

I'm not a robot



Etl testing tools

ETL testing is performed before moving data into a production data warehouse system. It is also known as table balancing or production reconciliation. The main goal of ETL testing is to identify and mitigate data defects. I have spent 85+ hours researching 47 ETL automation testing tools to bring you the top 7. This well-researched and trusted guide highlights key features, pros and cons, and pricing. For a comprehensive understanding of these tools, read the full article. My professional insights offer the credible information you need to make an informed decision. Best ETL Automation Testing Tools: Top Picks 1) iceDQ iceDQ is a DataOps platform for Data Testing and Monitoring. I tested its rules-based auditing, which allows you to automate ETL Testing, Cloud Data Migration Testing, and Production Data Monitoring. It might be helpful to look into this for your projects. Features: Rule Creation: I could create rules to test transformations, duplicates, schema, and Type II Dimension. This ensures data integrity. Reporting: It lets me use the built-in dashboard or enterprise reporting tools like Tableau, Power BI, and Qlik for generating insightful reports. This may help in better decision-making. DataOps Integration: It enables DataOps by allowing users to integrate with any Scheduling, GIT, or DevOps tool. I benefited from using this seamless integration. Scalability: The highly scalable in-memory engine evaluates billions of records efficiently. It allowed me to handle massive datasets convincingly. Connectivity: It lets me connect to databases, files, API, and BI Reports with over 50 available connectors. This provides extensive integration options. Enterprise Product Integration: The tool integrates smoothly with different enterprise products like Slack, Jira, ServiceNow, Alation, and Manta. It is perfect for enhancing collaboration. Security: It comes with single sign-on, advanced RBAC, and encryption features, ensuring robust data security. Deployment Flexibility: Deploy it anywhere, whether on-prem, AWS, Azure, or GCP. It allowed me to choose the most suitable environment. iceDQ provides seamless integration with various ETL tools and databases Detailed error reporting helps identify and resolve issues quickly As per my experience, it offers an intuitive user interface for ease of use Initial setup and configuration can be time-consuming Requires regular updates to maintain compatibility with new technologies Pricing Plans: Free Trial: 30-day free trial Live Demo: Available Visit iceDQ >> 2) DataTrust During my evaluation, I found that RightData's DataTrust platform provides the tools to trust my data for decision-making. As a no-code data observability platform, it detects anomalies, generates business rules, and validates data. It works for both one-time migrations and ongoing data operations. Features: Data Reconciliation: It allowed me to seamlessly reconcile data across single or multiple sources for enhanced data integrity and consistency. Machine Learning Business Rules: I could effortlessly generate business rules using advanced ML algorithms. It allowed me to seamlessly accept, modify, or discard them as per my requirements. Table Comparison: DataTrust compares row counts and checksums across multiple tables at the schema level. It helped me ensure data accuracy. Bulk Data Validations: It offers optimized validation across tables using source compute without ingestion. I benefited from using this feature for large datasets. Executive Reporting: DataTrust offers interactive executive reports with quality insights and personalized drill-down filters. The tool's robust data quality assurance features helped me ensure my data remains intact and accurate Comprehensive data audit capabilities enhance compliance Continuous data quality control improves long-term data reliability Highly rated with strong user feedback Some users might find the learning curve steep I have noticed that occasional updates can sometimes disrupt my ongoing tasks Pricing Plans: Price: Contact salesperson Free Trial: 30-day free trial Visit DataTrust >> 3) BIG EVAL BIG EVAL automates testing tasks in data projects such as data analytics, data warehouses, ETL/ELT, and data migrations. I particularly appreciate the unique automation approach and simple user interface that ensure quick results. According to my review, it is perfect for achieving efficient outcomes quickly. BIG EVAL Integration: Zapier, Make, Script and WebhookReal-Time Reporting: YesSupported Platforms: WindowsFree Trial: 14 Days Free Trial Visit BIG EVAL Features: Autopilot Testing: Using the autopilot testing feature driven by metadata, I could leverage test coverage massively, ensuring comprehensive evaluations. Customizable Test Algorithms: It offers fully customizable test algorithms, rules, and test behavior, adapting to your specific needs. I could tailor the tests to my requirements. Best Practices Templates: A gallery with hundreds of best practices test case templates is ready to be used, simplifying the setup process. I could save much of my time by using these project templates. Insightful Dashboards: Deep insight analysis with clear dashboards and alerting processes informs you of all critical metrics. It offered me a great way to track progress easily. Tool Integration: Its integration with hundreds of tools (e.g., Jira, ServiceNow, Slack, Teams) ensures seamless workflow and collaboration. DataOps and DevOps Integration: Embeddable into DataOps processes and DevOps CI/CD flows, it supports continuous integration and delivery pipelines. Data