erwin Data Intelligence – erwin Data Quality

Steward Playbook

Legal Notices

Contact erwin

Understanding your Support

Review support maintenance programs and offerings.

Registering for Support

Access the erwin support site and click Sign in to register for product support.

Accessing Technical Support

For your convenience, erwin provides easy access to "One Stop" support for erwin Data Intelligence (erwin DI), and includes the following:

- Online and telephone contact information for technical assistance and customer services
- Information about user communities and forums
- Product and documentation downloads
- erwin Support policies and guidelines
- Other helpful resources appropriate for your product

For information about other erwin products, visit http://erwin.com/.

Provide Feedback

If you have comments or questions, or feedback about erwin product documentation, you can send a message to distechpubs@erwin.com.

erwin Data Modeler News and Events

Visit www.erwin.com to get up-to-date news, announcements, and events. View video demos and read up on customer success stories and articles by industry experts.

Contents

Introduction to the Playbook

Welcome to **the erwin DQ User Playbook**. This guide is crafted to help you effectively utilize the erwin DQ platform, empowering you to consistently maintain and enhance Data Quality (DQ) and Data Observability across your organization. Whether you're managing small projects or large-scale enterprise data initiatives, this playbook provides structured, clear, and actionable instructions to ensure your data meets the highest standards.

This playbook aligns explicitly with the structured end-to-end Data Quality Management Flow, providing seamless integration of all platform capabilities from discovering critical data elements to configuring quality rules, continuous monitoring, managing issues, and generating insightful reports.

Purpose of this Playbook

The main purpose of this playbook is to offer a clear, structured, and practical approach for leveraging the erwin DQ platform. By utilizing this playbook, you will:

- Understand the core functionalities and capabilities of erwin DQ.
- Learn to set up and manage Data Quality and Data Observability workflows efficiently.
- Clearly define roles and collaboration processes between Domain and Data Owners, Data Delivery Leads, Business Data Stewards, and Data Engineers.
- Receive practical guidance with relevant, real-world examples.
- Have a reliable and quick-reference document for everyday tasks.

Who Should Use This Playbook?

This playbook is specifically tailored for two essential roles involved in data governance and management

Business Data Stewards

As a Business Data Steward, your responsibilities include defining and maintaining data quality standards to align with business objectives. This playbook supports you to:

- Identify and profile Critical Data Elements (CDEs).
- Configure, manage, and monitor data quality measures.
- Understand and respond effectively to data quality alerts and issues

Data Engineers

As a Data Engineer, your role is focused on ensuring the reliability and integrity of data infrastructure and pipelines. This playbook guides you to:

- Integrate and manage diverse data sources within erwin DQ.
- Proactively monitor data observability metrics and address alerts.
- Conduct thorough Root Cause Analysis (RCA) and resolve data issues at their source.

How to Navigate this Document?

To efficiently utilize this playbook:

- **Structured Chapters**: Each chapter is logically structured, progressively building knowledge. It is recommended for new users to follow chapters sequentially.
- **Step-by-Step Instructions**: Clearly defined navigation paths and instructions help simplify complex processes.
- **Screenshot Placeholders**: Visual placeholders (e.g., [Insert Screenshot – Measure Dashboard]) are provided to facilitate clear understanding and practical visualization.
- **Real-World Examples**: Relevant and generalized examples are included to demonstrate practical application across various scenarios.
- **Terminology Explained**: Key terms and concepts are clearly defined within their context.
- **Quick Reference**: A detailed Table of Contents and clearly defined sections enable easy, targeted reference during daily operations.

Following these guidelines will ensure you can efficiently manage data quality and observability within the erwin DQ platform, irrespective of your organization's specific domain.

What is Data Quality and Data Observability?

Ensuring data accuracy and reliability is critical for any organization's success. This chapter provides an overview of two foundational concepts: Data Quality (DQ) and Data Observability. Understanding these concepts is essential for effectively using the erwin DQ platform and maintaining robust data governance practices.

What is Data Quality?

Data Quality refers to the condition of data based on factors such as accuracy, completeness, consistency, reliability, and timeliness. High-quality data meets the requirements of intended use, enabling organizations to make informed, confidential decisions.

Key Dimensions of Data Quality

1. **Accuracy**: Data accurately reflects reality.
2. **Completeness**: All necessary data is available without gaps.
3. **Consistency**: Data remains uniform and consistent across different datasets and time periods.
4. **Validity**: Data conforms to predefined standards and rules.
5. **Timeliness**: Data is available when needed to support decision-making.
6. **Uniqueness**: Each record is distinct, with no duplicates present.

Importance of Data Quality

High-quality data supports critical business processes, reduces errors and risks, and enhances decision-making. Poor data quality can lead to misguided decisions, increased costs, and operational inefficiencies.

What is Data Observability?

Data Observability is the practice of monitoring, tracking, and understanding the health of your data systems over time. It allows you to proactively detect anomalies, issues, and changes in data patterns or behaviors, ensuring continuous data reliability.

Pillars of Data Observability

1. **Freshness**: Data arrives and updates within expected timeframes.
2. **Distribution**: Data maintains expected statistical distributions and patterns.
3. **Volume**: Data quantity remains consistent and within expected thresholds.
4. **Schema**: Data structure remains stable and consistent without unexpected changes.
5. **Lineage**: Clearly understanding where data originates, how it moves, and how it transforms across different systems.

Importance of Data Observability

Data Observability allows organizations to maintain data integrity and reliability proactively. By identifying and resolving issues early, organizations can prevent data downtime and reduce potential negative impacts on business operations.

Data Quality and Observability in erwin DQ:

erwin DQ integrates Data Quality and Data Observability into one comprehensive platform, enabling users to:

- Automated monitoring and tracking health data.
- Quickly detect, alert, and respond to data anomalies and issues.
- Implement continuous data quality checks across diverse datasets.
- Maintain data reliability through systematic issue tracking and resolution.

How erwin DQ Supports Data Quality and Observability

- **Profiling and Analysis**: Automatically profiles data to identify anomalies and irregularities.
- **Quality Rules**: Provides out-of-the-box and custom data quality rules to consistently monitor data.
- **Alerts and Notifications**: Proactively notifies users of anomalies and data quality issues through configurable alerts.
- **Issue Management**: Centralizes issue tracking, prioritization, and resolution processes.
- **Dashboards and Reporting**: Offers real-time insights into data quality metrics and observability status.

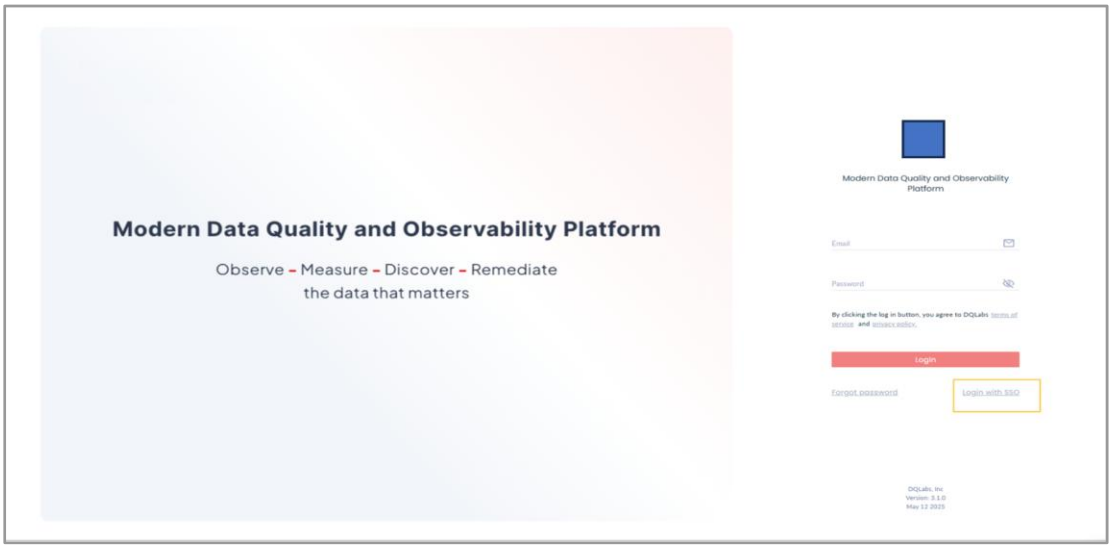
Understanding Data Quality and Data Observability ensures that you can fully leverage erwin DQ capabilities, enhancing your organization's ability to maintain high data standards and reliability.

Getting Started with erwin DQ Platform

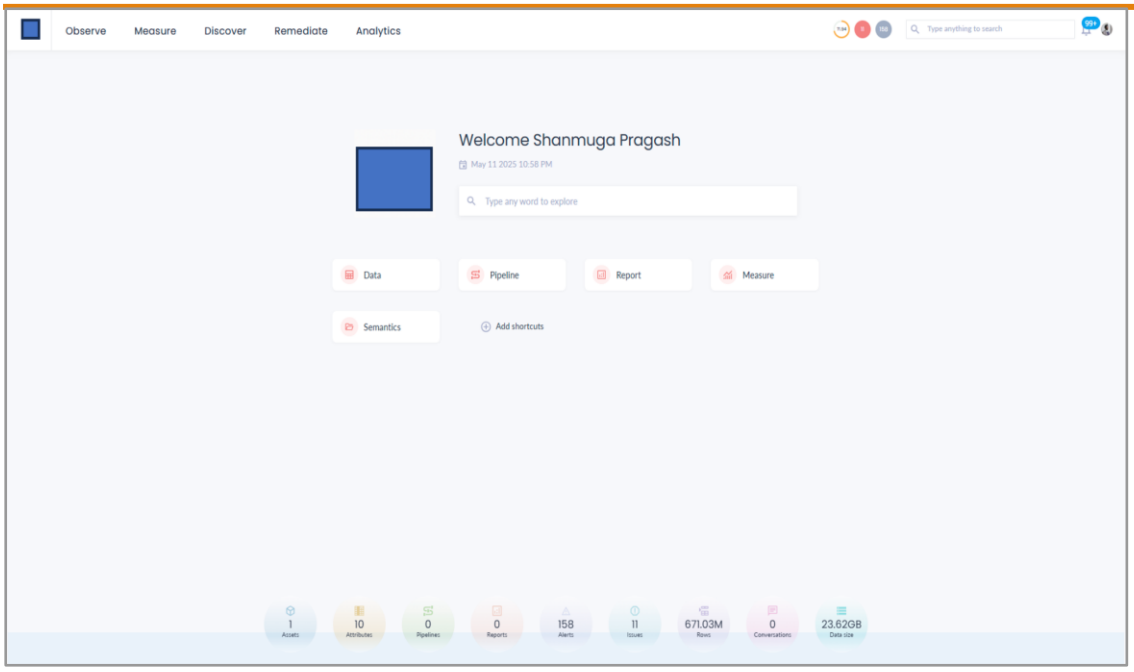
This chapter introduces new users to the erwin DQ platform by walking through how to log in, configure user preferences, and understand the overall interface.

Accessing the Platform

- How do I log in, and where do I set my user preferences?
- Step-by-Step Login Instructions:
- Access the Platform
    - Open your web browser
    - Navigate to the provided erwin DQ environment URL
- Enter Your Credentials:
    - Enter your email and password, or if you are having SSO-based authentication, then select the “Login with SSO” option.

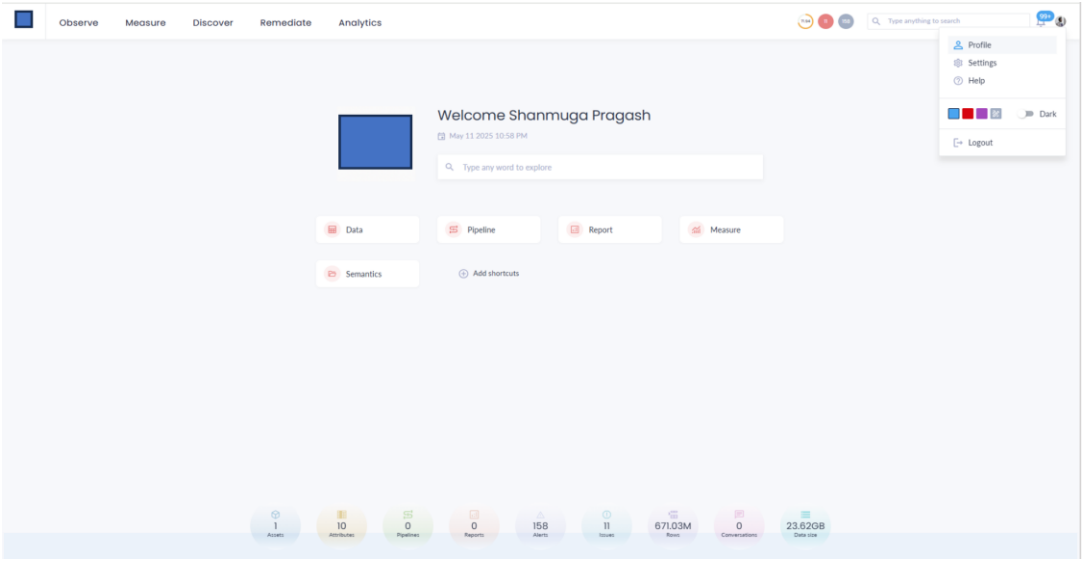


- Click "**Login**"
    - This will take you to the erwin DQ Home Dashboard.

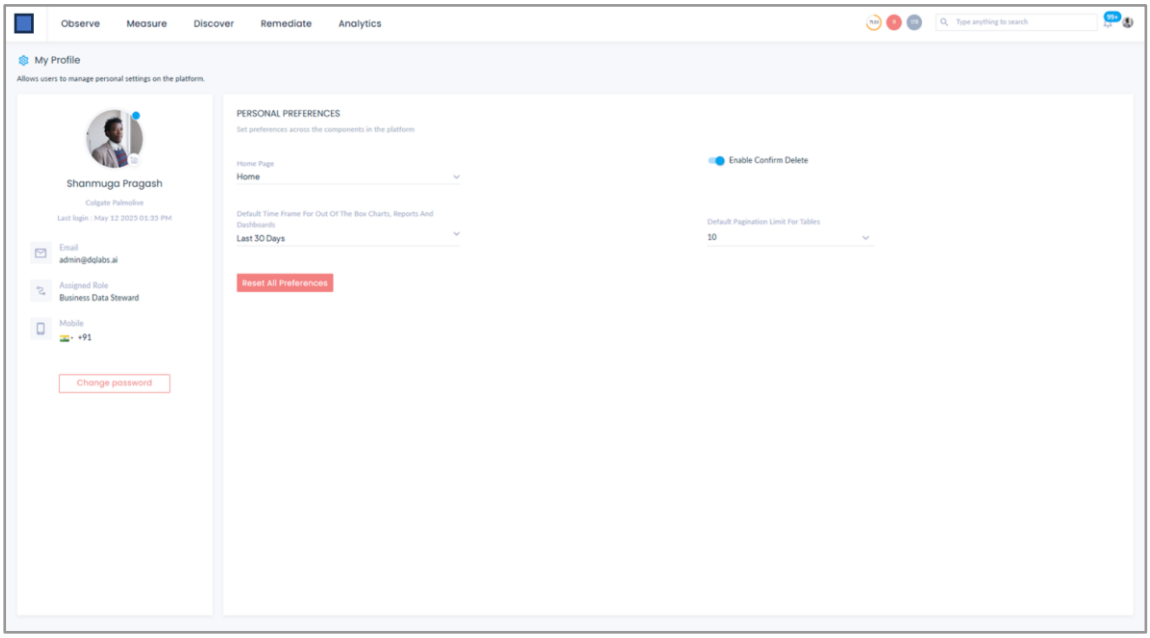


- Setting Personal Preferences:
  - After logging in, click the **Profile** icon in the top-right corner.

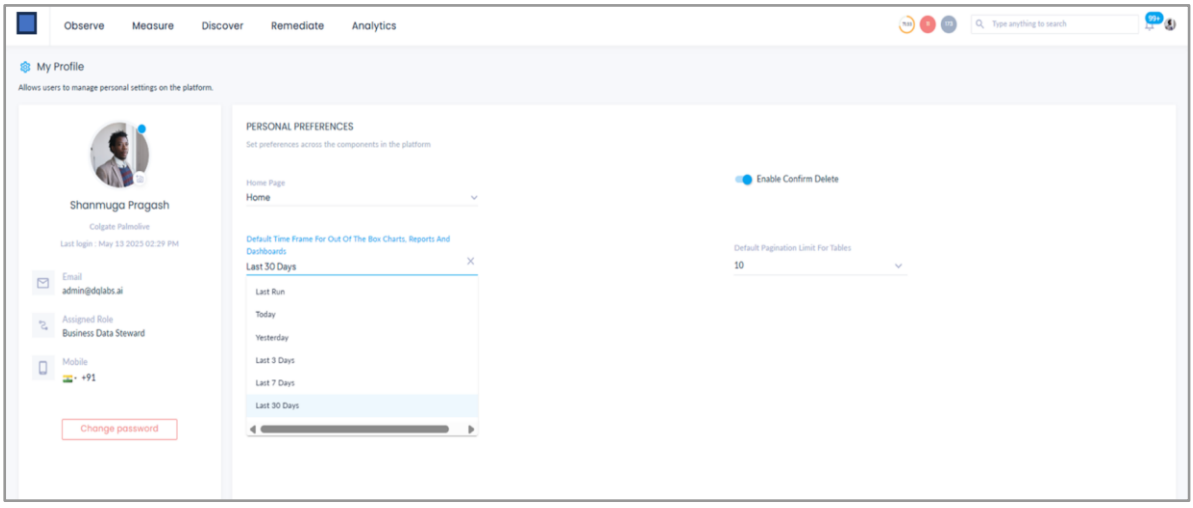
  - Select "**PERSONAL PREFERENCES**" from the dropdown menu.



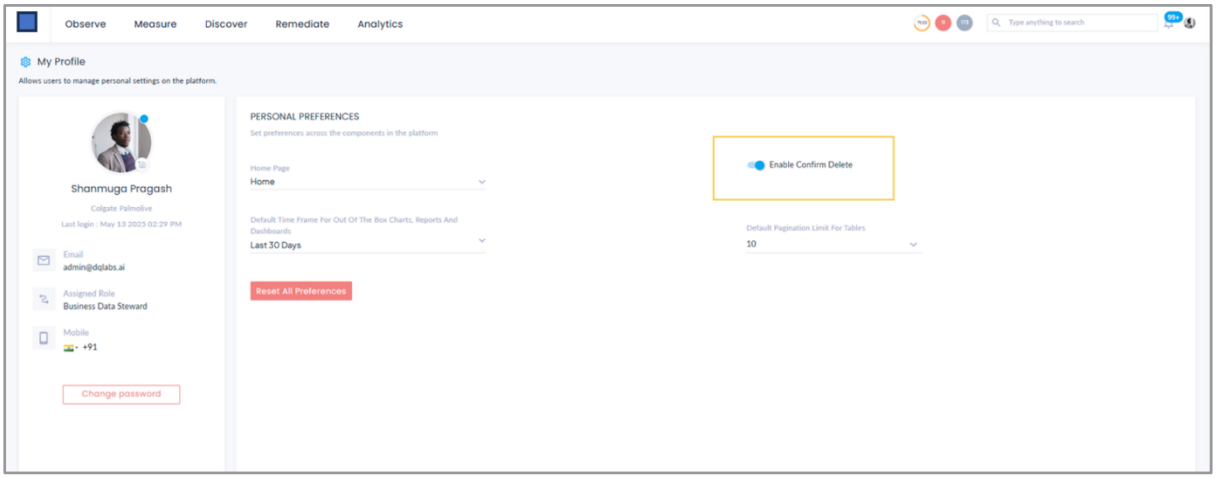
- Configure the following settings:
  - **Home Page**: Select the preferred module to load when logging in (e.g., Data, Pipeline, Report, Usage, Measure).



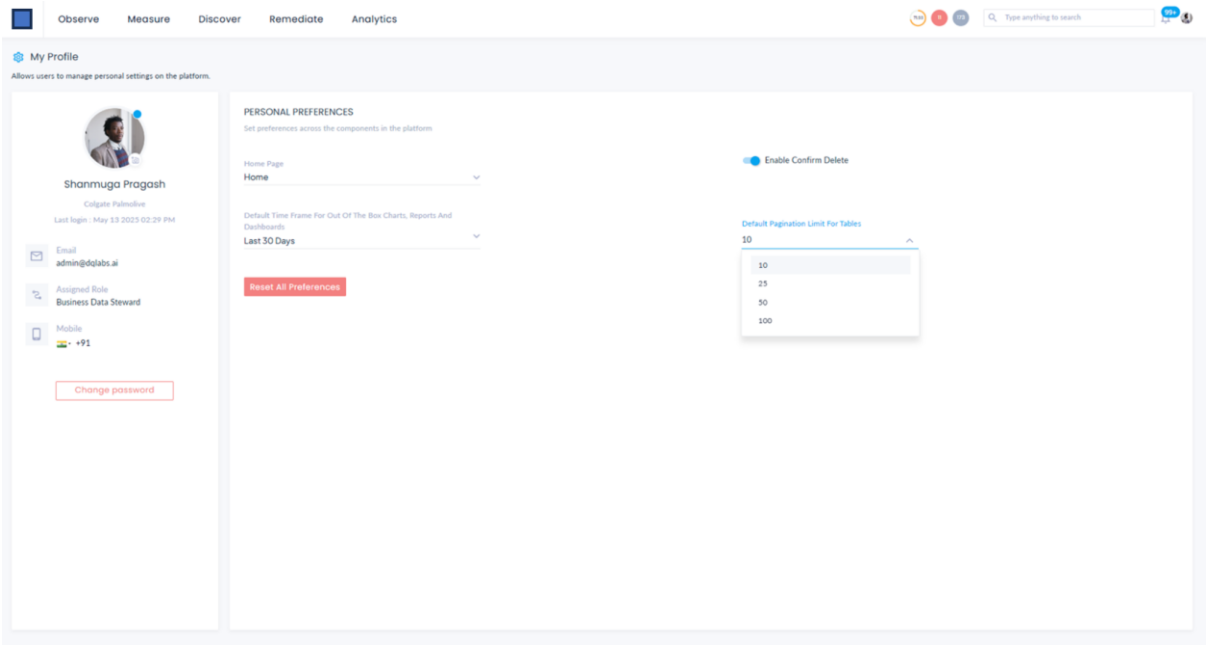
- **Default Time Frame**: Choose the default time filter for dashboards, charts, and reports:
  - Last Run
  - Today
  - Yesterday
  - Last 3 Days
  - Last 7 Days
  - Last 30 Days



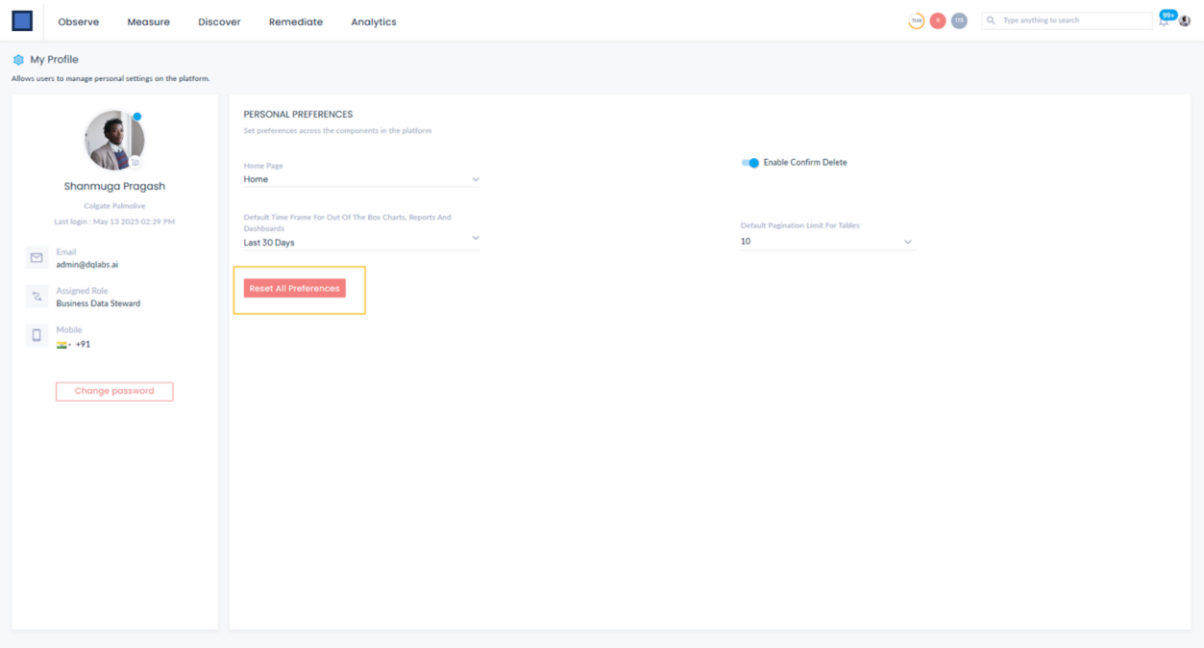
- **Enable Confirm Delete**: Toggle this option ON to receive a confirmation prompt before permanently deleting any object.



- **Default Pagination Limit for Tables**: Set the maximum number of records to be displayed per page on list views (e.g., asset list pages).



