



Data3Sixty™ DQ+

Release Notes



Version 3.1



Table of Contents

1. What's New in 3.1?	1
<i>Usability Improvements</i>	1
Search for stage in tree picker	1
Preserve Memory Settings on Import and Promote.....	1
<i>Analysis Improvements</i>	1
Prevent retention from deleting files that are in use by running analysis.....	1
Add support for paging REST APIs in REST node in analysis	2
<i>Data Store and Data View Improvements</i>	2
Basic Profiling of Data Store.....	2
Ability to query file based (S3 and HDFS) data views and data stores using SQL	3
<i>Reusable Rules, a.k.a. the "Rule Library"</i>	4
Reusable JavaScript Library Node	4
<i>Visualization Improvements</i>	5
Tooltips to Display Multiple Values	5
Single or Multiple Select Enforcement for Filter Values	5
Dimension Swap in Dashboard	5
Autofit Table Width to Entire Dashlet Frame	6
Ability to Sub Total within a Pivot Table	6
Ability to Export True XLSX files from dashlets	6
Configure Font Size, Color and Style in Dashboards	6
Left/Right Targeting of Dashboards.....	6
Default Value for Export Number of Fields Filled In	7
<i>Scripting Improvements</i>	7
Add Javascript function to convert Input object in Script node to a JS object	7
<i>Security/Auditing Improvements</i>	7
Improve identification of begin/end session in Audit Trail.....	7
<i>On-premises Installation and Infrastructure Improvements</i>	7
Add Support for Hortonworks 3 and Cloudera 6	7
Upgrade Tomcat to latest 9.0.x.....	7
Support multiple versions of Hadoop distributions in deployment process	8
2. Bug Fixes	8
Environment ID displayed for Environments instead of Environment Name	8
Upgrade to latest Tomcat Patch Level	8
Log file not generated for verifyEnvironment	8
Sorting is not working as expected for numeric fields in Show Sample Data screen when a Data Store is tested	8
UI Hangs after using Show Sample Data	8
Empty columns are created while creating Microsoft Excel Data Store even when 'Skip Empty Columns' checkbox is enabled	8
Error displayed while evaluating scored classification model in Analysis	8
Script Group or Script loses connection with Rule Library in Govern when Rule Library duplicated/copied	

and pushed	9
System is not capturing data due to timezone issue	9
verifyEnvironment : No such property: org for class: check_java_version_match.....	9
"Delete all Data" in View Content does not prompt user to enter "DELETE" like "Delete All Data" in View Stage in production environment	9
Tooltips do not display dimension field values when dimension is selected to display as tool tips in tree map	9
Paging Support for RestAPI: In Enterprise Environment Analysis execution is failing when REST API node is used	9
Issue when exporting Dashboard containing gauge dashlet to JPEG/PNG/PDF.....	9
Data Last Loaded display in UI is not sync with application timezone.....	10
Error when editing date field using Mass Edit both in Case Store and Data Store	10
DQ+ Environment properties do not work correctly in the S3 channel fields	10
Specified bucket is not valid error shows if bucket value is parameterized for S3 Output Data Store upon user trying to view the content of output S3 Data Store.....	10
For Big Integer type the Show All option for Mask Field is not persisting for secure fields.	10
Analysis failing while writing to Case Store	10
Reset on Analysis Setting for override execution adding properties that are not applicable for cluster type shared in enterprise version	10
Mass edit button is enabled for Data Store record in view screen of case, when record is edit restricted	
Disclaimer.....	12
Contact us.....	12

1. What's New in 3.1?

The *Data3Sixty DQ+ Release Notes* briefly describes enhancements and modifications included in Data3Sixty DQ+ 3.1. For more detailed and functional descriptions of newly added features, please consult the Data3Sixty DQ+ User Guide.

Usability Improvements

Search for stage in tree picker

With this release, the ability to search for a Data Stage by name has been added to multiple places where a Data Stage selection can be made with the tree picker UI component. This has been added to remove the need for manually traversing the tree to look for a Data Stage when the user knows the Data Stage by name.