Connectors: It has hundreds of connectors to various data sources (RDBMS, APIs, Flatfiles, Business applications, SaaS). These connectors ensure flexibility in handling different types of data. Covers comprehensive data quality testing and management Needs trusted by well-known brands, ensuring reliability I am really satisfied with the software, and the high user ratings back that up Free trials allow you to evaluate before purchasing Advanced features might require extensive training Troubleshooting complex issues might require dedicated resources Pricing Plans: Price: BIG EVAL LAUNCH monthly plan starts at \$89 Free Trial: 14-day free trial Visit BIG EVAL >> 4) QuerySurge QuerySurge is an ETL testing solution developed by RTTS. After analyzing it, I noticed it was built to automate testing for data warehouses and big data. According to my review, it ensures the extracted data remains intact in the target systems. Features: Data Quality & Governance: I could improve data quality and governance for more reliable and accurate datasets. Data Delivery Cycles: It lets me accelerate my data delivery cycles, reducing time to market significantly. Automated Testing: The tool helped me automate manual testing efforts, increasing efficiency and reducing errors. Platform Testing: It provides testing across different platforms like Oracle, Teradata, IBM, Amazon, and Cloudera, ensuring compatibility. Testing Speed & Coverage: QuerySurge speeds the testing process up to 1,000x, offering up to 100% data coverage. DevOps Integration: The tool integrates an out-of-the-box DevOps solution for most Build, ETL, and QA management software. Email Reports & Dashboards: It delivers shareable, automated email reports and data health dashboards for better insights. Automates data testing, increasing efficiency I love how it is optimized for data warehouses, Big Data, and BI reports Quickly identifies and analyzes data discrepancies Tailored solution for complex data environments Without dedicated support, integrating with existing tools was difficult for me Pricing Plans: Price: Full user SUBSCRIPTION for \$5,747 Free Trial: 15-day free trial Visit QuerySurge >> 5) Informatica Data Validation: Informatica Data Validation is a popular ETL tool. I reviewed and found it to be a top data testing tool that works with the PowerCenter Repository and Integration Services. It allows developers and analysts to create rules for testing mapped data. Features: Comprehensive Validation: Informatica Data Validation provided me with a complete data validation and integrity solution. Reduced Programming Effort: The tool's intuitive user interface and built-in operators reduced my overall programming efforts. Productivity Boost: It lets me identify and prevent data issues, increasing business productivity. SQL-Free Testing: The tool allows wizards to create test queries without writing SQL, which I found incredibly helpful. Design Library: This automated ETL testing tool offers a design library and reusable query snippets. Rapid Analysis: It can analyze millions of rows and columns of data in minutes, which I greatly benefited from using. Data Comparison: Informatica helps compare data from source files and data stores to the target data warehouse. Informative Reporting: The tool can efficiently produce informative reports, updates, and auto-email results. Robust ETL capabilities streamline data integration The scalable architecture fits both small businesses and enterprises These data quality tools have significantly improved my data integrity User-friendly interface simplifies complex tasks Some users find the licensing model expensive Integrations with newer SaaS applications can be limited Pricing Plans: Price: Contact the sales and support team Free Trial: 30-day free trial Download link: 6) QualiDI QualiDI helps clients cut costs, increase ROI, and speed up their time to market. I found that this ETL tool automates all parts of the testing lifecycle. In my experience, it is one of the best ETL automation tools that helps clients cut costs, boost ROI, and speed up time to market. Features: Data Quality and Integration: It helps you efficiently find bad and non-compliant data. QualiDL allowed me to perform comprehensive data integration testing across various systems. Test Cycle Management: Managing test cycles through dashboards and reports was made easier, offering me great insights. I found that the testing process worked seamlessly across multiple platforms. Automated Test Case Generation: Automated test case generation is available for direct mappings. Centralized Test Case Repository: A central test case repository allows test schedules for regression testing. Dashboard Reports: The test execution results are available in dashboards and reports at a click. Defect Tracking: Its built-in defect tracking and monitoring interfaces with a third-party defect tracking tool impressed me. Centralized platform for testing multiple ETL tools It improved accuracy and consistency in my data testing Facilitates integration across various enterprise systems Advanced automation capabilities streamline processes Requires specialized knowledge for optimal use May have limited support for non-ETL data testing Pricing Plans: Price: Contact support team to get the best quotes Free Trial: Not available Download link: 7) ETL Validator Datagapp ETL Validator is a Data warehouse testing tool. As I evaluated it, I found it is a top-rated tool for testing data migration, data integration, and data warehouse projects. It helped me with its built-in ETL engine that can compare millions of records. Features: