# INCREMENTAL REFRESH AND HYBRID TABLES IN POWER BI

Shabnam Watson

# Shabnam Watson

## BI Consultant


/ShabnamWatson

@shbWatson

info@abicube.com



https://shabnamwatson.wordpress.com

### Work

20 + years of data warehouse and business intelligence solutions development. Power BI and Azure Synapse Analytics

### Speaking/Community

PASS Summit, SQL Saturdays, PASS Virtual Chapters, .NET South, BI and SQL Server user groups. SQL Saturday Atlanta BI Organizer. Data Weekender

### Background

Bachelor's degree in Computer Engineering. Master's degree in computer science. Microsoft Data Platform MVP.

August 2022

# Hybrid Tables

Only available in Power BI Premium or Premium Per user

Combine two storage modes:

- Import: Improve Query Performance for historical data
- Direct Query: Near real time capability for recent data

Incremental Refresh

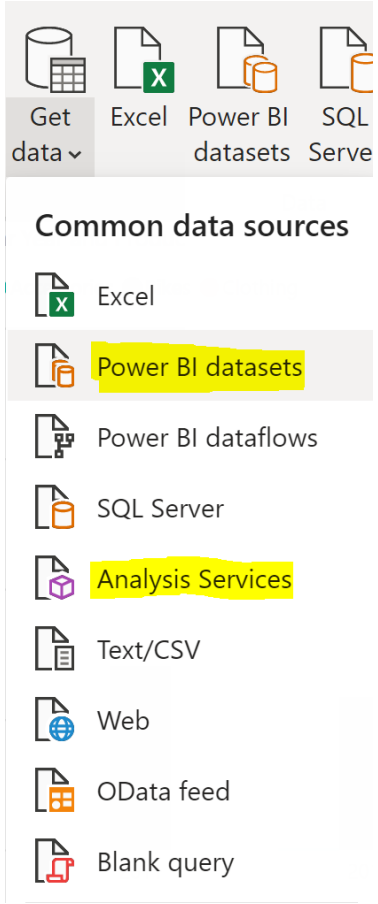
# Agenda

Table Storage Modes

Incremental Refresh

Hybrid Tables

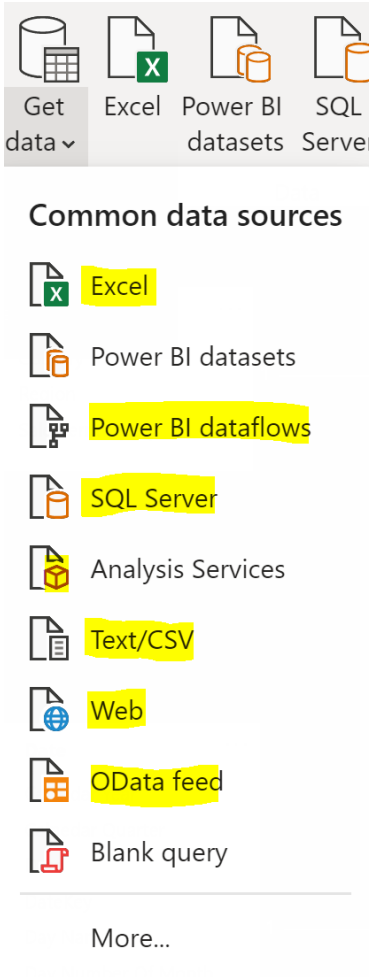
# Power BI Desktop: Reporting tool

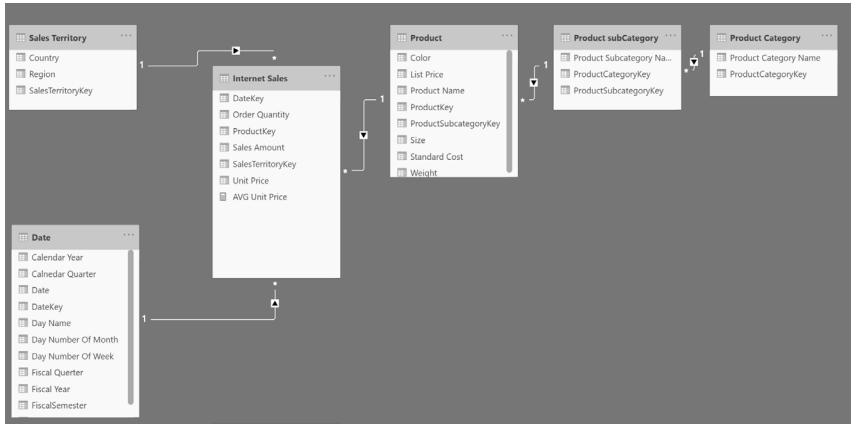



