SAS2PY

SAS2PY Platform Overview

SAS2PY can be deployed on-premise or in the cloud – You Decide! Originally designed for SAS modernization, SAS2PY has evolved into a comprehensive enterprise solution supporting large-scale code and ETL migrations across **SAS, IBM DataStage, Teradata BTEQ, Alteryx, Informatica, Qlik/Talend**, and more.

With unmatched automation, security, and performance, SAS2PY converts legacy code into scalable, cloud-native formats optimized for **Snowflake, Databricks, Google BigQuery, Azure Fabric, Redshift, and DBT** helping enterprises reduce costs, eliminate risk, and accelerate time-to-cloud.

What Makes SAS2PY Different

End-to-End AI Migration Leverage AI across the entire migration lifecycle analyze, convert, execute, and validate delivering fast, consistent, and fully auditable results.

Secure On-Premise Deployment Fully containerized; deploy in your VPC, data center, or behind your firewall. Nothing leaves your environment.
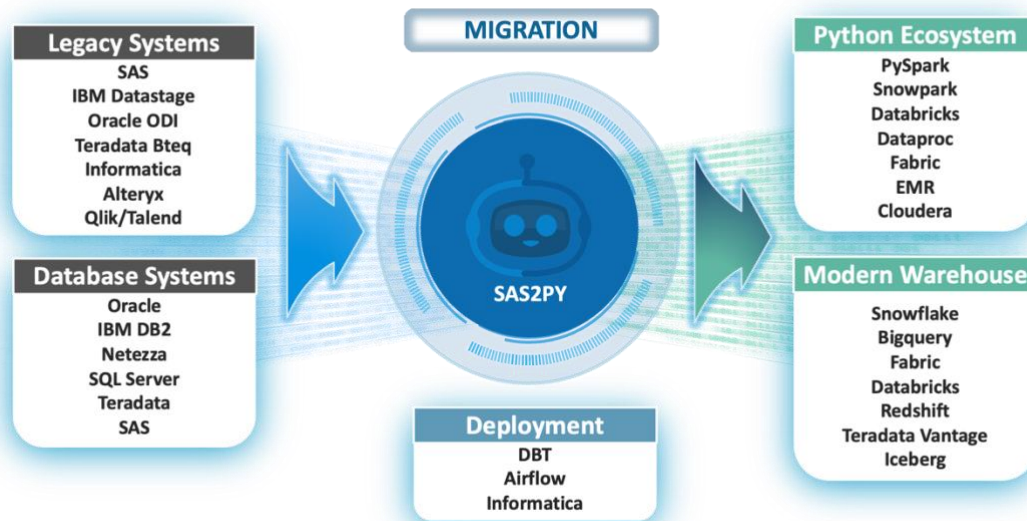
Multi-Platform Support Not just SAS. Migrate from **DataStage, BTEQ, Alteryx, Informatica, Oracle ODI**, and more.

Built-In Execution Layer Visually execute converted code in **Snowflake** and **Databricks** clusters.

(optional) AI Developer Assistant Merlin AI offers real-time guidance, debugging, and optimization during and post-migration.

Full Validation and Data Matching Guarantee consistency of logic and data before, during, and after migration.

SAS2PY

Legacy Systems/Database Systems/Python Ecosystem/Modern Warehouse/Deployment

MIGRATION

Legacy Systems
SAS
IBM Datastage
Oracle ODI
Teradata Bteq
Informatica
Alteryx
Qlik/Talend

Database Systems
Oracle
IBM DB2
Netezza
SQL Server
Teradata
SAS

SAS2PY

Python Ecosystem
PySpark
Snowpark
Databricks
Dataproc
Fabric
EMR
Cloudera

Modern Warehouse
Snowflake
Bigquery
Fabric
Databricks
Redshift
Teradata Vantage
Iceberg

Deployment
DBT
Airflow
Informatica

Simply upload or connect to file folders

1. Analyze Out-of-the-box
• Inventories legacy systems, ETL workflows, and codebases
• Automatically maps dependencies across SAS, DataStage, Alteryx, etc.
• Prioritizes migration sequences and creates resource timelines

2. Convert Out-of-the-box
• High-speed, parser-based translation of SAS, BTEQ, Informatica, and more
• Converts code logic into modular, modern, PEP8-compliant formats
• Preserves macros, metadata, and ETL logic with minimal refactoring

3. Validate
• Tests logical equivalence across platforms
• Validates schema, data types, structure, and results
• Supports row-level and aggregate comparison

4. Execute
• Execute transformed pipelines directly in Snowflake, Databricks & more
• Visual workflow control for job packaging, scheduling, and debugging
• DAG interface and versioning support included

Migrate Smarter. Analyze Faster. Transform with SAS2PY.
www.sas2py.com   sales@sas2py.com

SAS2PY

5. Data Matching
• Compare source and target data for metric and structure alignment
• Partition-based validation (e.g., by date, region, customer)
• Automated reconciliation of schema and formats
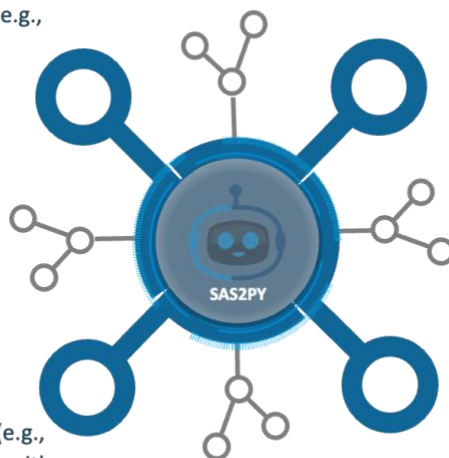
Automated Schema Mapping

Automatically maps source schemas (e.g., **SAS**, **Oracle**, **Teradata**) to the target system for fast and efficient setup.

