CELEBAL TECHNOLOGIES

Best Practices for Dataset and Report Design in Power BI Desktop

Microsoft Partner

Microsoft

Gold Data Analytics
Gold Data Platform
Gold Cloud Ops
Azure AI Inner Circle

ICICI Bank

Best Practices for Dataset and Report Design in Power BI Desktop

- Managing Dataset Authorship
- Version Control for Power BI Content
- Separating Datasets from Reports
- Using Row-Level Security for Securing Data
- Using Templates to Speed Up and Standardize Report Development
- Using Power BI Data Source Files
- Minimizing the Amount of Data Loaded
- Reducing Dataset Size
- Naming of Tables, Columns, and Measures
- Minimizing the Amount of Data Displayed on a Page

Best Practices for Dataset and Report Design in Power BI Desktop

- Focusing on Quality of Content
- Focusing on Information Design Principles
- Verifying Accessibility on Reports
- Using Certified Custom Visuals
- Using Custom Data Connectors
- Creating Dashboards to Unify Key Metrics from Various Reports

Power BI

Managing Dataset Authorship

Managing Dataset Authorship

Monitoring of who builds datasets, where these datasets are published, and how many datasets are published, are essential if you are going to avoid thorny issues later in your project. If you allow unrestricted dataset authorship, you will find that duplicate datasets are created. Often that diminishes many of the beneficial aspects of self-service BI. Hence, it is important to ensure that:

✓ All dataset and report authors are properly trained.

✓ Authors follow your organization’s documented best practices and standards when building datasets and reports.

✓ Duplicate datasets and reports are avoided whenever possible, and authors understand why this is crucial.

✓ Datasets and reports are thoroughly tested before they are published.

✓ Datasets and reports always have an owner identified who is responsible for maintenance, questions, and data quality issues.

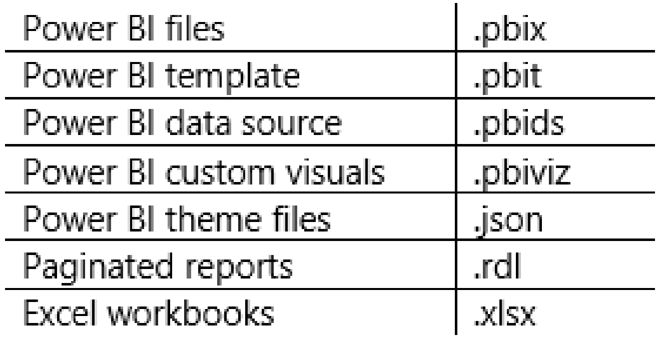
Many considerations such as these relate more to the ‘process’ and ‘people’ side of the equation, rather than strictly a ‘technology’ issue

Power BI

Version Control for Power BI Content

Version Control for Power BI Content

A significant amount of work goes into designing and validating content. Therefore, all original .pbix files should be stored in a safe location. This includes various Power BI file types, such as:



Power BI files	.pbix
Power BI template	.pbit
Power BI data source	.pbids
Power BI custom visuals	.pbiviz
Power BI theme files	.json
Paginated reports	.rdl
Excel workbooks	.xlsx

Use consistent, unchanging file names. This technique of changing the file name does not work well in Power BI because data refresh schedules, pinned tiles, and related reports are dependent upon use of a consistent file name.

Source Control Systems

Source control, such as Git, is a common requirement for IT and centralized BI teams. Consider using Power BI template files (.pbit files) instead of .pbix files so that:

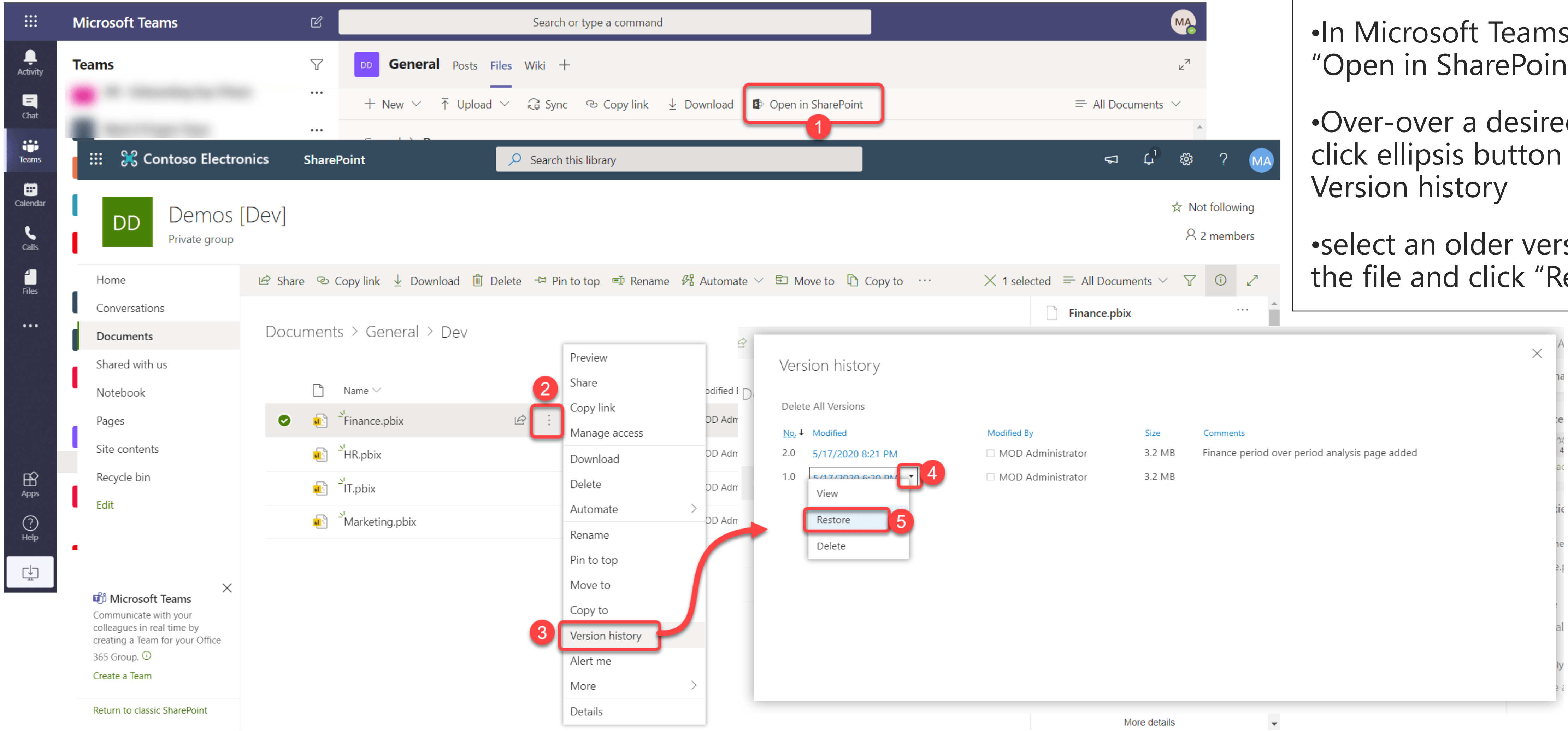
- The files you are storing in version control are as small as possible.
- Data is not stored in the source control system.

The ability to connect to Power BI datasets stored in Premium capacity, using the XMLA endpoint, allows tools that are currently used for Analysis Services development and deployment to be used for developing Power BI datasets as well. This capability is dependent on read/write capability for the XMLA endpoint (which is a Premium feature) offers another alternative for storing the definition of a dataset in source control.

File Versioning Systems

It is strongly recommend storing .pbix files in a file sharing solution which saves historical copies of files. The number of versions retained is typically configurable. This allows access to the previous versions, which is particularly helpful if a change needs to be reverted. Recent advancements in OneDrive file size limits, and the efficiency of sync operations, make this a very seamless way to work.

