**How to Use Azure Synapse with SAP Data Intelligence**

Guide created: June 3, 2020

Last updated: June 3, 2020

Author: Jim Giffin, SAP GCOE

Table of Contents

[Create a managed connection in DI 1](#_Toc42098398)

[Create a graph to pull data from Synapse: 2](#_Toc42098399)

[Custom Azure Synapse Operator: 5](#_Toc42098400)

[Create the Dockerfile 5](#_Toc42098401)

[Create the Operator 5](#_Toc42098402)

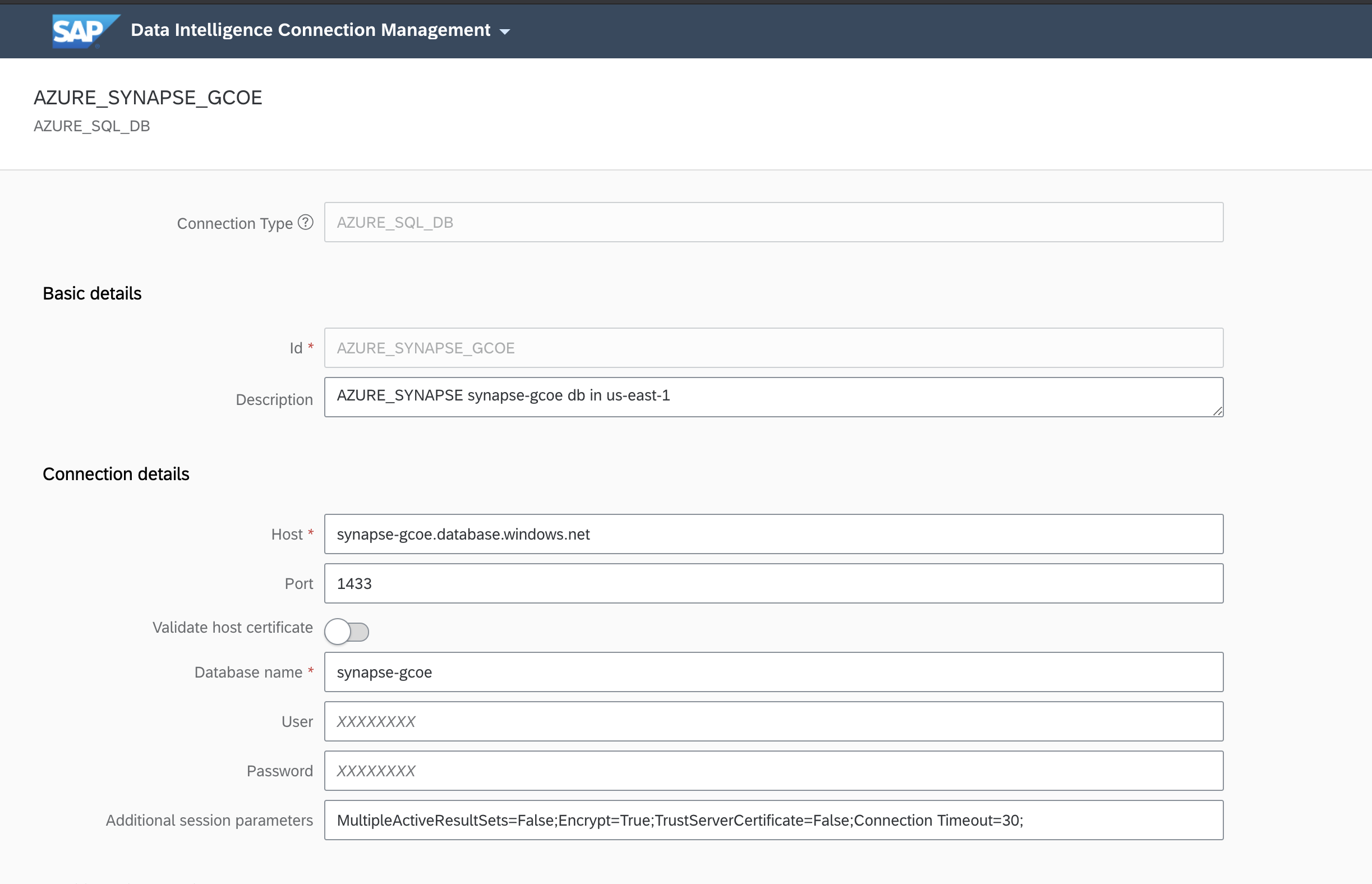
[Build a graph to insert data into Synapse using the new operator 12](#_Toc42098403)

[Target Table: 13](#_Toc42098404)

How-To document to use Data Intelligence to pull data from Azure Synapse to Flat File in ADLv2 and insert data into Synapse from the same flat file.

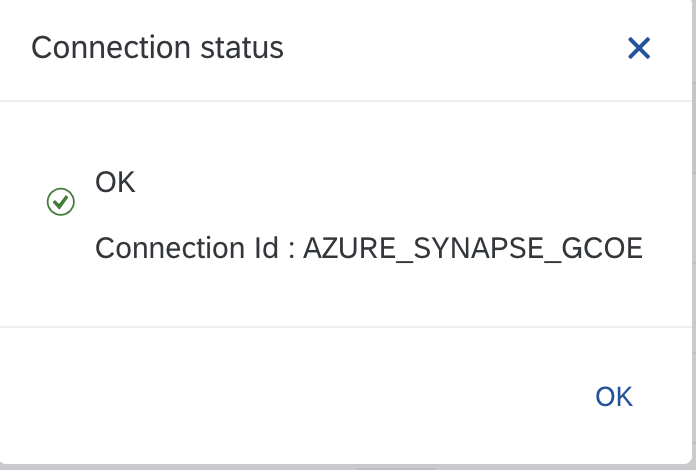
# Create a managed connection in DI

Use the connection type AZURE\_SQL\_DB



Use the parameters from the Azure Synapse details in Azure.

Confirm connectivity:



# Custom Azure Synapse Operator:

For writing to Synapse, we created a custom operator:

## Create the Dockerfile

First, create the dockerfile and tags in the Modeler repository to support the operator:



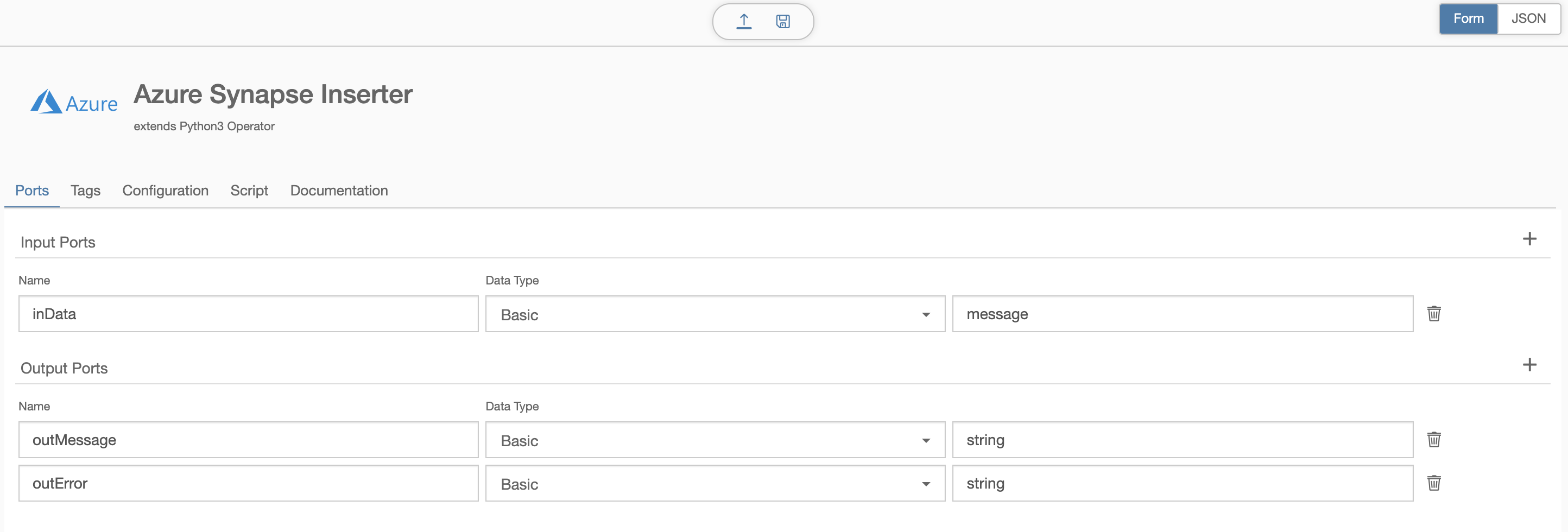
The Dockerfile contents are in the file: **synapse\_operator\_dockerfile.txt**

