



# Transforming Legacy Reporting to Cloud-based BI Applications

---

**Sabrina Ng**

TD Bank Group, Enterprise Data & Analytics

# Agenda

01. Overview

---

02. Results

---

03. Solution Framework

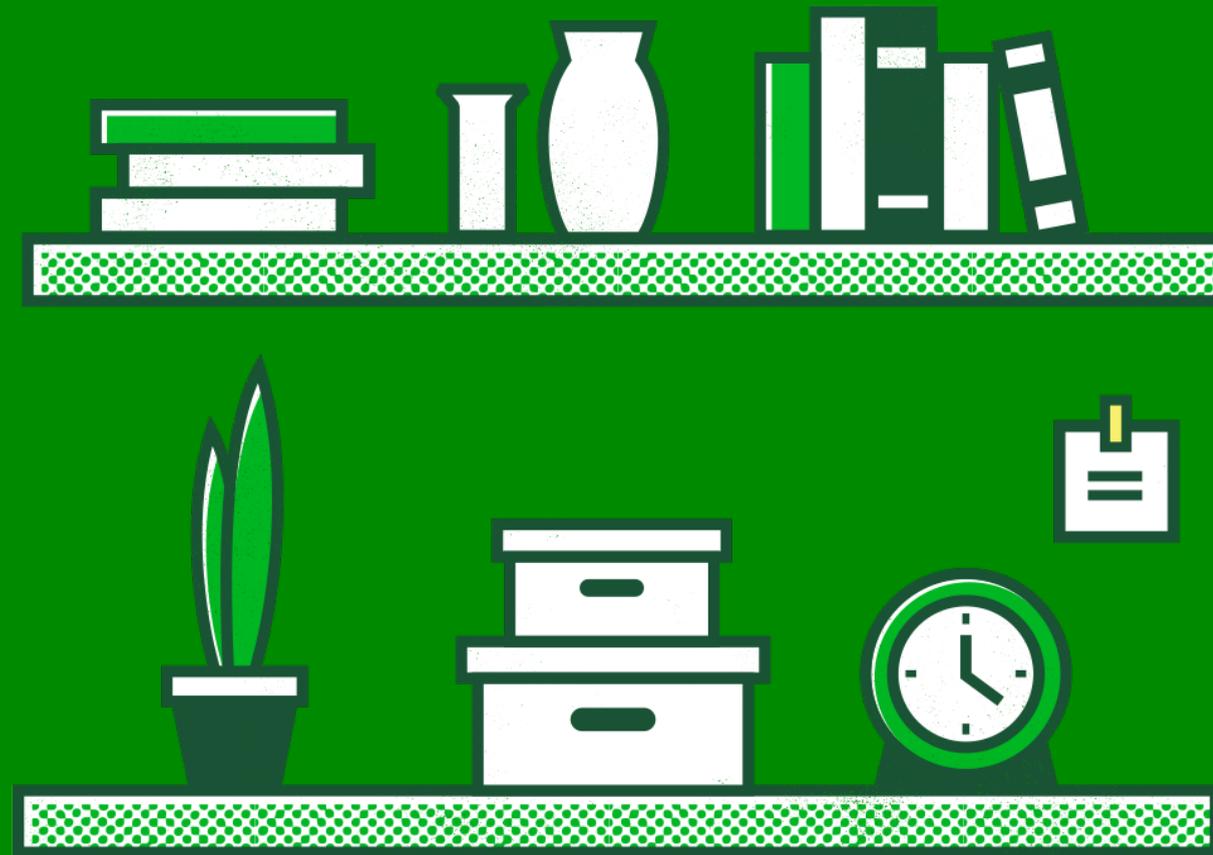
---

04. Databricks Partnership

---

05. Q&A

---



# What was the business problem?

Within a limited timeframe we had to find a strategic solution for all Business Intelligence (BI) challenges.