Preserve Memory Settings on Import and Promote

With this release, options have been added to the Import and Promote dialogs that allow the user to preserve execution settings in target Analyses and Data Views. These options allow the following memory settings for an Analysis or Data View to be preserved: Caching and Record Count settings, Override Execution Size settings, and Execution Property settings. Prior to this enhancement, imported or promoted Analyses and Data Views would overwrite pre-existing definition's Memory settings in target environments and there was no way to preserve the memory settings of the definition in the target environment.

Analysis Improvements

Prevent retention from deleting files that are in use by running analysis

This enhancement prevents an analysis from reading any data that is past its retention period. To handle the possibility of deletion due to data being within an analysis's retention period when the analysis begins executing but outside of the retention period as time passes while the analysis executes, such data will now not be deleted until 24 hours after its retention period.

Prevent analysis from reading data that is past retention period

When a data store is configured with a load range of All, New since last execution with no previous execution, New since last execution on a rebuild, or By Date with start time of null, the system will now check any data being read to make sure that it is within the data store's retention period at the time the analysis starts execution.

Delete data during retention only after a delay

The system retention process will now delete data that is past the retention period only after at least 24 hours have passed since the data was eligible for retention. This will allow any processes that are still reading that data to complete execution.

Add support for paging REST APIs in REST node in analysis

Previously, the REST API node did not handle APIs that support paging. This enhancement adds support for the most common paging variations.

Data Store and Data View Improvements

Basic Profiling of Data Store

With this release, a number of changes have been introduced to Data Store profiling. This includes:

- Profiling at the Data Store level, without needing to create an Analysis. This has been enabled by the addition of a new Profile tab in Data Store definitions. Once specifications are made in the Profile tab, a Data Store may now be executed and its profile may be viewed in a newly added Metrics tab.
- Profiling changes to support data store profiling. A number of profiling capabilities have been added or modified.

Changes include:

- Change name of nullsCount field to nullCount.
- Change name of uniqueCount field to distinctCount.
- New field uniqueCount that is the number of distinct values with only one instance.
- New field duplicateCount that is the number of distinct values occurring more than once.
- New field called distinctPercent ($\text{distinctCount}/\text{totalCount}$).
- New field called uniquePercent ($\text{uniqueCount}/\text{totalCount}$).
- New field called duplicatePercent ($\text{duplicateCount}/\text{totalCount}$).
- New field called completePercent ($(\text{totalCount} - (\text{nullCount} + \text{emptyCount}))/\text{totalCount}$).
- New field blankCount that is the number of values which are blank strings i.e., contains only whitespace.
- New field minLength which computes the minimum length of non-null and non-empty string fields. The minimum length of blank values is 0.
- New field maxLength which computes the maximum length of non-null and non-empty string fields. The maximum length of blank values is 0.
- New field patternPercent to the patterns array.
- New field valuePercent to values array.
- New field patternRegex in the patterns array that contains a regular expression for the pattern.
- Changes to profiling node UI in Analyses. The changes mentioned above now take the form of an entirely new Profile Data node, available within an Analysis. To support profiling performed in previous versions of the product, the original

Profile Data node has been renamed Profile Data Legacy and it is still available when editing an Analysis. The product thus now has 2 Profile nodes.

Ability to query file based (S3 and HDFS) data views and data stores using SQL

Previously, support for file based data views using Hive and Impala in the enterprise edition of the product was supported, yet there was no similar capability in the cloud edition of the product. Amazon Athena allows SQL queries on data that is stored in Amazon S3. This enhancement adds support for data views that are based on Amazon Athena. Data Views based on Athena/Hive/Impala are suitable for use with large amounts of data where using RedShift/Vertica could be cost prohibitive. The file based data views trade off performance for lower cost.

Additionally users may now directly perform SQL queries on files in data stores without having to create data views. This makes it easier to understand the data before building analyses.

The details of when to use this new functionality are detailed in the product's User Guide.

Big Data storage type in Data View UI

A new storage type called Big Data has been added to the data view UI. This new storage type applies to both cloud and enterprise editions of the product. For the enterprise edition, the Big Data option will only be shown if Hive/Impala is configured. Additionally, the Regular Performance option has been removed from the enterprise edition of the product. This means that users will need to manually edit any enterprise edition data view that was set to Regular Performance, selecting a supported storage type.

Allow no sort criteria to be specified in query criteria

To support the use of window functions on a data view that is performing its query using Amazon Athena, custom sort criteria may now be set in a dashlet's query criteria without requiring sort fields to be selected.

