

## **KNIME – HP Vertica Connector Node**

KNIME – HP Vertica Connector Node .....	1
License.....	2
Versions .....	2
Installation .....	2
KNIME Analytics Platform – HP Vertica Analytic Database Server Connection.....	2
Connect to a HP Vertica database.....	3
Query a HP Vertica database.....	5
In-database Processing .....	8

## License

Starting with Version 2.1, KNIME is released under the GNU General Public License, Version 3 (including certain additional permissions according to Sec. 7 of the GPL). It is also available – through the dual licensing model – under customized licenses. If you wish to receive KNIME under a different license than the GPL, please contact us to discuss licensing arrangements.

## Versions

	Linux	Windows	MacOS X
KNIME (32bit)	✓	✓	
KNIME (64bit)	✓	✓	✓
KNIME Developer Version (32bit)	✓	✓	
KNIME Developer Version (64bit)	✓	✓	✓

## Installation

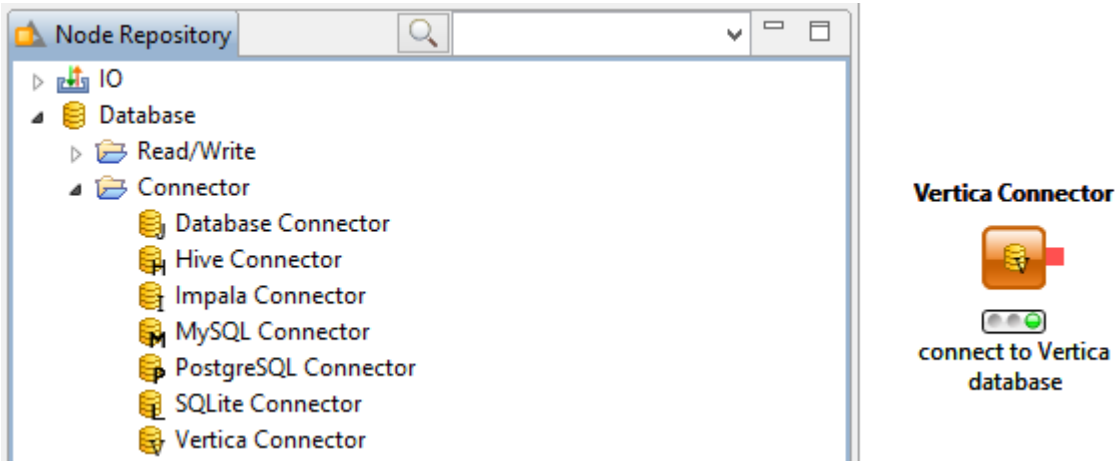
Download one of the above versions, unzip it to any directory for which you have write permissions. If you are using Windows click the *knime.exe* file, and for Linux click *knime* in order to start KNIME.

## KNIME Analytics Platform – HP Vertica Analytic Database Server Connection

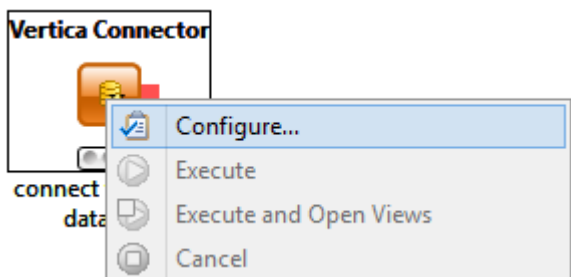
We assume that the HP Vertica Analytic Database Server is up and running and the database to which you want to connect can be accessed by a specific database user. Furthermore we assume that KNIME is installed as well as the feature *KNIME Connectors for Common Databases*. For details about how to install KNIME features, see <http://www.knime.org/downloads/update>. The feature *KNIME Connectors for Common Databases* provides connector nodes together with the associated JDBC drivers for common databases, such as HP Vertica, MySQL, PostgreSQL, etc. The nodes allow you to connect to databases easily without needing to register the JDBC driver manually.