Over 900+ artifacts

Over 3000+ metrics



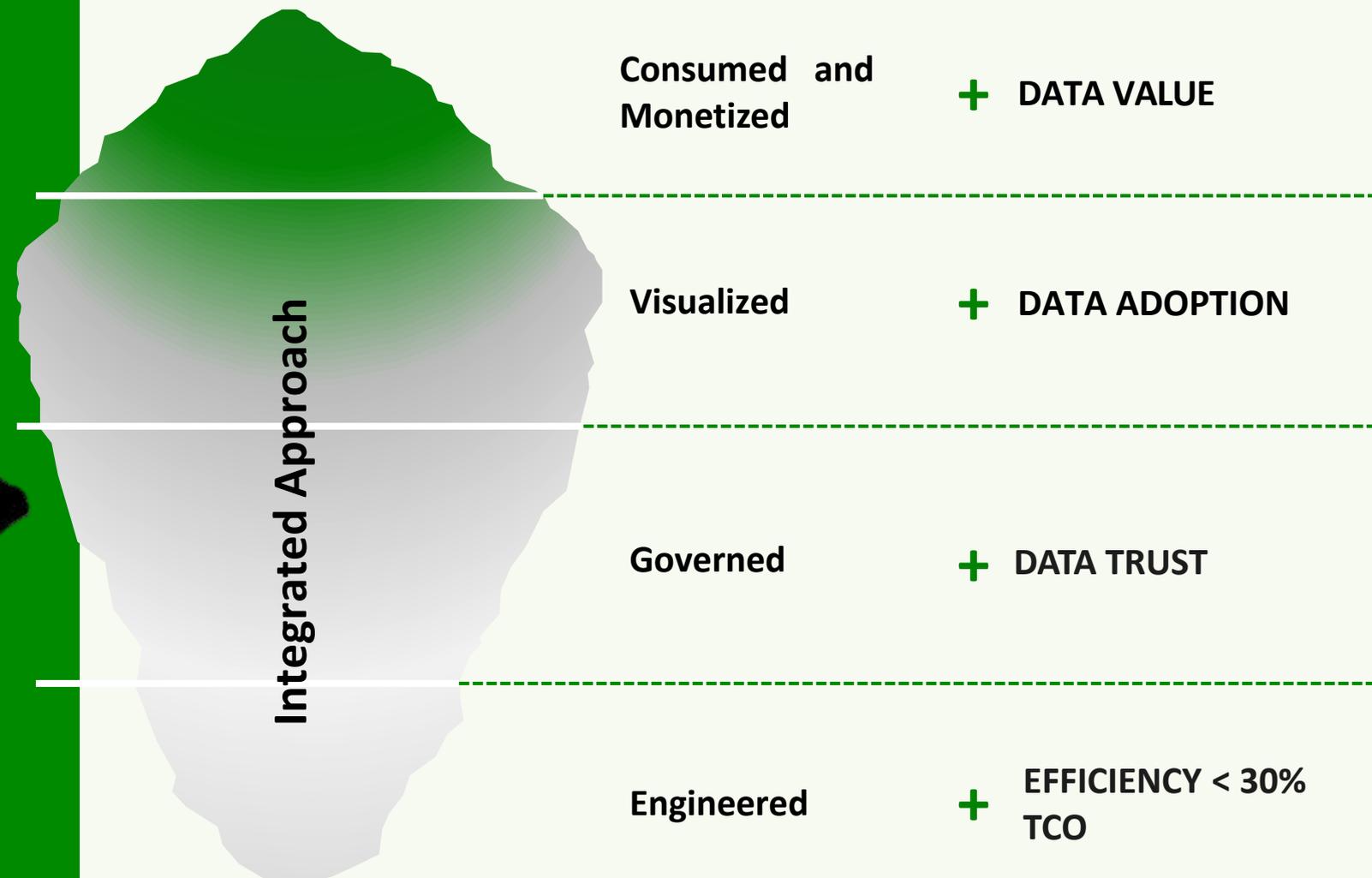
<12 months to self-migrate

Cloud Data-migration delays



No additional resources

Resolution with new Technology



## Why is this different?

- ✓ Business-led and prioritized
- ✓ Solving the data problem from the ground-up pragmatically

# What were the results?

This is more than a reporting initiative – it's a transformation of how we engineered, governed, visualized, and consumed business metrics and data; curated metrics prototyped differently and properly governed.