Data Comparison: The tool lets me create a baseline for accurately comparing data to identify any differences. As a result, I could access precise data discrepancies. Automatic Data Validation: It lets me define rules for automatically validating data in every column in the incoming file. Database Schema Comparison: ETL simplifies the database schema comparison across different environments. Test Plan Assembly and Scheduling: It allowed me to assemble and schedule test plans with ease. Web-Based Reporting: I would rate it as one of the best data warehouse testing tools for web-based reporting. REST API and CI Features: The tool includes REST API and continuous integration features to streamline processes. Data Quality and Integration Testing: It offers robust data quality and data integration testing capabilities. Benchmarking Capabilities: ETL provides comprehensive benchmarking capabilities for performance analysis. Automates validation of ETL data flows with minimal manual intervention Provides detailed and comprehensive reporting for ETL processes Reduces overall testing time significantly As per my experience, it enhanced data quality with thorough validation checks High costs may not be suitable for small businesses Pricing Plans: Free Trial: 15-day free trial Live Demo: Available Download link: What is ETL Testing? ETL testing is a software testing process that ensures accurate data is loaded from a source to the destination after business transformation. It also involves verifying data at various middle stages used between source and destination. ETL stands for Extract-Transform-Load. If you are looking for solutions to streamline this process, you might be interested in exploring some of the best ETL tools available in the market. Why is ETL Testing required? While data transportation from extraction to loading, there are some chances of human or system errors that could result in poor information being transferred throughout the system. ETL testing helps ensure such errors do not occur and eliminates defects that can harm the whole system. Utilizing automated testing tools can further streamline this process and improve accuracy. How to choose the best ETL Automation Testing Tool? Choosing the best ETL automation testing tool is essential for efficient data integration. It helps you ensure data accuracy and consistency. Consider these factors to make an informed decision. Compatibility: Ensure the tool is compatible with your data sources and target systems. It might be helpful to check for support with databases, cloud services, and big data platforms. Usability: Aim to choose a tool with an intuitive interface. This may help reduce the learning curve for your team. Automation Features: Look for automation capabilities that streamline testing processes. The more features, the better. Scalability: Considering how well the tool can scale with your growing data needs is important. Pay attention to performance under load. Integration: Select a tool that integrates well with your existing CI/CD pipelines. This is typically required for continuous testing. Support and Documentation: One of the easiest ways to avoid issues is to choose a tool with robust support and comprehensive documentation. Cost: Keep in mind your budget and the total cost of ownership. It is best to weigh the features against the price. How to do ETL Testing? Below is a process on how to do ETL testing: Step 1) Identifying data sources and requirements Step 2) Data acquisition Step 3) Implement business logic and dimensional modeling Step 4) Build and populate data Step 5) Build reports Why trust Guru99? Guru99 provides detailed and precise content created by industry professionals. We are deeply committed to credibility, ensuring you get the most reliable and current information. Our rigorous content creation and review process make Guru99 a trusted source for thorough guidance. Explore our editorial policy. Verdict If you are deciding on the best ETL automation testing tools, check my verdict. In my experience, the best way to resolve issues and ensure smooth data integration is by leveraging robust and reliable ETL tools. iceDQ: It is an ideal ETL testing tool offering comprehensive data validation and testing solutions. The robust features and stable performance makes it a reliable option for enterprise-level data integration projects. DataTrust offers features that can assist in this area. Its ability to simulate hundreds of thousands of concurrent users allows you to test your ETL systems under significant stress, identifying potential performance issues. The tool's integration with continuous integration (CI) pipelines enables automated performance testing, ensuring that your ETL processes maintain efficiency as they evolve. Additionally, WebLOAD's support for various protocols and environments provides flexibility in testing diverse ETL scenarios.Standout features & integrations:Features include an integrated development environment (IDE) for recording, editing, and debugging load test scripts, automatic correlation of dynamic values like session IDs, and real-time analytics with customizable dashboards for performance monitoring.Integrations include Selenium, Jenkins, AppDynamics, Dynatrace, Perfecto Mobile, New Relic, Ranorex, Original Software's Test Drive, Git repositories, and cloud services like AWS and Azure. LEARN MORE ABOUT WEBLOAD: Check out WebLOAD on their website Pros: Offers real-time performance analytics Integrates seamlessly with CI/CD pipelines Supports complex load testing scenarios Cons: Limited support for non-web protocols Resource-intensive