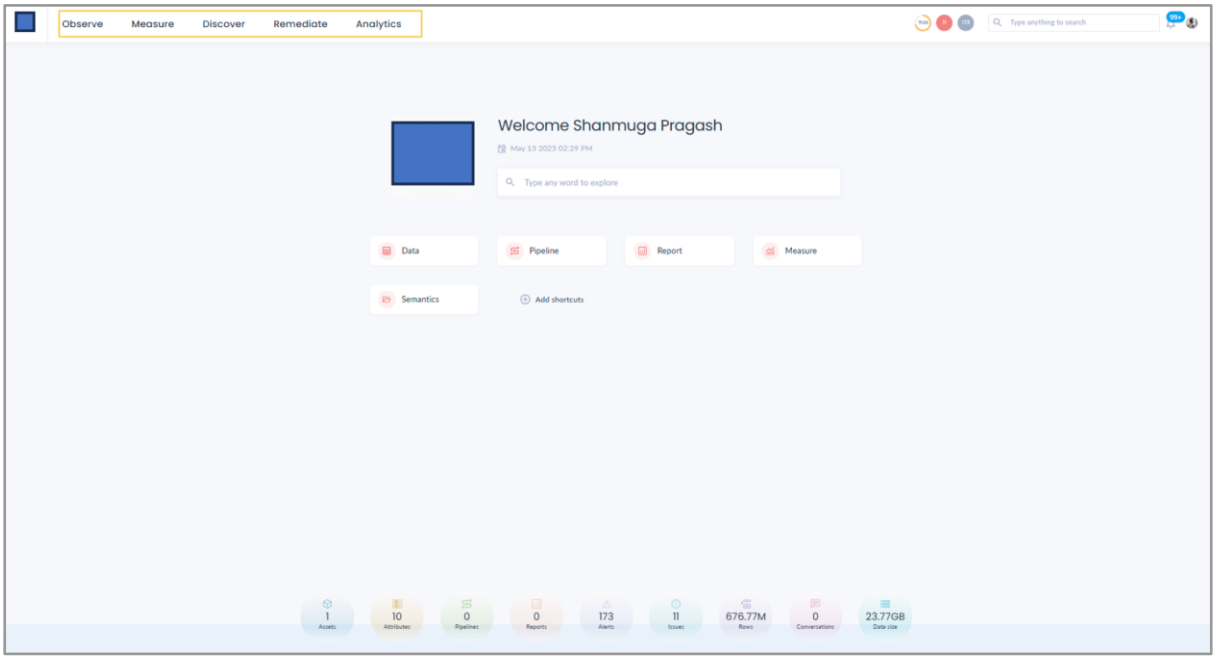
- Click "**Reset All Preferences**" to revert all changes to the default if needed.



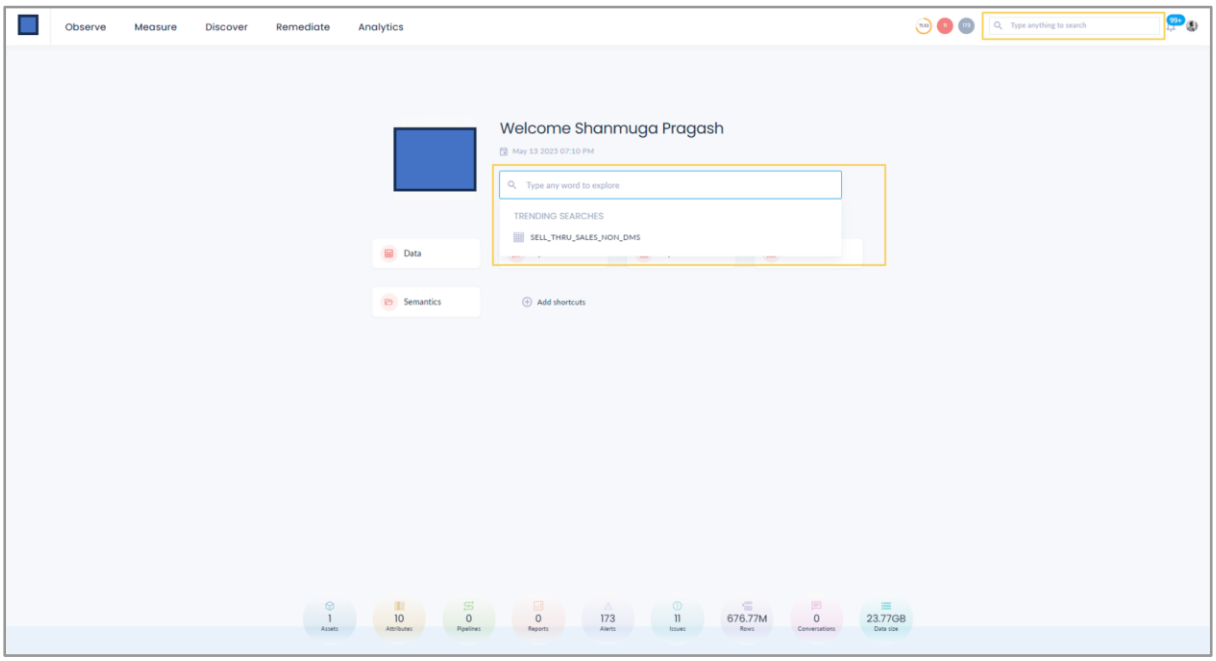
Understanding the Platform Interface

The erwin DQ user interface is designed for intuitive use, with key modules accessible from the top header menu bar and account settings grouped in the upper-right corner.

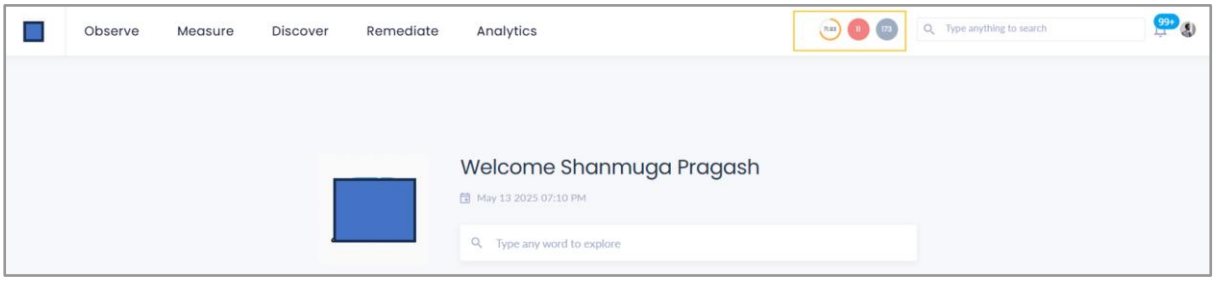
- Overview of Key Interface Components:
- Top Header Menu Bar
- The primary navigation menu runs across the top of the interface and includes quick access to core platform modules:
    - **Observe** – Monitor data health metrics like freshness, volume, and schema changes.
    - **Measure** – Configure and manage data quality rules (measures).
    - **Discover** – Search, browse, and manage data assets and critical data elements (CDEs).
    - **Remediate** – Review alerts, create, track, and resolve data issues.
    - **Analytics** – Access curated dashboards and insights.



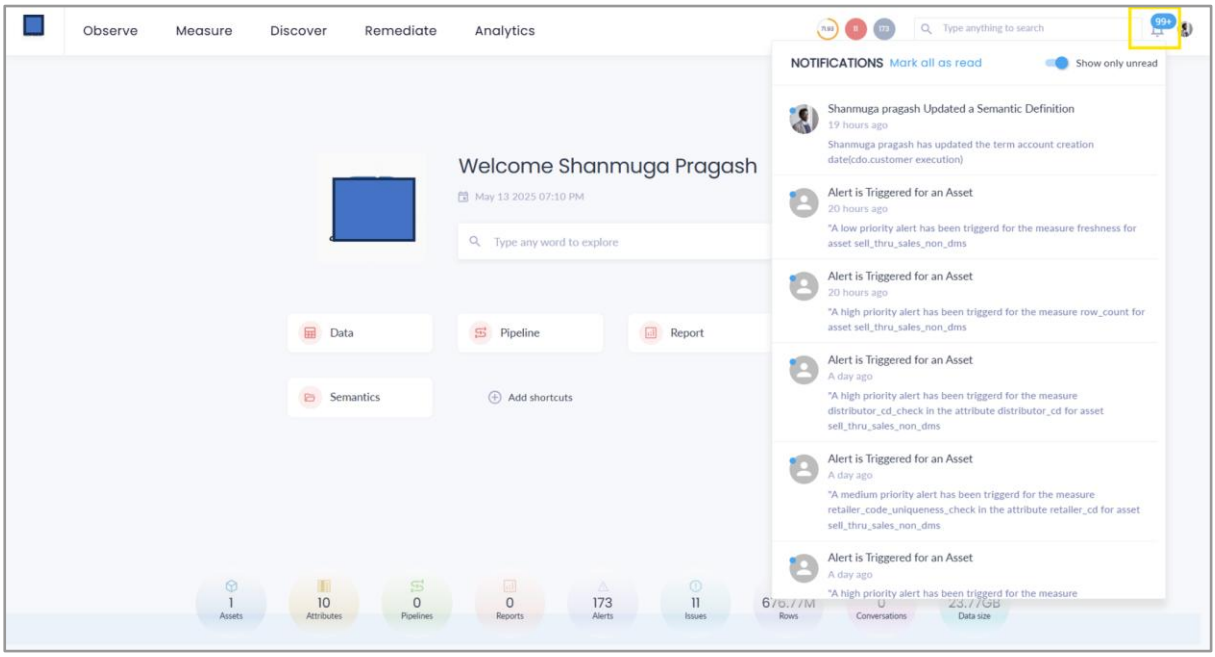
- **Header Bar Components**
- **Search Bar** – The Search bar is available in two places on the platform. Located prominently in the top right side of the landing page and in the middle of the landing page, allowing users to quickly search for assets or attributes.



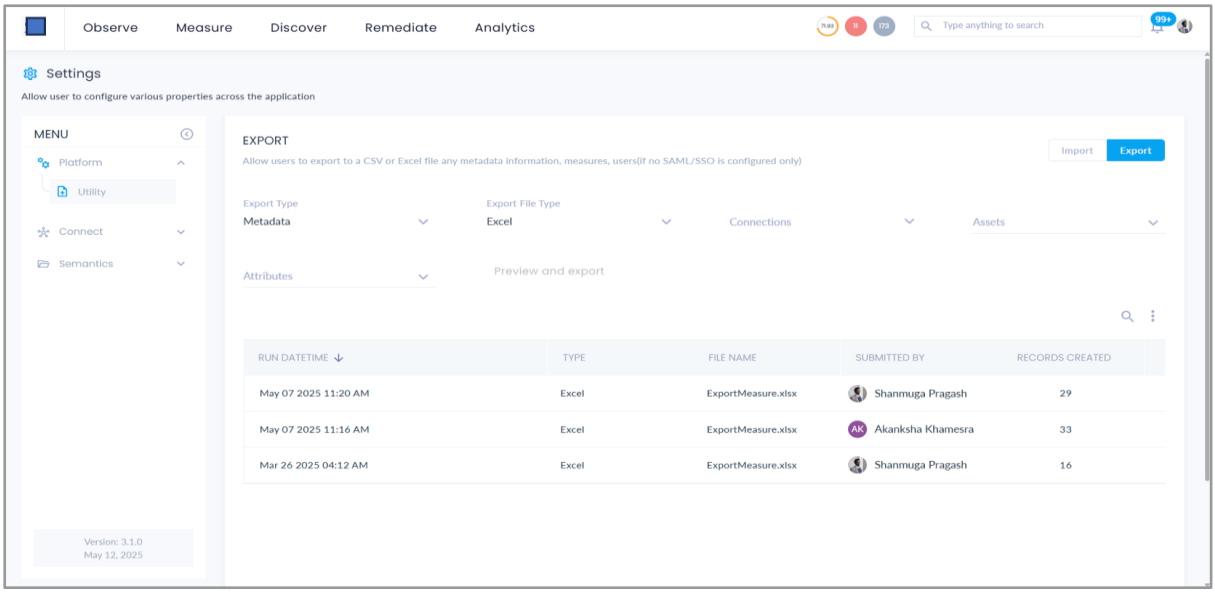
- **Overall Insights** – To view the overall organization-level DQ Score, Total DQ Issues, and Total DQ Alerts in the platform.



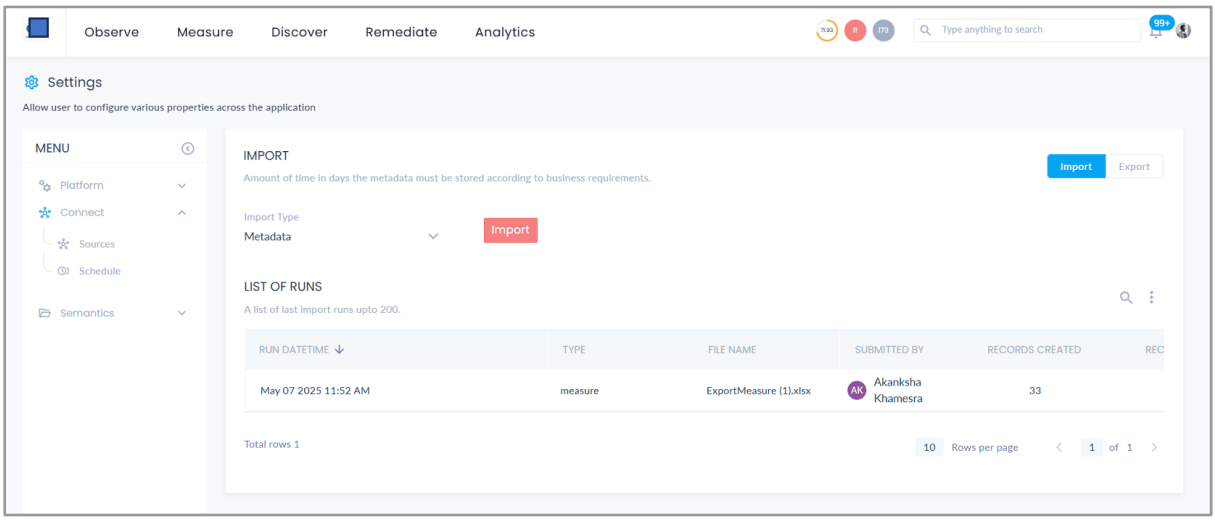
- **Notifications Icon** – Bell icon in the header to view platform alerts and data quality notifications.



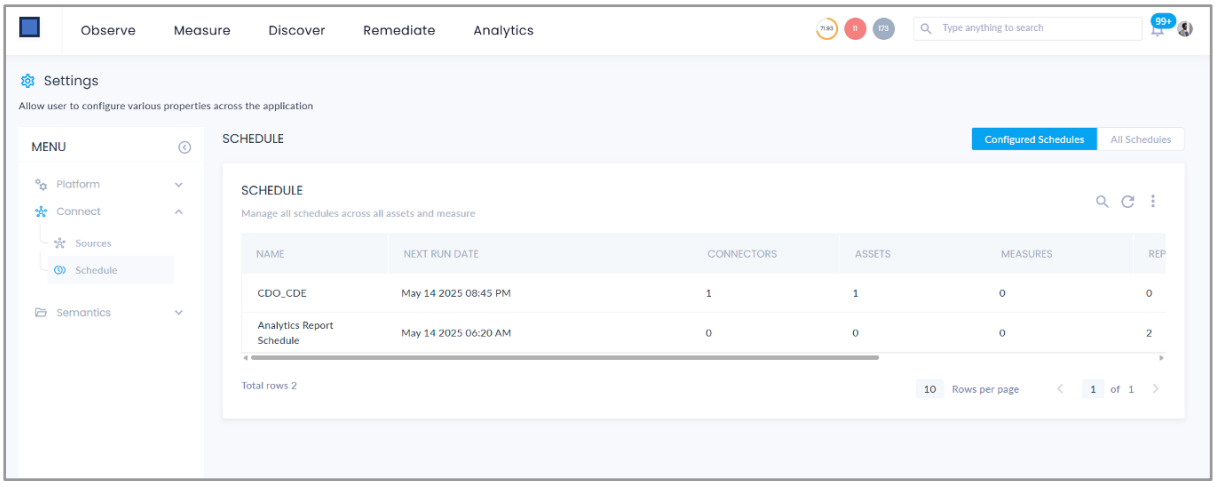
- **User Profile Dropdown** – Found at the top-right corner of the screen, with the following options:
    - **Profile** – To view and set the user preferences.
    - **Settings** – To view the platform-level settings page.
        - **Utility** – Features used to bulk import metadata, measures or users and export metadata, measures, or users' data from the platform.
        - **Export** – Allow users to export to a CSV or Excel file any metadata information, measures, users.



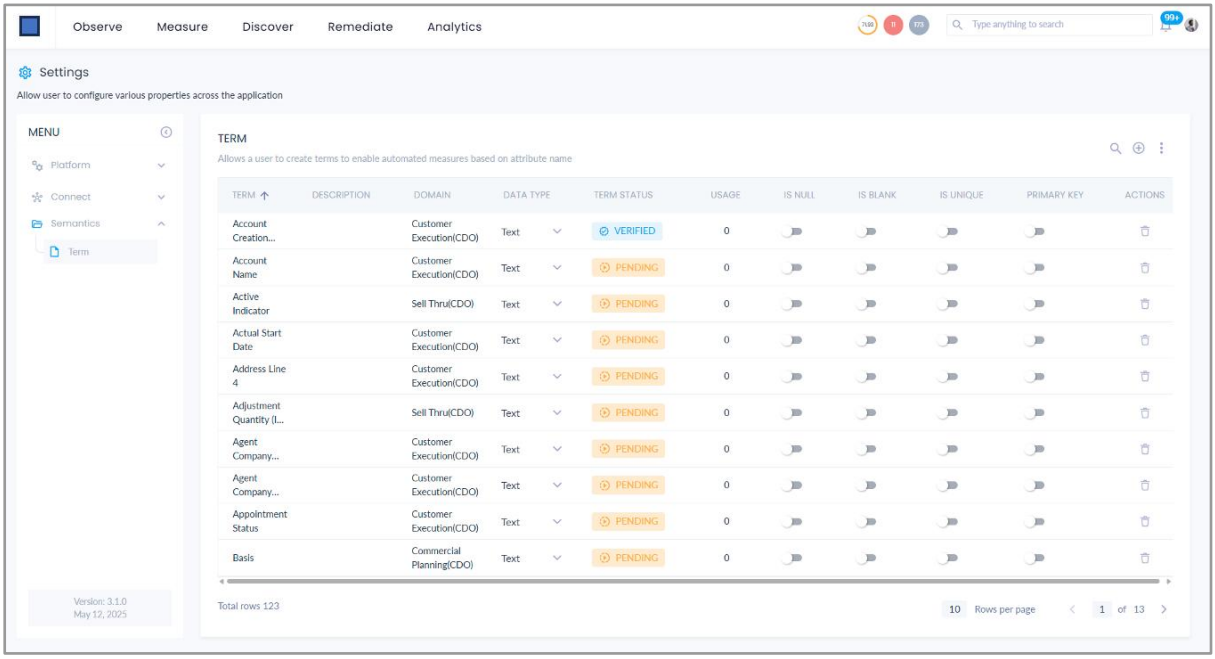
- **Import** – Allow users to bulk import a CSV or Excel file with any metadata information, measures, and users.



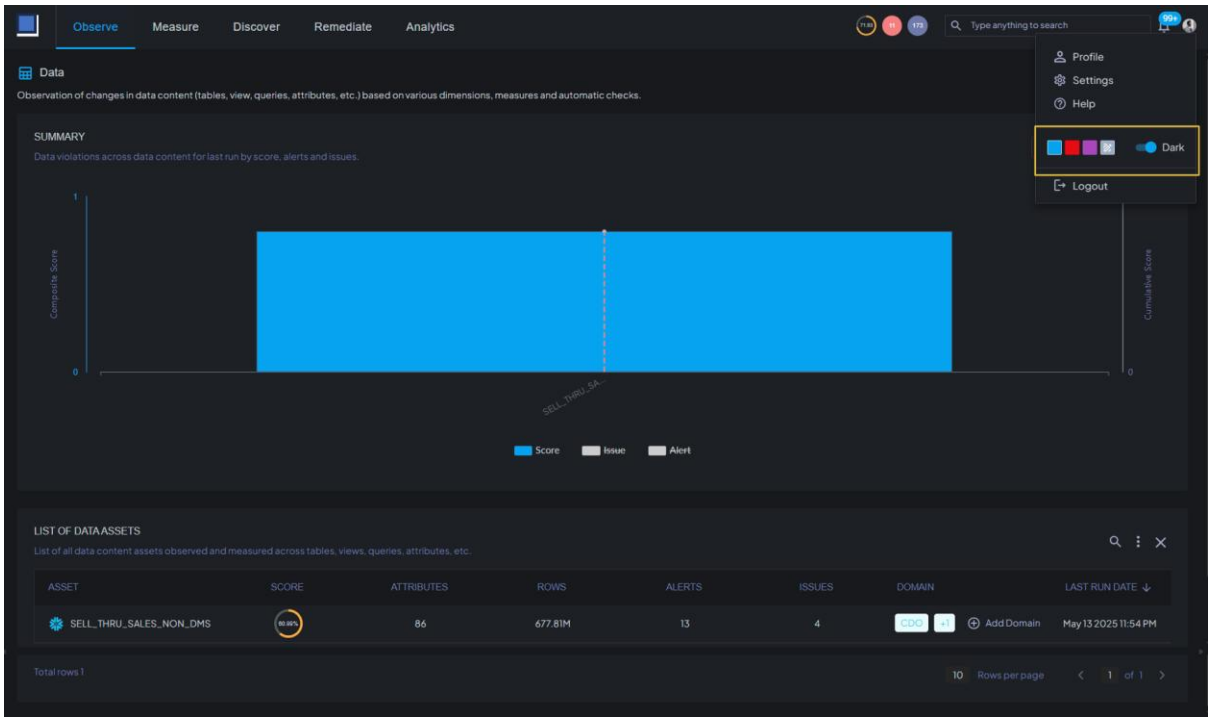
- **Connect > Source** – To view the configured data sources, connections assets.
- **Schedule** – To view all the schedules across all assets and measures.



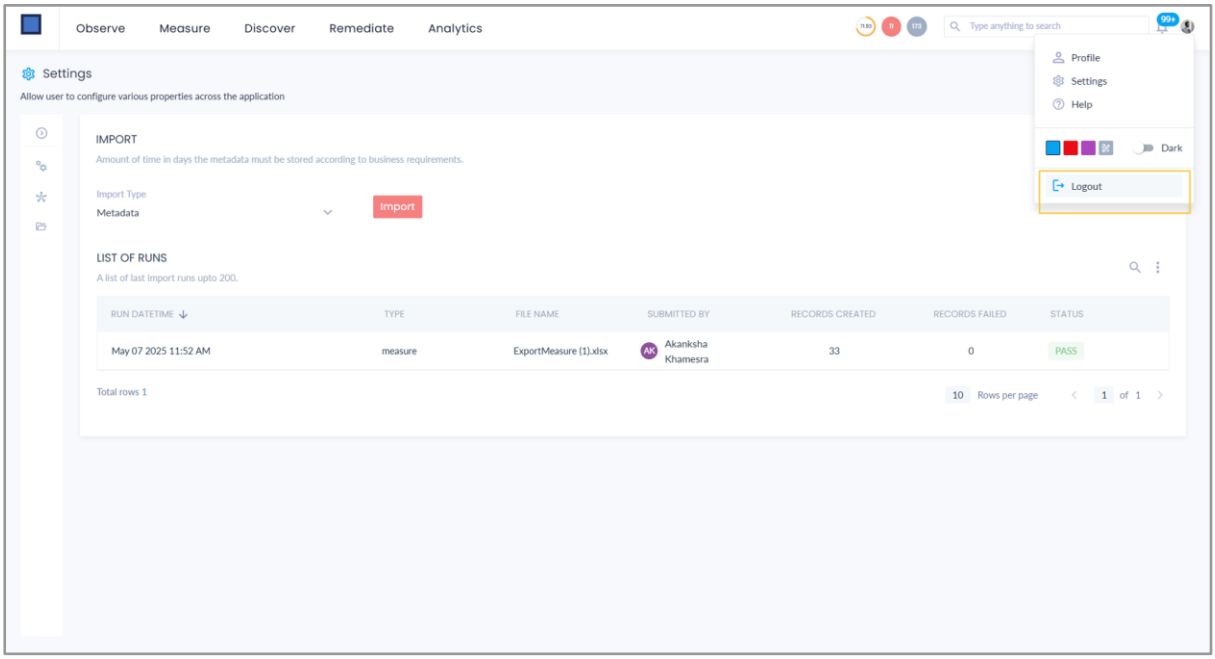
- **Semantics** – Allows a user to create and review business terms to enable automated measures based on attribute name.



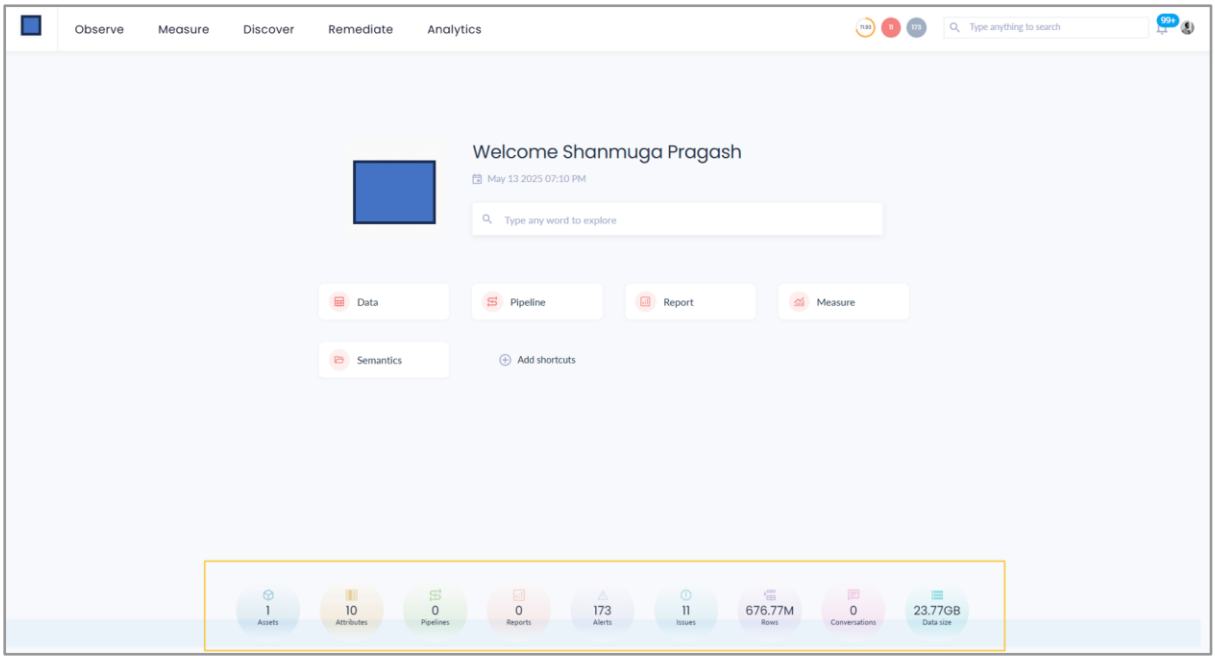
- **Dark Mode Toggle** –Users can toggle on the Dark Mode Toggle button and change the platform mode to dark mode.