**Better**

**Increased effective capacity:** freeing up 20-30% of reporting capacity by automating data feeds and streamlining the reporting suite



**Faster**

**Reduce time to deliver:** reduce reports time-to-market by 40-50% through pre-defined and curated metrics for every LOBs



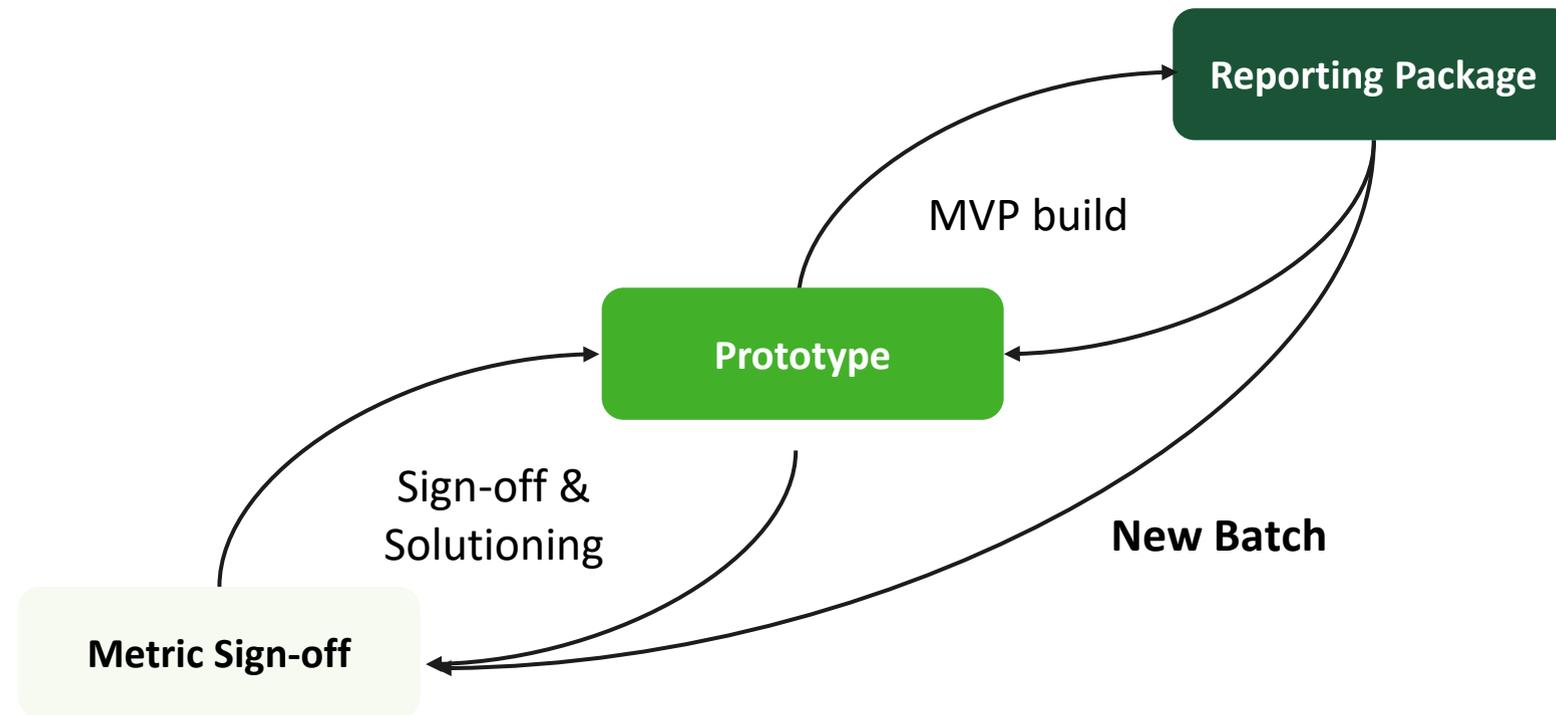
**Safer**

**Trusted and accessible:** lowering operational risks and improved adherence to enterprise standards with data residing in governed platforms

# What was the process?

The solution created a new governance framework and allows the development to go through an iterative process.

