



Release Notes

Legal notice

Copyright

© 2016 LAVASTORM ANALYTICS, INC. ALL RIGHTS RESERVED. THIS DOCUMENT OR PARTS HEREOF MAY NOT BE REPRODUCED OR DISTRIBUTED IN ANY FORM WITHOUT THE WRITTEN PERMISSION OF LAVASTORM ANALYTICS, INC.

Apache Hive, Hive are trademarks of The Apache Software Foundation.

Apache Spark, Spark, Apache, and the Spark logo are trademarks of The Apache Software Foundation.

Microsoft and SharePoint are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

MongoDB and Mongo are registered trademarks of MongoDB, Inc

Qlik®, Qlik Tech®, QlikView® and the Qlik Tech logos are trademarks or registered trademarks of Qlik Tech International AB.

Salesforce, SALESFORCE.COM and others are trademarks of salesforce.com, inc. and are used here with permission.

Tableau and Tableau logo are registered trademarks of Tableau Software, Inc.

TIBCO® Enterprise Runtime for R are either registered trademarks or trademarks of TIBCO Software Inc. and/or its subsidiaries in the United States and/or other countries.

The names of actual companies and products mentioned herein may be trademarks or registered trademarks of their respective owners.

Disclaimer

No representation, warranty or understanding is made or given by this document or the information contained within it, and no representation is made that the information contained in this document is complete, up to date or accurate. In no event shall LAVASTORM ANALYTICS, INC. be liable for damages or losses of any kind in connection with, or arising from its use, whether or not LAVASTORM ANALYTICS, INC. was made aware of the possibility of such damage or loss.

Contact us

If you encounter any technical issues, we recommend that you visit the Dataverse Forums at <http://help.lavastorm.com/>. If your query has not been discussed previously in the forums, you can create a new topic and receive answers from our Dataverse experts. Alternatively, you can log a ticket by emailing support at help@lavastorm.com.

Our product is constantly evolving and input from you is highly valued. If you have any suggestions, please contact the product team at product@lavastorm.com.



Tip: We recommend that you check our website for the latest documentation as minor updates or improvements may be made to the Help between releases.



Note: Note that the images in this Help are used purely for illustrative purposes.

Table of contents

1. Welcome to Dataverse	1
1.1 Supported platforms	3
1.2 System requirements	3
1.3 Accessible databases	3
2. What's new in 3.0.2?	4
2.1 New features and enhancements	4
2.2 Corrected issues	6
3. Known issues and limitations	7
3.1 Third parties	7
3.2 Web application	8
4. Product support	9

1. Welcome to Dataverse

Dataverse is a desktop application that allows you to build analytic applications that can be shared across teams with other Dataverse users. This document provides an overview of the latest release of Dataverse.

For information on installing Dataverse, please see the Setup Guide. For details about working with Dataverse, please see the Getting Started guide in the integrated product Help by clicking the Help icon in the corner of the screen:



The main features of Dataverse are listed below:

Build data flows

The Dataverse Designer enables you to easily build data flows, as follows:

- You can quickly start building your data flow by filtering the display of available nodes to show only the Dataverse Favorites which is a curated set of nodes to help beginner and intermediate users get started.
- If you are looking for our more advanced nodes, you can easily switch to display all nodes. You can also filter the list of nodes to help you to find a specific node.
- Using one of the Input Connector nodes, you can import data from sources such as Microsoft Excel or CSV, or from big data platforms such as Hadoop or MongoDB.
- You can then build out your data flow by connecting nodes and configuring their properties to customize the analysis that is performed on the underlying data at each step through your data flow.
- The tool bar gives you quick access to commonly used tools for working with your data flow, including: Running and re-running a data flow, Save, Cut, Copy and Paste, Undo and Redo.
- Using one of the Output Connector nodes, you can publish the results to a file or another system, for example to Excel or to a data visualization tool such as Tableau or Qlik.

Simplify, re-use and share

- You can simplify complex data flows by adding composites and grouping subsets of your application into a single node.
- You can use the Library Node Designer to create your own custom nodes which can be reused in many data flows.

- You can use the export and import feature to share your data flows, and custom nodes, with other Dataverse users.

Analyze data

- At each step through your data flow, you can view the underlying data in the data viewer. Large data sets can be filtered and sorted to display only relevant information.
- Data from any point within the data flow can be exported from the data viewer to an external CSV file for external viewing and sharing.

1.1 Supported platforms

This is a desktop only release for Windows 7 and Windows 10, 64-bit. The following browsers are supported:

- Chrome
- Internet Explorer 11

1.2 System requirements

The minimum hardware specification for this release is as follows:

- 4GB RAM minimum (8GB highly recommended). You will need 1GB of disk space to download and install the software, and then you will need additional space for the temporary file system used by the execution engine.
- Intel Core i5 or 4-core equivalent processor minimum (i7 recommended).

