3Cloud | Microsoft Gold Partner

Doing Power BI The Right Way

for Database Developers

Paul Turley

Principal Consultant,
3Cloud Solutions

MVP Microsoft® Most Valuable Professional

Level: Intermediate

What’s This All About?

From the perspective of a DBA and database professional, this session will provide prescriptive guidance to plan and build an enterprise BI solution correctly using best practice design. The mindsets of self-service BI and database development are often different, so using common language and a best-of-breed approach is a crucial step for designing futureproof BI solutions, and “doing Power BI the right way”.

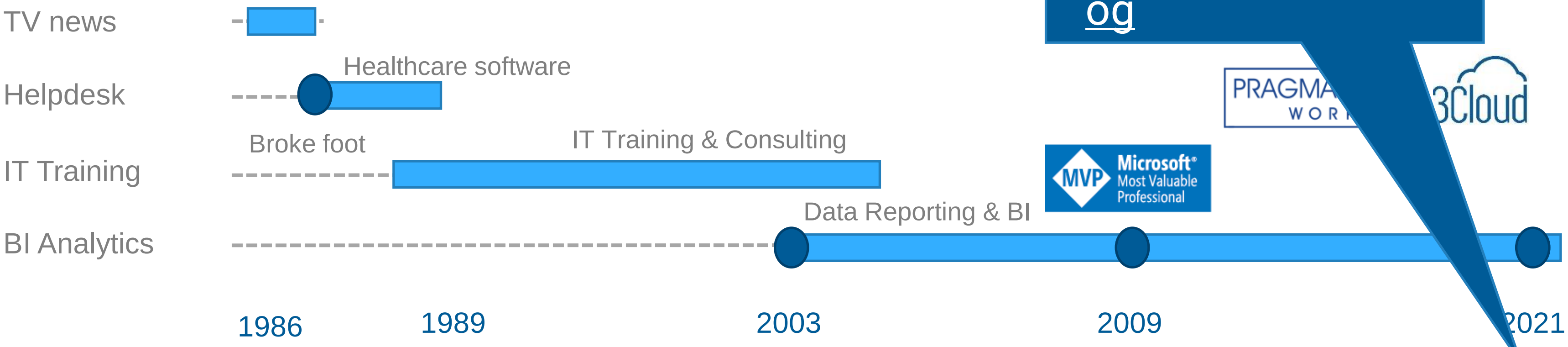
Based on 25 years of industry experience and hundreds of durable Power BI solutions, you will see how to create and maintain analytic reporting projects supporting data governance, self-service reporting; and agile, iterative development practices.

Paul Turley

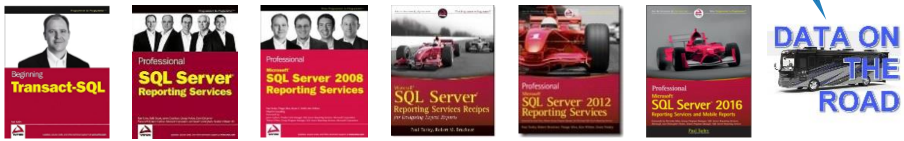
Principal Consultant, Microsoft Data Platform MVP

~25 years in IT, data platform,
Business Intelligence & data analytics

SqlServerBI.blog
DataOnTheRoad.blog


TV news
Helpdesk
Healthcare software
Broke foot
IT Training & Consulting
IT Training
Data Reporting & BI
BI Analytics
Microsoft® Most Valuable Professional
1986
1989
2003
2009
2021

14 books


Beginning
Transact-SQL
Professional
SQL Server Reporting Services
Professional
Microsoft SQL Server 2008 Reporting Services
Microsoft SQL Server Reporting Services Recipes
Professional
Microsoft SQL Server 2012 Reporting Services
Professional
Microsoft SQL Server 2016 Reporting Services and Mobile Reports
DATA ON THE ROAD

Agenda

- Enterprise Readiness
- Futureproofing Power BI solutions
- Decision points:
    - Managing Project Files
    - Accessing Source Data
    - Object Naming
    - Model Design
    - Calculations & Measures
    - Data Security
    - Deployment & Updates

Typical Power BI Use Cases

- Ad hoc analysis
- Quick reporting
- Transformation & shaping as-you-go
- Data model evolution
- Visualize: try-and-see
- Standards? Who needs standards?
  - Naming conventions
  - Modeling & relationships
  - Measure logic
  - Documentation
  - Version control

Database Development Discipline

- Data quality is enforced in the database
- Database is single version of the truth
- Avoid data copies & silos
- All requirements are documented
- Version control
- DevOps: DEV, TEST & PROD
- Scripted & automated
- Naming conventions: PascalCase, camelCase

Audience, Size & Scale

One size doesn't always fit all

With planning, informal projects can be adapted to meet enterprise needs

Know when to keep and update a design

Know when to toss a design and start over

Planning for Scale

3Cloud

Power BI Pro License

1 GB max

1 license per report user

Premium Per User

100 GB

Premium Capacity

400 GB

Premium Capacity

Hybrid Data Models

No stated size limits

No licenses needed for report viewers

How Much Data?

Power BI Tabular Engine is optimized for in-memory scans & calculations

Data volume:

- Millions of rows ?
- Billions of rows ?

Data model size:

- 1 GB
- 10 GB
- 100 GB
- 400 GB
- Terabytes & Petabytes

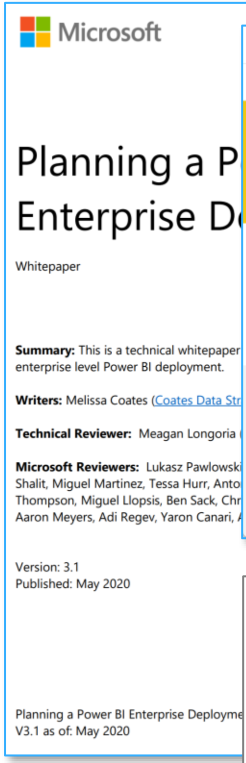
Licensing Power BI

(guidelines)

Power BI Pro	$10/user
Premium per User	$20/user
Premium Capacity (P1)	$5k/Month
A1 – A6	$735 - $3k
P2 – P5	$10k - $80k

Is Power BI Enterprise Ready?

