



Empower your teams with the latest Prophecy innovations



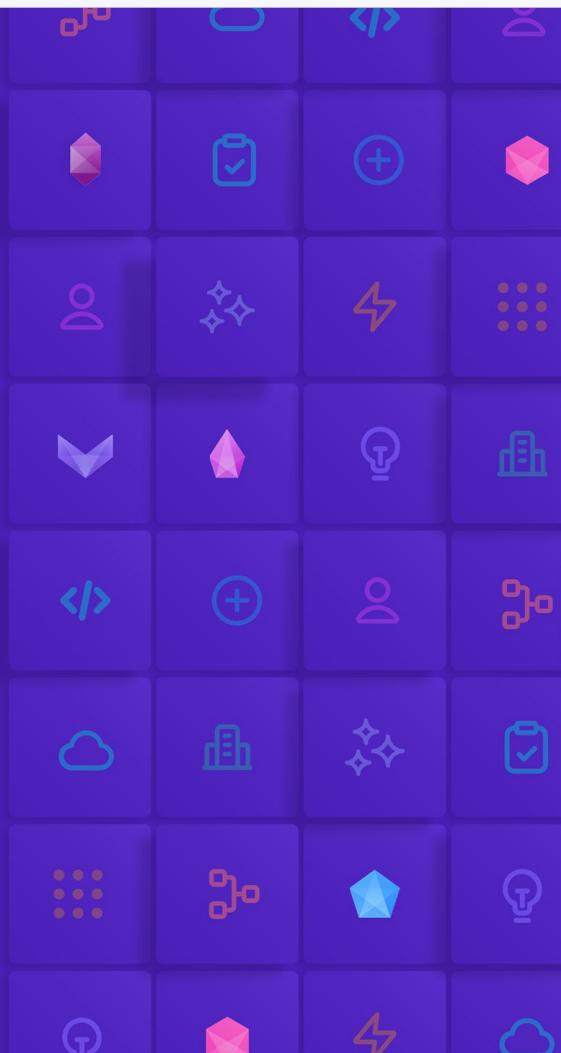
Nathan Tong
Sales Engineer



Anya Bida
User Advocate



Nimbus Goehausen
Fullstack Software and Data Engineer



Agenda

Today's main objective: Prophecy Platform Update for all Data Users

Will there be
demos?

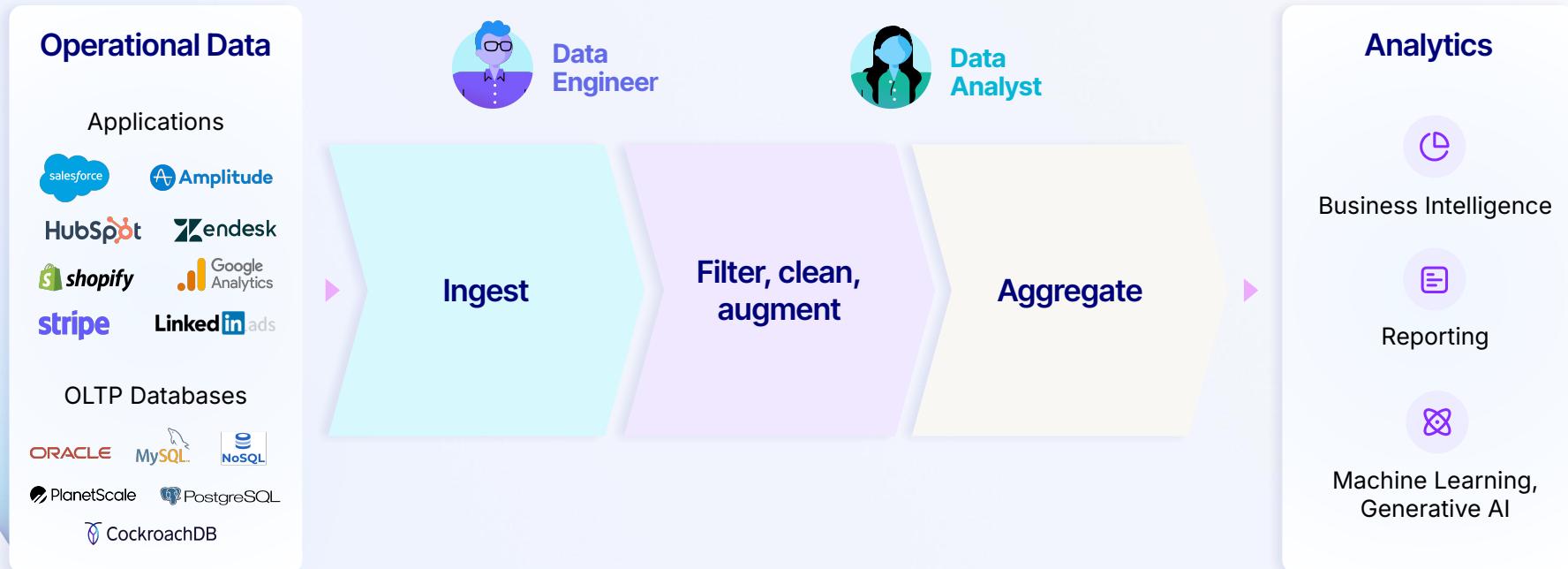
Yes! In 5 mins!

