precisely

Data360 Analyze

Release Notes

3.12.3

Table of contents

1. Setup requirements

	Server	Desktop

System requirements	The minimum hardware specification for the server product is as follows:
	
	• 32GB minimum (Web Server - 12GB, Database server - 8GB, Execution Server - 4GB base + 2GB per core)
	
	• Minimum 4 cores for a server instance
	
	Note: As the number of users and/or the number of scheduled jobs increases, you should look to increase the number of cores and thus memory.
	
	• At least two separate drives or partitions:
	  ◦ One for the Data directory.
	  ◦ One for temporary execution data.
		The minimum hardware specification for the desktop product is as follows:
		
		• 8GB RAM.
		
		• Intel Core i5 or 4-core equivalent processor minimum (i7 recommended).

	Server	Desktop
System requirements	For desktop installations and installations on Windows Server, you must install the required Visual C++ Redistributable Packages (choose the x64 variant of the .exe file): • Visual C++ Redistributable Packages for Visual Studio 2008 • Visual C++ Redistributable Packages for Visual Studio 2013 For Windows Server installations, you must also install: • Visual C++ Redistributable Packages for Visual Studio 2019 **Note:** The Power R node and the nodes in the Statistical and Predictive Node Pack process data in-memory. Additional RAM will be required when processing data sets with a large volume of data. Similarly, if the R node is used, the machine hosting the R environment must have sufficient available RAM to process the data.	
Supported platforms	The following operating systems are supported on the Data360 Analyze server product: • Windows Server Core 2019 • Windows Server 2019 (Desktop Experience) • Windows Server 2016 64-bit (Server with Desktop Experience) • RedHat Enterprise 7.X 64-bit • RedHat Enterprise 8.X 64-bit Requires the `libnsl-2.28-151.el8.x86_64` RPM package or later, and the `compat-libgfortran-48` package. The following browsers are supported on both server and desktop: • Google Chrome - Latest stable version • Microsoft Edge - New Chromium-based version	The following operating systems are supported on the Data360 Analyze single-user desktop product: • Windows 10 64-bit

	Server	Desktop
App server databases	Postgres	H2
Authentication servers (server only)	• Active Directory • OpenLDAP	N/A
Authentication protocols (server only)	• LDAP • LDAPs	N/A
Single Sign-On (server only)	SAML 2.0	N/A
App servers	Tomcat 9.0.83	

Execution Sandboxing

In order to use the Execution Sandboxing feature, the Analyze server must already have an appropriate container runtime. On Windows, Mirantis Container Runtime (formerly Docker Enterprise) is required to use the Execution Sandboxing feature. Data360 Analyze has been verified on Windows server with the Execution Sandboxing feature against Mirantis Container Runtime version 20.10.9. In order to use the Execution Sandboxing feature on Linux, docker-ce is required. Data360 Analyze has been verified on Linux with the Execution Sandboxing feature against docker-ce version 20.10.17.

2. Latest release - 3.12.3

This document provides release information for all releases of Data360 Analyze from version 3.12.0 onwards. If you require release information for an older version of Data360 Analyze, please contact us at support.precisely.com.

For information on installing Data360 Analyze, please see the installation guide. Once you are up and running, you can access the integrated product help by clicking the Help button, positioned towards the top right corner of the screen.

The availability of the following new features, enhancements and corrected issues is dependent on the installed edition of the product and licensed features.

2.1 What's new

Upload Tableau node

A new "AsyncUpload" property has been added to the Upload Tableau node.

Setting this property to "True" enables you to upload files asynchronously. The node output provides the job URL that you can use outside of the node to get the status of the async file upload.

By default, this property is set to "False" and files are uploaded synchronously.

Publish to DQ+ node

The previously experimental Publish to DQ+ node is now fully supported.

New Proxy and Retry connection properties have been added to the node.

ERP nodes

Updated the ERP nodes to support long ("Int8", ABAP Type "8") and large floating point ("DECFLOAT34", ABAP type "e") types which are present in newer SAP systems.

These changes require at least SAP JCO library version 3.1.8 to be installed. The ERP nodes have been verified against SAP JCO 3.1.8.

When upgrading a Data360 Analyze instance which was using the ERP nodes, the JCO library in use will also need to be updated. The locations and file names for the libraries required by the node have not changed.