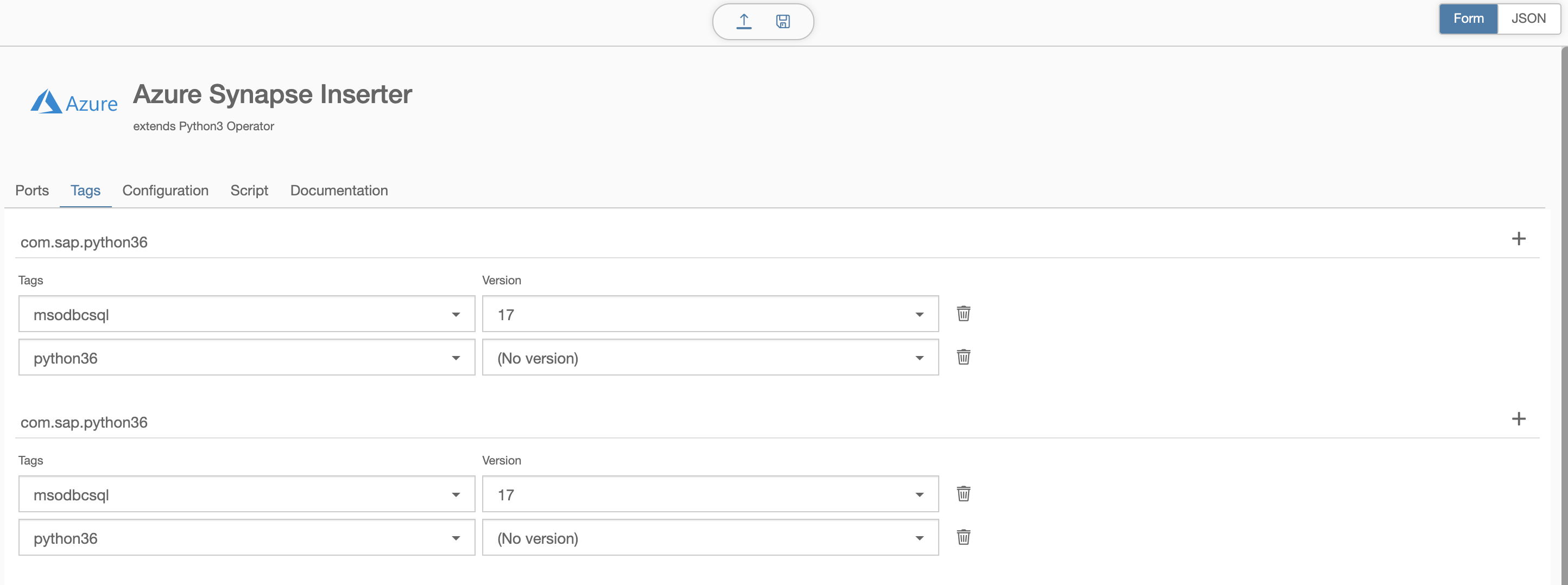
Build the dockerfile and confirm that it completes.

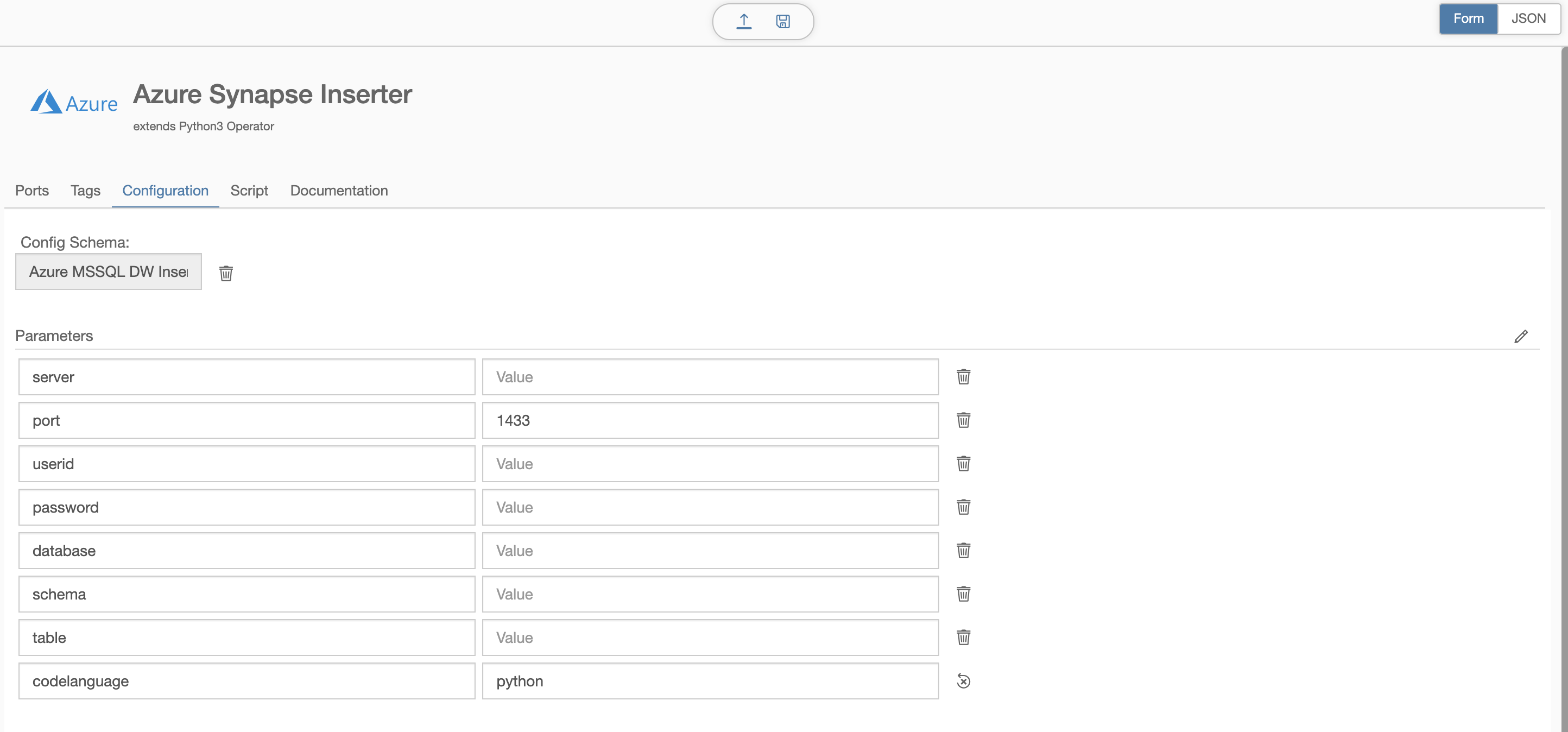
## Create the Operator

Create an operator in the operator section:

The config and script for the operator is in the json file: **synapse\_operator.json**