Support Hive for HDFS based data stores in the enterprise edition

With this release, the product now supports Hive for HDFS based data stores in the enterprise edition of the product.

Reusable Rules, a.k.a. the “Rule Library”

Reusable JavaScript Library Node

This enhancement encompasses multiple changes to the product that support the use of reusable JavaScript code blocks wherever scripting can be performed. These changes include:

- The addition of a Script Rule Group and Script Reusable Rule within the Rule Library Data Stage.
- Changes to the Transform tab in external data stores that allow Script Reusable Rules to be used.
- Changes to the Script node in analyses that allow Script Reusable Rules to be used.
- The ability to push reusable JavaScript rules to Govern.
- A new restriction that Execute Rule Library nodes may only reference DQ rule groups – not Script Rule Groups.

Visualization Improvements

Tooltips to Display Multiple Values

This enhancement provides the ability to add configurable tooltips to all chart types where tooltips are supported. This new ability includes the following features:

- The ability to create a custom tooltip layout for each measure in a chart, using HTML.
- The ability to show a dimension in a tooltip and hide that dimension in the x-axis of the chart.
- The ability to show a measure in a tooltip and hide that measure in the y-axis of the chart.
- The ability to show a series (bottom dimension) in a tooltip and hide that series in the x-axis of the chart.

Single or Multiple Select Enforcement for Filter Values

This enhancement adds the ability to restrict selection to a single item within a list-based filter dashlet. This feature is now present within the Add/Edit Filter dialog as a selection type option. Users may select Multiple or Single selection. By default, Multiple will be selected.

Dimension Swap in Dashboard

This enhancement introduces new 'Card Dashlet' functionality. A Card Dashlet may contain multiple Chart/Table dashlets. The display of these dashlets can then be controlled by a Filter Dashlet, such that certain cards will be displayed when users click on certain values within the Filter Dashlet.

Allow pre-defined values as filter values

In support of the Dimension Swap functionality, users may now create Static filter dashlets that use a static field name and static field values.

Dimension List for Static Filter

Also in support of the Dimension Swap functionality, users may now create Static filter dashlets that use dimension names from the dashboard's data source as the contents of the filter. This allows dashboard users to perform filtering on dashboards using dimension names.

Autofit Table Width to Entire Dashlet Frame

Each dimension, measure, and series field in a data grid dashlet now contains a new menu item called "Flex Width". When the data grid takes up less space than the dashlet, any columns marked "Flex Width" will now be expanded to fill the empty space.

Ability to Sub Total within a Pivot Table

This enhancement adds the ability to create Custom measures for Pivot Grids. Custom measures may reference other aggregation measure fields within the Pivot Grid to create derived fields, using a set of functions and operators.

Ability to Export True XLSX files from dashlets

With this release, users may now use the MS Excel option within the More->Export menu of a dashlet to export the dashlet's data in an .xlsx file.

Configure Font Size, Color and Style in Dashboards

- With this release, the ability to customize the font settings of dashboard components has been added. Users may now set a global Font Family for an entire dashboard. They may also customize fonts at the dashlet level, controlling font family, weight, color, style, and size.
- This release also includes the addition of a customer requested font family, Lato, which can be used when configuring fonts for dashboards.

Left/Right Targeting of Dashboards

- Changes have been made to the user experience of filtering in dashboards. By default, selecting components within a chart will now trigger a filtering event that is applied to all other chart/table dashlets in the dashboard. Once the selection has been made, a new overlay will appear above the chart/table where the selection was made, which allows the user to further control how the filter is applied. In addition, the concept of targeting has been introduced. With targeting, the user may override the global application of a filter created by a chart/table selection so that the filter is only applied to specific dashlets within the dashboard. This targeting is configurable by clicking the Criteria button that is available while editing a dashlet (previously known as the Query Criteria button), and using the Targets tab.
- In the case of making selections on components that have been configured to Drill to another Data Stage, behavior remains as it has been in previous releases. That is, the target stage is immediately drilled to and none of what is described above occurs.

- These changes also include a new option when clicking on a dimension while editing a dashboard called Include in Filter, which controls whether a dimension is included when filtering is applied.