Error messages

The text for all unexpected error messages returned by REST APIs, and thus by the UI, can be configured to be redacted for security purposes. The error codes for the errors will still be reported, but the error messages themselves will be replaced by a message indicating to check the application logs. Errors for data flow execution related problems will still be reported correctly to the Errors panel.

To enable redaction of all error messages, set the following in the <site-dir>/cust.prop file and restart the web application:

```
ls.framework.ls.web.rest.restErrorMessageRedactorName="allText"
```

Note that by enabling this feature, some UI features, especially with regard to configuration, may be affected and require the ability to consult the application logs to diagnose problems, for example authentication configuration.

Third party library updates

A number of third party libraries have been updated to take advantage of any security and feature updates, including:

- Amazon Corretto JDK - 11.0.21.9.1

- PostgreSQL Database - 13.13

- Tomcat - 9.0.83

Other changes

- Updated a number of HTTP headers to improve application security.

- Added `laeConfig` subcommand for managing key pairs in the secure store.

- When using execution containers, Data360 Analyze will wait longer by default for the container to start. You can now configure the retry and wait options.

- Updated to new version of General SQL Parser (GSP) to fix parsing issues in the DeepSQL node.

- Added a keep alive page at /html/keepalive.html which renders the text "alive" and can be used to ensure that the web application is running.

- Robots.txt is now accessible from the root path /robots.txt.

2.2 Corrected issues

Issue summary	Issue number
Fixed an issue that prevented you from enabling or disabling a schedule that was created by another user.	LAE-32009
Fixed an issue that prevented the Excel File node from importing files with unsupported characters in the headers.	LAE-32026
Fixed an issue that caused file descriptor leaks on Linux when using a non-default property to specify socket communications.	LAE-31974
Fixed an issue where setting invalid IdP Metadata in the SAML SSO authentication settings returned an uninformative error message.	LAE-31911
Fixed an issue that in some cases caused null pointer exception errors during the execution of a data flow.	LAE-31990
Fixed an issue where the input field on the "Rename" dialog was blank and was not pre-populated with the current name of the object when renaming a data flow, library node or folder from the directory page.	LAE-32462
Fixed an issue where it was not possible to exactly horizontally align a connection point pin to another pin or bend point.	LAE-32461
Fixed an error that occurred when you dropped a node on a data flow, ran the node, then tried to open the pop-out data viewer before saving the data flow.	LAE-32435
Fixed an issue where the Download Logs menu option was not available from the node context menu in read-only mode.	LAE-32431
Fixed an issue where clicking an error on the Errors panel did not navigate to the node on the canvas when the node was at a different composite level to the current view.	LAE-31772
Fixed an issue where nothing was displayed in the node log dialog if an error occurred when retrieving logs that no longer exist.	LAE-30613
Fixed an issue where, if the Download for SharePoint 365 node was run on a Linux environment, then the "_Output" field was incorrectly returning Windows style paths.	LAE-32167
Fixed an issue that sometimes prevented the Upload Tableau node from publishing the datasource. Now, all hyper files should be consistently published.	LAE-31965

Issue summary	Issue number
Fixed an issue where SharePoint 365 nodes created in 3.8.5 caused errors after upgrading to 3.10.2.	LAE-30539
Fixed an issue where specifying multiple email addresses in the Schedule "On Failure" notification list could cause no emails to be sent for some SMTP servers.	LAE-31906
For Data360 Analyze instances with a very large number of historical runs, the scheduled backup process could possibly fail with an OutOfMemory error. This has been fixed.	LAE-32432
Fixed an issue where the Database Profiler node could not be executed due to missing library entries in its Classpath property.	LAE-31785
Fixed an issue where some scheduled runs would not be executed when the system was under high load with lots of schedules starting at the same time. The missed scheduled runs were not visible in the system.	LAE-32165
Fixed an issue where the use of certain characters in node outputs and connection points would lead to incorrect results or errors on connection filters applied against those names. For instance, applying a connection filter to select an output with parentheses in its name would incorrectly select data sets which matched the name without parentheses.	LAE-32080
Fixed an issue that caused the Azure Datalake Storage Delete, Azure Datalake Storage Get and Azure Datalake Storage Put nodes to stop responding when authenticating.	LAE-31881
Fixed a Jython installation issue. The Java that is packaged with the install is now configured in the Java home environment variable for any processes created from the installer.	LAE-31809
Fixed an issue where the license file could not be accessed when setting up a multi-user configuration.	LAE-31771
The Execute Data Flow node now outputs the full path to the run state in the RunState field, not just the run state name.	LAE-27698
Fixed an issue that could cause backups to fail.	LAE-32268
Fixed an issue that caused connection errors on Azure Datalake Storage nodes.	LAE-32191