Metadata Comparison

Compares metadata (e.g., **table structures, indexes**) between the source and target systems to ensure complete structural alignment.

Data Type Validation

Ensures accurate translation of column types (e.g., numeric, string, date) into formats compatible with the target system, eliminating type mismatches.
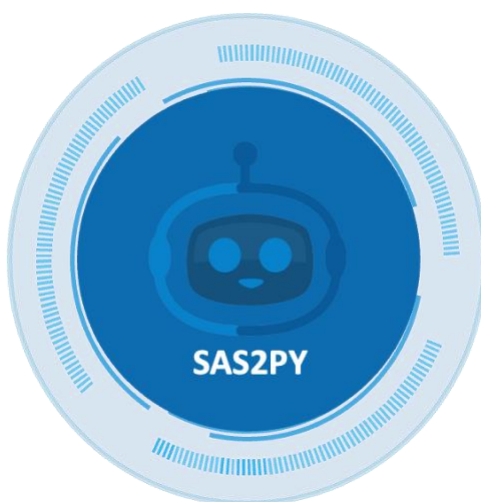
Metrics Comparison

Validates key metrics (**counts**, **sums**, **averages**, etc.) between source and target systems to confirm data accuracy.

Partitioned Validation

Performs aggregate checks by **partitions** (e.g., **date**, **region**) to ensure data consistency within subsets.

6. Document
• Auto-generates migration reports, inline code comments, and lineage maps
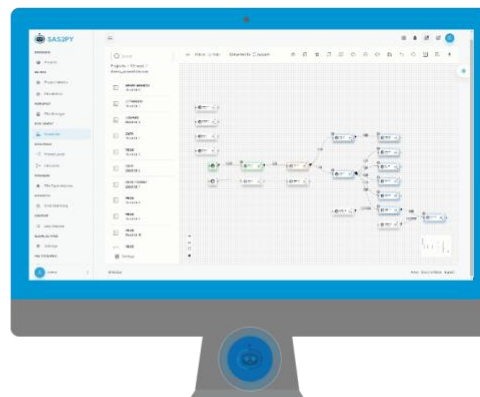• Preserves legacy code view for audits and rollback

7. Code/Data Lineage

End-to-End Code Lineage: SAS2PY **automatically maps dependencies** across all workflows, ensuring **100% traceability before, during, and after migration.**

Clickable Lineage Exploration: Users can **click on any dataset, transformation, or script** to immediately view the **original code alongside its migrated equivalent.**

Legacy Code Preservation & Full Auditability: Every **original script** remains **fully viewable, traceable, and auditable,** ensuring **long-term historical reference and compliance validation.**

Merlin AI – The Developer Assistant (Optional)

Merlin AI is a GenAI developer assistant for developers and data engineers working with migrated code. Built directly into the SAS2PY platform, Merlin supports:

• Real-time debugging and transformation assistance
• Auto-rewrite of SQL and Python blocks for optimization
• Error tracebacks, logic flow explanations, and quick fixes
• GenAI-enabled context-aware chat for specific code inquiries
• On-premise optional for security-sensitive environments

Enterprise-Grade Data Matching & Validation

SAS2PY includes a powerful data validation engine to ensure consistency and trust throughout your migration:

• **Automated Schema Mapping**: Aligns source and target columns and formats
• **Data Type Normalization**: Ensures compatibility across platforms
• **Metadata Comparison**: Verifies structures, indexes, and dependencies
• **Metric Validation**: Confirms row counts, aggregates, and KPI accuracy
• **Partition-Level Checks**: Detects mismatches by date, region, or business slice

Deployment Options & Infrastructure

- Deploy within your infrastructure or in the cloud – you decide
- Docker-based delivery; compatible with AWS, Azure, GCP, on-prem VMs
- Standard Instances: AWS m5.2xlarge, Azure D8s v3, GCP n2-standard-8
- No data or logic ever leaves your environment

Licensing & Engagement Options

- **Volume Pricing** Tiered by LoC (lines of code) and platforms migrated
- **Proof of Concept (PoC)** 4-week paid PoC
- **Partner-Friendly** We work with many SI's, technology partners and cloud migration partners

Take the Next Step in Digital Transformation

SAS2PY is the world's most secure, flexible, and automated legacy modernization platform—built to meet the needs of large enterprises across regulated, data-intensive industries. Whether modernizing SAS code, rebuilding ETL pipelines, or migrating entire platforms to the cloud, SAS2PY accelerates your path forward while reducing risk and cost.

Schedule a Demo: www.sas2py.com | sales@sas2py.com