1.3 Accessible databases

Within the Data Flow Designer, you can connect your analytic application to a number of databases. The following accessible databases are supported:

- Oracle 11g, 12c
- Teradata 14.10
- MySQL
- MS SQL Server 2012
- MongoDB 2.4.9
- Spark 1.5.0
- Hadoop 2.6.0
- Hive 1.1.0

2. What's new in 3.0.2?

- ★ NEW [File picker](#)
- ★ NEW [Keyboard shortcuts](#)
- ★ NEW [Rename nodes](#)
- ★ NEW [Toggle Fullscreen](#)
- ★ NEW [Database drivers](#)
- ★ NEW [Add filters and sorts from the data viewer](#)
- ★ NEW [Sample data flows](#)
- ★ NEW [Starting and stopping application processes](#)
- ✓ FIX [Data viewer display](#)
- ✓ FIX [DB Execute node](#)
- ✓ FIX [DB Store node](#)
- ✓ FIX [Error messaging](#)
- ✓ FIX [Script property](#)

2.1 New features and enhancements

User experience enhancements

Feature	Description
File picker	When working with the input connector and output connector nodes, you can now browse to the files that you want to import by using a file picker. This removes the need to type the file path, enabling you to more quickly configure the node to import or export data.
Keyboard shortcuts	The set of Dataverse keyboard shortcuts has been extended to allow you to interact with dialog boxes by using the Enter and Esc keys.
Rename nodes	Nodes can now be renamed directly from the canvas.

Feature	Description
Toggle Fullscreen	<p>There is a new toggle button in the top right corner of the screen that allows you to quickly swap to and from a fullscreen view.</p> <p>Additionally, you can now use the F11 keyboard shortcut to toggle the Script property field to be fullscreen or to return it to its normal size.</p>

Database connection enhancements

Feature	Description
Database drivers	<p>The process of connecting to some common databases has been simplified, as the following drivers are now included with your Dataverse installation:</p> <ul style="list-style-type: none"> • MariaDB • MS Access • Oracle • Postgres • SQL Server

Additional enhancements

Feature	Description
Add filters and sorts from the data viewer	If you have filtered or sorted data in the data viewer, you can now add these filters and sorts to your data flow. This allows you to quickly build your data flow without needing to use Script to configure nodes to filter or sort data.
Sample data flows	Sample data flows are now included as part of your installation, so you do not need to download them from the help.
Starting and stopping application processes	You can now start and stop the application processes directly from the Windows Start menu.

2.2 Corrected issues

Issue summary	Issue number
<p>Data viewer display</p> <p>Resolved data viewer display issue where the content in cells that contained empty strings was misaligned.</p>	LAE-7329
<p>DB Execute node</p> <p>It is now possible to define multiple outputs on the DB Execute node.</p>	LAE-7319
<p>DB Store node</p> <p>The DB Store node has been enhanced to support Unicode data when the LoadMethod property is set to optimized.</p>	LAE-7122
<p>The performance of the DB Store node has been greatly improved when the LoadMethod property is set to insert.</p>	LAE-7153
<p>Resolved issue where the DB Store node did not honor the option set in the ExtraFieldBehavior property. Previously, independent of the value that was set in the ExtraFieldBehavior property, the node behaved as if the property were set to log.</p>	LAE-7221
<p>Error messaging</p> <p>The property validation error messaging has been improved.</p>	LAE-7210
<p>Script property</p> <p>When configuring a node Script property, previously when you inserted multiple emit statements, the text became increasingly indented and it was not possible to remove the extra characters. This issue has now been fixed.</p>	LAE-7321

3. 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, please get in touch with us. We recommend that you visit the [Dataverse Forum](#). If your query has not been discussed previously in the forums, you can create a new topic and receive answers from our Dataverse experts. Alternatively, you can log a ticket by emailing support at help@lavastorm.com.

3.1 Third parties

The following table lists third party known issues and limitations:

Feature	Description
Apache	<p>The Spark SQL Query node has highlighted some Apache issues in the following scenarios:</p> <ul style="list-style-type: none"> Selecting a field with binary type fails with the exception "UnresolvedUnionException: Not in union ["bytes","null"]". This is already raised on Apache JIRA: https://issues.apache.org/jira/browse/AVRO-1401 Describe operation returns 3 fields (col_name, col_type, comment) however the comment field is handled as not "nullable" but returns a NULL value. <p>The Spark SQL Query node processes against Hive tables. When Hive tables are processed by the cluster, the minimum memory requirement is higher compared to the memory required to run other Spark nodes that do not access Hive. The DriverMemory and ExecutorMemory both have a minimum 5G threshold. We recommend that you increase this for larger environments.</p>
Avro	<p>The Avro 1.7.7 specification lists some supported metadata constraints. Specifically, it places restrictions on the names of fields, as follows:</p> <ul style="list-style-type: none"> The field names must start with [A-Za-z_] The field names must only contain [A-Za-z0-9_] <p>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 Dataverse nodes, these fields will need to be converted to string data types.</p>

Feature	Description
Hadoop Hive Cluster	<p>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.</p> <p>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.</p> <p>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.</p>
Logistic Regression node	The Logistic Regression node does not support Unicode for categorical data.

3.2 Web application

The following table lists Dataverse known issues and limitations:

Feature	Description
Data viewer	<p>A sample of up to the first 1000 records of node data can be displayed in the data viewer.</p> <p>The data viewer only shows the first line of multi-line values. You can hover over the cell to show the full multi-line value in a tooltip. Selected records can also be copied from the data viewer to another application (e.g. Notepad).</p>

4. Product support

Lavastorm recommends that you use the latest version of the product. To download Dataverse or Dataverse libraries, please go to <http://www.lavastorm.com/>.



© 2016 LAVASTORM ANALYTICS, INC.

www.lavastorm.com

Contact support: help@lavastorm.com

Document ID: DV-RN-3

Date of publication: Wednesday, November 9, 2016