## Incremental MVPs for Reporting Hierarchy



## Business Alignment on Metrics by Batches

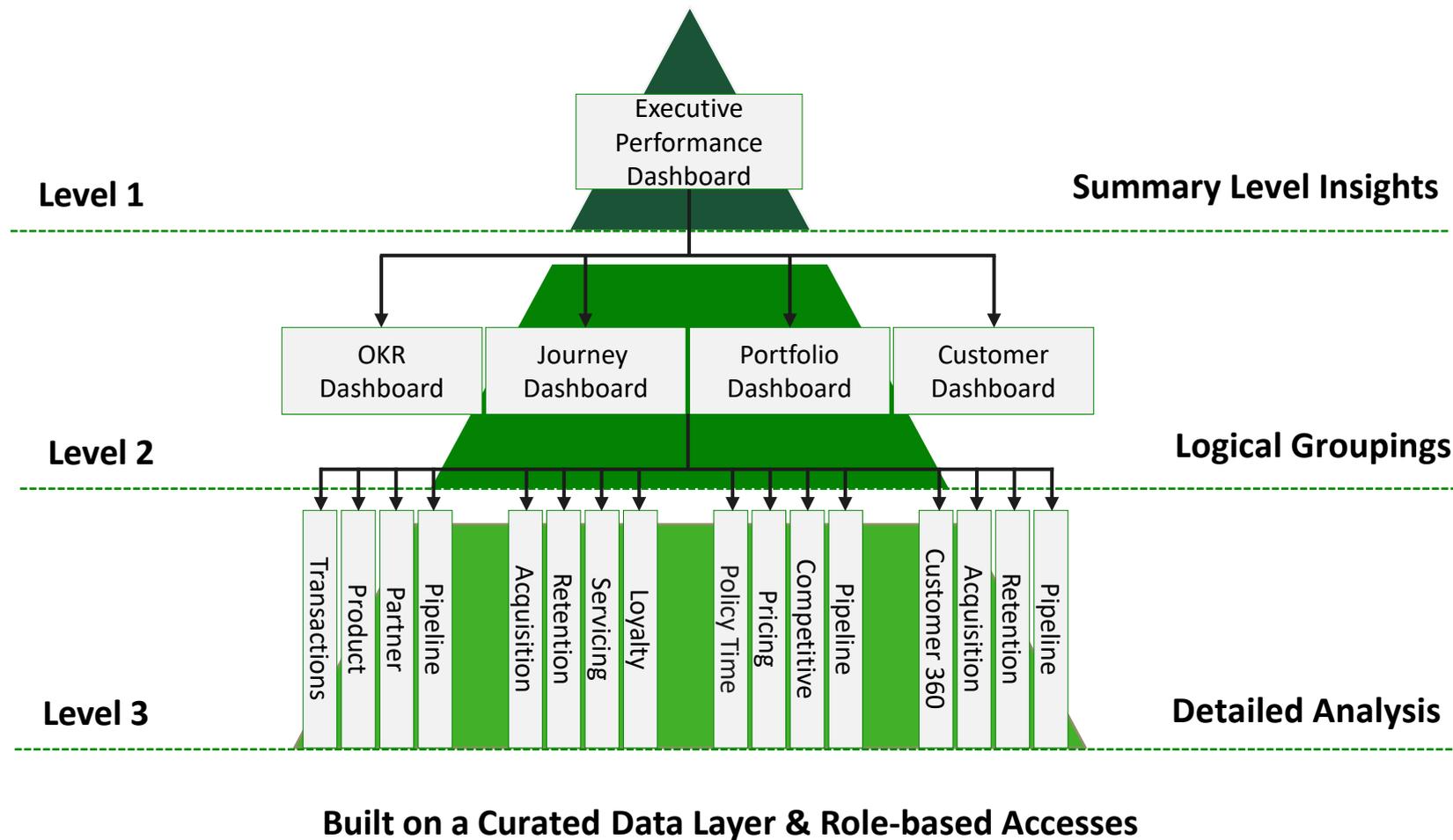


This transforming approach is a game-changer for our business partners in accessing ready-to-use curated data in a faster and safer manner.