**Live connect** to PBI Datasets published to PBI Service or Analysis Services.

**Note**: Live Connect is not the same as **Direct Query**.

# Power BI Desktop: Modeling tool




Create a semantic model with or w/o data

Define Table Storage modes: Import/Direct Query/Dual

Incremental Refresh

# Table storage modes and more

August 2022

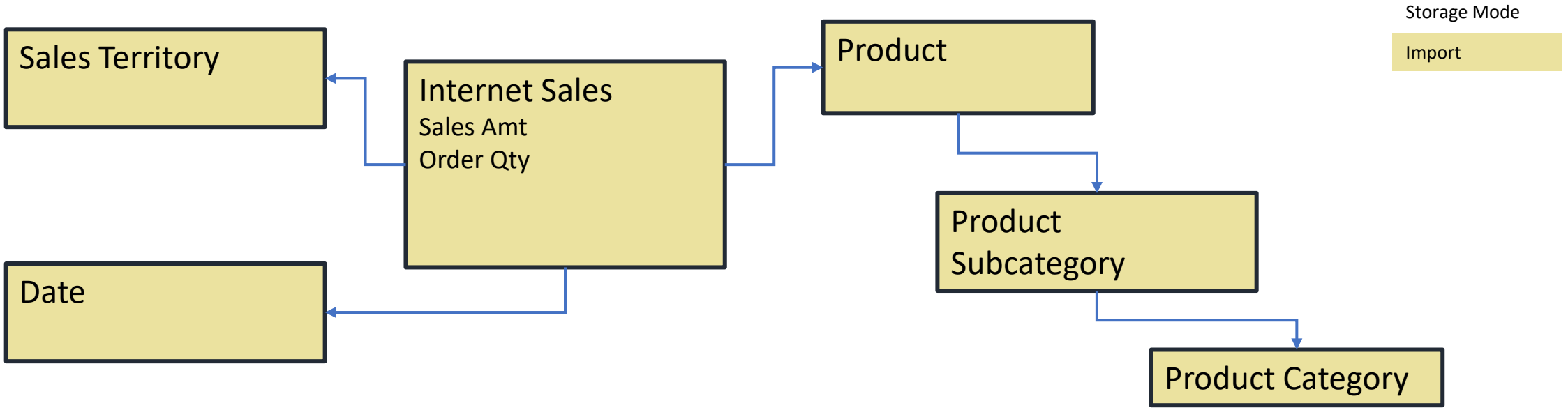
- Import
- Direct Query
- Composite Models/ Mixed Mode: Import, Direct Query, Dual
- Hybrid Table: One table, part Import, part Direct Query

August 2022

# Import vs. Direct Query

# Import Storage Mode

August 2022

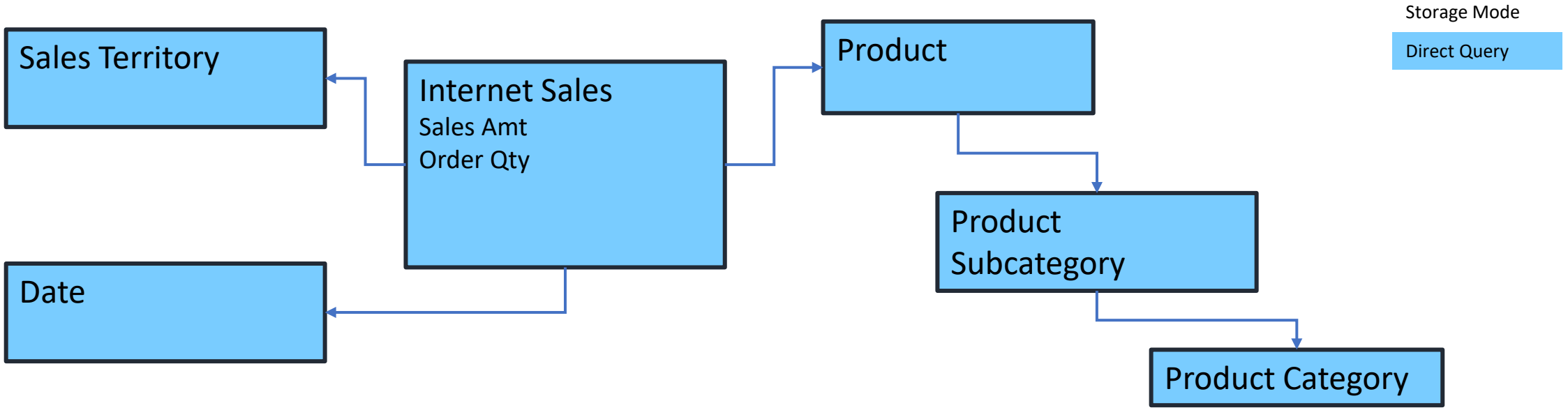



VertiPaq: columnar memory based storage

All queries answered from memory.

# Direct Query Mode

August 2022




All queries answered by data source, for example SQL Server.

# Import vs. Direct Query

## Import:

- Best for most models
- In memory Vertipaq (columnar storage) engine
- Super fast
- Load/processing time
- Refreshes can be slow

## Direct Query

- PBI model: Semantic model
- Near real time (low) latency
- Large datasets that do not fit into memory
- No processing time
- Queries can be slow with large tables

