### **Overview**

The **Power BI Data Retention Service** is a self-service platform designed to support the **retention of DAX query results over time**, enabling Power BI reports to access **pre-computed, historical data** without re-executing the logic. Users can register their queries, define how frequently they should run (e.g., daily or weekly), and specify how long the results should be retained in a centralized **Databricks-backed Data Store**.

This service is ideal for reports that require **trend analysis**, or **daily/weekly historical snapshots**—ensuring consistent and reliable data availability without putting pressure on live query performance.

In addition to retention, the service also helps optimize performance for **complex or resource-intensive DAX queries** by executing them asynchronously—either on a schedule or in response to events. Once completed, the results are stored and made instantly accessible to Power BI, significantly improving responsiveness, scalability, and the user experience.

### **🎯 Why This Service Is Planned**

The **Marvel Team** is developing a **Trend Report** that requires access to approximately **250 dates** of historical data. Currently, the Databricks model supports only about 70 days, and the associated DAX queries are already showing signs of **reduced performance**.

As more dates are added daily, these queries will become increasingly **resource-intensive**, impacting report load times and scalability. The Power BI Data Retention Service was created to address this issue by **precomputing and storing daily DAX results**, making it possible to deliver large date-range reports instantly and reliably.

### **✅ How the Service Works**

Marvel Team and others can:

* Register their **DAX queries** via an API
* Specify:
  + 📅 **Retention frequency** (e.g., daily, weekly)
  + 🗃️ **Retention period** (e.g., 30, 60, 90 days)
  + ⚙️ **Trigger mechanism**:
    - **Scheduled** (e.g., every morning at 7 AM)
    - **Event-based** (e.g., after data pipeline completion)

Once registered, the **Power BI Data Retention Service** performs the following:

1. **Executes the DAX query** on the Power BI model (typically for the **current date**)
2. **Stores the result** in a **Databricks table** that acts as the retention store
3. **Retains historical data** based on the retention policy
4. **Makes the data available for consumption in Power BI**, typically through a Direct Query connection
5. **Notifies users** once the data is ready and available for use in reports

This approach ensures a clean separation between DAX execution and Power BI reporting, enabling high-performance report development and seamless access to retained historical insights.

### **🔁 Handling DAX Query Updates**

If a team like Marvel **updates the DAX query** (e.g., adding or removing columns), they simply re-submit the updated definition via the API. The service will:

* Detect structural changes to the query
* Automatically **reprocess and regenerate** the retained data, if required
* Update the underlying retention table schema
* **Invalidate and reload** historical data if the structure has changed
* Notify users once the updated data is available and consistent

This ensures that retained data always reflects the latest business logic without requiring manual cleanup or intervention.