VP, Canadian Personal Banking Data & Analytics

# What is the Business Intelligence package?

This is a shift in paradigm from fit for purpose, ad-hoc, multiple artifacts to an integrated and well coordinated BI solution that is curated for specific personas.



## What is the main difference?

1. Standard BI packages built on a single source of truth reporting layer, consumed through reports or cubes.
2. Report visualization with enhanced functionalities, shared best practices and better tooling.
3. Reusable metric inventory across all Personal Banking reports will drive consistence in definition and usage.

## Why are they important?

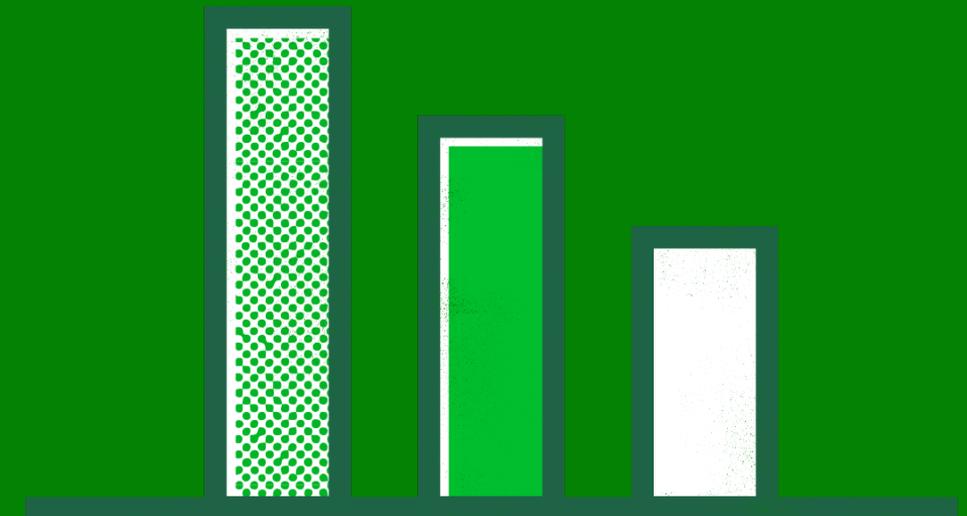
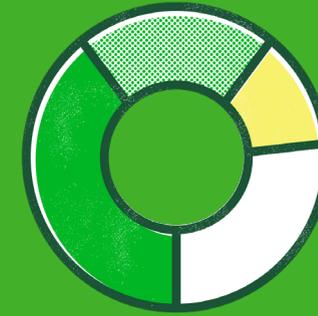
1. Enhancing data-driven culture by enabling accessible insights for business users across all levels.
2. Enabling a sustainable and scalable solution based on leading practices in report development.
3. Driving efficiency by eliminating duplicate and redundant reports and metrics.