Can I ask
questions?

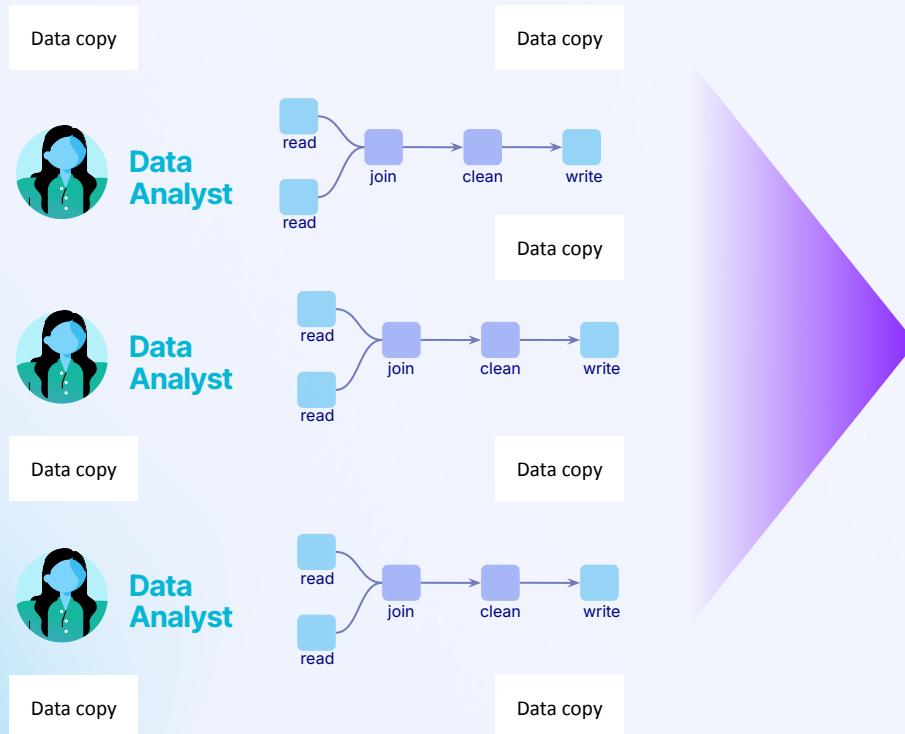
Best question in Q/A
tab wins a Bluetooth
Speaker!

Clean, trusted data fuels AI & analytics

Organizations that use high-quality, well-governed data are 3x more likely to see AI success (source: Gartner)



Analysts can prepare data on their own... But now you've introduced more risk.



Governance risk



Cost

Engineering rework, as analysts use tools not built for production.



Quality

Are analysts using the right data?
Following company standards?



Security

Data copies make permissions management a nightmare.