August 2022

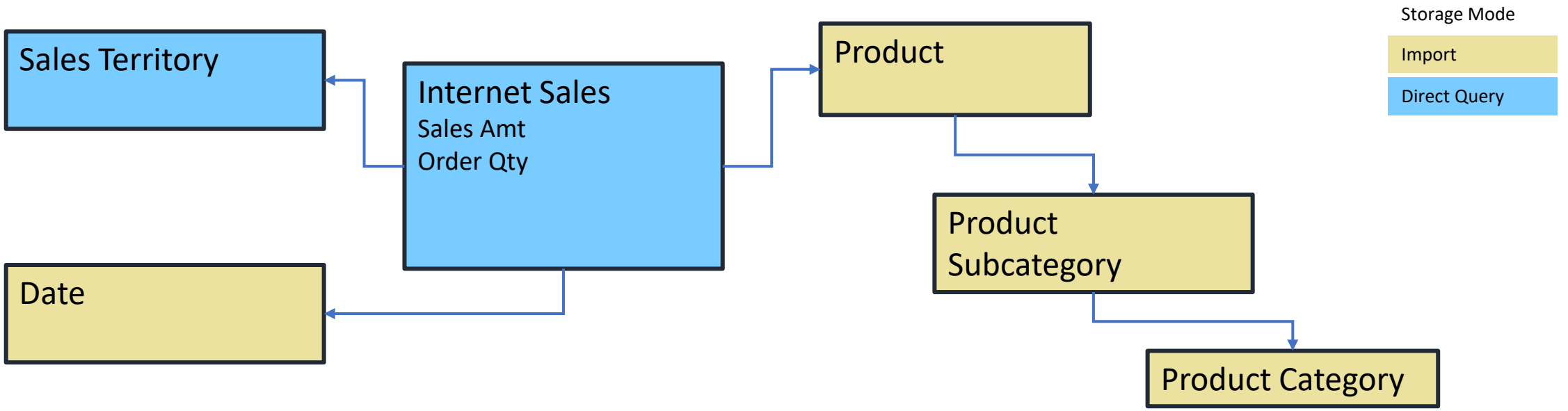
# Composite/Mixed Mode Models

August 2022

# Composite/Mixed Mode Models

Select storage mode for each table.

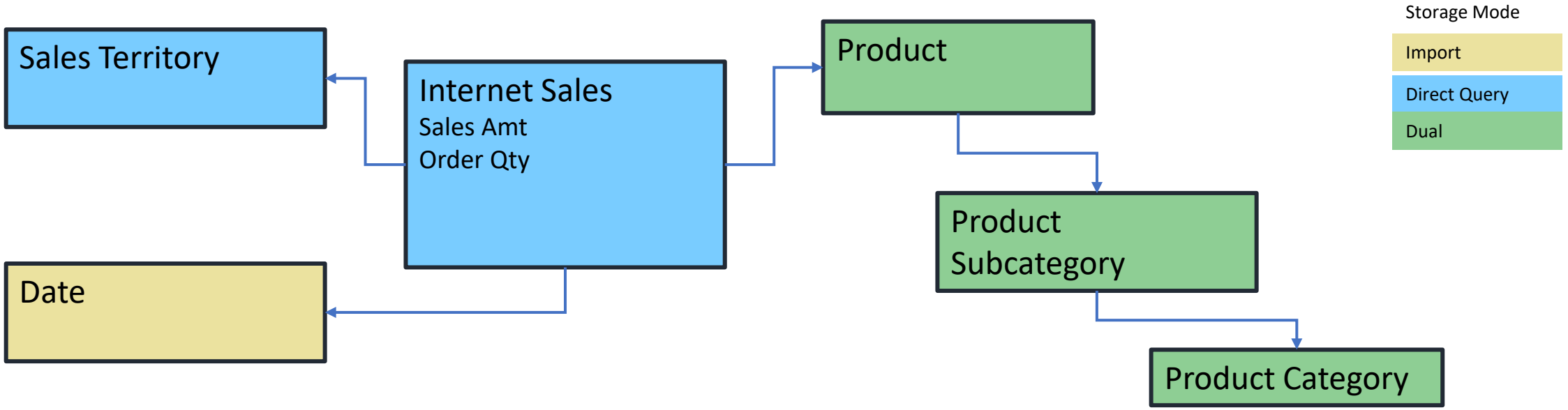
Leave big tables in the data source, load smaller tables into memory




Queries answered either from memory or the data source.

# Dual Storage mode

August 2022




Dual tables can flip back and forth depending on the query. Avoid having joins in PBI.