Default Value for Export Number of Fields Filled In

With this release, the system will now default to exporting 10,000 rows of data when exporting .CSV data or Data for Excel from a dashlet. Previously, no default number of rows was provided.

Scripting Improvements

Add Javascript function to convert Input object in Script node to a JS object

This enhancement adds a function called convertInputToJSObject to the Analysis Script node and the Transform script editor for Data Stores. This function allows the user to convert the script's input object from a set of Java objects to a native Javascript object, in order to allow the user to use the object more easily in native Javascript functions, such as JSON.stringify.

Security/Auditing Improvements

Improve identification of begin/end session in Audit Trail

With this release, improvements to the Audit Trail have been made to more accurately detect when a user session begins and ends. In addition, these changes are intended to support a future enhancement to prevent multiple sessions from the same user.

On-premises Installation and Infrastructure Improvements

Add Support for Hortonworks 3 and Cloudera 6

With this release, Hortonworks 3 and Cloudera 6 are now supported on enterprise deployments of the product.

New Hortonworks 3.1x Cluster with Edge Nodes

As part of the newly added support for Hortonworks 3, the enterprise version of the product may be installed and use a Hortonworks 3 cluster.

New Cloudera 6 Cluster with Edge Nodes

As part of the newly added support for Cloudera 6, the enterprise version of the product may be installed and use a Cloudera 6 cluster.

Upgrade Tomcat to latest 9.0.x

With this release, Tomcat was upgraded from 9.0.4 to 9.0.14.

Support multiple versions of Hadoop distributions in deployment process

With this release, a new deployment property, HADOOP_VENDOR_VERSION, has been added to the enterprise version of the product. This property allows the system to distinguish between Hortonworks and Cloudera.

2. Bug Fixes

Environment ID displayed for Environments instead of Environment Name

During testing, it was discovered that Environment IDs could be displayed in the UI instead of their names, if an environment was added by a user during another user's session. This issue has been fixed.

Upgrade to latest Tomcat Patch Level

In this release, the product has been upgraded to using the latest Apache Tomcat patch level. This upgrade was from 9.0.4 to 9.0.14.

Log file not generated for verifyEnvironment

During testing of the Enterprise version of the product it was discovered that a log file was not generated for the verifyEnvironment task. This issue has been fixed.

Sorting is not working as expected for numeric fields in Show Sample Data screen when a Data Store is tested

During testing, it was discovered that Date/Time fields did not properly sort when using a column's sort functionality in a Data Store's Show Sample Data screen. This issue has been fixed.

UI Hangs after using Show Sample Data

During testing, it was discovered that the UI could hang when using a Data Store's Show Data Sample feature and there was no data in the Data Store. This issue has been fixed.

Empty columns are created while creating Microsoft Excel Data Store even when 'Skip Empty Columns' checkbox is enabled

During testing, it was discovered that the Skip Empty Columns functionality available when creating a Data Store from a Microsoft Excel file would not work when the empty columns were in the middle of the data set. This issue has been fixed for .xlsx files, however it still exists for .xls files where only skipping empty columns at the beginning or end of the data set is supported.

Error displayed while evaluating scored classification model in Analysis

During testing, it was discovered that Spark would throw an error when encountering previously unseen labels that were not in an original classification data set. This option has been fixed by taking advantage of the newly added Spark keep option for

handleinvalid. As such, any classification models that encountered this bug need to be retrained after the application of this bug fix.

Script Group or Script loses connection with Rule Library in Govern when Rule Library duplicated/copied and pushed

During testing, it was discovered that copying or duplicating a Rule Library could cause it to lose its connection with an associated Rule Library in Govern. This issue has been fixed.

System is not capturing data due to timezone issue

During testing it was discovered that timezone mismatches could occur when using SQL to query an external database, between the timezone of the external database being queried and the timezone of the system's web application server. To account for this difference and fix this bug, the following new substitutable SQL variables have been added:

fromTimestampUTC/toTimestampUTC

fromTimestampOrgTz/toTimestampOrgTz.

For more information on these variables, see the User Guide.

verifyEnvironment : No such property: org for class: check_java_version_match

During testing of the product's Enterprise Installation, an error was found with the verifyEnvironment task. To fix this error, the version match check between the installation host and Gradle's JVM has been discontinued.