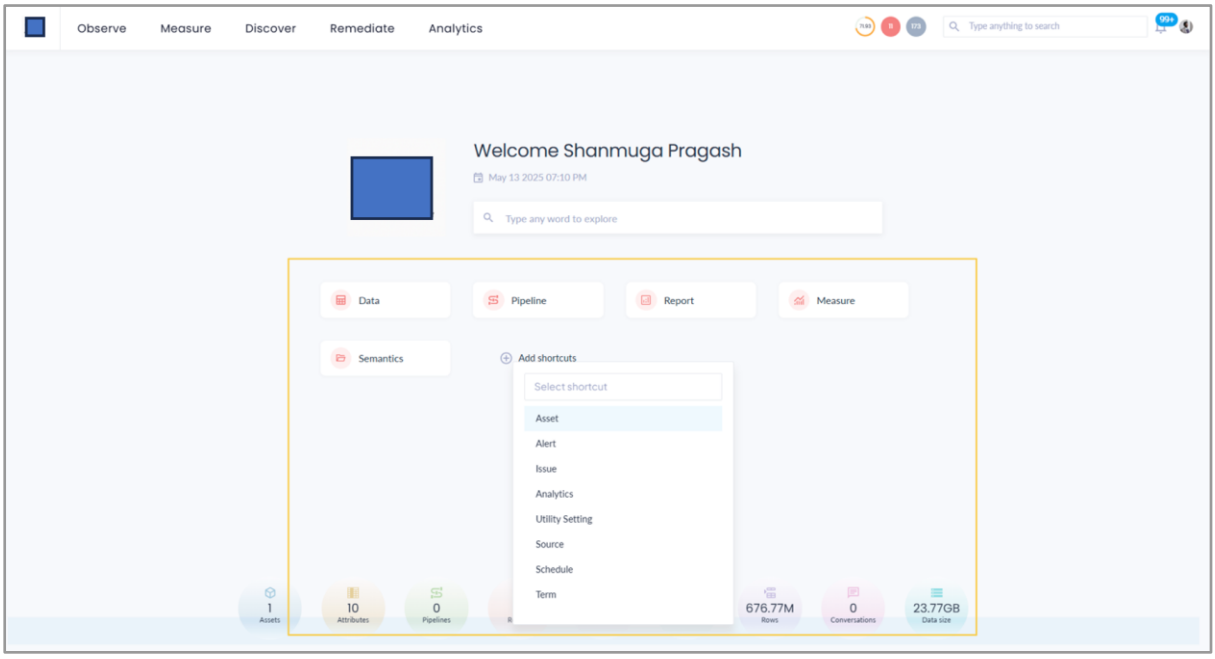
- **Logout** – Click Logout to exit the platform.



- **Dashboard Widgets**
  - After the user logs in, the default landing page will provide options to view the overall metrics configured in the platform, such as the:



- Asset count
- Attribute count
- Pipeline count
- Reports count
- Alerts –Number of Data Quality Alerts
- Issues –Number of DQ Issues
- Rows - Volume of records processed
- Shortcuts to different sections of the platform can be set up on the homepage to reduce the number of clicks to multiple pages.



With this foundational understanding of how to access and navigate erwin DQ, you are now ready to begin using the core modules, starting with data discovery and profiling in the next chapter.

Discovering Critical Data Elements (CDEs)

The first step in managing data quality is identifying the most important data elements that drive key business processes. These are known as **Critical Data Elements (CDEs)**. This chapter will walk you through how to search, explore, and find CDEs within the erwin DQ platform.

What are Critical Data Elements (CDEs)?

Critical Data Elements (CDEs) are high-value fields within datasets that have a direct impact on business processes, reporting, compliance, or analytics. Ensuring the quality of these specific elements is a key priority in any data quality initiative.

Examples of CDEs might include:

- Customer ID
- Invoice Number
- Product SKU
- Transaction Date

Once tagged within erwin DQ, CDEs receive focused data quality monitoring and appear prominently in dashboards, alerts, and reporting workflows.

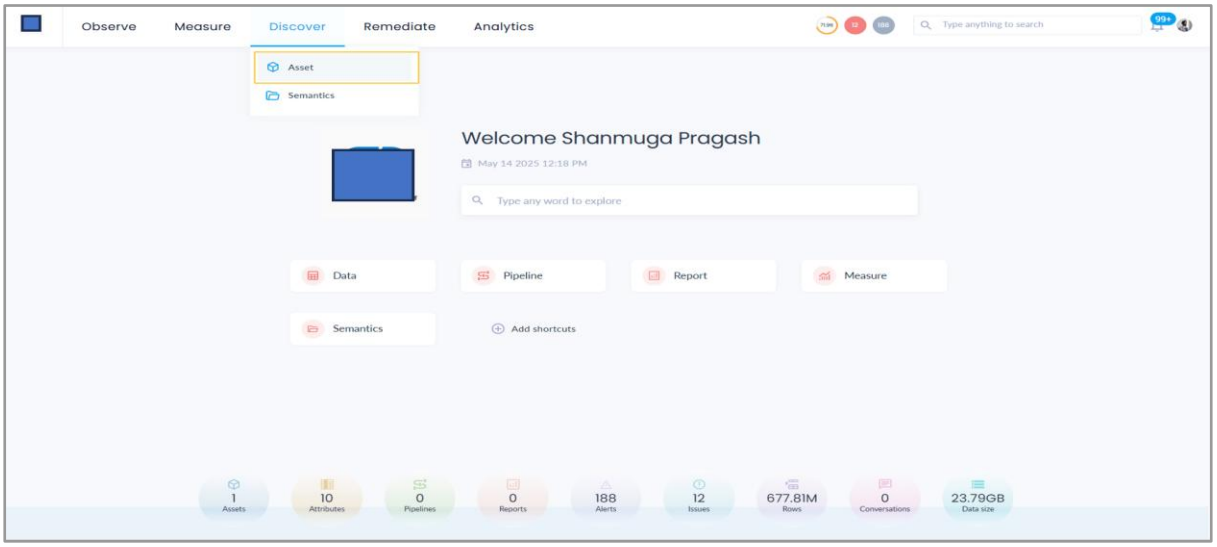
Finding and Managing Your Data Assets

Where and how do I search for my CDE or assigned data asset

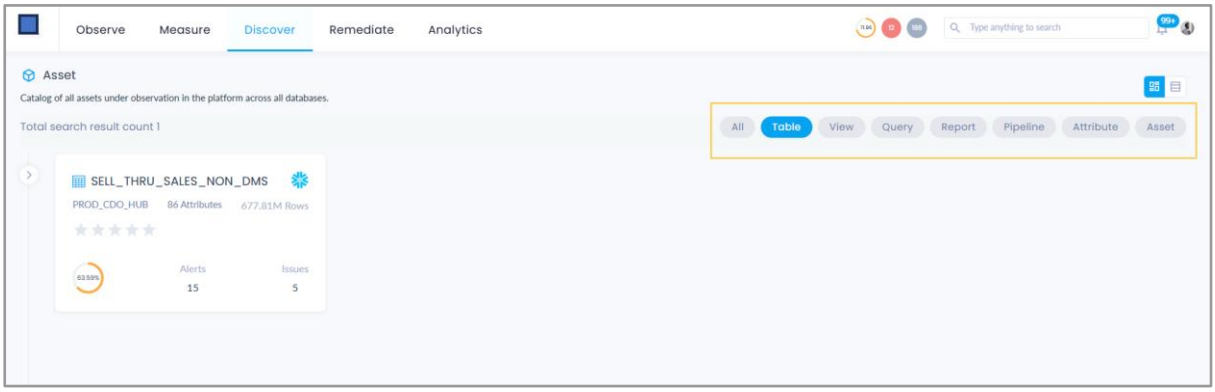
There are 3 ways to discover your assets and find the CDEs in the erwin DQ Platform.

User Journey CDE Discovery 1

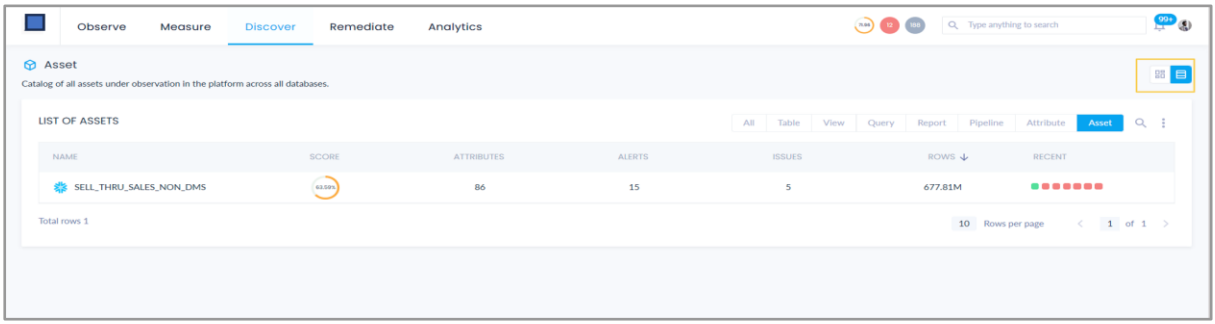
- Navigate to **Discover** from the top header menu and select the Asset from the menu.



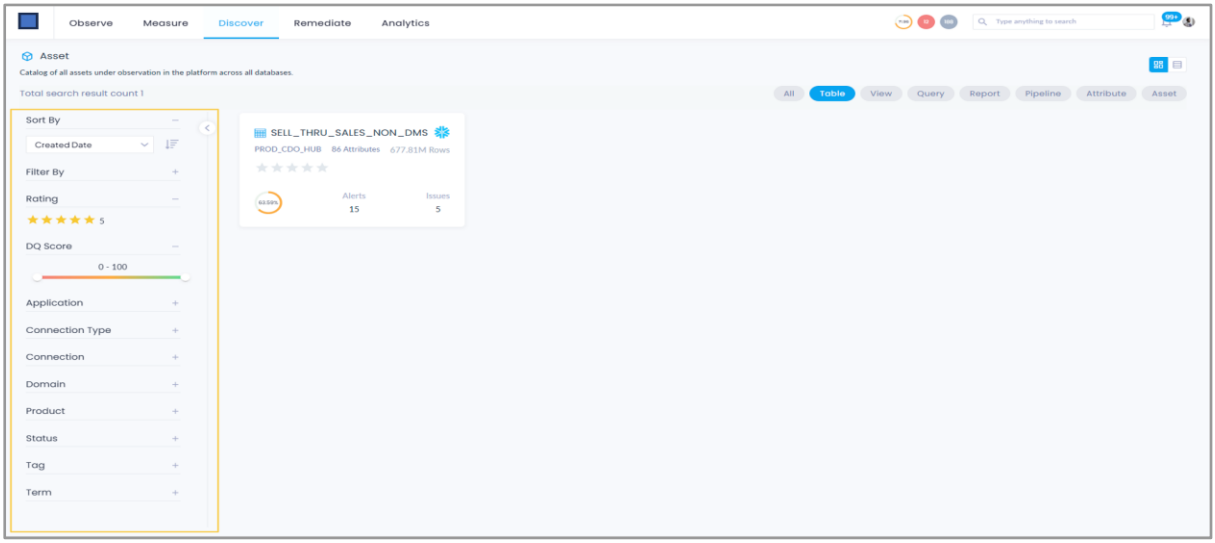
- Use the right-side top filters and apply the filters by clicking each one of them to view the assets category by Tables, Views, Query, Reports, Pipelines, Attributes, Assets, or All.



- Toggle between **table view** and **card view**.



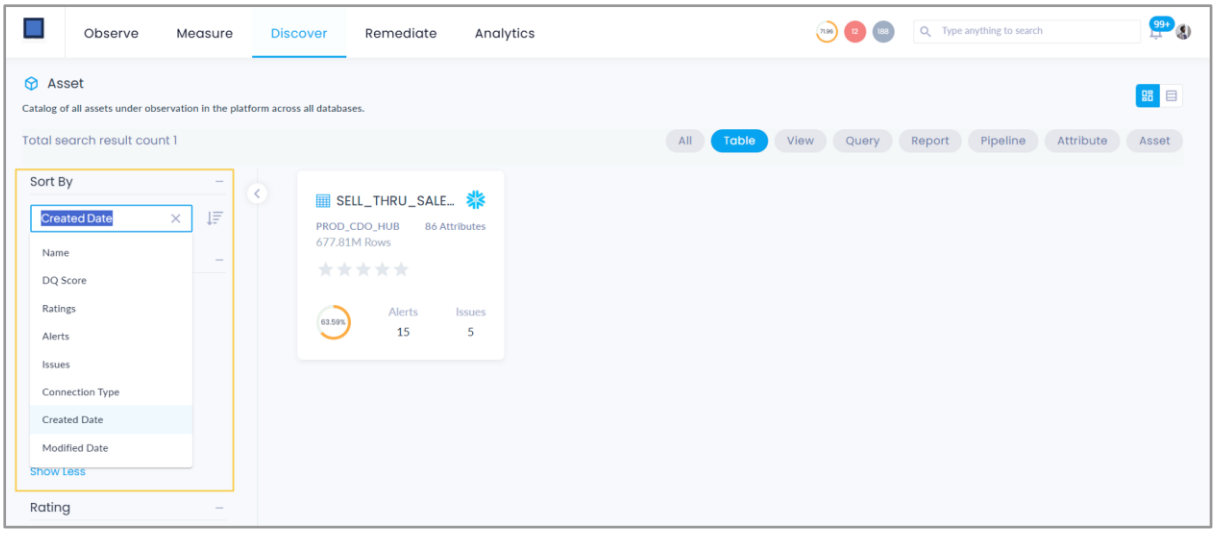
Navigate to the left side of the screen and click on the Search Filter Arrow and expand the Filters. You can sort the assets by applying various filters,



Sort By:

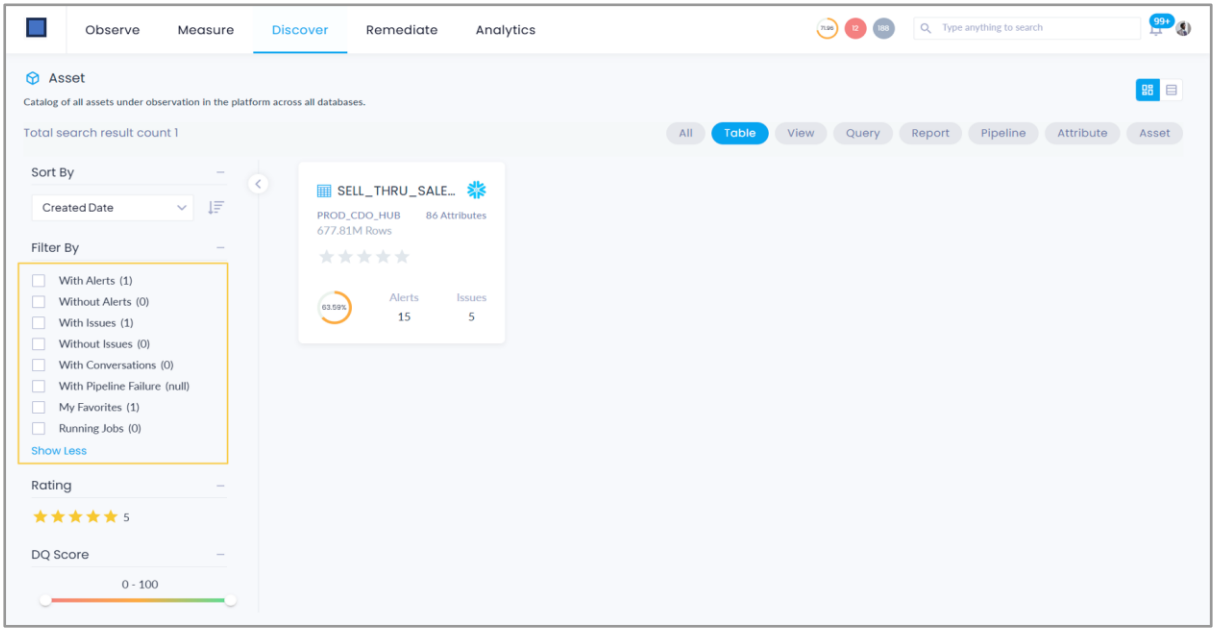
- Created Date
- Name
- DQ Score
- Ratings
- Alerts
- Issues
- Connection Type

- Created Date
- Modified Date



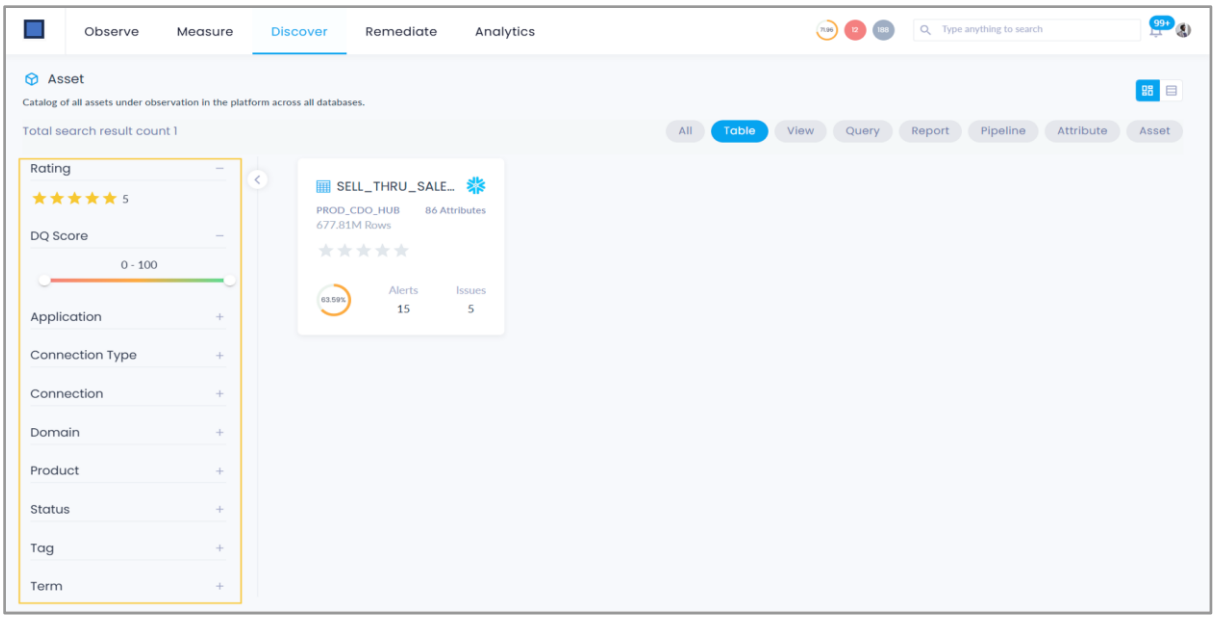
Filter By:

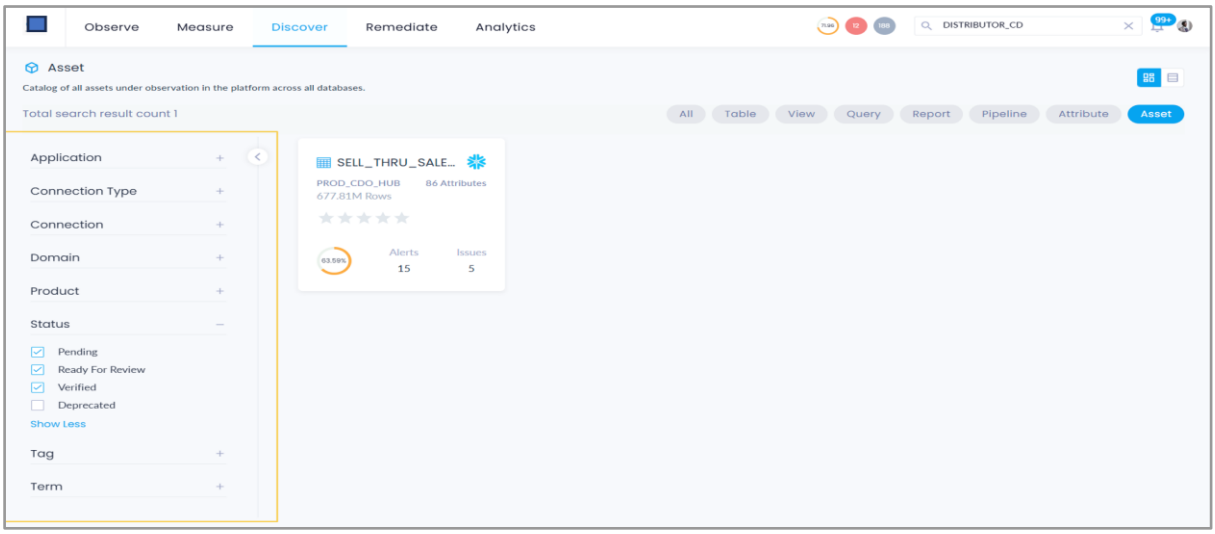
- With Alerts
- Without Alerts
- With Issues
- Without Issues
- With Conversations
- With Pipeline Failure
- My Favourites
- Running Jobs



- Ratings – Filter by Ratings (1 to 5)
- DQ Score – Filter by DQ Score between (0 to 100)
- By Application –Filter by Application Names (ERP, DMS, CRM, etc)

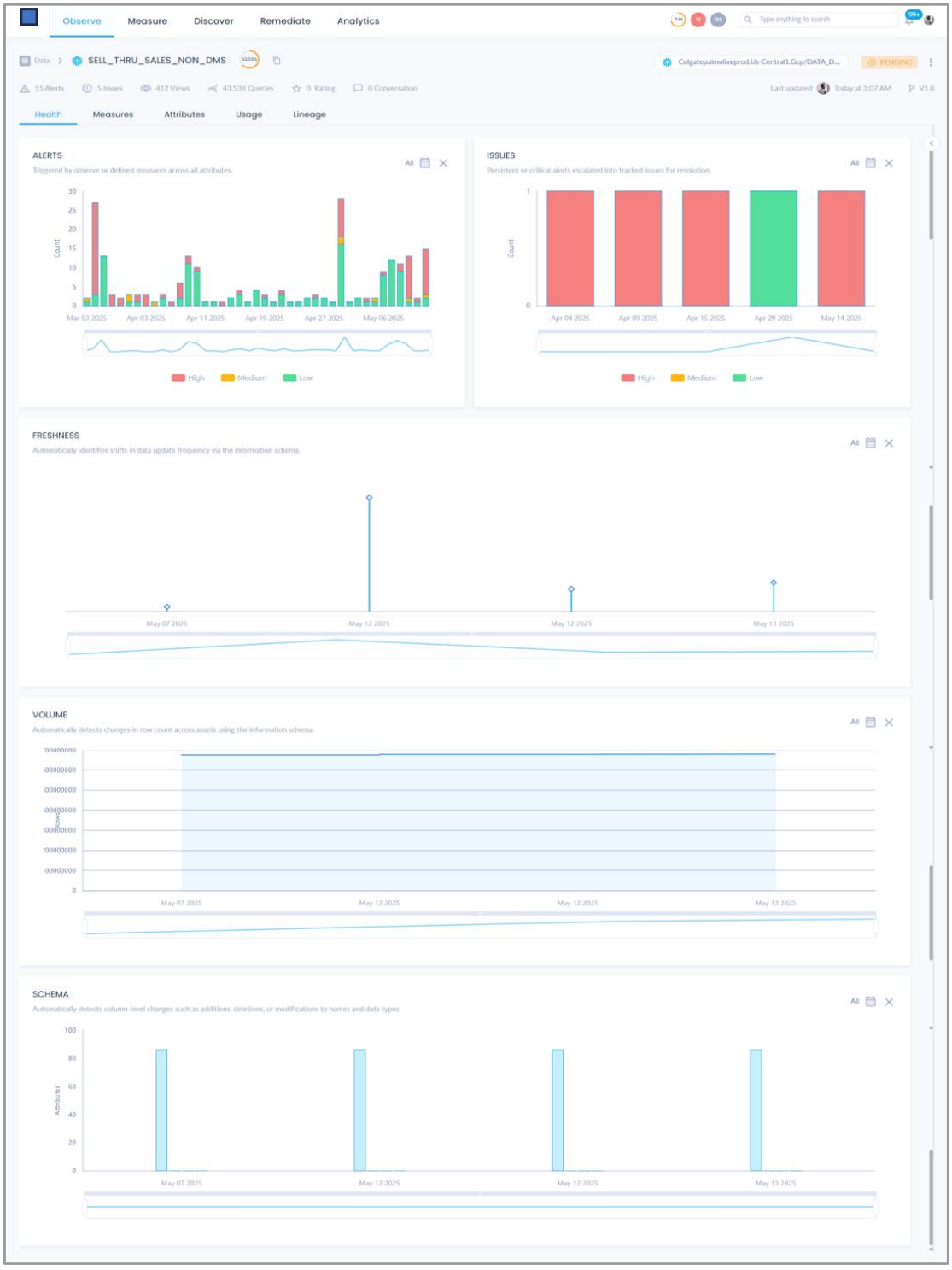
- By Connection Type – Filter by Data Source Type (Snowflake, MSSQL, Oracle, etc.)
- By Connection - Filter by data source Connection names
- By Domain – Filter by Business Domains
- By Product – Filter by Data Products
- By Status – Filter by Asset Status like Pending, Ready for Review, Verified, and Deprecated
- By Tags – Filter by Tags
- By Terms – Filter by Business Term Name



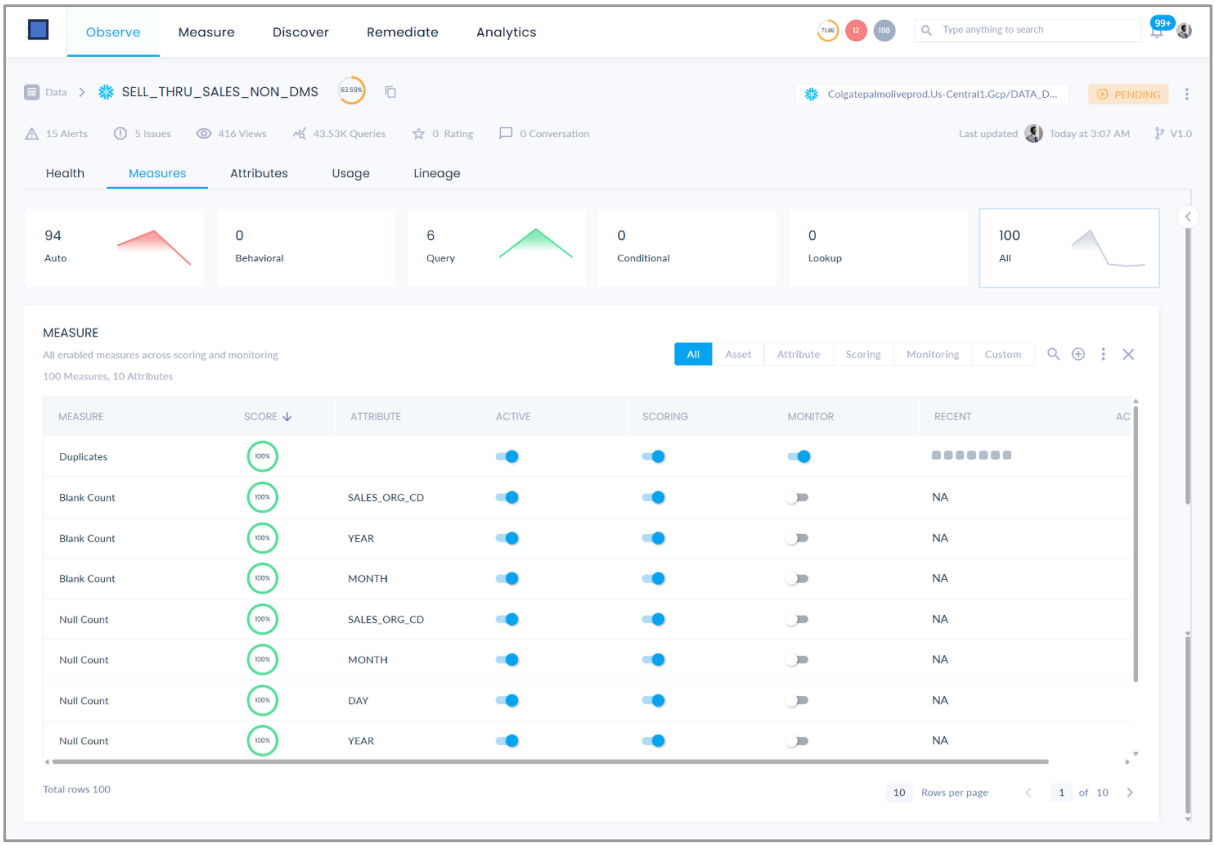


- Click on the asset name (Table, View, or Attribute) to open the **Asset Detailed** page to view the **Asset (table)** and **Attributes (CDE's Columns)**. This detailed Asset page gives you 5 tabs with detailed insights about the **Asset (table)**.
- **Health Tab** – To view the overall health of the table, which provides you the insights about:
  - **Alerts** -Triggered by observe or defined measures across all attributes.
  - **Issues** -Persistent or critical alerts escalated into tracked issues for resolution.