Version Control for Power BI Content



No.	Modified	Modified By	Size	Comments
2.0	5/17/2020 8:21 PM	MOD Administrator	3.2 MB	Finance period over period analysis page added
1.0	5/17/2020 6:20 PM	MOD Administrator	3.2 MB	

Restoring an Older Version

- In Microsoft Teams click “Open in SharePoint”
- Over-over a desired file -> click ellipsis button -> Version history
- select an older version of the file and click “Restore”

Power BI

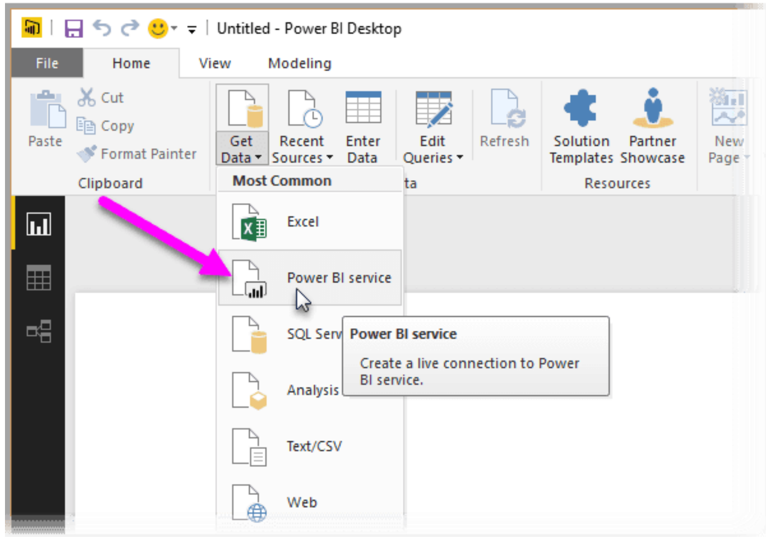
Separating Datasets from Reports

Separating Datasets from Reports

1. Shared Power BI Dataset (a .pbix file with dataset only)

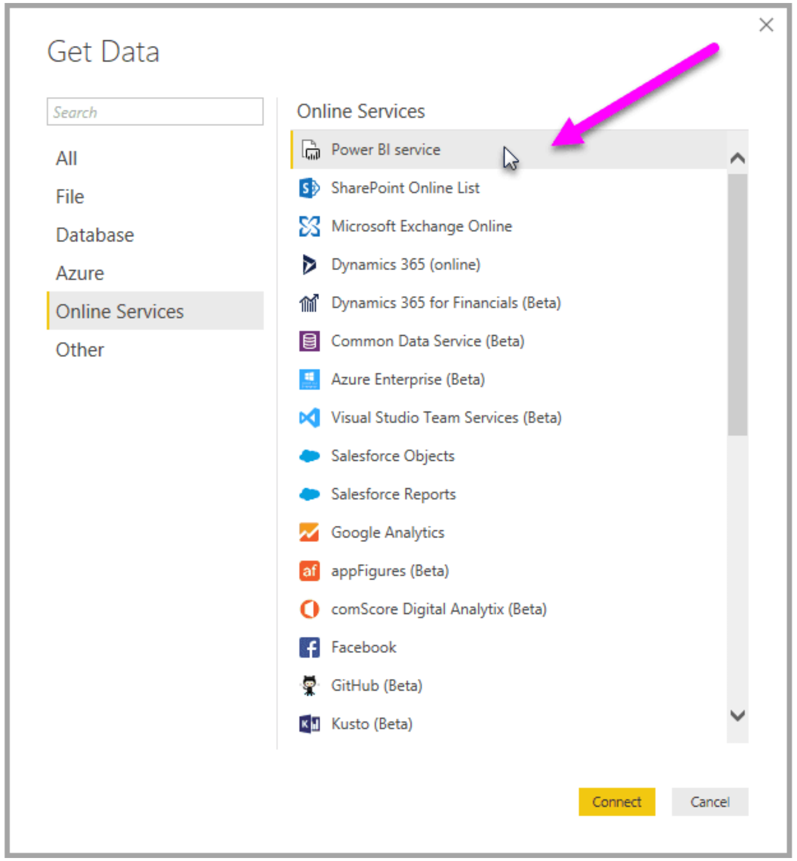
Create a dataset in a .pbix file without any analytical reports in that file.

This .pbix file is published to the Power BI service which creates a new dataset. Report creators can create new reports in Power BI Desktop which connect to the published dataset. This creates what is called a Power BI service live connection, and behaves in a very similar way to creating a live connection to Analysis Services.



Power BI service

Create a live connection to Power BI service.



We recommend you use both shared and certified datasets to the extent it is practical.

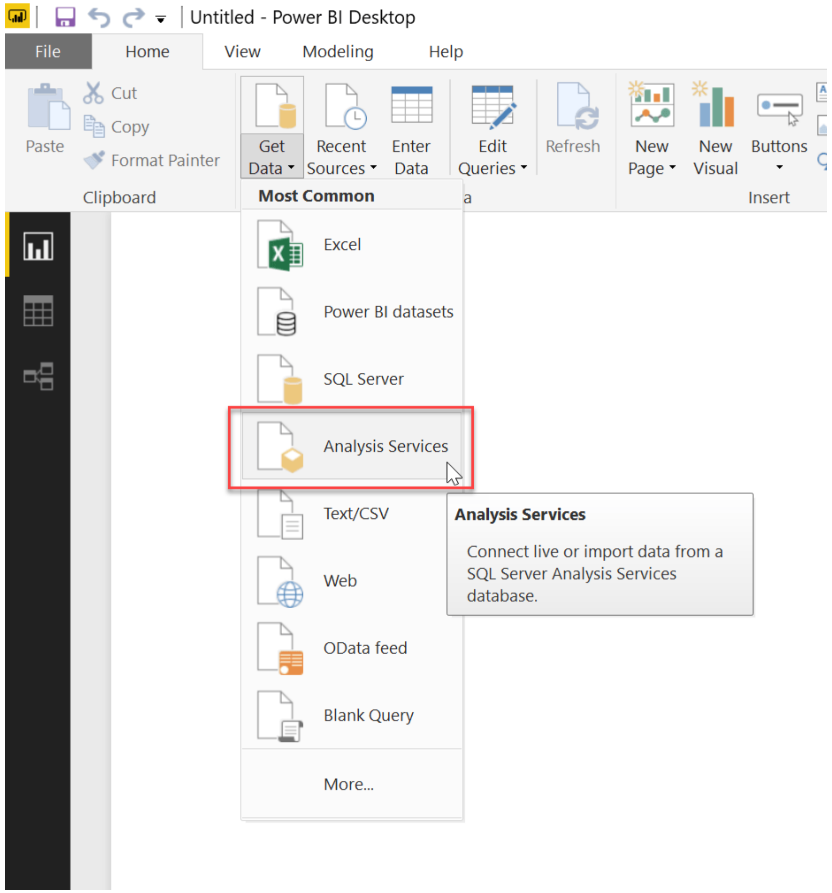
QUICK TIPS

Instead of an empty report, you can use the report pages for dataset documentation, data quality testing, or a change log. This type of report would only be seen by dataset authors.

Separating Datasets from Reports

2. Use of Analysis Services

Another option is to use Azure Analysis Services or SQL Server Analysis Services as a semantic model for storage of data.

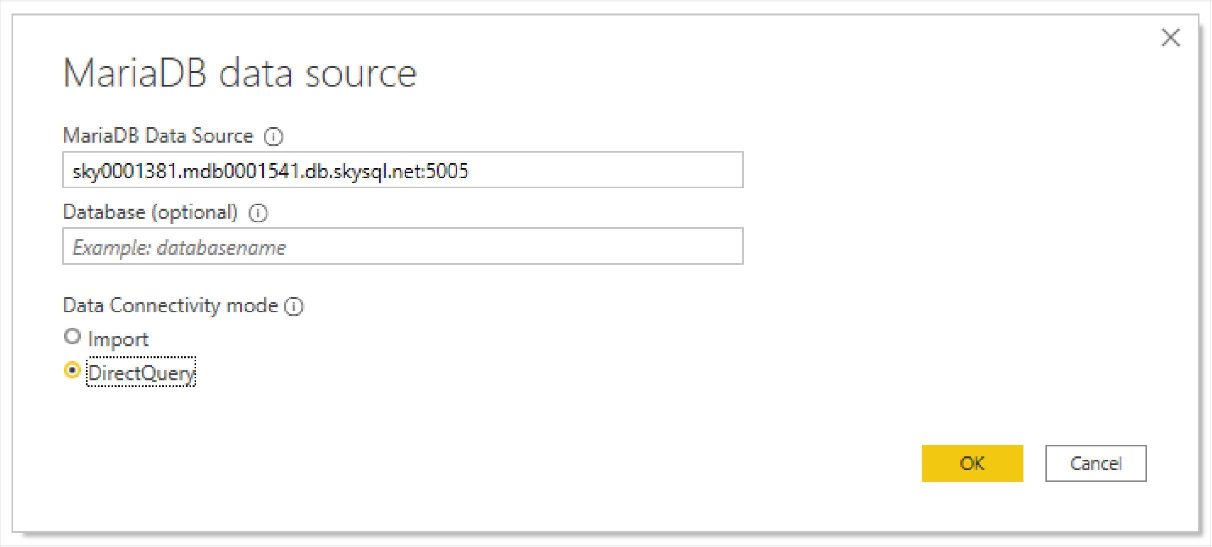


Analysis Services

Connect live or import data from a SQL Server Analysis Services database.

3. Use of DirectQuery

Another option is to use DirectQuery. Conceptually, this is like the Live Connection options— however, the data is sourced from a relational database instead.