## Connecting to a HP Vertica database

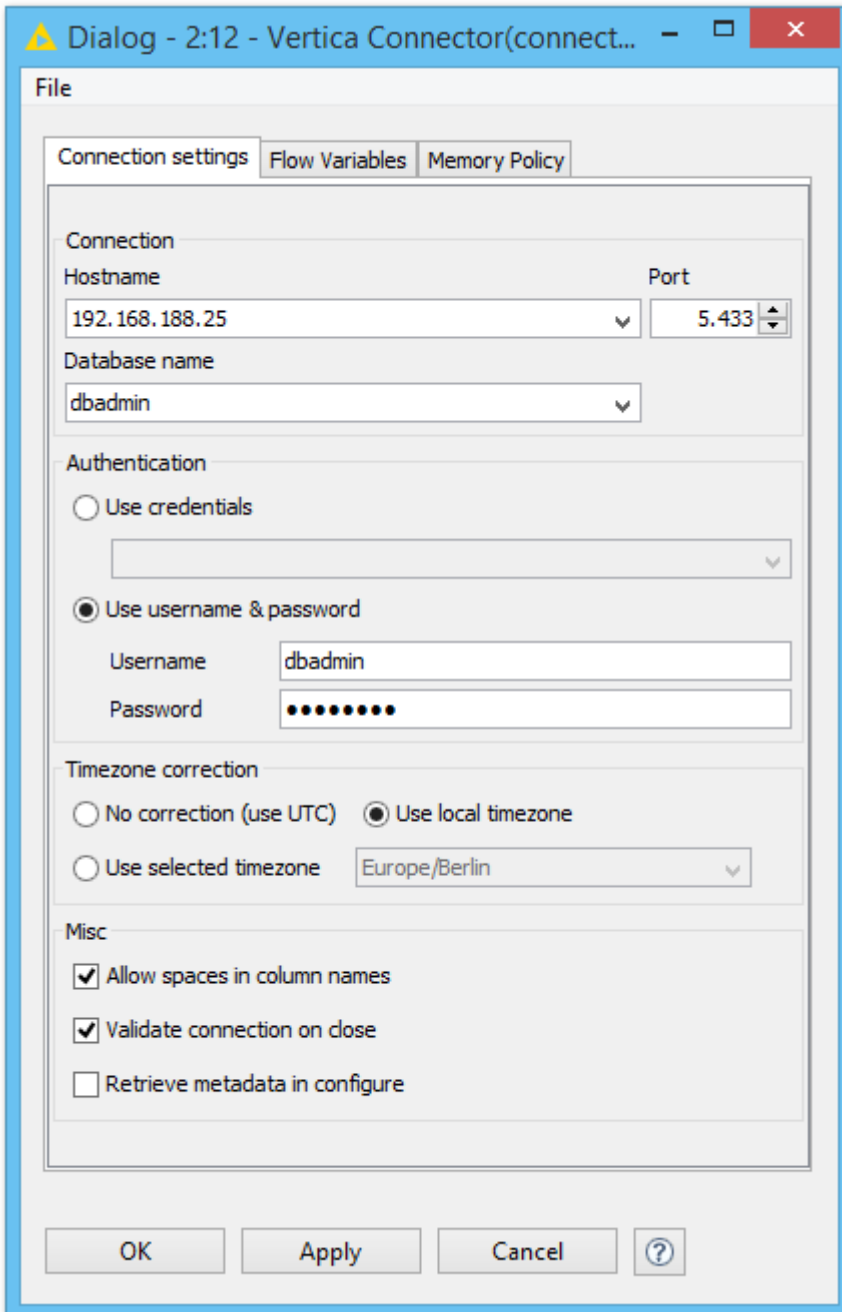
To connect to an HP Vertica database, use the *Vertica Connector* node in the *Database->Connector* category of the KNIME Node Repository.



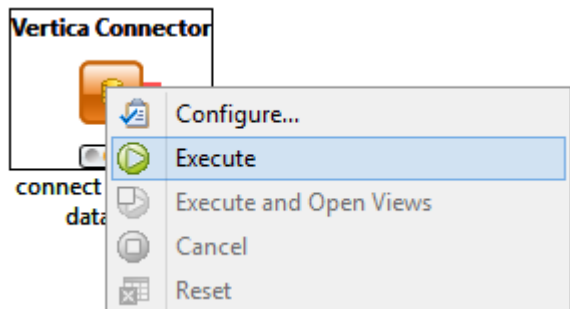
To configure the node and specify to which database the node should connect, open the node dialog by right clicking the node and selecting “Configure”.



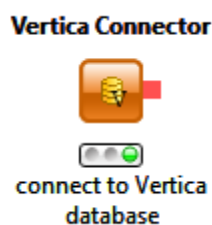
In this dialog window you now need to specify the server name or IP address of the HP Vertica database server, as well as the database username and password. You can also specify the port you want to connect to as well as any timezone corrections.



Next, right-click the node again and select “Execute” to establish the connection to the Vertica database.

