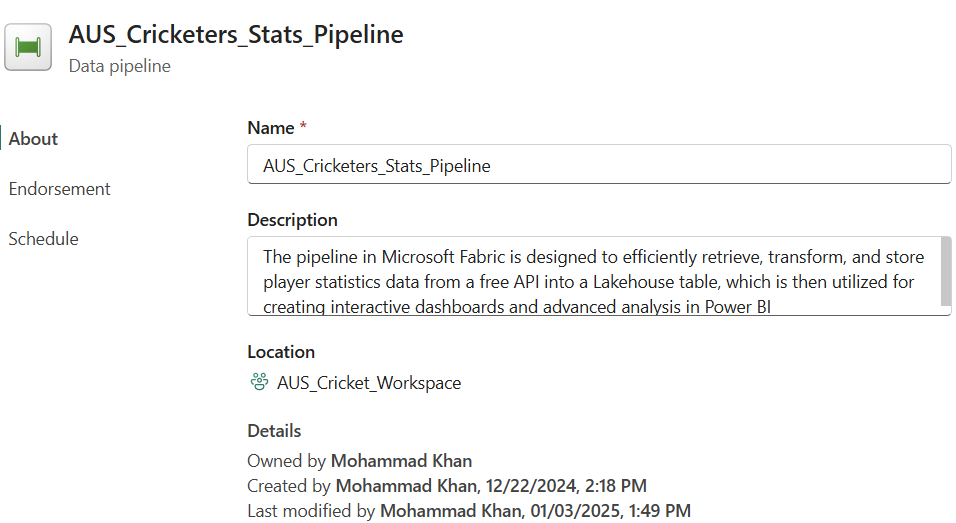
**AUS\_CRICKERERS\_STATS\_PIPELINE**



### Microsoft Fabric Pipeline Description for Player Stats Data Processing and Power BI Integration

The pipeline in Microsoft Fabric is designed to efficiently retrieve, transform, and store player statistics data from a free API into a Lakehouse table, which is then utilized for creating interactive dashboards and advanced analysis in Power BI. Below is a detailed description of each stage of the pipeline:

#### ****Step 1: Data Ingestion****

* The pipeline begins by accepting an array parameter containing player\_ids, where each player\_id corresponds to a unique player whose stats need to be fetched.
* For each player\_id, the pipeline dynamically queries a free API, sending the player ID to retrieve detailed player statistics.
* The API's response, containing raw player stats data, is stored in a DataFrame. This DataFrame serves as an intermediate storage format, ensuring that the data is structured for easy processing and transformation in subsequent steps.

#### ****Step 2: Data Transformation****

* After the raw data is ingested into the DataFrame, the pipeline proceeds to the data transformation phase, which processes the data to meet the structure required for analysis and reporting.
* Key operations in this step include:
  + **Data Cleansing**: Identifying and correcting incomplete or erroneous data entries to ensure data accuracy.
  + **Data Enrichment**: Calculating new fields, aggregating data, or deriving additional statistics from the existing raw data, enhancing the dataset’s value for analysis.
  + **Data Transformation**: Standardizing formats, converting data types, and filtering out irrelevant or unnecessary records to align with reporting requirements.
* The transformed data is saved into a new DataFrame, ensuring that it is structured and ready for storage and downstream use.

#### ****Step 3: Data Storage in Lakehouse Table****

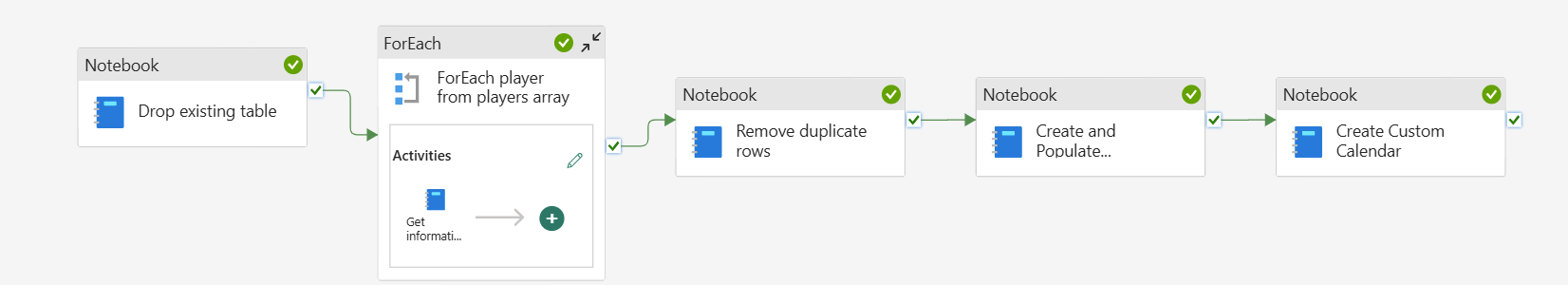
* Once the data has been transformed, it is stored in a **Lakehouse table** named AUS\_CricketerStatus.
* The Lakehouse table serves as a centralized storage solution, offering optimized access to the player statistics data and supporting both batch and real-time data processing.
* Storing data in a Lakehouse table ensures scalability, making it easier to manage large volumes of data and maintain performance over time.
* Additionally, a **semantic model** is created, establishing relationships between various tables to support more advanced data modeling and querying.