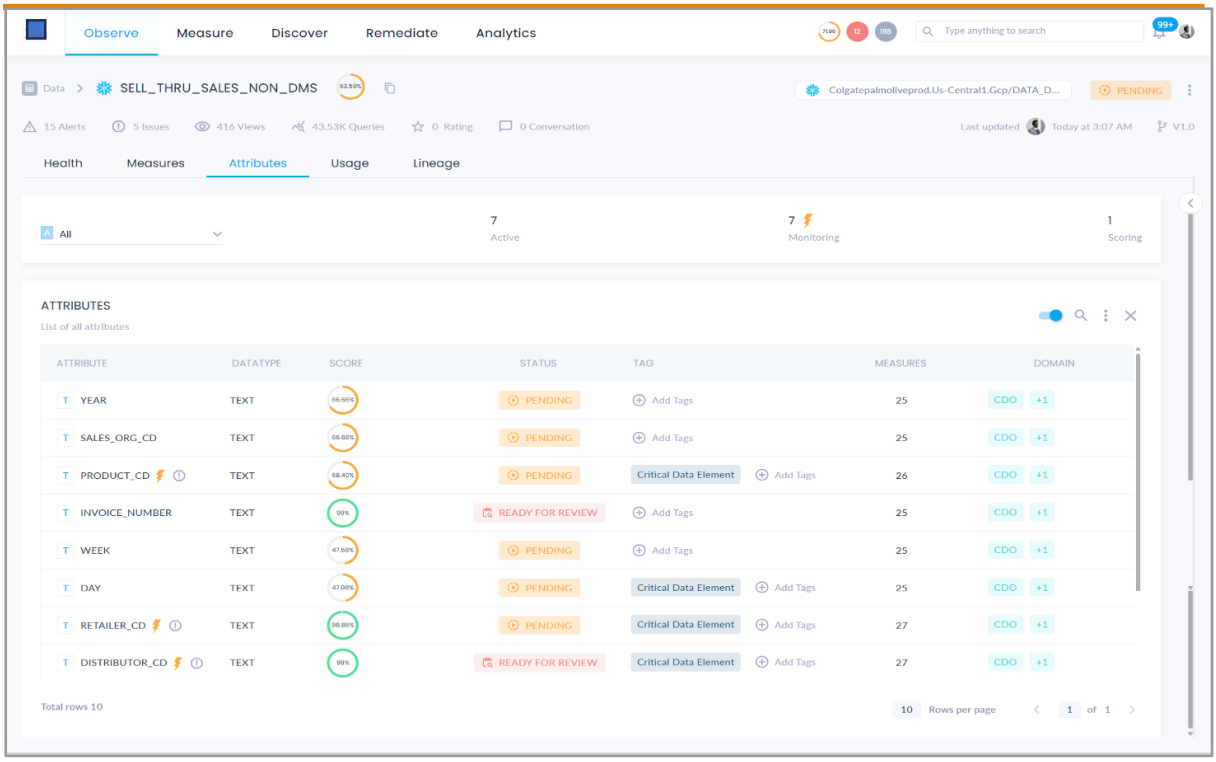
- **Freshness** -Automatically identifies shifts in data update frequency via the information schema.
- **Volume** -Automatically detects changes in row count across assets using the information schema.
- **Schema** - Automatically detects column-level changes such as additions, deletions, or modifications to names and data types.



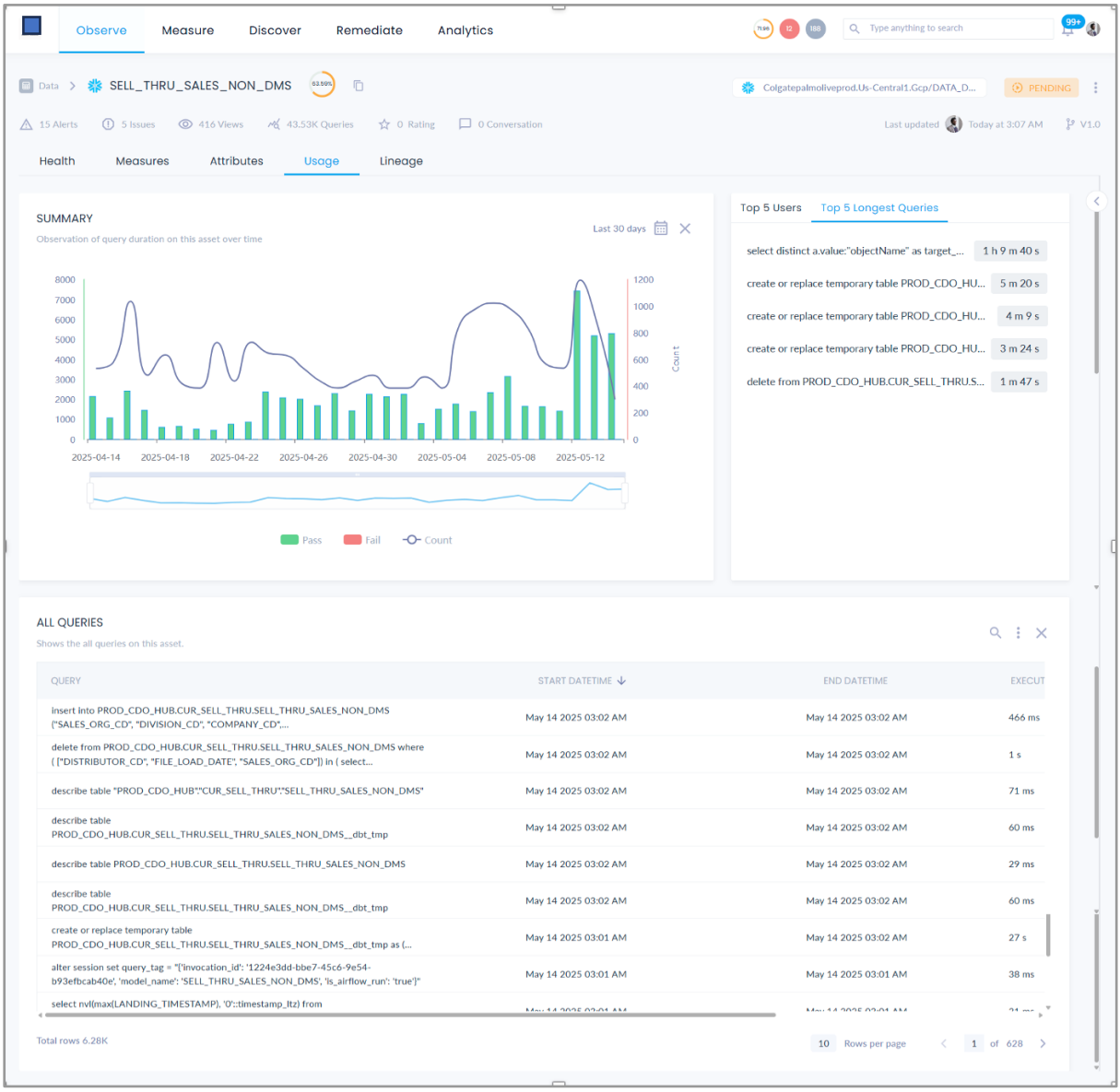
- **Measures Tab**– To view all the types of configured measures (Auto, Behavioral, Query, Conditional, Lookup) on this table. You can view all enabled measures across scoring and monitoring in this tab.



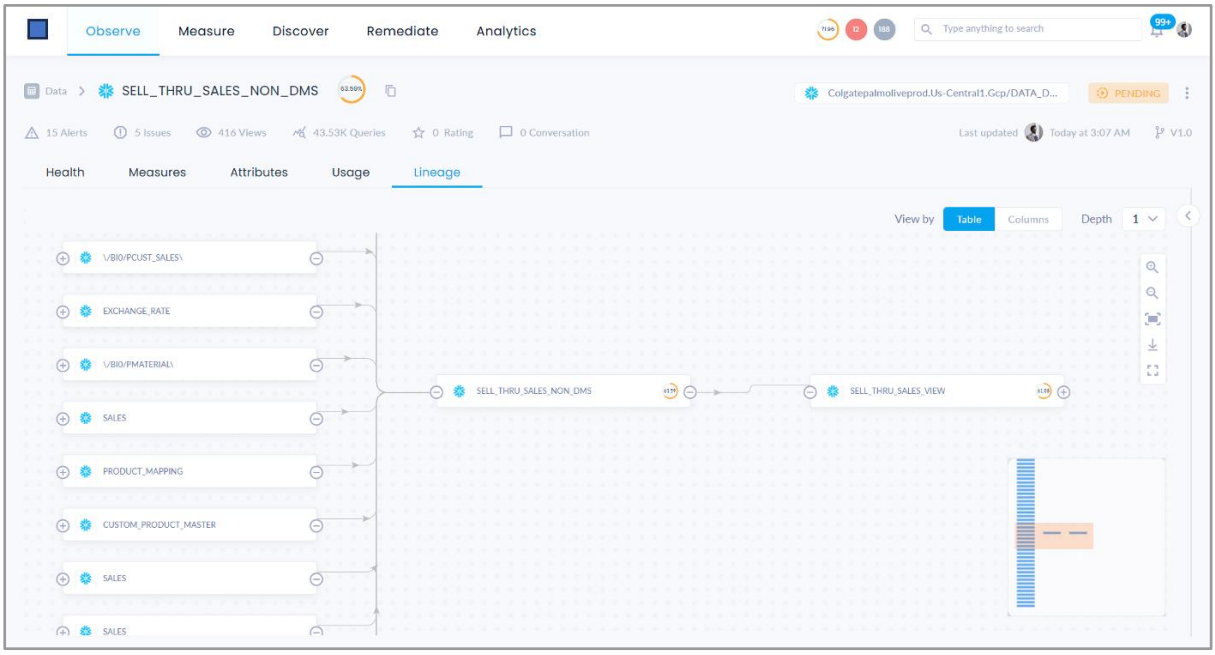
- **Attributes Tab**- You can view all the Attributes (CDE's Columns) with metadata properties and DQ Score.



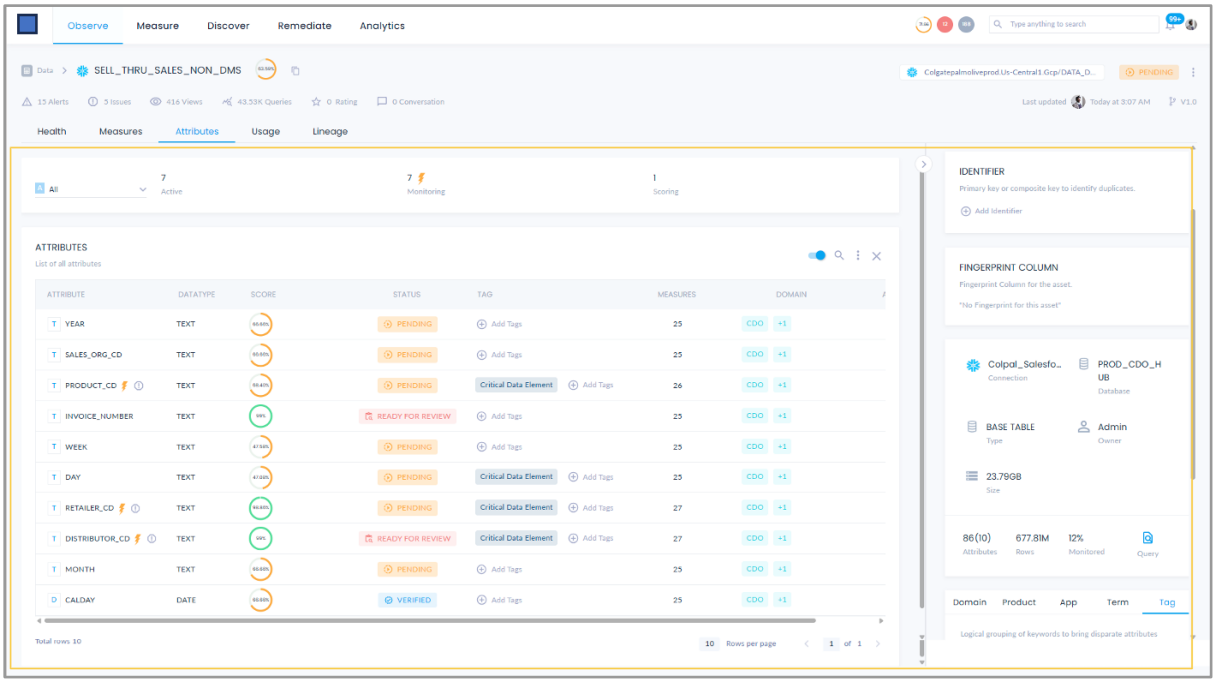
- **Usage Tab** - You can view the information about data sources like Query execution time, all queries, Top 5 users, and top 5 longest queries. This information is fetched from the data source's catalog table and made available to view in erwin DQ so that



- **Lineage** - On the lineage page, you can view the schema and attribute information of the connected asset. This page provides what other objects are present and the data source where the asset is present. The user can also view the DQ score of the associated objects in erwin DQ.

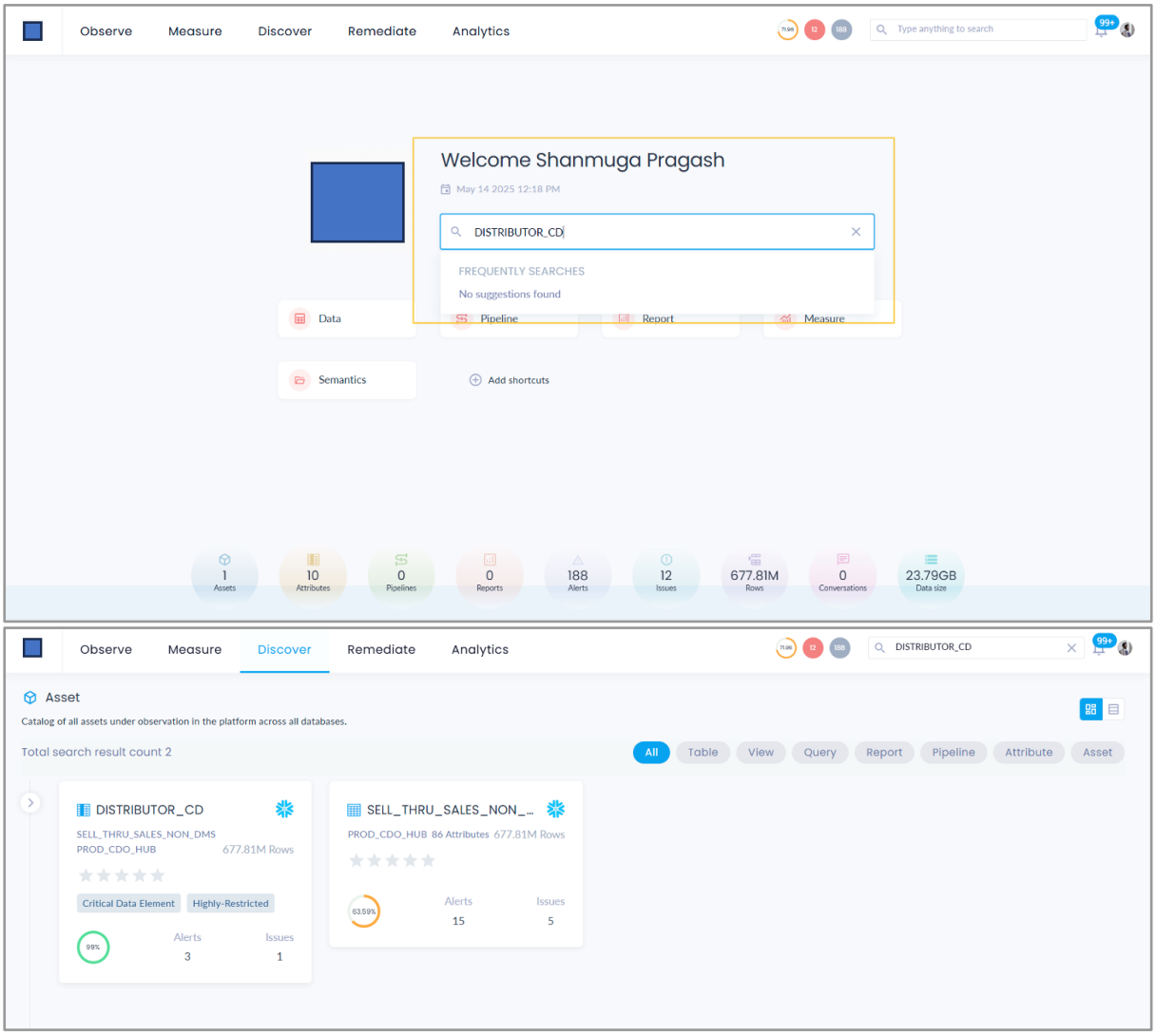


- Click on the attribute tab to navigate to the Attributes (CDE's Columns) and start reviewing the CDE's.



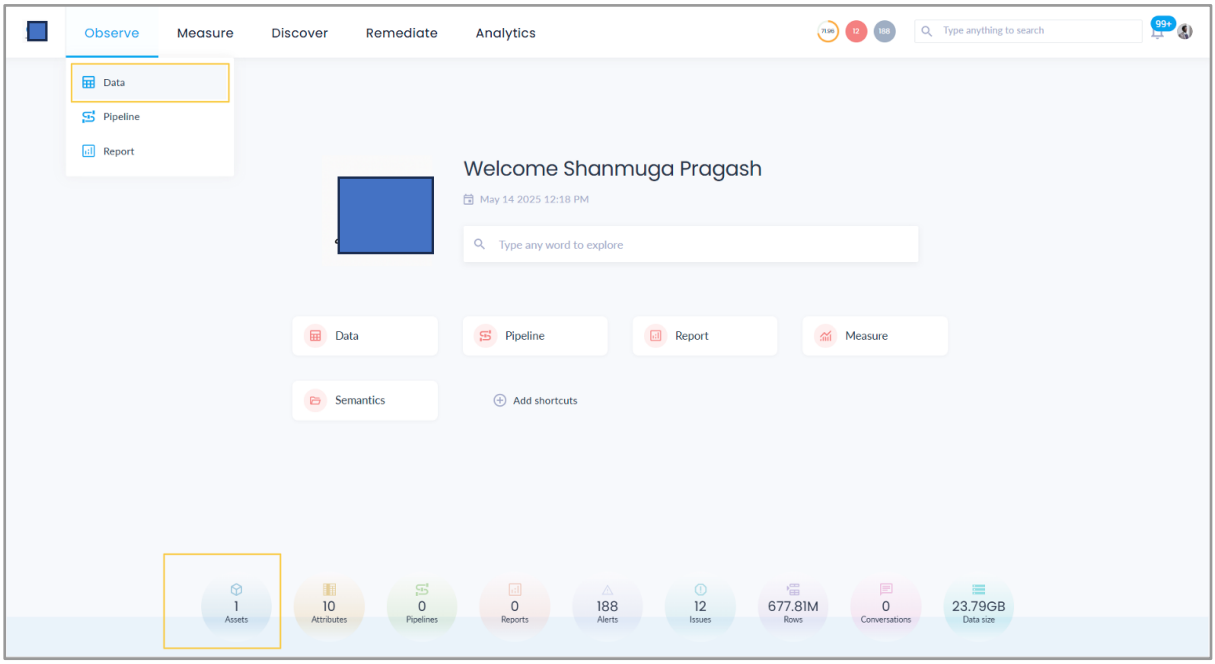
User Journey CDE Discovery 2:

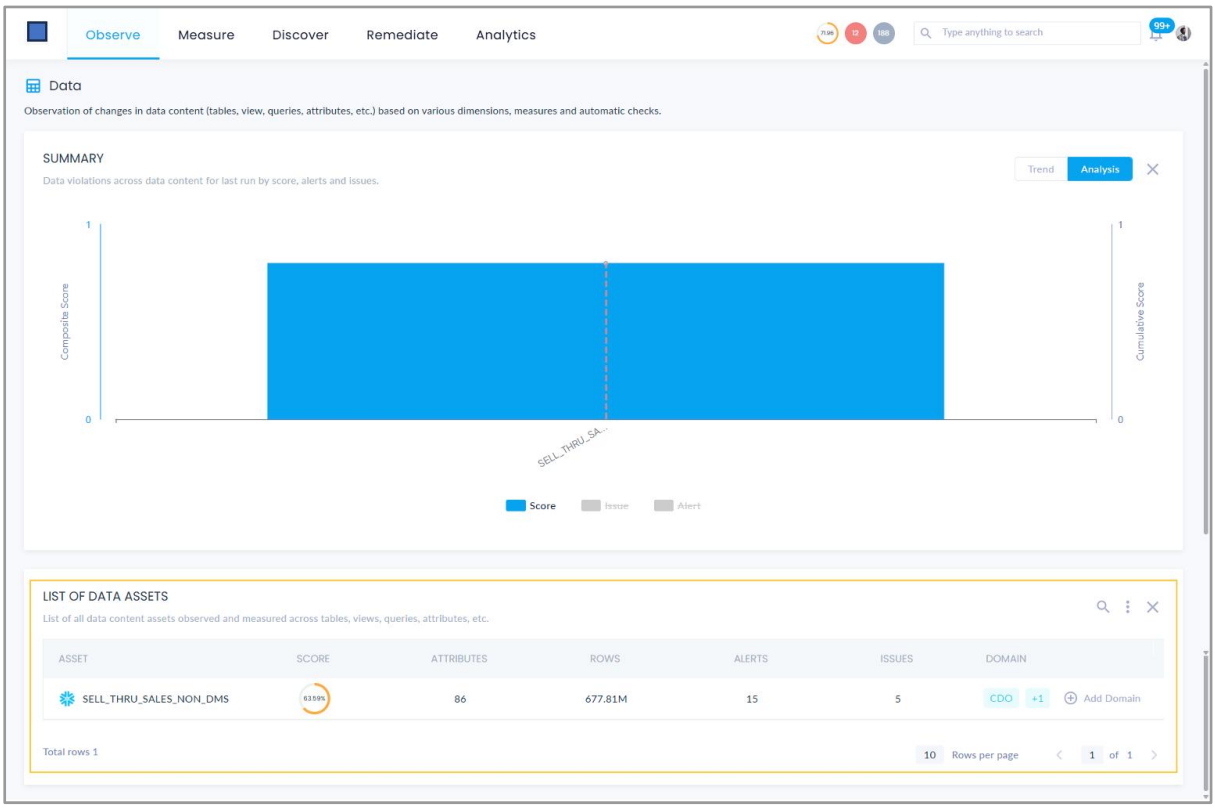
Go to the Home Page and then use the search bar in the middle of the page to provide the ability to search for a data asset in the organization by using a keyword. The user can search for a table or view or report or pipeline, or CDE attributes by using the search option on the homepage.