- ✓ Large-scale Data
- ✓ Broad user support
- ✓ Team Development & Collaboration
- ✓ Data Governance Policies
- ✓ Data Management Policies
- ✓ Lifecycle & Deployment
- ✓ Versioning
- ✓ Embedding & Integration
- ✓ Automation
- ✓ Administration
- ✓ Security:
  - User Access & Data Protection
  - Role-based & User-based filtering



Microsoft

Planning a P
Enterprise D

Whitepaper

Summary: This is a technical whitepaper
enterprise level Power BI deployment.

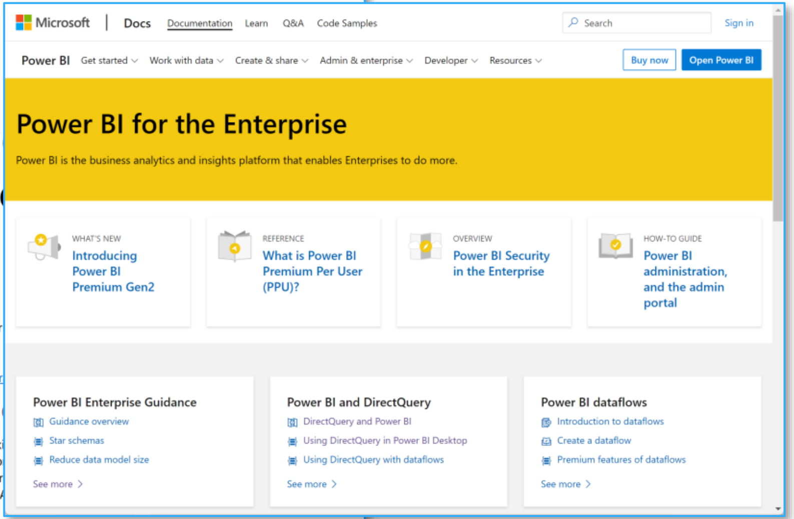
Writers: Melissa Coates (Coates Data Str

Technical Reviewer: Meagan Longoria (

Microsoft Reviewers: Lukasz Pawlowski
Shalit, Miguel Martinez, Tessa Hurr, Anto
Thompson, Miguel Llopsis, Ben Sack, Chr
Aaron Meyers, Adi Regev, Yaron Canari, A

Version: 3.1
Published: May 2020

Planning a Power BI Enterprise Deployme
V3.1 as of: May 2020


Microsoft | Docs Documentation Learn Q&A Code Samples Search Sign in
Power BI Get started Work with data Create & share Admin & enterprise Developer Resources Buy now Open Power BI

Power BI for the Enterprise

Power BI is the business analytics and insights platform that enables Enterprises to do more.

WHAT'S NEW
Introducing Power BI Premium Gen2

REFERENCE
What is Power BI Premium Per User (PPU)?

OVERVIEW
Power BI Security in the Enterprise

HOW-TO GUIDE
Power BI administration, and the admin portal

Power BI Enterprise Guidance
- Guidance overview
- Star schemas
- Reduce data model size

See more >

Power BI and DirectQuery
- DirectQuery and Power BI
- Using DirectQuery in Power BI Desktop
- Using DirectQuery with dataflows

See more >

Power BI dataflows
- Introduction to dataflows
- Create a dataflow
- Premium features of dataflows

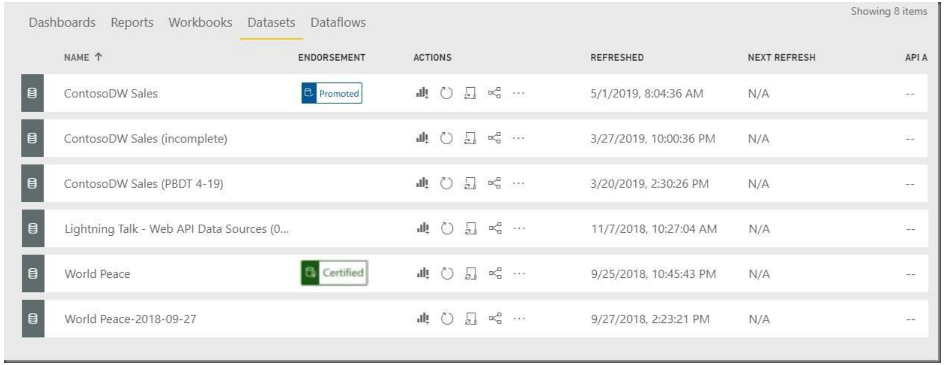
See more >

New Power BI Enterprise Hub:
https://docs.microsoft.com/en-us/power-bi/admin

White papers for Power BI - Power BI | Microsoft Docs:
https://docs.microsoft.com/en-us/power-bi/guidance/whitepapers

Certified & Shared Datasets

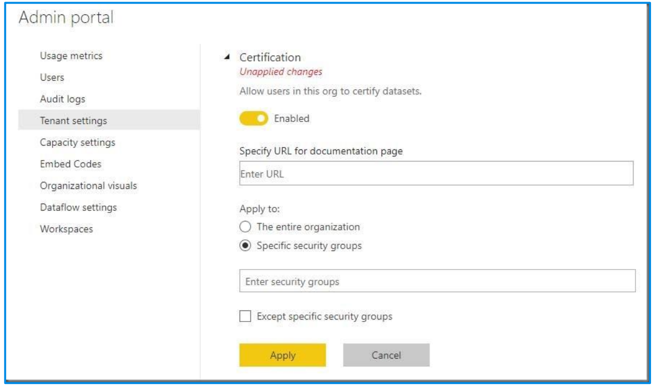
- Use Dataset endorsement & certification in the service
- Certification can be managed by security group
- Access to datasets can be restricted to certified datasets
- Organization defines certification policy & provides documentation



Dashboards Reports Workbooks Datasets Dataflows

Showing 8 items

NAME ↑	ENDORSEMENT	ACTIONS	REFRESHED	NEXT REFRESH	API A
ContosoDW Sales	Promoted	⋯	5/1/2019, 8:04:36 AM	N/A	--
ContosoDW Sales (incomplete)		⋯	3/27/2019, 10:00:36 PM	N/A	--
ContosoDW Sales (PBDT 4-19)		⋯	3/20/2019, 2:30:26 PM	N/A	--
Lightning Talk - Web API Data Sources (0...		⋯	11/7/2018, 10:27:04 AM	N/A	--
World Peace	Certified	⋯	9/25/2018, 10:45:43 PM	N/A	--
World Peace-2018-09-27		⋯	9/27/2018, 2:23:21 PM	N/A	--



Admin portal

Certification

Unapplied changes

Allow users in this org to certify datasets.

