Unifying Big Data Analytics

Databricks Tableau Connector

User Guide

v0.1.0 released: April 24, 2019

Introduction

The Databricks connector in Tableau is developed by Databricks based on the Tableau connector SDK. The Databricks connector provides full compliance with the TDVT testing suite while bringing to the business analysts a simpler user interface to connect to their Databricks clusters. In addition, the connector has several built-in connection optimizations.

IMPORTANT: Both the Tableau connector SDK and the Databricks Tableau connector are beta software. This library is in limited release private preview.

Requirements

1. Windows or MAC OS
2. Tableau Desktop version 2019.1 or above
3. Simba ODBC Driver version 2.6.4 or above

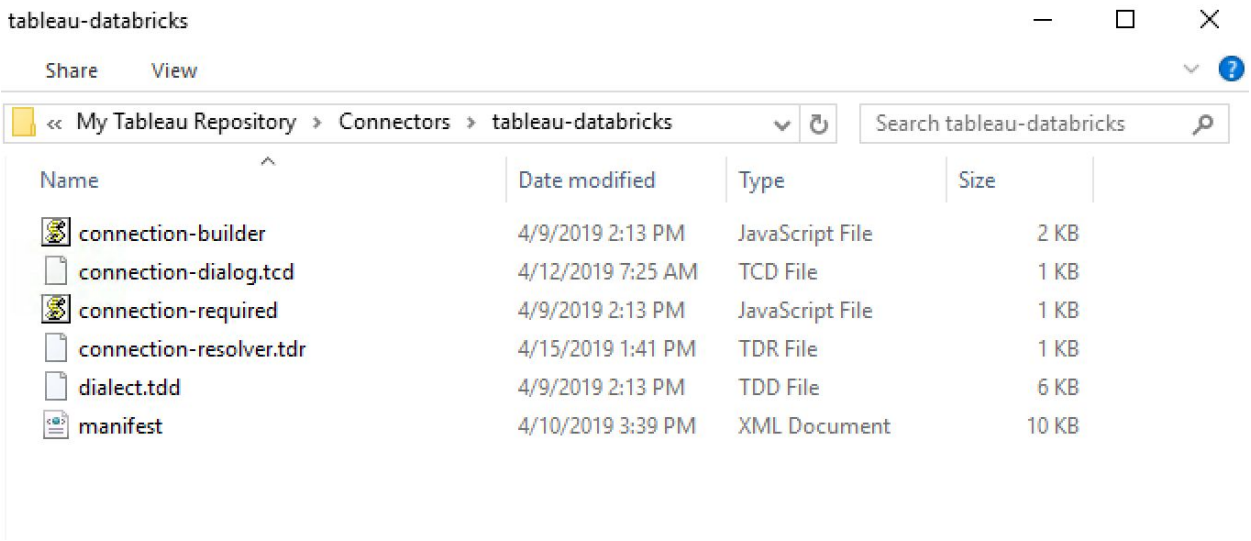
Configuration

You can deploy the Databricks connector using the Tableau Connector SDK, which is currently experimental and can only be launched from the command line.

Step 1. Extract the connector

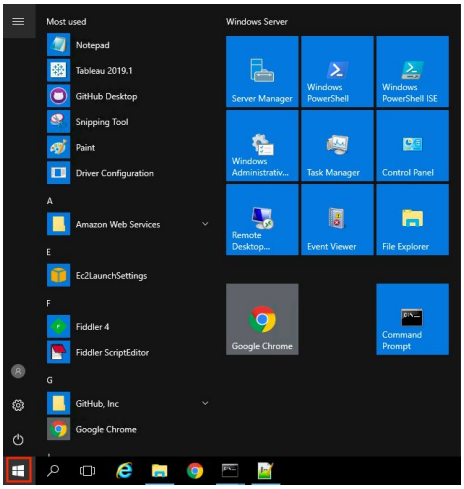
Tableau creates a directory named `My Tableau Repository` which is typically placed at `C:\Users\<username>\Documents\My Tableau Repository\` on Windows and at `/Users/<username>/Documents/My Tableau Repository` on Mac. Create a directory named `Connectors` inside the `My Tableau Repository` directory on your machine (if it does not already exist). Extract the ZIP archive with the Databricks connector in the `Connectors` directory.