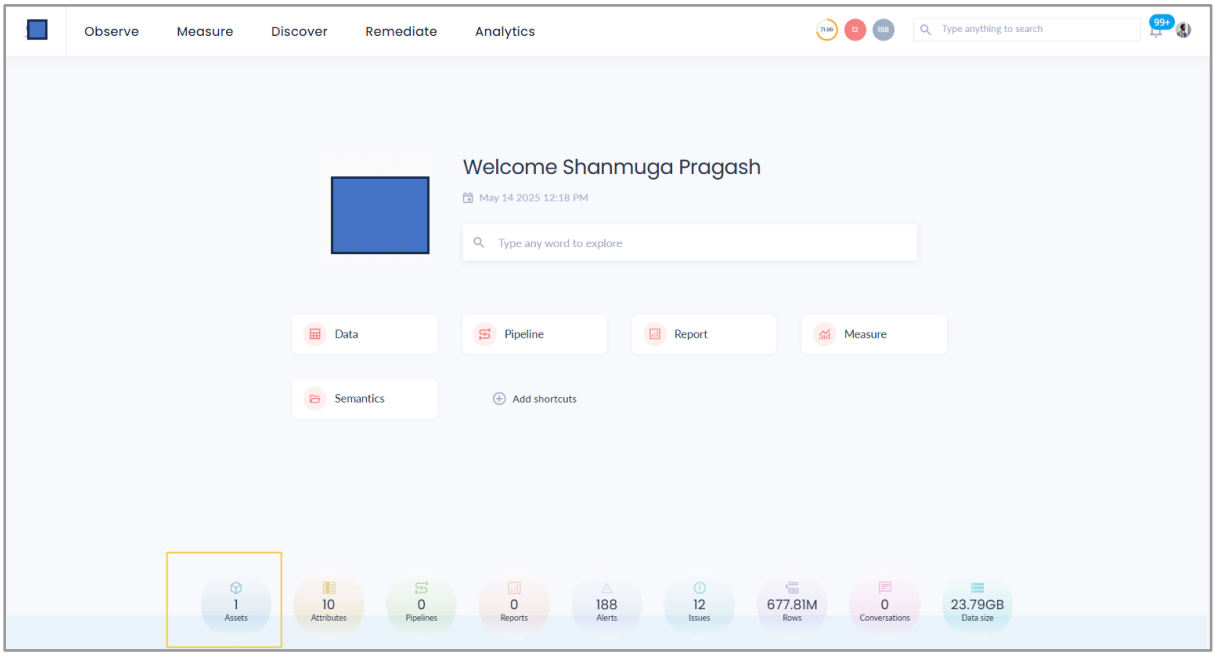
User Journey CDE Discovery 3:

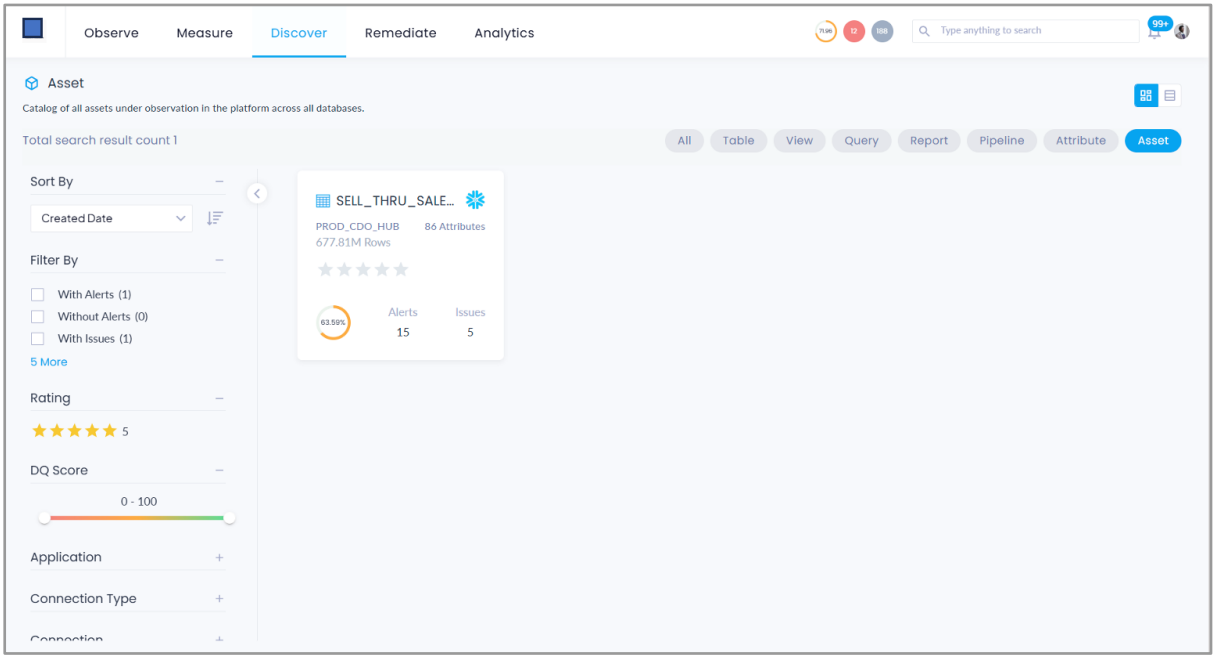
- Navigate to the top header menu Observe > Data, and it will take you to the detailed Data Asserts page where you can view a list of all data content assets observed and measured across tables, views, queries, attributes, etc.





- Navigate to the home page and click on the Assets button at the bottom of the page, and it will take you to the Asset catalog page.


Observe
Measure
Discover
Remediate
Analytics
Type anything to search
99+
Welcome Shanmuga Pragash
May 14 2025 12:18 PM
Type any word to explore
Data
Pipeline
Report
Measure
Semantics
Add shortcuts
1
Assets
10
Attributes
0
Pipelines
0
Reports
188
Alerts
12
Issues
677.81M
Rows
0
Conversations
23.79GB
Data size


Observe
Measure
Discover
Remediate
Analytics
Type anything to search
99+
Asset
Catalog of all assets under observation in the platform across all databases.
Total search result count 1
All
Table
View
Query
Report
Pipeline
Attribute
Asset
Sort By
Created Date
Filter By
With Alerts (1)
Without Alerts (0)
With Issues (1)
5 More
Rating
5
DQ Score
0 - 100
Application
Connection Type
SELL_THRU_SALE...
PROD_CDO_HUB
86 Attributes
677.81M Rows
63.59%
Alerts
15
Issues
5

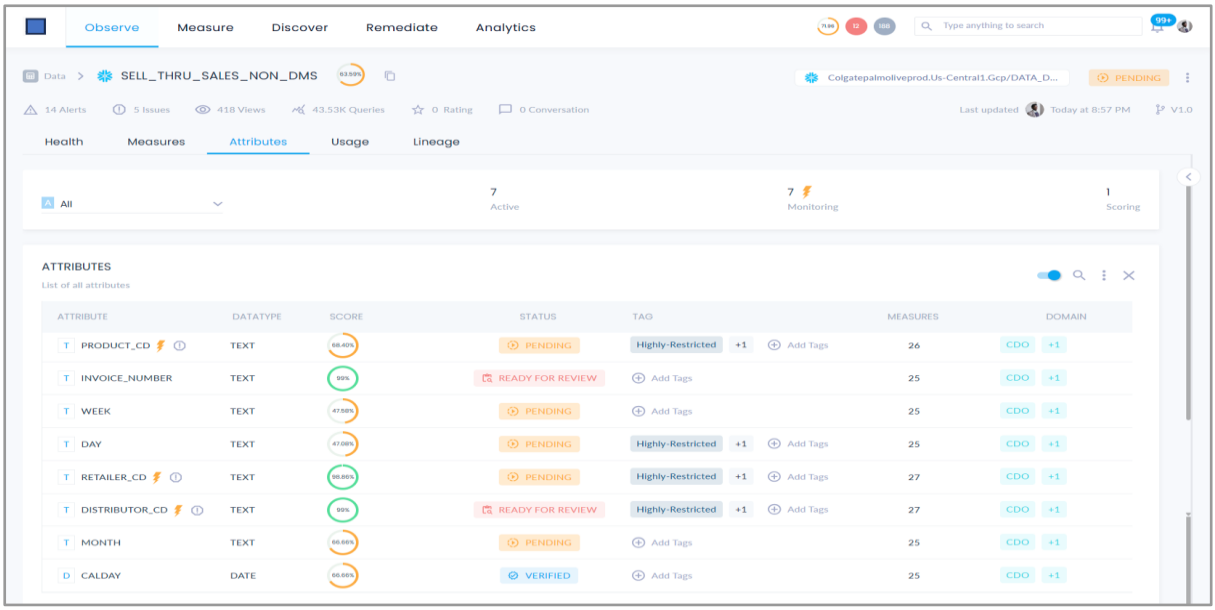
Profiling Data Assets

Data profiling is a critical first step in understanding the structure, content, and quality of your datasets. Before applying any data quality rules, users should review profiling insights to make informed decisions about what rules and thresholds to configure. This chapter covers how to initiate data profiling in erwin DQ and how to interpret the results.

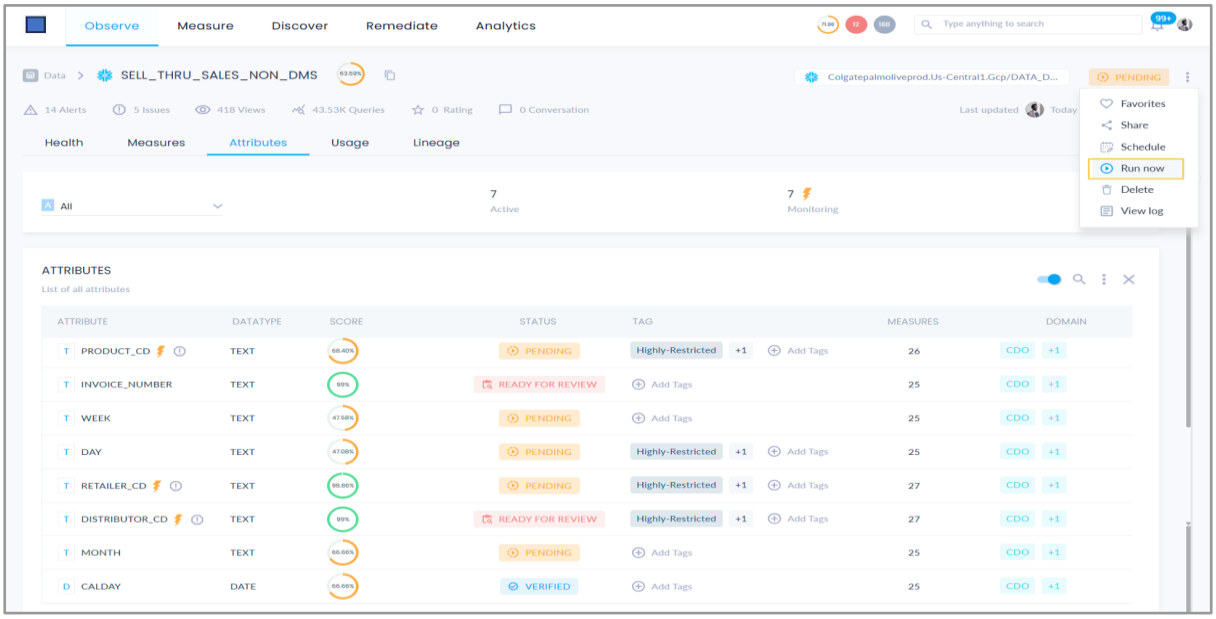
Performing Data Profiling

Where and how do I profile my data asset to understand data distribution?

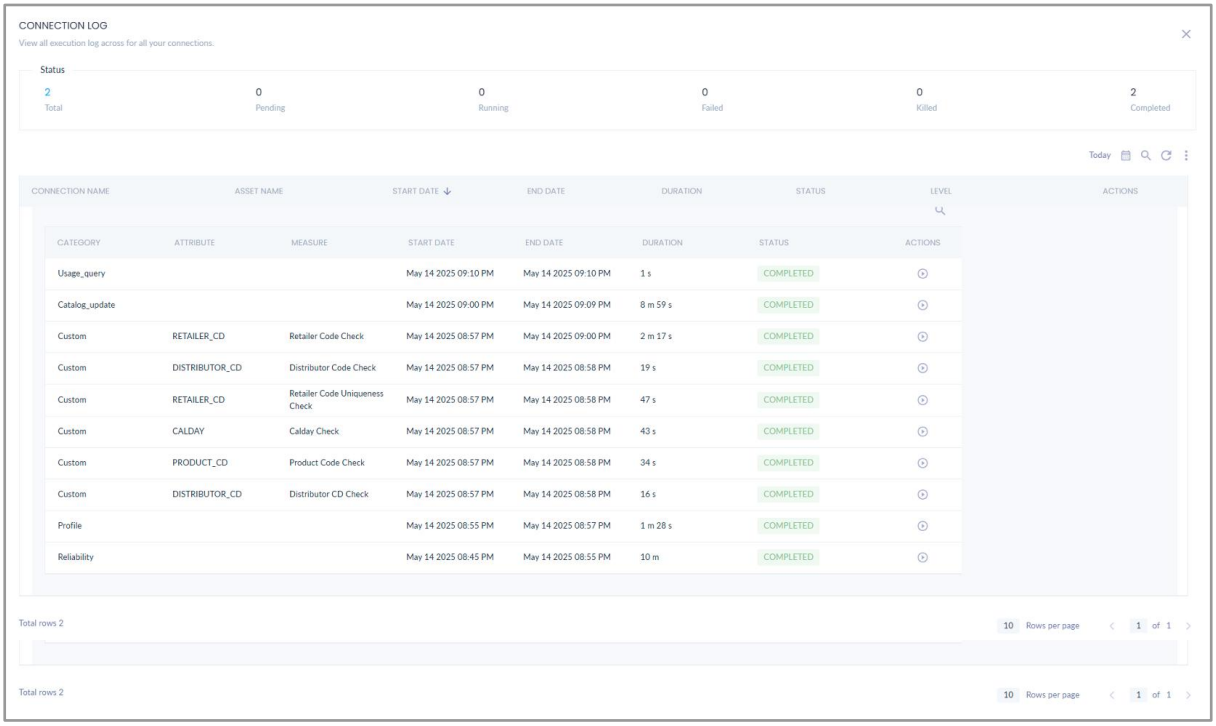
- Navigate to the **Discover** module, search, and select the asset you wish to profile.
- By default, the first time when the Data Governance team configures the table and CDE Columns for you, the platform will automatically start the profiling with the preconfigured out-of-the-box rules and perform the table-level and column-level checks, and generate the Base Data Quality Score.



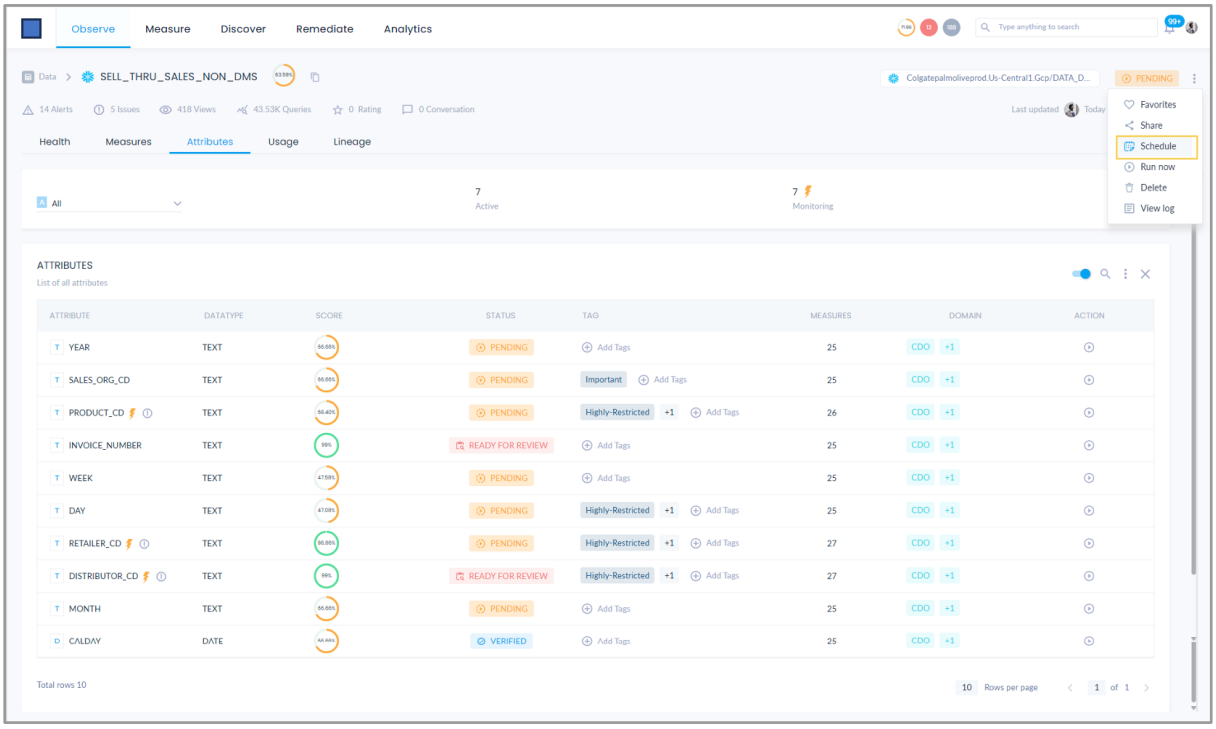
- You can review the initial base profiling results of the table and CDE columns, and then activate or Deactivate or add more OOB measures to the CDE columns.
- After your asset review, if you want to reprofile the asset, then you can run the on-demand profiling job by going to the Top right of the Asset Details page and select the triple dots to expand and click the "**Run Now**" button to initiate the profiling job.

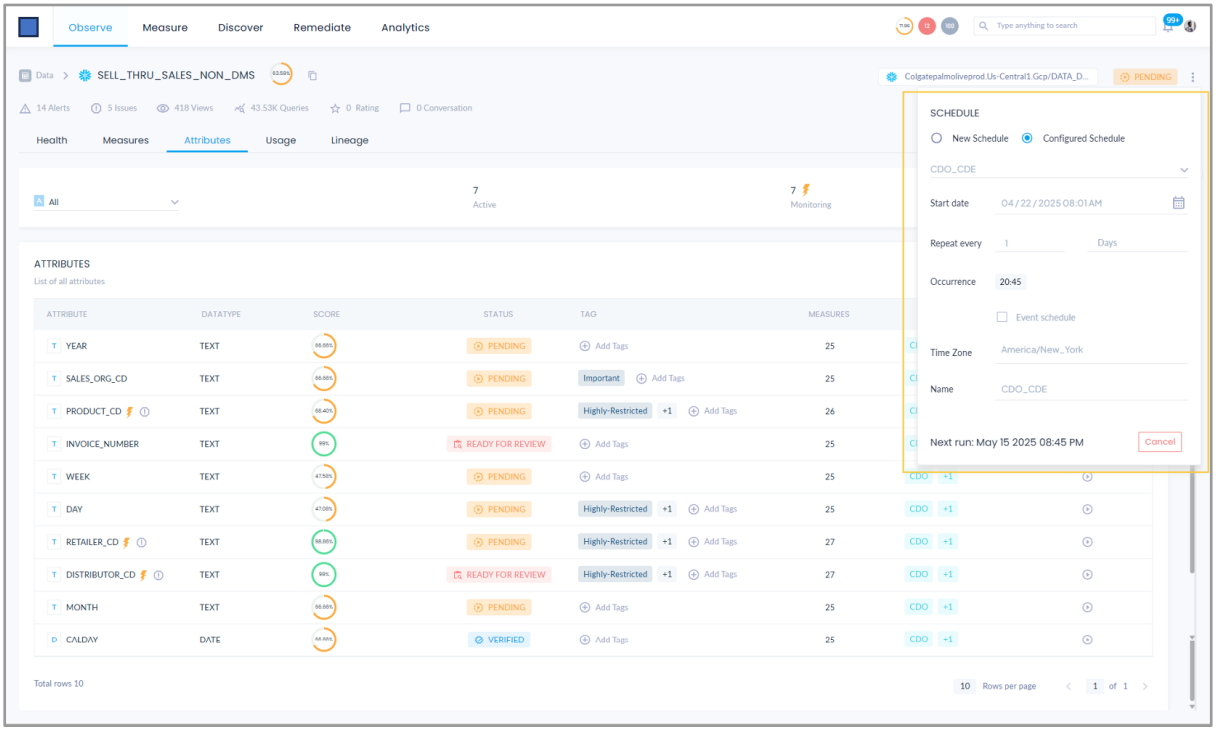


- You can click on the **View Log** and monitor the profiling process.



- Once profiling is complete, a summary of key metrics is displayed for each attribute.
- If you want to view the auto profiling schedule, you can click on the schedule and view the schedule frequency. The profiling is automatically run based on the master schedule defined by the Data Governance Team or whenever manual execution is triggered.

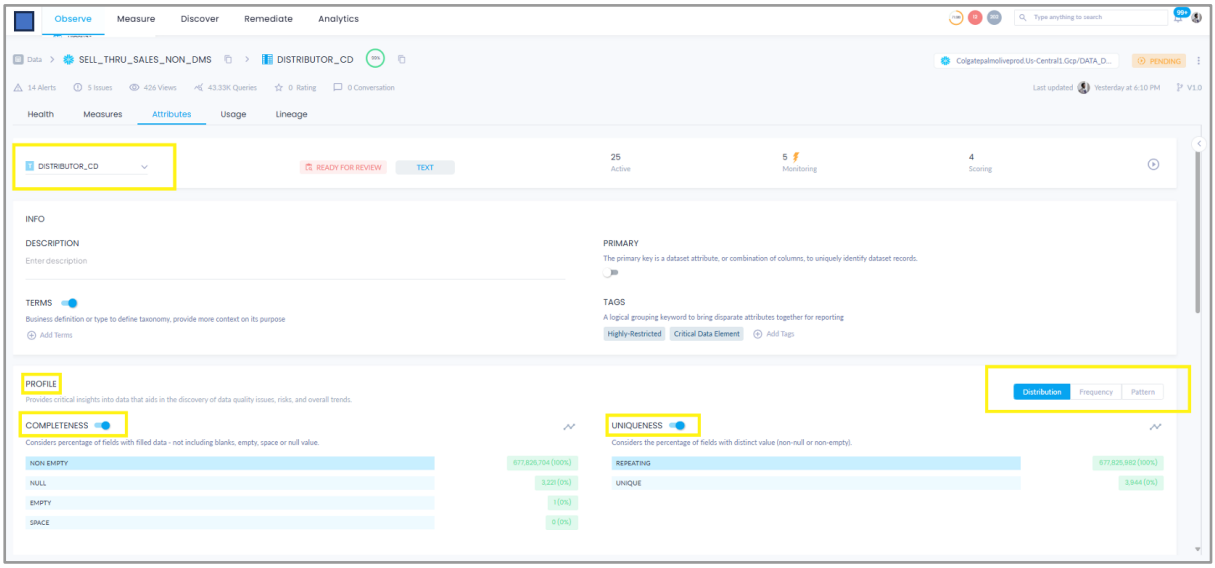




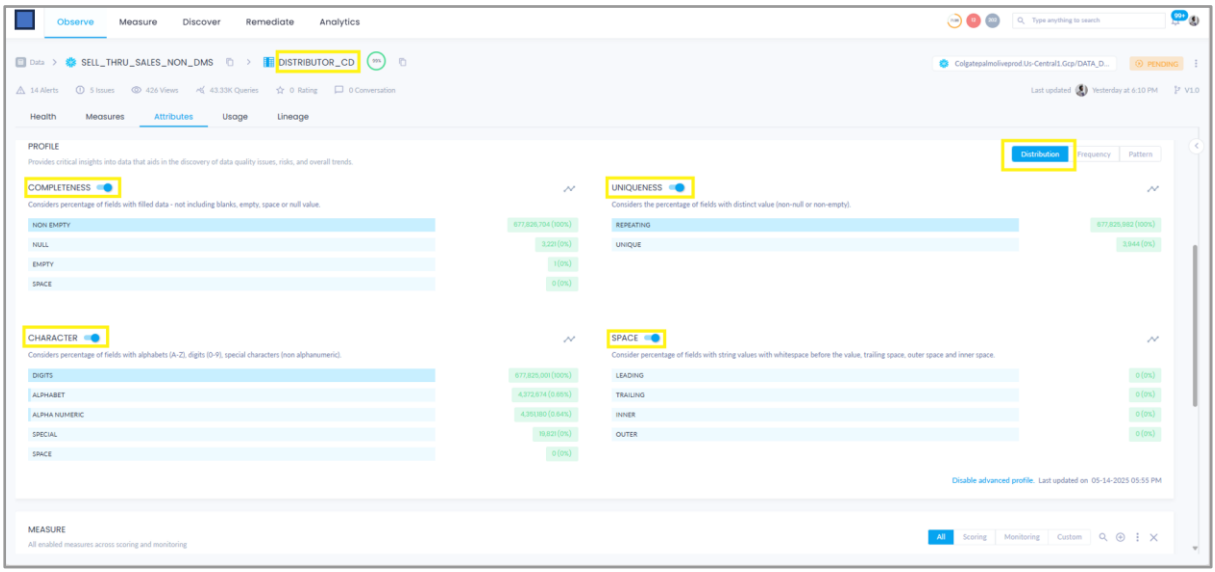
Interpreting Profiling Results

How do I interpret profiling results before applying DQ rules?

Profiling results provide a comprehensive summary of each attribute. Reviewing these results helps in identifying data anomalies, patterns, and quality gaps before defining DQ rules.



- Key profiling metrics explained:
- **Distribution**:
  - Visualizes the frequency of values or value buckets in the column.
  - Helps detect skewed data, dominant values, or gaps.
- **Null Counts / Completeness**:
  - Shows the percentage of null or missing values.
  - Key metric for evaluating the completeness of a field.
- **Uniqueness**:
  - Percentage of unique values in a column.
  - Useful for primary keys or fields that are expected to be unique.
- **Patterns**:
  - Detects common formatting or value patterns (e.g., dates, phone numbers, codes).
  - Help identify formatting inconsistencies or potential validation needs.
- **Min/Max/Mean/Standard Deviation**:
  - Useful for numeric fields to identify outliers or invalid ranges.



Profiling enables you to make informed decisions before rule creation. It ensures DQ rules are relevant and aligned with the actual data structure and behavior. Accurate profiling leads to better targeting rules, fewer false alerts, and more actionable insights.

Managing Semantics and Tags

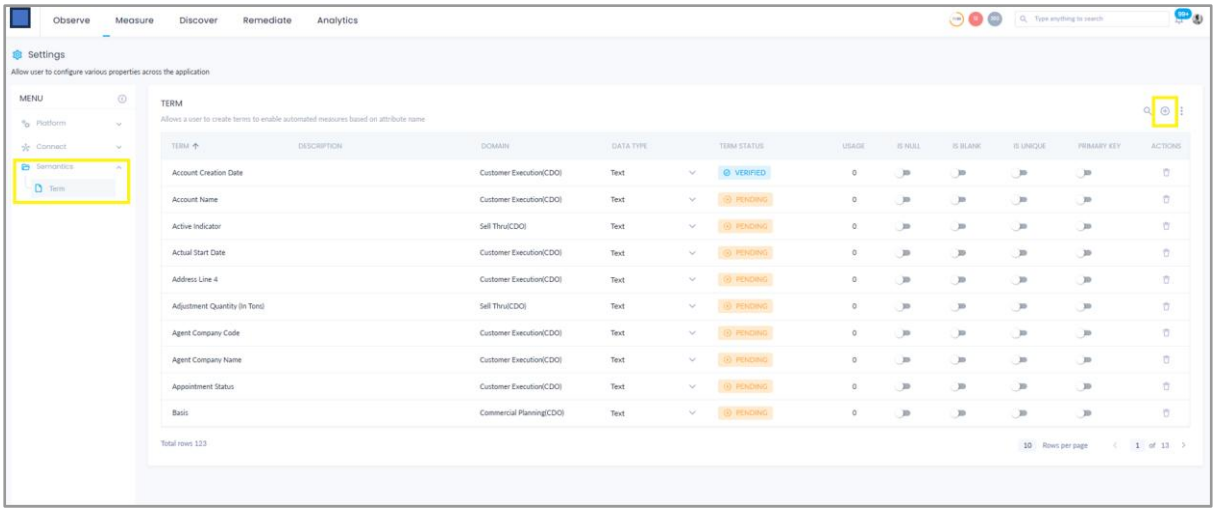
erwin DQ allows users to enhance data context and governance by applying semantic tags and business terms. This not only aids search and filtering but also improves alignment between technical metadata and business meaning. In this chapter, we'll explore how to apply and manage semantics effectively.