**Note:** as we built the new BI packages, we decommissioned the old reports, resulting in efficiency.

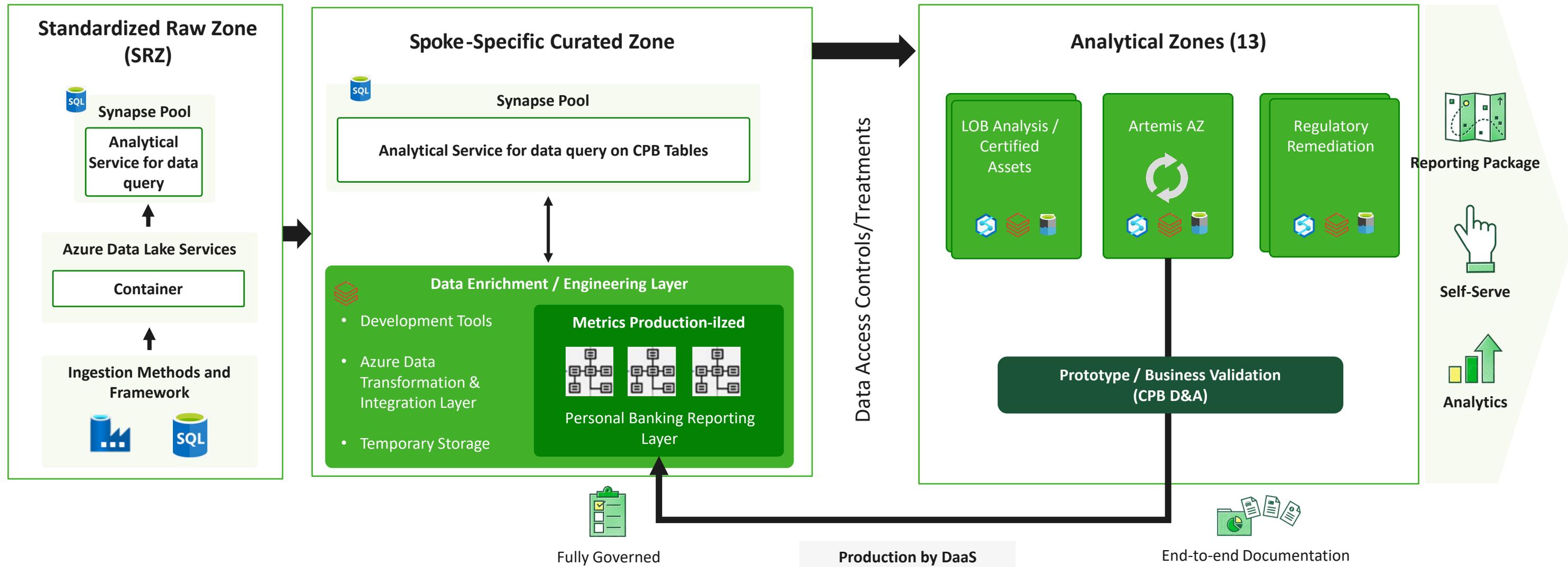
# What were the key curation guidelines?

- 1 Prototype & Production Approach
- 2 End-to-end Data Lineage Documentation
- 3 Data Readiness with Comprehensive Validation
- 4 Curated Metrics are Reusable Assets
- 5 Centralized Design & Delivery Team



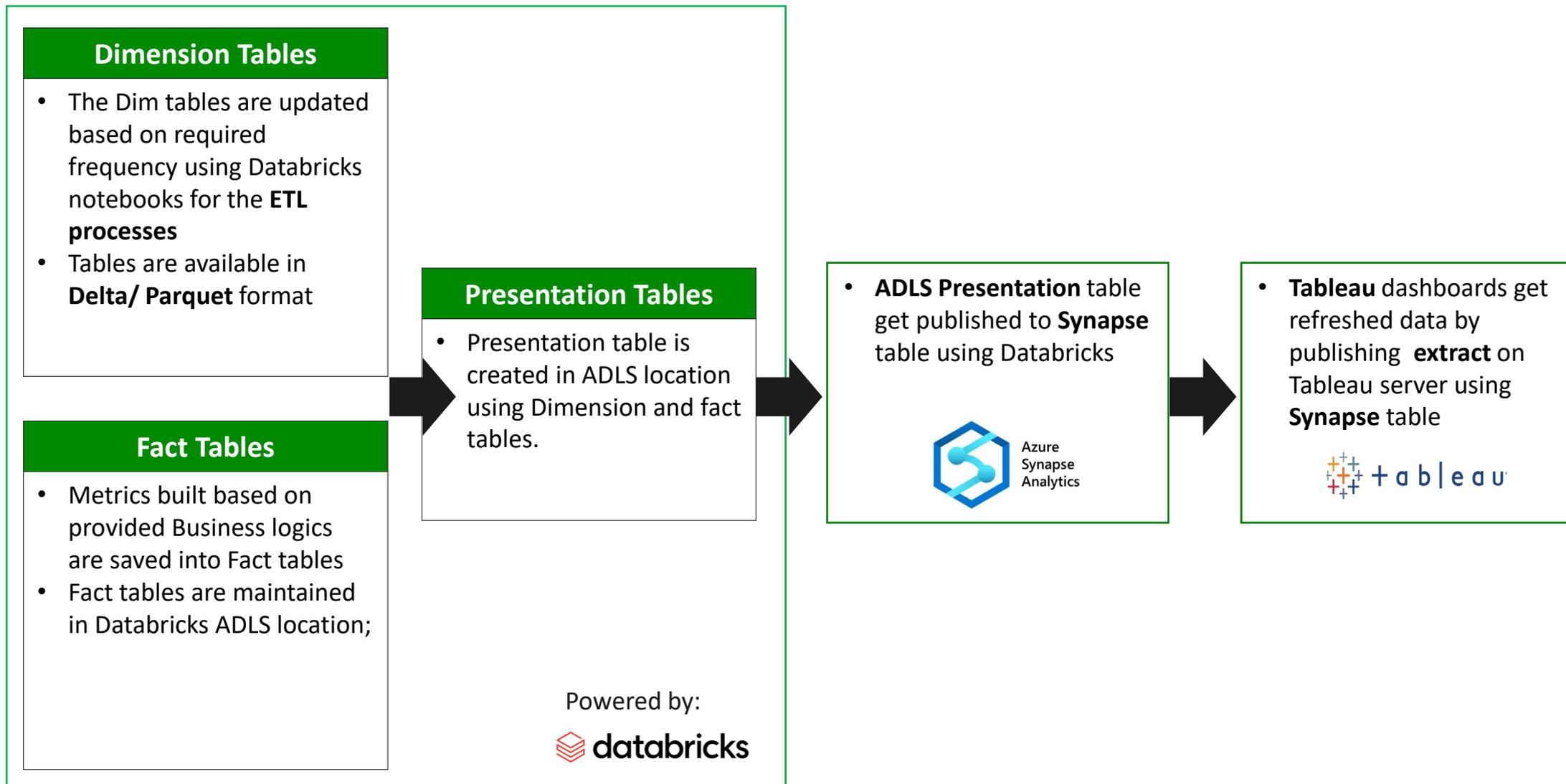
# How did the Target-state Architecture work?