August 2022

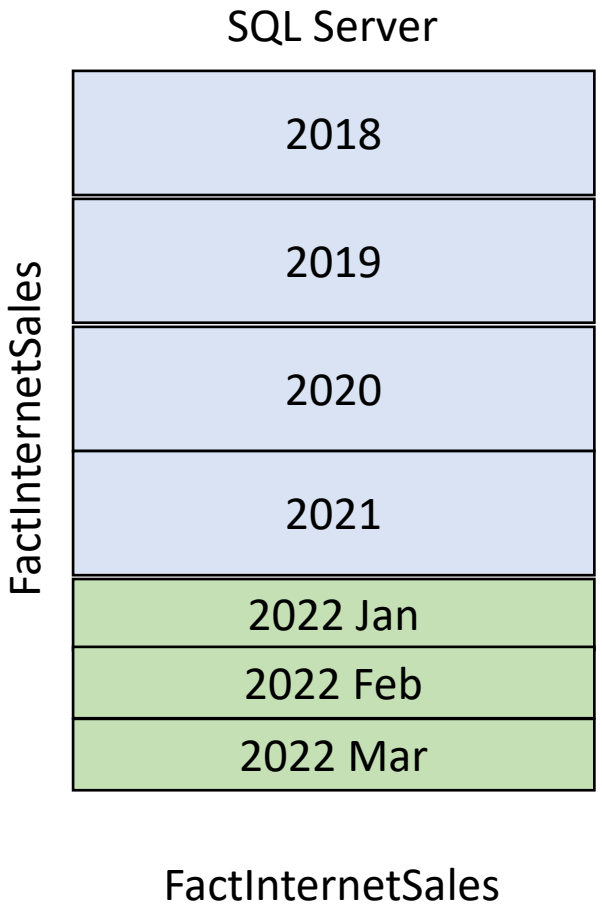
# Data Refresh

# Importing a table into PBI

August 2022

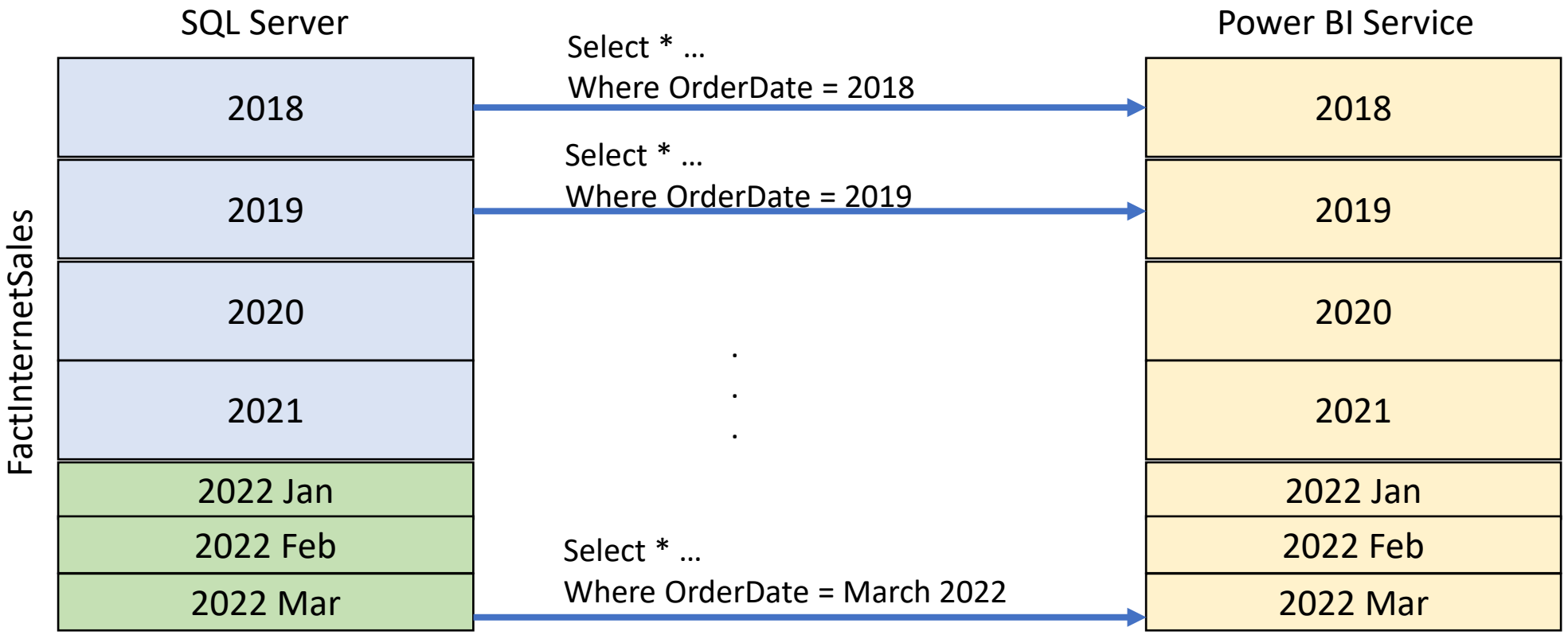
# Incremental Refresh

# Partitioned Table

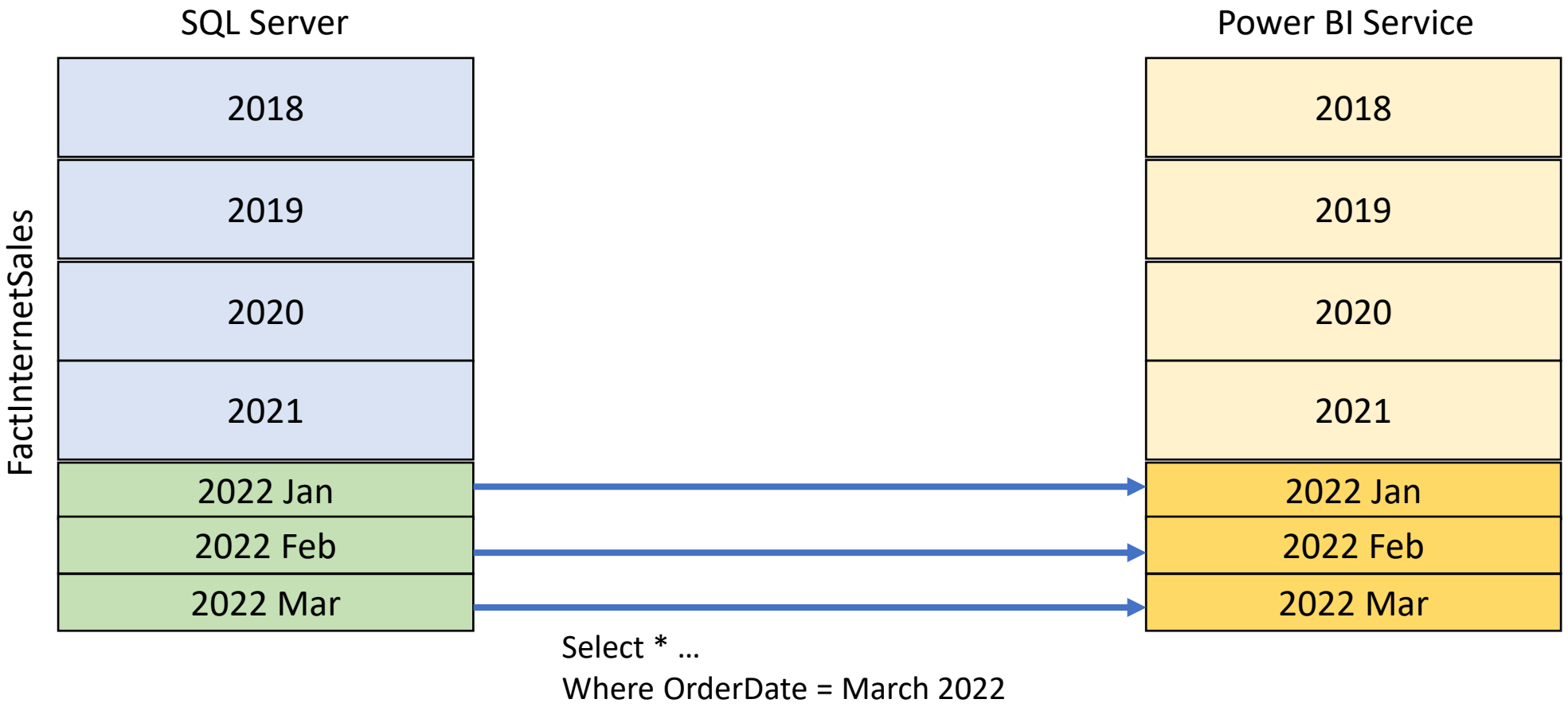



- Historical Data does not change
- Only last 3 months get new data