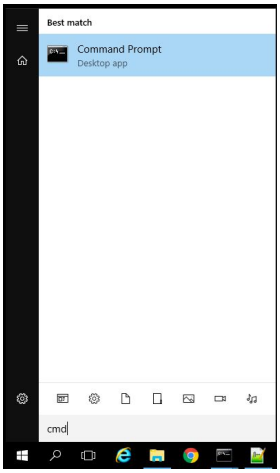
Make sure that the extracted connector files are placed directly inside the extracted `tableau-databricks` directory:



Step 2. Start a command prompt

Windows. Open the program in **Start Menu** by clicking the **Start** button, enter the text **cmd** in the search box, and click **Command Prompt** in the results.





Mac OS. Go to the **Applications > Utilities** folder, and open the **Terminal** application.

Step 3. Start Tableau with the connector plugin

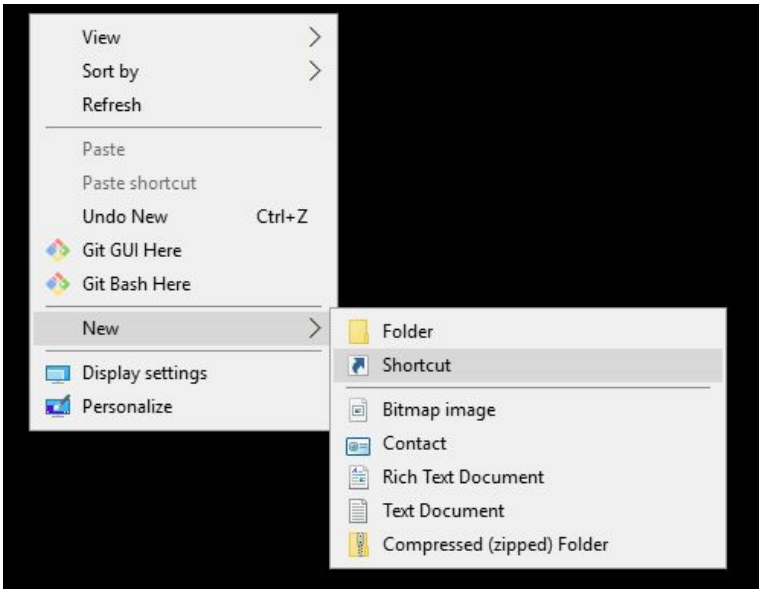
When starting up Tableau, you must pass the location of the new Databricks connector. Run the command below with the `-DConnectPluginsPath` option in the Command Prompt to launch Tableau Desktop with the Databricks connector plugin loaded. Replace `<username>` with your login username to the system.

Windows:

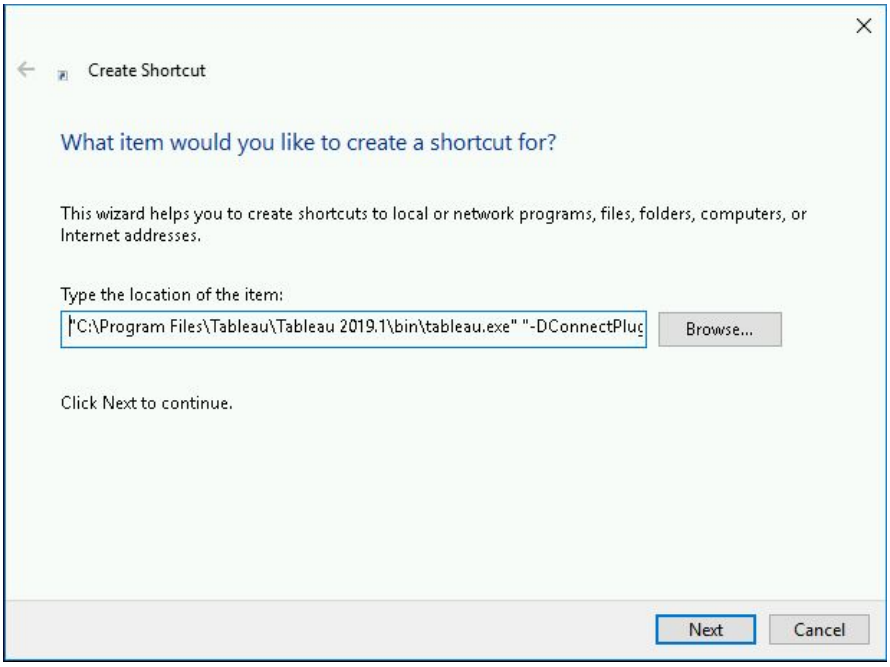
```
"C:\Program Files\Tableau\Tableau 2019.1\bin\tableau.exe"
"-DConnectPluginsPath=C:\Users\<username>\Documents\My Tableau
Repository\Connectors"
```