Enabled

Specify URL for documentation page

Enter URL

Apply to:

○ The entire organization

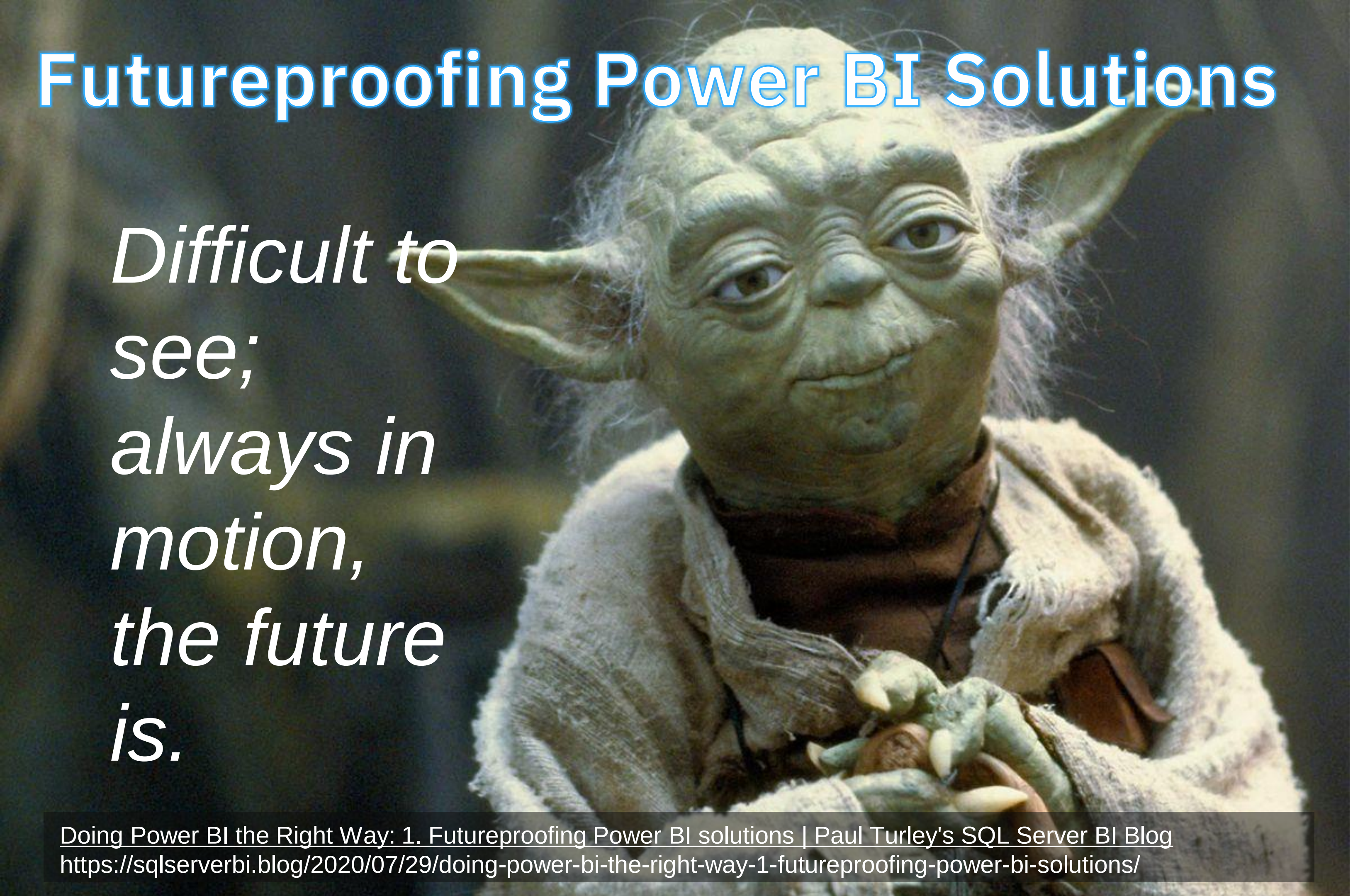
● Specific security groups

Enter security groups

□ Except specific security groups

Futureproofing Power BI Solutions

3Cloud



Difficult to see;
always in motion,
the future is.

Doing Power BI the Right Way: 1. Futureproofing Power BI solutions | Paul Turley's SQL Server BI Blog
https://sqlserverbi.blog/2020/07/29/doing-power-bi-the-right-way-1-futureproofing-power-bi-solutions/

Reality of BI Projects

3Cloud

You will never ever, ever..

EVER EVER

have ALL requirements defined ahead of time

Iterative Design

- Define solution scope
- Work in iterative cycles
- Evaluate design
- Keep what works
- Discard what doesn't
- Learn & improve
- Evaluate to solution scope

Managing Project Files

Database Developer

- Use source code repository
- Automated versio
- Manage everythir
- Automate build &

BI Solution Designer

- Shared file folder
- control
- sktop

Balance

- Shared storage of some kind
- Roll-back & disaster recovery
- Manual deployment is easy
- Automated deployment is possible
- Don’t merge & split PBIX files

Manage Power BI Desktop Files

- **Store files in a centrally managed network-assessable folder**
  The storage folder should support automatic backup and recovery in the case of storage loss.
- **Report and dataset developers must open files from the Windows file system**
  Files must either reside in or be synchronized with the Windows file system.
- Files containing imported data typically range in size from 100 to 600 MB. Any shared folder synchronization or disaster recovery system should be designed to effectively handle multiple files of this size.

Options:

- OneDrive For Business (shared by team, with folder synchronization).
- SharePoint or SharePoint Online (with folder synchronization).
- GitHub and/or VSTS with local repository & folder synchronization. If used, Git must be configured for large file storage (LFS) if PBIX files are to be stored in the repository.