during large-scale tests QuerySurge is an ETL testing automation tool aimed at testers, data architects, ETL developers, and BI analysts. It focuses on data validation and analysis, offering automated testing to improve data quality and optimize data processes.Why I picked QuerySurge: It specializes in continuous testing with features like automated testing across various platforms and a no-code/low-code solution. The tool enhances ROI by simplifying test creation and provides analytics for data optimization. It integrates easily with existing systems, making it versatile for diverse user needs. QuerySurge's focus on data quality improvement makes it a solid choice for continuous testing environments.Standout features & integrations:Features include automated testing across platforms, data quality improvement, and analytics for optimizing data. It also provides a no-code/low-code solution to enhance ROI and simplify test creation.Integrations include Apache Hadoop, Amazon Redshift, Microsoft SQL Server, Oracle, IBM DB2, SAP HANA, Salesforce, Snowflake, Teradata, and Google BigQuery. LEARN MORE ABOUT QUERYSURGE: Check out QuerySurge on their website Pros: Integration with existing systems Effective data quality improvement Extensive platform compatibility Cons: Limited customization options Requires technical expertise RightData is a comprehensive data product platform designed to transform raw data into actionable business insights. It caters to organizations aiming to enhance data quality, observability, and governance. RightData offers solutions for data integration, validation, and reconciliation, ensuring data accuracy and reliability across various systems.Why I picked RightData: If your team seeks a self-service ETL testing tool, RightData's DataTrust platform could be a suitable choice. Its no-code interface allows your team to automate data validation and reconciliation without extensive coding knowledge. Features like continuous data quality monitoring and integration with CI/CD pipelines enable your team to maintain data accuracy efficiently. Additionally, DataTrust's AI-driven anomaly detection assists in proactively identifying data issues, empowering your team to address them promptly.Standout features & integrations:Features include automated data reconciliation to ensure data consistency across systems, business rule validation for enforcing data quality standards, and integration with defect tracking tools like JIRA to streamline issue management.Integrations include popular data platforms such as SAP, Salesforce, data lakes, Oracle, Snowflake, SQL databases, cloud data warehouses, NoSQL databases, Excel, and CSV. LEARN MORE ABOUT RIGHTDATA: Check out RightData on their website Pros: Integrates with CI/CD pipelines Supports a wide range of data sources AI-driven anomaly detection enhances accuracy Cons: Potential learning curve for complex features Limited advanced analytics capabilities ETL Validator is a data testing automation tool designed for enterprises needing to ensure data accuracy and compliance. It automates data reconciliation and ETL/ELT testing, making it valuable for businesses in regulated industries.Why I picked ETL Validator: It focuses on compliance checks with features like data quality monitoring and synthetic data generation. The tool supports low-code/no-code options, allowing teams to automate testing processes with ease. It integrates with major cloud platforms, enhancing its utility across various industries. The ability to automate big data testing simplifies compliance for complex data environments.Standout features & integrations:Features include data warehousing automation, migration testing, and big data testing capabilities. It also offers data quality monitoring and synthetic data generation to improve compliance efforts.Integrations include Snowflake, Databricks, Azure Synapse, Salesforce, Power BI, Tableau, Oracle Analytics, Amazon Redshift, Google BigQuery, and IBM DB2. LEARN MORE ABOUT ETL VALIDATOR: Check out ETL Validator on their website Pros: Integrates with major cloud platforms Extensive data quality monitoring Supports low-code/no-code automation Cons: Limited user community Learning curve for complex features Rivery is a data integration platform designed for cloud-based applications, targeting data engineers and analysts. It automates data workflows and pipelines, enhancing data management and analysis across cloud environments.Why I picked Rivery: It excels in cloud data integration with its ability to manage and automate data workflows. The platform supports real-time data ingestion, making it suitable for dynamic data environments. It also offers a user-friendly interface for building and deploying data pipelines. Rivery's scalability is ideal for teams looking to grow their data capabilities without extensive infrastructure investments.Standout features & integrations:Features include real-time data ingestion, an intuitive interface for pipeline creation, and automation of data workflows. It also allows for easy scaling of data operations to match business growth.Integrations include Amazon Redshift, Google BigQuery, Snowflake, Microsoft Azure, Salesforce, HubSpot, Facebook Ads, Google Ads, Shopify, and MySQL. Check out Rivery on their website Pros: Cloud-native design Automation of data workflows Real-time data ingestion Cons: Requires cloud expertise Limited offline capabilities TestBench is a test management tool designed for agile teams and quality assurance professionals. It facilitates test planning, execution, and evaluation, offering solutions that enhance software quality