Semantic Tags & Business Terms

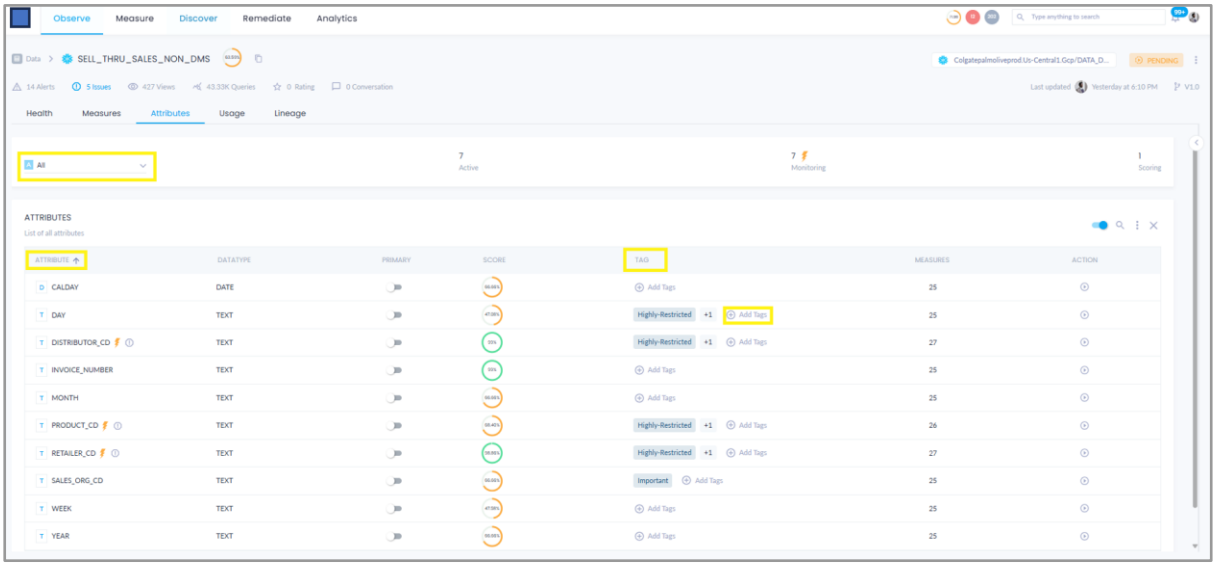
How do I create and manage business terms and apply predefined semantic tags?

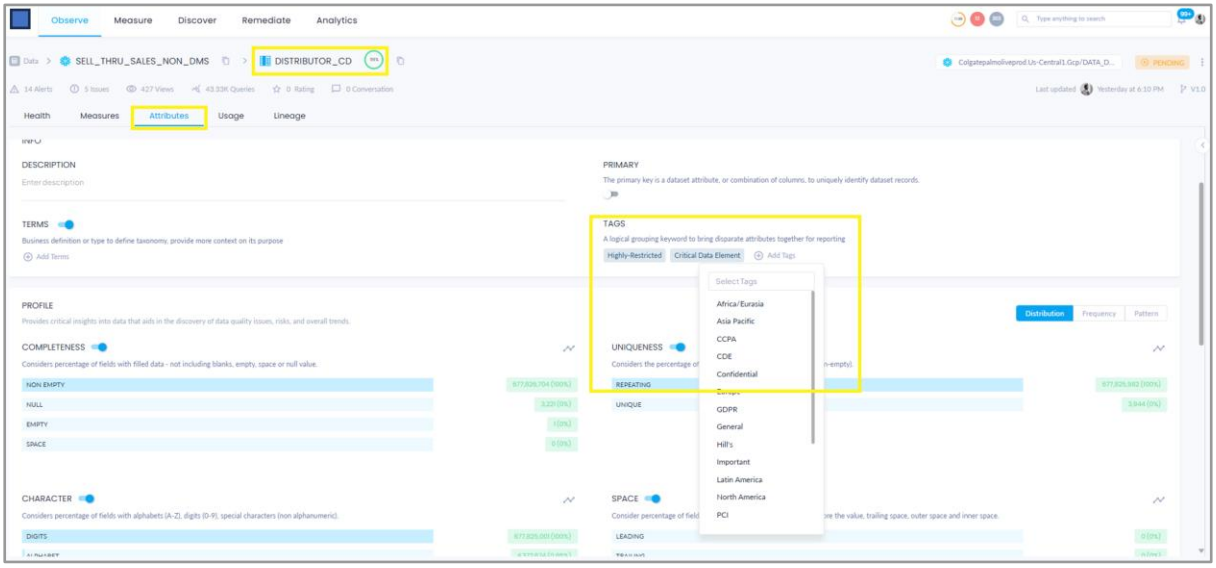
Semantic tags classify attributes with context (e.g., PII, Financial, Contact Info), while business terms align attributes with glossary definitions.

- Creating and Managing Business Terms:
  - Navigate to **Settings** > **Semantics** > **Term**.
  - Click on "**Add Term**" to define a new business term.
  - Provide a term name, description, and assign it to a domain.
  - Save the term. It will now appear in the glossary and be available for assignment on the platform.

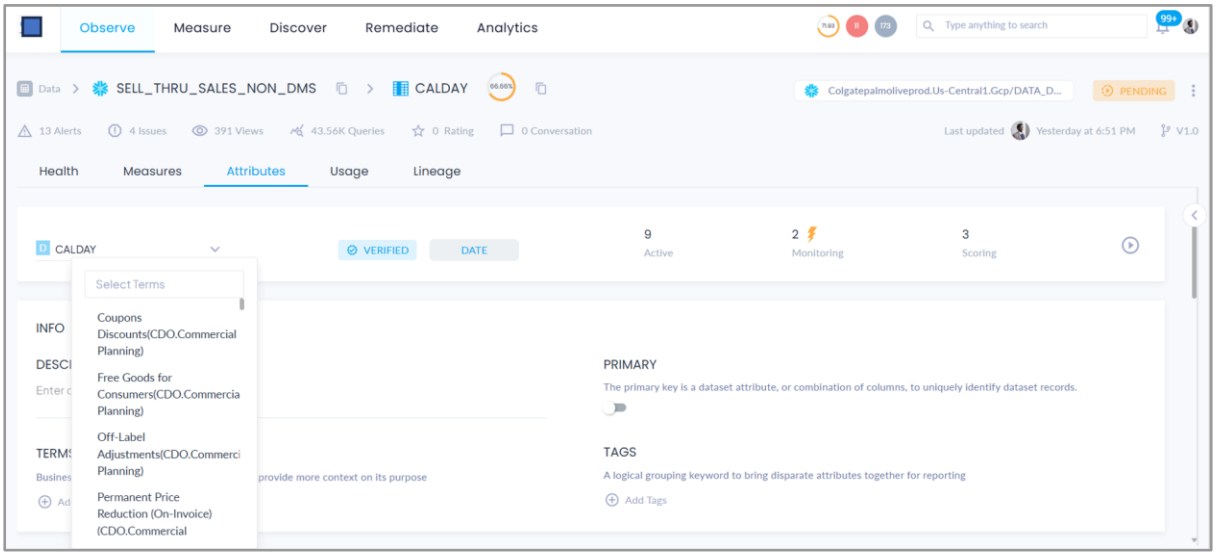


- Applying Predefined Semantic Tags:
  - Navigate to the Discover module and open an asset.
  - Click the Attributes tab to view the column-level metadata.
  - Use the "Semantic Tag" dropdown next to each attribute to apply a predefined tag.
  - Tags are centrally maintained by the Data Governance team or platform admins.





- Assigning Business Terms:
    - From the same Attributes tab, use the "Term" column to assign glossary terms.
    - Select from the list of predefined business terms in the dropdown.



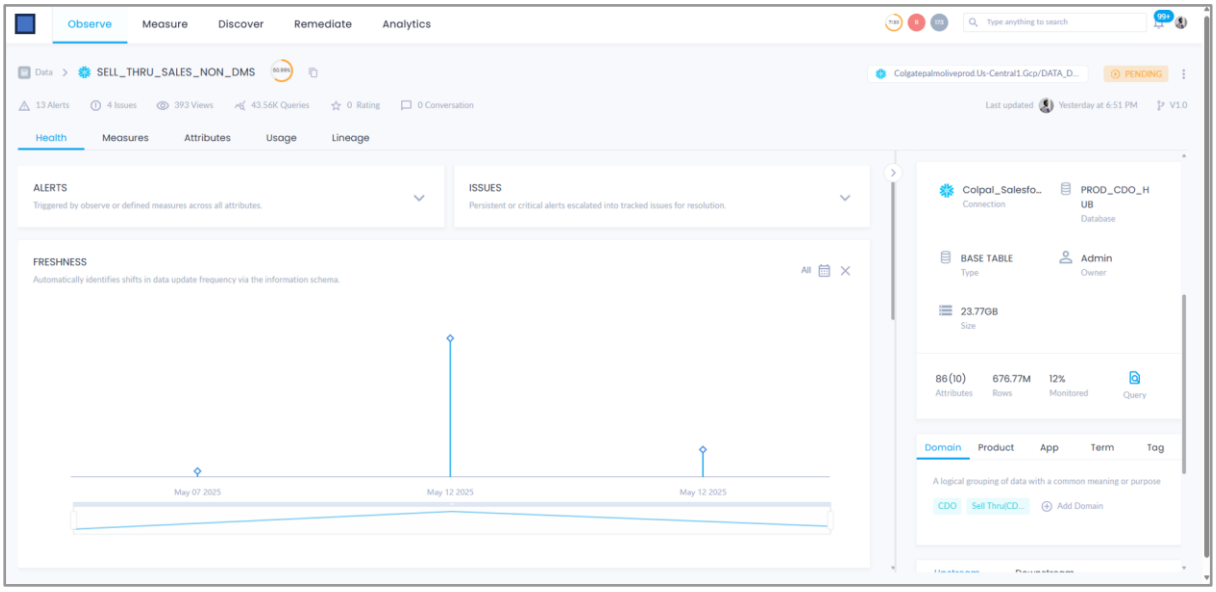
- Applying semantic tags and terms helps:
  - Categorize and prioritize attributes for monitoring.
  - Align technical columns with business meaning.
  - Improve searchability and clarity in dashboards and reports.

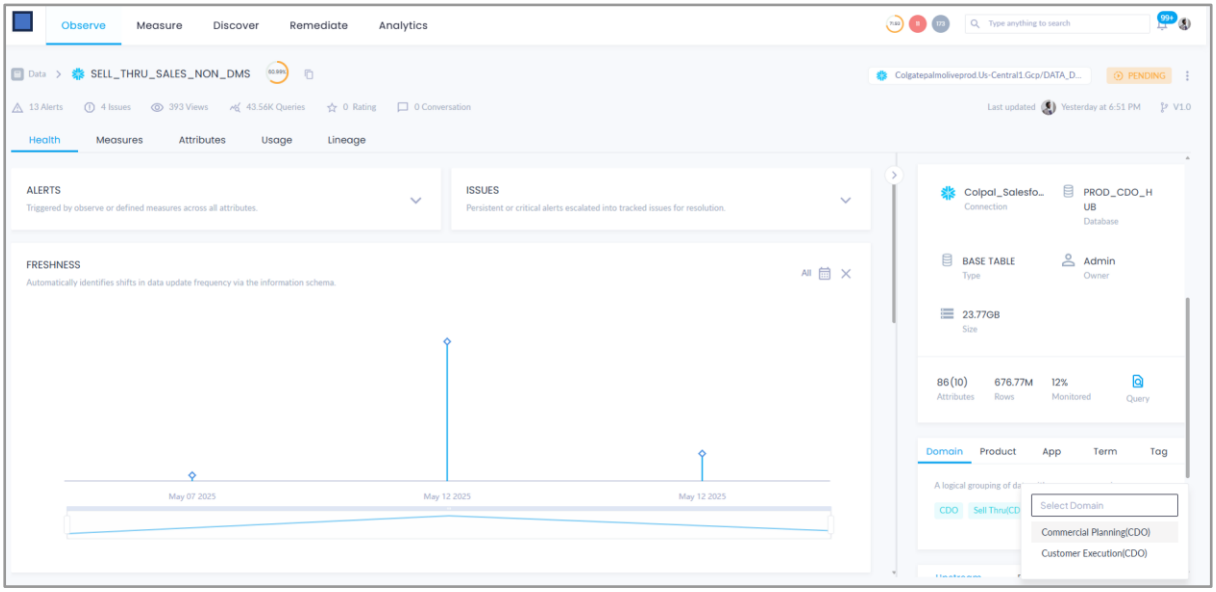
Domain Mappings

Where and how do I review domain mappings for consistency?

Semantic consistency across domains ensures that tagging is uniform and meaningful. erwin DQ provides a centralized view to review tag-domain mappings.

- To review Domain mappings:
  - From the Asset, use the "Domain" column to assign a domain, sub-domain or Application, or product.
  - Select from the list of predefined business domains, applications, products in the dropdown.
  - Review, update, or manage mappings as needed.





- Maintaining consistent domain mappings:
    - Supports unified DQ rule applications.
    - Enables domain-specific alert routing.
    - Enhance clarity in reporting by ensuring tagged elements match governance expectations.

Using semantics and tags effectively creates a shared understanding of data, streamlines governance practices, and ensures targeted data quality monitoring across domains.

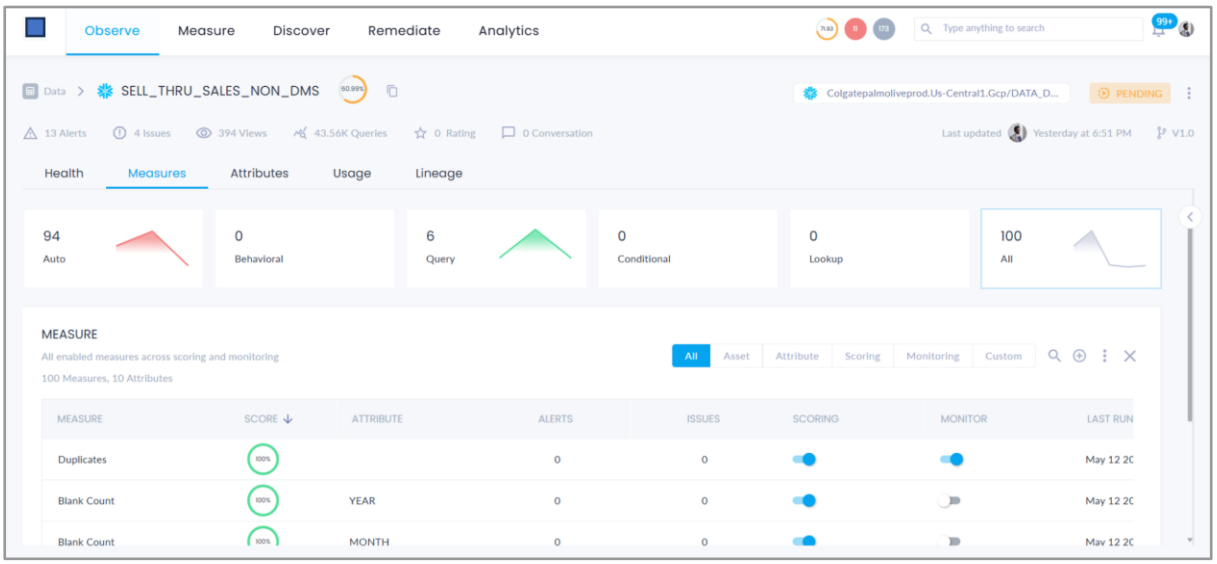
Configuring Data Quality Rules

Data Quality (DQ) rules are the foundation of automated data monitoring in erwin DQ. They help detect inconsistencies, errors, or anomalies in critical data attributes. This chapter walks through how to configure, define, and activate different types of DQ rules, particularly focusing on tasks handled by Business Data Stewards.

Setting Up Rules (Business Data Stewards)

Where and how do I configure DQ rules using out-of-the-box options?

- Navigate to **Measure** in the top header menu.
- Select **Measure List** to view all existing measures.
- Click **"Add Measure"** or **"Auto Measures"** to configure new rules.
- Choose an **out-of-the-box (OOB)** rule from the list:
  - Completeness
  - Uniqueness
  - Validity
  - Pattern Check
  - Format Check
  - Length
- Provide a name and optional description, then assign it to an attribute.
- Configure thresholds as needed (see section below).

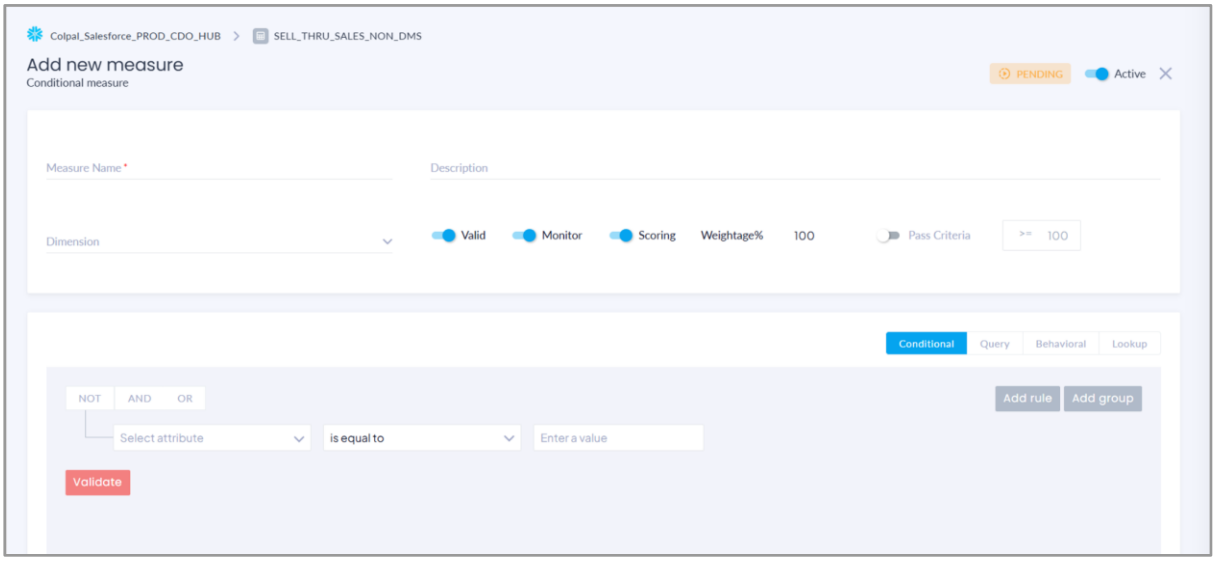


Where and how do I create conditional, behavioral, or query-based rules?

- Go to **Measure** > **Add Measure**.
- Select the **Rule Type** as Conditional or Query or Behavioral, or Lookup.

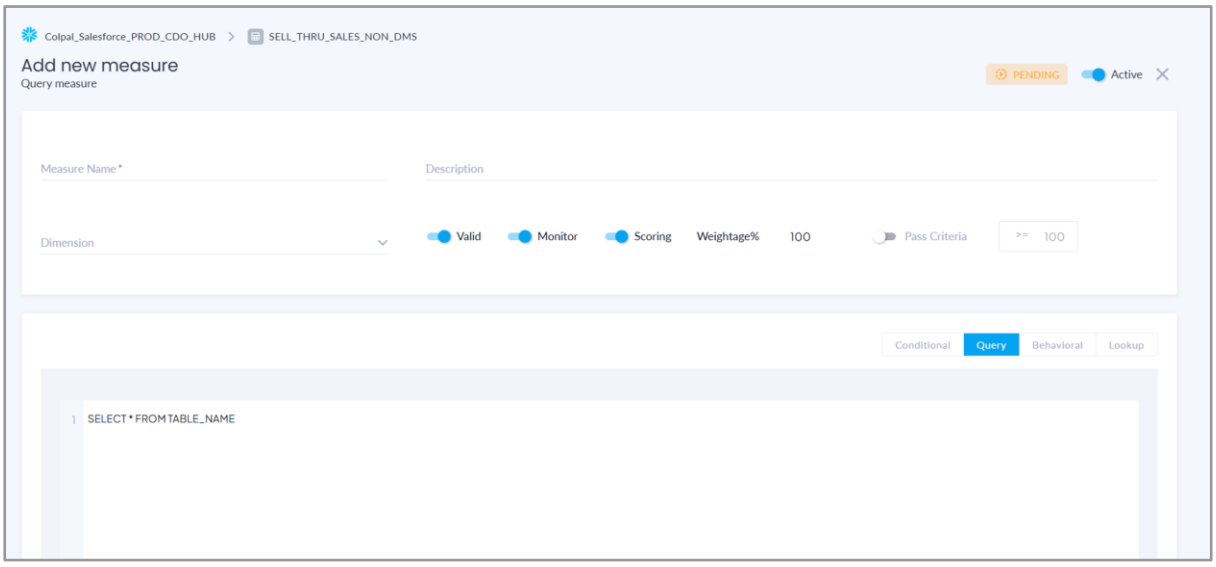
Conditional & Behavioral Rules:

- Go to **Measure** > **Add Measure**.
- Select the **Rule Type** as Conditional or Behavioral.
- Enter the rule logic using dropdown-driven conditions.
  Example: "If Status = 'Active', then Email must not be null."



Query-Based Rules:

- Navigate to **Measure** > **Add Measure**.
- Choose **Rule Type: Query**.
- Write a custom SQL WHERE clause to define the failed condition.
  - Example: WHERE email NOT LIKE '%@%' OR email IS NULL
- Test the query before saving.



Tip: Use query-based rules when validation logic involves multiple fields or complex calculations.

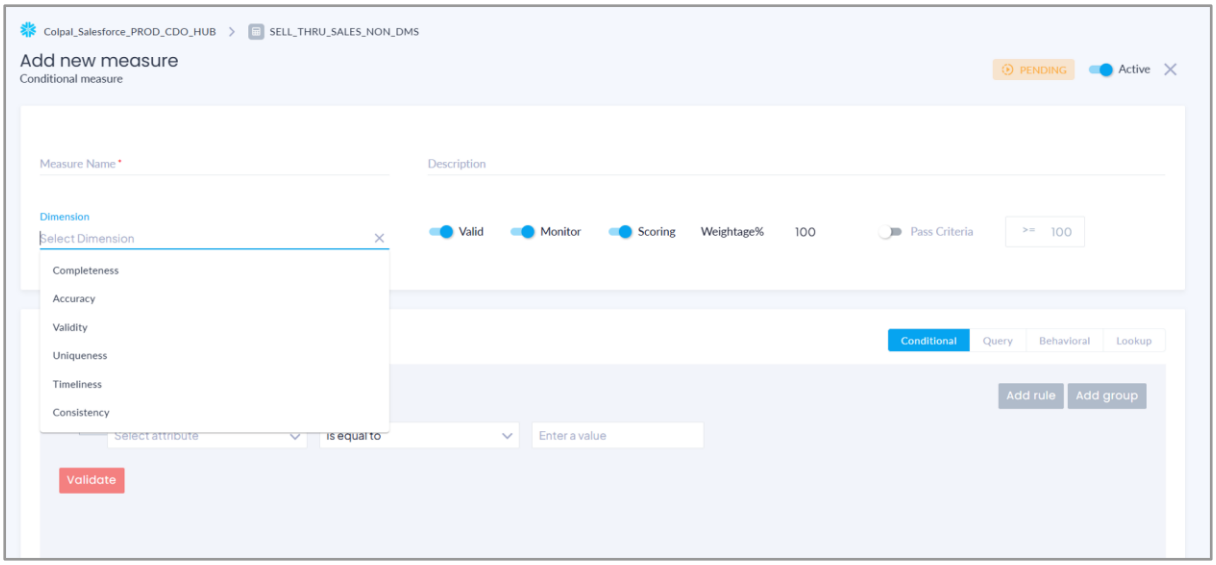
Defining Dimensions & Thresholds

How do I define DQ dimensions and thresholds for each rule?

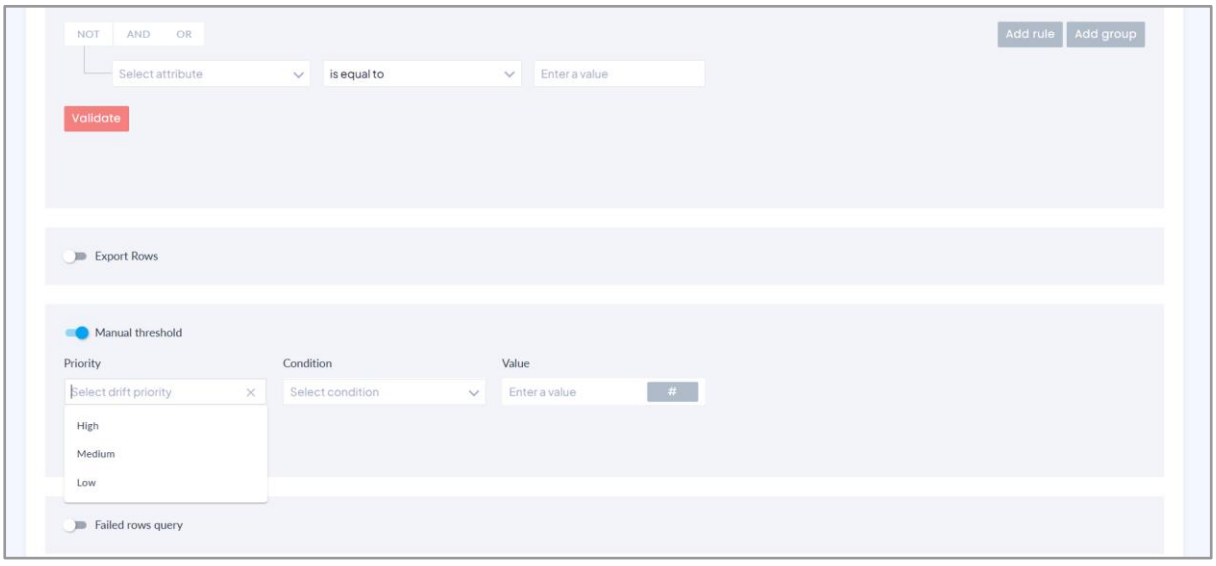
When configuring a rule, select the **DQ Dimension**:

- Completeness
- Accuracy

- Consistency
- Validity
- Uniqueness



- Define the **Threshold**:
  - Specific thresholds and conditions ensure relevant and timely notifications, making it easier to address the most critical issues first.
  - E.x. Pass if more than 95% valid
  - Warn between 80%–95%
  - Fail if below 80%
- Assign **Rule Criticality** (High, Medium, Low) to influence alert priority.