We'll talk about this a bit later

Connecting Source Data

3Cloud

Database Developer

- Write SQL
- Process querie
- Real time conn

BI Solution Designer

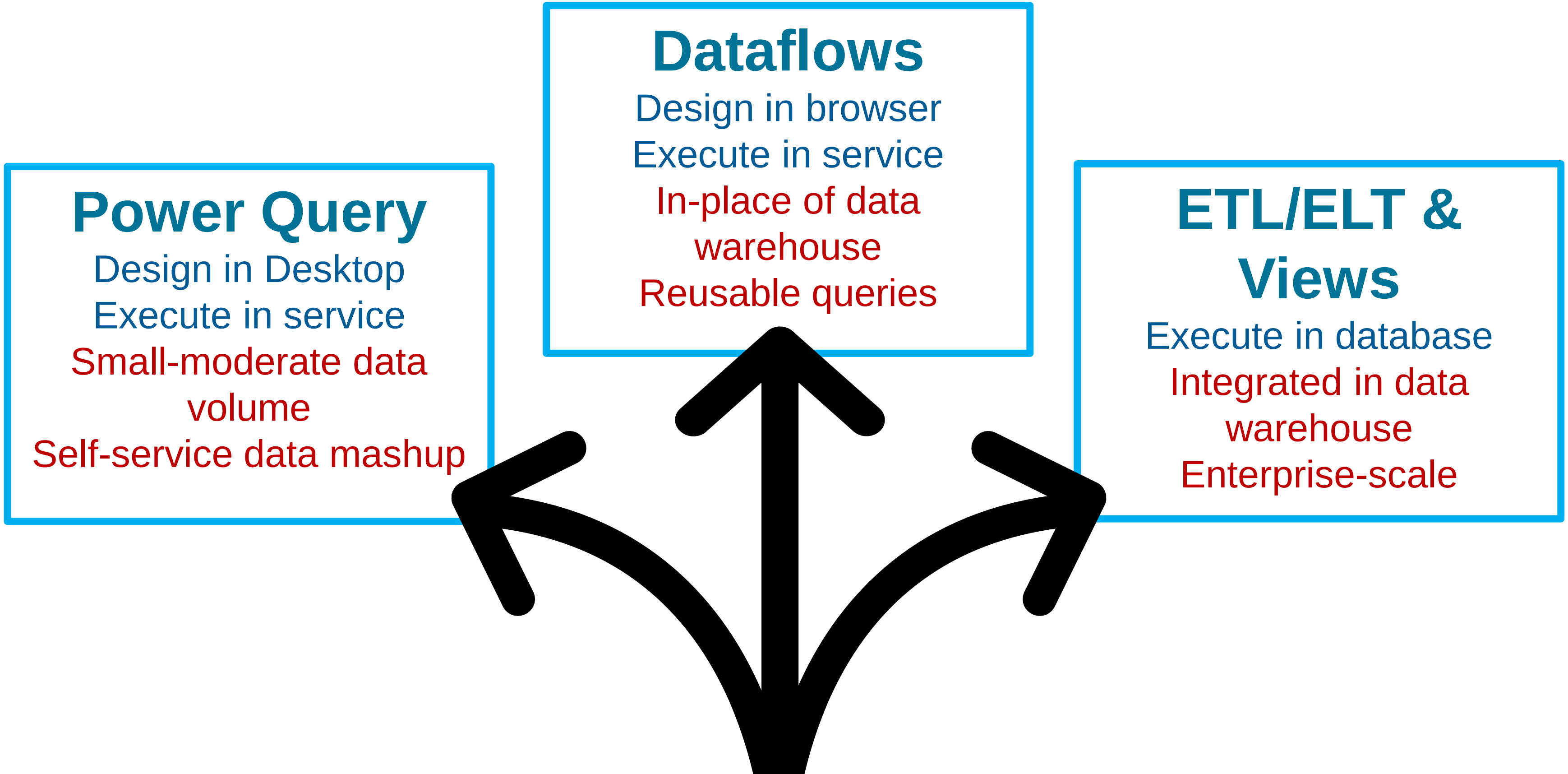
model

ions in-memory

Balance

- Use views to encapsulate SQL queries
- Ensure query folding works on large tables
- Use Import by default
- Use DirectQuery only when required

Transformation Options



Power Query

Design in Desktop
Execute in service
Small-moderate data volume
Self-service data mashup

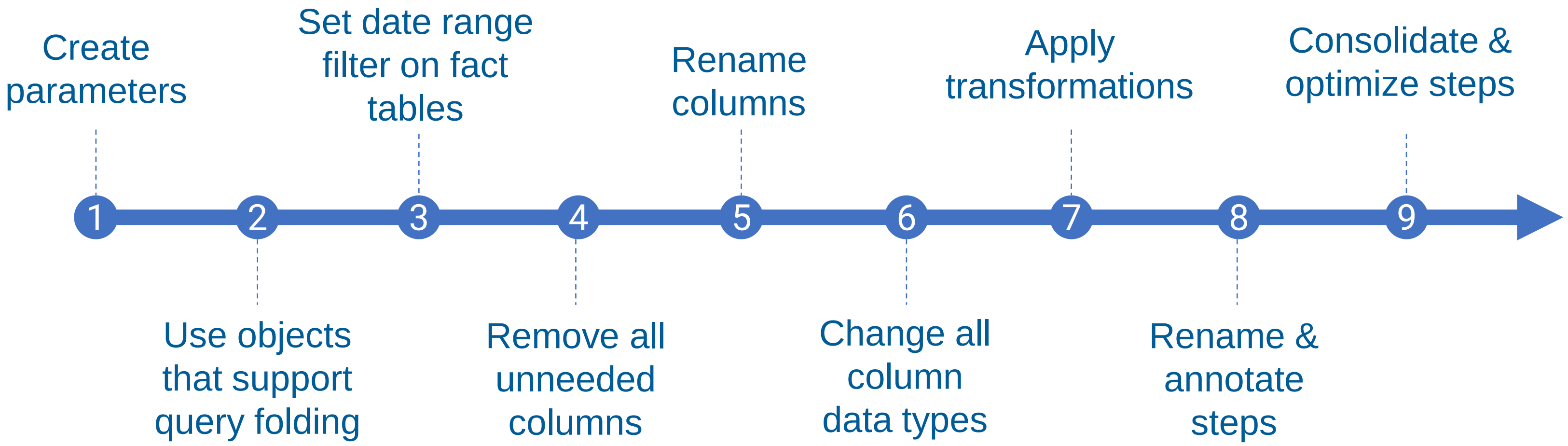
Dataflows

Design in browser
Execute in service
In-place of data warehouse
Reusable queries

ETL/ELT & Views

Execute in database
Integrated in data warehouse
Enterprise-scale

Preparing, shaping & transforming source data using Power Query



1. Create parameters
2. Use objects that support query folding
3. Set date range filter on fact tables
4. Remove all unneeded columns
5. Rename columns
6. Change all column data types
7. Apply transformations
8. Rename & annotate steps
9. Consolidate & optimize steps