MariaDB data source

MariaDB Data Source: sky0001381.mdb0001541.db.skysql.net:5005

Database (optional): *Example: databasename*

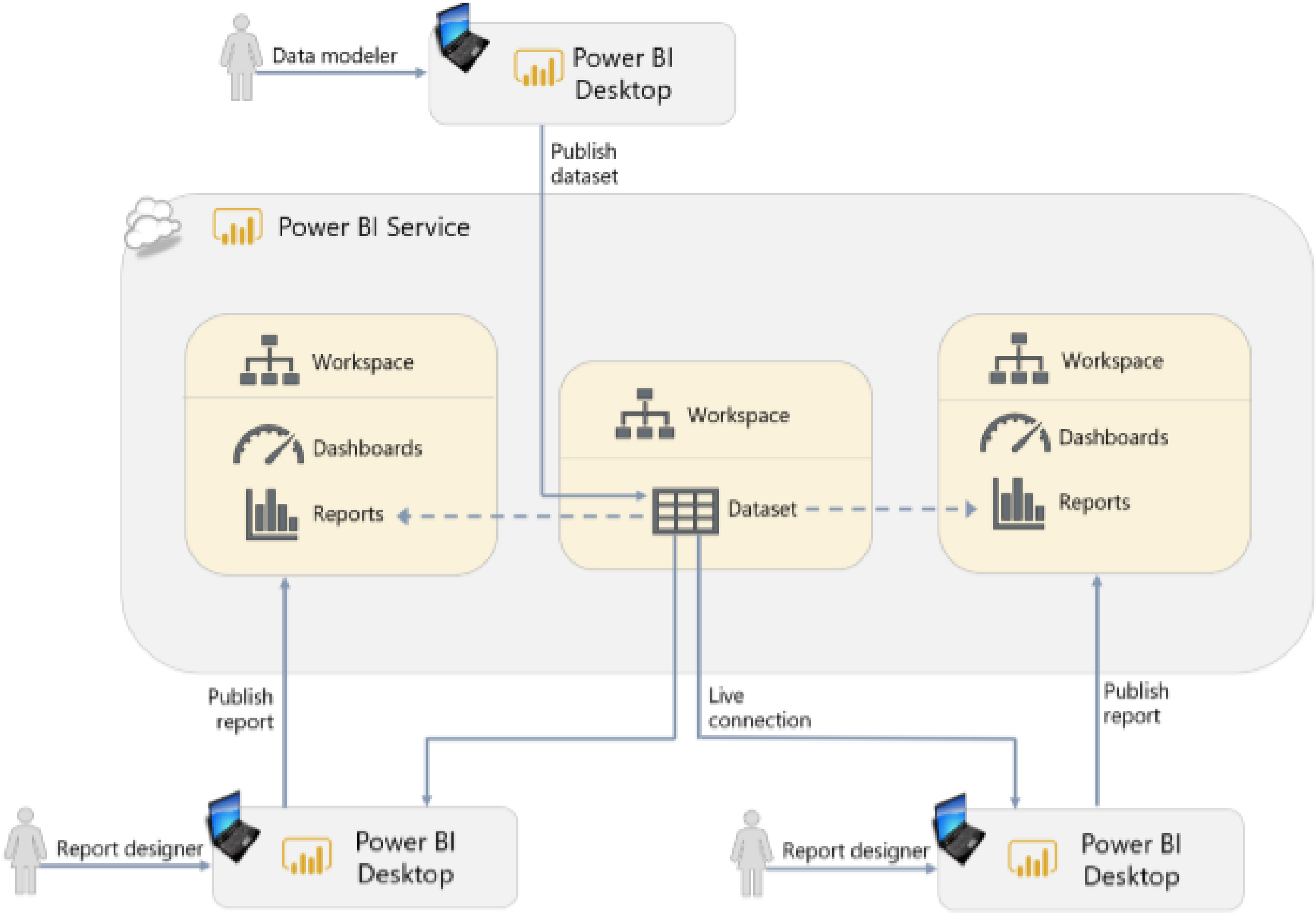
Data Connectivity mode

○ Import

● DirectQuery

Separating Datasets from Reports

The following graphic depicts of use of a shared dataset



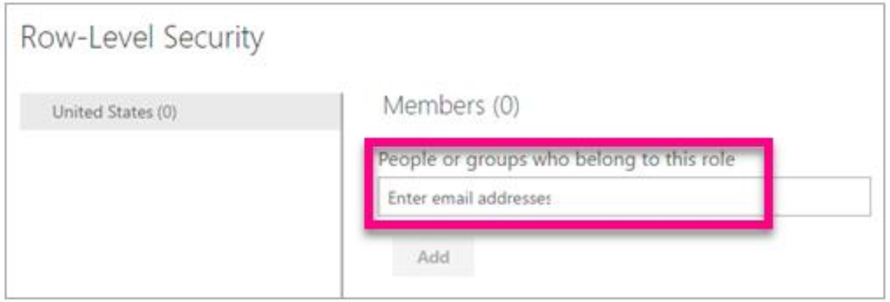
Depiction of a shared dataset which is used by multiple separate reports

Power BI

Using Row-Level Security for Securing Data

Using Row-Level Security for Securing Data

Use of row-level security (RLS) ensures the data in the dataset is secured based on each report consumer's job role and need-to-know. When RLS has been defined in a dataset, the same report can be distributed to multiple audiences which saves time and effort. With RLS defined on the underlying dataset, when a report is viewed by users in the USA, the report displays USA sales; when that same report is viewed by users in Germany it only shows sales for Germany.



Row-level security is defined in Power BI Desktop, and invoked in the Power BI service, for users who do **not** have edit permissions. This includes users with: Permission to view an app, or Viewer permissions in a workspace

Without RLS, if a user may access a dataset, they are permitted to see all the data in the dataset – not just the data shown in data visualizations.

QUICK TIPS

It is strongly recommended to use RLS when appropriate rather than creating multiple versions of the same dataset or report.

Pros and Cons on using Row-Level Security in Import vs Direct Query

Parameters	Import Query mode	Direct Query mode
RLS implementation	✓	✓ (with limitations)
Hierarchical RLS level	✓	×
Create Hierarchies at different levels	✓	×
Column specific RLS level	✓	✓
Use DAX Expression	✓	✓ (with limitations)
Time Intelligence Functions	✓	×
Example:	Can create organizational hierarchies and implement RLS on different levels using functions like PATH and PATHITEM.	Can create RLS on a specific column, for example ‘Country’ and assign it to specific users through ‘Manage Roles’ → Column value name(US, India, Russia).

