



Connecting to Amazon Athena using Qlik®

Connecting to Amazon Athena

Amazon Athena is an interactive query service that makes it easy for Qlik customers to analyze data in Amazon S3 (Amazon Simple Storage Service) using standard SQL. While Qlik can connect to S3 in different ways (i.e. using Qlik's Rest API connector), Amazon Athena makes it easier by providing an ODBC driver. Given that Amazon Athena is serverless, there is no infrastructure to manage so customers only pay for the queries that they run via the ODBC connection. With Athena, there's no need for complex ETL jobs to prepare the data for analysis or to move the data from S3. To use Qlik with Athena, simply point Qlik to the table schema representing the data in S3. Qlik will create the SQL queries to start interacting with the data in an associative way regardless of the size of the dataset and make it very easy to combine the S3 data with many other data sources. With Amazon Athena and Qlik, you can go from data stored in S3 to gorgeous visualizations and associative applications in just minutes.

Some of the key features of Amazon Athena include the following:

- **Data formats:** support for a variety of standard data formats such CSV, JSON, Apache ORC, Apache Parquet and Avro
- **Standard SQL:** behind the scenes, Amazon Athena uses Presto so it comes with full support of ANSI SQL
- **Managed Data Catalog:** Amazon Athena uses a managed Data Catalog to store metadata, including schemas about the databases and tables for your data stored in Amazon S3. Amazon Athena benefits from schema-on-read so that no data loading or transformation required
- **Secure:** Amazon Athena leverages AWS Identity and Access Management (IAM) as well as S3 ACLs and bucket policies to control access to the data. Users can also query data encrypted either via S3 Server Side Encryption or Client Side Encryption with keys in AWS KMS
- **Availability:** Amazon Athena is highly available and durable using compute resources across availability zones

To learn more about Amazon Athena visit: <https://aws.amazon.com/athena/faqs/>

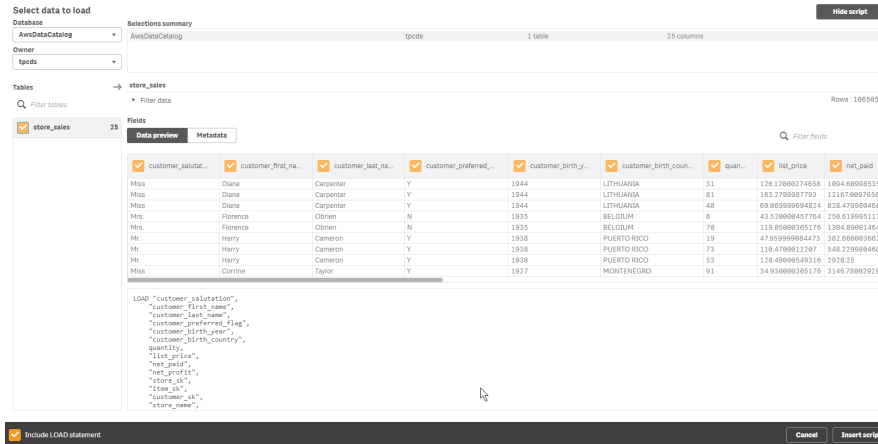
Getting started with Amazon Athena and Qlik

Qlik (Qlik Sense and QlikView) connect to Amazon Athena via ODBC. To establish connectivity, follow these steps:

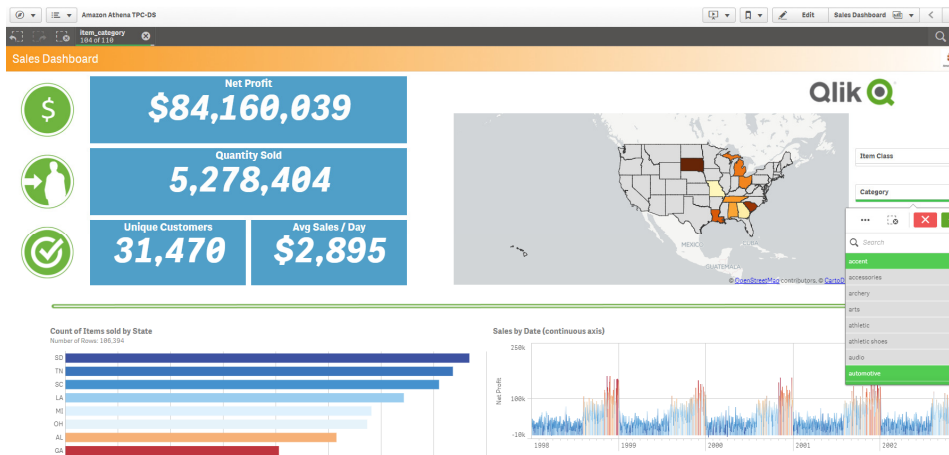
1. Download the latest version of the Amazon Athena 64-bit ODBC driver:
<http://docs.aws.amazon.com/athena/latest/ug/connect-with-odbc.html>
The link to ODBC Configuration Guide (how to establish connectivity to Amazon Athena) is:
<https://s3.amazonaws.com/athena-downloads/drivers/ODBC/Simba+Athena+ODBC+Install+and+Configuration+Guide.pdf>
2. Install the ODBC driver on the Qlik Server (Qlik Sense and/or QlikView) machine. More information on how to install an ODBC driver on the Qlik Sense Server is available here:
<http://help.qlik.com/en-US/sense/November2017/Subsystems/Hub/Content/DataSource/ODBC.htm#anchor-2>

The Configuration Guide covers in detail the different options on how to configure the Authentication. For example, one option is to leverage AWS IAM in which case the Amazon Access Key can be used as the username and the Secret Access Key ID as the password. Please, also notice that the exact steps how to create a connection on Qlik Sense and QlikView are slightly different but overall the approach is the same. For this post, the steps will be based on Qlik Sense.

- In a new or existing Qlik application, create a new ODBC connection pointing at the DSN that was created using the Amazon Athena ODBC driver. To learn how to create an ODBC connection in a Qlik Sense app, visit: <http://help.qlik.com/en-US/sense/November2017/Subsystems/Hub/Content/DataSource/ODBC.htm>
- Once the ODBC connection has been created in the Qlik app, select the data (by clicking on the 'Select data' button of the connection, a new window pops up where you can select the database, the owner and table(s) that you want to use for the application. The user has full control in terms of what and how to load the data.



- Once all the tables have been captured, load the data into the Qlik application. For more information on how to create a Qlik app, visit: <http://help.qlik.com/en-US/sense/November2017/Subsystems/Hub/Content/Introduction/create.htm>
- Through the Qlik application, users can analyze data coming from Amazon S3 in a visual and associative way



There is more documentation on Amazon Athena that can help you get started. Here are some recommendations:

- Amazon Athena User Guide: <https://docs.aws.amazon.com/athena/latest/ug/athena-ug.pdf>
- Amazon Athena Set-up Guide: <http://docs.aws.amazon.com/athena/latest/ug/setting-up.html>
- Amazon Athena Tuning Guide: <https://aws.amazon.com/blogs/big-data/top-10-performance-tuning-tips-for-amazon-athena/>

Solution Highlights

- Amazon Athena and Qlik allow you to easily analyze data stored in Amazon S3 without the need for complex ETL or managing infrastructure
- Leverage Qlik and the easy access to data stored in Amazon S3 to make faster and better-informed decisions