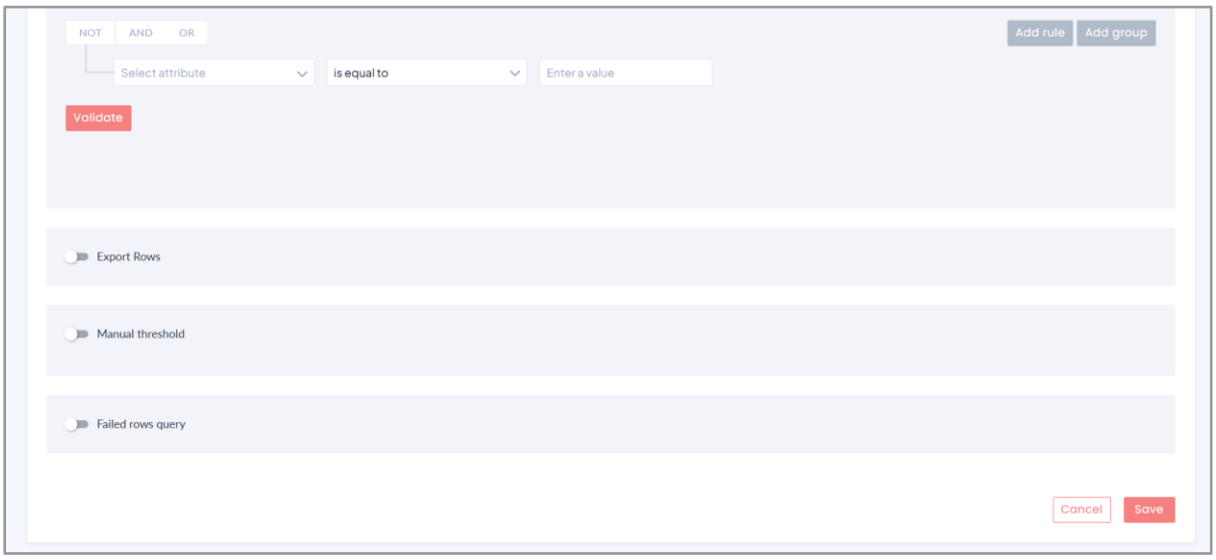


Thresholds help segment rule outputs and trigger alerts with appropriate severity levels.

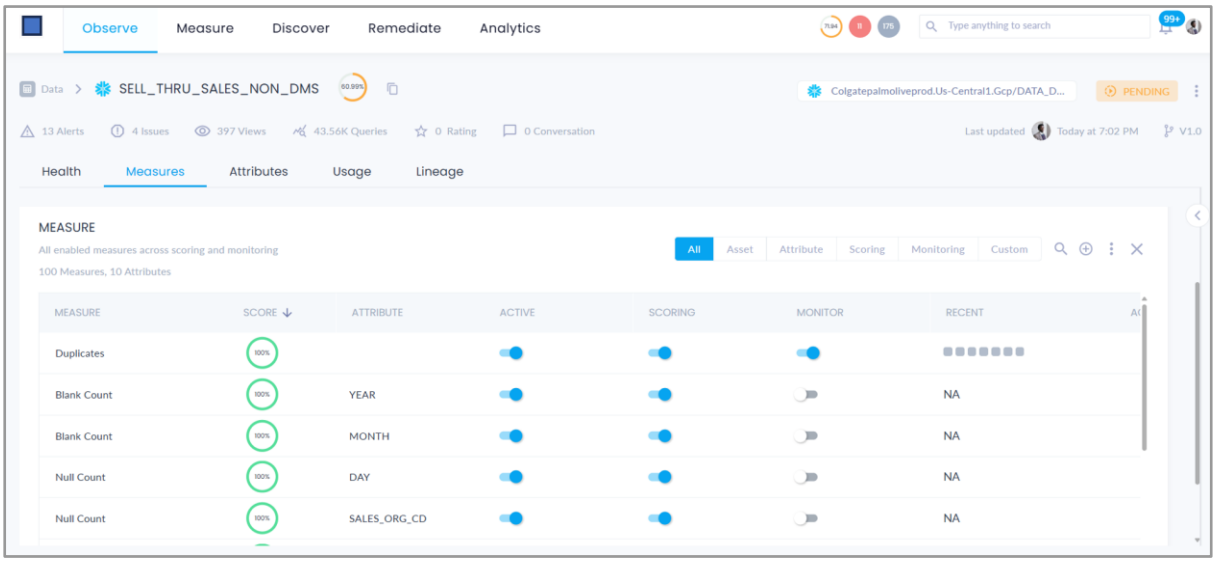
Enabling and Activating Rules

How do I save and enable rules on the target assets?

- After configuring the custom rule, click “**Save**”.



- Navigate to the **Measure List** and locate your Out of Box rule.
- Toggle the switch under the **Active** column to enable it.
- Once activated, the rule will be applied to the asset during profiling jobs.



Active rules are executed during asset profiling and contribute to data quality scoring, dashboards, and alerts.

With these configurations in place, Business Data Stewards can continuously monitor data integrity across critical fields, ensuring that quality expectations are met and actionable alerts are generated.

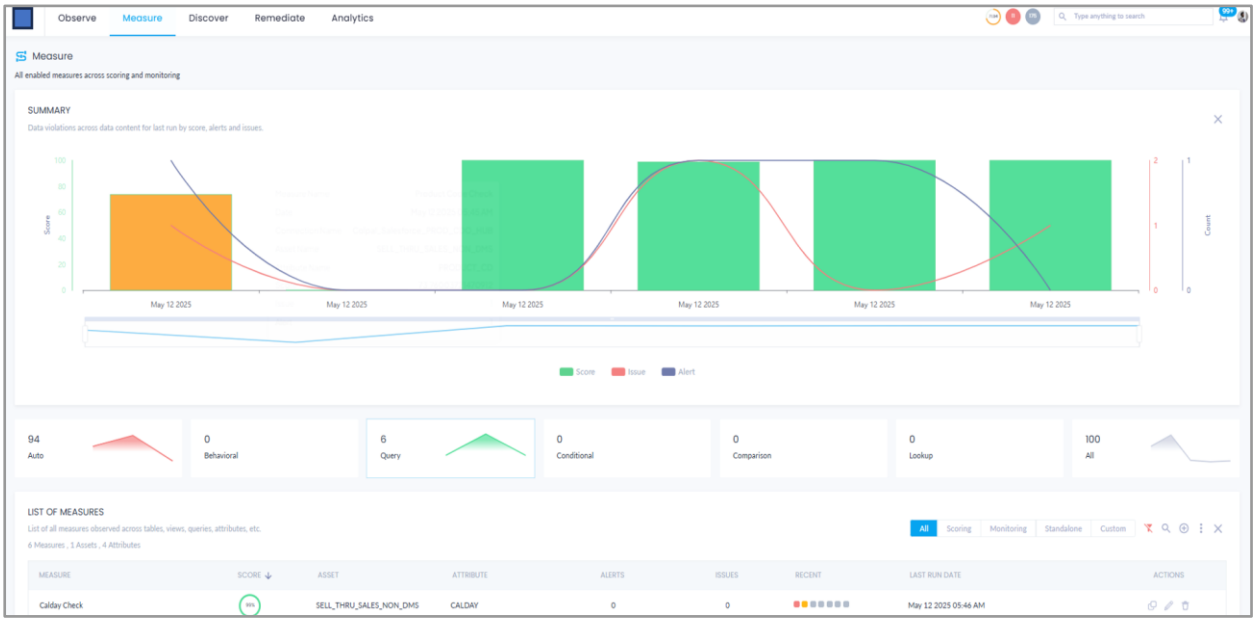
Monitoring and Scheduling

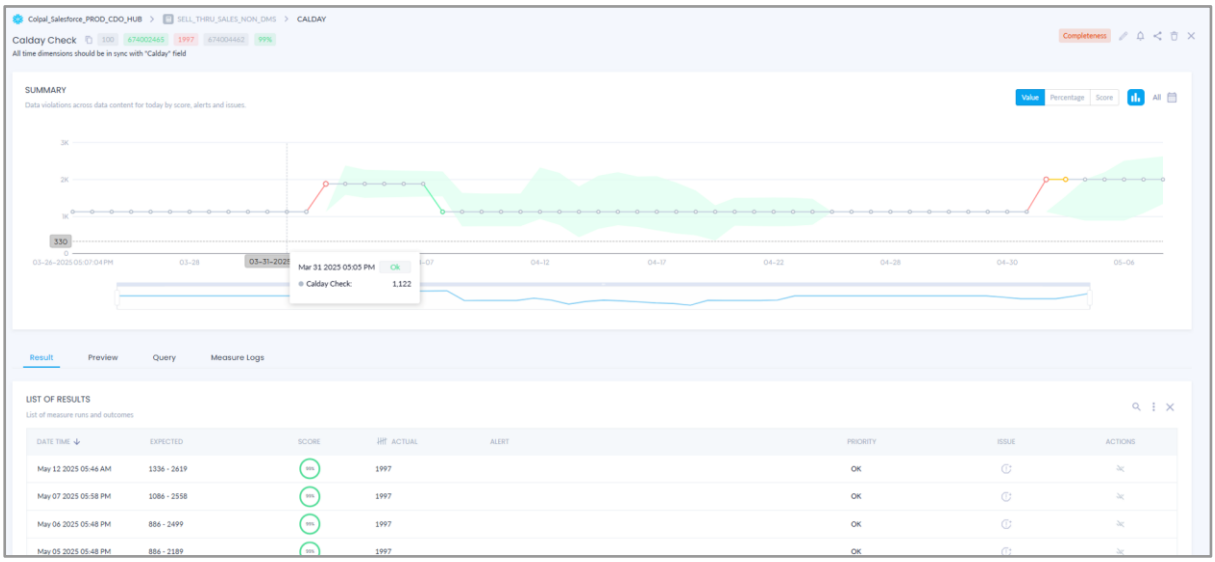
Once Data Quality (DQ) rules are configured, the next step is to monitor their performance, schedule recurring profiling jobs, and review logs. This ensures continuous visibility into the health of your data and automated execution of checks without manual intervention.

Data Quality Scores & Metrics

Where and how do I monitor data quality scores?

- Navigate to the **Measure** module in the top menu.
- Key metrics displayed include:
    - Rule pass/fail count by asset or domain
    - % Score per DQ dimension (Completeness, Accuracy, Validity, etc.)
    - Trends over time for active rules



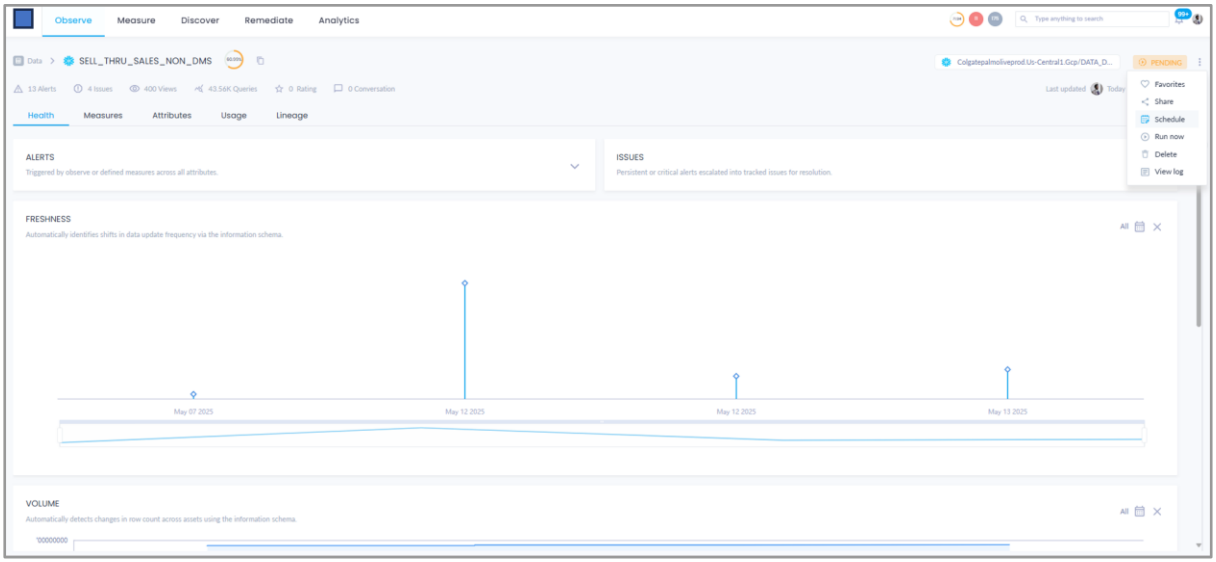


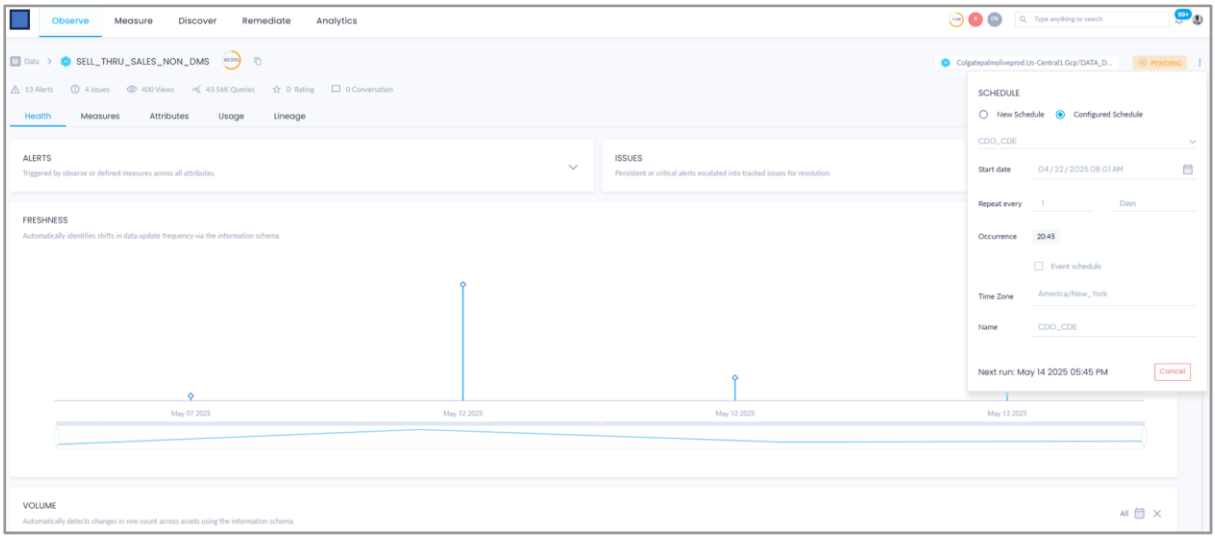
These scores provide an at-a-glance summary of your data quality health.

Scheduling DQ Rules & Profiling Jobs

How do I schedule DQ rule execution or profiling jobs?

- Go to the desired **Asset** from the Discover module.
- In the asset view, click the "**Schedule**" tab.
- Click "**Add Schedule"** to configure a new job.
- Define the frequency:
    - One-Time
    - Daily / Weekly / Monthly
    - Cron Expression (Advanced)
- Optionally, set notification preferences for job success/failure.
- Click **Save** to finalize the schedule.



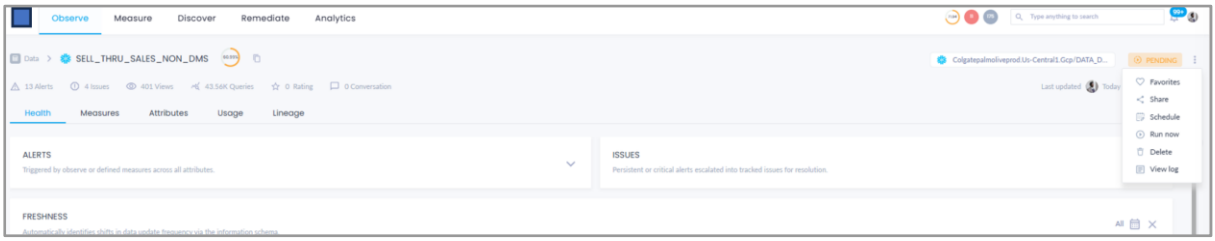


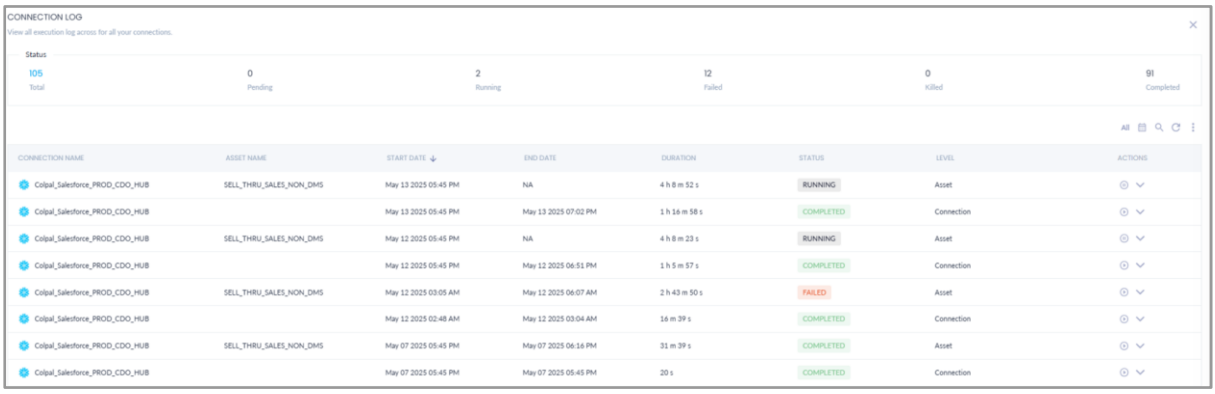
Scheduled jobs help automate DQ enforcement and free up manual effort.

Viewing Scheduled Jobs & Logs

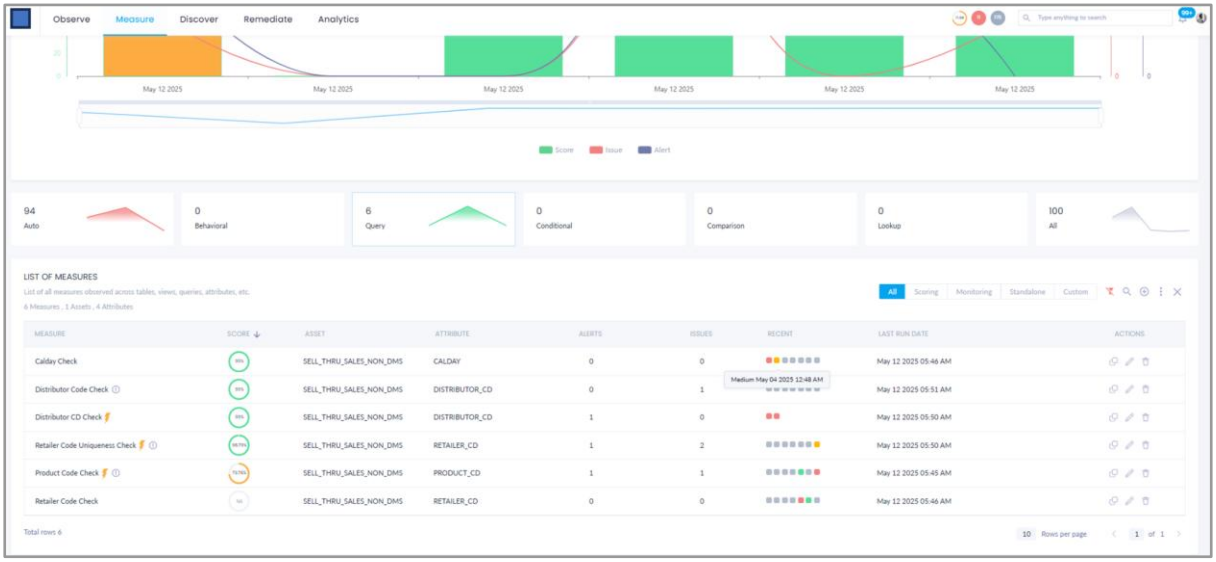
Where do I find scheduled job run history and profiling logs?

- Open the **Asset** from Discover.
- Click on the **Schedule** tab.
- Click "**View Logs**" next to any scheduled job.
- Logs include:
- Status (Success, Fail, In Progress)
- Run Date & Time
- Number of records processed
- Errors or warnings if any





Reviewing logs regularly ensures issues are caught early and jobs are running as expected.



By combining monitoring with scheduled automation, erwin DQ empowers teams to continuously enforce high data quality standards and stay on top of emerging issues.

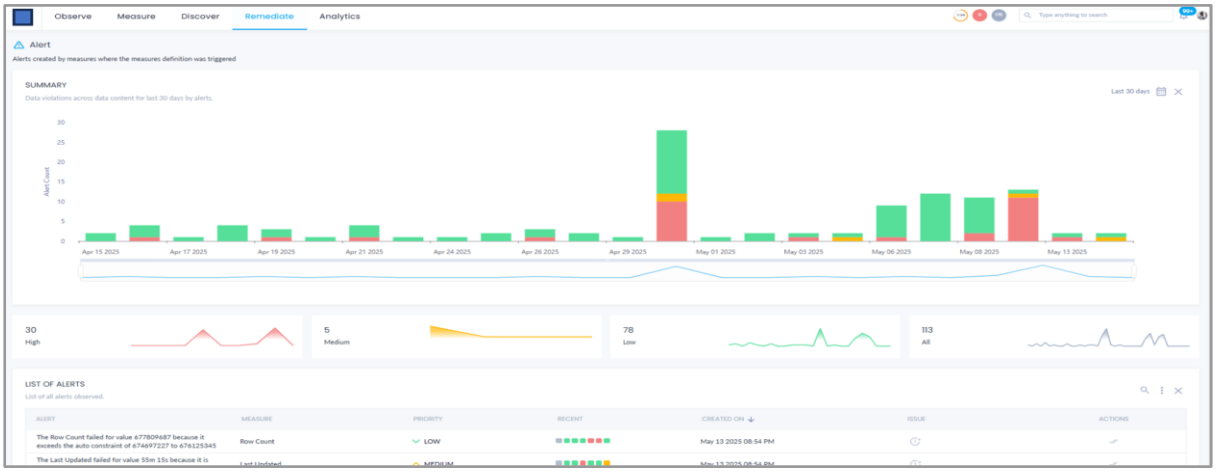
Alerts & Notifications

erwin DQ provides proactive alerts and real-time notifications to inform users of data quality issues and anomalies as soon as they occur. These alerts help users take immediate action and maintain trust in their data. This chapter covers where to access, interpret, and manage these alerts, and how to configure personal alert preferences.

Managing and Interpreting Alerts

Where do I access alerts, and how do I interpret them?

- Go to the **Remediate** module from the top header menu.
- Select the **Alerts** tab to view a list of current and past alerts.
- Each alert includes:
  - **Alert Type** (Threshold breach, Profiling failure, Issue creation)
  - **Associated Asset or Rule**
  - **Timestamp and Severity** (High/Medium/Low)
  - **Linked Measure or Issue ID** for quick triage
- Click an alert to view detailed context and actions taken or pending.



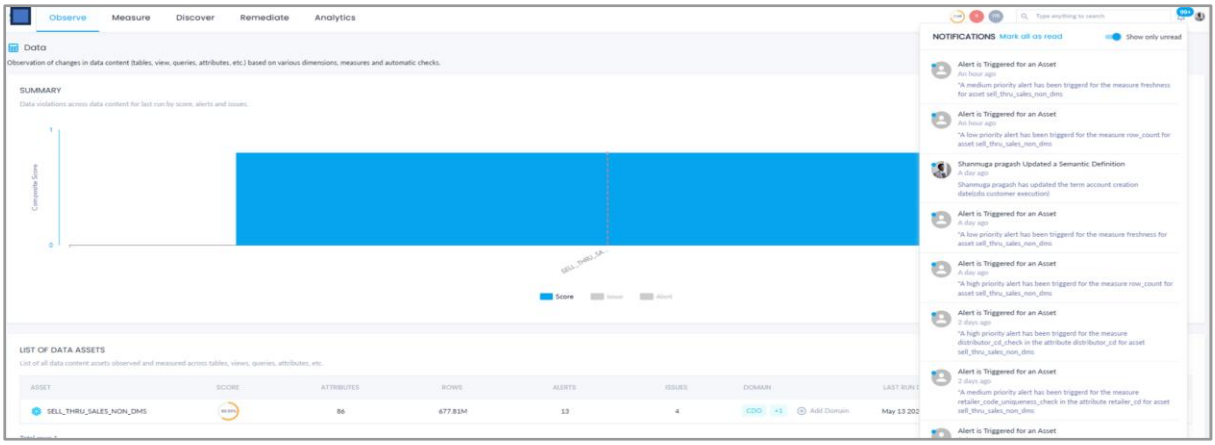
Alerts help users identify which rules or assets need attention and reduce the time to resolution.

- Alert types explained:

erwin DQ supports multiple delivery channels for alerts:

- **Email Alerts**:
  - Sent to users based on platform notification settings and rule ownership.
  - Include direct links to failed measures, issues, or affected assets.
- **Google Chat / Slack / MS Teams**:
  - Real-time notifications can be pushed to collaboration platforms.
  - Useful for team-based monitoring and triage.
- **Platform Notifications**:
  - Viewable via the **bell icon** in the top header bar.

- Includes success/failure of scheduled jobs and rule execution status.