Issue summary	Issue number
Fixed an issue where file patterns on the Azure Datalake Storage List node incorrectly applied to the entire file path, rather than applying to the name only.	LAE-31845
Fixed a timezone issue in the Jython script where the datetime calculations were not returning the correct time offset.	LAE-30042
Fixed an issue where the Azure Datalake Storage List node was not configured correctly for format detection for CSV parsing.	LAE-31788
Fixed an issue that prevented you from unlocking documents locked by a different user.	LAE-30967

3. New in 3.12.2

3.1 What's new

Parquet - ADLS Data Lake File Metadata Reader

Azure Data Lake Storage List node now has the ability to identify file types by choosing the "Identification" option on the "MetadataMode" property,

The node can also harvest metadata from both CSV and Parquet files by choosing the "Structure" option from the "MetadataMode" property.

New Filtering options are also available on the node to allow the user to include or exclude certain files or directories.

SharePoint 365 Nodes

A new method of filtering has been added in the Download for SharePoint, List for SharePoint and Delete for SharePoint 365 nodes that allow for file/ folder retrieval based on values entered in the new properties:

- FileName

- FileNameMatchingCriteria:

- Equals

- StartsWith

All SharePoint 365 Nodes nodes now have three new output fields _SitePath, _DocumentLibrary and _Folder, which, along with _FileName allow users to connect to a downstream SharePoint 365 node using the _from field property options to drive what gets listed/downloaded/deleted/uploaded etc...

The following properties are deprecated:

- Pattern property on Download for SharePoint, List for SharePoint and Deleted for Sharepoint nodes

- We recommend using the new FilaName and FileNameMatchingCriteria properties instead, which will give you stable and predictable results

- ResourcePath property on the Download for SharePoint node

- ServerPath property on the Upload for SharePoint Node, its functionality replaced by the new FolderPath property

Additionally, on the Upload for SharePoint Node, the Input group was replaced with 2 new groups:

- Source group that will contain the LocalPath and Pattern properties

- Target group that will contain the FolderPath (a new property replacing ServerPath) , ServerPath (deprecated and hidden) , SharePointUrl, SitePath, and Document Library properties

The nodes output a _ServerRelativeUrl field, this is being marked as deprecated and will be removed in a future feature release, as it currently provides inconsistent results between listed folders and files.

Other changes

- The Database Profiler now supports profiling of Azure Synapse databases.

- The TERR version used by Statistical and Predictive Analytics Nodes is upgraded to version 6.0.3. Note that while these nodes will work as before, scripts and packages used by the Power R may have to be updated.

3.2 Corrected issues

Issue summary	Issue number
Fixes an issue where you are not able to select text in the code editor when the Review tab is selected	LAE-30859
Fixes an issue whereby Govern paging in the HTTP node would fail if the response that no items.	LAE-30865
Fixes an issue whereby the user received a permission denied error when they have selected "Search Sub-Folders" when Editing Library Paths for a data flow and they do not have access to the sub folder.	LAE-30873

Issue summary	Issue number
Fixes an issue where the Scheduled Run Cleanup process could hang if a scheduled run was locked.	LAE-30880
Fixes an issue where the Scheduled Run Cleanup process could hang if a scheduled run was locked.	LAE-30997
Fixed an issue where when using execution sandboxing where the execution containers could be brought down by the container cleanup task immediately after they were created. This was especially a problem when a schedule which brought up the container would start at the same time as the container cleanup task.	LAE-31011
Fixed an issue where issues on scheduled runs could occasionally result in deadlocks.	LAE-31381
Fixes an issue whereby property definitions for properties in a group were hidden but not deleted when the group was deleted.	LAE-31383
Fixes an issue whereby the adhoc runs clean up process could hang and prevent system backups from taking place.	LAE-31385
Fixes an issue whereby data flows that had been unlocked but not executed after unlocking could not be opened after upgrading to 3.12.0 or 3.12.1.	LAE-31403
Fixes an issue where an authorized user could access restricted parent directories when uploading, downloading or listing files.	LAE-31420
Fixes an issue whereby if jython nodes which were executed within the node container failed early on in their processing (e.g. compilation failure on the script in ConfigureFields in a Transform node), they were not releasing the file descriptors for the temporary data files (input and output) meaning that these temporary data files could not be deleted until the node container was restarted.	LAE-31449
Application now displays a user friendly message when unable to export because you don't have access to all dependencies	LAE-31451
Fixes an issue where when moving a canvas element on a composite library node within a composite, the element would return to its original position.	LAE-31458