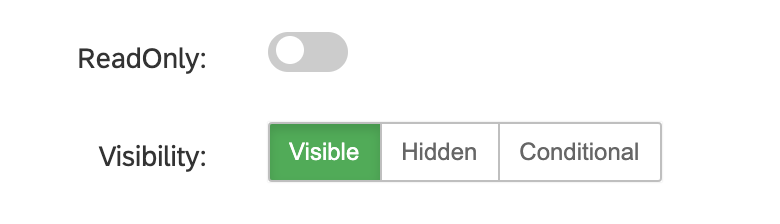


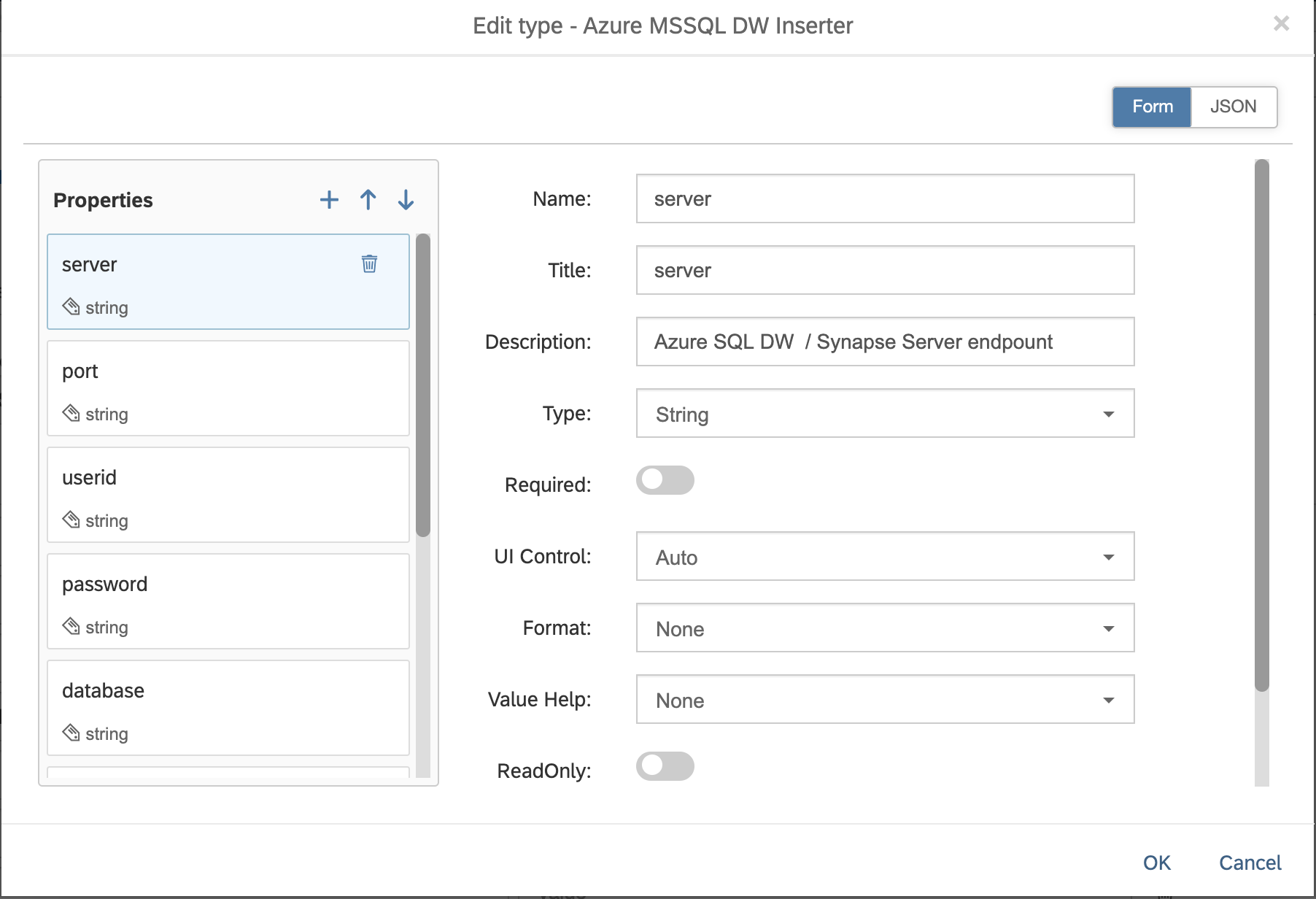


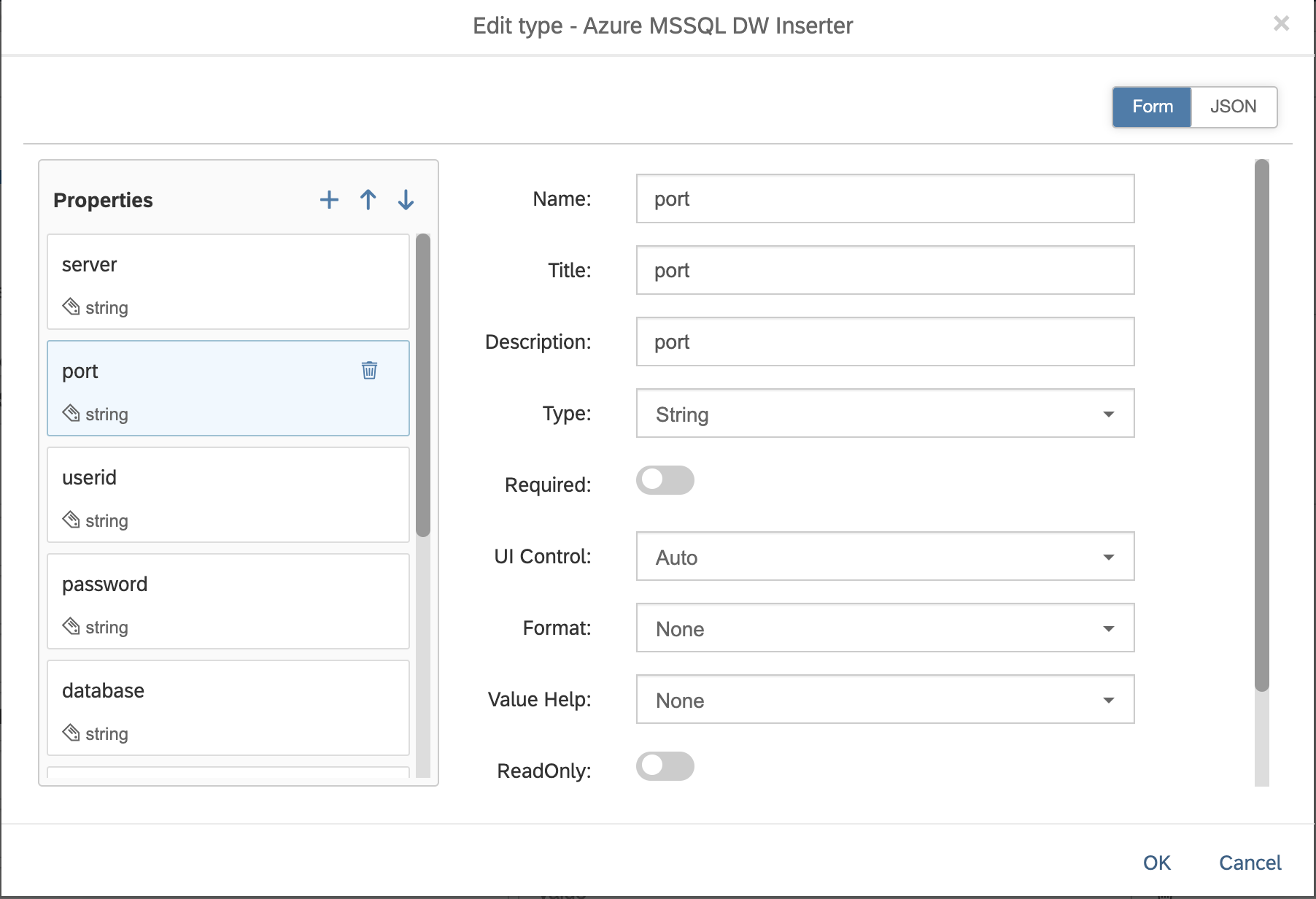


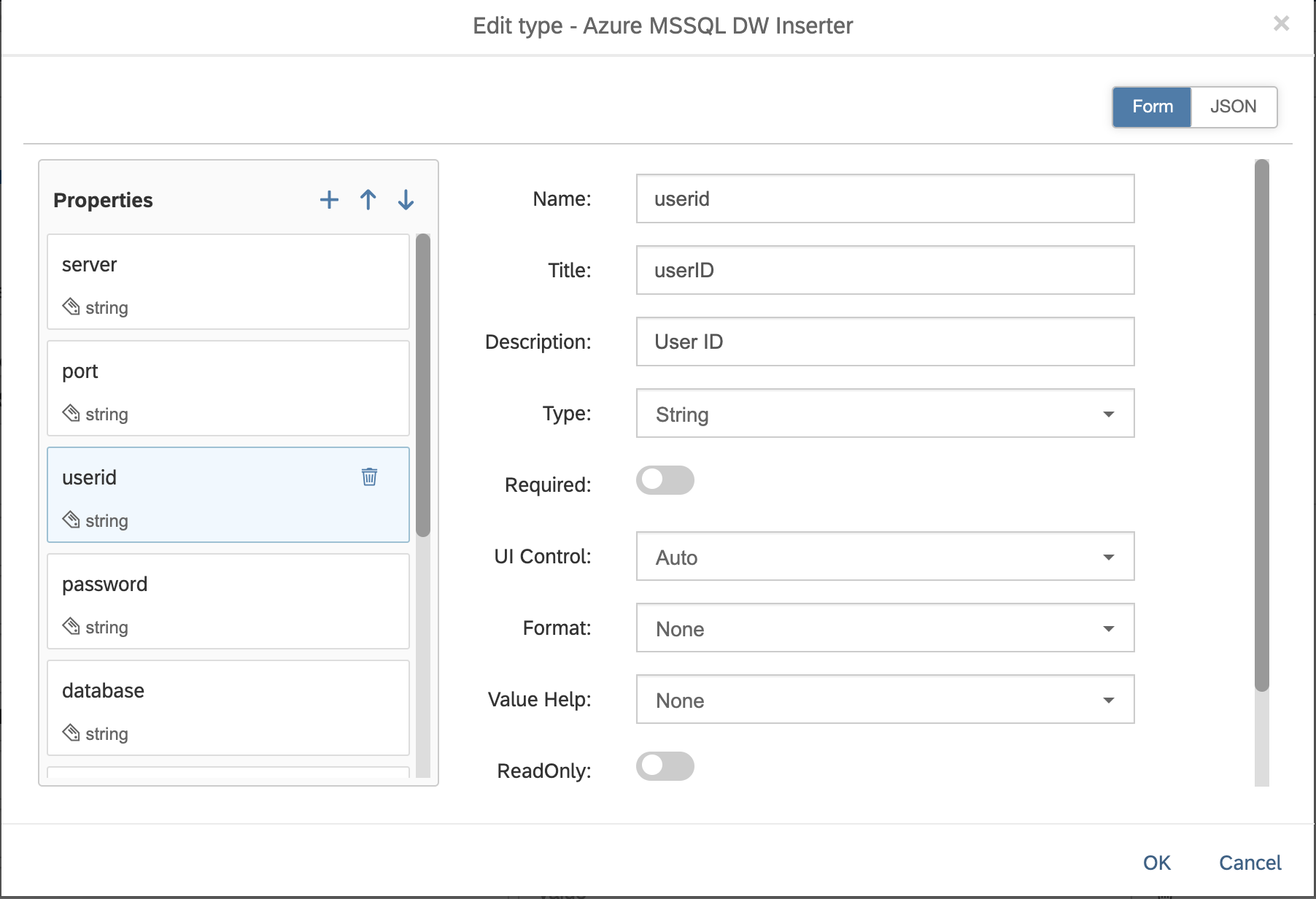
Detailed configuration:

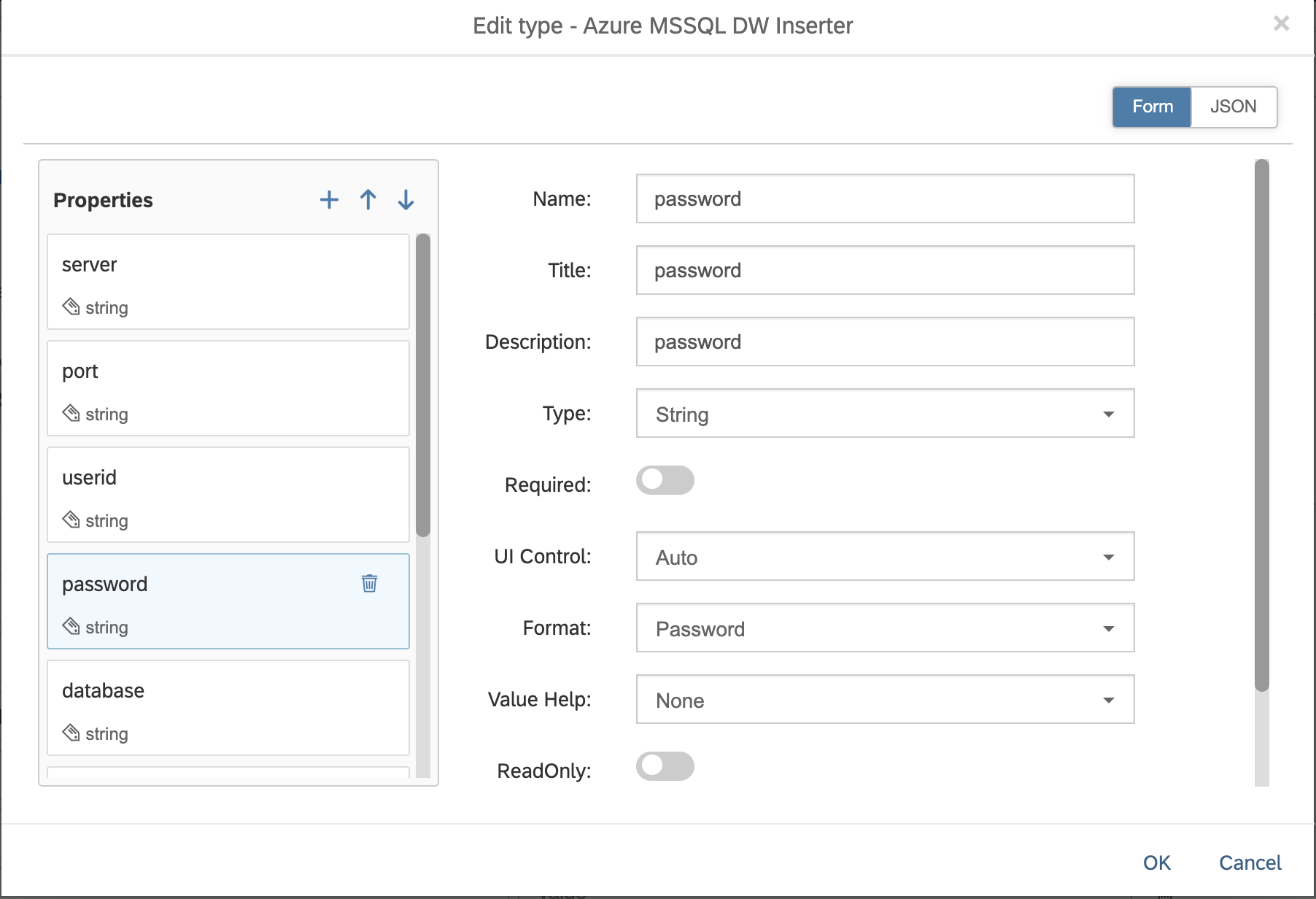
NOTE: for each of the configuration screens, Visible is set to True like below:

