

Getting Started

In order to begin using the SDK for Python, there are a few prerequisite criteria that must be met:

- The StreamSets SDK for Python must be [installed and activated](#).
- A Python 3.4-3.9 interpreter and the pip3 package manager must both be installed on the machine where the SDK will be used.
- Have at least one [StreamSets Data Collector](#) instance running and accessible. If you're planning to use the SDK with Control Hub, then the Data Collector instance will also need to be [registered](#) with the Control Hub instance.
- Have access to a [StreamSets Control Hub](#) instance with a user account in your organization. Note: Make sure the user account has proper access within the Control Hub organization.

Once you have satisfied the above requirements, you can utilize the SDK by launching a Python3 interpreter shell and importing the relevant modules you wish to use:

```
$ python3
Python 3.9.6 (v3.9.6:db3ff76da1, Jun 28 2021, 11:49:53)
[Clang 6.0 (clang-600.0.57)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>> from streamsets.sdk import ControlHub,DataCollector
```

You now have a Python3 interpreter running with the ControlHub and DataCollector modules imported from the StreamSets SDK! To find out more about the SDK's implementation and usage, check out the corresponding [Control Hub Usage](#) and [Data Collector Usage](#) pages.

For some examples of scripts using the SDK, check out the next section: SDK Sample Scripts.