Doing Power BI the Right Way: 2. Preparing, shaping & transforming source data | Paul Turley's SQL Server BI Blog
https://sqlserverbi.blog/2020/08/16/doing-power-bi-the-right-way-2-preparing-source-data/

SQL Isn’t Always The Answer

Import tables and views rather than hand-written SQL queries, so Power Query can fold queries and execute them at the source

- Prepare dimensional tables in the source database
- Create database views

Never Forget This

Transform Data as Far Upstream as Possible

Downstream transformations may be more flexible for self-service models but may not scale

Object Names

Database Developer

- Table, field/colu
- Measure names
- Code-friendly n
- Cryptic object n
  Pascal case, Ca
  Hungarian nota

BI Solution Designer

es

d

Balance

- All objects exposed to users should have friendly names
- Hide key columns & utility objects
- Hide numeric columns & create friendly named measures

Remember One Thing

Name Objects for Usability

Data model consumers are business users, not developers. Name tables, fields & measures so they are intuitive & easy to understand.

ACCT_TYP
customerPhoneNbr

Account Type
Customer Phone #

Model Design

Database Developer

- Flattened result
- Transform in SC
- Wide tables
- Pre-aggregated
- Import summar

BI Solution Designer

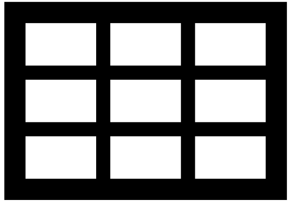
wer Query

Balance

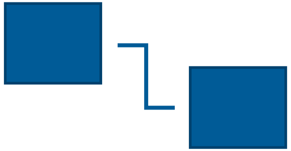
- Build dimensional star schema
- Avoid wide tables
- Remove long text fields
- Remove unused fields
- Tall tables are OK if they are optimized for storage & analytics

Data modeling essentials & best practices in Power BI

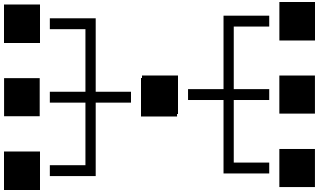
Flat Model



Master/Detail

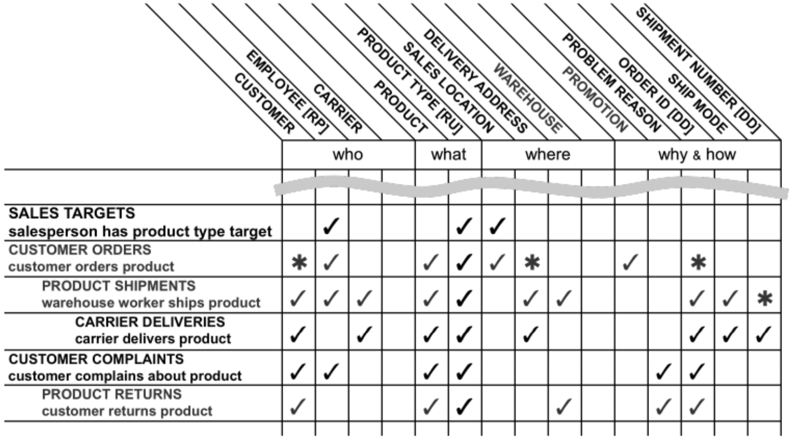


Dimensional

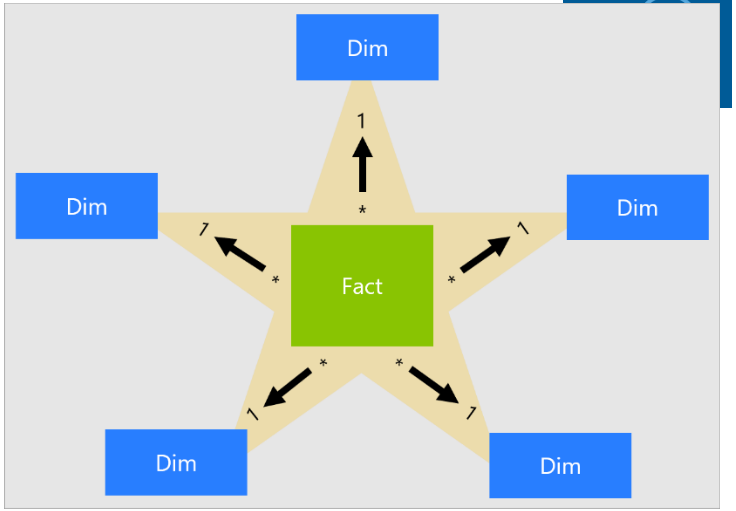


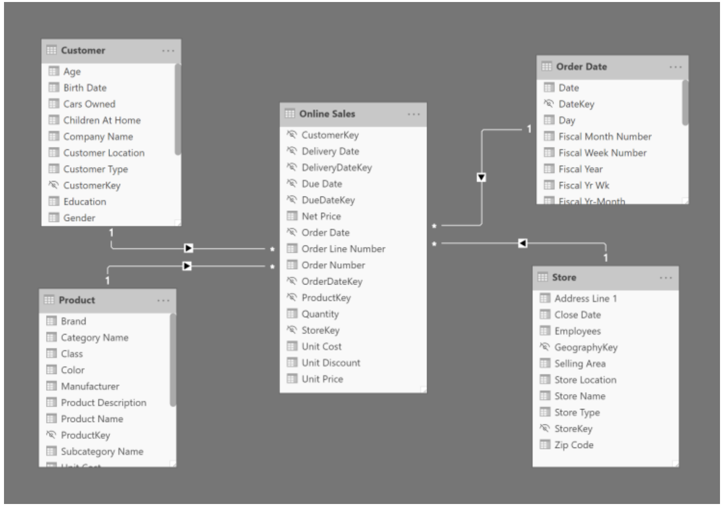
Doing Power BI the Right Way: 5. Data Modeling Essentials & Best Practices (1 of 2) | Paul Turley's SQL Server BI Blog
https://sqlserverbi.blog/2020/12/25/doing-power-bi-the-right-way-6-data-modeling-essentials-best-practices-1-of-2/