What the business demands

Self-service

Data analysts prepare data
on their own

Data
Analyst



Governance

Working within guardrails set by the
central data platform team

Data
Engineer



Prophecy apps

Business-facing apps for advanced analytics



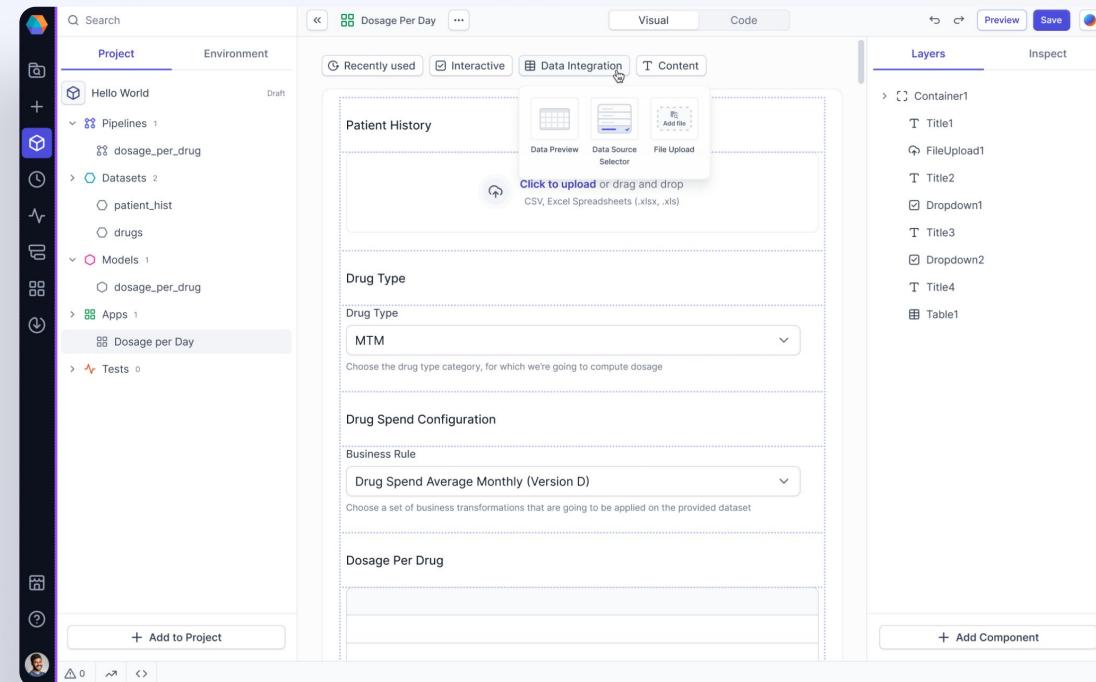
Higher quality

Easy access to common, potentially complex analytics through standard, reusable components.



Cost control

Reduce repetitive effort and eliminate suboptimal pipelines by enabling experts to build pre-configured workflows for the business.