Alternatively, on Windows you can create a Desktop shortcut to your Tableau application which will load the Databricks connector without needing to run it from the Command Prompt every time. Below, there are three configuration steps that you need to do only once in order to create a shortcut that you can use anytime you want to run the Databricks connector:

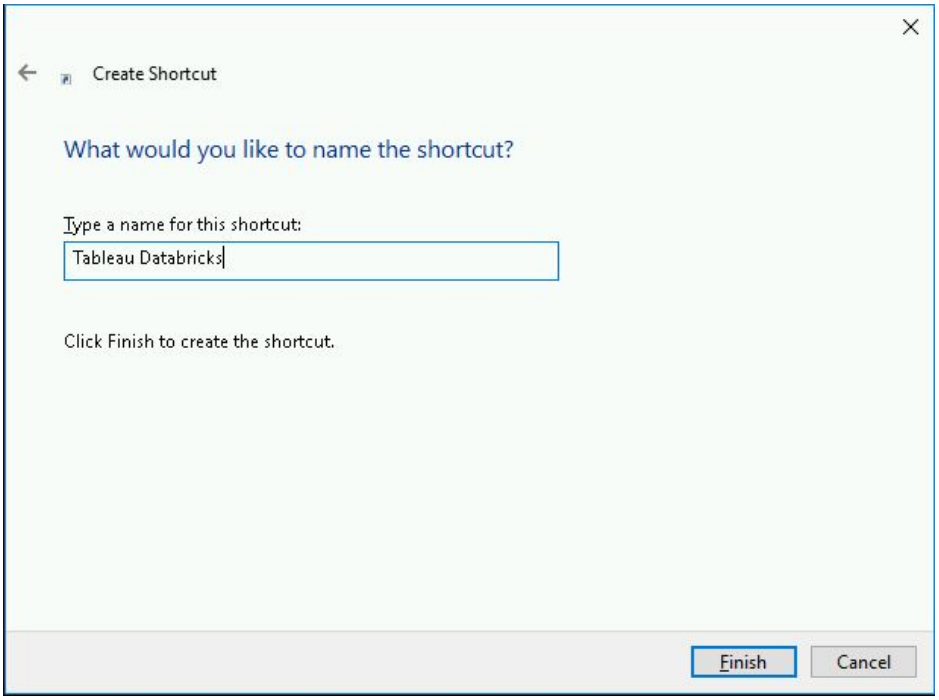
1. Right-click the Tableau Desktop shortcut and select **New > Shortcut.**



