



SECTION 4

Dataflows Gen2 & Pipelines

Modern Data Integration



Dataflows Gen2 & Pipelines

Modern Data Integration

"Power Query you know. Destinations you need."

The Supplier Wants More

"We're updating pricing weekly now. We added new columns. And we need confirmation when you've loaded the data."

Transform

Add calculations, clean values

Orchestrate

Schedules, error handling

Notify

Success/failure alerts

Shortcut Transformations

Just land the data

Dataflows + Pipelines

Transform, orchestrate, notify

Demo: Automate the Supplier Refresh

Act	What Happens
1	Understand the new requirements
2	Build a Dataflow Gen2 (Power Query!)
3	Output directly to Lakehouse
4	Create a Pipeline with notifications
5	Monitor and troubleshoot

~42 minutes, live demo

DEMO



You Already Know This

Dataflows Gen2 = Power Query in the Cloud

Power BI Desktop

Power Query Editor

M code

Transformations

Connectors

Output → .pbix

Dataflow Gen2

Same interface

Same language

Same 300+ operations

Same + more

Output → **Lakehouse, Warehouse, SQL, KQL**

Your skills transfer. The destinations upgrade.

Copilot in Dataflows

Natural Language → M Code



Your prompt:

"Add a new column called Margin Potential that calculates UnitPrice minus SupplierCost"

Copilot writes:

```
Table.AddColumn(#"Previous Step", "Margin Potential", each [UnitPrice] - [SupplierCost])
```

What Copilot Does

- Writes M code
- Suggests transformations
- Speeds up authoring

It's not magic. It's faster typing.

What You Still Do

- Review the output
- Verify the logic
- Know the manual way

Gen1 vs Gen2

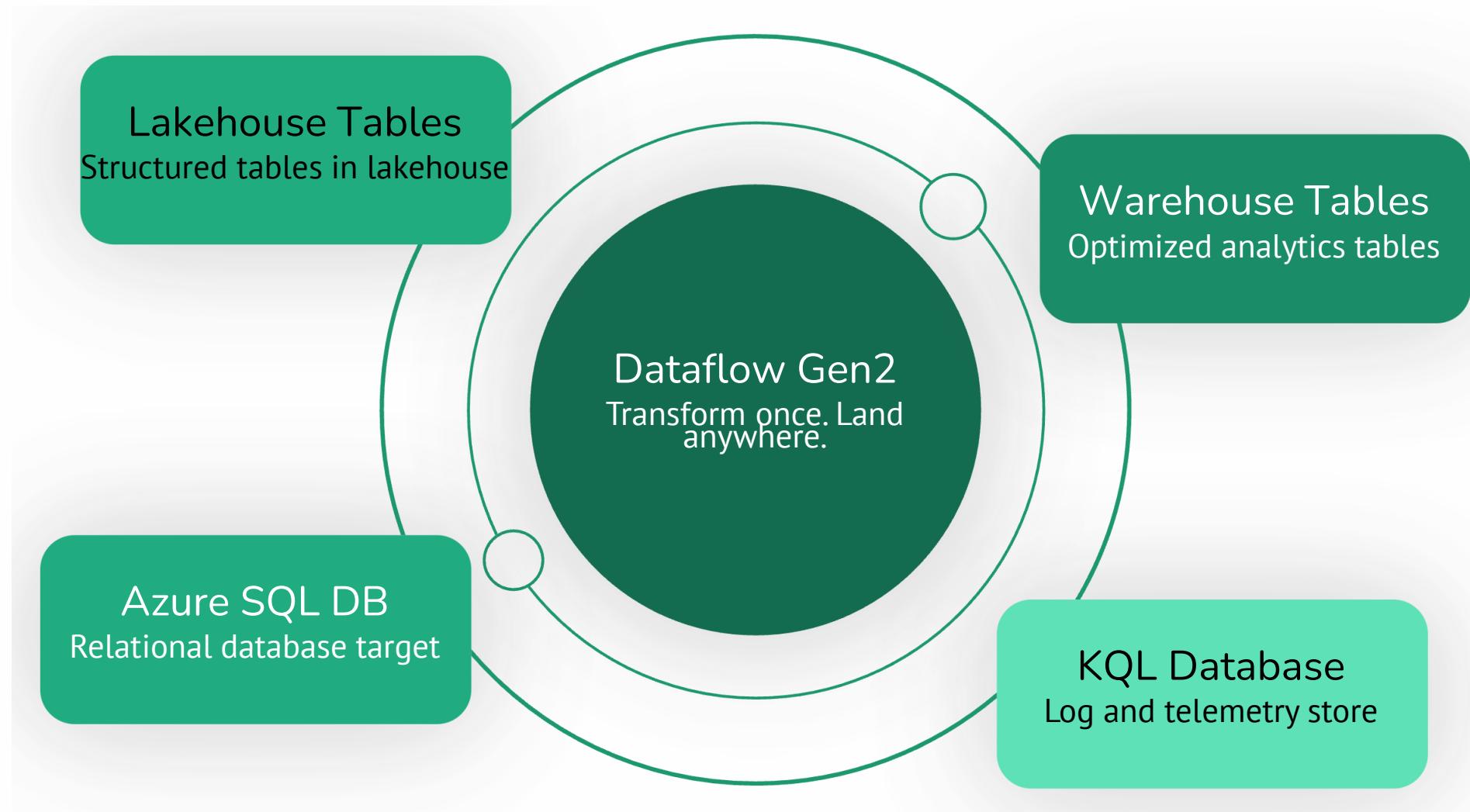
What Changed?