# Partitioned Table – Initial Refresh/Load



# Partitioned Table – Next Refreshes

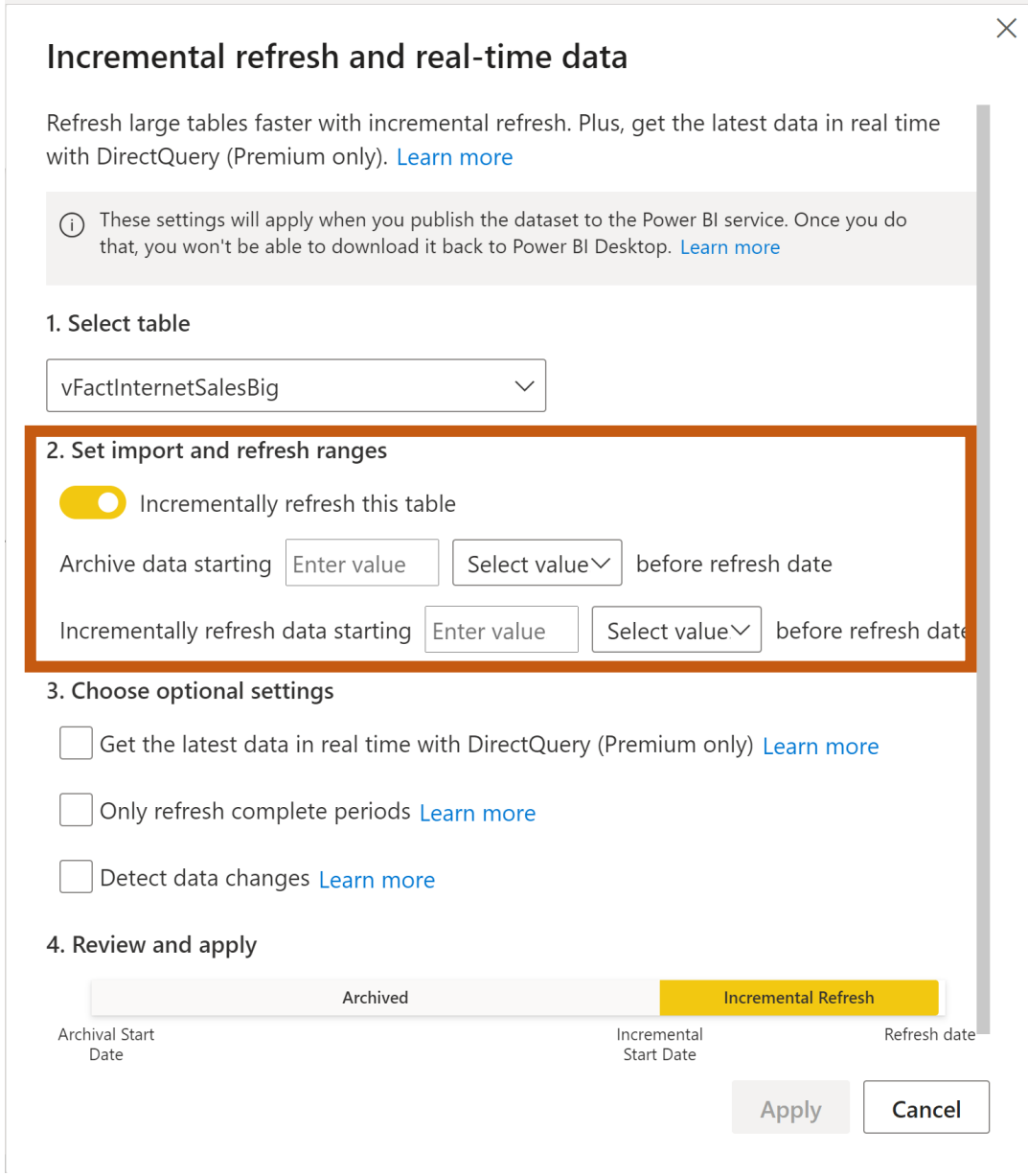


# Incremental Refresh Benefits

- Faster Refreshes:
  - Only recent data is pulled
  - Even entire loads may be faster
- More Reliable: Shorter queries avoid long-running connections
- Reduce back-end and PBI resource consumption
- Makes it possible to work with large datasets