Dimensionalize It!



	who				what		where				why & how				
	CUSTOMER	EMPLOYEE [RP]	CARRIER		PRODUCT	PRODUCT TYPE [RU]	SALES LOCATION	DELIVERY ADDRESS	WAREHOUSE		PROMOTION	PROBLEM REASON	ORDER ID [DD]	SHIP MODE	SHIPMENT NUMBER [DD]
SALES TARGETS **salesperson has product type target**		✓				✓	✓								
CUSTOMER ORDERS customer orders product	✱	✓			✓	✓	✓	✱			✓		✱		
PRODUCT SHIPMENTS warehouse worker ships product	✓	✓	✓		✓	✓		✓	✓				✓	✓	✱
CARRIER DELIVERIES **carrier delivers product**	✓		✓		✓	✓		✓					✓	✓	✓
CUSTOMER COMPLAINTS **customer complains about product**	✓	✓			✓	✓						✓	✓		
PRODUCT RETURNS customer returns product	✓				✓	✓			✓			✓	✓		





Customer
- Age
- Birth Date
- Cars Owned
- Children At Home
- Company Name
- Customer Location
- Customer Type
- CustomerKey
- Education
- Gender

Product
- Brand
- Category Name
- Class
- Color
- Manufacturer
- Product Description
- Product Name
- ProductKey
- Subcategory Name

Online Sales
- CustomerKey
- Delivery Date
- DeliveryDateKey
- Due Date
- DueDateKey
- Net Price
- Order Date
- Order Line Number
- Order Number
- OrderDateKey
- ProductKey
- Quantity
- StoreKey
- Unit Cost
- Unit Discount
- Unit Price

Order Date
- Date
- DateKey
- Day
- Fiscal Month Number
- Fiscal Week Number
- Fiscal Year
- Fiscal Yr Wk
- Fiscal Yr-Month

Store
- Address Line 1
- Close Date
- Employees
- GeographyKey
- Selling Area
- Store Location
- Store Name
- Store Type
- StoreKey
- Zip Code

- Nearly any reporting data can be shaped into a star schema & optimized for analysis
- Ad hoc models never evolve into a dimensional model

Model Design Guidelines

- Dimensional design concepts haven’t changed in 20 years & are as true as ever
- Dimensional modeling “rules” should be followed but can be relaxed for Power BI in certain cases, such as:
  - Leaving some dimensional attributes in fact tables
  - Use natural keys rather than generating surrogate keys
- The art of dimensional modeling ranges from simple to complex. Start with the basics.
- Flattened “spreadsheet” models are OK for small, informal projects but have significant limitations
- As models grow in size & complexity, data quality challenges will surface that can be solved by implementing proper governance controls

The Kimball Method: https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques
Lawrence Corr, Model Storming Agile method: https://modelstorming.com/hierarchy-map

Planning for Separation – data models & reports

The Thick and Thin of Reports

Separate reports and data models can be:

- Versioned
- Deployed & managed separately
- Central "Golden" dataset
- Use Live Connect when creating reports

Doing Power BI the Right Way: 7. Planning for separation – data models and reports | Paul Turley's SQL Server BI Blog

5 Tips for Separating Power BI Datasets and Reports — Coates Data Strategies

Hot Swap tool demo: PowerBI.tips
PowerBI.tips - Hot Swap Connections Webinar – YouTube https://www.youtube.com/watch?v=sygIUPMlgi0

Enterprise Scale Options

In many ways, Power BI has now surpassed the capabilities of SQL Server Analysis Services. Microsoft are investing in the enterprise capabilities of the Power BI platform by enhancing Power BI Premium Capacity, adding Paginated Report and features to support massive scale specialized use cases. Consider the present and planned capabilities of the Power BI platform; before, choosing another data modeling tool such as SSAS.

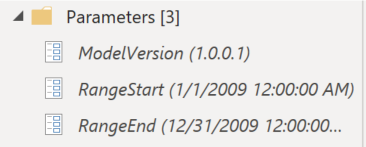
Resources:

https://sqlserverbi.blog/2018/07/27/power-bi-for-grownups

https://sqlserverbi.blog/2018/12/13/data-model-options-for-power-bi-solutions

Managing Dataset Size with Parameters

- Use parameters whether implementing incremental refresh or not
- **RangeStart** & **RangeEnd** parameters must be date/time type
- Apply range filter on date/time column in Power Query

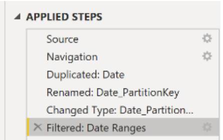


Parameters [3]

ModelVersion (1.0.0.1)

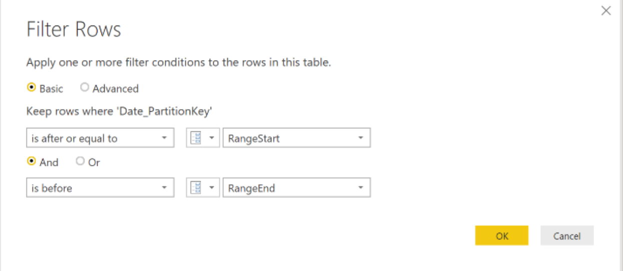
RangeStart (1/1/2009 12:00:00 AM)

RangeEnd (12/31/2009 12:00:00...



APPLIED STEPS

- Source
- Navigation
- Duplicated: Date
- Renamed: Date_PartitionKey
- Changed Type: Date_Partition...
- Filtered: Date Ranges



Filter Rows

Apply one or more filter conditions to the rows in this table.

