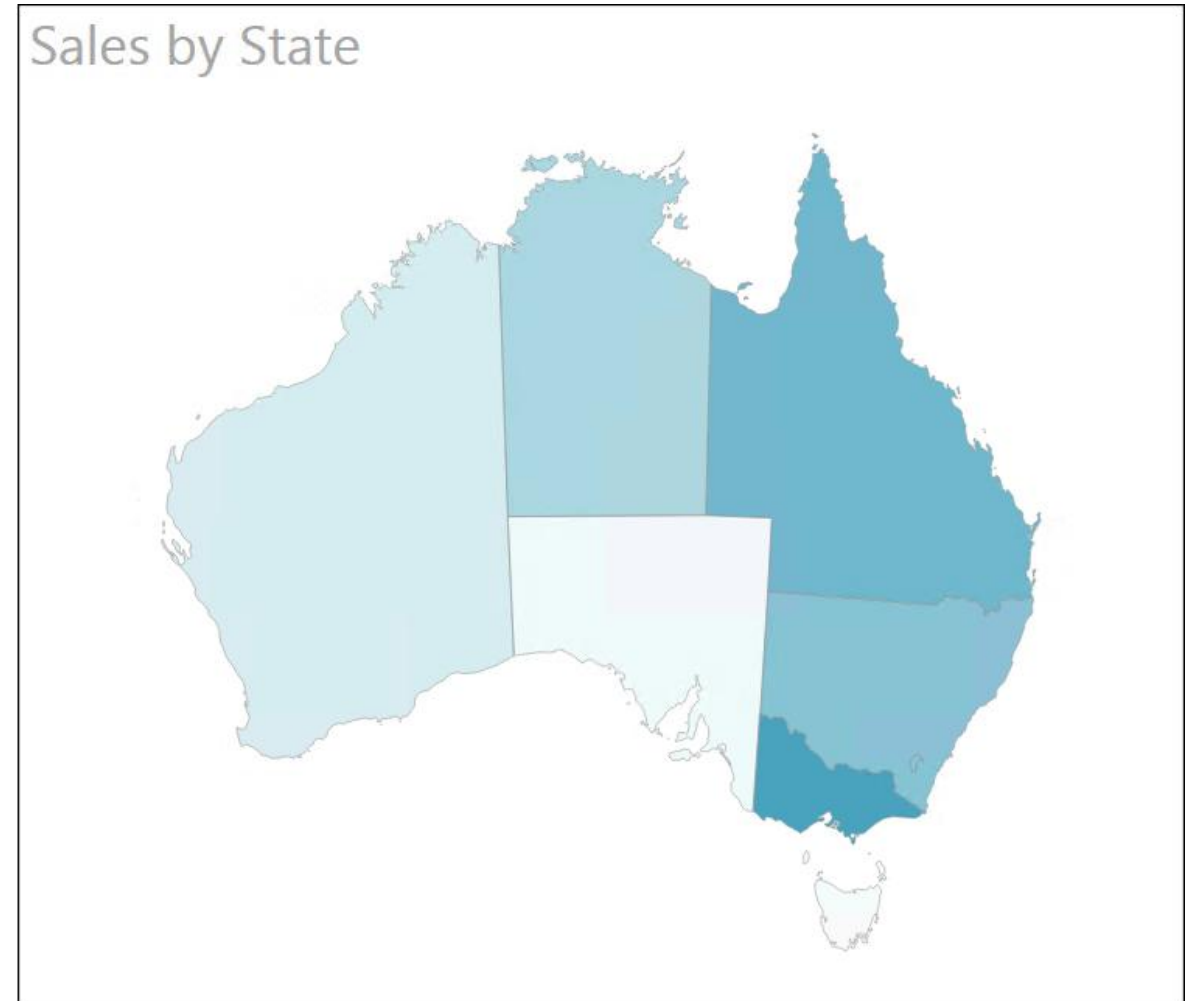
---

- Shape maps (Preview)
- A **Shape Map** visual shows relative comparisons of regions on a map by applying different colors to different regions
- In contrast to the **Map** visual, **Shape Map** cannot show precise geographical locations of data points on a map; instead, its main purpose is to show relative comparisons of regions on a map by coloring them differently

# Describing additional capabilities

---

- Shape maps (Preview): Continued
- Custom maps can be used provided they are in the TopoJSON format
- Possibilities: Display geographical, seating arrangements, floor plans, etc.



# Additional Capabilities of Power BI

- Exploring with Q&A

Ask a question about the data on this dashboard

- In the dashboard canvas, use Q&A to ask natural language questions
- The dashboard must contain at least one tile to connect to a dataset, and
- Only cloud-based cached datasets are supported
- Q&A helps formulate an appropriate question
- Q&A will format your question with suggestions, auto-complete, and even spelling corrections
- Featured questions can be added to datasets
  
- Only English is supported

# Additional Capabilities of Power BI

- Exploring with Q&A
  - Begin by asking “Show...”
  - For example:
    - Show <table>
    - Show <aggregate function><numeric column>
    - Show <measure> by <column>
    - Show <column> where <column> is before <date value>
    - Show <measure> where <column> is before <date value> sorted by <measure>
    - Show <measure> where <datetime column> is [this | last] [year | month | hour | minute | second]
  - Optionally, Q&A responses can be:
    - Customized by using the Visualizations or Fields pane
    - Pinned to the dashboard