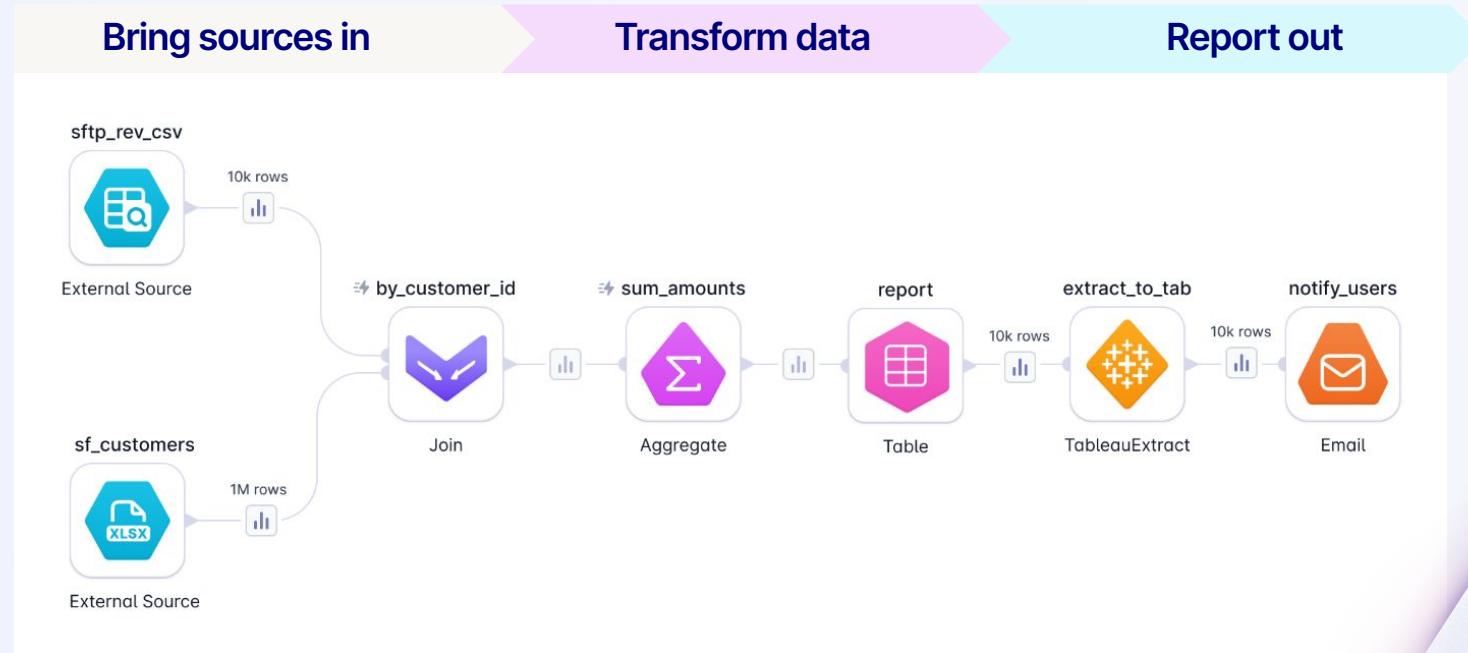


The screenshot shows the Prophecy app builder interface. On the left, the Project sidebar lists components: Pipelines (dosage_per_drug), Datasets (patient_hist, drugs), Models (dosage_per_drug), Apps (Dosage per Day), and Tests. The main workspace displays the 'Dosage per Day' app configuration. It includes sections for 'Patient History' (with a file upload selector), 'Drug Type' (set to MTM), 'Drug Spend Configuration' (Business Rule: Drug Spend Average Monthly (Version D)), and 'Dosage Per Drug'. The right side shows the 'Layers' panel with components like Container1, Title1, FileUpload1, etc., and the 'Inspect' panel.

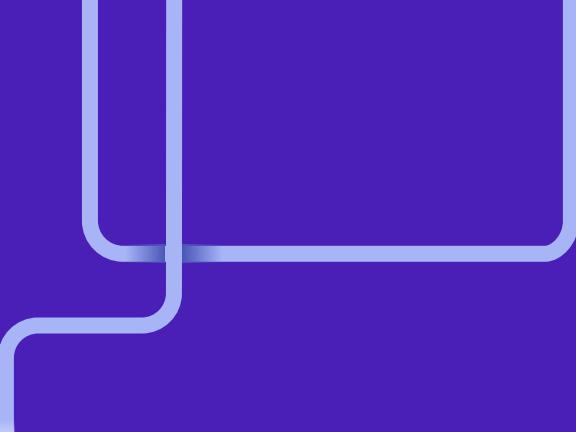
Demo



A familiar, drag-and-drop interface covering end-to-end data prep



Demo



Streamline governance to reduce risk

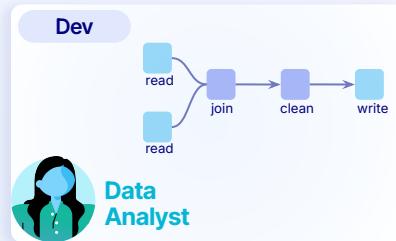


Demo



Collaborative & Simplified version control

All the power of Git, with none of the complexity



The screenshot shows a data pipeline management interface with several windows open. One window displays a pipeline named 'revenue_per_customer' with stages: 'sf_customer' (Source), 'sf_rev_csv' (Source), and 'revenue_per_customer' (Processor). Another window shows a 'Models' section with entities like 'rev_per_cust', 'customers', and 'revenue'. A third window is a 'Save to version history' dialog box, which includes fields for 'Version description' and a table showing changes since the last save, such as 'CustomersOrders' created, 'GoldReport' modified, and 'Orders' deleted. The interface also includes standard file operations like 'Save', 'Show version history', 'Undo', and 'Redo'.

Higher quality

Analysts can easily share their pipelines with colleagues, including expert data engineers. Engineers review and refine, ensuring pipelines meet performance and quality standards before going into production.

Cost control

Git serves as a “control point”. Data engineers can promote pipelines to production after they’re fully tested and optimized.

Prophecy Packages

Enforce pipeline best practices

Higher quality



Central data teams share standard pipeline operations with data analysts.

Analysts build compliant, high-quality pipelines.

- **Simplify discovery** of standard, reusable pipeline components through a single, searchable hub of packages
- **Facilitate collaboration** as platform leads & architects share standard types of transformations with data engineers & analysts
- **Enable best practices** by promoting a standardized & common approach to building data pipelines

| Type | Count |
|-----------|-------|
| All | 32 |
| Pipelines | 10 |
| Datasets | 1 |
| Gems | 16 |
| Functions | 3 |

| Fabric Provider | Count |
|-----------------|-------|
| Spark | 24 |
| Apache Airflow | 24 |
| Spark SQL | 4 |
| Snowflake SQL | 4 |