● Basic ○ Advanced

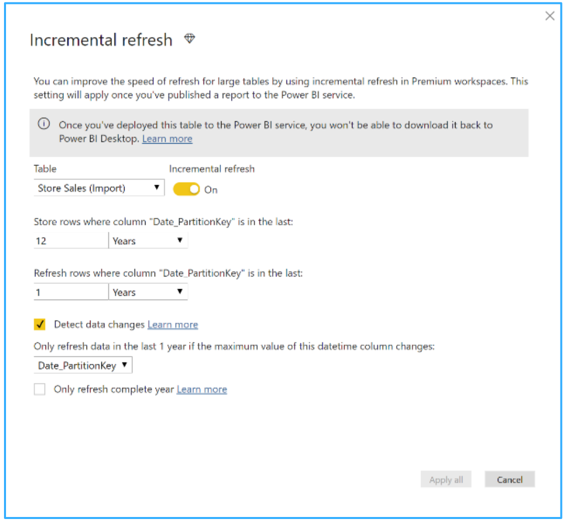
Keep rows where 'Date_PartitionKey'

is after or equal to | RangeStart

● And ○ Or

is before | RangeEnd

Compare row count & refresh history – 1 mo & incr refresh



Incremental refresh

You can improve the speed of refresh for large tables by using incremental refresh in Premium workspaces. This setting will apply once you've published a report to the Power BI service.

ⓘ Once you've deployed this table to the Power BI service, you won't be able to download it back to Power BI Desktop. Learn more

Table: Store Sales (Import)

Incremental refresh: On

Store rows where column "Date_PartitionKey" is in the last: 12 Years

Refresh rows where column "Date_PartitionKey" is in the last: 1 Years

☑ Detect data changes Learn more

Only refresh data in the last 1 year if the maximum value of this datetime column changes: Date_PartitionKey

□ Only refresh complete year Learn more

*Incremental Refresh is a Premium [capacity|user] feature but parameters may be used in any data model

Remember This

Use Parameters to Filter Results

Query parameters are used to reduce the model size in development

Implement Incremental Refresh

Scale model size & data volume

Calculations & Measures

Database Developer

- Perform calcula
- Store calculate
- Process calcula database

BI Solution Designer

ions in DAX

report context

ion logic in the

Balance

- Understand & embrace row & filter context
- Store detail-level calculations
- Design measures to work with grouped fields & hierarchies

Data Security

Database Developer

- Lock it all down
- Keep data in th
- Manage securit

BI Solution Designer

data for me

in the data

Balance

- Manage privacy & security according to data governance policy
- Apply row-level security in data model
- Create & maintain user mapping table to support row-level security

Deployment & Updates

Database Developer

- Rigid developm
- Documented re
- Change reques
- Long release cy
- Tidy/discipled/p

BI Solution Designer

ment lifecycle

requirements

s needed

cle

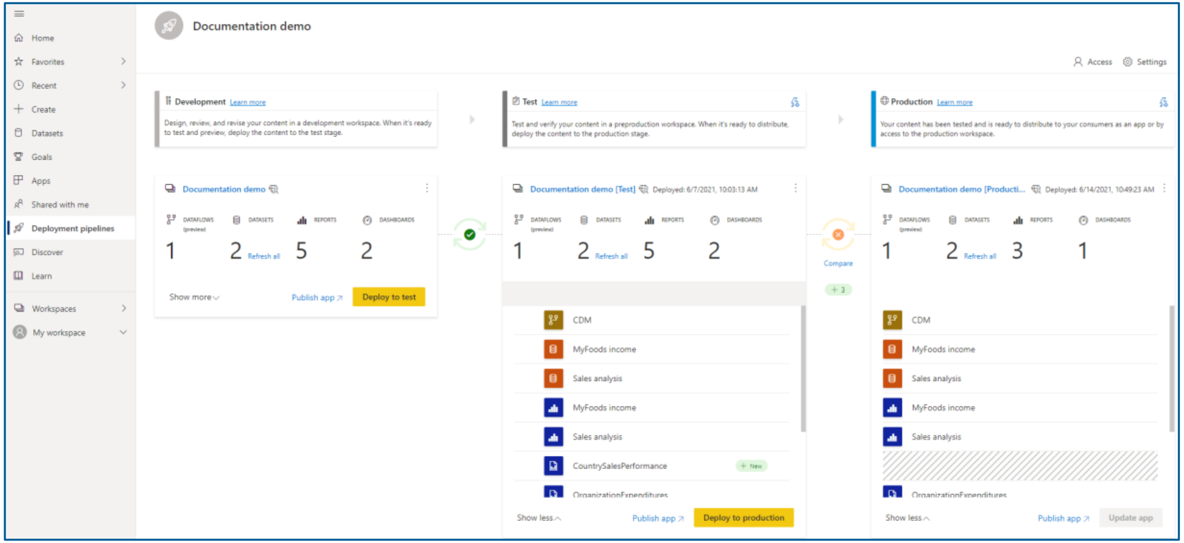
ned/chaotic

Balance

- Right-size process for the audience & business
- Sandbox self-service models & reports in workspaces
- Apply rigid controls to certified models & reports using deployment pipelines & workspace apps

Deployment Pipelines

- Deployment pipelines manage the promotion of workspace app content through lifecycle stages
- Versioning with regression



Documentation demo

Development Learn more

Design, review, and revise your content in a development workspace. When it's ready to test and preview, deploy the content to the test stage.

Test Learn more

Test and verify your content in a preproduction workspace. When it's ready to distribute, deploy the content to the production stage.

Production Learn more

Your content has been tested and is ready to distribute to your consumers as an app or by access to the production workspace.