and streamline testing processes.Why I picked TestBench: It caters to agile environments with features like keyword-driven and data-driven testing, which simplify test case creation. The tool supports low-code/no-code automation, making it accessible for teams with varying technical skills. Its integration with existing systems boosts efficiency in managing test tasks. TestBench's robust variant management and versioning capabilities ensure comprehensive coverage for systems with multiple variants.Standout features & integrations:Features include test automation with flexible execution control and efficient test design, allowing for reuse of data and test steps. It also provides convenient evaluation options for test planning and process mapping.Integrations include connections with requirements and defect management systems, enhancing overall testing efficiency. LEARN MORE ABOUT TESTBENCH: Check out TestBench on their website Pros: Variant management and versioning Low-code/no-code automation options Supports agile testing methodologies Cons: Requires technical expertise for advanced features May need customization for complex setups Codoid's ETL Testing Services specialize in validating data extraction, transformation, and loading processes to ensure data accuracy and integrity. These services cater to businesses seeking to maintain high-quality data for informed decision-making. The primary functions include data mapping validation, transformation logic verification, and data quality checks throughout the ETL lifecycle.Why I picked Codoid's ETL Testing Services: If your team requires comprehensive testing without the hassle of managing it internally, Codoid offers a tailored solution. They provide automated metadata testing, checking aspects like data length and type, which helps maintain consistency. Their application upgrade testing ensures your data warehouse stays compliant with the latest standards. Additionally, their data completeness testing verifies that all data is accurately loaded, reducing the risk of missing information. Standout features & integrations:Features include data quality testing to prevent issues like invalid characters, metadata testing to ensure data type and length accuracy, and production data validation to confirm business information is correct and consistent.Integrations include various data sources and platforms, allowing seamless testing across different environments. Pros: Advanced reporting capabilities Automated testing reduces manual effort Comprehensive data validation processes Cons: May require onboarding time to align processes Potential communication challenges with remote teams SSIS Tester is a framework built for unit and integration testing of SQL Server Integration Services (SSIS) packages. It targets developers and QA teams who need to ensure the functionality and reliability of their data integration tasks.Why I picked SSIS Tester: It's tailored for SQL Server users with features like integrated data taps and fake source/destination components, which eliminate external dependencies. The tool supports test-driven development using Visual Studio and ReSharper. It also provides a project template for easy test implementation. Its compatibility with .NET languages makes it accessible for developers with varying skill levels.Standout features & integrations:Features include real-time monitoring and reporting of test executions, exportable results to HTML, and support for SQL Server 2016. It also offers a ready-to-use project template and facilitates test-driven development with Visual Studio.Integrations include SQL Server 2016, SQL Server 2014, SQL Server 2012, SQL Server 2008, and .NET Framework 3.5/4.0. LEARN MORE ABOUT SSIS TESTER: Check out SSIS Tester on their website Pros: Easily integrates with Visual Studio Supports test-driven development Tailored for SQL Server Cons: Dependency on .NET Framework Limited to SQL Server environments Informatica PowerCenter is an enterprise data integration platform designed to facilitate the extraction, transformation, and loading (ETL) of data across various systems. It serves organizations seeking to manage complex data integration initiatives, such as analytics, data warehousing, application migration, and data governance. Why I picked Informatica PowerCenter: If your team requires a scalable solution to handle large-scale data integration projects, Informatica PowerCenter offers features that align with these needs. Its role-based tools promote collaboration between business and IT teams, enabling self-service data access and delivery. The platform's graphical, code-less tools come with pre-built transformations, enhancing reusability and automation. Additionally, its universal connectivity supports integration across diverse data sources and targets, ensuring comprehensive data management. Features include rapid prototyping, profiling, and validation capabilities that allow for quick iteration and validation of data integration processes. The platform also offers business and IT collaboration tools, facilitating agile processes and self-service data access. Furthermore, Informatica PowerCenter provides universal connectivity, enabling integration with a wide range of data sources and targets. Integrations include Oracle, Microsoft SQL Server, IBM Db2, SAP, Salesforce, Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform, Hadoop, and Teradata. Pros: Offers extensive customization options Scales effectively for large data volumes Supports complex data transformations Cons: Requires significant system resources Complex initial setup and configuration