| Team | Count |
|----------|-------|
| Prophecy | 24 |
| Workday | 4 |



Data analysts can get started now, securely

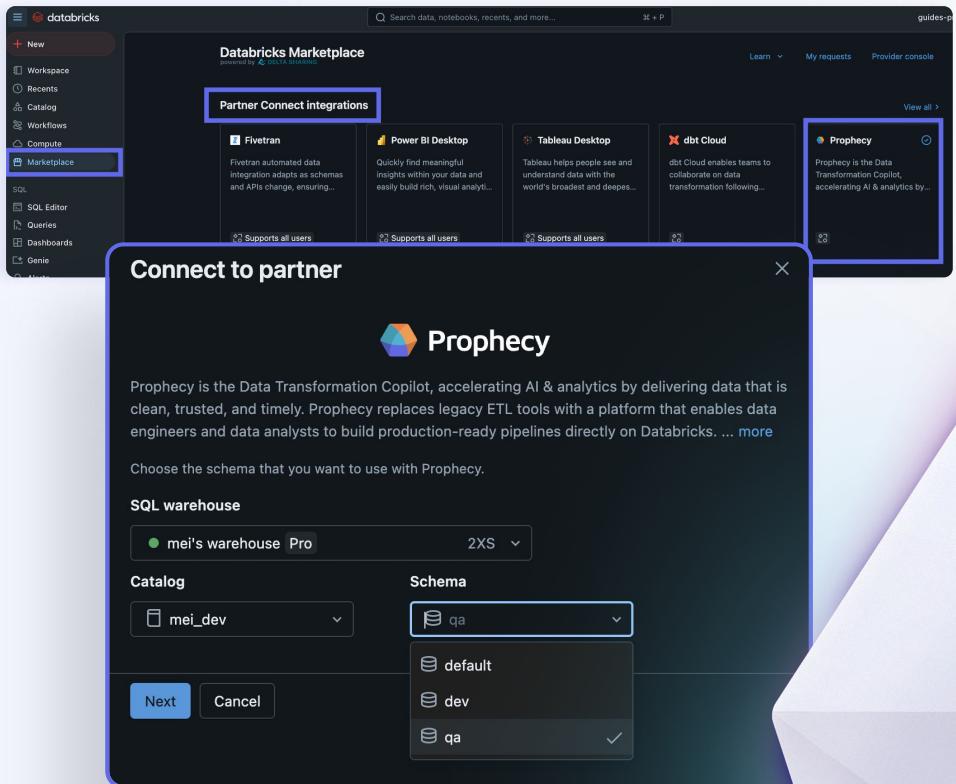
Databricks Partner Connect w/ OAuth

Before Partner Connect with OAuth:

- Required manual setup
- Admin-only access: Most users had no access to Prophecy via Partner Connect

Benefits of Partner Connect with OAuth

- Use their existing environment on Prophecy without additional setup.
- Access for all users: admins and business users can integrate with Prophecy without requiring special privileges.
- PAT tokens are no longer needed since users can utilize their existing Databricks environment without additional setup.





Prophecy Secure by design

Prophecy was designed to meet the rigorous security requirements of the world's largest enterprises, including those in highly regulated industries

Industry leaders trust Prophecy to support their data operations.



RALPH LAUREN



Authentication

Authentication methods, including Single Sign-On (SSO) and Multi-Factor Authentication (MFA), ensure only authorized individuals can access the platform.

Authorization

Role-Based Access Control (RBAC) enforces a "least privilege" model, granting users only the permissions they need.

Auditing

Detailed auditing capabilities log all user and admin activities for complete transparency and accountability.

Data protection

Data is protected through encryption at rest and in motion, with compliant key management practices to safeguard sensitive information.

Self service and Governance

...together, at last!

Recap

Prophecy 4.0 Release

** MAJOR IMPROVEMENTS IN 4.0 WITH ANALYST PROJECTS

DEVELOP PIPELINES



COLLABORATE



PACKAGES



NEW IN 4.0

MONITOR WITH PROPHECY AUTOMATE



NEW IN 4.0

ORCHESTRATE WITH PROPHECY AUTOMATE



LINEAGE



AI COPILOT



NEW IN 4.0

BUSINESS APPS





We'd love to hear what you think!

Best question in Q/A tab
wins a bluetooth speaker!



That concludes our webinar.

Thank you for joining us!



Feature detail slides



Schedule pipelines with built-in automation

✓ All-in-One Solution.