Documentation demo

DATAFLOWS (preview)	DATASETS	REPORTS	DASHBOARDS
1	2 Refresh all	5	2

Documentation demo [Test] Deployed: 6/7/2021, 10:03:13 AM

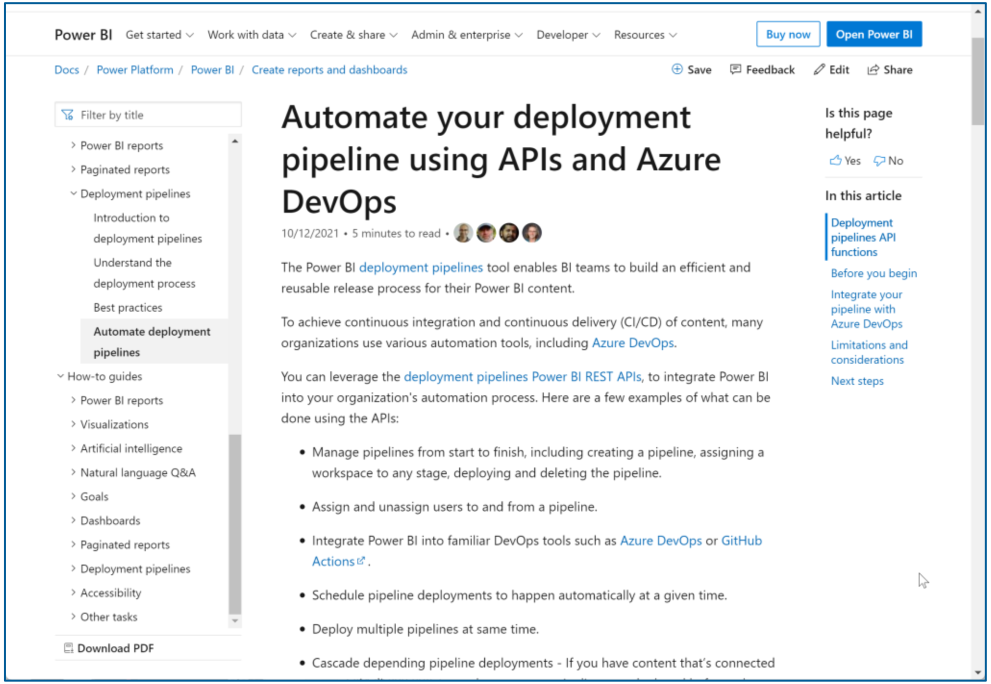
DATAFLOWS (preview)	DATASETS	REPORTS	DASHBOARDS
1	2 Refresh all	5	2

Documentation demo [Producti... Deployed: 6/14/2021, 10:49:23 AM

DATAFLOWS (preview)	DATASETS	REPORTS	DASHBOARDS
1	2 Refresh all	3	1

DevOps Integration

- Deployment pipelines API functions:
  - Deploy all
  - Selective deploy
  - Backward deploy
  - Update App
- REST APIs may be called with PowerShell or other automation code
- Supports OAuth



Automate your deployment pipeline using APIs and Azure DevOps

10/12/2021 • 5 minutes to read •

The Power BI deployment pipelines tool enables BI teams to build an efficient and reusable release process for their Power BI content.

To achieve continuous integration and continuous delivery (CI/CD) of content, many organizations use various automation tools, including Azure DevOps.

You can leverage the deployment pipelines Power BI REST APIs, to integrate Power BI into your organization's automation process. Here are a few examples of what can be done using the APIs:

- Manage pipelines from start to finish, including creating a pipeline, assigning a workspace to any stage, deploying and deleting the pipeline.
- Assign and unassign users to and from a pipeline.
- Integrate Power BI into familiar DevOps tools such as Azure DevOps or GitHub Actions .
- Schedule pipeline deployments to happen automatically at a given time.
- Deploy multiple pipelines at same time.
- Cascade depending pipeline deployments - If you have content that's connected

https://docs.microsoft.com/en-us/power-bi/create-reports/deployment-pipelines-automation

Multi-developer & Lifecycle Management for Power BI

DATA MODEL

- Separate data model PBIX from Report PBIX
- Deploy and manage separately
- Migrate large-scale or enterprise data models to Model.BIM file
- Manage enterprise models in Tabular Editor or Visual Studio
- Use Deployment Pipelines for Apps

REPORTS

- Treat report PBIX files as binary files
- Use OneDrive or code repo versioning
- Don’t use merging and differencing

Take Aways

- Store source files in shared storage
- Separate data model from reports divide work
- Understand and promote Query Folding
- Use SQL views rather than in-line SQL
- Use parameters to reduce working data model file size
- Consider using Incremental Refresh to manage large tables
- Use Deployment Pipelines to manage delivery

In Summary: Model Design Checklist

- □ **Model for the user experience, not for developers**
- □ **Build star schemas**
  Wherever possible, reshape data into fact a dimension tables with single key, one-to-many relationships from dimensions to fact.
- □ **Enforce dimension key uniqueness**
  Just because a key value "should" be unique, there is no guarantee that it will be unless enforced at the data source. Perform grouping and duplicate reduction in the data source views or Power Query queries to guarantee uniqueness. Duplicate record count checks and other mechanisms can be applied to audit source data for integrity but do not allow the data model to violate these rules.
- □ **Avoid bi-directional filters & unnecessary bridging tables**
  These data modelling patterns adversely affect performance.
- □ **Consider using DAX measures** rather than complex & inefficient relationships
- □ **Create custom columns in Power Query**
  Rather than DAX calculated columns wherever possible for row-level derived columns. This maintains a consistent design pattern for maintainability.
- □ **Annotate code**
  Use in-line comments and annotations in all code including SQL, M and DAX; to explain calculation logic and provide author and revision information.
- □ **Remove all unused fields** – if in doubt, take it out
- □ **Hide all fields not used directly by users**
  primary and foreign key columns, numeric columns used to create measures, and columns used to specify the sort order of other fields.
- □ **Use friendly field names**
  Rename all visible columns (in Power Query) to short but user-friendly names with mixed case and spaces.
- □ **Set to <u>Do Not Summarize</u>**
  Any non-hidden numeric columns that are not intended to roll-up or summarize values. Columns set to summarize are indicated with a Sigma icon.

Questions?

- Q&A
- Discussion

Resources, Contact & Network

- **Blogs:**
  - SqlServerBi.blog
  - DataOnTheRoad.blog
- LinkedIn.com/in/Pturley
- @paul_turley

- **New Power BI Enterprise Hub:** https://docs.microsoft.com/en-us/power-bi/admin
- **Guy In A Cube:** https://www.youtube.com/channel/UCFp1vaKzpfvoGai0vE5VJ0w
- Saturday morning livestream 10:30 AM EST