Feature	Gen1	Gen2
Output	Internal storage only	Lakehouse, Warehouse, SQL, KQL
Access data	Dataflow connector	Direct table access
AutoSave	✗	✓
Pipeline integration	Limited	Full
Compute	Mashup engine	Enhanced (faster)
Monitoring	Basic	Monitoring Hub
Requires	Pro or Premium	Fabric capacity

Gen2 is not a version bump. It's an architecture change.

Data Destinations

Output Anywhere



Transform once. Land where you need it.

Pipelines = Orchestration

When You Need Control

Dataflows

Transform data

Single task

Run manually or scheduled

No logic

No notifications

Pipelines

Coordinate activities

Multiple steps

Run on triggers, schedules, events

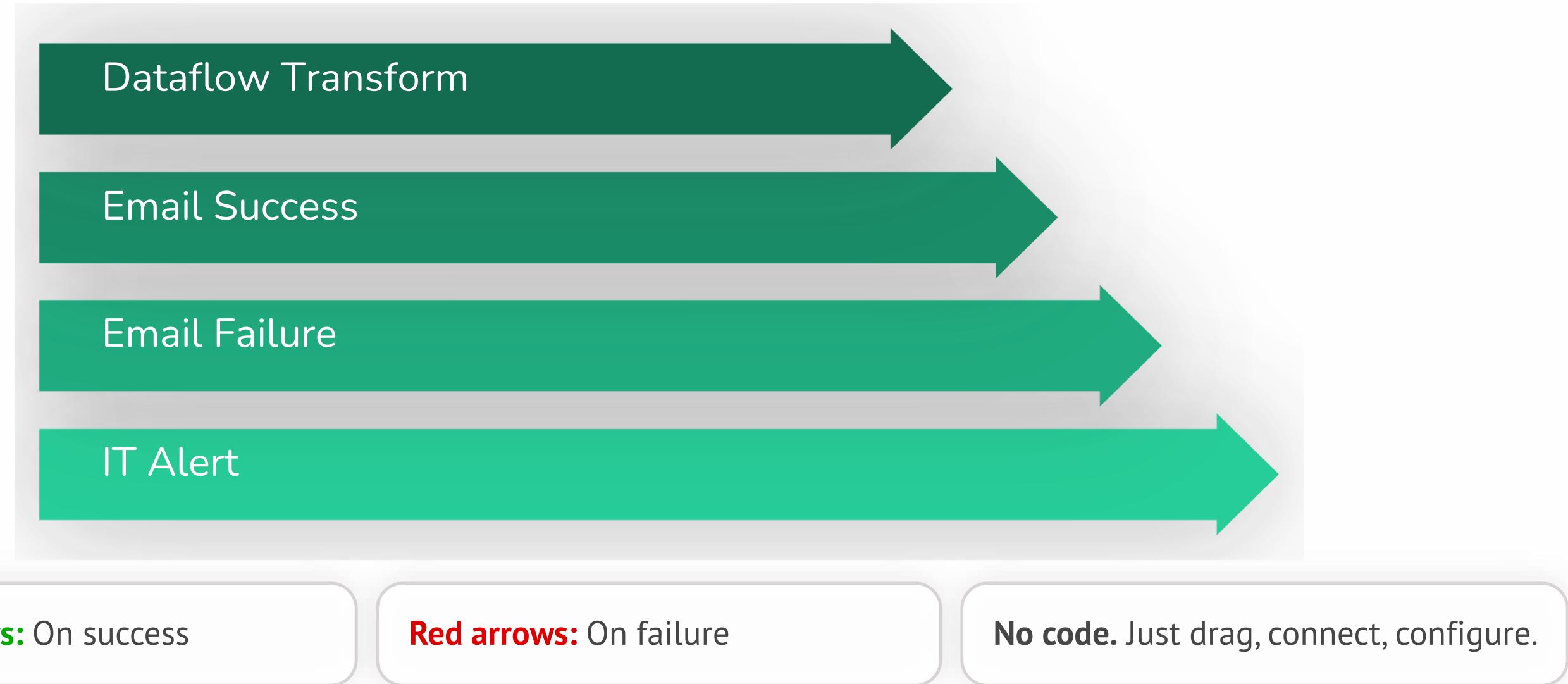
If/ForEach/Until

Email, Teams, webhooks

Dataflows do the work. Pipelines run the show.

Pipeline Flow

Visual Orchestration



Same Engine as Azure Data Factory

Enterprise-Grade

Capability	Fabric Pipelines
Activities	30+ built-in
Triggers	Schedule, event, on-demand
Control flow	If, ForEach, Until, Switch
Parameters	✓ Dynamic values
Monitoring	Full run history
Lineage	Data lineage tracking

This is Azure Data Factory, integrated into Fabric.

When to Use What

Decision Matrix

Scenario	Recommended Tool
Simple ETL, know Power Query	Dataflows Gen2
Complex orchestration	Pipelines
Advanced transforms, ML prep	Notebooks
Just land external files	Shortcut Transformations
Low-code, business user	Dataflows Gen2
Code-first, data engineer	Notebooks

Not either/or. Combine them.

What About Notebooks?

When Code-First Makes Sense