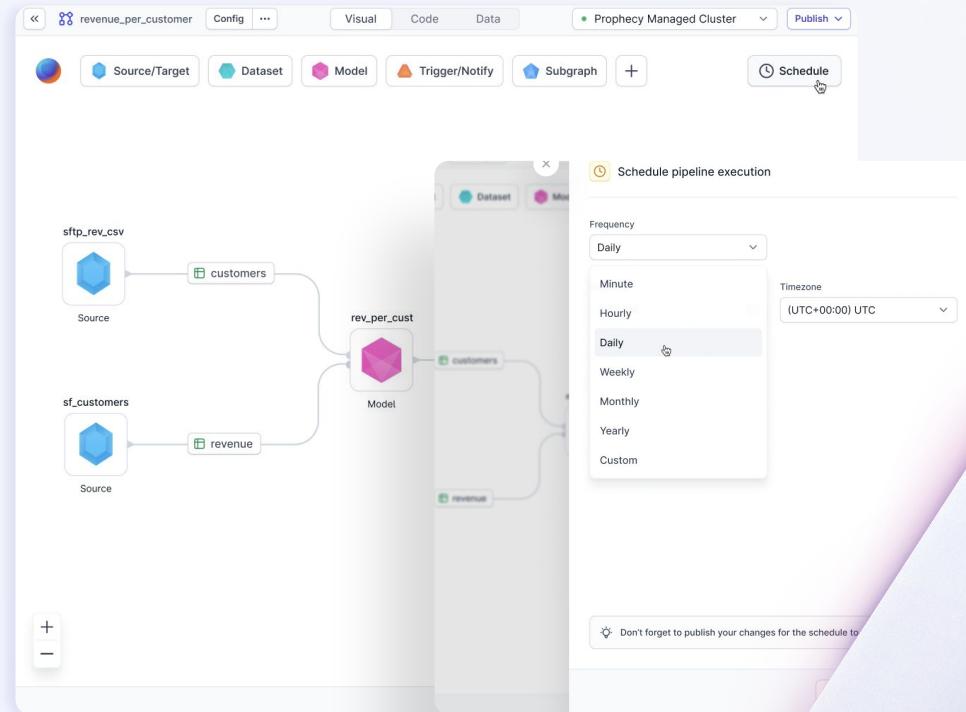
No need for separate orchestration tools. Schedule, manage, and monitor pipelines directly within Prophecy.

✓ Easy for Analysts.

A simple drag-and-drop interface makes scheduling effortless, so analysts can validate and run pipelines on their own.

✓ Seamless Workflow Management.

Build, schedule, and monitor data pipelines in one place, ensuring smooth execution and faster insights.



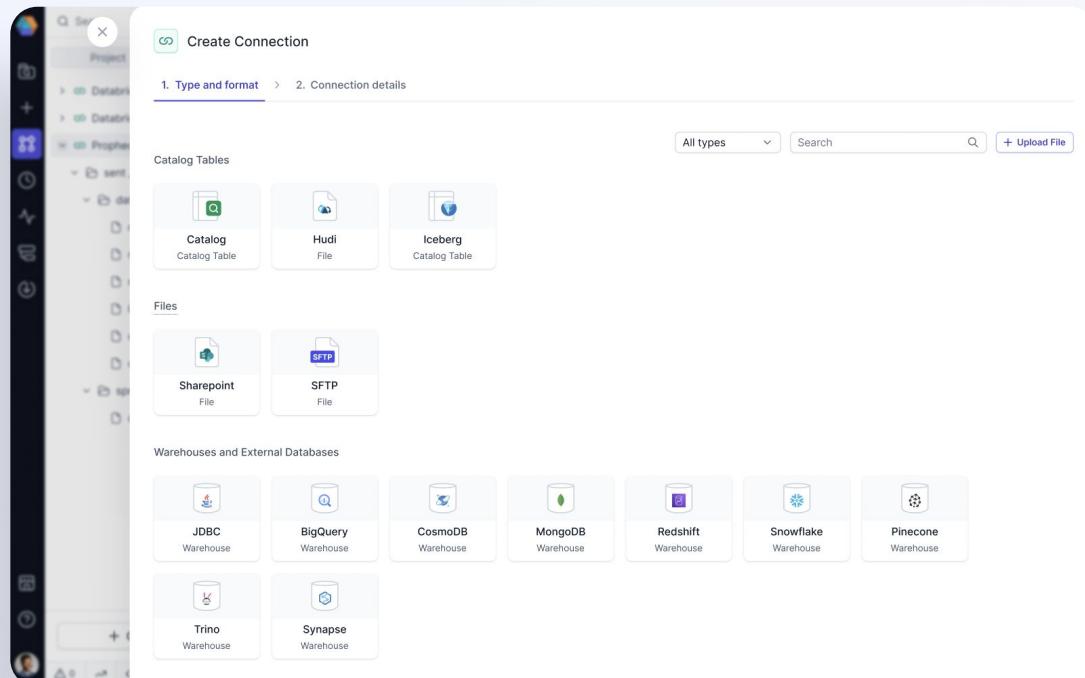
Integration with Databricks Unity Catalog

Honor the permissions you already have in place



Greater security

Reduce risk and simplify audits by ensuring analysts see only the data they're authorized to see.