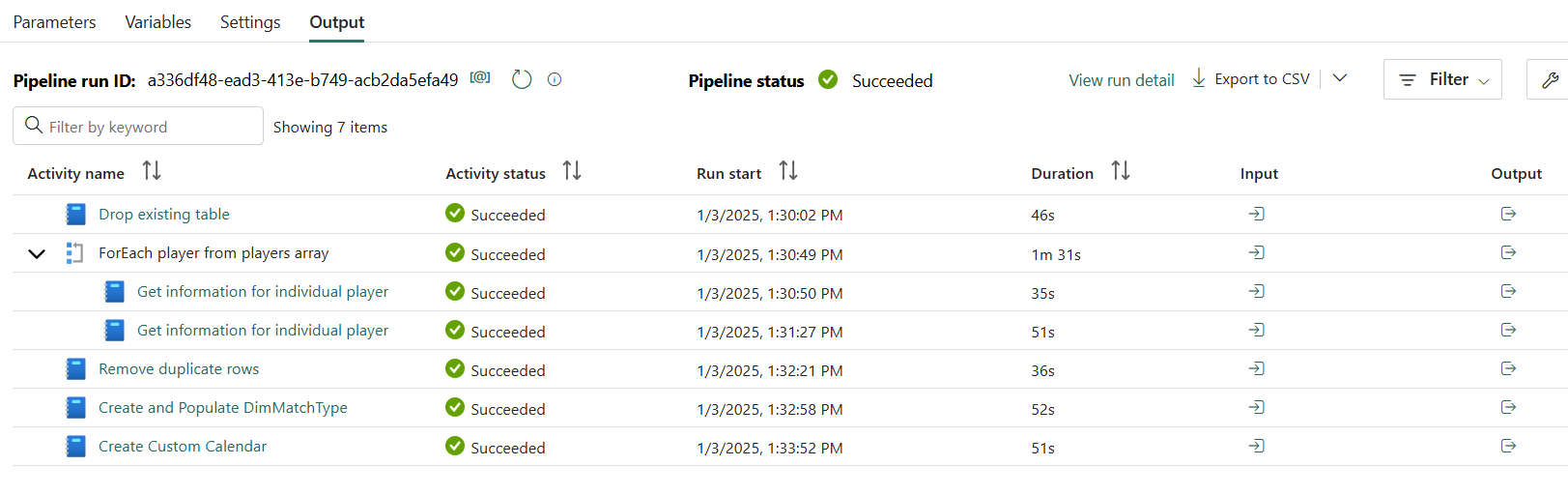
#### ****Step 4: Data Utilization for Dashboards and Power BI Analysis****

* The AUS\_CricketerStatus Lakehouse table is used as the primary data source for creating dynamic and interactive dashboards within Microsoft Fabric.
* These dashboards are designed to visualize key player statistics, such as performance metrics, rankings, and other relevant insights.
* The dashboards are then integrated with **Power BI**, allowing for real-time reporting, detailed data analysis, and business intelligence insights.
* This integration helps decision-makers and analysts explore the player data interactively, enabling more informed decisions and a deeper understanding of player performance.

#### ****Conclusion****

This pipeline provides an end-to-end solution for ingesting, transforming, and storing player statistics data, followed by the creation of insightful and interactive dashboards in Power BI. By ensuring seamless flow from data ingestion to real-time visualizations, the pipeline offers a comprehensive view of player stats, empowering stakeholders with actionable insights for decision-making and performance analysis.





GitHub repository link : <https://github.com/mzaidikhan/AUS_CRICKET>