Issue summary	Issue number
Ensures that the laeConfig commands to enable and disable execution sandboxing work correctly after HTTPS has been enabled. Fixes an issue where execution sandboxing would not work correctly in all cases when the hostname used for the webapp required an entry in the hosts file and was not resolvable without that entry.	LAE-31520
Fixes copy to clipboard function on Node Log dialog	LAE-31522
Fixes an issue whereby the run state associated with a data flow would be lost if it was run while the data flow was in an edited state and then saved as a new data flow.	LAE-31537
Fixed an issue with the SharePoint 365 nodes, where paged results were not all being retrieved.	LAE-31538
Fixed an issue where issues on scheduld runs could occassionally result in deadlocks.	LAE-31556
Fixed an issue where the Database Profiler node would sometimes report NullPointerExceptions when no regular expression could be determined for a column being profiled.	LAE-31697
Fixes an issue where changes to the Home Directory of imported LDAP users are not being persisted.	LAE-31724

3.3 Known issues

Legacy C++ based nodes

Legacy nodes which are built upon the C++ technology, on a Windows installation of Analyze, are unable to deal with files where there is a long file path. For example, the superseded "CSV File" node will throw an error when trying to acquire data from a file which is situated within a long path on the Windows file system.

In this scenario, users are recommended to move their file up a number of folder levels, in order to ensure the node can successfully read it.

3.4 Third Party library upgrades

A number of 3rd party libraries and components have been upgraded to take advantages of the latest features and security updates:

- Java - 11.0.20.8.1

- Postgres - 13.11

- ICU - 73.2 (Linux only)

- OpenSSL 1.1.1u (Linux only)

4. New in 3.12.1

4.1 What's new

Other changes

- Cached resolution entry was incorrectly used when virtual nodes inheriting all attributes from an instance in a library composite were updated leading to incorrect results.

- Improved user interaction performance when the Smart Guides are enabled.

- Upgraded a number of third party libraries to take advantage of any new features and security fixes.

- Improved UI performance when using the application over an extended period of time.

4.2 Corrected issues

Issue summary	Issue number
Fixed an issue where the javaException information is missing from the node error log dialog.	LAE-30460
When the interactive run cleanup was configured to only cleanup temporary data only, temporary data and logs from failed nodes was not getting cleaned up. This has been fixed so that data from failed nodes will also be cleaned up on the interactive run cleanup.	LAE-30625

Issue summary	Issue number
Ensures that the tasks for cleaning up scheduled and interactive runs are tolerant in the face of errors. Previously, the run cleanup would be aborted if the cleanup task encountered invalid data flow search paths. Now this issue will be logged, however will not prevent cleanup of unrelated run data. Rather than one big transaction, the cleanup operation has also been modified to commit its changes in batches to ensure that if there is a large backlog of execution runs to cleanup (for example due to the issue of bad data flow search paths causing the cleanup to be aborted) this will not cause issues with the transaction size getting too large.	LAE-30626
Fixed a number of memory leaks in the application UI.	LAE-30656
Fixes an issue whereby the upgrade fails if it was previously upgraded from a 3.8.5 or earlier to 3.10.x.	LAE-30687
Fixes an issue where the connections between nodes were not showing in the UI.	LAE-30689
Fixed issue where changes to library paths were not always being correctly respected after saving the data flow or library node on which the paths were changed.	LAE-30765
Fixed an issue where the javaException information is missing from the node error log dialog.	LAE-30800

5. New in 3.12.0

5.1 What's new

Save ad-hoc runs of your Data Flow

New collaboration feature, allowing users to save their ad-hoc runs of Data Flows into the Directory and share the results with other Analyze users.