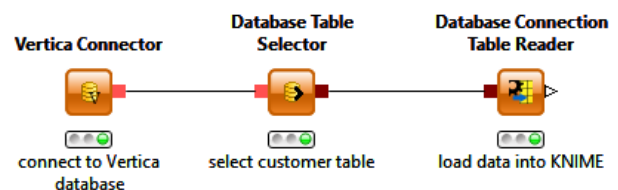
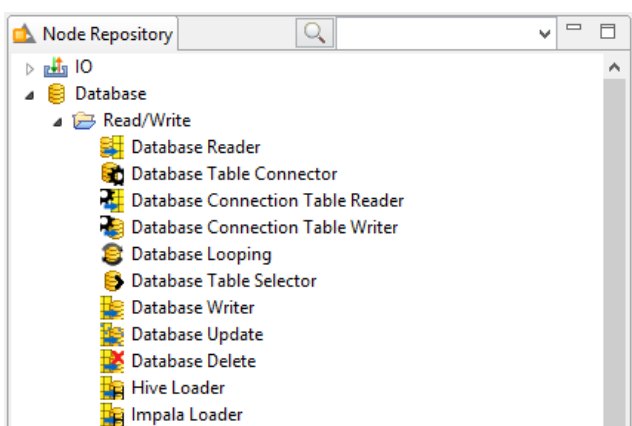


If the settings are correct and the node can connect to the database successfully the traffic light will switch to green.



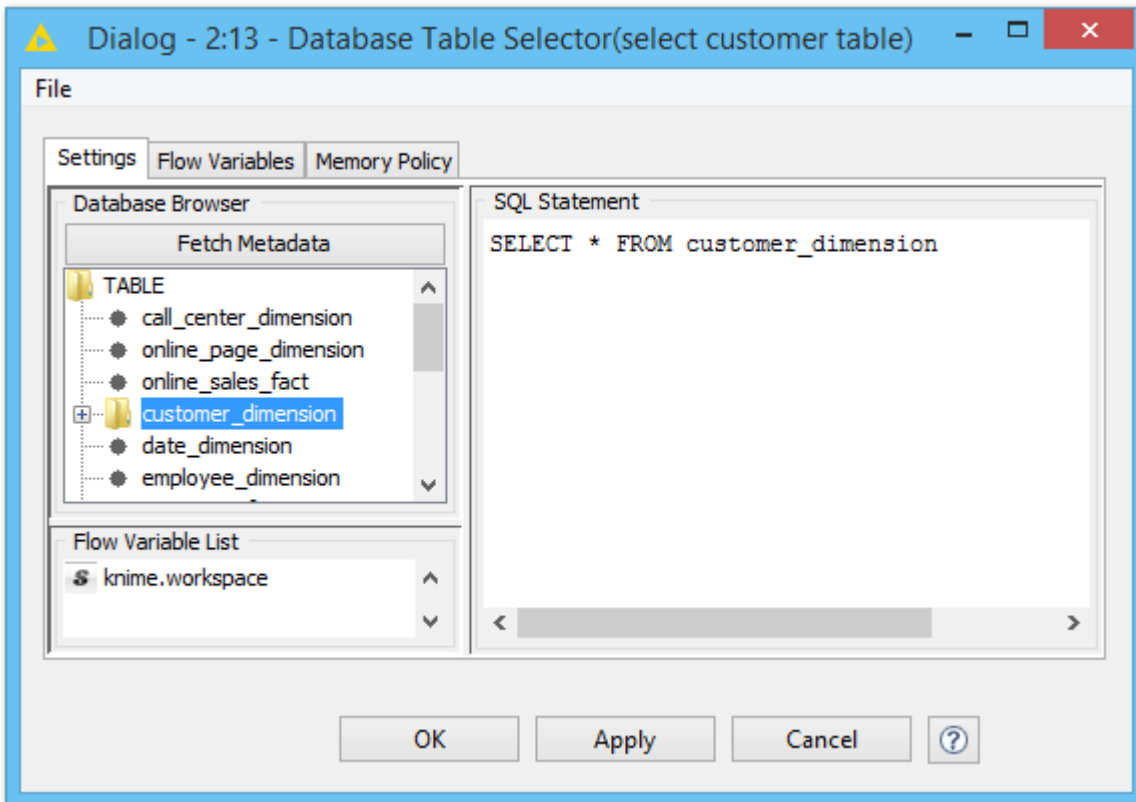
## ***Query a HP Vertica database***

Once the *Vertica Connector* node is configured and executed successfully a connection to a Vertica database is established. In order to select data from a table and load the data into KNIME, first a database table has to be selected. Therefore the node *Database Table Selector* can be used.

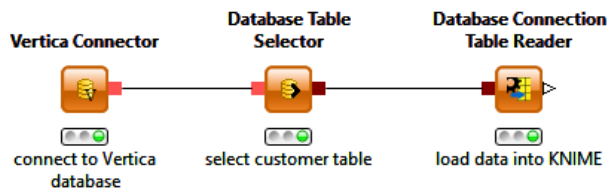


The *Database Table Selector* node can be found in the *Database->Read/Write* category.

Add the *Database Table Selector* node to the workflow by dragging and dropping it onto the workbench and connect it to the *Vertica Connector*. The node will use the already established database connection of the *Vertica Connector*. In the *Database Table Selector* dialog, you can specify from which table data should be selected through an SQL query.



