**Workflow Overview**

1. **Ingestion (Bronze Layer)**
   * Pull Fingrid hydro power API data (JSON/CSV) using Spark notebook.
   * Store directly in Fabric Lakehouse (bronze tables).
2. **Cleaning & Transformation (Silver Layer)**
   * Normalize timestamps, units, missing values.
   * Join with **hydro plant metadata** (capacity, type).
3. **Analytics (Gold Layer)**
   * Aggregations: hourly/daily averages, regional summaries.
   * Calculate load factor = actual output / capacity.
   * Store as clean fact tables for dashboards.
4. **Streaming Simulation**
   * Use Fabric Eventstream (or Python generator) to push real-time power readings.
   * Spark Structured Streaming job consumes → stores into Lakehouse.
5. **CI/CD**
   * GitHub repo with all notebooks & SQL scripts.
   * GitHub Actions workflow runs unit tests (Spark job validation, SQL checks).
   * Deploy updated pipelines automatically into Fabric workspace.
6. **Visualization (Power BI)**
   * Simple dashboard with:
     + Hydro production trends (hourly, daily)
     + Regional breakdowns
     + Anomaly detection (e.g., sudden dips in output)