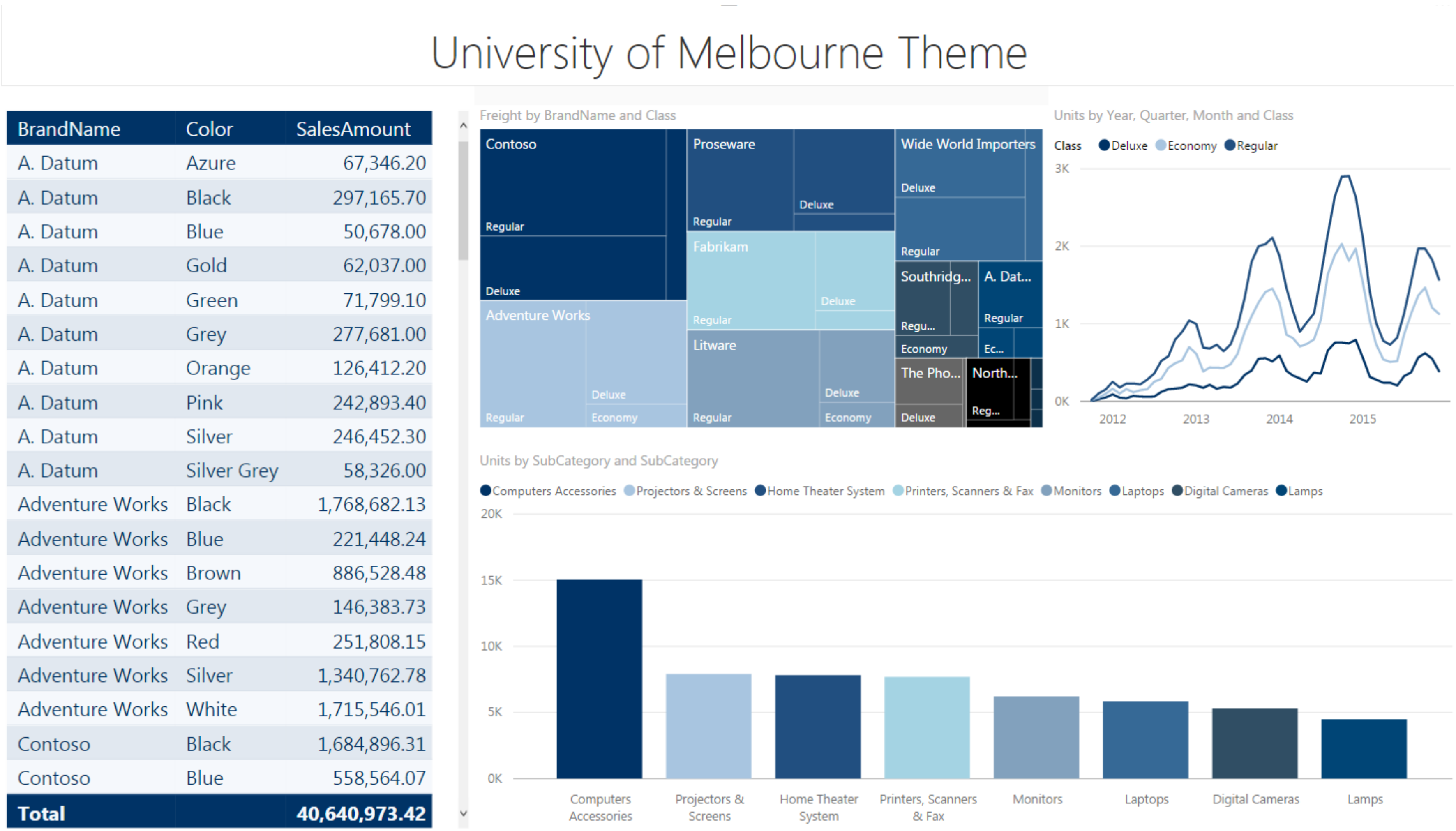
Power BI

Using Templates to Speed Up and Standardize Report Development

Using Templates to Speed Up and Standardize Report Development

Instead of starting with an empty .pbix file, templates allow report creators to create a new file with:

- A corporate color scheme already in place (it is also possible to store a color scheme as a Report Theme which can be imported into Power BI Desktop).
- Corporate branding already applied to pages.



University of Melbourne Theme

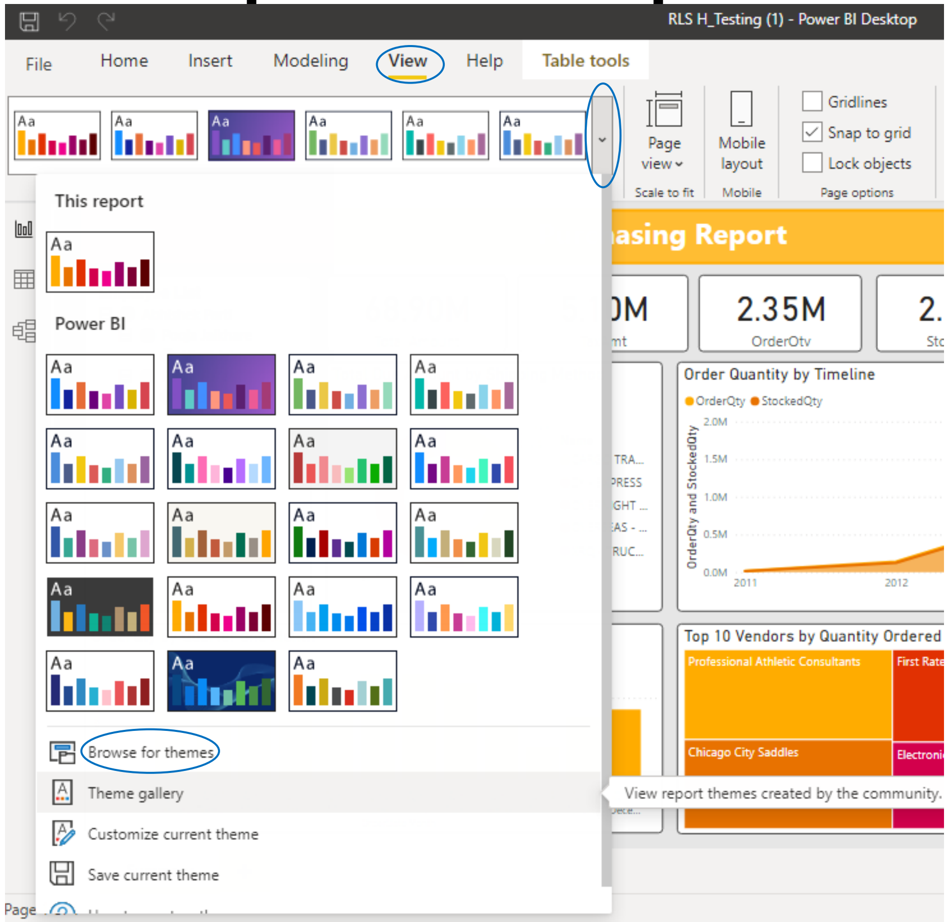
BrandName	Color	SalesAmount
A. Datum	Azure	67,346.20
A. Datum	Black	297,165.70
A. Datum	Blue	50,678.00
A. Datum	Gold	62,037.00
A. Datum	Green	71,799.10
A. Datum	Grey	277,681.00
A. Datum	Orange	126,412.20
A. Datum	Pink	242,893.40
A. Datum	Silver	246,452.30
A. Datum	Silver Grey	58,326.00
Adventure Works	Black	1,768,682.13
Adventure Works	Blue	221,448.24
Adventure Works	Brown	886,528.48
Adventure Works	Grey	146,383.73
Adventure Works	Red	251,808.15
Adventure Works	Silver	1,340,762.78
Adventure Works	White	1,715,546.01
Contoso	Black	1,684,896.31
Contoso	Blue	558,564.07
Total		**40,640,973.42**

Using Templates to Speed Up and Standardize Report Development

- Connections to commonly used data sources already created.
- Parameterized queries connected to a data source. This is particularly powerful because it allows your users to reuse queries that someone else has created yet customize the way these queries behave.
- Commonly used DAX measures already created inside the dataset.
- Sample visuals which employ data visualization and accessibility best practices.

Tips for delivering templates. Templates can be saved by creating JSON files and setting the color codes in accordance to the brand standardization. A sample JSON is shown below:

```
{
      "name": “ICICI standard",
      "dataColors": ["#568410",
"#3A6108", "#70A322", "#915203",
"#D79A12", "#bb7711", "#114400",
"#aacc66"],
      "background":"#FFFFFF",
      "foreground": "#3A6108",
      "tableAccent": "#568410"
 }
```



Include template links directly within a report or app that you think users are likely to want to customize. This helps you manage the experience. It is also more secure because it prevents users from downloading data unnecessarily.

Power BI

Using Power BI Data Source Files

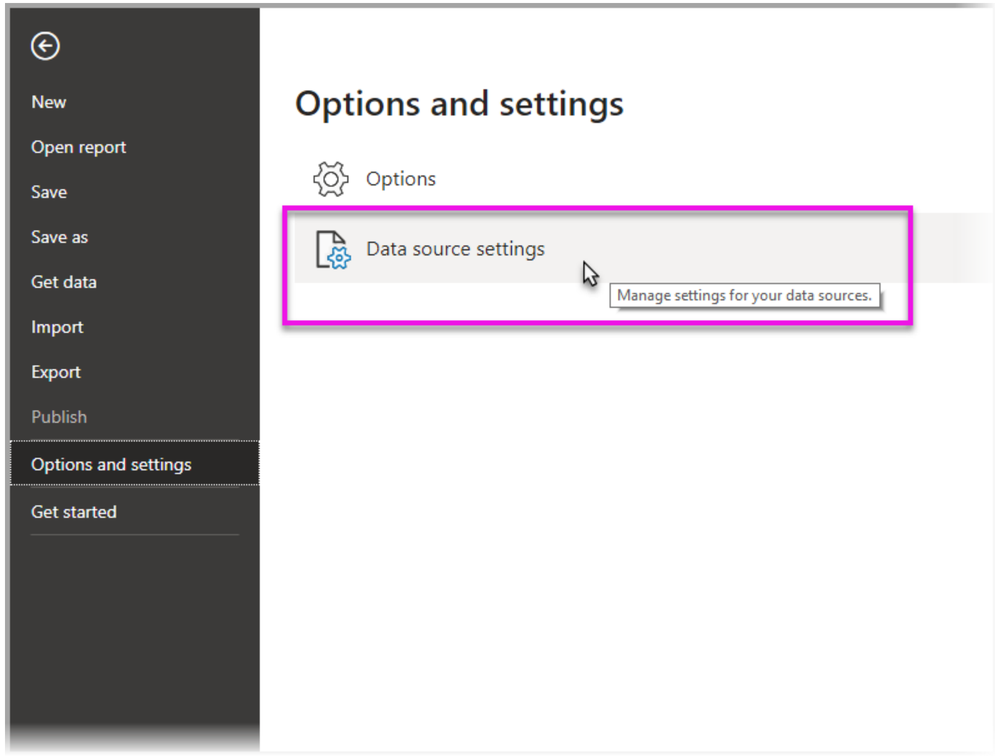
Using Power BI Data Source Files

To help report creators save time, connection information for commonly used data sources may be made available through Power BI data source files (.pbids file format).