Curated metrics prototyped differently, and properly governed.



# What was the Databricks reporting layer?

Our objective was to have a centralized place to maintain dimensions, fact (metrics), and presentation tables in ADLS (Azure Data Lake Storage).



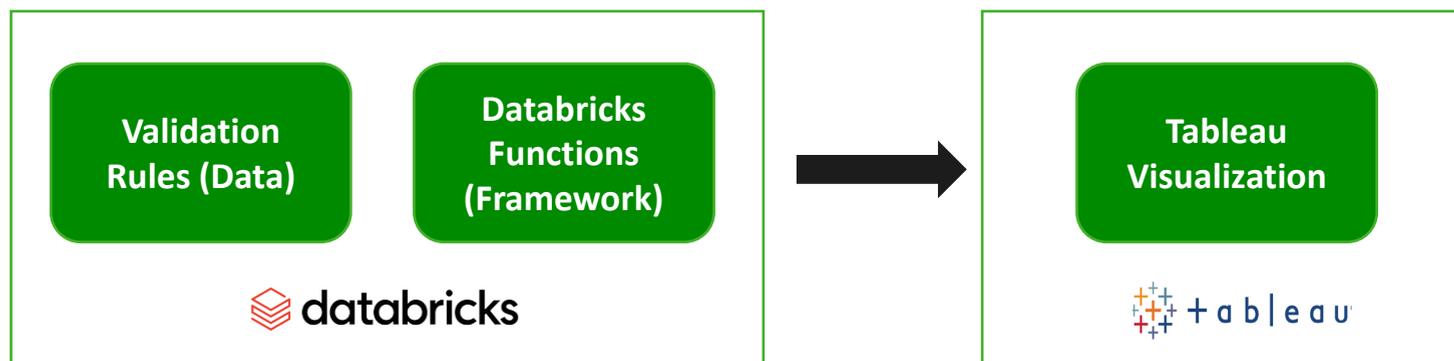
## Benefits:

- 1 Using single/common dimensional table instead of using separate ones for each process within Business AZ space.
- 2 Reusability of the metrics.
- 3 New presentation table can be ready quickly if there is a change in dimension table fields.
- 4 Upsert functionality used by Fact tables does easier update and insert of required records.
- 5 Saves Synapse storage cost by only keeping the final table in synapse required for the dashboard.