Execution Data Sandboxing on Windows

Issues with the Execution Data Sandboxing feature on Windows have been fixed, such that the feature will now work correctly when Analyze components are installed across different Windows drives. For example: Analyze installed onto an I: drive, the "site" installed onto S: drive, the temporary execution data installed onto T: drive, and the "data" installed onto D: drive. In order to support these fixes, when Execution Sandboxing is enabled on Windows, the file system locations for the data, site, temporary execution and shared data directories in the containers may be different than the locations on the host system. Similarly, the controlling Analyze server that the web application communicates with will be run within a container on windows when execution sandboxing is enabled on Windows.

Security Vault

A new security vault feature has been introduced, allowing users to connect their Analyze instance to an Azure Key Vault to safely store and retrieve both the Analyze database and keystore passwords.

New Code Editor

New code editor available for users when editing code or script on nodes including input field code completion, search, and find and replace.

ADLS List node improvements

New features of the ADLS List node including file type detection and CSV field metadata detection.

Other changes

- Upgraded Apache Tomcat to version 9.0.70

- Configuration changed on Linux such that the latest tag is used in the Dockerfile for retrieving the base image when using execution containers.

- Keeping the docker image used for execution sandboxing up-to date has been made easier with a new laeConfig command.

5.2 Corrected issues

Issue summary	Issue number
Fixed a number of help links from various nodes.	LAE-29955
Improved debug logging for network errors when using the HTTP node.	LAE-29810
Fixes an issue whereby the upgrade could fail when run property sets with parent run property sets have been defined in Analyze.	LAE-29775
Fixed an issue where in properties with a File chooser widget, the UI was previously displaying the selected paths with a substitution of the ls.lae.shareRootDir, but persisting the property values with an absolute file path and auto-converting user-typed file paths using that property substitution into an absolute file path.	LAE-29786
Fixes an issue where Analyze could not start up correctly if the user directories specified at installation contained spaces.	LAE-30182
Fixed an issue when generating a Filter node from the Data Viewer where the node would fail when being run.	LAE-30184
Fixes and issue whereby a non-Admin user could not copy and clear a run state for a data flow.	LAE-30191
Fixes the unwanted 'Changes that you made may not be saved.' dialog when you refresh the directory page after discarding data flow edits.	LAE-30492
Fixed an issue where, under certain conditions, a failed scheduled Data Flow was not reporting any errors when a user would open the run.	LAE-29583

Issue summary	Issue number
Fixes bug that prevents data viewer opening when field names contain quotes.	LAE-29797
Fixed an issue with Data Flow properties, where if one was created and had a value and the group it was in was then deleted, it would not show up as an unresolved property.	LAE-29910
Fixes an issue whereby the Analyze Desktop does not start if the JRE_HOME environment variable is set to a Java version other than 11.	LAE-29850
When installing using the multi-user install instructions on Linux, the permissions for the <site-dir>/shared/Public folder were set incorrectly and dataflows were not able to write to the directory.	LAE-29847
Fixes an issue whereby profiling certain columns and certain tables would result in an error in the Detailed Analysis field on the node's detail output: "No value for: 'limitgroups'."	LAE-29978
Fixes an issue whereby the LDAP Authentication Settings are corrupted if the password field is empty when the settings are updated.	LAE-30021
Fixes an issue in the Excel files whereby for XLSX files on some shared network drives, permissions were required for the full path to the file itself and not just the file and its container directory.	LAE-30003
Fixed an issue when upgrading Analyze on Windows, the Tomcat service would not always be removed correctly during the uninstall part of the upgrade.	LAE-29440
Fixed an issue where the system would report an error when a user tried to set a property value to {{^}}	LAE-30173
Added in connection retries to the Database Metadata node when connecting to Azure Synapse, to overcome connection closed issues.	LAE-30014
When the webapp is creating temporary files for download (e.g. when using the "Export to CSV" option in the data viewer), these are now created in a temp folder in the site directory on linux rather than in the installation directory. When running using a multi-user deployment, this directory is then assigned the correct group ownership and permissions and fixes an issue whereby CSV export did not work on linux when the installation was configured for a multi-user deployment.	LAE-30422

Issue summary	Issue number
Fixed an issue in the Database Metadata node which caused errors and incorrect results when run against objects which contained special characters such as "_" on a Synapse database	LAE-30441
The Execute Data Flow node was not working on desktop installations. This has been fixed.	LAE-30465

5.3 Security Consideration

CORS is not enabled on the Analyze application, therefore cross-origin requests are prevented by default from the browsers same-origin policy. There is no requirement for the Analyze application to accept cross-origin requests, and resources are protected by strict same-site cookie based authentication, therefore CORS is not enabled.