A .pbids file is unlike a template: it does not contain table information. It merely contains connection string information to simplify connectivity for self-service dataset authors.

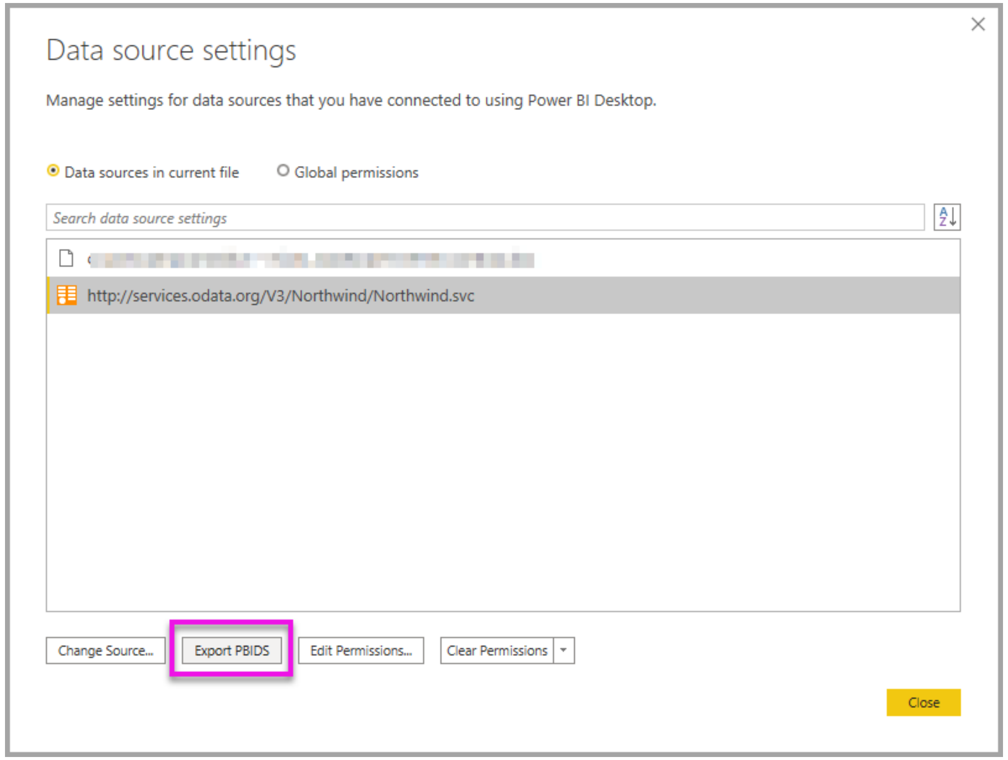
To create the PBIDS file, select **File > Options and settings > Data source settings**

In the dialog that appears, select the data source you want to export as a PBIDS, and then select **Export PBIDS**.



Options and settings

Manage settings for your data sources.



Data source settings

Manage settings for data sources that you have connected to using Power BI Desktop.

● Data sources in current file ○ Global permissions

Search data source settings

[redacted]

http://services.odata.org/V3/Northwind/Northwind.svc

Power BI

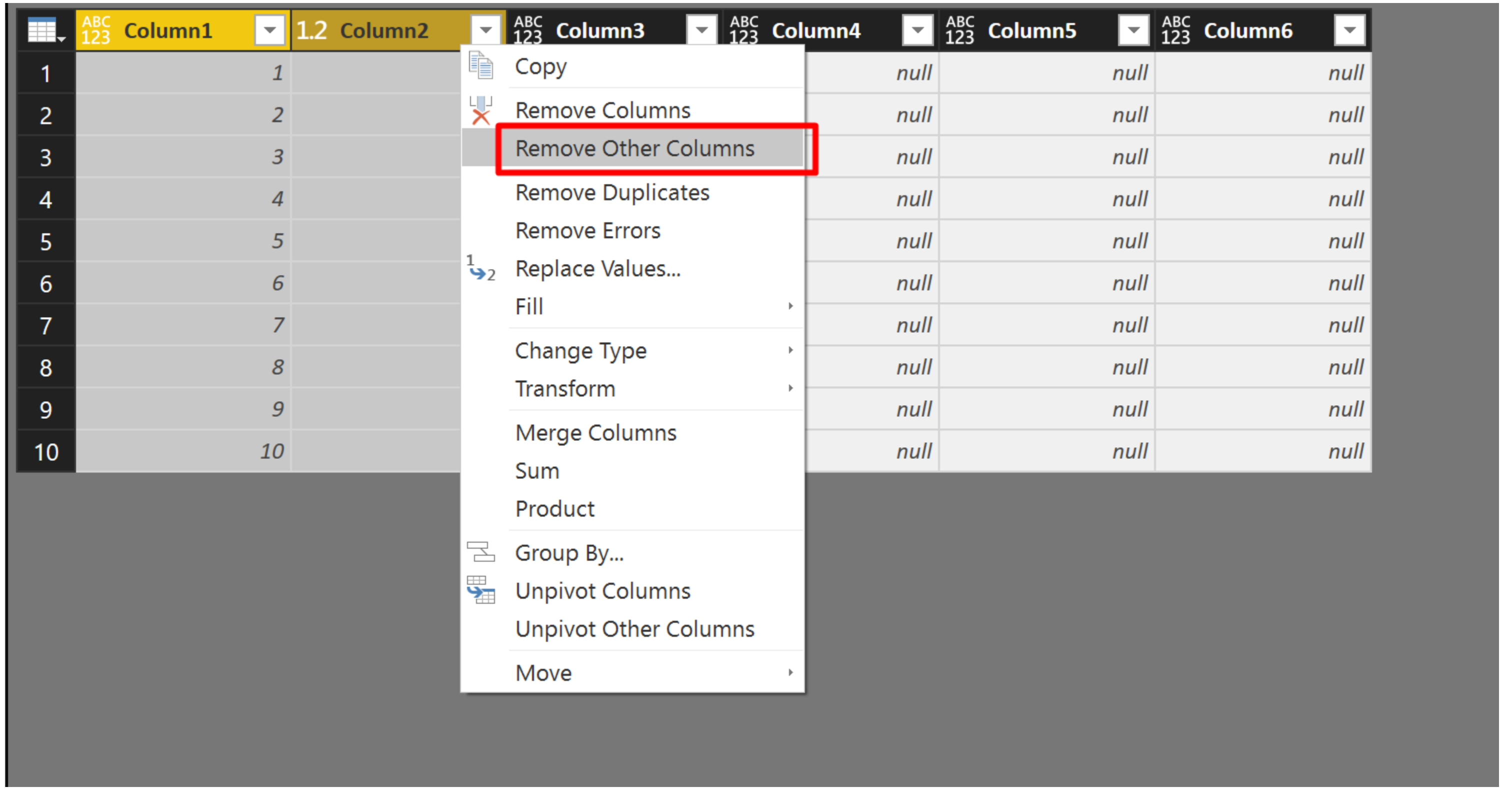
Minimizing the Amount of Data Loaded

Minimizing the Amount of Data Loaded

Two of the main ways to reduce the amount of data stored in a dataset include:

1. Removing columns from source tables in the Query Editor.
2. Filtering rows from the source tables down to just the subset of data which is required.

This can be done with hard-coded filters in the Query Editor or with query parameters.



	Column1	Column2	Column3	Column4	Column5	Column6
1	1			null	null	null
2	2			null	null	null
3	3			null	null	null
4	4			null	null	null
5	5			null	null	null
6	6			null	null	null
7	7			null	null	null
8	8			null	null	null
9	9			null	null	null
10	10			null	null	null

Make a habit of minimizing dataset size The best policy is to restrict the amount of data loaded as much as possible at an early stage. Add additional data, if it is necessary, rather than loading everything you possibly can initially.

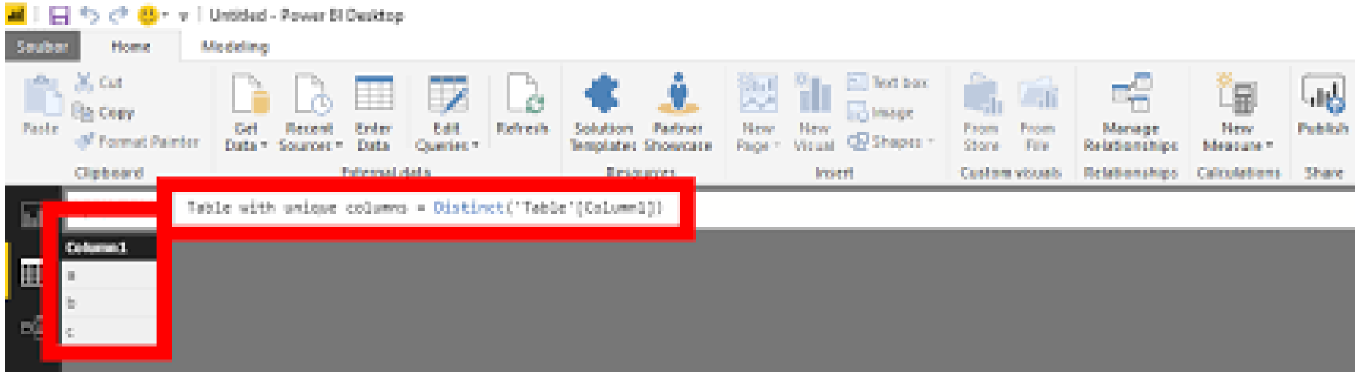
Power BI

Reducing Dataset Size

Reducing Dataset Size

Techniques which can reduce dataset size include:

- Reduce the number of distinct values in a column. One way to do this is to eliminate the time element from date/time columns if time is not relevant. This vastly reduces the number of possible combinations, thus reducing the dataset size and memory consumption.