2. Enter the command in the dialog box that asks for the location of the item

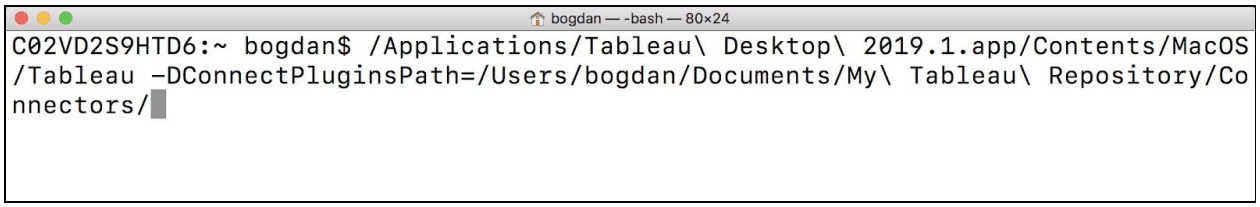


3. Enter a name for of your choice for the new shortcut, for instance `Tableau Databricks`.



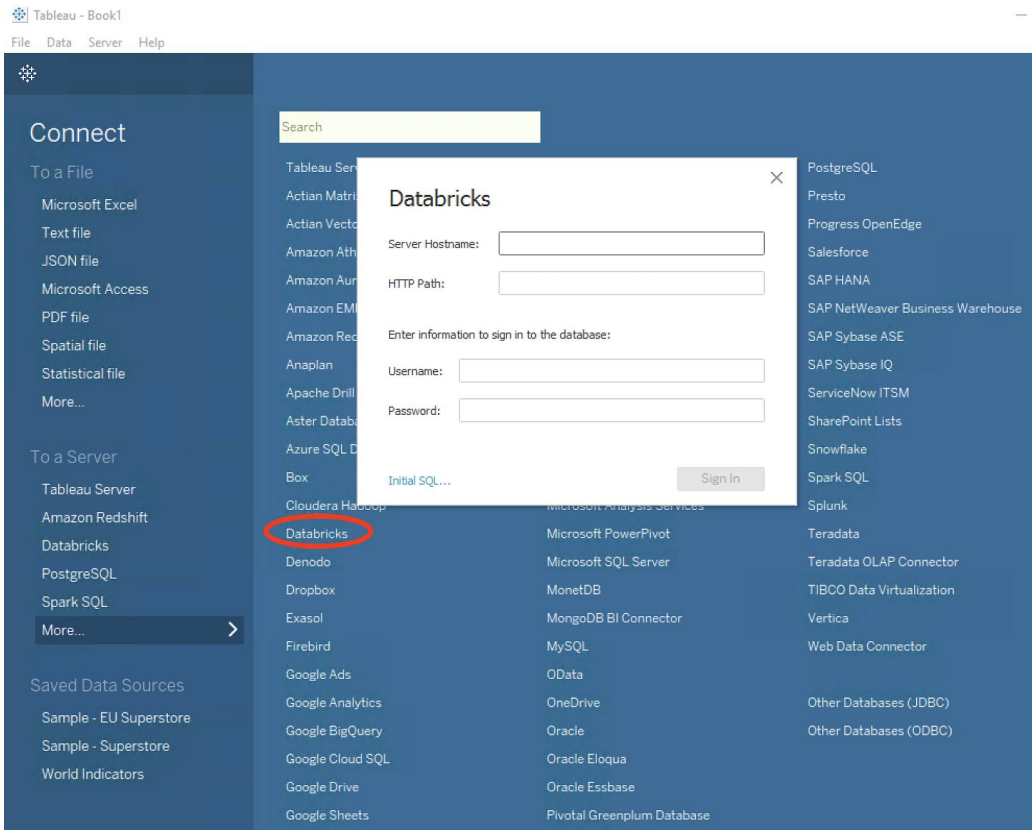
Mac OS:

```
/Applications/Tableau\ Desktop\ 2019.1.app/Contents/MacOS/Tableau
-DConnectPluginsPath=/Users/<username>/Documents/My\ Tableau\
Repository/Connectors/
```

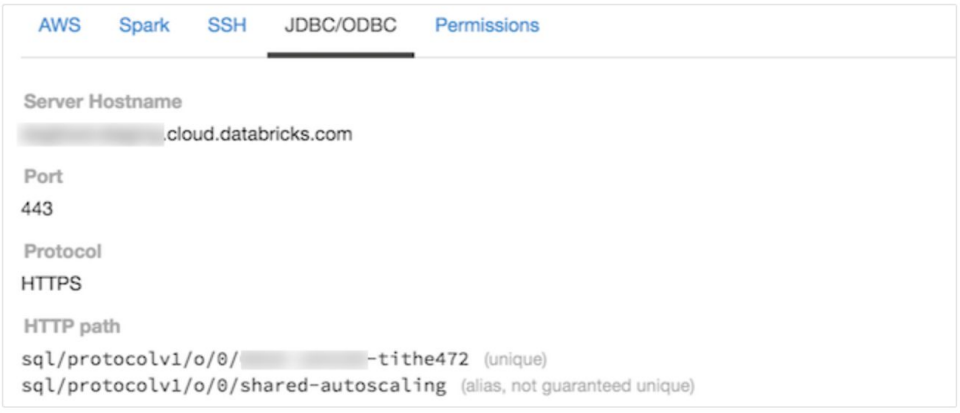


Connect to a Databricks Cluster

Step 1. After opening the Tableau Desktop application, go to the **To Server** menu on the left-side of the interface and select the Databricks connector. The connection dialog consists of four fields that require input: the Databricks server hostname, the cluster HTTP address, and authentication details.



Step 2. Configure the address of the Databricks cluster. You can retrieve the address from the cluster detail page by clicking the **JDBC/ODBC** tab. Enter the Server Hostname and the HTTP path found on the cluster detail page.



Step 3. To authenticate, use your Databricks username and password. You can also use `token` as the username and a personal access token as the password.

Step 4. Click **Sign in**.