To load the selected data into KNIME use the *Database Connection Table Reader* node, connect it to the *Database Table Selector* and execute it. The *Database Connection Table Reader* will reuse the established connection, run the SQL query – specified by the *Database Table Selector* – and load the resulting data into KNIME. The data is now available as a KNIME data table and can be further processed by other KNIME nodes.

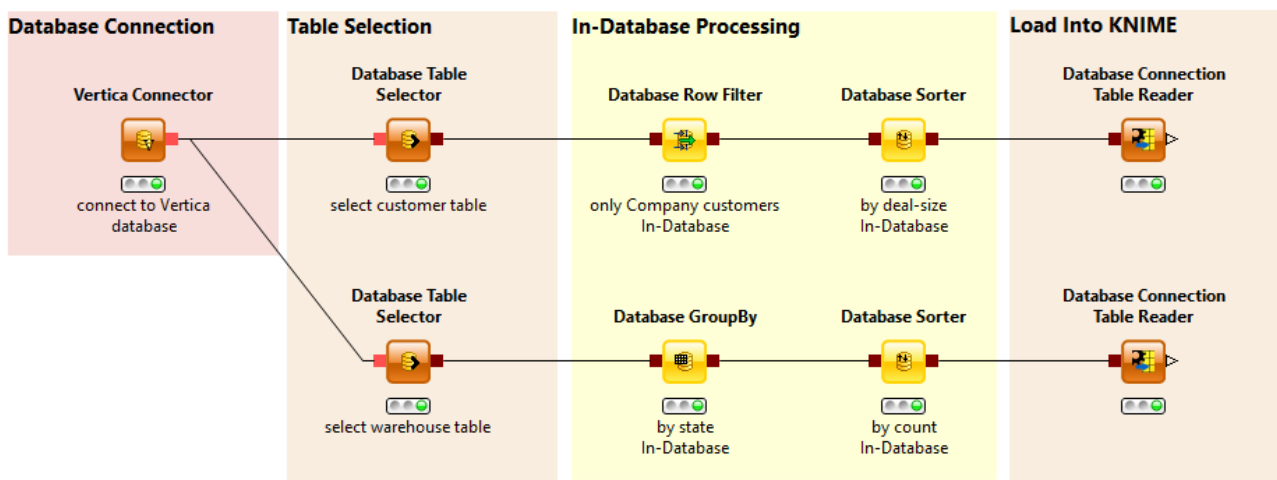


Data read from Database - 2:14 - Database Connection Table Reader(load data into KNIME)

Row ID	L custom...	S custom...	S custom...	S custom...	S title	L househ...	S custom...	S custom...	S custom...	S c
Row0	1	Individual	Kevin J. Dobisz	Male	Mr.	17454	430 Essex St	Denver	CO	South
Row1	2	Individual	Anna G. Li	Female	Dr.	9937	240 Maple St	Burbank	CA	West
Row2	3	Individual	Lauren P. Vu	Female	Dr.	19925	498 Essex St	Peoria	IL	MidW
Row3	4	Individual	Daniel I. Fortin	Male	Dr.	4209	111 Main St	Grand Prairie	TX	South
Row4	5	Individual	Craig S. Rob...	Male	Sir	10858	138 Alden Ave	Fayetteville	NC	East
Row5	6	Company	Foodshop	?	?	?	373 Brook St	Allentown	PA	East
Row6	7	Individual	David H. Gr...	Male	Dr.	15175	467 Green St	Lowell	MA	East
Row7	8	Individual	Wendy S. Y...	Female	Dr.	14028	63 Alden Ave	Chicago	IL	MidW
Row8	9	Individual	Theodore X...	Male	Mr.	22803	444 Cherry St	Cambridge	MA	East
Row9	10	Individual	Amy C. King	Female	Ms.	1753	140 Alden Ave	West Covina	CA	West
Row10	11	Individual	Sarah . Vogel	Female	Ms.	9758	483 Cherry St	Costa Mesa	CA	West

## In-database Processing

To make use of a powerful database server or cluster, KNIME offers a set of nodes that delegates the actual work to the server or cluster. Manipulation of data, such as filtering, sorting, or grouping is processed in-database. Once the preprocessing and transformation of the data has been completed by the server or cluster, the filtered or aggregated results can be loaded into KNIME. These nodes can be found in the *Database->Manipulation* category of the Node Repository. The *Database Table Selector* node has to be applied after the *Vertica Connector* node in order to use any of the in-database manipulation nodes. An example workflow, in which certain rows are filtered and aggregated, is illustrated below.



The data from the nodes with a database connection port (i.e. red square) are located and processed within the database. The data from the nodes with a small triangular in- or output port (e.g. *Database Connection Table Reader* node) are located and processed within KNIME. In the upper example workflow the data are not loaded into KNIME as long as the *Database Connection Table Reader* node has not been executed.