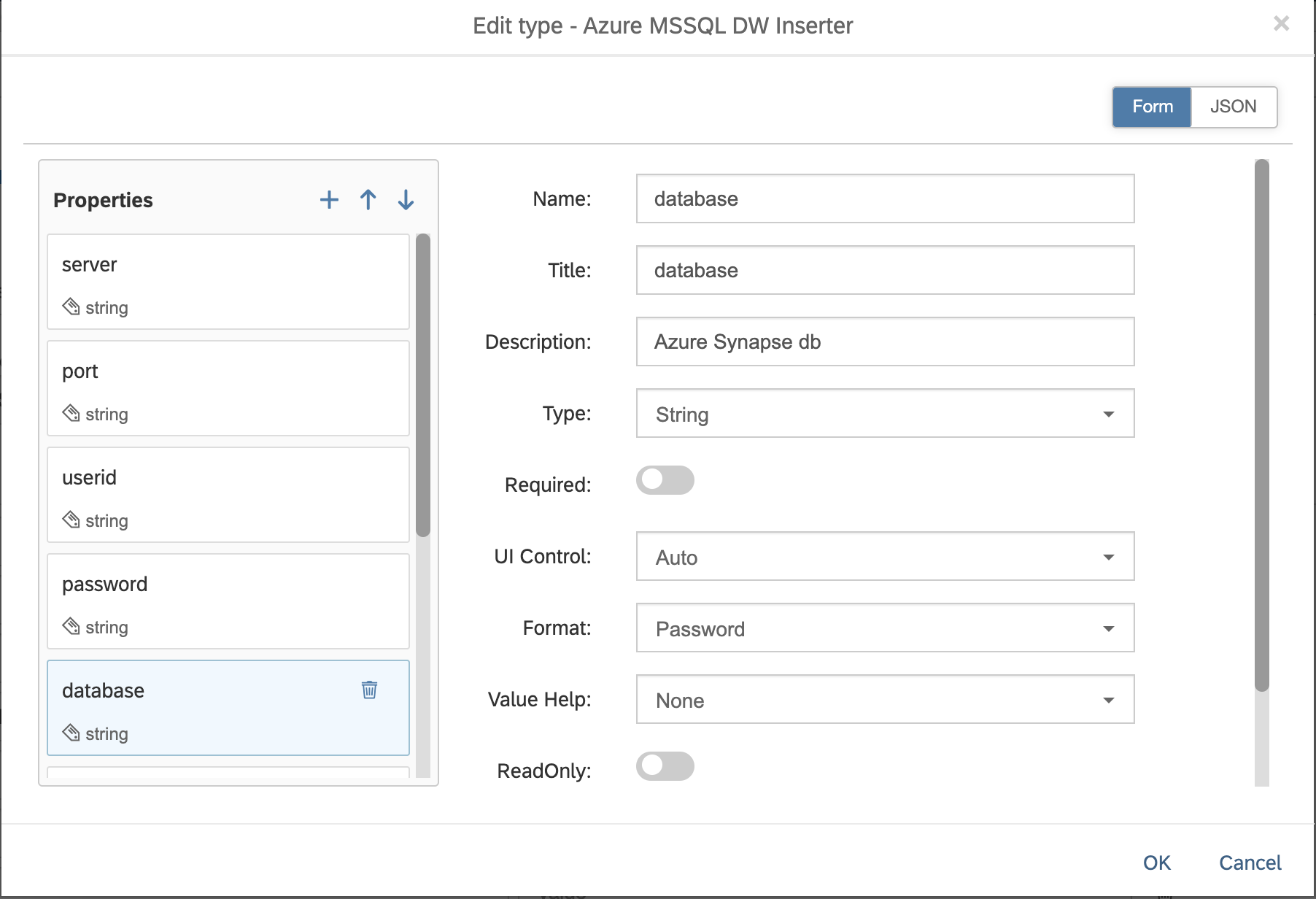


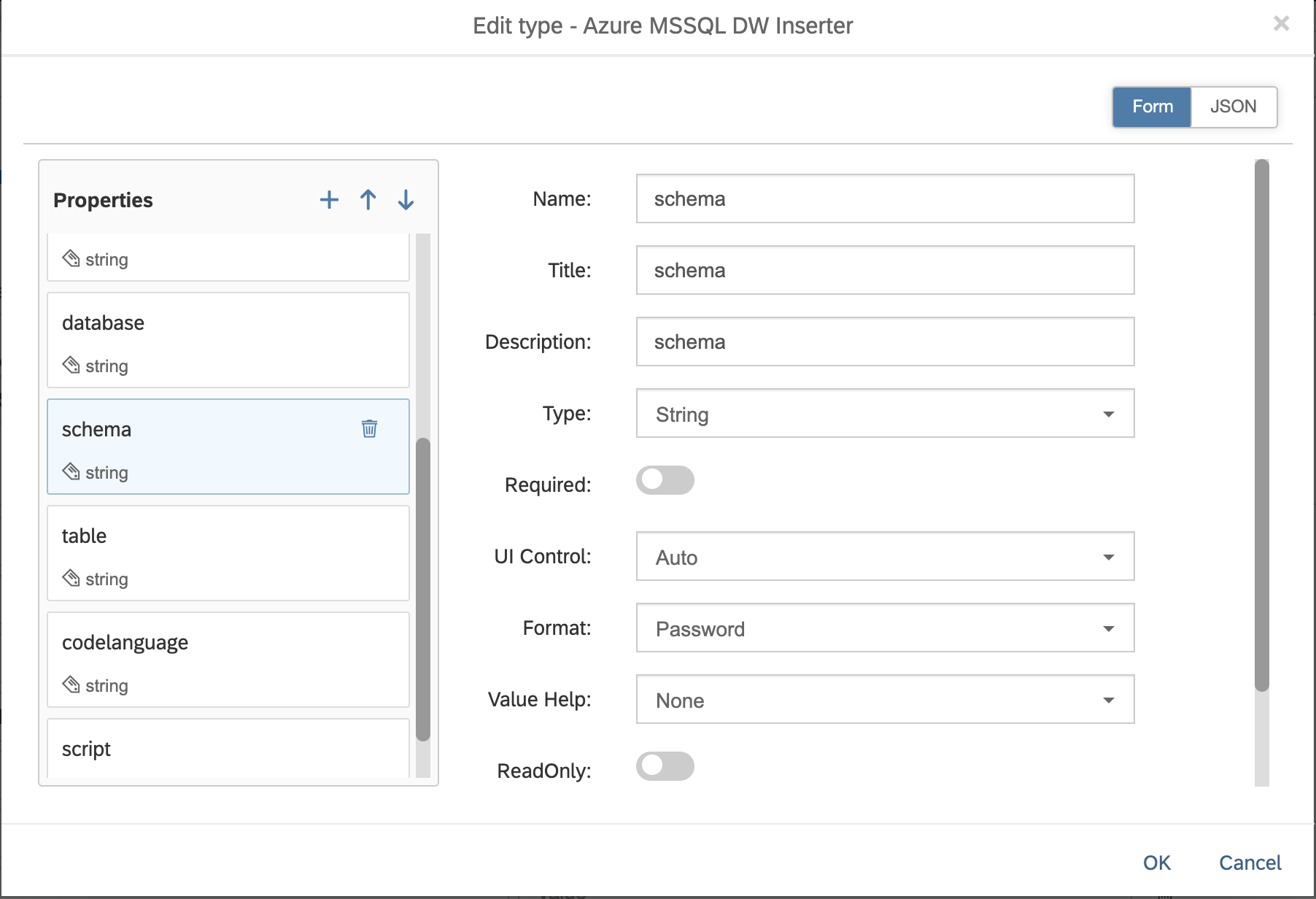


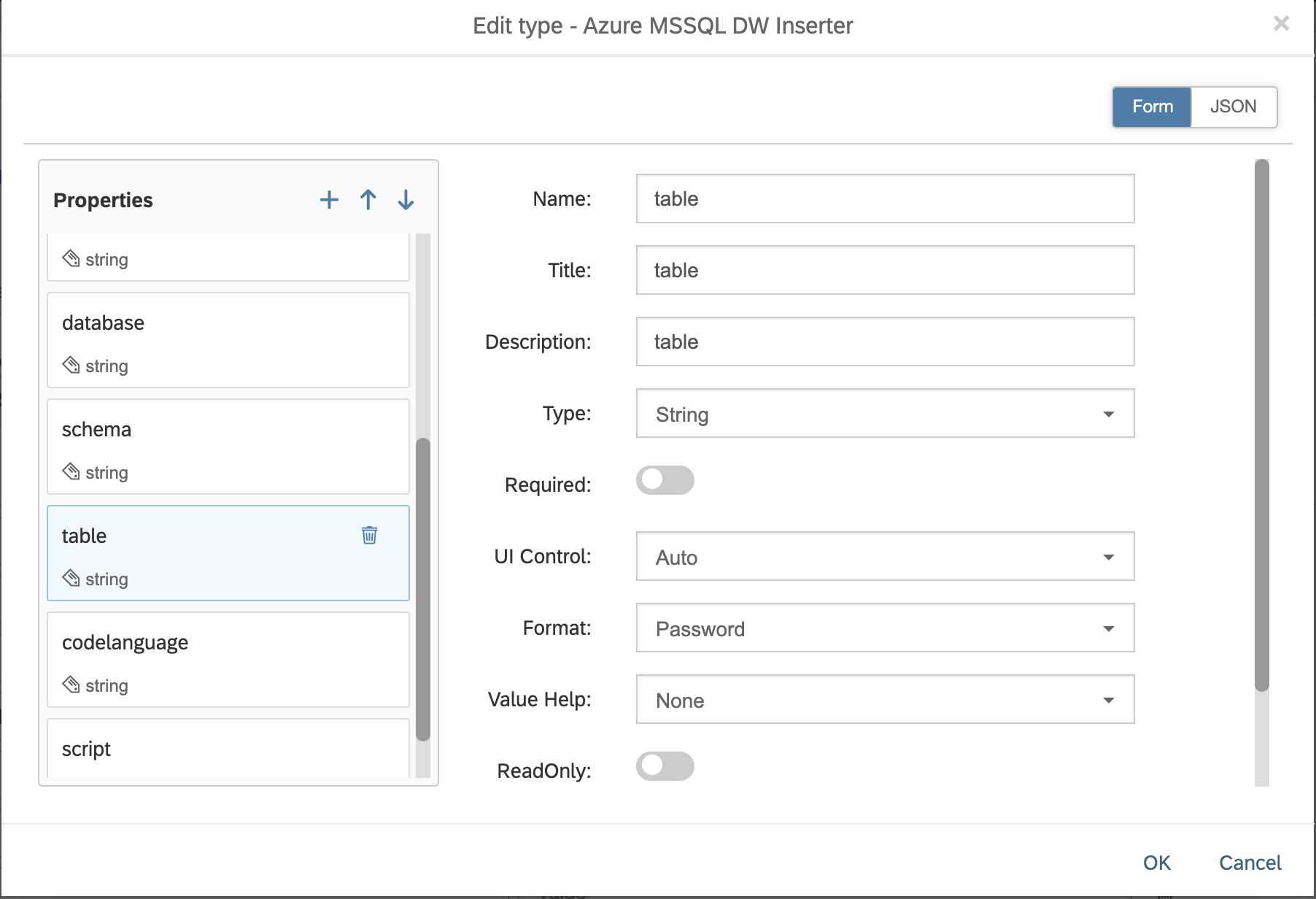


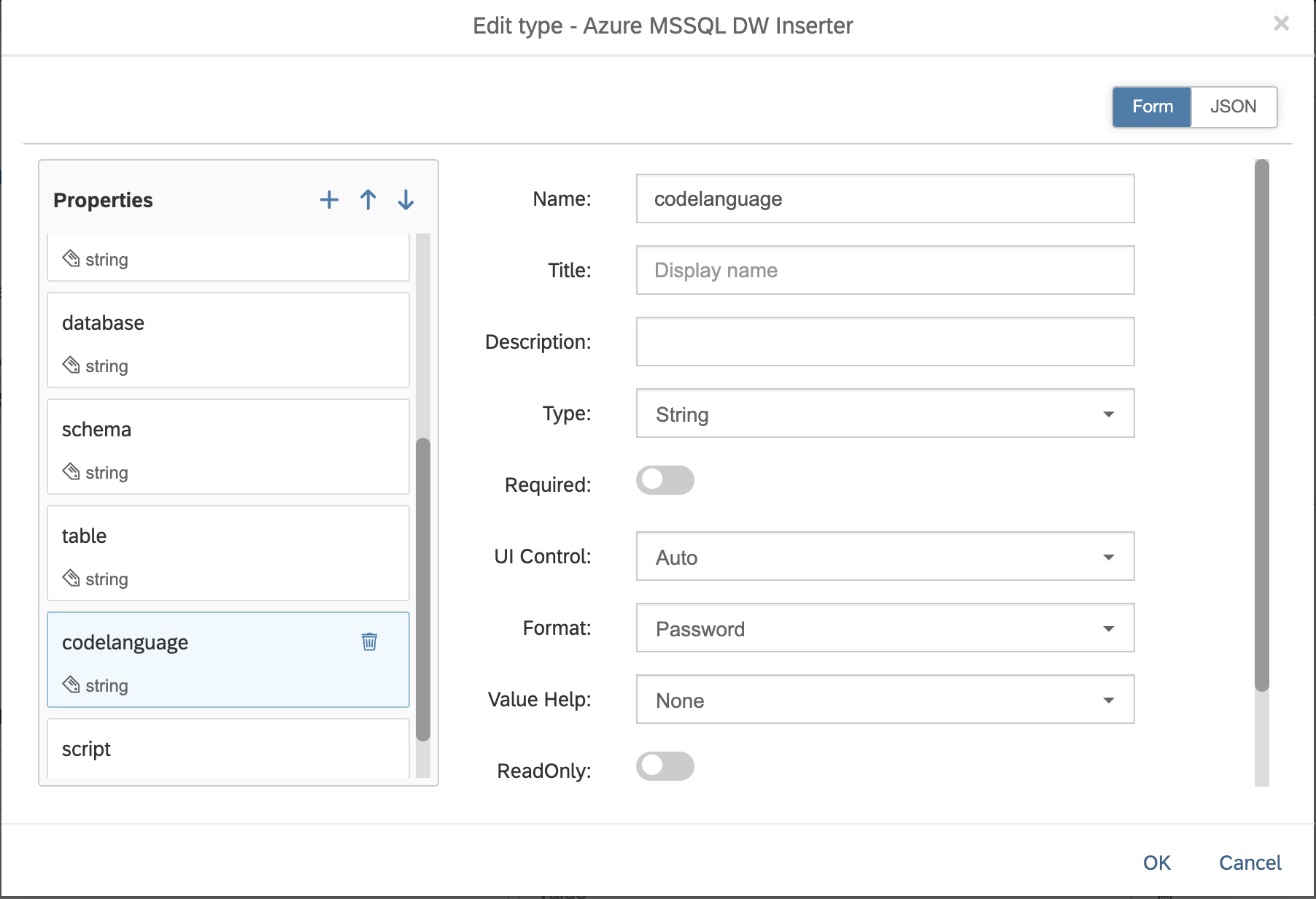


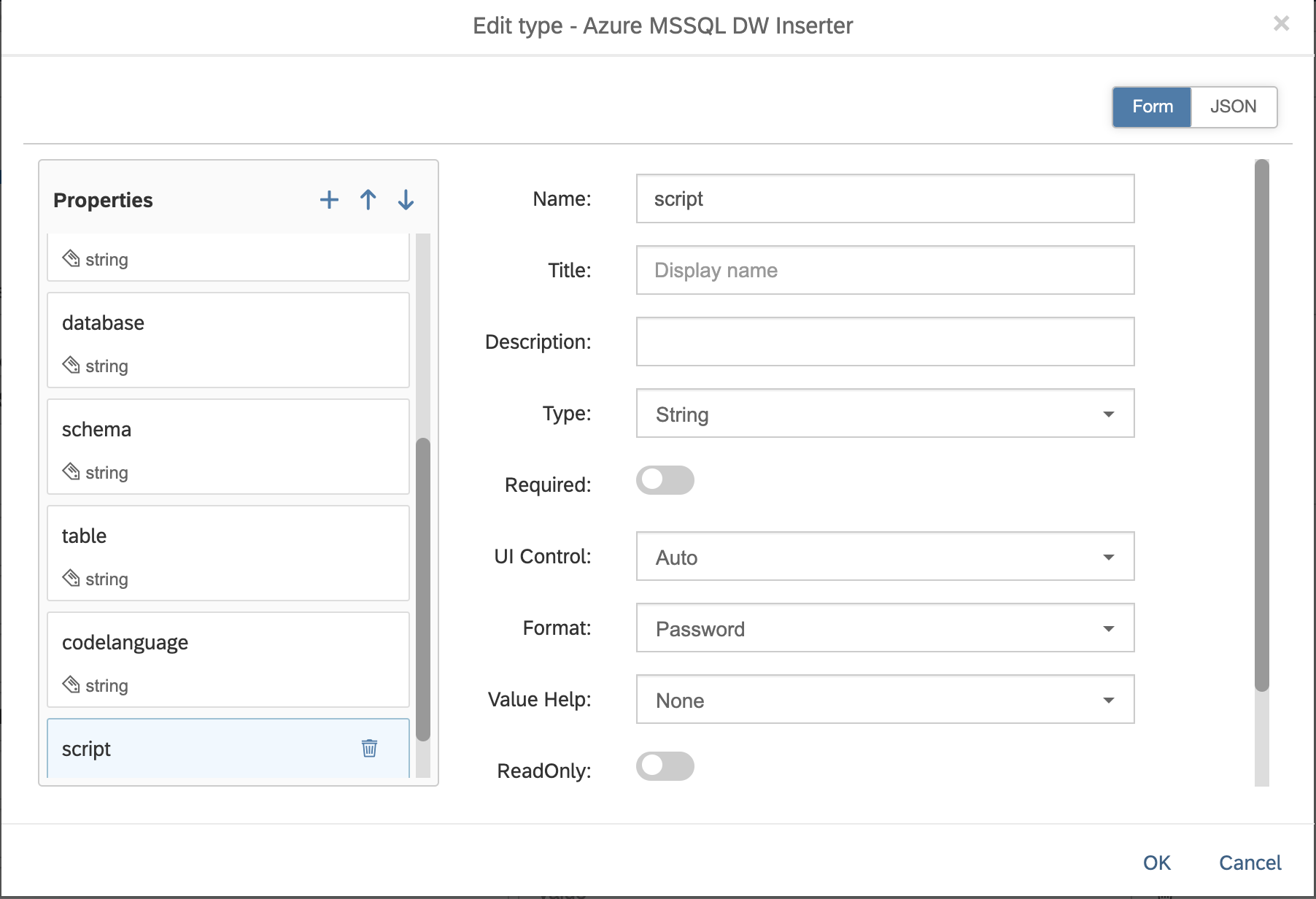


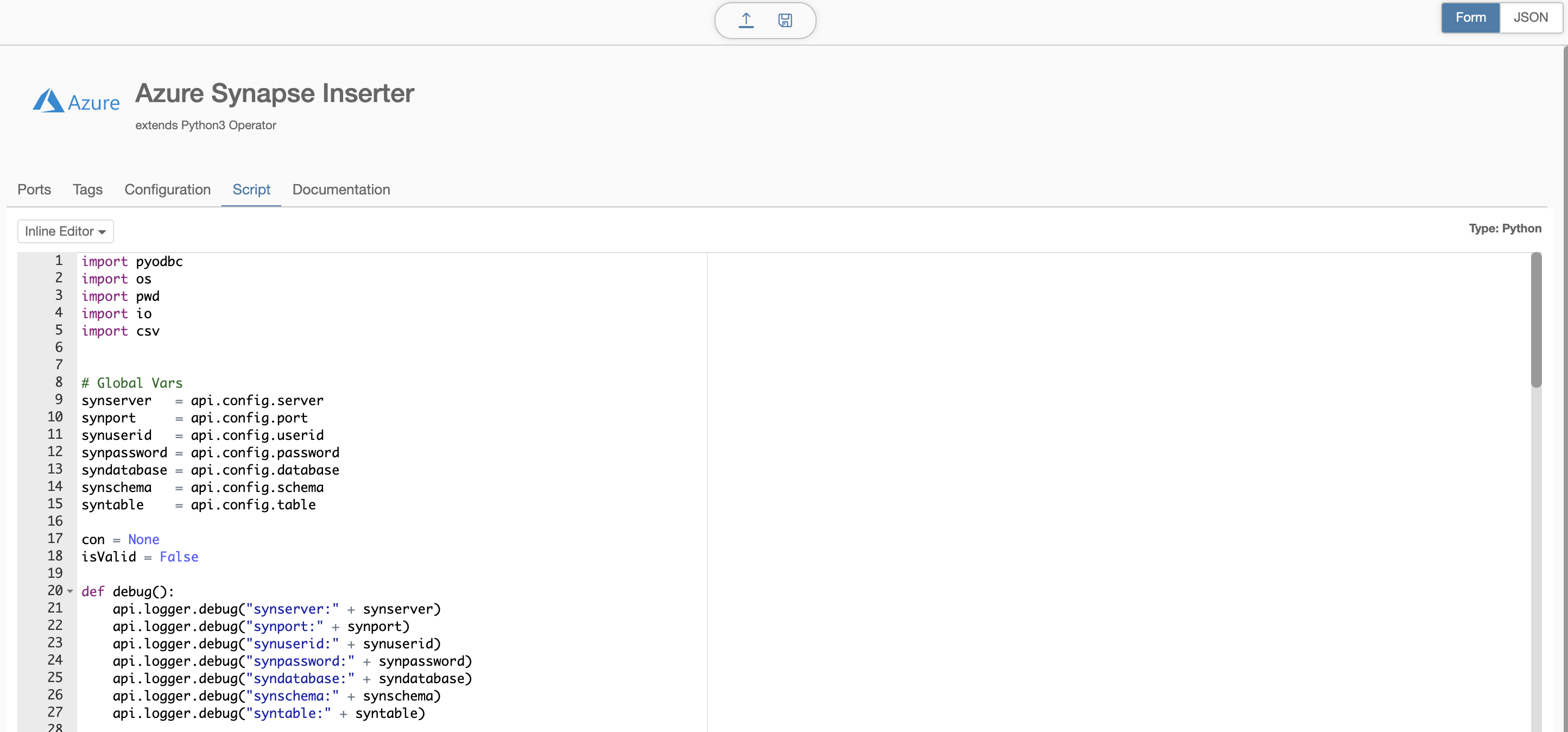








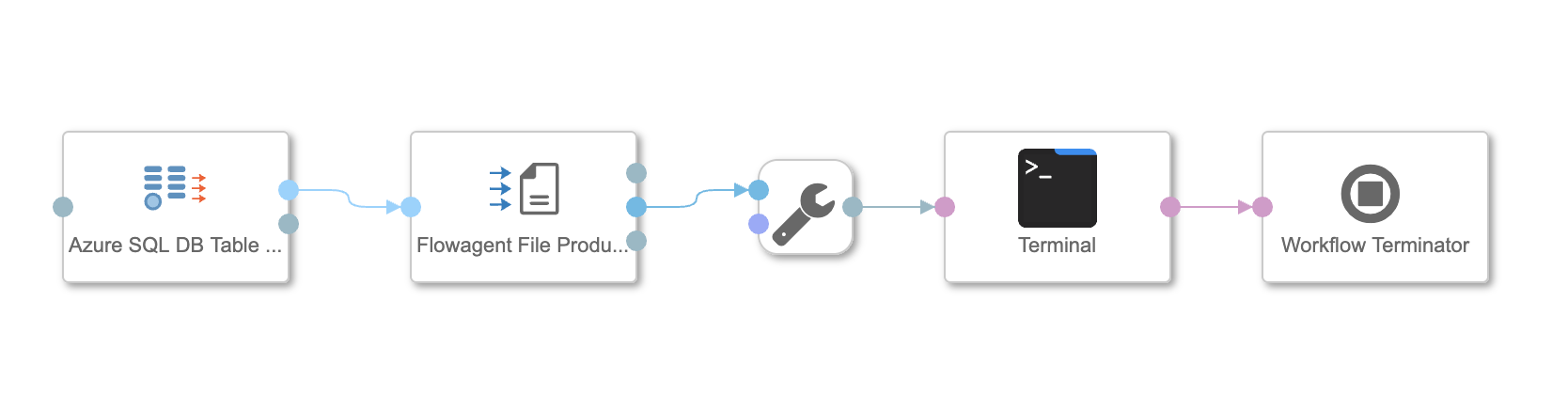




Note that the script here takes the message input from the CSV and builds an insert statement for each row using the header row as columns. It performs quite slow in initial testing, roughly 600 rows/minute. We are looking at an alternative to use the COPY method documented here: <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/quickstart-bulk-load-copy-tsql-examples>

