

Snowflake

- [Introduction](#)
- [Configuration](#)
- [Sample in Snowflake](#)
- [Collibra Results](#)

Introduction

Snowflake is a popular enterprise data warehouse in the cloud, typically compared to Redshift and gaining traction as a much easier, cheaper alternative to traditional, on-prem solutions from Oracle and SAP. Companies are adopting Snowflake incrementally, so aligning Collibra with a warehouse migration project will help reduce risk and give the overall project a greater chance of success. This solution approach will leverage the Catalog generic jdbc connector to ingest metadata directly from the Snowflake storage tier. If more complex use cases are required, such as synchronizing meta between Collibra and Snowflake, a [Connect template approach](#) should be considered.

Trial instances can be provisioned from their website: <https://www.snowflake.net> and by clicking the [start for free link](#).

Configuration

Since snowflake is a data warehouse, we will ingest metadata from the database tier, using the Snowflake JDBC driver available here:

<https://docs.snowflake.net/manuals/user-guide/jdbc-configure.html>

The driver used to test the config for this article was snowflake-jdbc-3.6.8.jar.









In Catalog, the source registration requires the following steps.

NOTE: the schema parameter passed is NOT part of the URL (i.e. the connection string) but it is required. Schema means schema and not database, typical schemas created in a Snowflake cloud trial are:

- TPCDS_SF10TCL
- TPCH_SF10
- INFORMATION_SCHEMA
- WEATHER (shown in example below)

Register data source ✕

Select the type of data source:

 Amazon RedShift	 Cloudera Hive
 Csv	 Excel
 Generic	 Hortonworks Hive
 HP Vertica	 IBM DB2

Showing 14 of 14

[< Back](#)

Here is the Driver creation page:



Generic

Create a JDBC driver for version Snowflake368

Connection *

jdbc:snowflake://\$host

Driver Class Name *

net.snowflake.client.jdbc.SnowflakeDriver

Connection properties

Add property

Label	Property	Mandatory	
Schema	schema	×	
Host	host	×	

< Back

Update

Register data source



Generic

Create a JDBC driver

JDBC Driver Version Name *

snowflake368

Driver files

Upload

Name	Type	
snowflake-jdbc-3.6.8.jar	jar	

and here is the connect info screen

Register data source



Generic

JDBC driver version

Snowflake368

Connection details

Schema

TPCH_SF100

Host

te86826.snowflakecomputing

Login information

Store credentials

Username*

bhardaway

Password

.....

Schedule data refresh

Cron pattern

Time zone

[< Back](#)

[Next](#)

Sample in Snowflake

An example of a schema being ingested to Catalog:

Snowflake view:

The screenshot shows the Snowflake web interface. At the top, there's a navigation bar with icons for Databases, Warehouses, Worksheet, and History. Below that, the breadcrumb path is "Databases > SNOWFLAKE_SAMPLE_DATA > WEATHER_14_TOTAL (WEATHER)". Underneath, there are tabs for Tables, Views, Schemas, Stages, File Formats, and Sequences. The "Tables" tab is active. Below the tabs, there's a "Load Ta..." button. The main content is a table with the following columns: Column Name, Ordinal, Type, Nullable, Default, and Comment. The table contains two rows of data.

Column Name	Ordinal ▲	Type	Nullable	Default	Comment
V	1	VARIANT	true	NULL	
T	2	TIMESTAMP_NTZ(9)	true	NULL	

Collibra Results

WEATHER		Generic				
WEATHER_14_TOTAL						
WEATHER_14_TOTAL > V	1	V	VARIANT	0	×	
WEATHER_14_TOTAL > T	2	T	TIMESTAMPNTZ	0	×	
DAILY_14_TOTAL						
DAILY_14_TOTAL > V	1	V	VARIANT	0	×	
DAILY_14_TOTAL > T	2	T	TIMESTAMPNTZ	0	×	

TPCDS_SF10TCL		Generic					Schema	Candidate
STORE_SALES						Table	Candidate	
CUSTOMER						Table	Candidate	
CUSTOMER > C_CUSTOMER_SK	1	C_CUSTOMER...	NUMBER	38	×	×	Column Candidate	
CUSTOMER > C_CUSTOMER_ID	2	C_CUSTOMER_ID	VARCHAR	16	×	×	Column Candidate	
CUSTOMER > C_CURRENT_CDEMO_SK	3	C_CURRENT_C...	NUMBER	38	×	✓	Column Candidate	
CUSTOMER > C_CURRENT_HDEMO_SK	4	C_CURRENT_H...	NUMBER	38	×	✓	Column Candidate	
CUSTOMER > C_CURRENT_ADDR_SK	5	C_CURRENT_A...	NUMBER	38	×	✓	Column Candidate	
CUSTOMER > C_FIRST_SHIPTO_DATE_SK	6	C_FIRST_SHIPT...	NUMBER	38	×	✓	Column Candidate	
CUSTOMER > C_FIRST_SALES_DATE_SK	7	C_FIRST_SALES...	NUMBER	38	×	✓	Column Candidate	
CUSTOMER > C_SALUTATION	8	C_SALUTATION	VARCHAR	10	×	✓	Column Candidate	
CUSTOMER > C_FIRST_NAME	9	C_FIRST_NAME	VARCHAR	20	×	✓	Column Candidate	
CUSTOMER > C_LAST_NAME	10	C_LAST_NAME	VARCHAR	30	×	✓	Column Candidate	
CUSTOMER > C_PREFERRED_CUST_FLAG	11	C_PREFERRED...	VARCHAR	1	×	✓	Column Candidate	
CUSTOMER > C_BIRTH_DAY	12	C_BIRTH_DAY	NUMBER	38	×	✓	Column Candidate	
CUSTOMER > C_BIRTH_MONTH	13	C_BIRTH_MON...	NUMBER	38	×	✓	Column Candidate	
CUSTOMER > C_BIRTH_YEAR	14	C_BIRTH_YEAR	NUMBER	38	×	✓	Column Candidate	
CUSTOMER > C_BIRTH_COUNTRY	15	C_BIRTH_COU...	VARCHAR	20	×	✓	Column Candidate	
CUSTOMER > C_LOGIN	16	C_LOGIN	VARCHAR	13	×	✓	Column Candidate	
CUSTOMER > C_EMAIL_ADDRESS	17	C_EMAIL_ADD...	VARCHAR	50	×	✓	Column Candidate	
CUSTOMER > C_LAST_REVIEW_DATE	18	C_LAST_REVIEW...	VARCHAR	10	×	✓	Column Candidate	
SHIP_MODE						Table	Candidate	
CUSTOMER_DEMOGRAPHICS						Table	Candidate	
CATALOG_SALES						Table	Candidate	
CATALOG_SALES > CS_SOLD_DATE_SK	1	CS_SOLD_DATE...	NUMBER	38	×	✓	Column Candidate	
CATALOG_SALES > CS_SOLD_TIME_SK	2	CS_SOLD_TIME...	NUMBER	38	×	✓	Column Candidate	
CATALOG_SALES > CS_SOLD_DATE_SK	3	CS_SOLD_DATE...	NUMBER	38	×	✓	Column Candidate	