# What was the Data Validation framework?

Perform automatic validation on data readiness check, process execution monitor, and output quality control (includes business validation) on multiple processes in Cloud for various line of businesses; it is flexible, scalable, and qualifies for one framework fits all validation rules.

## High-level Design



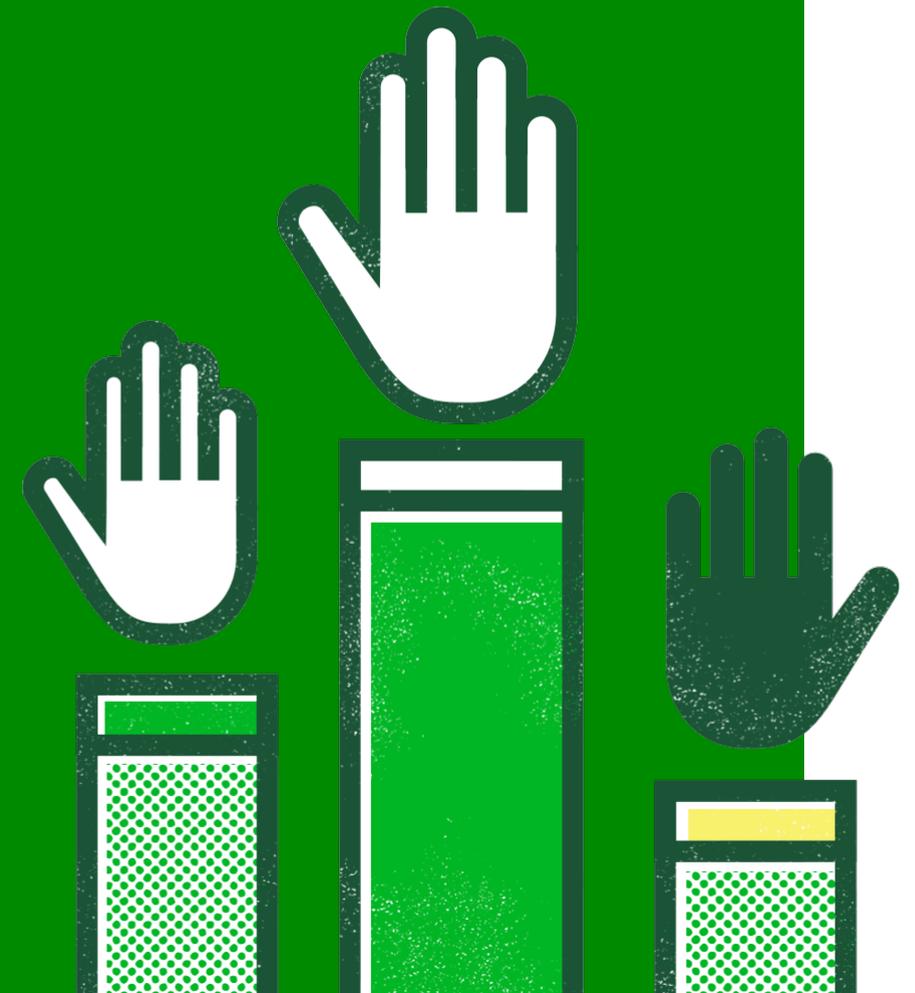
- 1 **Validation Rules:** Users add rules and config execution scenario using the Databricks Notebook in order to automate triggered by the framework.
- 2 **Databricks:** Databricks functions will execute the rules in the validation framework and save the results in a **delta** table, later pushed to a Synapse table.
- 3 **Tableau Visualization:** Tableau dashboard is used to review the validation results and give users ability to monitoring the health of the process and ensure the delivery quality.



## User Insights Derived from the Tableau Dashboard

- Each Line of Business will have their own dashboard view to check the number of successes and failures
- Difference between expected vs current validation results
- No of times the rules re-run by validation framework
- Which data source is bringing the incorrect results and who to reach out for the same
- History tracking to build improvement plan

# Questions?





# Thank you.

**Sabrina Ng**  
TD Bank Group  
Enterprise Data & Analytics