# Partitions Considerations

- Partition filters should be able to be pushed back to the data source: Query Folding
  - Get Data: View vs Select a,b,c from View
- Supported by indexes and statistics


Incremental refresh and real-time data
Refresh large tables faster with incremental refresh. Plus, get the latest data in real time with DirectQuery (Premium only). Learn more
These settings will apply when you publish the dataset to the Power BI service. Once you do that, you won't be able to download it back to Power BI Desktop. Learn more
1. Select table
vFactInternetSalesBig
2. Set import and refresh ranges
Incrementally refresh this table
Archive data starting
Enter value
Select value
before refresh date
Incrementally refresh data starting
Enter value
Select value
before refresh date
3. Choose optional settings
Get the latest data in real time with DirectQuery (Premium only) Learn more
Only refresh complete periods Learn more
Detect data changes Learn more
4. Review and apply
Archived
Incremental Refresh
Archival Start Date
Incremental Start Date
Refresh date
Apply
Cancel

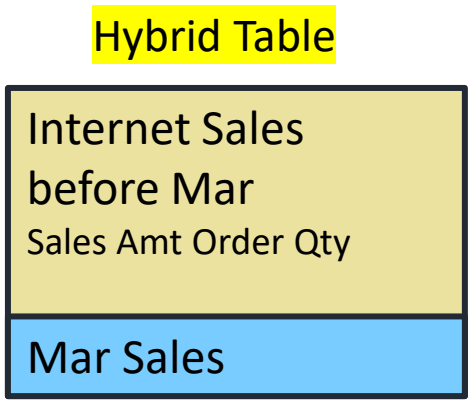
August 2022

# Hybrid Tables

# Hybrid Tables

August 2022

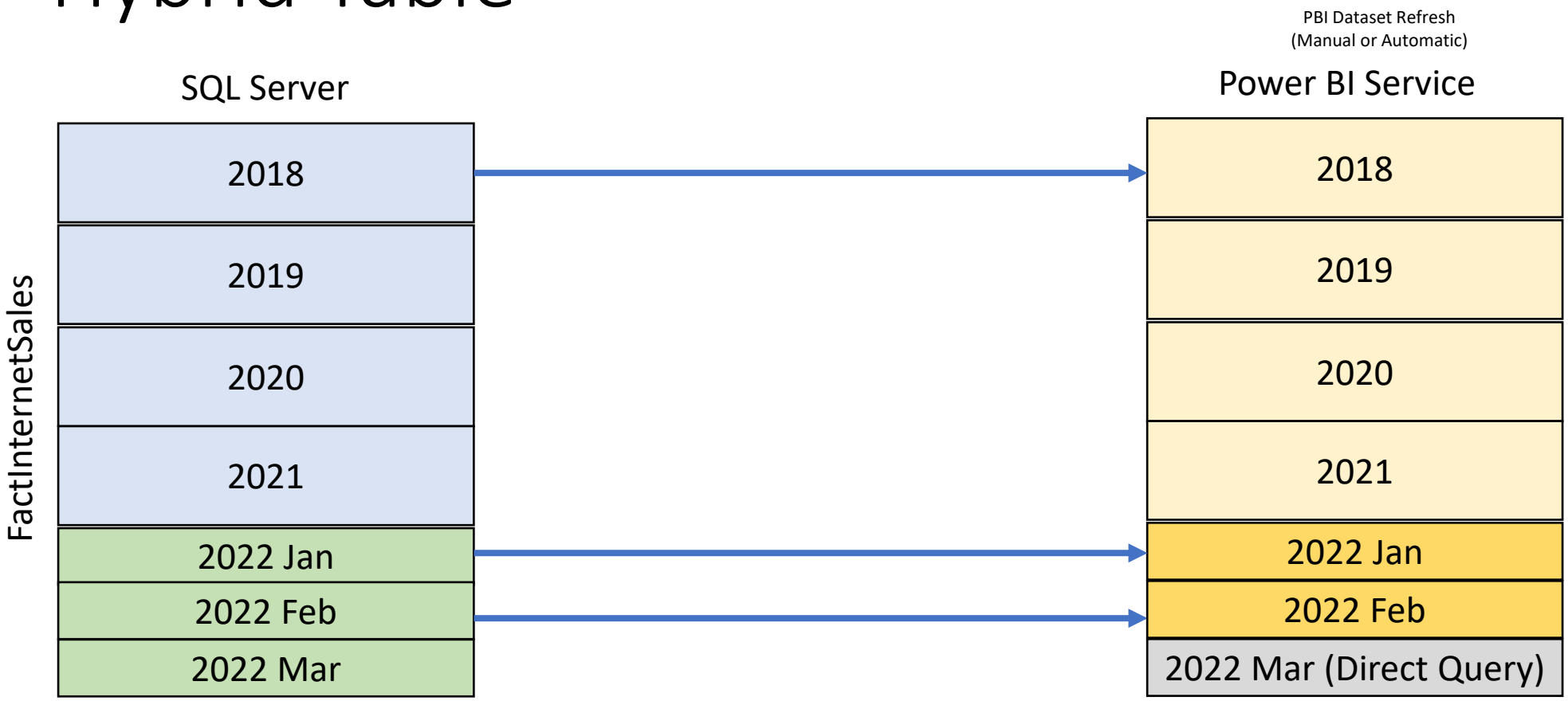
- One table with two storage modes at the same time! (Not Dual)
- Older data in Import: Super fast performance
- Newer data in Direct Query: Near Real time