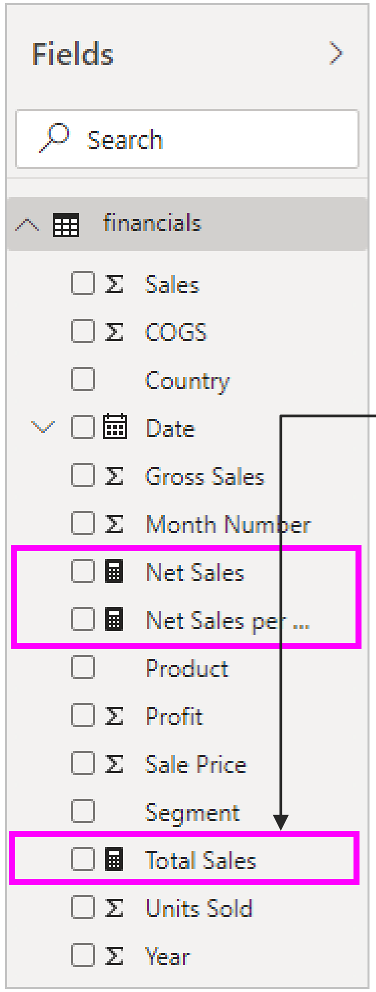


```
Table with unique columns = Distinct('Table'[Column1])
```

Column1
a
b
c

Distinct function measure

- Use calculated measures rather than calculated columns when possible.



Fields

Search

financials

- □ Σ Sales
- □ Σ COGS
- □ Country
- □ Date
- □ Σ Gross Sales
- □ Σ Month Number
- □ Net Sales
- □ Net Sales per ...
- □ Product
- □ Σ Profit
- □ Σ Sale Price
- □ Segment
- □ Total Sales
- □ Σ Units Sold
- □ Σ Year

Power BI

Naming of Tables, Columns, and Measures

Naming of Tables, Columns, and Measures

The objective for naming tables, columns, and measures is to avoid the misinterpretation of results. We recommend having naming conventions, including:

- **Clear, friendly names** which will display well on reports. Ex: “Product Name” looks far nicer on a report than ProdNm.
- **Clarity over brevity.** Ex: “Product Code” is far clearer for users than Product.
- **Prefixes for clarity**. Ex: if the dataset contains Product Name, Customer Name, and Sales Rep Name, they should all contain prefixes. This makes it clear which name is being used. Though the report creator can tell which name it is from the table it comes from, a report consumer does not have the same luxury if only “name” is displayed on the report. This also reduces the frequency of visual-level renaming.
- **Minimize use of acronyms and jargon**. If use of an acronym is required, a key or a tooltip can be of great help for others to understand the meaning.

Expose a column once and only once in the dataset

If a column exists multiple times in a dataset (perhaps for relationships), it should be exposed for reporting only once. Allowing the same column to be visible in a model multiple times is confusing for anyone creating reports from that dataset, and it can lead to incorrect results if the wrong column is used in a visual.

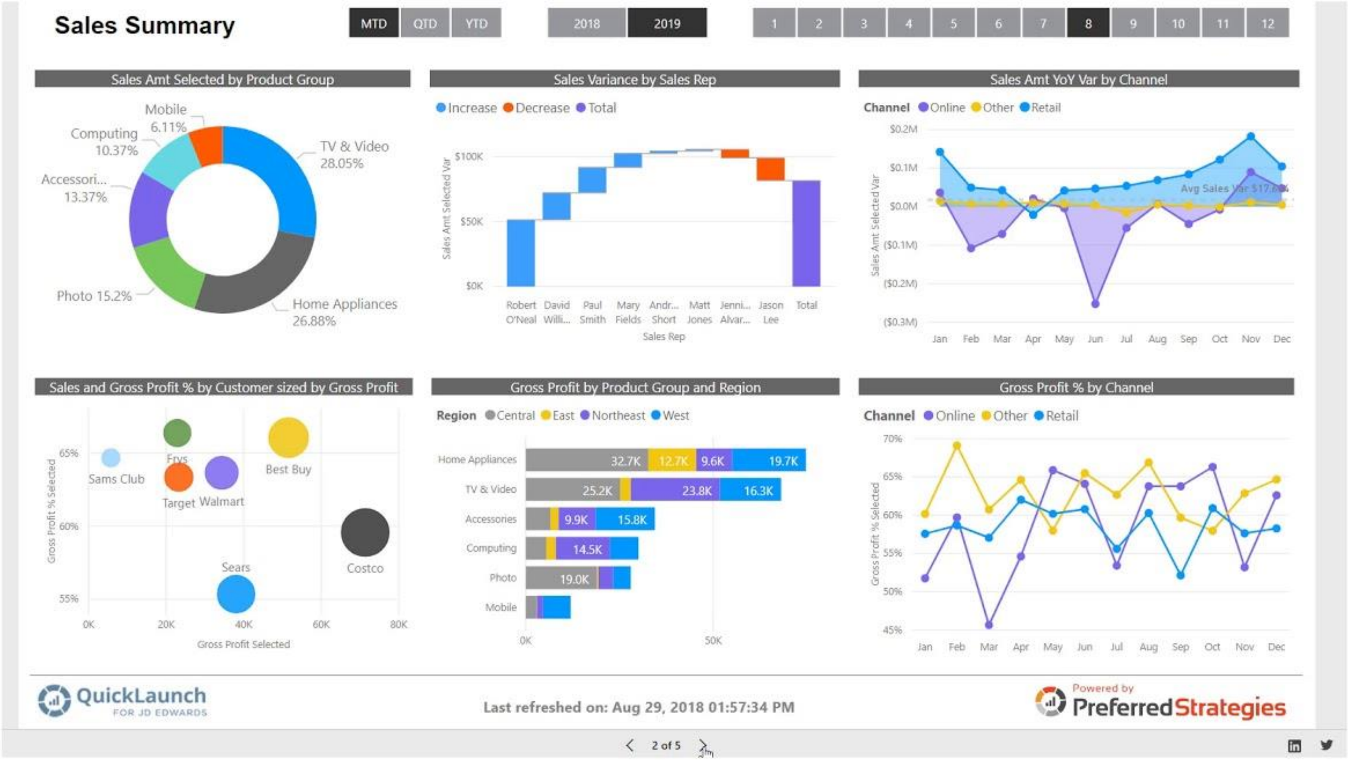
QUICK TIPS

The Description property for fields can be used as a data dictionary to help report authors understand the data model. This type of information can be invaluable.

Power BI

Minimizing the Amount of Data Displayed on a Page

Minimizing the Amount of Data Displayed on a Page



Sample report of an ideal Power BI Report

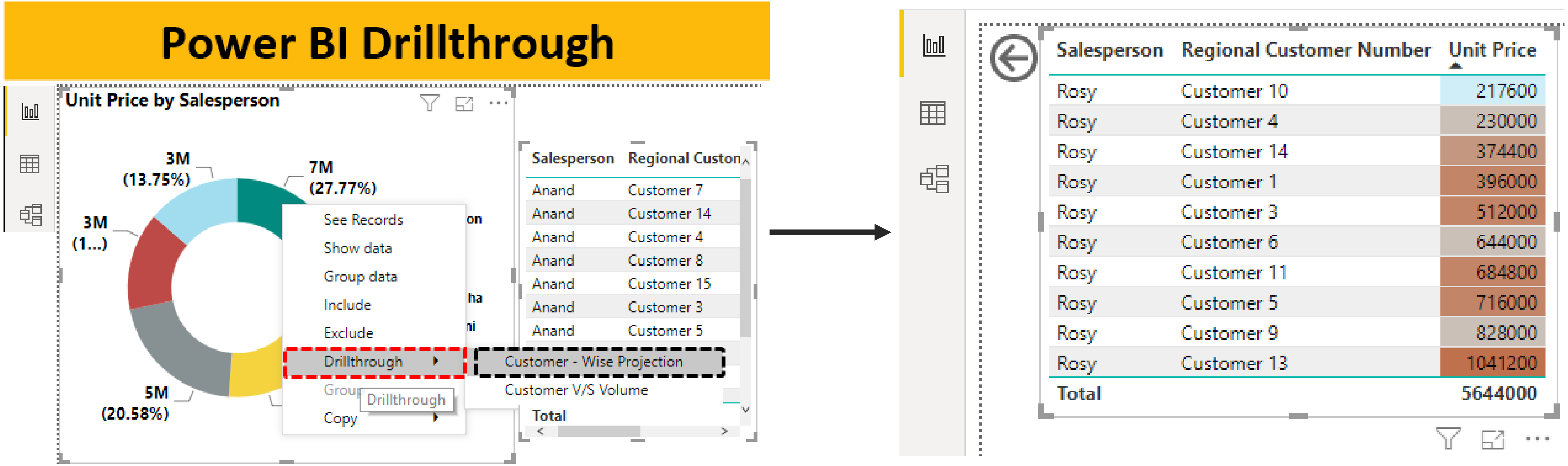
As the number of visuals on a page increases, so do the number of DAX queries requesting data from the dataset. Inevitably each query that is executed will increase the overall render time of the report.