"Delete all Data" in View Content does not prompt user to enter "DELETE" like "Delete All Data" in View Stage in production environment

During testing, it was discovered that the system was not properly confirming deletion of all data when in the View Content context in an environment marked as Production. This issue has been fixed.

Tooltips do not display dimension field values when dimension is selected to display as tool tips in tree map

During testing it was found that the Show only in Tooltip feature for dimensions was not working properly for Tree Maps in Dashboards. This issue has been fixed.

Paging Support for RestAPI: In Enterprise Environment Analysis execution is failing when REST API node is used

During testing it was discovered that Analysis execution was failing when using the REST API node in Enterprise deployments that used the Cloudera 5 distribution of Hadoop, due to a missing .jar file. This issue has been fixed.

Issue when exporting Dashboard containing gauge dashlet to JPEG/PNG/PDF

During testing it was discovered that when exporting a dashlet containing a solid gauge to a JPEG/PNG/PDF file, the gauge's needle would point to zero in the exported file, even if it was actually pointing to a non-zero value in the dashlet. This issue has been

fixed.

Data Last Loaded display in UI is not sync with application timezone

During testing it was discovered that in the Enterprise version of the product, the Data Last Loaded date displayed in Dashboards and Data Views could be out of sync with the system's timezone. This issue has been fixed.

Error when editing date field using Mass Edit both in Case Store and Data Store

During testing it was discovered that date fields would not be properly updated when changing data in Case Stores and Data Stores using the Mass Edit feature. This issue has been fixed.

DQ+ Environment properties do not work correctly in the S3 channel fields

During testing it was discovered that Environment Property placeholders were not being substituted in file paths when 'Use traditional folder structure when used for output' was set to true. This issue has been fixed.

Specified bucket is not valid error shows if bucket value is parameterized for S3 Output Data Store upon user trying to view the content of output S3 Data Store

During testing it was discovered that a UI error would occur if a user tried to use the View Content button on an External S3/HDFS Data Store in the execution details screen of an Analysis. This issue has been fixed by removing the button from this context, as it should not be shown.

For Big Integer type the Show All option for Mask Field is not persisting for secure fields.

During testing it was discovered that the mask of fields that were not of the String data type was always being set to Show None in the UI, even when another option, such as Show All, was chosen. This issue has been fixed.

Analysis failing while writing to Case Store

During testing it was discovered that, in certain scenarios, Analyses were failing to execute when writing to Case Stores. It was discovered that in these scenarios Key Fields for Update were not making their way into the definition that was executed. A fix for this issue has been added.

Reset on Analysis Setting for override execution adding properties that are not applicable for cluster type shared in enterprise version

During testing it was discovered that when using the Reset feature in Analysis settings, the Cluster type was incorrectly being reset to Dedicated. This issue has been fixed.



Mass edit button is enabled for Data Store record in view screen of case, when record is edit restricted

During testing it was discovered that the Edit/Mass Edit functionality for case records could be incorrectly enabled when it should not be – such as when a user does not have permission to edit. This issue has been fixed.

Copyright

© Copyright 2019 Infogix, Inc. All rights reserved.

Confidential—Limited distribution to authorized persons only, pursuant to the terms of Infogix, Inc. license agreement. This document is protected as an unpublished work and constitutes a trade secret of Infogix, Inc.

This document and the information contained herein are the property of Infogix, Inc. Reproduction or use in whole or in part of this document and the information contained herein by anyone without prior written consent of Infogix, Inc. is prohibited.

Disclaimer

Infogix, Infogix Assure, Infogix Insight, ACR, ACR/Detail, ACR/Summary, ACR/Workbench, ACR/Connector, ACR/Instream, ACR/File, Infogix ER, and Infogix Perceive are registered trademarks of Infogix, Inc. The Infogix logo, Data3Sixty, Data3Sixty Analyze, Data3Sixty Govern, and Data3Sixty DQ+ are trademarks of Infogix, Inc. Any other trademarks or registered trademarks are the property of their respective owners.

Contact us

If you need assistance, submit a case online at support.infogix.com, or contact Infogix Customer Support at +1-630-505-1890.

If you are experiencing issues contacting us by phone or by our website, please email support at support@infogix.com

Visit our website at www.infogix.com