# Create a graph to pull data from Synapse:

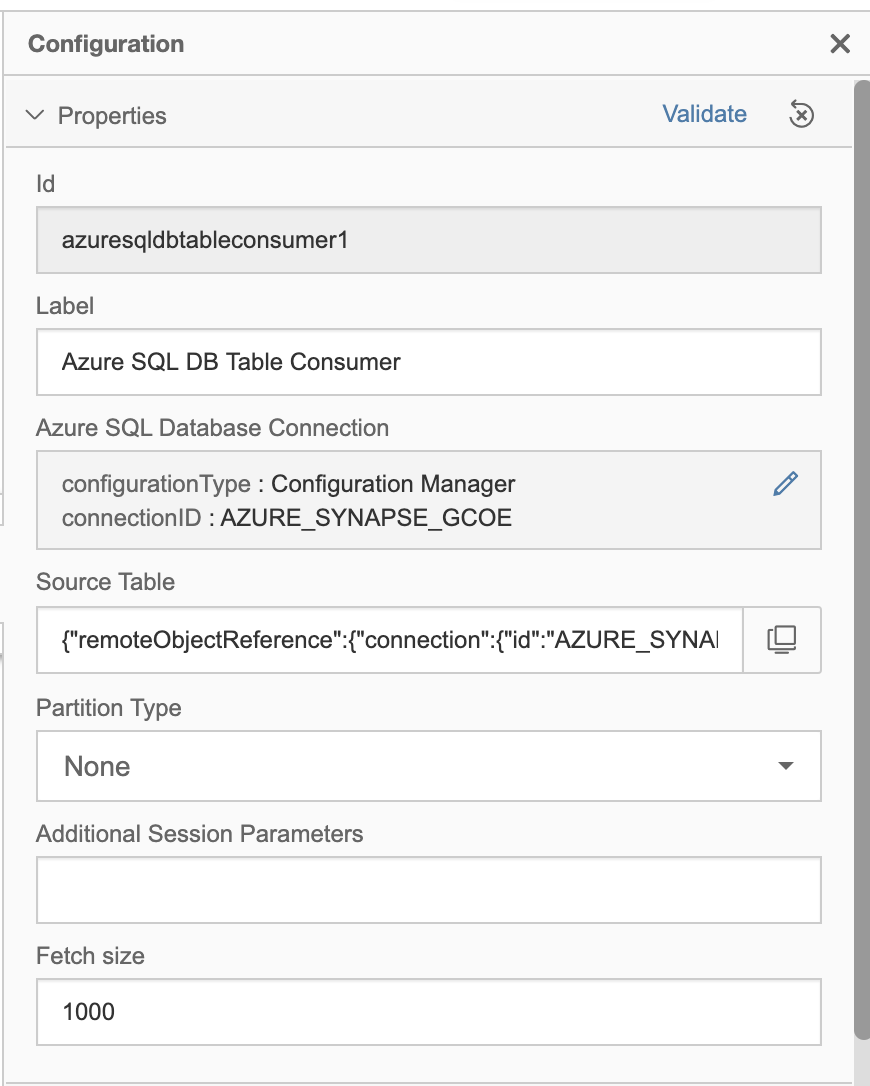
Switch to Modeler panel



Operators:

Azure SQL DB Table Consumer

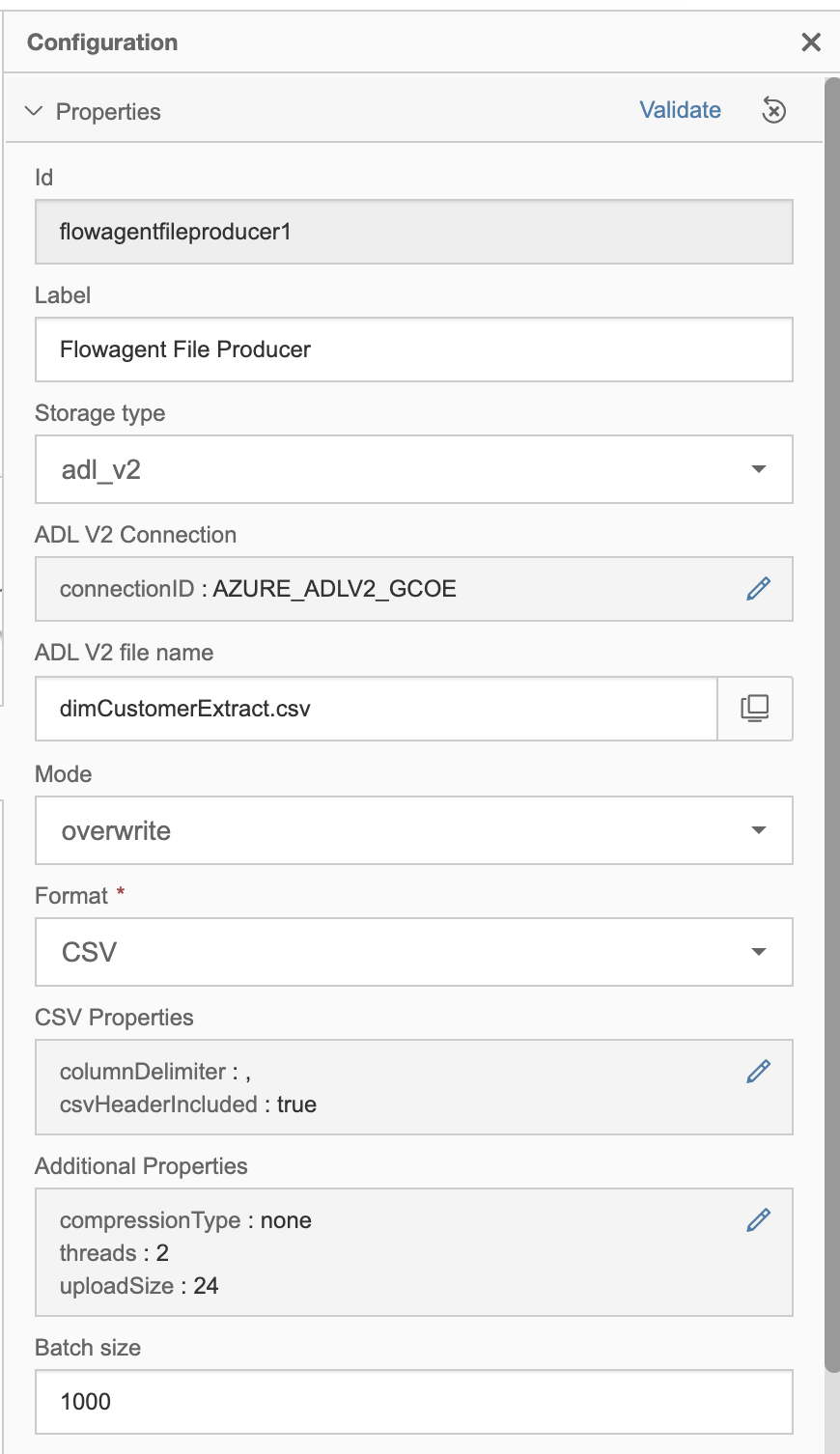
Config:



Use the connection created in the connection manager

The Source Table is a dialog box based on the permissions the user in the connection has to the database. In this demo, we’re using the table dbo.DimCustomer

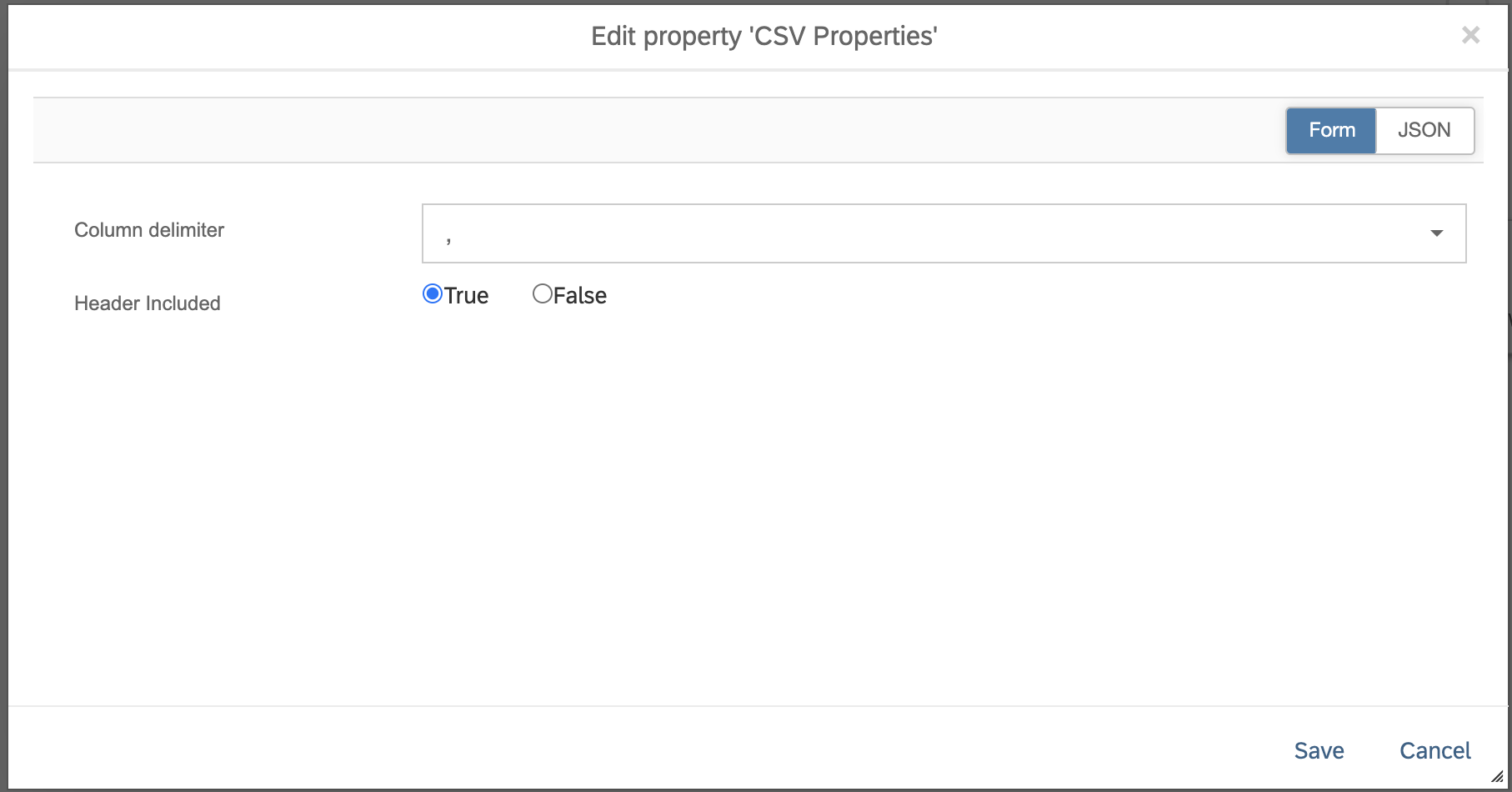
Flowagent File Producer



We have setup a storage account under ADLv2 also, we’ll select it as our storage location.