5.4 Deprecation notice

As of 3.12.0 the following nodes will be marked as deprecated, with a view to removal in a future Analyze release. These nodes will remain deprecated yet available for the duration of 3.12.x:

Output Tableau (Superseded)

Due to the drop in support of the underlying Tableau SDK that is used by the node, and the issues seen when running the node to upload .tde files to a Tableau server in a Linux Analyze environment (see "Known Issues" section), this node is being deprecated as of the 3.12.0 release, with a view to removal in the future 3.14.0 release.

Customers are recommended to move to using the newer "Output Tableau" and "Upload Tableau" nodes which use the .hyper file format.

Python2 based nodes

The following Python2 based nodes have been deprecated, with a view to removal in a later release:

- Dummy Input (Superseded)

- Fixed Format File (Superseded)

- Trim Fields (Superseded)

- URL Query Builder(Superseded)

- Output Delimited (Superseded)

- Output Fixed (Superseded)

- Output HTML (Superseded)

- Output XML (Superseded)

- Python

6. Known issues and limitations

We would like to make you aware of the following list of issues and limitations.

If you encounter any other technical issues, contact support at support.precisely.com.

6.1 Third parties

The following table lists third party known issues and limitations:

Feature	Description
Avro	The Avro 1.7.7 specification lists some supported metadata constraints. Specifically, it places restrictions on the names of fields, as follows: • The field names must start with [A-Za-z_] • The field names must only contain [A-Za-z0-9_] Avro 1.7.7 does not support date, time and datetime data types. As a result, if you want to upload data and use the Data360 Analyze nodes, these fields will need to be converted to string data types.
Hadoop Hive Cluster	When downloading files from the Hadoop Hive Cluster, the WebHDFS API automatically encodes files to base64 format. As a result, it is not always possible to view the contents of the download in the fields on the output. For example, if the **DataOutputMode** property is set to **Field**, due to the automatic base64 encoding, the encoded result will be visible instead of the contents. To view the contents, set the **DataOutputFieldEncoding** property to **None**. However, this is not always possible due to invalid characters in the original file; in this case, the workaround is to set the **DataOutputMode** to **File** and then import the data using one of the input connector nodes.

6.2 Web application

The following table lists Data360 Analyze known issues and limitations:

Feature	Description
Composite library nodes created in previous versions	When importing or running a data flow that was created in an older version of the product, you may see error messages if the data flow contains composite library nodes that have been upgraded since the data flow was first created. If the data flow did not previously show these errors, you can resolve the issues as follows: 1. Open the data flow and select all nodes. 2. Choose **Apply Auto-Fixes**. 3. Save the data flow, then return to the Directory before reopening the data flow.
Links from tooltips to help	Although it is not currently possible to open the integrated help from the links in node property tooltips, you can manually navigate to the help by pressing **F1** then searching for the relevant topic.
Logistic Regression node	The Logistic Regression node does not support Unicode for categorical data.
Output Tableau (Superseded) node	The Output Tableau (Superseded) node, which is now deprecated, has issues on a Linux installation of Data360Analyze when configured to upload files to a Tableau Server.
Legacy C++ based nodes	Legacy nodes which are built upon the C++ technology, on a Windows installation of Analyze, are unable to deal with files where there is a long file path. For example, the superseded "CSV File" node will error when trying to acquire data from a file which is situated within a long path on the Windows file system. In this scenario, you are recommended to move your file up a number of folder levels, in order to ensure that the node can successfully read it.

7. Contact us

If you encounter any technical issues, we recommend that you visit the support portal at
support.precisely.com.

Download

Precisely recommends that you use the latest version of the product. To download Data360 Analyze,
please go to https://data.precisely.com/software/product-details/data360-analyze.

Our product is constantly evolving and input from you is highly valued. If you have any suggestions,
please contact the product team by submitting a feature request on the Community.

Notices

Trademarks