# Additional Capabilities of Power BI

- Example of exploring with Q&A



Visualizations

Fields

Search

Date

Manager

Sales

Axis

Month

Legend

Drag data fields here

Values

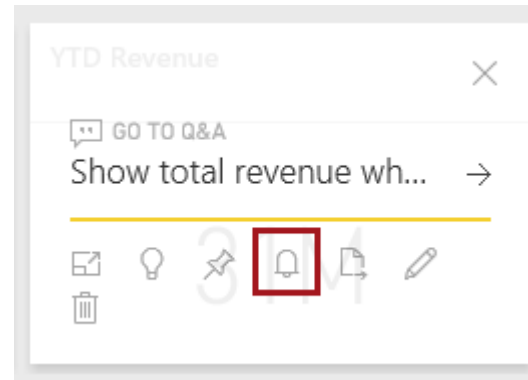
Revenue

# Additional Capabilities of Power BI

- Data alerts
  - notify users of changes beyond set limits
  - configured on dashboard tiles
    - Fully synchronized across platforms; set and view data alerts in the mobile apps, and in the Power BI service
    - Alert notifications provide details and links to view the tile or dashboard, and insights
  - Limitations:
    - Alerts are currently not supported for card tiles with date/time measures
    - Alerts only work with numeric data types
    - Alerts only work on data that is refreshed; they do not work on static data

# Additional Capabilities of Power BI

- Example of data alert



YTD REVENUE  
Manage alerts

+ Add alert rule

Alert for YTD Revenue

Active  
 On

Alert title  
Alert for YTD Revenue

Set alerts rule for  
Revenue

Condition Threshold  
Above 0

Maximum notification frequency  
 At most every 24 hours  
 At most once an hour

Alerts are only sent if your data changes.

By default, you'll receive notifications on the service in the notification center.  
 Send me email, too

# Additional Resources

- Power BI Whitepapers
  - <https://docs.microsoft.com/en-us/power-bi/whitepapers>
- Power BI Desktop knowledge base
  - <https://support.powerbi.com/knowledgebase/topics/68530-power-bi-desktop>
- Tips and tricks for creating reports in Power BI Desktop
  - <https://support.powerbi.com/knowledgebase/articles/464157-tips-and-tricks-for-creating-reports-in-power-bi-d>
- Power BI User Group (PUG)
  - Webinars <https://www.pbusergroup.com/engage/training>
  - Webinar Recordings <https://www.pbusergroup.com/engage/webinarrecordings>
  - Pug Exchange <https://www.pbusergroup.com/engage/redirectopenforum>
  - San Antonio Chapter <https://www.pbusergroup.com/communities/sanantonio>



# Additional Resources

- YouTube channels
  - Guy in a cube <https://www.youtube.com/channel/UCFp1vaKzpfvoGai0vE5VJ0w>
  - Microsoft Power BI <https://www.youtube.com/channel/UCy--PYvwBwAeuYaR8JLmrfg>
  - Avi Sing <https://www.youtube.com/channel/UCRNmSv7mAPYiC0Y40TJijAw>
  - Excel Is Fun <https://www.youtube.com/user/ExcelsFun>
- Other Web sources
  - DAX basics <https://docs.microsoft.com/en-us/power-bi/desktop-quickstart-learn-dax-basics>
  - Data Analysis Expressions (DAX) Language <https://social.technet.microsoft.com/wiki/contents/articles/677.power-bi-data-analysis-expressions-dax-language.aspx>

# Summary

- Power BI Desktop enables simple and quick visualization of data
  - Creating queries
  - Configuring relationships
  - Enhancing the model with hierarchies and calculations
  - Designing reports, including the use of custom visuals
  - Publishing to Power BI
- Power BI Pro (online) – trial for 60 days

# Summary

Feature	Power BI (Free)	Power BI Pro
Data capacity limit	1 GB/user	10 GB/user
Consume content that is scheduled to refresh	Daily	Up to 8 times per day
Data streaming rate	10K rows/hour	1M rows/hour
Consume live data sources with full interactivity	No	Yes
Access on-premises data with gateways	No	Yes
Share dashboards	Yes	Yes
All other collaboration features	No	Yes
Row-level security for users and groups	No	Yes

# Credits

- Power BI User Group – presentation Creating a Power BI Report
  - <https://www.pbusergroup.com/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=8429432a-66a8-6567-7c42-6dc53b426465&forceDialog=1>
- Microsoft Training Power BI Materials – Introduction
  - <https://cp-mlxprod-static.microsoft.com/015710-1000/en-us/downloads%2Fpowerbi-presentation.pptx>
- WUG Lectures
  - Jana Babáčková – Příprava ke zkoušce DA-100: Analyzing Data with Microsoft Power BI, SQL Server Bootcamp 2021.
    - <https://www.wug.cz/zaznamy/729-SQL-Server-Bootcamp-2021-Priprava-ke-zkousce-DA-100-Analyzing-Data-with-Microsoft-Power-BI>
  - Štěpán Rešl – Konec nudných reportů v Power BI, SQL Server Bootcamp 2021
    - <https://www.wug.cz/zaznamy/727-SQL-Server-Bootcamp-2021-Konec-nudnych-reportu-v-Power-BI>
  - Karel Rejthar – Power BI a mapové vizualizace, SQL Server Bootcamp 2020
    - <https://www.wug.cz/zaznamy/647-SQL-Server-Bootcamp-2020-Power-BI-a-mapove-vizualizace>