- ✓ Avoid "data dump"-style reports with tables containing hundreds of columns and thousands of rows, usually created when the user expects to export this data to Excel.
- ✓ Consider using slicers and filters in your reports to allow the user to select just the data that is needed to be displayed at any one time.

The performance of a report will depend significantly on storage of the dataset: imported, Analysis Services, or DirectQuery.

Minimizing the Amount of Data Displayed on a Page

✓ You may also consider using techniques such as bookmarks and drillthrough pages to reduce the amount of data displayed on a page. The drillthrough feature delays the rendering of detailed data until the user clicks on the link to drillthrough to an additional page (for instance, from a sales summary report to a sales detail report). In addition to improving performance, drillthrough can also simplify the higher-level reports thus improving the end user experience.



Salesperson	Regional Custon
Anand	Customer 7
Anand	Customer 14
Anand	Customer 4
Anand	Customer 8
Anand	Customer 15
Anand	Customer 3
Anand	Customer 5
Total	

Salesperson	Regional Customer Number	Unit Price
Rosy	Customer 10	217600
Rosy	Customer 4	230000
Rosy	Customer 14	374400
Rosy	Customer 1	396000
Rosy	Customer 3	512000
Rosy	Customer 6	644000
Rosy	Customer 11	684800
Rosy	Customer 5	716000
Rosy	Customer 9	828000
Rosy	Customer 13	1041200
Total		5644000

Power BI

Focusing on Quality of Content

Focusing on Quality of Content

- Have a hyper-focus on usability of the report for end users.
- Have a specific purpose for each page, and for each chart on each page.
- Include clear interactions and instructions for report consumers.
- Have a process for verifying accuracy of data presented, and accuracy of measures with different filter selections.
- Use bookmarks to create 'help' information to guide users.
- Carefully review the data displayed to assess if it can be easily misinterpreted.
- Solicit feedback from consumers for both effective data presentation and data quality issues.

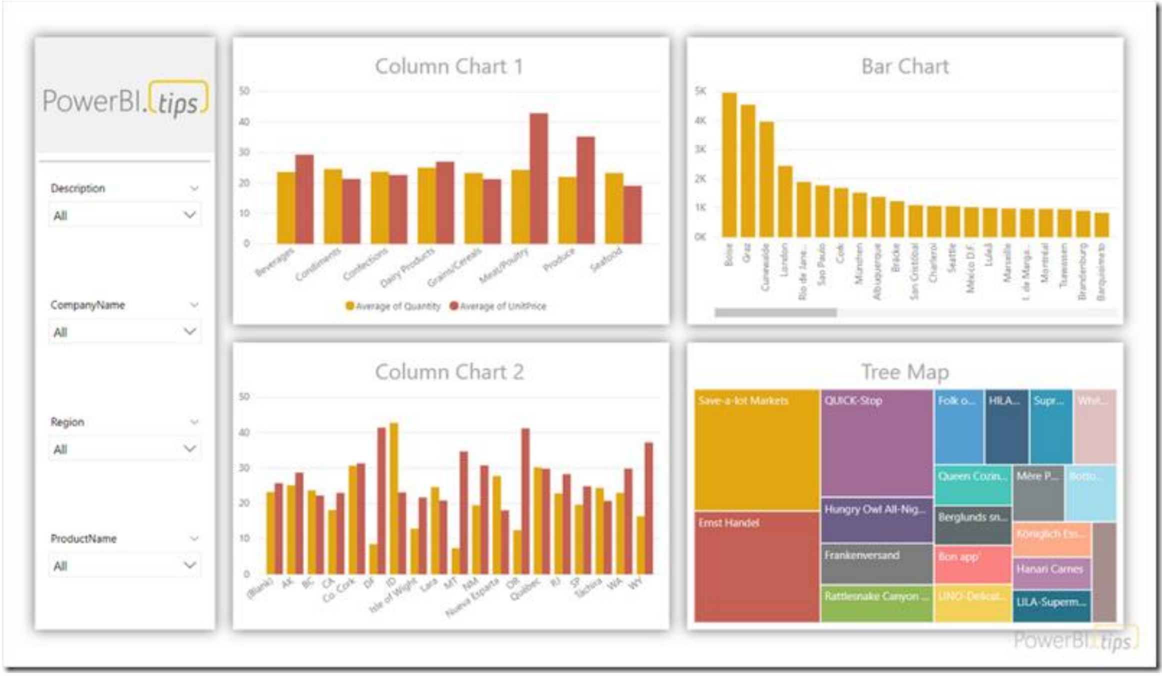
One way to get feedback from report consumers is with a survey link available alongside the report or within an app. This allows report creators to get structured feedback on the quality of content, and understand how reports are, or are not, meeting user needs.

Power BI

Focusing on Information Design Principles

Focusing on Information Design Principles

- Be intentional with selections for each visual. Focus on its ability to convey information, resisting the urge to use a variety of visual types for aesthetic appeal only.
- Use careful placement and a clear hierarchy for displaying information on the page.
- Use informative text so consumers can find subtle features like drill-through and/or find more information.
- Be consistent with placement of items which appear on multiple pages.
- Use a layout that focuses on telling the story you want to tell.
- Use color sparingly, and consistently, with a specific intention in mind.
- Use a color palette that follows visualization best practices. If you are required to follow a vibrant organizational color palette, investigate if you can mute the colors.
- Avoid clutter on the page to allow report consumers to focus on what is important.
- Address accessibility early in the design phase.



Description: All

CompanyName: All

Region: All

ProductName: All

Power BI

Verifying Accessibility on Reports

Verifying Accessibility on Reports

When designing reports, it is important to consider the accessibility needs of your users.

- Using very **clear titles, labels, and descriptions** can be extremely helpful to aid comprehension by all report consumers, including those who are visually impaired.
- Adding **alternative text descriptions** (alt text) for visualizations in a report helps users who use screen readers understand what a visualization is showing. Alt text should be set for every object that conveys meaningful information.
- Choosing a **color theme** that has a high degree of contrast between backgrounds and text helps make reports easier to read, especially for users with vision impairment. There are many desktop apps and online sites to help with choosing colors with appropriate contrast ratios, including Colour Contrast Analyser and WebAIM Contrast Checker. There are also helpful sites to simulate what colors look like to someone who experiences color deficiency, such as Color Blindness Simulator and WhoCanUse.
- Using **visual indicators**, in addition to color, helps report consumers who are colorblind, and therefore unable to understand good and bad with color alone. Red and green can be particularly difficult to distinguish. This is important to keep in mind, especially when using conditional formatting.
- Setting the **tab order for visuals** assists visually impaired report viewers with how to go about viewing and processing the report in a logical order. Any images or objects which are purely decorative can be omitted to aid with this navigation.