Issue Management and Remediation

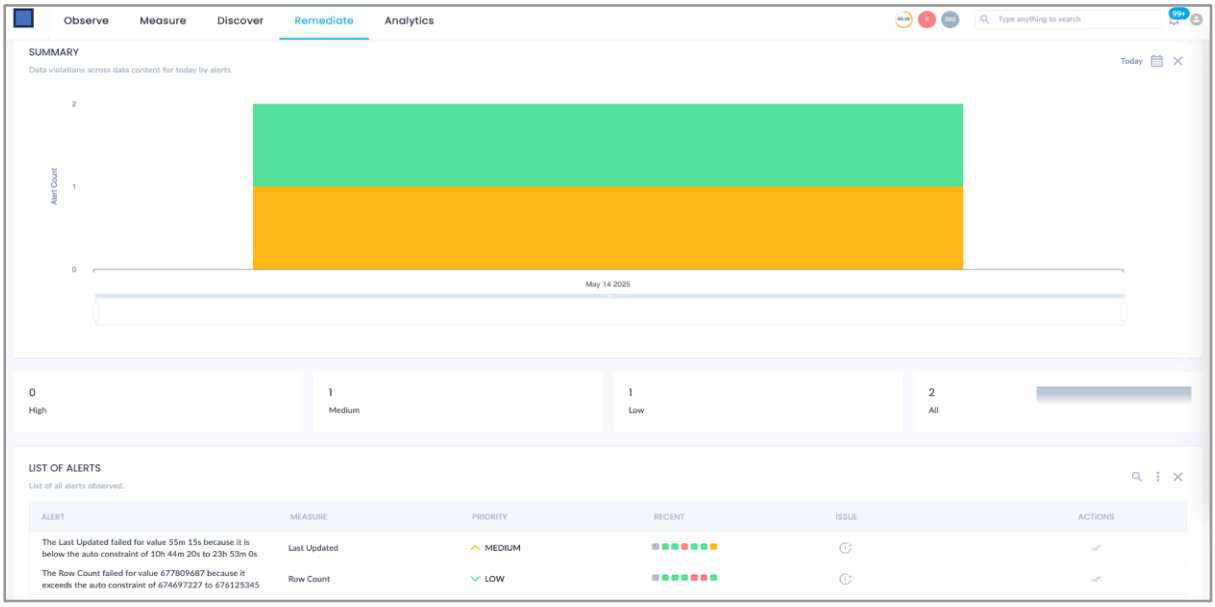
erwin DQ provides a centralized issue management framework to help users log, track, and resolve data quality issues. Business Data Stewards typically raise and prioritize issues, while Data Engineers perform root cause analysis and remediation. This chapter outlines the complete lifecycle of issue handling in DQLab.

Raising Data Issues

How and when do I raise a data issue in the platform?

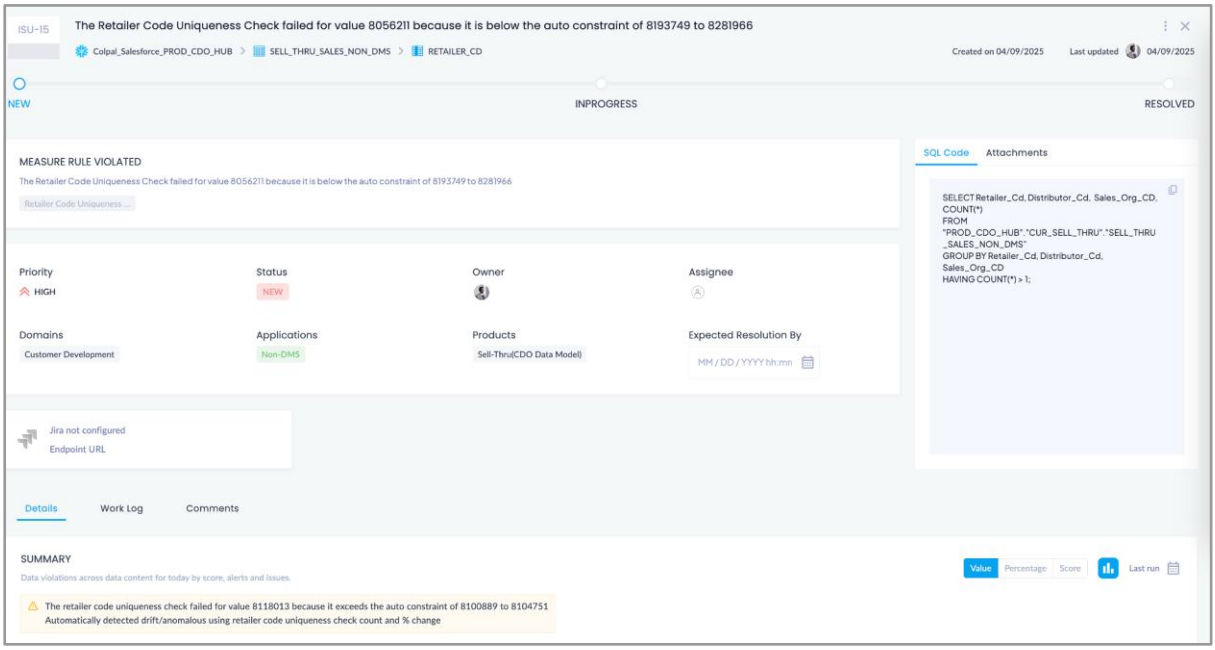
Data issues are raised when a rule breach is detected or during profiling/validation of CDEs.

- Navigate to the **Alert** module under Remediate > Alert or from the measure page
- Click "**Create Issue**".



Data Stewards raise issues when they detect quality deviations or suspect data problems that require remediation.

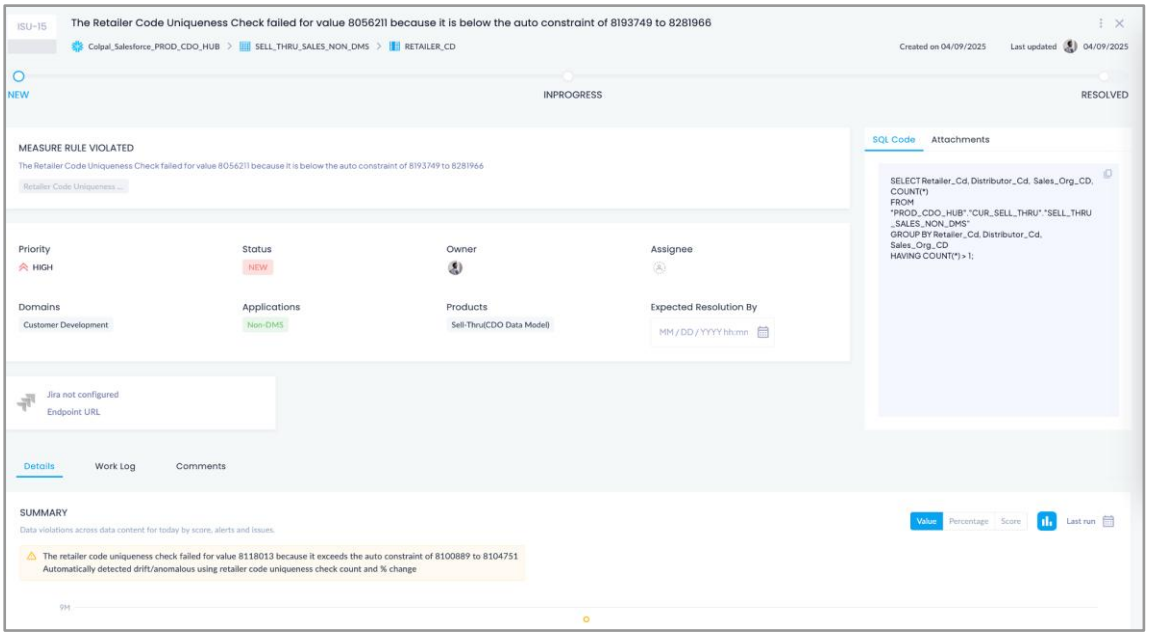
- Assign the issue to the respective Data Engineer, who is responsible for fixing these issues
- Monitor and track the issue until resolved
- Once resolved, verify if the scores are updated and invalid records are reduced from the previous run



Managing & Tracking Issues

Where do I view and manage assigned issues (Platform or Jira)?

- Go to **Remediate** > **Issues**.
- View the list of open, in-progress, and resolved issues.
- Use filters for Domain, Status, Assigned To, and Priority.
- Click any issue to open the detail view and history.
- If Jira integration is enabled:
  - A Jira ID will be linked to the issue.
  - You can track updates directly from erwin DQ or within Jira.



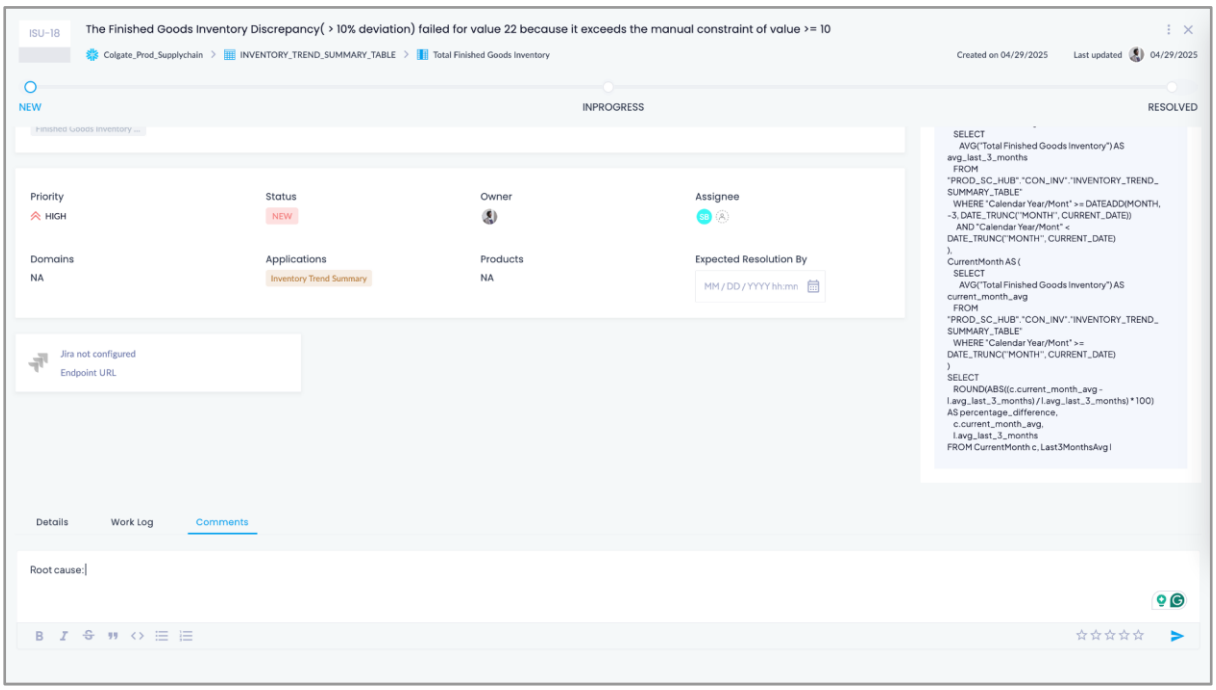
erwin DQ also supports external ticketing system integration (e.g., Jira), ensuring teams can use their existing workflows for tracking and collaboration

Root Cause Analysis & Resolution (Data Engineers)

How do I triage and resolve data issues as a Data Engineer?

Data Engineers play a key role in diagnosing and resolving root causes.

- Receive alerts or assigned issues from Stewards.
- Review the issue context, failed rule logic, and profiling results.
- Identify the upstream source or pipeline failure.
- Resolve the issue at the data source or pipeline level.
- Update issue status and add resolution notes.



- SLA management and best practices for resolving issues on time
- Issues are categorized by **Priority** (High, Medium, Low).
- SLAs can be configured at the Domain level (e.g., resolve high priority within 1 day).
- Regularly monitor issue dashboards and alerts.
- Collaborate with data owners and platform admins to validate fixes.

Adhering to SLAs helps maintain trust in data quality processes and supports proactive governance

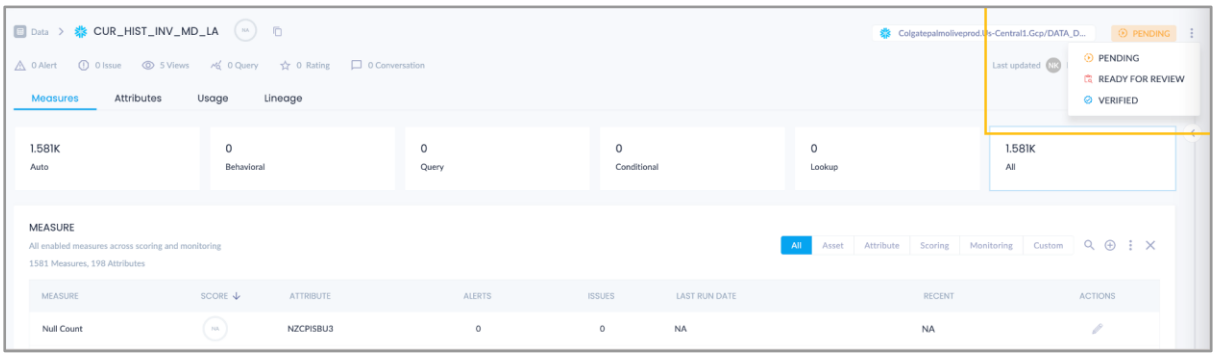
Certifying and Approving Data Assets Playbook

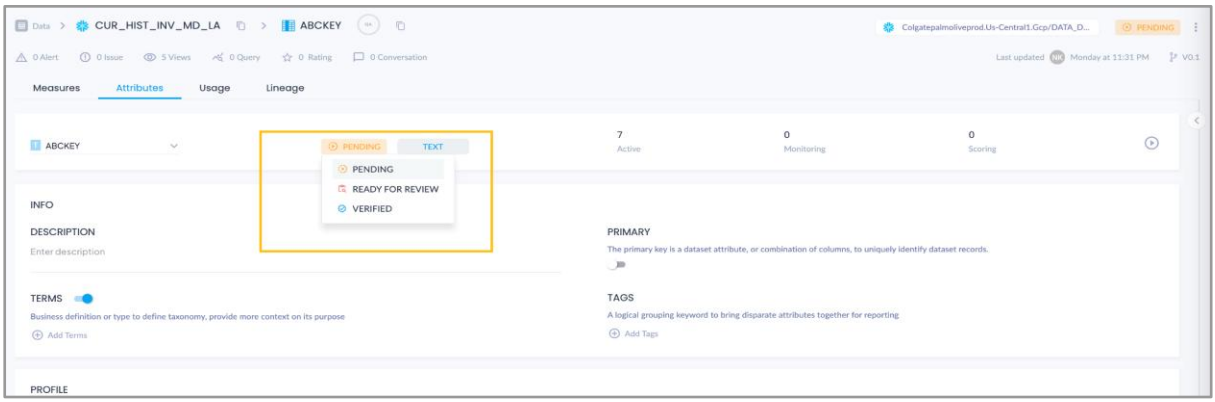
Once data assets and Critical Data Elements (CDEs) have successfully passed profiling and rule validations, they can be certified in erwin DQ. Certification provides an added layer of trust and signifies that an asset meets the required data quality standards. This chapter outlines how to certify assets and explains the meaning of certification and approval statuses.

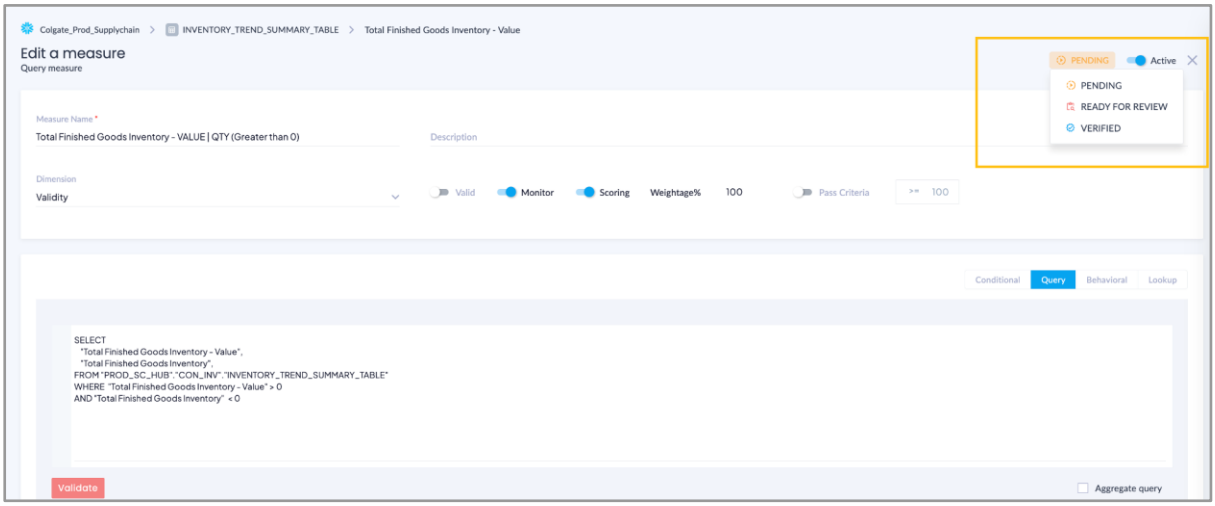
Understanding Approval Status & Impact

Each attribute and asset may have a status indicating where it stands in the certification and approval lifecycle.

- **Key Status Types (The status applies at all 3 levels – Asset, Attribute and Measure)**:
- **Verified**: The asset or attribute has been approved after passing validation.
- **Pending**: The asset is not yet certified; further review or rule execution is needed.
- **Ready for Review**: The asset is under assessment by a steward or domain owner.
- **Deprecated**: The asset is outdated or no longer valid for use.







Why It Matters:

- Verified data assets are prioritized in dashboards, reports, and scorecards.
- Approval status ensures that only reviewed, high-quality assets are used for analytics and reporting.
- Helps teams build trust and ensure governance alignment.

Certifying assets reinforces governance best practices, streamlines approval processes, and establishes confidence in the quality and reliability of key data assets.

Dashboarding & Reporting

erwin DQ offers powerful dashboarding capabilities to visualize data quality metrics and observability trends. Custom dashboards allow users to monitor specific assets, domains, or rule performance and share actionable insights with stakeholders. This chapter covers how to create, customize, and distribute dashboards across teams.

Creating Custom Dashboards

How do I create dashboards to monitor my domain's data quality?

- Navigate to the **Analytics** module from the top header menu.
- Click "**Dashboards**" and select "**Create Dashboard**".

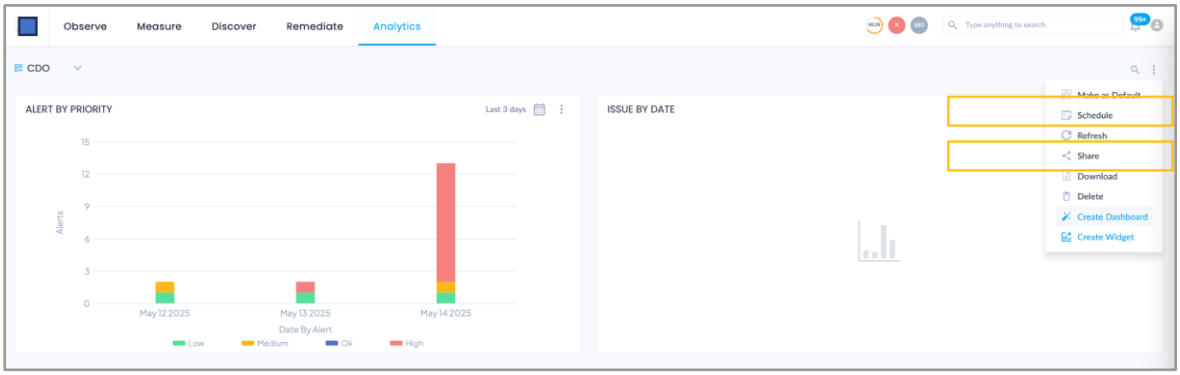
- Provide a name, description, and assign a domain or visibility level (Private or Shared).
- Choose from available widgets to populate your dashboard:
  - Data Quality Scores (by domain, asset, or rule)
  - Trend charts (pass/fail rates over time)
  - Alert summaries
  - Issue status by priority or owner
- Drag and arrange widgets on the canvas as needed.
- Click "**Save**" to publish the dashboard.

Tip: Create separate dashboards by persona (e.g., Stewards vs. Engineers) for role-specific visibility.

Sharing Insights with Stakeholders

How do I share dashboards with team members or leadership?

- Open the desired dashboard from **Analytics** > **Dashboards**.
- Click the **Share or schedule icon** (top-right corner of the dashboard).
- Choose sharing options:
  - Share with individual users or roles
  - Public within organization
  - Export to PDF or Image for offline sharing
- Use filters or time-frame selectors before exporting to tailor the snapshot.
- Shared dashboards will be accessible from the recipient's dashboard list.



Shared dashboards enhance collaboration and ensure everyone has a unified view of data quality performance.

Dashboards in erwin DQ provide real-time visibility into key data health indicators and help users stay informed, aligned, and proactive in managing data quality.

Terminologies & Definitions (Quick Reference Guide)

This chapter serves as a quick reference glossary of key terms used throughout the erwin DQ platform. Understanding these terms ensures consistency in communication, smoother platform navigation, and better alignment between business and technical users.

Key Platform Terms Explained

- **Asset** - A data source or table ingested into erwin DQ for analysis. Assets can include data tables, files, pipelines, reports, etc.
- **Attribute** - A column or field within an asset. Attributes are individually profiled, tagged, and monitored with data quality rules.
- **Domain** - A logical grouping of assets based on business context, such as Finance, Sales, Customer, or Operations.
- **Product** - A semantic layer used to represent a data product, curated dataset, or group of datasets designed for a specific analytical or operational use case. Data Products often serve specific consumer needs and are managed with associated SLAs, ownership, and quality metrics. This helps in aligning data management with product thinking and domain-driven design principles.
- **Application** - A semantic layer used to represent applications or systems that generate or own the data. Helps with organizing assets for SLA, lineage, and issue tracking.
- **Pipeline** - A data processing flow from source to target system. erwin DQ monitors pipeline reliability and observes performance and volume metrics.
- **Report** - A structured output, usually from BI tools, that is tracked and observed for data freshness and delivery.
- **Lookup Table** - A reference dataset used to validate or enrich incoming data in rules. Lookups can be linked to OOB or query-based rules.
- **Comparison Measure** - A rule that compares values across columns or datasets. Used to validate alignment between source and target systems.
- **Standalone Measure** - A rule not attached to a specific asset or attribute but used to evaluate broader or cross-asset conditions.
- **Critical Data Element (CDE)** - An attribute identified as business critical. These fields are prioritized for data quality rule application and monitoring.
- **Data Quality (DQ) Dimension** - A category for classifying a DQ rule. Common dimensions include:
    - **Completeness**: Ensures required values are present.
    - **Uniqueness**: Ensures values are not duplicated.
    - **Validity**: Ensures data adheres to expected formats or ranges.
    - **Consistency**: Ensures data aligns across systems or sources.
    - **Accuracy**: Ensures data reflects real-world values.
- **Profiling** - Automated analysis of attributes to understand distribution, nulls, patterns, and uniqueness. Used to inform rule creation.
- **Semantic Tag** - A label applied to attributes to classify their role (e.g., PII, Financial). Managed by administrators for consistent governance. Semantic tags help standardize attribute classification across domains and are used in reporting, filtering, and SLA configuration.

- **Semantics** - The semantic layer in erwin DQ organizes and governs metadata across **Domains**, **Applications**, and **Products**. Semantics define context-driven categorization, ownership, SLA targets, and issue routing strategies across the organization. It is also where semantic tags and business terms are configured, reviewed, and managed. A label applied to attributes to classify their role (e.g., PII, Financial). Managed by administrators for consistent governance.
- **Business Term** - A glossary term with a formal business definition. Mapped to attributes to clarify purpose and standardize usage.
- **Workflow** - A platform capability that enables users to automate issue creation, notifications, and task execution based on alert conditions. Workflows reduce manual intervention by linking alert triggers to tasks such as ticket creation or messaging via collaboration tools. They can be configured per domain, product, or priority.
    1. Workflow Setup Steps:
    2. Navigate to **Settings** → **Platform** → **Workflow**
    3. Click the "+" icon to create a new workflow
    4. Drag and drop events to configure the workflow visually
    5. Apply trigger conditions to tasks
    6. Save the workflow to activate it

Execution logs can be reviewed after task execution, and role-based permissions govern access to workflow creation and management. A series of steps or logic is configured to automate data quality or data observability processes within the platform. Commonly used for rule chaining, scheduling, and trigger-based execution.

- **Lineage** - A visual and metadata-driven representation of how data flows from source to target, including transformations and dependencies.
- **Import / Export** - Features used to bulk import metadata, measures, or users and export metadata, measures, or users' data from the platform.
- **Impact Analysis** - Assessment feature that evaluates which downstream processes or data assets are affected when changes occur to an upstream source or rule.
- **Schedule** - Functionality to define and manage recurring profiling or rule execution jobs, based on a defined time or event.
- **Libraries** - A centralized repository in erwin DQ to manage shared business rules, SQL queries, or templates that can be reused across domains or teams.
- **Measure** - A data quality rule applied to an attribute or dataset. Measures evaluate data against configured thresholds.
- **Out-of-the-Box (OOB) Measure** - Predefined rule types available in erwin DQ, such as completeness or length checks.
- **Conditional Rule** - A rule that evaluates an attribute based on specific criteria or conditions (e.g., "If status is active, email must not be null").
- **Behavioral Rule** - A rule that tracks expected behavior of data over time (e.g., weekly trend in transactions).
- **Query-Based Rule** - A custom SQL WHERE clause that defines failed rows directly.
- **Threshold** - The pass/fail criteria for a measure, such as "Pass if > 95% complete".
- **Run Now** - An on-demand execution trigger for profiling, rules, or pipelines that initiates processing immediately rather than waiting for a schedule.
- **Alert** - A notification triggered by a failed measure, schema change, or volume anomaly. Alerts can be delivered via email, chat, or the platform.

- **Issue** - A logged and tracked event resulting from a failed rule or identified data problem. Issues can be manually or automatically created.
- **Triage** - The process of reviewing, assigning, and prioritizing issues for resolution.
- **Remediate** - The act of resolving a data issue, often performed by a Data Engineer at the source level.
- **SLA (Service Level Agreement)** - Time-based commitment to resolve data issues based on priority levels. Configurable per domain or application.
- **Connections** - Configurations that define how erwin DQ connects to various source systems or databases to pull data and metadata.
- **Observe** - erwin DQ module that monitors the health of assets based on freshness, volume, schema changes, and lineage.
- **Discover** - erwin DQ module is used to explore and manage assets, CDEs, tags, and business terms.
- **Measure Module** - Used to configure and manage data quality rules, track measure status, and schedule jobs.
- **Remediate Module** - Used to manage alerts, issues, assignments, and remediation workflows.
- **Analytics Module** - Used for creating dashboards and reports to visualize data quality trends, scores, and issue metrics.

This glossary should be used as a reference throughout the platform and during onboarding to ensure a consistent understanding of platform functions and terminology.