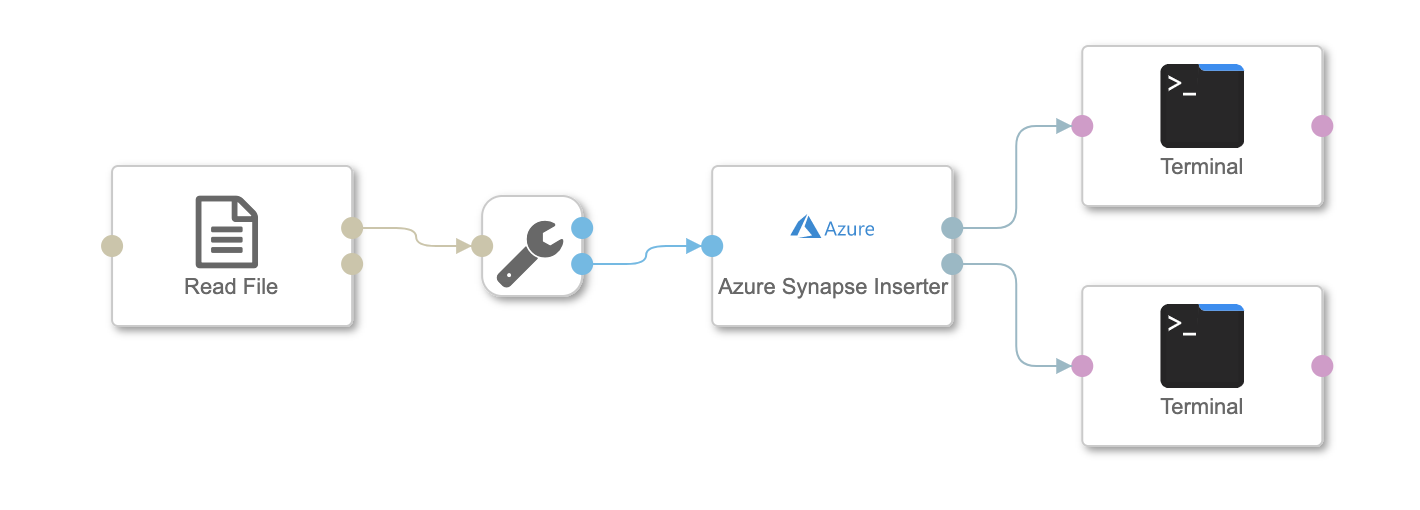
Specify the file name, overwrite and format CSV.

The key item for pulling the data into Synapse is that we use the header row, be sure to toggle it on in the config:

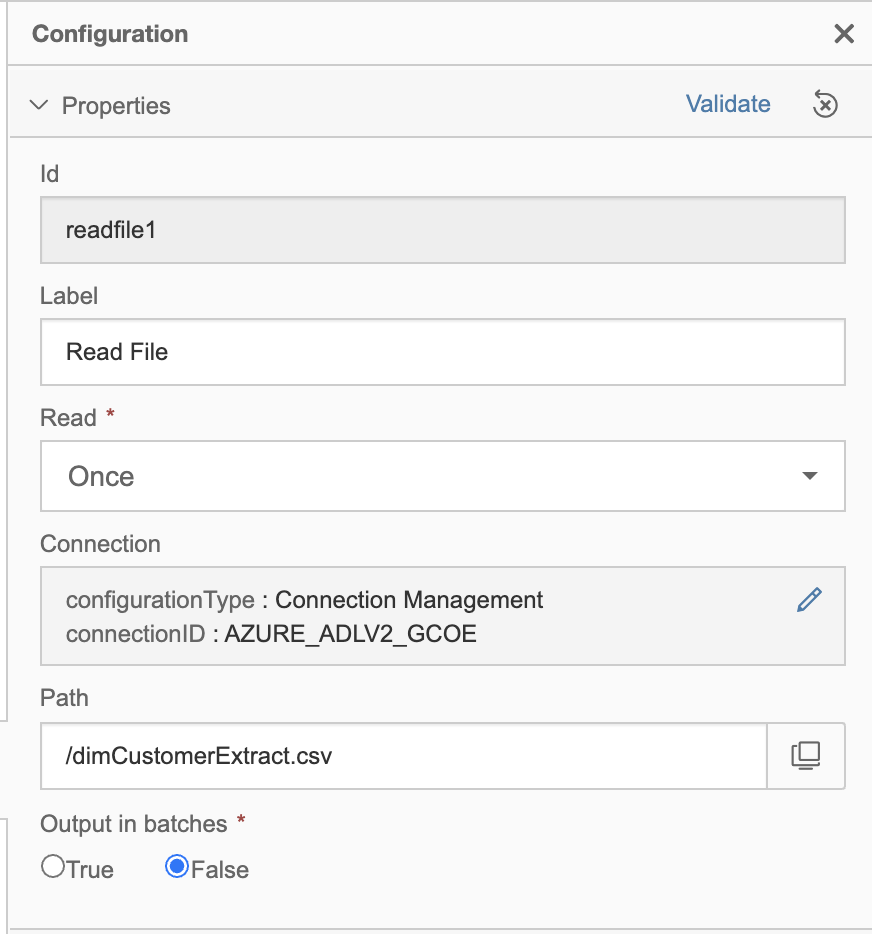


In the pipeline, I’ve added a terminal operator to monitor the progress, but it is not mandatory.

# Create graph to insert data into Synapse using the new operator



Use the **Read File** operator to read the CSV created in the previous pipeline



Connect the Read File to the **Azure Synapse Inserter** Operator and configure it to write to the Synapse connection and table.

## Target Table:

\*\*\*NOTE\*\*\*. Be sure that the target table is already created in Synapse as the operator is only an inserter at this time. I used the following SQL to create the test table:

create table [dbo].[DimCustomer\_test]

WITH

(

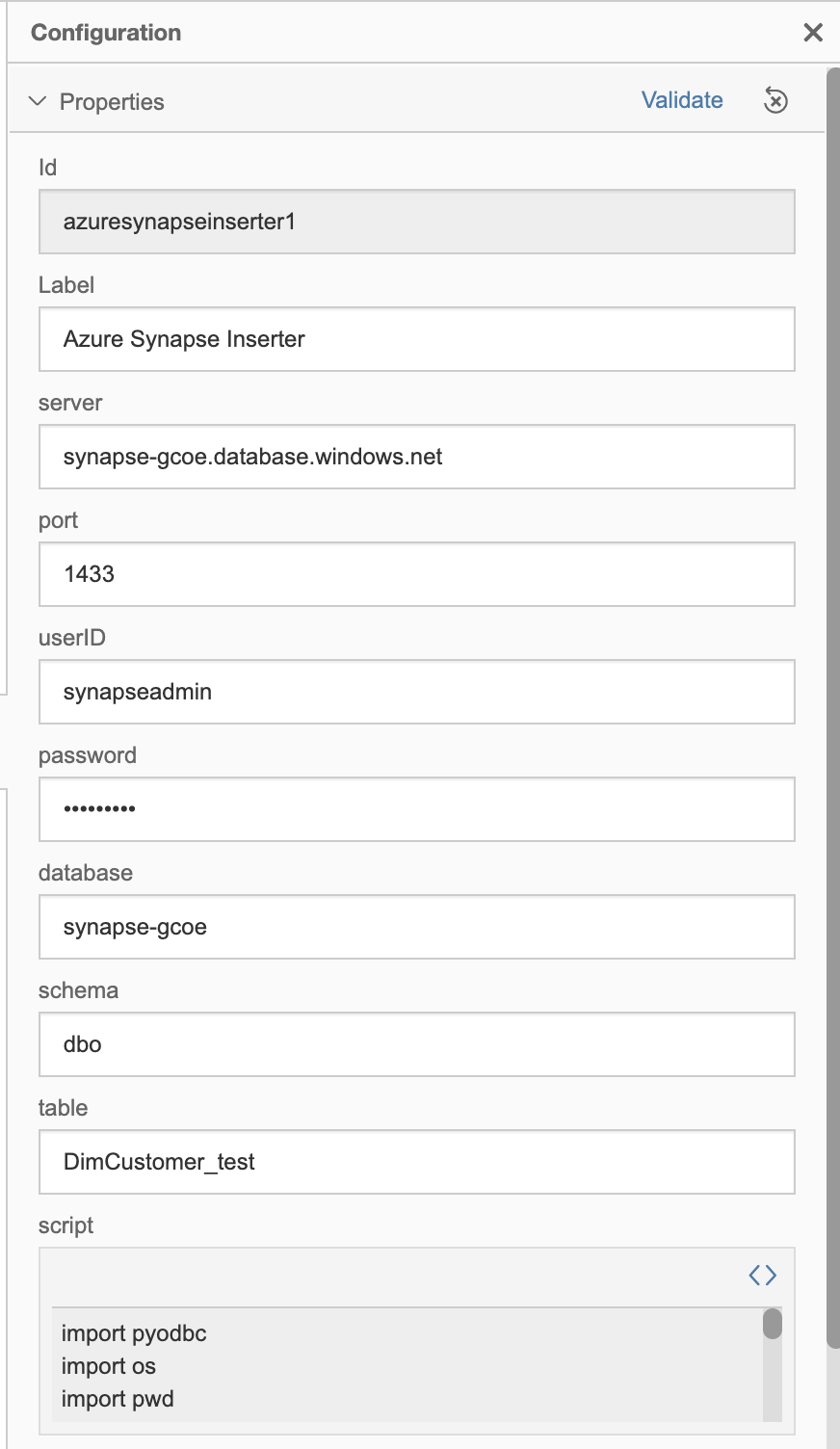
    DISTRIBUTION = ROUND\_ROBIN,

    CLUSTERED COLUMNSTORE INDEX

)

as select \* from [dbo].[DimCustomer]

delete from [dbo].[DimCustomer\_test];



Add the terminals to the outbound ports for monitoring