Get data

Analyze

Visualize

Publish

Collaborate

Power BI

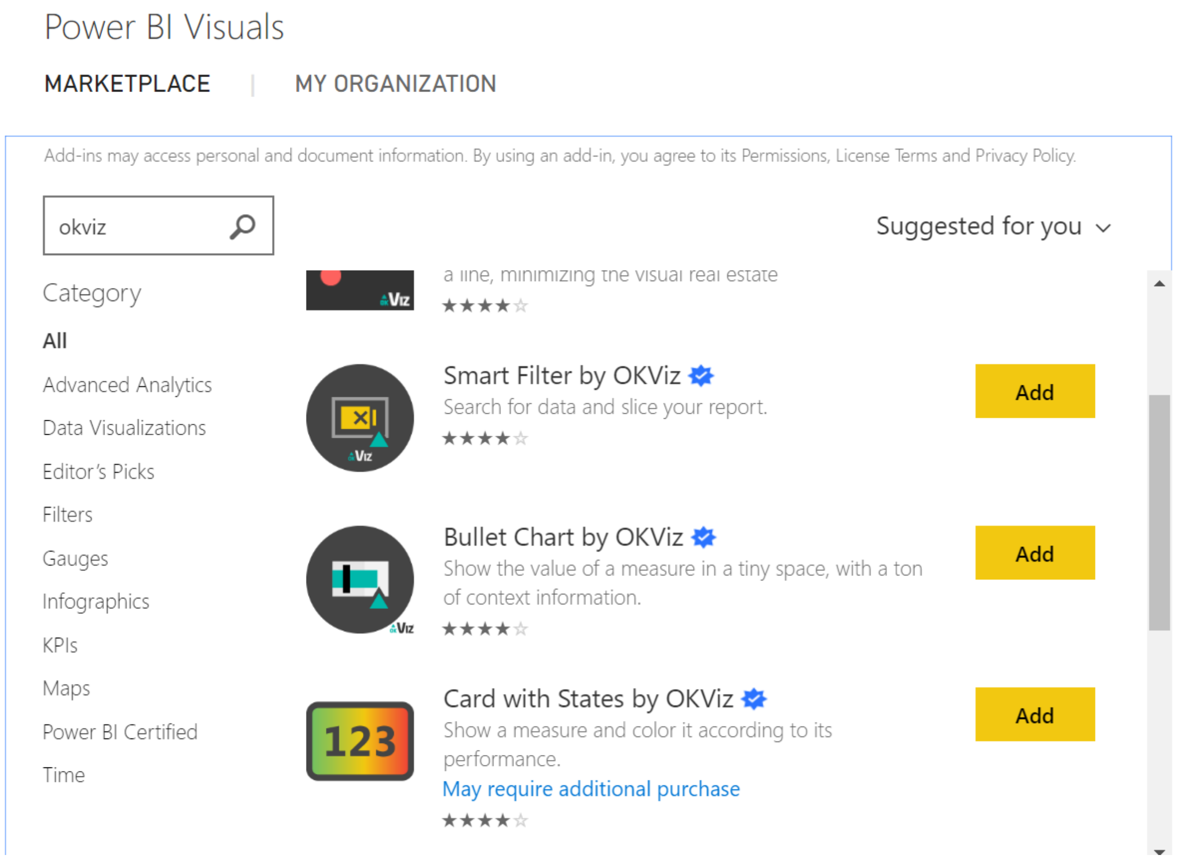
Using Certified Custom Visuals

Using Certified Custom Visuals

Custom visuals published to AppSource can optionally go through a validation process by Microsoft to become certified.

Administrators have the following three capabilities for management of custom visuals which will affect the experience for report creators:

1. Block all custom visuals. This is done with a tenant setting in the Power BI service and a group policy setting.
2. Allow certified visuals only. This is done with a tenant setting in the Power BI service and a group policy setting.
3. Deploy specific approved visuals through the organizational visuals store. This is done in the admin center in the Power BI service.



Power BI Visuals

MARKETPLACE | MY ORGANIZATION

Add-ins may access personal and document information. By using an add-in, you agree to its Permissions, License Terms and Privacy Policy.

okviz

Suggested for you

Category

All

Advanced Analytics

Data Visualizations

Editor's Picks

Filters

Gauges

Infographics

KPIs

Maps

Power BI Certified

Time

a line, minimizing the visual real estate
★★★★☆

Smart Filter by OKViz
Search for data and slice your report.
★★★★☆
Add

Bullet Chart by OKViz
Show the value of a measure in a tiny space, with a ton of context information.
★★★★☆
Add

Card with States by OKViz
Show a measure and color it according to its performance.
May require additional purchase
★★★★☆
Add

QUICK TIPS

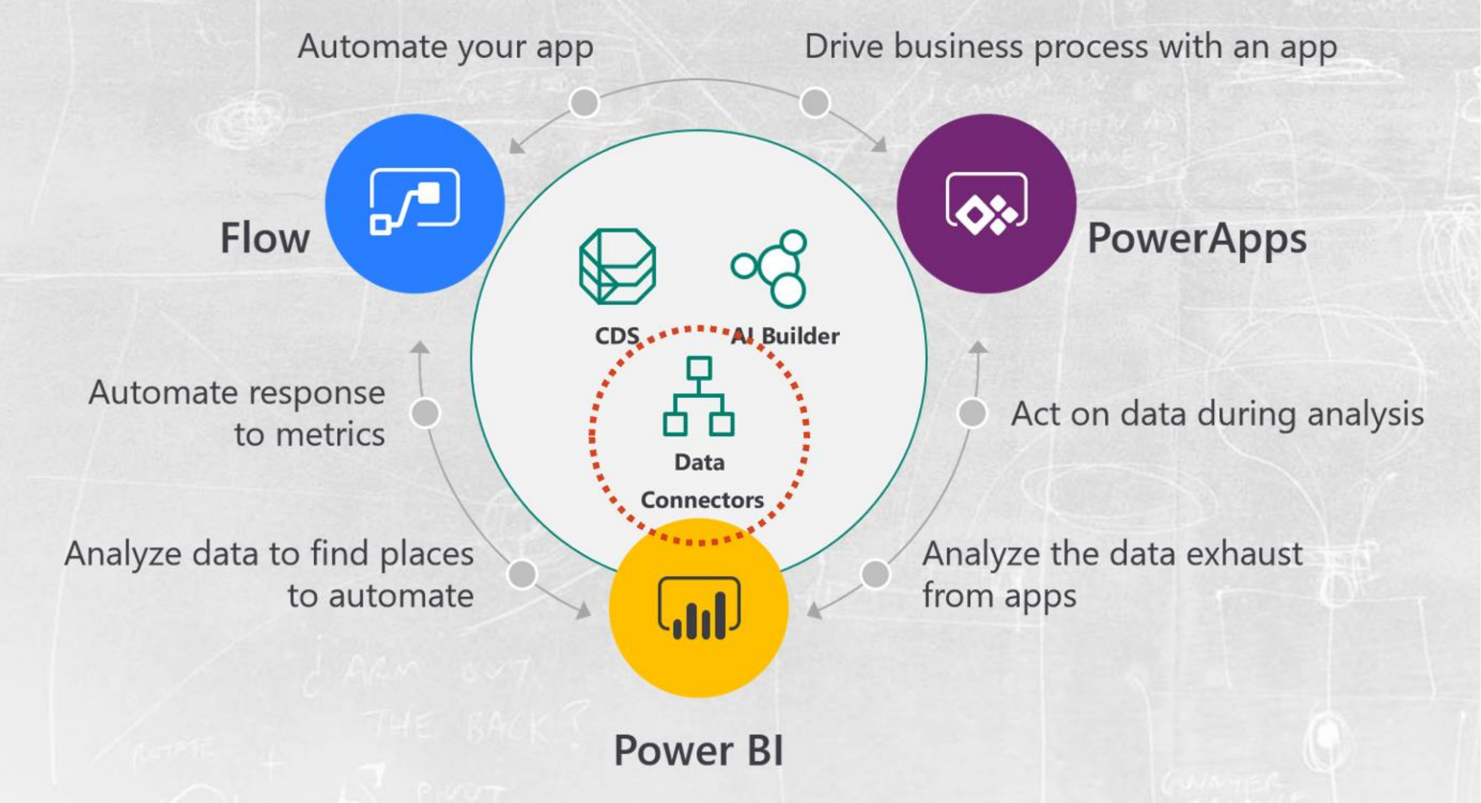
Be certain to only use custom visuals from sources that you know and trust. It is important to understand that custom visuals have access to the data in Power BI datasets, and they are not blocked from sending data out of the Power BI service

Power BI

Using Custom Data Connectors

Using Custom Data Connectors

It is possible to extend the number of data sources that Power BI can connect to by creating custom data connectors. The same warnings that apply to using custom visuals (discussed in the preceding section) also apply to using custom data connectors created by third parties: they may not be properly tested or supported, and as a result you should have a policy for evaluating and testing the use of custom connectors.


Automate your app
Drive business process with an app
Flow
PowerApps
CDS
AI Builder
Data Connectors
Automate response to metrics
Act on data during analysis
Analyze data to find places to automate
Analyze the data exhaust from apps
Power BI

Power BI

Creating Dashboards to Unify Key Metrics from Various Reports

Creating Dashboards to Unify Key Metrics from Various Reports

Pinning visualizations to a dashboard allows report creators to provide an “at-a-glance” view of the most important metrics. A dashboard could be considered something like a start page or executive summary: if all looks well, the report consumer may not need to review the underlying reports, but if necessary the user can click on a visualization and navigate to the underlying report to see it in context.

A dashboard can be created by pinning individual visualizations from multiple areas:

- Report visuals published to the Power BI service
- Q&A results in the Power BI service
- QuickInsights results in the Power BI service
- Visuals within reports published to Power BI Report Server
- Visuals within reports published to SQL Server Reporting Services
- Visuals within Excel workbooks published to the Power BI service

Usage of dashboards is how to provide a single view of data which spans multiple reports and/or datasets and/or tools.