Use Dataflows Gen2 When...

You know Power Query

Standard transformations

Business user friendly

Low-code preferred

Quick wins

Consider Notebooks When...

You know Python/Spark

Complex logic, loops

Data engineer/scientist work

Code-first preferred

ML feature engineering

Notebooks aren't "better." They're different.

 If you're a Power BI user:

- Start with **Dataflows Gen2** (your skills transfer)
- Graduate to **Notebooks** when you hit limits

Monitoring Hub

See Everything

What You Can See

All pipeline runs

All dataflow refreshes

Activity-level details

Duration metrics

Error messages

Why It Matters

Track automation health

Catch failures early

Pinpoint issues

Optimize performance

Debug problems

One place for all your data operations.

Key Gotchas

Watch Out For

Gotcha

Gen2 requires Fabric capacity

Staging Lakehouse is automatic

Pipeline ≠ Dataflow

Email activities need connection

First publish creates staging items

What to Know

Gen1 works on Pro, Gen2 does not

Don't delete the DataflowsStaging items

Pipelines orchestrate, Dataflows transform

Set up Office 365 connection first

Normal behavior, not an error

Your Action Items

Getting Started

01

Create a Dataflow Gen2

02

Apply transformations

03

Output to Lakehouse table

04

Wrap it in a Pipeline

05

Add notifications and schedule

Key insight: Transform with Dataflows. Orchestrate with Pipelines.



Resources:

- MS Learn: "Dataflow Gen2 overview"
- MS Learn: "Pipeline overview"

End of SECTION 4

Thank You