The screenshot shows the Databricks UI. On the left is a sidebar with navigation links like 'Project', 'Databases', 'Catalogs', and 'Proprietary'. The main area is titled 'Create Connection' with a sub-step '1. Type and format'. It lists several connection types:

- Catalog Tables:** Catalog, Hudi, Iceberg
- Files:** Sharepoint, SFTP
- Warehouses and External Databases:** JDBC Warehouse, BigQuery Warehouse, CosmoDB Warehouse, MongoDB Warehouse, Redshift Warehouse, Snowflake Warehouse, Pinecone Warehouse, Trino Warehouse, Synapse Warehouse

Greater confidence in AI & analytics with data profiling

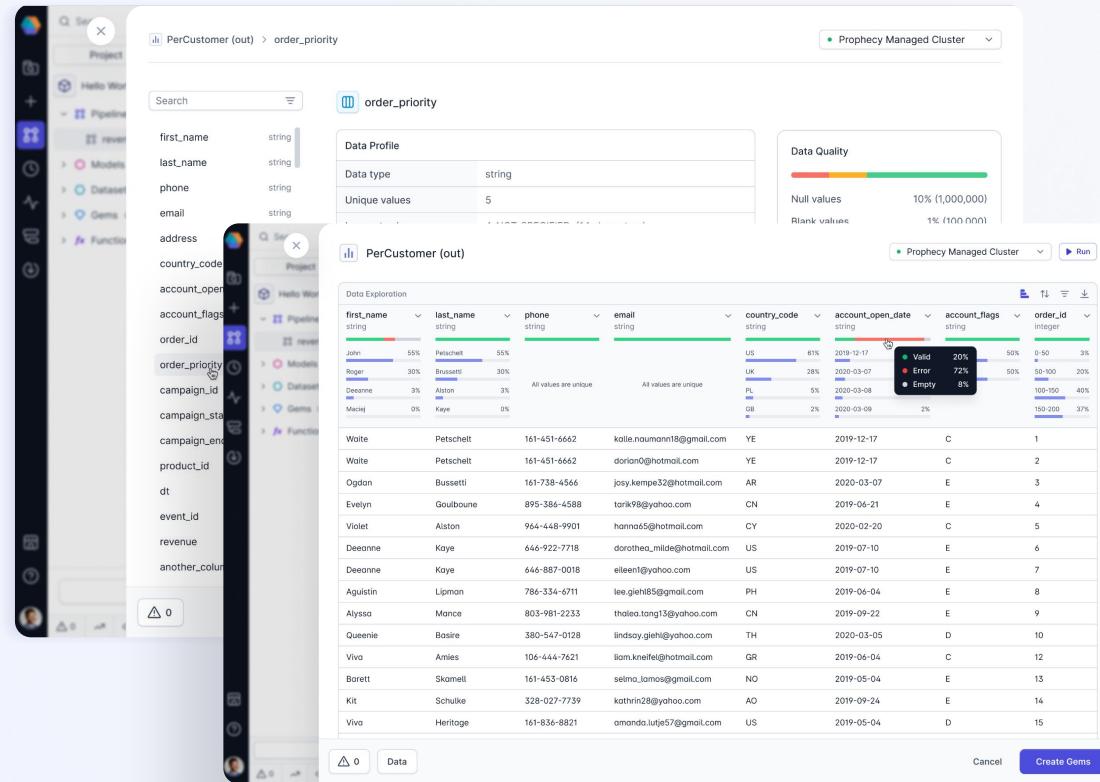
Quality data leads to quality decision-making



Higher quality

Analysts know they're using good data (or not) with an at-a-glance view of the data profile, including value distributions, data completeness.

- Instant Data Insights.** Quickly understand value distributions, completeness, and null values at a glance, ensuring your data is ready for analysis.
- Better Data Quality, Fewer Surprises.** Detect anomalies and inconsistencies early, reducing errors in downstream analytics and AI models.
- Faster, Smarter Decision-Making.** With clear visualizations of data health, teams can confidently transform and use high-quality data without second-guessing.