Data before Mar in Import mode and IR
Data from Mar is in Direct Query

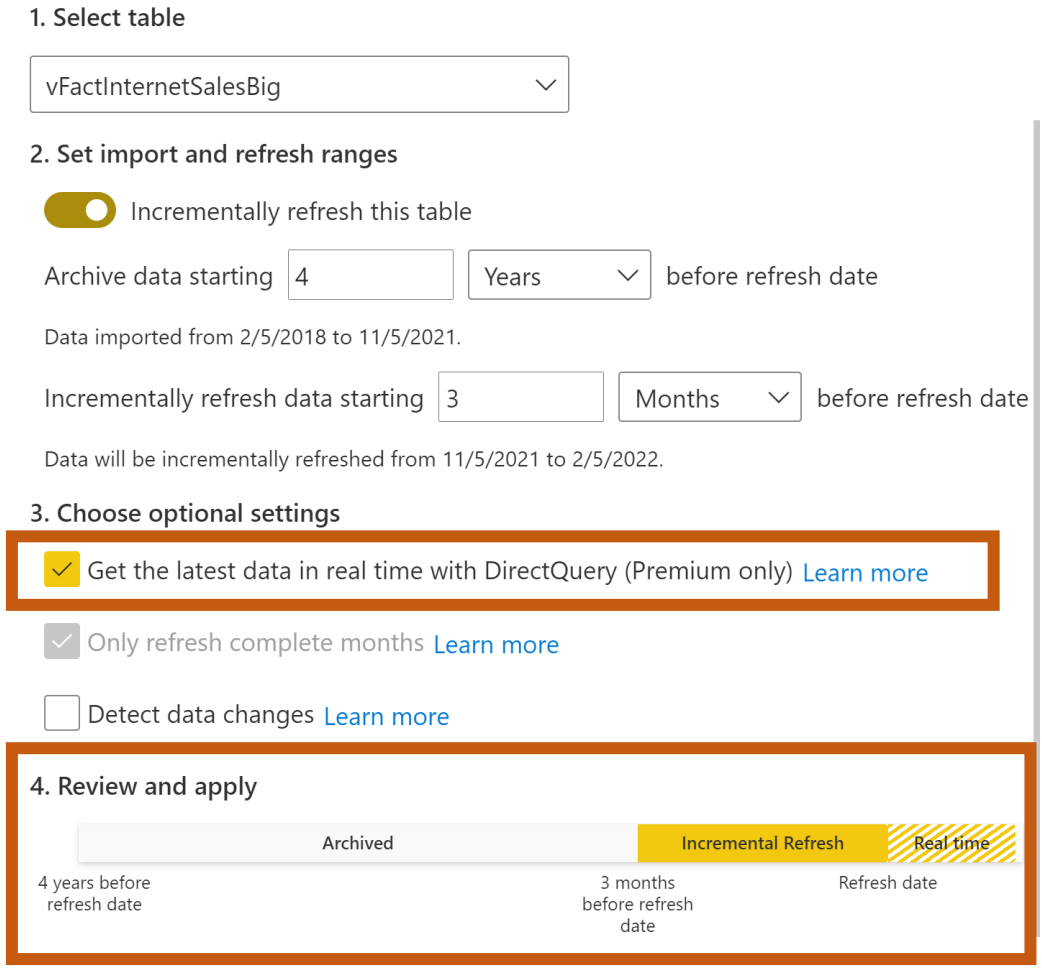
August 2022

# Hybrid Table



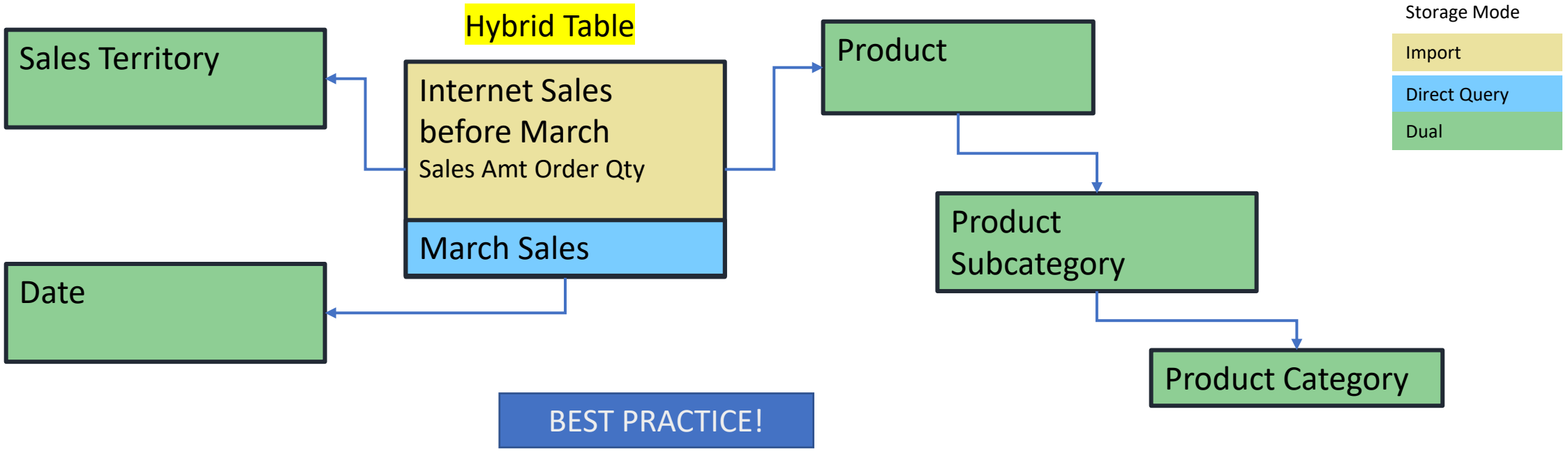
# Enabling Hybrid Table

August 2022



# Hybrid Tables + Dual

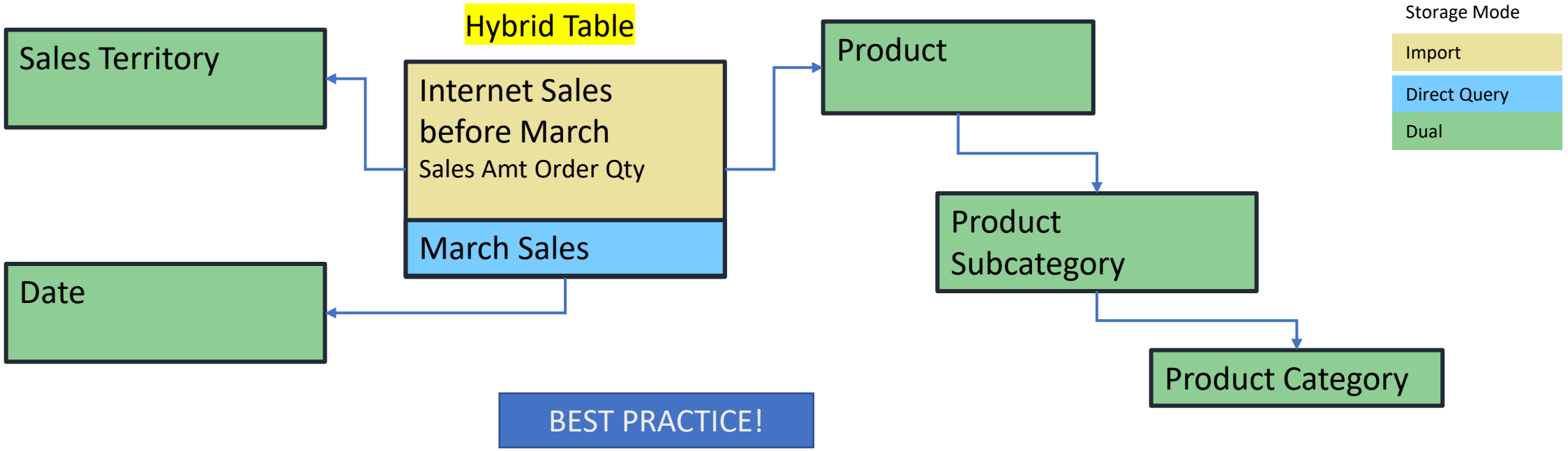
August 2022



# Demo

# Hybrid Tables + Dual

August 2022



# Limitations

- For high cardinality columns may have poor performance:
  - **Distinct Count** may
  - **Min/Max over string column**
  - Analyze in Excel

- ApproximateDistictCount not supported
- A Hybrid table cannot have an aggregation tables and cannot be an aggregation table itself
- A Hybrid table cannot have calculated columns

# Summary

August 2022

- Hybrid Tables are only available in **Power BI Premium** or **PPU**
- Easy to set up, PBI will create and maintain the partitions
- Advanced custom partitions possible with XMLA/TMSL:
    - Direct Query partition to a table for old historic data
    - Newer Data in Import
- Enable near real time data with great performance for older data
- Make all tables that connect to a Hybrid table dual storage.

# Resources to study more



- Announcing Public Preview of Hybrid Tables in Power BI Premium | Microsoft Power BI Blog | Microsoft Power BI
- Power BI Hybrid Tables – Prologika
- Hybrid Tables, Incremental Refresh and Table Partitioning in Power BI | Paul Turley's SQL Server BI Blog
- Incremental refresh for datasets and real-time data in Power BI - Power BI | Microsoft Docs
- Power BI Hybrid Tables with Synapse Analytics Serverless SQL Pools – Serverless SQL
- Automatic page refresh in Power BI Desktop - Power BI | Microsoft Docs

August 2022

# Thank you.

Please leave feedback for this session.

## Connect

/ShabnamWatson

@shbWatson

info@abicube.com

BLOG: https://shabnamwatson.wordpress.com

Microsoft® MVP Most Valuable Professional