**Config-Driven Data Pipeline with Exception Reporting & Rule-Oriented Audit**

**Technology used :** Google BigQuery, SQL, JSON, Orchestration using workflows

**Solution Summary :**

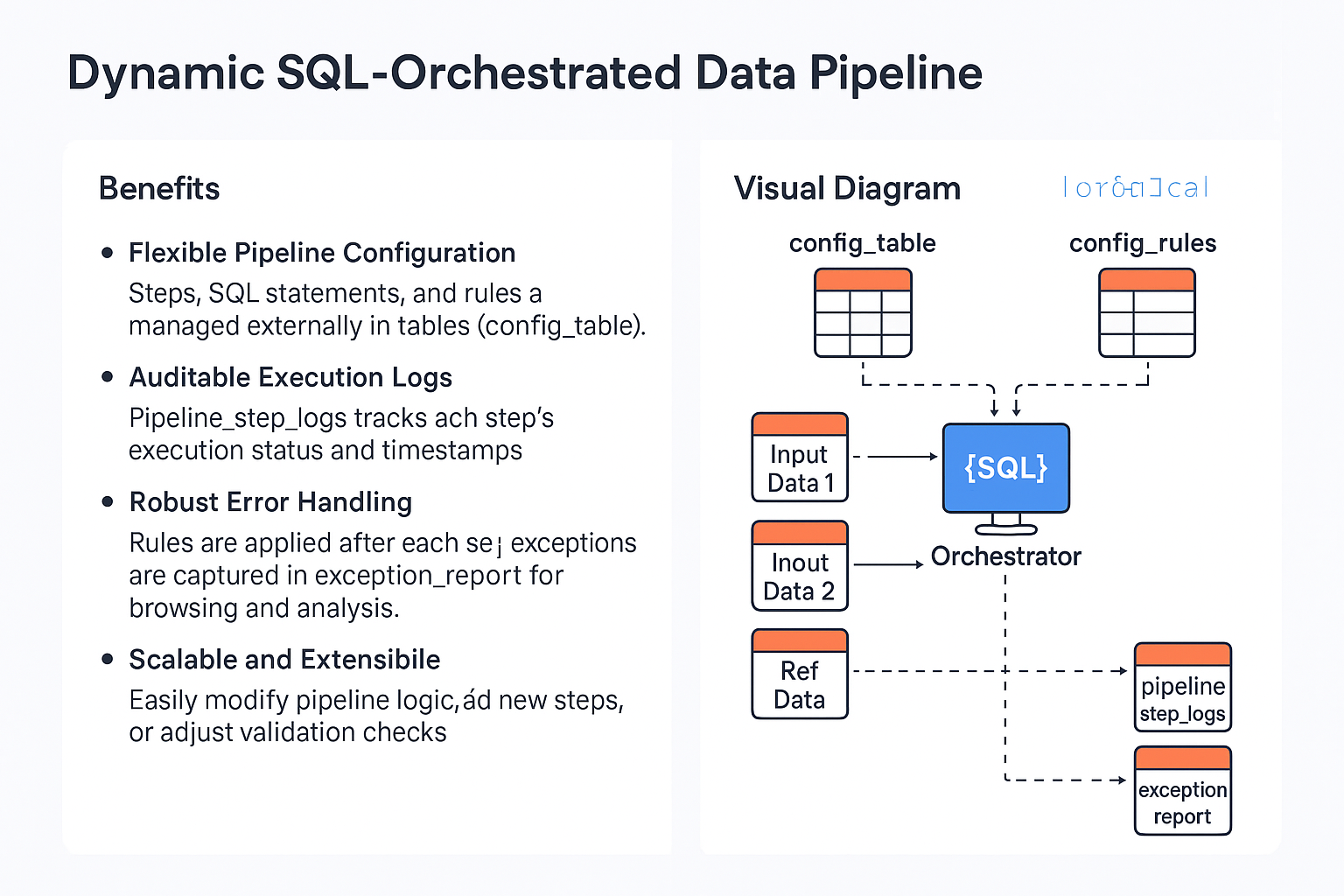
A fully orchestrated SQL pipeline that processes business data through staged transformations and joins, with embedded rule checks and anomaly capture for data quality analysis.

**Key Benefits :**

* **Reusability**: Logic and rules are externalized—no SQL rewrites needed for new datasets.
* **Transparency**: Every step is logged, and every anomaly gets captured in a structured exception report.
* **Governance-Ready**: Rule execution enables real-time validation for business and data quality teams.
* **Modularity**: Inputs, joins, calculations, and aggregations are broken into stages with isolated logic.
* **Flexibility**: Config JSON drives step execution; adding steps or rules doesn’t require pipeline rewrite.
* **Auditability**: Pipeline steps and rule failures are timestamped and linked to execution ID—perfect for compliance and diagnostics.

**Logical components :**

* config\_table for step orchestration
* config\_rules for dynamic validations
* exception\_report to track anomalies
* pipeline\_step\_logs for lifecycle visibility
* Temporary stages (stage1 to stage4)
* Final\_Output for aggregated insights



**Summary**

* **Scalable**: Easily add steps or rules without changing orchestration logic
* **Transparent**: Logs and exceptions are visible and traceable per execution
* **Business-Ready**: Anomaly tracking enables non-technical teams to review data gaps
* **Reusable**: Templates and config structure support cross-domain adoption
* **Governance-Friendly**: Timestamped validations support audit and compliance frameworks