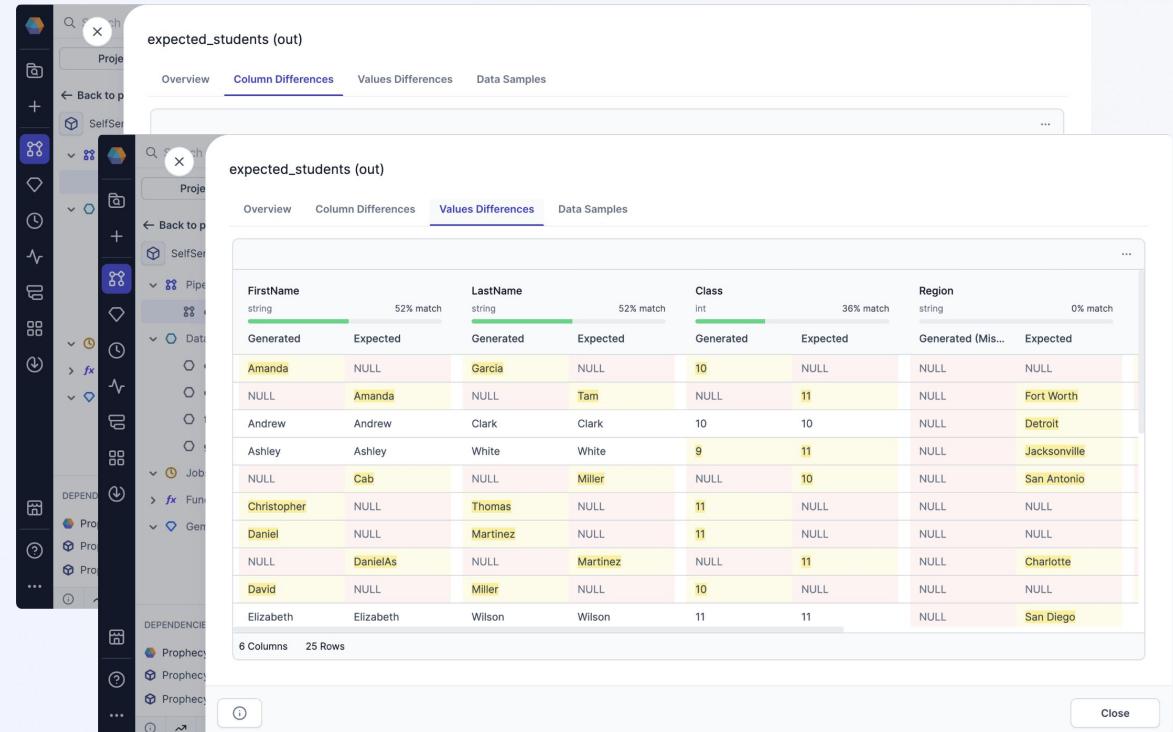
Validate Datasets using Data Diff



Higher quality

Compare generated and expected data.

- At a glance comparison.** Quickly compare differences across columns and values.
- Validate data at key steps.** Useful for promoting pipelines to production.
- Migrate pipelines with confidence.** Easily validate logic by comparing datasets created across different systems.



The screenshot shows the Apache Spark UI interface. On the left, there's a sidebar with various project and pipeline management icons. The main area displays two tabs for comparing datasets: "Column Differences" and "Values Differences". The "Values Differences" tab is currently selected, showing a comparison between "Generated" and "Expected" data for a dataset named "expected_students (out)". The table has columns for FirstName, LastName, Class, and Region. The FirstName and LastName columns show a 52% match, while Class and Region show a 36% match. The table contains 25 rows of student data, with some rows showing discrepancies between generated and expected values.

| FirstName | string | 52% match | LastName | string | 52% match | Class | int | 36% match | Region | string | 0% match |
|-------------|-----------|-----------|-----------|----------|-----------|-----------|----------|-----------|-------------------|--------------|----------|
| Generated | Expected | | Generated | Expected | | Generated | Expected | | Generated (Mis... | Expected | |
| Amanda | NULL | | Garcia | NULL | | 10 | NULL | | NULL | NULL | |
| NULL | Amber | | Tam | NULL | | 11 | NULL | | NULL | Fort Worth | |
| Andrew | Andrew | | Clark | Clark | | 10 | 10 | | NULL | Detroit | |
| Ashley | Ashley | | White | White | | 9 | 11 | | NULL | Jacksonville | |
| NULL | Cab | | Miller | NULL | | 10 | NULL | | NULL | San Antonio | |
| Christopher | NULL | | Thomas | NULL | | 11 | NULL | | NULL | NULL | |
| Daniel | NULL | | Martinez | NULL | | 11 | NULL | | NULL | NULL | |
| NULL | DanielAs | | Martinez | NULL | | 11 | NULL | | NULL | Charlotte | |
| David | NULL | | Miller | NULL | | 10 | NULL | | NULL | NULL | |
| Elizabeth | Elizabeth | | Wilson | Wilson | | 11 | 11 | | NULL | San Diego | |
| 6 Columns | 25 Rows | | | | | | | | | | |

Close